Motion Estimation from Sensor Images

Submitted By

Anupa Shah

13MCEC18

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2015

Motion Estimation from Sensor Images

Major Project

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering

Submitted By

Anupa Shah

(13MCEC18)

Guided By

Mr. Amit Dave(SAC-ISRO)

Prof. Manish Chturvedi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2015

Certificate

This is to certify that the major project entitled **"Motion Estimation from Sensor Images"** submitted by **Anupa Shah (Roll No: 13MCEC18)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering of Institute of Technology, Nirma University, Ahmedabad, is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this project, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Mr Amit Dave
External Guide & Scientist(SG),
SEDA-PCEG-PCSVD,
Space Applications Center,
ISRO(SAC), Ahmedabad.

Prof. Manish Chturvedi
Guide & Associate Professor,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Prof. Vijay Ukani
Associate Professor,
Coordinator M.Tech - CSE,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. Sanjay Garg
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. K Kotecha
Director,
Institute of Technology,
Nirma University, Ahmedabad.

भारत सरकार
अंतरिक्ष विभाग
अंतरिक्ष उपयोग केन्द्र
आंबावाडी विस्तार डाक धर,
अहमदाबाद-380 015. (भारत)
दूरभाष : +91-79-26912000, 26915000
फैक्स :

इसरो isro

Government of India
Department of Space
SPACE APPLICATIONS CENTRE
Ambawadi Vistar P.O.
Ahmedabad - 380 015. (INDIA)
Telephone : +91-79-26912000, 26915000
Fax

May 5, 201

CERTIFICATE OF PRACTICAL TRAINING

Name of the student	:	SHAH ANUPA MUKESHKUMAR
Name of the institution	:	INSTITUTE OF TECHNOLOGY, NIRMA UNIVERSITY, AHMEDABAD
Period of training	:	03/06/2014 to 30/04/2015
Name of the division where training was taken	:	PCSVD/PCEG/SEDA

DETAILS OF TRAINING

"MOTION ESTIMATION FROM SENSOR IMAGES"

PROFICIENCY SHOWN BY THE CANDIDATE

Attendance	:	REGULAR
Conduct	:	VERY GOOD
Practical ability	:	She is good at SIFT algorithm and implementation using MATLAB. She is sin and hardworking.

(J RAVISANK
HEAD, HRD

Statement of Originality

I, **Anupa Shah**, Roll. No. **13MCEC18**, give undertaking that the Major Project entitled "**Motion Estimation from Sensor Images**" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Anupa Shah

Date:23-5-15

Endorsed by

Manish Chturvedi

Acknowledgements

I would like to thank Mr.Ashish Mishra Head of PCSVD for giving chance to pursue my dissertaion under his guidance.A special thank you to **Mr.Amit Dave**,whose constant gudiance and willingness to share the vast knowledge help me in enhansing my knowledge completeing the tasks to perfection. Without his effort and full support,this project may not have succeeded and special thanks to ISRO(SAC) for providing me data of Chandrayan-1 MIP dataset.I would also like to thank Mr.Jitendra Shrma and Mr. Hiren Rambhia for helping me in completion of my project.

It gives me immense pleasure in expressing thanks and profound gratitude to **Prof. Manish Chturvedi**, Associate Professor, Computer Science Department, Institute of Technology, Nirma University, Ahmedabad for his valuable guidance and continual encouragement throughout this work.

It gives me an immense pleasure to thank **Dr. Sanjay Garg**, Hon'ble Head of Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his kind support and providing basic infrastructure and healthy research environment.

A special thank you is expressed wholeheartedly to **Dr K Kotecha**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout course of this work.

I would also thank the Institution, all faculty members of Computer Engineering Department, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

Anupa Shah
13MCEC18

Abstract

Recently, with the advances in computer vision algorithm, an alternative mechanism to navigate the lander to its precise location on interplanetary mission to moon and mars has been studied. Keeping up with the current trend, ISRO has planned to use vision based navigation in its next mission to moon.

This report presents horizontal velocity, angular velocity and altitude changes for accurate landing of lander based on the extraction and tracking of Scale Invariant Feature Transform features through image sequences. Implementation of SIFT operator is analyzed and adapted to be used in a feature-based motion estimation algorithm. Nearest Neighbor approach is implemented to match SIFT features and RANSAC is applied to remove matched outlier features. After that Transformation matrix is used to find rotation and scale information. The algorithm was tested on images acquired by Moon Impact Probe onboard ISRO's Chandrayan-1. For its robustness and accuracy, the results obtained were found to be comparable with simulated dataset.

Abbreviations

SIFT	Scale Invariant Feature Transform
DOG	Difference of Gaussian
LOG	Log of Gaussian
RANSAC	Random Sample Consensus
SSD	Sum of Squared Difference
MIP	Moon Impact Probe
TMC	Terrain Mapping Camera
DIMES	Descent Image Motion Estimation System
KLT	Kanade-Lucas-Tomasi
SURF	Speed Up Robust Features

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

In the domain of Computer Vision and Virtual reality,Motion Estimation is one of focal research topics.The Objective of motion estimation algorithm is to reliably and correctly model motion in the scene. This information is significant for human motion analysis,video Surveillance,robotics,vehicle motion analysis,automation of vehicles etc. The presented report provides brief comprehensive of literature survey of motion detection methods.In this research work , the vision based motion estimation system for landing is proposed.The algorithm uses multiple images data taken from camera during the decent phase.Here,Motion Estimation methods such as Block Match et al[1],Optical Flow et al[2][3],phase correlation et al[4],SIFT et al[5] methods are studied and applied to MIP dataset.

Extensive researches have already been done to explore many different approaches to solve those problems for Motion Estimation problems.Block Match is used for linear motion and it works fast on small motion.Optical Flow gives motion vectors in X and Y direction where object moves.phase correlation is used to estimate the relative translative offset between two similar images.But for dataset of different scale and rotation variant,Scale invariant feature transformation is necessary.

Here we took images of Chandrayan-1's MIP (Moon Impact Probe) sensor dataset. MIP was a lunar probe, discovered the presence of water on the Moon. It amassed close range images of the surface of the Moon during descent and before impact.To check algorithm correctness and reliabilty,we have also used simulated images that are generated in Lab.

1.2 Problem Definition

To develop an algorithm that take in images and other sensor data taken during decent phase of landing and output a velocity.The algorithm should be invariant to illumination, scale and rotation properties.The algorithm should be robust and computationally efficient. To achieve high level of robustness, the algorithm should track multiple features and apply check on features relation across images.

1.3 Purpose of research and Motivation

The purpose of this research to develop an algorithm for motion parameter estimation using vision based methods. The motivation behind this research is to develop an algorithm for low-cost and low-power system consisting of single camera. The vision based approach greatly reduces the hardware complexity and increases the overall reliability compared to other inertial based measurements.

1.4 Scope and Objective

This adaptive algorithm can be useful in other landing,Automation of vehicle and also useful in object detection and tracking algorithms.Main objective of our research is to develop robust algorithm for feature tracking,matching.So it will reduce error of landing as well as accurate finding motion parameters for satellite and spacecraft.

1.5 Challenges and Requirements

The images acquired during descent phase vary greatly in scale and orientation. The large changes between two consecutive images pose a great challenge in tracking features. This requires the algorithm to be invariant to the scale and affine transformation. The accurate motion estimation requires the algorithm to handle illumination variation, motion blur and sensor inaccuracies.

Chapter 2

Literature Review

[1]

Title of Research Paper	Distinctive Image Features from Scale-Invariant Key points
Name and Affiliation of Authors	David G. Lowe
Publication Year	January 5, 2004
Publish By	International Journal of Computer Vision,

SUMMARY

In this paper David et al [5] has derived algorithm of SIFT.In that he has shown step by step computation of SIFT. He has also shown which value is chosen for threshold and other parameters by experiment values and by graph.He has described how to find highly distinct features which are invariant to rotation and scale.He has also descibed how to match features of image using fast nearest neighbor approach.He has shown that Scale-space extrema detection, Key point localization, Orientation assignment, Key point descriptor are the key steps for SIFT(Scale Invariant Feature Transform).

[2]

Title of Research Paper	Design Through Operation of an Image-Based Velocity Estimation System for Mars Landing
Name and Affiliation of Authors	Andrew Johnson, Reg Willson, Yang Cheng, Jay Goguen, Chris Leger And Larry Matthies
Publication Year	January, 2007
Publish By	Springer Science

SUMMARY

In this paper, they have shown Descent Image Motion Estimation System (DIMES) et al [6],which is used for horizontal velocity estimation. They have used altitude and attitude measurements to rectify images to a level ground that is used to work with scale and orientation.Feature selection and tracking is used to compute horizontal motion.DIMES has combined multiple data from different sensors to create low cost and robust horizontal motion estimation.

[3]

Title of Research Paper	Autonomous Craters Detection from Planetary Image
Name and Affiliation of Authors	Ding Meng,Cao Yun-feng and Wu Qing-xian
Publication Year	2008
Publish By	IEEE

SUMMARY

In this paper,They have described vision based navigation system for autonomous landing system et al [7].optical landmark navigation algorithm are built on the craters detection and tracking.This paper is focused on crater detection, Ellipse Fitting and crater tracking algorithm.In this paper they have found robust approach for crater detection using candidate area selection.Then they applied region growing approach, edge detection, ellipse fitting.For candidate area selection they applied KLT(Kanade-Lucas-Tomasi) feature tracking algorithm.

[4]

Title of Research Paper	A Comparison of various Edge Detection Techniques used in Image Processing
Name and Affiliation of Authors	G.T.Shrivakshan and Dr.c.Chandrasekar
Publication Year	2012
Publish By	IJCSI

SUMMARY

In this paper they have compared various edge detection technique like Sobel, Robert, Prewitt ,Laplacian based edge detector ,Canny edge detector using Shark Fish image dataset.At end they have proved that Canny edge detector is best for that dataset. They also showed drawback of gradient based detection that is sensitive to noise.

[5]

Title of Research Paper	Determining Optical Flow
Name and Affiliation of Authors	Berthold K.P. Horn and Brian G. Rhunck
Publication Year	1981
Publish By	Artificial Intelligence

SUMMARY

This is a basic paper for optical flow motion detection.In this paper ,Horn and Schunk et al [8] have shown how to solve one optical flow equation with two unknowns.That is under estimation problem.In that paper they have assumed that brightness pattern varies smoothly in the image.

[6]

Title of Research Paper	Block Matching Algorithms For Motion Estimation
Name and Affiliation of Authors	Aroh Barjatya
Publication Year	2004
Publish By	IEEE

SUMMARY

In this paper ,he has compared various block matching algorithm like Exhaustive Search, Three step Search, Simple and Efficient Search ,Four step search.

[7]

Title of Research Paper	A COMBINED CORNER AND EDGE DETECTOR
Name and Affiliation of Authors Mike Stephens	Chris Harris
Publication Year	1988
Publish By	Alvey vision conference. Vol. 15.

SUMMARY

In this paper et al[9],He has solved one of the problem of Computer vision and object tracking, to detect object using corner detection.He described corner as the intersection of two edges and as a point for which there are two dominant and different edge directions in a local neighborhood of the point.This can be used in motion analysis of a monocular image sequence from a mobile camera.For that he has used KLT feature Tracker to find discontinuty in image.

