INVESTIGATION OF SEASAT-A SYNTHETIC APERTURE RADAR (SAR) FOR TOPOGRAPHIC MAPPING APPLICATIONS

By

ABDALLA ELSADIG ALI

B.Sc.(Hons.) (Civil Eng. (University of Khartoum))

A Thesis Submitted for the Degree of

Doctor of Philosophy (Ph.D.) in

Remote Sensing

to the

University of Glasgow

May, 1982

Volume 1

TO

My mother whose support and encouragement
throughout my academic life was never failing;

My wife Afaf and my son Alla-eldin;

The soul of my late grandfather Ali Wad Alhissain
who first taught me the meaning of persistence.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to his supervisor, Professor G. Petrie for suggesting this research topic and for his continuous advice, encouragement and supervision throughout the period in which this research work was carried out. Without his assistance, this work would not have been in its present form.

The author also acknowledges the help given by Mr. B.D.F. Methley for using his general-purpose least squares computer program for the checking of the author's own least squares program, and for his valuable answers to the questions raised by the author during the computational stages of the thesis.

The freedom granted by the Head of the Department of Geography, Professor I. Thompson, to use the facilities available in the Department is deeply appreciated.

Sincere thanks are also due to the following individuals and organizations:-

The Space Unit of the Royal Aircraft Establishment (RAE) at Farnborough for providing the imagery used in this research work and for supplying the necessary information on the images that made their analysis feasible. In particular the very full collaboration of Mr. D.W.S. Lodge and Dr. G. Keyte is greatly appreciated.

The Goodyear/Aerospace Corporation at Litchfield Park, Arizona, U.S.A. for supplying valuable material in the very early part of this research work about the tests they had carried out with their GEMS-1000 radar system.

Dr.B. Makarovic of the Department of Photogrammetry, International Institute for Aerial Survey and Earth Sciences (ITC), The Netherlands, for supplying some material on the work they had carried out on digital monoplotting from conventional aerial photography.

The technicians of the Department of Geography, University of Glasgow, especially Mr. I. Gerrard, Mr. M. Shand and Mr. L. Hill for their assistance and advice in photographing and drawing the figures used in this thesis.

Mrs. A. Strachan of the Department of Philosophy for neatly typing the thesis.

The University of Khartoum, Sudan, for financial support during the period in which this research work was in progress.

And finally special recognition is due to the author's wife Afaf, son Alla-eldin and family: mother, father, brothers, sisters, father- and mother-in-law for their spiritual support and understanding and who endured many years of neglect while this thesis was in preparation..

CONTENTS

next page

SUMMARY

The thesis is concerned with an investigation of the possibilities of generating metric information and carrying out topographic mapping operations from side-looking radar images acquired from Earth-orbiting satellites, as exemplified by the synthetic aperture radar (SAR) system flown onboard the Seasat satellite. Besides the theoretical analysis of the problem, several images covering test areas with different topographic characteristics have been used for extensive and comprehensive tests of the geometric accuracy of the SAR system; for experiments with digital monoplotting techniques applied to the SAR images; and for tests concerned with the detection and interpretation of objects appearing on these images.

The results show that metric information of a limited accuracy can be obtained from satellite SAR images. This could act as the basis for reconnaissance-type mapping at scales of 1:250,000 and smaller. The geometric accuracy actually achieved does, however, depend heavily on the method used initially to process the SAR image data. In this respect, the results obtained with the digitally processed images are superior to those obtained with the optically processed images. The influence of the topographic relief present on the ground is also noticeable and various techniques have been devised and used to eliminate or substantially reduce this effect.

The use of digital monoplotting techniques did not produce as good or as complete a rectification as expected due to the difficulties experienced with the interpretation of the terrain objects recorded on the SAR images. These result partly from the constraints in imaging direction that are an inherent feature of SAR imaging and which make the detection and interpretation of certain objects on an SAR image rather

arbitrary. A further difficulty is the presence of background clutter on all the Seasat SAR images but which is especially noticeable on the optically processed images tested. At the present stage of the development and application of satellite SAR imagery for mapping, the limitations are centered around shortcomings in the image resolution and quality rather than the geometric characteristics of the imagery or the rectification techniques which have been devised and implemented for mapping purposes.

CHAPTER I

INTRODUCTION

The subject of side-looking radar (SLR) and, in particular, one of its variants, synthetic aperture radar (SAR), forms part of the field of remote sensing in which there is so much interest at the present time. As a result, a considerable amount of research work has been undertaken into its applications, especially to the field sciences. However, the use of synthetic aperture radar is beset by many operational problems and, as a result, it has not attracted the same degree of attention as multi-spectral imagery derived from scanners or cameras. Nevertheless, synthetic aperture radar has some special features such as its all-weather, day and night capability and its unique set of responses derived from its unusual geometry and from its operation in the microwave part of the spectrum. These offer potential advantages to certain groups of users - not least those in tropical equatorial or temperate maritime regions of the world where optical sensors have obvious limitations derived from persistent and heavy cloud cover.

Synthetic aperture radar was originally developed in the United States in the late 1950's under military sponsorship. For a considerable period, because of its importance for military reconnaissance, its characteristics and operation were largely hidden under a security blanket and, in particular, access to actual imagery by civilian scientists was highly restricted. From 1970 onwards, SAR imagery began to become available to non-military

researchers and from 1973 onwards, the use of SAR images for mapping purposes began in earnest with extensive coverage of certain developing countries located in the tropical equatorial regions of the world, especially those in South America. All of this activity has been a spur to the development of the fundamental theory, equipment, techniques and procedures required to assist scientists and the mapping community to make the best use of this unique imagery.

Initially, the SAR imagery was acquired from aircraft. In late 1978, with the launch of Seasat, the first opportunity arose for civilian scientists to acquire satellite SAR data. Because of the very different operating conditions, particularly the high orbital altitude of the satellite, the Seasat SAR images have many very different characteristics to those encountered in airborne SAR work. In particular, the rapid and wide coverage of a satellite SAR and the ease of repetition of imagery over a particular area or site appeared to offer many attractions to users. Furthermore, many special problems would be expected to arise with satellite SAR imagery, especially those associated with data processing and the geometric rectification of individual images.

It is the purpose of this dissertation to report on an analysis of the basic geometry of the Seasat SAR imagery and on a series of experiments with the imagery carried out over test sites in the United Kingdom to establish its potential for topographic mapping. There are still very large and remote parts of the world which have only the most rudimentary small-scale cartographic coverage. These include areas of heavy cloud such as those experienced in parts of the present author's home country (Sudan).

The possibility of its exploitation for use in the Sudan certainly gave an additional stimulus and interest to this subject quite apart from the obvious and intrinsic value of conducting a research topic which is of great interest to the surveying and mapping community as a whole.

The next four chapters of this dissertation constitute an introduction to the work which has been undertaken on the Seasat imagery. In Chapter II, the basic principle and operation of side-looking radar (SLR), together with a detailed consideration of its resolution are outlined, based initially on its simpler real-aperture form. This is followed by a simplified introduction to the basic concepts of synthetic aperture radar (SAR) based on both the holographic and doppler points of view of the subject. This leads immediately to Chapter III and to a discussion of the principles of SAR data processing both by optical and digital methods which form a fundamental and integral part of the total SAR system needed to generate the images required by the users. In Chapter IV, the characteristics of the Seasat satellite, its operation and its sensors are first outlined, followed by a more detailed discussion of the Seasat SAR system itself, including the reception of the image data telemetered from the satellite. The next chapter (V) gives an account of the optical and digital processing carried out to produce the images actually used in the author's experimental work.

In Chapter VI, a detailed analysis of the geometrical characteristics of SLR imagery is presented with special reference to the particular characteristics of satellite-borne SAR images.

This leads directly to the derivation of the mathematical basis for the procedures used for the rectification of the processed SAR images carried out by the author. This is followed by a chapter (VII) in which a review of topographic mapping from SLR imagery is given, together with a summary of the experimental work on the geometrical aspects of this imagery which has been carried out to date.

The procedures devised and used to test the geometric accuracy of a representative set of Seasat SAR images are then outlined in Chapter VIII, including an account of the computer programs written specifically for the purpose. The actual results from these tests, together with a detailed analysis of the effectiveness of the different procedures used, are presented in Chapter IX. This is followed in Chapter X by an account of an experiment concerned with the plotting of topographic detail for one of the test areas using digital monoplotting techniques, including a description of the procedures and programs devised for the purpose and of the results achieved in the test. In Chapter XI, the results of a series of studies concerned with the interpretation of the test images are presented and discussed. Finally, in the concluding chapter (XII) the overall results of the experimental work carried out within the present projects are summarized and various recommendations for possible future research work are made.

CHAPTER II

THE PRINCIPLES OF SIDE-LOOKING RADAR

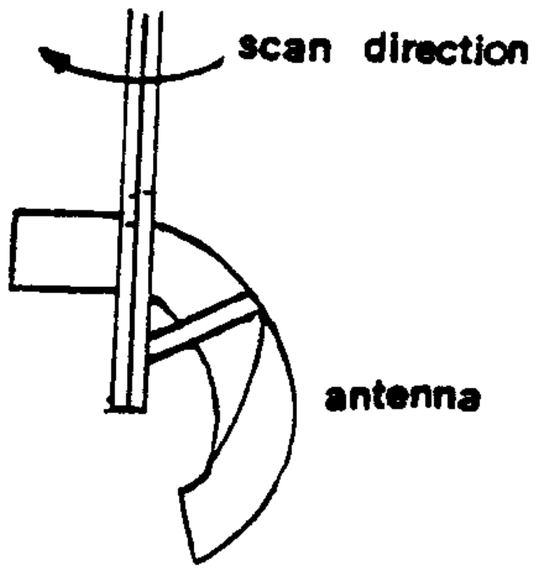
2.1. Introduction

Investigations into the use of radar imagery for mapping began during the Second World War and continued through the late nineteen forties and early fifties. The basis for these original investigations was the imagery generated by airborne Plan-Position Indicator (PPI) radars.

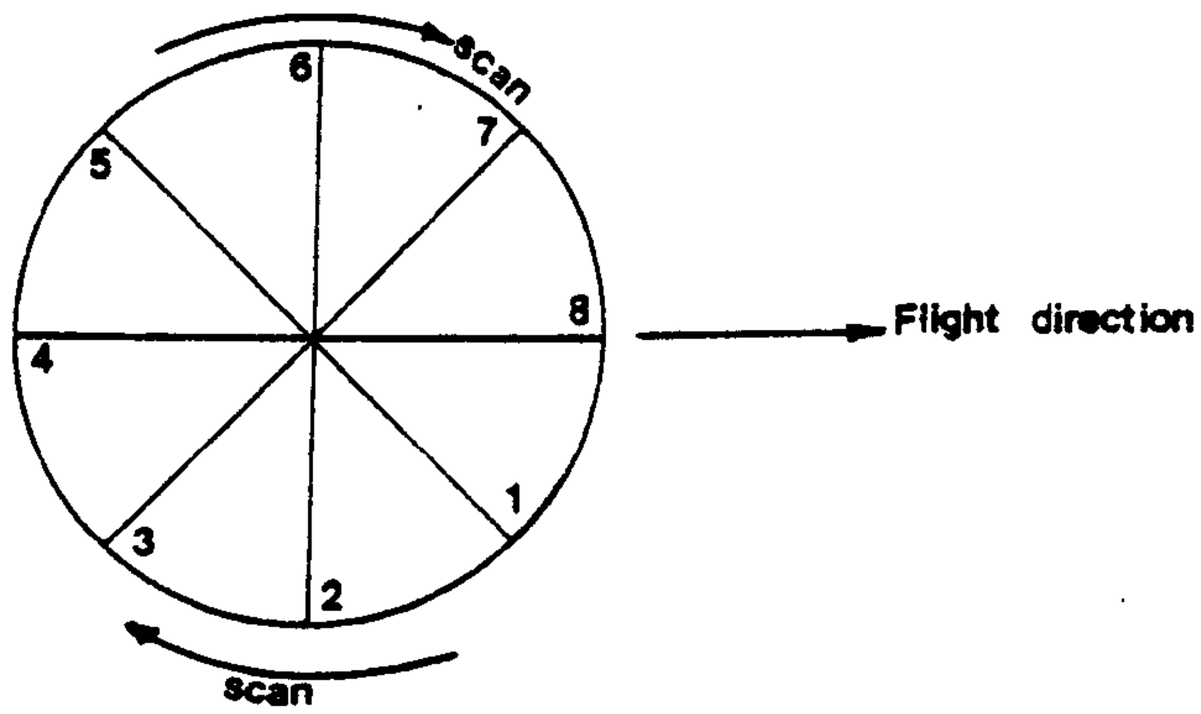
2.1.1. Plan-Position Indicator (PPI) Radar

In PPI radars, a circular scan of the terrain is produced by the rotation of an antenna which both emits pulses of electromagnetic energy towards the ground and receives back the reflected signals (Fig.2.1.a). The time intervals between these emitted and received pulses are measured and translated into distances. The successive positions of reflections received during the circular scan are used to build up an image which is displayed and stored on the face of a cathode ray tube (CRT) screen. Frame photographs of the CRT display are taken by synchronisation of the camera shutter in relation to the rotation of the antenna.

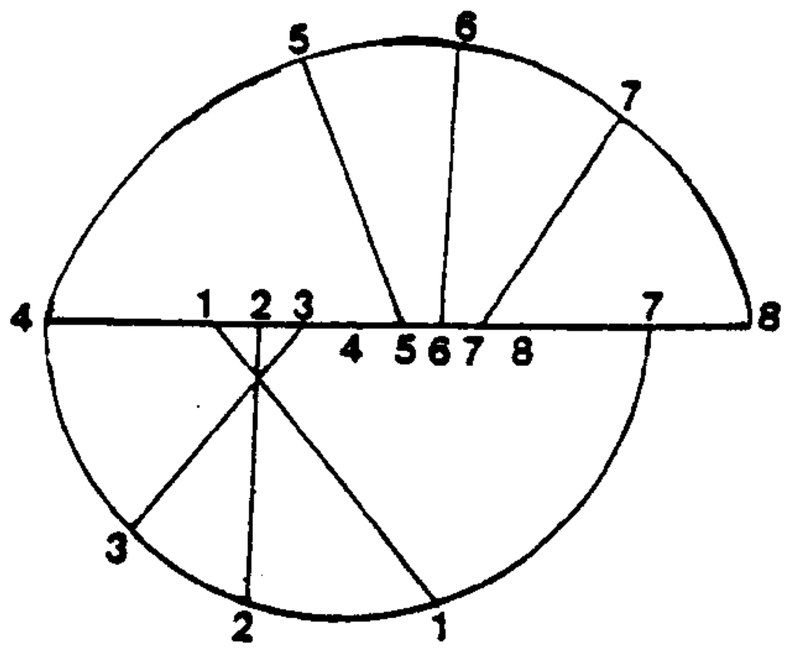
A helical distortion results from the forward motion of the aircraft combined with the 360° rotation of the PPI antenna (Fig.2.1.b). This results in a cycloidal image rather than a circular one. This distortion has to be removed (Fig.2.1.c) if any mapping application is to be contemplated. Further difficulties result from changes in yaw (K), roll (w) and pitch (Ø) of the aircraft and its radar resulting in gaps and double imaging. A great deal of effort went into experiments to



(a) PPI radar



(b) uncorrected cycloidal image on CRT



(c) corrected image

Fig. 2.1 The geometry of the PPI radar

exploit the possibilities of planimetric (two-dimensional) mapping from PPI radar imagery. These are well summarised in the book by Levine (1960).

2.1.2. Side-looking Radar

Because of the limitations of the PPI radar imagery, attention was focussed on an alternative type of radar known as side-looking radar (SLR) both for reconnaissance purposes and to provide an all-weather topographic mapping capability.

Basically, there are two types of side-looking radar - real-aperture and synthetic-aperture radar. The former was developed first and the latter somewhat later. Although this present study has been conducted using synthetic-aperture imagery, a brief account of the principle of real-aperture radar acts as a very suitable and necessary introduction to synthetic-aperture radar which has been derived and developed from it.

2.2. Principle of Real-aperture Radar

Fig. 2.2 shows the geometry of a side-looking radar of the real-aperture type. A short burst of microwave energy is created by a pulse generator, passed to a transmitter and emitted from a long, rectilinear antenna situated parallel to the longitudinal axis of the platform (normally an aircraft) on which it is mounted. The construction of such an antenna is shown in Fig. 2.3.

Every transmitted pulse illuminates a very narrow strip of the ground located to the side of the aircraft and at right angles to the aircraft track. When this pulse hits the terrain, it is reflected. Parts of this reflected energy return to the radar antenna in time sequence according to the distances of the reflecting elements from the

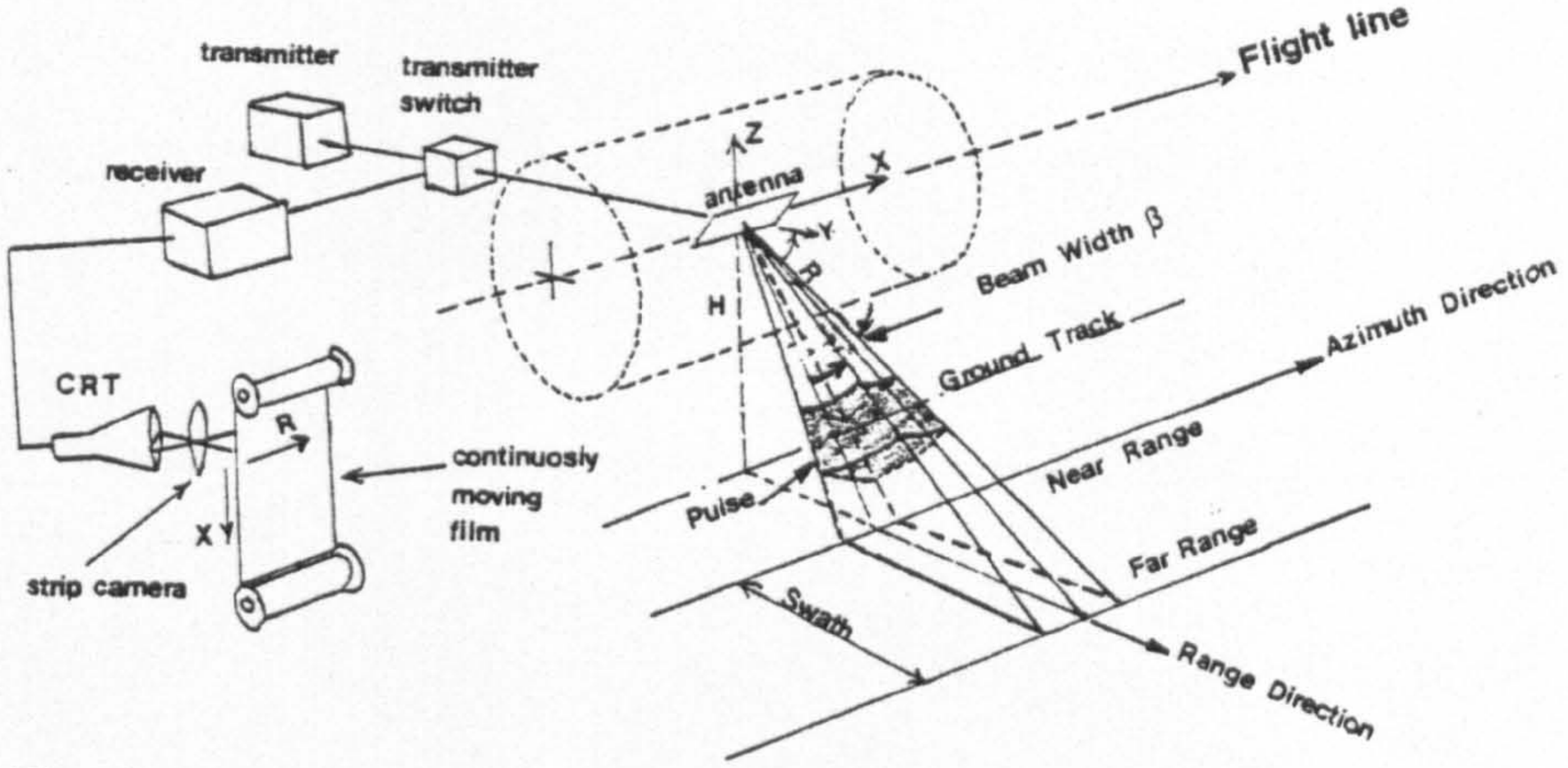


Fig. 2.2 Real-aperture SLR geometry

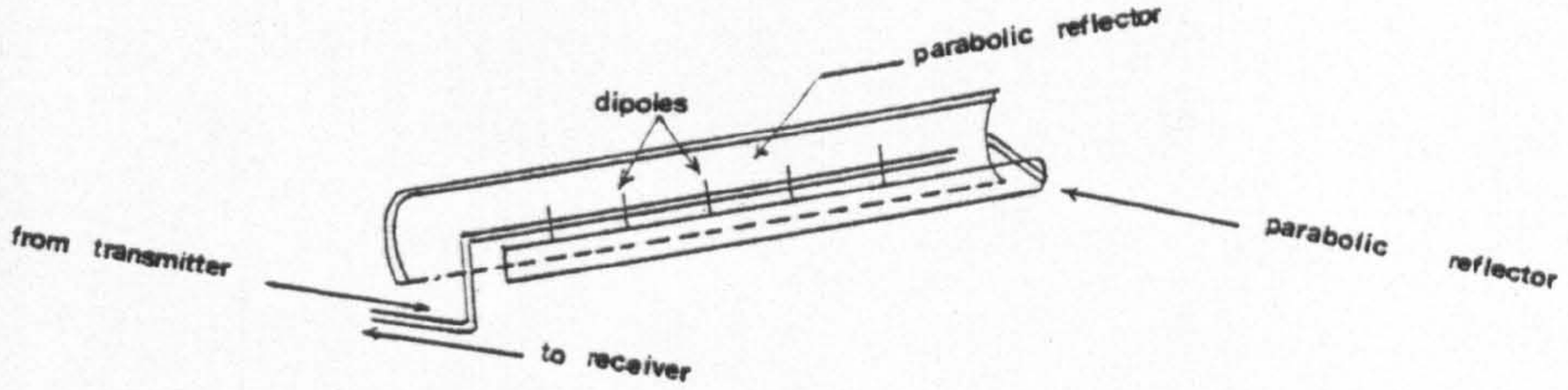


Fig. 2.3 Antenna particulars in SLR

antenna (Fig.2.4). The stream of reflected pulses is picked up by a sensitive radio receiver (Fig.2.2). The time between the original emitted pulse and each of the returning pulses can be measured very precisely so that the range of each reflection is known. The received pulses are amplified, detected and transformed into a series of electric analogue signals. The intensities of these analogue (or video) signals are then used to control the brightness of the flying spot of a cathode ray tube (CRT), the position of each spot on the face of the tube being determined by the times of the successive returns. Each time the transmitter emits a new pulse then, after a short time delay (known as sweep delay), the flying spot starts a new sweep across the tube and a fresh image is generated based on the new set of returns. By means of a camera, the linear image appearing on the CRT screen is recorded on to photographic film.

Due to the forward motion of the aircraft successive lines on the CRT cover adjacent areas, thus forming a continuous strip image of the terrain. This requires an exact synchronisation of the film movement both with the appearance of the image on the CRT screen and with the movement of the aircraft itself. The information regarding the aircraft speed is normally provided by Doppler radar since the standard type of airspeed indicator does not provide a sufficiently accurate determination of the velocity of the aircraft for the purpose of generating the continuous strip imagery.

The coordinates (x,y) of an individual point on a radar image correspond to the two linear coordinates of a cylindrical coordinate system whose axis is the flight path (Graham, 1974b). The former (x) coordinate is proportional to the distance travelled along the flight path and the latter (y) to the distance from the flight path to the target.

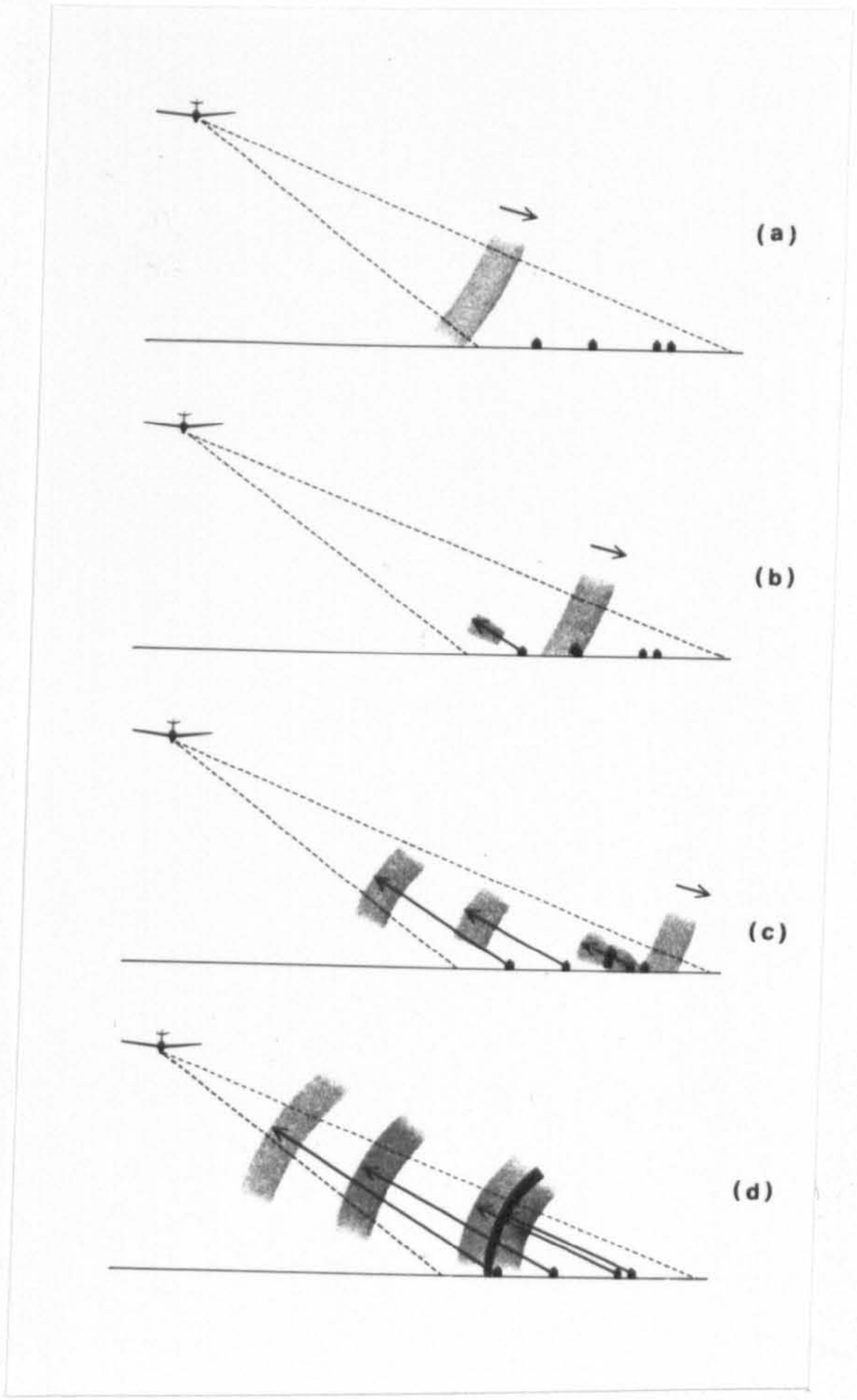


Fig 2.4 Range of objects is determined by the time between the emitted and the returning pulse for each object

There are two modes of presenting the SLR data. These are termed the slant range and the ground range presentations respectively.

(i) In a <u>slant range</u> presentation, the speed of the flying spot across the CRT is constant. Thus the distance of the spot at any moment from its initial position corresponds to the actual range or distance between the antenna and the ground object.

(ii) In a <u>ground range</u> presentation, the speed of the flying spot is varied in such a way that the position of the spot at any point indicates the ground distance between the reflector producing the point image and the orthogonal projection of the flight line on the Earth's surface. From Fig.2.5, a ground range (Sg) can be computed from the slant range (S) and altitude (H) above a flat terrain surface by the simple expression:-

$$Sg = (S^2 - H^2)^{\frac{1}{2}} \qquad 2.1$$

It will also be seen from Fig.2.5 that the SLR is measuring the successive distances OA, OB, OC and OD so that it is the differences in slant range (OB-OA), (OC-OA), (OD-OA) that are being generated in the slant range presentation and not the corresponding ground distances (NB-NA), (NC-NA), (ND-NA). This results in a <u>compression of the scale</u> in the range or cross-track direction in a slant range image. Its correction to a ground range image is achieved by incorporating a suitable time-delay mechanism on the CRT sweep circuitry which converts the travel time differences to give an orthogonal geometry for flat terrain.

In practice, <u>terrain relief</u> will cause deviations from the nominal altitude, and from the simple situation depicted in Fig.2.5. Therefore the ground range image only approximates the true horizontal across-track separation between each ground object and the antenna.

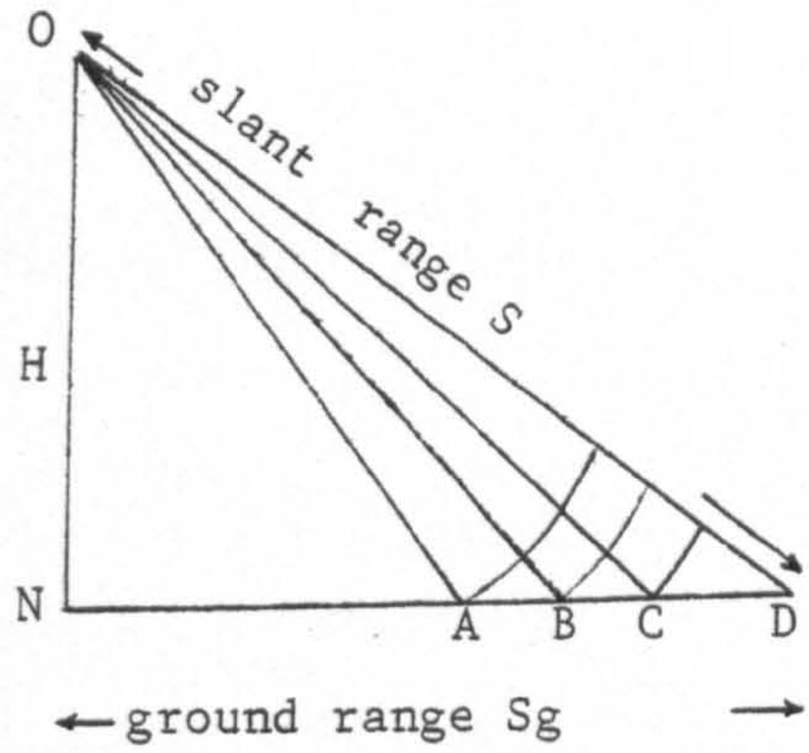


Fig 2.5(a) Relation between slant and ground ranges

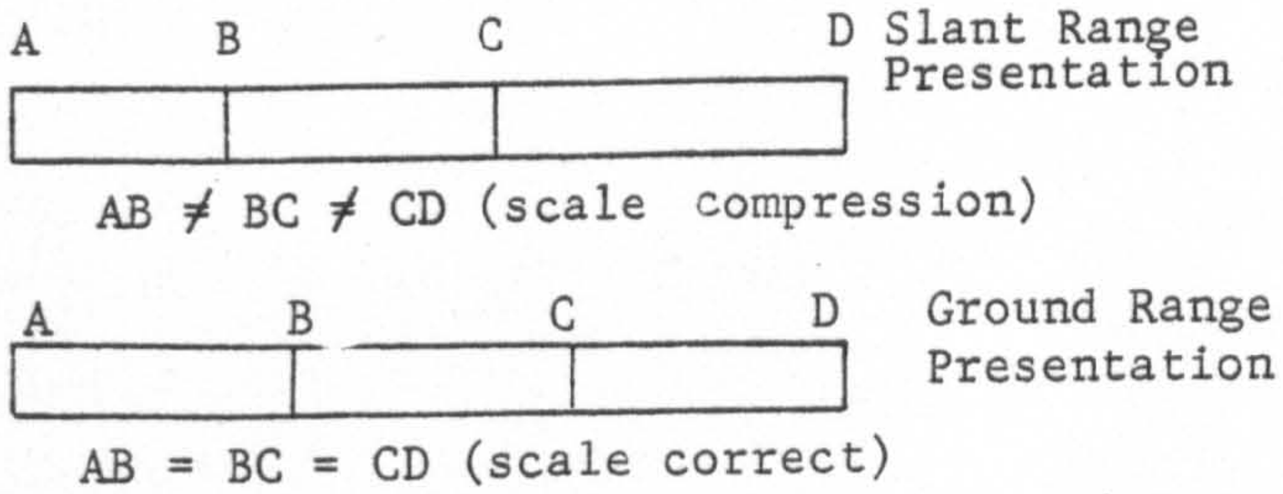


Fig 2.5(b) Scale compression in slant range presentation

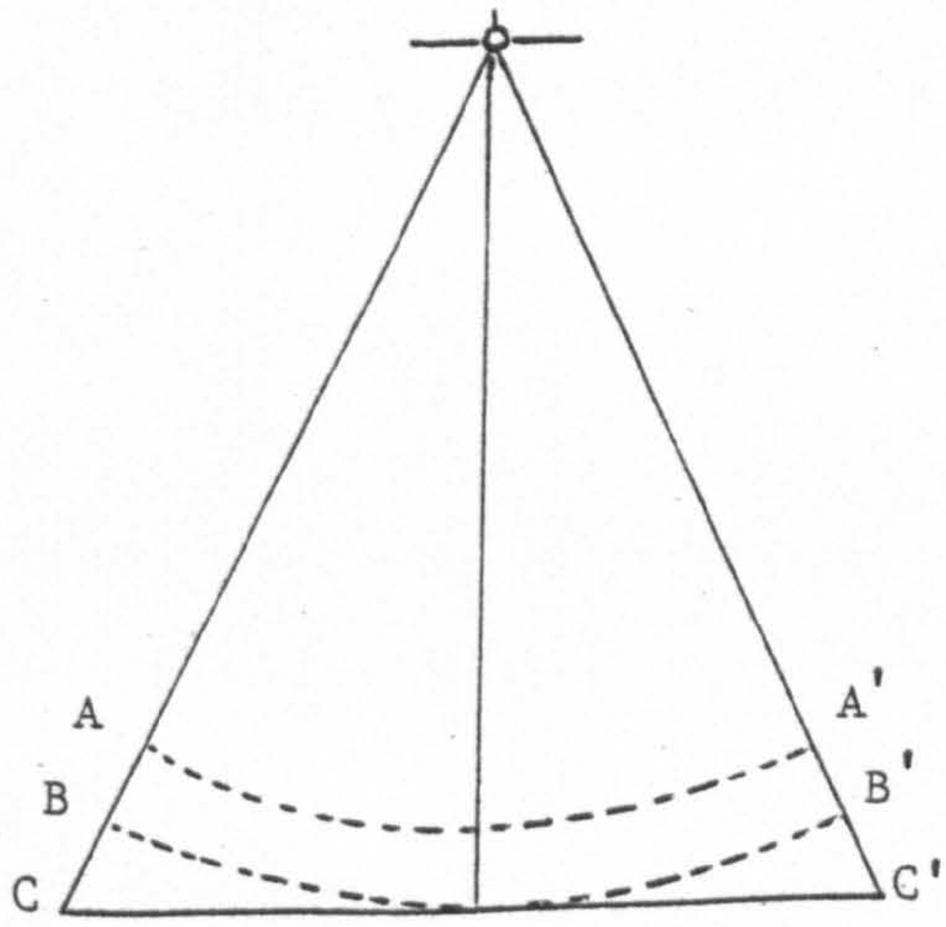


Fig 2.6 Simultaneous returns from either side of aircraft cause ambiguities

The side-looking character of the SLR will be noted and results in the distinctive feature of lack of coverage below the aircraft. This is to ensure that the reflected pulses from the terrain are separated in time and that no ambiguity can result. Thus, if the transmitted pulse was directed downwards, there would be simultaneous returns from objects equidistant on either side of the flight path (e.g. C and C′) and the system would be unable to discriminate between them (Fig.2.6).

2.2.1. Real-aperture Radar Resolution

An important element to consider in any imaging system is its resolution. In the context of side-looking radar, resolution is defined as the minimum distance in the object space at which two objects can be recorded separately on the radar image. Another definition relates to the ability of the radar system to detect and record small objects in the presence of other larger objects.

For flat terrain, the image resolution in the range (or cross-track) direction (Rr) is provided by the duration or length of the pulse (τ) and is theoretically related to half the pulse length ($\tau/2$). The pulse length is normally measured in microseconds (10^{-6} second) and can be converted to length by multiplying it by the velocity of electromagnetic radiation (c). For example, a typical SLR pulse might last 0.1 μs (10^{-7} second) which means that it is $(0.1 \times 10^{-6}) \times (3 \times 10^{8})$ metres = 30 metres long. In theory, this should yield a range resolution (Rr) of 15 metres.

The system range resolution is therefore a linear function of pulse duration τ. Fig. 2.7 shows this linear relation graphically. The actual resolution of the terrain must take account of the elevation angle (θ) (Fig. 2.8). Thus, $Rr = \frac{\tau}{2} \frac{c}{\sin\theta}$. For the example already

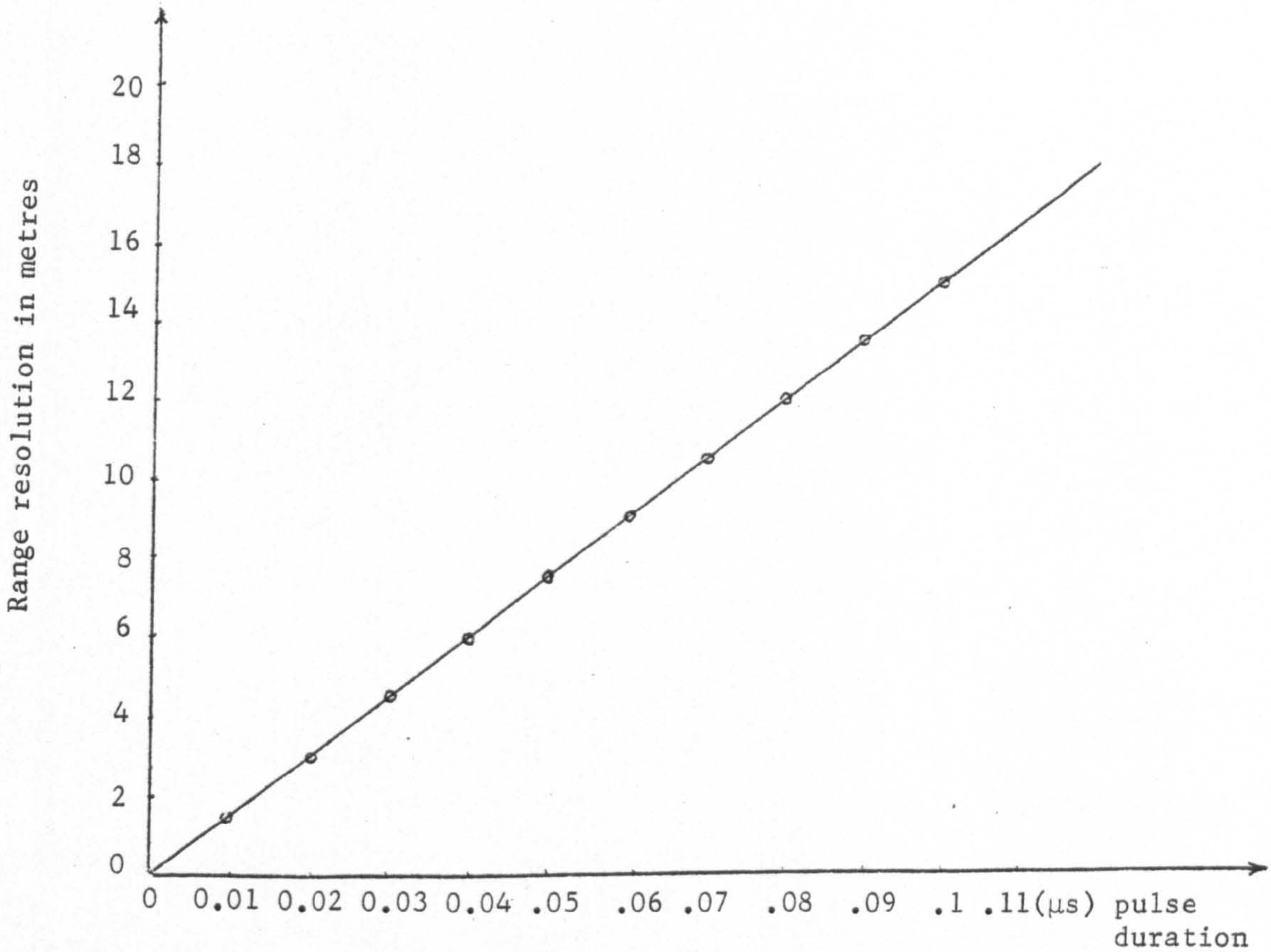


Fig 2.7 Variation of across-track resolution with pulse duration in a noncoherent radar system

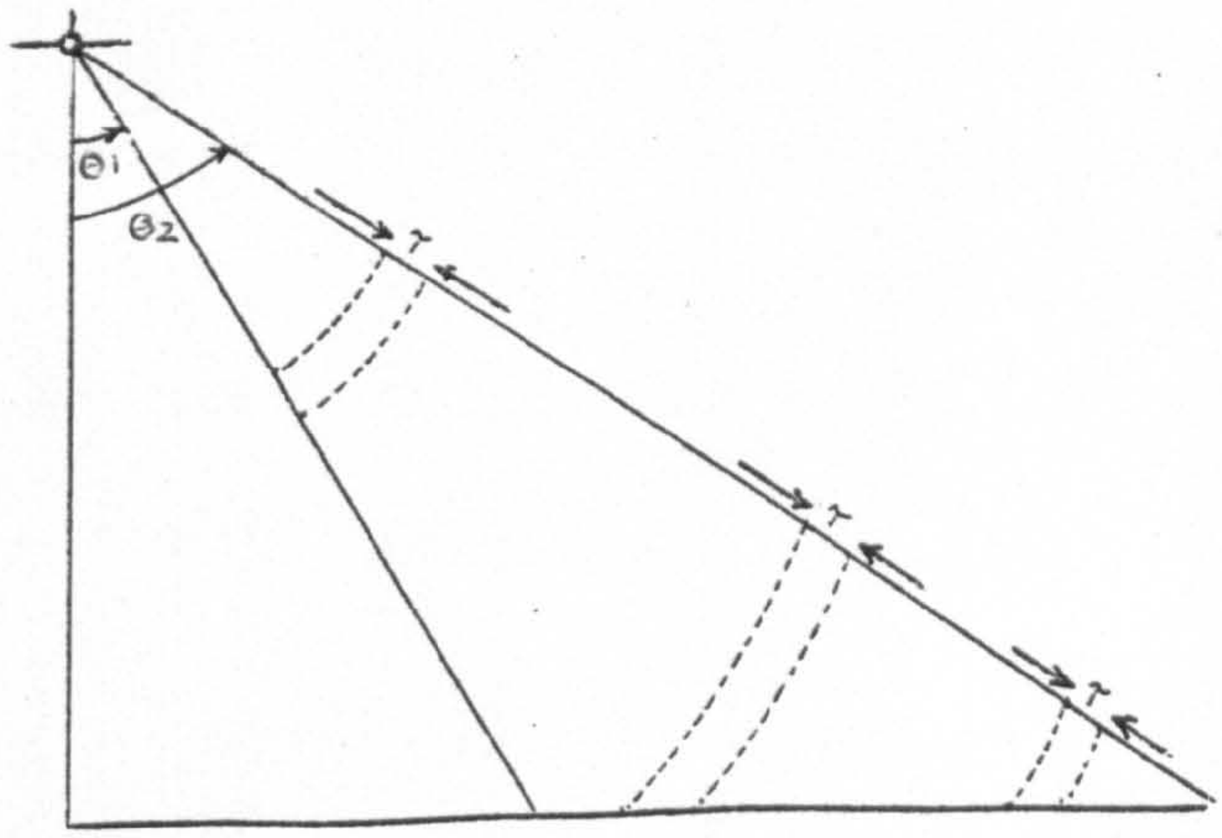


Fig 2.8 Ground range resolution is a function of elevation angle θ

quoted, where the pulse length was 0.1 μs:-

(i) in the near range, $\theta = 40^{\circ}$,

$$\therefore Rr = \frac{(0.1\times10^{-6})\times(3\times10^{8})}{2 \text{ Sin } 40^{\circ}} = {}^{30}/_{2} \text{ Sin } 40^{\circ}$$

$$= {}^{30}/_{1.286} = 23 \text{ m}; \quad \text{and}$$

(ii) in the far range, $\theta = 55^{\circ}$,

$$\therefore Rr = \frac{30}{2 \text{ Sin } 55^{\circ}} = \frac{30}{1.64} = 18 \text{ m}.$$

This has the result that the range or cross-track resolution is actually better in the far range than the near range. Fig.2.9 represents this graphically. To improve the range resolution (Rr), the pulse length must be shortened. However, this will reduce the amount of energy in each pulse transmitted by the antenna so weakening the return signal from the terrain.

