AD-785 172

COMPUTER COMPARISON OF PICTURES

Lynn H. Quam

Stanford University

Prepared for:

Advanced Research Projects Agency
National Aeronautics and Space Administration

May 1971

STANFORD ARTIFICIAL INTELLIGENCE LABORATORY
MEMO AIM-144

STAN-CS-71-219

COMPUTER COMPARISON OF PICTURES

BY

LYNN H. QUAM

SUPPORTED BY
THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AND
ADVANCED RESEARCH PROJECTS AGENCY
ARPA ORDER NO. 457

MAY 1971

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

STANFORD ARTIFICIAL INTELLIGENCE LABORATORY
MEMO AIM-144

MAY 1971

COMPUTER SCIENCE DEPARTMENT
REPORT NO. STAN-CS-71-219

COMPUTER COMPARISON OF PICTURES

by

Lynn H. Quam

ABSTRACT: This dissertation reports the development of digital computer techniques for detecting changes in scenes by normalizing and comparing pictures which were taken from different camera positions and under different conditions of illumination. The pictures are first geometrically normalized to a common point of view. Then they are photometrically normalized to eliminate the differences due to different illumination, camera characteristics, and reflectance properties of the scene due to different sun and view angles. These pictures are then geometrically registered by maximizing the cross correlation between areas in them. The final normalized and registered pictures are then differenced point by point.

The geometric normalization techniques require relatively accurate geometric models for the camera and the scene, and static spatial features must be present in the pictures to allow precise geometric alignment using the technique of cross correlation maximation.

Photometric normalization also requires a relatively accurate model for the photometric response of the camera, a reflective model for the scene (reflectance as a function of the illumination view, and phase angles) and some assumptions about the kinds of reflectance changes which are to be detected.

(cont'd)

This research was supported (in part) by the Advanced Research Projects Agency of the Office of the Secretary of Defense under Contract SD-183, and (in part) by the National Aeronautics and Space Administration under Grant No. NGR-05-020-508. The work by Bob Tucker was supported (in part) by JPL sub-contract No. 952489.

Reproduced in the USA. Available from the National Technical Information Service, Springfield, Virginia 22151.

These techniques have been incorporated in a system for comparing Mariner 1971 pictures of Mars to detect variable surface phenomena as well as color and polarization differences. The system has been tested using Mariner 6 and 7 pictures of Mars.

Although the techniques described in this dissertation were developed for Mars pictures, their use is not limited to this application. Various parts of this software package, which was developed for interactive use on the time-sharing system of the Stanford Artificial Intelligence Laboratory, are currently being applied to other scenery.

ACKNOWLEDGEMENTS

The research described in this dissertation would have been difficult, if not impossible, without the contributions of many people and, of course, the PDP-10 computer system of the Stanford Artificial Intelligence Project.

I would first like to thank my advisor, Professor John McCarthy, who added inspiration and guidance to this project. His efforts made possible the Artificial Intelligence Project at Stanford.

Dr. Elliot Levinthal contributed greatly by providing the interface to the Mariner 1971 Mars Mission, and the Variable Features Team. His interest and the interest of Jim Pollack and Carl Sagan, who are members of that team, provided the most of the motivation to develop this system.

I wish to thank Bob Tucker for his help in developing software to convert Mariner 6 and 7 images from the Jet Propulsion Laboratory (JPL) format to our local format, and in developing software to compress images. Bob also served as an interface with people at JPL.

Special thanks goes to Tom Thorpe and other people of the Image Processing Laboratory at JPL who made the Mariner '69 data available.

Without the programming support of many other people at the Artificial Intelligence Laboratory, this system would never have been written. In particular, I am indebted to Dan Swinehart and Bob Sproull, the co-developers of SAIL.

I must thank all of the hardware and software people, whose efforts have produced a valuable, although sometimes unreliable, system for research.

This project would never have started without the initial inspiration and interest of Bruce Baumgart, who made the first contacts with people of the Mariner '71 Mission and JPL.

I wish to extend particular thanks to Marsha Jo Hannah, who worked patiently, for long hours, reading and helping to revise this dissertation.

TABLE OF CONTENTS

APPENDIXES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

The general picture comparison problem is to geometrically and photometrically normalize and register images so that true differences in the scene can be determined, rather than differences in conditions of viewing, and to analyze these differences.

The geometric analysis of photographs belongs to the science of analytical photogrammetry (Doyle [1966]), which generally deals with the triangulation of aerial photographs to generate topographic maps. Most results in this field, however, are approximations which solve suitably linearized versions of the equations. Even neglecting this deficiency, these techniques are of little use except when extremely accurate models exist for the camera.

Military aerial reconnaisance probably represents the major current application of picture comparison technology. Most of the known techniques in this field rely on manually operated, analog image correlators, and various other analog hardware. There is little in the literature to describe what digital computer techniques, if any, have been developed for this purpose.

The system described in this dissertation is the only system known to solve the problems of image normalization and registration.

I.A THE MARINER 1971 MARS MISSION [1]

The primary objectives of the Mariner Mars 1971 Project are the observation and mapping of Mars by two orbiting spacecraft beginning in November of 1971 and continuing for at least 90 days (JPL [1970]).

An orbiter has the advantage over flyby missions, such

---

[1] Mariner 8 was launched from Cape Kennedy on May 8, 1971, and failed to make Earth orbit. Therefore, the nominal Mariner '71 mission plans must be changed. It is not known at this time whether the Variable Features Mission, which is the purpose of this research, will be attempted. Since most of this dissertation was written before May 8, many of the verb tenses referring to the Mariner '71 Mission should be changed to reflect the current situation.

as the Mariner Mars 1964 and 1969 missions, that observations from orbit can be made over an extended period of time, thus permitting the study of temporal changes on the Martian surface.

The Mariner IV, VI, and VII missions could make only limited observations of the surface of Mars because they were in the planet's vicinity for a very short time. To accumulate more extensive data it is necessary to orbit Mars as planned in the Mariner '71 Mission, or orbit and land as planned in the 1975 Viking mission.

The two identical Mariner '71 spacecraft will perform separate missions designated as missions A and B. Mission A is primarily devoted to routine mapping, attempting to view a large portion of the surface of Mars with the highest possible resolution.

Mission B is primarily devoted to studying time variable features of the Martian atmosphere and surface. It will utilize an orbit which will give repeated coverage of several different surface areas under essentially constant illumination angle, view angle, and spacecraft altitude (Sagan [1969]). With such an orbit, it is possible to study these areas of the surface to detect changes during the 90-day mission.

The types of changes which are expected are either transient or diurnally or seasonally recurrent. Transient phenomena that have been observed on Mars include clouds, hazes, and bright spots. White clouds have been seen of all sizes and shapes, from terminator haze lasting a few hours to dense, 1200-mile giants lasting days or weeks. Yellow clouds vary from small, dense, orange or yellow objects lasting from one to a few days, to objects which start large and grow larger until they become a yellow veil covering most of Mars and lasting a month or more. These yellow clouds are almost universally accepted as being dust clouds. The Mariner '71 television observations will yield data on atmospheric circulation by following cloud movement.

The most obvious seasonal changes on the Martian surface involve the polar caps and the wave of darkening. The surface also exhibits seasonal changes in color and even changes in the size, shape, and internal appearance of the various dark areas on the planet.

The polar caps are believed to be deposits of solid carbon dioxide condensing during the fall and winter in each hemisphere and then subliming during the spring and summer. From many biological points of view, the receding polar cap is a locale of great interest, and will be observed during the Mariner '71 mission.

The wave of darkening is probably Mars' greatest enigma and is the most dynamic event on the planet. It has been described as a progressive decline in the reflectivity of the dark surface areas (and increase in contrast with surrounding bright areas) starting in local springtime from the edge of the vaporizing polar cap, and moving toward and across the equator. Whether the darkening actually occurs as a "wave" from the pole has been argued. This darkening might also be accompanied by color and polarization changes. According to the biological explanation for these seasonal changes, Martian organisms inhabit the dark regions, and their springtime growth in response to the increased temperatures and humidities is the cause of the darkening events. Several alternative non-biological hypotheses have been proposed, including one in which seasonal changes in wind patterns (due either to meridional circulation or dust-devils) redistribute the particle sizes in the bright and dark areas and produce the albedo changes.

In selecting scientific objectives for the Mariner '71 Mission, the study of the wave of darkening was singled out for special consideration, since during the 1970 to 1980 decade, the wave of darkening in the southern hemisphere can only be observed in 1971. The southern hemisphere contains most of the permanently dark regions and is considered the more interesting hemisphere in which to study this phenomenon.

The two spacecraft are identical and contain a variety of scientific instruments including two television cameras, an ultraviolet spectrometer, an infrared interferometer, and an infrared radiometer. In terms of volume of data collected, the television experiment data will outweigh all other data by several orders of magnitude.

The two television cameras are designated as cameras A and B. The A camera has a focal length of 50 mm, giving it an 11 by 14 degree angular field of view. The B camera has a focal length of 500 mm giving it a 1.1 by 1.4 degree field of view and hence 10 times the resolution. Camera A has a collection of 8 color and polarization filters which can be arbitrarily selected. The B camera has a single fixed (minus blue) color filter. The wave of darkening and other variable surface features will be studied primarily using A camera pictures which are taken before periapsis when the sun and spacecraft vectors are nearly vertical to the surface in the area being photographed, thus minimizing uncertainties due to light scattering. Most of these pictures will be taken under an orange color filter which, on the basis of ground observations, should make it easiest to detect changes in albedo.

The television cameras are vidicons, which are shuttered at speeds ranging from 6 to 192 milliseconds. The image on the vidicon target is digitized in 42 seconds in a format of 9 bits per point, 832 points per line, and 700 lines per frame. Digitized samples are stored on a digital magnetic tape recorder which can hold approximately 36 pictures until it is possible to transmit the pictures to Earth.

Each spacecraft has a radio receiver which is used to control such things as propulsion, scan platform slewing, color filter selection, camera shuttering, and tape recorder playback (and transmission). Each spacecraft also contains a 10 or 20 watt 2300 MHz radio transmitter which during most of the mission will be able to transmit experiment and spacecraft data to the 210 foot Goldstone antenna at a 16 kbps (kilo-bits per second) rate. At this rate, each picture requires about 5 minutes transmission time. A daily tape recorder load consists of about 36 pictures, which will require a transmission time of 3 hours. This is about the longest period of continuous communication between the spacecraft and Goldstone which can be guaranteed on every orbit. (Briggs [1971]).

To maximize the scientific return from the Mariner '71 mission, it is necessary to control the picture taking sequence to concentrate surface coverage in those surface areas which show the greatest variability. This requires that pictures received at Earth be analyzed as quickly and sensitively as possible to detect differences from previous pictures of the same area. This dissertation is the result of research in the development of digital computer techniques to perform such analysis on pictures from the Mariner '71 mission.

I.B A SOLUTION TO THE MARS PICTURE COMPARISON PROBLEM

In response to the needs of the Mariner '71 Variable Features Team, image processing techniques have been developed by the author at the Stanford University Artificial Intelligence Project, utilizing a PDP-10 interactive time-sharing computer system. These techniques have been integrated into a system which compares pairs of images taken at different times, from different spacecraft positions, and perhaps even different spacecraft.

Figure 1-1 shows the structure of the system. In this figure, the rectangular blocks indicate programs which implement the techniques described in this paper. Where appropriate, the block contains the chapter or section number where the technique is described. Oval blocks indicate data (usually pictures).

GEOMETRIC NORMALIZATION: Since the images are different perspective views of a spheroid, with no surface elevations

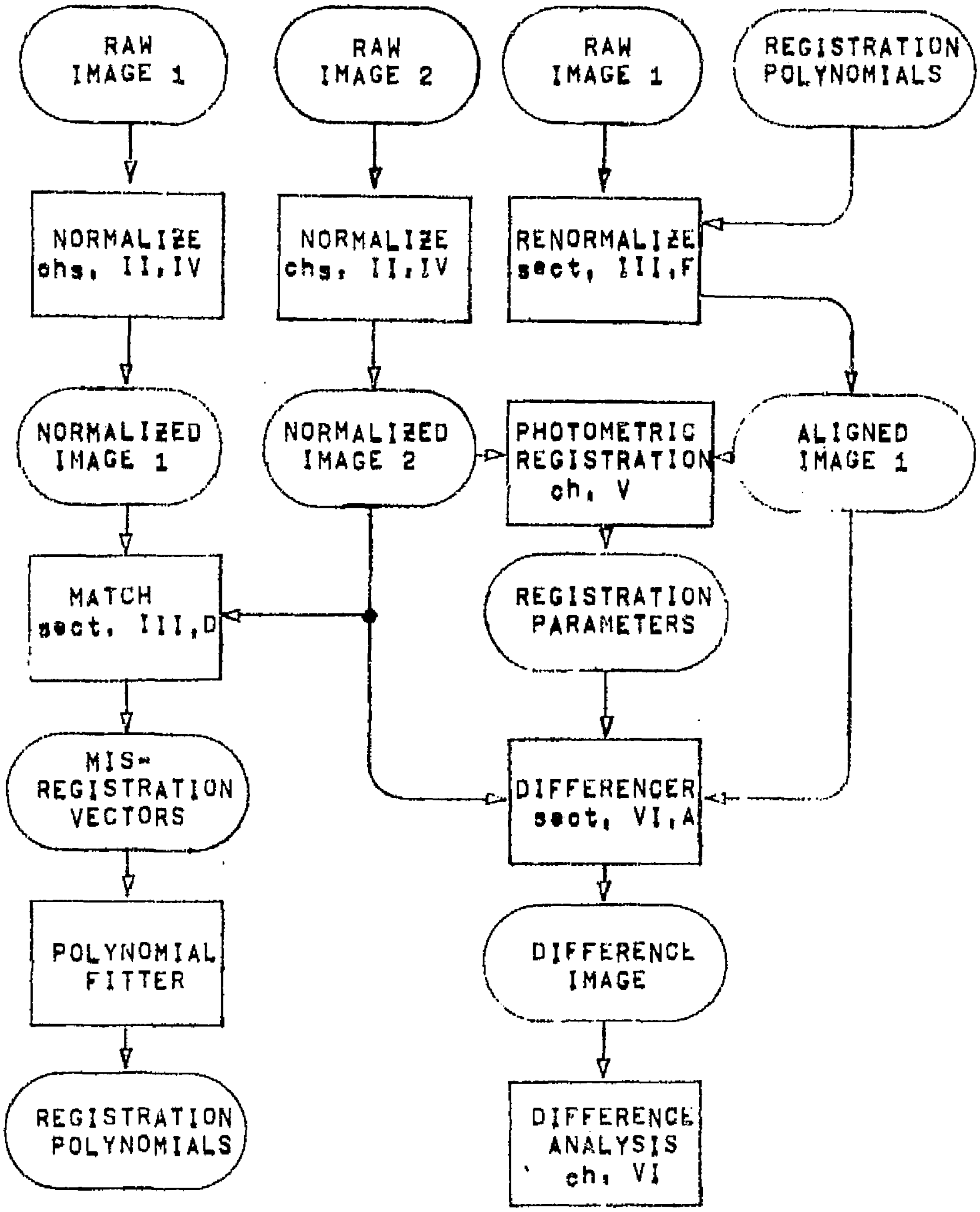


Figure 1-1 Block Diagram of System

assumed, a geometric normalization is required to relate common surface points. In particular, each image is transformed to the same orthographic projection of the spheroid. If there were no errors in our knowledge of the spacecraft position and orientation, and no geometric distortions in the optical and electronic systems of the TV camera, and if our object were a perfect spheroid, then the two normalized images should be in exact geometric correspondence. Chapter II describes geometric normalization in detail.

GEOMETRIC REGISTRATION: Unfortunately, many sources of geometric errors exist, with spacecraft orientation contributing the largest error. In order to remove these geometric alignment errors, it is necessary to align geometrically corresponding areas (features) in the two images. A technique was developed which displaces one image relative to the other, searching for a (dx, dy) translation vector which maximizes the cross correlation of the images over a specified area. Maximum cross correlation occurs when the images are properly registered in that area.

If we know the (dx, dy) translation necessary to properly register every point in the geometrically normalized images, then we can analyze the images point by point for differences.

The assumption is made that the misregistration vector as a sgfunction of position in the image, is a smooth continuous function, which can be modeled by low order polynomials in two variables, This assumption is realistic if the model of the object is accurate (i,e, surface elevations are small or view angles are similar) and if optical and electronic distortions are smooth and continuous,

In practice, when we perform a least squares fit of low order polynomials in two variables to a set of optimal translation vectors on normalized Mariner 6 and 7 images, we get very small residual errors, In particular, when fitting 1st order polynomials (which have 3 degrees of freedom) to from 10 to 100 data points, we usually get residual alignment errors of less than one picture unit (pixel) standard deviation, This, at least empirically, indicates that the predominant sources of error cause a smooth and continuous misalignment between the two images, For Mariner 6 and 7, the known geometric errors are primarily in spacecraft (camera) orientation, which would cause primarily a translational, and to a lesser degree, rotational misalignment between two images, and therefore 1st order polynomials in two variables are a good approximation,

Having an accurate model for the misalignment between the two images, we can recalculate the orthographic projection for one of the images, taking into account the misalignment model. The resulting images will be registered to the accuracy of the misregistration model. Chapter III describes geometric registration in detail.

PHOTOMETRIC NORMALIZATION: Since the images are taken under different illumination and view angles, a photometric correction is required to relate light intensity levels received at the camera to albedo on the surface. If there were no errors in our knowledge of the light scattering function at each location on the planet, and no errors in the photometric response of the vidicon, then, in theory, we should be able to precisely determine the albedo at each point in the images and perform albedo comparisons to detect variable features. Chapter IV describes photometric normalization in detail.

PHOTOMETRIC REGISTRATION: Detecting variable features necessitates a high degree of photometric accuracy, since some of the variations anticipated are relatively small (5% or less) albedo changes over rather large areas of the planet. These albedo changes may increase the contrast between two areas, or may only change the average absolute light level in both areas. If the errors and/or noise in the calibration of the camera system are larger than the albedo changes to be detected, then we must improve the calibration using information in the images.

The known repeatable sources of photometric errors are such things as vidicon shading (non-uniform response) and residual images, both of which can be handled in a systematic manner [1]. The primary sources of error which cannot be predicted are an absolute gain error, and an absolute offset error resembling scattered light. These errors, which may be caused either by errors in camera calibration, or by errors in the light scattering model for the surface, give a system output Y as a function of light input X as: Y = aX + b.

Using this model for the photometric misregistration function, we can choose areas in the two images which are assumed to be photometrically equivalent and solve for a combined gain and offset error which will minimize the difference between the areas using conventional least squares techniques. Chapter V describes photometric

---

[1] JPL has processed the Mariner '69 images to reduce these errors. Non-uniform vidicon response is corrected from extensive calibration tables for the vidicon. Residual images are the contributions of previous images to the vidicon output for a given image (i.e. the vidicon surface has a memory). Residual images are partially removed using calibration tables and previous images.

registration in detail.

DIFFERENCE ANALYSIS: Once two images are both geometrically and photometrically aligned, the analysis of differences between the images can begin. Certain classes of known differences are expected. These include albedo differences due to errors in the photometric model of the planet, variations of the photometric function from place to place on the planet, and errors due to the effects of slopes (such as crater rims) on the photometric function.

Given images which were taken with approximately the same illumination, view and phase angles, the above sources of error should be minimized. The Mariner '71 Mission B images near periapsis are intended to satisfy the above requirements, and hence these sources of error (except for regional variation of the photometric function) can be largely ignored.

The remaining albedo differences can be attributed to temporal variations of the photometric function due to such phenomena as clouds, dust storms, etc.

Analysis of these albedo variations requires that the pixel by pixel albedo differences be reduced to area differences and other graphical representations. One useful form of area difference display is a drawing of albedo difference contour lines. If the albedo difference has a well defined outline than a drawing of this outline is useful. Another useful display of differences is a graph of the size of an albedo difference area versus time (for a fixed difference level) or a graph of the magnitude of the albedo difference versus time (for a fixed area).

Chapter VI describes these difference analysis techniques.

I.C NOTATION

To facilitate the printing of this dissertation by the line printer at the Artificial Intelligence Project, some compromises in mathematical notation were made. In particular, superscripts and subscripts are not possible, and many standard mathematical symbols do not exist. Consequently, the scholarly use of an abundant collection of Greek letters other than α, β, and ε will not be found here.

ZERO: Zero is printed with a slash through it:

zero = 0

EXPONENTS: Exponents are denoted by the ALGOL 60 notation. A to the power N is written as:

A↑N.

INDICES: Indices of vectors and matrices are written in the ALGOL 60 notation as:

A[i,j] , V[k]

VECTORS: Vectors which are formed from a collection of scalar expressions are surrounded by parentheses:

(u,v) , (x,y,z)

When it is necessary to specify a column vector, the transpose notation is used:

(x,y,z)'

MATRICES: Matrices which are formed from a collection of scalars expressions are written using as many lines as there are rows in the matrix. Thus, the 3x3 identity matrix would be written:

$$I = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

The transpose of a matrix M is written:

transpose (M) = M'

The inverse of matrix M is written:

inv(M)

INNER PRODUCT: The inner product of two vectors V1 and V2 is written:

<V1,V2>

EXPRESSIONS: Arithmetic operations between scalars, vectors and matrices are denoted in conventional mathematical notation whenever possible. Consequently, a particular symbol, such as *, can mean different operations depending upon the context.