Chapter 3

System Model

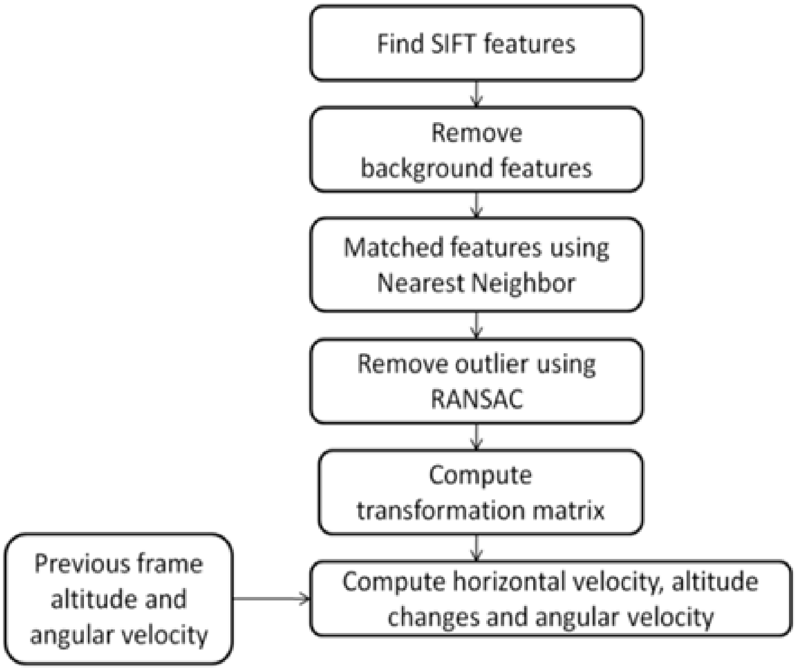

Figure 3.1: System Flow Diagram

In this research work,Chandrayaan-1 MIP dataset is used.Features are identified and tracked in two consecutive images.Background removal is done from identified features from both images and then apply k-nearest neighbor approach is used to match SIFT features. After matching phase,RANSAC algorithm is applied to remove outliers.The matched features are fed to the transformation block to find translation, scale and rotation information. Previous frame is taken as a reference frame and transformation information is used to find altitude, angular velocity and horizontal velocity changes.

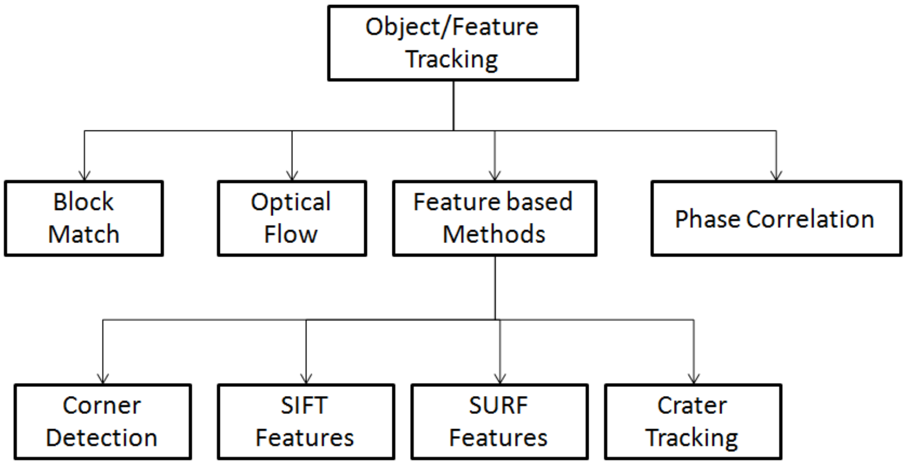

Figure 3.2: Object/Feature tracking Algorithms

For motion Estimation, Block match, Optical Flow and Feature based method are the basic methods.Block Match is basically used to identify linear motion.For Object detection and tracking, Optical Flow is generally used.We have checked many algorithm for our dataset but according to dataset SIFT is best for this dataset. We merge other algorithms and try to find out adaptive algorithm for altitude, angular velocity and horizontal velocity estimation.

Chapter 4

Motion Estimation Methods

Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence. The motion vectors may relate to the whole image (global motion estimation) or specific parts, such as rectangular blocks, arbitrary shaped patches or even per pixel. The methods for finding motion vectors can be categorized into pixel based methods (”direct”) and feature based methods (”indirect”).

Direct Methods:-
[1]Block Match
[2]Optical Flow
[3]phase Correlation

Indirect Methods:-
Indirect method uses features such as corner detection,edge detection,SIFT to estimate motion. This features are matched using Nearest Neighborhood procedure or data structures.Data structure includes k-d trees, a variation of binary trees that recursively divide the data space into smaller hyper-boxes to speed-up computation. Then statistical function such as RANSAC is used to remove matches that do not correspond to the actual motion.

4.1 Block Match

The purpose of a block matching algorithm is to find a matching block from a frame i in some other frame j, which may appear before or after i.In this,current block is compared with each block in search window and the best match is obtained (based on one of the comparison criterion).There are many two match one block to other blocks in second image.Matching one micro block to other is based on a cost function such as Mean Difference or Mean Squared Error.

Many algorithms are developed but some of most basic have been discussed below.

- Exhaustive Search
- Three step search
- Two Dimensional Logarithmic Search
- Binary Search
- Four Step Search
- Orthogonal Search Algorithm

In Exhaustive Search method, whole image is divided into blocks and every block of image with all blocks of another image is matched.Good match can be by Cross Correlation Function, Mean Absolute Difference, Mean Squared Difference etc.Here Searching window size is Full image.It is time consuming method.In this approach, computation time can be decreased by reducing searching window size.

In Three Step Search, eight blocks at a distance of step size from the centre (around the centre block) are picked for comparison. Then step size is reduced by two and the same procedure is done till Three step.Here we can also compare method on the basis of Cross Correlation Function, Mean Absolute Difference, Mean Squared Difference.TSS algorithm is illustrated graphically in figure 4.1.

As we described in Figure 4.1 for Two Dimensional Logarithmic Search method, some step size is taken and four blocks at a distance from this on the X and Y axes are matched.

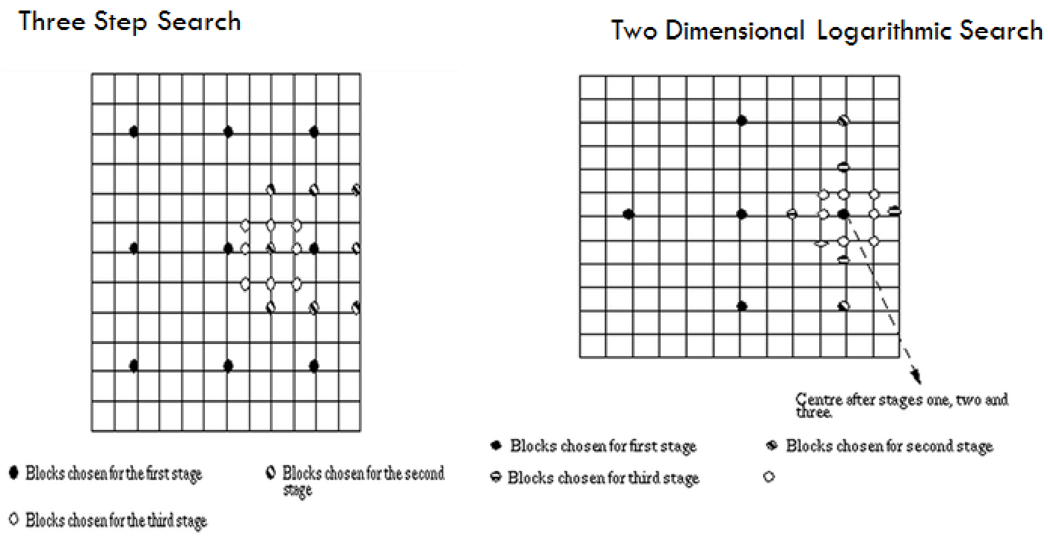

Figure 4.1: Block match algorithms

After Finding best match,step size is reduced by half.And that best point is taken as center point and will continue above process till step size becomes 1.

In Four Step Search method,step size is 2. Then nine points around the search window centre is taken and found the point with the smallest distortion. Then that point is taken to be the centre of the searching area .Then step size is reduced to 1 and all nine points around the centre of the search are examined for next process.If the previous minimum point is located at the corner of the previous search area, all five points are picked as good match.If the minimum point is at horizontal or vertical position then three points are taken as good match.

In Binary Search method,points are devided into search window and nine blocks are evaluated and examined based on MAD.Smallest MAD point is taken as center point and reevaluated this process until step size will become smaller than 1.

Block Matching algorithms work best in Linear Motion.But when dataset is rotation and scale varationt.Then it is not work properly.

4.2 Optical Flow

Optical flow or optic flow is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (an eye or a camera) and the scene.Optical flow mostly useful in object tracking, motion detection, lumination.The optical flow methods try to calculate the motion between two image frames which are taken at times t and t+Delta t at every pixel position.

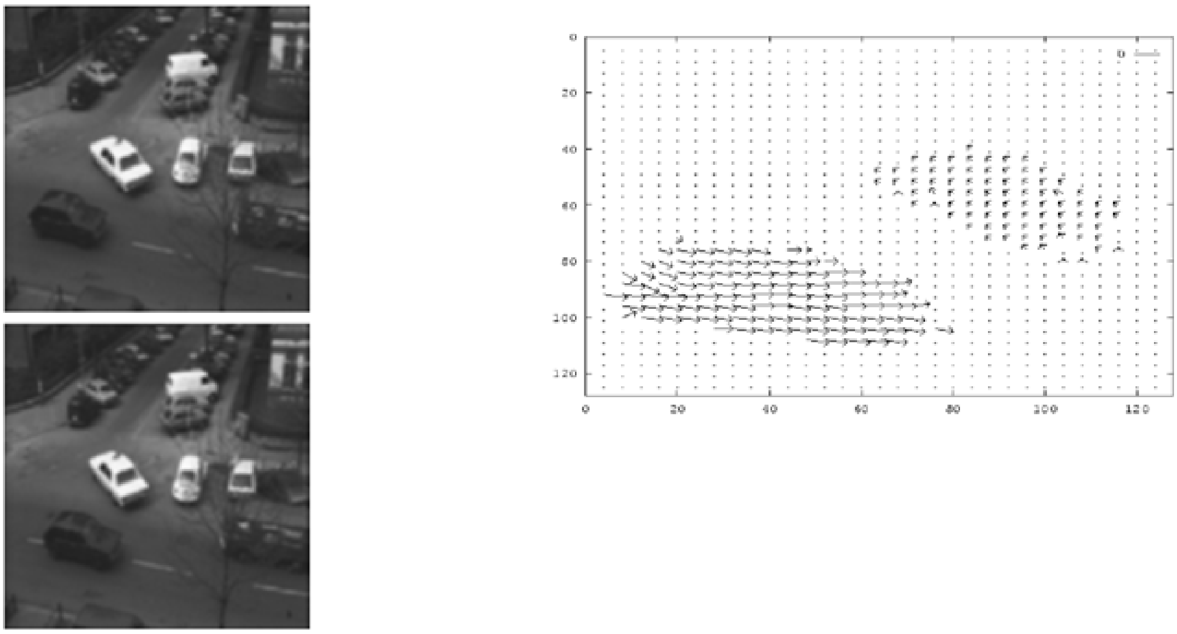

Figure 4.2: Optical Flow Results

Here we assume intensity is constant means images are taken in near by time.