In the along-track (or azimuth) direction, the situation regarding resolution is totally different. As can be seen from the diagram (Fig.2.10), the resolution in the along-track direction (Ra) is determined by the angular width of the terrain strip illuminated by the beam. For two targets to be resolved, they must be separated by a distance greater than the beam width at that point on the terrain. Since the beam width will be narrower in the near range than in the far range, so the along-track resolution is higher in the near range than in the far range. Thus, the value of Ra deteriorates with range. An antenna of length D illuminated at a wavelength λ has a far field angular width $\beta \simeq \lambda/D$. The along-track resolution $Ra = k\beta S = k\frac{\lambda}{D}S$ 2.2

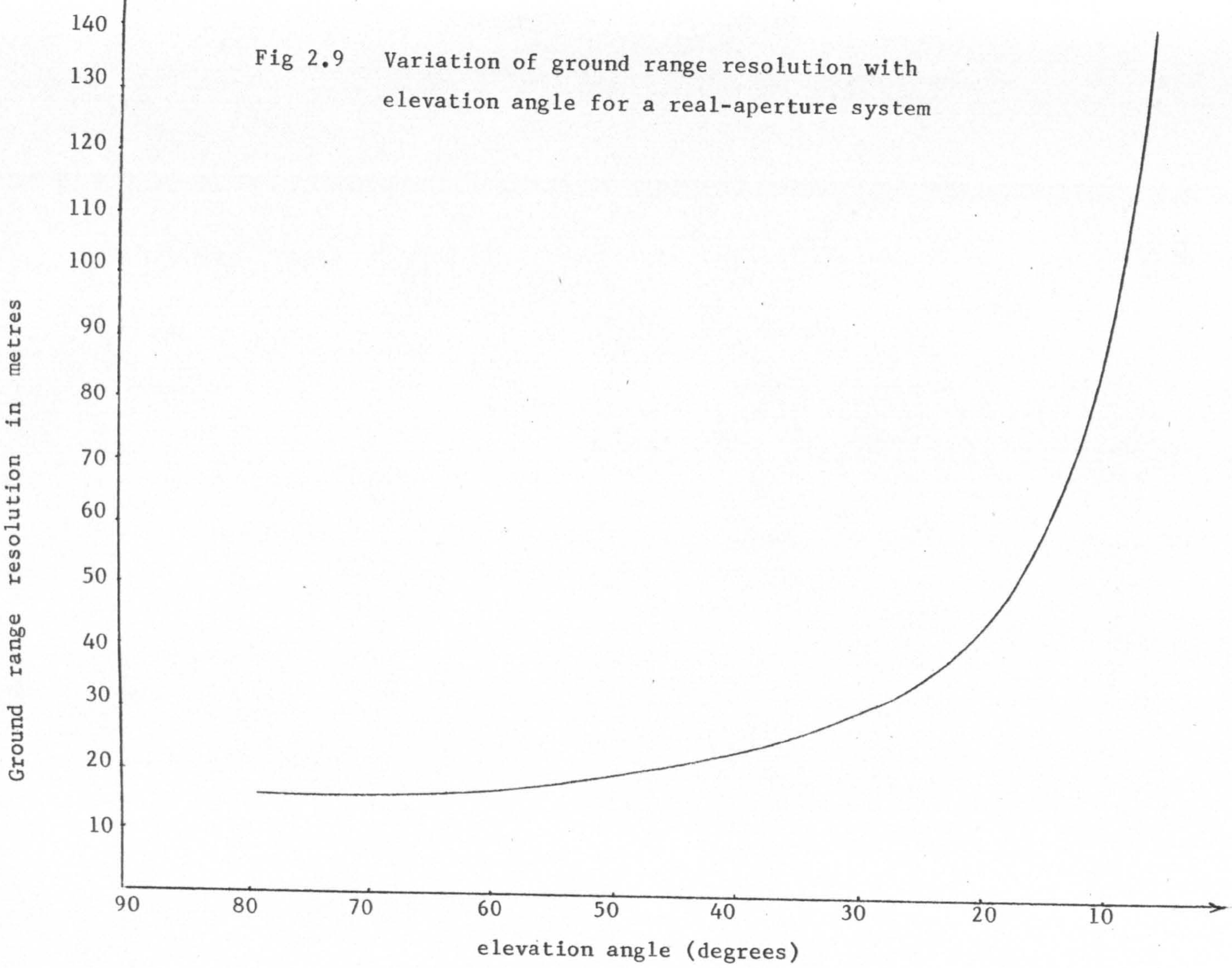
where:

λ = wavelength of the radar transmission;

D = physical (or real) length of antenna;

S = slant range to a reflecting object P; and

Fig 2.9 Variation of ground range resolution with elevation angle for a real-aperture system



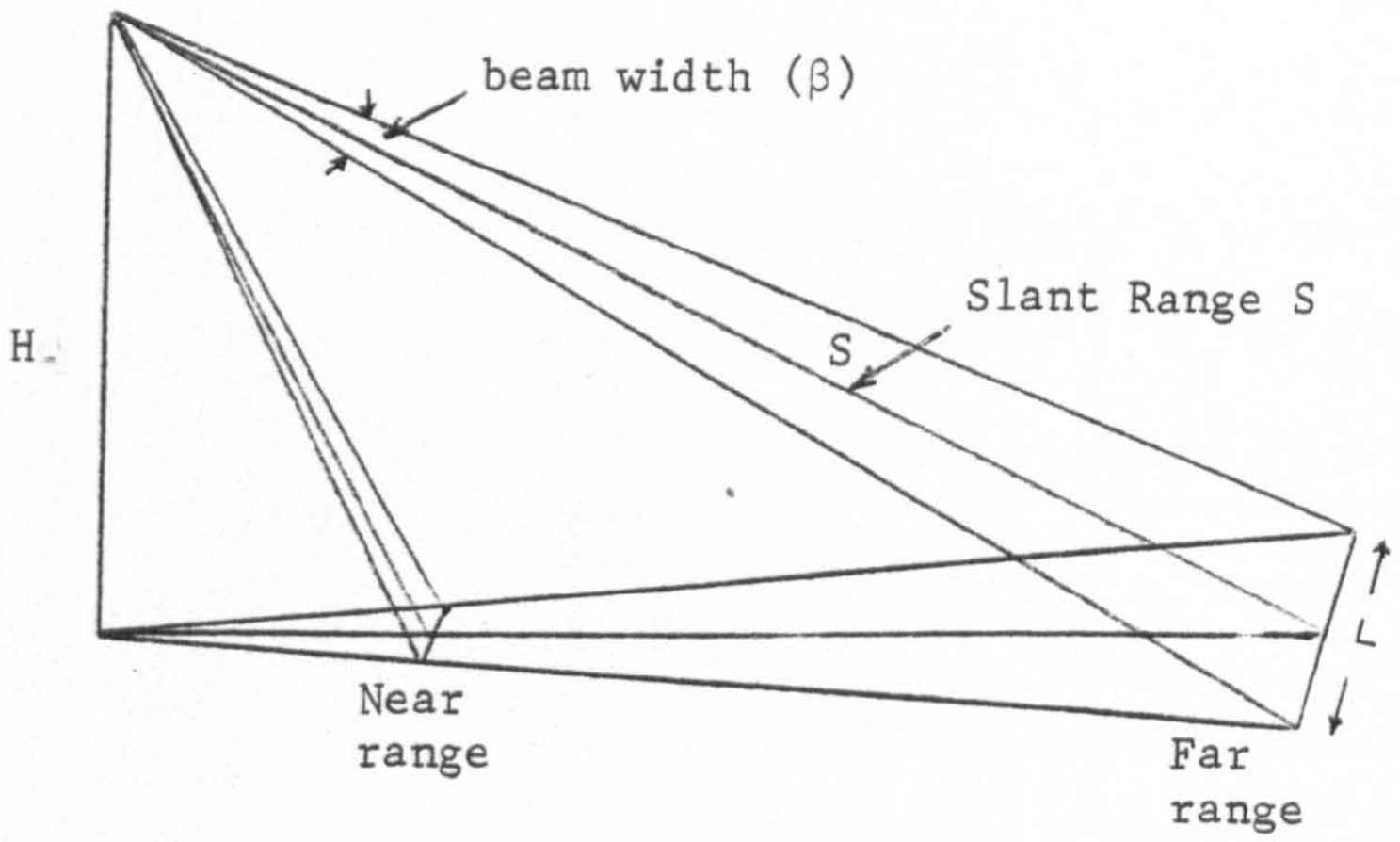


Fig 2.10 Azimuth resolution deteriorates with range

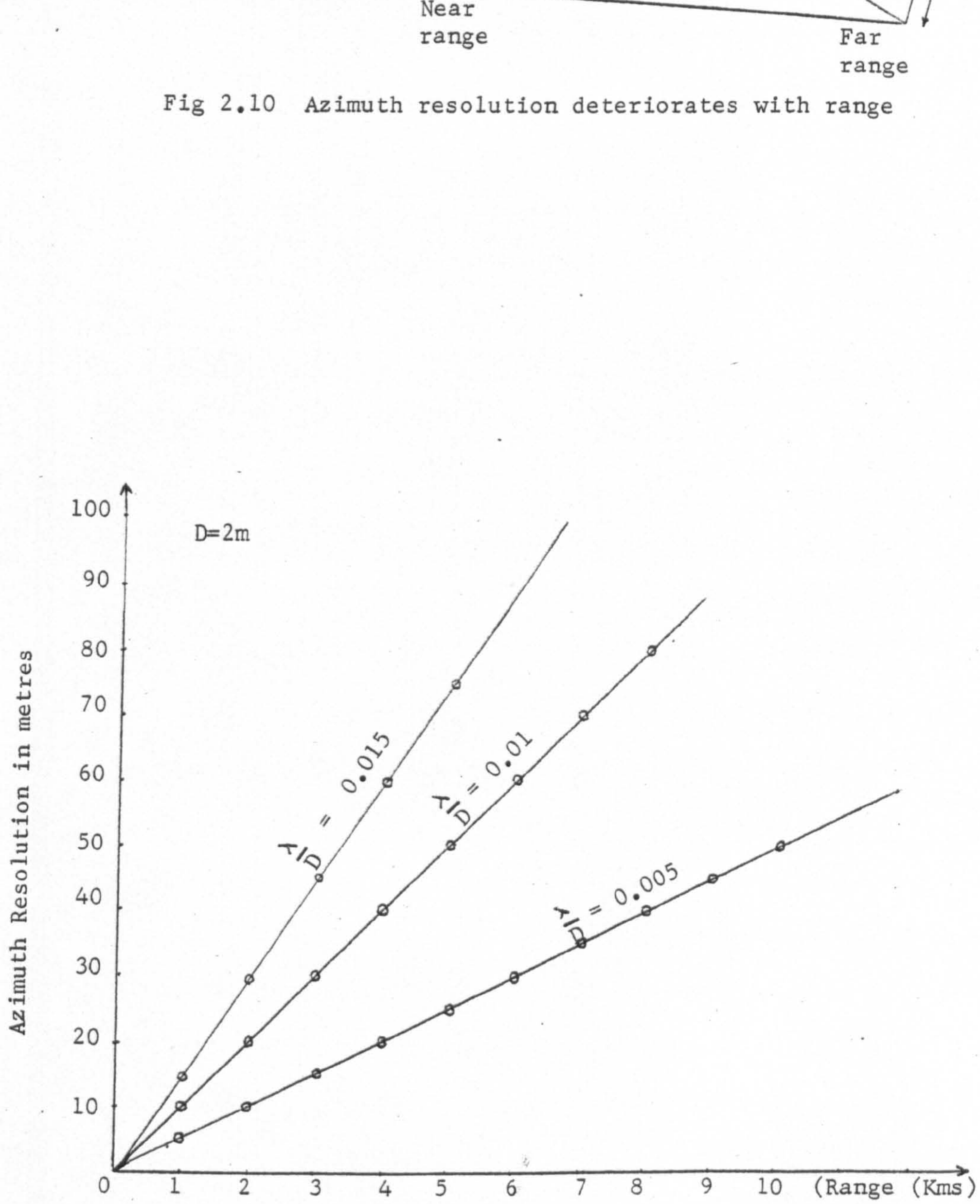


Fig 2.11 Variation of along-track resolution with range in a noncoherent radar system

k = an illumination constant.

Taking typical values for a short-wavelength, high resolution SLR (λ = 0.86 cm), a very long antenna (D = 4.9 m) and an illumination constant (k) of 0.7, then

(i) in the near range (s = 8 km)

$$Ra = 0.7 \times \frac{0.86}{490} \times 8000 = 9.8 \text{ m}; \quad \text{and}$$

(ii) in the far range (s = 20 km)

$$Ra = 0.7 \times \frac{0.86}{490} \times 20000 = 24 \text{ m}$$

It will be noted that this results in the rather contradictory situation that the along-track resolution (Ra) is best at near ranges, whereas the cross-track resolution (Rr) is highest in the far ranges. Fig.2.11 is a graphical representation of equation 2.2 and shows the variation of azimuth resolution with range.

Equation 2.2 shows that Ra gets better with decreasing wavelength of transmission (λ) and/or increasing the size of the antenna aperture (D). Therefore to effect an improvement in the along-track resolution of a real-aperture radar system, it is necessary to:

(i) reduce the wavelength (λ) of the transmitted signal. However, this increases the effect of atmospheric attenuation, and also increases the likelihood of back-scatter due to rain, snow and sleet which is a serious limitation in an all-weather imaging system. A further difficulty is that power requirements also increase with a decrease of wavelength (Deane, 1973).

(ii) The alternative is to increase the antenna aperture size (D). This, certainly, has a limit. For instance at a range of 30 km, using an x-band (λ = 3 cm) radar system with a required resolution of 15 metres and an illumination constant of 1.00, an antenna size of 60 metres is needed (Ra = 1 x $\frac{3 \text{ cm}}{6000\text{cm}}$ x 30,000 m =

15 m). Obviously, this is physically too large to be carried on any aircraft.

From what has been discussed above, the limitations inherent in real-aperture side-looking radar systems are obvious. The basic conclusion is that, if a usefully high resolution is to be obtained, then the real-aperture SLR must be operated at low altitudes and at short ranges. For many airborne applications both in the military and in the civilian fields, this is not too much of a drawback and, as a result, the real-aperture type of SLR has found widespread application. In the mapping context where high resolution is usually a paramount requirement, the situation can be very different and the characteristics of real-aperture radar are unfavourable in this respect, as they are too in high-altitude long-range reconnaissance work. It was these factors which lèd to the development of the alternative synthetic-aperture type of side-looking radar.

In the context of the present thesis which is concerned with spaceborne applications of side-looking radar, it is obvious that the altitudes and ranges (200 to 1,000 km) are such that real-aperture radar has no part to play whatsoever.

2.3. The Principle of Synthetic Aperture Radar (SAR)

The originator of synthetic-aperture side-looking radar appears to have been Wiley of the Goodyear Aircraft Corporation (now the Goodyear Aerospace Corporation) in the early fifties (Sherwin et al, 1962; Brown and Porcello, 1969). Wiley noticed that, when an object was linearly swathed by a microwave beam, a one-to-one correspondence existed between its along-track coordinate and the instantaneous Doppler shift of the signal reflected to the antenna by that object. Wiley concluded that a frequency analysis of the reflected signals would make

it possible to obtain finer resolution than that allowed by the along-track width of the real antenna.

2.3.1. The Basic Concept of SAR

Basically, a synthetic aperture radar (SAR) system differs from a real aperture radar in its synthesis and utilisation of an effectively long antenna using a method of signal processing of the reflected microwave pulses to achieve fine azimuth resolution rather than using a physically long antenna.

The main objective of a SAR system is the improvement of azimuth resolution. This necessitates recording not only the reflected microwave energy, but the phase of these reflected signals as well. This results in a form of radar signal which, if converted to an image in its raw form, is virtually uninterpretable. Thus, the signals received by the SAR require special optical and/or digital processing techniques to transform this data record into an image that can be interpreted and used in different ways.

As is the case with the real-aperture system, the antenna transmits and receives microwave pulses but of a coherent nature. Each pulse produces a series of reflections from terrain objects in a direction transverse to the direction of the motion of the platform. Each transmitted and then reflected signal is compared with a coherent reference wave derived from a stable oscillator.

A much simplified graphical view of this basic concept is provided in Fig.2.12:-

(i) In Fig.2.12a, a terrain object enters the radar beam and reflects part of the incident radiation back to the antenna. In this particular example, the object is eleven wavelengths away from the aircraft at this point. The received signal is then combined

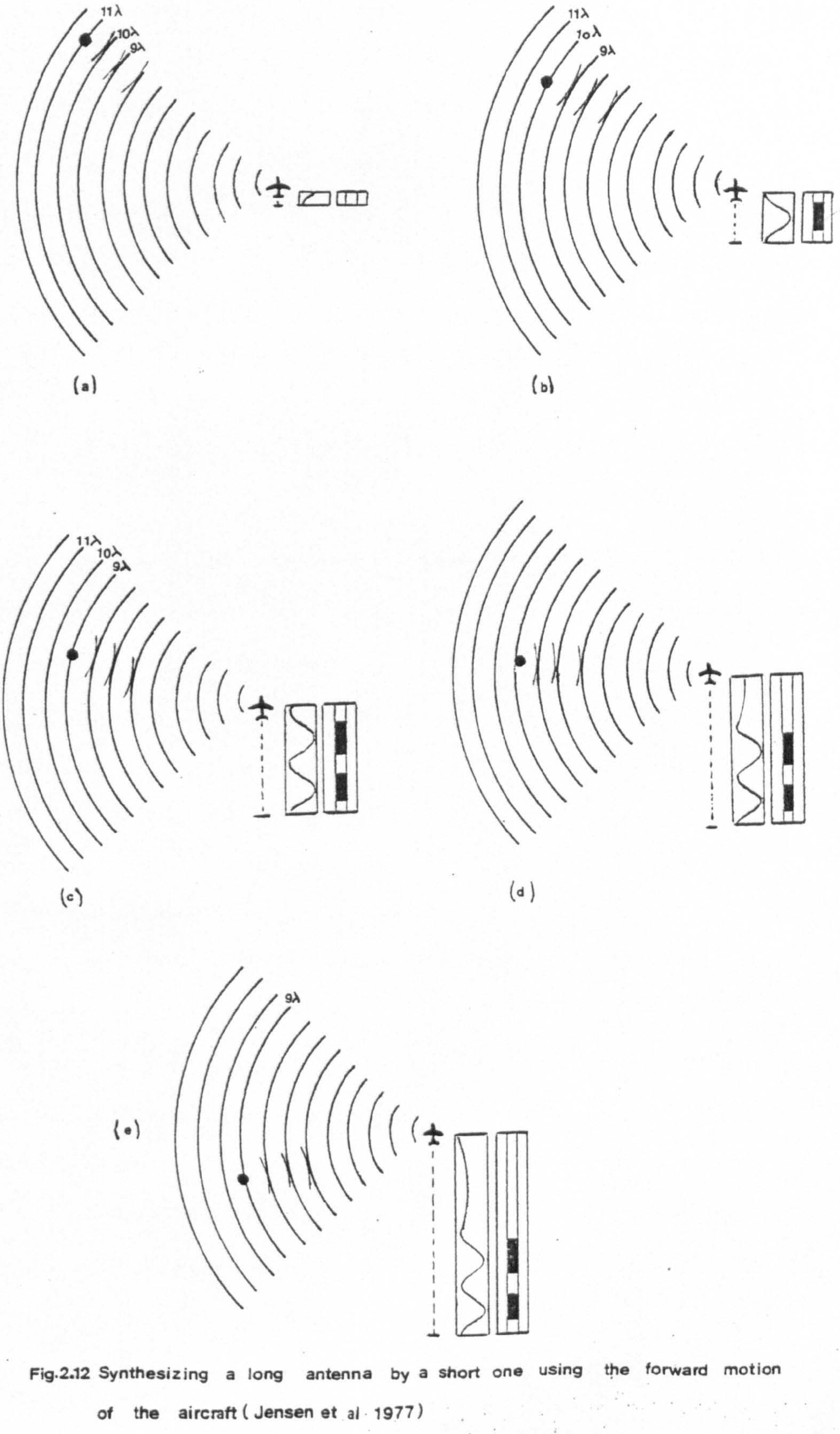


Fig.2.12 Synthesizing a long antenna by a short one using the forward motion of the aircraft (Jensen et al 1977)

electrically with the reference wave so that the two interfere. The actual interference pattern produced will depend on the phase comparison between the two signals. Thus, when the phase angle of the reflected signal coincides with that of the reference wave, the interference is constructive. Obviously when it is 180^{o} out of phase, it is destructive.

(ii) As the aircraft flies forward (Fig.2.12b), the object will reach a position where the object is ten wavelengths away. Once again, when an integral number of wavelengths is reached, the received and the reference signals are again in phase and they again interfere constructively. This will happen again at nine wavelengths as shown in Fig.2.12c.

The interference signal can be interpreted as a voltage which can be used to control the brightness of the flying spot scanning across the face of a CRT. When the interference is constructive, the voltage is high and the spot is at its brightest. Equally, when the two signals reach their maximum out-of-phase position (180^{o} apart) they interfere destructively so that the voltage is at its minimum and the spot at its dimmest.

(iii) In Fig.2.12d, it will be seen that the object is now at its closest position to the aircraft ($8\frac{1}{2}$ wavelengths away) at which position it is at right angles to the antenna. From then on, the aircraft is again steadily increasing its range from the object and will again pass through a series of positions where it is an integer number of wavelengths away from the object (Fig.2.12e).

The record of the brightness of the CRT spot can be recorded on a film strip in somewhat the same manner as with the real-aperture SLR with the forward movement of the film being related to the aircraft

velocity. While the series of diagrams illustrate the process for a single object, it will be realized that signals are being received back from the multiplicity of objects present on the terrain. All of these are being compared with the reference signal to create not a single interference pattern but a whole series of these patterns superimposed on one another. Thus what is being created on a film record are these sets of interference patterns instead of the direct image of the terrain which is generated by a real-aperture SLR system. It is these interference patterns which need further processing to recover the actual terrain image.

2.3.2. Analogy between SAR Records and Optical Holograms

Still taking this much simplified view of the synthetic-aperture radar process, it will be seen that a close analogy exists between the SAR record, albeit produced at microwavelengths, and the hologram produced optically at visible wavelengths. In holography, the light waves from a coherent source of light (usually a laser) illuminate an object. Each point on this object reflects back light waves which are made to interfere with a reference beam, so producing the multitude of circular interference patterns which can be recorded photographically on film to produce the familiar hologram. If this photographic film record is later illuminated by a beam of coherent light, the multitude of interference patterns reconstructs the original set of reflected waves to form a three-dimensional image of the object.

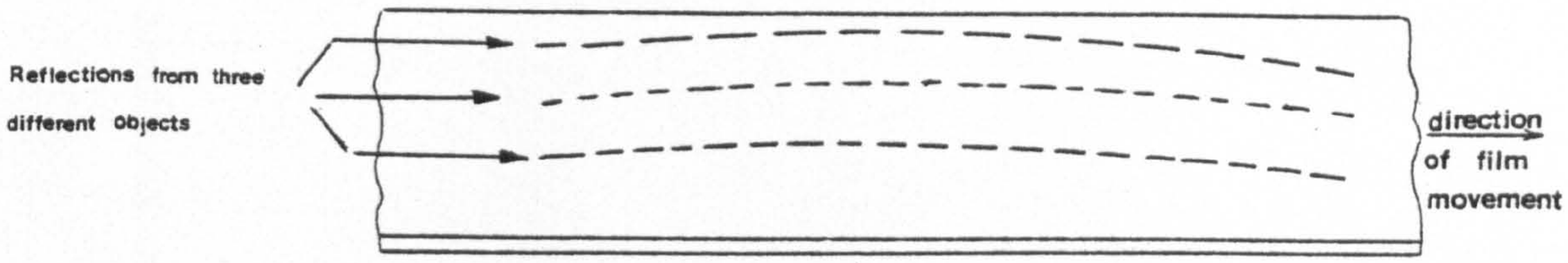
Something of the same process is taking place with synthetic aperture radar. Coherent microwaves illuminate the terrain, the receiver collecting the reflected waves and combining them with a reference set of waves to produce sets of interference patterns which are recorded on the signal film via the CRT. This can be regarded as a form of radar hologram.

However, the analogy is not a complete one. Because of the system characteristics, each object's reflections are recorded along a single line on the signal film instead of over the complete photographic format as in a true hologram. Typically (Fig.2.13a) this will appear as a series of broken lines parallel to the edge of the film. This can be regarded as corresponding to a cross-section through the circular type of interference pattern. This leads to its description as a one-dimensional interference pattern or hologram by Cutrona et al (1960, 1966), Leith (1971), Larsen et al (1972),etc. . A typical example of such a hologram is given in Fig.2.13b. In this way, the range of the object is recorded as the distance of the signal record from the near edge of the signal film. However, the transformation of the complete signal record requires further processing.

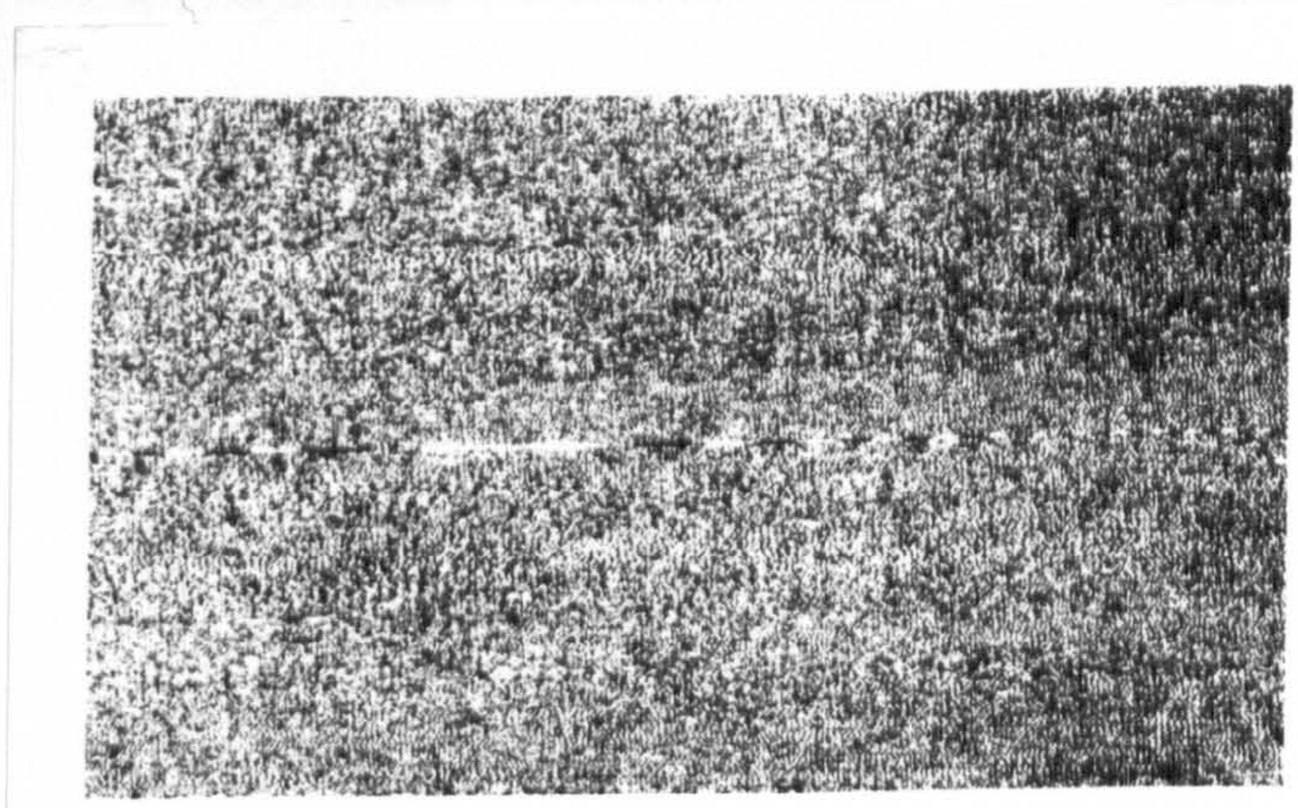
In this highly simplified introduction (a more detailed description of the actual process will be given later) it will be seen (Fig.2.14) that, if the signal film is illuminated by a coherent light source, then the reconstruction of an image of the terrain will take place behind the film. For the example given above, there will be a single point position (Fig.2.14a) where all the recorded interference patterns interfere constructively as follows:-

- at a distance 11 wavelengths from the corresponding point on the record (a);
- at a distance 10 wavelengths from the corresponding point on the record (b);
- at a distance 9 wavelengths from the corresponding point on the record (c);
- and so on.

At this point therefore, light from the whole length of the interference pattern is combined and focussed to form an image of the original

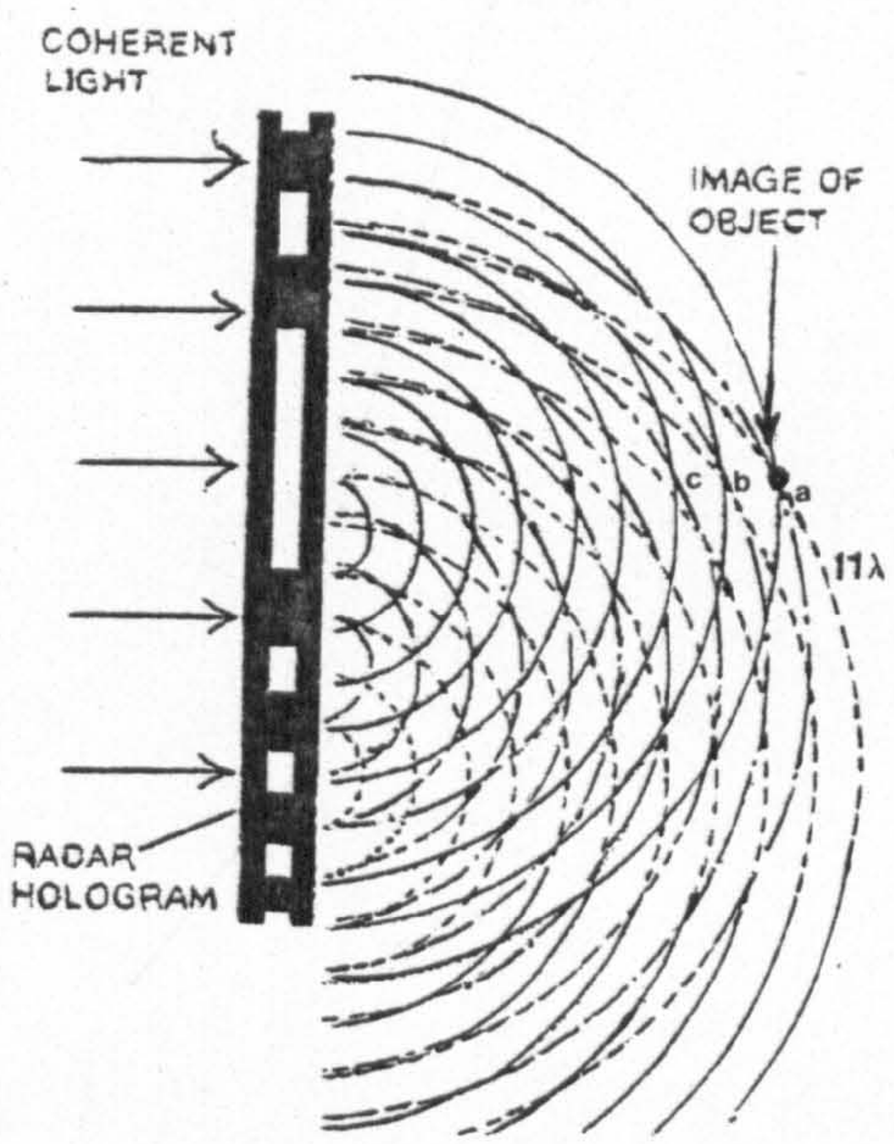


(a) Objects' reflections are recorded along a single line on the signal film.

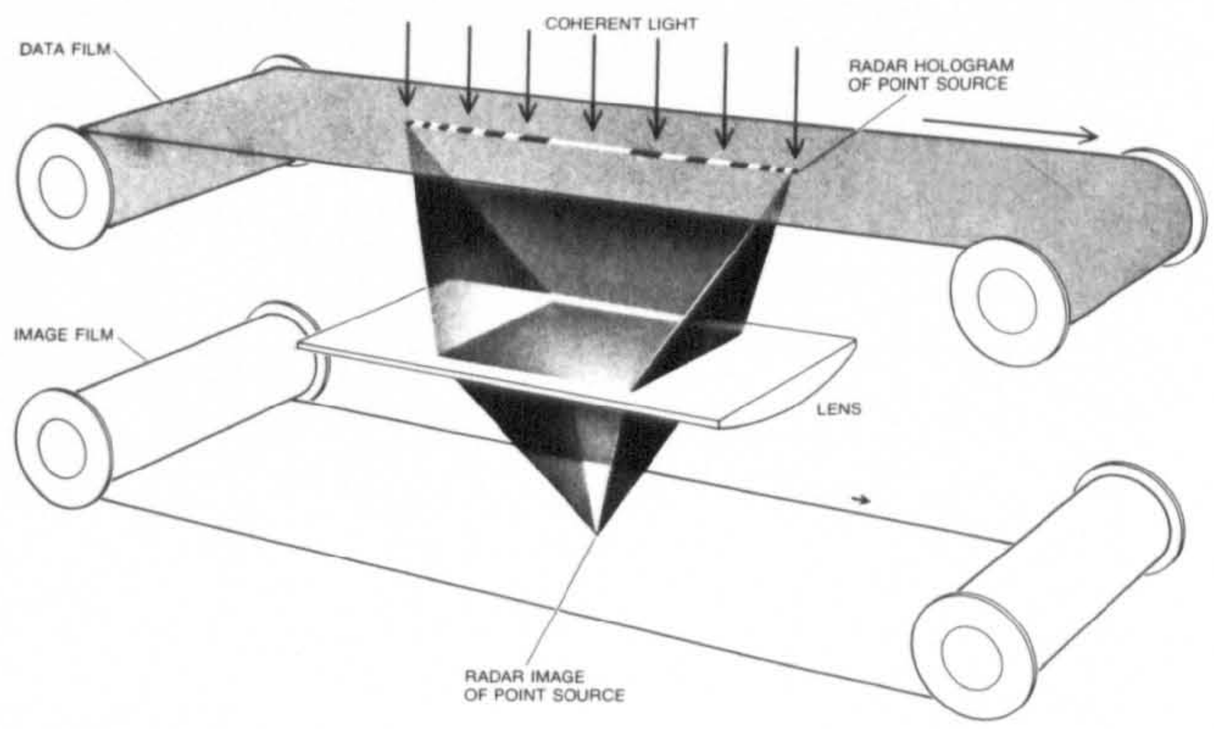


(b) Section of SAR data film showing the interference patterns produced by the returning trains of microwaves.

Fig 2.13 Radar Hologram



(a)



(b)

Fig 2.14 Reconstruction of object's image by illuminating the signal film by a coherent light source.

object (Fig.2.14b). Using optical processing techniques, this process can be carried out simultaneously for all the objects which have been recorded on the signal film.

As already mentioned, the holographic record exists only in the along-track direction of the film, the signals having already been located in the cross-track direction according to their position - corresponding to range - across the film. This leads, as will be seen in detail later, to the need for specialized optical components, e.g. conical and cylindrical lenses, to allow the focussing of both the along-track and cross-track records to form the image of the object.

2.4. Synthetic-aperture Radar Resolution

Since the above discussion has centred on the generation of the synthetic aperture and the processing of the data in the along-track direction in particular, the resolution in this direction will be discussed first. The expression beam sharpening is frequently used for the increased resolution provided by the use of a synthetic aperture. Because of the very long lengths of synthetic antennae which can be achieved, it is not possible to ignore the curvature of the wavefront or phase fronts of the received signals. Without focussing of the antenna, it would be impossible to achieve the theoretical angular resolution of a SAR. Since the curvature of the phase front varies with range, the data processing must include a focussing function that varies with range. The processing of the synthetic aperture is achieved through the insertion of a phase shift which will vary according to the position along the synthetic aperture. This eliminates the effects of the curvature of the reflected wavefront. This focussing operation being range dependent complicates the data processing somewhat.

The fullest degree of resolution is achieved by processing

all the Doppler band-width and by compensating for the phase differences on a pulse-to-pulse basis. The processing operation therefore represents a phase adjustment of the received signals applied to each element of the array so that the contributions from all array elements can be combined and used in the final processing. The image resolution resulting from the application of this preliminary correction is termed the <u>focussed SAR azimuth resolution</u>. As with the real-aperture case

$$\beta = \lambda/D \quad \text{(see Fig.2.15a)}$$

where β = beamwidth;

λ = wavelength and

D = actual antenna length.

Also Ls (the synthetic aperture) = $S\beta$

where S = slant range.

The angular resolution after processing

$$\beta s = \frac{\lambda}{2Ls} \qquad 2.3$$

The corresponding linear resolution $Ra = \frac{\lambda S}{2Ls}$

$$= \frac{\lambda S}{2S\beta}$$

$$= \frac{\lambda}{2(\lambda/D)} = \frac{D}{2} \qquad 2.4$$

It will be seen that the focussed azimuth or along-track resolution Ra is independent of the transmitted frequency or wavelength (λ) and of the range (S) and is wholly a function of the physical aperture (D). This is of course of prime importance when considering satellite-borne SARs where inevitably the ranges used are extremely long. In theory, therefore, the resolution resulting from the use of a 1 m antenna would be 0.5 m irrespective of whether it is mounted in an aircraft or a spacecraft. In practice, however, the questions of operating range, power requirements, etc. will loom large and so the selection of the

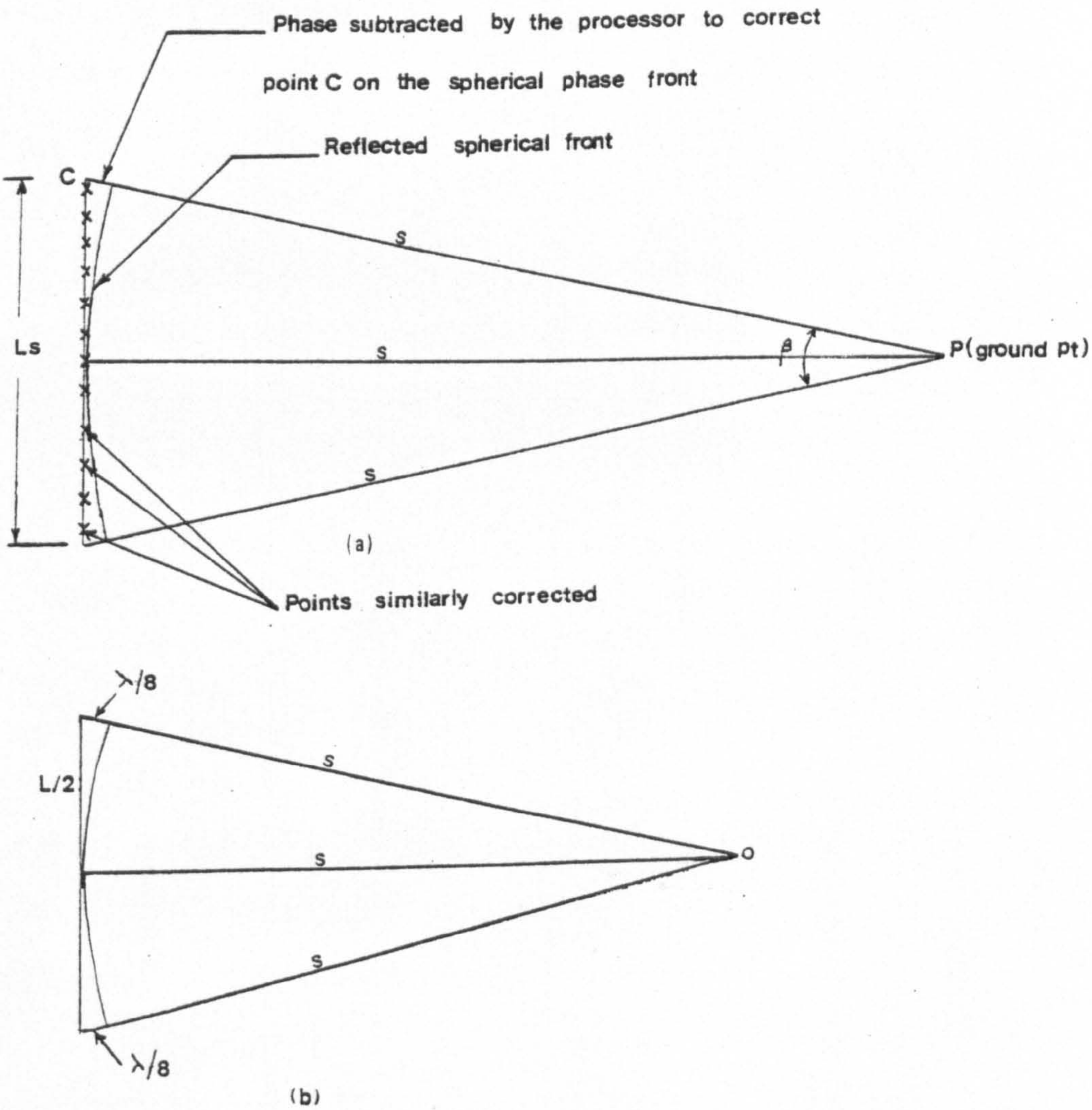


Fig 2.15 The geometry of the focussed and unfocussed SAR systems.

actual antenna size used is the result of trade-offs against these other system parameters.

If all the phase variations resulting from curvature of the wavefront are not compensated for in the Doppler record, the case of the so-called unfocussed synthetic antenna will be realized. The lack of phase adjustment results in a maximum length to the synthetic aperture which can be generated and used. This occurs for a given range when the round-trip distance from an object to the centre of the synthetic aperture differs by $\lambda/4$ from the round-trip distance between the object and the extremities of the synthetic aperture (Cutrona and Hall, 1962). In turn, this results in a lower level of azimuth or along-track resolution. From Fig.2.15b:-

$$(S + \lambda/8)^2 = S^2 + (L/2)^2 \quad \text{(by Pythagoras)}$$

$$\therefore\ S^2 + \frac{\lambda S}{4} + \frac{\lambda^2}{64} = S^2 + L^2/4$$

The term $\lambda^2/64$ will be very small and can be ignored.

$$\therefore\quad \frac{L^2}{4} \approx \frac{\lambda S}{4}$$

$$\therefore\quad L \approx \sqrt{\lambda S}$$

$$Ra = \frac{\lambda S}{2L}$$

$$= \frac{\lambda S}{2\sqrt{\lambda S}} = \tfrac{1}{2}\sqrt{\lambda S} \qquad 2.5$$

This shows that the unfocussed azimuth or along-track resolution (Ra) is a function of range (S) and wavelength (λ) and is independent of the physical length of the aperture or antenna (D).

It is important to realize that the azimuth resolution (Ra) of even an unfocussed SAR system will be a considerable improvement on that achievable on a comparable real-aperture radar. Since the much improved azimuth resolution (Ra) meets the needs of many users, unfocussed

SAR systems have a widespread use in airborne applications since the resulting processing is less complex. However, for spaceborne applications, the resolution of the images produced from unfocussed SAR systems would fall far short of the needs of users and a fully focussed SAR system is quite essential for such missions (Fig.2.16).

Another important point is that the along-track resolution (Ra) is constant for all ranges. From Fig.2.17, it will be seen that the greater the range of the object, the longer it remains in the beam of the antenna. Thus for an individual terrain feature, the synthetic aperture will be longer if it is located at a greater distance (position Q) from the platform on which the SAR is mounted and shorter if it is closer (position P) to the platform. In this way, the effective length of the synthetic aperture is directly proportional to the range at which the object is located. The equation $Ra = \frac{\lambda}{2} \frac{S}{Ls}$ 2.6 shows that the along-track resolution (Ra) is determined by the values of S (range) and Ls (synthetic aperture). Therefore, if the two are related in direct proportion to one another, the term $\frac{S}{Ls}$ becomes a constant. Thus the resolution of the image in the along-track direction is the same at all ranges and high resolution is possible at long ranges (Fig.2.17).

To achieve an increase in resolution in the range direction (Rr) comparable to that produced in the azimuth direction by the synthetic aperture either a high powered pulse of very short duration must be used or pulses of a longer duration and moderate power are employed which have frequency modulation applied to them. In the latter technique, the transmitted pulse (τ) incorporates a linear change in frequency e.g. from f_1 and f_2, over the pulse length (τ) (Fig.2.18). From this frequency coding of the transmitted pulse combined with proper phase matching in the correlator i.e. the use of a matched filter, a pulse

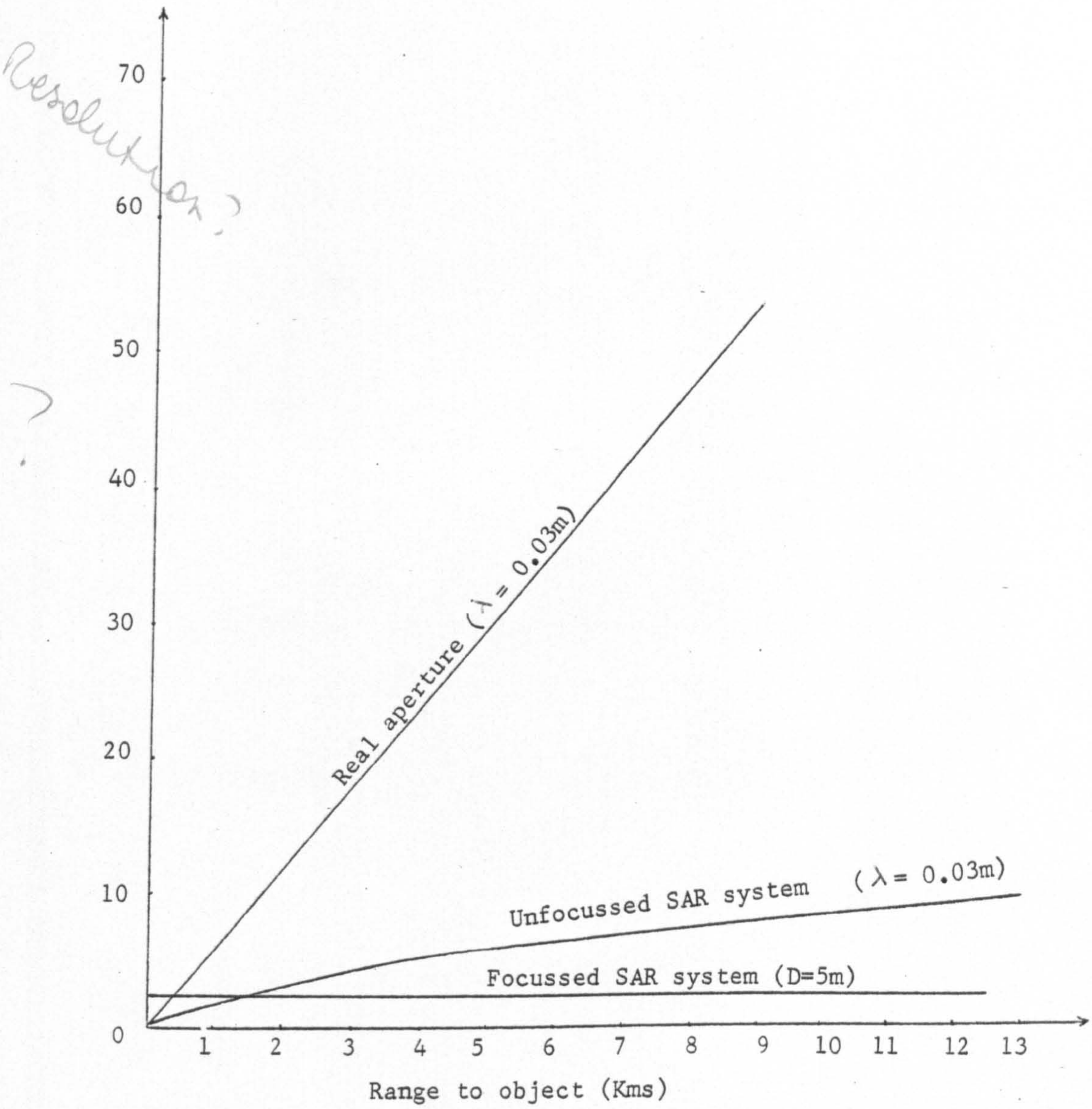
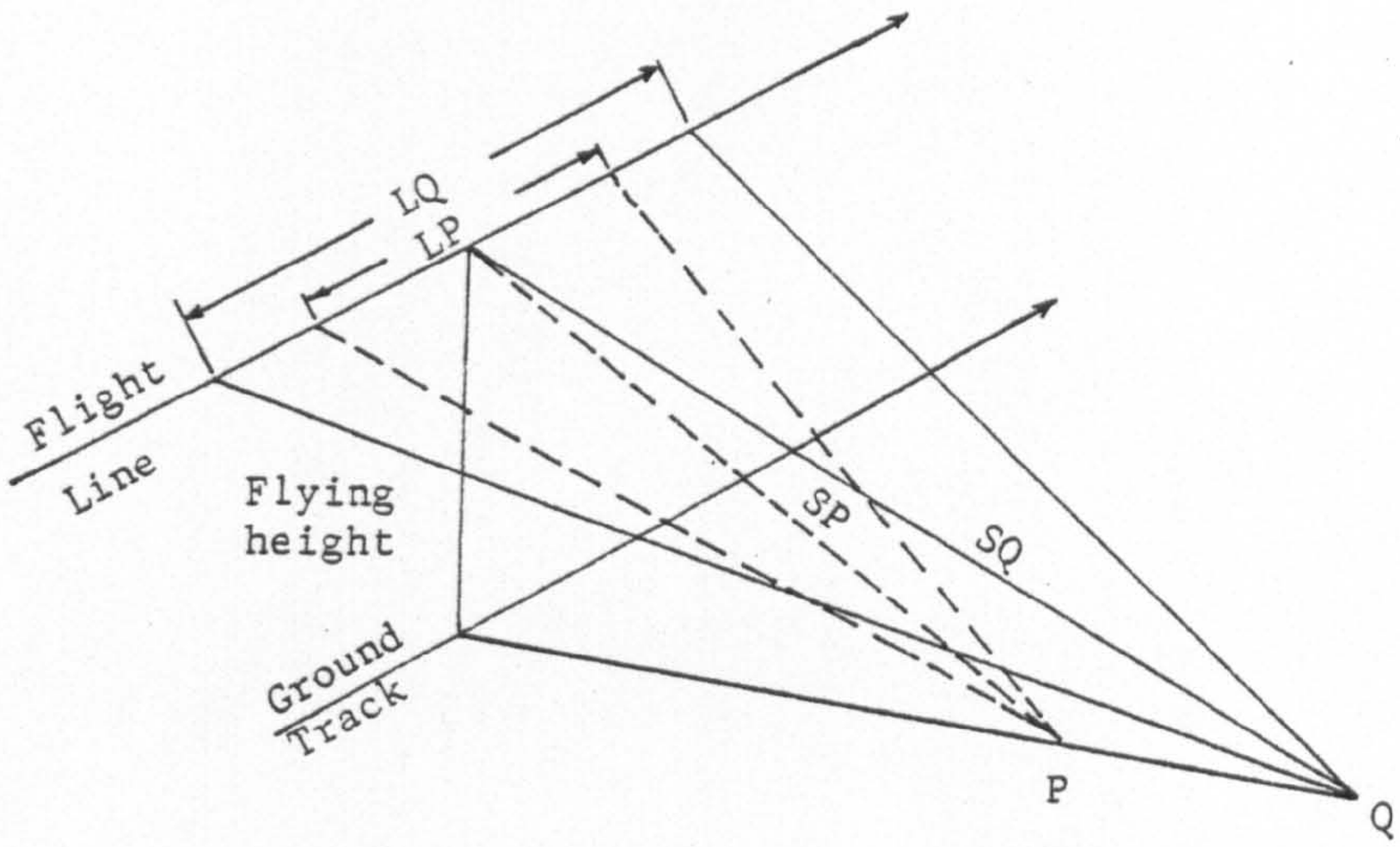
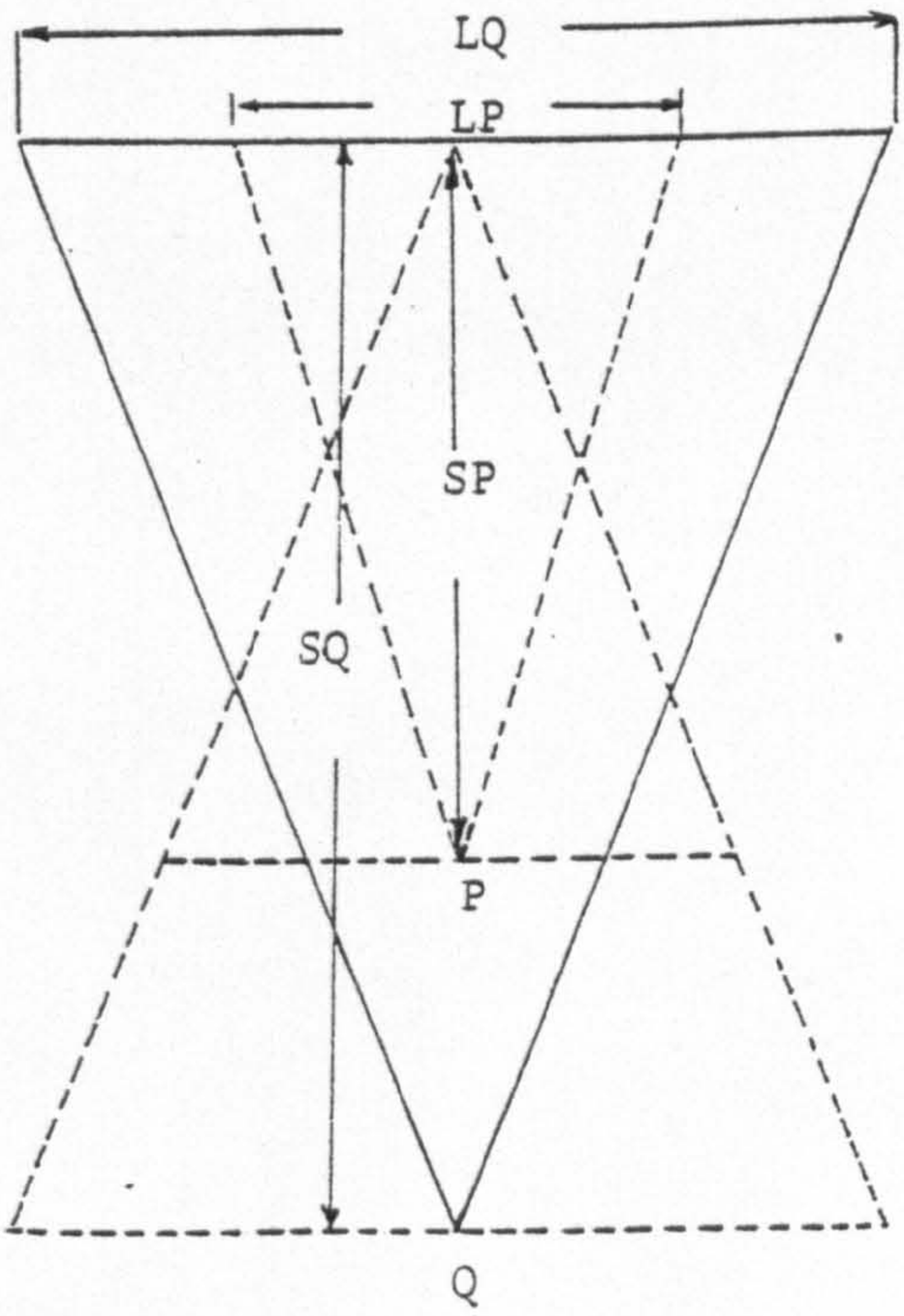


Fig 2.16 Dependence of system azimuth resolution (Ra) on range

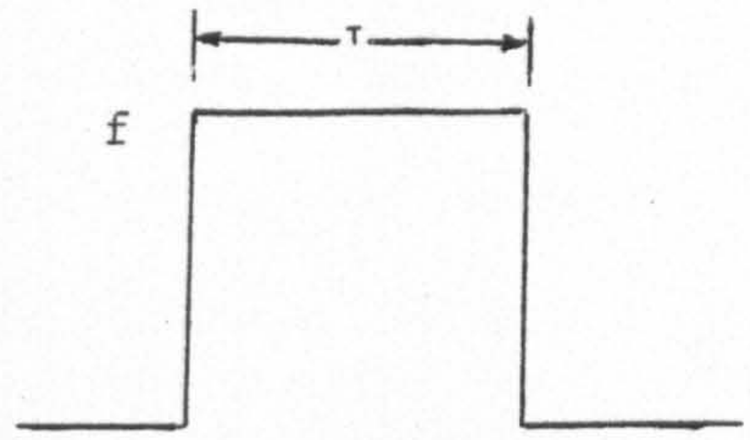


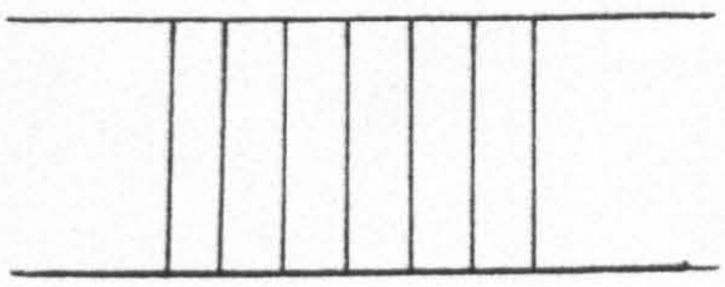
(a) Three-dimensional view



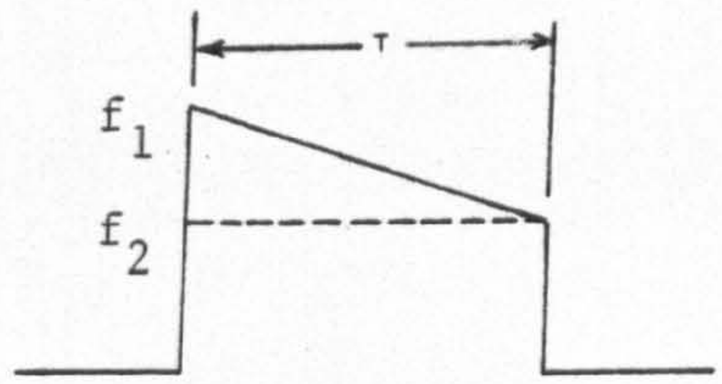
(b) View in slant range plane

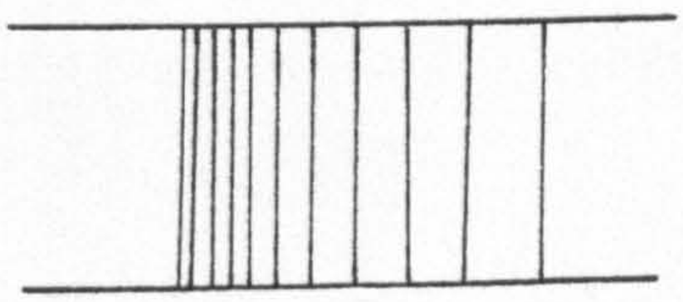
Fig 2.17 The synthetic aperture length is directly proportional to the range to an object.