EUCLIDEAN NORM: The length of a vector V (Euclidean norm) is written:

||V||

INTEGER FUNCTION: The integer part of a real number X (i.e. the greatest integer not exceeding X) is written:

[X]

SUMMATION: The sum of an expression over an index variable is written using two lines: the first containing the word "sum" followed by the expression, and the second line containing the name and range of the index of summation under the word "sum". Using this notation, the product of two matrices A and B would be written:

```
C[i,j]  =  sum A[i,k]*B[k,j]
           1≤k≤n
```

INTEGRATION: The integral of an expression is written analogously to a sum. The integral of function f(x) over the range (-1,1) is written:

$$\underset{-1 \le x \le 1}{\text{integral}}\ f(x)$$

I.D VOCABULARY

Some of the words used in this dissertation are uncommon, some are localisms, and some are highly technical. Therefore, some such words are defined:

albedo - The ratio of the reflected light to the incident light normal to a surface. A white Lambertian surface, normal to the incident sunlight, would have an albedo of one.

apoapsis - apogee - The point in the orbit most distant from the planet.

periapsis - perigee- The point in the orbit closest to the planet.

pixel - Abbreviation for "picture element", referring to the light value at a point in a picture.

The next four chapters describe in detail the techniques for normalizing and registering images.

CHAPTER II

GEOMETRIC IMAGE NORMALIZATION BY "DEAD RECKONING"

Image normalization by "dead reckoning" refers to techniques which geometrically and photometrically register images using only calibration information such as camera position and orientation, camera sensitivity, etc, rather than information contained within the images themselves. The quality of such normalization techniques is determined by the quality of the calibration information. For Mariner 6 and 7, errors in the calibration result in a geometric misregistration of 5 to 10 pixels (25 to 50 kilometers on the surface).

PROBLEM: Given two images taken from different spacecraft (camera) positions and orientations, possibly from different spacecraft (cameras), geometrically transform the region in each image which is common in the scene (on the surface of the planet) so that the two images can be compared pixel by pixel.

DEFINITION: An *image* is a real function of two variables F(u,v) representing the light intensity received at point (u,v) on the image plane of some imaging device (eg, a camera).

All of the image scanning systems of interest, however, are discrete systems which quantize the light level at a fixed number (eg, 512) of intensity levels over a fixed rectangular array of points (eg, 945 x 702).

Such discrete imaging systems, furthermore, encode the light level integrated over an area rather than at a point. This integration can be formalized as:

$$F(x,y) = \underset{-\infty<u,v<+\infty}{\text{integral}} \left(f(u,v) * g(x-u,y-v) \right)$$

where f is the intensity function at a point and g is the "point spread" function of the imaging device (Rosenfeld [1969],p 44), which is usually adjusted by optical or electronic defocusing to minimize the errors due to aliasing introduced by discrete sampling.

In terms of scene coordinates, there exist two geometric transformations T1 and T2 which map coordinates of points in the images I1 and I2 into scene coordinates, such that when (u1,v1) and (u2,v2) represent the same point (x,y,z) in the scene, the following is true:

$$T1(u1,v1) = (x,y,z) = T2(u2,v2) \tag{2.1}$$

The next sections derive transformations T1 and T2 using geometric models for the scene and camera.

II.A GEOMETRIC MODELS

PROJECTIONS OF SCENES: A projection of a scene is some geometric mapping T(x,y,z) which maps the coordinates of points in the 3-D scene into coordinates of the 2-D projected image. Some points in the scene do not map to the projected image (e.g. points on the opposite side of an object being viewed by the eye).

The most familiar projection is the perspective projection (Fig. 2-1a), which maps the point Ps in the 3-D scene into the point Pi which is the intersection of the image plane and the straight line through the lens center O and Ps. If there is any other point Pq on the line between O and Ps, then Ps is said to be "occluded" by Pq. Occluded points are not mapped to the image plane.

Another useful projection is the orthographic projection (Fig. 2-1b) which maps the point Ps in the 3-D scene into the point Pi such that Ps is on the line normal to the image plane at Pi.

CAMERA MODEL: The camera can be geometrically modeled as a lens center and an image plane (Fig. 2-1a). At each

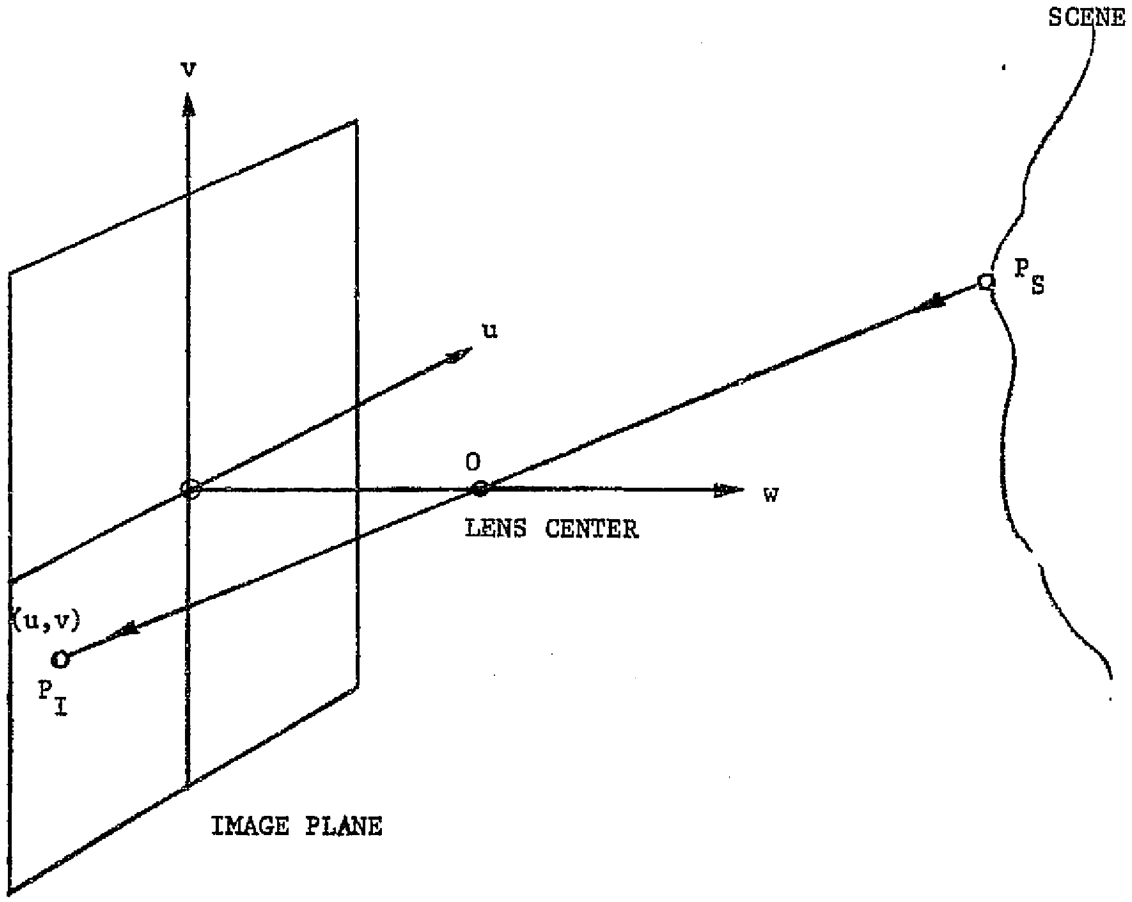


Figure 2-1a. Perspective Projection

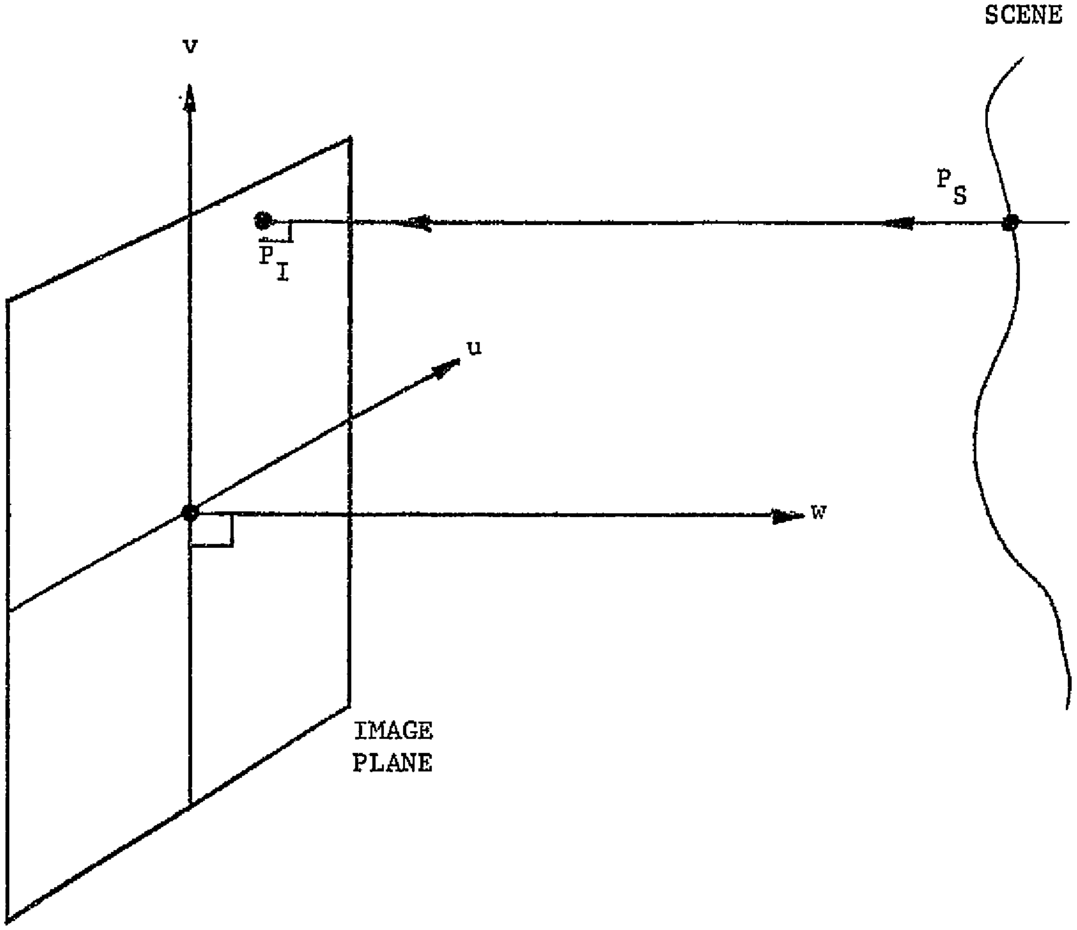


Figure 2-1b. Orthographic Projection

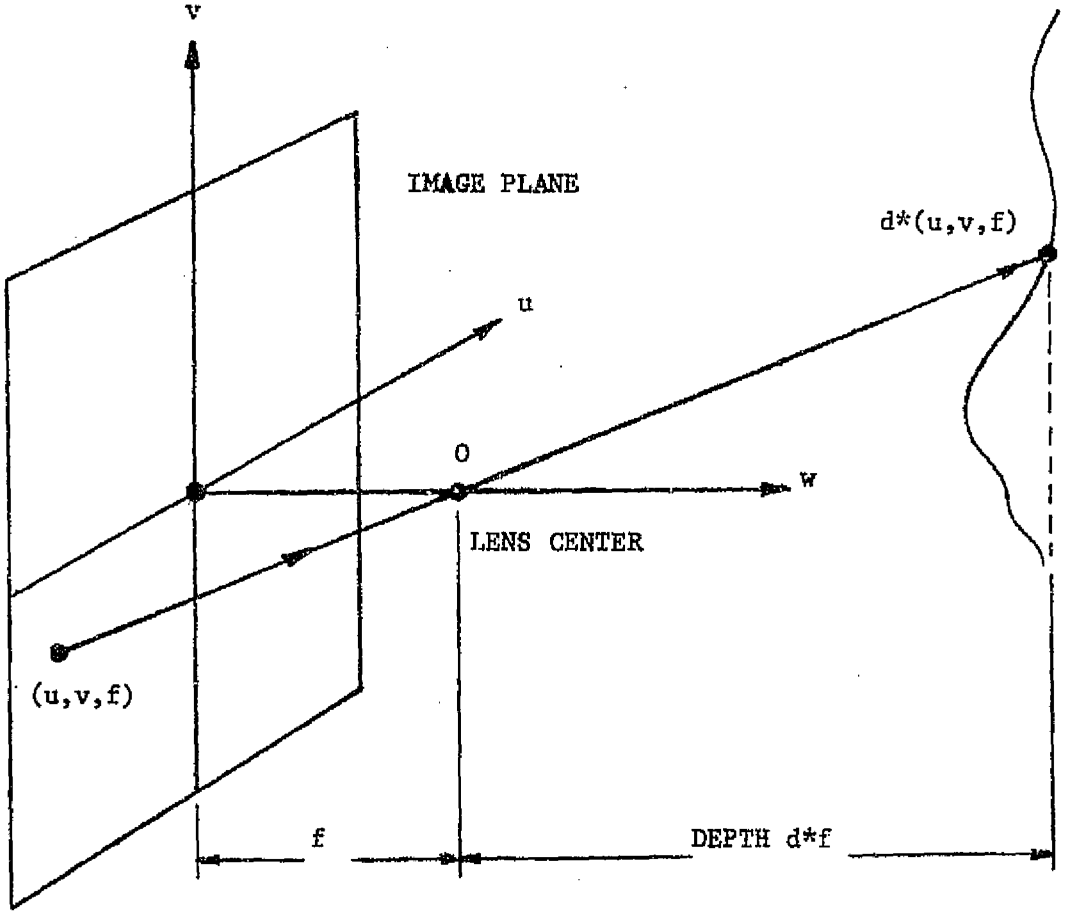


Figure 2-1c. Inverse Perspective Projection

point (u,v) on the image plane, the camera records the amount of light received from the scene along the line through the lens center and point (u,v). Thus, the camera is said to generate a perspective image of the scene. In order to normalize images taken from different camera positions, it is necessary to determine the scene coordinates of points in the perspective images.

SCENE MODELS: Theoretically, general scenes are easy to model geometrically, but in practice this can be very difficult to accomplish. A very general class of scenes may be modeled by assuming that the scene is made from objects whose surfaces are opaque and are described by a function $H(x,y,z)=0$. A scene consisting of a sphere of radius r centered at the origin can thus be modeled:

$$H(x,y,z) = x^2 + y^2 + z^2 - r^2 = 0 \qquad (2.2)$$

The planet Mars can be approximated by an oblate spheroid with equatorial radius $Req=3393.4$ km (equator in the x-y plane, with $z=0$) and polar radius $Rpol=3375.6$ km (poles along the z-axis). With this model, H becomes:

$$H(x,y,z) = x^2 + y^2 + (z*Req/Rpol)^2 - Req^2. \qquad (2.3)$$

II.B THE INVERSE PERSPECTIVE PROJECTION

Inverse projections map coordinates of points in images which are projections of scenes, back into coordinates in the scene. Geometric models for the scene and the projection are required in order to invert the projection.

The inverse perspective image projection is the projection from two dimensional perspective image coordinates to three dimensional spatial coordinates. This projection maps point (u,v) in the perspective image plane to the closest point (x,y,z) in the 3-D scene which is along the line through the lens center and (u,v). Any point along that line is given in camera relative coordinates (the lens center is the origin, with axes u,v, and w) as a function of the depth parameter d as $d*(u,v,f)$, where f is the image distance of the camera (Fig. 2-1c).

To transform camera relative coordinates into scene relative coordinates, a linear transformation L is defined:

$$L(u,v,w)) = R(u,v,w)' + Pc \qquad (2.4)$$

where:

1) R is a rotation matrix from camera space orientation to scene orientation. Appendix A derives matrix R from geometric calibration data.
2) Pc is the position of the camera relative to the origin of the scene coordinate system.

Using these definitions, the inverse perspective projection T of point (u,v) is defined in scene coordinates by finding the unknown depth parameter d such that:

$$T(u,v) = L(d*(u,v,f)') \qquad (2,5a)$$

$$= R(d*(u,v,f)') + Po \qquad (2,5b)$$

$$= d*R(u,v,f)' + Po \qquad (2,5c)$$

such that $d>0$, $d=min$, and (2,6)

$$0 = H(\ T(u,v)\ ) \qquad (2,7a)$$

$$= H(\ d*R(u,v,f)' + Pc\ ) \qquad (2,7b)$$

where the primed vectors denote column vectors. Equation (2,5) requires that T(u,v) be a perspective view of the scene from a camera at position Po with orientation defined by R and with image distance f. Equation (2,7a) requires that T(u,v) be a point in the scene, and (2,6) requires that it be the point closest to the camera.

Using the above definitions, the perspective projection T' is the inverse of T and can be defined as follows:
Let the transformation T" be defined:

$$T''(x,y,z) = inv(R)*((x,y,z)'-Po) \qquad (2,8)$$

where inv(R) is the inverse of matrix R. Note the following:

$$T''(T(u,v)) = T''(R(d*(u,v,f)') + Pc) \qquad (2,9a)$$

$$= inv(R)*(R(d*(u,v,f)') - Po + Pc) \qquad (2,9b)$$

$$= inv(R)*(R(d*(u,v,f)')) \qquad (2,9c)$$

$$= d*(u,v,f)' \qquad (2,9d)$$

If

$$(u,v,w)' = T''(x,y,z) \qquad (2.10a)$$

then, the perspective projection T' is defined as:

$$T'(x,y,z) = f * \frac{(u,v)}{w} \qquad (2.10b)$$

Note that w=d*f.

II.C SPECIAL CASES OF THE INVERSE PERSPECTIVE PROJECTION

Equations (2.5-2.7) are in general very difficult to solve because of the complexity of H. This section derives their solution for the special cases of spherical and spheroidal scene models.

SOLUTION FOR A SPHERE: The inverse perspective transformation problem can be easily solved for known camera position and orientation with respect to a sphere. In particular, for a sphere of radius r centered at the scene origin, we have for a given point P on the sphere:

$$||P|| = r, \text{ i.e. } Px{\uparrow}2+Py{\uparrow}2+Pz{\uparrow}2 = r{\uparrow}2. \qquad (2.11)$$

So, given the perspective projection of P in image coordinates (u,v), we know that:

$$
\begin{aligned}
0 &= ||T(u,v)||^2 - r^2 && (2,12a)\\
&= ||R(d*(u,v,f)')+Pc||^2 - r^2 && (2,12b)\\
&= ||R||^2*||d*(u,v,f)'||^2 && (2,12c)\\
&\quad + 2*\langle R(d*(u,v,f)'),Pc\rangle && \\
&\quad + ||Pc||^2 - r^2 && \\
&= d^2 * ||(u,v,f)||^2 && (2,12d)\\
&\quad + 2*d * \langle R(u,v,f)',Pc\rangle && \\
&\quad + ||Pc||^2 - r^2 &&
\end{aligned}
$$

d is found by solving the quadratic equation:

$$d = (-b - sqrt(b^2-4*a*c))/(2*a) \qquad (2,13)$$

where

$$
\begin{aligned}
a &= ||(u,v,f)||^2 = u^2+v^2+f^2 && (2,14a)\\
b &= 2*\langle R(u,v,f)',Pc\rangle && (2,14b)\\
c &= ||Pc||^2-r^2 && (2,14c)
\end{aligned}
$$

The smaller solution (- sqrt) of the quadratic equation is the only meaningful one since the point on the sphere corresponding to the larger solution is at a greater distance from the camera and is therefore occluded by the point corresponding to the smaller solution. Complex solutions correspond to camera rays which do not intersect the sphere.

SOLUTION FOR AN SPHEROID: The solution for a sphere easily generalizes to a spheroid by introducing an eccentricity constant e along the z-axis (North to South pole axis) such that:

$$x^2 + y^2 + (e*z)^2 = r^2 \tag{2,15}$$

If the eccentricity matrix E is defined as:

$$E = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e \end{matrix} \tag{2,16}$$

then d is found by solving the following quadratic equation:

$$0 = ||E*T(u,v)||^2 - r^2 \tag{2,17a}$$

$$= ||E*(R(d*(u,v,f))+Pc||^2 - r^2 \tag{2,17b}$$

$$\begin{aligned} &= d^2 * ||E*R(u,v,f)||^2 \\ &\quad +2*d*\langle E*R(u,v,f),E(Po)\rangle \\ &\quad +||E(Po)||^2 - r^2 \end{aligned} \tag{2,17c}$$

II,D EXAMPLES OF GEOMETRIC NORMALIZATION

This section contains examples of the geometric normalization techniques applied to far-encounter Mariner 7 pictures of Mars, These pictures were taken when the spacecraft was sufficiently far from Mars to see the entire disk of the planet,

Figure 2-3 shows the disk of the planet (the circle) with the north pole at the top as seen from an orthographic projection in the direction 0 latitude, 0 longitude (i,e, (0 lat., 0 long.) is the center of the circle). The dotted lines surrounding the circle indicate the outlines of two pictures which are designated 7F75 and 7F78. The vectors in the circle indicate the direction to the spacecraft and the sun from the central points of the two pictures. The square indicates the area on the surface for which an orthographic projection of each picture will be generated.