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \tag{4.1}$$

If we expand right hand side part and equate with left hand side part than

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\sigma I}{\sigma y}\Delta x + \frac{\sigma I}{\sigma y}\Delta y + \frac{\sigma I}{\sigma t}\Delta t \tag{4.2}$$

$$\frac{\sigma I}{\sigma y}V_x + \frac{\sigma I}{\sigma y}V_y + \frac{\sigma I}{\sigma t}V_t = 0 \tag{4.3}$$

now it is become

$$I_x V_x + I_y V_y = -I_t \tag{4.4}$$

Now we can solve this optical flow methods using following method

1]Lukas-Kanade Method

2]Horn -Shunk Method

In lukas kanade method ,they apply same equation for nine neighboring pixel so now there are nine equation now it is not under estimate problem but it will become over estimate problem. So we can take any two equation and solve velocity vectors.

$$I_x(q1)V_x + I_y(q1)V_y = -I_t(q1) \tag{4.5}$$

$$I_x(q2)V_x + I_y(q2)V_y = -I_t(q2) \tag{4.6}$$

$$I_x(q9)V_x + I_y(q9)V_y = -I_t(q9) \tag{4.7}$$

Horn -Shunk et al [4] has taken this problem as a line problem and solve using optimation method

$$V_x = -\frac{f_x}{f_y}u + -\frac{f_t}{f_y} \tag{4.8}$$

But Optical flow gives motion vectors in x and y direction where object moves.And it is also not applicable in the case of rotation and scale variant images.

4.3 Phase correlation

Phase correlation provides straight-arrow estimation of rigid translational motion between two images, which is based on the well-known Fourier shift property that a shift in the spatial domain of two images results in a linear phase difference in the frequency domain of the Fourier Transforms (FT).

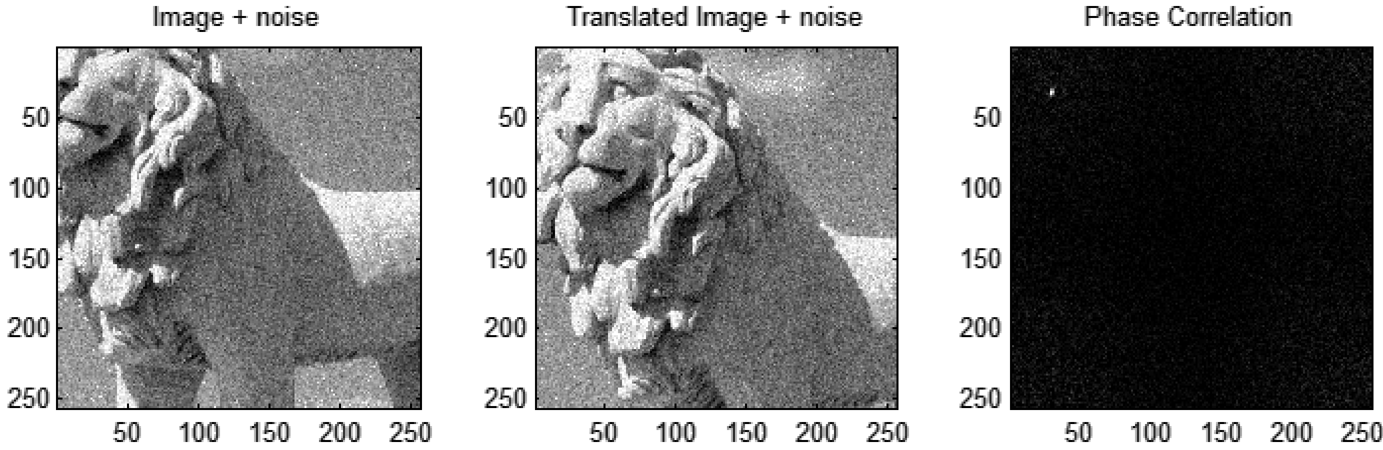

Figure 4.3: Phase correlation

In this method ,images ga and gb are converted into Fourier transform. And then cross power spectrum is taken by the complex conjugate of the second result, multiplying the Fourier transforms with element wise, and normalizing this product elementwise.

$$Ga = \mathcal{F}(ga) \quad and \quad Gb = \mathcal{F}(gb) \tag{4.9}$$

$$R = \frac{Ga \circ Gb^*}{|Ga \circ Gb^*|} \tag{4.10}$$

Then normalized cross-correlation is generated by applying the inverse Fourier transform.peak in r shows translation of second image.

$$r = \mathcal{F}^{-1}(R) \tag{4.11}$$

$$(\triangle x, \triangle y) = argmax_{x,y}(r) \tag{4.12}$$

4.4 Object Detection and Tracking

Object Detection is technology related to Machine vision and image processing that deals with detecting instances of semantic objects of a certain class such as humans, buildings, craters, or cars in images and videos. We can detects objects using various modules such as Feature-based object detection, SVM classification with histograms of oriented gradients (HOG) features, and Image segmentation and blob analysis etc.Here

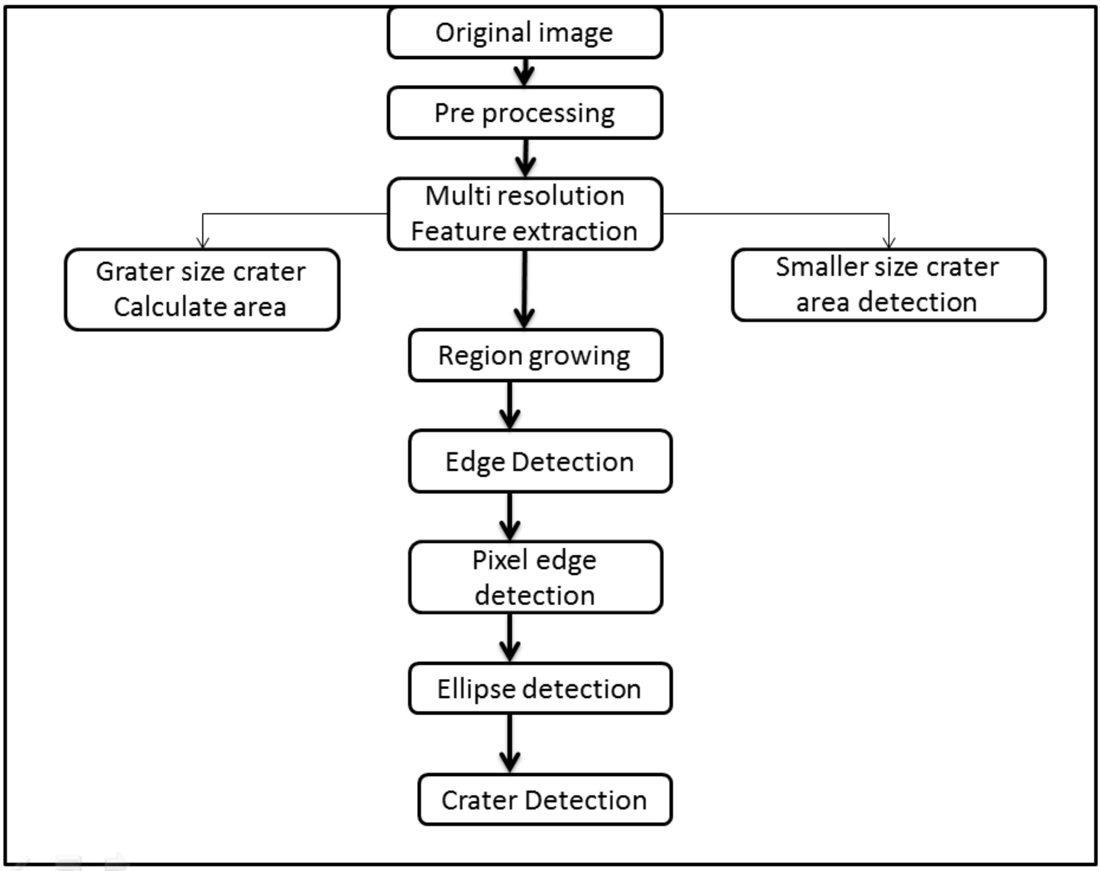

Figure 4.4: Crater Detection

we have used MIP Dataset. For this dataset craters are the objects. The crater detection algorithm is broken down into six steps: Edge Detection, Crater anchor point detection, edge grouping, Ellipse fitting, Ellipse refinement and Crater confidence evaluation. In the First stage, Different size of window is selected and KLT algorithm is applied to Intensity variations are strong in the whole region of crater. This procedure is done till smallest size of Window to get all size craters.After finding feature patch, edges are extracted and delete the sideline of crater which is a boundary between the shaded part and light part .After that ellipse detection is applied to fit on crater. This crater edge map is stored in database and this will use to find match in next image . But Crater detection and tracking is not useful because when scale and rotation change than that craters may be not in next image and it may be not good point to land in crater. so It is not good technique for scale and rotation variant satellite images dataset.

4.5 Feature Detection and Tracking

In computer vision and pattern recognition, features are non redundant, informative and useful in human Interpretations. This features can be used in motion estimation algorithms.It is one of the part of object detection. Feature extraction is used to convert large set of dataset into a reduced set of features. This features contains relevant information from input dataset so desired task can be performed on reduced set instead of large set. For homologous images, simple corner detection and edge detection are used for feature matching. A corner can be described as the intersection of two edges. A corner can also be described as a point for which there are two dominant and different edge directions in a local neighborhood of the point. Generally, Harries algorithm is used for corner detection. That is described below.Corner shows variation in the gradient in the image.

$$E(x,y) = \sum_{x,y} W(x,y)[I(x+u,y+v) - I(x,y)]^2 \tag{4.13}$$

Here, w(x,y) is the window at position (x,y).I(x,y) is the intensity at (x,y).I(x+u,y+v) is the intensity at the position (x+u,y+v).So to find variation in intensity,we have to maximize the equation 4.13.

$$\sum_{x,y} [I(x+u,y+v) - I(x,y)]^2 \tag{4.14}$$

Using Taylor expansion

$$E(u,v) \approx \sum_{x,y} [I(x,y) + uI_x + vI_y - I(x,y)]^2 \tag{4.15}$$

$$E(u,v) \approx \sum_{x,y} u^2 I_x^2 + 2uvI_xI_y + v^2 I_y^2 \tag{4.16}$$

Which can be expressed in a matrix form as,

$$E(u,v) \approx [u,v] \left(\textstyle\sum_{x,y} w(x,y) \quad \begin{pmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{pmatrix} \right) \begin{pmatrix} u \\ v \end{pmatrix} \tag{4.17}$$

$$H = \sum_{x,y} w(x,y) \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \tag{4.18}$$

To determine if it can possibly contain a corner

$$R = det(H) - k(trace(H))^2 \tag{4.19}$$

a window with a score R greater than a certain value is considered a "corner".Then feature descriptor algorithms are used to describe features. In Feature descriptor generation square window centered at the key point into a 1D descriptor is used to describe features. That features are matched using nearest neighbor approach. Harris Corner Detection has rotation invariance, partial intensity affine invariance, but it is not scale invariant.

4.6 Scale Invariant Feature Tracking

The SIFT key point detector and descriptor have proven exceptionally successful in number of application in image stitching, object reorganization, visual mapping etc. SIFT (Scale Invariant Feature Transform) algorithm is developed by David Lowe et al [10]. He extended his work with more in-depth development and analysis of earlier work in et al [5] ,which is shown in Figure 4.5. Here, Each feature is highly distinctive so that single feature can be correctly matched with high probability against a large database of features and finding candidate matching features based on Euclidean distance of their feature vectors.