(a) Constant carrier frequency





(b) Linear frequency modulation of carrier frequency

Fig 2.18 The transmitted pulse (τ) incorporates a linear change in frequency to achieve a fine range resolution.

compression is achieved which gives increased resolution in the range direction. Radars of this type are known as "chirped radars".

The compression time width $\tau = \frac{1}{\Delta f}$ where Δf is the frequency bandwidth $(f_1 - f_2)$. Multiplying both sides of this equation by $\frac{c}{2}$ where c is the velocity of microwave energy and the factor of 2 is required for the two-way range distance,

$$\frac{c\tau}{2} = \frac{c}{2\Delta f}$$

As before, $\frac{c\tau}{2}$ will be the range resolution (Rr)

$$\therefore Rr = \frac{c}{2\Delta f} \qquad 2.7$$

Once again, this matter is of great importance when attempting to achieve high resolution imagery from a spaceborne SAR.

2.5. Doppler View Point of SAR

For the purpose of an introduction to aperture synthesis and to SAR systems and processing, the much simplified holographic point of view followed above is quite appropriate. However, other quite different viewpoints can be taken and alternative explanations made of these processes which will be more appropriate in other circumstances. These include the highly mathematical analyses which appear in the literature concerned with electronics and electrical engineering. One aspect of this approach which will be helpful in the explanations of SAR processing which will follow later in this thesis is to view the matter of the received signals in terms of the Doppler shift in frequency experienced after reflection of the emitted pulse by terrain objects (Fig.2.19).

Fig.2.20 shows a block diagram of a coherent radar system. A stable oscillator feeds its signal to the transmitter using the carrier frequency (fc). After reflection from the terrain it will be

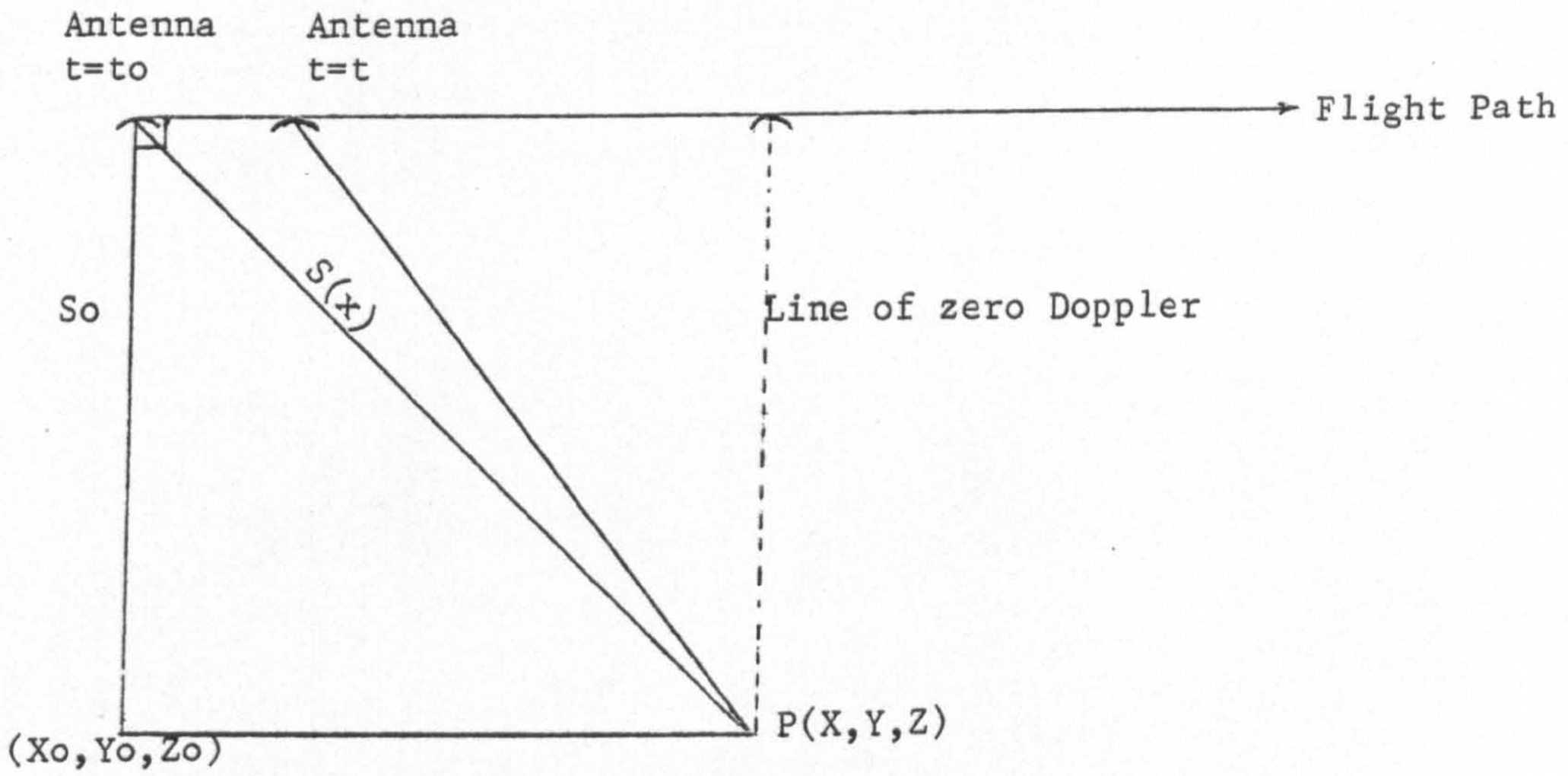


Fig 2.19 Doppler shift of object point p

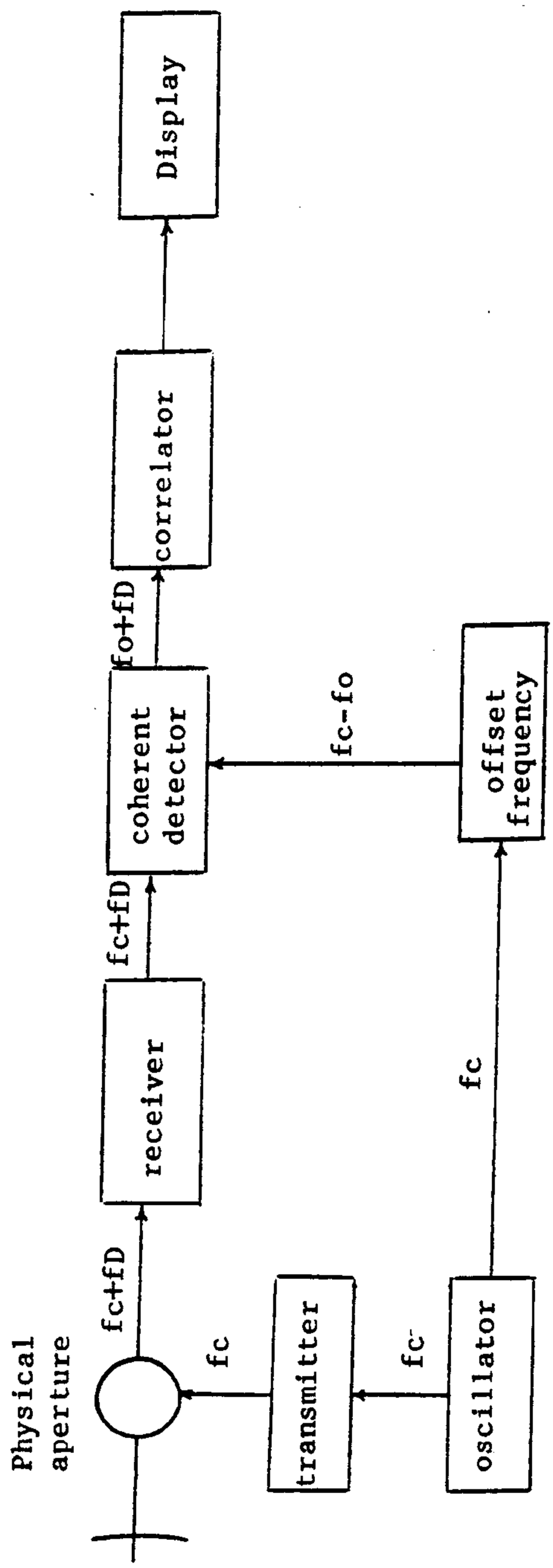


Fig 2.20 Doppler view point of SAR (Kovaly, 1976)

received back with a shift in frequency i.e. as $(f_c + f_D) = f_r$ (the received frequency). This is passed to a comparator or coherent detector where a comparison takes place between the received frequency $(f_c + f_D)$ and the original carrier frequency (f_c) from which an offset frequency (f_o) has been subtracted to avoid frequency fold-over (Kovaly, 1976). The two frequencies are subtracted so that the output from the coherent detector (f_{CD}) is as follows:-

$$f_{CD} = (f_c + f_D) - (f_c - f_o)$$
$$= f_D + f_o \qquad 2.8$$

The Doppler frequency f_D is normally defined by the equation:-

$$f_D = -(2\frac{ds}{dt})/\lambda \qquad \text{(Skolnik, 1970; Kovaly, 1976)}$$

where s = instantaneous slant range to the point P at time t.

Thus the slant range to P is time varying and the resulting Doppler shift in frequency is equal to the time varying range multiplied by the factor $2/\lambda$. From Fig.2.19, an expression can be derived for the Doppler frequency f_D as follows:-

$$S(x) = \sqrt{S_o^2 + (x - xo)^2}$$
$$\simeq S_o + \frac{1}{2}\frac{(x - xo)^2}{S_o}$$

But $(x - xo) = v(t - to)$ where v is the velocity of the vehicle.

$$\therefore S(t) \simeq S_o + \frac{v^2(t-to)^2}{2So} \qquad 2.9$$

$$\therefore \frac{ds(t)}{dt} \simeq \frac{2v^2(t-to)}{2So} = \frac{v^2(t-to)}{So}$$

$$\therefore f_D = \frac{-2v^2(t-to)}{\lambda So}$$

substituting in equation 2.8 above

$$f_{CD} = f_o - \frac{2v^2(t-to)}{\lambda So} \qquad 2.10$$

The output from the coherent detector (f_{CD}) will be based on the set of returns from the object P and is referred to as the "Doppler frequency history" of point P. This frequency history for each object is generated from the time that the object enters the radar beam to the time that it leaves the beam. The output frequency (f_{CD}) is therefore a function of time. Each object has the same Doppler history which goes from some high frequency to some low frequency (Kovaly, 1976). The azimuth bandwidth will be $(2v^2 / \lambda So)T$ where T is the time during which the object P stays within the beam.

Similar concepts may also be used to describe the behaviour of the transmitted and received pulses in the range direction. Conventional SAR devices generate a pulse τ with an amplitude A. As already discussed above, if average power requirements are critical as in the case of spaceborne radars, then the microwave energy of the radar pulse must be transmitted with a linear frequency modulation (FM) sweep of the carrier frequency i.e. from frequency f_1 to frequency f_2 over the pulse length τ . Mathematically, the linear FM transmitted pulse has the form

$$f(t) = \cos(2\pi f_c t - \frac{2\pi\alpha t2}{2}) \qquad - \frac{\tau}{2} \leq t \leq \frac{\tau}{2}$$

where f_c = carrier frequency;

α = rate of frequency sweep $\frac{(\Delta f)R}{\tau}$;

$(\Delta f)R$ = sweep spectrum bandwidth; and

τ = uncompressed pulsewidth.

The detected pulse frequency is given by:-

$f_{CD} = (fo - \alpha t)$, αt being the bandwidth of the time interval τ (Skolnik, 1970; Kovaly, 1976).

As previously discussed, the Doppler frequency histories (resulting from the generation of the synthetic aperture) and the pulse

compression (resulting from the use of FM transmitted pulses) are two independent operations which can be combined just as they were in the holographic viewpoint. The output from the coherent detector is a series of pulses (each lasting a few microseconds) and a bandwidth (which is of the order of MHZ). Thus the pulses are the backscattered signals picked up by the receiver and stored in the correlator over a period of time (which is of the order of seconds). Over this period of time, a bandwidth is generated.

The two bandwidths are generated by the linear frequency modulation over time intervals τ and T respectively. At this stage an optical correlator can focus this information in each dimension to create the final image. In the range direction, the resolution (Rr) is inversely proportional to the RF bandwidth, i.e. $Rr = \frac{c}{2(\Delta f)R}$. In the azimuth direction, the resolution (Ra) is inversely proportional to the Doppler bandwidth, i.e. $Ra = \frac{W}{(\Delta f)a}$. Fundamentally therefore in this concept, the required resolution is obtained in both directions by generating a suitable bandwidth. It will be noticed that each of the two expressions for resolution has a numerator related to velocity - the velocity of microwave energy (c) in the range direction and the velocity of the aircraft or spacecraft (V) in the azimuth direction.

CHAPTER III

SYNTHETIC APERTURE RADAR DATA PROCESSING

3.1. Possible Approaches to Processing of SAR Imagery

A real-aperture or brute-force radar system has the merit of producing a direct image of the terrain on a CRT display or on a photographic film to give a permanent record. The image is directly interpretable and the only corrections required are those arising from geometric considerations or due to radiometric biases.

A synthetic-aperture radar (SAR) record, on the other hand, is not directly interpretable and special processing techniques are required to reconstruct the image so that it can be viewed by those wishing to make use of it. Therefore, some more detailed discussion of the different approaches that have been used so far to achieve the transformation of the data recorded from both airborne and spaceborne SAR systems appears appropriate.

Essentially, there are four main ways of processing the recorded SAR returns as follows:-

(i) The data may be recorded as electronic analogue data and processed using hard-wired electronic elements.

(ii) The data may be recorded in photographic form and then processed by optical means (optical processing).

(iii) The data may be recorded in digital form which is then processed in a digital computer (digital processing).

(iv) A combination of optical and digital processing may be carried out (hybrid technique).

Approaches (ii) and (iii) are those which have been used most widely in

actual practice. They will therefore be discussed in some detail in the account which follows.

3.2. Initial Studies and Developments

In 1954, the theoretical studies mentioned earlier in Section 2.3 which indicated the possibility of synthesizing the multiple reflections produced by objects in the terrain from the microwave pulses generated by a short physical antenna with a broad beam, were the basis of research carried out by several groups who tried to convert them into practice. The crux of the whole problem was to solve the difficulties posed by the data processing of the multitude of signal returns. Three early major efforts were made by:- (i) the Goodyear Corporation;

(ii) a group in the University of Illinois; and

(iii) a group in the University of Michigan.

The Goodyear approach placed an emphasis on the development of a range-gated filter bank. However, this was discovered to be impractical since it could not incorporate the required focussing operation and so the idea was completely abandoned (Brown, 1967).

The University of Illinois group pursued three different approaches, all based on purely electronic methods of data processing and again limited to the unfocussed case of SAR.

The University of Michigan group suggested the possibility of optical processing of the synthetic-aperture radar (SAR) data with the use of photographic film both as the intermediate storage medium of the received radar signals and for the final output. This suggestion resulted from a comprehensive theoretical study of the parameters of synthetic-aperture radar systems. These included the achievable

lengths of synthetic antennae; the curvature of the phase front of received signals; the achievable theoretical angular resolution with focussed and unfocussed antennae; and the requirement for a focussing function that would vary with range.

Once it was realized that all of these varied demands could be satisfied with a suitably designed optical processor, it was soon devised by Cutrona. This had the special property that it permitted the generation of a synthetic-aperture image which was properly focussed at all slant ranges. The processor was demonstrated experimentally in 1957 and fully reported in the open literature in 1960 (Cutrona et al, 1960).

It is particularly interesting to notice that digital processing techniques for generating images from synthetic-aperture data were also considered even at this very early time (1954). However, they were thought to be unpromising because of the enormously large amounts of data which had to be handled and the further complication caused by the difference in focussing functions for each segment of a radar range.

An early prototype of the University of Michigan optical data processor was carried on-board a C-46F aircraft of the U.S. Air Force with the object of achieving in-flight processing and display of the terrain image. It was soon realized that the weight, bulk and need for a stable vibration-free platform for the processor made this impractical. The result was that the optical processing had to be carried out on the ground as a post-flight operation, which meant that, unlike real-aperture radar, no real-time in-flight display of the image was possible. A whole generation of ground-based optical data processors came into use to process images obtained by the AN/UDP-1 and

L-23D radars (Cutrona et al, 1961). Early images produced by these devices can be found in Sherwin et al (1962), Cutrona et al (1960) and Cutrona et al (1961).

The first optical processor devised by Cutrona was designed on the basis of being a multi-channel cross-correlator. It was viewed as performing cross-correlation of the radar returns against a reference function having the form of the expected return from a point target. This is very much a viewpoint based on communication theory. This concept was adopted because of its multi-channel capability which is needed when processing an extended range interval in which a separate function is required for each resolvable range element.

Immediately after the development of this processor, it was recognized that the processing of synthetic antenna radar returns carried out with the coherent optical system could equally well be interpreted as a process of "wavefront reconstruction" (i.e. akin to holography). This holographic viewpoint was a more flexible one than the cross-correlation viewpoint (Leith and Ingalls, 1968). On this basis, forms of optical data processing came into being utilizing either special cylindrical and conical lenses or tilted signal and recording planes (Kozma et al, 1972). These new concepts offered significant improvements in versatility and flexibility, with improved image quality.

3.3. Optical Processing of SAR Data

With optical processing, the SAR returns are normally recorded on photographic film, usually known as the signal film or data film. This is an analogue record of the returns recorded photographically from the face of a CRT display. The record produced on the data film is a quite unintelligible set of interference

patterns. It is these which require further processing so that an image is generated that can be interpreted by human beings.

Basically, coherent light is used to illuminate the data film and the projected image is then passed through a series of specially designed and constructed lenses which effect the necessary optical processing required to produce the final image (see Fig.2.14). This final transformed image is then recorded on photographic film (the image film) in negative form. From this, further positive copies may be made by contact or by projection onto film or paper which can then be used by the image interpreter.

3.3.1. A Basic Optical Processor

A more detailed diagram of an optical processor is shown in Fig.3.1. Mathematically speaking, this processor operates on the following basis:-

$$k(y) = \int q_1(\alpha,y).q_2(\alpha,y)d\alpha, \qquad 3.1$$

where α is a variable of integration;

y is a variable which takes on a range of discrete values, each corresponding to an independent data channel; and the functions

q_1 and q_2 are light amplitude transmission functions which are incorporated as transparencies placed in planes P1 and P3 respectively.

The processing operation will go through the following procedures:-

(i) The recorded signals on the data film are passed through the aperture plane P1 where they are illuminated by the coherent collimated beam from a laser. This projects the signals through the lens L1 into the plane P2. If the two planes P1 and P2 are placed equidistant from lens P1 at the focal length distance (f1) of the lens i.e. at the

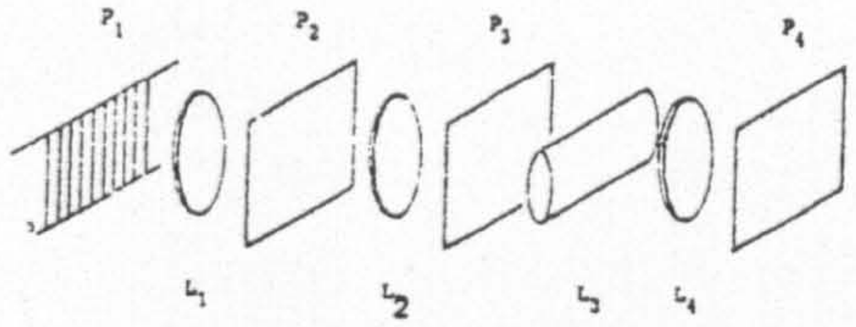


Fig 3.1 A basic cross-correlation system for processing radar data

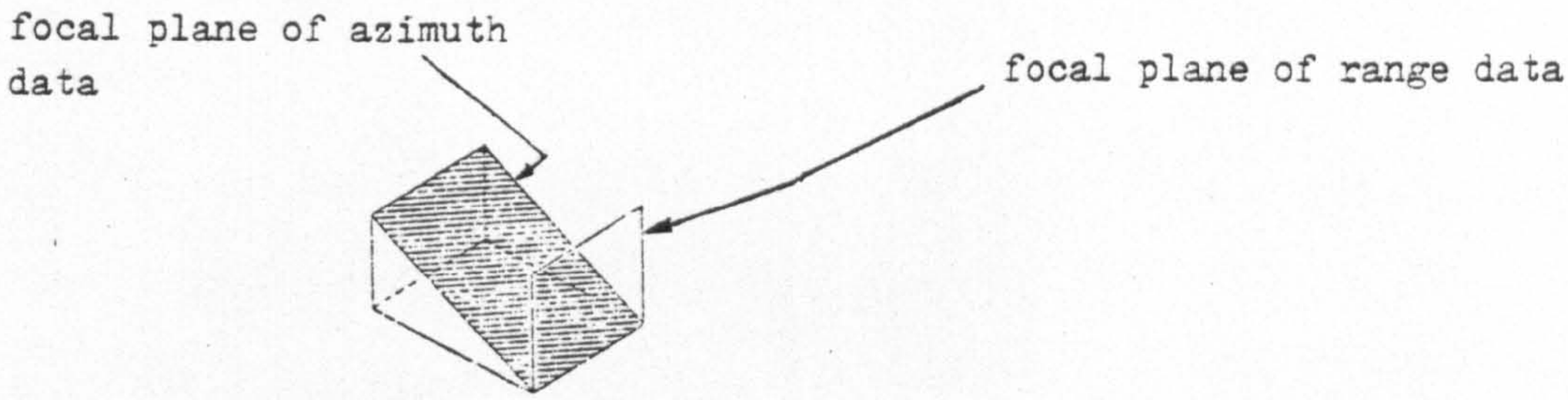


Fig 3.2 Range and azimuth focus planes geometry after reconstruction.

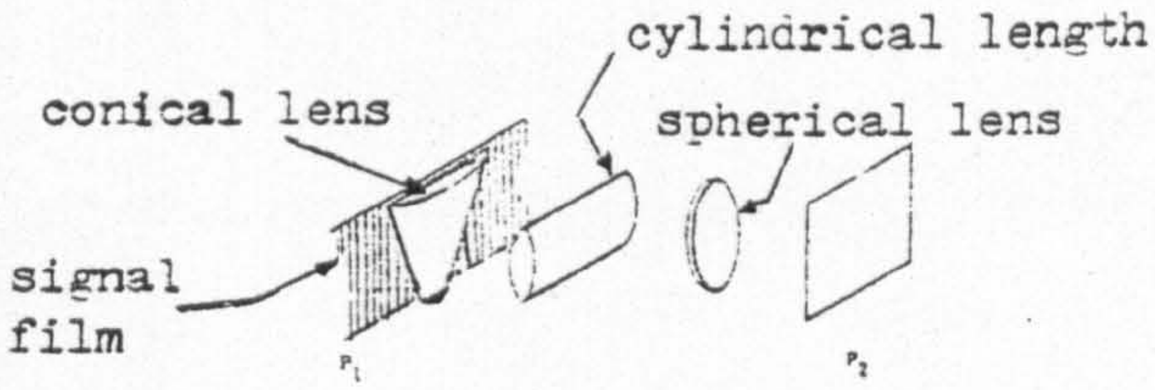


Fig 3.3 An optical processor using a conical lens

front and back focal lengths of the lens, then a property of such a coherently illuminated optical system is that the light distribution in the P2 plane is the Fourier Transform of the signals contained in the P1 plane. If required, an amplitude weighting can be accomplished by inserting a shaded transparency with uniform phase thickness adjacent to the film. Also phase shifting can be carried out by inserting suitably-shaped pieces of glass adjacent to the data film.

(ii) A second similar transformation takes place between planes P2 and P3 using lens L2 which reverses the coordinates of the original amplitude distribution.

(iii) In plane P3, a mask or transparency is placed which contains a transmittance factor representing a matched filter for the signal in the range direction. A narrow slit is a suitable approximation to the required matching filter. The slit passes the optical frequency corresponding to the target, other optical frequencies corresponding to other targets being rejected. As the signal record is moved through the plane P1, different target positions assume in turn the proper frequency for passage by the filter. This image is being projected by the lens L2 into the plane P3. It is of course possible to insert shaded transparencies or optical phase shifters instead of the slit at plane P2 to filter and modify the image in other ways.

(iv) A further transformation (equivalent to the required integration process) is affected between planes P3 and P4 through the combined effects of a cylindrical lens (L3) and a spherical lens (L4). The use of the cylindrical lens ensures that the transformation of the projected image takes place in only one dimension. The final processing is carried out using the spherical lens (L4).

(v) As the signal film continuously moves through the input aperture (P1) a second film moving continuously in synchronisation records the illumination through a slit located in the output plane (P4) (i.e. along the line x=o). The result of this transformation is a fine-resolution image produced from the uninterpretable raw data film.

3.3.2. Optical Processing by Wavefront Reconstruction

The processor described above was certainly a successful design and a useful device. However, there were some disadvantages in the system described above. In particular, the reference function transparency (P3) could absorb about 50% of the incident light which resulted in a large attenuation and loss of light (Leith, 1971). The second disadvantage of the system was that the need to combine the processing of the signal and the reference function one onto the other is complicated and indeed rather bewildering (Cutrona et al, 1966). If by some means the two functions could be multiplied i.e. combined by directly superimposing them, the inconvenience caused by imaging one onto the other could be eliminated.

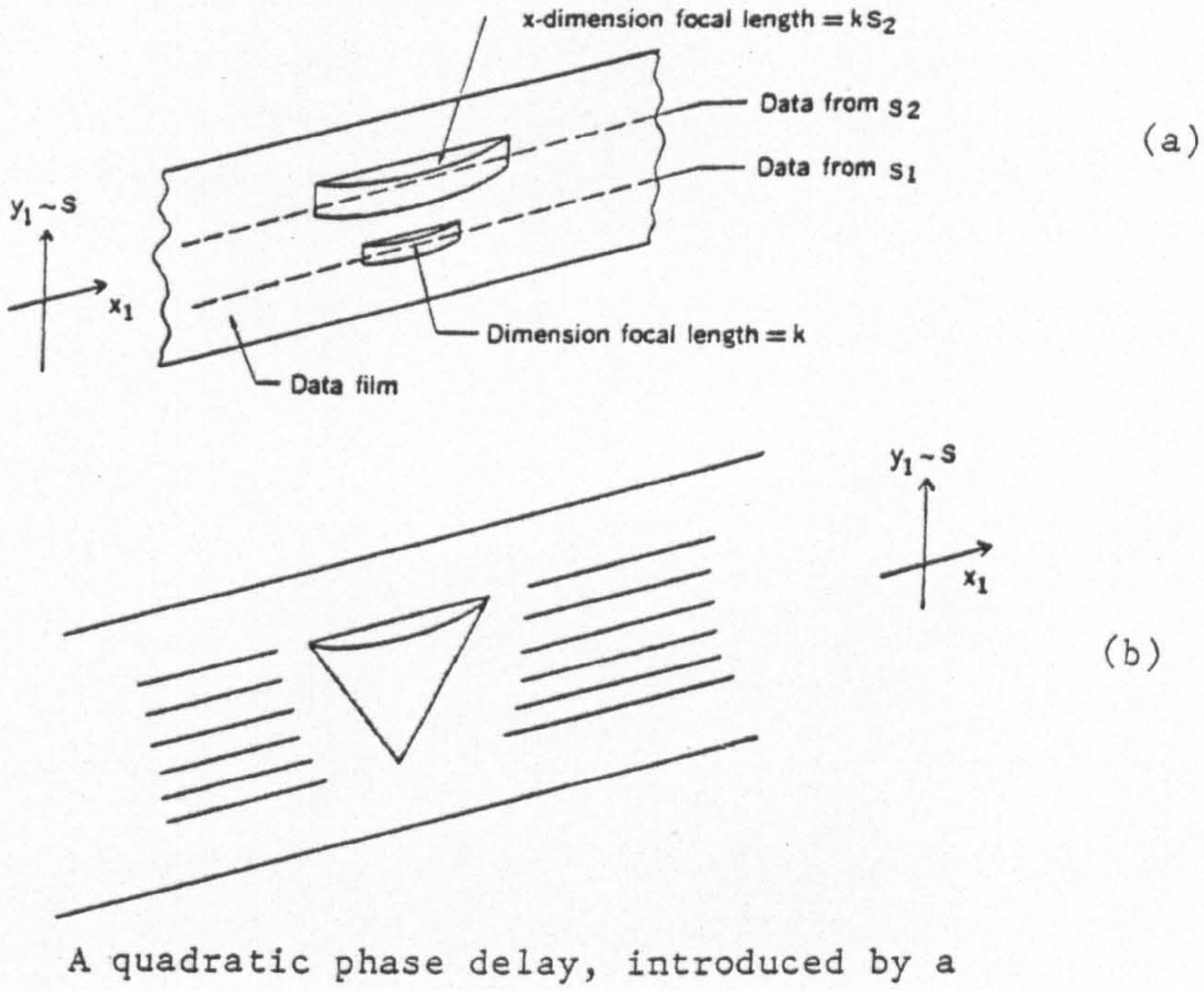
In the system previously described (Fig.3.1) the locus of azimuth focus is a plane tilted with respect to the optical system axis, while the plane of range focus is the signal record itself (Fig.3.2). These two focal planes need to be brought into coincidence. The optical correlator shown in Fig.3.3 accomplishes this task.

The concept of this type of processor evolved as early as 1955. It is based on an optical viewpoint built around Gabor's theory of wavefront reconstruction (Leith and Ingalls, 1968) which is essentially a holographic one. This means that the synthetic-aperture is realized through conversion into an equivalent real-aperture of a holographic optical system. However, imagery formed in such a way

will have various special characteristics e.g. the received signals have focal lengths proportional to the range to the object, thus the locus of focal positions is a tilted plane (Leith and Ingalls, 1968). This may be compensated for by use of a conical lens. Its effect is to compensate for the quadratic phase delay which occurs when the curved wavefront reflected from an object in the terrain is received along the synthetic aperture. The correction may be provided by a simple plano-convex lens. However, the correction will be different for each range which is recorded on the film so requiring a different focal length for each range. The use of the conical lens gives this required property (Fig.3.4).

Since the signals from any individual object are recorded in the x-direction only, all of the above operations take place in the azimuth or along-track direction only. Provision must also be made to focus the image in the range direction and so a cylindrical lens has to be provided to focus the image in this direction. Expressed in another way, since the radar system is anamorphic with respect to its recorded image data, a compensating anamorphic property has to be utilized in the optical system which processes it.

The optical system therefore consists of three lenses, one spherical, one cylindrical and the third a conical lens. The conical lens with a variable focal length matched to each line of the signal record is placed against the signal film (Pl). This re-images the tilted azimuth plane at infinity and also erects it while doing so. Next, the cylindrical lens placed one focal length from the signal record performs the anamorphic transformation required, operating in the range direction only, and brings the two planes in coincidence (Figs.3.5, 3.6 and 3.7). Finally a conventional spherical lens then



(a)

(b)

Fig 3.4 A quadratic phase delay, introduced by a plano-convex lens (a) is used to focus abscissa information for a particular slant range S. To bring entire area between two slant ranges into focus, therefore requires a conical lens (b) (Brown and Porcello, 1969).

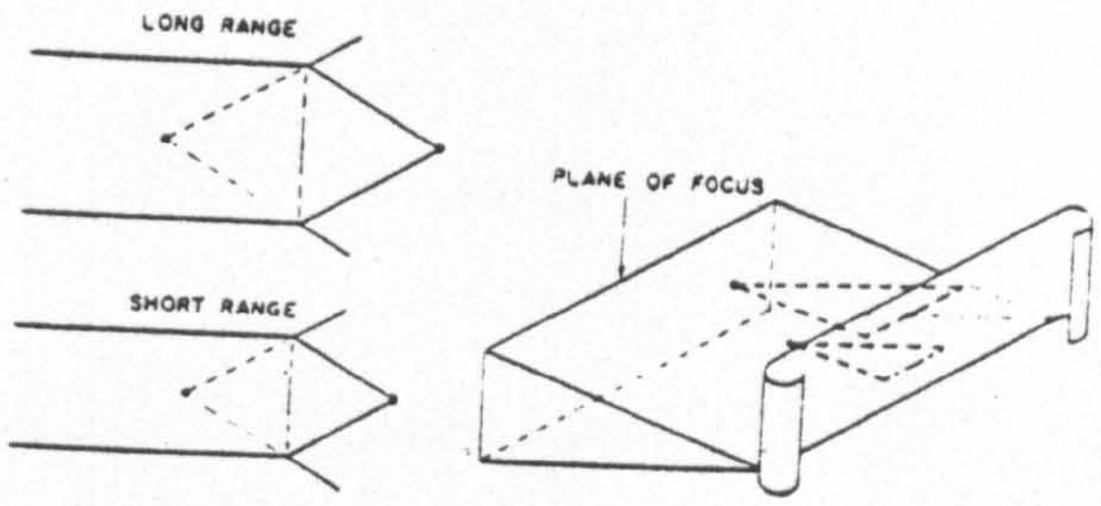


Fig 3.5 Focal properties of the signal record.All targets have focal points lying in the tilted azimuth plane when the signal record is illuminated at normal incidence with a collimated beam.Targets have focal lengths proportional to range(Leith and Ingalls,1968)

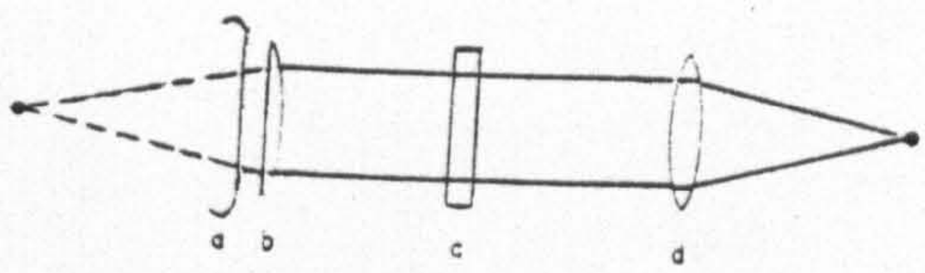


Fig 3.6 The performance of the anamorphic system in azimuth direction, the conical and spherical lenses re-image the virtual image formed from the signal record.

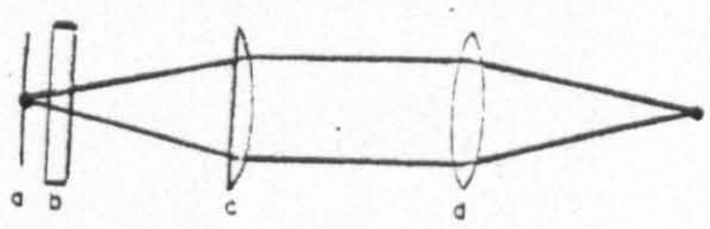


Fig 3.7 The performance in the range direction , the cylindrical and spherical lenses re-image the signal record at unity magnification(Leith and Ingalls,1968)

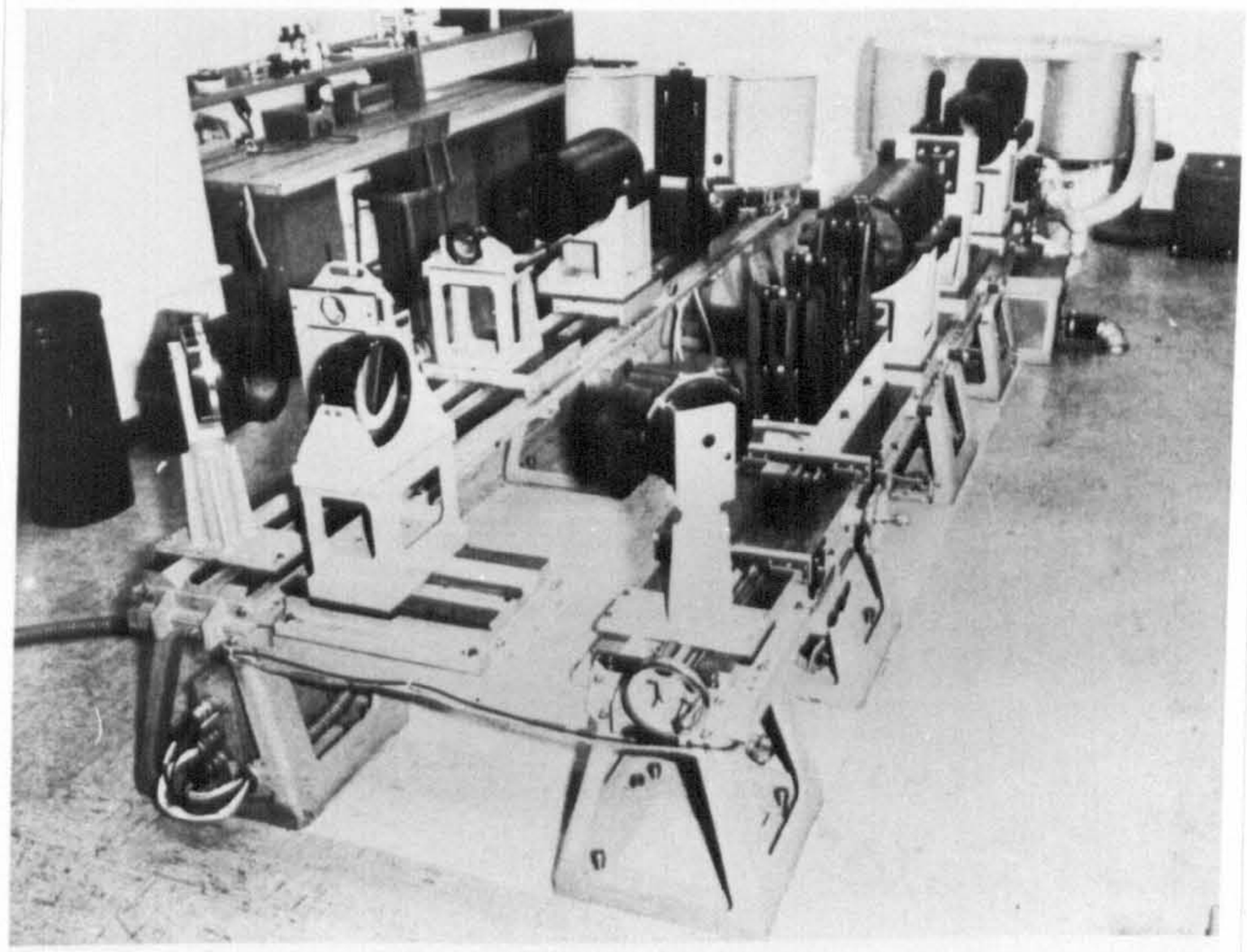
forms a real image free from the original defects. This operates in both directions i.e. azimuthal and range, and takes the transformed image at its infinity focus and re-images it at its focal plane where the image can now be sharply focussed in each direction.

This system was twenty times better in its light utilization than that previously described (Leith, 1971). It is also much simplified. Since there is no bias term to remove as in the previous system, the conical lens may be placed in direct contact with the signal film, thereby eliminating two lenses and a spatial filter. The system, of course, comprises a most unlikely combination of lenses, but accomplishes a task that would have been very difficult to achieve by any other means. An optical processor built along these lines is the Universal Radar Signal Processor of the U.S. Army Engineer Topographic Laboratories (Fig.3.8).

3.3.3. Optical Processing Using the Tilted Plane Concept

While the processor described in the previous section was a great success and formed the basis of successful production-type instruments, it did require the production of very special optical components such as the conical lenses which are difficult to manufacture and are correspondingly expensive. An alternative type of optical processor which eliminates the need for the conical lens is the so-called tilted plane type. As the name suggests, this uses two tilted planes, the one for the signal film and the other for the output image film, to eliminate the need for the conical lens. The optical arrangement is shown in Fig.3.9.

The signal film is illuminated by collimated monochromatic light which, in a modern optical system, is readily generated by a laser. A two-lens spherical telescope (lenses L1 and L2) of unity



Universal Radar Signal Processor

Fig 3.8 The Universal Radar Signal Processor

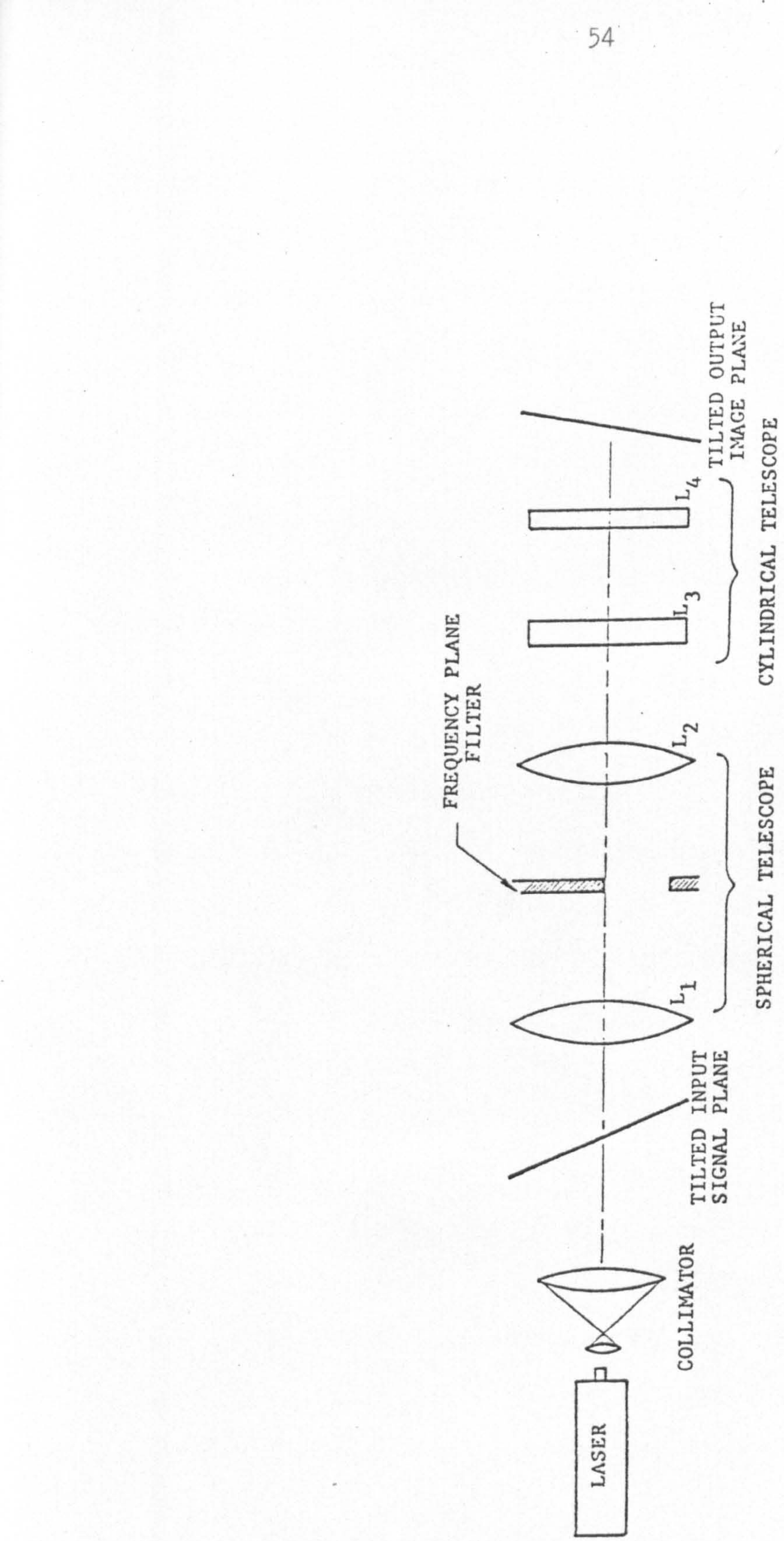


Fig. 3.9 Atypical tilted plane optical processor configuration

magnification is combined with a tandem lens cylindrical telescope of variable magnification consisting of lenses L3 and L4. Taken together, these form an output image which is sharply focussed in both directions. A tilt angle exists which brings the two image planes into focus in the output plane. As before, both the input and output films must be translated in a synchronised fashion, through the processor, thus forming a continuous correlation in the azimuth dimension.

The ERIM tilted-plane processor can readily be adapted to accommodate signal films from a variety of imaging SARs and offers an excellent degree of flexibility in handling a variety of ranges, depression angles and resolution requirements. Images of the River Tay area in Scotland shown as Figs. 8.2, and 8.3 were optically processed using this kind of SAR optical correlator.

The obvious advantage of the processor is that it can be modified to accommodate signal films of various aspect ratios. This can be done by replacing the two lens cylindrical telescope consisting of lenses L3 and L4 by an alternative three-lens cylindrical telescope with the capability of allowing variable magnification of the final image. The system has also the advantage of being simple and easy to understand. However, requirements regarding (i) the dimensional tolerances of the film drive and its stability during operation of the correlator,and (ii) the still demanding lens performance requirements associated with rigorous applications make the construction of such a device technically challenging and often expensive (Ausherman, 1977).

3.4. SAR Digital Processors

Although the basic idea goes back to the early fifties (e.g. see Sherwin et al, 1960), digital techniques of SAR data processing are

relatively new as far as their actual implementation is concerned. Most implementations of the digital approach take advantage of the range and azimuth separatability associated with the required two-dimensional correlation process. Fig.3.10 shows a general SAR digital processing system showing the relationship between range and azimuth processing (Ausherman, 1977).

Using the digital signal notation commonly used in processing, the actual processing operations to be carried out by the digital SAR processor can be expressed mathematically as:-

$$O(n,m) = \sum_{xy}\sum S(x,y)w(x-n,y-m)fm(x-n,y-m) \quad \text{(Ausherman, 1977)}$$

$$(S \text{ and } f \neq o) \qquad 3.2$$

where S(x,y) is the sampled SAR signal in range (y) and azimuth (x);

fm(x-n, y-m) is the sampled version of the reference function, where n and m represent different values of azimuth and range respectively;

w is a two-dimensional weighting function used for impulse response control in a manner equivalent to the frequency filtering in the optical case; and

O(n,m) is the sampled output radar image.