Figure 2-4a shows the area of the picture 7F75 (the dotted area) which is specified by the square in figure 2-3. The rectangle is the 945x702 pixel outline of picture 7F75. Similarly, figure 2-4b shows the area of 7F78 which is to be orthographically projected.

Figures 2-5a and 2-5b show the pictures 7F75 and 7F78, as seen from the spacecraft. Note that the crater (Nix Olympica), which is in the upper middle part of 7F75, is in the upper right part of 7F78. This is because the planet rotated by 34 degrees between the two pictures. A scale factor difference is also obvious.

Figures 2-6a and 2-6b show the orthographic projection of each image as specified by figures 2-3, 2-4a, and 2-4b. In these normalized pictures, the crater looks about the same, and various features match. However, there is a definite geometric error in the registration of these images.

CHAPTER III

ACCURATE GEOMETRIC REGISTRATION OF NORMALIZED IMAGES

III.A FORMAL DEFINITION OF GEOMETRIC MISREGISTRATION

Let us suppose that we have two geometrically normalized images (F1,T1) and (F2,T2) where the Ti are transformations from image coordinates to actual 3-space coordinatees, rather than those predicted by Chapter II, which contain errors due to the camera and scene models. The Fi are the intensity functions in the images. The images are said to be "geometrically misregistered" at point (u,v) if:

$$T1(u,v) \neq T2(u,v) \qquad (3.1)$$

We will define two functions du(u,v) and dv(u,v) which represent the misalignment of two images in the u and v directions respectively as a function of the position in the images as:

$$(du(u,v),dv(u,v)) = \{(du,dv): T1(u,v)=T2(u+du,v+dv)\} \qquad (3.2)$$

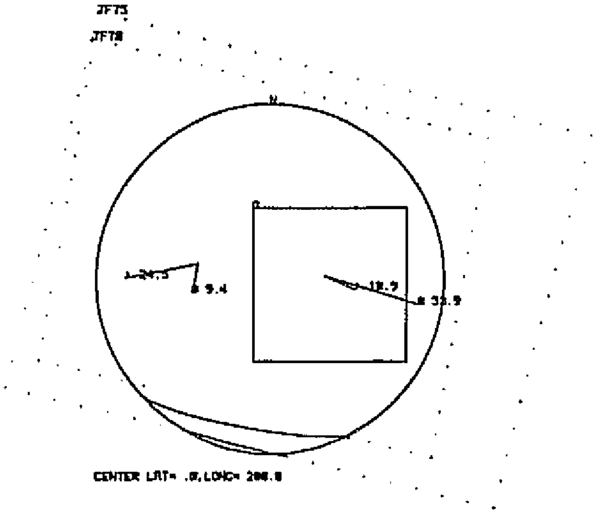


Figure 2-3
Orthographic Projection
of Mars

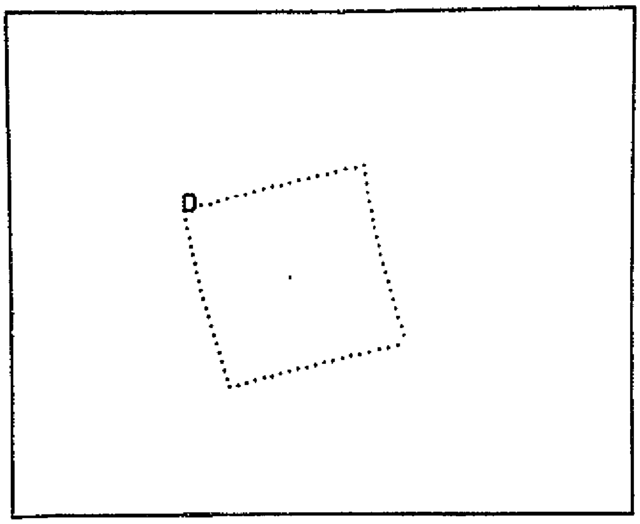


Figure 2-4a
7F75 Window

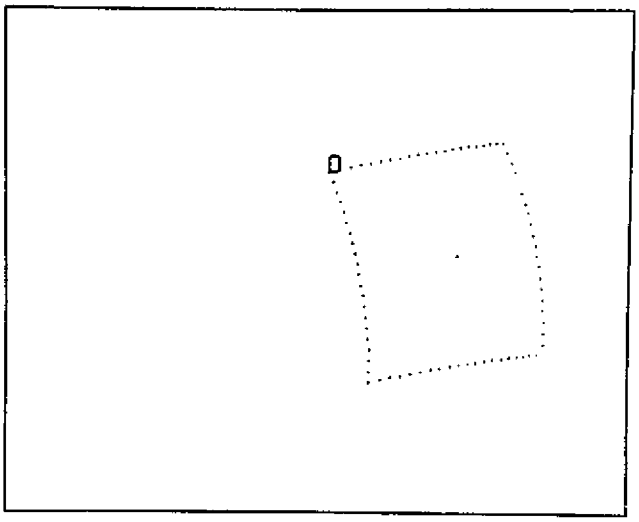


Figure 2-4b
7F78 Window

Figure 2-5a
7F75 Original

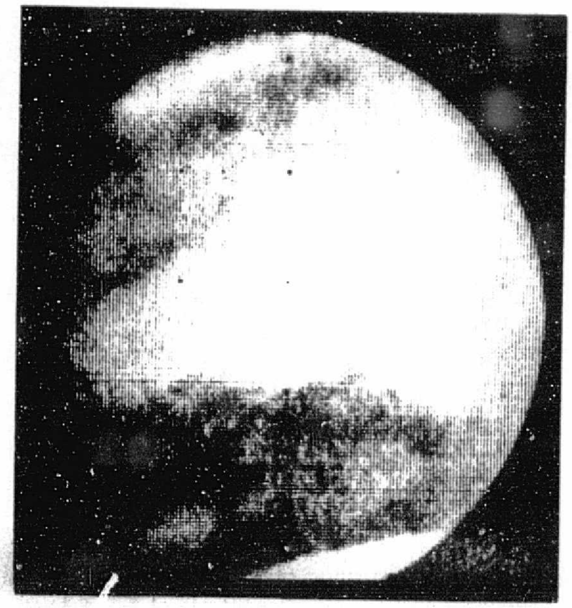


Figure 2-5b
7F78 Original

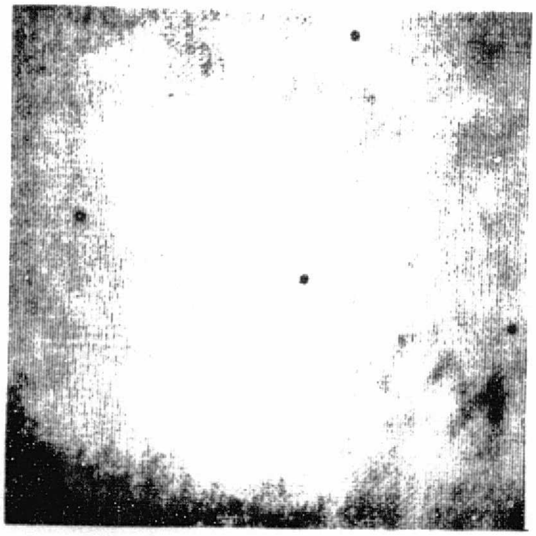


Figure 2-6a
7F75 Normalized

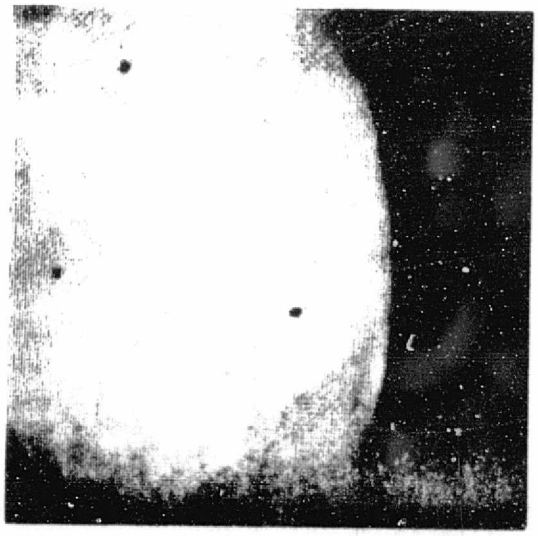


Figure 2-6b
7F78 Normalized

These misalignment functions tell us that, given a point (u,v) in image 1, this same point in the scene is at (u+du,v+dv) in image 2.

If we have the misalignment functions defined at every point in the discrete images, then we can define a new image (F3,T3) which is exactly registered with image 1 as follows:

$$F3(u,v) = F2(u+du(u,v),v+dv(u,v)) \qquad (3.3)$$

(assuming an adequate model for F2 between data points)

$$T3(u,v) = T2(u+du(u,v),v+dv(u,v)) \qquad (3.4)$$
$$= T1(u,v)$$

The next sections empirically derive the misalignment functions for a pair of images.

III.B MODEL FOR MISREGISTRATION

Because of geometric errors primarily in the Mariner '69 camera model (camera position and orientation), images which have been geometrically normalized by "dead reckoning" are not exactly registered. It can be shown that small errors in the camera model cause primarily a translational misalignment between geometrically normalized images. There are also small rotational and scale factor errors. Higher order errors might exist due to residual distortions in the optics and errors in the 3-D model for the scene.

One might observe that given two images which are approximately aligned, people very easily match most features (such as craters) which correspond. However, the mechanisms used in image matching by people are not necessarily those best for automatic image alignment by computer.

The predominantly translational nature of the misalignment of images suggests that, when local areas in images are appropriately translated, the images will match. If the two images were taken under sufficiently similar photometric conditions (such as illumination, color filters, etc.) then the quality of the match can be measured by statistical cross correlation. The registration technique which has been developed is to maximize the local cross correlation of the images as a function of the translation of one image with respect to the other.

An alternative procedure would be to search one image for a distinct feature such as a crater or some other topographic feature, and then search the second image in a predicted area for the same feature. Such feature recognition, however, requires that each image which is to be aligned contain the necessary number (and quality) of the types of features which can be recognized. Experience with a crater finder [1] shows that in order to reliably locate craters, the contrast of the crater rims must be good, and that the signal-to-noise ratio must be high. Even when these conditions are met, the cross correlation alignment procedure seems to be superior, because of its relative insensitivity to different types of terrain and to noise. There also appears to be no inherent speed advantage for a feature detector. One outstanding virtue of a crater finder would be the ability to locate craters and develop a 3-D depth model so that the photometric errors due to crater slopes could be removed. Such slopes, however, cause significant errors only when the illumination and viewing angles are quite different between the pictures being registered.

Cross correlation maximization was selected for geometric alignment, because of its simplicity and applicability to varying terrains.

---

[1] A crater finder was developed which looks for high contrast edges which contain large circular arcs. Many craters are recognized by the fraction of their rim which is a high contrast crescent. The outside curve of the crescent is a circular arc with approximately the same radius and center as the crater.

III.C CROSS CORRELATION AS A MEASURE OF THE QUALITY OF REGISTRATION

Correlation is a statistical measure of agreement between two functions. For two random variables X and Y defined on a discrete set of n points, their normalized cross correlation (cor) is defined in terms of the means (E) and standard deviations (sd) as:

$$\mathrm{cor}(X,Y) = \frac{E(X*Y) - E(X)*E(Y)}{sd(X) * sd(Y)} \qquad (3.5)$$

where

$$E(X) = \underset{1\le i\le n}{\mathrm{sum}} (X[i]) \,/\, n \qquad (3.6)$$

and

$$sd(X) = \mathrm{sqrt}\,(\, E(X\uparrow 2) - E(X)\uparrow 2 \,) \qquad (3.7)$$

To apply cross correlation to an area of two images with intensity functions F1 and F2, the one dimensional sum in equation (3.6) becomes two dimensional, and we make the following definitions for the variables X and Y. Given (u,v) as the center of a (2n+1)x(2n+1) area to be correlated and a (du,dv) translation vector, define X and Y as:

$$X[i,j] = F1(i+u-n, j+v-n) \quad 0\le i,j\le 2n \qquad (3.8a)$$

$$Y[i,j] = F2(i+u-n+du, j+v-n+dv) \qquad (3.8b)$$

The normalized correlation function has the property that it does not vary with the size of the correlation area, or with the means and standard deviations of X and Y. Subjectively this guarantees that the correlation function is not affected by the average intensity level and level of contrast in the window. The quality of the correlation measure is, however, strongly related to signal-to-noise ratio. Increased noise will increase the standard deviations in equation (3.5) without increasing the product term E(X*Y).

This effect can be derived by assuming that the two windows are identical except for uncorrelated noise. Then they can be represented:

$$X = Z + N1 \; , \; Y = Z + N2 \qquad (3.9a)$$

Where Z is the "true" value of window. If N1 and N2 are uncorrelated noise, and have the same standard deviations, then, the cross correlation of X and Y is:

$$cor(X,Y) = \frac{E((Z+N1)*(Z+N2)) - E(Z+N1)*E(Z+N2)}{sd(Z+N1) * sd(Z+N2)} \qquad (3.9b)$$

$$= \frac{E(Z*Z) - E(Z)*E(Z)}{sd(Z+N1) * sd(Z+N2)} \qquad (3.9c)$$

$$= var(Z)/var(Z+N) \qquad (3.9d)$$

$$= 1 - var(N)/var(Z+N) \qquad (3.9e)$$

where var is the variance. High uncorrelated noise levels consequently will decrease the value of the correlation function. This suggests that the correlation function should be evaluated only in areas where the standard deviations of the windows are significantly greater than the standard deviation of the noise, and that the search for maximum correlation should find a correlation value approximately equal to the value of equation (3.9e).

For a desired level of correlation quality, the window size can be adjusted according to the local standard deviations of the two images. Increasing the window size (i.e. increasing n) excessively will cause poor results due to non-translational (rotation and scale factor) distortions between the two images (Appendix B).

III.D LOCAL CROSS CORRELATION MAXIMIZATION TO DETERMINE LOCAL MISREGISTRATION

Analyses of the errors in the Mariner 6 and 7 geometric model show that images which are normalized by "dead reckoning" should contain primarily translational errors in alignment with smaller rotational and scale factor errors. The results in tables and figures 3-1 to 3-7 confirm this prediction, and show that constant translational errors dominate other errors by a factor of more than 10:1. The presence of non-translational errors puts an upper limit on the size of the (2n+1)x(2n+1) correlation window.

Therefore, to align two images using correlation, it is sufficient to find the (du,dv) translation of one image with respect to the other which maximizes the value of the correlation function around each point (u,v) in the images. This (du,dv) translation is called the "misalignment vector" for the area of correlation centered at (u,v). It has not been necessary to maximize the correlation with respect to other parameters such as rotation and scale factor changes, because the magnitudes of these errors are small relative to translational errors, and it is possible to find the local correlation maximum using only translations. These higher order errors are found when many different misalignment vectors are modeled over the entire image as described in III.E.

SEARCH FOR MAXIMUM - STRATEGIES: The subject of this section is to describe two different strategies to search for the translation vector which produces the maximum correlation between two images. The predominant constant translational misalignment suggests that the search for maximum correlation should be confined to the neighborhood of that predominant error vector.

The primary reason for considering these search strategies is to reduce the number of evaluations of the correlation function and thus improve the performance of the image registration system.

The alignment strategies are:

1) Global strategy: determine the predominant global misalignment of the two images.

2) Local strategy: determine the local misalignments of the two images by searching in limited neighborhoods predicted by the global misalignment model.

GLOBAL SEARCH STRATEGY: The purpose of the global search strategy is to search for the translation vector which produces the maximum correlation between two images without knowledge to limit the search to a small neighborhood. Because the correlation function is evaluated only at integer values of the (du,dv) displacement vector, the global search strategy could be implemented by exhaustive evaluation of the correlation function at all integer translations. However, some analysis of the correlation function, and the areas of the images being correlated, shows that one can limit the number of evaluations of the correlation function considerably.

If one knows the sharpness of the correlation peak and its amplitude relative to other relative peaks, then one can determine a search grid spacing which will guarantee finding the absolute peak. If for instance, the absolute peak is k units wide at the level of any other relative peak, then the global search can be carried out on a kxk grid, reducing the number of evaluations of the correlation function by a factor of k*2 (see figure 3-0a). Various parameters (such as window size, spatial averaging, etc.) can be adjusted to broaden the peak sufficiently to allow flexible choice of the grid spacing (Figs. 3-0c and 3-0d show the effects of spatial filtering on correlation).

If one wants to search for the correlation maximum by evaluating the correlation function only at points on a kxk pixel grid, then the sampling theorem says that one should limit the frequency spectrum of the correlation function to spatial wavelengths longer than 2k pixels. This is successful only if the power spectra of the two images have sufficient power at spatial wavelengths longer than 2k pixels. The disadvantage of low pass spatial filtering is that high spatial frequency information, which usually best characterizes topographic features, is lost. Therefore, a search for the local correlation maximum of unfiltered images is desirable in a limited neighborhood around the low frequency maximum.

LOCAL SEARCH STRATEGY: When the predominant translational misalignment between two images is known, the search for the local misalignment can be limited to a small neighborhood, thus limiting the number of evaluations of the correlation function. The local search strategy begins searching for maximum correlation starting with a translation vector which is predicted by a misregistration model based on the predominant global misregistration vector and previous local registration vectors. The strategy searches a limited distance in 8 different directions looking for a maximum. A hill-climb search then follows, until a (du,dv) point is found such that all 8 neighbors give a lower correlation value.

MODELING THE CORRELATION FUNCTION AT NON-INTEGRAL TRANSLATIONS: Given the values of the correlation function at a discrete set of integer translations, it is useful to model the function at non-integral translations using interpolation or, equivalently, fitting linear combinations of functions (such as polynomials) to the data.

In particular, one can fit a polynomial in 2 variables to the correlation function at known translations, and solve for the maximum of the polynomial. The method chosen was to use a least squares fit of an Nth order polynomial in 2 variables, and solve for its maximum using a 2-dimensional generalization of Newton's method (Kowalik [1968] pp. 65-71) in the vicinity of the maximum empirical value of the correlation function.

Although it is difficult to prove that the correlation function can best be approximated in this fashion, there is

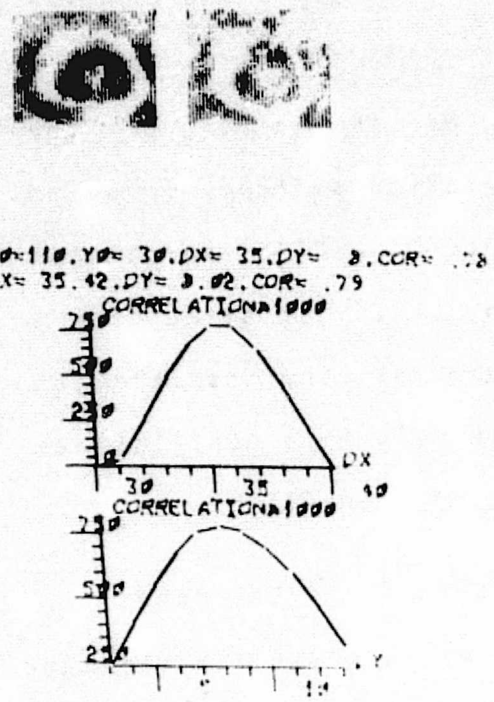


Figure 3-1a
Nix olympica

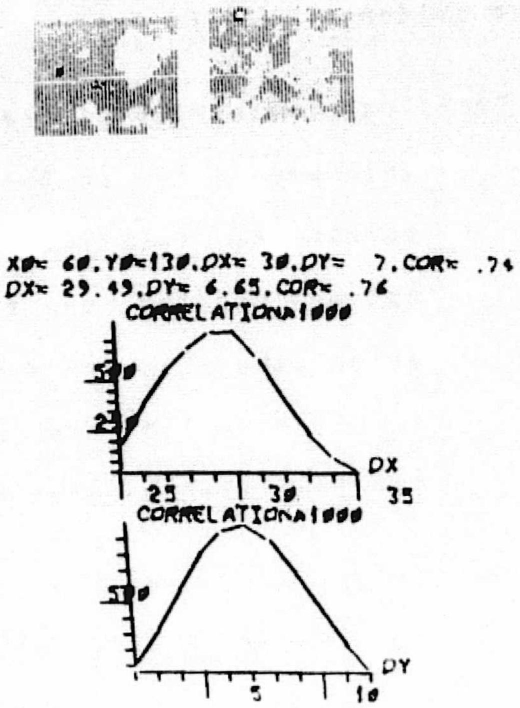


Figure 3-1b
Random Area

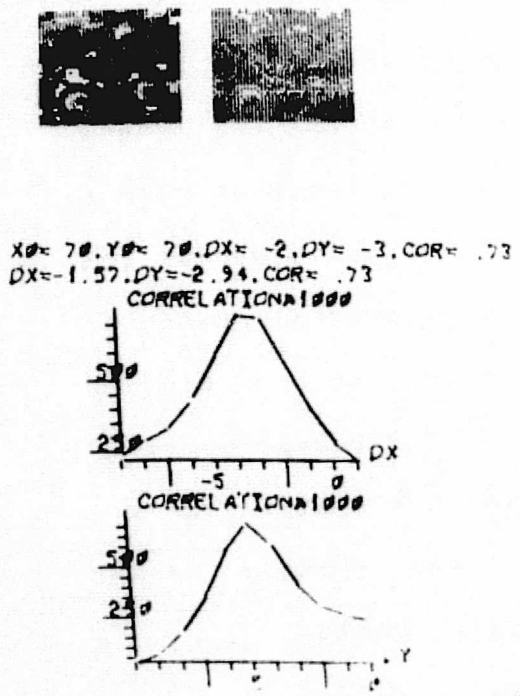


Figure 3-1c
Craters

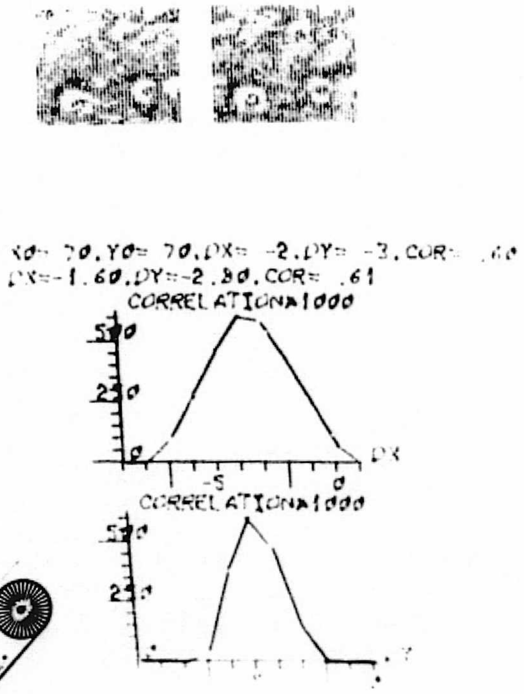


Figure 3-1d
Hipass Craters

some evidence that it can:

1) Empirical evidence: Increasing N, the order of the polynomials fit to the correlation surface, produces quintuples (u[i],v[i],du[i],dv[i],c[i]) which are better modeled by section III.E in terms of RMS error (the c[i] are the values of the correlation function which are used as weighting coefficients for the least squares fit in section III.E).