Goal of Scale Invariant Feature tracking

- Extracting distinct invariant feature
- Invariant to scale and rotation
- It is robustness to affine transformation and 3D view point

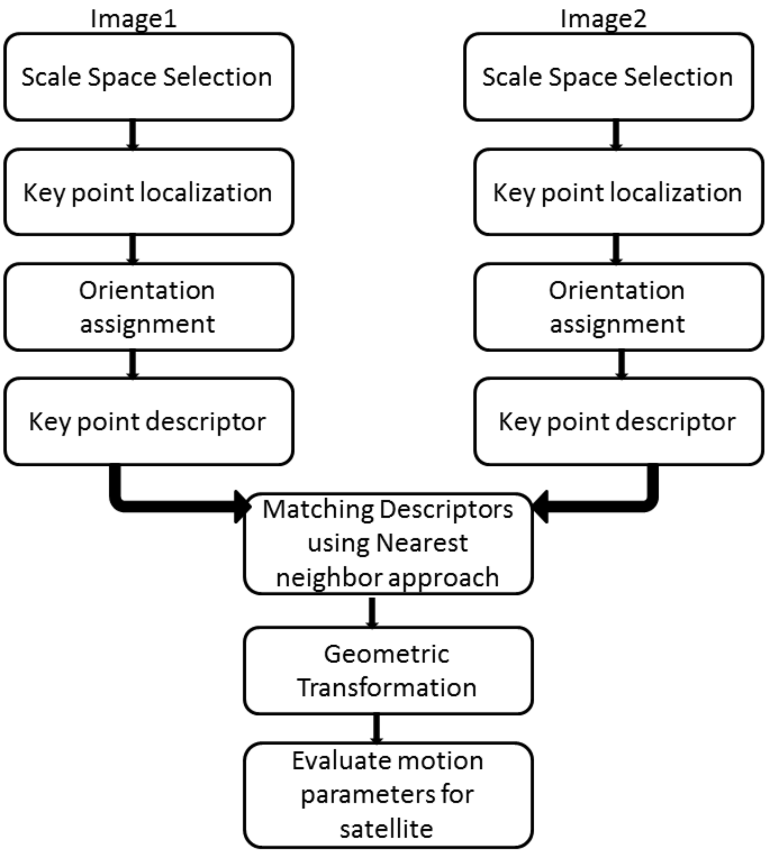

Figure 4.5: Work Flow of SIFT

In SIFT algorithm,Scale-space Selection,key point localization,Orientation assignment and key point descriptor are the major stages of computation in generating the set of image features.This Stages are explained in the next Section.

Chapter 5

System Implementation

5.1 Scale Invariant Feature Tracking(SIFT)

Scale Invariant Feature tracking algorithm is used to describe and detect local features in the images. It is key point extractor and detector.Harries et al [9] is also interest point detector but it is not scale invariant.SIFT (Scale Invariant Feature Transform) algorithm is developed by David Lowe et al [5].SIFT features are reasonably invariant to scaling, translation, and rotation, and partially invariant to illumination changes, affine distortion, addition of noise and even partial occlusion.

key stages for finding SIFT features are given below

- Scale space peak selection
- Key point localization
- Orientation Assignment
- Key point descriptor

1]Scale Space Peak Selection
Image I(x,y) is convoluted with a different scale Gaussian filter G(x,y,σ) to project the signal in to what is known as scale-space.Image is convolved with Gaussian to find discontinuity in the image, that is called edge.It is also used to suppressed the noise.when sigma increase noise is more suppressed.In SIFT algorithm, output is divided in Octave and stack to get more accurate result and to detect key point in each scale. Here, output

is divided into three octave and each octave has five stack.

Here for expriment purpose ,I have taken images form dataset of MIP. Second is scale by 1.5 of first image which is shown in Figure 5.1.

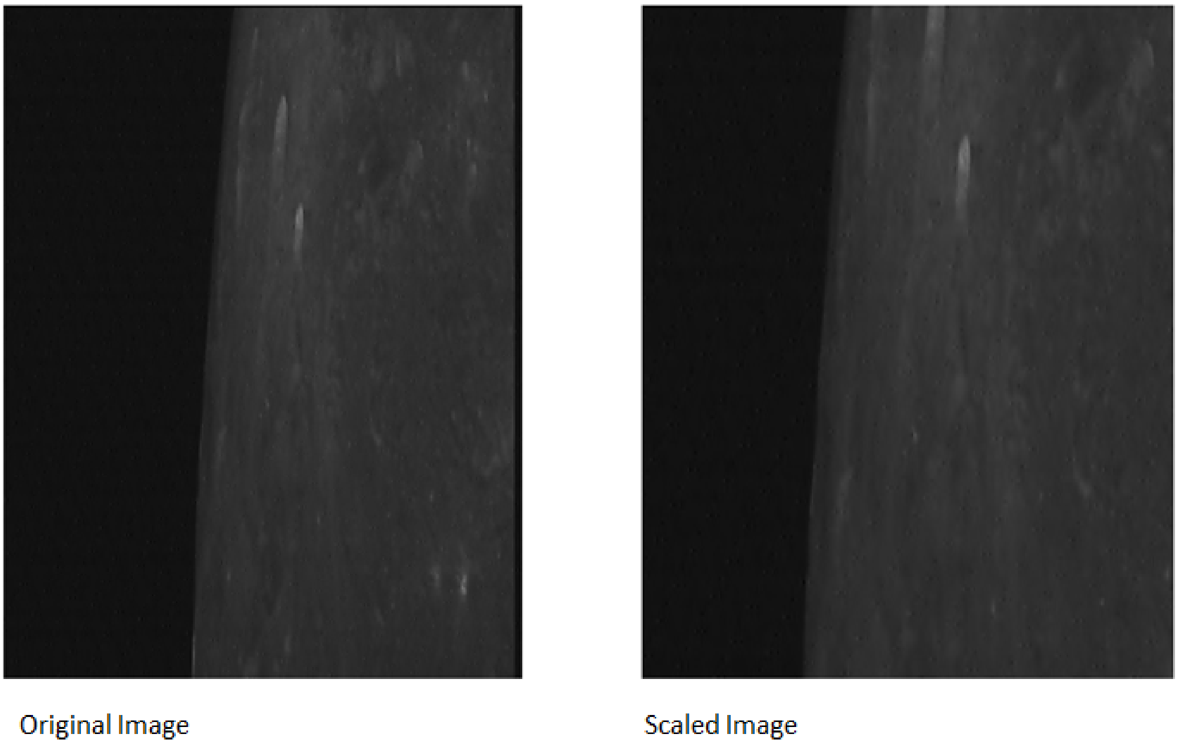

Figure 5.1: Input Images

In this stage ,Image I(x,y) is convoluted using 2D Gaussian function G(x,y,σ) to generate an image pyramid. This creates a 3D representation with x and y representing the image axis, and representing the scale axis.With increasing σ ,image becomes increasingly more blurred as the higher spatial frequencies are filtered out.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{5.1}$$

where is the convolution operation in x and y with Gaussian

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} * e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{5.2}$$

In above equation, sigma shows standard Deviation which is vary Octave to Octave and Stack to Stack. Here sigma is vary sqrt(2) for stack to stack and for Octave to Octave 2 power of(1/s).s is Octave no.

Figure 5.2 shows output after applying Gaussian function in first image for first octave and first stack.

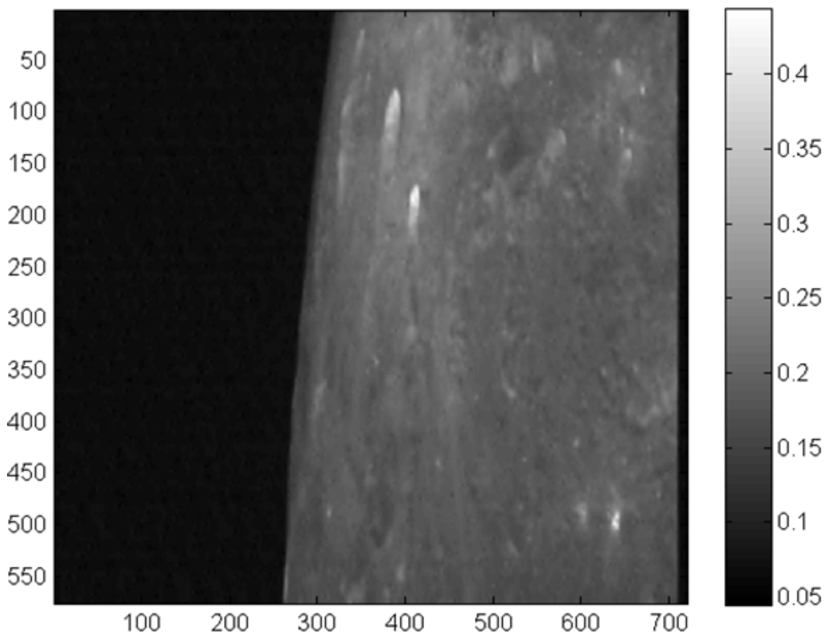

Figure 5.2: Result of Gaussian of first image for first octave and first stack

Figure 5.3 shows output after convolution of Image for each and every octave.

Figure 5.3: Gaussian Scale Space for each octave

Then Difference of Gaussian function D(x, y,σ) can be computed from the difference of two Gaussians L(x, y,σ), which is up and down in the stack (scale space).The DOG function D(x, y,σ) of the convoluted image L(x, y,σ) is the subtraction of two nearby scales in the Gaussian scale-space pyramid separated by constant factor k.

$$D(x, y, k\sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \tag{5.3}$$

$$D(x, y, k\sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{5.4}$$

After every octave, or doubling of sigma , the image is down sampled by a factor of two and the blurring iterations are re-started.Figure 5.4 shows output after applying Difference-of-Gaussian function in first image for first octave and first stack.

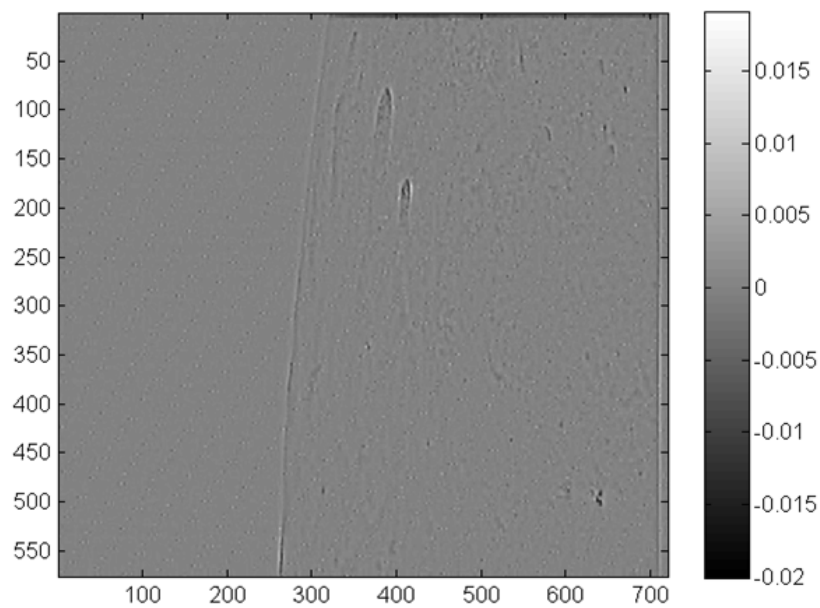

Figure 5.4: Result of Difference-Of-Gaussian for first octave and first stack

Figure 5.5 shows output after Difference-Of-Gaussian for each octave.

