

# **Robust Feature based Automatic Image Registration for Image Fusion and Change Detection in Remote Sensing Data**

**THESIS**

Submitted in partial fulfilment of the  
requirements for the degree of  
**DOCTOR OF PHILOSOPHY**

by

**Indranil Misra**

Under the Supervision of

**Prof. Mukesh Kumar Rohil**

Department of Computer Science and Information Systems, BITS Pilani, Pilani Campus

**&**

Under the Co-Supervision of

**Dr. S. Manthira Moorthi**

Space Applications Centre, Indian Space Research Organisation (ISRO), Ahmedabad



**BITS Pilani**  
Pilani | Dubai | Goa | Hyderabad

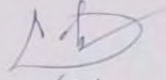
**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**

**2024**

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

## CERTIFICATE

This is to certify that the thesis entitled **Robust Feature based Automatic Image Registration for Image Fusion and Change Detection in Remote Sensing Data** and submitted by **Indranil Misra** ID No. **2021PHXF0301P** for award of Ph.D. of the Institute embodies original work done by him under my supervision.



(Mukesh Kumar Rohil)

Signature of the Supervisor

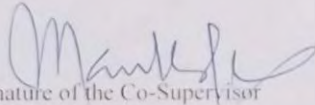
**Prof. Mukesh Kumar Rohil**

Professor, Department of Computer Science & Information Systems

Birla Institute of Technology and Science, Pilani,

Pilani Campus

Vidya Vihar, Pilani – 333031 (Rajasthan) INDIA



Signature of the Co-Supervisor

**Dr. S. Manthira Moorthi**

Scientist/Engineer 'G', Group Director,

Space Applications Centre, Indian Space Research

Organisation (ISRO), Jodhpur Tekra,

Ahmedabad – 380015 (Gujarat) INDIA

**डॉ. एस. मंथिरा मूर्ति / Dr. S. Manthira Moorthi**

समूह निदेशक / Group Director

ओडीपीसी/सिपा/सेक / ODPG/SIPA/SAC

अंतरिक्ष व्यवसाय केंद्र (इसरो)

Space Applications Centre (ISRO)

अहमदाबाद / Ahmedabad-380015

Date *Oct 08, 2024*

# Acknowledgements

First and foremost, I extend my heartfelt appreciation to my supervisor, Prof. Mukesh Kumar Rohil, Professor in the Department of Computer Science & Information Systems (CSIS) at Birla Institute of Technology and Science (BITS), Pilani, Pilani Campus, Rajasthan, India. His guidance, motivation, support, and vast knowledge have been invaluable to me. I am equally grateful to my co-supervisor, Dr. S. Manthira Moorthi, Scientist/Engineer ‘G’ and Group Director at the Space Applications Centre (SAC), Indian Space Research Organisation (ISRO), Ahmedabad, for his unwavering guidance, motivation, support, and extensive expertise.

My sincere thanks goes to Prof. V. Ramgopal Rao, Vice Chancellor, BITS Pilani; Prof. Sudhir Kumar Barai , Director, BITS, Pilani, Pilani Campus; Col Soumyabrata Chakraborty (Retd), Registrar, BITS, Pilani; Prof. Vamsi Krishna Venuganti, Dean, Academic–Graduate Studies & Research Division (AGSRD), BITS, Pilani; Prof. Srinivas Krishnaswamy, Former Dean, AGSRD, BITS, Pilani; Prof. Shamik Chakraborty, Associate Dean, AGSRD, BITS, Pilani, Pilani Campus, for providing me an opportunity to pursue and complete my PhD work from this esteemed institution.

I express my gratitude to Prof. Yashvardhan Sharma, Head of the Department (HoD), CSIS Department; Prof. Navneet Goyal, former HoD, CSIS Department; Prof. Sudeept Mohan, former HoD, CSIS Department; Prof. Shashank Gupta, Doctoral Research Committee (DRC), Convenor, and Prof. Amit Dua, former DRC Convenor in the CSIS Department, BITS, Pilani, Pilani Campus. Their guidance and support have been instrumental in meeting the diverse requirements of the Ph.D. program.

Additionally, I appreciate the input of the members of the Doctoral Advisory Committee: Dr. Karunesh Kumar Gupta, Professor in the Electrical & Electronics Department, BITS, Pilani, Pilani Campus; and Dr. Sundaresan Raman, Associate Professor in the CSIS Department, BITS, Pilani, Pilani Campus. Their valuable feedback has been greatly improved my research work.

I extend my gratitude to Mr. Debajyoti Dhar, Deputy Director, Signal and Image Processing Area (SIPA), SAC, ISRO, Ahmedabad, for providing resources, supervision, and necessary approvals for my research. I also acknowledge the encouragement and support from the Director, SAC, ISRO, Ahmedabad. My sincere thanks also go to my colleagues in the Optical Data Processing Division (ODPD) at SAC, ISRO, Ahmedabad, for their constructive feedback on my research outputs.

I am deeply indebted to my parents, Mr. Pradyot Kumar Misra and Mrs. Arati Misra, whose constant encouragement and assistance have been invaluable throughout my personal and academic journey. I am also grateful to my wife, Mrs. Shreya Misra, and my son, Mr. Rudranil Misra, for their unwavering love and moral support during the entirety of my research work.

**(Indranil Misra)**

# Abstract

Remote sensing image registration is the process of aligning two or more images of the same area to the same coordinate system, enabling accurate comparison, analysis, and integration of data. Remote sensing image registration is crucial for earth observation data analytics. In this aspect, georeferencing corrects the systematic geometric degradation in the image. However, achieving sub-pixel geometric accuracy across multi-temporal scenes is challenging. The research introduces a novel method, which employs hierarchical feature detection and motion smoothness constraint optimization to enhance geometric fidelity at the sub-pixel level. The proposed methodology utilizes Patch Affine Oriented Fast and Brief (PA-ORB) along with Mode Guided Tiled Scale Invariant Feature Transform (MT-SIFT) techniques in a coordinated multistage processing architecture. The experimental evaluations on Indian Resourcesat multispectral camera images demonstrate proposed image registration technique capability to handle significant geometrical errors, achieving a sub-pixel level RMSE on multi-temporal image scenes. The research work also discusses the challenges of geometric image registration, particularly for images with scanty features, such as those captured over deep ocean water or sparse island regions. Presenting a novel method for automatic multi-satellite island image registration, the approach enhances images using anisotropic coherence and Principal Component Analysis (PCA) to improve feature matching accuracy. Evaluated on Indian Resourcesat and Landsat images, the proposed approach demonstrates sub-pixel registration accuracy, with a 2.6% improvement in Correct Matching Ratio (CMR) compared to existing techniques, achieving an average RMSE of 0.45 pixel post-geometric correction.

Band-to-Band Registration (BBR) is essential for remote sensing multispectral sensors, aiming to align spectral wavelength channels accurately. This research presents an innovation in BBR technique utilizing Co-occurrence Scale Space (CSS) and Spatial Confined RANSAC (SC-RANSAC), along with a Segmented Affine Transformation (SAT) model to reduce distortion. The experimental evaluation on Nano-MX multispectral images showcases superior performance over existing techniques, particularly in scenarios dominated by cloud pixels, enhancing Normalized Difference Vegetation Index (NDVI) accuracy. The research also provides insights into direct planetary feature extraction and introduces a novel methodology for extracting features directly from Bayer Pattern raw planetary images, aiming to streamline object recognition and image co-registration tasks. By extracting gradient information using standard edge operators on Indian Mars Color Camera (MCC) Bayer Pattern raw images, the method achieves sub-pixel image co-registration accuracy below 0.5 pixel. Feature matching is enhanced using Gradient Intensity induced Scale Invariant Feature Transform (GI-SIFT) and Feature Similarity Score guided Random Sample Consensus (FSS-RANSAC) estimation technique. Image mosaicking is crucial for generating large area coverage maps from multiple adjacent overlapping satellite images. The research presents a novel mosaicking workflow utilizing Scale Invariant Feature Transform (SIFT) for strip geo-registration and Mode Biased Random Sample Consensus (MB-RANSAC) for outlier removal. Evaluations on

Resourcesat images demonstrate the system's ability to achieve sub-pixel registration accuracy and maintain spectral fidelity across different classes.

Remote sensing image fusion is the process of integrating information from multiple remote sensing images of the same scene, acquired from different sensors or at different times, to generate a composite image with enhanced spatial, spectral, and temporal characteristics. The research presents a balanced and robust image fusion method aimed at preserving spectral characteristics while merging high and low-resolution multispectral images. Utilizing Holistic Nested Edge Detection (HNED) and minnaert parameter in a Spectra Preserving Bayesian Probabilistic model, outperforms existing fusion methods both visually and quantitatively. Evaluation on various datasets confirms its effectiveness in maintaining spectral fidelity and producing accurate surface reflectance values. In addition, the research on Infrared-Visible remote sensing image fusion is explored and a novel methodology is proposed for fusing thermal infrared and multispectral visible images, enhancing spatial resolution and feature delineation. The method combines boundary-preserving information from thermal infrared images with enhanced visible images using CLAHE and inverse transformation. The experimental evaluations on Indian Nano Satellite (INS) and MODIS images demonstrate superior feature demarcation compared to individual images, outperforming existing fusion techniques visually and quantitatively. A comprehensive workflow for processing Venus' visible images is also presented in the research work, encompassing denoising, enhancement, registration with radar images, and fusion using a guided filter. The evaluation confirms the Venus fused image's effectiveness in delineating planetary morphological features, validated against radar nomenclature maps.

Remote sensing change detection involves identifying and analyzing alterations in land cover, land use, or environmental conditions over time using remotely sensed imagery. The automatic historical change detection using multi-decadal time-lapse remote sensing images is a challenging problem due to lot of surface changes. The research work devised a novel framework for change detection that utilizes Mode Improved Scale Invariant Feature Transform (M-SIFT) for spatial alignment and Guided Image Filter Enhanced Multivariate Alteration Detection (GIF-MAD) for change detection, achieves high accuracy, validated against ground truth data. The proposed approach demonstrates effectiveness in detecting changes using different satellite images over extended periods at Earth and Mars, achieving an overall accuracy of 90.9% with a kappa value of 0.81.

# Table of Contents

<b>Title Page</b>	<b>I</b>
<b>Certificate</b>	<b>II</b>
<b>Acknowledgements</b>	<b>III</b>
<b>Abstract</b>	<b>IV-V</b>
<b>Table of Contents</b>	<b>VI-IX</b>
<b>List of Tables</b>	<b>X-XII</b>
<b>List of Figures</b>	<b>XIII-XIX</b>
<b>List of Abbreviations</b>	<b>XX-XXIII</b>
<b>List of Symbols/Notations</b>	<b>XXIV-XXVI</b>
<b>Chapter 1 : Introduction</b>	<b>1-23</b>
1.1 Feature-based Image Registration	2
1.1.1 Scanty Feature Island Image Registration	7
1.1.2 Band-to-Band Registration (BBR)	8
1.1.3 Country-Level Image Mosaicking	8
1.1.4 Planetary Feature Extraction and Co-Registration	9
1.2 Remote Sensing Image Fusion	12
1.2.1 Infrared-Visible Image Fusion	14
1.2.2 Venus Visible Image Processing and Multi-Modal Fusion	16
1.3 Multi-Temporal Change Detection	17
1.3.1 Change Detection over Mars Terrain	19
1.4 Research Agenda	21
1.5 Organization of the Thesis	22
<b>Chapter 2 : Literature Review</b>	<b>24-36</b>
2.1 Introduction	24
2.2 Review of Feature Based Image Registration Methods	25
2.3 Review of Band-to-Band Registration (BBR) Techniques	27
2.4 Survey on Remote Sensing Image Fusion Methods	31
2.5 Review of Multi-Temporal Change Detection Techniques	33
2.6 Problem Definition	34
2.7 Objectives of the Research	36
2.8 Summary	36
<b>Chapter 3 : Related Theory</b>	<b>37-71</b>
3.1 Image Registration Methods using Features	37
3.1.1 Harris Corner Detector	39

3.1.2	SIFT Detector and Descriptor	40
3.1.3	SURF Detector/Descriptor	44
3.1.4	ORB Feature Detection/Description	45
3.1.5	BRISK Detector/Descriptor	48
3.1.6	Accelerated KAZE (A-KAZE) Feature Detection/Descriptor	49
3.2	Outlier Removal Techniques	52
3.2.1	RANSAC based Outlier Removal Technique	52
3.2.2	RANSAC Algorithm Steps	53
3.2.3	RANSAC Phases	68
3.3	Remote Sensing Image Fusion Theory and Types	56
3.3.1	Component Substitution (CS)	57
3.3.2	Multi Resolution Analysis (MRA)	58
3.3.3	Color Based (CB) Techniques	59
3.3.4	Variational Optimization (VO)	60
3.3.5	Deep Learning (DL) based Techniques	61
3.3.6	Hybrid Based Techniques	62
3.4	Change Detection Components and Methods	63
3.4.1	Image Differencing	64
3.4.2	Image Ratioing	65
3.4.3	Principal Component Analysis (PCA)	65
3.4.4	Multispectral Indices	67
3.4.5	Fuzzy Logic and Expert Systems	68
3.4.6	Deep Learning based Change Detection	69
3.5	Overall Methodology	70
3.6	Summary	71

**Chapter 4 : Feature based Image Registration Techniques** 72-177

4.1	Multi-Stage Feature Detection and Improved Outlier Removal Method	72
4.1.1	Introduction	72
4.1.2	Methodology Developed	72
4.1.3	Experimental Results	81
4.1.4	Summary	97
4.2	Multi-Satellite Scanty Feature Island Image Registration Methodology Development	99
4.2.1	Introduction	99
4.2.2	Methodology Developed	100
4.2.3	Experimental Results	107
4.2.4	Summary	120
4.3	Band-to-Band Registration of Multispectral Images	121

4.3.1	Introduction	121
4.3.2	Methodology Developed	121
4.3.3	Experimental Results	129
4.3.4	Summary	141
4.4	Direct Feature Extraction and Co-Registration of Bayer Pattern Raw Planetary Images	143
4.4.1	Introduction	143
4.4.2	Methodology Developed	144
4.4.3	Experimental Results	155
4.4.4	Summary	177
<b>Chapter 5 : Novel Country-Level Image Mosaicking System using Optical Remote Sensing Imagery</b>		178-196
5.1	Introduction	178
5.2	Methodology Developed	178
5.3	Experimental Results	187
5.4	Summary	195
<b>Chapter 6 : Spectra Preserving Remote Sensing Image Fusion Framework</b>		197-257
6.1	Spectra Preserving Pan-Sharpening Technique	197
6.1.1	Introduction	197
6.1.2	Methodology Developed	198
6.1.3	Experimental Results	211
6.1.4	Summary	221
6.2	Infrared-Visible Image Fusion	223
6.2.1	Introduction	223
6.2.2	Methodology Developed	224
6.2.3	Experimental Results	229
6.2.4	Summary and Avenues for Future Research	240
6.3	Multi-Modal Image Fusion and Processing using Venus Remote Sensing Images	241
6.3.1	Introduction	241
6.3.2	Methodology Developed	242
6.3.3	Venus Processed Image Quality Assessment	249
6.3.4	Conclusions and Future Research Prospects	257
<b>Chapter 7 : Enhanced Procedure for Multi-Decadal Change Detection using Planetary Remote Sensing Images</b>		258-281
7.1	Introduction	258
7.2	Proposed Method	258
7.3	Experimental Results	269
7.4	Summary	280
<b>Chapter 8 : Conclusions</b>		282-283



<b>Specific Contributions</b>	284-285
<b>Future Scope of Research Work</b>	286-287
<b>List of References</b>	288-316
<b>List of Publications</b>	317-318
<b>Brief Biography of the Candidate</b>	319
<b>Brief Biography of the Supervisor</b>	320

# List of Tables

<b>Table</b>	<b>Caption</b>	<b>Page No.</b>
1.1	Resourcesat Sensor Specifications	4
4.1	Details of LISS-4 Scenes and Pixel Error Before and After Correction	84
4.2	Absolute Location Accuracy in LISS-4 image before and after accurate Geometric Correction	87
4.3	Details of Multi-Satellite Images for Evaluation	92
4.4	Sub-Pixel Displacements between Registered Resourcesat-2 LISS-4 and Landsat-8 OLI	93
4.5	Details of LISS-3 Images for Co-Registration	94
4.6	Comparative Analysis with Feature-Based Methods	96
4.7	Details of AWiFS Image Scenes	97
4.8	Details of Satellites and Sensors Utilized for the Evaluation of Island Image Registration	108
4.9	Details of Multi-Satellite Images encompassing Lakshadweep Islands, Tamil Nadu Coast, and Gujarat Coast	108
4.10	Comparison of RMSE for Island Scenes employing different techniques for Remote Sensing Image Registration	111
4.11	Comparison of CMR for Island Scenes employing different techniques for Remote Sensing Image Registration	112
4.12	Processing Time (in Seconds) for Image Registration Methods across Various Island Image Scenes	112
4.13	Performance and Comparative Analysis of Outlier Removal Techniques	113
4.14	Shifts in Multi-Satellite Images Prior to and following Correction at Ground Control Points (GCPs)	117
4.15	RMSE Performance Comparison of MIRACLE <sup>NOENH</sup> , MIRACLE <sup>NOMAH</sup> , and MIRACLE across Island Scenes	118
4.16	CMR Performance Comparison of MIRACLE <sup>NOENH</sup> , MIRACLE <sup>NOMAH</sup> , and MIRACLE across Island Scenes	119
4.17	Comparison of Matched points after pruning and the corresponding matching time (in seconds) with and without the enhancement-based data pre-processing stage	119
4.18	Satellite Image Details used for analysis of the impact of Band Misregistration Effect	134
4.19	Estimation of Band-to-Band Misregistration before Correction	135
4.20	Estimation of Band-to-Band Misregistration Estimation after Correction	135
4.21	RMSE for the Comparison of BBR	139
4.22	Computation of CMR for the Comparison of BBR	139
4.23	NCM computation for BBR Comparison	140
4.24	Computation Time (in seconds) for Comparison of BBR	140

4.25	RMSE among different combinations of BBR Components	141
4.26	Details of MCC Bayer Pattern Raw Images	156
4.27	Experimental Parameters Configuration for the Proposed Method	156
4.28	Comparative Analysis of Quantitative Gradient Maps Between Raw image using the MCC Bayer pattern and Image demosaicked with MCC	166
4.29	Performance of Feature Extraction Across Various Levels of Noise	167
4.30	Comparison of RMSE and CMR Accuracy for Image Co-Registration	168
4.31	Co-Registration Performance of Images for Arisa Mons Across Iterative Executions of the Proposed Method	169
4.32	Co-Registration Performance of Multi-temporal Images in the Presence of Complex Backgrounds and Diverse Illumination Conditions	169
4.33	Comparative Assessment of RMSE Utilizing Different Image Registration Techniques on Bayer Pattern Raw Planetary Images	175
4.34	Comparative Assessment of CMR Utilizing Different Image Registration Techniques on Bayer Pattern Raw Planetary Images	175
4.35	Comparative Assessment of NCM Utilizing Different Image Registration Techniques on Multi-Temporal Bayer Pattern Raw Planetary Images	176
5.1	Parameters of Map Projection for the Indian Terrain	180
5.2	Introduced Errors in Input Image Strips/Scenes	188
5.3	Root Mean Square Error (RMSE) in Pixel Values Prior to and Following Correction with SIFT+RANSAC and SIFT+MB-RANSAC	190
5.4	Comparison of Strip and Mosaic Reflectance Measurements in the Overlapping Region	194
6.1	Comparison of SNR (Signal-to-Noise Ratio) and Noise Estimation	201
6.2	Metadata details for the panchromatic/multispectral image pairs utilized in the fusion evaluation	212
6.3	Evaluation of Image Fusion Metrics in Urban Landscape (Fusion Pair No. 1)	216
6.4	Evaluation of Image Fusion Metrics for Stadium Structure (Fusion Pair No. 2)	216
6.5	Evaluation of Image Fusion Metrics for Convention Centre (Fusion Pair No. 3)	217
6.6	Image Fusion Metrics Assessment over Solar Power Plant Structure (Fusion Pair No. 4)	217
6.7	Assessment of Fused Image Quality without Reference	218
6.8	Specifications of Thermal and Visible Remote Sensing Images for Fusion Evaluation	230
6.9	Assessment of Reference Quality Parameters for Thermal-Visible Image Fusion in the Pair of Thermal Infrared from INS-2TD and Multispectral Visible Images from MODIS	236
6.10	Evaluation of Reference Quality Parameters for Thermal-Visible Image Fusion in the Pair of Thermal Infrared Images from Landsat-8 and Multispectral Visible Images from MODIS	236

6.11	Assessment of Reference Quality Parameters for Thermal-Visible Image Fusion in the dataset containing Long-Wave Infrared (LWIR) and Visible Images from TNO	237
6.12	Evaluation of No-Reference Image Quality Metrics for Fusion of INS-2TD Infrared MODIS Visible Images	238
6.13	Assessment of No-Reference Image Quality Metrics for Fusion of Landsat-8 Infrared-MODIS Visible Images	238
6.14	Processing Time (in seconds) of Image Fusion Methods for Infrared-Visible Image Pairs	239
6.15	Comparison of Fusion Performance for Landsat-8/MODIS Visible Image Pair among CLIM <sup>NOENH</sup> , CLIM <sup>NOCOF</sup> , and CLIM	240
6.16	Comparison of SNR and Noise Estimation	250
6.17	Performance of Venus' Visible Image Quality without Reference	251
6.18	Evaluation of Registration Accuracy for Venus' Visible Co-Registered Image with Respect to Venus' Radar Image Reference	253
6.19	Comparative RMSE Analysis of Planetary Image Registration Methods	254
6.20	Evaluation of Venus' Image Fusion Quality	254
7.1	Details of IRS-1C/1D Scenes and Corresponding RMSE in Comparison to RS-2A Reference	262
7.2	Change Detected in No Change Region for IRS Data	266
7.3	Metadata Specifications of the Remote Sensing Datasets Employed in the Experiment	270
7.4	Alterations in Healthy Vegetation within a Specific Area near Gandhinagar, India	273
7.5	Error Matrix for Change/No-Change Classes in Bi-Temporal Images No. 1	275
7.6	Error Matrix for Change/No-Change Classes in Bi-Temporal Images No. 2	275
7.7	Error Matrix for Change/No-Change Classes in Bi-Temporal Images No. 3	275
7.8	Comparison of Change Detection Results for Bi-Temporal Images No. 3	276
7.9	Error matrix for change/no-change classifications in Bi-Temporal Images No. 4	277
7.10	Error matrix for change/no-change classifications of Bi-Temporal Images No. 5	280

# List of Figures

## Figure Caption

	<b>Page No.</b>	
1.1	Workflow for Automatic Feature-Based Image Registration	3
1.2	Multi-Temporal Remote Sensing Image Pair	5
1.3	(a) The three-tier imaging mechanism of Resourcesat. (b) The reference layer of LISS-4 covering India and its adjacent areas	5
1.4	Misalignment of multi-temporal images over (a) agricultural land and (b) wetland areas	6
1.5	Lakshadweep Islands, predominantly situated in the deep ocean, are captured by the Indian Resourcesat LISS-4, along with the geographic coordinates of the scene center	7
1.6	Band-to-Band Misregistration in Multispectral Remote Sensing Images.	8
1.7	Mars Full Disc Image captured by Indian Mars Color Camera (MCC)	10
1.8	MCC captured Pital Crater (a) Raw Bayer Pattern Image (b) Radiometrically Processed Image	12
1.9	Remote Sensing Image Fusion Procedure	14
1.10	Panchromatic and Multispectral Image Pair for Image Fusion	14
1.11	Multi-Modal Remote Sensing Image Pair: (a) Thermal Infrared Broadband Image with Relatively High Spatial Resolution, and (b) Multispectral Visible False Color Composite (FCC) Image with Low Spatial Resolution	15
1.12	Venus' Multi-modal Image Pair (a) Visible Raw Image (b) Radar Topography Image	17
1.13	Multispectral Images from Indian Remote Sensing (IRS) satellite for Historical Change Detection (a) Gandhinagar Region (b) Ahmedabad Airport and Surrounding	18
1.14	Indian Mars Color Camera processed images captured different Mars Terrain	20
1.15	Mars Global Mosaic using recent MCC Full Disc Images	20
1.16	Viking Merged Color Mosaic showing Global Mars Landscape	21
3.1	Classification of Points using Eigenvalues of Matrix M	40
3.2	DoG Pyramid Creation and Localization of extrema	41
3.3	Feature Detection using ORB	48
3.4	(a) Matched control points, including a significant number of outliers for line fitting. (b) RANSAC based line fitting process, impervious to the influence of outliers	52
3.5	RANSAC Flowchart	54
3.6	Loss Function	55
3.7	Remote Sensing Image Processing Workflow to generate value added information	71
4.1	Pruning of Matched Points Utilizing Motion Smoothness Constraint	76
4.2	Joint Scatter Plot Illustrating Pruned Matched Points with MSC	77

4.3	Workflow for Coarse Image Co-Registration	77
4.4	(a) Segmented Affine Transform Estimation through Tiled SIFT Matching	
	(b) Workflow for Fine Image Co-Registration	79
4.5	Framework for Precision Correction using FIRM	81
4.6	Geographical Extent of Resourcesat LISS-4 Scenes Employed in Co-Registration	82
4.7	Location Discrepancy (in pixels) of Input LISS-4 Data Products	86
4.8	Location Discrepancy (in pixels) of Output LISS-4 Data Products	86
4.9	Ground Control Points Identified on the LISS-4 Image	88
4.10	(a) Sub-Pixel Alignment between LISS-4 on April 3, 2018, and Reference	
	(b) Sub-Pixel Registration between LISS-4 on October 11, 2019, and Reference	89
4.11	Co-Registration Performance over Wetland	90
4.12	Co-Registration Performance over Agricultural Land	91
4.13	(a) Coverage of Resourcesat-2 on the Same Day over Landsat-8 (b) Overlapping Image Chips for Evaluation	92
4.14	Co-Registration Performance of Resourcesat-2 and Landsat-8	93
4.15	Co-Registration Performance of LISS-3 in Various Image Regions	95
4.16	Geometric Consistency in Overlap Areas in the AWiFS Image Mosaic Covering Bangladesh and its Surroundings	97
4.17	Lakshadweep Islands, primarily located in the vast expanse of the deep ocean, as captured by the Indian Resourcesat LISS-4, along with the geographical coordinates (latitude and longitude) of the central point of the scene	99
4.18	Outlier Removal process using Mahalanobis Distance Guided Marginalization Optimization	105
4.19	Workflow of the MIRACLE Island Image Registration process	106
4.20	Evaluation of Image Registration across the Lakshadweep Islands, Tamil Nadu, and Gujarat Coast Regions	109
4.21	Quality Assessment of Co-registration of Amini Island: (a) Vertical adjustments prior to and following correction, (b) Horizontal shifts prior to and following correction	115
4.22	Quality Assessment of Co-Registration of Bitra Island (a) Vertical shifts prior to and following correction, (b) Horizontal shifts prior to and following correction	115
4.23	Quality Assessment of Co-Registration of Kiltan Island: (a) Vertical shifts prior to and following correction, (b) Horizontal shifts prior to and following correction	116
4.24	Evaluation of Checkerboard Pattern in Co-registration of images between the red channel registered image from Resourcesat LISS-4 and the Landsat-8 OLI PAN image (a) Amini, (b) Bitra, (c) Kiltan	116
4.25	Scale Space Construction for Feature Detection and Description (a) Input Image for Co-occurrence Scale Space (CSS) (b) CSS at Layer-1 (c) CSS at Layer-2 (d) CSS at Layer-3 (e) Input Image for SIFT Scale Space (f) SIFT Scale Space at Layer-1 (g) SIFT Scale Space at Layer 2 (h) SIFT Scale Space at Layer-3	123

4.26	Pruned Matched Points between two Image Bands (a) RANSAC (b) SC-RANSAC	126
4.27	Proposed Workflow for Band-to-Band Registration (BBR)	128
4.28	INS-2B Nano-MX Images (No of Image Scenes =220) utilized for BBR evaluation, blended over the boundary of Indian States	129
4.29	BBR Performance over Water Body, Sparse Urban Region, and Vegetation with a 4X Zoom	131
4.30	BBR Performance over Snow and Glacier Region with a 4X Zoom	131
4.31	BBR Performance over Cloudy Region with a 4X Zoom	132
4.32	Checkerboard Pattern based Visual BBR Evaluation (a) Band-1 w.r.t. Band-3 (b) Band-2 w.r.t. Band-3 (c) Band-4 w.r.t. Band-3 (d) Band-1 w.r.t. Band-3 before BBR Correction (e) Band-1 w.r.t. Band-3 after BBR Correction (f) Band-2 w.r.t. Band-3 before BBR Correction (g) Band-2 w.r.t. Band-3 after BBR Correction (h) Band-4 w.r.t. Band-3 before BBR Correction (i) Band-4 w.r.t. Band-3 after BBR Correction	132
4.33	Band Misregistration Effect on NDVI (a) Nano-MX FCC before BBR (b) Nano-MX FCC after BBR (c) Landsat-9 FCC (d) NDVI Map from Nano-MX FCC before BBR (e) NDVI Map from Nano-MX FCC after BBR (f) NDVI Map from Landsat-9 FCC	134
4.34	Band-1 wrt. Band-3 BBR Performance before and after correction	136
4.35	Band-2 wrt. Band-3 BBR Performance before and after correction	136
4.36	Band-4 wrt. Band-2 BBR Performance before and after correction	137
4.37	Edge Operators: (a) Central Difference Operator (b) Sobel Operator	144
4.38	Feature Extraction Pipelines (a) Traditional Feature Extraction Pipeline (b) Novel Pipeline for the Direct Extraction of Features from MCC Bayer Pattern Raw Image	146
4.39	Extraction of Gradient Magnitude from MCC Raw Bayer Pattern Image	146
4.40	(a) Image in MCC Bayer Pattern (b) Image Obtained from MCC Demosaicking (c) Gradient Magnitude Map Derived from Bayer Pattern (d) Gradient Magnitude Map Obtained from Demosaicking (e) Image of Gradient Differences	148
4.41	(a) Bayer pattern MCC Intensity Raw Image (b) Gradient Magnitude Map (GMM) obtained directly from the MCC Bayer Intensity Raw Image	151
4.42	Workflow for Direct Co-Registration of Bayer Pattern Raw Images	154
4.43	Extraction of Features Directly at Arisa Mons (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image	158
4.44	Extraction of Features Directly in the Vicinity of the Gale Crater Region (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image	159
4.45	Extraction of Features Directly at the Pital Crater (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer	

	Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image	160
4.46	Extraction of Features Directly at Tharsis Tholus (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image	161
4.47	Extraction of Features Directly from a Full Disc Image of Mars Captured by the Viking Orbiter Mission (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image	162
4.48	Extraction of Features Directly from an Image of Mars Captured by the Mars Global Surveyor Mission (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image	163
4.49	Evaluating feature extraction performance on the Pital Crater under different levels of Gaussian Noise (a) Raw MCC Bayer Pattern Image with no noise. (b) Image in its raw form with the addition of Gaussian Noise $\sigma = 2.0$ (c) Image in its raw form with the addition of Gaussian Noise $\sigma = 5.0$ (d) Image in its raw form with the addition of Gaussian Noise $\sigma = 10.0$ (e) Image with extracted features and no noise (f) Image with extracted features and the presence of noise $\sigma = 2.0$ (g) Image with extracted features and the presence of noise $\sigma = 5.0$ (h) Image with extracted features and the presence of noise $\sigma = 10.0$	164
4.50	Transformed Region of Interest (ROI) reference extracted from MDIM 2.1 (a) Arisa Mons (b) Gale Crater (c) Tharsis Tholus	166
4.51	Visual Assessment of Sub-Pixel Alignment between MCC Bayer and MDIM Reference for Different Mars Craters (a) Crater-01 (b) Crater-02 (c) Crater-03 (d) Crater-04	170
4.52	Results of Co-Registration for Planetary Images Across a Martian Crater (a) Prior to Registration (b) Post Registration	171
4.53	Results of Co-Registration for Planetary Images Across Kinkora Crater (a) Prior to Registration (b) Post Registration	171
4.54	Performance of Outlier Removal Utilizing FSS-RANSAC Near Gale Crater (a) MCC Bayer Intensity Image (b) MDIM 2.1 Reference Image	172
4.55	Performance of Outlier Removal Utilizing FSS-RANSAC Near Martian Craters (a) MCC Bayer Intensity Image (b) MDIM 2.1 Reference Image	173
4.56	Performance of Outlier Removal Utilizing FSS-RANSAC Near Valles Marineris (a) MCC Bayer Intensity Image (b) MDIM 2.1 Reference Image	173
4.57	Performance in Matching Features Between Multi-temporal Planetary Images (a) Harris+RANSAC (b) SIFT+RANSAC (c) ORB+MSC (d) SURF+MLESC (e) KAZE+MAGSAC (f) Semantic DNN (g) MU-NET (h) Proposed Method	176
5.1	Proposed System for Integrated Image Mosaicking at the Country Level	179
5.2	Workflow for Fine Geo-Registration of Image Strips	184



5.3	Workflow for Automatic Image Stitching	186
5.4	Large Area Mosaic of Multi-Temporal LISS-3 and Identification of RoIs (Yellow, Green, and Blue) for Assessing Overlap Region Performance of Remote Sensing Image Strips	188
5.5	Full-Resolution View of Overlapping Region within the Yellow ROI Box (Sparse Vegetation)	188
5.6	Full-Resolution View of Overlapping Region within the Green ROI Box (Urban/Water Bodies)	189
5.7	Full-resolution perspective of the overlap region within the Blue ROI Box (Hilly Region).	189
5.8	(a) Complete Mosaic of India (Image Layer: RS-2A LISS-3 False Color Composite (FCC) utilizing Near Infrared, Red, and Green Wavelength Channels) (b) Enlarged View of Gujarat Region Demonstrating a Seamless Mosaic	192
5.9	Reflectance Image Mosaic of Bangladesh at the Country Level for Various Years (Band Combination: SWIR, Near Infrared, and Red) (a) 2017 (b) 2018 (c) 2019 utilizing Resourcesat-2A LISS-3 Data	192
5.10	Image Blending Comparison (a) Image Blending using Mean Statistics in the Overlapping Area (b) Image blending using region-based statistics as proposed	193
5.11	Spectral Classes in the Overlapping Region for Comparing Mosaic Reflectance (Image Layer displayed in False Color Composite)	194
5.12	Schema of AWiFS Country-Level Mosaic	195
6.1	Performance of co-registering panchromatic and multispectral images	199
6.2	Computation of medoids from a multispectral image	200
6.3	Performance of the Bilateral Filter (a) Original Panchromatic Image (b) Output from the Bilateral Filter (c) Zoomed-in View (4X) of the Original Panchromatic Region (d) Zoomed in View (4X) of the Denoised Region after Applying Bilateral Filter	202
6.4	Architecture of the The Holistic Nested Edge Detection Network implemented on an image with medoid intensity matching	204
6.5	(a) Angles used in the calculations for solar illumination, (b) Digital Elevation Model (DEM), and (c) Surface normal derived from the DEM	207
6.6	Processing Workflow of SPRINT	210
6.7	Visual Evaluation of Urban Landscape Fusion Pair No. 1 [(a) Cartosat-1 PAN (b) Resourcesat LISS-4 Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]	213
6.8	Visual Evaluation of Stadium Fusion Pair No. 2 [(a) IKONOS PAN (b) IKONOS Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]	214
6.9	Visual Evaluation of Convention Centre Fusion Pair No. 3 [(a) Ground Truth (b) Degraded Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]	214