2) Theoretical evidence: The correlation function tends to be smooth and radially symmetric about its maximum value (see figs. 3-0a to 3-0d). From the Fourier analysis point of view, the Fourier transform of the correlation surface is the product of the Fourier transforms of the two images (complex conjugate of one of them). If the two images correlate well, then the correlation surface is approximately equivalent to the autocorrelation surfaces of either of the images, which tend to be radially symmetric. If both images have most of their spectral power at low spatial frequencies, then the correlation surface will also have most of its spectral power at low spatial frequencies, and the correlation surface will be correspondingly smooth and broadly peaked around the nominal misregistration vector.

EXPLANATION OF FIGURES 3-0a,b,c,d: These figures show windows from pairs of pictures, and cross sections of the correlation surface in the vicinity of the correlation maximum. These cross sections are graphs of the value of the correlation function with one of the variables dx and dy changing and the other constant.

Figure 3-0a shows the correlation function for a window containing the crater Nix Olympica. The crude 30x30 pixel windows are data from the pictures 7F75 and 7F78. The first graph is the correlation function with dx changing and dy = 8. The second graph shows the correlation function with dy changing and dx = 35. The correlation surface is modeled by fitting a second order polynomial in dx and dy to points around the correlation peak (dx = 35, dy = 8). The maximum point on this surface is found to be (dx = 8.08, dy = 35.38) attaining a correlation value of .75, using a two-dimensional Newton's method.

Figure 3-0b shows an area which does not contain any particular features. The correlation function in this area is also smooth and symmetric.

Figures 3-0c and 3-0d show the behavior of the correlation function when low frequency information is removed from a picture (these areas are from different Mariner 6 and 7 pictures). Figure 3-0c is the correlation function for an area containing several craters. Figure 3-0d shows the correlation function for the same area after low spatial frequencies have been removed (subtracting the local average intensity computed over a 10x10 area from each point). Removing low spatial frequencies clearly reduces the width of the correlation peak.

III.E MODELING THE MISREGISTRATION AS A FUNCTION OF POSITION IN THE IMAGE

There are several alternative methods available for modeling the misalignment functions. One method is Mth order interpolation (usually M=2) in 2 variables in local regions of the images. Another method (which was chosen) is to fit polynomials in 2 variables to the entire set of misalignment vectors, and then minimize the mean squared error between the polynomials and the empirical data points (the method of least squares). Interpolation is good if the errors in the empirical data points are small (that is, the correlation search always works well), and if there are enough data points to adequately cover the entire image. Polynomial fits to the entire images were chosen because they have good smoothing properties on noisy data.

The misregistration functions du(u,v) and dv(u,v) are empirically derived from n quintuples of the form (u[i],v[i],du[i],dv[i],c[i]), where (du[i],dv[i]) is the translation vector which maximizes the correlation c[i] between the 2 images over a window centered at (u[i],v[i]), For simplicity, these functions were chosen to be Nth order polynomials in 2 variables of the form:

$$du(u,v) = \sum_{0 \le i+j \le N} A[i,j]*u^{i}*v^{j} \qquad (3,10)$$

The traditional least squares approach is to choose du(u,v) such that:

$$\sum_{k} (du(u[k],v[k])-du[k])^{2} = \min \qquad (3,11)$$

$$= \sum_{k} \left(\sum_{i,j} A[i,j]*u[k]^{i}*v[k]^{j} - du[k]\right)^{2} \qquad (3,12)$$

This least squares problem degenerates to the solution of an (N+1)*(N+2)/2 order system of linear equations, The polynomial for dv is constructed similarly,

The RMS error of the polynomial fit to the empirical data points is a good measure of whether the order N of the polynomials is high enough, and if in fact polynomials are the appropriate functions to fit to the data. To account for the number of degrees of freedom in the polynomial fit, the following measure is used:

$$\mathrm{sqrt}\left(\sum_{k} \frac{(du(u[k],v[k])-du[k])^{2}}{Nk - (N+1)*(N+2)/2} \right) \qquad (3,13)$$

where Nk is the number of data points,

Since the empirical data points represent a sampling of a surface, we would like the polynomials to model that surface everywhere as well as sample points. A useful test is to fit polynomials to one set of sample points, and then measure the RMS error between these polynomials and another sampling of the surface.

The above least squares bi-variate polynomial approximations have been used successfully to model the misregistration vectors between images. Tables 3-1 to 3-5 and figures 3-1 to 3-7 show some actual misregistration vectors and the results of fitting polynomials of various orders.

The tables contain the misalignment vectors (DU,DV) which maximize the correlation function over a 21x21 pixel window centered at points (U,V). COR is the maximum value of the correlation function. The residual error vector (DUerr,DVerr) is the difference between the empirical (DU,DV) vector and the vector predicted by the polynomial model. ||error|| is the Euclidean norm of the residual error vector.

The figures graphically show the data in the corresponding tables. The square on the left represents the entire area of the 200x200 pixel pictures. The vectors in the square start at points (U,V) and have length and direction (10*DUerr,10*DVerr), The single digit at (U,V) represents the value of the correlation maximum (i,e, 20% to 29% is represented by 2). The circle represents residual errors of one standard deviation. The endpoints of the residual vectors (DUerr,DVerr) are plotted relative to the center of the circle.

Tables and figures 3-1 to 3-3 show residual errors before bad misregistration vectors have been eliminated by "blunder removal" (explained later), Table and figure 3-1 show the residual errors from zero order polynomial fits to 14 data points, The predominant translational error is DU=34 and DV=9 pixels. The weighted RMS error is 5,25 pixels, where data points are weighted according to their correlation value COR, The corrected error of equation (3,13) is 5,65 pixels, To compare the quality of the fit, these corrected errors will be used, Table and figure 3-2 show the residual errors from first order polynomial fits to the same data points, The corrected error, which is 4,47 pixels, has not been reduced much, Table and figure 3-3 show the residual errors from a second order fit, The corrected error is 4,25 pixels, which is not a significant reduction,

Tables and figures 3-4 and 3-5 show residual errors after 4 "blunders" have been removed, Blunders are misregistration vectors which do not fit any systematic model. These vectors result from cross correlating areas where these is some form of systematic difference between the pictures such as clouds, or crater "shadows" which have moved due to changes in illumination. These blunders are removed by finding residuals (DUerr, DVerr) which are large, and do not cluster with the other residuals.

Table and figure 3-4 show the residual errors after a first order fit. The corrected error is now 1.03 pixels. The second order fit of table and figure 3-5 has a higher corrected error of 1.18 pixels.

Figures 3-6 and 3-7 show the residual errors from first and second order polynomial fits to a much larger (about 60 points) set of misalignment vectors, after blunder removal. The first order fit has a corrected error of 1.26 pixels, which is fairly consistent with figure 3-4. The second order fit has a corrected error of 1.02 pixels, which is somewhat better than the first order fit.

TABLE 3-1

POLYNOMIALS FIT TO ORDER 0 BEFORE BLUNDER REMOVAL

```
DU(U,V)= 34,34
DV(U,V)= 8,991

 U     V        DU        DV        COR     DU err   DV err  ||error||
-60   -60     30,16     10,73       ,68    -4,17     1,73     4,52
-60   -20     32,01     20,73       ,87    -2,33    11,74    11,97
-60    20     32,00     10,00       ,00    -2,34     1,01     2,55
-60    60     32,88      9,81       ,74    -1,46      ,82     1,67
-20   -60     33,92      4,89       ,25     -,41    -4,10     4,12
-20   -20     31,77      8,70       ,88    -2,56     -,29     2,58
-20    20     31,07      7,50       ,82    -3,27    -1,49     3,60
-20    60     30,87      6,66       ,87    -3,47    -2,33     4,18
 20   -60     37,17      7,68       ,77     2,83    -1,31     3,12
 20   -20     36,13      8,69       ,79     1,79     -,30     1,81
 20    20     35,65      6,68       ,71     1,31    -2,31     2,66
 20    60     35,18      6,57       ,85      ,85    -2,42     2,56
 60   -60     38,00      7,00       ,00     3,66    -1,99     4,17
 60   -20     41,11      6,00       ,53     6,77    -2,99     7,40
 60    20     41,63      9,23       ,51     7,29      ,24     7,29
 60    60     39,10      5,46       ,30     4,76    -3,54     5,93
-----------

TOTAL WEIGHTED RMS ERROR=  5,25
```

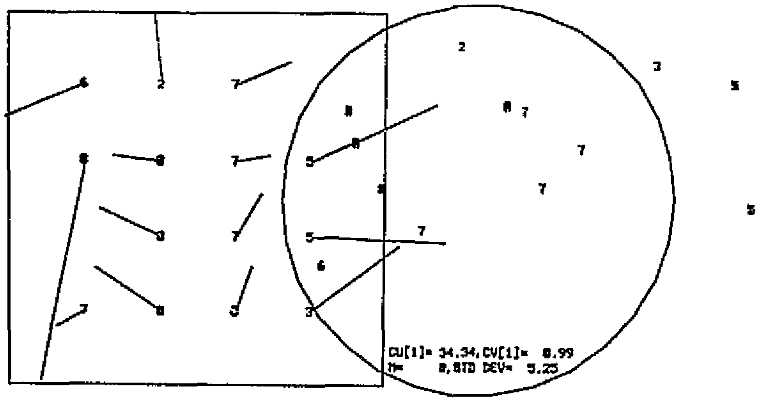


FIGURE 3-1 RESIDUAL ERRORS FROM TABLE 3-1,

TABLE 3-2

POLYNOMIALS FIT TO ORDER 1 BEFORE BLUNDER REMOVAL

```
DU(U,V)= 34,78 + ,7638@-1*X -,8829@-2*Y
DV(U,V)= 8,783 - ,5617@-1*X -,1997@-1*Y

 U     V        DU        DV        COR    DU err    DV err  ||error||
-60   -60     30,16     10,73      ,68     -,56     -2,63      2,69
-60   -20     32,01     20,73      ,87     1,64      8,18      8,34
-60    20     32,00     10,00      ,00     1,98     -1,75      2,64
-60    60     32,88      9,81      ,74     3,21     -1,15      3,41
-20   -60     33,92      4,89      ,25      ,14     -6,21      6,21
-20   -20     31,77      8,70      ,88    -1,66     -1,61      2,31
-20    20     31,07      7,50      ,82    -2,01     -2,01      2,84
-20    60     30,87      6,66      ,87    -1,86     -2,05      2,76
 20   -60     37,17      7,68      ,77      ,33     -1,18      1,22
 20   -20     36,13      8,69      ,79     -,36       ,63       ,73
 20    20     35,65      6,68      ,71     -,48      -,58       ,76
 20    60     35,18      6,57      ,85     -,60       ,11       ,61
 60   -60     38,00      7,00      ,00    -1,90       ,39      1,94
 60   -20     41,11      6,00      ,53     1,56       ,19      1,58
 60    20     41,63      9,23      ,51     2,44      4,21      4,87
 60    60     39,10      5,46      ,30      ,26      1,24      1,27
-- --------

TOTAL WEIGHTED RMS ERROR=  3,51
```

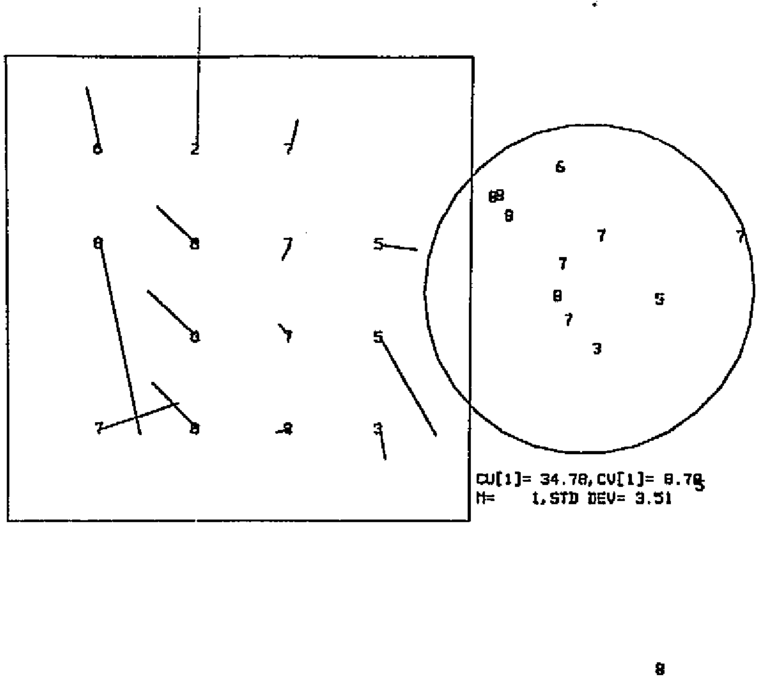


FIGURE 3-2 RESIDUAL ERRORS FROM TABLE 3-2

TABLE 3-3

POLYNOMIALS FIT TO ORDER 2 BEFORE BLUNDER REMOVAL

```
DU(U,V)= 33,35+,08510*X+,8858@-3*X2
          -,01250*Y-,3867@-3*X*Y+,6157@-4*Y2

DV(U,V)= 8,861-,05670*X+,9649@-3*X2
          -,7672@-2*Y+,2504@-3*X*Y-,9130@-3*Y2
```

U	V	DU	DV	COR	DU err	DV err	||error||
-60	-60	30,16	10,73	,68	-,84	-3,09	3,20
-60	-20	32,01	20,73	,87	,77	4,90	4,96
-60	20	32,00	10,00	,00	,33	-4,92	4,93
-60	60	32,88	9,81	,74	,58	-1,28	1,41
-20	-60	33,92	4,89	,25	1,42	-2,96	3,28
-20	-20	31,77	8,70	,88	-,35	-1,57	1,61
-20	20	31,07	7,50	,82	-,86	-2,26	2,42
-20	60	30,87	6,66	,87	-1,07	,33	1,12
20	-60	37,17	7,68	,77	,33	2,69	2,71
20	-20	36,13	8,69	,79	,29	,89	,94
20	20	35,65	6,68	,71	,63	-1,02	1,19
20	60	35,18	6,57	,85	,77	1,91	2,06
60	-60	38,00	7,00	,00	-6,01	1,80	6,27
60	-20	41,11	6,00	,53	-1,27	-2,42	2,73
60	20	41,63	9,23	,51	,67	,51	,85
60	60	39,10	5,46	,30	-,62	-,63	,89

TOTAL WEIGHTED RMS ERROR= 2,43

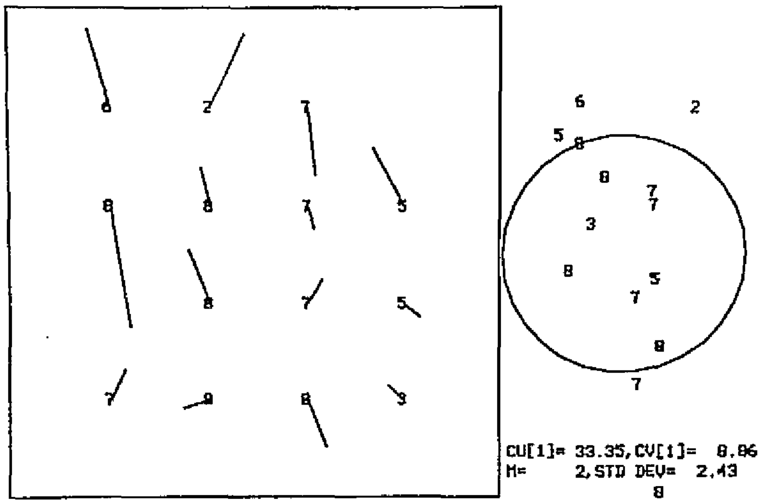


FIGURE 3-3 RESIDUAL ERRORS FROM TABLE 3-3

TABLE 3-4

POLYNOMIALS FIT TO ORDER 1 AFTER BLUNDER REMOVAL

DU(U,V)= 34.04 + .1013*X - .02093*Y
DV(U,V)= 7.738 - .0256*X - .01850*Y

U	V	DU	DV	COR	DU err	DV err	||error||
-60	-60	30.16	10.73	.68	.95	.34	1.01
-60	-20	32.01	20.73	.00	3.63	11.08	11.66
-60	20	32.00	10.00	.00	4.46	1.10	4.59
-60	60	32.88	9.81	.00	6.17	1.64	6.39
-20	-60	33.92	4.89	.00	.65	-4.47	4.51
-20	-20	31.77	8.70	.88	-.66	.08	.67
-20	20	31.07	7.50	.82	-.53	-.38	.66
-20	60	30.87	6.66	.87	.11	-.48	.49
20	-60	37.17	7.68	.77	-.16	-.66	.68
20	-20	36.13	8.69	.79	-.36	1.10	1.15
20	20	35.65	6.68	.71	-.00	-.18	.18
20	60	35.18	6.57	.85	.37	.46	.59
60	-60	38.00	7.00	.00	-3.38	-.31	3.39
60	-20	41.11	6.00	.53	.56	-.57	.80
60	20	41.63	9.23	.00	1.92	3.40	3.90
60	60	39.10	5.46	.30	.23	.36	.43