As will have been apparent from the discussion in Section 2.5, the two-dimensional sampling frequency of these functions must be adequate to unambiguously sample the RF bandwidth in the range dimension and the Doppler bandwidth (as determined by the real antenna beamwidth) in the azimuth dimension. The range sampling rate must be met by the system analogue to digital (A/D) converter and the azimuth sampling is initially determined by the radar pulse repetition frequency (PRF). The first two components of the digital processor shown on Fig.3.10 - the PRF Buffer and Azimuth Presummer - carry out certain pre-processing

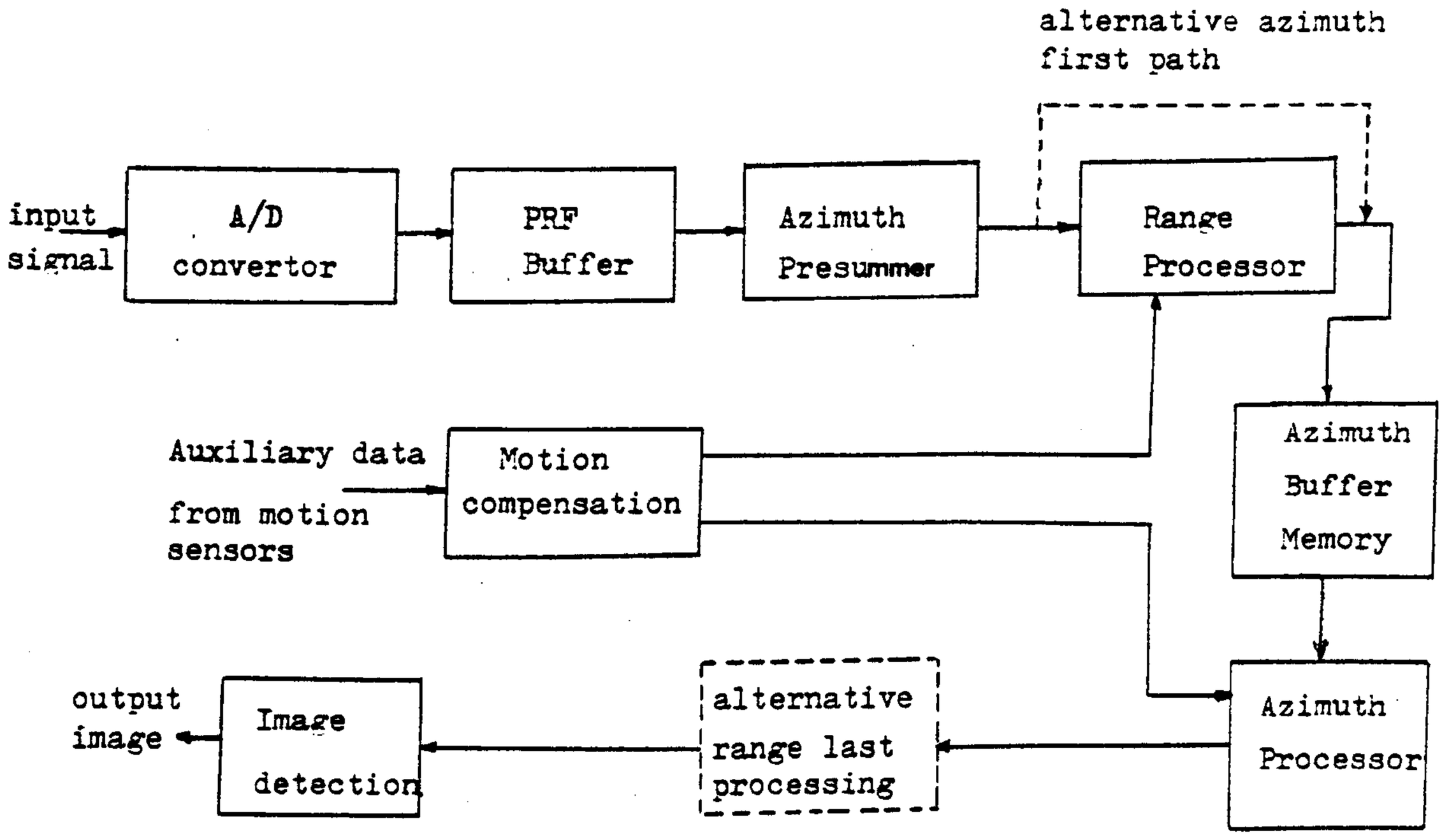


Fig 3.10 SAR image-formation digital processing system.

functions which minimize the data rate through the rest of the processor. The PRF Buffer serves to stretch the occurrence of the A/D output samples over the time interval available between subsequent pulses. The azimuth pre-summer serves to reduce the Doppler bandwidth to the minimum required for the desired azimuth resolution. It is therefore a low pass digital filter in the along-track direction capable of outputting samples at a lower rate than input.

These preliminary operations are followed by the range and azimuth processors each of which performs a one-dimensional pulse compression in the appropriate direction only. Since the range processor works in the range direction only, the actual processing merely requires access to an entire digitized radar pulse. Since the azimuth processor must carry out repetitive integration in the along-track direction, it must have access to a large number of along-track samples in each range sample interval. Thus it must have access to many parallel channels corresponding to the range resolution intervals. In addition, since the azimuth compression is range-dependent, a different reference function has to be used at each range.

The general procedure for azimuth compression illustrated in Fig.3.11 illustrates these points very well. The range compressed data are fed into the memory of the processor. This data is tapped at intervals corresponding to the number of range elements which comprise the swath. The values which are obtained are then weighted by appropriate reference function values and then integrated to give an output image element (or pixel) which is recorded for later display. The output pixels are obtained sequentially in the range direction at the same rate as the range compressed data samples are being fed into the memory. After a full cycle across the image, the next output

Source?

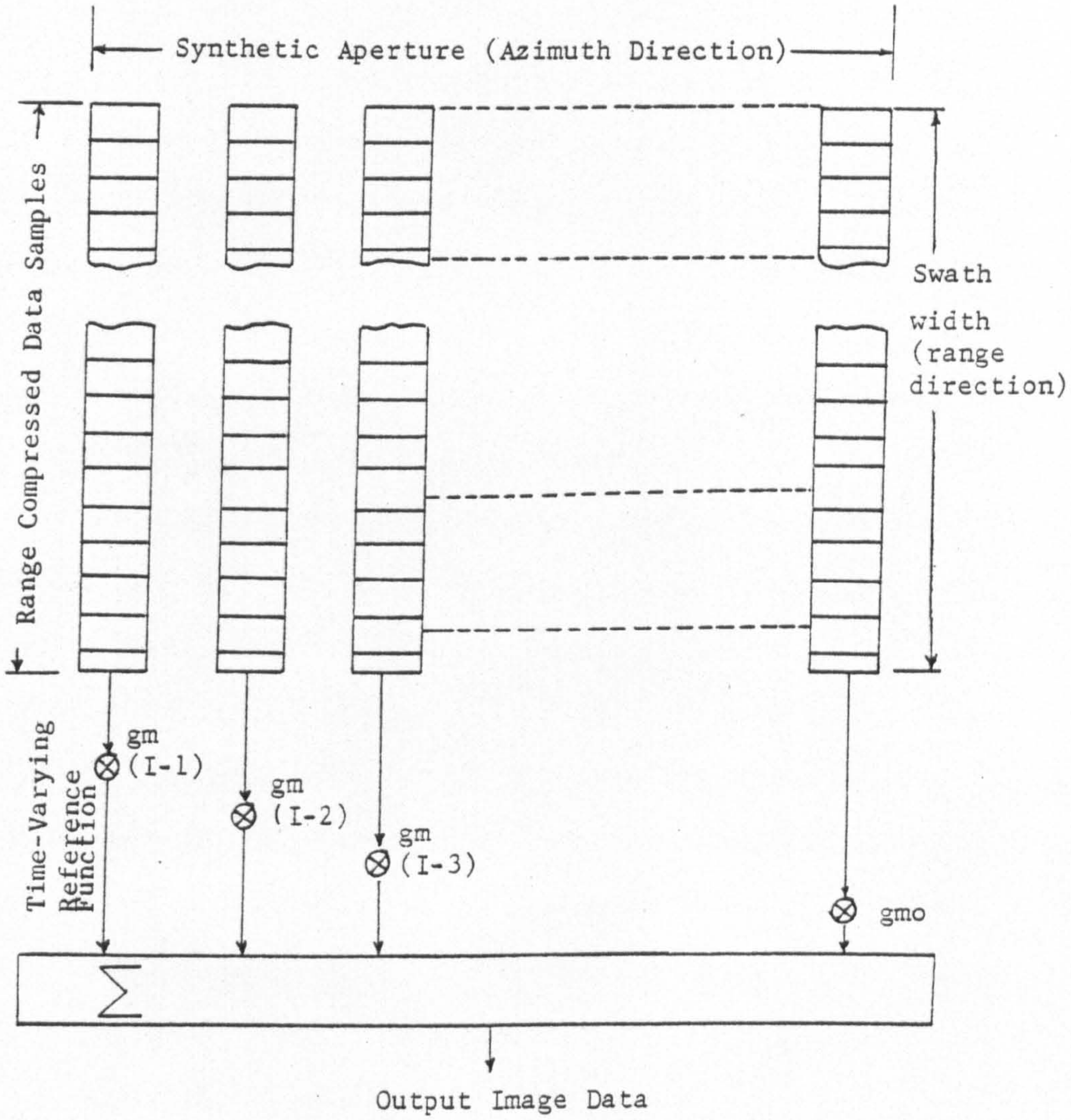


Fig 3.11 Convolution Azimuth Processor

sample which is formed is the first range element in the subsequent azimuth scan line position.

The reference function values g_m (I-1),....., g_{mo} correspond to changes in the subscript of g_m to accommodate the range-dependent azimuth focus. The values of these coefficients will continually change. To reduce the computational burden, various methods may be used. One is to use a nominal or average value of g_m which will provide an adequate focus over a number of range-sampled intervals. Another approach is to compute the values of g_m in advance and store them in a ROM (Read Only Memory) which allows them to be read out at high speed. Yet another approach is to compute the coefficients using specially built high-speed arithmetic units.

Returning to equation 3.2 above, the subscript on the function labelled fm represents the range dependence of the system reference function in the equation. In the absence of range curvature effects or when its effects are negligible (as is the case with many airborne SAR surveys), and taking some liberty from the constraints imposed by range dependence of fm, it can be seen that this operation can be performed as two simple sequential one-dimensional operations since

$$fm(x,y) = g_m(x)h(y)$$

where $g_m(x)$ represents the sampled, range-dependent quadratic phase term in azimuth;

h(y) represents the sampled quadratic phase term in range; and where the amplitude of the reference function has been selected as unity.

The same argument applies to the impulse response control function w where:

$$w(x,y) = wg(x)wh(y)$$

which once again allows for independent control of the impulse response in both directions.

Carrying out range-compression first, one obtains the expression:

$$O(n,m) = \sum_x w_g(x-n)g_m(x-n)\sum_y s(x,y)w_h(y-m)h(y-m),$$

which allows the processing to be carried out as two independent, one-dimensional operations, the one in the x-direction, the other in the y-direction.

As discussed above, the execution of range-compression as the initial stage of processing is generally preferred due to the fact that the azimuth focus is range-dependent.

Not only is the prior execution of range-compression desirable due to the range-dependent azimuth focus, it is also necessary to permit simple methods of range curvature correction. In a mathematical sense, the appropriate expression for the function g_m can then be used in the final formation of the sampled output image (i.e. the function $O(n,m)$). One result of adopting the procedure of executing range-compression first is that it does result in an increase in the required wordlength prior to the azimuth processing due to the growth in the dynamic range of the range-compressed data (Ausherman, 1977).

It should be noted that many other approaches can be followed to implement digital processing of SAR data based on different computational strategies and algorithms to that described above. A comprehensive account of these is given in Ausherman (1977). Each can be considered for implementation and each has its advantages and disadvantages, depending upon system requirements, the financial constraints and the available processing facilities.

Hardware architectures which can be used for the implementation of digital processing can range from hardwired systems utilizing purpose-built electronic components through high-speed special-

purpose programmable signal processors to conventional general-purpose computers. The last type, which are of course widely available nowadays can readily be applied to the digital processing of SAR data as long as time is not a crucial element. Therefore this process has been implemented recently both on mainframe and on mini-computer installations.

3.5. Combined Optical and Digital Procedures

This image processing technique combines the sheer processing power of two-dimensional optical processing techniques which can simultaneously process huge amounts of SAR data and the flexibility and versatility of digital techniques which will allow slightly different operations to be carried out on adjacent sets of image data. The best known example of this category is the ERIM hybrid optical-digital processor (Fig.3.12) which consists of a coherent optical processor connected to a digital computer through a digitally-controlled image dissector (Bayma et al, 1977).

The ERIM system was constructed so that the image occurring in any plane in the precision optical processor can be converted into a digital recording using the dissector. The latter is equipped with a special photomultiplier tube designed so that the photo cathode area can be scanned in a series of lines to convert the light energy in any plane of the optical processor into an electric analogue signal. These signals are digitized by means of an analogue to digital (A/D) convertor and processed on a mini-computer for final recording on digital magnetic tape. Subsequent digital processing can be carried out using a large mainframe computer such as an IBM 360/75 (Bayma et al, 1977).

3.6. Optical Versus Digital Techniques in SAR Data Processing

Each of the two main SAR processing techniques - optical and

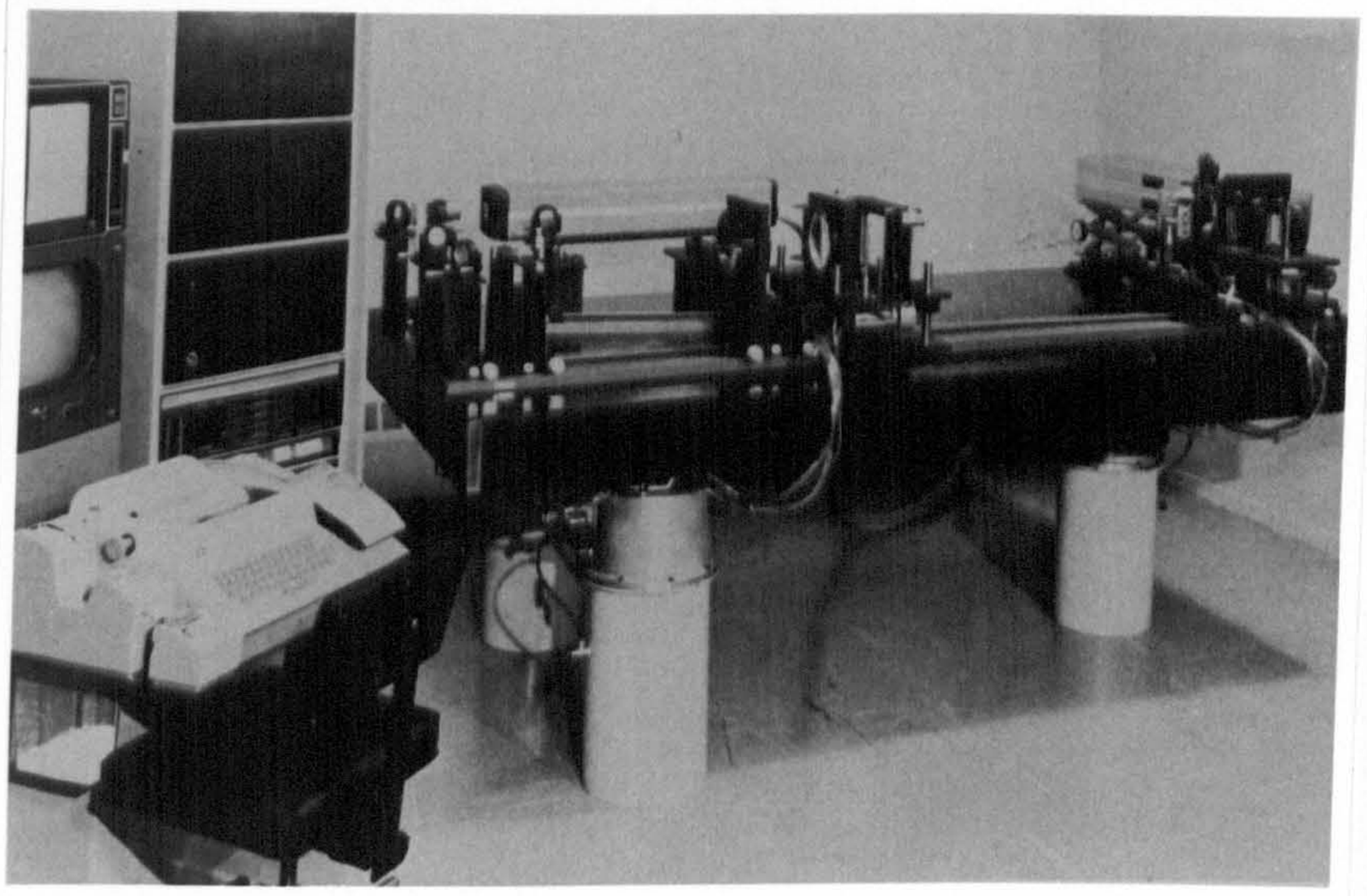


Fig 3.12 Hybrid optical-digital SAR processing facility. The optical system performs high speed parallel operations; the digital system performs slow-speed serial operations.

ERIM

digital - have relative advantages and disadvantages. In general, one can say that SAR digital techniques are not nearly as mature as their optical equivalents. Perhaps the strongest argument for the selection of the optical technique at the moment is speed. Its parallel processing capability is at the moment unequalled and can not be matched by digital processors, although there are now certain digital processors which have closed the gap to some extent. However, as already mentioned, the optical processor is essentially an off-line device with a need for expensive optical components mounted on stable bases and operating in an optical laboratory type of environment with very skilled personnel. The erection, calibration and maintenance of the optical type of processor is a task for highly skilled and well-trained specialists.

With digital processors, the situation varies from one type to another. Hard-wired processors offer the highest speed, certainly near real-time, which is attractive for airborne applications. However the quality and resolution of the output of these devices does not appear to be compatible with the needs of many users. On the other hand, the digital processor has excellent flexibility, though some degree of flexibility can also be obtained from the tilted-plane type of optical processors.

The manner in which the dynamic range of the image can be maintained is another point of discussion. To allow good representation of the ground returns, a sufficient dynamic range has to be preserved. The basic limitation on dynamic range for a digital processor relates to the wordlength used, a factor which is dependent on the characteristics of the computer and somewhat related to cost. However, with the advent of comparatively inexpensive and fast 32-bit word mini-

computers and matching array processors, this is no longer a matter of such high concern as it was previously. The dynamic range associated with recorded image output for either approach may also be limited by the characteristics of the input/output (I/O) recording devices and media. For example, the data from airborne and spaceborne SARs is often collected on high density digital tape (HDDTs). However this must be converted to another digital form or to an optical form to allow processing to take place. The usual input and output storage medium for digital approaches is computer-compatible magnetic tape. This is a comparatively bulky form of storage when used to store large quantities of SAR data. In addition, the search for the requisite data is sequential, which is inconvenient for processing SAR data so that it will normally be stored on some form of magnetic disc for fast input/output to the computer memory. For optical processors, there is less difficulty, since film serves as a sufficiently dense input/output medium.

In terms of actual use, virtually all production SAR image processing has been optical to date. Although high performance digital SAR processors have been demonstrated experimentally, e.g. the American, British and Canadian digital SAR systems used with Seasat data, no production digital processors are known to exist, at least to the present author. However, inevitably, most current research has been conducted in the digital area. Table 3.1 summarizes the merits and demerits of each of these main approaches.

Finally, the comparison of optical and digital techniques can only be conducted and the final selection made depending on the user requirements, the development time-frame, the available funding and the operational purpose for which the system is intended.

Table 3.1. Relative Advantages and Disadvantages of the Optical and Digital SAR Data Processors

Parameter	Digital processor	Optical processor
Laboratory size and environment	Normal computer room size and environment. Special digital processors possible for use in aircraft.	Current versions large with requirements for clean dust-free laboratory environment. Airborne versions unlikely.
Flexibility	Very good since its programming capability permits modifications to the processing to be incorporated relatively easily.	Rather limited flexibility (though the tilted plane is better in this aspect).
Calibration	Relatively easy, although not so in calibrating transmitter and receiver antenna.	Initially difficult though should be good thereafter.
Dynamic range	Since it depends on wordlength, it may be costly to keep the required range. Also the output image may be display limited.	Very good when high quality optics are used.
Availability	Digital computers are widely available but digital processing is virtually a research tool at present.	Relatively few optical processors exist, but most processing is done optically.
Focussing	Manual focussing impossible.	Manual focussing is fast and simple.
Signal and Image Storage	Current CCT's are bulky (HDDTs less so, but not computer-compatible). Magnetic discs needed for intermediate storage. Large processor memories needed. (32-bit machines help in this respect.)	Film is a very efficient form of storage both for input and output and can serve as the processor memory also.
Processing time per frame	Very long and costly.	Relatively inexpensive.
Quality of output image	Can be excellent.	Good, though somewhat poorer than highly refined digital processing.

However, the new type of hybrid image processing facility which combines the capabilities of coherent optical processors and digital computers, may well have a profound effect on the development of SAR data processors during this coming decade. For the first time, operators and users will be able to utilize the best features of each type of processor. The optical processor can continue to perform the relatively cheap and fast routine processing, while the digital component of the system can be used to carry out more complicated tasks over selected parts of the image data set.

CHAPTER IV

SEASAT SYNTHETIC APERTURE RADAR (SAR) EXPERIMENT

4.1. Overall Mission Objectives

Seasat was the first satellite to be dedicated primarily to the study of the sea surface and so most of its sensors and experiments were designed with oceanographic objectives in mind. According to Teleki et al (1978), the following main goals were to be pursued:

(i) to study the ocean wave patterns present in the deep ocean, on the continental shelf and in coastal waters;

(ii) to derive wave spectra;

(iii) to obtain information on land-water interaction in coastal areas;

(iv) to study sea- and fresh-water behaviour; and

(v) to obtain information concerned with the study of various phenomena associated with land surfaces, e.g. roughness, materials, vegetation, etc.

4.1.1 Seasat Sensors

To meet these objectives, the satellite carried a total of five sensors, each one complementing the others to provide an array of instrumentation capable of measuring and interrelating various ocean surface phenomena on a continuous, all-weather and world-wide basis. Four of these instruments - the scatterometer (SASS), the scanning multi-frequency microwave radiometer (SMMR), the visible/infrared radiometer (VIRR) and the radar altimeter (ALT) were concerned primarily with global ocean dynamics and were designed to obtain global data sets which were stored on-board the satellite and replayed later to ground receiving stations. The fifth sensor, the synthetic-

aperture radar (SAR), with which this study is concerned, acquired data only in real-time over a network of ground stations limited geographically to the viewing capability of the stations' receiving antennae. The SAR coverage included both oceanic and land areas.

The satellite itself was constructed by the Applied Physics Laboratory of John Hopkins University (Lodge, 1980a) which has of course gained considerable experience of earth satellite operation through its major role in the Transit series of navigational satellites. Fig.4.1 illustrates the bus and payload configuration and the actual placement of each of the five sensors on board the satellite.

Seasat was the first civilian spacecraft to carry an imaging radar system in an orbit around the Earth. However, one can be fairly certain that it has been preceded by military SARs mounted in reconnaissance satellites such as Big Bird. The experience gained with these has doubtless been utilised in the design of the Seasat SAR. Previously the only orbital side-looking radar data available to non-military users was that acquired during the Apollo-17 mission to the Moon in December 1972 (e.g. see Porcello et al, 1974; Leberl, 1976a; Tiernan et al, 1976).

4.2. Seasat Orbital Parameters

The satellite was launched on the 26th of June 1978 as Seasat-A since it was intended to be the first of a series of these satellites - Seasat-B and operational Seasat were to follow. At the time of writing, it is still uncertain as to whether these intended successors will be orbited or not. If the latter, then Seasat will then be a unique spacecraft.

Seasat was launched into a non-synchronous, nearly

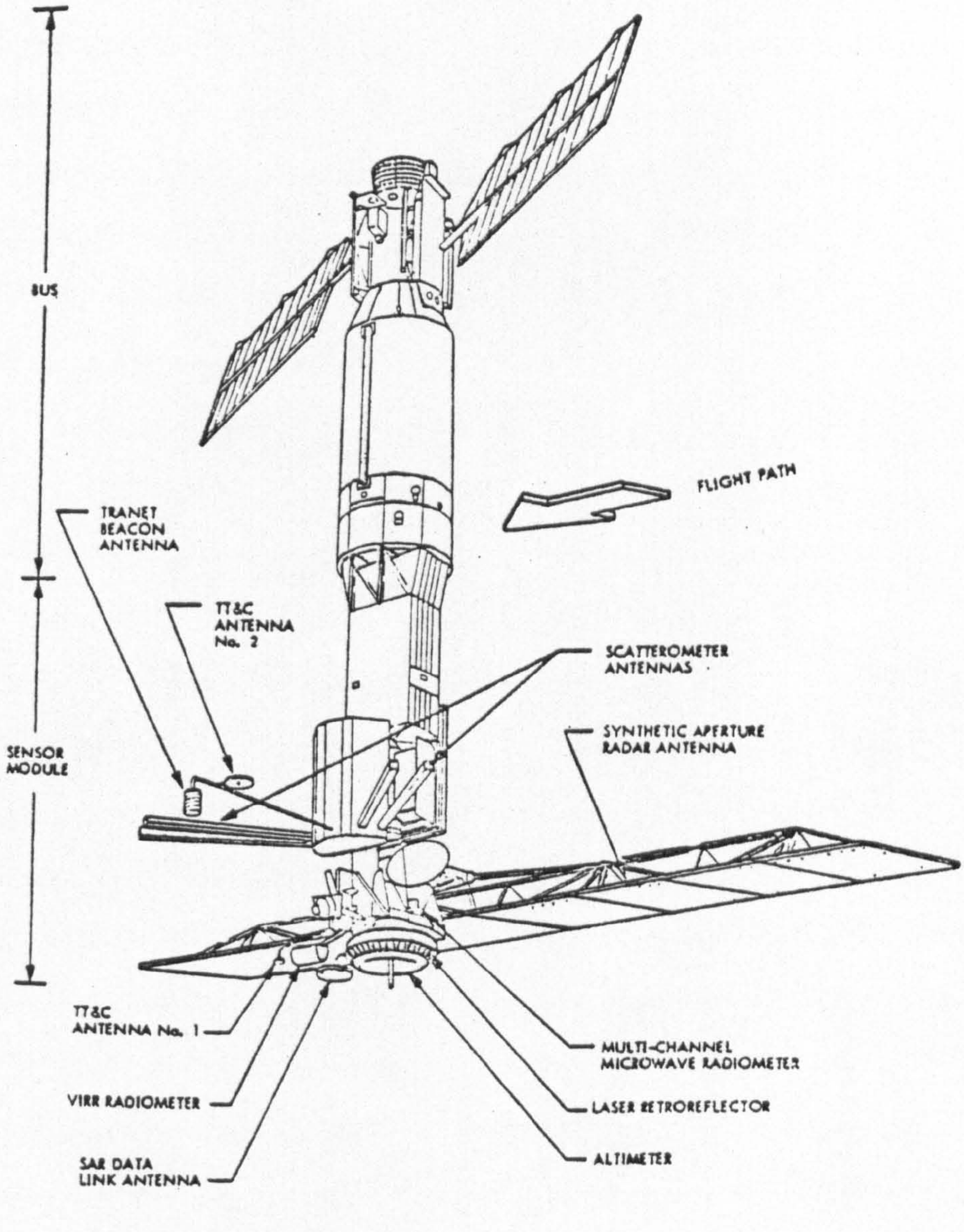


Fig 4.1 SEASAT-A bus and payload configuration (Teleki et al, 1978).

circular orbit with an eccentricity of less than 0.006 and an orbital altitude of approximately 800 kms. The orbital inclination was 108° (i.e. retrograde) to the equator so that its maximum latidudinal coverage was 72° (180° - 108°) (Fig.4.2). The retrograde orbit dictated that the launch itself had to be made across the Pacific Ocean from Vandenberg, California where a clear firing direction would be obtained. The orbital period of Seasat was of the order of 100 minutes and 45 seconds, resulting in about 14.3 orbits being completed per day. During the period between the satellite passing over the equator on successive orbits, the Earth will have rotated 25.2° to the east (15° rotation in 60 minutes; therefore 25.2° in 101 minutes). This means that each equatorial pass is some 2,800 kms to the west of the previous one.

The satellite was placed into what is called a "frozen orbit" for a period of some weeks. The characteristics of this type of orbit are such that the ground track does not progress around the Earth to ensure contiguous systematic cover as with the well-known Landsat series of satellites but simply repeats itself on a three-day cycle (Fig.4.3). This means that the coverage of large areas is not possible nor was it intended, the emphasis being on the repetitive coverage of selected areas. These comprised certain polar ice and coastal and oceanic areas of North America and North-West Europe which would be in the view of the few receiving stations able to receive the great volume of data generated and transmitted by Seasat.

It will be noted that, since all the sensors except one (the visible/infrared radiometer) operate in the microwave part of the spectrum, considerations of sun elevation and shadow which play such an important part in the orbital design of most earth resources satellites (resulting in sun-synchronous orbits for example) played

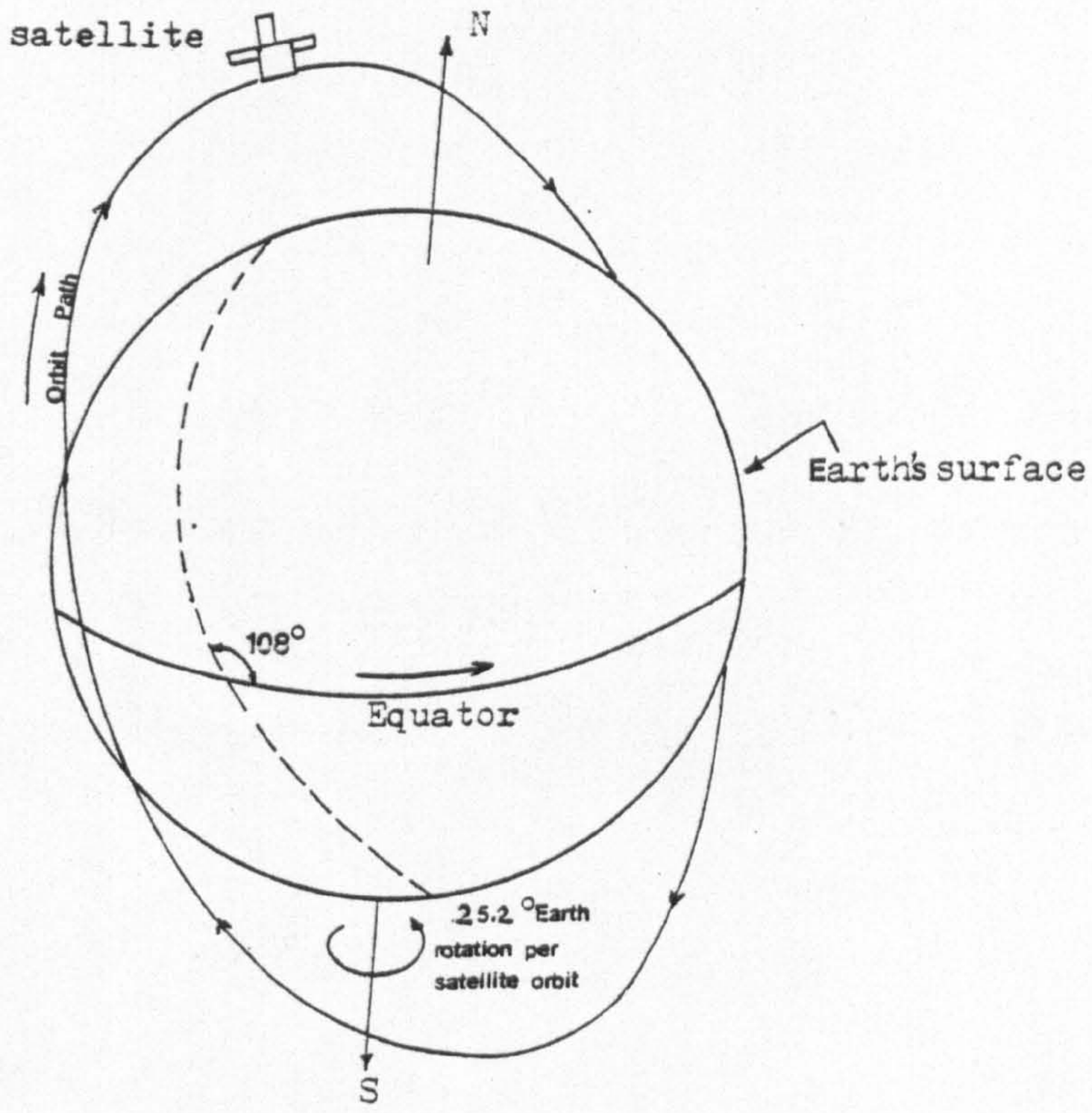


Fig 4.2 Seasat-A orbit geometry.

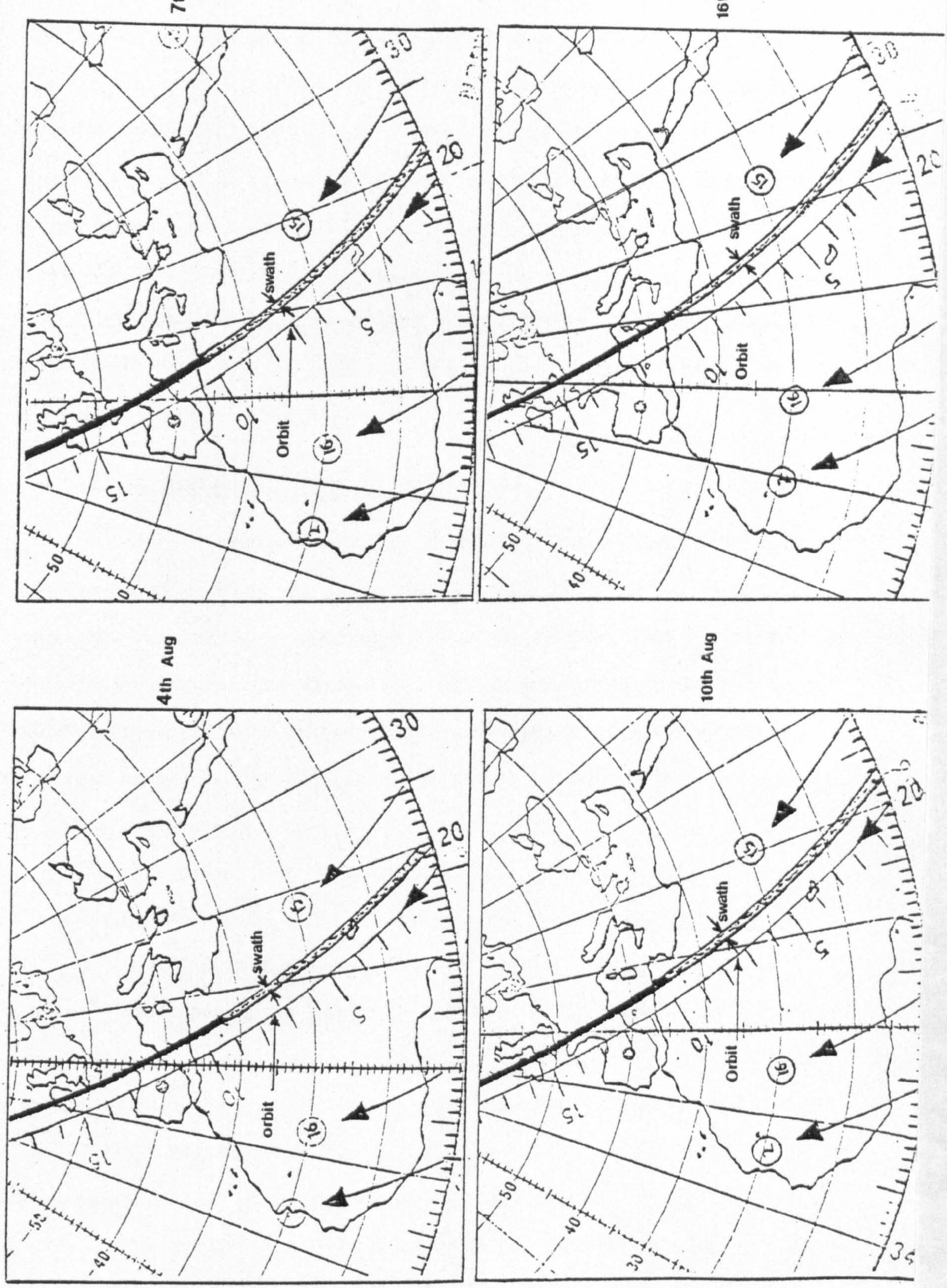


Fig 4.3 Seasat repeats itself every three days.

little part in the design of the Seasat orbit. Furthermore, the all-weather, day and night capabilities of microwave sensors meant that the satellite could acquire data on both the ascending and descending parts of an orbit, whereas a satellite such as Landsat is restricted to the collection of data on the descending part of the orbit and then only if the terrain is in daylight. Such considerations help to explain the very large data stream generated by the sensors which posed considerable problems both in the reception and in the subsequent processing of the data.

4.3. Characteristics of the Seasat SAR

In the design stage of the mission, NASA sought the involvement of a large number of the intended users of Seasat data to ensure that the needs of each particular application were met as far as possible (M cCandless et al, 1975). Obviously a compromise had to be sought between the many conflicting requirements; and the design and operation of the SAR itself cannot be viewed in total isolation from the other sensors.

The Seasat SAR operated at a wavelength (λ) of 23 cm (i.e. in the L-band UHF) with an 11 m long antenna (D) and a beam-width (β) of 1^{o} in azimuth (see Fig.4.4). The radar used a transmitted band-width of 19 MHZ and pulse length (τ) of 33.9 μs. A frequency modulated (FM) chirped pulse was used to improve the range resolution (Rr). Four pulse repetition frequencies (PRFs) were provided ranging from 1,464 HZ to 1,647 HZ. The actual PRF selected will depend on the type of object (land, ocean, ice, etc) which is being imaged.

The choice of a long wavelength (23 cm) for the Seasat SAR was dictated by two main factors. The first is that the power

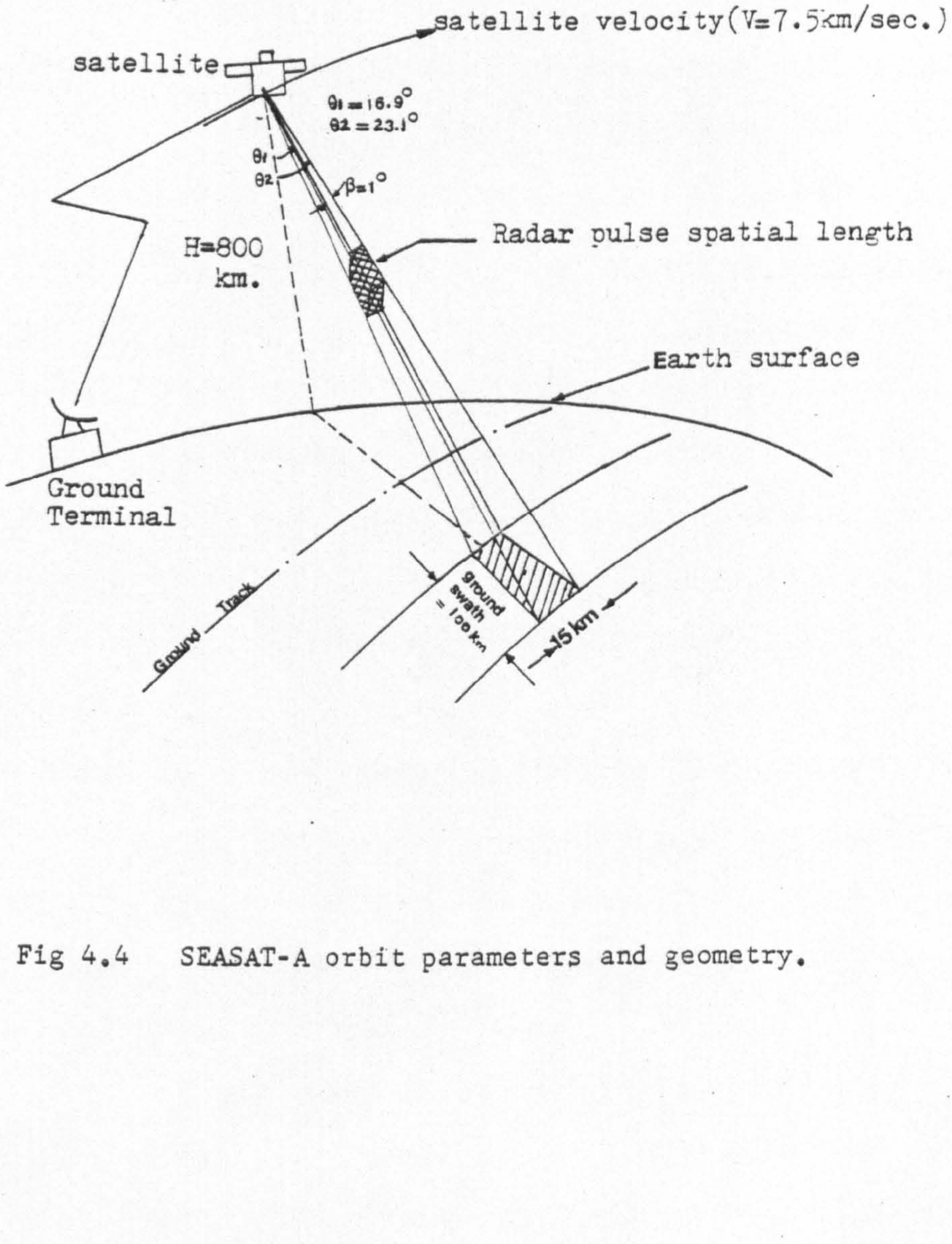


Fig 4.4 SEASAT-A orbit parameters and geometry.

requirements for a SAR are inversely proportional to wavelength - in the case of the Seasat SAR, the average power consumption was only 250 watts. The other important factor is that signal losses resulting from the attenuation caused by the atmosphere and by clouds and the backscattering from precipitation tend to increase sharply at short wavelengths. By contrast, L-band radar is relatively immune to these effects - a most important consideration in an orbital radar.

Elevation angles (θ) from the vertical varied from 16.9° at near range to 23.1° at far range so that a steep look angle is a characteristic of the imagery. The minimum and maximum slant ranges were 836 km and 879 km respectively and a ground swath of approximately 100 km resulted located to the right of the satellite path. The range of look angles (6.2° in the vertical plane) is rather small compared with conventional airborne systems. However, this is directly related to the very high orbital altitude of the satellite and to the energy requirements which limit the vertical width of the microwave beam.

As with any other type of SAR, the Seasat example has two independent resolutions in the azimuth and in the range directions. The along-track (or azimuth) resolution (Ra) of the SAR system is of the order of 7 m and cross-track (or range) resolutions (Rr) are 8 m in slant range and 25 m in ground range.

As previously discussed, the cross-track or range resolution $(Rr) = \frac{\tau}{2} c$ where τ is the pulse length and c is the velocity of electromagnetic energy. Since the pulse length (τ) is 33.9 μs (microseconds) this would result in a value of $Rr = \frac{33.9 \times 10^{-6}}{2} (3 \times 10^{8}) = 5085$ m which certainly would be of little use for most remote sensing applications. To obtain a much finer resolution, the transmitted pulses are frequency modulated or "chirped". The return signal of a

chirped radar has to be processed in an optical or digital correlator in order to compress the signal and obtain a slant range resolution equivalent to that of an unmodulated pulse of much shorter duration. For Seasat, this process yields a fine range resolution of about 6 m in slant range. In practice, this will be degraded due to atmospheric effects and slight departures of the electrical components from their designed performance to give a figure of 8 metres. This will then have to be converted to actual ground resolution taking into account the steep look angles. At the minimum elevation angle of $\theta = 16.9^{o}$,

$$Rr = \frac{8}{\text{Sin } 16.9} = 8 \text{ x cosec } 16.9^{o} = 27.5 \text{ m.}$$

Since the Seasat SAR is a fully focussed synthetic-aperture system, its azimuth resolution (Ra) is theoretically equal to half the real antenna aperture (D) i.e. 5.5 m. This value is affected by atmospheric conditions and slight degradations in the performance of the electronic components. Therefore, a more realistic figure would be 7 m. In fact, the 7 m azimuth resolution is artificially degraded during subsequent processing to 25 m to achieve equal resolution both along- and across-track. However, the final image resolution will depend greatly on the type and method of processing used to produce the synthetic radar image data e.g. whether survey optical processing, precision optical processing or digital processing was employed for the purpose.

Because of the very high data rate (110 Mbits/s) generated by the SAR, this sensor was only operated in real time while within sight of a tracking station equipped to receive and record its data. Furthermore, because of the limited power which can be generated by the solar cells of the satellite, it was only possible for the SAR to be used to generate imagery for a maximum period of about 10 to 12 minutes per orbit.

At a ground speed of 7 km/second, the images would thus cover a swath of about 4,000 kms in length and 100 km in width.

The need for very rapid scanning to be carried out by the SAR resulting from the ultra-high speed of the spacecraft results in the need for a very rapid pulse repetition frequency (PRF). Also, the extreme velocity will lead to discontinuities in the aperture. Obviously, the distance (D_s) travelled by the real aperture, i.e. the antenna, during an interpulse period will be equal to the satellite orbital velocity divided by the radar PRF (Entres, 1970). With the high satellite velocity, the values of D_s will become many times larger than the physical length of the antenna (D). This results in a discontinuous aperture leading to the possibility of angular ambiguities. To eliminate this effect, the value (D_s-D) must be decreased to the point where the first major sidelobe of the synthetic aperture falls outside the mainlobe of the real aperture (Greenberg, 1967). Combined with this is the requirement for a narrow beam width from orbital altitudes. Both of these reasons led to the choice of the long antenna (11 m) used in the Seasat SAR. Such a length would have been rather impracticable if mounted on an aircraft, but once it has been deployed it will of course cause far fewer problems in the vacuum of space where atmosphermic turbulence is unknown.

4.4. The Telemetry and Reception of Seasat SAR Data

In theory at least, the ground returns received by the satellite SAR receiving antenna can be processed either on-board the spacecraft or at a ground terminal. In both cases, data has to be telemetered from the satellite to the Earth. The attraction of on-board processing is of course that only the final processed data would have to be transmitted to the ground receiving station, thus

reducing greatly the extent of the data flow. However, it will be readily apparent from the previous discussions on processing that there is utterly no way at present or in the foreseeable future that the type of complex optical or digital processing described in the previous chapter can be carried out on-board a satellite for reasons of weight, bulk, complexity, power requirements and reliability.

Furthermore, the very large data sets generated by a SAR precluded any form of on-board recording, especially when experience with Landsat and other earth resources satellites has shown that recorders are the weakest point in all of these operations. Thus, real-time acquisition and transmission of the data was inescapable, especially when linked to the limited period (10 to 12 minutes) per orbit when the SAR could be operated. It was originally planned that Seasat SAR data would be collected through modified Landsat receiving stations. The additional receiving equipment required and the considerable changes that would have had to be made to the receiving antenna would have required massive expenditures. Portable stations to receive the SAR data were also considered by NASA but were ruled out as impracticable.

In the end, NASA adopted the solution of receiving Seasat SAR data through five existing satellite receiving stations - Fairbanks (Alaska), Goldstone (California), Merrit Island (Florida), Shoe Cove (Newfoundland) and Oakhanger (England) (Teleki et al, 1978). The three American stations form part of NASA's Satellite Tracking and Data Network (SDTN) operated from the Goddard Space Flight Centre (GSFC) and are equipped with large antennae capable of dealing with the high data rates of the SAR and with the rapid movements required to track an Earth-orbiting satellite.

4.4.1. Oakhanger Receiving Station

The only European receiving station for Seasat data was set up by the Remote Sensing Unit of the Royal Aircraft Establishment, Farnborough (U.K.) on behalf of the European Space Agency (ESA) and was operated by RCA Ltd. The receiving dish, located at Oakhanger not far from Farnborough, had previously been used for the control and reception of data from military geo-synchronous communications satellites. This had two significant characteristics regarding its use to collect data from a near-earth orbiter. The first was a very favourable one, namely that the diameter of its receiving dish was 18.3 m which would allow the reception of large amounts of data in a short time period. The second characteristic was a highly unfavourable one - that it had a limited azimuth tracking rate of only 2^{o} per second. This would have prevented data acquisition from passes reaching elevations greater than about 70^{o}. Accordingly, and with some trepidation (Lodge, 1979) the azimuth drive system was modified to achieve a maximum rate of 10^{o} per second, so enabling data to be collected to within about 5^{o} of the zenith. Lodge comments that these very rapid movements of such a massive dish provided an awe-inspiring sight, even when viewed from a safe distance.

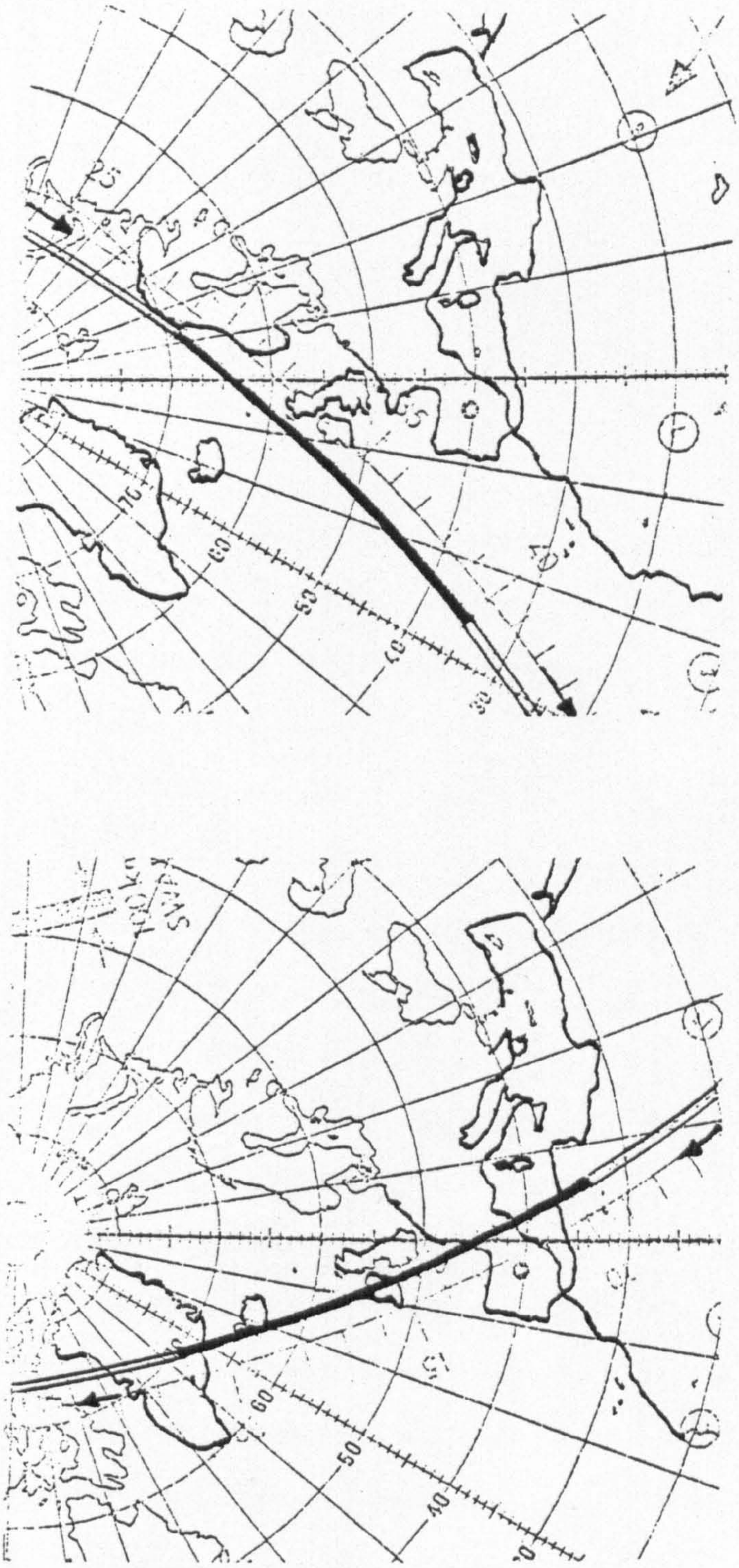
A feature of the Seasat SAR System is that all the essential timing marks and frequencies used throughout the system from the received signals to the final processing are derived from the stable oscillator on board the satellite. Therefore, along with the SAR radar data, the clock signal and an encoded marker were transmitted to enable the pulse repetition frequency and the other essential parameters to be reconstructed on the ground.