6.10	Visual Evaluation of Fusion Pair No. 4 featuring Solar Power Plant Structure [(a) Ground Truth (b) Degraded Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]	214
6.11	Charts depicting image quality metrics for the comparison of image fusion	218
6.12	Performance of SPRINT over hilly terrain for Fusion Pair No. 5 (a) Ground Truth Reference, (b) Cartosat-1 Degraded Ortho Rectified Tile, (c) LISS-4 Degraded Multispectral, (d) SPRINT, (e) Ground Truth NDVI, (f) SPRINT NDVI, and (g) Comparison of NDVI between SPRINT and the reference across various classes	219
6.13	(a) Visual comparison of SPRINT reflectance fusion over an urban area (Fusion Pair No. 6). (b) Visual comparison of SPRINT reflectance fusion over water and vegetation (Fusion Pair No. 7).	221
6.14	Comparison of spectral reflectance profiles between the original interpolated reflectance from LISS-3 and SPRINT-generated reflectance across various spectral classes	221
6.15	(a) Original Broadband Thermal Image (b) Output of Co-occurrence Filter	225
6.16	(a) Original Thermal Infrared Image (b) Sharpened Thermal Infrared Image using <i>LoG</i>	226
6.17	Comparative Analysis of CLAHE for enhancing multispectral visible images (a) Input Image (b) Enhanced with $w = 64$ (c) Enhanced with $w = 256$ (d) Enhanced with $w = 512$	227
6.18	CLIM Workflow for Infrared-Visible Fusion	228
6.19	Registration Performance Comparison for Multi-Modal Images of Infrared and Visible Channels (a) Before Registration (b) After Registration	230
6.20	Region of Saso Dam and Refinery Complex (a) INS-2TD Broadband Thermal Image (b) Resampled MODIS Visible Multispectral Image (c) CLIM Fused Image	231
6.21	Areas with Ship Docking Zones (a) Thermal Infrared Image from Landsat-8 (b) Resampled Multispectral Visible Image from MODIS (c) CLIM Fused Image	231
6.22	Comparison of Fusion Results for INS-2TD Broadband Thermal Infrared and MODIS Visible Images across Different Feature Targets (a) Thermal Infrared Image (b) Resampled Multispectral Image (c) BT (d) IIHS (e) PCA (f) SOE (g) FGIF (h) CLIM	233
6.23	Comparison of Fusion Outcomes for Landsat-8 Thermal Infrared and MODIS Visible Images across Different Feature Targets (a) Thermal Infrared Image (b) Resampled Multispectral Image (c) BT (d) IIHS (e) PCA (f) SOE (g) FGIF (h) CLIM	234
6.24	Assessment of CLIM Fusion on TNO Dataset (a) Infrared Image (b) Visible Image (c) BT (d) IIHS (e) PCA (f) SOE (g) FGIF (h) CLIM	234
6.25	Performance of Venus' Visible Image Denoising (a) Raw Venus' Visible Image (b) Denoised Venus' Visible Image	243
6.26	Visual Assessment of Venus' Image Enhancement (a) Denoised Venus' Image (b) Enhanced Venus' Image	244
6.27	Multi-modal Pair of Venus' Images (a) Denoised and Enhanced Visible Image (b) Radar Image from Magellan Mission (c) Simulated Venus' Image from Radar Data for Image	

	Registration	245
6.28	Venus' Image Fusion Results (a) Enhanced Co-Registered Visible Image (b) Fused Image	248
6.29	Proposed Advanced Workflow for Venus' Image Processing	249
6.30	Quality Evaluation of Venus' Visible Image Region of Interest (a) Preview of Denoised Image (b) 4X Zoomed Raw Visible Image (c) 4X Zoomed Denoised Visible Image	250
6.31	Venus' Visible Cropped Image Pair for Entropy and BRISQUE Evaluation (a) Denoised Image (b) Enhanced Image	252
6.32	Horizontal Swipe Visual Assessment of Venus' Image Co-Registration at Different Regions of Interest (RoI) (a) RoI-1 Swipe (b) RoI-2 Swipe (c) RoI-3 Swipe (d) RoI-4 Swipe	252
6.33	Validation of Venus' Fused Annotated Image with Venus' Radar Nomenclature Map	255
6.34	Fuzzy Segmentation on Venus' Visible Image (a) Venus' Fused Image (b) Fuzzy Segmented Image (c) Color-coded Fuzzy Segmented Map	256
7.1	Image Co-Registration Performance of M-SIFT Across Various Features	262
7.2	Workflow for change detection using GIF-MAD	265
7.3	Visual Quality Evaluation of Various Change Detection Techniques	267
7.4	EPOCH Processing Framework	268
7.5	(a) Study Regions (Highlighted in Yellow Boxes) depicted on the Indian Resourcesat LISS-3 Mosaic (Earth) (b) Study Regions (Highlighted in Yellow Boxes) presented on ISRO's MCC Global Mosaic (Mars) (c) Study Region (Highlighted in Yellow Box) displayed on the Blue Marble Image of the United States (Earth)	270
7.6	Expansion of Ahmedabad Airport depicted through Bi-Temporal Images No. 1	271
7.7	Identification of man-made features around Umsawli, India, depicted in Bi-Temporal Images No. 2	271
7.8	Change detection observed over Shillong Airport, India, depicted in Bi-Temporal Images No. 2	272
7.9	Detection of greenery in Gandhinagar using IRS data (Bi-Temporal Images No. 3)	274
7.10	Urban expansion in Sioux Falls, United States, as observed from Landsat images (Bi-Temporal Image No. 4)	277
7.11	Change Detection across Mars Quadrangles (Bi-Temporal Images No. 5)	278
7.12	Change Detection around Elysium Mons (Bi-Temporal Images No. 5)	279
7.13	Change Detection over Utopia Basin (Bi-Temporal Images No. 5)	279
7.14	Fresh Impact Crater Formation near Amenthes Fossae (Bi-Temporal Images No. 5)	279

# List of Abbreviations

<b>Abbreviation</b>	<b>Expanded Form</b>
CP	Control Points
ARD	Analysis Ready Data
LISS-3	Linear Imaging Self Scanner-3
LISS-4	Linear Imaging Self Scanner-3
AWiFS	Advance Wide Field Scanner
GAN	Generative Adversarial Network
UTM	Universal Transverse Mercator
BBR	Band-to-Band Registration
FCC	False Color Composite
NDVI	Normalized Difference Vegetation Index
CFA	Color Filter Array
MCC	Mars Color Camera
MOM	Mars Orbiter Mission
LTA	Long Term Archive
PDS	Planetary Data System
SNR	Signal to Noise Ratio
CS	Component Substitution
BT	Brovey Transform
IHS	Intensity Hue Saturation
MRA	Multi-Resolution Analysis
HPF	High Pass Filtering
SFIM	Smoothing Filter Intensity Modulation
VO	Variational Optimization
CNN	Convolutional Neural Networks
WISPR	Wide-Field Imager for Parker Solar Probe
NASA	National Aeronautics and Space Administration
ISRO	Indian Space Research Organisation
ESA	European Space Agency
SIFT	Scale Invariant Feature Transform
SURF	Speeded-Up Robust Features
ORB	Oriented FAST and Rotated BRIEF
UAS	Unmanned Aerial System
MODIS	Moderate Resolution Imaging Spectroradiometer
NMI	Normalized Mutual Information
FFT	Fast Fourier Transform
EFL	Effective Focal Length

UAV	Unmanned Aerial Vehicle
MPT	Modified Projective Transformation
RAC	Robust and Adaptive Correction
SRCA	Spectro Radiometric Calibration Assesmbly
FPA	Focal Plane Assesmbly
BDSB	Band Dependent Spatial Detail
PCA	Principal Component Analysis
GLP	Generalized Laplacian Pyramid
MTF	Modulation Transfer Function
MRA	Multiresolution Analysis
HQB	High-Quality Bayesian
SFA	Slow Feature Analysis
CVA	Change Vector Analysis
MAD	Multivariate Alteration Detection
RANSAC	Random Sample Consensus
RMSE	Root Mean Square Error
CMR	Correct Matching Ratio
LoG	Laplacian of Gaussian
DoG	Difference of Gaussian
FAST	Features from Accelerated Segment Test
BRIEF	Binary Rotated Independent Elementary Feature
SAR	Synthetic Aperture Radar
BRISK	Binary Robust Invariant Scalable Keypoint
GPU	Graphical Processing Unit
CB	Color Balanced
DL	Deep Learning
RNN	Recurrent Neural Networks
FIRM	Framework for Image Registration using Multi-Stage Feature Detection and Mode Guided Motion Smoothness Keypoint Optimization
MSC	Motion Smoothness Constraint
GCP	Ground Control Points
OLI	Operational Lander Imager
PAN	Panchromatic
MIRACLE	Multi-Satellite Island Image Registration using Anisotropic Coherence Locality Enhanced Nonlinear Diffusion and Mahalanobis Distance Guided Marginalization
FED	Fast Explicit Diffusion
MDGM	Mahalanobis Distance Guided Marginalization
LSS	Local Self Similarity
DNN	Deep Neural Network

ST	Semantic Template
CSS	Co-occurrence Scale Space
DoC	Difference of Co-occurrence
LoC	Laplacian of Co-occurrence
SC-RANSAC	Spatial Confined RANSAC
SAT	Segmented Affine Transformation
VNIR	Visible-Near Infrared
NIR	Near-Infrared
CDC	Color Difference Constancy
PSNR	Peak Signal to Noise Ratio
PRNU	Photo Response Non-Uniformity
GMM	Gradient Magnitude Map
FSS	Feature Similarity Score
CoV	Coefficient of Variation
MSSIM	Mean Structural Similarity Index Map
MSE	Mean Square Error
CD	Canberra Distance
SAD	Sum of Absolute Distance
GMSD	Gradient Magnitude Similarity Deviation
MDIM	Mars Digital Image Model
MAGSAC	Marginalization Sample Consensus
NCM	Number of Correct Matches
MB-RANSAC	Mode-Biased Random Sample Consensus
LCC	Lambert Conformal Conic
Geo-TIFF	Geo Tagged Image File Format
RoI	Regions of Interest
RTM	Radiative Transfer Modeling
SPRINT	Spectra Preserving Radiance Image Fusion Technique using Holistic Deep Edge Spatial Attention and Minnaert Guided Bayesian Probabilistic Model
SA-SIFT	Segmented Affine Scale Invariant Feature Transform
HNED	Holistic Nested Edge Detection
DEM	Digital Elevation Model
CLIM	Co-Occurrence with Laplacian Intensity Modulation
INS-2TD	Indian Nano Satellite Technology Demonstrator
CoF	Co-occurrence Filter
CLAHE	Contrast Limited Adaptive Histogram Equalization
ISSDC	Indian Space Science Data Center
SAC	Space Applications Center
NRSC	National Remote Sensing Center
FGIF	Fast Guided Image Filtering
LWIR	Long-Wave Infrared
SAM	Spectral Angle Mapper
ERGAS	Relative Dimensionless Global Error in Synthesis
UIQI	Universal Image Quality Index
QNR	Quality with No Reference
TSM	Tri-State Median

GIF	Guidance Image Filter
CR	Component Ratio
BRISQUE	Blind/Reference-less Image Spatial Quality Evaluator
M-SIFT	Mode Improved SIFT
GIF-MAD	Guided Image Filter Enhanced Multivariate Alteration Detection
LMD	Local Mean Difference
GLS	Global Land Survey
EPOCH	Enhanced Procedure for Operational Change Detection using Historical Invariant Features and PCA Guided Multivariate Statistical Technique
RAM	Random Access Memory
CPU	Central Processing Unit

# List of Symbols/Notations

Notation	Detail
$I(x, y)$	Image Intensity Component at coordinate position $(x, y)$
$I(x + u, y + v)$	Shifted Intensity Component at coordinate position $(x + u, y + v)$
$R$	Corner Response Measure
$M$	Matrix obtained from the Gradients of the image
$L(x, y, \sigma)$	Smooth Image at coordinate position $(x, y)$ and $\sigma$ is the standard deviation of Gaussian Distribution
$G(x, y, \sigma)$	Gaussian Filter
$DoG(x, y, \sigma)$	Difference of Gaussian
$Z$	Extremum Location
$m(x, y)$	Gradient Magnitudes at coordinate position $(x, y)$
$\theta(x, y)$	Orientation at coordinate position $(x, y)$
$H(x, \sigma)$	Hessian Matrix
$Sp \rightarrow x$	location state of $x$ with respect to point $p$
$H(p)$	Entropy of the set $p$
$c(x, y, t)$	Conductivity function at coordinate position $(x, y)$ with scale parameter $t$
$N$	Number of iterations to randomly select all samples of inliers at least once
$\gamma$	Inlier ratio
$F(i, j)$	pixel value at location $(i, j)$ in the merged image
$C(i, j, k)$	coefficients from the input images at scale $k$
$W(k)$	weight at scale $k$
$\alpha$	weighting factor that determines the contribution of each source image to the merged image
$E(F)$	Energy function that needs to be minimized for fused Image $F$
$D(i, j)$	Difference at location $(i, j)$ between the corresponding pixels in the two input images
$R(i, j)$	Ratio at location $(i, j)$ between the corresponding pixels in the two input images
$C$	Co-variance Matrix
$NDVI$	Normalized Difference Vegetation Index
$a_i$	Affine Parameters
$S_{ij}$	count of matches occurring between the input grid cell $I_i$ and the reference grid cell $R_j$
$J_p$	Semidefinite Matrix
$\mu$	Eigenvalues of the Matrix
$D_{mx}$	Mahalanobis distance in the horizontal direction
$PD_x$	Difference in sample matched points for point $P$ along the horizontal direction
$M_x$	average of the differences in the set of matched points along the horizontal direction



$SDx$	Deviation from the mean in the horizontal direction
$Dmy$	Mahalanobis distance in the vertical direction $Dmy$ for a sample point $P$
$PDy$	Sample matched point difference for point $P$ in the vertical direction
$My$	Mean of the matched point set differences in the vertical direction
$SDy$	Standard deviation in the vertical direction
$Dm$	Mahalanobis distance estimate for a sample point $P$
$CMR$	Correct Matching Ratio
$Cop$	Co-occurrence output pixel image value at the pixel index $p$
$DoC(x,y,\sigma)$	Difference of Co-occurrence
$Dx(x,y)$	Image gradients in the horizontal direction at pixel coordinates $(x, y)$
$Dy(x,y)$	Image gradients in the vertical direction at pixel coordinates $(x, y)$
$FSSp$	Feature Similarity Score for the matched pixel $p$
$ps$	strip reflectance measurement
$pm$	mosaic reflectance measurement
$SRD(\%)$	Spectral Reflectance Deviation in percentage
$\eta k$	medoid intensity value at a specific pixel location $k$
$\mu k$	average intensity value at a specific pixel location $k$
$n$	number of spectral bands
$xik$	intensity value of pixel at location $k$ for spectral band $i$ .
$BF(Ip)$	Bilateral Filter Output of Image $I$ at pixel location $p$
$Wp$	Normalization Factor at pixel location $p$
$lside$	Side Output Layer
$Lfuse$	Loss function of the fusion layer
$RF$	Radiance Factor
$k$	Minnaert Constant
$\varphi$	Geographic Latitude
$\omega$	Geographic Longitude
$\rho H$	radiance of a horizontal surface
$\rho T$	radiance of an inclined terrain
$\beta$	solar illumination angle
$\lambda$	Terrain slope
$k$	Minnaert Parameter
$\lambda$	Terrain slope
$\Phi$	solar zenith angle
$\theta a$	solar azimuth
$\theta b$	Topographic azimuth
$Y_S$	Pixel in the multispectral band of the image with low spatial resolution
$Y_P$	Pixel in the panchromatic band of the image with high spatial resolution
$Y_H$	Pixel representing holistic nested deep edge information from the medoid intensity-matched image

$E$	vector of errors originating from the panchromatic, multispectral, and holistic nested edge images
$I_{HPF}$	High Pass Image
$I_T$	Thermal input image
$I_{LPF}$	Co-occurrence Low Pass Filter Image
$I_{SH}$	Enhanced Thermal Image
$I_{CLIM}$	CLIM Fused image for Infrared and Visible Bands
$TSM_{ij}$	Tri-State Median Filter pixel value at image pixel coordinates (i,j)
$I_{Visible}$	Venus' Visible Image
$mode_x$	mode estimate in the horizontal direction
$mode_y$	mode estimate in the vertical direction
$\Delta x$	spatial displacement between bi-temporal images along the horizontal direction
$\Delta y$	spatial displacement between bi-temporal images along the vertical direction
$f_i$	Guided Image Filter

# 1. INTRODUCTION

Feature-based image registration is a fundamental process in the field of image analysis and computer vision, facilitating the alignment of different images depicting the same scene. By identifying and matching distinctive features such as corners, edges, or blobs between images, feature-based registration enables the accurate spatial alignment necessary for tasks like multi-temporal analysis, image fusion, and object tracking. This process involves detecting salient features in each image, describing them mathematically, and then matching corresponding features across images. Once matched, geometric transformations are estimated to align the images spatially. Feature-based image registration finds applications in various domains, including remote sensing, where precise spatial alignment is critical for subsequent analysis and decision-making.

Image fusion is a pivotal technique in the realm of image processing, serving to integrate information from multiple images of the same scene into a single, comprehensive representation. Through the amalgamation of complementary details from various sources, remote sensing image fusion aims to produce an enhanced and more informative image that preserves the most pertinent spectral features from each input. Whether operating in the spatial or transform domain, image fusion methodologies range from simple averaging to sophisticated algorithms designed to exploit specific characteristics of the input data. This process finds application across a myriad of domains, including remote sensing imagery, where the amalgamation of diverse data sources enriches the analysis and interpretation of images, facilitating informed decision-making and enabling a deeper understanding of complex visual data.

Remote sensing change detection is a vital methodology employed in analyzing alterations occurring within landscapes over time, utilizing satellite or aerial imagery to detect and quantify changes. By comparing multi-temporal images of the same geographic area, remote sensing change detection facilitates the identification of significant alterations in land cover, land use, infrastructure, and environmental conditions. This process involves several steps, including image registration to align images, pixel-wise differencing to highlight changes, and classification to categorize detected alterations. Remote sensing change detection plays a pivotal role in various fields such as urban planning, environmental monitoring, disaster management, and agricultural assessment, providing invaluable insights into land dynamics, natural phenomena, and human activities. Through its ability to monitor and analyze changes at different scales and temporal

resolutions, remote sensing change detection contributes to informed decision-making processes aimed at sustainable resource management and resilient development strategies.

## **1.1. Feature-Based Image Registration**

Remote sensing images are a valuable source of information regarding our planet, attracting attention from researchers globally (Campbell & Wynne, 2011). Remote sensing image registration is a crucial step in processing data for various space-based applications, such as monitoring wetland dynamics and understanding crop growth patterns (Le et al., 2011). Image registration involves aligning images of a particular scene taken at various times, angles, and through diverse sensors has been emphasized (Zitova & Flusser, 2003). When mapping the Earth using remote observation images, efficiently handling substantial volumes of satellite data is essential for prompt and precise registration, playing a vital role in applications of remote sensing such as merging images and detecting changes (Ghassemian, 2016; Singh, 1989).

Remote sensing image registration techniques are generally categorized into "area-based methods," "feature-based methods," and "deep learning methods." Among these, deep learning, particularly Generative Adversarial Network (GAN)-based feature matching, is emerging but not yet mature for operational deployment. Feature-based methods, relying on feature abstraction, often outperform other methods for most remote sensing pairs. Traditional registration based on features involves manually selecting Control Points (CPs), a laborious task requiring expert skills. The need for an automated registration technique relying on features, as suggested by Bentoutou et al. (2005), is highly advantageous, even when producing Analysis Ready Data (ARD) products (Dwyer et al., 2018).

Automatic feature-based image registration involves sequential and iterative phases, including feature detection, mapping agreements among corresponding feature points, parameter estimation, and resampling. The process requires selecting a suitable feature detector/descriptor, matching features, and removing outliers producing a pair of co-registered satellite images. Additionally, estimating transformation parameters using pruned matched control points and resampling the input image are crucial steps in generating the required registered image. Feature-based image registration is well-suited for satellite remote sensing images with diverse terrains, given the abundance of feature targets. Figure-1.1 illustrates the typical workflow for image registration based on features applicable to creating co-registered data products.

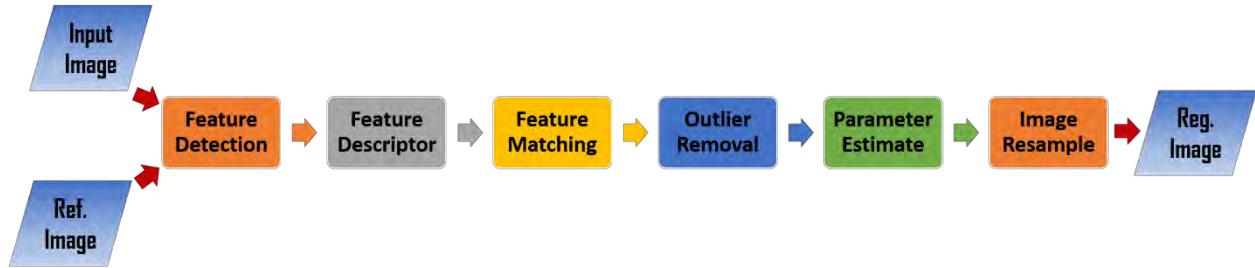


Figure-1.1. Workflow for Automatic Feature-Based Image Registration

Automated satellite image registration based on features poses a formidable challenge, involving the precise alignment of two images to achieve geometric conformity through the establishment of a transformation model. This process extracts unique feature points simultaneously from an image pair without any assistance (Misra et al., 2012a). Remote sensing images encompass various electromagnetic spectrum bandwidths, offering diverse information for analysis. Multispectral visible images, covering blue, green, and red wavelengths, resemble natural images, making them suitable for tasks such as image classification (Li et al., 2014) and object recognition (Inglada, 2007). Figure-1.2 illustrates a multi-temporal remote sensing image pair that necessitates accurate alignment at the sub-pixel level for varied remote sensing data analytics. The remote sensing images are captured by the LISS-3 sensor onboard the Indian Resourcesat-2A satellite. Resourcesat-2/2A satellites play a crucial role in providing essential multispectral data within the visible observable and Shortwave Infrared (SWIR) portions of the electromagnetic spectrum used for monitoring land resources. The Resourcesat spacecraft, equipped with AWiFS, LISS-3, and LISS-4 sensors (NRSC 2011; Moorthi et al., 2008), employs an exclusive three-tier imaging concept, illustrated in Figure-1.3 (a), designed to provide medium-resolution multispectral remote sensing images. Table-1.1 outlines the specifications of the Resourcesat-2/2A cameras.

A crucial data processing stage on the ground entails geometric calibration to ascertain the exact location observed by a remote sensing pixel (Wang et al., 2014). The radiometric measurement of a pixel is meaningless without knowledge of its viewing location. Geometric calibration and correction present challenges due to factors such as resolution, adjusting the tilt to encompass particular regions of interest and accommodating variations in acquisition times (Radhadevi et al., 2016).

The georeferencing procedure tackles geometric distortion at the system level and establishes the correlation of pixel-look angles between the image and the ground. Geometric modeling

involves internal sensor parameters (such as focal length, detector angles, and payload alignment angles) along with ancillary data (ephemeris and attitude), thereby improving the accuracy of geometric location to a certain extent. Yet, attaining sub-pixel level geometric registration across multi-temporal images is challenging, and this stage is essential for producing Analysis Ready Data (ARD) products (Dwyer et al., 2018). The availability of such accurately geo-referenced data products is essential for users applying scientific studies like object classification, change detection and time series analysis. Figure-1.4 shows multi-temporal image misregistration at different feature landscapes. It is found that agricultural land and water bodies at two different imaging acquisitions are not aligned. This makes it difficult to assess the crop pattern and perform wetland monitoring at pre-monsoon and post-monsoon seasons utilizing satellite remote sensing images across different time periods.

Table-1.1. Resourcesat Sensor Specifications

<b>Sensor</b>	<b>Wavelength (in <math>\mu\text{m}</math>)</b>	<b>Spatial Resolution (in m)</b>	<b>Swath (in km)</b>
LISS-4	0.52-0.59 (Green)	5.8	70
	0.62-0.68 (Red)		
	0.77-0.86 (NIR)		
LISS-3	0.52-0.59 (Green)	23.5	141
	0.62-0.68 (Red)		
	0.77-0.86 (NIR)		
	1.55-1.70 (SWIR)		
AWiFS	0.52-0.59 (Green)	56.0 at nadir	740
	0.62-0.68 (Red)		
	0.77-0.86 (NIR)		
	1.55-1.70 (SWIR)		

Satellite= Resourcesat-2A, Sensor = LISS-3



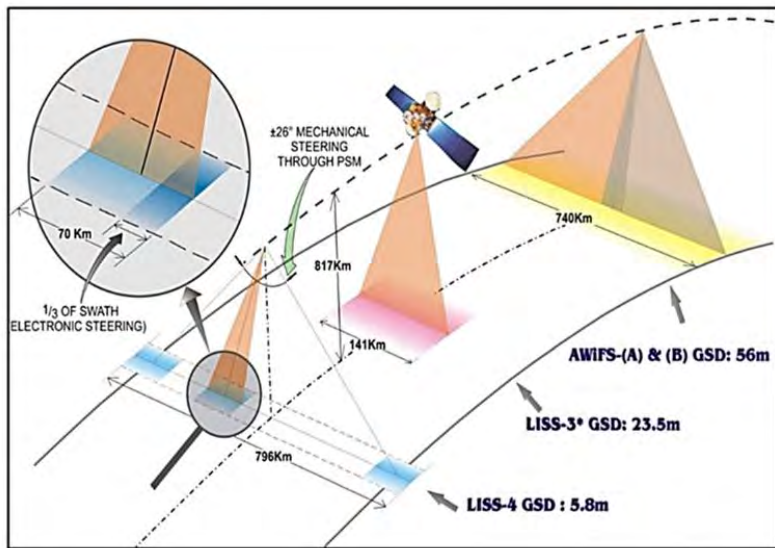
Date of Imaging: 25-Dec-2016



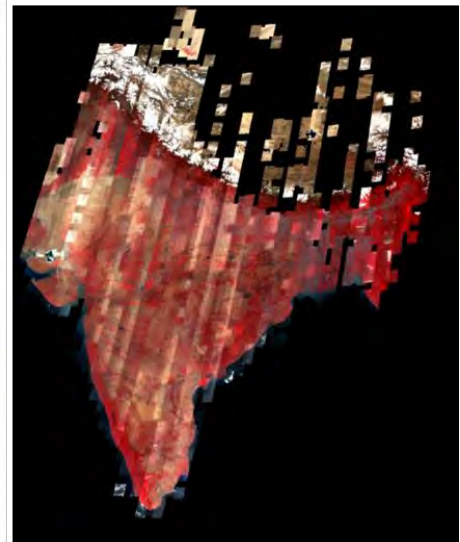
Date of Imaging: 18-May-2017

Spectral Band = Red, Region = Ahmedabad, Gujarat, India

Figure-1.2. Multi-Temporal Remote Sensing Image Pair



(a)



(b)

Figure-1.3. (a) The three-tier imaging mechanism of Resourcesat.

(b) The reference layer of LISS-4 covering India and its adjacent areas.

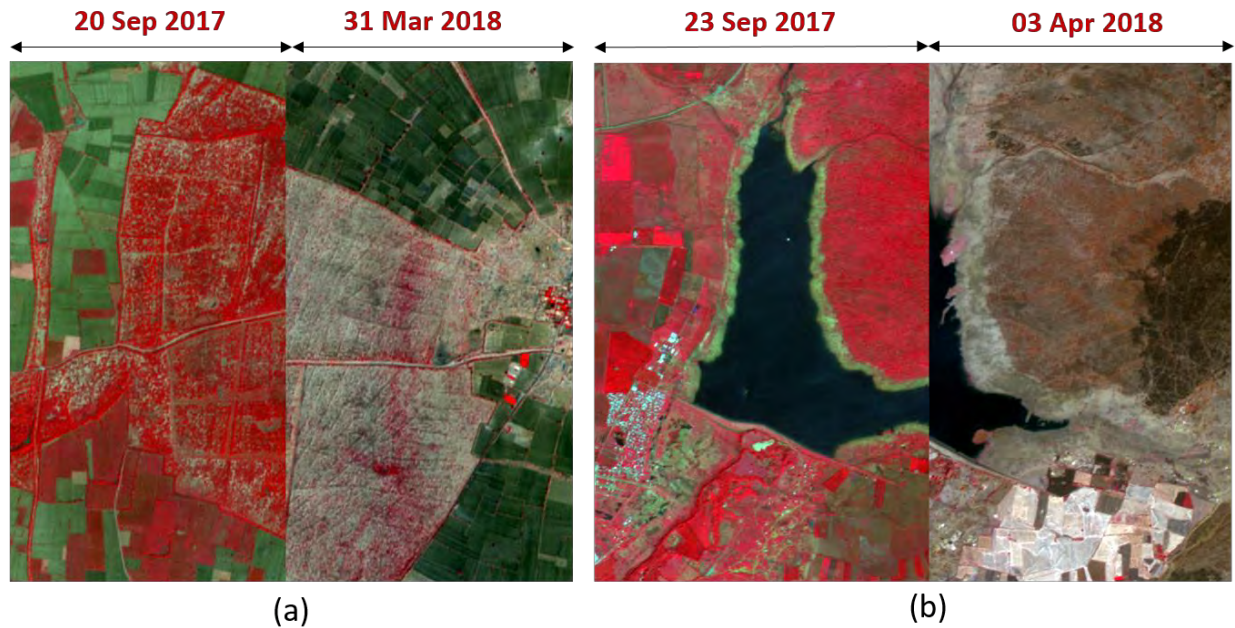


Figure-1.4. Misalignment of multi-temporal images over (a) agricultural land and (b) wetland areas

The necessity of multi-temporal co-registration is evident in various space-borne applications, ranging from monitoring snow/glaciers and flood impact assessment to observing variations in wetland conditions throughout various seasons and grasping the patterns of crop growth from planting to harvesting. Multi-satellite harmonization also demands co-registered data products to validate land surface measurements and construct a virtual constellation. In such scenarios, an effective automatic solution involves co-registering datasets with sub-pixel accuracy at equivalent spatial resolutions (Moigne et al., 2011; Bentoutou et al., 2005).

The land terrain of India has been effectively surveyed using Resourcesat-2/2A LISS-4 data through numerous acquisitions, resulting in the creation of a reference layer devoid of clouds. The geometry of the reference layer undergoes correction through bundle adjustment and is further refined with sub-meter precise ground control points to improve the absolute location accuracy of the stack (Dechoz et al., 2015; Misra et al., 2018). The resultant reference stack includes multispectral channels of LISS-4 are mapped using the Universal Transverse Mercator (UTM) map projection and sampled with an output spatial resolution of 5.0 meters. The information about the reference stack is stored in a dedicated database, enabling the retrieval of matching tiles according to the geographic coordinates of the input image. The reference stack is made uniform across all three Resourcesat sensors as specified in Table-1.1 and an extra procedure is employed



to adjust the resolution of the extracted reference set to match the native resolution of the input image being examined. Figure-1.3 (b) illustrates the coverage provided by the LISS-4 reference layer over the Indian land terrain and its surroundings, serving as the foundation for geospatial alignment within our processing framework.

### 1.1.1. Scanty Feature Island Image Registration

Registering images becomes notably more formidable when a scene lacks an abundance of features. Specifically, earth observation images capturing vast expanses of the deep ocean present a significant challenge due to minimal feature and texture information. Islands with scarce features, known as scanty feature islands refer to small, isolated land masses encircled by water, providing minimal features for the automatic registration of multi-satellite island images. Despite their limited features, these islands play a crucial role in different applications of earth observation, such as the mapping of coral reefs (Mumby et al., 2004), necessitating precise geometric mapping for advanced operations related to image processing.

Achieving co-registration of multi-spectral remote sensing images involves aligning them spatially with a reference image that possesses sufficiently accurate absolute location accuracy within a few meters, applicable to both high and medium-spatial resolution images (Misra et al., 2021a). Figure-1.5 illustrates Resourcesat LISS-4 multispectral images capturing ocean water with only a few islands, such as those in Lakshadweep, India, geo-located. Extracting accurately keypoint correspondences for automatic feature-based image registration becomes a complex challenge in remote sensing images with scanty features in the scene.



Figure-1.5. Lakshadweep Islands, predominantly situated in the deep ocean, are captured by the Indian Resourcesat LISS-4, along with the geographic coordinates of the scene center.

### 1.1.2. Band-to-Band Registration (BBR)

Correcting geometric misalignment among spectral bands in remote sensing images poses a significant challenge for earth observation applications. Band-to-Band misregistration occurs due to factors such as payload configuration, varying imaging times of wavelength bands, and considerable spacecraft perturbation (Jhan et al., 2016). Addressing this issue, Band-to-Band Registration (BBR) becomes a crucial step in data processing to rectify the spatial alignment of spectral bands obtained from space-borne imaging sensors (Moorthi et al., 2008). Feature-based remote sensing image registration offers advanced techniques to co-register multispectral bands with sub-pixel level accuracy (Misra et al., 2022a). Feature detection serves as the initial and crucial step in the BBR procedure, extracting stable and distinct control feature points. Figure-1.6 illustrates the misregistration among multispectral bands over man-made structures. This spatial misalignment leads to significant spectral distortion and blur in the False Color Composite (FCC) of multispectral images, formed by combining Green (G), Red (R), and Near Infrared (NIR), wavelength channels.

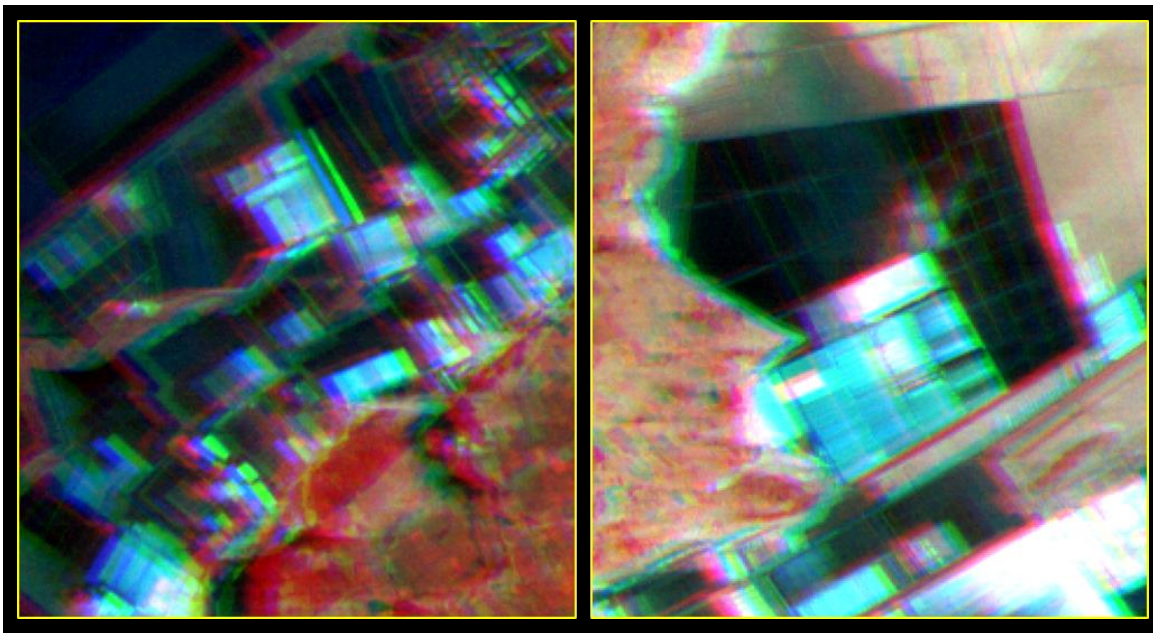


Figure-1.6. Band-to-Band Misregistration in Multispectral Remote Sensing False Color Composite Images (RGB Combination: Near Infrared, Red, Green)

### 1.1.3. Country-Level Image Mosaicking

Accurate image registration serves as a foundational step in creating country-level mosaic data products. Advancements in earth observation sensors now enable the monitoring of extensive

geographic areas through frequent time-lapse observations. Images obtained through multispectral remote sensing, capturing identical and adjacent areas at different acquisition instances, facilitate image mosaicking for studying regional and country-level processes related to monitoring of resources and the environment (Li et al., 2019).