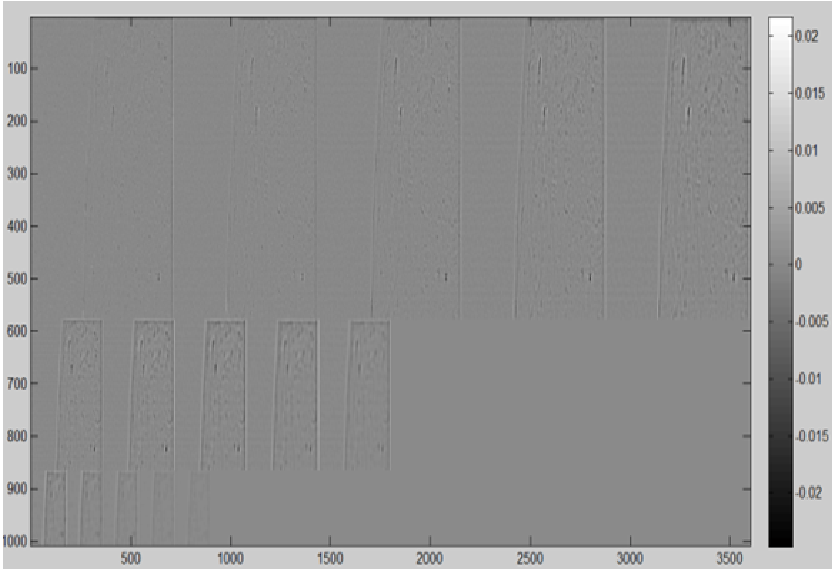

Figure 5.5: Difference-Of-Gaussian pyramid for each octave

Local minima and maxima of Difference-of-Gaussian(DoG) D are computed by comparing point with its eight neighbors in the current image and nine neighbors in the stack above and below (total of 26 neighbors), which is shown in Figure 5.6.

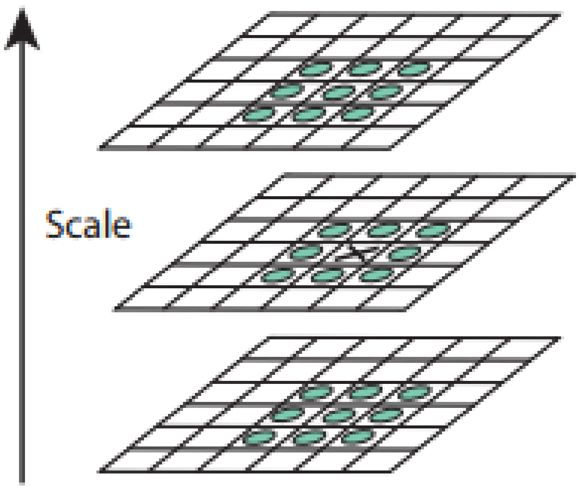

Figure 5.6: Minima-Maxima point selection

The feature point is called good if and only if it is larger or smaller than all of them.Figure 5.7 shows result of extrema points of first Octave and first stack.

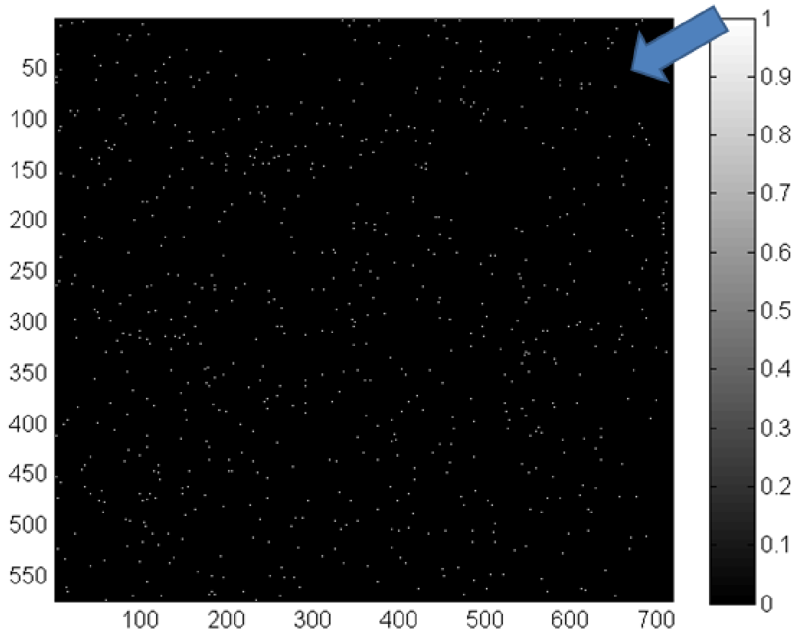

Figure 5.7: Result of Minima-Maxima point selection for first octave and first stack

[2]Key point localization

After extrema calculation,two criteria are used to detect unreliable keypoints. The first criteria is to find the value of absD at each candidate key point. If the value is below certain threshold, that means the structure has low contrast and that key point is removed.For this dataset we have taken 0.03 as a threshold. The second criteria evaluate the ratio of principal curvatures of each candidate key point to search for poorly defined peaks in the Difference-of- Gaussian function.

In figure 5.8, low contras points are removed using described method form first octave and first stack of minima maxima points.For stability ,edge points are also removed because it cause problem if edge is poorly determined and it may also due to noise. A poorly defined peak in Difference of Gaussian function will have a large principal curvature across the edge but a small one in the perpendicular direction.The principal curvature is found using Hessian matrix,

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{5.5}$$

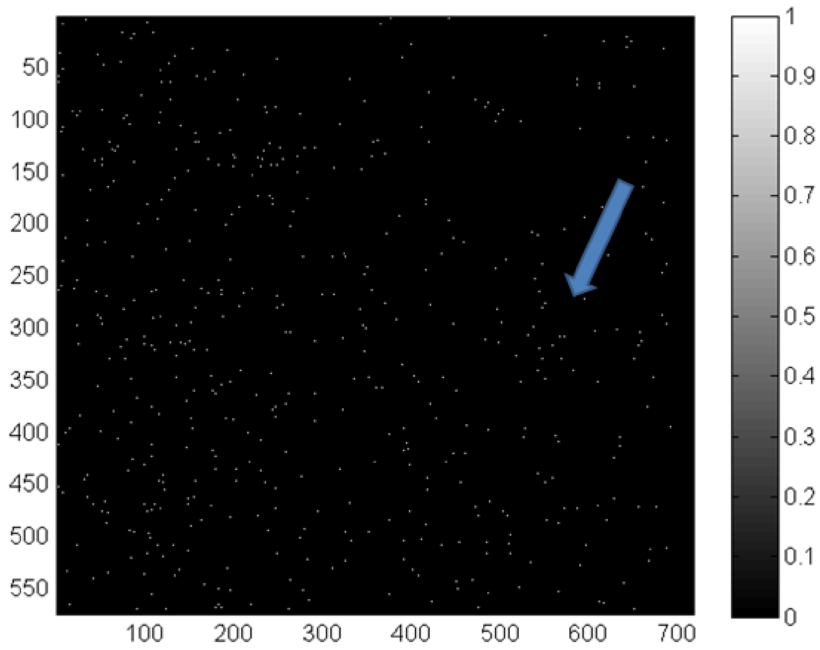

Figure 5.8: After removing low contras points for first octave and first stack

The Eigen values of H are proportional to the principal curvatures of D. Borrowing from the approach used by Harris and Stephens et al [9], we can avoid explicitly computing the eigen values, as we are only concerned with their ratio. Let be the eigen value with the largest magnitude and be the smaller one. Then, we can compute the sum of the eigen values from the trace of H and their product from the determinant:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \tag{5.6}$$

$$Det(H) = D_{xx}D_{yy} - D_{xy}^2 = \alpha\beta \tag{5.7}$$

which depends only on the ratio of the eigen values rather than their individual values. The quantity ratio is at a minimum when the two eigenvalues are equal and it increases with r. Therefore, to check that the ratio of principal curvatures is below some threshold r, we only need to check

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r+1)^2}{r} \tag{5.8}$$

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \tag{5.9}$$

From experiment it is found that value of r = 10, which eliminates key points that have a ratio between the principal curvatures greater than 10. Figure 5.9 shows output after removeing edge points using describe method form first octave and first stack after removing low contras points.

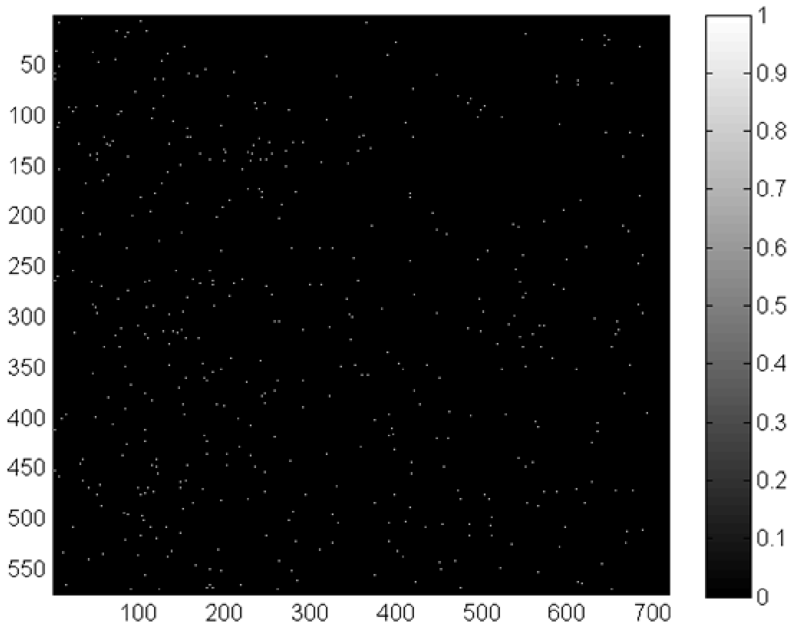

Figure 5.9: After removing edge points for first octave and first stack

Figures 5.10,5.11,5.12 ,show results after applying same procedure of finding minima-maxima points, removeing low contras points and edge points for last octave and last stack.

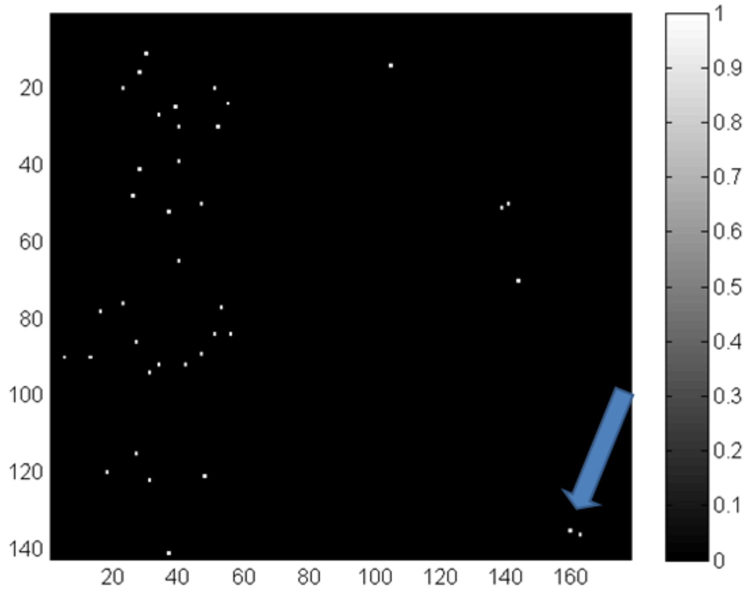

Figure 5.10: Result of Minima-Maxima point selection for last octave and last stack

Figure 5.11, shows that arrowed points that are remove in next image during edgepoints removal phase.