The signals were transmitted from the satellite and received

at the 18.3 m diameter dish aerial at Oakhanger using S-band frequencies. The data was digitized at a sampling frequency of 45.5 MHZ and a resolution of 5 bits per sample. A data formatter buffered and re-formatted the resulting bursts of digitised data to deliver a continuous data stream as output at approximately 110 Mb/second. These data were recorded on a purpose-built Martin-Marietta high data rate recorder operating at 380 cm per second and using 40 out of the 42 available tracks on the 25 mm tape (Lodge, 1980a).

The format used a major/minor frame structure. The actual SAR data was recorded within a constant minor frame size made up with fill data as required. A series of those minor frames were compiled into a major frame of 13,680 radar samples in which the timing and status information was also included. Sections of the recorded data were later extracted from the HDDT tape using a specially-built hardware and transformed using the Prime 300 mini-computer located at the Oakhanger ground terminal so that the data was produced in the form of CCTs which could be used as input to a digital processing system. A data tape containing the satellite orbit and altitude information necessary to compute the SAR processing parameters was also prepared off-line.

The Seasat satellite ceased sending data on the 10th of October 1978. In the short (3½ month) period of its operation, fifty three Seasat passes were received at Oakhanger representing 265 minutes of SAR data in total. The individual records are from one to twelve minutes long, with an average of about five minutes. They produced data over areas extending from North Africa to Greenland and from Northern Scandinavia to Mid-Atlantic (Fig.4.5).



(a)

(b)

Fig 4.5 Seasat coverage from Oakhanger Station extends from North Africa to Greenland (a), and from Northern Scandinavia to Mid-Atlantic (b).

4.5. Pre-orbital Aircraft Tests of the SAR System

Before the launch of the Seasat satellite, NASA organized a whole series of tests using similar SARs to that intended for use in Seasat but mounted in aircraft. These were intended to prove the operational characteristics of the SAR and its reliability before it was integrated into the satellite itself.

4.5.1. AIDJEX Experiments

The first was the series of NASA-AIDJEX experimental flights designed to analyse and demonstrate the applicability of radar imaging to mapping off-shore sea ice and to gain experience both with the SAR itself and its supporting processing. The main objectives were as follows (McCandless, 1975):-

(i) tests and check cut of the sensor design;

(ii) tests of the overall system design;

(iii) tests of the algorithm development;

(iv) tests for geophysical validity; and

(v) the development of the user-data interface.

This particular experiment started in 1971 (Thompson et al, 1971) as a pilot study carried out in collaboration with the Jet Propulsion Laboratory (JPL). SAR L-band imagery of Arctic ice floes covering an area of up to 10 km x 20 km was obtained and correlated with photographic coverage of the same area in order to establish the similarities and differences between the various image records. Preliminary radar mosaics were also constructed in order to study Arctic sea ice drifts (Thompson et al, 1971).

As a continuation of this program, three further flights over the North Slope of Alaska and Arctic Sea were conducted in the

winter, summer and fall-ice of 1975 (Leberl et al, 1976c) with the L-band SAR system mounted on-board a NASA CV-990 aircraft equipped with an Inertial Navigation System (INS). Single images were transformed to a geocentric coordinate system, and the accuracy with which sea ice drift can be measured was evaluated. The results of these tests are presented in Leberl et al (1976c).

The AIDJEX experiment also aimed at a comparison of the results of interpreting the SAR radar imagery of sea ice in the Arctic region with those obtained from Landsat (Hibler et al, 1975; Dunbar, 1975; Campbell, 1975; Parashar et al, 1974). The results of this comparison were very favourable to the L-band SAR images. The longer wavelength radar penetrates the snow cover and passes through part of the ice. Therefore, the images appear very different to those generated by shorter wavelength X-band radar for example. The results also showed that the mapping of sea ice with an L-band radar gave an overall positional accuracy of the same order as that produced from Landsat. The comparisons also showed that the resolution of the SAR images was much higher (Leberl, 1976c). Since it can be obtained in any weather conditions and day or night and not only in conditions of good lighting and absence of clouds, this would suggest that, as a tool for mapping and interpretation of sea ice distribution, side-looking SAR image data is superior to Landsat data (compare for example Hibler et al, 1975; Dunbar, 1975 and Leberl et al, 1976c). In the poor weather conditions and limited winter daylight of the area, these are matters of considerable importance.

4.5.2. ERIM Experiments

Investigations into the possibility of using a multi-channel radar system in orbital missions began in the mid-seventies (Drake and

Shuman, 1974). The ERIM 4-channel radar (which allows the use of two wavelengths and two types of polarizations) was flown over Florida and the south-eastern part of Michigan in April 1974. The results of the test showed that significantly more information for the mapping of vegetation communities and for water resources management was obtained from the combined X- and L-band SLAR imagery than could have been obtained from the imagery of either wavelength by itself and that the major differences in appearance on the multiplexed SLAR imagery of vegetation and water resources features were primarily wavelength differences and not polarization differences.

As a continuation to this test, in December 1975 ERIM collected dual wavelength and dual polarization SAR data of ocean waves off the eastern coast of Florida near Marineland together with simulated Seasat SAR data of coastal wetland areas adjacent to Marineland (Shuman and Lowry, 1977). Ground-truth for this study included colour and black and white aerial photography as well as surface ground truth photography and documentation. Each of the four SAR channels interacts with the roughness of surfaces in a different way to produce different degrees of diffusion and depolarization which may be exploited as aid to the interpretation of the imagery.

Some of the multichannel data was used to classify the coastal wetlands area in terms of vegetation types. An initial manual interpretation was followed by a machine-based classification of vegetation types. The digital computer-compatible tapes (CCTs) were prepared on the ERIM hybrid optical digital processor from the SAR data of the Marineland coastal zone. The CCTs were then analysed interactively on the Canada Centre of Remote Sensing (CCRS) Image Analysis System (CIAS) using a simple parallel piped classification scheme.

The results showed that, if geometric and radiometric corrections of digital SAR imagery were carried out, the resultant classification would correspond well with the ground truth. Interesting though these results are, it will be noted that this type of digital image processing based on the classification of images recorded at different wavelengths (which in this sense is similar to those of Landsat MSS) has little relevance to the single-channel SAR actually used in Seasat.

Other parts of the NASA Seasat/Marineland experiment were concerned with the analysis of the ocean waves recorded on the SAR imagery. For this purpose, CCRS developed a software package to perform geometric and radiometric corrections and two-dimensional transforms of radar imagery. From the analysis of the L-band imagery, it was concluded that, in case of Seasat, wavetrain movements would cause only minor defocussing problems at the 25 m resolution of the SAR system but might be more important at high resolutions. It was also found that, at Seasat look angles (maximum 23^{o}), the performance of the L-band radar could be severely reduced for the imaging of both waves and ships. Much of the rest of the experimental work was concerned with comparisons of imagery at different wavelengths and polarizations and,in the end, this was not really relevant to the operational Seasat SAR imagery. However, without doubt,the experience gained from these tests helped to decide the final rather simple form of single-channel imaging SAR which was used with Seasat.

CHAPTER V

SEASAT SAR DATA PROCESSING

5.1 Special Characteristics of Satellite-borne SARs

When a SAR is mounted on-board a satellite instead of an aircraft, a number of new problems appear which concern the flight design, the characteristics of the SAR itself and the subsequent processing. Many of these problems relate to the special parameters of the satellite orbit and their interaction with the SAR system. In particular, certain points have to be considered which are very different from those of an aircraft. The operating altitudes and the measured ranges are enormous - from 200 km to 1,000 km being typical for earth resources satellites. These can be compared with a maximum of 10 km for aircraft-borne SARs. The huge altitude also allows great areal coverage of the Earth, though this may be constrained by the orbital inclination which defines the latitudinal extent of the coverage. Also the speed of the satellite over the ground is 8 km per second which can be contrasted with the 300 km per hour (i.e. 5 km per minute) of an aircraft.

Additionally, there are certain other factors which must be taken into account in considering the problems of the data processing of satellite SAR data.

(i) As is the case with air-borne systems, the range to a target to the side of the satellite changes continually within the synthetic aperture. However, due to the very long effective synthetic aperture possible with satellite-borne radar systems, this change in range is appreciable, so that a requirement for larger and additional processing

facilities is almost certain to occur. However the synthetic aperture length can be sub-divided into separate sub-apertures, which results in substantial savings in the number of samples which have to be processed.

(ii) Furthermore, during the period in which the object is within the synthetic aperture, the rotation of the Earth will be moving the terrain objects themselves relative to the satellite so producing a transverse component of relative velocity. This produces a range-migration or range-walk which further shifts the frequency of the received signals by the Doppler effect (Barnett et al, 1978) and complicates the later ground processing.

(iii) Other difficulties arising from the extreme range over which a satellite-borne SAR has to operate, include the power requirements and the nature of the signals themselves. One solution which helps to overcome both of these difficulties is the halving of the pulse repetition frequency (PRF). This has the advantage of halving the mean transmitted power and the important consequence of halving the amount of data which has to be processed, though it also leads to the difficulties with a discontinuous aperture as already discussed.

(iv) The emitted signal is chirped, i.e. the signal is comparatively long with a frequency varying with time. Since unchirped pulses are used in most airborne systems this means that the signal returns from the seasat SAR must be processed in a special manner in order to extract and maintain the desired resolution.

(v) Lastly, an obvious point, but one which has considerable implications for the processing of the data, is that the SAR data has to be telemetered by radio link back to the ground. This means that a

special receiving station is necessary to receive and record the data. This also means that a signal film will never be available as the initial form of recording the satellite SAR data even if the subsequent processing is optical in character.

5.2 The Processing of Seasat SAR Imagery

While the general approaches to the processing of SAR data, e.g. whether it should be carried out by optical or digital techniques, have already been discussed in previous sections in Chapter 3, since the main concern of this dissertation is with Seasat SAR data, some more detailed discussion of the processing of Seasat data appears appropriate. As it is the first satellite-borne SAR data to which civilian users and researchers have had access, there has been much interest in developing new methods of processing, especially on the digital side. Therefore, an account of these will be made including certain proposals which have not in fact been realized but which nevertheless offer new approaches to SAR data processing.

5.2.1. Optical Processing of SAR Data Using the ERIM Tilted Plane System

As already mentioned, the main production systems for SAR data utilize optical methods. This is the case with Seasat also, the principal device used for the purpose being the ERIM tilted-plane optical processor.

The design and construction of this processor has been described in section 3.3.3. As applied by ERIM to process Seasat data, the satellite data is first recorded in digital form on high density digital tapes (HDDTs) at the receiving station. It is then transformed on to a signal film using a D/A convertor which drives a

precision graphic CRT display. From this the image can be transferred by projection via a recording camera on to photographic film. An electron beam recorder offers an alternative possibility. The signal film record is then placed in the input stage of the optical processor where it is illuminated with a collimated coherent beam. A precision film transport is used to move the signal film through the processor input plane at a selectable constant speed. In this way, an area encompassing the full range interval of recorded data and slightly more than the full azimuth length of a point target history may be processed.

The output film is exposed through an output aperture which covers an extended range interval (swath width) and an azimuth interval several resolution cells wide. The position of an image point on the output film is maintained over the output aperture by proper setting of the output film speed to match that of the input film. Output imagery can be provided in several forms - as a directly viewable projection display, as a film transparency or in the form of a digital tape. The processor system itself has a peak signal to noise ratio of 80 decibel (dB) or more and will readily accommodate the large dynamic range of the SAR imagery (> 50 dB).

The phase perturbation of the SAR signal caused by azimuth wander, range walk and range curvature is corrected in the processor. Azimuth wander is corrected by directing the processor input illuminating beam to an angle which shifts the azimuth signal spectrum to a prescribed fixed frequency location. Range walk is corrected by insertion of an optical phase function in the signal frequency plane which is equal but opposite in sign to that caused by range walk in the signal frequency spectrum. Range curvature correction is also accomplished

by using a compensating phase function in the signal frequency plane (Guignard, 1980).

Since the conversion from slant range to ground range was carried out inside the satellite to a first order only and since no higher term corrections are possible with this optical processor, residual scale variations on the image in the cross-track direction of about $\pm$ 5% relative to the centre of the swath are to be expected. This was indeed encountered on the two images processed by this system, which have been used by the author in the present tests.

The optical corrector elements of the processor can be set to certain mean values of the altitude, attitude etc on the basis that they are fixed for the processing of a whole path. This is termed survey processing. The images produced by this method will, of course, be slightly degraded when compared with the nominal performance of the SAR since the actual orbital parameters will vary somewhat from the average values used in the correlator. To quantify the extent of this defocussing is almost impossible since it varies from image to image but, as a very rough approximation, one could assume a ground resolution of about 50 m x 50 m (Lodge, 1980b).

An alternative procedure is termed precision processing. Here the optical correlator settings are maintained for only certain limited portions of the input SAR data. Typically, this amounts to along-track section of the orbit about 500 km to 700 km long. The settings of the corrector elements are made by operator observation of the signal frequency spectrum in the processor frequency plane and on the resultant image quality as judged at the processor output image plane. It was hoped that the precision processed images would have a spatial resolution limited more by the SAR system itself than by

the processing method; that is of the order of 10 m in the azimuth direction (Ra) and 25 m in ground range (Rr).

All the survey-processed images and most of the precision-processed images are single-look images, that is to say the optical processor used the full extent of the available synthetic aperture. The film imagery is generated in 28 km wide segments with a very extended azimuth length. This means that, four such swaths have to be generated to permit coverage of the 100 km swath of the Seasat SAR. The material used in the present study was 70 mm wide photographic film which gave a nominal scale of 1:685,000. The imagery was supplied in 6 m lengths which gave 4,000 km of coverage in the along-track direction. The output rate of the ERIM system amounts to 12,000 km of 100 km swath width in one day (Guignard, 1980).

5.2.2. Proposed Optical Processing of Seasat SAR Data Using the Imperial College System

Barnett et al (1978) of the Blackett Laboratory, Imperial College, London, put forward a proposal for an optical processor to be built specially for the processing of Seasat SAR imagery. A schematic diagram of this system is shown in Fig.5.1. The signal film is illuminated by a laser beam. The collimator lens, focussed on the virtual holographic image in the range section, produces a convergent beam in the azimuth section due to the long focal length of the signal in this direction. The image then passes through an objective lens which brings the reconstructed image to a focus and it can be recorded on the output film. As is the case with other optical processors, a cylindrical lens is used to bring the range and azimuth sections into the same plane of focus.

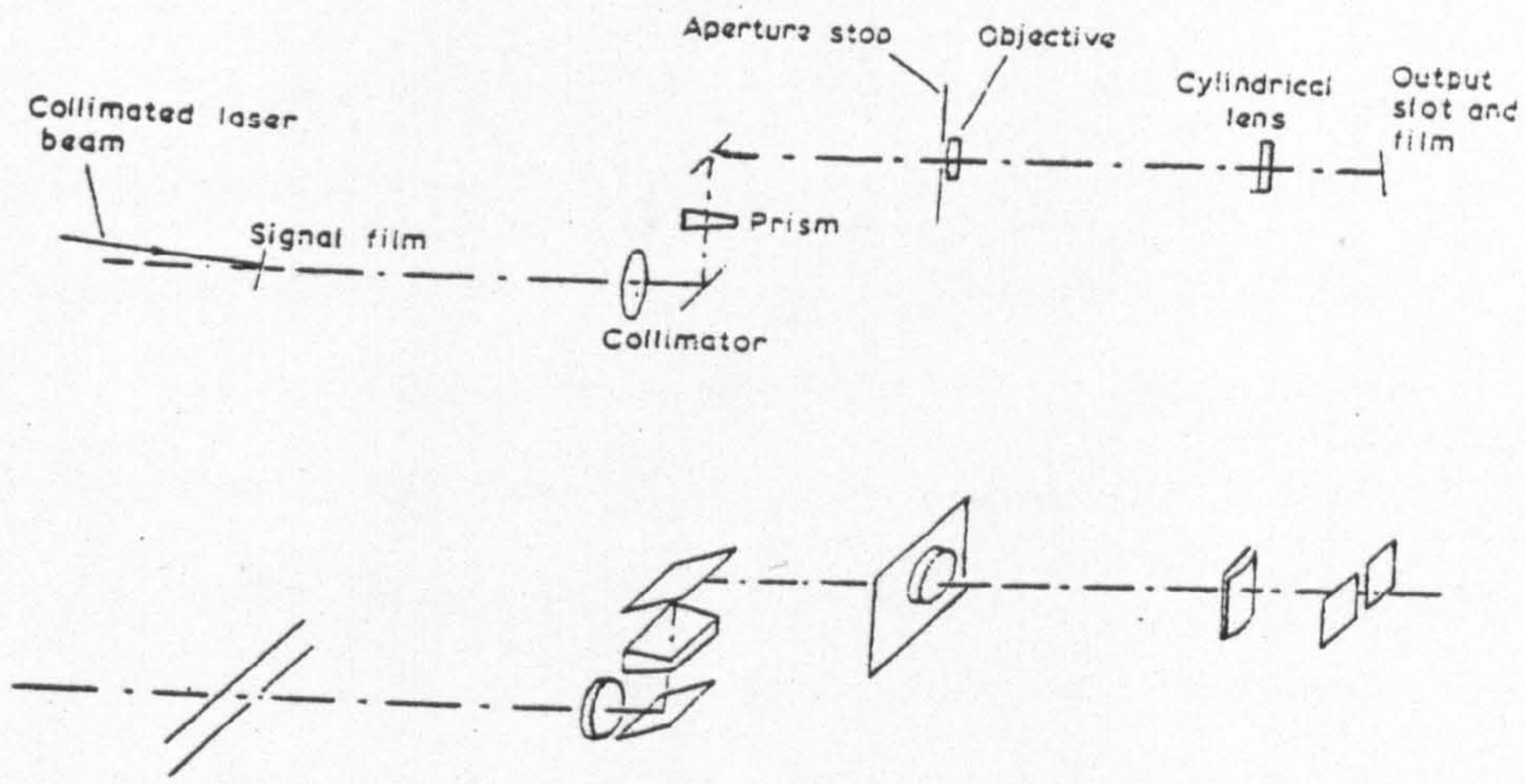


Fig 5.1 Schematic diagram for Barnett's et al optical processor (Barnett et al, 1978).

In this proposed system, the Doppler shifts in the signal histories induced by the Earth's rotation producing range-walk would be reduced to an acceptable level by offsetting the demodulating frequencies (to get rid of the carrier frequency) by an appropriate amount. Also the effect of curvature of range histories has to be catered for. Although this effect is also present in airborne systems, it is negligible because the total depth of curvature is usually less than one range resolution element. However, in spaceborne systems, since the extent of curvature is several range resolution elements, this effect must be rectified. In the Imperial College design, this would have been solved by placing a weak dispersing prism in the path of the beam after the collimator. This would neutralize the curvature of the range histories since the curvature of the prism would be equal and opposite to the curvature of the range histories.

The speckle effect, normally reduced by a "multi-look" technique (Fig.5.2), would have been tackled by dividing the length of the signal film corresponding to one synthetic aperture by four. Each sub-section would then be exposed onto the same strip of the output film, thus reducing the occurrence of the speckle effect by an appreciable amount.

The system as a whole was intended to operate under the control of a mini-computer which would allow the input of information about the attitude and position of the satellite which would be used to control the mechanical settings of tilt and focus. Residual metrical distortions on the output film would also have been rectified by optical elements controlled by the mini-computer and then produced as a photographic transparency or digitized and stored on tape.

There were certain other special features to this proposed

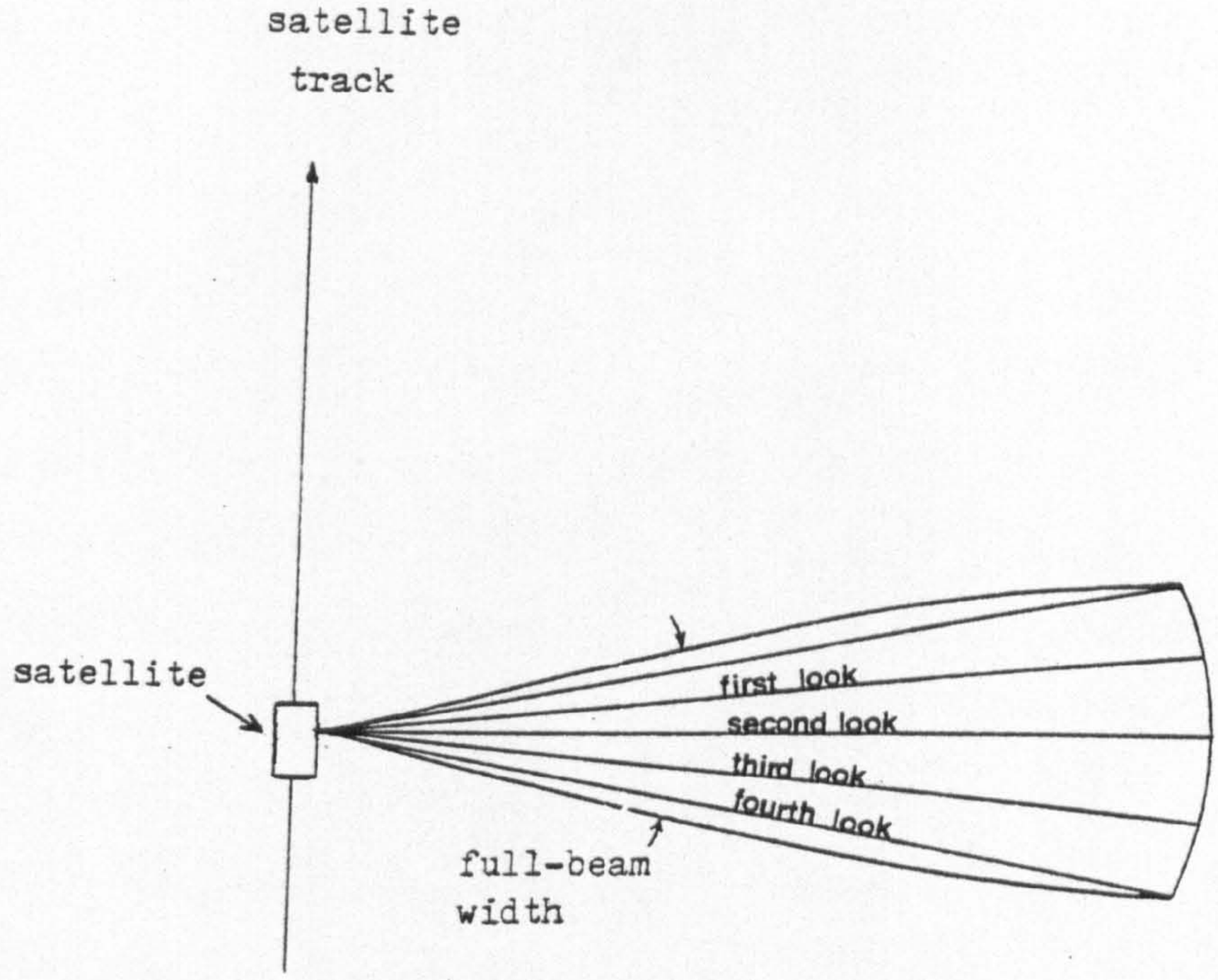


Fig 5.2 Multi-look addition technique for reducing image speckle in a coherent radar system.

system. Like other optical processors, it is of a comparatively simple design with very high throughput. It was proposed that the processing of the SAR signals would be carried out in near real-time as they arrived from the spacecraft without first recording them on magnetic tape. To achieve this, the incoming data would use a flying spot CRT to convert the record on to the signal film. This meant that a short interval would have to elapse for the photographic processing of the signal film.

The Imperial College proposal was evaluated by the European Space Agency (ESA) against several competing proposals based on digital processing. If accepted, the Imperial College proposal would have resulted in the construction of a sophisticated optical processor which would have given Europe a facility similar to that of the ERIM tilted-plane optical processor with which it has many features in common. In the end, it was decided not to implement the proposal for this system and to sub-contract the bulk production of the Seasat imagery of Europe received at Oakhanger to the existing optical facility at ERIM. The Agency (ESA) also decided to concentrate its remaining resources on the development of digital SAR processing which appeared to be a more worthwhile investment for the future. As a result, ESA has supported several different groups who have conducted the digital processing of Seasat SAR data on a research rather than a production basis.

5.2.3 Digital Processing of Seasat SAR Data

As previously discussed, it is possible to process SAR imagery in a digital computer. However, it is a very slow and costly approach for the processing of satellite imagery which exhibits such

high azimuth compression ratios (200:1 in the case of the Seasat SAR imagery).

The SAR data can be arranged as a two-dimensional array in which the range samples are stored as rows and the azimuth samples as columns. Processing is carried out one row at a time after which data is accessed by columns. Normally a SAR data set is so large that the array has to be partitioned into blocks which need to be stored on some secondary form of data storage such as magnetic disc. This means that the data in the computer memory has to be continuously passed to and from the disc with a consequent loss in speed. In addition to this, the arithmetic processing rates of general-purpose digital computers are too slow to allow efficient data processing of satellite-derived SAR data. The processing times for a single scene of a limited area may measure in hours if not days (Bayma et al, 1977). Therefore general-purpose digital computers can be ruled out of consideration for the production processing of orbital SAR imagery.

To speed up the processing of Seasat SAR data even on an experimental basis, some very large and fast mainframe machines such as the IBM 360/75 at the Jet Propulsion Laboratories (JPL) and the Amdahl V-6 of DFVLR have been used for the purpose. However, dedicated systems based on smaller mini-computers have also been employed to a considerable extent, their more limited processing power being augmented by devices such as array processors. Examples are the Prime 400 mini-computer used at the Royal Aircraft Establishment (RAE), the SEL 35/55 at JPL, the Interdata 8/32 at DFVLR and the Data General Nova at the Bendix Laboratories, all of which utilize an FPS AP-120 B array processor to handle the bulk of the SAR processing. As far as European users are concerned, the principal sources of

digitally processed Seasat SAR imagery are those produced by the RAE and DFVLR facilities. These will be discussed in the following sections.

5.2.3.1. Digital Processing Using the RAE Facility

The Royal Aircraft Establishment (RAE) of the Ministry of Defence has its Experimental SAR Processing Facility (ESPF) located at Farnborough. The software for the system has been written in collaboration with System Designers Ltd. (SDL), Frimley, Surrey. Fig.5.3 gives an overview of the Experimental SAR Processing Facility. Up to four independent looks are available from this processing system, each look being processed separately. Within the limits of the radar system, the azimuth resolution may be selected at will. Also the image may be arbitrarily aligned with the satellite line of sight. Adjacent images may be exactly registered with one another.

The actual processing carried out by the Facility is divided into two stages:-

Stage one: This takes place at the Oakhanger receiving station where a Prime 300 computer is located. The SAR data record (SDR) is collected for each pass by the large dish antenna and recorded on high density digital tapes (HDDTs). Using the Prime Computer, the data is first processed and transferred from the HDDTs frame by frame (typically 50 km x 50 km) on to a 300 M byte disc. When data transfer to this disc is completed, the Prime 300 is then used to transfer the data from the disc onto 1,600 b.p.i. computer compatible tapes (CCTs) which can then be used as input to the next stage.

Stage two: This utilizes a Prime 400 equipped with an FPS AP-120 B array processor situated at Farnborough. The input data from the previous stage is first checked and corrected for missing data and

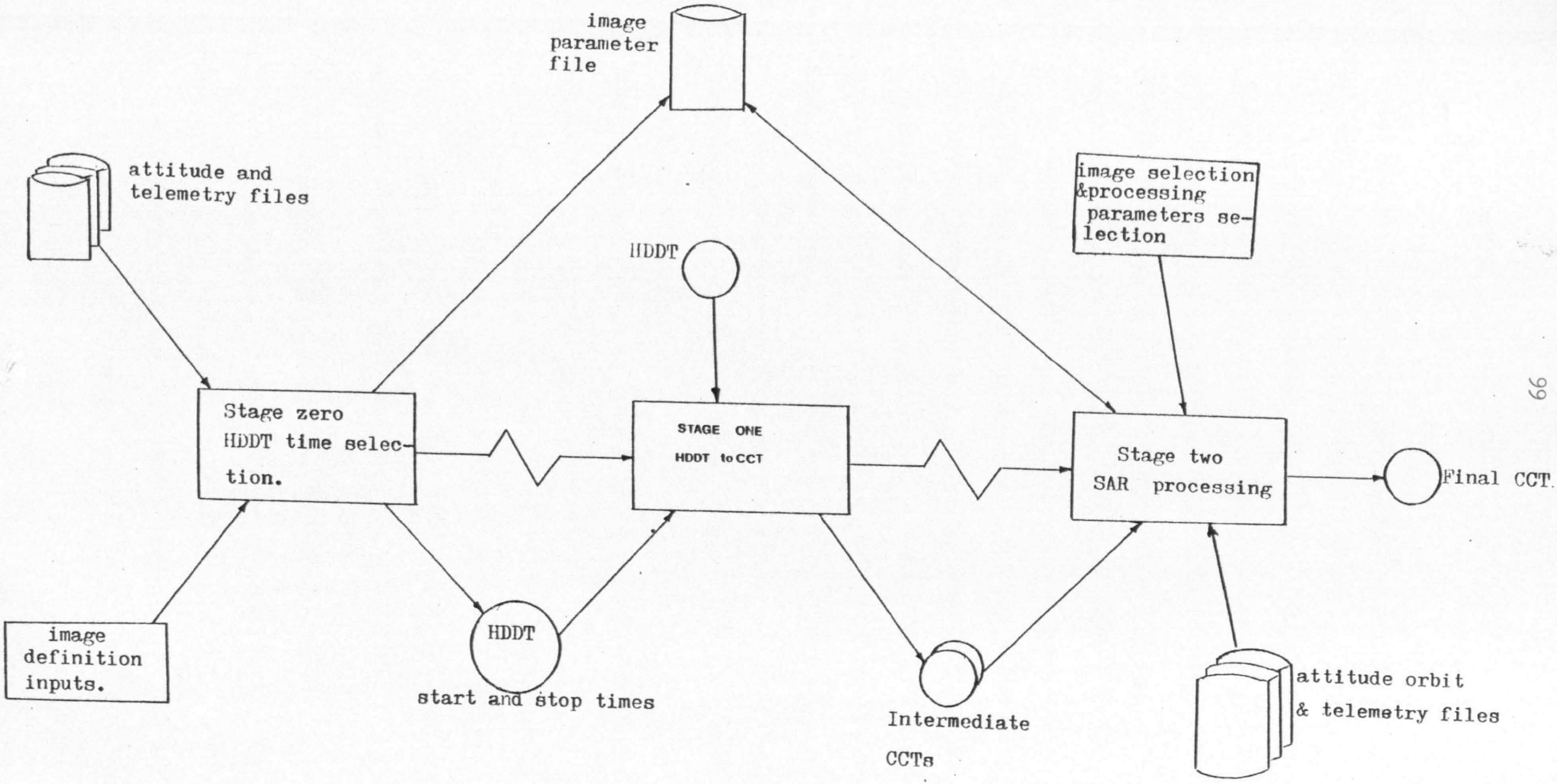


Fig 5.3 RAE/SDL Experimental SAR Processing Facility (ESPF) block diagram.

random errors. Next, the range compression is carried out by Fourier Transform techniques using the array processor. The data is then filtered digitally to reduce the Doppler bandwidth and the PRF - a reduction factor of 3:1 was reported by Corr and Haskell (1978). Next, the azimuth compression (coherent summation) is carried out together with the necessary radiometric and geometric corrections (see Fig.3.10). Since the Seasat data sets are too large to be held in the core memory of the computer, only a section at a time can be processed. These are systematically transferred to and from the disc into the computer. When azimuth compression is finished, the image pixels are sorted and a final image tape is produced in a line-by-line format. This may be supplied direct to users or it may be produced in hard copy form on film or paper using the RAE's Linoscan 204D raster drum plotter.

Generally speaking, the processing procedure is similar to that discussed in Section 3.4 in that range compression is performed first since the azimuth focus is largely range dependent. Furthermore, this procedure also allows simple methods of range curvature correction to be carried out (Ausherman, 1977).

A typical Seasat SAR product from the RAE Facility is a 50 km x 50 km image with 25 m resolution produced in about 30 hours (Lodge, 1980b) Thus the Facility has a very low throughput and must be viewed mainly as a research tool (Guignard, 1980).

5.2.3.2. Digital Processing using the DFVLR Facility

This facility, located in Oberpfaffenhofen in West Germany, operates along the same general lines as that at RAE. It utilises a very fast 32-bit Interdata 8/32 mini-computer equipped with an FPS array processor. The software was originally developed by Macdonald

Detweiler Associates (MDA) in Canada for use on similar Interdata machines both in MDA itself and at the Communication Research Centre (CRC). It was then purchased by ESA as a proven off-the-shelf item which would allow immediate implementation of Seasat SAR digital processing. A typical product with 40 km x 40 km coverage, 4 looks and 25 m resolution is produced in 20 hours.

Table 5.1 lists the main processor performance of the three operational correlators discussed in this section.

5.2.3.3. Proposed Digital Processing of Seasat SAR Data Using the Marconi Research Laboratories System

This proposal for a digital processor (Fig.5.4) was put forward by the Marconi Research Laboratories, Great Baddow, U.K. and was designed specifically for the processing of Seasat SAR data (Ellis, 1978). In this system, the radar information would have been processed in frames of approximately 100 x 100 km. The data required for one of these frames is selected from the HDDR and is transferred to a 300 M byte disc using a special hardware controller and formatter. Once on the disc, the data can be read down in blocks of an appropriate size to a computer compatible tape (CCT) making available the unprocessed SAR data for off-line processing. All of this activity would take place at the ground receiving station.

The SAR data written on the disc pack may also be transferred manually to a second 300 M byte disc drive attached to the SAR processor itself. Data is then read from this 300 M byte disc drive and range compression is carried out using the Fast Fourier Transform (FFT) method of correlation with the reference chirped waveform. This is followed by the azimuth processing or compression, again carried out

Table 5.1 Main Processor Performance (modified from Earthnet Review, 1979)

Processor origin	ERIM	RAE	MDA-DFVLR
Location	Michigan (USA)	Farnborough (U.K.)	Vancouver (Canada) Oberpfaffenhofen (W.Germany)
Resolution	20x60m (1-Look) 20x27m (3-Look)	25x25m (1-Look) 20x20m (4-Look)	25x25m (4-Look)
Number of Looks	up to 3	up to 4	4
Swath width	100 km (in segments 4x25 km)	50 km	36 km
Swath length	Survey: up to 4000 km Precision: up to 700 km	50 km	40 km
Output	Film or print	CCT, film or print	CCT, film or print
Interpixel distance	-	20 or 12 m	12.5 m

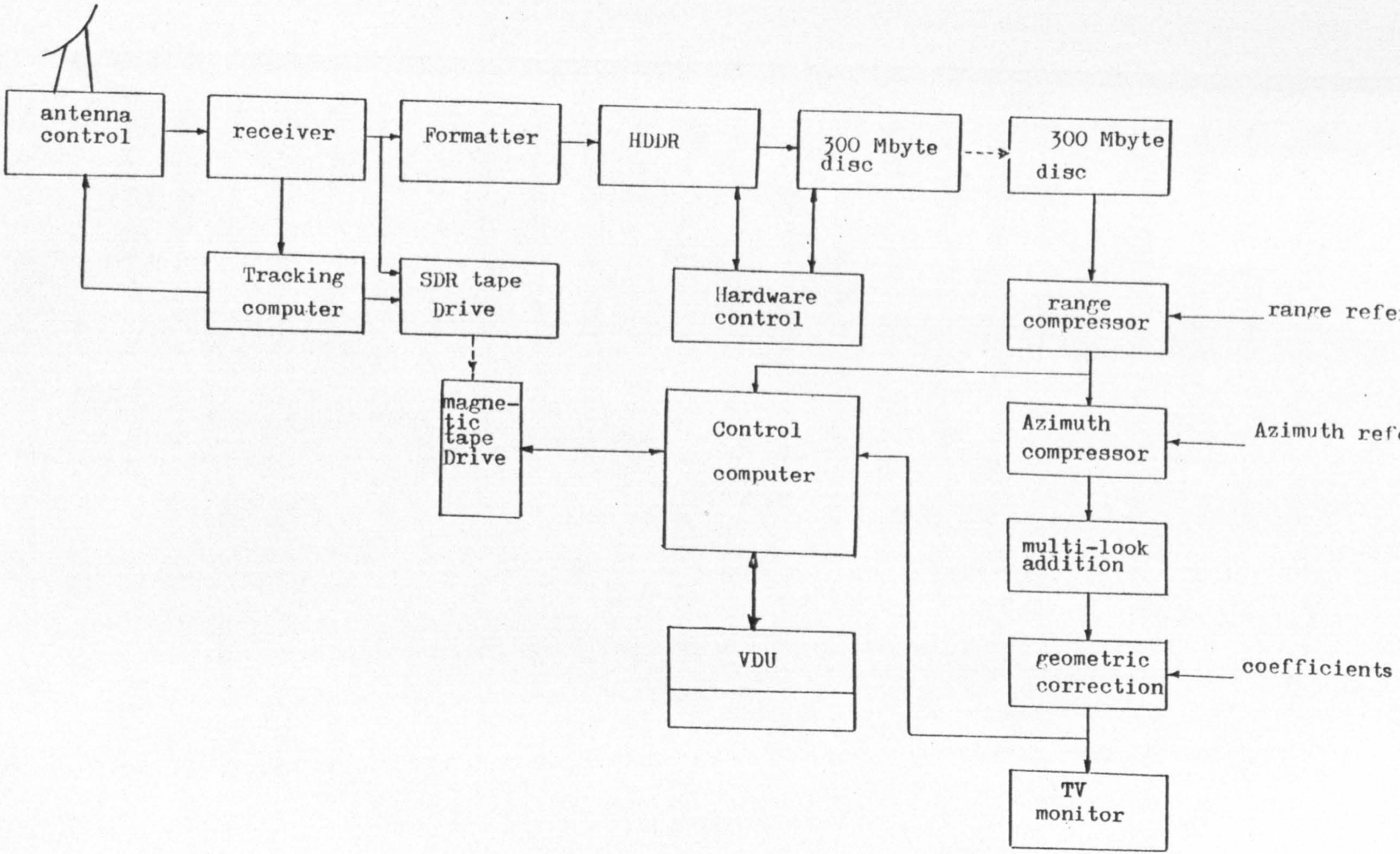


Fig 5.4 Marconi's Labs digital Seasat Processor block diagram (Ellis, 1978).

using the FFT technique. Both of these operations would be carried out using specially-built hardware processors operating under the control of a mini-computer which also generates the range and azimuth reference patterns and the geometric correction coefficients.

Range-walk would have been taken out in two operations. The coarse linear range-walk is removed when reading down from the 300 M byte disc, while the remaining linear and non-linear range-walk is allowed for in the azimuth processor. During the azimuth FFT operation, the data would have been filtered into four sets of data representing four individual looks at a section of the ground. These data are then combined in a multi-look addition process after which the data would be geometrically corrected, formatted and recorded on CCTs. A typical frame size (100 x 100 km) was estimated to take 12 hours to process. This can be compared with the 20 to 30 hours required to process much smaller areas (40 x 40 km and 50 x 50 km respectively) with the RAE and DFVLR processors based on large general purpose mini-computers equipped with array processors.

The Marconi hardware-based system was also evaluated by ESA against the other digital processors proposed for its programme. Eventually it was decided not to implement this proposal and instead it was decided to support the development of the more flexible RAE and DFVLR systems. The relative financial implications of these various rival projects are not known to the present author. However, obviously the Marconi proposal would have resulted in a more production-oriented system than the present RAE and DFVLR systems which are essentially research and development systems.

5.2.3.4. Proposed Digital Processing of Seasat SAR Data Using the French Thomson-CSF Digital Processor

Within the Seasat project, the French firm Thomson-CSF and

its affiliate CIMSA had conducted a study of several advanced systems for processing the SAR imagery (Le Gal and Gilbert, 1978). All of these were to be built around the Mitra 125 mini-computer and the PROPAL II associate parallel processor. One of these configurations (Fig.5.5a) was designed to perform processing of SAR data at a rate of two to five 100 x 100 km images per day.

Processing would have begun with the transfer of the raw SAR data of the images requiring processing from the HDDT on to an intermediate magnetic tape (CCT). In turn, these would have been transferred onto a 300 M byte magnetic disc. Processing parameters were then to be calculated in the computer from a sample of the data and the auxiliary data record from Seasat. Next, radiometric corrections would be applied followed by range compression consisting of the Fourier Transform, filtering, frequency shifts and an inverse Fourier Transform. This would be followed by azimuth processing using a direct convolution algorithm similar to that outlined in Section 3.4. Multi-look addition and geometric corrections in azimuth (including range-walk and Earth Curvature) would then have been performed. The final step would deal with the necessary geometric corrections in the range direction and the output of the final image.

The proposed use of the PROPAL II parallel processor controlled by the Mitra mini-computer is in a sense not dissimilar to that of the array processors attached to mini-computers used by JPL, RAE, DFVLR, MDA etc. The purpose in each case is to lighten the highly repetitive computations involved in the processing of SAR data. However, while the array processors speed this up by the use of pipe-lined processors specifically designed to work on arrays of data, the PROPAL II is a truly parallel processor with an array of processors

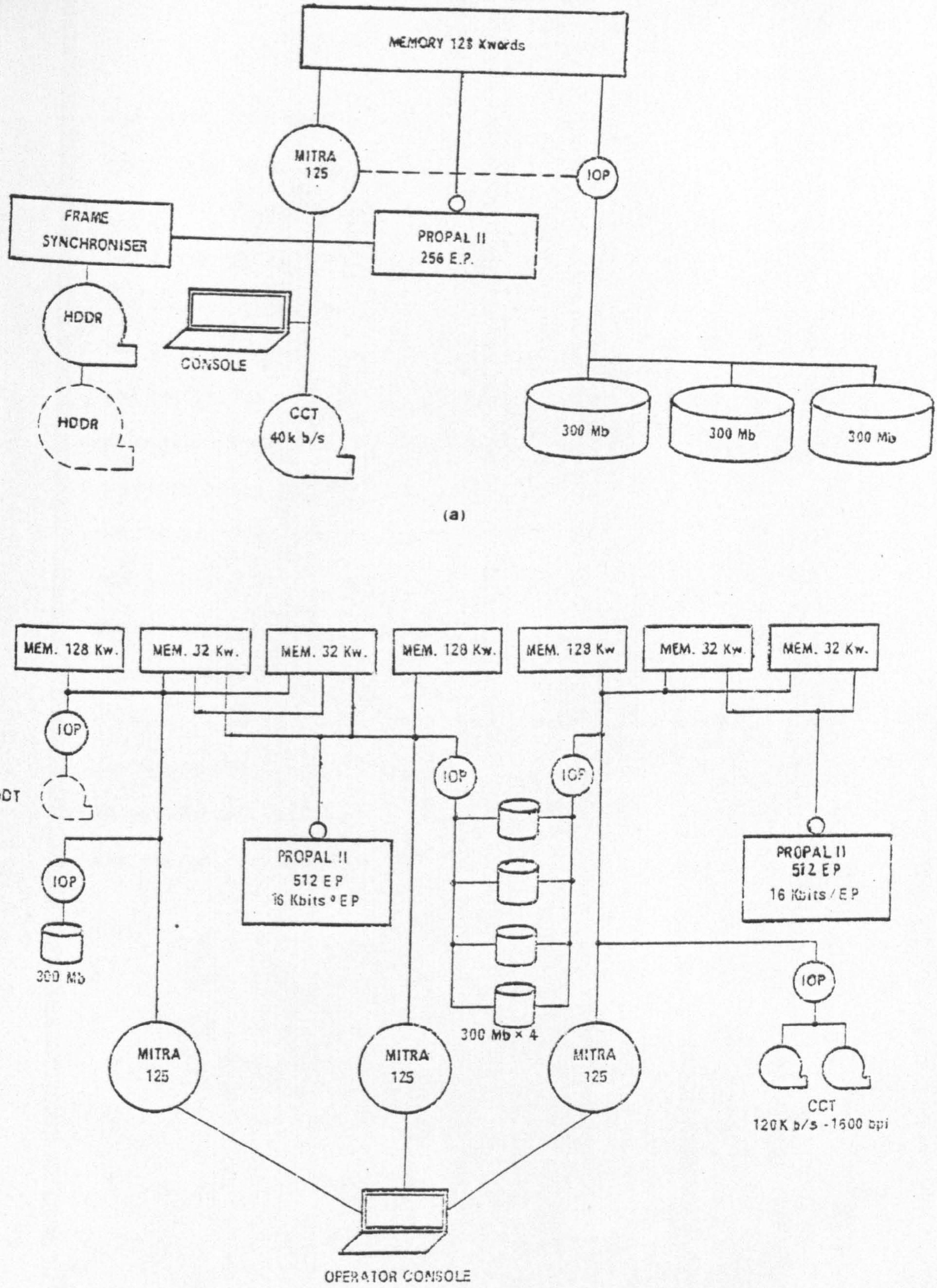


Fig 5.5 Configuration of Thomson-CSF Seasat digital processors (Le Gal et al, 1978).

and associated memories, each simultaneously executing the same operations under the control of the mini-computer.

The estimated speed of processing, amounting to 5 to 8 hours per 100 x 100 km image, would have been somewhat faster than that of the Marconi proposal. An even more ambitious proposal (Fig.5.5b) along the same lines would have utilized three Mitra 125 control computers and two PROPAL II parallel processors, each consisting of 512 separate elementary processors, instead of the single 128K 16-bit word Mitra computer and single PROPAL II device equipped with 256 elementary processors used in the first proposal. This second massive system was proposed for the processing of all the daily SAR data received from Seasat within 24 hours of its reception. This represents a daily throughput of 40 images, each of 100 x 100 km.

After being evaluated by ESA, it was decided not to implement these proposals. As mentioned above, the decision was made to concentrate the available resources on the less expensive DFVLR and RAE digital systems.

CHAPTER VI

GEOMETRICAL ANALYSIS OF SIDE-LOOKING RADAR (SLR) IMAGERY

6.1 Introduction

The basic geometry of SLR systems has been studied by a large number of investigators, especially in the early seventies when the imagery from these systems first started to become available to non-military users. The results of these rather varied studies, which concerned both the real-aperture and synthetic-aperture types of side-looking radar, have been published by Graham (1970, 1974); Hockeborn (1971); Konecny and Derenyi (1966); Laprade (1963, 1969, 1970); Leberl (1970, 1972) and Norvelle (1972).

From the geometric point of view, side-looking radar (SLR) imagery is rather complex, since it is being produced through the continuous interaction of mechanical, optical and electronic components functioning in a dynamic mode of operation. The geometric aspects of SLR imagery will first be investigated from a theoretical point of view. However, this will inevitably have its limitations, since the actual instrumentation will almost certainly operate rather differently in practice to what it was originally designed to do. Hence the matter of calibration (or the lack of it) will figure prominently in any discussion of the geometric characteristics of SLR system. To calibrate the system, i.e. to see how it actually performs as distinct from what it was designed to do, one or other of two different approaches can be implemented.

(i) System Technique: Here the whole instrument unit is subjected to a thorough calibration both after manufacture and during service. The application of this technique (which is equivalent to that followed

with photogrammetric cameras) should present a clear picture of the performance and distortion characteristics of the entire imaging system. On this basis, the geometrical rectification of the imagery produced by that particular system is possible without any additional control data, provided it can be combined with the continuous measurement and recording of the position and height and all motions of the platform while it is being operated. Such a procedure appears especially suitable for the rectification of a large volume of imagery such as that generated from a spacecraft equipped with a wide-swath imaging system.

Unfortunately, however, this is also the most complex and expensive technique to implement in practice. Consequently the implementation of a systematic calibration both after manufacture and regularly during service (corresponding to that of a camera calibration) does not seem to take place in practice. The reasons for this omission may be the complexity and expense of the procedure but an important factor is the lack of demand for it from the main body of military users. Furthermore, the development of techniques for calibrating and monitoring the actual performance of the SLR system does not appear to have reached a satisfactory operational level. Konecny (1976) comments that radar system designers are not in a position, nor do they seem to be concerned to ensure that the system operation and the actual platform orientation parameters should be measured and recorded to the desired accuracy. Therefore any analysis of the SLR imagery itself or the devising of techniques for its geometric rectification can only be carried out in the knowledge that the information normally available from regular systematic calibration (as occurs in the case of photogrammetric cameras) will not be available to the users of SLR imagery.

(ii) Image Testing Technique: As an alternative, representative samples of radar imagery may be subjected to measurement and analysis using a test field of measured control points as given information. Obviously, this approach has its limitations but nearly all reported work on the calibration and possible utilization of SLR imagery for mapping applications has employed this technique.

6.2 The Image Coordinate System of SLR

The basic geometry and principle of operation of side-looking radar (SLR) have been outlined in Chapter II using a simple descriptive and graphical approach. In this section, more emphasis will be placed on the mathematical relations of these systems with a view to devising and implementing numerical and analytical techniques for rectification of the imagery.