-- --------

TOTAL WEIGHTED RMS ERROR= .72

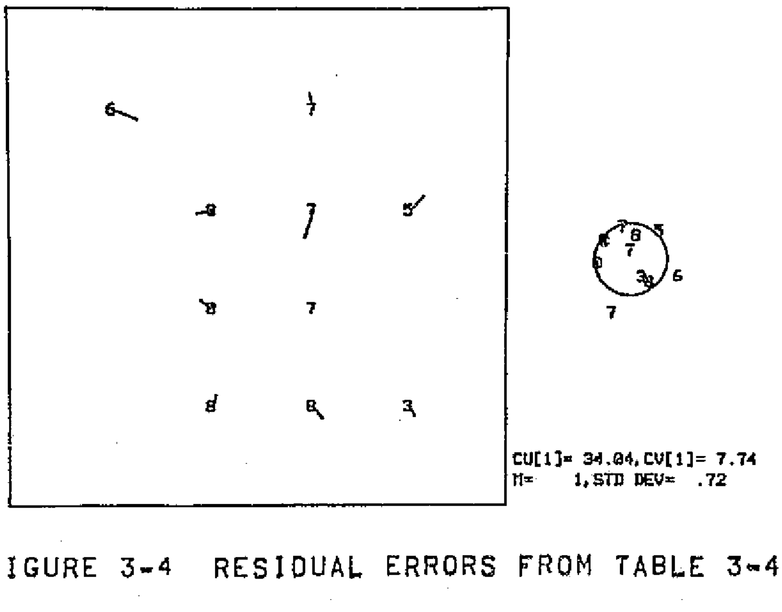


FIGURE 3-4 RESIDUAL ERRORS FROM TABLE 3-4

TABLE 3-5

POLYNOMIALS FIT TO ORDER 2 AFTER BLUNDER REMOVAL

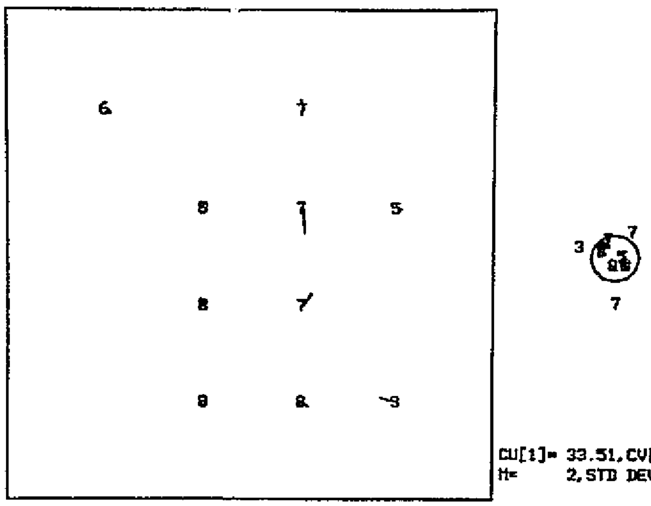
```
DU(U,V)= 33.51+.1015*X+.2499@-3*X2
          -.01951*Y+.1685@-5*X*Y+.1330@-3*Y2

DV(U,V)= 8.003-.02122*X-.1949@-3*X2
          -.01986*Y+.3108@-3*X*Y-.7619@-4*Y2
```

U	V	DU	DV	COR	DU err	DV err	||error||
-60	-60	30.16	10.73	.68	.19	.11	.22
-60	-20	32.01	20.73	.00	3.24	11.42	11.87
-60	20	32.00	10.00	.00	4.02	2.23	4.59
-60	60	32.88	9.81	.00	5.25	3.82	6.49
-20	-60	33.92	4.89	.00	.69	-4.75	4.80
-20	-20	31.77	8.70	.88	-.25	-.14	.29
-20	20	31.07	7.50	.82	-.18	-.30	.35
-20	60	30.87	6.66	.87	-.02	.15	.15
20	-60	37.17	7.68	.77	-.12	-.37	.39
20	-20	36.13	8.69	.79	.04	.95	.95
20	20	35.65	6.68	.71	.35	-.52	.62
20	60	35.18	6.57	.85	.23	.16	.29
60	-60	38.00	7.00	.00	-4.14	1.17	4.31
60	-20	41.11	6.00	.53	.17	-.02	.17
60	20	41.63	9.23	.00	1.46	3.25	3.57
60	60	39.10	5.46	.30	-.72	-.23	.75

-- --------

TOTAL WEIGHTED RMS ERROR= .47



CU[1]= 33.51,CV[1]= 8.00
N= 2,STD DEV= .47

FIGURE 3-5 RESIDUAL ERRORS FROM TABLE 3-5

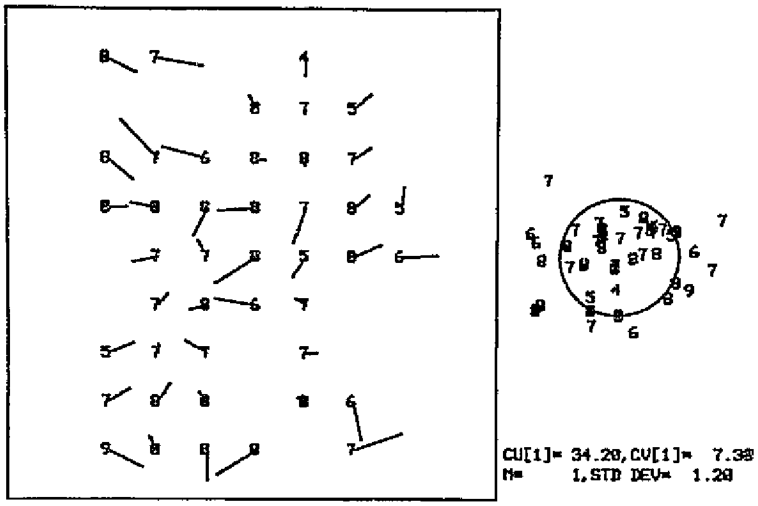


FIGURE 3-6 RESIDUAL ERRORS FROM POLYNOMIALS
FIT TO ORDER 1 AFTER BLUNDER REMOVAL

```
DU(U,V)= 34,20 + ,1083*X -,0141*Y
DV(U,V)= 7,297 - ,8098@-2*X -,6235@-2*Y
```

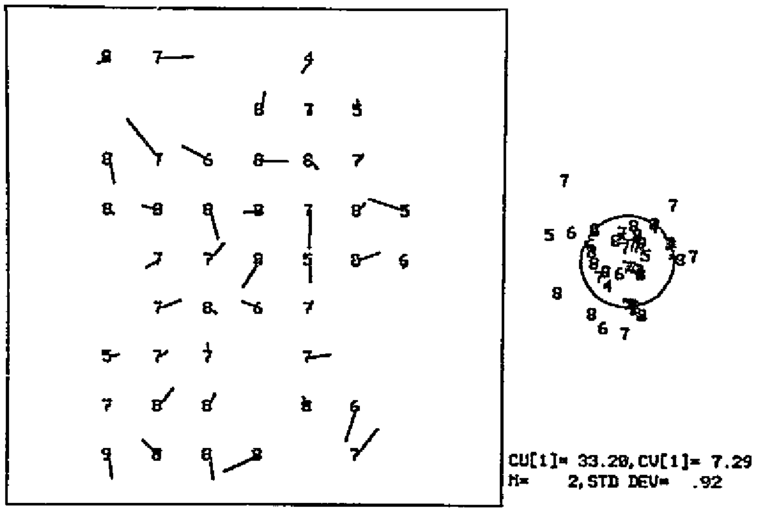


FIGURE 3-7 RESIDUAL ERRORS FROM POLYNOMIALS
FIT TO ORDER 2 AFTER BLUNDER REMOVAL

```
DU(U,V)= 33,20+,1174*X+,5789@-3*X2
            -,0150*Y+,1627@-4*X*Y+,1373@-3*Y2

DV(U,V)= 7,294-,8736@-2*X-,8608@-4*X2
            -,5937@-2*Y+,6122@-4*X*Y+,5104@-4*Y2
```

III.F RENORMALIZATION OF ONE IMAGE USING THE MISREGISTRATION MODEL

Once the geometric misregistration between two normalized images A and B has been detected and modeled, a new image B' can be generated using this model such that B' is in geometric alignment with image A according to equation 3.3. In practice, this requires the substitution of u+du(u,v) for u and v+dv(u,v) for v in equations (2.12) to (2.17).

Figures 3-9a to 3-10b show the results of the registration techniques presented in this chapter. Figures 3-9a and 3-9b are the same as figures 2-6a and 2-6b. Figure 3-10a is the same as figure 3-9a. Figure 3-10b shows the result of renormalizing 7F78 to align with 7F75 using the second order registration polynomials from figure 3-7.

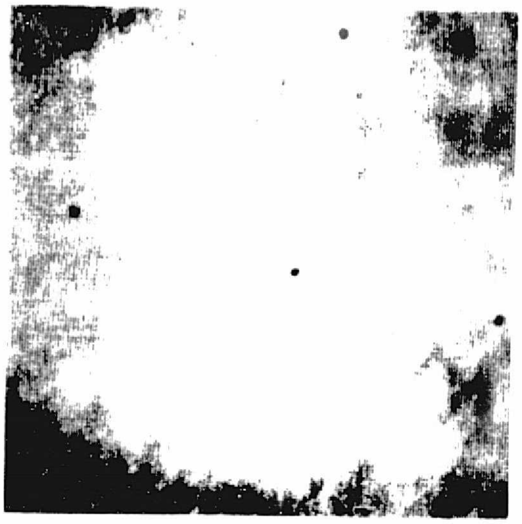


Figure 3-9a
7F75 Unaligned

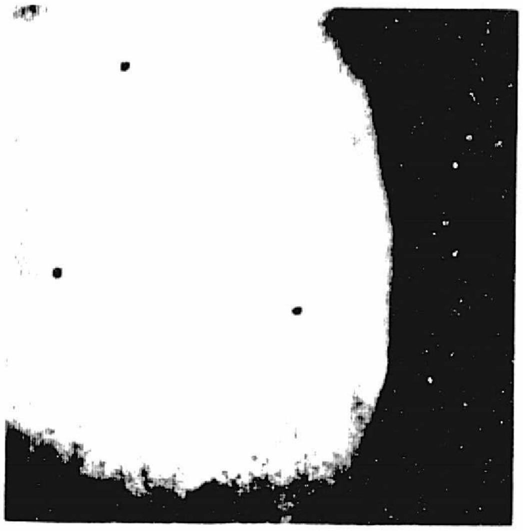


Figure 3-9b
7F78 Unaligned

Figure 3-10a
7F75 Aligned

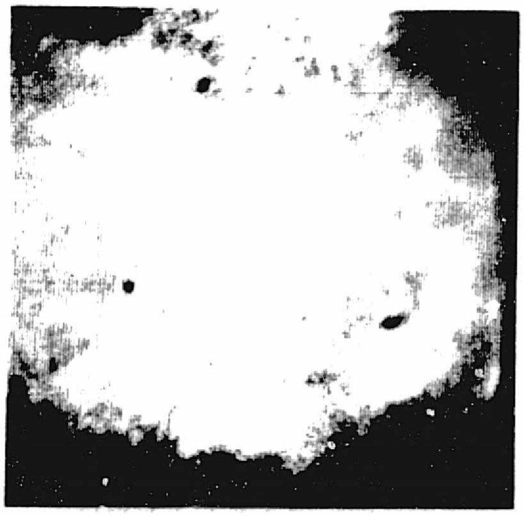


Figure 3-10b
7F78 Aligned

CHAPTER IV

PHOTOMETRIC NORMALIZATION BY "DEAD RECKONING"

The photometric normalization problem is to determine two functions R1(u,v) and R2(u,v) which define the reflectivity of the scene at point (u,v) in the respective images (and therefore at point T(u,v) in the scene). The reflectivity function can be determined by "dead reckoning" (from accurate calibrations and models) if the response of the vidicon and the reflectivity function of the scene as a function of the location in the scene and the illumination, view, and phase angles are precisely known. Images can be photometrically registered with respect to some type of errors in the photometric model. As for the geometric model, the combination of the dead reckoning and misregistration models is used to define the reflectivity function R(u,v).

Given the geometric and photometric normalization functions (R1,T1) and (R2,T2) for a pair of images, the picture comparison problem becomes one of comparing the reflectivities of corresponding points in the images, i.e., for all point pairs (u1,v1),(u2,v2) such that T1(u1,v1)=T2(u2,v2), compare R1(u1,v1) with R2(u2,v2).

IV.A PHOTOMETRIC MODEL FOR THE SCENE

The scattering of light by a surface is a very complex function of the spectral composition and polarization of the incident light, the composition of surface material, and the various angles between the incident light and the surface (angle i) the emergent light and the surface (angle ϵ) and between the incident light and the emergent light (phase angle). For some surfaces, with known light sources, the reflectance function is fairly well understood.

The reflectance function of the moon can be modeled by the Minnaert law (Minnaert [1961]):

$$Eout/Ein * \cos(\epsilon) = A * (\cos(\epsilon)*\cos(i))\uparrow k \qquad (4.1)$$

where A is the albedo of the surface, Ein and Eout are the light energy incident to and emergent from the surface, and k is a parameter particular to the surface and the phase angle.

Young has found (Young [1971]) that regions of Mars covered by the Mariner 6 and 7 pictures obey the Minnaert law fairly well; he found values for k ranging from .46 to

.71, and values for A ranging from .071 to .146.

Pollack has shown (Pollack [1969]) a phase angle dependence for the value of k. Although actual data exists for only a limited number of different phase angles, Pollack suggests that it is reasonable to interpolate the following table:

TABLE 4-1

phase angle	k
0 deg.	.5
50	.7
180	1.0

IV.B INTENSITY MODEL BETWEEN DATA POINTS

To compare two images (F1,T1) and (F2,T2) which were taken from from different camera positions and orientations it is necessary to compare pixels in image 1 with those pixels in image 2 which correspond in the scene. For discrete images, exact pixel correspondence does not exist since the projections of the pixels from the images onto the planet will not match in location, size and shape, and since the orientation of the image coordinate systems (scan direction) may be different.

Consequently, in order to compare such images it is necessary to model the intensity function between given pixels (Rosenfeld [1969], pp20-21).

ZERO ORDER MODEL: The simplest model for the intensity function F at non-grid points is the 0th order model F0(u,v):

$$F0(u,v) = F(i,j) \qquad (4.2)$$

where (i,j) is the closest sample point to (u,v). If one assumes that F is defined with (i,j) spaced on integer values, then the 0th order model F0 is equivalent to:

$$F0(u,v) = F([u+.5],[v+.5]) \qquad (4.3)$$

where [x] means the greatest integer which does not exceed x.

FIRST ORDER MODEL: The 1st order model F1(u,v) is bilinear interpolation which computes a weighted average of F at the four sample points which are nearest to (u,v).

$$\begin{aligned} F1(u,v) = {} & w1*F(i,j) + w2*F(i+1,j) \\ & +w3*F(i,j+1) + w4*F(i+1,j+1) \qquad (4.4) \end{aligned}$$

where

$$i=[u], j=[v] \qquad (4.5a)$$

$$p=u-[u], q=v-[v] \qquad (4.5b)$$

$$w1=(1-p)*(1-q) \qquad (4.5c)$$

$$w2=(1-p)*q \qquad (4.5d)$$

$$w3=p*(1-q) \qquad (4.5e)$$

$$w4=p*q \qquad (4.5f)$$

Note that $0 \le w[i] \le 1$ and $w1+w2+w3+w4=1$.

HIGHER ORDER POLYNOMIAL MODELS: If higher order surfaces are used for interpolation, instead of the bilinear model, one can require either that each piece of surface pass exactly through a large number of sample-point intensities, or alternatively, one can require that the pieces agree, not only in intensity, but also in the values of the various derivatives, along their curves of intersection, so that the resulting interpolated surface is not only continuous but also smooth.

SAMPLING THEORY MODEL: From a sampling theory point of view, a discrete image can be thought of as a bandwidth limited function of two variables, where the point spread function W determines the maximum frequency contained in the discrete image. Given this model, the function F can be exactly reconstructed from the sample points using a 2-dimensional generalization of the Shannon sampling theorem (Shannon [1949], Prosser [1966] pp. 574-584). However, this mathematically sound model for $\underline{F}$ is exceedingly expensive computationally, and of questionable practical improvement over simpler models for the intended application.

If it were necessary to measure the geometric positions of edges of features such as craters with errors of fractions of a pixel, then a higher order model would be necessary. For the Mariner '71 variable features problem, photometric accuracy is more important than such geometric accuracy, and consequently, the first order model (bilinear interpolation) was selected because of both its computational simplicity and its smoothing properties.

CHAPTER V

ACCURATE PHOTOMETRIC REGISTRATION OF NORMALIZED IMAGES

V.A FORMAL DEFINITION OF PHOTOMETRIC MISREGISTRATION

Assume that we have two photometrically normalized images with reflectance (albedo) functions R1(u,v) and R2(u,v), which are also geometrically normalized and registered as described in Chapters II and III. Assume also that the actual scene reflectivity is the same for the two images. The images are said to be "photometrically misregistered relatively" at point (u,v) if:

$$R1(u,v) \neq R2(u,v) \qquad (5.1)$$

If R(u,v) is the true scene reflectivity then an image R1 is said to be "photometrically misregistered absolutely" at point (u,v) if:

$$R1(u,v) \neq R(u,v) \qquad (5.2)$$

The following sections present a model for photometric misregistration and derive functions which describe this misregistration.

V.B MODEL FOR MISREGISTRATION

To photometrically register images more accurately than is possible with the "dead reckoning" photometric normalization requires that we have a model for photometric misregistration. There are many potential sources of photometric errors in images which have been "normalized" by Chapter IV. Some of these are:

1) Camera errors
   - a) Errors in light transfer function
   - b) Non-uniform sensitivity across vidicon surface
   - c) Residual images
   - d) Coherent noise
   - e) Random noise

2) Photometric model errors
   - a) Reflectance function variations with position on planet
   - b) Inappropriate values of k in equation (4.1)

A general model for photometric misregistration would be to assume that there is a function P which describes for each picture the combined photometric characteristics of the surface of the planet, the response of the vidicon, and the photometric model of Chapter IV. If the photometric errors are spatially uniform, that is, the errors do not depend on their location on the surface of either the vidicon or the planet, then the photometric registration function P is a function only of the reflectivity R(u,v) of each point. In general, P is a function of a set S of variables, such as the position on the vidicon surface, the position on the planet, and the illumination, view, and phase angles as well as R(u,v). Then, for two photometrically normalized images R1 and R2, we would have two photometric registration functions P1 and P2 such that:

$$P1(R1(u,v),S)) = R(u,v) = P2(R2(u,v),S) \qquad (5.3)$$

where R(u,v) is the true reflectivity (albedo) of point (u,v). If it is not necessary to obtain photometric registration to true reflectivities, then the images can be photometrically registered relative to one another using the function P12 which is defined:

$$P12(r,S) = P1'(P2(r,S),S) \qquad (5.4)$$

where P1' is the inverse of P1, or

$$R1(u,v) = P12(R2(u,v),S) \qquad (5.5)$$

The Mariner 6 and 7 pictures have been processed (Rindfleisch [1971]) by the Jet Propulsion Laboratory (JPL) to remove as many of these errors as possible. The high level of coherent noise present in the Mariner 6 and 7 pictures required that JPL develop some rather sophisticated selective filtering techniques. There is some question about the absolute photometric accuracy of these final pictures, and there are still errors due to the residual image processing in the area of the limb of the planet.

There does not presently exist an accurate model for the photometric errors present in the Mariner '69 pictures. However, after comparing images which were processed by the techniques of Chapters II, III, and IV, it was found that the predominant errors were due to errors in absolute camera sensitivity. Consequently, the following simple photometric registration function was used:

$$P12(r,S) = a*r + b \qquad (5.6)$$

where the parameter a adjusts for errors in camera sensitivity, and parameter b adjusts for any constant offset. This model assumes that a and b are constant for each point in images R1 and R2.

It might be better to allow a and b to be polynomials in the two variables u and v. Such a model would correct for errors which are spatially dependent. If it is possible to empirically derive a(u,v) and b(u,v) and if we could separate the contributions due to the camera from those due to the photometric model of the planet, then the empirical functions should be useful in refining our photometric model for Mars.

V,C A MEASURE OF PHOTOMETRIC MISREGISTRATION

Obviously, in any discussion of errors between sets points it is necessary to define some measure of the "tota error". The conventional RMS error measure was chosen primarily because the familiar least squares techniques minimize that measure.

The RMS error in photometric registration between two pictures R1 and R2 is then defined:

$$RMS(R1,R2) = \mathrm{sqrt}\ (\ \sum_{u,v} ((R1(u,v) - R2(u,v))\uparrow 2)\ /\ N) \qquad (5,7)$$

where N is the number of points in the pictures.

V,D CHOOSING A MODEL TO MINIMIZE PHOTOMETRIC MISREGISTRATION

The purpose of this section is to empirically derive the function P12 which "best" registers pictures R1 and R2 photometrically in terms of the RMS error. More precisely, we find P12(r,S) such that:

$$RMS(R1,P12(R2,S)) = \min \qquad (5,8)$$

or

$$\sum_{u,v} ((R1(u,v) - P12(R2(u,v),S))\uparrow 2) = \min \qquad (5,9)$$

For the simple model of equation (5.6), minimizing equation (5,9) is equivalent to minimizing:

$$E(R1\uparrow2) + a\uparrow2*E(R2\uparrow2) - 2a*E(R1*R2) - 2b*E(R1) + 2ab*E(R2) + b\uparrow2 = \mathrm{min} \qquad (5.10)$$

where E is the expected value. The solutions for a and b are:

$$a = \mathrm{cor}(R1,R2) * \mathrm{sd}(R1)/\mathrm{sd}(R2) \qquad (5.11a)$$

$$b = E(R1) - a*E(R2) \qquad (5.11b)$$

where sd is the standard deviation, and cor is the cross correlation defined by equation (3.5).

If a more complicated model is used, where a and b are functions of the (u,v) position in the image formed by linear combinations of functions Ak(u,v) and Bk(u,v), then we want to find coefficients a[k] and b[k] to minimize:

$$\operatorname*{sum}_{u,v} ((R1(u,v) - R2(u,v) * \operatorname*{sum}_{k}(a[k]*Ak(u,v)) - \operatorname*{sum}_{k}(b[k]*Bk(u,v)))\uparrow2 = \mathrm{min} \qquad (5.12)$$

where Ak and Bk are indexed by k. The values of the a[k] and b[k] are found in the traditional manner of least squares. Intuitively, first order polynomials in u and v appear to be useful for the functions A and B. Since this technique for photometric registration has not yet been used, the proper order and form for the A and B functions is not yet known. Only zero order models (equation (5.6)) have actually been tried.