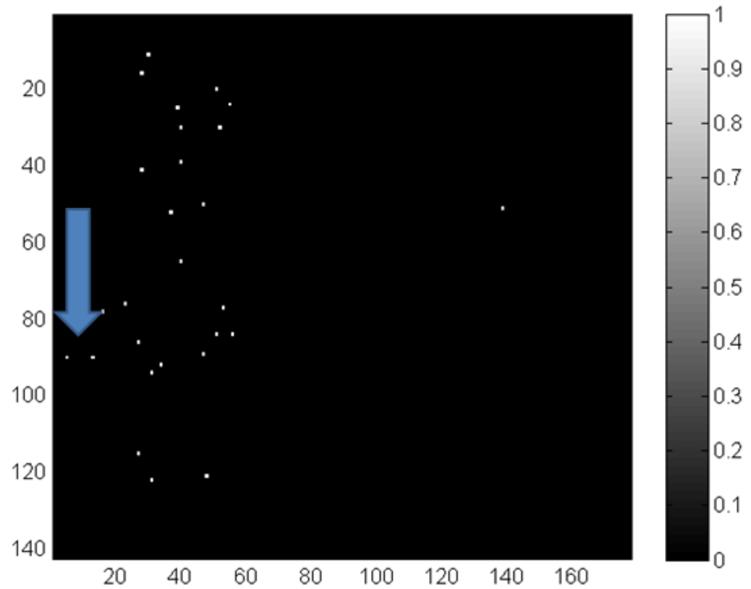

Figure 5.11: After removing low contras points for last octave and last stack

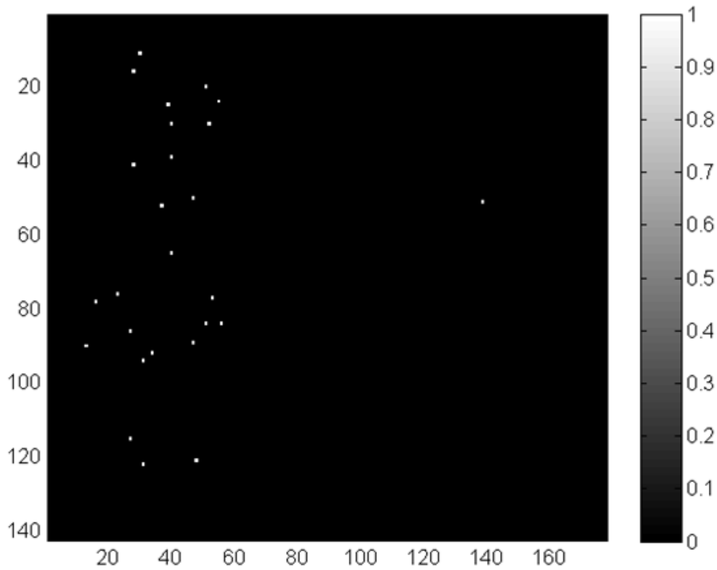

Figure 5.12: After removing edge points for last octave and last stack

[3]Orientation assignment

In this step ,orientation(magnitude and angle) is assigned to each key point using 16*16 window around key point.

$$p = L(x+1, y) - L(x-1, y) \tag{5.10}$$

$$q = L(x, y+1) - L(x, y-1) \tag{5.11}$$

$$magnitude = \sqrt{p^2 + q^2} \tag{5.12}$$

$$theta = \tan^{-1} \frac{p}{q} \tag{5.13}$$

In the orientation assignment, orientation histogram is generated by 8 bin covering 360 degree. Each point in 16*16 window is added to the histogram weighted by the gradient magnitude m(x,y). peak in that histogram has maximum value is taken as a magnitude and orientation of that key point. If any bin has 80% of highest peak than that peak is also used to create key point orientation.

[4]Key point Descriptor

After Doing the same process for second image,We find Descriptor of both the image.The descriptor of key point is generated by first computing the gradient magnitude and orientation at each image point of the 16xl6 key point neighborhood.Here each histogram has 8 bins; therefore each key point descriptor feature has 4x4x8 = 128 elements. The co-ordinates of the descriptor and the gradient orientations are rotated relative to the key point orientation to get orientation invariance and the descriptor is used to enhance invariance to changes in illumination. After finding the Descriptor we will match both images Descriptor using nearest neighbour approach.

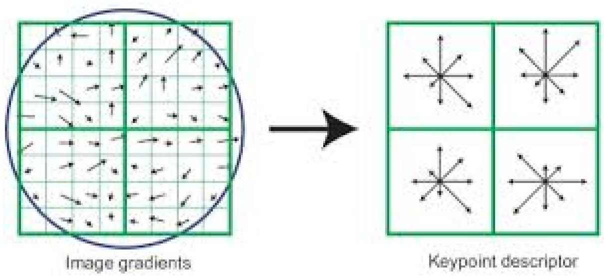

Figure 5.13: SIFT Descriptor Generation

Background feature points are not correct points.So background points are removed by applying mask.Here, mask is gotten by identifying average value of background points and after removing below that value that is shown in figure 5.14.

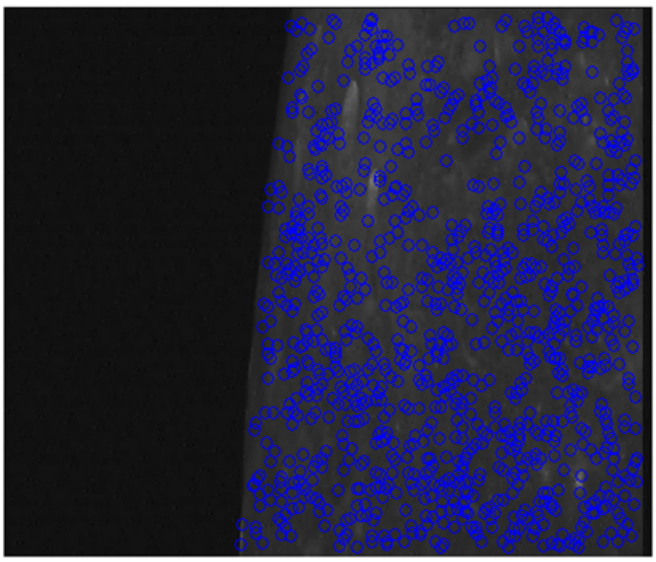

Figure 5.14: Masked Image

5.2 Random sample consensus

After computation of SIFT features, we use RANSAC(Random sample consensus) algorithm to remove outlier points.Random sample consensus (RANSAC) is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. Here I have described RANSAC algorithm. In the first step, a sample subset containing minimal data items is randomly selected from the input dataset. A fitting model and the corresponding model parameters are computed using only the elements of this sample subset. The cardinality of the sample subset is the smallest sufficient to determine the model parameters.

In the second step, the algorithm checks which elements of the entire dataset are consistent with the model instantiated by the estimated model parameters obtained from the first step. A data element will be considered as an outlier if it does not fit the fitting model instantiated by the set of estimated model parameters within some error threshold that defines the maximum deviation attributable to the effect of noise.

The set of inliers obtained for the fitting model is called consensus set. The RANSAC algorithm will iteratively repeat the above two steps until the obtained consensus set in certain iteration has enough inliers.

5.3 Transformation Matrix

To compute scale, rotation and translation information we have computed matrix. Here we required at least two correctly matched points to find exact scale, rotation and translation information. Let (u,v) represent the original image coordinates and (x,y) is consecutive image coordinates. Then translation information can be found by

$$x = u + t_x \quad y = v + t_y \tag{5.14}$$

So, matrix notation is

$$x = \begin{pmatrix} x \\ y \end{pmatrix}, u = \begin{pmatrix} u \\ v \end{pmatrix}, t = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \tag{5.15}$$

If there is scaling in image

$$x = s_x u, y = s_y v \tag{5.16}$$

So,matrix representation is

$$x = Su \quad where \quad S = \begin{pmatrix} S_x & 0 \\ 0 & S_y \end{pmatrix} \tag{5.17}$$

If there is rotation changes then

$$x = u\cos\theta - v\sin\theta, y = u\sin\theta + v\cos\theta \tag{5.18}$$

Its matrix notation

$$x = Ru \quad where \quad R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{5.19}$$

So from this geometry matrix ,scale and rotation information are found.Now we take first frame as a reference frame and from above scale and rotation information, altitude changes and angular velocity are measured.

Chapter 6

Results and Future work

6.1 Output

From Survey and experiment, we found difference between Block Match and Optical flow which is described below.

	Optical Flow	Block match
Motion Between Images is high		yes
Object detection and tracking	yes	
Linear motion		yes
Motion in direction of edge		yes
Local Information is used	yes	
Global Information is used		yes
Less time consume	yes	
Same object than also track	yes	

Figure 6.1: Differences between Blockmatch and Opticalflow

Without RANSAC,we got 61% accuracy Which is shown in Figure 6.2.

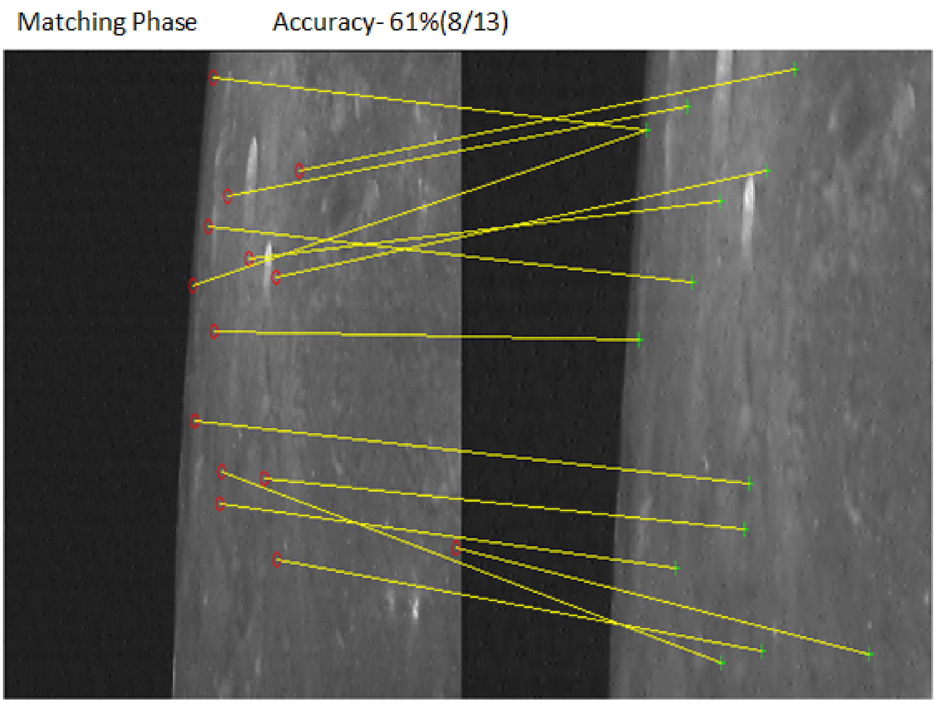

Figure 6.2: Result of Matched Features

After applying all the steps described in above section,we got 99% accuracy.Figures 6.3 to 6.12 shows output after Feature matching,RANSAC and geometric Transformation.