The information obtained from sensors on Indian Resourcesat., namely Resourcesat-2 and Resourcesat-2A, provide a substantial wealth of multispectral image information with enhanced temporal resolution due to the coordinated imaging operations of these two satellites. The Resourcesat AWiFS sensor provides a spatial resolution of 56.0 meters with a temporal frequency of 5 days. In contrast, the LISS-3 sensor on the Resourcesat satellite surveys identical areas every 24 days with a spatial resolution of 24.0 meters. (Resourcesat-2 Data User Handbook, 2011). The data from both sensors are perfect for creating a Pan-India mosaic data product. The mosaic data that encompasses extensive areas enables the evaluation of crop proportions at the state and country levels, along with the assessment of regions through geophysical parameters like NDVI thresholding methods. (Saxena et al., 2021; Misra et al., 2016).

Furthermore, the mosaic data from LISS-3 enables the creation of an inventory of wetlands at a national level for both pre-monsoon and post-monsoon seasons. This aids in detecting changes (Asokan & Anitha, 2019) and accurately delineating wetlands (Guo et al., 2017). Mosaic data products at the national level provide varied space-based applications for conducting extensive scientific studies and analyses. Additionally, they play a crucial role in strategic decision-making and contribute to achieving sustainable development goals.

#### **1.1.4. Planetary Feature Extraction and Co-Registration**

Extracting features from planetary remote sensing images to register morphological structures is a challenging task due to limited texture information. The features obtained play a fundamental role, providing essential information for the following phases of image co-registration and analysis (Nevatia & Babu, 1980). Computer vision algorithms are utilized to extract crucial data from satellite images with multiple spectral channels, usually consisting of a minimum of three channels.

To create a natural-color image of a particular area on a planet through remote sensing, an orbital camera captures unadulterated color samples for each pixel using Red (R), Green (G), and Blue

(B) wavelength channels. While a three-sensor camera is capable of producing high-quality color images for various space-borne applications, its substantial size and weight pose challenges. Weight is a crucial constraint for designing payloads or instruments, particularly for interplanetary missions. This limitation necessitates exploring alternative solutions, such as a mono-sensor camera, now commonly found in smartphones. In a single-sensor camera, images are taken using a solitary sensor that is overlaid with a Color Filter Array (CFA), creating the Bayer Pattern mosaic arrangement. (Bayer, 1976). In this arrangement, every pixel element records just a single sample from the three color components. The resultant image with a single channel is subsequently transformed into a color image through the process of demosaicking (Lukac & Plataniotis, 2005). Demosaicking is a digital technique used to produce color images from images with a Bayer Pattern Color Filter Array (CFA) by interpolating the absent color components at each pixel. Figure-1.7 shows Mars's full disc demosaic processed image captured by an Indian Mars Color Camera (MCC). The prominent features such as Olympus Mons, Gale Crater, and Martian clouds are easily visible in the MCC image.

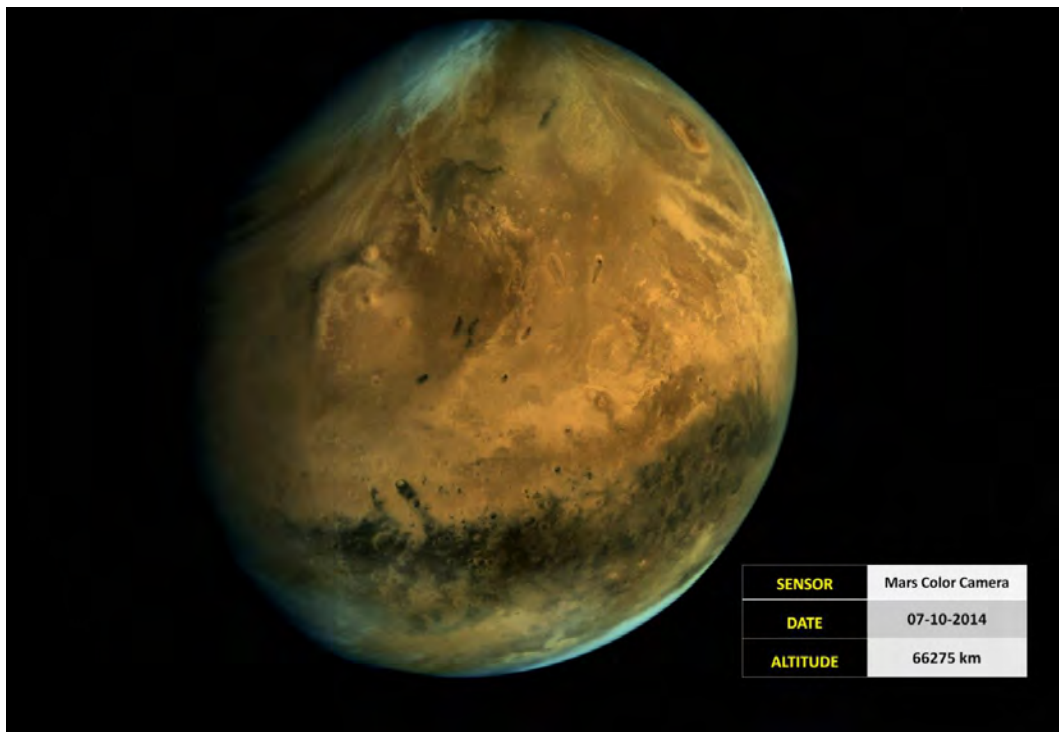


Figure-1.7. Mars Full Disc Image captured by Indian Mars Color Camera (MCC)

The Mars Color Camera (MCC) on board the Indian Mars Orbiter Mission (MOM) satellite, as discussed by Arya et al. (2015) and Arunan & Satish (2015) is a Bayer Pattern sensor capturing the Martian surface. Due to the strongly elliptical orbit of MOM, MCC captures Mars at varying spatial resolutions. Operating within the visible segment of the electromagnetic spectrum, MCC produces an RGB Bayer Pattern arrangement from a detector array frame consisting of 2048\*2048 elements, each with a pixel pitch of 5.5 $\mu$ . Serving as the "eye" of MOM, MCC records a complete disc image of Mars with a spatial resolution of 4.0 km when at apoapsis. Moreover, it can offer a more detailed observation of specific areo-morphological features with an increased spatial resolution of 15.0 m when at periapsis.

The MCC Long Term Archive (LTA) adheres completely to the standards of the Planetary Data System (PDS) for its collection of imagery, containing comprehensive information about each object. This repository is accessible to planetary scientists in the public domain (Moorthi et al., 2015). With over a thousand distinct MCC images, this collection offers a valuable resource for understanding recent Martian surface phenomena, including monitoring dust devils, observing alterations in the ice cap of Mars, and tracking the formation of new impact craters (Chevrier & Mathe, 2007).

However, extracting features directly from raw Bayer pattern planetary remote sensing images is a challenging task, particularly necessary for co-registering morphological structures acquired at different instances in time. Figure-1.8 shows the Pital impact Crater over Martian terrain captured by MCC. MCC raw Bayer pattern image looks like a grayscale image where the pixel is arranged in an RGGB pattern scheme at each sample location. The MCC demosaic image is also shown, a processed image where complete end-to-end radiometric processing takes place to create a composite MCC image with natural colors.

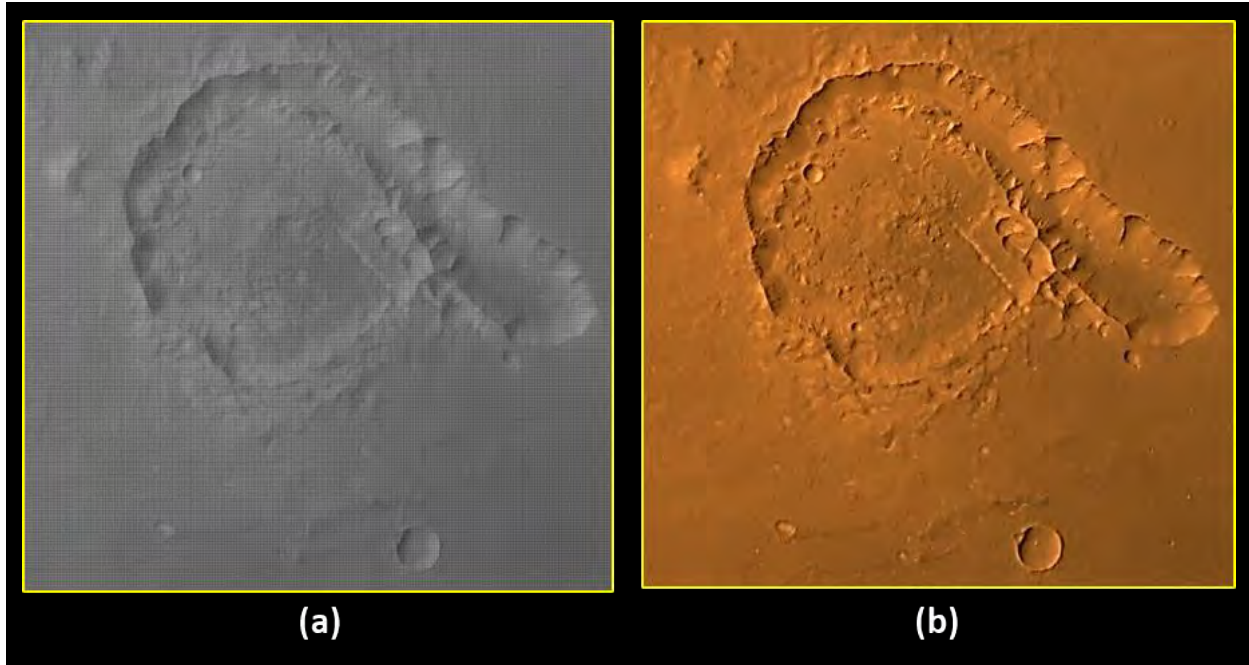


Figure-1.8. MCC captured Pital Crater (a) Raw Bayer Pattern Image (b) Radiometrically Processed Image

## 1.2. Remote Sensing Image Fusion

Remote sensing images hold immense potential for evaluating Earth's observational dynamics, observing land resources, tracking crop growth stages, city planning, researching glaciers, and analyzing time series data for vital environmental indicators (Hasan, 2016). Earth observation constellations involving multiple satellites generate extensive data archives for harmonized information extraction, contributing to sustainable environmental practices (Claverie et al., 2018). Nevertheless, it has been observed that data from a single sensor may be incomplete or lack the desired content. This limitation often arises because achieving higher spectral resolution frequently comes at sacrificing spatial resolution in favor of maintaining a specific Signal to Noise Ratio (SNR). Additionally, constraints related to onboard storage and bandwidth transmission pose challenges to obtaining optimal sensing information from space (Thomas et al., 2008).

To address these issues, Image fusion or pan-sharpening stands out as a widely used technique. This method entails combining multispectral remote sensing images with lower spatial resolution alongside panchromatic data with higher spatial resolution. The heightened temporal frequency among multi-sensor images encompassing the identical area presents a chance to improve

multispectral images by utilizing high spatial radiance data. Generally, bands with increased spectral resolution are paired with intricate spatial details from the panchromatic band to create a fused image that aptly captures both spatial and spectral characteristics. Yet, a significant obstacle in the image fusion process is to guarantee that improving spatial qualities does not distort the spectral attributes of the multispectral data (Ranchin & Wald, 2000). The main objective of the fusion process is to generate an improved image containing a comprehensive array of information, merging the best features of both the multispectral signature for identifying features and the spatial detail for detecting objects and textures.

In recent decades, various techniques for merging images have been formulated, investigated, and employed in multi-sensor remote sensing datasets. Conventional methods, such as Component Substitution (CS), have been utilized to tackle the challenge of image fusion (Pohl & Genderen, 1998). Two widely used fusion methods based on Component Substitution (CS) encompass the Brovey Transform (BT) (Gillespie et al., 1987) and the Intensity Hue Saturation (IHS) Transform (Carper et al., 1990). The Multi-Resolution Analysis (MRA) method improves spatial details by decomposing the finer panchromatic band into multiple resolutions for fusion (Gemine et al., 2014). High Pass Filtering (HPF) (Chavez et al., 1991) and Smoothing Filter Intensity Modulation (SFIM) (Liu, 2000) belong to the class of Multi-Resolution Analysis (MRA)-based techniques for image fusion. Nevertheless, while both Component Substitution (CS) and Multi-Resolution Analysis (MRA) methods are easier to execute and aesthetically pleasing, but they may encounter issues of spectral distortion in the resulting fused image. This problem becomes crucial when remote sensing applications depend on spectral patterns like lithology, soil, and vegetation.

Variational Optimization (VO) models have become more proficient in maintaining spectral features in fused images by regulating spatial enhancement (Ballester et al., 2006). Recently, Convolutional Neural Networks (CNN) (Zhong et al., 2016) and Generative Adversarial Networks (GAN) (Gastineau et al., 2021) have demonstrated enhanced fusion effectiveness with relatively minimal spectral distortion. Figure-1.9 shows the generic remote sensing image fusion procedure adopted for pan-sharpening multispectral images with panchromatic data. In general, multispectral images have low spatial resolution and high spectral resolution. Conversely, the panchromatic image band possesses superior spatial resolution but lower spectral resolution. Figure-1.10 shows panchromatic and multispectral image pairs for the remote sensing image fusion task. At the same map scale, the airstrip and surrounding region spatial features are demarked and visible in the

panchromatic image. In contrast, the corresponding features in the multispectral False Color Composite look blurry. So, the primary objective of image fusion is to attain optimal spatial and spectral resolution by amalgamating panchromatic images with multispectral data.

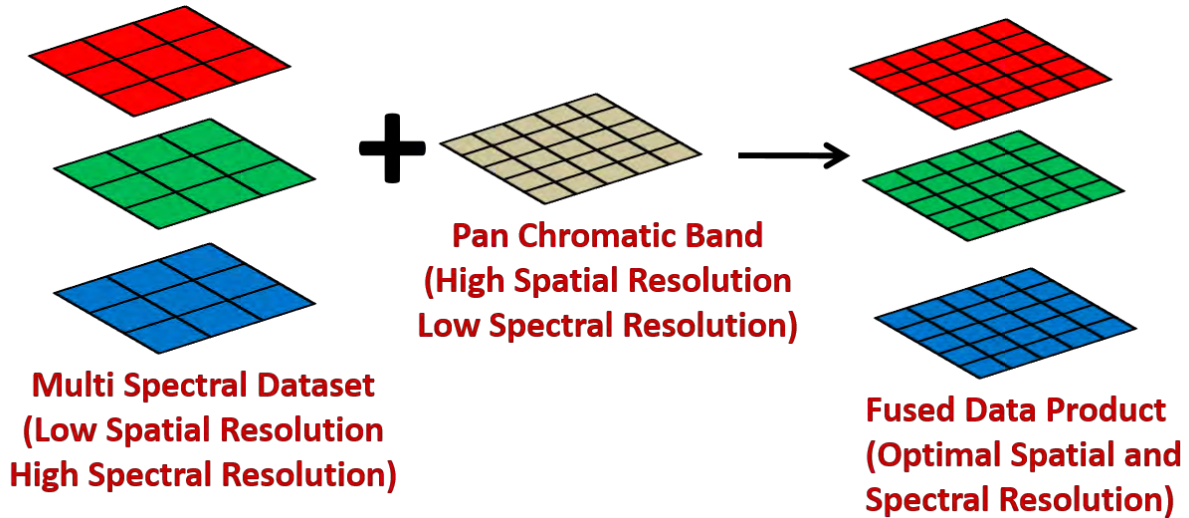


Figure-1.9. Remote Sensing Image Fusion Procedure



Figure-1.10. Panchromatic and Multispectral Image Pair for Image Fusion

### 1.2.1. Infrared-Visible Image Fusion

The challenge of image fusion becomes more pronounced when integrating multi-modal images to create a representative image, a crucial task in medical imaging (Du et al., 2016). In remote



sensing, one of the most suitable pairs for multi-modal image fusion consists of thermal infrared and multispectral visible images (Guo & Moore, 1998). Thermal infrared images stand out in identifying targets based on differences in thermal radiation information, while multispectral visible images retain excellent texture detail for classifying spectral classes in remote sensing imagery. With progress in sensor technology, thermal infrared cameras integrated on-board now deliver images with both high and medium spatial resolutions (Pearlman et al., 2020). Moreover, multispectral visible images are accessible at diverse spatial resolutions from various space agencies globally (Loveland & John, 2012; Drusch et al., 2012; Ranganath et al., 2010). This supplementary source of information provides comprehensive landscape details, requiring the creation of innovative techniques to integrate relatively high spatial resolution thermal infrared bands with corresponding low spatial resolution multispectral visible images. Figure-1.11 depicts the fusion of a thermal infrared image with relatively high spatial resolution and a corresponding multispectral visible image with lower spatial resolution.

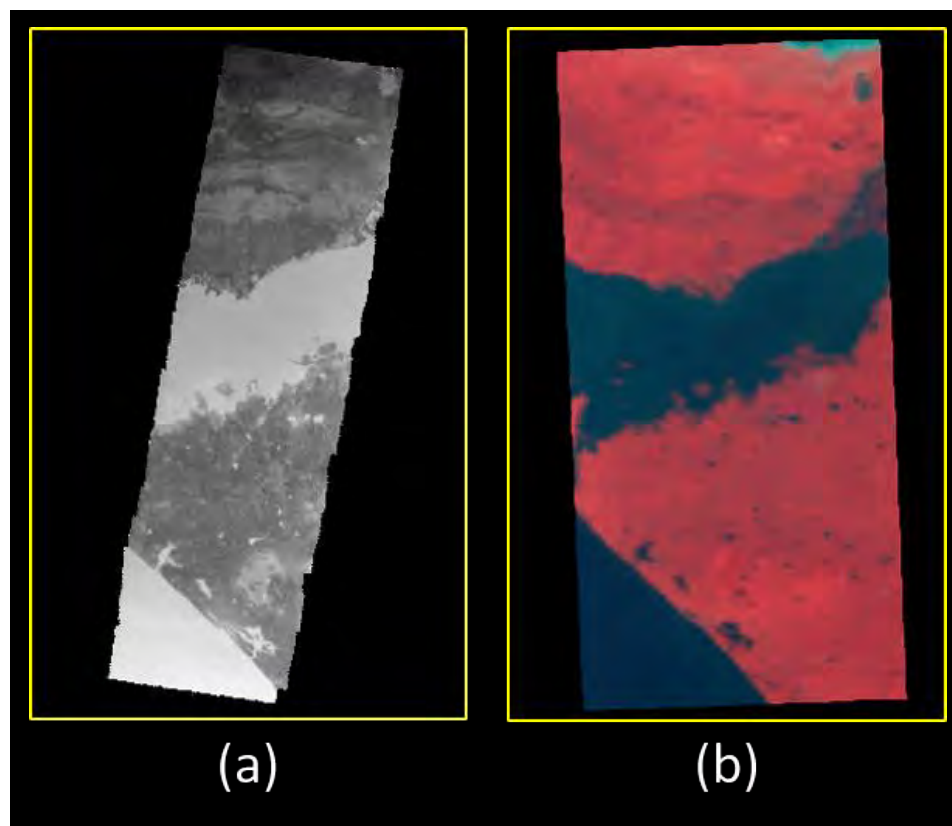


Figure-1.11. Multi-Modal Remote Sensing Image Pair: (a) Thermal Infrared Broadband Image with Relatively High Spatial Resolution, and (b) Multispectral Visible False Color Composite (FCC) Image with Low Spatial Resolution.

### **1.2.2. Venus Visible Image Processing and Multi-Modal Fusion**

Venus, the second planet from the Sun, has captivated human curiosity owing to its prominent visibility even in broad daylight. Numerous space science missions have been dedicated to studying Venus' surface characteristics, atmosphere, and historical evolution. One notable mission is NASA's Magellan, which commenced mapping Venus on September 15, 1990, providing remote sensing data on the planet's surface. The Magellan spacecraft's payloads compiled global maps of Venus, depicting tectonic and volcanic features, impact craters, and surface processes (Saunders et al., 1992).

The European spacecraft Venus Express was launched on November 5, 2005, and it successfully reached Venus on April 11, 2006. Equipped with seven instruments, Venus Express was dedicated to investigating the composition, structure, and dynamics of the atmosphere of Venus, as well as studying escape processes and the planet's interaction with the solar wind (Svedhem et al., 2007). Venus, characterized by extreme heat, numerous volcanoes, and thick clouds, prompted space agencies to conduct remote sensing missions with radar instruments to penetrate the planet's cloud cover and enhance surface feature interpretation.

In a significant development, The Parker Solar Probe, a NASA mission, performed close encounters with Venus in July 2020 and February 2021. The Wide-Field Imager for Parker Solar Probe (WISPR) captured the first visible night images of Venus from space (Wood et al., 2022). This rare event revealed a faint glow from Venus' surface, showcasing distinct features such as plains and plateaus, providing new insights into the study of Venus' planetary surface.

The visible raw image obtained over Venus using remote sensing instruments, as demonstrated in Figure-1.12, is a valuable yet challenging dataset. It exhibits different types of noise, and the morphological features appear faint, making it unsuitable for direct scientific image analysis. To enhance scientific interpretation and analysis, further processing of the visible raw image is necessary.

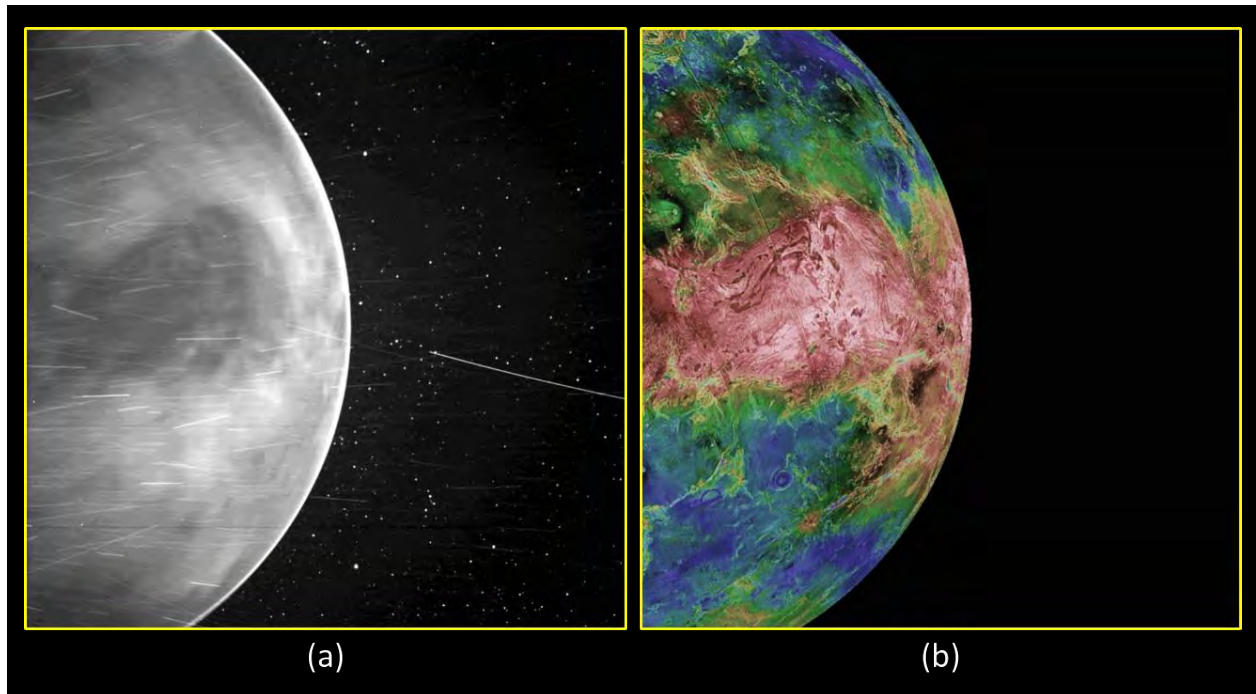


Figure-1.12. Venus' Multi-modal Image Pair (a) Visible Raw Image (b) Radar Topography Image

### 1.3. Multi-Temporal Change Detection

Examining remote sensing images captured over multiple decades is essential for investigating slow alterations in surface features that occur over extended periods. Earth observation satellites play a vital role in monitoring these changes, enabling the identification of urban sprawl, deforested areas, the location of new infrastructure, alterations in agricultural land, and the distribution of wetlands over inland water bodies. These observations play a crucial role in diverse space-based applications focused on safeguarding the habitability of planet Earth (Singh, 1989). The information derived from these changes serves as a vital input for land and urban planning, facilitating informed policy-level decisions for extensive geographic regions.

In this context, the Indian Remote Sensing (IRS) satellites, spanning from the launch of IRS-1A in 1988 to the recent Resourcesat-2A launched in 2016, boast extensive data archives (Kumar & Samudraiah, 2021; Kasturirangan et al., 1996). The Landsat satellite series systematically covers the entire globe, providing extensive image archives suitable for the analysis of time series data (Markham et al., 2004; Roy et al., 2014). The successful identification of historical changes depends on utilizing multi-temporal remote sensing data acquired over an extended period of

several decades. A significant difficulty in automated change detection is achieving sub-pixel level alignment of time-lapse images. The core challenge involves creating invariant features for image registration (Zitova & Flusser, 2003) to guarantee accurate co-registration of pixel data.

To achieve efficient change detection, it is crucial to perform radiometric normalization (Yuan & Christopher, 1996) on multispectral images taken at different times, potentially under diverse conditions. For radiometric normalization, correction of data necessitates knowledge of atmospheric parameters during the imaging process (Canty et al., 2004). However, extracting such atmospheric profiles and information from the archive for historical datasets proves challenging. Hence, the change detection technique should be capable of managing such diverse data and demonstrating resistance to linear and affine scaling. This eliminates the necessity for radiometric normalization as a preprocessing step in the data processing pipeline. Figure-13 shows historical IRS-1A and IRS-1C remote sensing images over the Gandhinagar and Ahmedabad regions that are almost three decades older. The recent acquisition of IRS images and legacy remote sensing data shown in Figure-1.13 can be used to find urban changes in this long timeline.

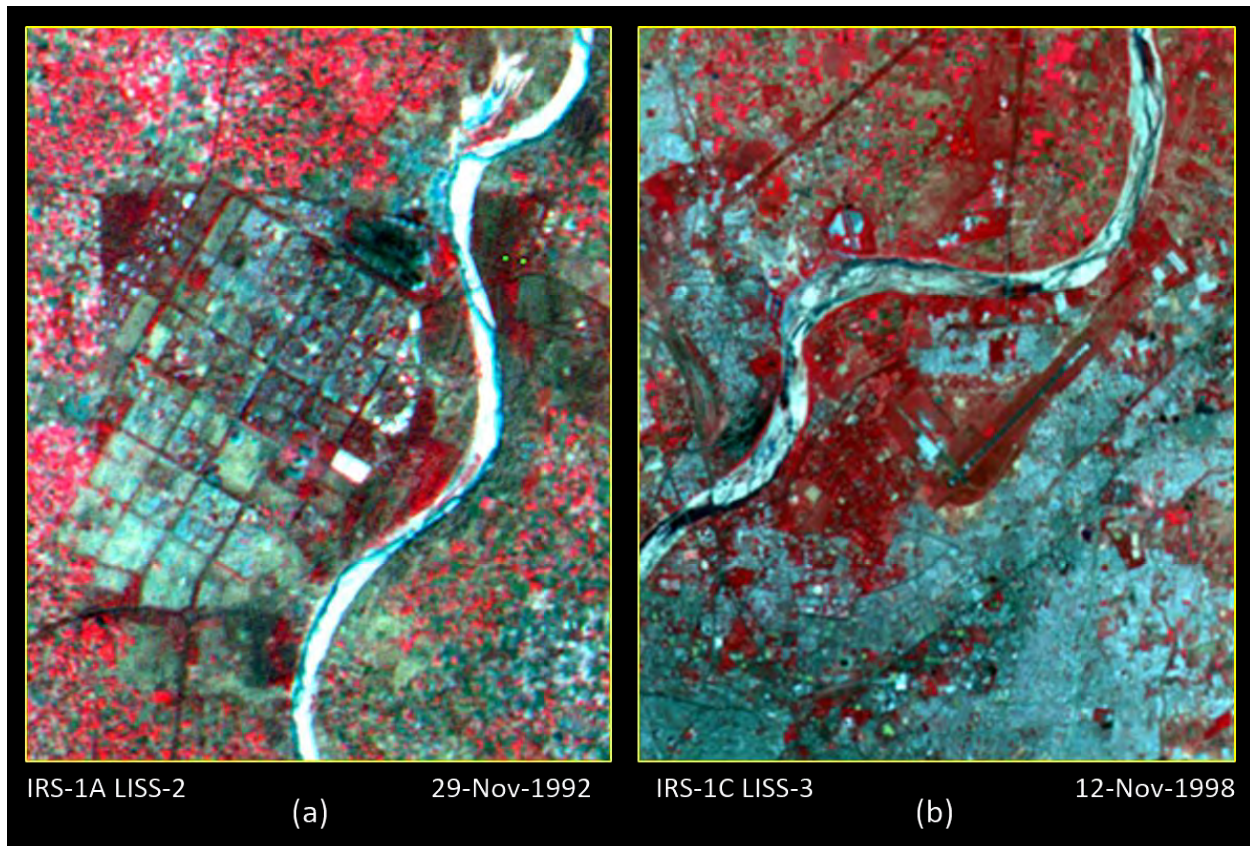


Figure-1.13. Multispectral Images from Indian Remote Sensing (IRS) satellite for Historical Change Detection (a) Gandhinagar Region (b) Ahmedabad Airport and Surrounding

### 1.3.1. Change Detection over Mars Terrain

The scope of change detection investigations has broadened to include Earth and other planets in our solar system, facilitated by the accessibility of extensive archives of remote sensing images. Mars, in particular, stands out as a dynamic and active planet continuously undergoing various geological events that reshape its surface (Mutch et al., 1976). Multiple Mars missions, including those led by Chicarro et al. (2004) and Malin & Edgett (2001), have delved into understanding areo-morphological processes occurring over decades, investigating change detection, studying dust devils and clouds in different locations, conducted crater counting, and investigated the climate and geological history of Mars. Detecting changes on Mars provides vital information about alterations in the structure of Martian mountains, the existence of outflow channels, and the creation of new impact craters.

The Mars Orbiter Mission (MOM) by the Indian Space Research Organisation (ISRO) successfully entered Mars orbit on September 24, 2014. The Mars Color Camera (MCC) onboard MOM is an RGB Bayer camera that operates in the visible range (0.4 to 0.7  $\mu\text{m}$ ). It captures images of Mars' surface with different spatial resolutions (Arya et al., 2015). Having transmitted more than a thousand images of Mars, the Mars Color Camera (MCC) plays a crucial role in detecting changes on the planet's surface (Moorthi et al., 2015). Figure-1.14 displays Mars remote sensing images captured by MCC over the Valles Marineris region, a vast tectonic crack along the equator, and the Mangala Valles region, illustrating a flow pattern featuring a channel bar in the MCC image. The MCC Bayer data in its raw form undergoes de-mosaicking to reconstruct a complete color image. Additionally, areographic coordinates are computed using orbit and attitude information to determine the imaging location on Mars. MCC images also undergo topographic correction to normalize radiance measures before scientific analysis, enhancing feature demarcation across different topographical slopes (Misra et al., 2015).

Globally, MCC captures full disc images in a perspective view, which are then employed for analysis. These images undergo geometric correction, including map projection, and an image registration procedure during mosaicking eliminates relative geometric differences between images. A seamless Mars full disc canvas, presented in Figure-1.15, was created by adjusting color differences between eight MCC full disc images obtained during December 2015 and January 2016.

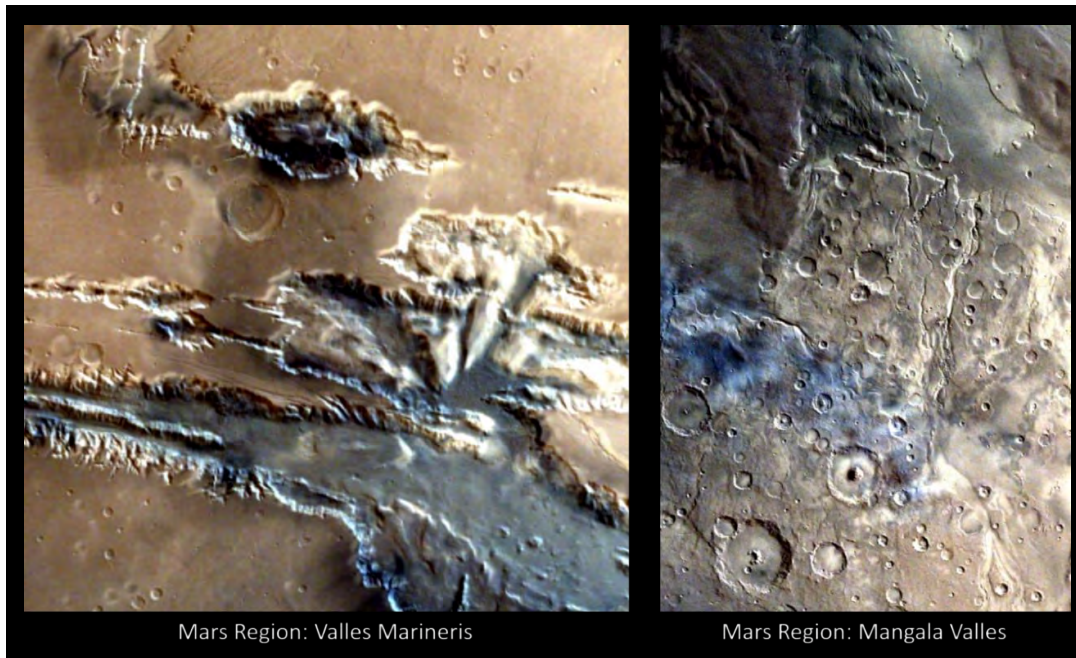


Figure-1.14. Indian Mars Color Camera processed images captured different Mars Terrain

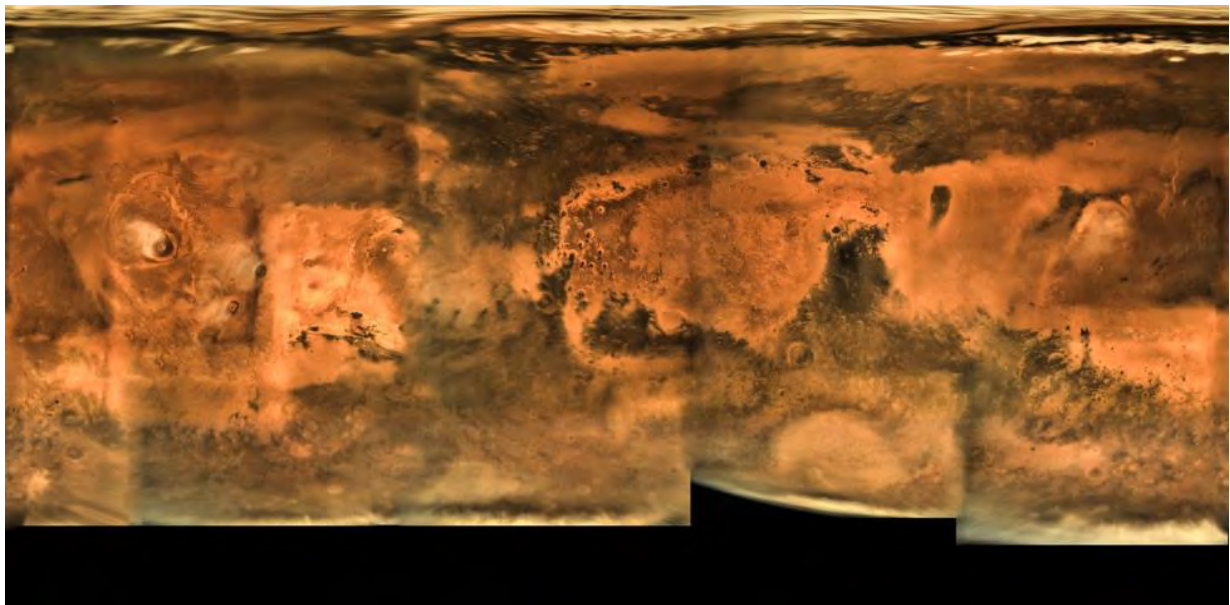


Figure-1.15. Mars Global Mosaic using recent MCC Full Disc Images

The pixel resolution of the MCC was consistently adjusted to 4 km pixels in the Mars Global mosaic image. The relevant historical images of the Viking mission from NASA (Klein et al., 1976), with a period difference of more than 30 years compared to ISRO's MCC images, aid in recognizing various surface changes on Mars. Figure-1.16 shows a Viking color merged Mars mosaic portraying Martian mons, impact craters, and valleys, which was imaged a few decades earlier.