With a side-looking radar, one-dimensional imaging takes place, normally using a C.R.T. as the image generating device. Instead of directions (as in aerial photography), double transit times of the radar pulse are measured for a whole set of reflectors located along a very narrow beam (approximating to a line) on the ground (Fig. 6.1a). Furthermore, the imaging takes place sequentially both along such a beam and along successive lines. This can be contrasted with the simultaneous exposure of all terrain points in a single image using a conventional frame camera. Essentially, the measured transit times are converted to slant ranges. For a given sensor altitude, the presentation of the image may also be made in the form of ground distances.

If Δt_i represents the double transit time towards a particular object point P with coordinates (X_p, Y_p) at the time t_i, the value of Δt_i is very small due to the very high speed of electromagnetic energy

(c = 300,000 km./sec). Thus, for an object located 6 km from the SLR set, a 12 km double transit distance will be completed in $\frac{1}{25,000}$ second. In practice, it is most unlikely that a measurable change in the orientation parameters can take place within such a short period. Thus, the orientation matrix and the coordinates of the exposure station can be considered to be constant for all points lying along a single line on the image with the same x-coordinate. For the computations of slant range (S) and ground range (Sg) in SLR systems, the following equations are used (Fig. 6.1a):

$$S = (Sg^2 + H^2)^{\frac{1}{2}} = \frac{c\Delta t_i}{2} \quad ; \quad \text{and}$$

$$Sg = (S^2 - H^2)^{\frac{1}{2}} = \left[\left(\frac{c\Delta t_i}{2}\right)^2 - H^2 \right]^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots 6.1$$

where H = flying height;

c = velocity of microwave propagation; and

t_i = time between transmission and reception of a radar pulse.

The actual image coordinate y_p in the cross-track direction corresponding to the slant range S to a point P on a <u>slant range presentation</u> is given by:

$$y_p = m_y \frac{c\,\Delta\, t_i}{2} - m_y \frac{c.e}{2} \quad \ldots\ldots\ldots\ldots\ldots\ldots 6.2$$

where m_y is the value of the image scale in the cross-track direction; and

e is the sweep delay constant whose value (in μs) should be known a priori as an inner orientation parameter.

<u>For ground range presentation</u>, and considering imaging over flat ground, one may write:-

$$y'_p = m_y \left[\left(\frac{c.\Delta t_i}{2} \right)^2 - H^2 \right]^{\frac{1}{2}} - m_y \frac{c.e}{2} \quad \ldots\ldots\ldots 6.3$$

where H is the nominal platform height above terrain.

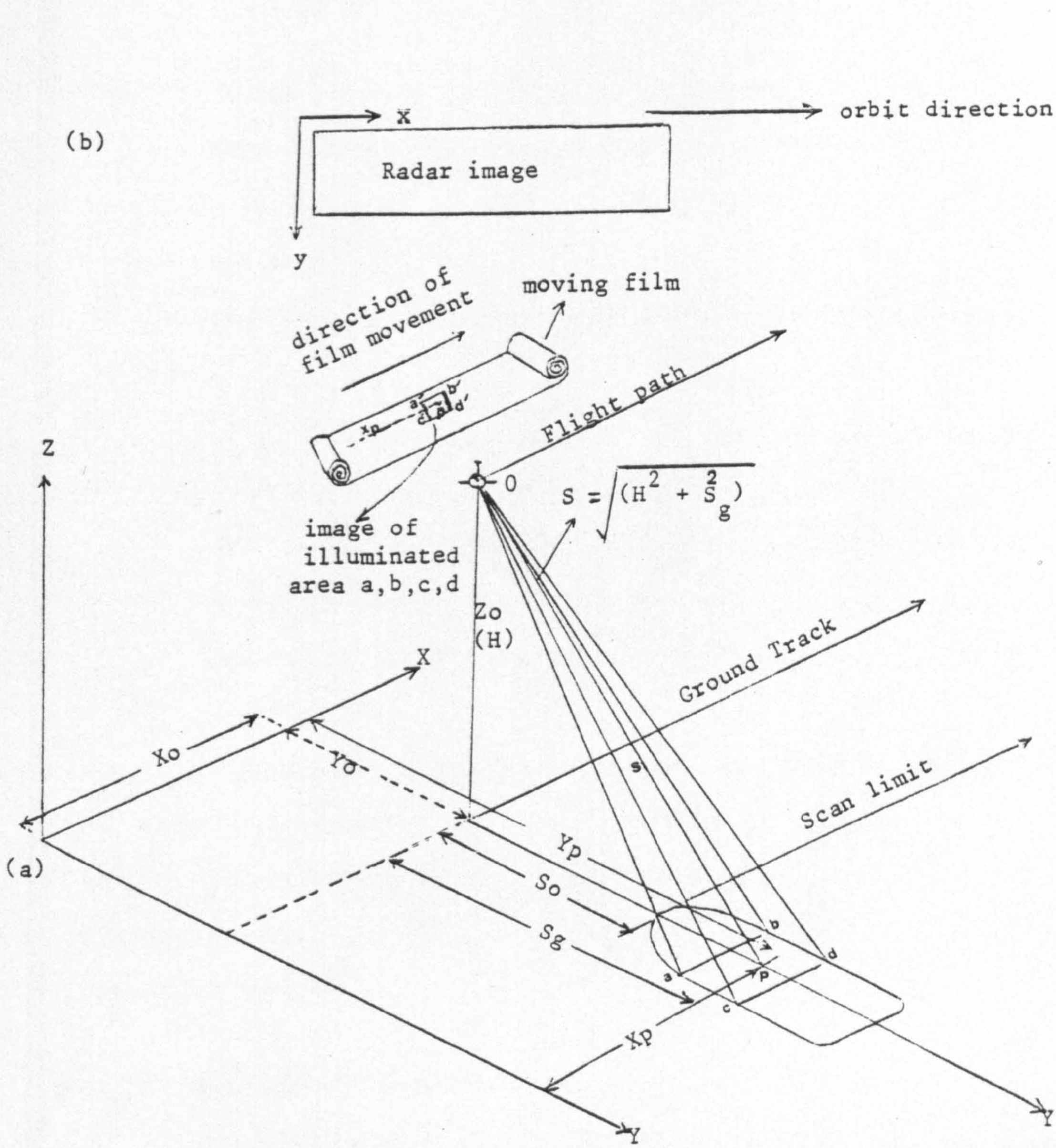


Fig 6.1 Geometry and coordinate system for a SLR image of a flat terrain.

In the <u>x-(along-track) direction</u>, the x-coordinate of an image point P is a function of the time t_j at which the object point P has been imaged. If t_{jo} is the time for which x = o, and t_{jP} the time at which $x = x_P$, then

$$x_P = \int_{t_j = t_{jo}}^{t_j = t_{jP}} \frac{dv}{dt_j}\, dt_j \simeq (t_{jP} - t_{jo}).v; \quad \ldots\ldots 6.4$$

where v is the velocity of the recording film movement whose value must be known from inner orientation parameters; and

dt_j is the usual variable of integration of the given integral function.

At the same time, x_P is a function of the carrying vehicle's velocity V so that:

$$x = m_x V\,(t_{jP} - t_{jo}); \quad \ldots\ldots 6.5$$

where m_x is the scale factor in the along-track direction which again must be known a priori as an inner orientation parameter; and V is obtained either from the system velocity sensor (such as an inertial navigation system or a Doppler radar) or a nominal value is assumed.

Thus a rectangular image coordinate system can be defined in the radar image such that the x-image coordinate points lie in the azimuth direction (i.e. in the flight or orbital direction) and the y-image coordinate in the direction normal to this (i.e. in the range direction) (see Fig. 6.1b).

6.2.1 <u>Inner Orientation</u>

The term inner orientation refers to the definition and determination of certain quantities, which must be known so that the image recorded by the SLR system may later be reconstructed either

geometrically or analytically. In conventional photogrammetry, these quantities are (i) the principal point, which is the orthogonal projection of the inner perspective centre on the image plane; (ii) the principal distance, which is the distance from the interior projection centre to the image plane; and (iii) the lens distortion characteristics of the taking camera. In an ideal geometrical system, the bundle of rays on the image side of the projective centre should be congruent to those on the object side. In reality, however, the ray path becomes distorted to a certain extent while travelling from the exterior to the interior projection centre. Several different calibration procedures are available to photogrammetrists to ascertain the occurrence and the extent of this distortion in frame cameras.

However, for a dynamic imaging device such as a side-looking radar, completely different considerations come into play. The inner orientation parameters which need to be considered are as follows:-

(a) Timing and range marks should appear which relate the x- and y-coordinates to the time t_j and the transit time Δt_j. The timing marks are generated on the image in the x-direction by the SLR system's crystal-controlled oscillator which acts as an electronic clock and generates small ticks on the image at pre-determined time intervals. Another corresponding set of marks known as range marks are also generated in the y-direction by the same oscillator. Both types of marks are recorded during the in-flight generation of the image and may be used later to check on the system's internal errors.

(b) The sweep delay constant (e) has already been defined and its value should be known as part of the inner orientation. Its value (in microseconds) should normally be determined during manufacture and checked periodically during its use in the field. It can then be

supplied to the user of SLR imagery.

(c) Where applicable, constant cathode ray tube deflection errors should also be determined. The voltage applied to the deflection plates of the CRT should ensure that the flying spot which writes the image moves across the face of the tube at a constant speed if a slant range presentation is required or, alternatively, at another pre-defined speed if a ground range presentation is required. Any deviation of the voltage from the required value will give rise to an error in the range direction. If no C.R.T. is used for imaging, this effect will of course be completely absent.

(d) The scale factors m_x and m_y, should be determined exactly for both x and y image directions. Since the scale of an SLR image along the y-axis (m_y) will change due to the slant range projection, only an average scale number can be defined. If a ground range presentation is made, then a more accurate determination of the scale number should be possible. This is the width of the imaged ground strip in terrain terms divided by the width of its record on film.

In the x-direction, the scale m_x may be viewed as being the ratio between the platform velocity and the film velocity. A constant value of this parameter depends heavily on good synchronization between the velocity of the film motion and the ground speed of the actual platform. A check for this is carried out using items of parameter (a) above.

(e) The number of ground resolution elements contained in a single scanned line should match the number of resolution elements which can be displayed on the C.R.T. The resolution of the actual film used to record the image should also match these requirements. A typical high quality C.R.T. contains 6,000 points in a line of 10 cm length

corresponding to a resolution of 30 lines/mm. It is of course impossible to incorporate a bar target with a range of low and high contrast targets for the measurement of SLR image resolution as can be done in camera calibration procedures. Instead, the resolution can only be checked by simple post-flight procedures based on measurement of objects of known dimensions.

(f) The calibration of the orientation of the antenna and its mount with respect to the platform reference system may need to be carried out. This should be done when the actual antenna is being mounted or attached to the platform and may need to be checked from time to time during the period of operational use of the system.

6.2.2 Effects of Errors in Inner Orientation Parameters on SLR Imagery

Due to certain imperfections in the radar equipment and its operation, the values of certain of the parameters of inner orientation discussed in the previous section will not always keep to their calibration values. In particular, the following errors in these parameters need to be considered.

(i) The time-basis: The sweep delay e could be erroneous by Δe. This gives rise to an error in the image position in the y-direction of

$$\Delta y = m_y (\Delta e) \frac{c}{2} \qquad \text{.....6.6}$$

(ii) The voltages applied to the CRT deflection coils should be such as to make the flying spot move with a constant speed (if a slant range presentation is required) or some other given speed (e.g. if a ground range presentation is required). Any deviation in the required voltage causes a constant error (Δq) in the flying spot speed q which leads to an error Δy in the image position produced on the C.R.T given by

$$\Delta y = - m_y \left(\frac{y}{q^2} \right) (\Delta q) \left(\frac{c}{2} \right) \qquad \text{.....6.7}$$

Furthermore, the C.R.T.-screen normally has a curved surface so that an image produced by direct photography of the screen results in geometric errors. These errors have been described by Wong (1968) as:

$$\left.\begin{aligned} \Delta x &= -y(a_1|r| + a_2|r|^3 + \ldots\ldots) \\ \Delta y &= \ \ y(b_1 + b_1|r|^2 + \ldots\ldots) \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots 6.8$$

where $|r| = \frac{(y - y_{max})}{2}$ (see Fig.6.2); and

$a_1, a_2, \ldots a_n$; $b_1, b_2, \ldots b_n$ are constants.

(iii) Film Transport and Ground Speed: An error Δv in the film velocity resulting from inaccurate synchronization with the platform velocity gives an error Δx in the image x-coordinate. If this is constant value, then

$$\Delta x = x \frac{\Delta v}{v} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 6.9$$

where v is the velocity of the film movement. Further internal geometrical errors of a similar nature may also arise in the optical processing of the SAR data. For example, an along-track scale error (Δx) will also be generated if the image film velocity is not matched exactly to the velocity of the signal film in the optical processor. Film deformations may also occur in either the signal or image films. Again these are likely to result in an affine scale change (which can also be viewed as a linear change in x (i.e. Δx)). So the effect of all of these errors is likely to be substantially the same, whatever their source.

From the discussion above, the overall formula for the description of the errors of interior orientation can be set out as follows:-

$$\left.\begin{aligned} \Delta x)_{total} &= x.\frac{\Delta v}{v} - y(a_1|r| + a_2|r|^3 + \ldots) \\ \Delta y)_{total} &= m_y(\Delta e)\ \frac{c}{2} - y(m_y)(\Delta q)\ \frac{(c)}{2} + y(b_1 + b_2\ r^2 + \ldots) \end{aligned}\right\} \ldots 6.10$$

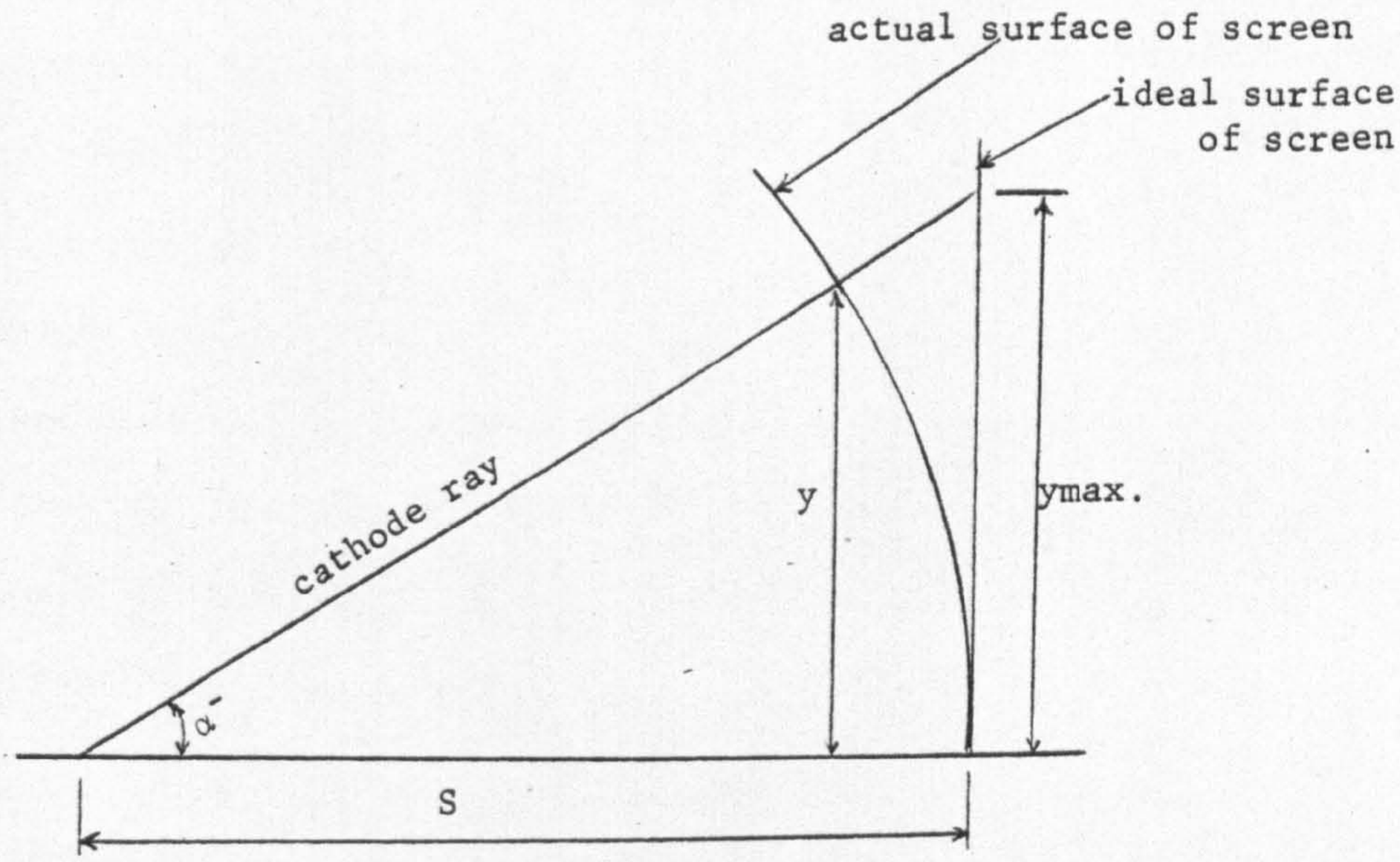


Fig. 6.2 CRT deflection errors

No quantitative information is, however, available to the present author to test the significance of this derived expression. It will be noted that these two equations can be simplified and written in the following general form:

$$\Delta x = a_o + a_1 x + a_2 y \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6.11$$

$$\Delta y = b_o + b_1 y$$

6.3 Exterior Orientation

So far, this discussion of the geometric characteristics of the image has omitted any consideration of the location of the antenna and its relationship with the points on the terrain which will be imaged by the side-looking radar system. Furthermore, any basic relationships established between the antenna and the terrain must take account of the fact that rotations and translations will be present during image formation, just as they are with photographic imaging systems. Thus, the effects of these deviations from the nominal or assumed position and direction of the antenna need to be analysed also.

Starting with conventional frame photography, the fundamental relations of analytical photogrammetry between the image coordinates (x_p, y_p) and the ground coordinates (X_p, Y_p, Z_p) of a particular terrain point (P) (Fig. 6.3) are given by the well-known formula:

$$\begin{bmatrix} x_p \\ y_p \\ -f \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} X_o - X_p \\ Y_o - Y_p \\ Z_o - Z_p \end{bmatrix} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad 6.12$$

where X_o, Y_o and Z_o are the coordinates of the exposure station; and

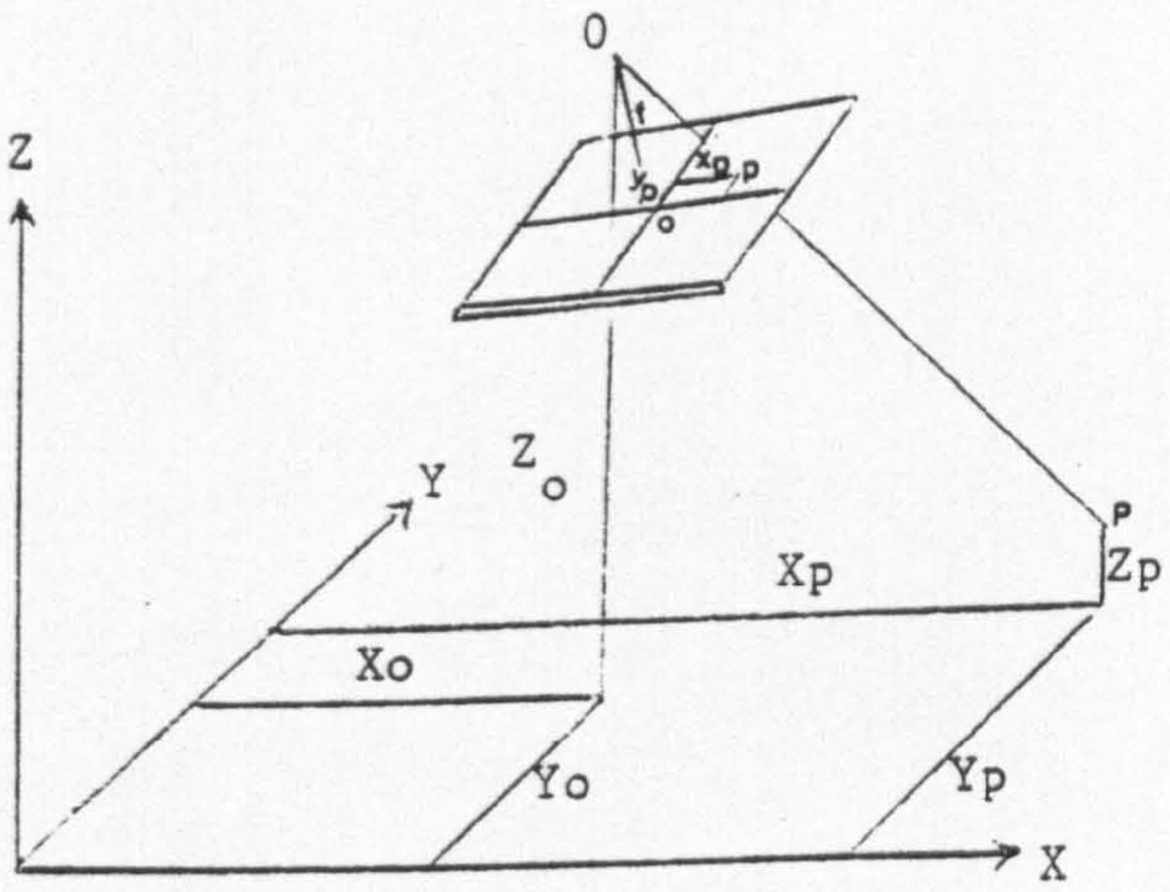


Fig 6.3 Relation between image coordinates and object coordinates for aerial photography.

a_{11} to a_{13} are the coefficients of the orthogonal matrix which contain the ω, ϕ and κ parameters (tilts) of the exterior orientation. As usual, f is the principal distance and λ is a scale factor.

6.3.1 Object Coordinate System

In undertaking the geometric analysis of side-looking radar images, it is first necessary to define the different coordinate systems that may be used to relate the radar system to the actual terrain. These relationships have been defined by Leberl (1972) (Fig. 6.4) and are given in modified form as follows:

(i) The Antenna Coordinate System: The electrical centre of the radar antenna constitutes the origin O_1 of the antenna coordinate system u, v, w. The w-axis points vertically upwards, the u-axis is the direction of the antenna's longitudinal axis and the v-axis lies in the direction of the actual line-scan.

(ii) The Co-ordinate System of the Antenna Mount: The origin O_2 of this system is located in the centre of the antenna mount with the u'-axis directed parallel to the antenna's longitudinal axis, while the w-axis again points vertically upwards along the normal and v'-axis occupies the third axis position in a left hand system. These three axes should be parallel to the u, v and w axes of the antenna and only slightly shifted (translated) in position from the antenna centre itself.

(iii) The Platform Coordinate System: The coordinate system $\bar{X}$, $\bar{Y}$, $\bar{Z}$ has its origin O_3 lying at the centre of gravity of the carrying vehicle. The $\bar{X}$-axis is defined as being the vehicle's longitudinal axis, the $\bar{Z}$-axis is again defined as lying along the normal in a vertical direction upwards from the origin and the $\bar{Y}$-axis occupies the position of the third axis according to the left-hand rule. Again, it must be remembered that the origin O_3 will lie only a short distance from the

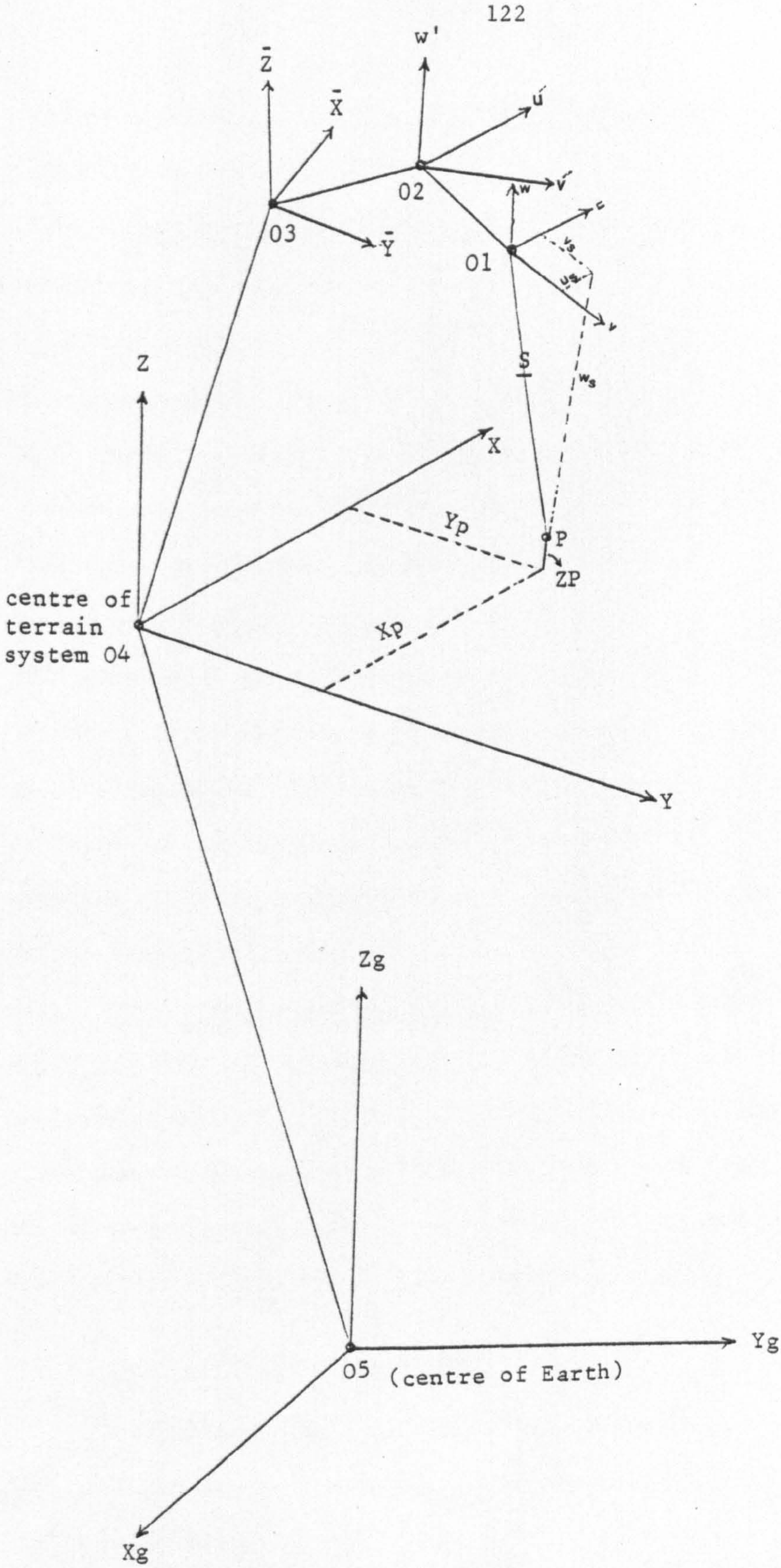


Fig 6.4 Coordinate Systems in the mathematical model of SLR.

origins O_1 and O_2 of the antenna and antenna mount respectively (as shown in Fig. 6.4).

(iv) The Terrain Coordinate System: The radar image coordinates of a point (x_p, y_p) must of course be expressed in terms of the ground coordinates (X_p, Y_p and Z_p) of the corresponding point on the terrain. This requires the definition of a terrain system with its origin at O_4. Normally, the values of X_p and Y_p will be expressed in terms of a particular grid on a specified projection. The origin of the particular grid system will have to be defined, as will the characteristics of the particular projection on which it is based. Also, the reference level for heights will need to be defined (e.g. mean sea level). Z_p is the height of point P above the reference surface. As an alternative, the system of geographical coordinates comprising latitude (b) and longitude (I) values can be used. Well-known formulae allow the transformation of geographical to grid coordinate values (and vice-versa).

(v) Geocentric Coordinate System: This system with coordinates X_g, Y_g, Z_g has its origin O_5 at the centre of the Earth. The Z_g-axis is directed along the Earth's polar axis of rotation while the $X_g Y_g$ plane lies in the equatorial plane with the X_g-axis passing through the intersection point of the Equator with the Greenwich meridian. Once again, the Y_g-axis completes a left hand system.

6.3.2 Coordinate Transformations

The relationships between these different coordinate systems have been set out by Akovetskiy (1968), Hockeborn (1971) and Leberl (1972) as follows:-

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix}_S = \underline{A} \cdot \left(\underline{B} \cdot \underline{C} \cdot \underline{D} \left(\begin{bmatrix} X_g \\ Y_g \\ Z_g \end{bmatrix} \right) \right) \quad \ldots\ldots\ldots 6.13$$

where:

$\underline{A}$ is a matrix which contains the rotations required to transform coordinates in the antenna system into the coordinate system of the antenna mount;

$\underline{B}$ is a rotation matrix which comprises the rotations required to bring the coordinates expressed in the system of the antenna mount into the coordinate system of the platform itself;

$\underline{C}$ is the matrix of rotation between the $\bar{X}$, $\bar{Y}$, $\bar{Z}$ platform coordinate system and the XYZ ground coordinate system; and

$\underline{D}$ is the matrix of rotations required for transformation from the X,Y,Z system into the geocentric system X_g, Y_g, Z_g; and

X_g, Y_g, Z_g are the geocentric coordinates of a point P.

Matrices $\underline{A}$, $\underline{B}$, $\underline{C}$ and $\underline{D}$ are rotational matrices only. For a full transformation from one coordinate system to another, additional terms which describe the translations or shifts between the coordinate origins should also be included.

Considering the rotational matrix $\underline{A}$, if the angles between the u, v, w and the u', v', w' are known (e.g. from calibration) and if the three translations are known (also from calibration), then the full transformation of coordinates from the u, v, w system to the u', v', w' system can be carried out. Obviously the rotational matrix $\underline{B}$, which relates the coordinate system of the antenna mount to the vehicle's coordinate system, can be derived in a similar manner, since the various rotations and translations are or should have been determined by calibration. In which case, strictly speaking, matrices A and B form part of the inner orientation of the side-looking radar set.

In practice, the coordinate axes of the u,v,w, u',v',w' and the $\bar{X}$ $\bar{Y}$ $\bar{Z}$ systems can be considered to be effectively parallel or very

nearly parallel, in which case, $\underline{A} \simeq \underline{B} \simeq I$ (a unit matrix). Furthermore, the translational components of the full transformations will be of the order of one or two metres. Therefore, in practice, these are also too small to be considered as important taking account of the limited resolution capabilities of most SLR systems. Thus the values of the translations can, in most circumstances, be considered to be zero.

While the translation and rotation values of the transformation matrices A and B are small, enabling the points O_1, O_2 and O_3 and the corresponding coordinate axes u, v and w; u', v' and w'; and $\bar{X}$ $\bar{Y}$ $\bar{Z}$ to be considered as being coincident for most situations, the same consideration cannot be applied to the remaining origins O_4 and O_5 and their respective coordinate systems. For example, large shifts between the coordinate values of the origins and in the dimension of the rotation values can be expected in the transformation between the platform/antenna mount/antenna coordinate system and the ground coordinate system. The same remarks will apply to the transformation of coordinates between the ground and the geocentric coordinate systems.

While the points discussed above are all relevant, once again in practice, the main consideration from the geometrical point of view will be the relationship between the position of the antenna at a given moment of time (defining the projection centre) and the corresponding set of points on the ground which can be imaged by the side-looking radar system at that moment. Due to the inherent geometry of the system, there are notable constraints which apply to this relationship between these points which will be discussed in the next section (6.3.3).

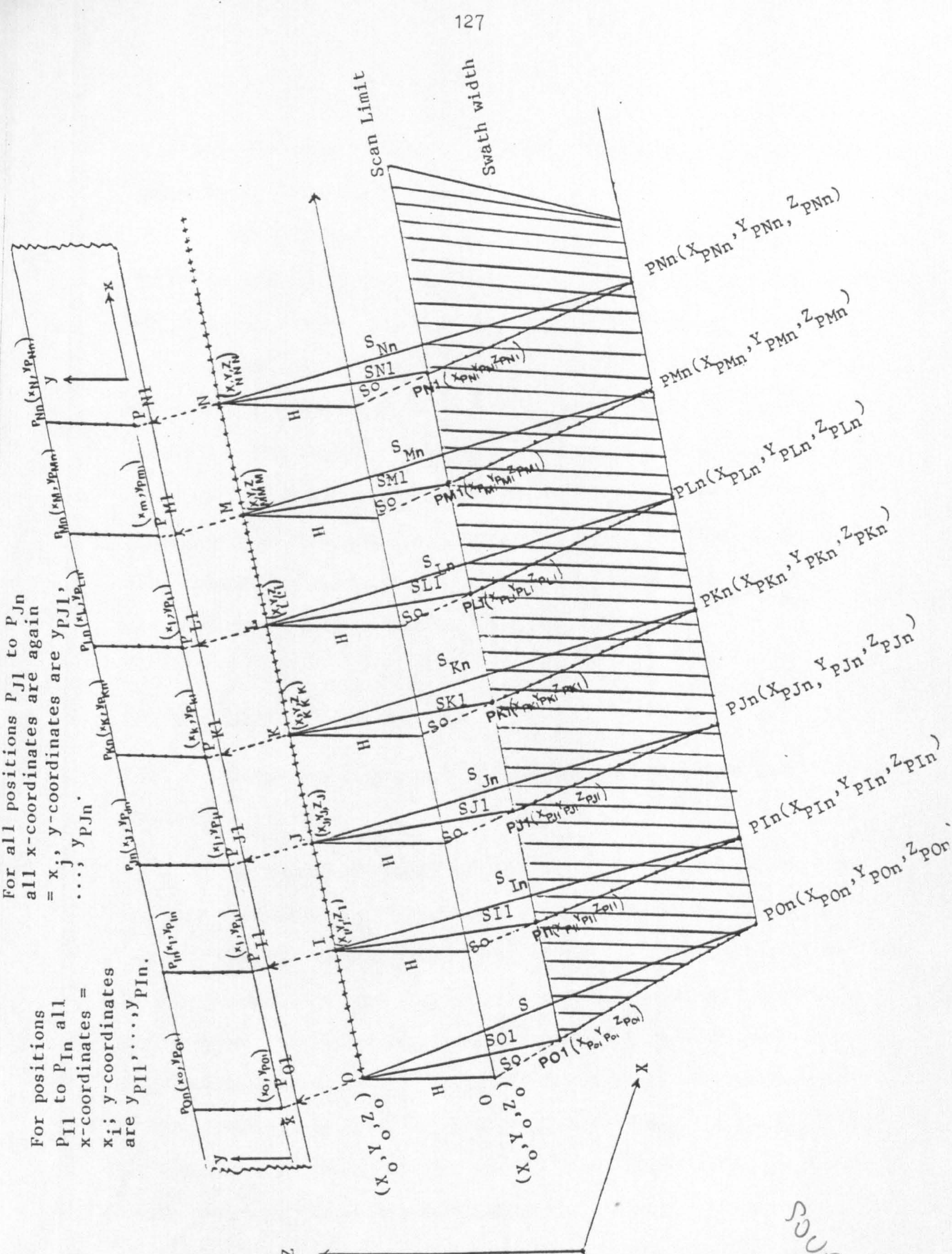
It is an obvious and welcome simplification to express the instantaneous position of the antenna in the air or in space (i.e. the projection centre) in the same ground coordinate system as that of the

terrain being imaged, in which case the necessity to determine and use the transformation parameters discussed above will no longer be necessary.

6.3.3 Constraints in Imaging Direction

In considering the constraints involved in SLR imaging, it is essential to note that, while equation 6.12 can be applied to the case of conventional photogrammetry using perspective photographs, this has no validity and cannot be applied to side-looking radar imagery. In conventional analytical photogrammetry, the exterior orientation parameters of equation 6.12 (X_o, Y_o, Z_o, ω, ϕ and κ) are constant for the whole of the imaged area in a single photograph. Furthermore, imaging can take place in any direction, i.e. along the line joining any point in the object space to the perspective centre (the optical centre of the taking camera) and then along a continuation of this line to the plane of the negative (which forms the basis of the well-known collinearity condition). This takes place simultaneously for all points located within the field of view for a single exposure.

However, in the case of side-looking radar, the sensor changes its position continuously during the exposure of the image of a corresponding area of the terrain. Therefore, there are a large number of projection centres located successively along the antenna vector (i.e. along the u-axis) during the imaging process (Fig. 6.5). Unlike aerial photography, the SLR imaging process is therefore highly constrained. Imaging can only take place in one direction only along a series of parallel vectors at right angles to the flight direction, i.e. along the v-axis and directions parallel to it. Consequently, each of the successive points I, J, K, L, M, N,..., Q along the antenna vector (i.e. each successive projection centre) has its own relationship



For all positions P_{J1} to P_{Jn} all x-coordinates are again $= x_j$, y-coordinates are y_{PJ1}, $\ldots$; y_{PJn}.

For positions P_{I1} to P_{In} all x-coordinates = x_i; y-coordinates are $y_{PI1}, \ldots, y_{PIn}$.

Fig 6.5 SLR imaging geometry(no tilts)

SOURCE,

with the object space which will be quite different to those of the other points lying along this vector.

The consequence of this is that, for each successive location of the antenna along the flight vector, imaging may take place in the vertical plane containing that location and situated at right angles to it and consequently along a single line on the ground. For each new position of the antenna, there is no possibility of imaging this particular line again. Instead, a new line located parallel to the first will be imaged. From the point of view of geometry, these considerations regarding the direction in which the imaging may take place constitute a fundamental difference in geometry between an aerial photograph and a side-looking radar image.

6.3.4. Basic Imaging Equations

Referring to Fig 6.5, the imaging process can be described as follows:

At a certain instant of time the platform is located at the exposure station I (with coordinates X_I, Y_I and Z_I in the ground system), which is some distance from the initial exposure station O. At this instant, a single line on the terrain lying cross-track at right angles to the line OI is imaged, extending from the point P_{I1} with coordinates (X_{PI1}, Y_{PI1}, Z_{PI1}) at the start of the imaged line in near range to the point P_{In} with coordinates (X_{PIn}, Y_{PIn}, Z_{PIn}) at the furthest position in the far range. This is the only line that can be imaged by the side-looking radar at this particular moment. The longest slant range recorded corresponds to the point P_{In} (X_{PIn}, Y_{PIn}, Z_{PIn}) and is designated by S_{In} while the shortest slant range

recorded corresponds to point P_{I1} (X_{I1}, Y_{I1}, Z_{I1}) and is designated by S_{I1}.

Turning next to the actual image of this line, the series of returns from the line P_{I1} P_{In} on the terrain will be represented by a straight line lying at right angles to the flight vector designated by P_{I1} P_{In} on the image. The representation of this line on the actual image can, as discussed in Section 6.2, be either in slant range or ground range. The x-image coordinate of each of the points contained in the line P_{I1} P_{In} will be constant, its actual position on the image being a function of the distance that the platform has flown from the initial position O with ground coordinate values Xo, Yo, Zo where imaging commenced (Fig 6.5). The instantaneous exposure station (or projection centre) I has ground coordinate values X_I, Y_I, Z_I, so that the image x-coordinates of all the ground points exposed from this point I lie along the line $p_{I1}p_{In}$ on the image and are given by:

$$x_{P_{I1}} = x_{P_{I2}} = x_i = m_x \left[(X_I - X_0)^2 + (Y_I - Y_0)^2 + (Z_I - Z_0)^2 \right]^{\frac{1}{2}} \quad .. \; 6.14$$

where m_x is the scale factor in the X-direction. This offers an alternative to the expression for the image coordinate x given in equation 6.5, expressed wholly in terms of the object space coordinates of the projection centres.

In the across-track direction, each point imaged along the terrain line $P_{I1}P_{In}$ will be at a unique slant range from the instantaneous antenna position I. These allow the generation of a series of discrete positions along the y-axis of the image. The first point on the image line (p_{I1}) corresponding to ground point P_{I1} will have a zero y-image coordinate, i.e. it will lie on the x-axis of the image; the

second point (p_{I2}), corresponding to ground point P_{I2}, will have a y-image coordinate y_{PI2} which will be a function of the difference in slant range between points P_{I1} and P_{I2}. Successive positions P_{I3}.... ....p_{In} will be generated on the image with y-image coordinates y_{PI3}, ..., y_{PIn}. Each of these will be a function of the difference between the slant range to the point concerned S_{In} and the initial slant range to the first point P_{I1} on the line (S_{I1}). This allows one to write the following equation for the y-image coordinate:-

$$y_{PIn} = m_y \left[S_{In} - S_{I1} \right]$$

but $$S_{In} = \sqrt{(X_{PIn} - X_I)^2 + (Y_{PIn} - Y_I)^2 + (Z_{PIn} - Z_I)^2}$$

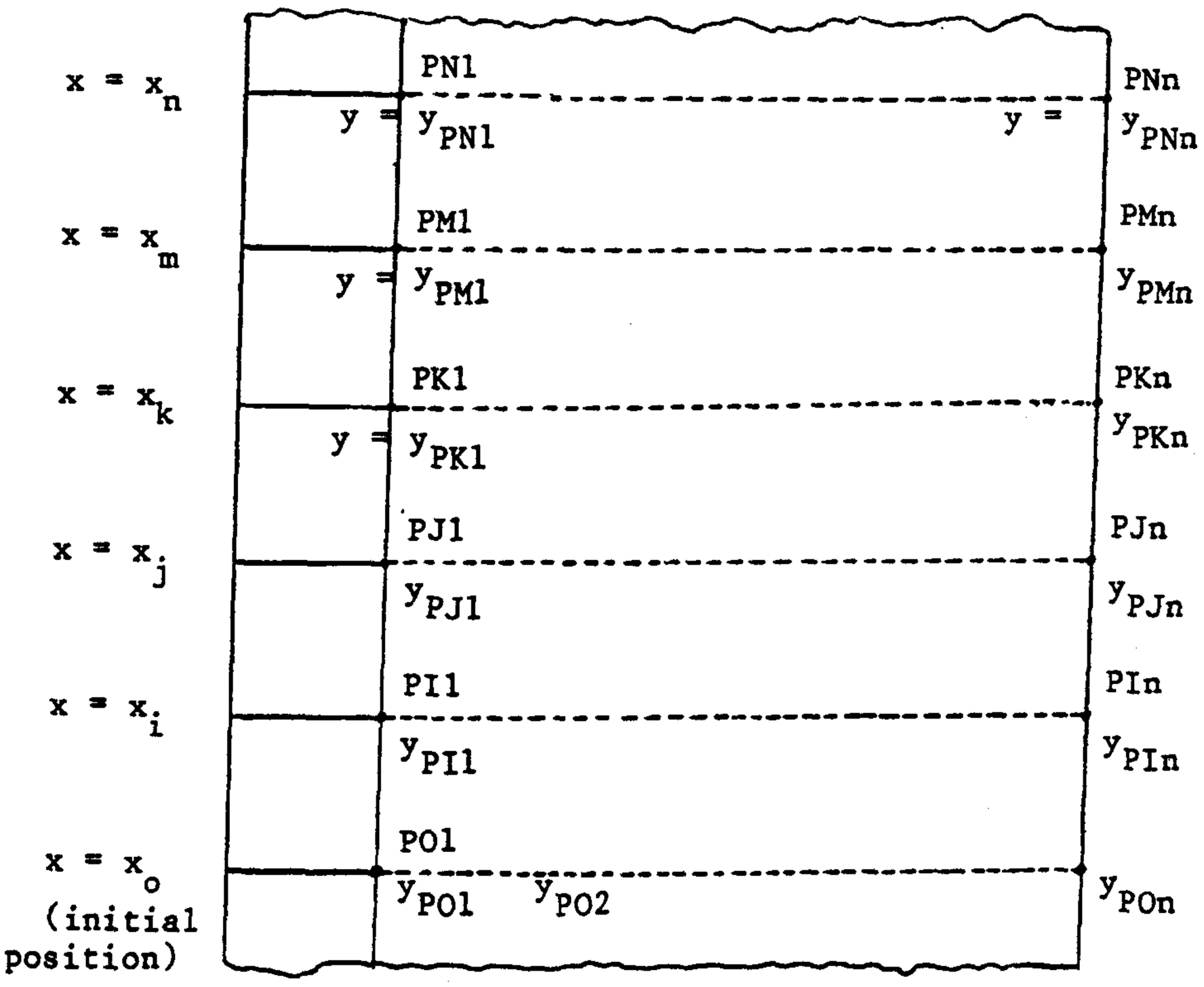
and $$S_{I1} = \sqrt{(X_{PI1} - X_I)^2 + (Y_{PI1} - Y_I)^2 + (Z_{PI1} - Z_I)^2}$$

$$\therefore y_{PIn} = m_y \left(\left[(X_{PIn}-X_I)^2+(Y_{PIn}-Y_I)^2+(Z_{PIn}-Z_I)^2 \right]^{\frac{1}{2}} - \left[(X_{PI1}-X_I)^2 + (Y_{PI1}-Y_I)^2 + (Z_{PI1}-Z_I)^2 \right]^{\frac{1}{2}} \right) \quad \ldots\ldots\ldots\ldots\ldots\ldots 6.15$$

where m_y is a scale factor in the cross-track direction; and

n = 1,2,3,...,n being the successive numbering for the points lying along the cross-track direction.

A few moments later, the platform will have moved to the exposure station J with object space coordinates X_J, Y_J, Z_J. At the exposure station J, a new line $P_{J1}P_{Jn}$ on the terrain lying in the across-track direction will imaged in the same manner as line $P_{I1}P_{In}$ at the exposure station I. All points $p_{J1},\ldots,p_{Jn}$ located along this newly imaged line will have the same x-image coordinate value which will of course be different to those of the preceding line $p_{I1},\ldots,p_{In}$ (see Fig 6.6). As already discussed, the y-image coordinates of all the



$$y_{PO1} = y_{PI1} = y_{PJ1} = y_{PK1} = y_{PL1} \cdots \quad y_{POn} = y_{PIn} = y_{PJn} = y_{PKn} = y_{PMn} = y_{PNn}$$

Fig. 6.6 Plan view of the imaged lines on film.

points on the line $P_{J1}P_{Jn}$ will be functions of the corresponding slant ranges S_{Jn} and S_{J1}. This process continues for all successive platform positions between and including J, K, L, M, N,..... etc until the required strip of the terrain is imaged in its entirety. The general set of relationships between the image coordinates and the object coordinates of the successive projection centres and the corresponding terrain coordinates during the continuous imaging process can be expressed by the following equations:-

$$\left.\begin{aligned} x_i &= m_x \left[(X_I-X_o)^2 + (Y_I - Y_o)^2 + (Z_I-Z_o)^2\right]^{\frac{1}{2}} \\ y_{PIn} &= m_y \cdot \left(\left[(X_{PIn}-X_I)^2+(Y_{PIn}-Y_I)^2+(Z_{PIn}-Z_I)^2\right]^{\frac{1}{2}} \right. \\ &\qquad \left. - \left[(X_{PI1}-X_I)^2+(Y_{PI1}-Y_I)^2+(Z_{PI1}-Z_I)^2\right]^{\frac{1}{2}}\right) \end{aligned}\right\} \quad \text{---6.16}$$

6.3.5 The Determination of the Elements of Exterior Orientation

The designed or nominal flight path, the specified constant altitude of the platform and the desired attitude of the antenna are never achieved in practice due to many external factors, e.g. navigational errors and atmospheric turbulence for airborne systems and perturbations resulting from drag forces and variations in the Earth's gravity field for spaceborne systems, especially at low orbital altitudes. Three different approaches are possible for the assessment of exterior orientation parameters for single images:

(i) By Platform Stabilization: This technique attempts to eliminate or minimize any departures from the designed or nominal flight pattern. While its successful implementation would have considerable advantages to

the users, to implement it fully is an enormously expensive undertaking. Furthermore, it may only be partially successful in fulfilling its declared aims, especially in airborne environments where substantial atmospheric turbulence may be prevalent, resulting in uncompensated departures from the planned flight parameters of an unpredictable nature. With a sophisticated synthetic aperture radar such as the Aero Service/Goodyear GEMS - 1000 airborne system, an equally sophisticated and expensive Inertial Navigation System (INS) gives continuous inputs to the autopilot to ensure that the aircraft is flown along the correct track and at the correct altitude. Furthermore, the antenna mount receives inputs from the INS which will hopefully result in a stabilized attitude of the antenna. Although a few results have been published (e.g. see Graham, 1970, Derenyi, 1975 and Leberl et al, 1976d), generally speaking, it is quite difficult to establish from the literature the degree to which this type of stabilization is successful in practice. In less sophisticated systems, only ω-stabilization of the antenna may be provided.

(ii) By In-flight Recording of the Orientation Parameters : An alternative approach is to measure the departures from the desired or designed flight path and parameters with a view to using this information later in the rectification process. Thus with an airborne system, an on-board Doppler radar can be used to track the actual flight path flown while an APR used in conjunction with a statoscope may record the terrain clearance and the variations in flying height. Various attitude sensors may be used to provide continuous in-flight measurements and monitoring of the tilts. However, the shortcomings of using such an approach are that it does not prevent the double-imaging, gaps, etc, which result from incorrect navigation, turbulence,etc.

With spaceborne imaging, there is a continuous flow of attitude information which is telemetered back to the ground receiving station together with the image data. There seems to be a lack of positive confirmation that this attitude information is used during the actual processing of the imagery.

(iii) By Use of Ground Control: This method is used when direct measurements of the exterior orientation parameters are not available. It may be carried out by employing an appropriate number of ground control points in conjunction with a mathematical model. The major difficulty is that the exterior orientation of each line imaged by the SLR device would need to be determined. This approach cannot be utilized in any practical terms - the demands for control and computation would be impossible to meet. Therefore,as with all other types of line scan imagery, the usual approach is, in the first instance, to treat a short section of the SLR strip image as having a constant orientation. Since the orientation parameters are continuously varying in their amount and direction with time over the area of a single SLR image, almost inevitably the coefficients can only be determined in an approximate manner which has inherent limitations in the accuracy of their determination. Attempts can then be made to improve the accuracy of determining the orientation parameters by modelling their behaviour across the image using an interpolative method.

The approach given by Leberl (1975a) for the determination of the orientation parameters uses the following unknown constants:

(i) the flying height H;

(ii) origin X_o, Y_o in object space corresponding to image coordinates $x = o$, $y = \left(\frac{c^2 e^2}{2} - H^2 \right) m_y$;

(iii) heading K_o ;

(iv) squint ϕ_o ;

(v) scale m_x along the x-coordinate axis; and

(vi) scale m_y along the y-axis.