CHAPTER VI

DIFFERENCING AND DIFFERENCE ANALYSIS

The title of this dissertation is "Computer Comparison of Pictures", but the previous four chapters have all dealt with image normalization and registration, and nothing has really been said about image comparison. Of course, all of these preliminary operations were necessary so that the differences between the two pictures represent actual changes in the scenes rather than differences due to the conditions of viewing and illumination.

This chapter describes techniques for image differencing and the analysis of these differences.

VI.A POINT BY POINT DIFFERENCES

The most obvious form of differencing of two normalized and registered pictures R1 and R2 is the point by point difference defined by:

$$D(u,v) = R1(u,v) - R2(u,v) \qquad (6.1a)$$

If the two images have not been photometrically registered, then the following point by point difference also performs the photometric registration:

$$D(u,v) = R1(u,v) - A(u,v)*R2(u,v) - B(u,v) \qquad (6.1b)$$

where A and B are the photometric misregistration polynomials derived in Chapter V. These difference pictures show where there are errors in the data due to noise of various forms, where there are errors in the photometric or geometric registration models, or, hopefully, where there are actual differences in the reflectivity of the scene.

Applying the latter differencing technique to the pictures in figures 6-1a and 6-1b (which are the same as 3-10a and 3-10b), we get the difference pictures which are shown in figures 6-2a and 6-2b. Figure 6-2a was produced by subtracting figure 6-1b from 6-1a (and adding a constant so that there are no negative numbers). Figure 6-2b is the negative of 6-2a.

The most obvious feature of these difference pictures is the large irregular blob in the lower-right corner, which is dark in 6-2a and bright in 6-2b, and whose cause will be explained later. There are numerous small circular light and dark spots which are due to reseau marks on the vidicons. (In future work, these areas will always be black). Reseau marks are deposits on the vidicon surface

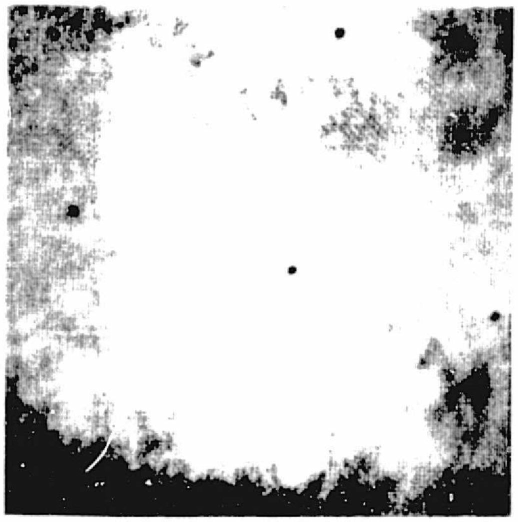


Figure 6-1a
7F75 Aligned

Figure 6-1b
7F78 Aligned

Figure 6-2a
7F75 - 7F78

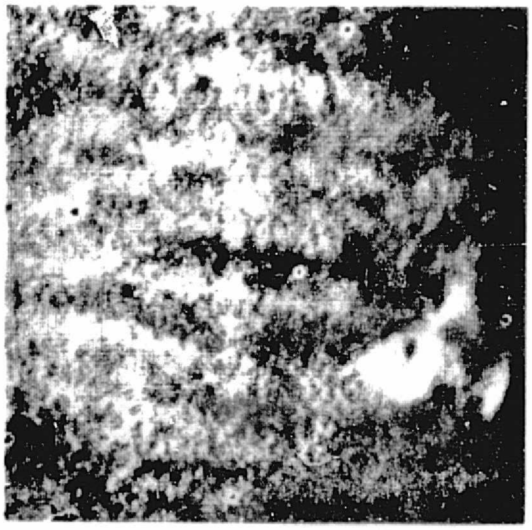


Figure 6-2b
7F78 - 7F75

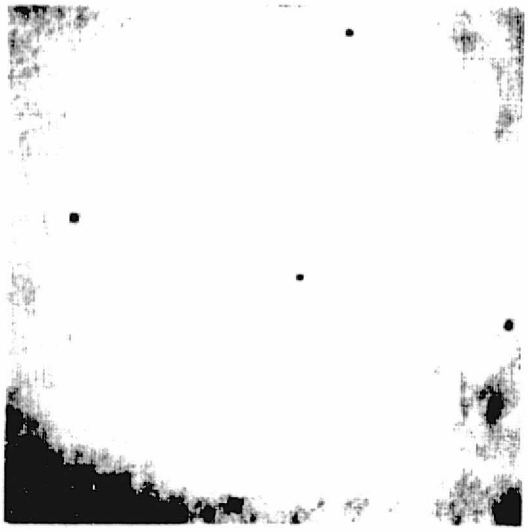


Figure 6-3a
7F75 Aligned

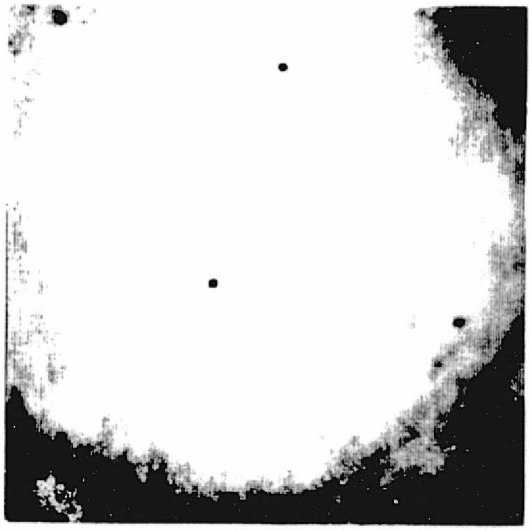


Figure 6-3b
7F77 Aligned

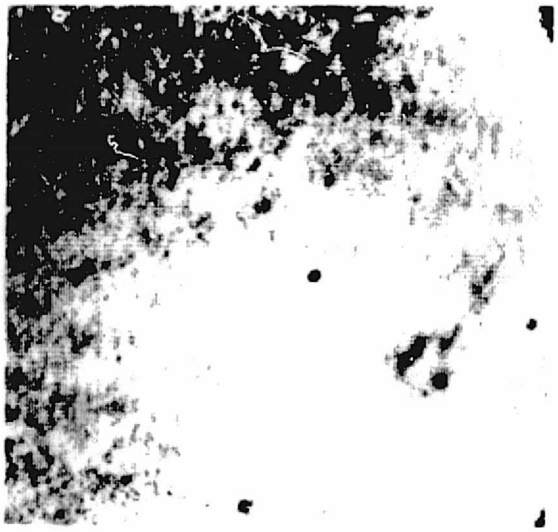


Figure 6-4a
7F75 - 7F77

Figure 6-4b
7F77 - 7F75

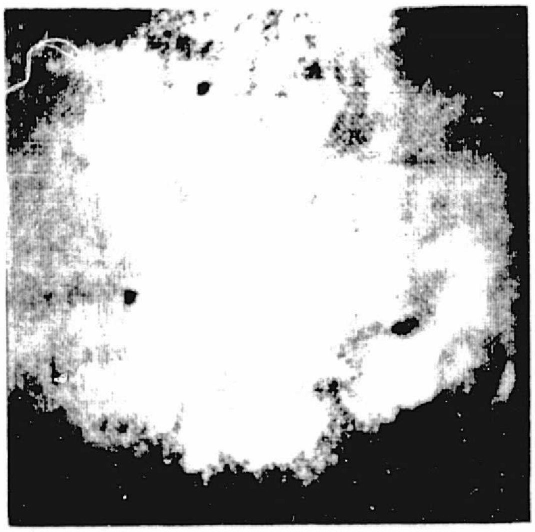


Figure 6-5a
7F78 Aligned

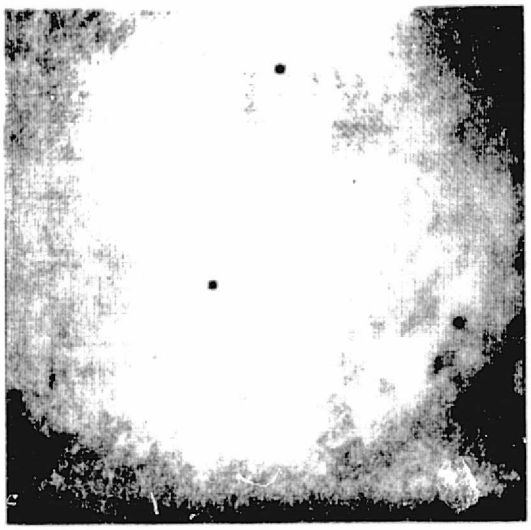


Figure 6-5b
7F77 Aligned

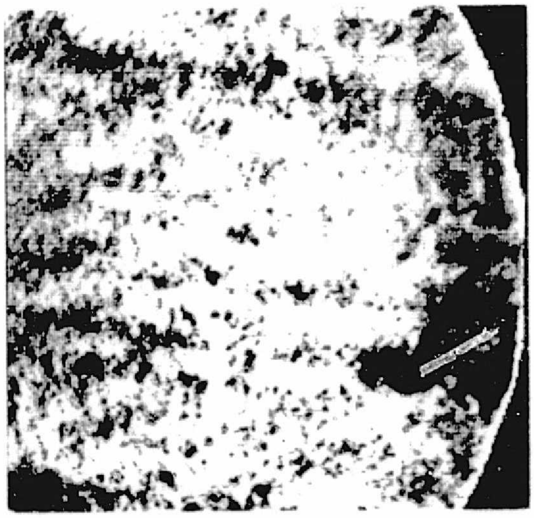


Figure 6-6a
7F78 - 7F77

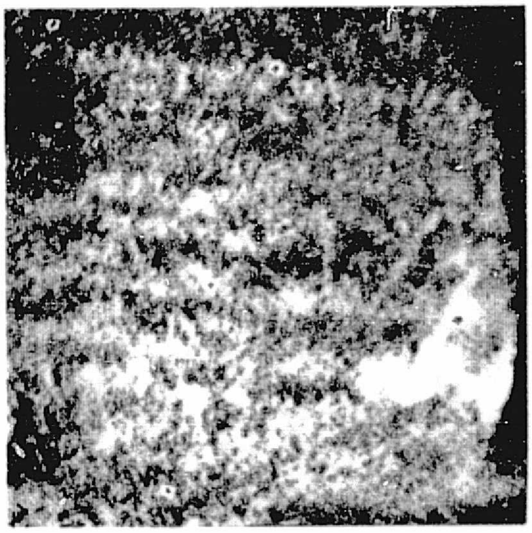


Figure 6-6b
7F77 - 7F78

for geometric alignment of the vidicon scan. There are also light and dark linear horizontal streaks which are oriented in the horizontal scan direction in the original camera coordinate system. These streaks are probably due to coherent noise in the camera system (Rindfleisch [1971]). Slight geometric misregistration is indicated around Nix Olympica, the large crater located at the top-center of the aligned pictures. A high level of visually uncorrelated noise is also present. The global dark to light shading from left to right in figure 6-2a can be explained by errors in the assumed dependence of the photometric function on viewing angle near the limb. A linear model for the photometric registration model of equation (5.12) would have reduced this error.

The only significant unexplained difference between figures 6-1a and 6-1b is the large blob which is brighter in 6-1b than in 6-1a. This blob has been attributed (Leovy [1971]) to a cloud consisting of water vapor. This difference is visible in figures 6-1a and 6-1b, but is very difficult to see in the original pictures (figures 2-5a and 2-5b).

Figures 6-3a through 6-6b show the results of introducing a third picture named 7F77, which was the picture taken immediately before 7F78 in the Mariner 7 far encounter sequence. Picture 7F77 was normalized to the same orthographic projection as was used for 7F75 and 7F78, and geometrically aligned with 7F75. The figures show differences between 7F77 and both 7F75 and 7F78. Since 7F77 was taken between 7F75 and 7F78 in time, it is reasonable that we see an intermediate state of the development of the cloud.

The higher level of noise in figures 6-6a and 6-6b is due to an increased gain factor introduced by a histogram "stretching" program. This stretching is used to adjust the intensity range of a picture to maximize the contrast when generating a photograph. These figures also show another camera related error across the top which is oriented in the horizontal scan direction of the camera. It should be emphasized that these pictures were never geometrically registered with one another, but both were registered to the same third picture (7F75). There seems to be little or no extra geometric misregistration due to cumulative errors as might be expected.

VI.B NOISE "REMOVAL"

The high levels of noise present in the difference pictures suggests that some form of noise removal is needed. The obvious problem with noise removal is that usually one cannot distinguish signal from noise, and consequently, when performing noise removal one also performs some signal removal.

If the noise is random, such that the noise in adjacent samples is uncorrelated, then one way to reduce it is by spatial filtering or averaging. Figures 6-7b, 6-8b, and 6-9b show the results of replacing each point in the difference pictures (6-7a, 6-8a, and 6-9a) by the local average computed over an 11x11-point window. This process effectively reduces the level of random noise, but the cloud, and the systematic errors in photometry still remain.

Other types of noise removal have been developed elsewhere but have not yet been tried here. One such technique is "salt and pepper" removal, which consists of replacing points which differ greatly from a surrounding set of approximately equivalent points. This technique (which is sometimes called "Custering" after the well known American general, who, when surrounded by Indians, was wiped out) would remove most of the random "shot" noise in the difference pictures without introducing the blurring created by spatial averaging.

VI.C CONTOURING ISO-DIFFERENCE LEVELS

The technique of contouring is often used in terrestrial map-making to emphasize lines of constant altitude. Recognizing the utility of this technique,

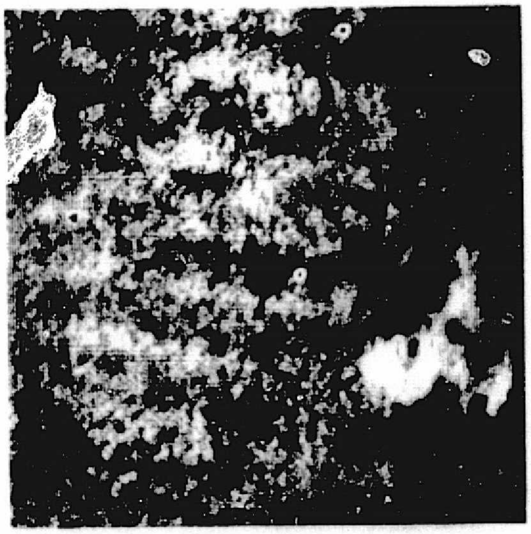


Figure 6-7a
7F78 - 7F75

Figure 6-7b
7F78 - 7F75 Averaged

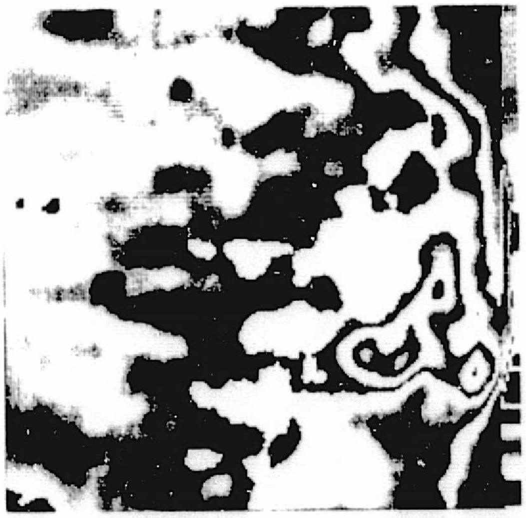


Figure 6-7c
7F78 - 7F75 Contoured

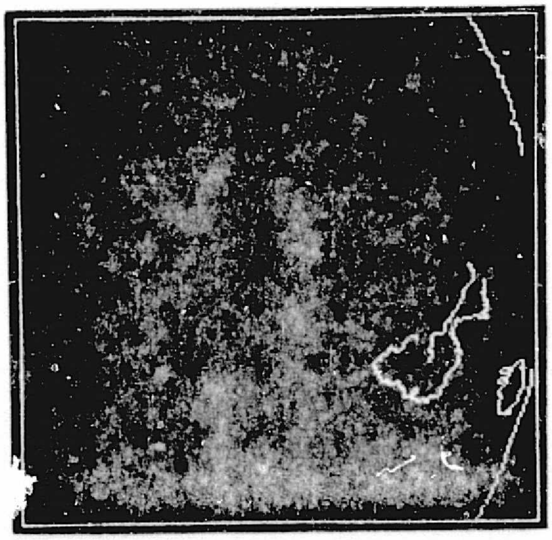


Figure 6-7d
7F78 - 7F75 Blobs

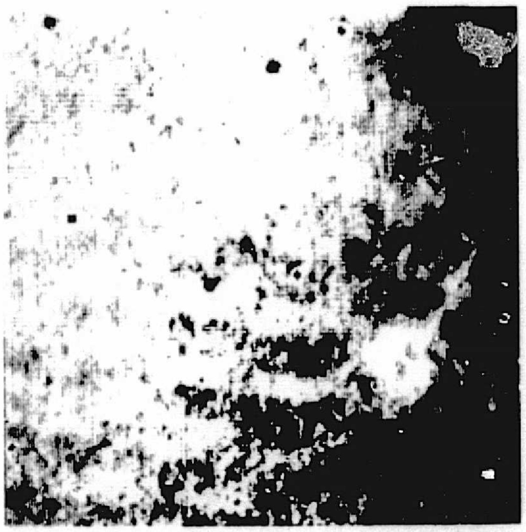


Figure 6-8a
7F77 - 7F75

Figure 6-8b
7F77 - 7F75 Averaged

Figure 6-8c
7F77 - 7F75 Contoured

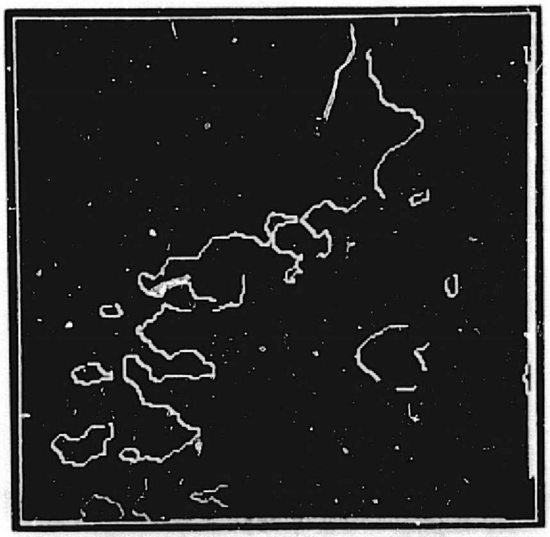


Figure 6-8d
7F77 - 7F75 Blobs

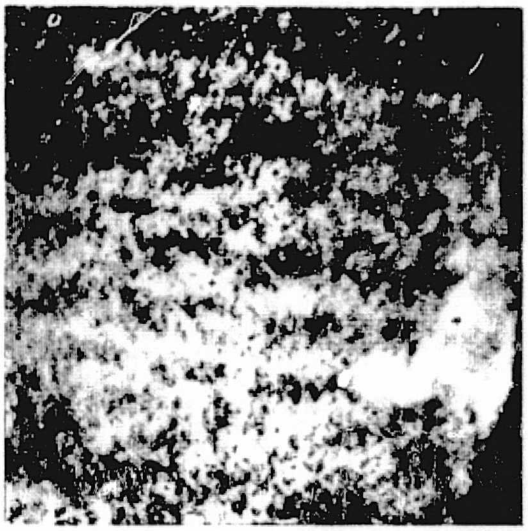


Figure 6-9a
7F78 - 7F77

Figure 6-9b
7F78 - 7F77 Averaged

Figure 6-9c
7F78 - 7F77 Contoured

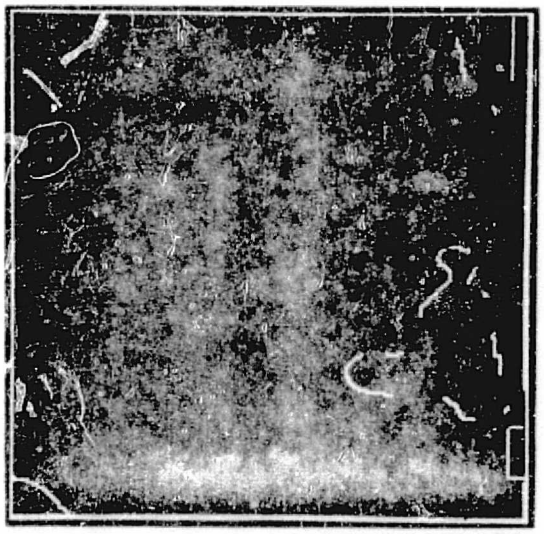


Figure 6-9d
7F78 - 7F77 Blobs

figures 6-7c, 6-8c, and 6-9c were produced by contouring the averaged pictures 6-7b, 6-8b, and 6-9b [1].

In these figures, the sharp dark to white edges occur at difference levels of 16, 32, and 48 in the corresponding spatially averaged pictures whose difference values range from 0 to 63. Computationally, these pictures were trivially generated by multiplying each difference sample by [illegible], and truncating the result to 6 bits (0 to 63 range).