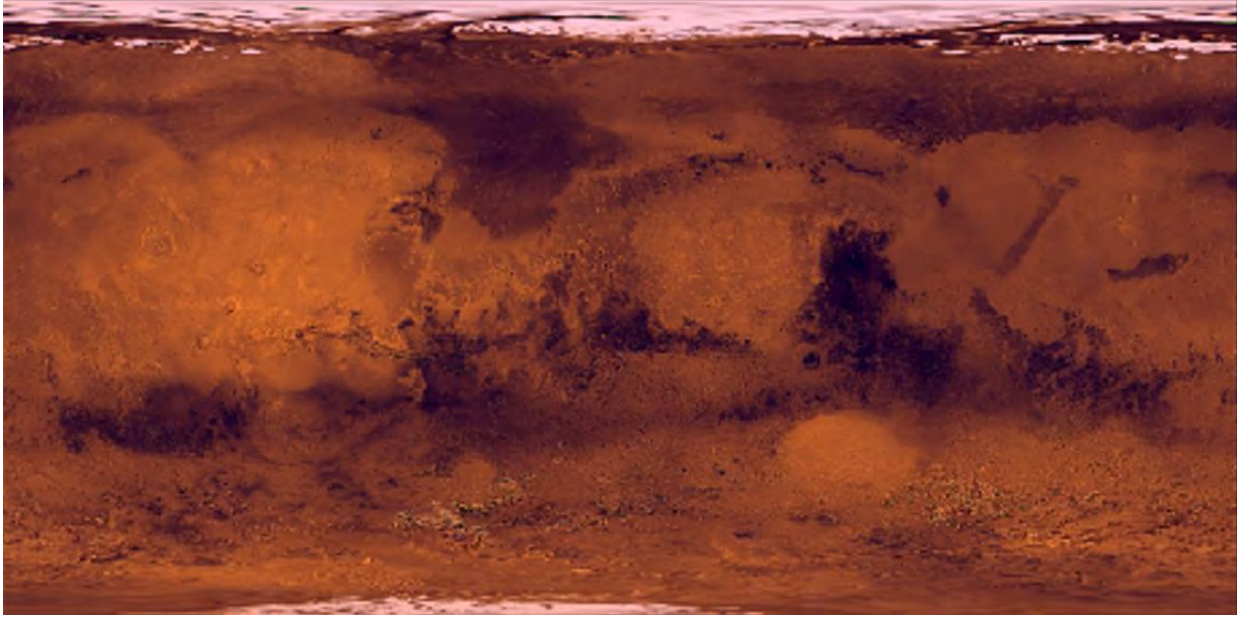


Figure-1.16. Viking Merged Color Mosaic showing Global Mars Landscape

#### **1.4. Research Agenda**

It has been observed that multi-temporal sub-pixel remote sensing image registration is a challenging problem because of changes in features that happened over a span of time. Stable key point detection and putative key point correspondences are challenging to achieve using current feature based image registration techniques. In addition, for planetary images, multitemporal co-registration is more challenging because of the absence of texture details, poor contrast, and uneven illumination characteristics. It is observed that the keypoints match set contains spurious matched points. It becomes critical to eliminate these spurious points to estimate transformation parameters in image co-registration effectively. To the best of our understanding, our research endeavors to create a novel multi-satellite image registration method specifically tailored for areas with sparse features, marking the initial exploration in this field. Attaining sub-pixel precision in registration accuracy, enhancing the geometric location accuracy across diverse datasets, introduces a new set of challenges in the realm of processing of remote sensing images.

Remote Sensing Image Fusion represents a challenging task focused on generating precise and high-quality data that incorporates both the multispectral information, aiding in object identification, and the spatial details crucial for object localization. The main goal of a proficient image fusion technique is to improve the spatial details in the fused images while minimizing radiometric distortion across diverse surface features, as assessed by low-resolution multispectral

images. It is still an inquisitive area of research where we need to combine image processing operations with machine learning models to generate a merged product that is qualitatively and quantitatively superior to state-of-the-art techniques.

To the best of our understanding based on current literature, there isn't an automated processing pipeline capable of efficiently handling multi-decadal time-lapse optical remote sensing images for the purpose of managing data pre-processing and generating a change detection map to directly interpret notable regional changes. The development of effective and accurate change detection techniques utilizing computer vision algorithms continues to be a persistent challenge in the realm of remote sensing data analysis.

## **1.5. Organisation of the Thesis**

Chapter-1 briefly introduces feature-based image registration, the need for sub-pixel registration accuracy, and earth observation datasets specification used for the experiment. The need for remote sensing image fusion is explained using representative panchromatic and multispectral images. The different types of image fusion methods are presented in brief. The applications of remote sensing, such as the detection of changes over multiple time periods are introduced in this chapter. The change detection scenarios for the Earth and other planets, such as Mars, are explained, and remote sensing imagery available for detecting surface change at different terrains is presented concisely.

Chapter-2 covers an exhaustive literature survey of feature based image registration techniques, Band-to-Band registration approaches, remote sensing image fusion methodologies, and multi-temporal change detection using remote sensing imagery over Earth and Mars. The problem definition is specified, and the research objective is clearly stated in this chapter.

Chapter-3 offers an overview of the relevant theory related to the methods used in feature-based remote sensing image registration for feature detection, description, and outlier removal. The chapter also focuses on remote sensing image fusion theory and its different types. The components and methods for change detection are described in the chapter. In addition, the overall methodology for remote sensing image processing workflow is illustrated.

Chapter-4 introduces the process of automatic remote sensing image registration through multi-stage feature detection and an enhanced outlier detection model. The focus of this chapter is on addressing the geo-registration challenges posed by images of areas with sparse features, particularly those obtained by remote sensing satellites with high and medium spatial resolutions



above deep oceanic expanses. It outlines the Band-to-Band Registration (BBR) technique designed for multispectral remote sensing images, offering a comparison with current state-of-the-art image registration methods across diverse land terrains. Furthermore, the chapter details a methodology developed for the direct extraction of features and co-registration of raw Bayer pattern planetary remote sensing images.

Chapter-5 demonstrates an innovative system for country-level image mosaicking utilizing optical remote sensing imagery. The chapter emphasizes a unique processing workflow for image mosaicking, where geo-referenced image strips with significant overlap are input to generate mosaic data products at the country level. The effectiveness of the mosaic system has been assessed using optical remote sensing images of medium resolution acquired by Resourcesat over the Indian subcontinent. At common regions, radiometric performance and geometric fidelity after correction are showcased.

Chapter-6 introduces the Spectra Preserving Remote Sensing Image Fusion Framework. This chapter introduces a well-balanced and resilient image fusion approach that addresses the limitations of current techniques, aiming to better retain the spectral attributes of multispectral images in the fused output. The chapter outlines an innovative methodology for fusing Infrared-Visible remote sensing images, enhancing the spatial resolution of coarser multispectral images through the incorporation of thermal infrared images with a relatively higher spatial resolution. In addition, the chapter also presents Venus visible image processing and multi-modal image fusion workflow.

Chapter-7 illustrates a framework for change detection spanning several decades using remote sensing images from both Earth and Mars. The chapter presents a method for automatic historical change detection using multi-decadal time-lapse remote sensing images. The effectiveness of this approach is assessed using images from Indian Remote Sensing (IRS) and Landsat over a period of three decades on Earth. Moreover, the suggested approach is tailored to identify changes in satellite images on Mars, one of the neighboring planets in our solar system.

Chapter-8 presents the conclusions of the work. After that, the specific contribution of the research work, suggestions for future research, and the list of references are appended.

Finally, a list of all the research papers published, accepted, or communicated for publication from this research work is given. This list follows a brief biography of the candidate (research scholar) and the supervisors.

## 2. LITERATURE REVIEW

### 2.1. Introduction

Image registration plays a crucial role in various fields such as medical image analysis (Oliveira et al., 2014), robotics (Zheng et al., 1993), and processing of remote sensing images (Yang et al., 2017). In our particular scenario, we are focused on remote sensing images, which exhibit numerous distinctive features. We utilize sophisticated techniques for detecting features to estimate the transformation between multispectral images. The registration of remote sensing images is a vital preprocessing step for the analysis of earth observation data. The georeferencing model is observed to correct systematic geometric distortions in the image. However, achieving sub-pixel geometric accuracy across multi-temporal scenes remains a challenge. Feature-based image registration emerges as a prominent method for overlaying multi-temporal remote sensing images and aligning datasets with sub-pixel level geometric precision. Additionally, thorough exploration of outlier removal algorithms is essential to generate reliable critical point correspondences and achieve accurate transformation parameter estimation.

The utilization of remotely sensed satellite imagery has gained significant importance in contemporary earth observation applications, encompassing tasks like resource monitoring, vegetation profiling, and snow/glacier studies. This reliance on satellite imagery is supported by diverse datasets obtained from various remote sensing satellites. However, in numerous space applications, the information from individual sensors may be inadequate or lacking in content. This limitation arises because achieving increased spectral resolution frequently results in reduced spatial resolution, all within a specific signal-to-noise ratio framework.

To address this issue, lower resolution multispectral channels are frequently supplemented by a single band with higher resolution, creating a "representation" that blends the desired spectral and spatial resolution. This process is referred to as remote sensing image fusion. One specific type of image fusion is pan-sharpening, where a high-resolution multispectral image is created by combining a panchromatic image with its corresponding multispectral counterpart of lower resolution. The main goal of pan-sharpening is to preserve the spectral features of multispectral images while improving the spatial contextual details derived from the panchromatic data.

Another form of image fusion is multi-modal fusion, such as Infrared-Visible image fusion, which involves creating a single hybrid representative image from two distinct modality satellite images. This approach enables the extraction of comprehensive information from remotely sensed datasets, facilitating the achievement of optimal spatial and spectral resolution.

The utilization of earth observation satellites to monitor gradual changes in land cover is essential for identifying urban sprawl, deforestation zones, new man-made constructions, alterations in regions of agricultural land and the arrangement of wetlands across inland water bodies. These space-borne applications contribute to the goal of creating a habitable planet Earth. Indian Remote Sensing (IRS) satellites, starting from the launch of IRS-1A in 1988 to the recent Resourcesat-2A (RS-2A) launched in 2016, possess extensive data archives. The multi-temporal IRS data spanning three decades serves as crucial input for land and urban planning, aiding in the formulation of accurate policy decisions for large geographic areas.

Change detection techniques in remote sensing can identify significant changes over time, producing comprehensive change detection maps. Notably, remote sensing datasets are now available for other planets within our solar system, including Mars. Numerous remote sensing missions over Mars have captured the Martian surface at various intervals, providing rich data to analyze changes in surface features over time. The Indian Mars Color Camera (MCC) onboard the Mangalyaan satellite, operational for over six years, has captured Mars surface images at different spatial resolutions. These extensive data archives can be utilized for planetary change detection studies. The developed change detection methodology should be robust enough to identify surface changes on both Earth and Mars.

## **2.2. Review of Feature Based Image Registration Methods**

To address the challenge of sub-pixel image registration, techniques relying on features and intensity have been developed and applied to align pixels effectively (Zitova & Flusser, 2003; Brown & Lowe, 2007). Image registration is a crucial step in analysis of medical images, robotics, and processing of remote sensing images (Maintz & Viergever, 1998; Guizar-Scicairos et al., 2008). In our context, we are working with optical remote sensing imagery, characterized by numerous distinct features. We use sophisticated methods for detecting features to determine the

transformation between multispectral images. The goal is to align multi-temporal images at the sub-pixel level through the image registration procedure.

The essential stages in image registration include identifying unique features, aligning vectors of these feature points, calculating transformation parameters, and reconstructing the input image to produce a co-registered pair of images. Feature detection entails extracting information from images and establishing correspondences between input and reference images. However, detecting and matching features in remote sensing images present challenges because of variations in imaging time, seasonal changes, and capturing from different viewing angles. The shift in datasets spanning multiple decades brings about additional complexities, such as changes in surface feature textures and difficulty in extracting invariant stable feature sets for matching.

For multi-sensor co-registration, feature shape context is extracted from airborne images (Huang & Li, 2010). Detection based on feature points, employing the construction of image pyramids, has been devised to register Landsat-8 and Sentinel-2 data products. (Yan et al., 2016). This method has also been modified to rectify multi-temporal misalignment in Sentinel-2A imagery (Yan et al., 2018). Given the higher spatial resolution with dynamic geometric location errors throughout the image and an increased time gap in acquiring image pairs, matching features in multi-temporal images introduces a new dimension. This highlights the need to enhance algorithms for detecting features and optimizing processes to attain accuracy at the sub-pixel level.

A feature in an image is defined as a unique pixel or a cluster of pixels. In remote sensing images, examples of unique features include a road cross-section, a building corner, or the edge of an airstrip. These features are generally stable and undergo minimal changes over time. Practical algorithms like Canny and Sobel edge detectors are commonly used to detect edge features in images captured through remote sensing, delineating object boundaries in the image (Ding & Goshtasby, 2001; Kittler, 1998; Davis, 1975). While edge detectors are valuable, corners are often preferred for satellite image spatial alignment. The Harris Corner measure has become widely recognized in recent decades as an effective method for detecting features in image registration (Harris & Stephens, 1988; Misra et al., 2012a).

More recently, powerful feature detectors such as SIFT (Lowe et al., 2004), SURF (Bay et al., 2006), ORB (Rublee et al., 2011), and KAZE (Alcantarilla et al., 2012) have been employed on remote sensing images captured at different times to attain co-registration accuracies at the sub-pixel level (Goncalves et al., 2011; Misra et al., 2019a). Among these, SIFT provides stable

keypoints with strong descriptors for matching; however, it demands significant computational resources and time, particularly for larger image dimensions. ORB is more computationally efficient but encounters the difficulty of a higher number of outliers in matched feature points, potentially introducing errors in estimated transformation parameters (Ma et al., 2016). Techniques like RANSAC (Bolles & Fischler, 1981; Brown & Lowe, 2007; Skakun et al., 2017) can be employed to remove outliers and improve the accuracy of model parameters within a defined threshold.

Feature detection based on phase congruency is a popular method for extracting meaningful image features in remote sensing image registration (Ma et al., 2018). More recently, Convolutional Neural Networks (CNNs) are created to establish spatial connections and produce consistent matching points (Ma et al., 2019). A novel framework for deep learning image registration in multimodal remote sensing images employing two task-specific deep models (Quan et al., 2023). A global–local consistency network aimed at enhancing feature matching, mitigating outliers' influence across different transformation patterns, and ensuring stable neighborhood support for the similarity metric of feature points (Liu et al., 2023). A hybrid matching approach that leverages attention-enhanced structural features, amalgamating the strengths of handcrafted-based and learning-based feature methods (Ye et al., 2024). Addressing the limitations of current feature-based methods, a deep transformer-based network is devised to overcome the challenges associated with constructing a deep local feature matcher (Xie et al., 2024). However, achieving sub-pixel precision registration accuracy to enhance geometric location error for most datasets remains a persistent problem in the domain of remote sensing image processing.

### **2.3. Review of Band-to-Band Registration (BBR) Techniques**

BBR, or Band-to-Band Registration, refers to the alignment of different spectral bands or images obtained from remote sensors to ensure they share precise spatial coordinates. This spatial alignment is crucial for the accurate analysis and interpretation of remote sensing data. The literature review will delve into essential aspects and methodologies employed to address BBR in remote sensing imagery.

In certain scenarios, BBR errors arise when different spatial resolutions are gridded before aggregation, leading to inaccuracies in characterizing coarser resolution bands using finer ones. These artifacts introduce biases during the creation of multi-date composites and the application

of spectral compositing criteria. Consistency in remote sensing records relies on uniform sensing scenarios and compositing methods (Tan et al., 2006).

For the MiniMCA multispectral sensor used in Unmanned Aerial Systems (UAS) for remote sensing applications, a precise BBR method is proposed due to significant band misalignment. This method involves a modified projective transformation model and error corrections to accurately align all sensor bands (Jhan et al., 2016). An image correlation method is employed for calculating BBR for the Moderate-Resolution Imaging Spectroradiometer (MODIS), providing good registration accuracy (Yang et al., 2000).

Addressing BBR challenges for near-equatorial orbit multispectral sensors, a technique was developed involving the conversion of image bands to grayscale, image compression, and the generation of critical correspondences using Scale Invariant Feature Transform (SIFT) (Dibs et al., 2015). Accurate BBR is achievable by regularly acquiring lunar images to characterize spatial alignment (Wang et al., 2015).

The BBR methodology for MISR imagery incorporates satellite orbit geometry, the rotating ellipsoid Earth, and the view angle separation between different spectral bands into a mathematical model describing band-to-band line and sample parallaxes (Zong et al., 1996). Another BBR approach involves calculating the Normalized Mutual Information (NMI) between pairs of shifted image bands and identifying the required shift to reach the maximum NMI value (Tilton et al., 2016).

Various techniques, including the Fast Fourier Transform (FFT) method, correlation techniques, Lagrange polynomials, and detector calibration methods based on histogram matching, are employed to achieve high registration accuracy for spatial alignment in different remote sensing systems (Barker & Seiferth, 1996; Desachy et al., 1985).

For hyperspectral images, where traditional registration methods may struggle, it is proposed to use undiffracted beams to enable accurate measurement of image motion in different bands (Zhao et al., 2013). An approach for measuring the Effective Focal Length (EFL) and BBR of specific spectral bands in satellite-borne whiskbroom imaging sensors is proposed using on-orbit data, yielding promising results for cloud-free subset images with distinct land features (Tilton et al., 2019). In the case of HJ-1A/1B CCD images, a parallel BBR method is employed, utilizing Open Multi-Processing (OpenMP) and founded on tiny facet differential rectification (Pan et al., 2011).

The challenges associated with band misregistration in low-altitude multispectral images captured by Unmanned Aerial Vehicles (UAVs) equipped with multispectral cameras are widely acknowledged. These cameras, featuring multiple lenses for each spectral band, present difficulties in achieving accurate registration. To mitigate this issue, the research proposes an image-based registration method aimed at minimizing the impact of relief displacement on registration errors. The approach involves dividing the image into patches, selecting appropriate local windows, and conducting local matching within each window to ensure proper point distribution (Hassanpour et al., 2019).

The Modified Projective Transformation (MPT) is utilized to synchronize the diverse image geometries of a multi-lens system into a cohesive sensor geometry. Furthermore, a procedure for Robust and Adaptive Correction (RAC) is integrated to correct systematic errors and attain reasonably accurate Band-to-Band Registration (BBR) (Xie et al., 2011). Emphasizing utilization of MODIS lunar observations to characterize Band-to-Band Registration (BBR) in both across-track and along-track directions aims to enhance the overall spatial alignment of multispectral bands (Xiong et al., 2011). On-orbit calibrators are utilized to observe misregistration between spectral bands in remote sensing imagery, which can lead to inaccuracies in science data products when combining measurements from different spectral bands. To address this, MODIS employs the Spectro-Radiometric Calibration Assembly (SRCA) throughout its mission. On-orbit results indicate minimal misregistration for Terra MODIS and relatively significant misregistration for Aqua MODIS between bands on warm and cold Focal Plane Assemblies (FPAs) (Montgomery et al., 2000; Choi et al., 2019).

A hybrid framework is proposed to tackle BBR errors in remote sensing images, particularly those acquired by multispectral push broom spectrometers like the Sentinel-2 Multispectral Instrument (MSI) (Chen & Liu, 2021). The accuracy of inter-channel relative geometric alignment depends on the precision of look angle characterization for the spectral channels (Nain et al., 2019). With the increasing significance of satellite jitter affecting imagery products, especially with high-resolution satellites, a method is introduced to analyze the impact of satellite jitter on multispectral images. This approach provides a means of detecting and compensating for jitter to enhance BBR accuracy, even in the absence of attitude sensor data (Zhu et al., 2014).

To model BBR errors for remote sensing camera systems, a polynomial-based co-registration method is developed. The method utilizes a fourth-degree polynomial for Band-to-Band Registration (BBR) across the track direction and a fifth-degree polynomial along the track direction to reduce BBR error (Pan et al., 2022). Dense vector matching, relying on a non-centered cross-correlation technique, achieves subpixel accuracy in the image registration process. Every line vector and column vector of the reference band are cross-correlated with adjacent vectors in the target band. Subpixel misalignments are estimated by applying a cubic polynomial fit and identifying the maximum correlation (Boukerch et al., 2018). Additionally, a Phase Correlation (PC) method is introduced for registering multiple bands acquired sequentially at different wavelengths. This process corrects scaling, rotation, and translation errors in images captured by an airborne hyperspectral imaging system (Erives et al., 2005).

The requirement to align three multispectral bands of the LISS-4 camera on the IRS-P6 satellite arises from their physical separation and the time gap in acquisition. A method for co-registration is outlined using collinearity equations, trajectory fit, and georeferencing (Radhadevi et al., 2009). The morphological band registration technique utilizes the quantile matching method, emphasizing matching pixel correspondence through radiometric distribution rather than geometric relationships (Kim et al., 2020). The study investigates the impact of Band-to-Band Registration (BBR) errors on geophysical parameters like NDVI using simulated images generated from Landsat TM images (Hashimoto et al., 1999).

Consideration is given to band registration aspects of a hyperspectral imager based on tunable filters, along with the development of a thorough and effective method for registering bands in intricate 3D environments, such as forests. The approach first establishes the orientations of chosen reference bands and reconstructs the 3D scene using structure-from-motion and dense image-matching technologies (Honkavaara et al., 2017). Image alignment in remote sensing using deep learning is highlighted as capable of handling BBR errors by optimizing processing through a learned-mapped function (Wang et al., 2018). In a deep learning framework, semantic template-based matching is introduced to generate higher BBR accuracy (Li et al., 2021).

The increasing prevalence of compact, consumer-grade multispectral systems with multiple cameras, especially for integration onto limited spacecraft platforms, is noted (Shahbazi et al., 2019). This highlights a common challenge of misalignment between spectral bands when



processing data from these cameras, impacting data quality and necessitating automatic handling. Despite advancements, BBR remains a challenging problem in remote sensing image processing due to sensor variability, atmospheric effects, complex terrain, and seasonal changes.

## **2.4. Survey on Remote Sensing Image Fusion Methods**

Improving the spatial quality of remote sensing images through the fusion of panchromatic-multispectral data pairs is a complex task (Kaur et al., 2021). The use of Component Substitution (CS) with partial replacement is one approach that combines high-resolution panchromatic images with corresponding multispectral images, aiming to mitigate a degree of distortion in the spectral characteristics (Choi et al., 2010). The Band Dependent Spatial Detail (BDSD) technique has shown promising outcomes in enhancing multispectral images, characterized by a comparatively faster implementation procedure (Garzelli et al., 2007).

Multivariate regression, as an Adaptive CS (ACS) image fusion technique, is widely employed to acquire intensity components that not only enhance spatial quality but also improve spectral quality (Aiazzi et al., 2007). An adaptive Intensity Hue Saturation (AIHS) technique incorporates a multiscale guided filter approach to handle the panchromatic image, extracting improved spatial information (Yang et al., 2016). The amalgamation of Adaptive Principal Component Analysis (APCA) and Contourlet has been noted to minimize spectral distortion in comparison to traditional PCA pan-sharpening (Shah et al., 2008).

A fusion technique utilizing joint-guided image filtering combines static and dynamic filters, demonstrating effective edge-smoothing properties and resilience against artifacts (Zhang et al., 2021). Fuzzy logic-driven image fusion can integrate prominent edges from both panchromatic and multispectral images using local salience measures (Aiazzi et al., 2006). The application of a Generalized Laplacian Pyramid (GLP) with a multiscale approach tailored to the Modulation Transfer Function (MTF) in remote sensing image fusion has shown promising outcomes in spatial enhancement, while maintaining spectral information from coarser data (Vivone et al., 2017).

The utilization of High Pass Filter (HPF) modulation with robust regression modeling represents an alternative Multiresolution Analysis (MRA) technique that produces superior fusion outcomes compared to current spectral matching methods (Chen et al., 2015). Moreover, object-based image analysis can categorize essential feature targets in multispectral images, and spatial enhancement

through panchromatic channels can be specifically applied to those regions containing objects (Seal et al., 2020).

The robust statistical model, treating image fusion as a convex optimization challenge aimed at minimizing the least square error using dynamic gradient descent regularization, effectively preserves spectral intricacies from multispectral images while incorporating crisp edges from panchromatic images (Zhang et al., 2015). However, it has been noted that despite their efficacy, statistical image fusion techniques may encounter spectral distortions due to rigid assumptions that fail to precisely mirror the realities governed by the principles of remote sensing physics.

To tackle the issue of spectral consistency, Bayesian data fusion utilizes a modification of the weight regularization parameter. This adjustment aims to strike a balance between spatial enhancement contributions and spectral preservation (Fasbender et al., 2008). Bayesian methods are classified under the Variational Optimization (VO) fusion category, providing flexibility to achieve adjustable outcomes tailored to meet user requirements. The posterior probability model, based on rational assumptions, has the ability to produce a High-Quality Bayesian (HQB) pan-sharpened image (Wang et al., 2018).

Over the past few years, deep learning has demonstrated notable success in diverse fields such as computer vision, image processing, detection of image forgery (Gurnulu & Ozturk, 2022), and pattern recognition problems (Sezer & Altan, 2021a) (Sezer & Altan, 2021b). The application of Convolutional Neural Networks (CNN) in remote sensing image fusion has led to the creation of impressive fused images (Shao et al., 2018). Deep Convolutional Neural Networks (CNNs) exhibit the capability to independently extract significant features from both multispectral and panchromatic images, leading to fused images with minimized spectral distortion (Huang et al., 2020), (Ye et al., 2019). The utilization of CNNs for pan-sharpening improve fusion performance by utilizing multiple feature maps of nonlinear radiometric indices significantly (Masi et al., 2019). The design of deep network architectures with domain-specific knowledge proves to generalize well across various satellite images without requiring retraining (Yang et al., 2017). Residual Learning deep CNNs have been identified to enhance fusion accuracy in contrast to less sophisticated flat networks (Wei et al., 2017).

Generative Adversarial Learning also influences pan-sharpening in remote sensing, integrating an attention-based mechanism to discover optimal solutions (Liu et al., 2020). In the domain of machine learning or deep learning methodologies, the training frequently depends on a process of

resolution downscaling to produce labeled data for supervised training. This could lead to spatial irregularities, emerging as a significant issue in methods relying on learning. However, a recent development is the design of an unsupervised pan-sharpening framework known as Pan-GAN, which relies on something other than reference data during the training phase (Ma et al., 2020). The addition of a position-agnostic style loss function, trained in conjunction with an unsupervised network, improves fusion performance by acquiring the style information of input images. (Cheng et al., 2022).

Despite the significant progress, it has been observed that the primary obstacle for deep learning models is in creating training datasets with proper data curation, a job that can be cumbersome and time-consuming. The computational demands and the stringent necessities of the training phase of deep learning techniques may encounter difficulties in achieving near-real-time generation of fused data products. Nonetheless, image fusion or pan-sharpening remains an ongoing research challenge, and innovative technical contributions have the potential to enhance fusion performance further in the realm of remote sensing data analytics.

## **2.5. Review of Multi-Temporal Change Detection Techniques**

Numerous techniques for change detection are commonly applied to analyzing remote sensing images to detect changes in different feature targets (Ridd & Liu, 1998 ; Jianya et al., 2008). The Simple Image Difference, a conventional approach for identifying changes in multi-temporal images, is responsive to the radiometric attributes of the images and presents difficulties for scientific analysis (Bruzzone & Prieto, 2000). Principal Component Analysis (PCA) is a robust technique for reducing the dimensionality of datasets and finds application in remote sensing change detection, particularly effective with data that is radiometrically normalized (Deng et al., 2008; Fung & LeDrew, 1987). Nevertheless, it is observed that PCA does not maintain invariance to linear scaling and affine variation in datasets from remote sensing. Change Vector Analysis (CVA) provides insights into spectral changes in remote sensing images captured at various time points (Johnson & Kasischke, 1998). The Normalized Difference Vegetation Index (NDVI) as a geophysical parameter is efficient for detecting changes in regions with sparse vegetation (Pu et al., 2008). Slow Feature Analysis (SFA) extracts unchanging components from multispectral images, converting the data into a new feature space (Wu et al., 2013). A transformation based on the Chi Square method is employed on multi-date images to detect changes in dense forests and urban landscapes (Lu et al., 2004; Vázquez-Jiménez et al., 2017).

Recently, there has been a surge in the popularity of utilizing machine learning and deep learning techniques for change detection in remote sensing (Yang et al., 2019; Khelifi & Mignotte, 2020). The PCA filter treated as a convolution network (PCANet) generates a promising change detection map by utilizing representative neighborhood features (Gao et al., 2016). The integration of a deep network with pre-detection using Change Vector Analysis (CVA) for preparing the training sample set demonstrates favorable outcomes in the context of remote sensing change detection (Du et al., 2019). Utilizing transfer learning, a Convolutional Neural Network (CNN) produces feature difference maps for change detection and is implemented on diverse resolution remote sensing datasets (Zhang & Shi, 2020). The Dense attention-based Convolutional Neural Network (CNN) adeptly captures pixel-level change information from bi-temporal images (Peng et al., 2020; Chen & Shi, 2020). However, it suffers from time-intensive aspects of data curation, labeling, and the training phase for models that heavily rely on learning. Multivariate Alteration Detection (MAD), relying on canonical correlation analysis, is a widely used algorithm for identifying changes in multispectral satellite images captured at different times. It remains invariant to linear scaling and demonstrates reduced susceptibility to noise contamination in the image (Nielsen & Simpson, 1998). However, it is noted that MAD variates may also identify some undesired changes in the ultimate output of change detection.

An effective and precise change detection technique development using computer vision algorithms continues to be an unresolved challenge for the remote sensing community. Moreover, as far as our awareness extends, there is currently no automated processing pipeline that can handle the ingestion of multi-decadal time-lapse optical remote sensing images, address the data pre-processing stage, and produce a change detection map for straightforward comprehension and interpretation of noteworthy changes in the region.

## **2.6. Problem Definition**

**Research Gaps Emerged from the Literature Review:** It has been observed from the literature review that multi-decadal remote sensing image registration is a challenging problem because of changes in features over a long time. Stable vital point detection and putative critical point correspondences are difficult to achieve using current feature based image registration techniques. In addition, for planetary images, multitemporal co-registration is more challenging due to a lack of texture information, low contrast, and uneven illumination characteristics. It is observed that the key points match set contains spurious matched points, and it becomes critical to eliminate

these spurious points to estimate transformation parameters in image co-registration effectively. The pursuit of sub-pixel highly accurate registration precision to enhance the geometric location accuracy for a majority of datasets presents fresh challenges within the domain of remote sensing image processing.

The task of Remote Sensing Image Fusion is intricate, seeking to generate data of high quality and accuracy that incorporates both the multispectral information (for object identification) and spatial details (for object localization). The primary objective of an effective image fusion technique is to enhance the spatial information in fused images while reducing radiometric distortion across diverse surface features observed by low-resolution multispectral images. This remains an area of ongoing research, prompting the exploration of combining image processing operations with machine learning models to create a merged product that surpasses current state-of-the-art techniques in both qualitative and quantitative aspects.

Based on our understanding of existing literature, there is presently no automated processing pipeline capable of handling input multi-decadal time-lapse optical remote sensing images to tackle both the data pre-processing stage and the generation of a change detection map. Developing effective and accurate change detection techniques using computer vision and machine learning algorithms remains a continuous challenge in the field of remote sensing image processing.

**Research Gaps this work tries to improve:** The proposed research will result in a robust automatic image registration system for multi-temporal, multi-sensor, and multi-decadal remote sensing images. Novel techniques will be developed to merge multi-sensor images to generate fused data products with optimum spatial and spectral resolution. The system will autonomously identify changes in surface features in multi-decadal remote sensing images, covering the entire process from data pre-processing to the creation of change detection maps over both Earth and Mars.

**Problem Definition:** An exhaustive literature survey on feature-based remote sensing image registration techniques is carried out to target the aims mentioned above. The different types of remote sensing images, such as multi-temporal, multi-sensor, and multi-modal datasets, are chosen for the image registration. In addition, planetary images are also considered to evaluate image processing tasks, which generally have low texture and are taken at varying solar illumination conditions. The co-registered data pairs at the sub-pixel level are ready for the next level of image processing operations, that includes image fusion and change detection. Remote Sensing Image

fusion is challenging, and a thorough literature review is performed to understand the fusion performance. In addition, multi-decadal change detection is another problem in the remote sensing world. As per the literature review, no central end-to-end automation has yet been performed to generate a change detection map. To solve this problem in the remote sensing image processing domain, novel algorithms, and frameworks need to be developed that can align multi-temporal remote sensing images at the sub-pixel level and merge remote sensing datasets in an optimum manner to enhance spatial information. At the same time, spectral characteristics should be preserved and accurately identify changes using multi-decadal remote sensing images over Earth and Mars automatically.

## **2.7. Objectives of the Research**

The proposed research aims to build automatic image registration techniques for multi-temporal and multi-sensor remote sensing images. The co-registered pair acts as input to develop a robust image fusion method and design an enhanced change detection approach in the remote sensing image processing domain. The proposed work aim is achieved by:

- 1) Designing an automation-intensive system for image registration tasks to generate co-registered data products at the sub-pixel level and improve geometric location accuracy.
- 2) Proposing multi-stage hierarchical feature detection/extraction algorithm and spurious key points outlier removal methods to refine spatial alignments.
- 3) Advanced image processing technique development to achieve Band-to-Band Registration (BBR) in multispectral images.
- 4) Proposing novel spectra preserving image fusion framework to merge panchromatic remote sensing images with multispectral datasets.
- 5) Proposing and developing techniques to handle multi-decadal remote sensing images over Earth and Mars from the data pre-processing phase to the creation of change detection maps.

## **2.8. Summary**

This chapter presents the literature review of methods for remote sensing image registration based on features, recent remote sensing image fusion techniques, and change detection approaches using remote sensing images. The chapter briefly describes the current methods' limitations for different remote sensing image processing operations. The next chapter presents the theoretical concepts related to the research work.

### **3. RELATED THEORY**

The related theory of image registration, fusion, and change detection in remote sensing constitutes a cornerstone of modern spatial analysis, enabling the extraction of meaningful insights from satellite or aerial imagery. Image registration serves as the foundational process, aligning multiple images of the same geographic area to a common coordinate system. This alignment facilitates accurate comparison and integration of data for subsequent analysis. Image fusion, on the other hand, synthesizes information from diverse sources, enhancing the quality and interpretability of imagery by combining complementary details. Lastly, change detection techniques leverage registered and fused images to identify and characterize alterations in the Earth's surface over time, shedding light on dynamic phenomena such as urban growth, deforestation, and natural disasters. Understanding the theoretical underpinnings of these processes is essential for harnessing the full potential of remote sensing data in addressing contemporary challenges related to land use, environmental management, and disaster response.

#### **3.1. Image Registration Methods using Features**

Feature-based satellite image registration is a critical process in remote sensing and computer vision that involves aligning several images of a particular region captured at various moments or perspectives of various sensors. The spatial alignment allows for the comparison and analysis of these images, facilitating tasks that include image fusion, change detection, and monitoring of land cover changes. The theory behind feature based remote sensing image registration is described as follows:

*i. Definition of Features:* Features in this context refer to distinctive points or patterns in the images that are easily identified and matched between the images. Standard features include corners, edges, or other high-contrast points. These features are picked because they are relatively stable across different images.