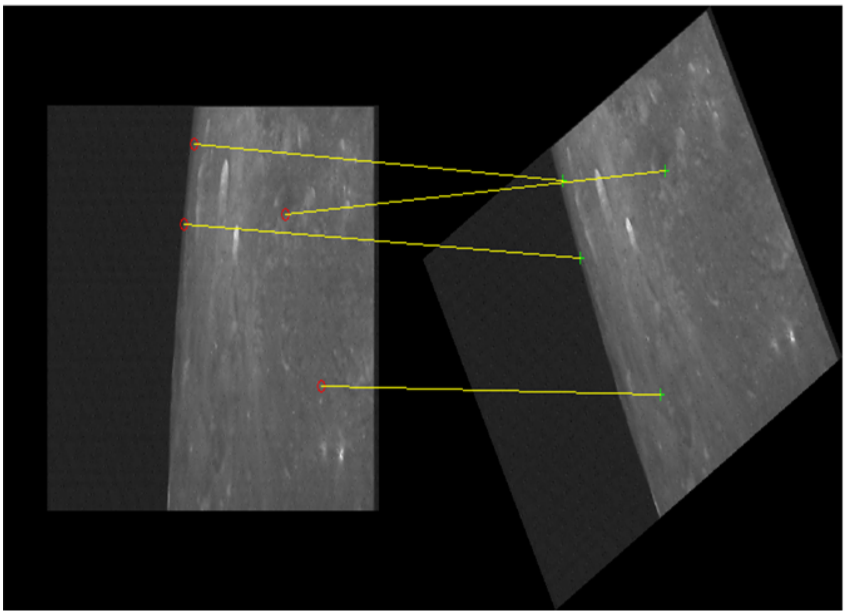

Figure 6.3: Angle=30

Here,scale recovered is 1.0047 and theta recovered is 29.8398

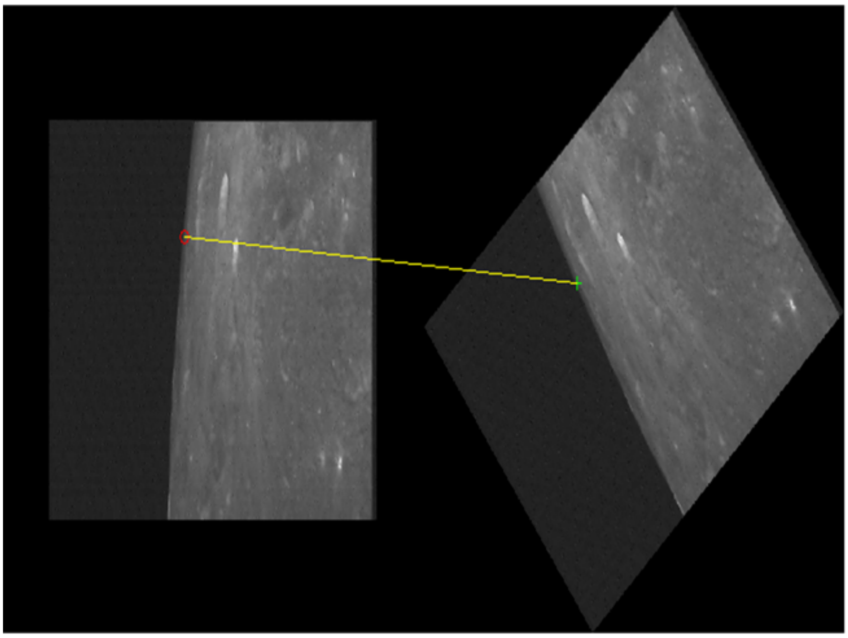

Figure 6.4: Angle=40

Here,scale recovered is 0.9999 and theta recovered is 45

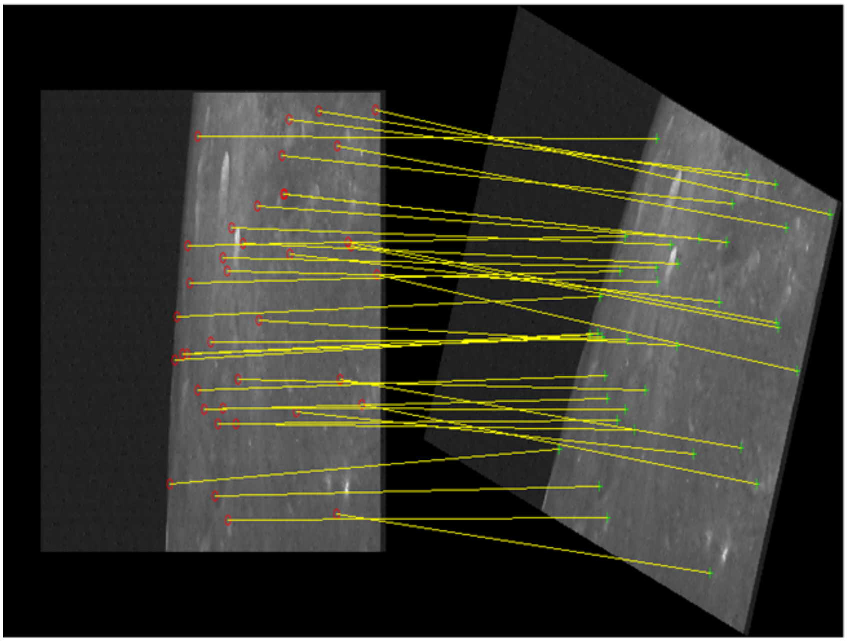

Figure 6.5: Angle=-20

Here,scale recovered is 0.9999 and theta recovered is -19.9787

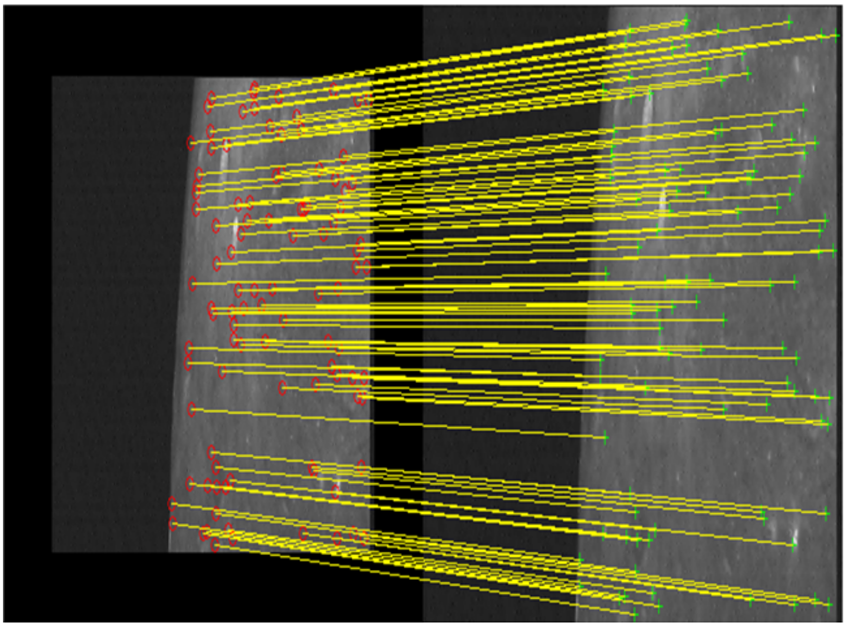

Figure 6.6: Scale=1.3

Here,scale recovered is 1.3003 and theta recovered is -0.0299

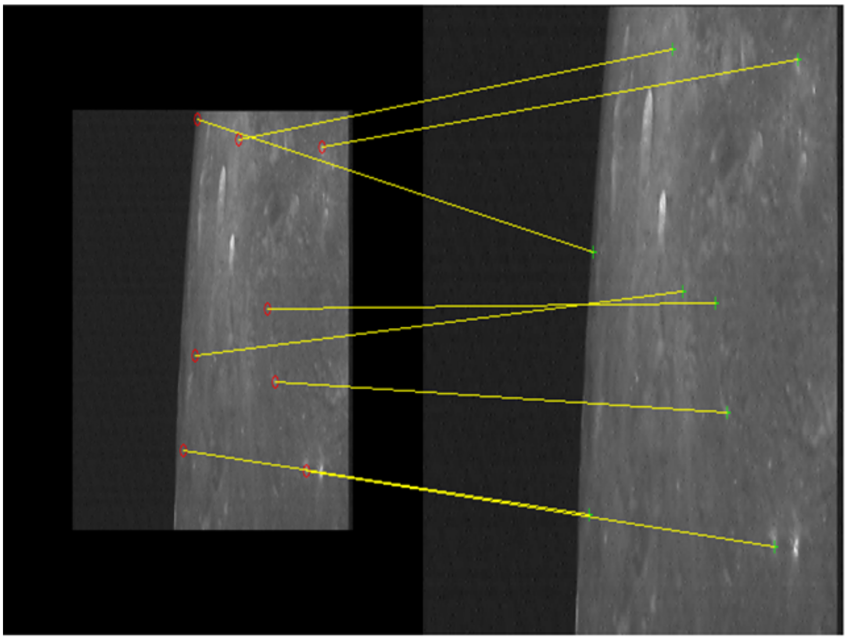

Figure 6.7: Scale=1.5

Here,scale recovered is 1.5006 and theta recovered is -0.0142

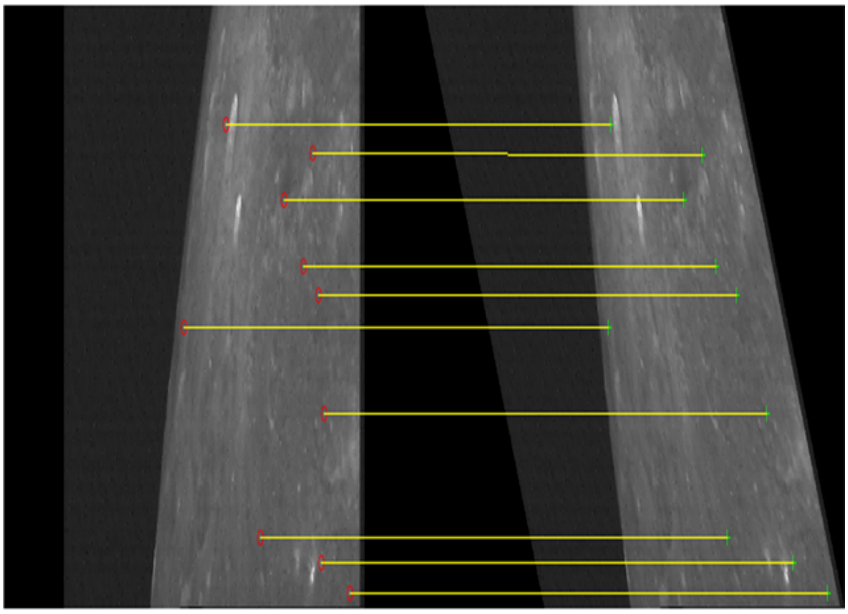

Figure 6.8: Shear

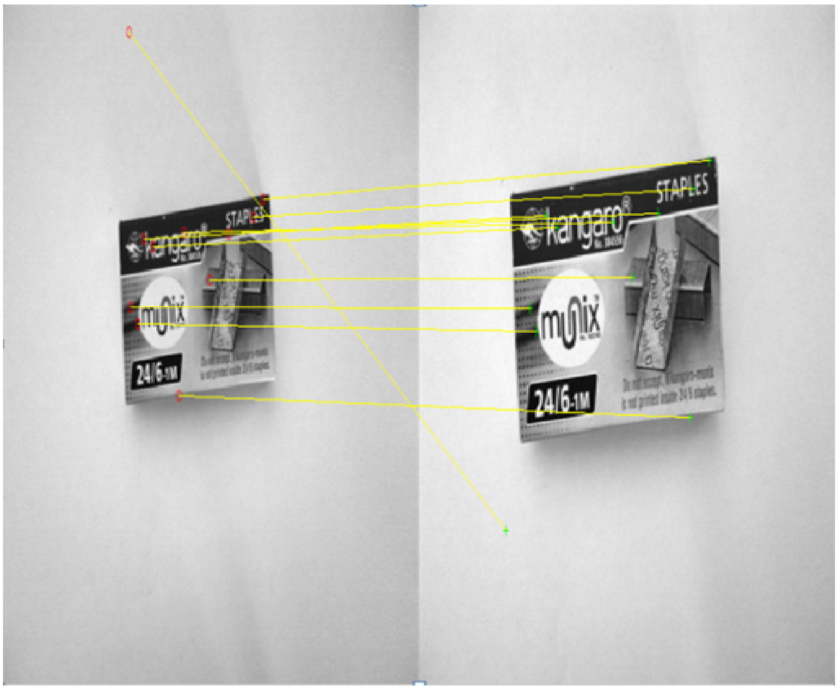

Figure 6.9: Scale=1.3

Here,scale recovered is 1.3556 and theta recovered is -0.4859

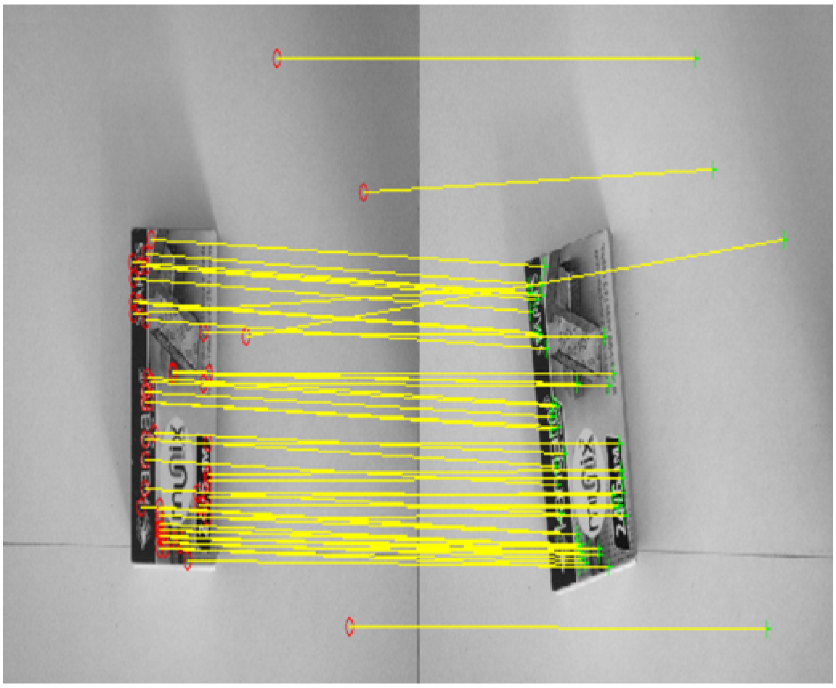

Figure 6.10: Angle=10

Here,scale recovered is 1.0012 and theta recovered is 10.9565

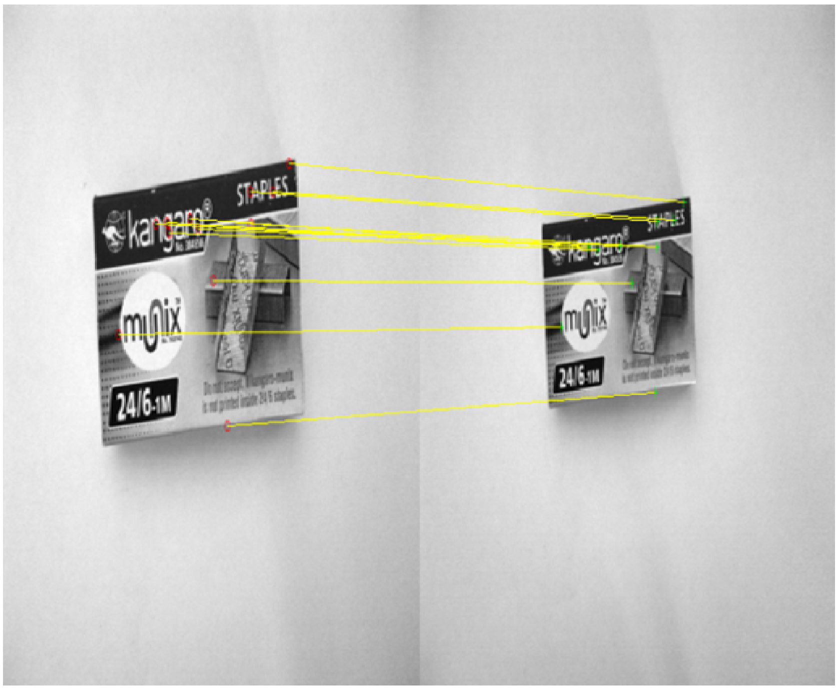

Figure 6.11: Scale=0.75

Here,scale recovered is 0.7508 and theta recovered is 1.2993

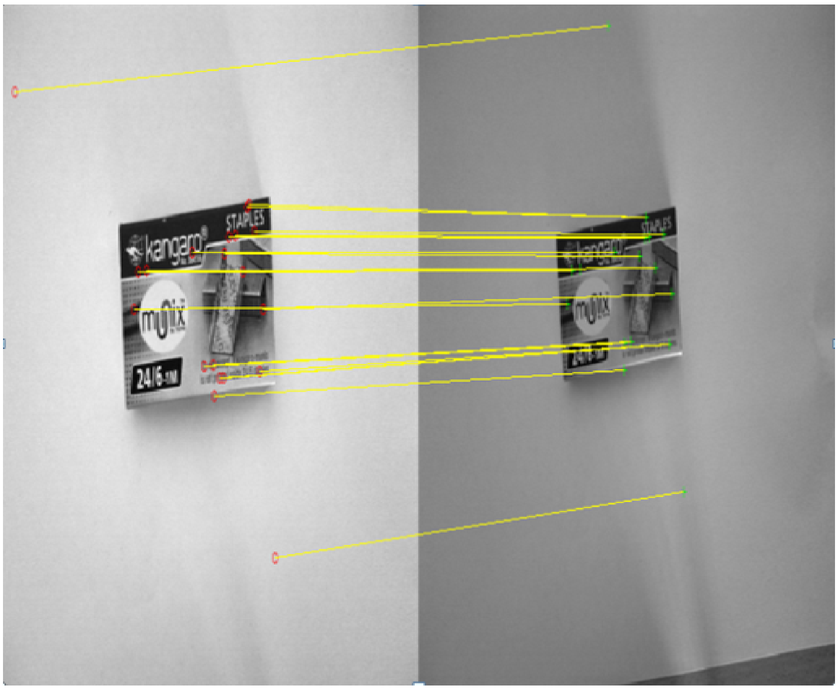

Figure 6.12: Scale=0.8

Here,scale recovered is 0.7968 and theta recovered is 1.5073

For algorithm verification ,known scaling and rotation are applied to two sets of images.First set contains the actual MIP images and second set of images are generated in lab.The given parameters are tabulated in Table 6.1.The result shows output after transformation matrix generation.

Angle	**Scale**	**Angle Obtained**	**Scale Obtained**
Dataset1			
-20	1	-19.9787	0.9999
30	1	29.8393	1.0047
40	1	45	0.9999
0	1.5	-0.0142	1.5006
Dataset2			
10	1	10.9565	1.0012
0	1.3333	-0.4859	1.3556
0	0.75	1.2993	0.7508
0	0.8	1.5073	0.7968

Table 6.1: Scale and Rotation Observation

Table 6.2 shows output of motion estimation of first six frames of MIP dataset. In Table 6.2, frame-0 is initial frame. Calculation of frame-1 is done with reference of frame-0 and follow the same procedure for all consecutive frames.

No	**Scale**	**Rotation**	**Angular Velocity (spinning)**	**Altitude**	**Altitude Changes(km)**	**Horizontal Distance travelled (km)**
0				97.8947		
1	1.0032	9.7512	81.26	98.2119	15.8602	2.385
2				97.8847		
3	1.0018	8.9124	74.27	98.0576	8.6452	1.965
4				97.8796		
5	0.9909	10.2262	85.22	96.9845	44.7563	0.61785

Table 6.2: Motion Estimation Results

Here time between frame-0 and frame-1 is 0.020 seconds and frame-1 and frame-2 is 1.7 seconds. one pixel equals to 0.015 km (= 10.85 km/720) on moon surface et al [11].Due to spinning and coninng dataset, Horizontal motion from first frame to fourth frame is 1.6732km/s.