These contour pictures fairly well outline the cloud, but they also show a lot of lines in areas which seem to contain no significant differences. In fact, one of the contour levels corresponds to the mean value of the difference picture, and therefore, one would expect many contours corresponding to small positive and negative deviations from the mean.

An improved contouring algorithm is planned which will select contour levels on the basis of the histogram of the difference picture, attempting to emphasize significant differences, but not small deviations from the mean.

---

[1] The spatially averaged pictures were used here because of the high random noise levels in the unaveraged pictures, which, when contoured, resulted primarily in garbage.

VI.D DIFFERENCE "BLOB" DETECTION

Another form of analysis of difference pictures is difference blob detection, which attempts to find the outlines of areas (blobs) of significant differences. The significance of a blob is determined by the magnitude of the differences, and by the length of the edge surrounding the difference blob.

Figures 6-7d, 6-8d, and 6-9d show the results of applying such a blob detector to the corresponding averaged pictures 6-7b, 6-8b, and 6-9b.

A further refinement of the blob detector (Hannah [1971]) attempts to determine which of the undifferenced pictures contains the difference blob. This determination is made by looking in the undifferenced pictures for edges in the vicinity of the edges of the difference blob. An edge is defined statistically in both pictures in terms of the means and variances both inside and outside of the difference blob. If the variance of one picture is significantly larger than that of the other (in the vicinity of the blob), then the blob is attributed to the picture with the larger variance. Otherwise, if the absolute difference in the means inside and outside of the blob is larger in one picture than in the other, then the blob is attributed to the picture with the larger difference. This technique is still being developed, and will be valuable to difference analysis.

Other difference analysis techniques are planned, but not yet implemented.

CHAPTER VII

COMPUTER IMPLEMENTATION

The techniques described in this dissertation were developed for use on a digital computer. Analog implementation of these techniques is very limited because of the difficulty in performing the necessary photometric and geometric normalizations and registrations using analog devices.

Except where otherwise noted, all of these techniques have been implemented by the author using the PDP-10 computer at the Stanford Artificial Intelligence Project. This computer is attached to 128K (1K = 1024) 36-bit words of core memory, and executes instructions at the rate of approximately one every 3 microseconds. The PDP-10 is "time-shared" among typically 20 users at a time, making it possible for many users to develop and debug programs interactively with very short "turn-around" times.

Most of the programs have been written in SAIL (Swinehart [1970]), a local dialect of ALGOL-60, and many of the time consuming inner loops have been written in the assembly language embedded in SAIL.

VII.A BLOCK DESCRIPTION OF THE PROGRAMS

The programs are divided both logically and physically into blocks which operate independently. Most of the major blocks of figure 7-1 correspond to the techniques described in Chapters II through VI.

Images are represented as packed arrays of light (or albedo) values, and are accessed using PDP-10 byte operators. As many samples as possible are packed in a word. Usually, either 6 or 9 bit samples are used, giving either 6 or 4 samples per 36-bit word. Thus, an image consisting of 200 samples per line, 200 lines, and 6 samples per word would require 6667 words of storage. Procedures are available to transfer pictures between core and disk, and to display images on either of two video synthesizers (section VII.B).

In the block diagrams, rectangular blocks indicate functional operations, and oval blocks indicate data. The description of each block follows:

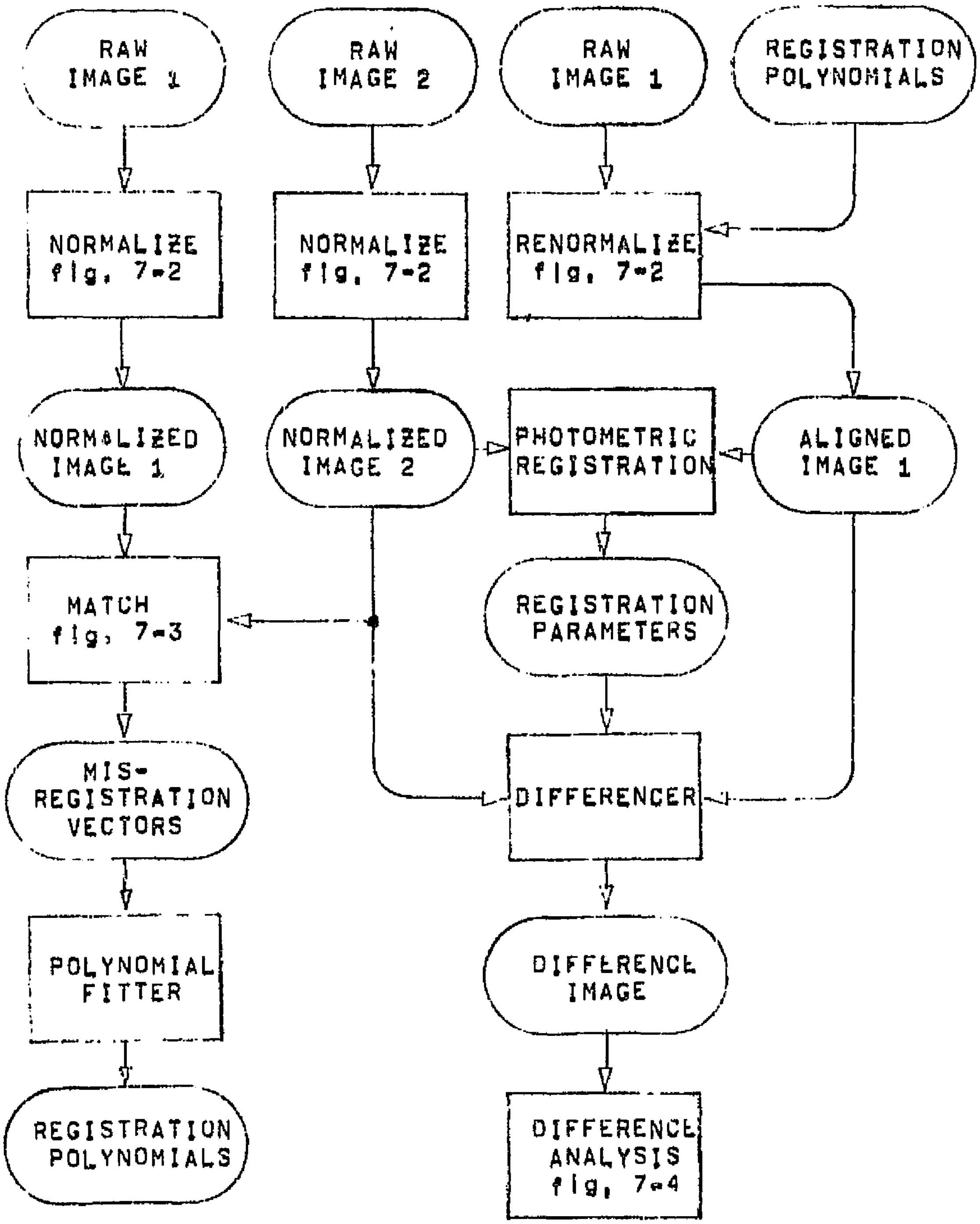


Figure 7-1 Block Diagram of System

VII.A.1 NORMALIZE

NORMALIZE is the name of the block which implements the normalization techniques of Chapters II and IV. The internal structure of NORMALIZE is shown in some detail in figure 7-2. The major functions of NORMALIZE are geometric and photometric normalization based on models for camera and scene geometry and photometry.

USER INTERACTION: The block labelled "user interaction" indicates control of the following parameters for the orthographic projection:

1) Direction (center latitude and longitude) of the projection.
2) Pixel size in the projected image (kilometers per pixel).
3) Horizontal and vertical size (in pixels) of the projected image.
4) Rotation of the projected image.
5) Position of the center of the projected image on the planet.

The combination of the models and the user interaction generates a set of projection parameters which control the actual normalization procedure labelled "PROJECTOR".

PROJECTION ANALYZER: This block analyzes the geometric calibration data for the spacecraft (i.e. spacecraft position and orientation at the time the picture was taken) to determine the area on the surface which is covered by the

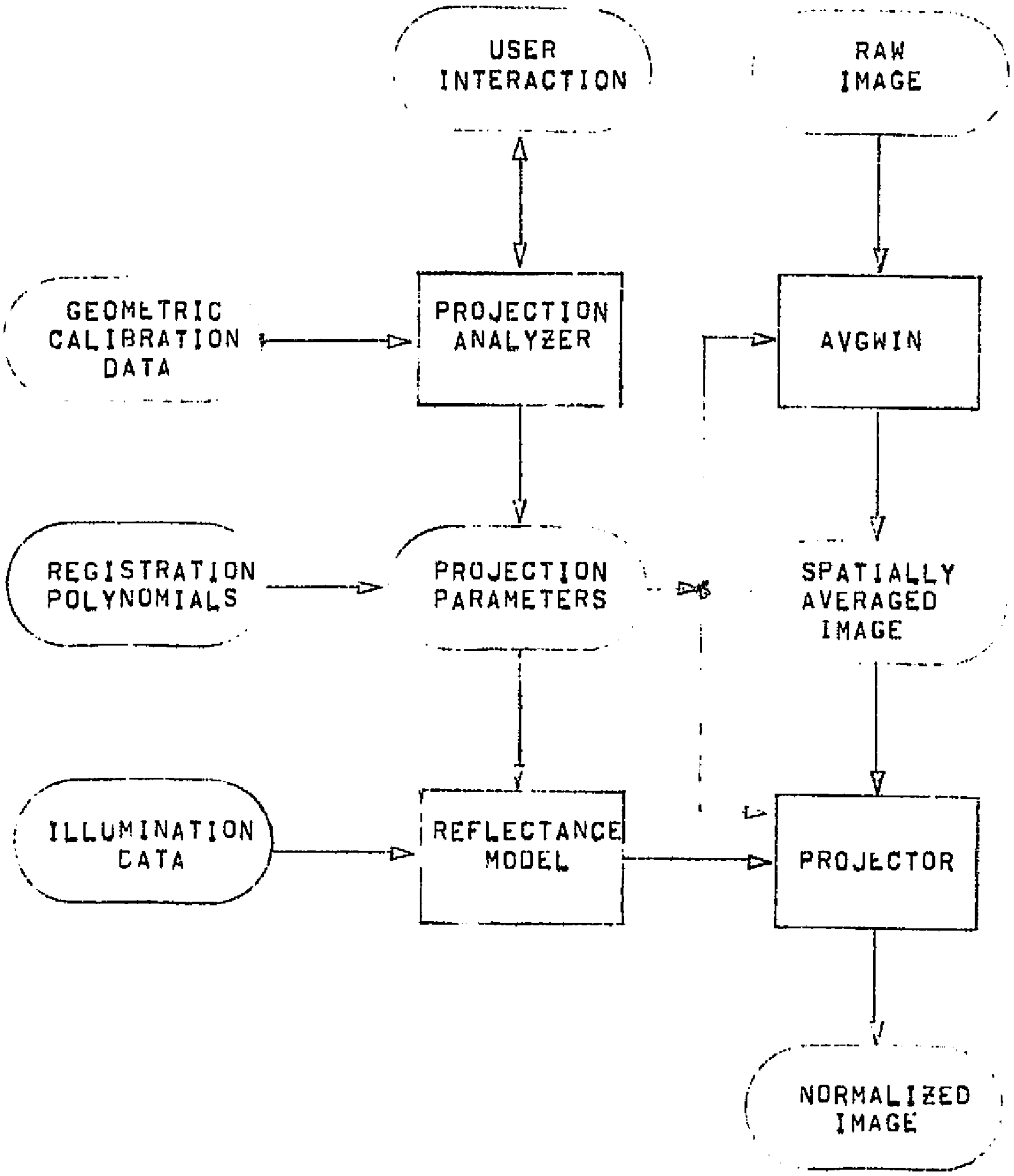


Figure 7-2 Block Diagram of NORMALIZE

picture. This analysis is usually done for pairs of pictures to determine the surface area which is included in both pictures. A graphical display of the spacecraft and planet geometry is presented to the interactive user to enable him to specify the desired projection.

AVGWIN: Since the Mariner '69 pictures are quite large (180K words) it is necessary to operate on windows of entire pictures, spatial averages of pictures, or combinations of windows and averages. AVGWIN computes an Nx by Ny spatial average of an arbitrary rectangular window of a picture, where Nx and Ny are the numbers of pixels averages in the vertical and horizontal directions. If pixels of the required projected image correspond in size to a very large number of pixels in the raw image, then spatial averaging is mathematically necessary because of the sampling theorem, and computationally useful to reduce the memory size required for the raw image.

PROJECTOR: This block generates a geometrically and photometrically normalized picture from a raw or averaged raw picture (which is hopefully correct in light values and camera geometry) using the normalization parameters and reflectance parameters. For each point in the projected image, PROJECTOR calculates its position in the raw image. Since this position is usually not exactly at a sample point, a bilinear interpolation of the light values of the four closest points is calculated. PROJECTOR also calculates the sun, view, and phase angles, which together with the light value, yield the albedo using equation (4.1).

REGISTRATION POLYNOMIALS: NORMALIZE performs the renormalization procedure described in section III.F using a pair of polynomials produced by the geometric misregistration model.

VII.A.2 MATCH

This block implements the techniques for determining geometric misregistration described in Chapter III. The internal structure of this block in shown in more detail in figure 7-3. Selected nxn windows of two normalized images are cross correlated according to equation (3.5). A search is made for the vector (du[i],dv[i]) which maximizes the value of the correlation function for integer values of du and dv. The correlation surface is modeled by polynomials in two variables in the vicinity of the maximum using a 2-D least squares program. The maximum value of this surface is found by computing partial derivatives finding the zero of the Euclidean norm of the partial derivatives of the surface, using a two-dimensional generalization of Newton's Method (section III.D).

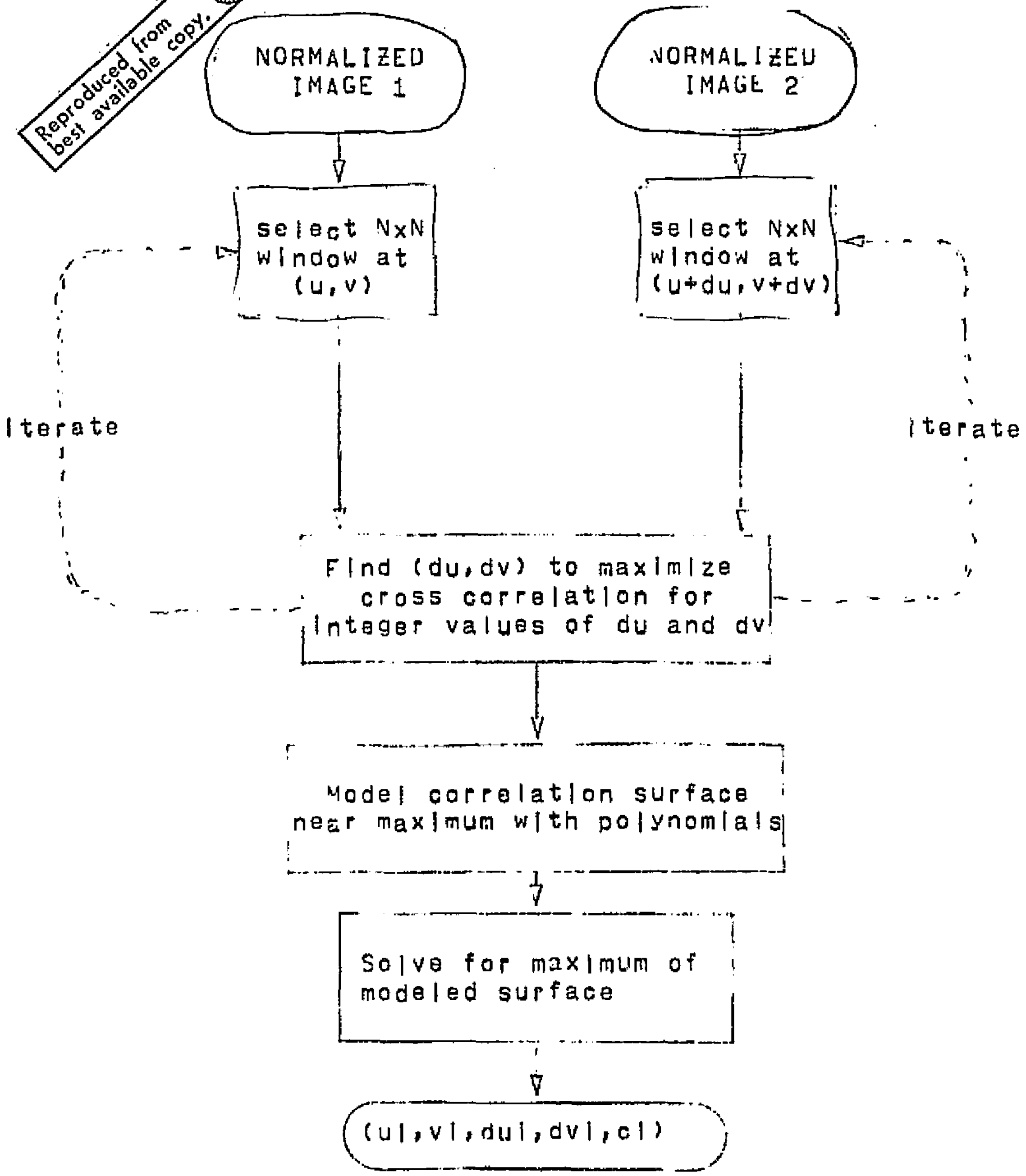


Figure 7-3 Block Description of MATCH

VII.A.3 POLYNOMIAL FITTER

This block performs a least square fit of two polynomials du(u,v) and dv(u,v) to the empirical correlation vectors (du[i],dv[i]) at sample points (u[i],v[i]). The values of the correlation peaks c[i] are used as weighting coefficients to the data points.

VII.A.4 PHOTOMETRIC REGISTRATION

This block implements the photometric registration techniques described in Chapter V. It computes the registration parameters which minimize the photometric misregistration between the geometrically aligned images.

VII.A.5 DIFFERENCER

The DIFFERENCER implements the pixel-by-pixel differencing described in section VI.A. The differencer uses the photometric registration model generated by the techniques of Chapter V to compute the difference picture defined by equation (6.1b).

VII.A.6 DIFFERENCE ANALYSIS

This block performs the difference analysis techniques described in Chapter VI, and is shown in more detail in figure 7-4.

VII.B IMAGE DISPLAY

Images which are generated by any of the image generating techniques of this dissertation can be displayed by either of two image display devices. Both of these hardware devices are controlled by programs which input pictures from the disk file system, and output the pictures on a selected area of the image generation surface, which in each case consists of a cathode ray tube.

VII.B.1 HIGH RESOLUTION VIDEO SYNTHESIZER

The high resolution video synthesizer consists of a collection of digital and analog circuitry to control the position and intensity of the beam of a Tektronix model 611 storage oscilloscope which is used in non-storage mode. This synthesizer generates grey level pictures one line at a time as follows:

A. The 611 oscilloscope beam is positioned at the (x,y) coordinates of first point on the line.

B. The 611 oscilloscope beam is turned on for a duration proportional to the intensity value for this point.

C. The beam is deflected to the right one sample position, and control is returned to step B if there are still points to display on this line.

D. When the line is finished control returns to step A.

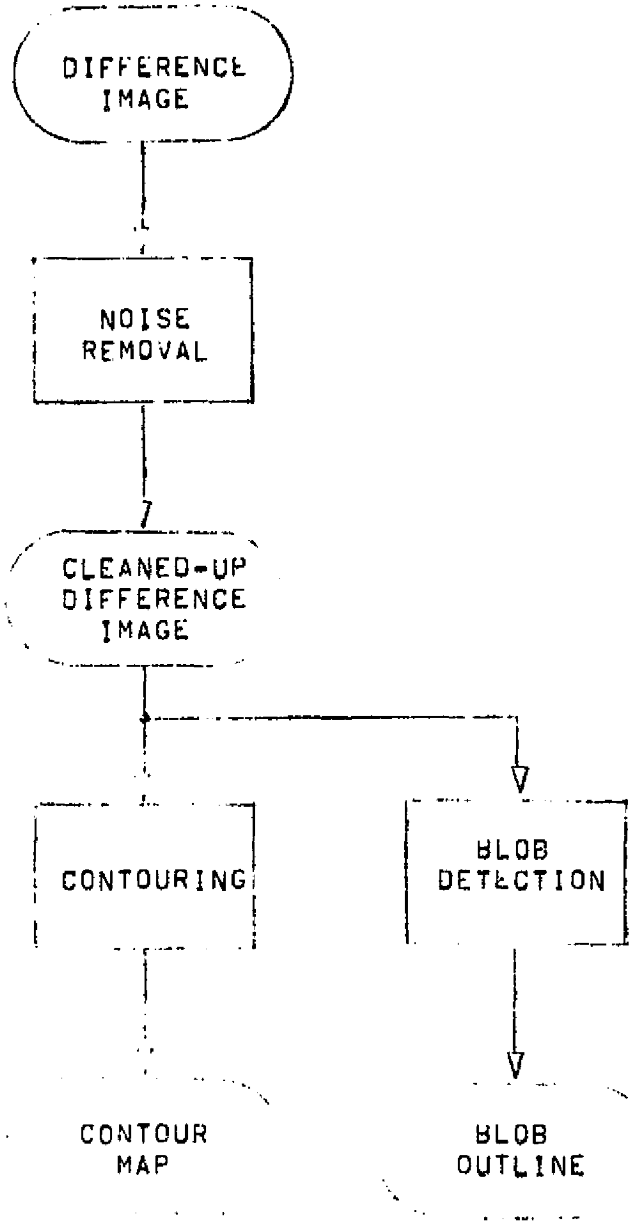


Figure 7-4 Difference Analysis

The hardware for the high resolution video synthesizer is theoretically capable of 4096- by 4096-point spatial resolution with 512 levels of grey. The system is currently limited by the capabilities of the 611 oscilloscope, which limits the effective resolution to approximately 700 by 700 points, with approximately 64 resolvable levels of grey.