*ii. Feature Extraction:* The initial processing step extracts distinct features from each image. Feature extraction involves identifying and describing the selected features invariant to changes in scale, rotation, and illumination. Feature extraction algorithms such as SIFT (Scale-Invariant Feature Transform) (Lowe, 1999), SURF (Speeded-Up Robust Features) (Bay et al., 2008), and ORB (Oriented FAST and Rotated BRIEF) (Rublee et al., 2011) are popular.

*iii. Feature Matching:* Once stable features are extracted, the following procedure is to find corresponding features between the image pair. Feature matching is typically done by comparing the feature descriptors obtained from the extracted features. Matching algorithms like the nearest neighbor or RANSAC (Random Sample Consensus) (Fischler & Bolles, 1981) can be used to establish correspondences.

*iv. Transformation Estimation:* With the matched feature correspondences, the aim is to estimate the geometric transformation (e.g., translation, rotation, scaling, and perspective transformation) that aligns one image with the other(s). Standard transformation models include affine transformations, projective transformations, and homographies.

*v. Transformation Application:* The estimated transformation is then applied to warp or resample one image onto the coordinate system of another image. This step ensures that the satellite images are correctly aligned for further analysis.

*vi. Registration Evaluation:* The registration quality is assessed using image quality metrics such as Root Mean Square Error (RMSE) and Correct Matching Ratio (CMR). These metrics help determine how well the registration process has aligned the images.

*vii. Refinement and Iteration:* Refinement steps may be applied to improve accuracy that depends on the initial registration quality. Iterative techniques like bundle adjustment or control points may enhance registration quality further.

*viii. Multi-Sensor and Multi-Temporal Considerations:* In remote sensing, images are acquired from different sensors with varying spatial and spectral characteristics or from different times. These differences may involve radiometric calibration, atmospheric correction, or temporal adjustment.

*ix. Applications:* Once the images are successfully registered, they are ready for space borne applications, which include change detection, image fusion, land cover classification, and environmental monitoring.



### 3.1.1. Harris Corner Detector

The effectiveness of the Harris Corner Detector relies on the auto-correlation function at the local level, which calculates signal variations locally by shifting patches with a small window in various orientations (Derpanis, 2004).

The Harris Corner Detector offers a mathematical algorithm to assess if a feature point exhibits a noteworthy change in all directions, warranting its classification as a distinct point (Yao et al., 2017). Recognition of pixel intensity values within a small window facilitates the identification, and subsequently, moving the window in any direction should result in a significant change in viewpoint (Misra et al., 2012a).

The change in intensity due to the shift  $[u, v]$  is expressed as:

$$I(x + u, y + v) \approx I(x, y) + uI_x(x, y) + vI_y(x, y) \quad (3.1)$$

In the above equation,

$I(x + u, y + v)$  represents the Shifted Intensity Component,  
 $I(x, y)$  represents the Intensity Component.

Employing the Taylor Series for a 2D function and making a first-order approximation

$$I(x + u, y + v) \approx I(x, y) + uI_x(x, y) + vI_y(x, y) \quad (3.2)$$

Through substitution and additional derivation, the matrix representation of the Harris Corner for a slight shift  $[u, v]$  can be represented as follows.

$$E(u, v) \cong [u \ v]M \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.3)$$

Here,  $M$  represents a 2 x 2 matrix obtained from the gradients of the image.

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3.4)$$

where  $w$  is a window function, and represents the components of the satellite image derivative. Following that, the calculation of the corner response measure takes place.

$$R = \det M - k (\text{trace } M)^2 \quad (3.5)$$

where  $\det M = \lambda_1 \lambda_2$  represents the determinant of matrix  $M$ ,  $\text{trace } M = \lambda_1 + \lambda_2$  denotes the matrix trace,  $\lambda_1$  and  $\lambda_2$  stand for the eigenvalues of matrix  $M$ , and  $k$  is a constant determined empirically.

A point is recognized as a "corner" when both  $\lambda_1$  and  $\lambda_2$  exhibit considerable magnitudes, illustrated in Figure-3.1. This indicates that the intensity change at the point is experiencing uniform growth in every direction.

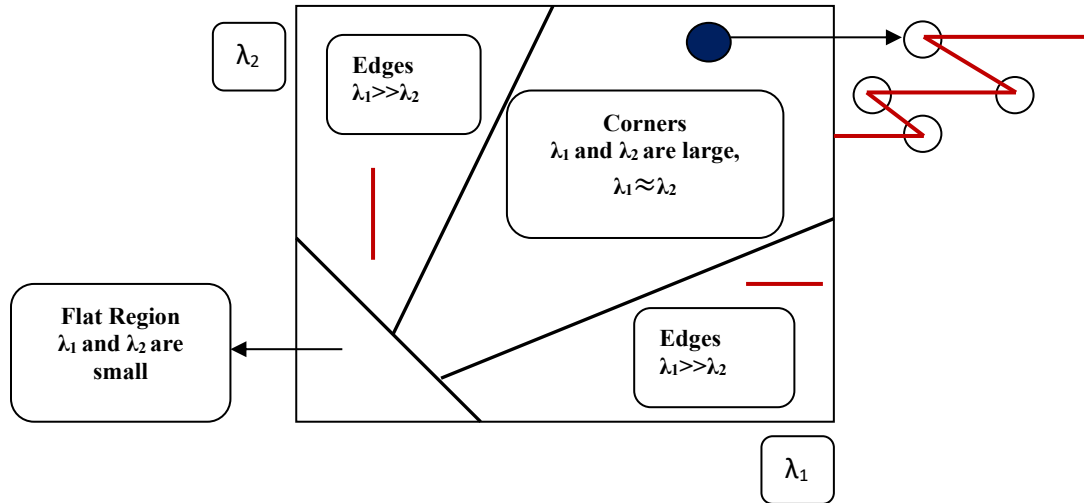


Figure-3.1. Classification of Points using Eigenvalues of Matrix M

The multi-scale Harris Operator is utilized to improve the detection of feature points. In the context of multi-source image registration, the Harris Corner is obtained through the utilization of a Partial Main Orientation Map (PMOM) in conjunction with a Generalized Location and Orientation Histogram (Gao et al., 2022). An approach based on Harris-affine for extracting points of interest combines mutual information with the goal of attaining sub-pixel registration accuracy (Liu et al., 2008). The Multiscale Harris method is utilized to detect stable feature points within the space of gradient amplitudes (Wang et al., 2020). The response function of the Harris Detector is expanded from a linear to a quadratic structure, resulting in a weighted function that incorporates both spatial and intensity details of pixels. This results in more accurate localization of corners (Shang et al., 2020). The Harris Corner Feature Detector algorithm offers the advantage of efficiently extracting numerous point features in images captured through remote sensing. Nonetheless, it also identifies numerous keypoints that lack adequate uniqueness and differentiation for feature matching.

### 3.1.2. SIFT Detector and Descriptor

The SIFT detection procedure begins by generating extrema in scale space through the utilization of a Gaussian Kernel, as outlined by Lowe in 1999. While the Laplacian of Gaussian (LoG) has

been identified as capable of delivering distinctive features and robust scale representations, its computational intensity prompts the use of the Difference of Gaussian (DoG) pyramid creation. This is regarded as a close approximation to the Laplacian of Gaussian (LoG). Equation-3.6 illustrates the connection between the Smooth Image ( $L$ ) and the Input Image ( $I$ ) at the coordinate  $(x, y)$ .

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3.6)$$

$$\text{where } G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma^2} \quad (3.7)$$

The subsequent step involves creating a pyramid of Differences of Gaussian (DoG) to detect the minimum and maximum values among all neighboring keypoints, encompassing various scales, as indicated by Ma et al. in 2016. Figure-3.2 demonstrates the formation of the DoG pyramid and the identification of extrema through various scales. The calculation of DoG is accomplished by determining the disparity between two consecutive scales, where the scales are spaced apart by a constant multiplicative factor  $k$ . The formulation for DoG is given by:

$$DoG(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (3.8)$$

Using equation (3.6), it is feasible to represent equation (3.8) in relation to the smooth image  $L$  as:

$$DoG(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3.9)$$

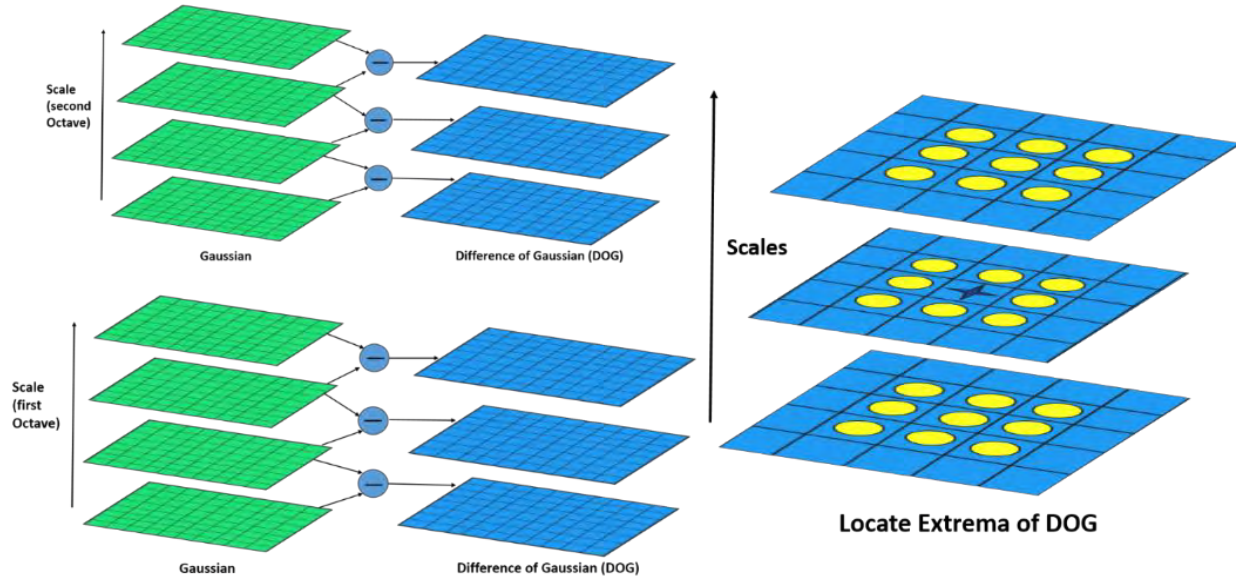


Figure-3.2. DoG Pyramid Creation and Localization of extrema

For the sub-pixel level localization of potential stable feature points to be accomplished and eliminate initial outliers, Taylor Series approximation is employed, as suggested by Goncalves et

al. in 2011. In this context, the Taylor Expansion, incorporating quadratic terms, of the scale-space function is modified to reposition the origin to the sample point.

$$DoG(Z) = DoG + \frac{\partial DoG^T}{\partial Z} Z + \frac{1}{2} Z^T \frac{\partial^2 DoG^T}{\partial Z^2} Z \quad (3.10)$$

In this scenario, the assessment includes  $DoG$  and the rates of change at the sampled point, where  $Z = (x, y, \sigma)^T$  signifies the deviation from this point. The determination of the extremum location,  $\hat{Z}$ , is achieved by equating the derivative to zero.

$$\hat{Z} = \frac{-\partial^2 DoG^{-1}}{\partial Z^2} \frac{\partial DoG}{\partial Z} \quad (3.11)$$

The orientation of keypoints is determined by analyzing the gradient of individual blurred images. This orientation assignment is effective in imparting rotation-invariant properties to the keypoints. The process primarily involves calculating the central derivatives, gradient magnitudes ( $m$ ), and orientation ( $\theta$ ) of the smooth image ( $L$ ) related to the keypoint located at coordinates ( $x$ ,  $y$ ). A weighted direction histogram is then generated within the proximity of each keypoint, serving as a bin. Ultimately, the direction of the keypoint is established by selecting the peak in this histogram.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3.12)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (3.13)$$

The process of constructing descriptors begins with the sampling of points around a given point. Afterward, the gradients and coordinates undergo rotation according to the previously established direction, and the area is divided into smaller sub-regions. A histogram is generated for every one of these smaller regions utilizing predetermined bins. The representation for each keypoint is stored as a vector element, and these descriptors are utilized for matching features in images captured at two different time points.

A revised variant of SIFT has been suggested to attain evenly features that are distributed and matched as well as enhance the reliability of matching through a robust criterion (Paul & Pati, 2016). For robust feature matching in satellite images, a scale-orientation joint restriction criterion has been introduced for SIFT matched points (Li et al., 2009). Furthermore, a pre-registration step is recommended employing SIFT, followed by refinement through an adjusted Marquardt-

Levenberg search strategy within a framework that incorporates multiple resolutions (Gong et al., 2013).

The improvement of matches is executed through the application of the SIFT technique based on Feature Slope Grouping (FSG), followed by an affine transform (Chang et al., 2019). SIFT is considered an appealing option for the registration of SAR images based on point features (Schwind et al., 2010). To enhance the efficiency of image registration tasks, there is a proposal for a parallel implementation of global optimization applied to block-wise SIFT utilizing reweighted least squares in an iterative mode (Huo et al., 2011).

Sophisticated algorithms based on SIFT are utilized for the matching of images captured through multiple sensing modes, including optical and SAR images (Xiang et al., 2018). The iterative form of SIFT incorporates a feedback loop for rectification, which is employed to update rectified parameters, producing consistent sets of feature points achieved through maximum similarity. This procedure is repeated iteratively until an automatic stopping criterion is satisfied (Chen et al., 2020).

The recommendation is for the use of Adaptive binning-based SIFT due to its resilience to local geometric transformations, leading to a notable enhancement in the distinctiveness and resilience of the SIFT descriptor (Sedaghat & Ebadi, 2015). The optical and SAR image registration utilizes a simultaneous segmentation using an iterative level set coupled with the SIFT methodology (Xu et al., 2015).

The cutting-edge methods for matching features in multi-modal spatial alignment encompass the utilization of SIFT alongside a polar frequency descriptor and employing cascading sample consensus in multi-modal image registration (Yu et al., 2021; Zhang et al., 2021). SIFT, in conjunction with a dynamic outlier elimination method, computes the average distances among each feature point, resulting in the generation of putative keypoint correspondences for image registration (Hosseini & Nasri, 2017). Although SIFT exhibits scale and rotation invariance, as well as accurate feature point detection, its effectiveness in matching features in multi-temporal microwave imagery is reported to be suboptimal. This is attributed to the challenges posed by multiplicative noise in SAR images during feature extraction.

### 3.1.3. SURF Detector/Descriptor

Speeded Up Robust Features (SURF) identify features by utilizing the Hessian matrix measure and utilize a descriptor based on a distribution approach (Bay et al., 2008). For a given point  $(x, y)$  in an image  $I$ , the computation of the Hessian matrix  $H(x, \sigma)$  in the  $x$  image space at a particular scale  $\sigma$  is expressed as follows.

$$H(x, \sigma) = \begin{pmatrix} L_{xx}(x, \sigma)L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma)L_{yy}(x, \sigma) \end{pmatrix} \quad (3.14)$$

where  $L_{xx}(x, \sigma)$  is the convolution of the Gaussian second order derivative  $\frac{\partial}{\partial x^2} g(\sigma)$  with the image  $I$  in point  $x$ , and similarity for  $L_{xy}(x, \sigma)$  and  $L_{yy}(x, \sigma)$ .

While Gaussians are widely regarded as effective for the analysis of scale-space, it has been found that issues arise when discretizing and cropping them. Notably, aliasing tends to occur with Gaussian filters when the images results are downsampled. Recognizing this limitation, SURF opts for approximations using box filters (Sedaghat & Mohammadi, 2019). Such estimations of integral images enable the swift computation of second-order Gaussian derivatives, and their application is not dependent on size. Utilizing box filters in conjunction with integral images removes the necessity for applying the same filter iteratively to the previous one. Moreover, filters of any size can be directly applied to the original image at the same speed, enabling parallel processing (Zhang et al., 2020).

The SURF descriptor shares similar characteristics with SIFT but with reduced complexity. The initial step involves determining a consistent orientation by utilizing data from a circular area surrounding the identified feature keypoint. Afterward, a square region oriented in accordance with the chosen orientation is generated, and the descriptor based on SURF is extracted from this region.

Expanding the SURF descriptor is demonstrated to surpass in effectiveness than the original SURF approach in multi-temporal image co-registration (Bouchiha & Besbes, 2013). The integration of the scale restriction method with SURF enhances the performance of registration of images in multiple spectral bands (Teke & Temizel, 2010). The modified SURF introduces a similarity measure function based on trajectories obtained from Lissajous figures, resulting in improved robustness and accuracy (Song & Zhang, 2010). Employing spatial relationships derived from Kernel Canonical Correlation Analysis (KCCA) in combination with the SURF descriptor,

results in a considerable number of accurate matches (Yan et al., 2014). The normalized SURF significantly increases the count of potential correspondences of keypoints between images characterized by differences in spectral characteristics (Jhan & Rau, 2021).

To address common deformations in images, a suggested SURF approach includes co-registering to integrate topology structure for the correlation of aerial images with satellite images (Brook & Ben-Dor, 2011). SURF is used to pick Ground Control Points (GCPs) from UAV and LiDAR images, improving the accuracy of feature matching (AL-NAJJAR, 2019). The combination of SURF with efficient outlier removal techniques enhances the accuracy of image co-registration, particularly for high-resolution satellite images (Anzid et al., 2017). Employing a linear transformation in segments in conjunction with SURF speeds up the co-registration process, achieving higher accuracy and meeting real-time demands (Guo et al., 2010). Leveraging hyperspectral information, SURF contributes to more effective image co-registration in hyperspectral image alignment (Ordonez et al., 2019).

Addressing the challenges of SAR image registration, SURF, combined with an effective elimination of outliers, that is utilized to obtain improved matched features for estimating transformation parameters (Durgam et al., 2016). Even though it extracts keypoints from remote sensing imagery more quickly than SIFT, SURF faces limitations in detecting reliable keypoints for feature matching in multi-modal image pairs.

#### **3.1.4. ORB Feature Detection/Description**

ORB (Rosten & Drummond, 2006; Rublee et al., 2011) comprises fusion of methods, namely, the Features from Accelerated Segment Test (FAST) for detection and the Binary Rotated Independent Elementary Feature (BRIEF) for creating descriptors (Calonder et al., 2010). The ORB algorithms have the capability to identify unique and consistent correspondences between keypoints in remote sensing imagery. The FAST method involves a single parameter, the intensity threshold denoted as ' $t$ .' This parameter is assessed by considering the center pixel and its surrounding pixels arranged in a circular ring. In our particular scenario, we choose 16 adjacent pixels positioned along the circular perimeter to establish the threshold for intensity between the central pixel and those positioned on the circumference of the circle.

In each instance position  $x$  within the  $P \{1, \dots, 16\}$  set around the circumference, the intensity of the pixel  $I$  at that specific location in relation to the point  $p$  exhibits three distinct states:

$$S_{p \rightarrow x} = \left( \begin{array}{l} d, I_{p \rightarrow x} \leq I_p - t(\text{darker}); \\ s, I_p - t < I_{p \rightarrow x} < I_p + t(\text{similar}); \\ b, I_p + t \leq I_{p \rightarrow x}(\text{brighter}) \end{array} \right) \quad (3.15)$$

where  $S_{p \rightarrow x}$  represents location state of  $x$  with respect to point  $p$

The described procedure segregates the initial set  $P$  into three categories:  $P_d$  (Darker Subset),  $P_s$  (Similar Subset), and  $P_b$  (Brighter Subset). The final decision regarding whether a candidate pixel qualifies as a corner is made by calculating the entropy of the set  $P$ .

Consider  $K_p$  as a Boolean variable, taking the value true if  $p$  is a corner and false otherwise. The entropy  $H(p)$  of the set  $P$  is then expressed as:

$$H(p) = (c + \bar{c}) \log_2(c + \bar{c}) - c \log_2 c - \bar{c} \log_2 \bar{c} \quad (3.16)$$

where  $c = \{p | K_p \text{ is true}\} = \text{corner points}$ ,  $\bar{c} = \{p | K_p \text{ is false}\} = \text{non-corner points}$

FAST, lacking corner measures, tends to generate numerous edge responses. Consequently, the FAST response undergoes filtration using the Harris Corner metric to obtain consistent keypoints. Figure-3.3 depicts the filter points identified by ORB, which exhibit a higher brightness compared to the neighboring pixels on the circular ring. The BRIEF feature descriptor utilizes binary tests between pixels within a smoothed image patch. Its efficacy has been assessed in various challenging scenarios, demonstrating robustness in images taken at different times, the application of blurring, and various viewpoints.

The image matching method based on ORB is utilized in diverse image processing fields, such as navigation, panorama stitching, object recognition, and the registration of satellite images (Luo et al., 2019). For improved precision in matching, a modified ORB algorithm has been introduced, relying on the Non-Subsampled Contourlet Transform (NSCT) domain (Ma & Lai, 2019). The Affine-ORB algorithm addresses the challenging task of SAR with optical feature matching



extracting consistent features in images with multiple modalities to facilitate efficient feature matching (Najafi & Ivanov, 2021). To attain precise image registration, ORB creates a cost function to assess similarities among points and acquire accurately matched pairs (Lu et al., 2018). The tracking speed and the quality of matching keypoint pairs are enhanced by the ORB-based tracking algorithm (Fanqing & Fucheng, 2013).

In vital image processing operations like image mosaicking for large-scale panoramic images, Feature detection and description using the ORB technique are used for image co-registration (Guiqin et al., 2019). The performance of blending algorithms is improved by the ORB descriptor through sub-pixel level co-registration, accomplished by utilizing stable feature points extracted from overlapping regions (Tian & Shi, 2014). The ORB matching technique is employed to generate attractive panoramic UAV mosaics (Chen et al., 2018). The utilization of ORB feature points, characterized by binary vectors for feature descriptors, enables real-time image stitching, significantly accelerating the processing steps for extracting and matching features (Wang et al., 2017). Utilizing three-patch and regional variations in grayscale, ORB demonstrates superior performance in the registration of remote sensing images over multiple time periods (Ma et al., 2020). In image stitching applications, PCA-driven ORB diminishes the dimensionality of conventional ORB descriptors, simplifying the intricacy of descriptor data for feature points. This enhances the speed of image merging, making it suitable for instantaneous and extensive-scale spaceborne applications (Zhu et al., 2020).

Under various illumination circumstances, ORB demonstrates resilience as well as utilized for practical tasks involving identifying features (Li et al., 2022). The speed and comparative reliability of ORB make it advantageous for image registration when compared to contemporary feature detection techniques. ORB has the capability to identify a multitude of evenly distributed keypoints across a remote sensing image. However, the drawback is that the number of spurious keypoints is higher, potentially leading to errors in transformation parameter estimation.

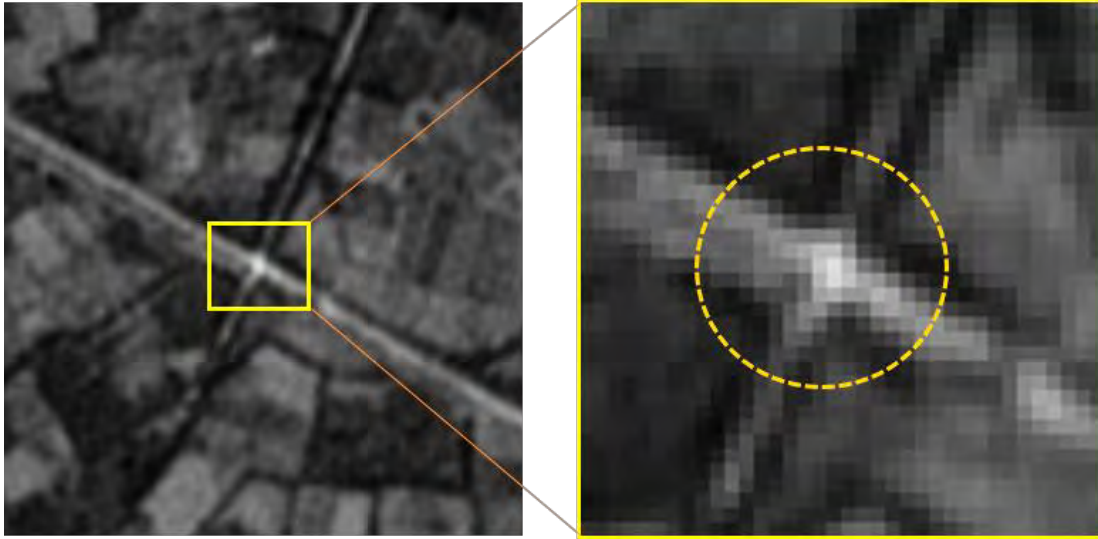


Figure-3.3. Feature Detection using ORB

### 3.1.5. BRISK Detector/Descriptor

BRISK, the Binary Robust Invariant Scalable Keypoint, is capable of detecting points across the entire image and based on a saliency criterion in the scale dimension (Leutenegger et al., 2011). Enhancing the efficiency of keypoint detection involves detecting feature points in both the image pyramid octave layers and the intermediate layers (Mair et al., 2010). The keypoint positions and scale information are derived in the continuous domain through the fitting of quadratic functions (Schwind & Angelo, 2015).

Within the BRISK framework, the scale-space pyramid comprises  $n$  octaves  $c_i$  and  $n$  intra-octaves  $d_i$  for  $i = \{0,1,2,3, \dots, n-1\}$  with  $n$  typically being 4. The octaves are generated by iteratively halving the size of the original image ( $c_0$ ). Each intra-octave, denoted as  $d_i$ , is positioned between layers  $c_i$  and  $c_{i+1}$ . The initial intra-octave,  $d_0$ , is generated by reducing the size of the original image  $c_0$  by a factor of 1.5. Subsequent intra-octave layers are then produced by successively halving the scale.

In BRISK, the FAST criterion is satisfied when a minimum of 9 successive pixels within the 16-pixel circular region around the central pixel are either significantly brighter or darker. The sampling pattern used in the BRISK keypoint descriptor includes points distributed on concentric

circles at suitable scales. This pattern is utilized around each keypoint collects gray measurements, examines local intensity rate of change in values, and determines the orientation of distinctive attributes. In the end, the BRISK sampling pattern with orientation is employed to acquire pairwise brightness assessment outcomes, which are aggregated to construct the BRISK descriptor in binary form. The benefit of keypoints based on BRISK lies in the effective feature matching made possible due to the descriptor being binary.

RGB-D camera-captured imagery utilize the BRISK algorithm, wherein the scale factor of matched points is directly calculated using depth information obtained from the image (Liu et al., 2018). Enhancing the accuracy of co-registration between Hyperspectral SWIR and VNIR images, the BRISK algorithm is used in combination with a robust outlier removal method (Schwind et al., 2014). Enhancing BRISK's precision involves integrating it with the SIFT algorithm and employing the Hamming distance to measure the likeness of feature descriptors (Shi et al., 2021). By incorporating statistics derived from a bilateral grid, BRISK achieves an increased correct rate, a wider matching range, and enhanced robustness (Zhang & Li, 2019). A synthetically color-enhanced variant of BRISK identifies the most indicative feature keypoints and improves the descriptor by integrating color data, leading to matching with greater precision (Cheng & Matsuoka, 2021). Although BRISK may not directly identify accurate feature points in remote sensing images from multiple modes, in situations such as registering Visible-SAR image pairs in multi-modal image registration, BRISK is capable of capturing feature points that are locally prominent. This is achieved by employing a variable and elliptical bilateral filter to eliminate speckle noise (Salehpour & Behrad, 2017).

### **3.1.6. Accelerated KAZE (A-KAZE) Feature Detection/Descriptor**

Techniques for detecting features at multiple scales like SIFT detect features by building a Gaussian kernel within a pyramidal framework. Nevertheless, the application of Gaussian filters may result in blurring, potentially jeopardizing the preservation of feature boundaries. The adverse impact of this blurring effect on object localization and uniqueness, especially in images captured through remote sensing, can compromise accuracy. To tackle this concern, the KAZE approach introduces feature detection and description within a non-linear space through the application of filtering through non-linear diffusion (Alcantarilla et al., 2012). The establishment of this non-

linear space is accomplished utilizing methods of additive operator splitting that include diffusion with variable conductance (Alcantarilla et al., 2011). The traditional non-linear diffusion equation is expressed as follows.

$$\frac{dl}{dt} = (c(x, y, t) \cdot \nabla L) \quad (3.17)$$

Here,  $div$  and  $\nabla$  represent the gradient and divergence operators, respectively, and  $L$  represents the image luminance. The definition of the conductivity function  $c$  is as follows.

$$c(x, y, t) = g(|\nabla L_\sigma(x, y, t)|) \quad (3.18)$$

In this context, the function  $\nabla L_\sigma$  represents the gradient of an image obtained by smoothing the original image  $L$  with a Gaussian filter.

KAZE attains acceleration by incorporating Fast Explicit Diffusion (FED), which merges the benefits of both explicit and partially implicit approaches to enhance the computational process. The idea behind FED schemes comes from decomposing box filters into explicit methods, which are subsequently integrated into a pyramidal framework. This significantly speeds up the process of detecting features in spaces that are nonlinear (Pourfard et al., 2021). Utilizing iterative box filters is straightforward, and they provide a high-quality approximation of Gaussian kernels. The FED is structured to perform  $m$  cycles of  $n$  explicit diffusion steps, utilizing various step sizes  $t_j$  derived from the factorization of the box filter.

$$t_j = \frac{t_{max}}{2\cos^2\left(\pi\frac{2j+1}{4n+2}\right)} \quad (3.19)$$

Here,  $t_{max}$  denotes the maximum step size that adheres to the stability requirement of the explicit scheme.

Improving matching efficiency, the KAZE method is strengthened by integrating composite filtering through non-linear diffusion with a multi-scale pyramid space that gradually decreases in resolution (Liu & Xiao, 2020). The utilization of GPU-accelerated KAZE for spatial alignment of hyperspectral images, thereby decreasing the processing time (Ordonez et al., 2020). Enhancements to KAZE include utilization of modified cosine vectors, which improves the rate of matching (Yang & Li, 2018). The combination of SIFT descriptors with KAZE features proves to

be a highly effective pairing for remote sensing images captured at different time points, showcasing exceptional efficiency in matching of such images (Liu et al., 2016). A-KAZE features support feature tracking for sea ice drift retrieval, surpassing the performance of other cutting-edge techniques for feature detection (Demchev et al., 2017).

The KAZE method distinguishes itself by extracting pertinent features extracted from prechosen bands and proves to be suitable for registering hyperspectral images, particularly in challenging scenarios characterized by significant variations in image scales (Ordonez et al., 2018). Detection, extraction, and matching relying on A-KAZE are applied in sophisticated driver assistance systems, interpreting the nearby environment (Mentzer et al., 2019). The implementation of PCA and A-KAZE contributes to the creation of robust underwater image mosaicking systems, incorporating optimal seamline methods (Abaspur et al., 2020). The A-KAZE algorithm effectively extracts stable feature points in oblique remote sensing images exhibiting alterations in viewpoint (Liu et al., 2016). The A-KAZE algorithm is successful in accurately extracting visual features of buildings for converting 2D building images into 3D reconstructions (Seong et al., 2018). To seamlessly stitch multiple images and eliminate fractures in the overlap area, matching keypoints extracted using A-KAZE are employed (Tengfeng, 2018).

The feature extractor of A-KAZE is also employed in aerial images to accelerate the image mosaicking process, producing geo-referenced mosaic from a UAV in near-real-time (Avola et al., 2017). A-KAZE, combined using a framework based on conditional probability, is employed for the radiometric correction of unmatched multisensory images (Moghimi et al., 2021a). It has been noticed that, when it comes to the adjusting the radiometric properties of two-temporal images, A-KAZE provide higher accuracy in comparison to corner-detection algorithms like BRISK and ORB (Moghimi et al., 2021b). A significant benefit of A-KAZE is its capability to identify consistent key points in Synthetic Aperture Radar (SAR) images, which commonly experience multiplicative noises. Nevertheless, it is crucial to acknowledge that A-KAZE relies on non-diffusion filtering, which may lead to longer execution times for feature extraction when compared to ORB and Harris methods.

### **3.2. Outlier Removal Techniques**

Addressing the task of estimating model parameters from data when outliers are present remains notable challenge, particularly in the domain of automatic satellite image registration. This area is

actively explored by researchers in image processing who are focused on finding effective transformation parameters. The global research community has extensively studied various approaches for estimating models from data. In practical scenarios, the challenge arises from outliers, which deviate from the expected pattern exhibited by the majority of the data. To tackle this issue, Random Sample Consensus (RANSAC) is commonly utilized because of its straightforward execution and resilience.

### 3.2.1. RANSAC based Outlier Removal Technique

RANSAC, a versatile parameter estimation technique (Fischler & Bolles, 1981), exhibits the capability to handle numerous outliers within input data. In contrast to commonly adopted robust estimation methods like least-median squares and M-Squares in the computer vision community, RANSAC has been acknowledged for its superior performance when outliers are present. The RANSAC algorithm accepts sets of observed data values, a model with parameters, and defined parameters of confidence as its inputs. The RANSAC procedure consists of two iterations: creating a conjecture from randomly selected samples and then validating it against the data. Importantly, it doesn't necessitate sophisticated optimization algorithms or extensive memory resources. This algorithm finds applications in various computational problems.

In our scenario, RANSAC was used to infer transformation parameters and eliminate outliers from matched keypoints in the automated satellite image alignment process (Li et al., 2014). The traditional method in image registration typically involves estimating model parameters using the least squares estimate, but it faces challenges when dealing with outliers (Figure – 3.4 (a)). The utilization of RANSAC has proven effective in enhancing the precision of model parameter estimation, effectively filtering out outliers (Figure – 3.4 (b)).

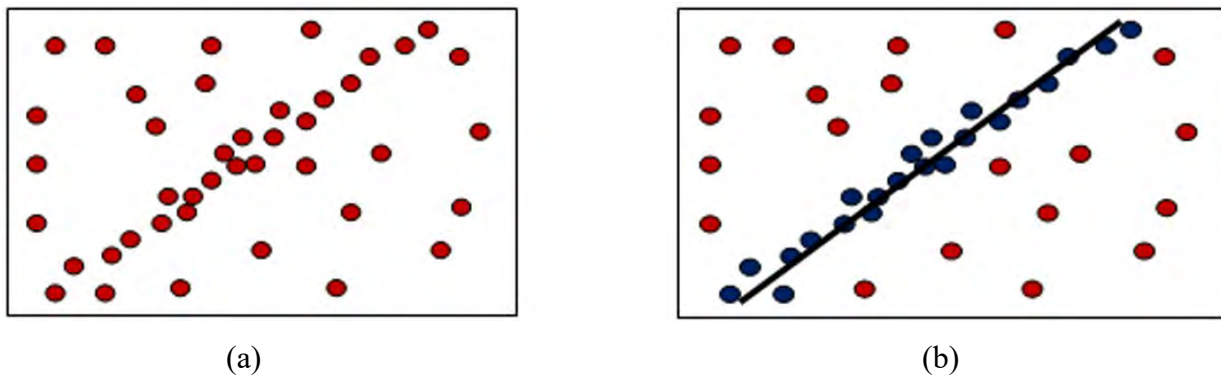


Figure-3.4. (a) Matched control points, including a significant number of outliers for line fitting  
(b) RANSAC based line fitting process, impervious to the influence of outliers

### 3.2.2. RANSAC Algorithm Steps

The RANSAC estimation method produces potential solutions utilizing the minimum necessary minimal number of observations (control points) for deducing the underlying model parameters. In contrast to traditional sampling methods that aim to employ as much data as possible to obtain an initial solution and subsequently eliminate outliers, RANSAC starts with the smallest set and gradually expands it by incorporating reliable data points (Li et al., 2018).