6.2 Conclusion

The approach described above has been implemented in MATLAB.Results show a close match with the values calculated manually and given known values.It proves that the algorithm can be utilized on the set of images obtained from lander for object tracking and estimating motion related parameters.

This developed algorithm gives accurate output in range of -25 to +25 in angle and 0.75 to 1.4 in Scale. The future work is to improve the algorithm and we will use DSP based hardware to provide real time performance/feedback to the lander. However for implementation in hardware resource estimation needs to be critically carried out.

Chapter 7

References

[1]Barjatya, Aroh. "Block matching algorithms for motion estimation." IEEE Transactions Evolution Computation 8.3 (2004): 225-239.

[2]Aslani, Sepehr, and Homayoun Mahdavi-Nasab. "Optical Flow Based Moving Object Detection and Tracking for Traffic Surveillance." International Journal of Electrical, Electronic Science and Engineering VoJ 7.

[3]Aslani, Sepehr, and Homayoun Mahdavi-Nasab. "Optical Flow Based Moving Object Detection and Tracking for Traffic Surveillance." International Journal of Electrical, Electronic Science and Engineering VoJ 7.

[4]http://en.wikipedia.org/wiki/Phase correlation

[5] D.G. Lowe, "Distinctive image features from scale invariant keypoints," International Journal of Computer Vision, vol. 60, pp. 91-1 10,2004.

[6]Johnson, Andrew, et al. "Design through operation of an image-based velocity estimation system for Mars landing." International Journal of Computer Vision 74.3 (2007): 319-341.

[7]Meng D., Yun-Feng C., Qing-xian W (2008)., Autonomous craters detection from planetary image, Int. conf. On Innovative Computing Information and Control, Dalian.

[8]Horn, B.K.P., and Schunck, B.G., Determining Optical Flow, AI(17), No. 1-3, August 1981, pp. 185-203.

[9]C. Harris and M. Stephens. A Combined Corner and Edge Detector. In Fourth Alvey Vision Conference, 1988.

[10]Lowe, David G. "Object recognition from local scale-invariant features."Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Vol. 2.

Ieee, 1999.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[11]Srivastava, Vinay K., Tapas Kumar Nandy, and Anurag Verma. "Evaluation of Associated Motion Parameters of Descending Moon Impact Probe (MIP) of Chandrayan-1 from Time Sequential Images." Advances in Signal Processing 1.4 (2013): 63-67.