The amount of time required to display a picture depends on both the size of the picture and its content. Bright points require that the beam be turned on longer than dark points. Typically, a picture requires 25 microseconds per point. This means that a 200 by 200 picture would be generated in one second. In order to view pictures which are generated so slowly, a photograph must be taken. A Polaroid camera is used for this purpose.

VII.B.2 REAL TIME VIDEO SYNTHESIZER

The real-time video synthesizer consists of a different collection of digital and analog circuitry which generates a television picture, using as input 8 channels from a 10 MHz digital disk (Data Disk). These 8 digital channels contain the bits of the binary coded sample values, which, when used as inputs to a digital to analog convertor, generate one of 256 analog voltages.

The picture consists of 480 scan lines of 512 points per line, with 256 levels of grey (approximately 128 usable levels). This picture can be either viewed directly or photographed.

VII.C EXECUTION TIMES

This section describes the CPU time required for some of the important and time-consuming parts of the system (CPU time refers to actual processor time, rather than elapsed time). These times are averages over several passes through the system, and are only approximate.

The CPU times to generate the 200x200 pixel pictures shown in this dissertation are as follows:

program	CPU time	comments
AVGWIN	10 sec.	
PROJECTOR	60 sec.	
MATCH	45 sec.	for 16 correlation windows
	220 sec.	for 81 correlation windows
POLYFIT	4 sec	for 1st order

There are the only programs whose execution times have been measured. None of the other programs require times comparable to either PROJECTOR or MATCH.

CHAPTER VIII

CONCLUSION

This final Chapter describes what has been accomplished towards solving the Mariner '71 Variable Features Problem, how this contributes to the general science of image processing, and plans for future refinements and applications of these techniques.

VIII.A ACCOMPLISHMENTS

The primary goal of this research was the development of image processing techniques to compare pictures of Mars which were taken from space under different conditions of viewing and illumination. The solution was to normalize the pictures to the same projection using geometric and photometric models for the camera and the scene. Geometric and photometric registration techniques were developed to remove the effects of errors in these models.

After pictures have been normalized and registered, they are analyzed for differences. Most of the differences which have been found between the Mariner 6 and 7 pictures are due to noise in the cameras. Some noise reduction and difference analysis techniques have been developed to aid the analysis of the difference pictures.

The examples presented in this paper show the detection of a real difference, due to a cloud on Mars. The system has been tested with other Mariner 6 and 7 pictures, which have not displayed such dramatic and recognizable differences.

The Jet Propulsion Laboratory devised a test of the capabilities of people and computers to detect variable features. JPL introduced artificial differences of 5% in reflectance into a Mariner 6 picture. This modified picture was then geometrically distorted to what JPL claimed was a different perspective view of Mars. The test was to see if people (or computers) could determine the differences between the original picture and the modified, distorted picture.

The result was that people were unable to see the differences, but the system described in this dissertation was able find most of the differences. The only differences that were difficult to detect occured because the JPL distorted picture was not a true perspective view of the planet, and, after geometric normalization and registration, some high order residual distortions remained.

No systematic difference analysis of the Mariner 6 and 7 pictures has been attempted yet. However, this analysis is planned, to test the Mars global model which is described in section VIII.C.

VIII.B CONTRIBUTIONS

The major contributions of this research to the science of picture processing are summarized as follows:

1) Accurate geometric registration based on cross correlation: Although cross correlation has been used in character recognition and scene congruence (Fischler [1971]), apparently no one has derived a model for the misregistration between two pictures and used this model to produce two pictures which are in geometric alignment.

2) Photometric normalization using the Minnaert scattering model.

3) Development of a system which combines a large variety of image processing techniques: Some of the techniques described in this paper have been used elsewhere, but have never been combined in a coherent system which

actually works with real (rather than synthetic) pictures,

4) The development of an image processing system for interaction with a planetary exploration mission: Although the system has not actually been tested in its intended application, it should satisfy most of the needs of variable features detection. The ability to rapidly detect changes on Mars provides the ability to increase the scientific return from the mission.

VIII.C FUTURE PLANS

Further research is planned to refine the techniques in this paper, to apply these techniques to new image processing problems, and to develop new capabilities for the Mariner '71 Mission.

REFINEMENTS:

Several improvements in the correlation maximization search technique are planned which will use information about the signal-to-noise ratios and autocorrelation functions of the pictures to adjust various search parameters.

The higher order photometric misregistration model of equation (5.12) will be implemented. This model should help to refine the parameter (k in equation (4.1)) of the

Minnaert function.

Some new picture difference operators are planned which would minimize the effects of one pixel misregistrations. This technique finds the difference between a point in picture A and a point in picture B which is closest in intensity value and a neighbor. An attempt will be made to introduce symmetry into this operator.

Tests of "significance" of differences need to be developed. Presumably, these tests would be based on a noise model for the camera, and some knowledge of the types of differences to expect.

OTHER APPLICATIONS:

Research is planned which would use the cross correlation registration technique to match areas of two parallax views of a scene. Using a set of match points, a depth model for the scene will be derived.

NEW CAPABILITIES FOR THE MARINER 1971 MISSION:

Additional capabilities are required for the Mariner '71 Mission. To detect variations on Mars which occur over long periods of time, it is necessary to maintain a "library" of previous pictures. New pictures would be compared with previous pictures of the same area on the surface.

Each picture contains approximately 5 million bits of data, which can be compressed by 3:1 using Huffman coded pixel differences. At least 1000 pictures with controlled sun and view angles are to be processed for differences. With the data compression, this amounts to 1.7 billion bits, which is about three times the storage capacity of our present disk file system. This quantity of data obviously will require a carefully designed data management and retrieval system.

A data structure is planned which will be organized by area on the surface and by original picture.

The surface of Mars will be divided into convenient areas, probably 10 degrees latitude by 10 degrees longitude. For each area a list will be maintained (on the disk file) of the following information:

1) A list of all pictures which cover this area.
2) A reference image for this area. This image represents this area to the best of our knowledge. The reference image is probably created by averaging pictures which do not demonstrate significant differences in this area. An image which shows the pixel by pixel variances of these pictures should probably be associated with the reference image.
3) Control points. Control points are features, such as

craters, whose positions are accurately known. These points are useful for registering images to fixed surface coordinates. For each control point, the data structure should include its planet coordinates, and a small image of the control point for use by the correlation procedure.

4) A list of any significant differences seen in this area. For each such difference there might be a difference image, a contour map, a blob outline, or some other representation of the difference. The actual difference information need not be currently stored on the disk file system, but information must be present to retrieve it from magnetic tape.

For each picture the following list will be maintained:

1) Picture identification (e.g. 7F75).
2) Calibration data, i.e. spacecraft position and orientation, and sun angle.
3) List of surface areas covered by this picture.
4) Location in the magnetic tape library of:
    a) Original version of the picture.
    b) Locally processed version of the original.
    c) A list of all other images which were derived from this picture, and the necessary retrieval data.

VIII.D PHILOSOPHICAL REMARKS

Some general remarks about the image processing philosophy are an appropriate conclusion to this dissertation. There are many difficult problems in this field, which can be made either harder or easier depending on the approach taken to solve them. While it is recognized that models for the camera and scene are important, one cannot depend totally upon accurate calibration of these models to eliminate the registration problems, whether the working environment is the surface of Mars, the surface of a flat table, or a highway.

It is important that techniques be developed which are sufficiently general that they can be applied to unanticipated environments. It is believed that most of the techniques in this paper can satisfy this requirement by suitably changing the various models.

The unmanned space exploration programs should provide excellent application areas for computer image processing, since, as the spacecraft probe deeper into space, the needs for onboard image processing will increase.

All exploration of the surface of other planets within the foreseeable future must be done by unmanned landing vehicles. Any intelligent exploration of other planets will require some level of onboard image processing capability in order that the vehicle can navigate and interact with its environment.

It is hoped that the results of this and other research in the application of computers will be used by man to better understand the universe around him, and to free man from the trivia of daily life, rather than to provide men the tool to enslave others.

APPENDIX A.

DERIVATION OF MATRIX R

This appendix gives the derivation of the matrix R of equation (2.4), given calibration data which describe the position and orientation of the camera.

The rotation matrix R is derived in several steps from the geometric calibration data. The first step derives a rotation matrix R0 which rotates the principal ray Pp of the camera, given in scene coordinates, into point (0,0,w) which is assumed to be the center of the image plane, given in camera coordinates.

$$(0,0,||Pp||)' = R1(Pp) \qquad (A.1)$$

This rotation matrix is easily constructed by first rotating Pp in the x-y plane by matrix R1 into (x',0,z)', and then rotating (x',0,z)' in the x-z plane by matrix R2 into (0,0,z')'. The matrices R0, R1, and R2 are defined:

$$R1 = \begin{matrix} x/k & y/k & 0 \\ -y/k & x/k & 0 \\ 0 & 0 & 1 \end{matrix} \qquad (A.2)$$

where $k=\mathrm{sqrt}(x\uparrow 2+y\uparrow 2)$,

$$R2 = \begin{matrix} z/h & 0 & -k/h \\ 0 & 1 & 0 \\ k/h & 0 & z/h \end{matrix} \qquad (A,3)$$

where $h = \mathrm{sqrt}(k\uparrow 2+z\uparrow 2) = \mathrm{sqrt}(x\uparrow 2+y\uparrow 2+z\uparrow 2)$.

Then we define $R0 = R1*R2$. $R0$ is not unique, since a rotation around the principal ray is still possible. This ambiguity is solved by specifying the direction of one other vector in both scene and image coordinates. Given vector V in the scene and the angle α between the u-axis and the projection of V on the image plane the rotation matrix R3 is defined:

$$R0(V) = (x,y,z), \qquad (A,4)$$

If the direction of V in the image plane is β, then $\cos(\beta) = x/||x,y||$, and $\sin(\beta) = y/||x,y||$. To orient the vector $R0(V)$ in the direction α, it is sufficient to rotate $R0(V)$ by an angle of $\alpha-\beta$. Rotation matrix R3 is then defined:

$$R3 = \begin{matrix} \cos(\alpha-\beta) & \sin(\alpha-\beta) & 0 \\ -\sin(\alpha-\beta) & \cos(\alpha-\beta) & 0 \\ 0 & 0 & 1 \end{matrix} \qquad (A,5)$$

Finally, the matrices R4 and R are defined:

$$R4 = R3 * R0 \qquad (A,6)$$

$$R = \mathrm{inv}(R4) \qquad (A,7)$$

For Mariner 6 and 7 pictures of Mars, the camera position is derived from the latitude and longitude of the sub-spacecraft point on the surface, and the altitude of the spacecraft. The orientation of the camera is determined from the latitude and longitude of point Pp, the intersection of the principal optic ray with the surface, and the angle in the image plane between the u-axis and a north-pointing tangent vector at point Pp projected on the image plane. If Pp is given in spherical coordinates as:

$$\begin{aligned} Pp = r * (\ & \cos(long)*\cos(lat) \\ & ,\sin(long)*\cos(lat) \\ & ,\sin(lat))' \end{aligned} \qquad (A.8)$$

where lat is latitude, long is longitude, and r is the radius of the sphere (Mars), then the north-pointing tangent vector to the sphere at this point is:

$$\begin{aligned} V = (\ & -\cos(long)*\sin(lat) \\ & ,-\sin(long)*\sin(lat) \\ & ,\cos(lat))' \end{aligned} \qquad (A.9)$$

APPENDIX B

THE EFFECTS OF NON-TRANSLATIONAL GEOMETRIC DISTORTIONS ON CORRELATION

This appendix analyzes the effects of non-translational geometric distortions between the windows of two images on the value of the correlation function. An upper limit on the size of the correlation window is derived as a function of the magnitudes of the higher order errors and the spatial frequency information in the windows. It is shown that the effects of non-translational geometric distortion will be less if the images contain primarily low spatial frequencies rather than high spatial frequencies.

A.1 ROTATION AND SCALE CHANGE

Rotation and scale factor errors between correlation windows can be modeled with a first order misalignment model:

$$u' = a*u + b*v + c \;=\; Su * (u*\cos(\alpha) + v*\sin(\alpha)) \quad (B.1a)$$

$$v' = d*u + e*v + f \;=\; Sv * (v*\cos(\beta) - u*\sin(\beta)) \quad (B.1b)$$

where

$$\alpha = arctan(b/a) \qquad (B.2a)$$

$$\beta = arctan(d/e) \qquad (B.2b)$$

$$Su = sqrt(a\uparrow 2+b\uparrow 2) \qquad (B.2c)$$

$$Sv = sqrt(d\uparrow 2+e\uparrow 2) \qquad (B.2d)$$

Presumably, the rotation angles α and β are small and the scale factors Su and Sv are approximately unity.

The analysis of the alignment errors between points in two windows whose centers are properly aligned (i.e. $c=f=0$) will be presented for the special case of:

$$S = Su = Sv \quad \text{and} \quad \alpha = \beta$$

$$\begin{aligned} & ||u'-u,v'-v||\uparrow 2 \\ & \quad = (u*(S*cos(\alpha)-1) + v*S*sin(\alpha))\uparrow 2 \qquad (B.3a) \\ & \quad\quad +(v*(S*cos(\alpha)-1) - u*S*sin(\alpha))\uparrow 2 \\ & \quad = ||u,v||\uparrow 2 * [\ (1-S)\uparrow 2 + 2*S*(1-cos(\alpha))\ ] \qquad (B.3b) \end{aligned}$$

Hence we have a linear model, i.e.

$$||u'-u,v'-v|| = Q(\alpha,S) * ||u,v|| \qquad (B.4)$$

$$Q(\alpha,1) = 2*sin(\alpha/2) \qquad (B.5a)$$

$$Q(0,S) = |\ 1-S\ | \qquad (B.5b)$$

If the centers of two correlation windows are properly aligned, then the misalignment error due to rotation and scale change at a point n pixels from the center is Q*n.

Thus far we have only estimated the magnitude of the error in the position of corresponding pixels based on our knowledge of the distortion process. We now want to know the effect of this misalignment on the value of the correlation function.

Fischler assumes (Fischler [1971]) that adjacent pixels are statistically independent, so that there is no contribution to the product term E(X*Y) by multiplying pixels which do not geometrically match. Fischler then derives limits on the size of the correlation window as a function of the rotation and scale change. His derivation is reinterpreted as follows:

Assume that we have two arrays of pixels X[i,j] and Y[i,j] which are geometrically in alignment. Suppose that we have a geometric distortion of the coordinate system such that a new array of pixels Z is created, where each pixel Z[i,j] is derived by bilinear interpolation (the same derivation holds for higher order interpolation and weighted averages in general) from a point at (u,v) in array Y.
Let k=[u],l=[v]; then

$$Z[i,j] = a*Y[k,l] + b*Y[k+1,l] + c*Y[k,l+1] + d*Y[k+1,l+1] \quad (B.6)$$

where a,b,c, and d satisfy the conditions of equations (4.5a-f).

Let us assume for simplicity (with no real loss of generality) that the images have been normalized to zero mean intensity values. Fischler's assumption of independence than implies that if $|u-i|\geq 1$ or $|v-j|\geq 1$ then $Z[i,j]$ was derived from 4 points of Y which are independent of $Y[i,j]$ and $X[i,j]$. Hence, such points contribute $E(X)*E(Y)=0$ to the sum of the $X[i,j]*Y[i,j]$ in equation (3.5).

If $|u-i|<1$ and $|v-j|<1$ then

$$\sum_{i,j} X[i,j]*Z[i,j] = \sum_{i,j} a[i,j]*X[i,j]*Y[i,j] \qquad (B.7a)$$

$$+ (1-a[i,j])*X[i,j]*(3 \text{ independent points of } Y)$$

$$= \sum_{i,j} a[i,j]*X[i,j]*Y[i,j]. \qquad (B.7b)$$

and there exists an A such that $0\leq A\leq 1$ and

$$\sum_{i,j} X[i,j]*Z[i,j] = \sum_{i,j} (a[i,j]*X[i,j]*Y[i,j]) \qquad (B.8a)$$

$$= A * \sum_{i,j} (X[i,j]*Y[i,j]) \qquad (B.8b)$$

$$E(X*Z) = A*E(X*Y). \qquad (B.9)$$

Therefore the only effect of the distorted coordinate system is to attenuate the $E(X*Y)$ term which attenuates the value of the correlation function. Of course, the value of A depends on the magnitude of the geometric distortion errors of equation (B.3).

A limit on the size of the correlation window is derived from equation (B.4) by limiting the maximum error to one pixel:

$$n = \|u,v\| \qquad (B.10a)$$

$$= \|u'-u,v'-v\| \,/\, Q(\alpha,S) \qquad (B.10b)$$

$$= 1/Q(\alpha,S) \qquad (B.10c)$$

For Mariner 6 and 7 pictures, the maximum combined rotation errors between two images are about 2 degrees and the scale changes are less than about 2%. Therefore,

$$n^2 < 1/(.02^2 + 2*(1.02)*(1-\cos(2))) \qquad (B.11)$$

$$\text{or } n < 25$$

Fischler's independence assumption essentially says that the autocorrelation functions for X and Y are one at u=0 and v=0 and zero elsewhere. We know that reasonable pictures do not have such autocorrelation surfaces. In particular, the autocorrelation surfaces have central peaks whose width depends on the low frequency power of the power spectra of the images. Consequently, knowing the autocorrelation function tells how large the errors in the alignment of two images can be before the cross correlation is adversely changed. If one has a choice of which areas in the two images to align, then the areas can be chosen on the basis of having broadly peaked autocorrelation functions.

REFERENCES

Briggs, G. [1971], "Mariner Mars '71 Mission B - A Proposed Change in the Orbital Period", Bellcom, Inc, Jan. 26, 1971.

Cutts, J., et al.[1970], "The Surface of Mars III - Light and Dark Markings", J. Geophysical Research, Vol. 76, No. 2, Jan. 10, 1971.

Doyle, F. [1966], "Manual of Photogrammetry", Vol. 1, Chapter X, American Society of Photogrammetry, Falls Church, Va.

Fischler, Martin A. [1971], "Aspects of the Detection of Scene Congruence", Proceedings of the Second Annual Joint Conference on Artificial Intelligence.

Hannah, M. [1971], "Edge Verification", Stanford University CS-225, CS-293 Term Project Report.

JPL [1970], "Mariner Mars 1971", Jet Propulsion Laboratory Technical Memo No. TM 33-449.

Kowalik J. and Osborne M. R. [1968], "Methods for Unconstrained Optimization Problems", New York: American Elsevier Publishing Co.

Lange, F. H. [1967], "Correlation Techniques", Van Nostrand, Princton, New Jersey.

Leovy, C. B., et al. [1971], "Mariner Mars 1969: Atmospheric Results", J. Geophysical Research, Vol. 76, No. 2, Jan. 10, 1971.

Minnaert, M. [1961], "Photometry of the Moon", in Planets and Satellites, chapt. 6, edited by G. P. Kuiper, Univ. of Chicago Press, Chicago Ill.

Pollack, J., Sagan, C. [1969], "An Analysis of Martian Photometry and Polarimetry", Vol. 9, p. 243, Space Sciences Review.

Prosser, R. T. [1966], "A Multidimensional Sampling Theorem", J. MATH. ANAL. APPL. 16.

Rindfleisch, T. [1966], "Photometric Method for Lunar Topography", Photogrammetric Engineering, March 1966.

Rindfleisch, T., et al, [1971], "Digital Processing of the Mariner 6 and 7 Pictures", J. Geophysical Research, Vol. 76, No. 2, Jan. 10, 1971.

Rosenfeld A. [1969], "Picture Processing by Computer", New York: Academic Press.

Sagan, C. and Pollack, J. [1969], "Optimization of the Mariner Mars 1971 Mission for the Analysis of Time-Variable Phenomena", Variable Features Team Memo, Cornell University.

Shannon, C. and Weaver, W. [1949], "The Mathematical Theory of Communication", University of Illinois Press.

Swinehart, D., Sproull, R. [1970], "SAIL", Stanford Artificial Intelligence Project Operating Note No. 57.1.

Young A. T. and Collins, S. A. [1971], "Photometric Properties of the Mariner Cameras and of Selected Regions on Mars", J. Geophysical Research, Vol. 76, No. 2, Jan. 10, 1971.