The fundamental steps of the algorithm are described as follows:

- i. Randomly choose the smallest number of points needed to establish the model parameters.
- ii. Calculate the model parameters.
- iii. Identify the number of points from the complete set that align within a predetermined tolerance.
- iv. If the ratio of inliers to the total number of points surpasses a predetermined limit, readjust the model parameters using all the identified inliers and conclude the process.
- v. If not, iterate through steps i to iv a maximum of  $N$  times.

The number of iterations,  $N$ , is set to be sufficiently high to guarantee a probability,  $p$ , that at least one among the randomly chosen sample sets does not contain an anomalous data point (Misra et al., 2012a).

### 3.2.3. RANSAC Phases

RANSAC functions through two successive phases: "Hypothesis Generation" and "Hypothesis Evaluation." The process of estimating optimal transformation parameters is illustrated in the flowchart shown in Figure-3.5.

#### *i. Hypothesis Generation*

RANSAC employs a random selection of a subset of data to derive parameters, creating a proposed assumption of the truth (Li et al., 2018a). The algorithm generates multiple hypotheses during its iterations. In contrast to regression techniques like the least squares method and support vector machine, RANSAC does not function as a regression method. Instead, it leverages the data to formulate a hypothesis, refining and reinforcing it, even though this process may experience reduced accuracy if specific data points are outliers. RANSAC employs only a subset of the data, rather than the entire dataset (Zheng et al., 2021). If the chosen data exclusively consists of inliers, it has the capability to generate a conjecture that closely approximates the truth. This presumption

necessitates a sufficient number of iterations to randomly select all samples of inliers at least once, with a likelihood of failure denoted as  $\alpha$ :

$$N = \frac{\log \alpha}{\log(1-\gamma^m)} \quad (3.20)$$

In this context,  $m$  denotes the number of data points required to formulate a hypothesis, and  $\gamma$  represents the probability of selecting an inlier, specifically the proportion of inliers to the entire set of sample data (inliers ratio). Nevertheless, the inliers ratio  $\gamma$  is frequently unknown in various practical scenarios, necessitating users to ascertain it. RANSAC transforms a continuous-domain estimation problem into a discrete-domain selection problem. The current difficulty lies in choosing the most appropriate pair from numerous pairs (Misra et al., 2021a).

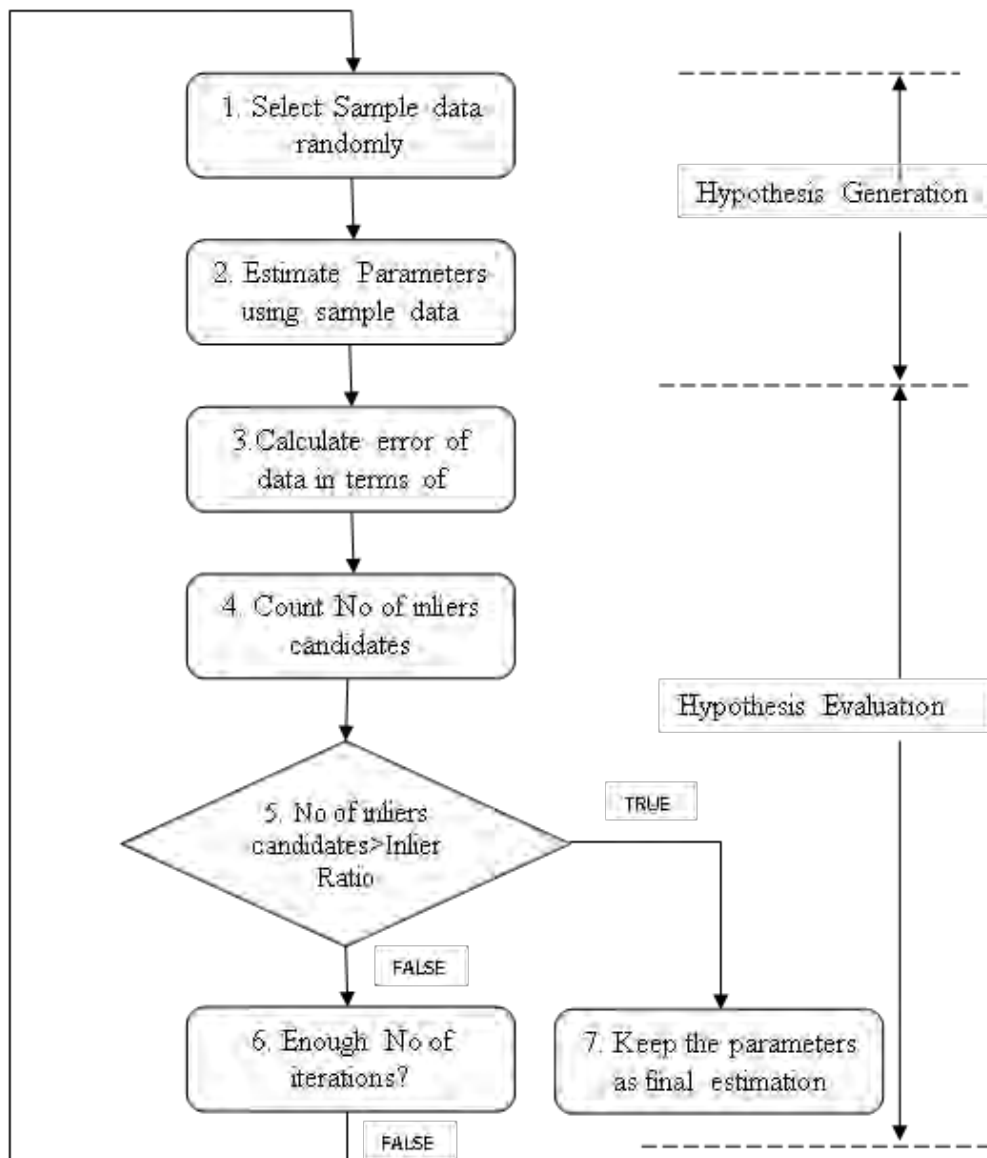


Figure-3.5. RANSAC Flowchart



ii. Hypothesis Evaluation

RANSAC selects the hypothesis that finds support from the candidates identified as inliers. An observation is considered an inliers' candidate if its deviation from a hypothesis is below a predefined threshold. In the context of fitting a line, the error can be described as the geometric distance between the observation and the estimated line. The threshold serves as the secondary tuning parameter, closely linked to the magnitude of noise that affects inliers, or simply put, the noise magnitude. RANSAC tackles the optimization formulation for the selection problem, which is expressed as:

$$\hat{M} = \operatorname{argmin}\{\sum_{d \in D} L(\operatorname{Err}(d; M))\} \quad (3.21)$$

In this context,  $D$  represents the data,  $L$  is a loss function, and  $\operatorname{Err}$  is a type of error function, for example, geometric distance.

The depiction of the loss function in the least squares method is given by  $L(e) = e^2$ .

On the other hand, RANSAC employs

$$L(e) = \begin{cases} 0, & |e| < c, \\ \text{constant}, & \text{otherwise} \end{cases} \quad (3.22)$$

In this representation,  $c$  denotes the threshold. Figure – 3.6 illustrates the contrast between the two loss functions. RANSAC shows a consistent loss for large errors, whereas the least squares method experiences a significant loss. Outliers disturb the least squares method, typically introducing substantial errors.

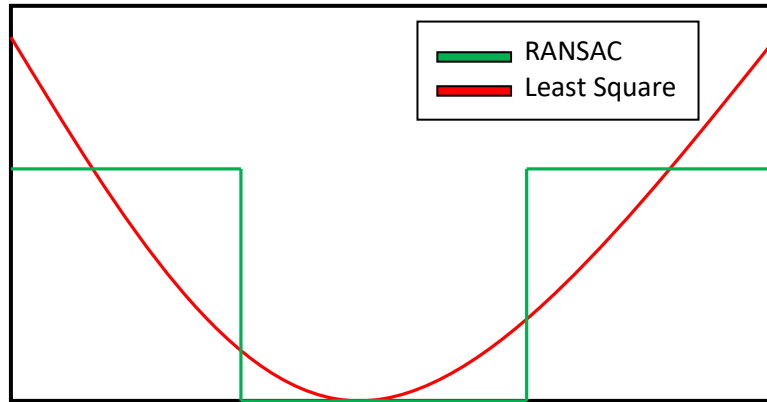


Figure-3.6. Loss Function

The RANSAC framework incorporates Stratified Random Selection to remove mismatched features in the initial matched point set, thereby improving the performance of registration of images based on features (Zhang et al., 2020). The fusion of applying RANSAC to the Harris

Corner Detector produces encouraging outcomes in the registration of deformable remote sensing images (Colleu et al., 2006). Integration of RANSAC with affine invariant feature extraction results in improved accuracy in image matching across different remote sensing scenes (Cheng et al., 2012). Combining SIFT with RANSAC rejects outliers, estimates the transformation model, and has the capability to generate seamless panoramic images (Redzuwan et al., 2015). In the case of fusion of panchromatic and multi-spectral images, sub-pixel co-registration is achieved by employing the SIFT with RANSAC (Li & Zhang, 2012). RANSAC is utilized to establish potential correspondences between keypoints in multi-temporal images, aiding in the superimposition of images (Safy et al., 2013). In the accurate rectification of geometric distortions in remote sensing images, RANSAC is employed to optimize ground control points, resulting in a substantial quantitative enhancement in the accuracy of co-registration (Hosseini-Nejad & Nasri, 2022; Tan et al., 2018). In SAR image registration, the amalgamation of SURF and enhanced Random Sampling Consistency effectively removes mismatched feature points, thereby improving registration effectiveness (Liao et al., 2020). The integration of region growing technique, RANSAC and the Hough transform is employed for processing LiDAR data in 3D building modeling (Tarsha-Kurdi et al., 2007). Utilizing Spatial Consistency RANSAC enhances both the accuracy and execution time of RANSAC when calculating the transformation parameters of a geometric model (Fotouhi et al., 2019).

### **3.3. Remote Sensing Image Fusion Theory and Types**

Image fusion in remote sensing is the amalgamation of information from several images acquired by different sensors or at different times to create a single composite image that enhances the overall quality and provides comprehensive information for analysis. The main objective of image fusion is to preserve essential details from each source image while reducing redundancy and improving the interpretability of the fused image. The underlying theory of remote sensing image fusion can be summarized as follows.

*i. Data Sources:* Typically, remote sensing image fusion entails amalgamating data from various sources, which can include:

*a. Spectral Data:* Images captured with varying spectral bands or wavelengths, such as multispectral and hyperspectral imagery.

*b. Spatial Data:* Images captured at differing spatial resolutions, ranging from high-resolution optical imagery to relatively low-resolution thermal or radar imagery.

*c. Temporal Data:* Images captured at different intervals to monitor changes over time.

*ii. Preprocessing:* Before fusion, it is essential to preprocess the input images. This could encompass radiometric calibration, geometric correction, atmospheric correction, and noise reduction to ensure that the images are in a compatible and high-quality format.

*iii. Image Registration:* Image registration is a crucial step to align all input images onto the same coordinate system. This ensures that corresponding pixels in different images represent the exact location on the Earth's surface.

Merged remote sensing images find applications in diverse fields, such as land cover classification, change detection, disaster monitoring, environmental analysis, urban planning, and military reconnaissance. Image fusion often involves trade-offs between spatial and spectral resolution, as well as between information preservation and data redundancy reduction. The choice of fusion method depends on the specific requirements of the application. Image fusion techniques in the remote sensing domain are classified into six prominent classes: i. Component Substitution (CS) ii. Multi Resolution Analysis (MRA) iii. Color Based (CB) Techniques iv. Variational Optimization (VO) v. Deep Learning (DL) vi. Hybrid Methods. The theoretical principles underlying various methods of remote sensing image fusion are outlined below.

### **3.3.1. Component Substitution (CS)**

CS-based techniques are widely used in popular image fusion approaches. These methods include transforming the source multispectral image into a domain that allows the separation of color and spatial features into distinct components. The component containing spatial features is substituted with the source panchromatic data. Before the substitution, histogram matching is used to minimize spectral distortion between the source panchromatic image and the component preserving spatial features. This process becomes crucial, especially when the source imagery is captured by varied sensors. Ultimately, the data is projected back to acquire the fused image in the initial color domain (Tu et al., 2001).

The Brovey Transform is a widely used method in CS based image fusion that operates on multispectral and panchromatic images (Misra et al., 2012b). It combines the high spatial

resolution from panchromatic images with the color information from multispectral images. The equation for the Brovey Transform-based image fusion is defined as follows:

$$F_R = I_{PAN} * \left(\frac{M_R}{M_{avg}}\right) \quad (3.23)$$

$$F_G = I_{PAN} * \left(\frac{M_G}{M_{avg}}\right) \quad (3.24)$$

$$F_B = I_{PAN} * \left(\frac{M_B}{M_{avg}}\right) \quad (3.25)$$

where,  $F_R$ ,  $F_G$ , and  $F_B$  represents the red, green, and blue channels of the fused image respectively,  $M_R$ ,  $M_B$ ,  $M_G$  represents the red, green, and blue bands of the multispectral image respectively,  $M_{avg}$  is the mean of the multispectral channels, and  $I_{PAN}$  represents the panchromatic image.

### 3.3.2. Multi Resolution Analysis (MRA)

MRA-based techniques are also among the frequently employed methods for image fusion. Generally, these approaches involve three main steps (Vivone et al., 2017). In the initial stage, the source images undergo decomposition into various scale levels through multiresolution decomposition algorithms like wavelet, curvelet, pyramid transform, and so on. Then, in the second step, spatial/spectral features are blended within the various decomposition levels using a fusion rule. Finally, in the third step, applying an inverse transform to produce the merged data in the original color domain (Chen et al., 2015; Ghassemian, 2016). Fusing images through multi-resolution analysis involves combining the information from different images at various resolutions to create a single, fused image that preserves the essential features from each input image. The process typically includes the following steps:

- i. Decomposition: Split the source images into distinct levels or scales using a multi-resolution analysis technique like the wavelet transform.
- ii. Fusion Rule: Establish a fusion rule that dictates how to merge the corresponding coefficients or components from each image at different scales.
- iii. Reconstruction: Reconstruct the merged image by combining the fused coefficients or components from each level.

The specific equation for multi-resolution analysis image fusion may differ based on the chosen fusion rule. One joint fusion rule is to take the maximum, minimum, mean, or weighted mean of

the corresponding coefficients from the input images at each scale. The fusion equation can be described as follows:

$$F(i, j) = \sum_k W(k) * C_1(i, j, k) + (1 - W(k)) * C_2(i, j, k) \quad (3.26)$$

where,  $F(i, j)$  represents the pixel value at location  $(i, j)$  in the merged image,  $C_1(i, j, k)$  and  $C_2(i, j, k)$  symbolize the corresponding coefficients from the input images at scale  $k$ ,  $W(k)$  is the weight at scale  $k$  that determines the contribution of each image at that scale.

### 3.3.3. Color Based (CB) Techniques

Color balanced image fusion is a method used to combine multiple images while maintaining a consistent color balance across the merged image. This is especially crucial when combining images with varying lighting conditions or color casts to ensure a visually pleasing and natural-looking final result. The specific equation for color-balanced image fusion can vary depending on the chosen method and the specific requirements of the application. However, one common approach involves adjusting the color channels of the provided images to achieve a balanced color representation in the fused image. The general equation for color balanced image fusion might involve the following steps:

- i. Color Transformation: Convert the color representation of the source images to a common color space, such as RGB, LAB, or YUV, to facilitate color balancing.
- ii. Color Adjustment: Apply color adjustments or corrections to the color channels of the source images to achieve a consistent color balance across the fused image.
- iii. Fusion Rule: Combine the color balanced color channels of the input images using a fusion rule, which may involve blending or averaging the color values.
- iv. Inverse Color Transformation: Convert the fused image back to the original color space if necessary.

The equation for color balanced image fusion can be represented as:

$$F(i, j, c) = \alpha * I_1(i, j, c) + (1-\alpha)*I_2(i, j, c) \quad (3.27)$$

where,  $F(i, j, c)$  represents the pixel intensity value of color channel  $c$  at location  $(i, j)$  in the fused image,  $I_1(i, j, c)$  and  $I_2(i, j, c)$  represent the corresponding color channel values from the input images,  $\alpha$  is a weighting factor that determines the contribution of each source image to the merged image. The value of  $\alpha$  may depend on the specific color balancing technique used.

The color balance methods employ certain methods to retain the color details of the source multispectral data (Yilmaz et al., 2019; Yilmaz et al., 2021). Color balance methods like Local Mean Variance Matching (LMVM) (de B'ethune et al., 1998), Local Mean Matching (LMM) (de B'ethune et al., 1998), and Smoothing Filter-Based Intensity Modulation (SFIM) (Liu, 2000) use sliding windows on the source data to preserve color characteristics throughout the image fusion process. Other color balance techniques such as Hybrid Colour Mapping (HCM) (Zhou et al., 2016), Coupled Nonnegative Matrix Factorization (CNMF) (Yokoya et al., 2012), and criteria-based (Gungor, 2008) techniques are also part of the CB methods that yield satisfactory outcome in image fusion.

### 3.3.4. Variational Optimization (VO)

Variational Optimization (VO) method-based image fusion is a technique used to combine multiple images by minimizing a defined energy or cost function. This approach formulates the image fusion problem as an optimization task, where the objective is to discover an optimal solution that minimizes the difference between the merged image and the source images while satisfying certain constraints. The specific equation for Variational Optimization based image fusion typically involves the definition of an energy functional and the application of optimization techniques to find the optimal solution. The general equation for Variational Optimization based image fusion can be described as follows.

$$E(F) = \int (\sum_{i=1}^n \lambda_i \cdot \rho_i (W_i(F) - I_i)^2 + \sum_{j=1}^m \eta_j \cdot \varphi_j(W_j(F))) \quad (3.28)$$

where,  $E(F)$  is the energy function that needs to be minimized,  $F$  represents the fused image,  $I_i$  denotes the source image,  $W_i$  is a transformation operator or a fusion rule utilized for the source images,  $\lambda_i$  and  $\eta_j$  are weighting coefficients that govern the importance of the data fidelity term and the regularization term, respectively,  $\rho_i$  and  $\varphi_j$  represent data fidelity terms and regularization terms, which capture the difference between the merged image and the source images, and any desired constraints or priors on the fused image.

The optimization process involves finding the optimal fused image  $F$  that minimizes the energy functional  $E(F)$ . This can be accomplished by using various optimization techniques, such as gradient descent, conjugate gradient, or other iterative optimization algorithms. The specific formulation of the energy functional and the choice of the data fidelity and regularization terms may vary based on the requirements of the image fusion task and the attributes of the source images. Variational Optimization based image fusion algorithms aim to balance data fidelity and

regularization to produce high-quality fused images while preserving essential features derived from the source images. Variational optimization (VO) based methods approach image fusion as an optimization problem. These methods utilize an objective function within a Variational optimization scheme to find an appropriate solution. Techniques based on VO typically produce creating pansharpened images using sensor models that establish a correlation between input images and pansharpened images (Vivone et al., 2020). Bayesian methods, which include a robust probability framework, can also fall into the category of techniques based on VO (Fasbender, 2008; Zhang et al., 2009), providing them with a distinct advantage.

### 3.3.5. Deep Learning (DL) based Techniques

Image fusion based on Deep Learning entails the utilization of deep neural networks for learning the fusion mapping directly from the input images. The neural network is trained on a dataset of paired input images and their respective fused images to learn the complex mapping between the source images and the desired output. The specific equation for image fusion based on deep learning depends on the architecture and design of the neural network used for the fusion task. However, the general framework for the DL image fusion equation can be described as follows.

$$F = NN(I_1, I_2, \dots, I_n) \quad (3.29)$$

where,  $F$  represents the fused image produced by the deep neural network,  $NN$  denotes the utilized deep neural network architecture for image fusion,  $I_1, I_2, \dots, I_n$  represents the provided images used as input to the neural network for the fusion task.

The deep neural network is trained using a dataset of source image pairs and their corresponding ground-truth fused images. The training process involves optimizing the network parameters to minimize a defined loss function that measures the discrepancy between the predicted fused image and the ground-truth fused image. This involves forward propagation of the input images through the network, computation of the loss function, and backpropagation to update the network parameters utilizing optimization techniques such as stochastic gradient descent or its variants. The architecture of the deep neural network employed for image fusion can vary, including Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), or other custom-designed architectures tailored for the specific requirements of the image fusion task. The specific layers, connections, and neural network parameters are established based on the complexity of the fusion task and the attributes of the source images.

The methods for image fusion based on deep learning (DL) have become prominent due to their achievements in recent times. These techniques examine the hierarchical features of input image patches using deep neural networks (Xing et al., 2018). The initial effort in this direction regarded image fusion as a super-resolution challenge and employed a three-layer Convolutional Neural Network (CNN) structure to tackle it (Masi et al., 2016). However, this uncomplicated network architecture was discovered to possess restricted non-linear capacity, resulting in notable pansharpening challenges (Yuan et al., 2018), including the loss of a significant amount of spatial detail content (Zhang et al., 2019; Jiang et al., 2020). To address this issue, several enhanced network structures were developed, including the architecture of deep residual networks (Wei et al., 2017; Benzenati et al., 2020), denoising auto-encoder architecture (Azarang et al., 2019; Azarang et al., 2020), generative adversarial network architecture (Ozcelik et al., 2021; Peng et al., 2020), and dense network architecture (Deb et al., 2002; Peng et al., 2020).

### 3.3.6. Hybrid based Techniques

Hybrid based techniques in image fusion often combine the advantages of multiple algorithms or methods to enhance the quality and efficiency of the merged image. The specific equations for hybrid-based image fusion methods can vary depending on the particular methods being combined and the desired outcomes. One common approach is integrating the outcomes of various fusion methods using a weighted average or a more sophisticated fusion rule. The equation for a generic hybrid-based image fusion algorithm can be described as follows.

$$F(i, j) = \sum_{k=1}^n w_k \cdot F_k(i, j) \quad (3.30)$$

where,  $F(i, j)$  is the pixel value at location  $(i, j)$  in the fused image,  $F_k(i, j)$  represents the pixel value obtained from the  $k^{th}$  fusion method or algorithm,  $w_k$  is the weight designated for the  $k^{th}$  fusion method, which determines its contribution to the final fused image. These weights can be predefined or determined dynamically depending on the attributes of the source images and the specific fusion methods being used.

Hybrid pansharpening techniques integrate diverse approaches from various methods to attain optimal performance in image fusion (González-Audicana et al., 2004; Zhang et al., 2005; DadrasJavan et al., 2018). The hybrid methods Wavelet-Based PCA (W-PCA) (Yusuf et al., 2013), IHS-DWT (Abd El-Samie et al., 2012) and IHS-based Discrete Wavelet Frame Transform (IHS-DWFT) (Abd El-Samie et al., 2012) blend the benefits of the CS and MRA image fusion techniques.



### 3.4. Change Detection Components and Methods

Remote sensing change detection is a process used to analyze and identify alterations on the Earth's surface or other remote environments over time using remotely acquired data (Asokan & Anintha, 2019). This technique finds extensive application in diverse fields, including environmental monitoring, urban planning, agriculture, forestry, and disaster management. The theory for change detection in remote sensing involves several key components and methods.

- i. *Data Acquisition*: Remote sensing images can be acquired through various sensors, such as satellites, aerial photography, drones, or ground-based instruments. These sensors capture information in different spectral bands, such as visible, infrared, thermal, and microwave.
- ii. *Multi-Temporal Data*: Change detection relies on comparing two or more images or datasets taken at different times. These images are referred to as "multi-temporal data" and are essential for identifying changes over time.
- iii. *Preprocessing*: Preprocessing steps include radiometric and geometric correction to ensure the images are correctly calibrated and aligned. Atmospheric and topographic corrections may also be necessary to remove unwanted effects.
- iv. *Image Registration*: To compare different images accurately, they must be spatially registered, meaning that corresponding pixels in the images correspond to the exact geographic locations.
- v. *Image Differencing*: Among the prevalent methods for change detection entails the subtraction of pixel values between two images. The resulting difference image emphasizes regions where changes have taken place.
- vi. *Thresholding*: After image differencing, a thresholding process is often applied to the difference image to separate changed from unchanged areas. Pixels exceeding a specific threshold value are deemed changed.
- vii. *Post-Classification Comparison*: Another approach involves classifying the land cover in each image independently and then comparing the land cover categories across the two temporal periods. Discrepancies in land cover classifications identify changes.
- viii. *Change Detection Algorithms*: Various change detection algorithms exist, ranging from simple methods like image differencing to more advanced techniques like object-based image analysis, spectral mixture analysis and machine learning-based approaches.

- ix. *Change Analysis*: Once changes are detected, they can be further examined to determine the type and extent of the changes. This may involve thematic mapping, spatial statistics, and time series analysis.
- x. *Validation*: It is vital to validate the outcome of change detection to ensure accuracy. This can be accomplished through ground truth data, field surveys, or comparisons with higher-resolution data.
- xi. *Interpretation and Visualization*: The final step involves interpreting the detected changes and visualizing the results using maps, charts, or other graphical representations.

Remote sensing change detection techniques refer to the methodologies utilized for identifying and analyzing changes in the Earth's surface or environment over time, employing remote sensing data. These techniques can be categorized based on the approach and data used. Below are some typical kinds of remote sensing change detection techniques.

### 3.4.1. Image Differencing

Image differencing is among the most straightforward change detection techniques. It includes subtracting the pixel intensities of two images captured at distinct times to generate a distinct image (Lu et al., 2004). Pixels with values distinct from zero in the difference image indicate areas of change. This method is straightforward but sensitive to registration errors and image radiometric differences. The equation for image difference-based change detection can be represented as follows:

$$D(i, j) = |I_1(i, j) - I_2(i, j)| \quad (3.31)$$

where,  $D(i, j)$  represents the computed difference at location  $(i, j)$  between the corresponding pixels in the two input images,  $I_1(i, j)$  and  $I_2(i, j)$  denote the pixel values at location  $(i, j)$  in the first and second input images respectively.

After calculating the differences, further analysis or processing steps can be utilized for the resulting difference image  $D$  to identify and delineate the regions where significant changes have occurred. This may involve thresholding, clustering, or other segmentation techniques to distinguish between areas with minor variations and regions that represent significant changes. The specific implementation of image difference-based change detection may vary based on the characteristics of the provided images, the characteristics of the changes being detected, and the

application's requirements. Various pre-processing and post-processing steps might be incorporated into the workflow to improve the precision and dependability of the outcomes in change detection.

### 3.4.2. Image Ratioing

Image ratioing calculates the proportion of pixel values between two images. The ratio image highlights areas with significant changes, and it is less sensitive to radiometric differences (Singh, 1989). This approach is commonly used in remote sensing and land cover analysis to detect changes over time, especially in environmental monitoring, urban development, or agricultural assessment. It is frequently employed for vegetation health monitoring, as it can be sensitive to changes in vegetation properties. The equation for image ratioing-based change detection can be described as follows.

$$R(i, j) = \frac{I_1(i, j)}{I_2(i, j)} \quad (3.32)$$

where,  $R(i, j)$  represents the computed ratio at location  $(i, j)$  between the corresponding pixels in the two input images,  $I_1(i, j)$  and  $I_2(i, j)$  denotes the pixel values at location  $(i, j)$  in the initial and subsequent input images respectively.

After computing the ratio image, further analysis or processing steps are generally employed to detect regions with notable changes using the ratio values. This might involve setting thresholds, performing statistical analysis, or applying classification techniques to differentiate between regions exhibiting substantial changes and areas where changes are insignificant. The specific implementation of image ratioing-based change detection can vary based on the characteristics of the input images, the specific application, and the nature of the changes being detected. Pre-processing and post-processing steps may be integrated into the workflow to enhance the accuracy and reliability of the change detection results, such as noise reduction, normalization, or filtering.

### 3.4.3. Principal Component Analysis (PCA)

PCA is a technique for reducing dimensionality that can be applied to multispectral or hyperspectral data. It transforms the original data into a set of uncorrelated principal components. Change detection is performed on the principal component images, as changes are often more evident in this reduced-dimensional space. Principal Component Analysis (PCA) is applicable to

change detection in multi-temporal remote sensing images. The technique involves analyzing the principal components derived from the images to identify changes in the spectral or spatial characteristics between the images (Deng et al., 2008). The equation for PCA-based change detection can be represented as follows.

- i. Compute the mean of the multi-temporal images.

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (3.33)$$

where,  $\bar{X}$  is the mean image and  $X_i$  are the individual images.

- ii. Subtract the mean from each image to obtain the anomaly image.

$$A_i = X_i - \bar{X} \quad (3.34)$$

- iii. Stack the anomaly images into a matrix

$$A = [A_1, A_2, \dots, A_N] \quad (3.35)$$

- iv. Compute the covariance matrix of the stacked anomalies.

$$C = \frac{1}{N-1} AA^T \quad (3.36)$$

- v. Perform eigenvalue decomposition on the covariance matrix to obtain the eigenvectors and eigenvalues.

$$C = V\Lambda V^T \quad (3.37)$$

where,  $V$  is the matrix of the eigen vectors and  $\Lambda$  is the diagonal matrix of eigenvalues.

- vi. Analyze the eigenvalues and eigenvectors for identifying the principal components representing the dominant change information.

PCA-based change detection typically involves analyzing the principal components and their corresponding eigenvalues to determine the significant changes between the images. The principal components with the largest eigenvalues represent the most significant changes, while those with lower eigenvalues correspond to less significant or noise-related changes. The specific thresholding or decision-making process for change detection depends on the application and the characteristics of the images under analysis.

### 3.4.4. Multispectral Indices

Multispectral indices are often used in remote sensing to highlight specific features or characteristics of the land's or vegetation cover. They can alternatively be helpful in change detection analysis, where environmental changes are identified based on alterations in these indices between different periods. Multispectral indices, such as the Normalized Difference Vegetation Index (NDVI) or the Enhanced Vegetation Index (EVI), are computed for each image in a time series. Changes are detected by comparing the index values between images. These indices are especially valuable for monitoring changes in vegetation cover (Lu et al., 2004). The equation for multispectral indices-based change detection typically involves the calculation of the index for each image at different periods and comparing the various index values to identify significant changes. Here is the general representation of the equation for this objective, using the Normalized Difference Vegetation Index (NDVI) as an example.

- i. Calculate the NDVI for each period.

$$NDVI = \frac{(NIR-Red)}{(NIR+Red)} \quad (3.38)$$

where, NIR represents the reflectance in the near infrared band and Red represents reflectance in the red band.

- ii. Compute the difference in NDVI values between these two time periods.

$$Change = NDVI_{t_2} - NDVI_{t_1} \quad (3.39)$$

The resulting change values can be utilized to recognize areas where significant changes in vegetation cover or other environmental factors have taken place between the two temporal intervals. Thresholding or other analysis techniques can be applied to determine the magnitude and location of these changes. The specific implementation of multispectral indices-based change detection can vary depending on the particular indices used and the attributes of the data being analyzed. Different indices may be selected based on the specific properties or interest changes in the study area.

### 3.4.5. Fuzzy Logic and Expert Systems

Fuzzy logic and expert systems incorporate domain knowledge and uncertainty into change detection processes. These methods can be effective in handling imprecise or uncertain data. Fuzzy logic and expert systems are often employed in change detection applications to handle the uncertainty and imprecision frequently linked to remote sensing data. They can aid decision-making by integrating domain knowledge and linguistic variables into the change detection process. Integrating fuzzy logic and expert systems allows for the modeling of complex and imprecise relationships between input variables. It facilitates a more nuanced comprehension of the change detection problem. The specific equations for fuzzy logic and expert systems-based change detection can involve formulating fuzzy rules, membership functions, and decision-making processes (Metternicht, 1999). The terms used for fuzzy logic-based change detection are as follows.

- i. Fuzzification: Convert the crisp input data (e.g., spectral values, indices) into fuzzy sets via membership functions, such as triangular, trapezoidal, or Gaussian functions.
- ii. Rule-based System: Define a set of fuzzy rules that describe the relationships between the input variables and the output (change or no-change) based on expert knowledge or empirical observations. These rules can appear as "if-then" statements that capture the linguistic variables and their interdependencies.
- iii. Inference Mechanism: Apply fuzzy logic inference rules, such as Mamdani or Sugeno methods, to aggregate the fuzzy rule outputs and determine the overall degree of change or no-change based on the input data.
- iv. Defuzzification: Convert the fuzzy output into a crisp value or decision using defuzzification methods, such as centroid or mean of maximum.

The fuzzy logic and expert systems-based change detection process can include the definition of membership functions, the calculation of fuzzy rule outputs, and the aggregation of fuzzy outputs to generate the final decision. These equations incorporate linguistic terms and domain knowledge, enabling a more interpretable and human-like approach to change detection that can handle the inherent uncertainties in remote sensing data.

### 3.4.6. Deep Learning based Change Detection

Change detection based on deep learning involves using deep neural networks to detect and identify changes between two or more images. The specific equation for deep learning-based change detection depends on the architecture and design of the neural network used for the task (Shafique et al., 2022). Typically, the process involves training a Convolutional Neural Network (CNN) or another deep learning model on a dataset of paired input images and their corresponding change labels.

Here is a general framework for deep learning-based remote sensing change detection.

- i. **Input Representation:** Represent the input images as tensors and feed them into the deep learning model. The input tensors may undergo preprocessing, normalization, or augmentation depending on the particular requirements of the model.
- ii. **Convolutional Layers:** The deep learning model comprises several convolutional layers that extract hierarchical features from the input images. These layers perform operations such as convolutions, activation functions, pooling, and normalization to learn relevant patterns and features.
- iii. **Fully Connected Layers:** The extracted features are subsequently flattened and passed through fully connected layers that perform classification or regression tasks to predict the existence or nonexistence of changes in the input images.
- iv. **Loss Function:** Specify an appropriate loss function, such as binary cross-entropy, mean squared error, or another relevant loss function, to quantify the disparity between the predicted change labels and the actual ground truth labels.
- v. **Optimization:** Use an optimization algorithm, such as stochastic gradient descent, Adam, or RMSprop, to reduce the loss function error and update the model's weights and biases during the training process.
- vi. **Training and Evaluation:** Train the deep learning model on the labeled dataset and evaluate its performance on a separate validation or test dataset. Assessing the model's performance can involve using metrics like accuracy, precision, recall, or F1 score.

The specific architecture, hyper-parameters, and optimization algorithms used in deep learning-based change detection may differ depending on the complexity of the task, the characteristics of

the input data, and the particular aims of the change detection application. The equations used in deep learning-based change detection are tailored to optimize the model's performance in detecting and identifying changes in the input images. Deep learning techniques, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have witnessed increased adoption in change detection (Wang et al., 2018; Mou et al., 2018). These methods have the capability to automatically learn intricate patterns and features from multi-temporal image data. The selection of a change detection technique is influenced by various factors, such as the characteristics of the data, the scale of the changes, the available computational resources, and the specific objectives of the analysis. Frequently, the application of these techniques aims to improve the accuracy and reliability of change detection outcomes.

### **3.5. Overall Methodology**

The comprehensive approach to remote sensing image processing workflow involves seamlessly integrating image registration, fusion, and change detection to extract valuable information. Image registration aims to determine a geometric transformation that aligns one image with another, ensuring that corresponding pixels in both images accurately represent identical ground features. The co-registered data pair is input to the image fusion process that combines two or more images of the same scene, taken from different sensors, to create a merged image that holds more information than any of the individual images. In the final phase of the remote sensing image processing pipeline, the change detection process identifies the changes from the multi-temporal remote sensing fused images that have occurred in a particular region of interest. Figure-3.7 below shows the remote sensing image processing workflow proposed to monitor land cover changes, urban development, natural disasters, and other environmental changes.