Neglecting Earth curvature, one obtains:

$$X_P = X_0 + x \cos K_o \cdot m_x + S. \sin\phi_o$$

$$Y_P = Y_0 + x \sin K_o \cdot m_x + S. \left\{ \frac{(H - Z_P)^2}{S^2} - \sin \phi_o \right\}^{\frac{1}{2}}$$

where $S = y.m_y + \frac{ce}{2}$

Each control point produces two equations with seven unknowns. The equation system in these seven unknowns is non-linear. The simplest solution therefore is an iterative one using linearized form of the above equations:

$$dX_P = dX_o + x\,(\cos K_o.\, dm_x - \sin K_o.m_x\, dK_o) +$$

$$\sin\phi_c(y.dm_y + \tfrac{c}{2}\, de) + \cos\phi_o(y.m_y + \tfrac{c}{2}\, e)\, d\phi_o \; ; \quad \text{and}$$

$$dY_P = dY_o + x(\sin K_o\, dm_x + \cos K_o.\, m_x.dK_o)$$

$$- 2\sin^2\phi_o(y.m_y + \tfrac{ce}{2})\,(y.dm_y + \tfrac{c}{2}\, de)$$

$$+ (ym_y + \tfrac{ce}{2})^2.(2\sin\phi_o\, \cos\phi_o)\, d\phi_o + 2(H-Z_P)dH-2(H-Z_P)dZ_P$$

Once these values have been determined for an individual image, a more refined model can be established by applying time-varying functions to these parameters. This can be done in various ways. Two of these are outlined below.

(a) <u>Interpolation by Polynomials representing the Flight Path and Attitude Variations</u>

The choice of the terms contained in the polynomials depends entirely on the assumptions made regarding the time-varying behaviour of the parameters. Polynomials of the type:-

$$\left.\begin{aligned} dbx &= a_o + a_1x + a_2x^2 + a_3x^3 + ... ; \\ dby &= b_o + b_1x + b_2x^2 + b_3x^3 + ... ; \\ dbz &= c_o + c_1x + c_2x^2 + c_3x^3 + ... ; \\ d\phi &= d_o + d_1x + d_2x^2 + d_3x^3 + ... ; \text{ and} \\ dK &= e_o + e_1x + e_2x^2 + e_3x^3 + ... \end{aligned}\right\} \text{---6.17}$$

have been suggested by Derenyi (1974a,b). This assumes that the variations in the orientation parameters are functions of the image x-coordinates. Initially, their values are set to predetermined values for $x = o$. When $x \neq o$, the equations can be used to set up observation equations for the image coordinates observed. These allow the computation of the polynomial coefficients, which can then be used to estimate the situation for any other point in the image. In turn, these can be used to correct the errors in position either in terms of the change in image or ground coordinates induced by the deviations according to equations 6.24 (see later).

(b) Parameter Restitution by Modelling of Flight Path and Attitude

The basic idea in this approach is to assume a periodic variation for the parameters in terms of Fourier series expansions (Dowideit, 1975). According to Dowident (1975) and Konecny (1975, 1976), this model takes the following form:

$$X_I = X_I \text{ approx.} + a_o + a_1 \cos \frac{x}{t_m} + a_2 \sin \frac{x}{t_m} + a_3 \cos \frac{2x}{t_m} +$$

$$a_4 \sin \frac{2x}{t_m} + ...$$

$$Y_I = Y_I \text{ approx.} + b_o + b_1 \cos \frac{x}{t_m} + b_2 \sin \frac{x}{t_m} + b_3 \cos \frac{2x}{t_m} + b_4 \sin \frac{2x}{t_m} + \ldots$$

$$Z_I = Z_I \text{ approx.} + c_o + c_1 \cos \frac{x}{t_m} + c_2 \sin \frac{x}{t_m} + c_3 \cos \frac{2x}{t_m} + c_4 \sin \frac{2x}{t_m} + \ldots$$

$$\phi = \phi \text{ approx.} + d_o + d_1 \cos \frac{x}{t_m} + d_2 \sin \frac{x}{t_m} + d_3 \cos \frac{2x}{t_m} + d_4 \sin \frac{2x}{t_m} + \ldots$$

$$K = K \text{ approx.} + e_o + e_1 \cos \frac{x}{t_m} + e_2 \sin \frac{x}{t_m} + e_3 \cos \frac{2x}{t_m} + e_4 \sin \frac{2x}{t_m} + \ldots \qquad 6.18$$

where t_m is a constant time interval appropriately chosen for the frequency range of data.

These equations will also allow (after linearization) the setting up of observation equations based on the image coordinates in order to compute the coefficients and can be used in the same manner as (a) above.

Whichever of these methods is used to determine the elements of the exterior orientation, the extent of the displacements produced by the rotations $d\phi$ and dK, and the translations dbx, dby and dbz is a matter of fundamental importance in considering the geometric aspects of SLR imagery.

6.3.6 The Geometrical Effects of the Rotations

(i) dK (swing or yaw): A displacement dK results from an uncompensated crab in the orientation of the antenna relative to the

flight path so that it no longer points in the desired direction at right angles to the path. If this pointing is maintained over a short section of the flight line, the result is a shearing deformation of the imaged terrain features (Fig 6.7),i.e. a square-shaped object on the terrain will take the form of a rhombus on the image. Furthermore, variations in the value of dK will also cause double imaging and gaps. Such phenomena are especially noticeable in airborne real-aperture SLR imagery. In the case of synthetic aperture radar, these effects are often minimized since frequently an INS is used in conjunction with a rotational mount for the antenna. Furthermore, any remaining residual misalignment of the antenna will largely be compensated for in the processing of the Doppler frequency shift of the signal returns.

From Fig 6.7, it can be seen that the effect of the dK in a single point is a lever arm displacement. The positional error induced in the X-direction $(\Delta X)_K$ by an angular rotation dK is given by:

$$(\Delta X)_K = Y\ dK \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 6.19$$

where Y is the across-track coordinate ; and

dK is expressed in radians.

Thus if a yaw of 0.1° ($\simeq$ 2 milliradians) is present, the image will have a displacement of about 30 m in the X-direction at the centre of an SLR image located 15 km from the antenna. It can be seen that serious positional errors arise from this source if the dK error is allowed to arise or remains uncompensated for, especially when space-borne imagery is being considered.

(ii) <u>$d\phi$ (pitch or tip)</u>: Again, this is a lever arm displacement but, in this case, in the vertical (X-Z) plane. A rotation $d\phi$ applied to the antenna (Fig 6.8) will give rise to a positional error $(\Delta X)_\phi$ in

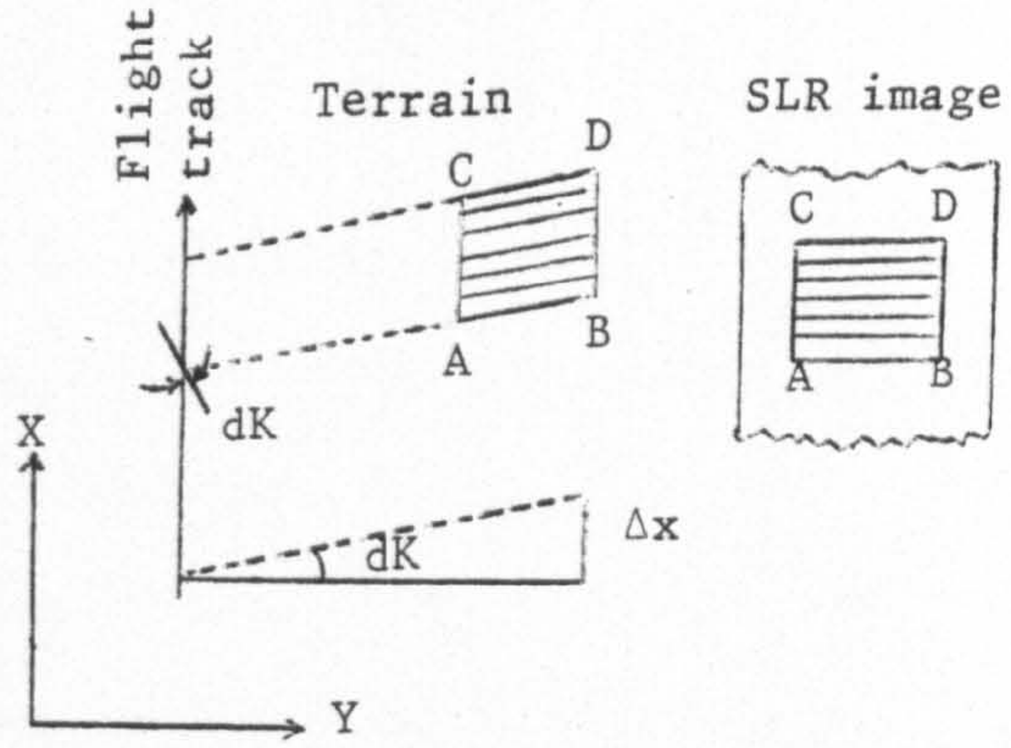


Fig. 6.7 Geometrical effect of a dK rotation (Petrie, 1978)

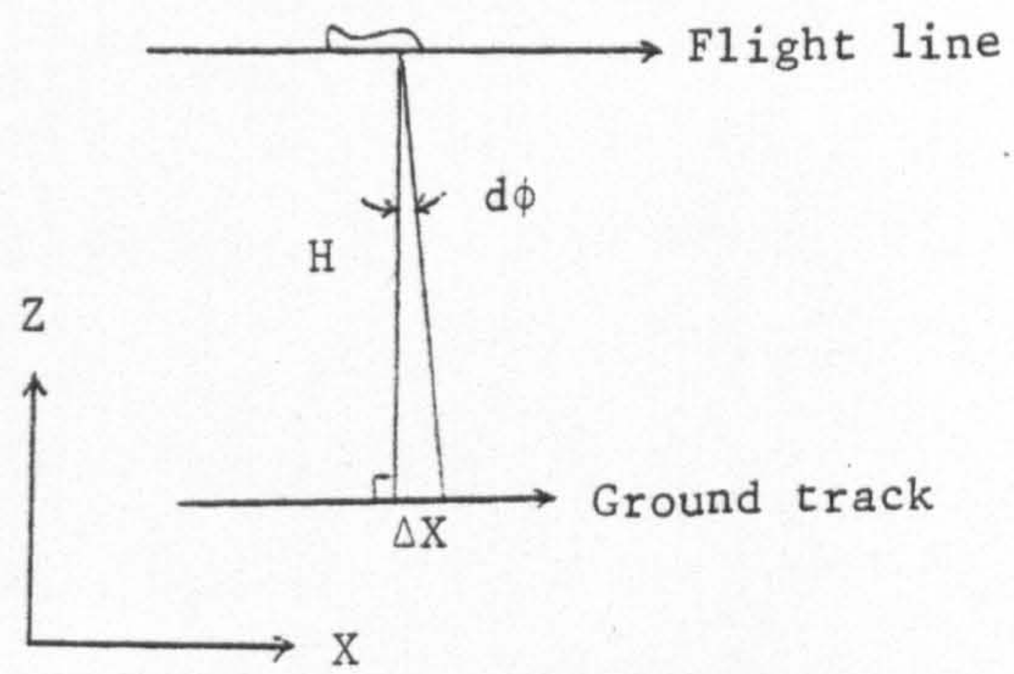


Fig. 6.8 Geometrical effect of a dϕ rotation

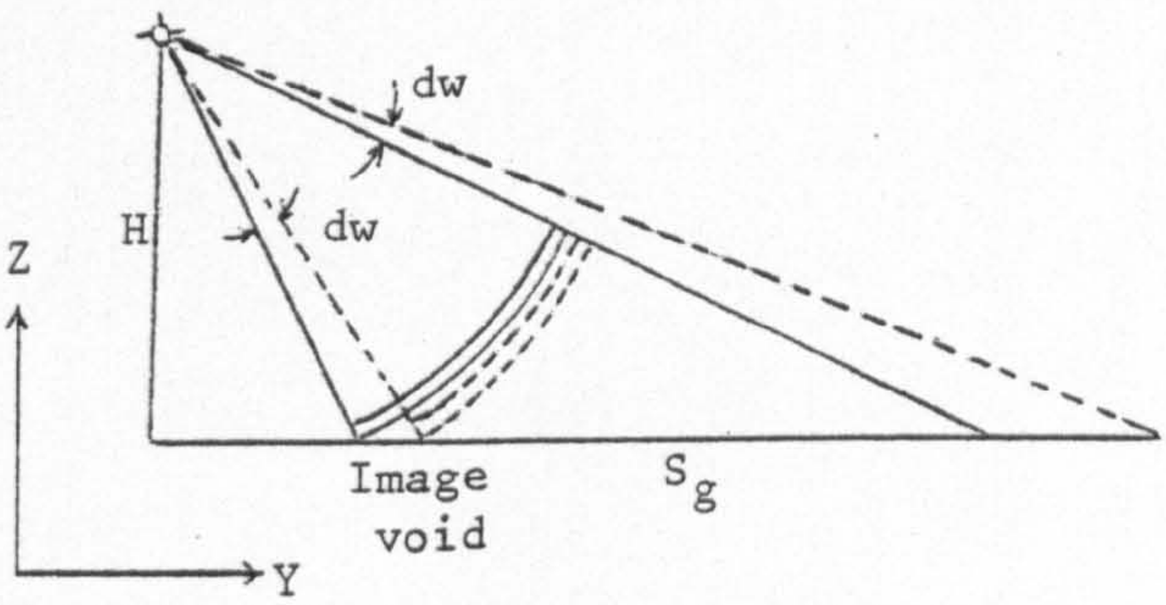


Fig. 6.9 dw rotation has no geometrical effect on SLR

the X-direction as follows:

$$(\Delta X)_{\phi} = H\, d\phi \qquad \text{..... 6.20}$$

where H is the flight (or orbital) altitude. Therefore, at an altitude of 12 km, a $d\phi$ rotation of 2 mrad will result in an along-track displacement of 24 m in the location of an SLR image. However, in the case of the Seasat radar system (H = 800 km), the same $d\phi$ error of 2 mrad will give rise to an along-track error of 1.6 km if it is allowed to occur or has not been compensated for. This, again, shows that serious positional errors will arise from this source if the $d\phi$ rotation remains uncompensated for.

(iii) <u>$d\omega$ (tilt or roll)</u>: The occurrence of this angular displacement is as shown on Fig 6.9. An across-track tilt $d\omega$ does not in fact produce a geometric displacement on the image since it has no effect on the measured radar range to an object. However, any excessive roll will cause a loss of coverage at either the near range or far range depending on whether the tilt is inwards or outwards with respect to the ground track. The result will be an image void at one edge of the film.

6.3.7 <u>The Geometrical Effects of the Translations</u>

The translational displacements are caused by the deviation of the carrying vehicle from its designed track or altitude.

(i) <u>dbx (deviation in the along-track direction)</u>: As previously discussed, this type of displacement (Fig 6.10) takes place when an erroneous platform velocity is determined or when the velocity of the vehicle and that of the film are wrongly synchronized. It is obvious therefore that this displacement affects only the image coordinates in the x-direction, i.e.

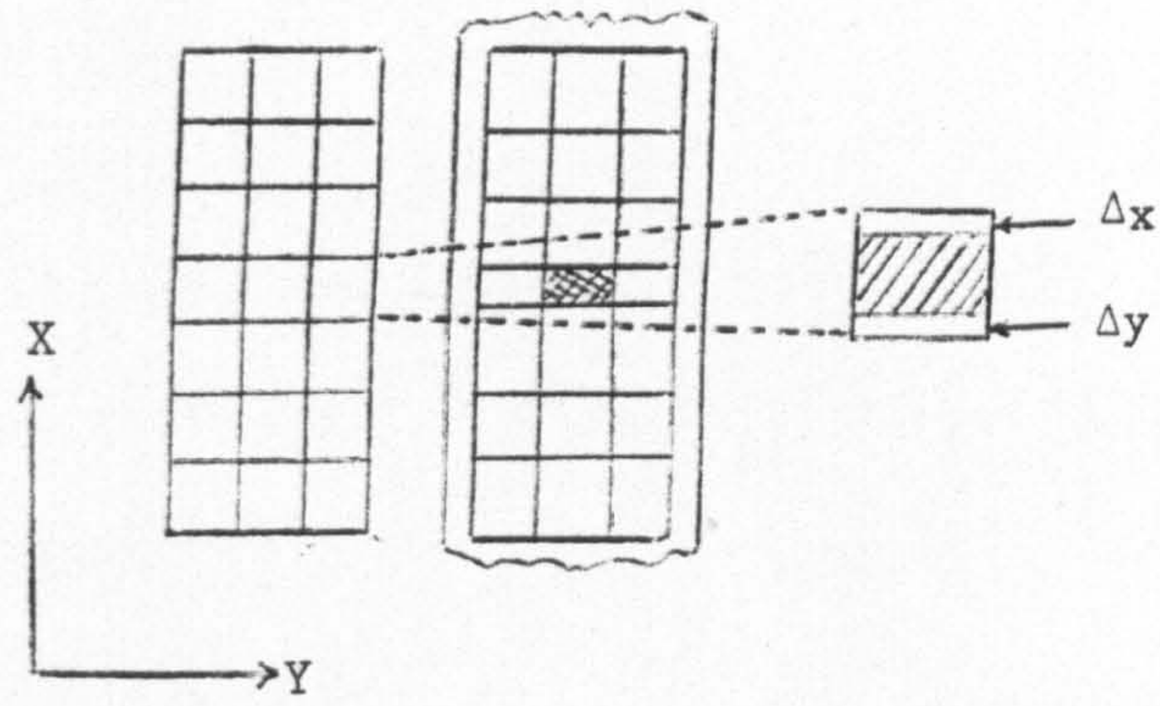


Fig. 6.10 Geometrical effect of dbx displacements

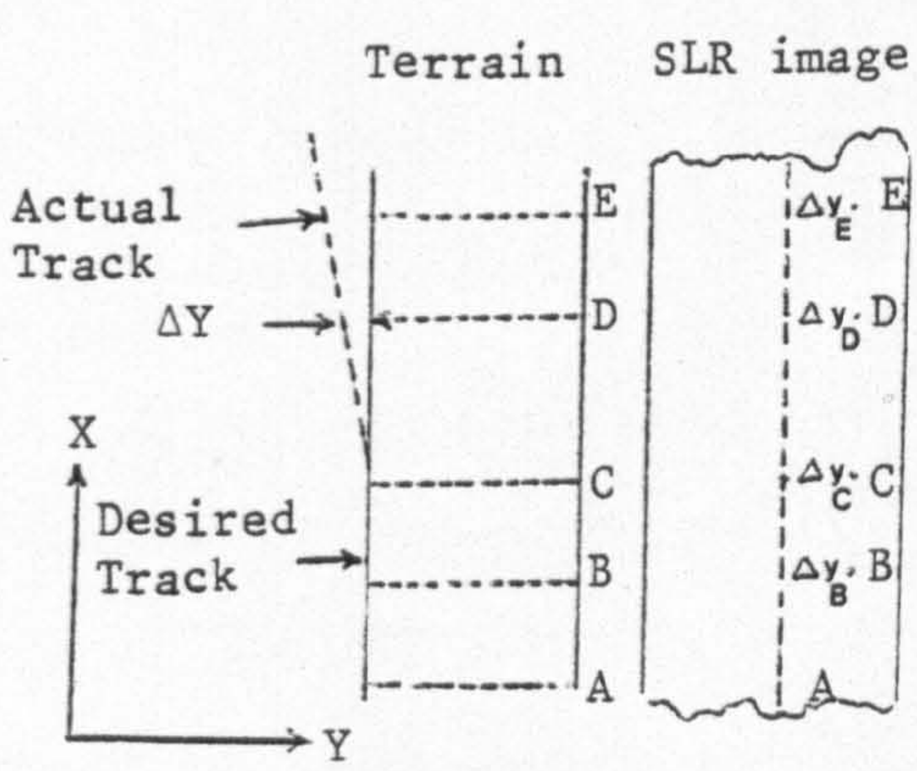


Fig. 6.11 Geometrical effect of dby displacements

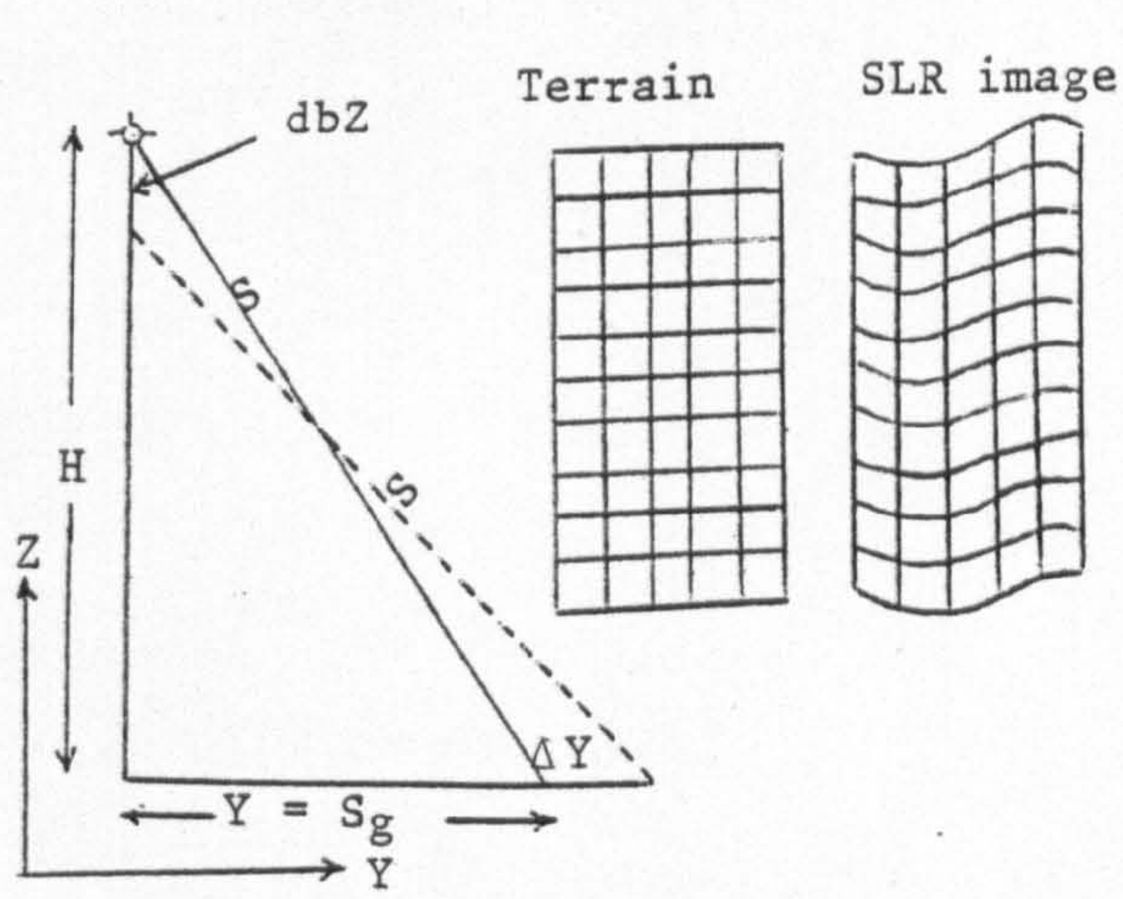


Fig. 6.12 Geometrical effects of dbz displacements

$$\Delta X_{dbx} = dbX \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6.21$$

For example, for a vehicle moving at 300 m/sec, a velocity error of + 10 m/sec (i.e. 3 %) continuing for two seconds will give rise to an along-track error dbX of + 20m, i.e. the platform has moved 620 m instead of 600 m. For Seasat, with a velocity of 7500 m/sec over the ground a 1% error in this velocity continuing for two seconds will result in an along-track error of 150 m.

(ii) dby (deviation in the cross-track direction): The cause of this type of displacement (Fig 6.11) is the deviation of the track from its desired or assumed value. It takes the form of a range error, which means that it affects the y image coordinate only. The resulting error in position can be derived by differentiation of equation 6.1. Putting ds $\simeq \Delta S$, $\Delta Sg \simeq \Delta Y$, one gets:

$$\Delta S_{dby} = \frac{Sg}{S} \quad Sg = \frac{Y}{S} dbY \qquad \ldots\ldots\ldots\ldots\ldots\ldots 6.22$$

(iii) dbZ (altitude variation error): This displacement is sometimes referred to as the altitude uncertainty and is caused by the variation in the flying altitude of the vehicle above the reference surface. As can be seen from Fig 6.12, this error affects the cross-track position only, i.e. the y-image coordinate. The formula for this error can also be derived from equation 6.1 in a similar way to (ii) above and this gives:

$$\Delta S_{dbZ} = \frac{H}{S} \cdot dbZ \qquad \ldots\ldots\ldots\ldots\ldots\ldots 6.23$$

Examination of equation 6.23 shows that the effect of this translation is smaller at longer ranges, i.e. in the far-range part of the image, and larger at short ranges, i.e. at narrow elevation angles in the near-range part of the image.

6.3.8 Combined Effect of the Elements of Exterior Orientation on SLR Imagery

The combined effect of the changes in the elements of the exterior orientation on an SLR image can be derived by summing up the individual equations for image displacements derived in sections 6.3.6 and 6.3.7 above.

The individual equations 6.19, 6.20 and 6.21 may be summed up to give the overall displacement caused by the elements dK, dϕ and dbX. Similarly, in the Y-direction, equations 6.22 and 6.23 may be summed up to give the overall displacement caused by dbY and dbZ. The resulting equations are as follows:

$$\left.\begin{aligned} \Delta X &= Y dK + H d\phi + dbX \\ \Delta Y &= \frac{Y}{S}\, dbY + \frac{H}{S}\, dbZ \end{aligned}\right\} \qquad \text{..... 6.24}$$

6.3.9 The Effect of the Earth Curvature on SLR Imagery

The shape of a segment on the Earth's surface imaged by a SLR is not a plane but part of a sphere. However, for many situations e.g. especially with aircraft imagery at low altitudes, the surface can be considered to be a plane, so that the basic imaging equations discussed above are derived on this basis. With SLR imagery taken from high-flying aircraft or from a satellite such as Seasat, this simplification is no longer valid and it is necessary to consider the effect of the curvature of the Earth on the imaging process and to derive corrections to the measured ranges to ensure that they correspond to the values that would have resulted if the Earth had been a flat surface.

For flat terrain, the error ΔS induced by the curvature of

the Earth on the measured range S to a point P on the surface of the Earth is given by (Fig 6.13):

$$\Delta S = (S - S')$$

where S = the measured (or computed) slant range; and

S' = the equivalent slant range to the point P' located on the horizontal surface through the ground analogue of the antenna position O_1.

But $S' = (Sg^2 + H^2)^{\frac{1}{2}}$

$$\therefore \quad \Delta S = S - (H^2 + Sg^2)^{\frac{1}{2}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots 6.25$$

Now, in triangle O_1EP, one can write the following equation:

$$S^2 = R^2 + (R+H)^2 - 2R(R+H)\cos\varepsilon$$

where ε is the angle subtended at the centre of Earth;

R = radius of the Earth; and

H = flying height.

$$\therefore \quad \cos\varepsilon = \frac{R^2 + (R+H)^2 - S^2}{2R(R+H)}$$

But from Taylor's Series, $\cos\varepsilon \simeq 1 - \frac{\varepsilon^2}{2}$

$$\therefore \quad 1 - \frac{\varepsilon^2}{2} = \left[\frac{R^2 + (R+H)^2 - S^2}{2R(R+H)}\right]$$

$$\therefore \quad \frac{\varepsilon^2}{2} = 1 - \left[\frac{R^2 + (R+H)^2 - S^2}{2R(R+H)}\right]$$

$$= \frac{2R(R+H) - \left[R^2 + (R+H)^2 - S^2\right]}{2R(R+H)}$$

$$= \frac{-H^2 + S^2}{2R(R+H)}$$

$$\therefore \quad \frac{\varepsilon^2}{2} = \frac{S^2 - H^2}{2R(R+H)}$$

$$\therefore \quad \varepsilon^2 = \frac{S^2 - H^2}{R(R+H)}$$

$$\therefore \quad \varepsilon = \sqrt{\frac{S^2 - H^2}{R(R+H)}}$$

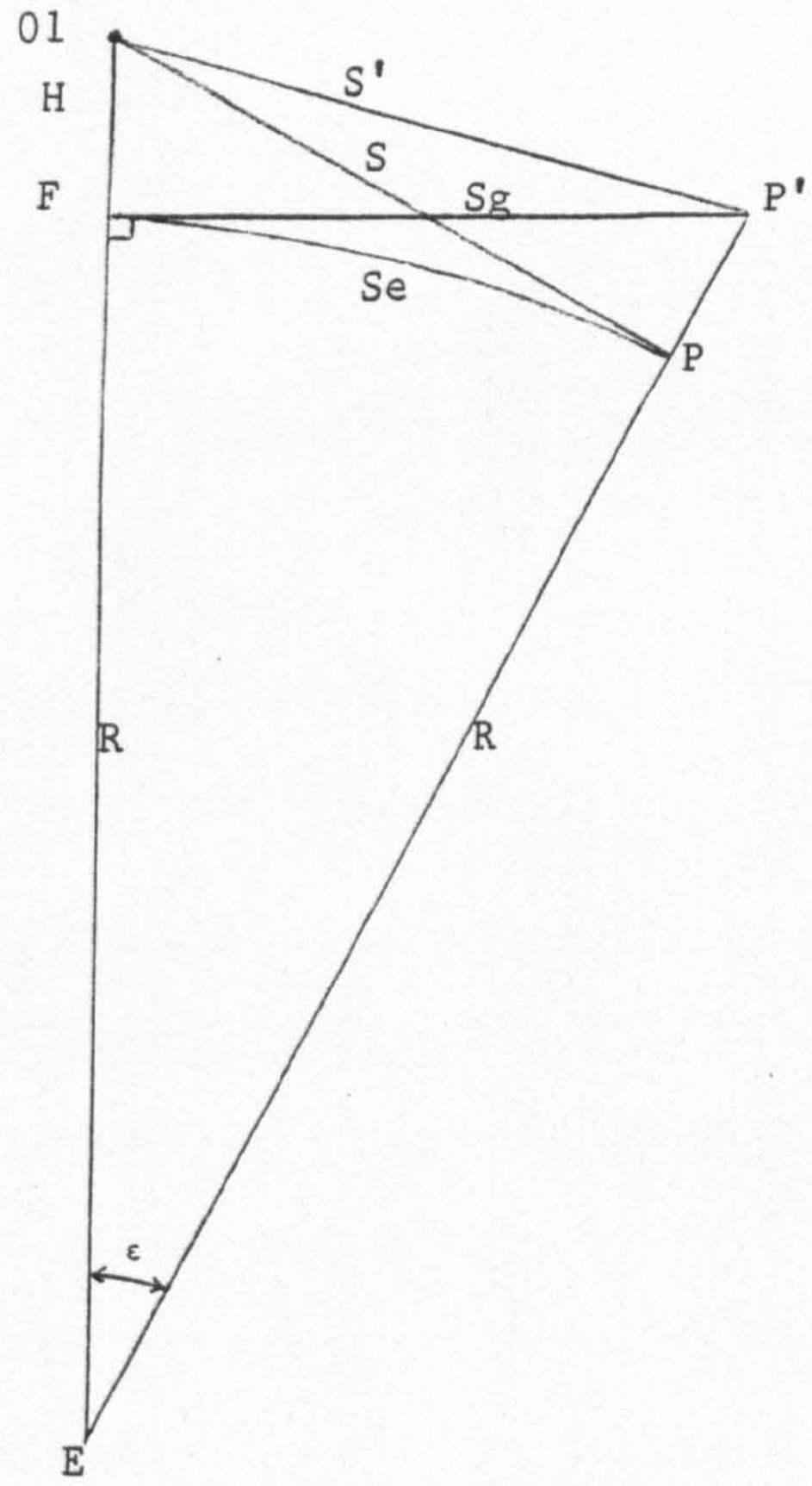


Fig 6.13 Effect of Earth Curvature

But $\frac{Sg}{R} = \tan \varepsilon$

For small values of ε, $\tan \varepsilon \simeq \sin \varepsilon \simeq \varepsilon$

$$\therefore \frac{Sg}{R} = \sqrt{\frac{S^2 - H^2}{R(R+H)}}$$

$$\therefore Sg = R \sqrt{\frac{S^2 - H^2}{R(R+H)}}$$

$$\therefore Sg^2 = R^2 \left(\frac{S^2 - H^2}{R(R+H)} \right) = R \left(\frac{S^2 - H^2}{(R + H)} \right)$$

$$\therefore Sg^2 = \frac{S^2 - H^2}{1 + H/R}$$

Substituting this equation into 6.27 gives

$$\Delta S = S - \left[H^2 + \frac{S^2 - H^2}{1 + H/R} \right]^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots\ldots \quad 6.26$$

Since R is a constant and H and S are measured or computed, equation 6.28 gives the correction for Earth curvature directly. Taking R = 6370 km, then taking the Seasat SAR range of 836 km, and with H = 800 km, this error amounts to 3,940 m which shows clearly that the Earth curvature effect must be corrected for in the case of spaceborne radar imageries.

6.3.10 The Effect of Range Migration on SLR Imagery

Range migration (or range walk) is a term describing the effects that would cause a stationary ground object on the surface of the Earth to appear to describe a locus other than a straight line parallel to the flight path (Raney, 1977) (see Fig 6.14). A number of factors contribute towards this effect, the most important ones being:

(i) the spherical wavefront; and

(ii) the rotation of the Earth.

With regard to the former effect, with satellite SARs, the length of the synthetic aperture is so great that the curvature of the

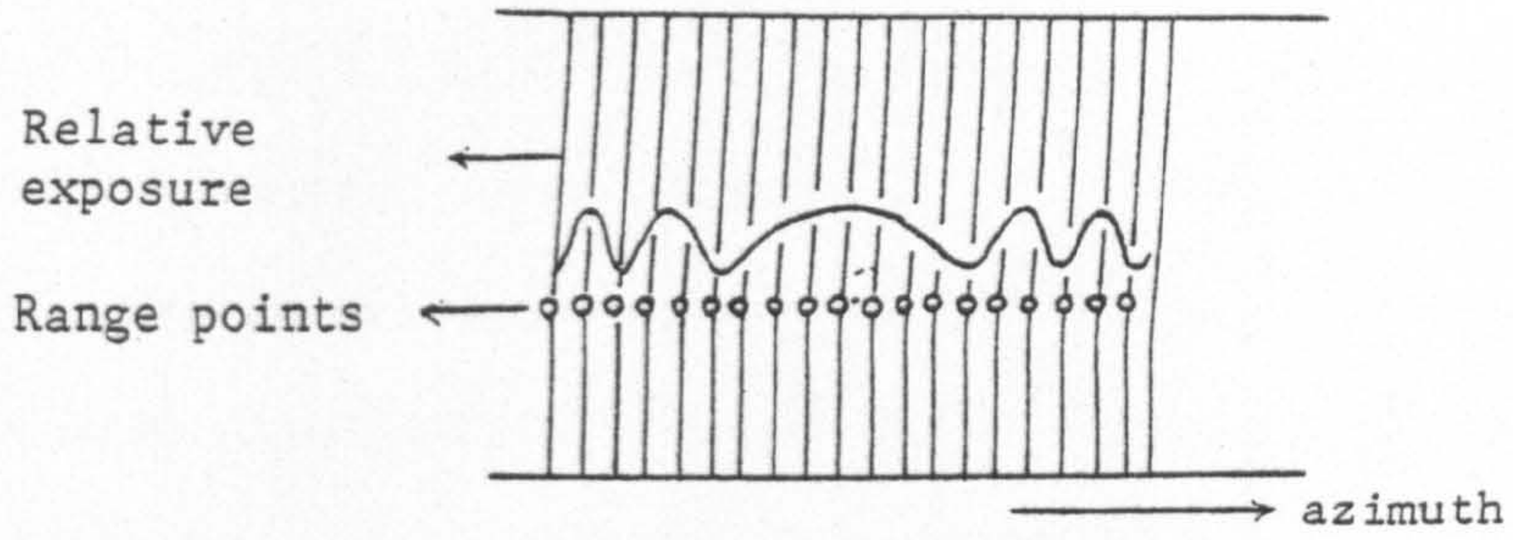


Fig 6.14a. The returns on a signal film from one scattering point in successive range shots.

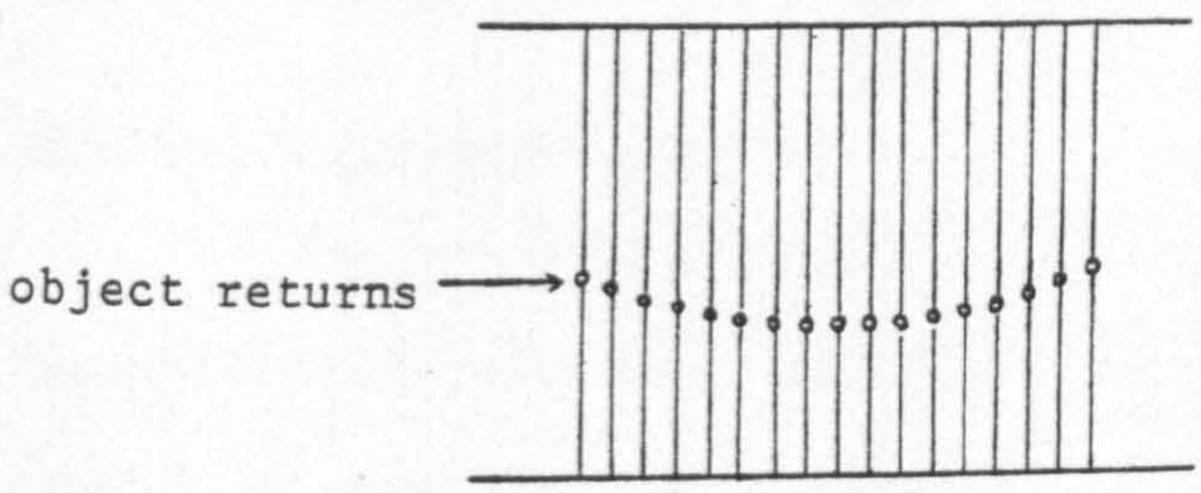


Fig 6.14b. Curvature of the range history from a single point.

Fig 6.14c. The appearance of the signal film when the radar pulses are chirpec

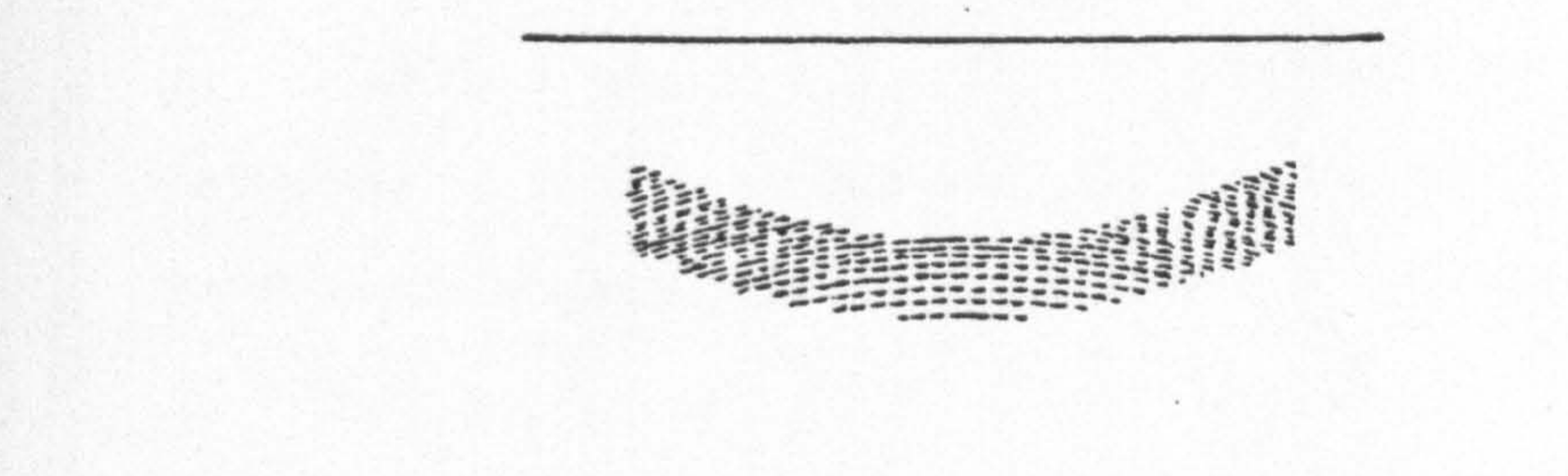


Fig 6.14d. The effect of the Earth's rotation.

successive range returns from a single point is quite appreciable amounting to several range elements (Fig 6.14b). Furthermore with chirped radar pulses (as normally used with a satellite-borne SAR) it will be seen (in Fig 6.14c) that a time-varying Doppler shift is introduced so that, as the satellite moves away from the shortest range position, the frequency range of the chirped signal changes giving rise to curved "fringes" in the returning signals. A fuller account of this phenomenon is given in Barnett, Smith and Welford (1978).

Correction for Range Migration

The extent of the total path curvature produced by these various effects will be considered significant if the slant range changes by more than one half the slant range resolution during the time that the point target is in the processed part of the beam (Bennett et al, 1979, Raney, 1977). If this situation exists, then range migration is said to be present, and must be corrected for before the cross-correlation operation is performed on the SAR data. In addition to causing a geometric error, failure to correct for this range migration will degrade the quality of the processed image. Bennett and Cumming (1979) have derived an expression to enable a decision to be taken as to whether a range migration correction is necessary or not as follows:

$$\Delta S = \frac{\lambda V}{2R_a} \cdot \delta t$$

where λ = radar wavelength of transmission;

V = platform speed;

R_a = azimuth resolution; and

δt = the range of offsets from zero Doppler in the processed beam in the maximum beam yaw situation.

Then, if $\Delta S > \frac{1}{2} R_r$, then a range migration correction is needed. Using typical values for δt of 3 seconds for spaceborne satellites (Bennet et al, 1979); $V = 8{,}000$ m/sec; $\lambda = 0.3$ m; and $R_a = 25$ m, an error ΔS of 144 m. will result which shows clearly that the range migration correction is needed for spaceborne radar imagery.

The contribution of the Earth's rotation to the problem of range migration stems from the fact that there is a transverse component of relative velocity between the satellite and a given object point. This produces a linear variation in range as the object progresses through the synthetic aperture, so that the range history curve will be tilted. It also produces a constant Doppler shift; this latter adds a further tilt term to the range history curve (Fig 6.14d).

Geometrically, this Earth movement produces a distortion of rectangles into parallelograms (Barnett et al, 1978).

The extent of the contribution of the Earth rotation on Seasat SAR images was evaluated by Kratky (1980) where it was found that, for a full synthetic aperture length, the geometrical error induced by Earth rotation in the across-track direction amounts to 16 metres while in the along-track direction, the error amounts to less than one metre and could be ignored.

6.3.11 The Effects of Atmospheric Refraction on SLR Imagery

There is a steady variation in the air density of the Earth's atmosphere as altitude increases. This gives rise to a variation in the refractive index of the atmosphere. This variation in the refractive index tends to bend the microwave rays reflected from objects while they are travelling towards the radar antenna. In

practice, each ray is refracted towards the normal so that the path of radiation may take the form of a spatial curve of higher order. The question arises as to whether the extent of this refraction is appreciable on SLR imagery. For airborne applications, consideration of this effect may take into account two factors:-

(i) the curvature of the electromagnetic path; and

(ii) the velocity error in the travel time of the electromagnetic wave.

(i) Curvature of the Electromagnetic Path

Referring to Fig 6.15, the range that should be considered is the actual value S and not the measured value S_s. The result of using S_s is an error ΔS in the slant range given by:-

$$\Delta S = \frac{S_s^3}{24\rho^2} \qquad \text{(Leberl, 1972)}$$

where ρ is the radius of curvature of the ray (Fig 6.15). In practice, it is difficult to acquire an accurate value for ρ. However, it appears from the little evidence available in the literature that the actual curvature is small. In the case of most airborne SLR applications, ρ may be expected to be very large compared with S_s, consequently this error is normally neglected in such cases.

(ii) Variation of the Electromagnetic Wave Velocity

If the velocity of microwave propagation (c) used to compute S_s deviates from the true value, then an erroneous value for S_s will be obtained. The slant range S_s is given by:

$$S_s = \frac{c \cdot t}{2}$$

where c is the velocity of the microwave energy in a vacuum, taking the refractive index of the atmosphere as unity.

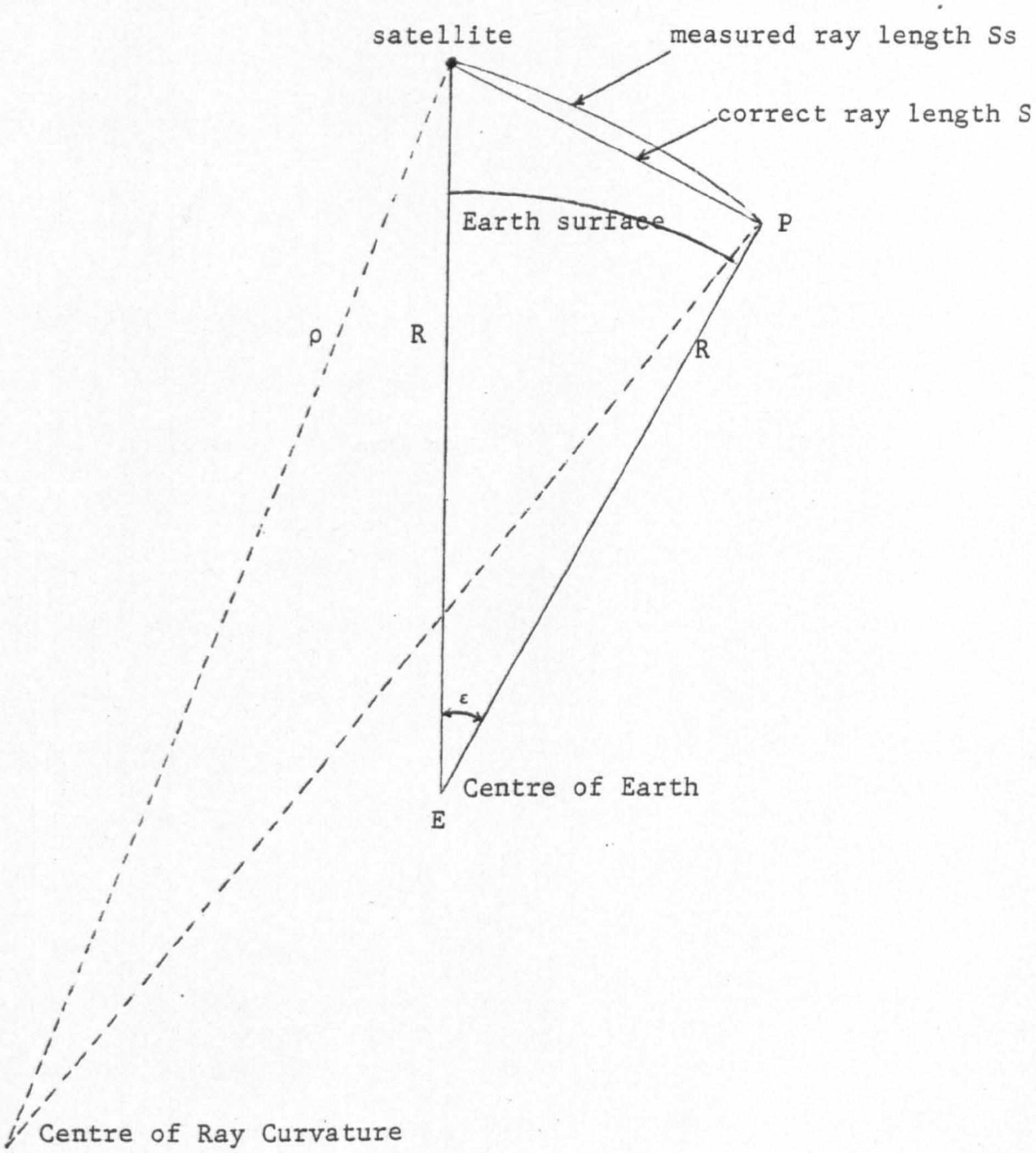


Fig 6.15 Geometry of Atmospheric Refraction

If c_o = true velocity of microwave energy to be used in the computation of S_s, then:

$$S_s = \frac{c_o}{n} \frac{t}{2}$$

where n is the true refractive index of the atmosphere ($\neq$ unity).

A deviation Δn in the value of the true refractive index will result in an error in the slant range ΔS given by:

$$\Delta S = \frac{-c_o}{n^2} \frac{t}{2} \Delta n = - \frac{S_s}{n} \Delta n$$

Again, in most airborne SLR surveys, Δn is a very small value so that the error resulting from this source is fairly small and can be neglected.

For spaceborne SLR systems, however, the situation is quite different. Since satellite radars will operate at ranges between 300 and 1000 kms, the SLR rays have to pass through the ionosphere which is present in the upper part of the atmosphere. This consists of a layer of charged particles - electrons and positive ions ionized by the ultraviolet radiation emitted by the sun. The profile of ion density in the ionosphere contains various concentrations in layers whose pattern is in a state of continual change from day to night, from season to season and from year to year. Solar activity such as Sunspots and solar flares can have a serious effect which is highly unpredictable in nature and extent. The path taken by the microwave energy emitted and received from the SAR will therefore differ substantially from the straight line geometric (or vacuum) path because of the considerable refraction induced by the ionosphere, thus causing a slant range error ΔS. At vertical incident angles, the refraction will be a minimum, but at the inclined angles which will be encountered in the case of an SLR, the refraction will increase depending on the angle of incidence (Fig 6.16).