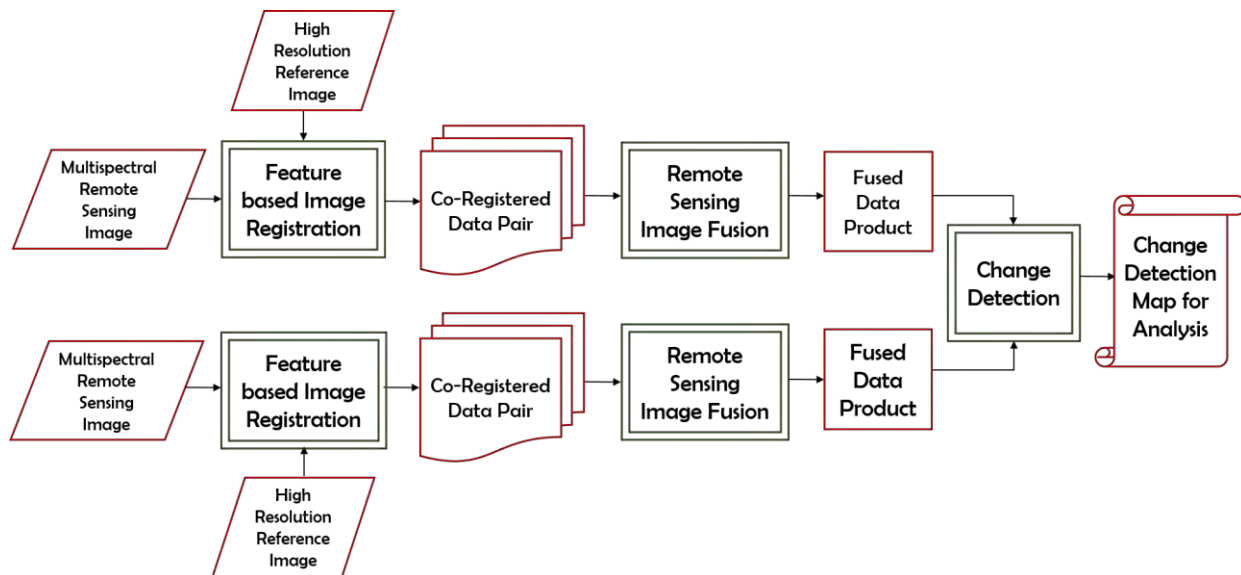


Figure-3.7. Remote Sensing Image Processing Workflow to generate value added information

### 3.6. Summary

In this chapter, the theory behind feature based image registration, which is a fundamental step in the analysis of remote sensing images, is described. It involves feature extraction, matching, outlier removal, transformation estimation, and alignment to enable the meaningful comparison of images for a wide range of space borne applications. Remote sensing image fusion is a critical process that enhances the utility of remotely sensed data by integrating information from multiple sources. It enables better-informed decision-making and analysis across various domains, contributing to improved understanding and management of the planet’s surface and its changes over time. The chapter also describes the theory behind remote sensing change detection, a powerful tool for monitoring and managing dynamic environments. It can detect land cover changes, vegetation health, urban expansion, deforestation, natural disasters, and other planetary exploration applications. Remote sensing image registration, fusion, and change detection are potent tools that can be amalgamated to extract valuable information from remote sensing images. By combining these techniques, we can gain a more comprehensive understanding of the changes occurring on the planet’s surface.

## **4. FEATURE BASED IMAGE REGISTRATION TECHNIQUES**

### **4.1. Multi-Stage Feature Detection and Improved Outlier Removal method**

#### **4.1.1. Introduction**

The following section introduces FIRM, a multistage image registration workflow specifically designed for Resourcesat images. FIRM, an acronym for Framework for Image Registration using Multistage Feature Detection and Mode Guided Motion Smoothness Keypoint Optimization, combines Mode Guided Tiled Scale Invariant Feature Transform (MT-SIFT) integrated with Patch Affine Oriented Fast and Rotated Brief (PA-ORB). This innovative approach incorporates the precision optimization of keypoints with a constraint on the smoothness of motion. Notably, FIRM excels in handling substantial location errors in the initial stages and achieving precise adjustments in later stages, ensuring sub-pixel co-registration accuracy.

The effectiveness of FIRM is demonstrated through testing on over fifty LISS-4 multi-temporal image scenes. The products exhibit an agreement of below a pixel in comparison to the reference, helps to monitor the dynamics of wetlands and predicting crop assessments. Additionally, the rectified result is validated through cross-referencing with an independent and widely accepted geospatial data platform, such as Google Earth. The proposed method FIRM is applied to the co-registration of multi-satellite images, specifically LISS-4 from Resourcesat-2 and OLI from Landsat-8 (Roy et al., 2014). Furthermore, FIRM is employed for aligning pairs of LISS-3 images with a decade-long time gap at a sub-pixel accuracy, with quantifiable outcomes intended for comparison with innovative feature-based approaches.

#### **4.1.2. Methodology Developed**

It is found from the input data that the spatial discrepancy of LISS-4 compared to the reference image can range from a few meters to more than one kilometer in specific instances due to platform instability. Co-registering LISS-4 data over rugged terrain poses a significant challenge because of varying tilt angles during imaging and the comparatively higher spatial resolution. In contrast, the LISS-3 camera on Resourcesat, which looks directly downward (nadir view), and typically exhibits superior geometric performance, operates within a systematic referencing scheme, systematically covering the identical region periodically.

However, achieving sub-pixel multi-temporal co-registration with LISS-3 for image pairs with a decade or more difference in acquisition time proves to be a laborious task. On the other hand, the AWiFS camera boasts a wide field of view, leading to extensive coverage of the area during the imaging process. Nonetheless, the predominant internal geometric distortion along the scan line in AWiFS could lead to feature misalignment in the overlapping region of two consecutive across-track multi-temporal scenes.

Handling the various spatial inaccuracies inherent in Resourcesat payloads poses a challenge in achieving sub-pixel level co-registration through geometric calibration alone. The subsequent steps outline the development of the FIRM framework, designed to enhance the geometric location accuracy of Resourcesat sensors and improve multi-temporal sub-pixel co-registration accuracy using image-based rectification techniques.

#### *A. Dynamic Patch Generation*

The geo-corrected image at the system level from the LISS-4 input has a substantial size of approximately 18,000 \* 18,000 pixels. Managing such a large image poses challenges in terms of feature matching and also significantly extends the execution time required for retrieving matched key points. In our methodology, we address this issue by extracting depth-wise patches of dimensions 1024 \* 1024 pixels from both the input and reference images, utilizing the spatial details linked to every pixel in both images. Matching features is then conducted, establishing correspondence of keypoints specifically within these patches. This approach effectively decreases the search radius for acquiring corresponding points, consequently minimizing execution time. The patch generation technique can be applied similarly to LISS-3 and AWiFS sensors owing to their substantial image dimensions.

#### *B. Gross Error Estimate using Patch Affine ORB*

The estimation of the gross spatial discrepancy in LISS-4 is accomplished through the utilization of the Patch Affine Oriented Fast and Rotated Brief (PA-ORB) feature matching technique. ORB, as introduced by Rosten and Drummond in 2006 and 2008, is a fusion of methodologies: the Features from Accelerated Segment Test (FAST) feature detector and the Binary Rotated Independent Elementary Feature (BRIEF) descriptor by Calonder in 2010. This combination demonstrates the capability to identify consistent and matched keypoints in remote sensing images. FAST operates with a single parameter, the intensity threshold between the center pixel and those

within a circular ring surrounding it. In our approach, we specifically select 16 contiguous pixels around the circle to assess the intensity threshold between the center pixel and those positioned on the circular ring.

The ultimately selected and refined matched points in image coordinates are subsequently converted into geographic coordinates through the application of a geometric transformation model, as indicated in Equation-4.2, which is established for geo-tagged source and reference images. Leveraging these corresponding geographical coordinates, affine transformation parameters are computed using Equation-4.3. The geographic grid of the input image is then updated, mitigating coarse geo-location errors by aligning the input image with the reference geospace. The algorithmic steps to reduce the gross geometric errors in Resourcesat images are outlined in Algorithm-4.1.

#### **Algorithm-4.1**

1. Traverse the image  $I_{x,y}$  depth-wise.

where  $x$  represents number of image pixels along the width axis,

and  $y$  represents the number of image pixels along the height axis.

2. Retrieve patches from the input Resourcesat image  $P_{m,n}$  and its corresponding reference  $Pr_{m,n}$ .

where  $m$  denotes the number of patch pixels along the width axis,

and  $n$  denotes the number of patch pixels along the height axis.

In our situation, the patch dimensions are set at 1024 by 1024 pixels.

3. Retrieve the corresponding control points using the patch ORB method.

$$CP_k = PMATCH(P_{m,n}, Pr_{m,n}) \quad (4.1)$$

where  $k$  denotes the count of matched control points in a patch correspondence.

4. Convert the matched control points from  $CP_k$  in patch image coordinates to  $GCP_k$  in geographic coordinates.

$$Geo-Control\ Points\ GCP_k = GEO\_TRANSFORM(CP_k) \quad (4.2)$$

5. Repeat steps 1 to 4 iteratively until the final input patch is reached.

6. Determine the parameters for the global affine transformation.

$$a_i = Aff\_Trans(GCP_l) \quad (4.3)$$

where  $i$  represents the count of affine parameters, and

$l$  represents the overall count of matched control points in the image.

### C. *Motion Smoothness Constraint (MSC)*

The utilization of ORB for matching feature points reveals a notable presence of false matches, introducing errors in the transformation parameters. Estimation based on motion smoothness, as outlined by Bian in 2017, proves effective in distinguishing between genuine and incorrect matches, ensuring precise relationship between the input and reference images. The concept of motion smoothness assumes that a small neighborhood near a genuine match should be recognized as the identical location in the other image, while the surroundings near a false match signify geometrically distinct locations. In cases of smooth motion, neighboring pixels and features exhibit cohesive movement.

The Motion Smoothness Constraint (MSC) is translated into a statistical model that discriminates against false matches. MSC, characterized by its simplicity, robustness, and superior performance, surpasses alternatives such as RANSAC (Derpanis, 2010) and other keypoint optimization techniques (Pizarro & Bartoli, 2012).

To apply the MSC model, the input LISS-4 patch and the reference patch are divided into grid cells measuring 32 by 32 pixels. Every cell within the input LISS-4 patch is examined, and the count of matched points in that cell is determined concerning every reference cell. The reference patch is also divided into equally spaced cells. Subsequently, the objective is to calculate the quantity of matched points that fall within each reference cell corresponding to the selected input cell and assign a score to each reference cell.

Choosing true match pairs relies on identifying the maximum number of matching combinations between the input and reference cells. If the maximum score for an input cell surpasses the threshold limit, only the matched feature points from that cell are identified as inliers. The choice of a 32\*32-pixel grid cell has proven to yield optimal results in FIRM, as it encompasses a sufficiently large neighboring area surrounding a matched feature point. The threshold limit is dynamically established from the image, depending on the average number of matched points within a grid cell.

In Figure-4.1, the input LISS-4 cell, outlined with a yellow border, displays the highest number of matches (depicted in green) with a particular reference cell (also outlined in a yellow border). On the flip side, the other reference cell has a lower count of matched points (depicted in red), suggesting the presence of false matches. The MSC technique can be employed on both LISS-3

and AWiFS images to eliminate outliers and enhance the accuracy of transformation parameter estimation. The procedural steps of the MSC model are outlined in Algorithm-4.2.

**Algorithm-4.2**

1. For every grid cell in the input LISS-4 patch,  
where  $i$  represents the index of the input grid cell.
2. Calculate the score  $S_{ij} = |I_i R_j|$  (4.4)  
where  $S_{ij}$  represents the count of matches occurring between the input grid cell  $I_i$  and the reference grid cell  $R_j$ , and  $j$  indicates the index of the reference grid cell.
3. Increase  $j$  from 1 to 36 to scan through all the reference grid cells in a patch and calculate the score  $S_{ij}$ .
4. Calculate the maximum score  $MS_{ij} = (\max_{j=(1,\dots,36)} S_{ij})$  (4.5)
5. If  $MS_{ij} > TL$ , choose the matched feature points as inliers.  
where  $TL$  denotes the threshold limit.
6. Increase  $i$  and iterate through steps 1 to 5 until reaching the final input grid cell.

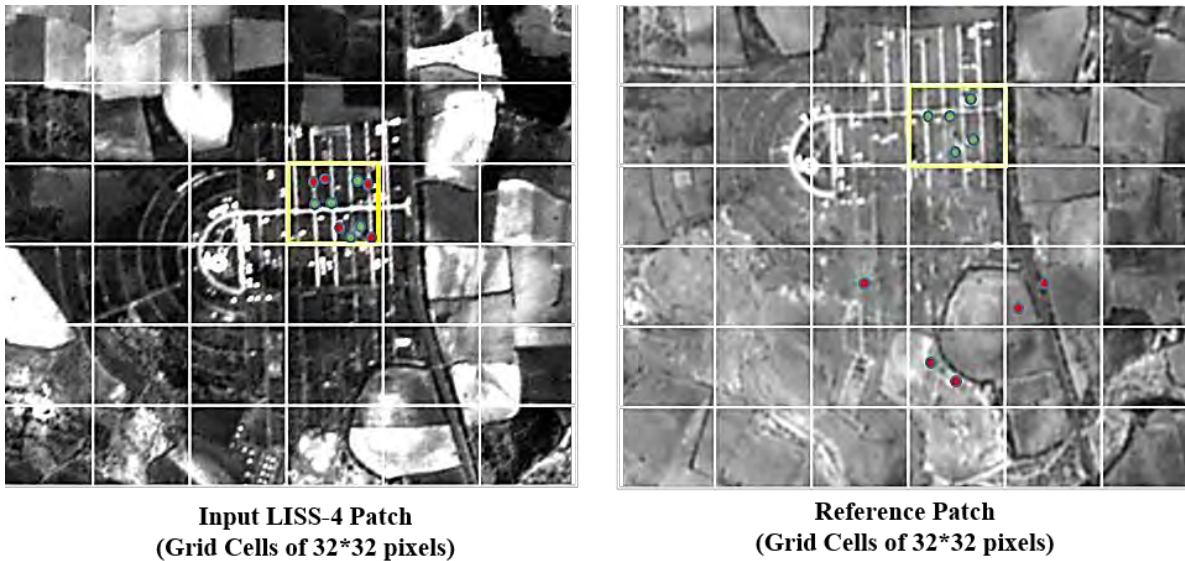


Figure-4.1. Pruning of Matched Points Utilizing Motion Smoothness Constraint

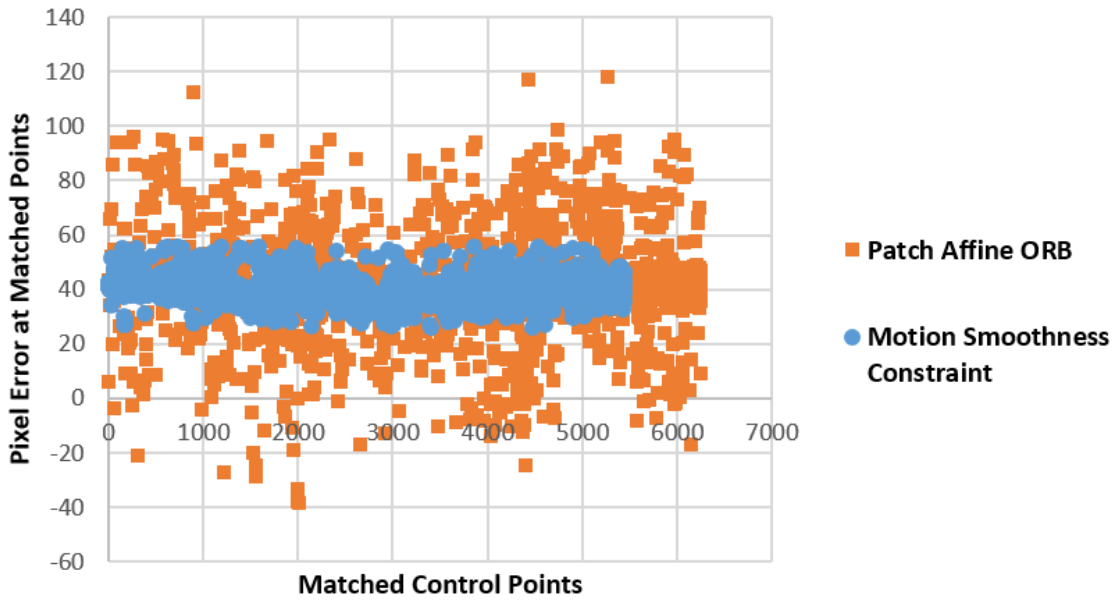


Figure-4.2. Joint Scatter Plot Illustrating Pruned Matched Points with MSC

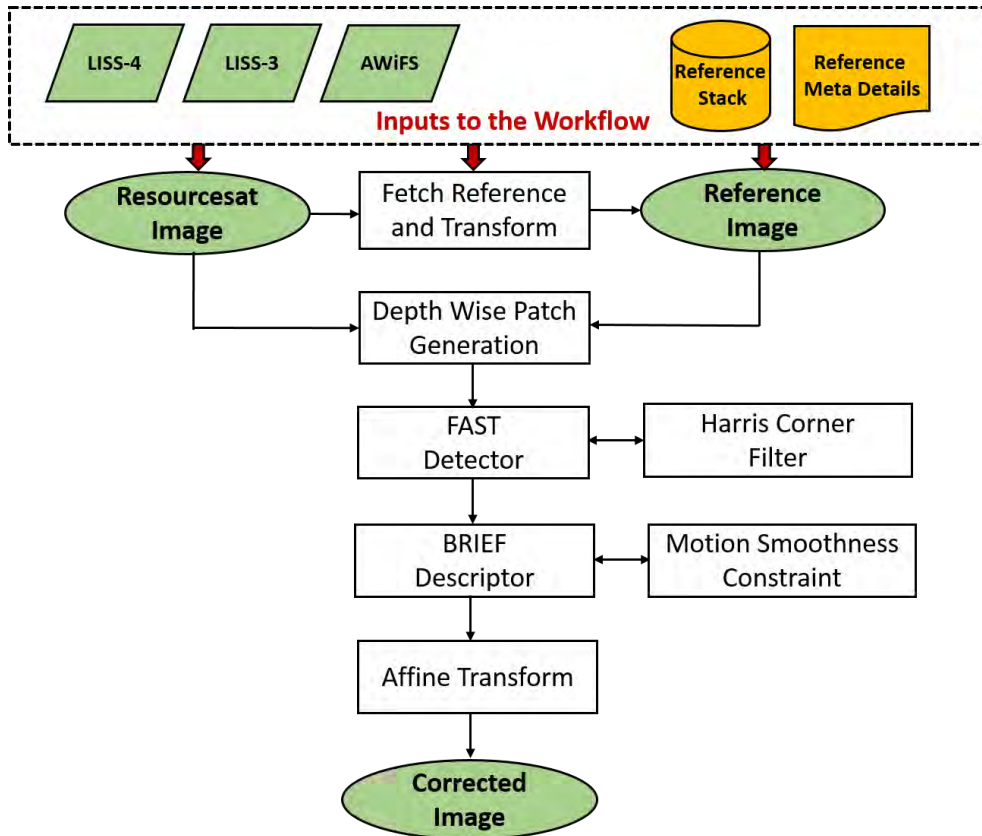


Figure-4.3. Workflow for Coarse Image Co-Registration

To evaluate the efficiency of MSC in LISS-4 feature correspondence, we apply the MSC estimation filter to the matched control points obtained through patch affine ORB. Figure-4.2 shows that patch affine ORB contains many outlier-matched points (highlighted in orange in the scatter plot), which might introduce undesired errors in the estimation of transformation parameters. MSC serves as a filter, removing the outlier points and generating refined matched control points (shown in blue in the scatter plot) for a more accurate transformation estimation.

The complete process for coarse image co-registration is illustrated in Figure-4.3. The inputs for the workflow include Resourcesat images captured by LISS-4, LISS-3, and AWiFS sensors. The workflow receives both the reference stack and its metadata to concurrently fetch the corresponding reference and adjust it based on the specific sensor's spatial resolution. A patch with depth information undergoes generation and proceeds through the FAST detector, followed by filtering based on the Harris Corner measure. The BRIEF descriptor is utilized for feature matching based on Euclidean distance, and keypoints undergo optimization with the motion smoothness constraint. An affine transformation model applied to the input Resourcesat image, producing a corrected image that aligns more closely with the spatial characteristics of the reference image.

#### *D. Fine Correspondence using Mode Guided Tiled SIFT*

The utilization of ORB and the motion smoothness constraint is noted to establish a general correspondence, bringing input Resourcesat images to a pixel-level discrepancy in comparison to the reference image. To enhance the precision of image feature matching to a sub-pixel level, the mode-guided tiled Scale Invariant Feature Transform (MT-SIFT) model is employed, as proposed by Kupfer et al. in 2014.

One advantage of SIFT lies in its generation of keypoints associated with a 128-element descriptor in the image. This descriptor vector is employed for matching features between the input and reference images, enabling the creation of dependable matched keypoints. In the fine-stage processing, the transformed image with reduced coarse errors and the corresponding reference image are aligned using SIFT. In our methodology, we utilize an interleaved tile-based strategy for feature matching, facilitating the extraction of reliable keypoints while reducing the matching time. A tile dimension of 128 by 128 pixels is selected, showcasing optimal performance for precise correspondence in LISS-4 images. Through experimentation, it has been established that



this identical tile size provides an ample number of keypoints for matching in LISS-3 and AWiFS images. Figure-4.4 (a) depicts the adopted scheme for tile interleaving in SIFT feature detection.

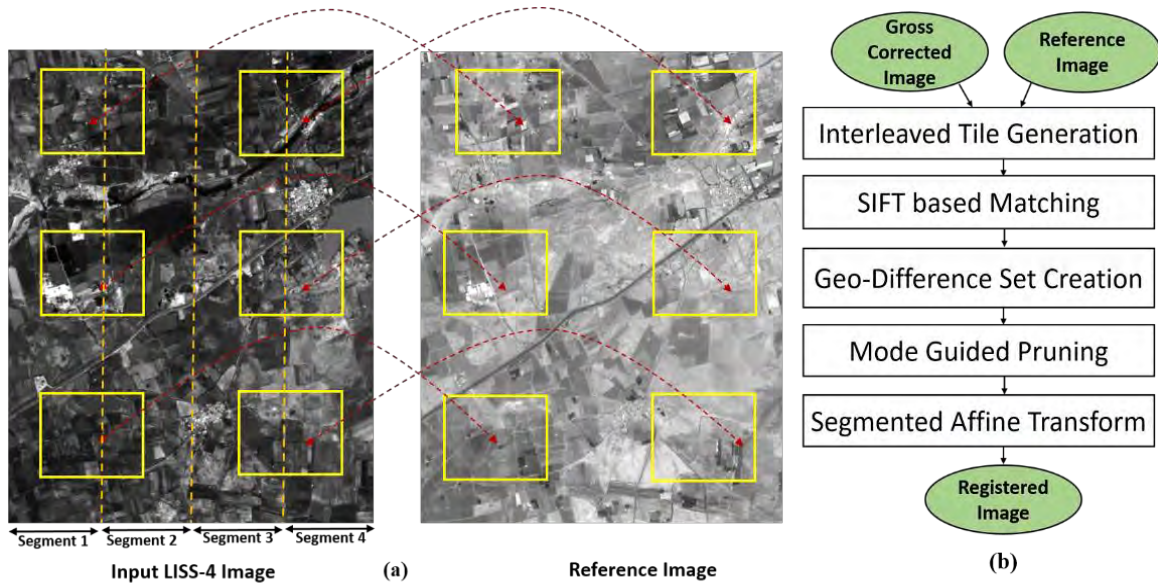


Figure-4.4. (a) Segmented Affine Transform Estimation through Tiled SIFT Matching

(b) Workflow for Fine Image Co-Registration

The sets of differences are generated for both the along-track and across-track directions for each corresponding keypoint. A method guided by the mode is utilized to eliminate outliers. The mode is determined from the set of difference image points, and a threshold limit is set to distinguish the authentic matched points. Based on experimental results, it has been observed that setting a threshold limit of 2 effectively identifies all the intended matched points in Resourcesat images. The keypoints that have been pruned are employed for the estimation of model parameters.

A segmented affine transformation model is employed to calculate affine values for individual image segments. This, in turn, updates the geometric grid for each segment, effectively mitigating local distortion. The processing workflow for fine image registration is illustrated in Figure-4.4 (b). The process begins with the interleaved tile generation for both the gross corrected image and the reference image. Subsequently, SIFT matching is executed, leading to the creation of a geometric difference set. The pruning module guided by the mode removes outliers from the set of keypoints. The image is divided into four segments, and distinct affine parameters are calculated to tackle local distortion in each segment. The calculated segmented affine transformation

parameters are subsequently utilized to update the geometric grid for each image segment. This process involves resampling the input image, resulting in a sub-pixel co-registered output. The expression for the affine transformation in a segment is as follows:

$$aff(k) = \{a_{0k}, a_{1k}, a_{2k}, a_{3k}, a_{4k}, a_{5k}\} \quad (4.6)$$

where  $k$  represents the segment identifier, and  $a_{ik}$  represents the affine parameters.

The calculated accurate geographical coordinates ( $lat_{cor}, lon_{cor}$ ) for the image point  $(x, y)$  derived from the input geographic coordinates  $(lat, lon)$  and the affine parameters  $a_{ik}$  for segment number  $k$ , are:

$$\begin{bmatrix} lat_{cor} \\ lon_{cor} \end{bmatrix} = \begin{bmatrix} a_{0k} & a_{1k} \\ a_{2k} & a_{3k} \end{bmatrix} \begin{bmatrix} lat \\ lon \end{bmatrix} + \begin{bmatrix} a_{4k} \\ a_{5k} \end{bmatrix} \quad (4.7)$$

The Resourcesat image after undergoing transformation in this second phase successfully attains sub-pixel level co-registration with a standard reference, elevating the accuracy of geometric location. The detailed steps of the fine image registration processing procedure are delineated in Algorithm 4.3.

### **Algorithm 4.3**

1. Traverse the image  $I_{x,y}$  tile-wise.
2. Retrieve the given tile  $T_i$  and the corresponding reference tile  $R_i$  using geographic info.  
where  $i$  denotes the tile number.
3. Utilize SIFT to match the tiles and create corresponding control points.

$$SCP_i = SIFTMATCH(T_i, R_i) \quad (4.8)$$

where  $SCP$  represents Control points identified by SIFT,

and  $l$  signifies the number of corresponding points.

4. Iterate through steps 1 to 3 until processing the final input tile.  $T_i$ .
5. Create sets of differences  $D_x$  and  $D_y$ .  
 $D_x$  denotes the Set of difference images in the along-track direction.  
 $D_y$  denotes the Set of difference images in the across-track direction.
6. Calculate the Mode  $M_{x,y}$  using the difference sets  $D_x$  and  $D_y$ .
7. Establish threshold values  $T_{x,y}$  to filter outliers and produce the pruned geometric set  $PGSet_m$ , where  $m$  denotes the number of pruned geometric control points.

8. Affine transformation parameters  $aff(k)$  for segmentation are estimated using the pruned geometric point set  $PGSet_m$ .

$$aff(k) = Aff\_Trans\_Seg(PGSet_m) \quad (4.9)$$

where  $k$  represents the image segment number (In this instance,  $k=4$ ).

The pivotal components for achieving sub-pixel co-registration in medium-resolution Resourcesat images lies in the accurate processing phase comprises of hierarchical feature detection and dependable optimization models. The gross processing workflow adeptly manages substantial geometrical errors, effectively enhancing the spatial positioning of pixels within the geographical space. The fine processing phase deals with any remaining errors in the image, ensuring alignment at sub-pixel level co-registration accuracy.

FIRM demonstrates proficiency in handling medium-resolution remote sensing images obtained from various sensors and excels in co-registering image pairs with significant acquisition time differences. The image registration process in FIRM utilizes the Near Infrared (NIR) channel from both the input and reference images. The transform parameters estimated are then applied to all wavelength channels to generate co-registered multispectral data products. Figure-4.5 depicts the structure of FIRM designed for aligning multi-temporal remote sensing images.

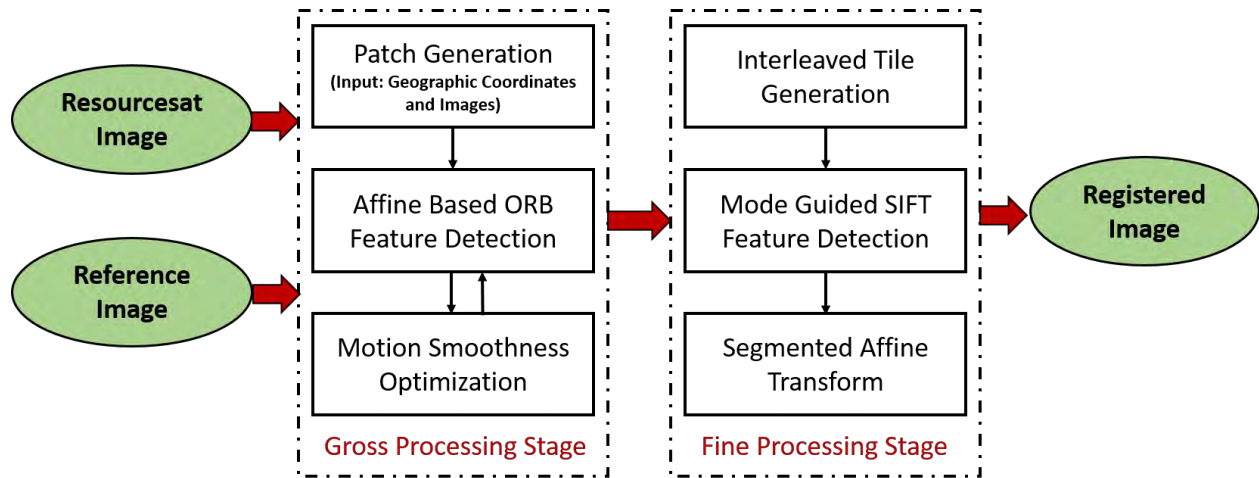


Figure-4.5. Framework for Precision Correction using FIRM

### 4.1.3. Experimental Results

#### A. Assessment of LISS-4 data over diverse landscapes

FIRM undergoes thorough testing using fifty one Resourcesat-2 LISS-4 scenes captured in October 2019, showcasing a variety of features like deserts, flora, aquatic areas, snow, elevated

terrains, and coastal zones. The georeferenced data products of LISS-4 images at the Level-1 system level are projected in the UTM map projection. Given their substantial geometric location errors, the gross processing stage aims to identify and address these errors, followed by further refinement at the sub-pixel level during the refined processing phase.

The reference stack, along with metadata, is utilized to acquire the corresponding reference image based on geographic corner coordinates of the input LISS-4 scenes. The process flows via the established framework depicted in Figure-4.5. The inclusion of varied landscapes facilitates the assessment of image co-registration performance across various feature targets. Figure-4.6 depicts the geometric extent of the LISS-4 scenes utilized to assess the software chain. Here in the figure explorer base map (NASA Visible Earth, 2020) overlays LISS-4 False Colour Composite (FCC) images used for evaluation. The yellow-tiled area on the map is magnified to emphasize specific regions of interest, encompassing nearly all notable target features. Moreover, cloudy LISS-4 scenes are taken into account to confirm the effectiveness of FIRM.

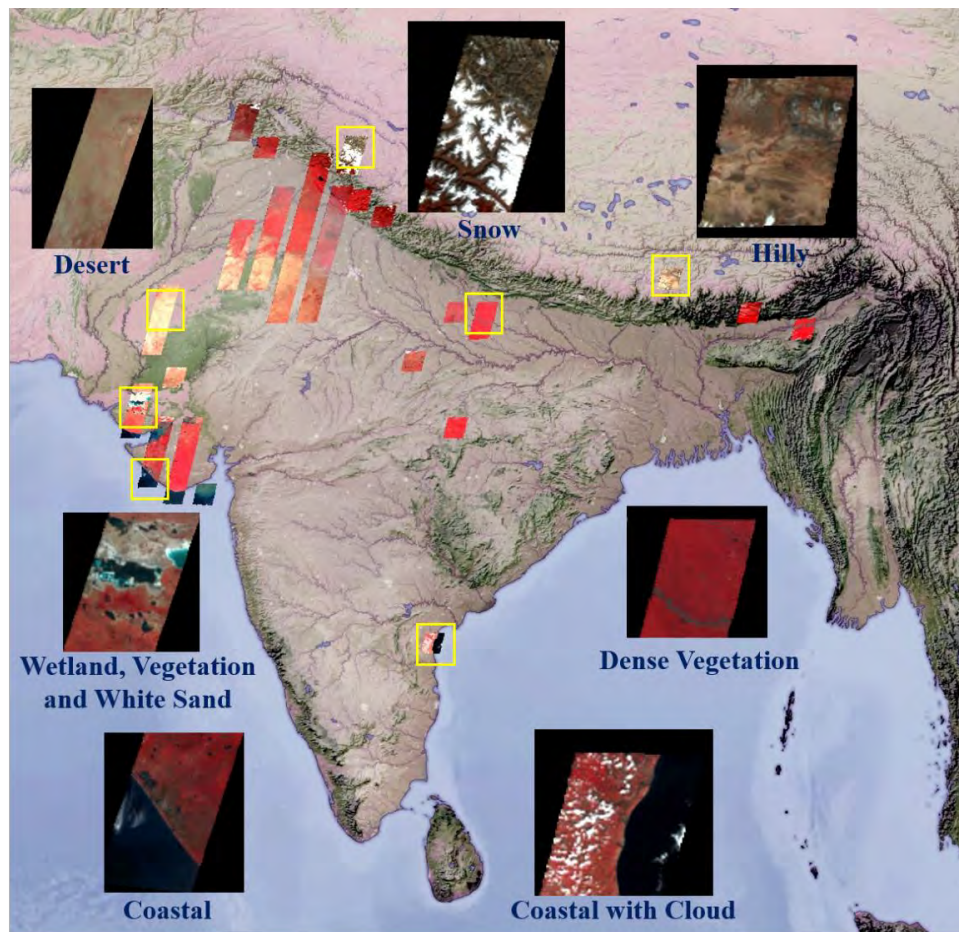


Figure-4.6. Geographical Extent of Resourcesat LISS-4 Scenes Employed in Co-Registration

Table-4.1 presents details of LISS-4 scenes, including sub-scene, acquisition date, and details regarding the path and row. The Root Mean Square Error (RMSE) stands as the preferred metric for evaluating the image co-registration performance of Resourcesat images, as delineated in our specific case:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|^2} \quad (4.10)$$

where,  $N$  denotes the overall count of corresponding points,

$X_i$  represents the coordinates  $(x_i, y_i)$  of the Reference image, and

$\hat{X}_i$  signifies the coordinates  $(\hat{x}_i, \hat{y}_i)$  of the input/output Resourcesat image.

The Root Mean Square Error (RMSE) is calculated in pixels both prior to and following the application of the correction to LISS-4 images using FIRM. It is evident that following accurate geometric correction, the RMSE consistently remains below one pixel for all LISS-4 scenes. For a more comprehensive understanding, the pixel error in relative location is computed for input LISS-4 scenes in both along-track and across-track directions.

The geometric location error at the system level prior to correction across multiple temporal acquisitions is depicted in Figure-4.7. The mean pixel error in the across-track direction is 24 pixels, whereas the mean pixel error along the track for the same group of LISS-4 scenes is 90 pixels. The considerable location discrepancy is due to the uncertainties in spacecraft perturbations in the Resourcesat-2 satellite. Resourcesat-2, launched in April 2011, has concluded its operational lifespan (Resourcesat, 2011). Nevertheless, the quality of radiometric data across diverse wavelength spectrums from Resourcesat-2 sensors remains consistently high across various target features.

Following accurate geometric adjustment, the accuracy of the output LISS-4 data products is re-evaluated using the same reference datasets. The results indicate the achievement of registration accuracy for LISS-4 scenes at a sub-pixel level. In Figure-4.8, the sub-pixel positional accuracy of LISS-4 scenes is depicted in both along-track and across-track directions for different acquisition dates. The mean Root Mean Square Error (RMSE) post-correction for the 51 LISS-4 scenes is 0.21 pixels. This finding serves as a robust indicator, demonstrating that the LISS-4 output products are aligned at a sub-pixel level aligned with the reference datasets. The strength of the suggested method, FIRM, is underscored by its ability to correct pixel shift errors and adjust the input image into sub-pixel alignment with the reference.