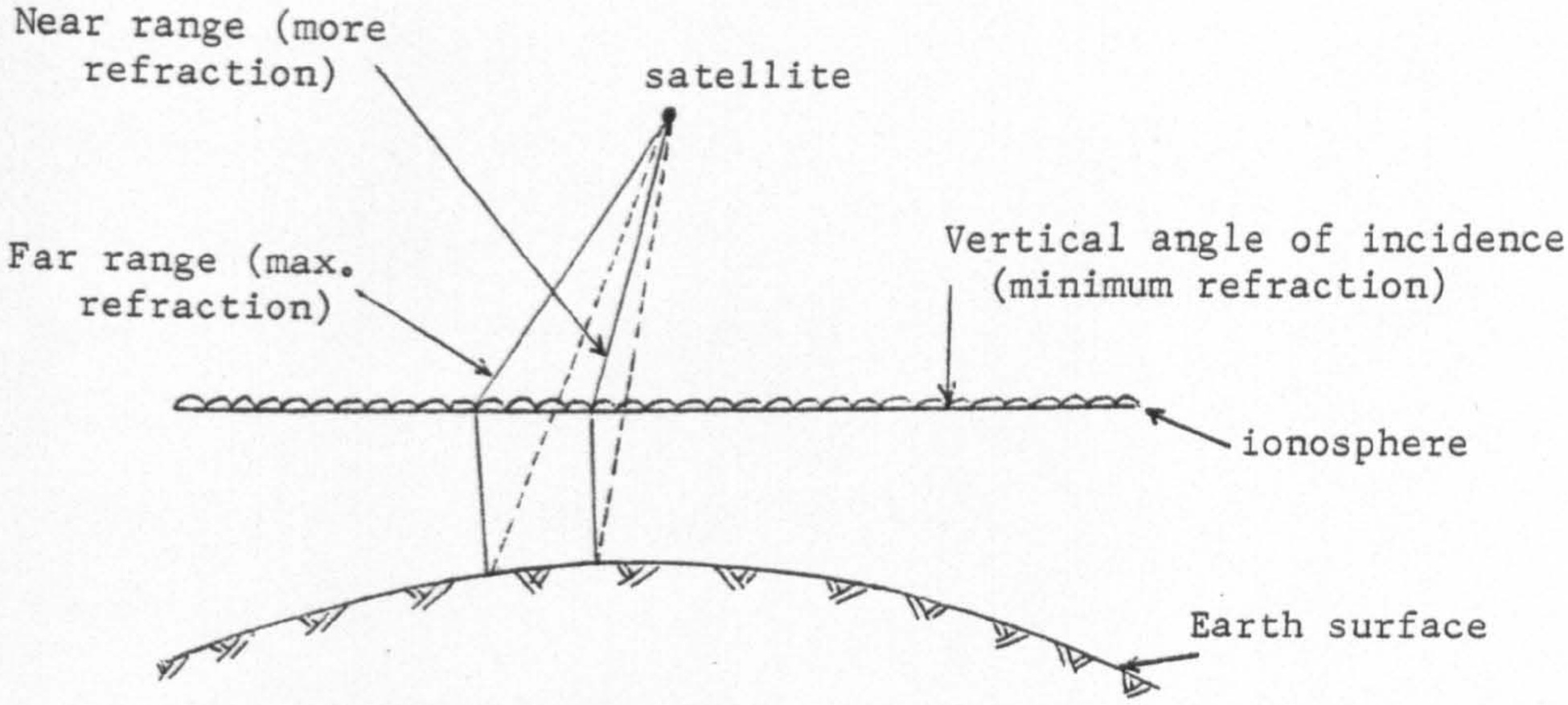


Fig 6.16 Geometry of Refraction on Spaceborne Radar Systems.

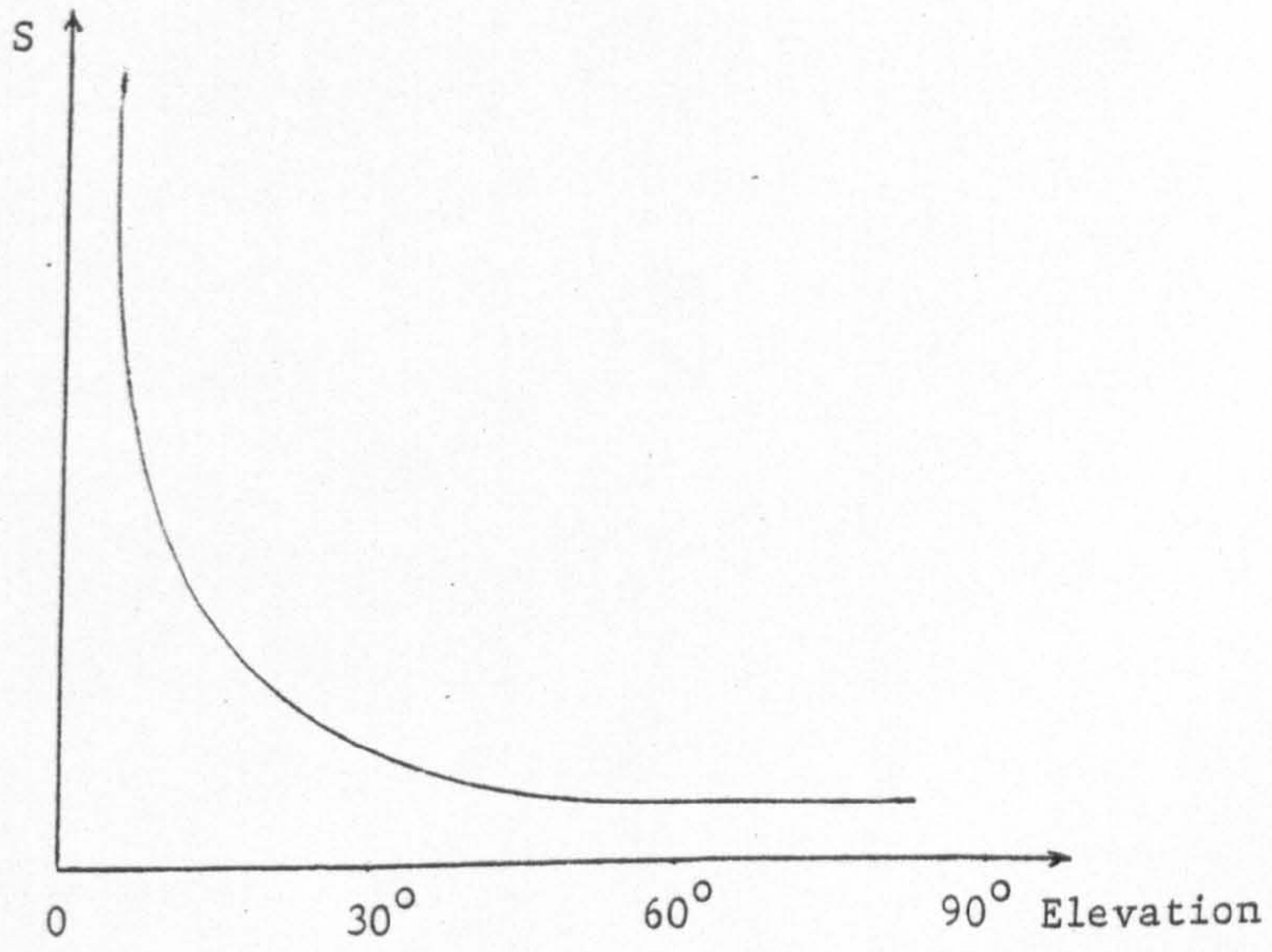


Fig 6.17 Range correction vs Elevation (not to scale)

(from Wells, 1974)

Two factors which will greatly influence the actual amount of refraction produced by the ionosphere are the frequency (and wavelength) of the microwave energy emitted by the SLR and the elevation. At low frequencies, e.g. 30 MHz, the ionosphere reflects the energy, while at higher frequencies, the microwave region propagation through the ionosphere is dispersive. Also at low elevation angles, the resulting error in range ΔS will be small, and larger at higher elevations (Fig 6.17).

The overall ionospheric contribution to the range error is given by Wells (1974) as follows:

$$\Delta S = \int (n-1)dS = \frac{b_1}{f^2} + \frac{b_2}{f^3}$$

where n is the ionospheric refraction index;

b_1, b_2 are variables which are not functions of frequency;

f = frequency of propagation; and

dS = the variable of integration.

The higher the frequency, therefore, the smaller will be the error in range. This is almost certainly one of the reasons for the choice of such a high frequency (1,000 MHz) for the Seasat SAR.

With SECOR and Doppler ranging carried out from geodetic satellites, measurements are normally made at a minimum of two frequencies so as to reduce the considerable errors resulting from the refractive effects of the ionosphere. The measurements resulting from each frequency can be compared and, since each is affected by refraction to a different extent, the comparison allows a much more accurate range to be determined. The improvement in the positional accuracy which results from such corrections to the ranges can be seen in the quoted figures for a single frequency satellite doppler device of $\pm$ 400 m for a single pass compared with $\pm$ 25 m for

a dual frequency device for a single pass. Since Seasat SAR is a single frequency ranging system, one must anticipate that large errors in the imagery will be induced by ionosphere refraction.

As is the case with Doppler satellite measurement techniques, the measured ranges must first be corrected for refraction. A number of formulae, corresponding to different approaches, exist for the correction of the range error caused by ionospheric refraction on single frequency systems (e.g. see Clynch, 1979, Wells, 1974). One of these is given by Wells (1974) as follows:

$$\Delta S = \frac{c}{f_g} D^{-c\left(\frac{f_g - f_s}{f_g}\right)} (t_2 - t_1) \quad \ldots\ldots\ldots\ldots\ldots 6.27$$

where c = velocity of propagation of electromagnetic waves in vacuum;

f_g = the reference frequency generated by the system;

D = an observed quantity corresponding to a phase difference measurement and contains the effect of the ionosphere;

f_s = the transmitted frequency of the radar system; and

$(t_2 - t_1)$ = a specified time interval.

The error is therefore a variable function of D and time interval $(t_2 - t_1)$.

The extent of this error on Seasat SAR imagery is not exactly known to the present author, but from experiences with other satellite ranging techniques, it would appear that it is not less than 50 m. Furthermore, as mentioned above, effects may vary widely and the extent of the refraction present during an individual flight is not an exactly predictable amount since it depends greatly on solar activity. Obviously, an estimated but substantial correction must be made to the measured ranges and this will have to suffice, since it is difficult to imagine that the necessary information will be available from

atmospheric scientists regarding the situation during a specific satellite orbit anywhere on the Earth.

6.3.12 Relief Displacement in SLR Imagery

In conventional photogrammetry carried out with frame cameras, the relief displacement is defined as the shift or displacement in the photographic position of an image caused by the relief of an object, i.e. by its elevation above or below a reference datum surface. With respect to this datum plane, relief displacement is outward for points whose elevations are above the datum and inward for points whose elevations are below the datum.

For SLR, however, relief displacement takes place in the directions opposite to those experienced in conventional perspective photography, i.e. inwards towards the radar antenna for points whose elevations are above the reference datum and outwards for points whose elevations are below it. This is because the pulse of electromagnetic energy emitted by the SLR will travel with a definite wavefront so that the top of a vertical object is the first point struck by the radar pulse. Since it gives a return signal before that produced from the bottom, it is imaged first and so gives the distinctive form of relief displacement of SLR in which the image is displaced inwards towards the flight line for an object lying above the datum plane.

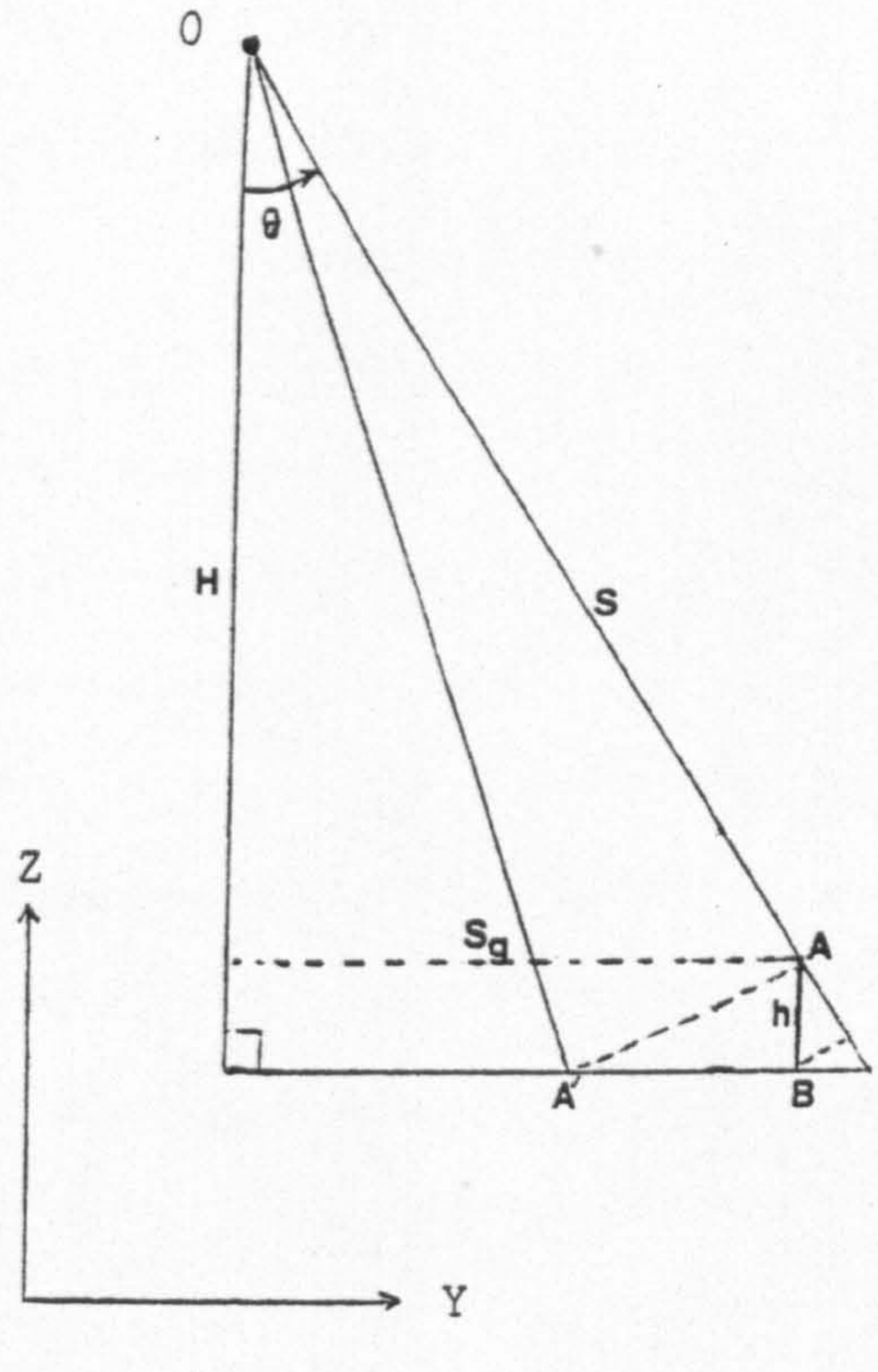
From Fig 6.18a, the image of point A on the vertical object AB will be displaced at A'. On a slant range presentation, the displacement is given by the distance $AB' = h \cos\theta \simeq h.\frac{H}{S}$.

where h is the height of object AB above the datum;

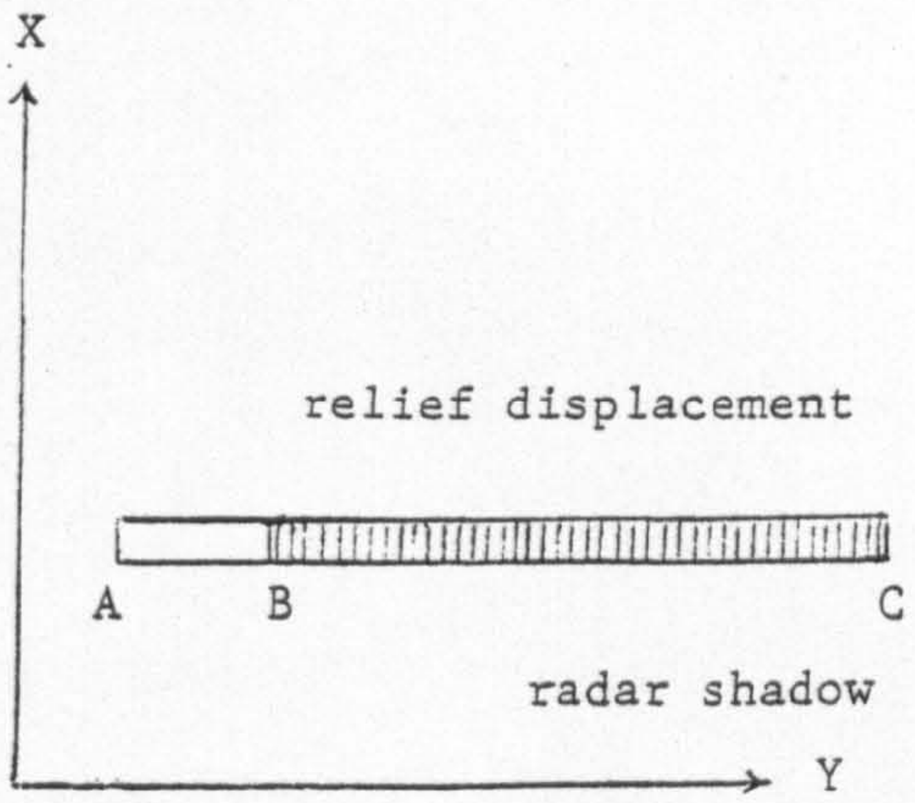
θ is the elevation angle; and

S is the slant range distance to the object.

With a ground range presentation, the displacement is given by BA' =



(a)



(b)

Fig 6.18 Geometry of SLR relief displacement and Radar shadowing

$$h \cot \theta \simeq h \cdot \frac{H}{Sg} . \qquad \text{---------- 6.28}$$

Thus in both cases, the displacement increases with decreasing elevation angles (and shorter ranges).

Associated with the relief displacement error is the radar shadow corresponding to the "dead areas" that the photogrammetrist encounters with terrain or buildings with wide angle or super wide angle aerial photography. The shadow length in SLR imagery is the distance BC = h tan θ (Fig 6.18b).

Thus it increases with increasing elevation angles and longer ranges. As can be seen, a radar shadow is not a geometric error but it creates a void in the image, which prevents any interpretation and measurement of the image in the area of the void. Needless to say, the presence of such a void in the image cannot be corrected for. For mountainous areas, radar relief displacements are usually the largest departures from a map-like geometry which are present on the radar image. They are a serious problem for the user since terrain height information has to be generated from some other source before any geometric correction can be made.

Certain configurations of side-looking radar have the capability of determining this height information from single radar images by using interferometric techniques (see e.g. Graham, 1974a). However, this has not yet been used operationally as far as civilian users are concerned. Recently, interest has been shown in the use of digital terrain models, which digitally store the elevation as a function of horizontal position. Height information from these data can be obtained to effect corrections for terrain relief on SLR imagery. While conceptually quite simple, this procedure is operationally quite difficult and attainable only with sophisticated computer

programs. Tests of this technique have been carried out by the present author and will be reported in Chapter X .

6.3.13 Radar Image Transformation Polynomials

Instead of attempting to compute the exterior orientation elements or their variations as discussed in section 6.3.5, another possible approach is to apply some form of interpolative procedure to allow rectification of the processed SAR image or to enable it to be fitted to the terrain coordinate system without directly involving either the imaging equations or the orientation parameters derived previously. Either quite arbitrary or specially selected transformation polynomials can be used to effect a direct transformation from the measured radar image coordinates to the terrain system in much the same manner as has been done with Landsat MSS and RBV imagery in recent years.

However, in the case of SLR imagery, some attempts should be made to select a set of polynomial parameters in the light of the knowledge of the basic imaging equations 6.16 and the known effects of the image displacements produced by dK, dØ, dbx, dby and dbz. By expressing these displacements as continuous functions of the image x-co-ordinates as in equations 6.17, applying them to equations 6.24 and taking the scale factor into consideration, the errors in the image coordinates dx and dy induced by these displacements will have the general form:

$$\left.\begin{aligned} dx &= \alpha_o + \alpha_1 x + \alpha_2 y + \alpha_3 xy + \alpha_4 x^2 + \alpha_5 x^2 y + \alpha_6 x^3 + \alpha_7 x^3 y \\ dy &= \beta_o + \beta_1 x + \beta_2 y + \beta_3 xy + \beta_4 x^2 + \beta_5 x^2 y + \beta_6 x^3 + \beta_7 x^3 y \end{aligned}\right\} \quad \ldots 6.29$$

Based on these equations, the relationship between the x,y image coordinates of the SLR system and the X,Y coordinates of the corresponding point in the terrain may be expressed as follows:-

$$\left.\begin{aligned} X &= n_o + n_1x + n_2y + n_3xy + n_4x^2 + n_5x^2y + n_6x^3 + n_7x^3y \\ Y &= m_o + m_1x + m_2y + m_3xy + m_4x^2 + m_5x^2y + m_6x^3 + m_7x^3y \end{aligned}\right\} \quad \ldots 6.30$$

where n_i, m_i are the polynomial transformation parameters;

X,Y are the terrain coordinates of a point; and

x,y are its corresponding image coordinates.

Equations 6.30 are the main polynomial transformations used in the experimental work described in this thesis. They have been programmed for the purpose of the digital rectification of Seasat imagery (described later in Chapter VIII). Their efficiency in transforming SLR imagery to the terrain has been evaluated by a comprehensive series of tests (described in Chapter IX).

CHAPTER VII

TOPOGRAPHIC MAPPING FROM SIDE-LOOKING RADAR (SLR)

7.1 Introduction

The period since 1972, when military security restrictions were lifted and SLR image data started to become available for civilian use, has been marked by a steady increase in the use of SLR imagery particularly for reconnaissance-type mapping carried out by field scientists at small scales (1:100,000 and smaller). On the one hand, very large areas in tropical equatorial countries have been imaged using both real and synthetic aperture radars. Thus the Amazon Basin area in particular has had extensive coverage, as have substantial parts of Central America, Nigeria, Indonesia and the Philippines. In all these cases, the areas imaged are those with continuous cloud cover for which it is difficult to acquire photographic cover for mapping. Thus the all-weather capability of the SLR is the feature which is primarily being exploited rather than the special responses produced by SLRs. To a large extent therefore, the radar imagery has been used as a map substitute, acting as the base for recording forest, vegetation, geological and other thematic information required for planning and development purposes. The techniques used for the rectification and mapping are often very simple or sometimes non-existent but, given the very small scale and the reconnaissance nature of the mapping, the lack of any alternative base and the urgency with which even these crude maps are needed, the results appear to meet the needs of many users.

At the other extreme, SLR imagery has been the subject of much research by a small group of photogrammetrists, especially Leberl,

Derenyi, Konecny and their associates, who have been concerned primarily with the development of techniques for geometrical rectification, usually digitally based and often reaching a high degree of sophistication and refinement. Associated with these developments have been a considerable number of tests carried out to establish the efficiency and accuracy of the algorithms and procedures devised for rectification. Usually however, these have been carried out over very small areas in North America or Western Europe where existing maps and control of a high quality are already available for the purposes of such tests.

Only in a series of projects in which Leberl has been involved - in Colombia and in the eastern United States - has there been a fusion of the requirement for rapid mapping of large areas with the use of techniques derived from the experimental research carried out by photogrammetrists.

As mentioned above, the largest application of side-looking radar has been the very small scale (1:100,000 and smaller) reconnaissance-type mapping of the cloud-infested areas of the world's equatorial belt.

7.2 Mapping Applications of Real Aperture SLR Imagery

The first extensive application of side-looking radar imagery to mapping was the work carried out in 1968 over the Darien Province of Panama. In spite of much effort, little photographic coverage of the area had been obtained over a period of many years. The area was covered by a real-aperture Ka band SLR which Westinghouse had developed for the U.S. Army, no less than 17,000 sq.km. being covered in six imaging hours (Crandall, 1969). The imagery was used to produce 1:250,000 scale radar mosaics.

In 1971, Westinghouse again conducted a large SLR survey in Central America, on this occasion over Nicaragua, in co-operation with Hunting Surveys. The whole country comprising approximately 148,000 sq.km. was covered in six days using a DC-6B aircraft flying at 6,000 m, the imaging often being carried out through three layers of cloud (Leatherdale, 1981). The acquisition scale of the imagery was 1:250,000, each strip being 20 km wide. These were used by Hunting to compile radar image mosaics at 1:100,000 scale for the 80,000 sq.km. covering the eastern and central parts of the country, together with a series of overlays to show the planimetric detail, land use, geology and geomorphology, vegetation, etc, for regional planning purposes. The whole work from flight to delivery of the mosaics and overlays took only seven months.

Since then, a number of real-aperture SLR systems have been used for mapping. In particular, the Motorola X-band system equipped with dual antennae allowing imaging on either side of the flight path has been used by Grumman Ecosystems and by Motorola Aerial Remote Sensing (MARS) Inc. A notable recent survey by MARS has been the coverage of the whole of Nigeria (967,000 sq.km.) resulting in mosaics based on the U.S. Joint Operations Graphics (JOG) 1:250,000 map series (Leatherdale, 1981; Parry and Trevett, 1979). Once again, the interpretation, principally for vegetation and forestry, has been carried out by Hunting Surveys with the aid of conventional aerial photography and Landsat MSS imagery wherever available.

7.3 Mapping Applications of Synthetic Aperture SLR Imagery

The greatest SLR coverage anywhere in the World is of the Amazon

Basin. The relevant parts of the Basin belonging to Brazil, Colombia and Peru have all been covered using the X-band synthetic aperture radar (GEMS-1000) operated jointly by the Goodyear Aerospace Corporation and the Aero Service division of the Western Geophysical Corporation and mounted in a Caravelle jet aircraft with a Litton INS navigation system. Large areas in the U.S.A. have also been covered by SAR imagery, taken both by the civilian Goodyear system and by the U.S.A.F.

The work undertaken in three extensive mapping projects - Project RADAM (Brazil), Project PRORADAM (Colombia) and one carried out over the area of West Virginia, Kentucky, Ohio, etc (U.S.A.) - will be outlined here as representative of the techniques employed. All three surveys were based on imagery acquired by the Goodyear/Aero Service system.

7.3.1 Project RADAM

Initially, an area of about 1,500,000 sq.km. covering the eastern part of the Amazon Basin was included in the project, but later this was extended to cover the whole of the Brazilian part of the Basin amounting to 4,500,000 sq.km (Van Roessel and Godoy, 1974). The success of these projects lead to the SAR survey being extended to cover the entire Brazilian territory of 9 million sq.km (Leberl, 1976b).

Ground control was provided mainly by the Inter-American Geodetic Survey (I.A.G.S.) using stations fixed by satellite doppler methods. Transponders were then placed at each station to allow the establishment of the air station coordinates of the aircraft every few seconds using the SHORAN trilateration method. Because of cost considerations, this method of continuous tracking was employed only in the initial stages of the RADAM project. In the rest of the project, the main North-South

flights were untracked, control being provided by East-West flown tie strips which were tracked using SHORAN. Thus, an individual block of North-South flights was controlled by the two tie strips, one located at each end of the block (Fig 7.1).

7.3.2 Project PRORADAM

This basically similar project was undertaken in Colombia to give coverage over the area of 360,000 sq.km. of that country's Amazonian territory, again to provide simple radar mosaics of the area (Leberl, 1974, 1975c, 1975e). However, under the influence of Leberl, a rather different procedure was followed to that used in the Brazilian project. The flying height was again 12.5 km with a North-South flying direction with the antenna looking west. However, 60% overlap was adopted so that stereo-viewing could be contemplated. Also, three East-West transverse tie flights were flown to control the whole block of imagery. A total of 70 flights was completed, 30 of which had to be re-flown. Using two CRTs to improve resolution, each strip was divided into two parts, the near range (NR) and the far range (FR) images with a 5% overlap between them to allow the selection of tie points to allow the re-combination of the separate parts of the strip into a single image (Fig 7.2). SHORAN tracking was not employed. Instead forty four ground control points were measured using satellite doppler techniques and traversing. The rest of the control was provided by radar-grammetric block adjustment techniques. In this, the measurements of common points in the overlap areas of adjacent strips were used to tie the individual strip images together into a block, and the block coordinates were then transformed into ground coordinates using the network of ground control points in a procedure similar to that of a photogrammetric block adjustment.

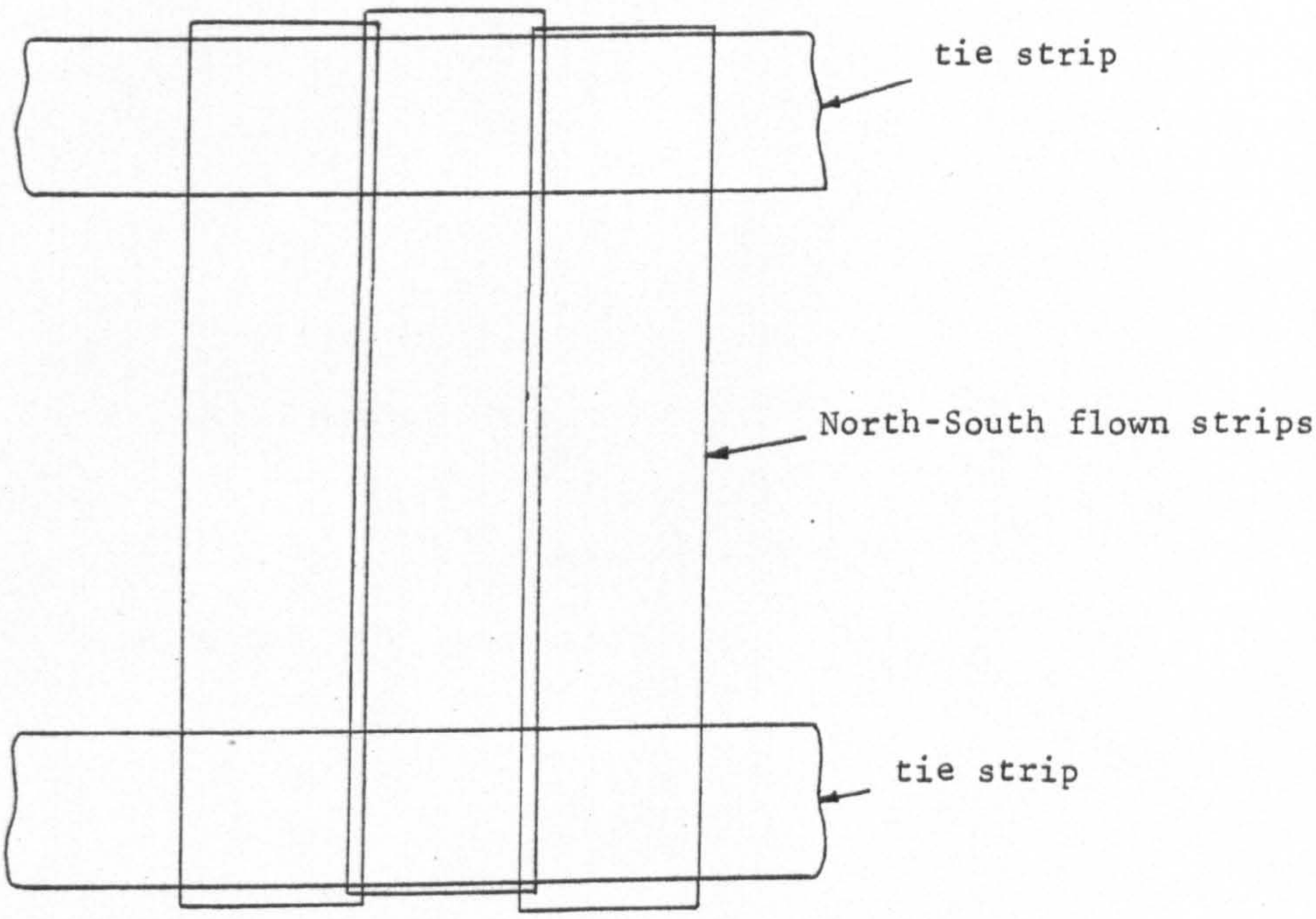


Fig 7.1 Tie strips to control an SLR block

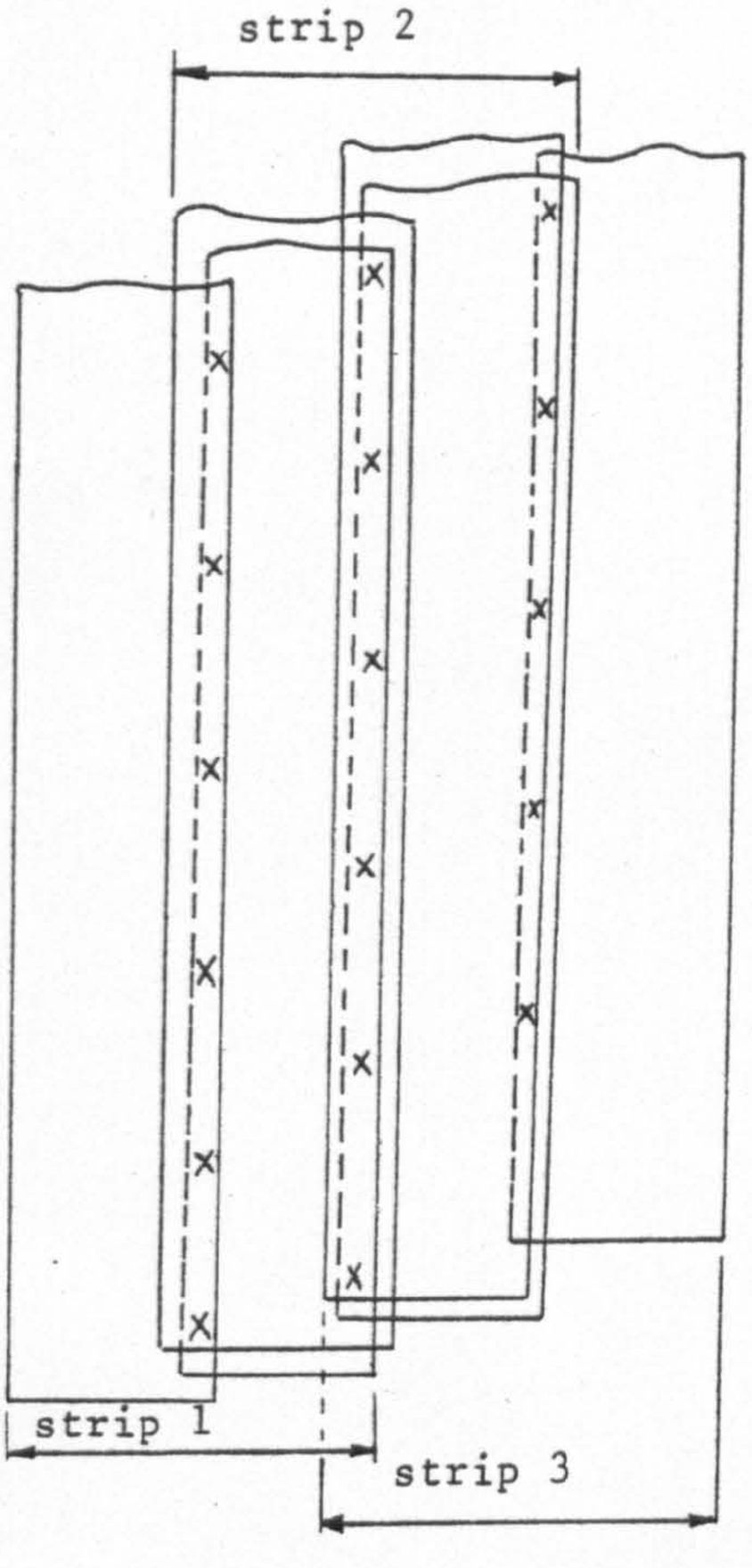


Fig 7.2a SLR strips with 60% overlap but split into near and far range views.

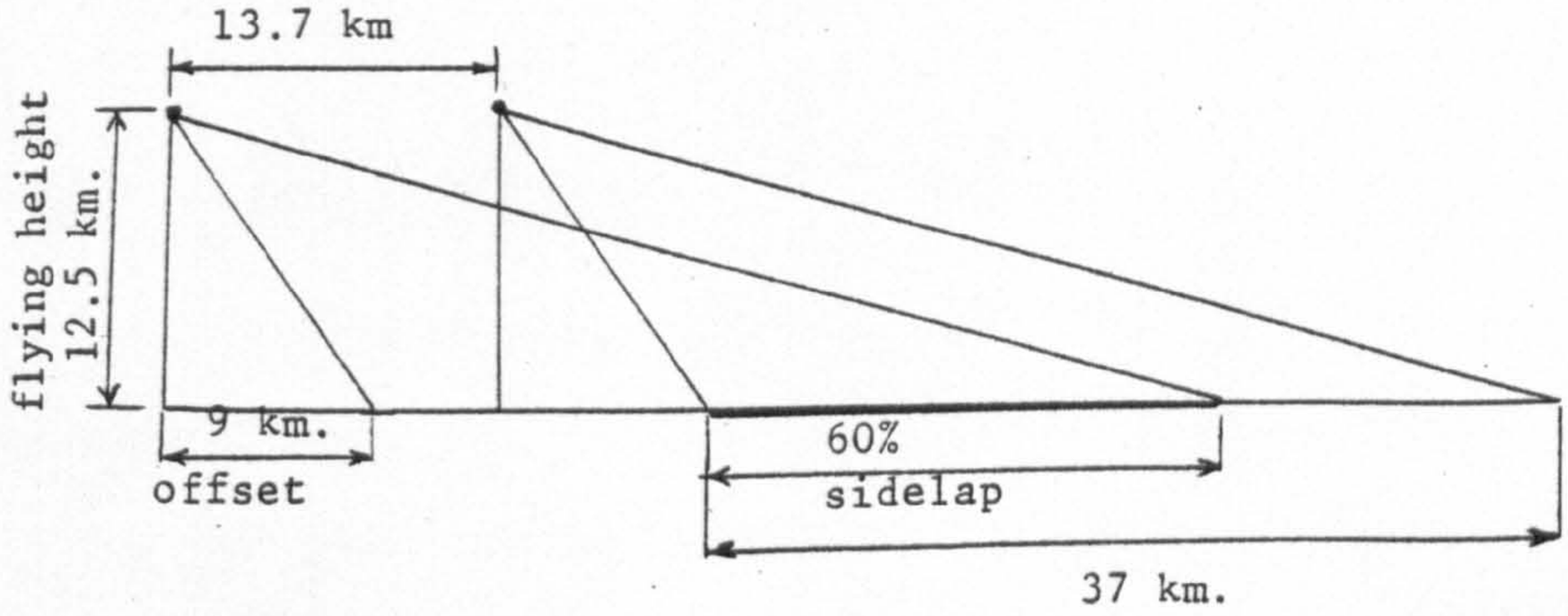


Fig 7.2b SLR imaging geometry to obtain 60% overlap

7.3.3 Appalachian Project (U.S.A.)

In this project, a block of twenty four overlapping SAR strips was flown over a well-mapped area of about 300 x 300 km^2 (90,000 sq.km.) covering parts of Ohio, West Virginia, Virginia and Kentucky for the specific purpose of evaluating the mapping accuracy achieved in radar mapping projects. The flying height used was again H = 12 km, but the sidelap used was only 25% (as compared with 60% in PRORADAM). Large numbers of control points and check points were available on the 1:24,000 and 1:62,500 maps, from which their coordinates were accurately scaled by coordinatograph. A variety of different techniques were evaluated using this large test block (see Leberl et al, 1976d).

7.4 Radar Mosaicing Procedures

Based on these large projects, it is possible to distinguish four different procedures by which the mapping can take place:-

(i) uncontrolled mosaics;

(ii) controlled mosaics based on SHORAN tracking;

(iii) semi-controlled mosaics using tie lines; and

(iv) controlled mosaics based on numerical radargrammetric block adjustment.

7.4.1 Uncontrolled Mosaics

The SAR ground range image is simply accepted in the form it is generated in the optical correlator. The individual strips are laid out on a base board and the detail in the common overlaps fitted as best as possible. Any control which is available is likely to be sparse and is used to give a rough scaling, rotation and positioning to

the mosaic in terms of the map coordinate system. The accuracy (r.m.s.e.) of check points determined by Leberl et al (1976d) using this procedure was $\sigma_X = \pm 1.59$ km. and $\sigma_Y = \pm 2.28$ km. respectively, equivalent to a planimetric accuracy of ± 2.8 km.

7.4.2 SHORAN Controlled Mosaics

The overall procedure used in project RADAM has been mentioned above in section 7.3.1. The X-coordinates of the control points are those derived from the SHORAN tracking; the Y-coordinates are the measured air station coordinates in Y to which the ground range offsets are added to give the coordinates of specific range marks generated at intervals on the image. These control points are plotted on a base board at 1:200,000 scale. Sections of the negative image are then projected to fit the plotted control - the projected image passing through an anamorphic lens to deal with the affine scale changes between the X and Y axes which are typical of SLR images. The accuracy (r.m.s.e.) derived from the tests carried out in the RADAM project was $\sigma_X = \pm 0.19$ km. and $\sigma_Y = \pm 0.31$ km.

7.4.3 Semi-Controlled Mosaics Using Tie Lines

This was the procedure followed in the later stages of the RADAM project using East-West flown tie lines to control the extremities of a block of North-South flown strips. The control points will be located along these tie strips derived either by SHORAN (as in RADAM) or from existing maps (as in the Appalachian tests). These points are first plotted on the base sheet and the tie strips fitted to them. Next, the main production strips are fitted to these by optical enlargement with the mosaicing attempting to ensure as good a correspondence and

fit as possible between the strips. The accuracy of this procedure as tested by Leberl et al is $\sigma_X = \pm 0.5$ km. and $\sigma_Y = \pm 0.33$ km. Obviously, it is difficult with this method to deal with differential along-track scale errors and with the curvature of the main strips, e.g. arising from systematic effects produced by the INS navigation system.

7.4.4 Numerical Radargrammetric Method

This technique devised by Leberl (1975e) will be described as employed in the Colombian PRORADAM project. Points located at intervals in the 5% overlap zone between the near range and the far range views are first used for unification of the individual strips. Once these have been formed, the same points may be used to connect the strips to form a single block of SAR imagery (Fig 7.2a), the coordinates of all tie points being given in a single image coordinate system. Each tie point will appear therefore on four individual images. Beside the tie points, all the available ground control points are also identified and brought into the block coordinate system. The result of the whole procedure is to produce a single set of coordinates for each tie point. Finally, the block image coordinates are fitted to the ground control points using at first a linear conformal transformation. The resulting discrepancies are then corrected for using an interpolative procedure such as polynomials, linear prediction or moving average techniques.

In the tests of the numerical radargrammetric method carried out by Leberl et al (1976d), different configurations of ground control gave rise to different patterns of errors. The accuracies obtained ranged from $\sigma_X = \pm 0.24$ km. and $\sigma_Y = \pm 0.37$ km. when only four control points were used (one at each corner of the whole block) to $\sigma_X = \pm 0.10$ km. and $\sigma_Y = \pm 0.13$ km. when perimeter control was available, supplemented by a few points located within the block.

Obviously, the most significant result of these experiments was that the use of a numerical radar block adjustment procedure carried out prior to mosaic compilation gave results vastly superior to those obtained using SHORAN control or tie lines and with a great economy in the provision of control points and a consequent saving in expense.

7.5 Numerical Techniques for SLR Rectification

As mentioned above, a considerable amount of experimental work has been undertaken to investigate the techniques of rectifying individual SLR images by measurement of their image coordinates and their subsequent transformation into terrain coordinates by numerical methods. This has been carried out using a number of points whose coordinates are known, some of these being used as the control points on which the transformation has been based, the remainder being used as check points to establish the accuracy of a particular transformation or procedure. This approach has been followed by Derenyi (1974a, 1975), Konecny (1972) and Leberl (1976a, 1980).

The transformations used vary from the simple linear conformal through affine to more complex polynomials such as those utilised by Konecny (1972) and Derenyi (1974b) which take the form:

$$\left.\begin{aligned} X &= a_o + a_1x + a_2y + a_3xy + a_4y^2 + a_5xy^2 + a_6x^2y; \quad \text{and} \\ Y &= b_o + b_1x + b_2y + b_3xy + b_4y^2 + b_5xy^2 + b_6x^2y \end{aligned}\right\} \quad \ldots\ 7.1$$

Once the parameters a_o, ..., a_6, b_o, ..., b_6 have been determined, they can be used to generate terrain coordinate values (X, Y) for any other image position (x, y). The transformed coordinates can then be compared with the known values (X, Y) to allow an assessment of the efficiency and accuracy of a particular transformation. The results

of tests using these various transformations on airborne SLR imagery are shown in Table 7.1.

It is noticeable that, in the years prior to 1974, all the tests have been carried out on real-aperture SLR imagery with simple linear conformal transformations used for image-to-map coordinate conversion. Using such a transformation, the planimetric accuracies obtained with images taken over flat terrain amount to nearly $\pm$ 100 m, while those obtained with images taken over hilly or mountainous areas range between $\pm$ 260 m to nearly $\pm$ 300 m.

From 1974 onwards, increasing attention has been given to the synthetic aperture (SAR) type of imagery. In most cases, the improvement in the planimetric accuracy obtained using the linear conformal transformation more than doubled as compared with those obtained using real-aperture imagery over flat terrain. It is also noticeable that the 6-term polynomial fit applied to the Goodyear GEMS-1000 imagery did not alter the results significantly.

Another approach which has been investigated utilizes <u>interpolative methods</u> of reducing errors in conjunction with the basic transformation. Thus, once the initial transformation has been carried out, the unknown parameters determined and the residual discrepancies at the control points computed, the interpolative method is used to provide corrections to the transformed coordinates of any other measured point. In particular, Leberl (1972) has pursued this approach vigorously using both

(i) point-wise; and

(ii) piece-wise interpolative methods.

The measured SLR image coordinates are first transformed to their terrain values using a transformation. The particular transformation selected and used has of course a great influence on the accuracy of the inter-

TABLE 7.1 MAPPING ACCURACIES OBTAINED WITH DIFFERENT TRANSFORMATIONS

Investigator	Type of radar system	Transformation technique used	σ_X (m)	σ_Y (m)	σ_P (m)	Type of terrain and area of investigation
Bosman (1971)	Real-aperture	Linear Conformal	±80	±46	±92	very flat (Arnhem-Nijmegen, Netherlands)
Leberl (1972)	"	"	±62	±76	±88	very flat (Noordoostpolder, Netherlands)
Konecny (1972)	"	"	±151	±255	±296	mountainous (New Guinea)
Derenyi (1974a)	"	"	±154	±211	±261	moderate relief (Washington D.C., U.S.A.)
Konecny (1974)	Synthetic aperture*	"	±46	±65	±80	moderate relief (Phoenix, Arizona, U.S.A.)
Goodyear (1974)	"	* "	±38	±30	±48	"
Derenyi (1975)	"	* "	±28	±30	±41	"
Derenyi (1975)	"	* 6-term polynomial	±28	±30	±41	"
Bosman and Clereci (1975)	"	* 6-term polynomial	±45	±60	±75	"
Derenyi (1980)	"	6-term polynomial	-	-	±19	moderate relief (Quebec, Canada)

*Imagery acquired by Goodyear/Aerospace GEMS-1000 system.

polative technique since quite different residual errors will be produced at the control points depending on which transformation has been used initially.

(i) The point-wise procedures involve the computation of the corrections to be applied at any new image point on the basis of its position (and therefore its respective distances) relative to the surrounding control points on which the interpolation is based. One of the widely used point-wise procedures is that of the arithmetic mean. To each pair of image coordinates x,y of a new point, corrections are applied which are a weighted mean of the given discrepancies v_{x_c} and v_{y_c} of the N surrounding nearby control points, the weighting being a function of their distances from the new point. The computation of corrections at a particular point i, is carried out using the formulae:

$$v_{x_i} = \frac{\sum_{i=1}^{N} Q_i \, v_{x_c}}{\sum_{i=1}^{N} Q_i}$$

$$v_{y_i} = \frac{\sum_{i=1}^{N} Q_i \, v_{y_c}}{\sum_{i=1}^{N} Q_i} \qquad \text{..... 7.2}$$

where Q_i is a weight function chosen according to the following formulae:

$$Q = \frac{1}{D^j} \qquad \text{or} \qquad Q = \frac{1}{(1+D)^j}$$

where D = the distance between the control point and the new point; and

j is a constant.

(ii) The piece-wise procedure can take various forms, the best known being the use of piece-wise polynomials. An individual polynomial is utilized for each section (e.g. a triangle or a quadrilateral) of the SLR image lying between a group of control points. To avoid sudden jumps or discontinuities in the corrections between one section and its neighbour, a joining function is employed so that a smooth transition will take place at the junctions between sections.

The particular polynomial used is chosen so that it best describes the errors on the SLR imagery. Normally it will take the form of one of the equations given in 6.29 but usually there is no need for it to be of very high order. If the area imaged is fairly flat, an individual strip can be divided in the along-track direction into sections defined at the edges by control points so that a number of pieces are formed. A whole series of polynomials are fitted, one to each section or piece with joining conditions imposed to ensure a smooth transition between two neighbouring sections. The joining functions may take the following form:

$$
\left.
\begin{aligned}
(a_o + a_1x + a_2x^2)_j &= (a_o + a_1x + a_2x^2)_{j+1} \quad ; \\
(a_3 + a_4x + a_5x^2)_j &= (a_3 + a_4x + a_5x^2)_{j+1} \quad ; \\
(b_o + b_1x + b_2x^2)_j &= (b_o + b_1x + b_2x^2)_{j+1} \quad ; \\
(b_3 + b_4x)_j &= (b_3 + b_4x)_{j+1} \quad ; \text{ and} \\
(b_5)_j &= (b_5)_{j+1}
\end{aligned}
\right\} \ldots 7.3
$$

where j denotes the number of a particular section on the image.

If, on the other hand, the area imaged is hilly or mountainous or if the mesh of control points has an irregular distribution then a surface fitting (as opposed to curve fitting above) may be employed.

These may take the form of planes, e.g. a triangular plane joining three points or a more complex curved surface. In this case, the terms describing the surface will be proportional to x ; y ; xy ; x^2 ; y^2 ; x^2y and xy^2 and more complex conditions will have to be imposed in order to effect the smooth transition from one surface to another.

The results obtained by Leberl (1972) over the very flat area of Noordoostpolder, Netherlands using (i) a linear conformal transformation followed by (ii) two of these interpolative procedures are as follows:

Interpolative Procedure	σ_X (m)	σ_Y (m)	σ_P (m)
None (Linear Conformal)	$\pm$ 62	$\pm$ 76	$\pm$ 98
Arithmetic Mean	$\pm$ 73	$\pm$ 48	$\pm$ 88
Piece-wise Polynomial	$\pm$ 62	$\pm$ 28	$\pm$ 68

This shows that, for the circumstances of this particular test, the implementation of the piece-wise procedure can result in a better accuracy which may justify the extra computational work required by the method.

It will be seen later that the general philosophy of the type of testing carried out and summarized in table 7.1 has been applied by the present author on representative samples of the Seasat SAR material. Thus the results given in this chapter offer a direct comparison with those produced as a consequence of the experimental work carried out in the present study.