Table-4.1. Details of LISS-4 Scenes and Pixel Error RMSE Before and After Correction

<b>S.No.</b>	<b>Image Details [Subscene-Date of Acquisition (DD-MON-YYYY)- Path-Row]</b>	<b>Input RMSE (in pixel)</b>	<b>Output RMSE (in pixel)</b>
1.	A-08-OCT-2019-101-053	156.01	0.38
2.	A-08-OCT-2019-101-056	152.64	0.09
3.	A-11-OCT-2019-092-050	105.75	0.80
4.	A-11-OCT-2019-092-051	108.46	0.06
5.	A-11-OCT-2019-092-056	60.95	0.12
6.	A-11-OCT-2019-092-057	20.80	0.18
7.	A-15-OCT-2019-112-053	77.20	0.15
8.	A-16-OCT-2019-093-049	89.67	0.73
9.	A-16-OCT-2019-093-050	105.30	0.03
10.	A-16-OCT-2019-093-051	110.00	0.13
11.	A-16-OCT-2019-093-058	69.18	0.83
12.	A-21-OCT-2019-094-048	74.24	0.18
13.	A-21-OCT-2019-094-049	93.02	0.74
14.	A-21-OCT-2019-094-050	95.00	0.14
15.	A-21-OCT-2019-094-051	89.00	0.75
16.	A-21-OCT-2019-094-052	93.08	0.15
17.	A-25-OCT-2019-090-052	104.40	0.10
18.	A-25-OCT-2019-090-053	125.31	0.14
19.	A-25-OCT-2019-090-055	145.91	0.02
20.	A-25-OCT-2019-090-056	148.51	0.22
21.	A-26-OCT-2019-095-048	32.64	0.79
22.	A-26-OCT-2019-095-049	82.60	0.09
23.	A-26-OCT-2019-095-050	110.31	0.36
24.	A-26-OCT-2019-095-051	137.78	0.49
25.	A-26-OCT-2019-095-052	141.32	0.01
26.	B-02-OCT-2019-095-049	38.27	0.42
27.	C-08-OCT-2019-101-052	141.00	0.66
28.	C-11-OCT-2019-092-050	106.60	0.04
29.	C-11-OCT-2019-092-051	111.61	0.13

30.	C-11-OCT-2019-092-055	70.03	0.09
31.	C-11-OCT-2019-092-056	50.21	0.45
32.	C-13-OCT-2019-102-062	46.17	0.18
33.	C-14-OCT-2019-107-051	119.85	0.07
34.	C-16-OCT-2019-093-049	105.92	0.47
35.	C-16-OCT-2019-093-050	120.06	0.15
36.	C-16-OCT-2019-093-051	120.00	0.23
37.	C-21-OCT-2019-094-048	86.00	0.16
38.	C-21-OCT-2019-094-049	96.02	0.16
39.	C-21-OCT-2019-094-050	99.00	0.06
40.	C-21-OCT-2019-094-051	87.09	0.24
41.	C-21-OCT-2019-094-052	66.75	0.80
42.	C-25-OCT-2019-090-052	115.52	0.04
43.	C-25-OCT-2019-090-053	123.98	0.17
44.	C-25-OCT-2019-090-055	148.39	0.10
45.	C-26-OCT-2019-095-049	99.00	0.08
46.	C-26-OCT-2019-095-050	121.45	0.11
47.	C-26-OCT-2019-095-051	146.29	0.12
48.	C-26-OCT-2019-095-052	154.54	0.40
49.	C-27-OCT-2019-100-052	111.28	0.51
50.	C-29-OCT-2019-110-052	37.00	0.08
51.	D-31-OCT-2019-096-049	39.56	0.11

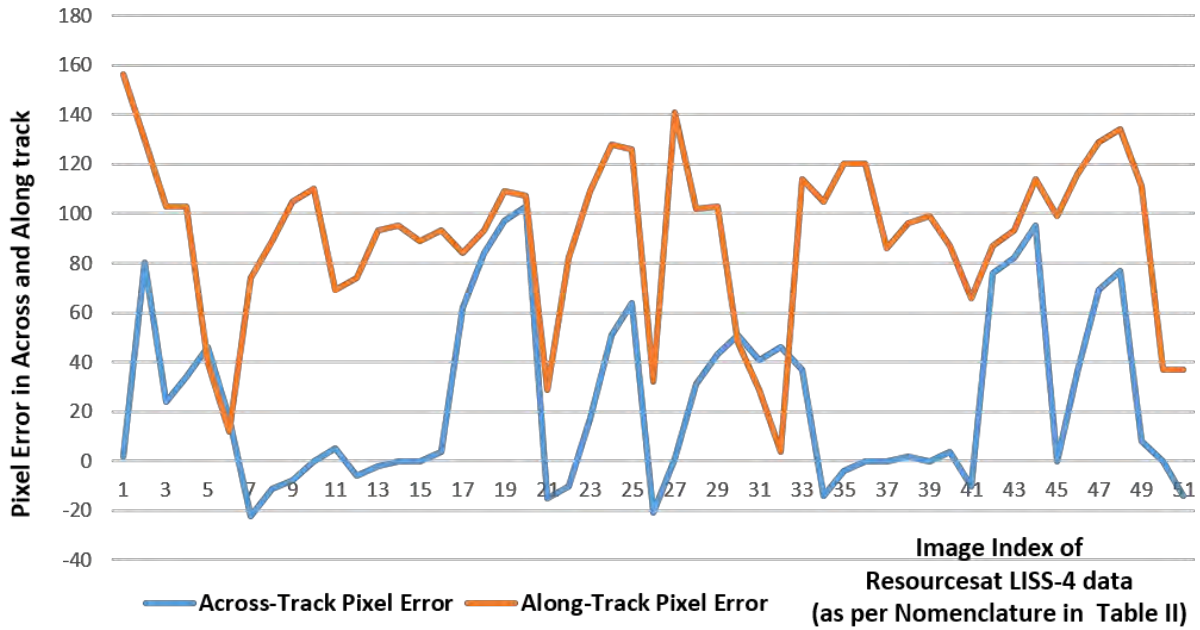


Figure-4.7. Location Discrepancy (in pixels) of Input LISS-4 Data Products

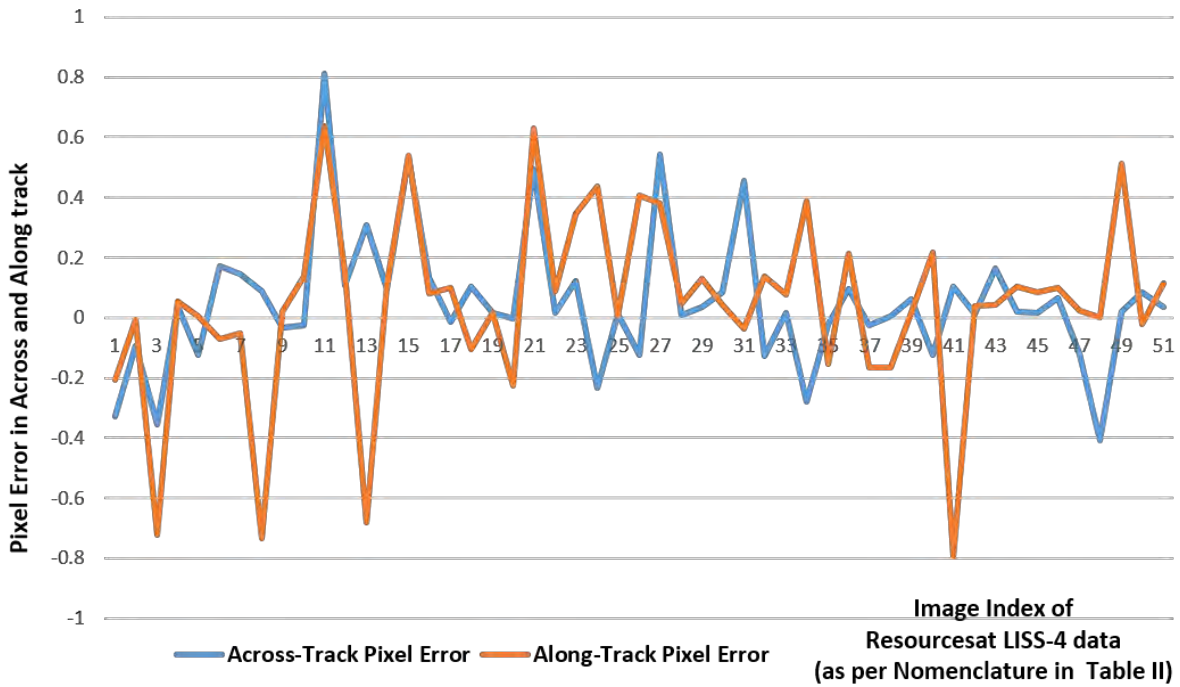


Figure-4.8. Location Discrepancy (in pixels) of Output LISS-4 Data Products



### *B. Validation using Google Earth*

To assess the accurate positional precision of the transformed LISS-4 image, we choose Google Earth as a reference data platform (Gorelick et al., 2017). In regions with low elevation, Google Earth satellite images exhibit an average horizontal positional accuracy exceeding 2 meters (Goudarzi & Landry, 2017), (Pulighe et al., 2016). Ground Control Points (GCP) are chosen in both the original LISS-4 image and the ultimately transformed LISS-4 image, with corresponding points identified in Google Earth. Figure-4.9 displays the selected Ground Control Points (GCP) in the Red Channel image of LISS-4 alongside their counterparts in Google Earth. These strategically selected Ground Control Points (GCP) allow for the evaluation of location errors across various regions in the image. Table-4.2 provides the absolute radial location discrepancy at the designated Ground Control Points (GCP) (in meters) for both the initial LISS-4 image and the transformed LISS-4 image. The outcomes indicate that the absolute location discrepancy of the ultimately co-registered image consistently remains under 5 meters.

Table-4.2. Absolute Location Accuracy in LISS-4 Image Before and After Accurate Geometric Correction

<b>GCP Points</b>	<b>Input Radial Location Error (in m)</b>	<b>Output Radial Location Error (in m)</b>
GCP-1	258.34	4.28
GCP-2	269.12	4.70
GCP-3	238.53	2.32
GCP-4	238.82	5.15
GCP-5	240.62	3.60

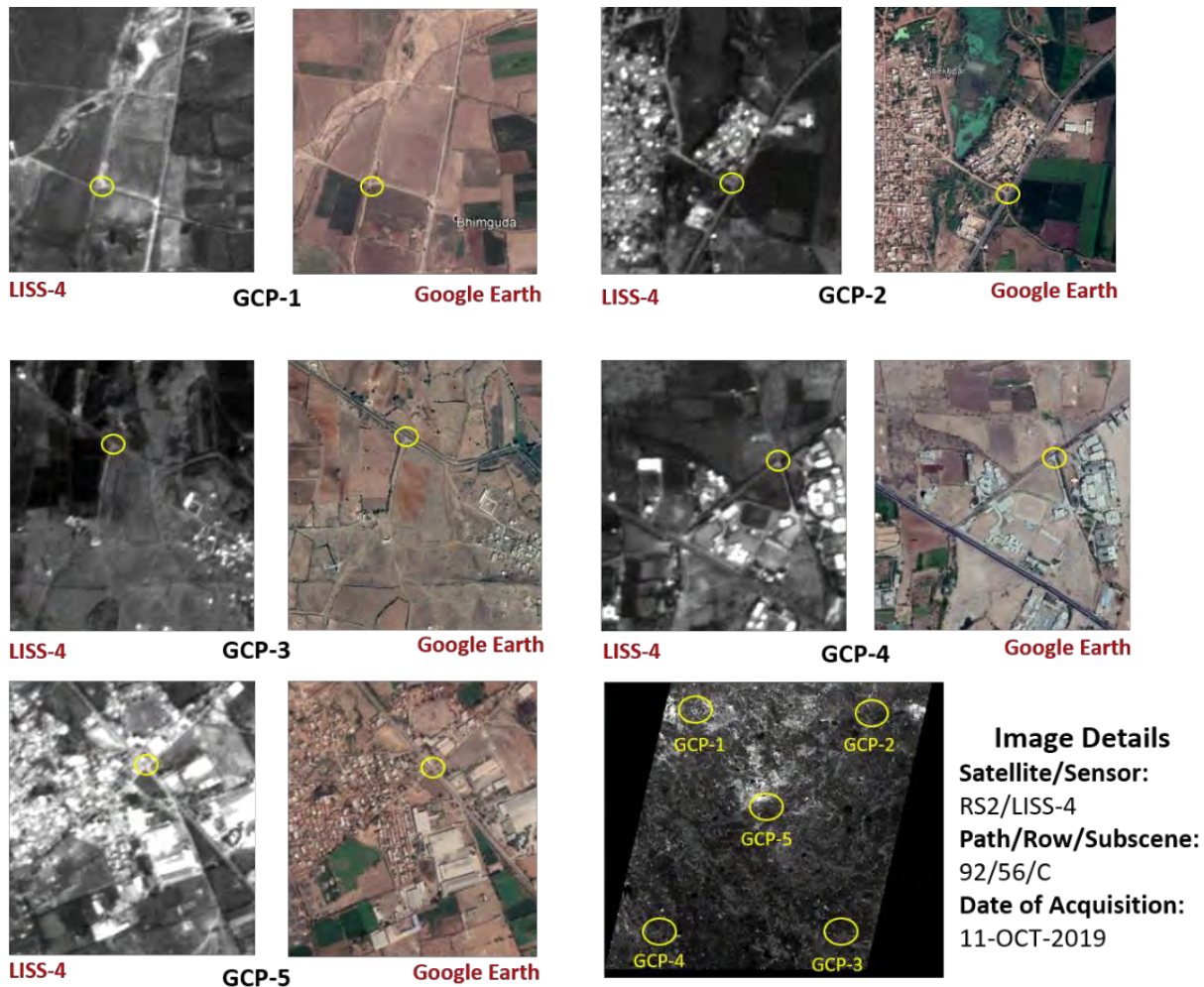


Figure-4.9. Ground Control Points Identified on the LISS-4 Image

### C. Evaluation of LISS-4 Quality: Improving Wetland Dynamics and Crop Prediction

The accurate geometric correction framework is utilized as a processing pipeline, functioning randomly across various LISS-4 images obtained on different dates. To visually assess the co-registration of the processed LISS-4 output, a checkerboard pattern image is created between the geo-registered LISS-4 image and its corresponding reference image. Different permanent features, such as roads, airstrips, and buildings in the LISS-4 images, match closely with the reference data. In Figure-4.10 (a) and (b), checkerboard pattern images are presented to qualitatively assess co-registration in various areas surrounding the scenes. The orange, green, and yellow tiles distinctly emphasize diverse regions of sub-pixel-level alignment in the output LISS-4 image with the corresponding reference layer.

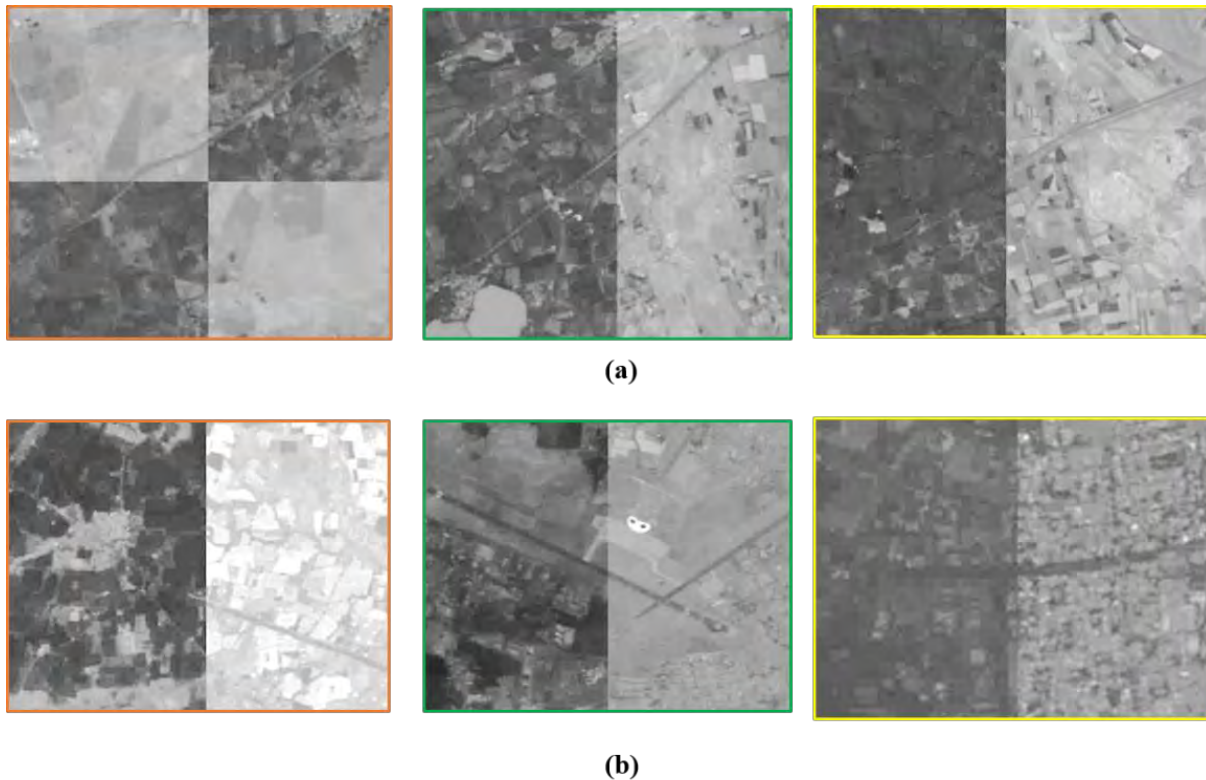


Figure-4.10. (a) Sub-Pixel Alignment between LISS-4 on April 3, 2018, and Reference

(b) Sub-Pixel Registration between LISS-4 on October 11, 2019, and Reference

Mapping and inventorying wetlands pose significant challenges (Panigrahy et al., 2012). Wetlands, crucial for maintaining Earth's ecosystem balance (Guo et al., 2017), necessitate the co-registration of multi-temporal remote sensing datasets as a prerequisite for accurate mapping. Typically, assessments of wetland dynamics involve considering pre-monsoon and post-monsoon datasets. Utilizing high-resolution LISS-4 data enhances our ability to gain insights into smaller wetlands and facilitates more effective change detection, resulting in improved map scale for change detection maps.

Figure-4.11 showcases the achieved co-registration accuracy through the accurate geometric correction of multi-temporal LISS-4 datasets using image swipe. The left side of Figure-4.11 clearly shows a misalignment in the water body and its surrounding area. Conversely, the right side of Figure-4.11 illustrates effective alignment, where the larger water body from post-monsoon LISS-4 data aligns with the smaller water body from pre-monsoon data. The FIRM design tackles geometric aspects by rectifying substantial geometric location errors in Resourcesat-2 LISS-4

datasets. Furthermore, it enhances the alignment to attain sub-pixel level co-registration accuracy with the most recent in-orbit Resourcesat-2A LISS-4 multispectral data.

The use of satellite imagery for crop evaluation has developed into a practical approach for predicting diverse crops at district, state, and national scales (Parihar & Oza, 2006; Oza et al., 2008; Moorthi et al., 2014). The LISS-4 instrument on Resourcesat-2 and 2A furnishes multi-temporal images with improved temporal frequency. Its elevated spatial resolution provides more intricate information about crop growth, facilitating enhanced segmentation of various agricultural lands at the town level. The registration of images across multiple dates is a crucial stage in predicting crops through remote sensing images. FIRM has been evaluated using multi-date images from Resourcesat-2/2A LISS-4. In Figure-4.12, the results before and after correction are presented over agricultural land, visually demonstrating the successful accomplishment of multi-temporal sub-pixel registration across the timeline.

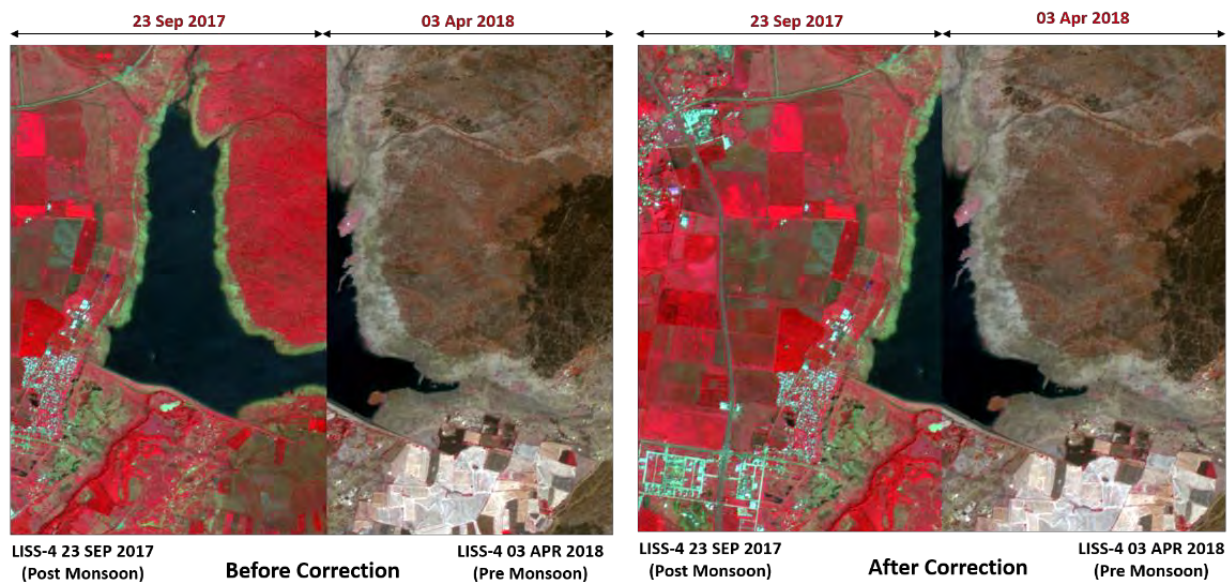


Figure-4.11. Co-Registration Performance over Wetland



Figure-4.12. Co-Registration Performance over Agricultural Land

#### D. Spatial Alignment of Resourcesat-2 with LANDSAT-8

To evaluate the efficiency of the FIRM approach, we employed Resourcesat LISS-4 data and Landsat-8 Operational Land Imager (OLI) data obtained on identical dates. Landsat-8 is a well-known Earth observation satellite recognized for monitoring land surfaces (Storey et al., 2014). Both satellites follow a systematic referencing scheme, aiding users in identifying the shared geographic area covered. The LISS-4 wavelength channels have a spatial resolution of 5.8 meters, whereas OLI has its highest spatial resolution channel at 15 meters, encompassing the panchromatic wavelength spectrum. Considering that OLI exhibits a radial RMSE superior to 12 meters for Level-1 Terrain Corrected data products (Claverie et al., 2018), we opted for OLI PAN as the reference and LISS-4 as the input image. Table-4.3 furnishes the metadata particulars for Resourcesat-2 and Landsat-8 data products. In Figure-4.13 (a), a composite of LISS-4 and OLI data encompassing a region in the Gujarat state, India, is showcased. Figure-4.13 (b) exhibits segments encompassing various overlapping regions of the image, employed for quantitative evaluation.

The co-registered LISS-4 data product is evaluated at distinct features across the image in contrast to OLI. The proposed FIRM approach precisely calculates the mis-registrations between the images. Figure-4.14 demonstrates the co-registration performance of LISS-4 and OLI both prior to and following registration. For processing, both LISS-4 and OLI images are interpolated to a spatial resolution of 5.0 meters. The analysis of the dataset pairs indicates that roads, bridges, vegetation boundaries, and building structures existing in LISS-4 align at a sub-pixel level relative to the OLI reference.

Table-4.3. Details of Multi-Satellite Images for Evaluation

S.No.	Satellite /Sensor	Path	Row	Date & Image Time
1.	Resourcesat-2 /LISS-4	93	56	01-JAN-2018; 05:50:54 UTC
2.	Landsat-8 /OLI PAN	148	44	01-JAN-2018; 05:32:57 UTC

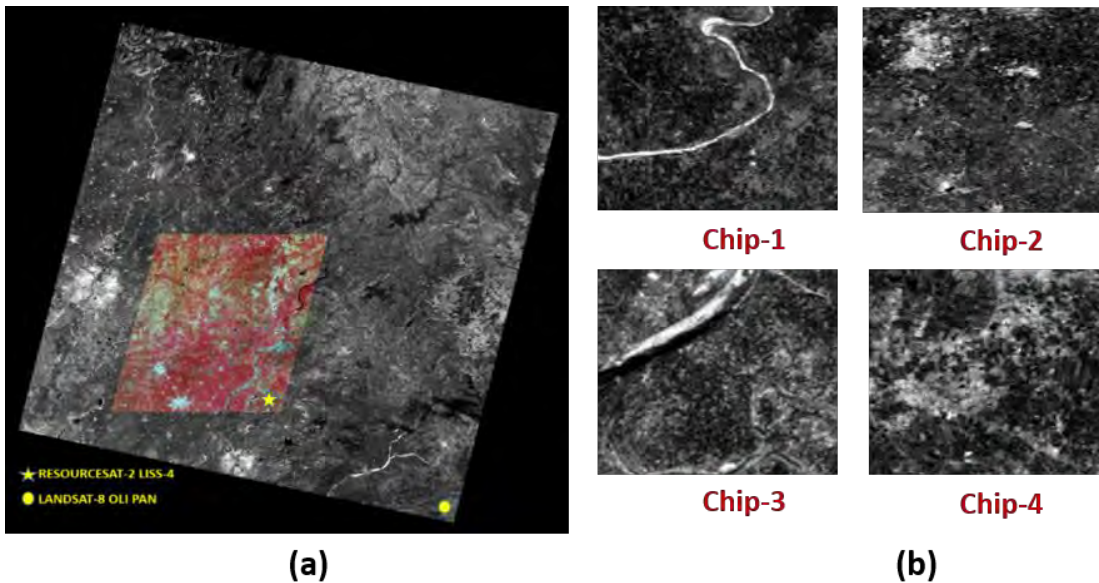


Figure-4.13. (a) Coverage of Resourcesat-2 on the Same Day over Landsat-8  
(b) Overlapping Image Chips for Evaluation

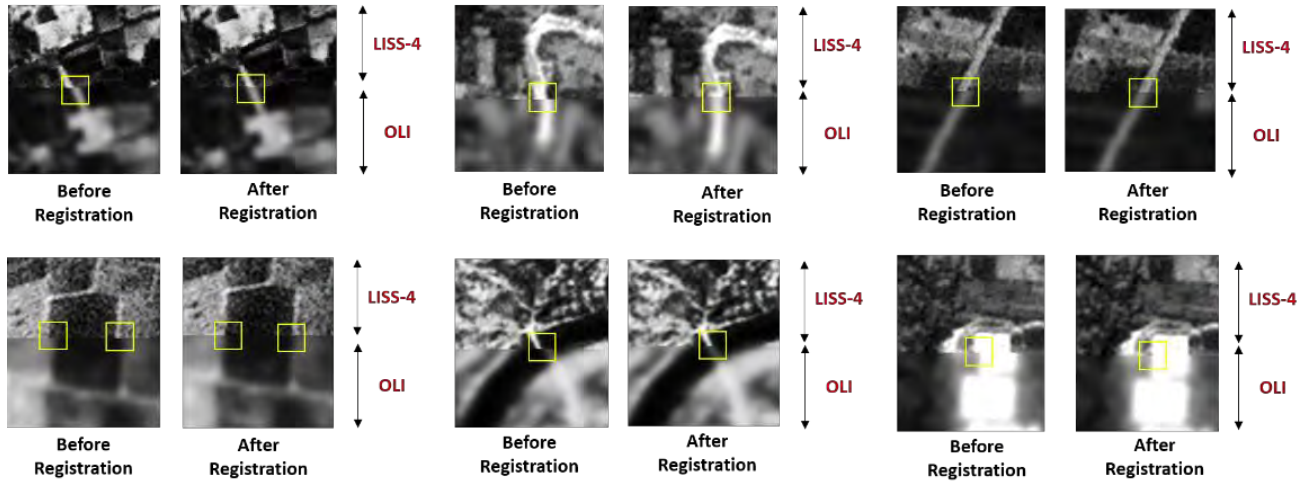


Figure-4.14. Co-Registration Performance of Resourcesat-2 and Landsat-8

Table-4.4. Sub-Pixel Displacements between Registered Resourcesat-2 LISS-4 and Landsat-8

OLI

S.No.	Chip Number	Shift in Width (in pixel)	Shift in Height (in pixel)
1.	Chip-1	0.33	0.01
2.	Chip-2	0.01	0.02
3.	Chip-3	-0.01	-0.11
4.	Chip-4	0.08	0.17

To quantitatively evaluate the co-registration performance, segments are extracted from registered LISS-4 images and their corresponding OLI references at different geographic locations, as shown in Figure-4.14. Afterward, pairs of image segments are matched using mode-guided SIFT to identify sub-pixel shifts in both the width and height directions. The computed sub-pixel shifts between multi-satellite images in both directions are presented in Table-4.4. It is noted that the Root Mean Square Error (RMSE) of LISS-4 compared to OLI is less than 0.12 pixels at a spatial resolution of 5 meters. The sub-pixel geographical alignment attained through the developed model signifies the initial phase in the harmonization of data between Resourcesat and Landsat (Claverie et al., 2018).

### *E. Alignment Performance and Comparison of LISS-3 Acquisitions a Decade Apart*

The framework for image co-registration undergoes testing using images captured by LISS-3 with a time difference exceeding 10 years. The co-registration performance is assessed by comparing Resourcesat-1 (Harinath et al., 2008) LISS-3 images from the years 2005/2006 using the most recent Resourcesat-2A LISS-3 images obtained in 2017. Over the course of a decade, numerous changes occur in features, but permanent features persist and remain distinct, necessitating efficient detection for accurate transformation estimation. Scene details for pairs of LISS-3 images employed in testing the multistage workflow are provided in Table-4.5. FIRM is applied to georeferenced LISS-3 data sampled with a spatial resolution of 24.0 meters, with the reference generated from a stack with corresponding resolution. The observed co-registration accuracy is at a sub-pixel level, as visually demonstrated in Figure-4.15. The road network and manmade structures align at a sub-pixel level, making the decadal co-registered LISS-3 product is an essential input for studies involving change detection using remote sensing data.

Table-4.5. Details of LISS-3 Images for Co-Registration

<b>S.No.</b>	<b>LISS-3 Image Pairs</b>	<b>Path/ Row</b>	<b>LISS-3 Pair Acquisition Dates</b>
1.	Image Pair-1	96/56	27-NOV-2006, 24-OCT-2017
2.	Image Pair-2	98/56	01-OCT-2005, 03-NOV-2017
3.	Image Pair-3	100/56	06-OCT-2006, 13-NOV-2017



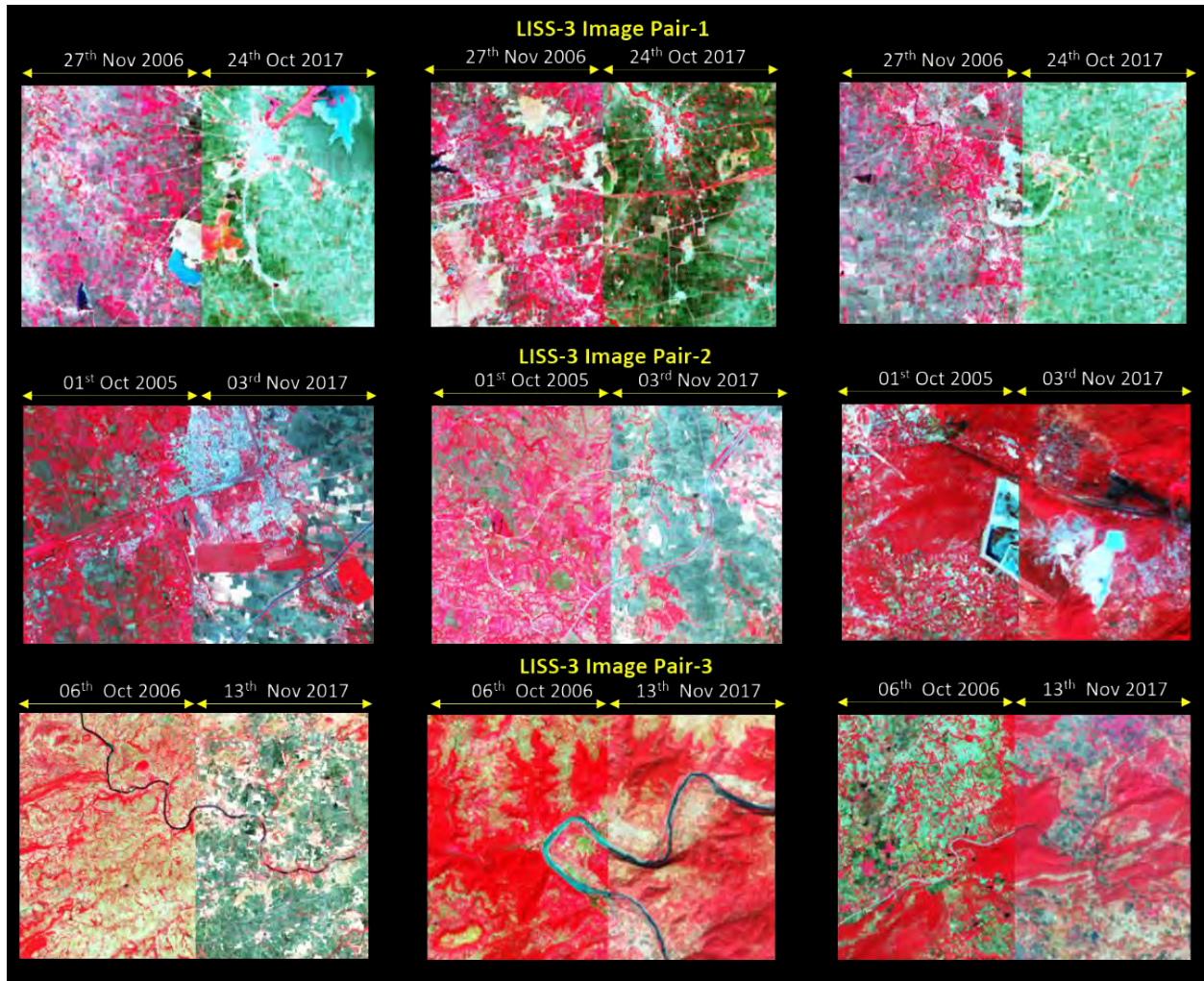


Figure-4.15. Co-Registration Performance of LISS-3 in Various Image Regions

The co-registration performance of the rectified LISS-3 output products is evaluated and compared with advanced feature-based image registration techniques. Harris Feature Detection, utilizing the RANSAC estimation model, emerges as a robust automated method for registering satellite images (Misra et al., 2012a). Another effective method for registering multi-temporal images is SIFT, which employs mode-seeking outlier rejection (Kupfer et al., 2014). Modified SIFT, known as PSO-SIFT, enhances feature matching and improves performance for multi-sensor images (Ma et al., 2016). Phase Congruency with spatial constraint (PC-SC) is another efficient feature-based method for aligning pairs of images at a sub-pixel level (Ma et al., 2018).

Table-4.6 displays the Root Mean Square Error (RMSE) in pixels computed for co-registered LISS-3 image pairs using different methods, including the proposed FIRM approach discussed in this paper. Mode Seeking SIFT outperforms Harris utilizing RANSAC optimization. The FIRM

approach demonstrates superiority over other methodologies in RMSE, achieving the closest sub-pixel image co-alignment among the LISS-3 images taken a decade apart.

Table-4.6. Comparative Analysis with Feature-Based Methods

<b>Feature Based Methods for Image Registration</b>	<b>RMSE (in pixel) Image Pair-1</b>	<b>RMSE (in pixel) Image Pair-2</b>	<b>RMSE (in pixel) Image Pair-3</b>
Harris + RANSAC	0.75	0.90	0.82
SIFT + Mode Seek	0.68	0.86	0.80
PSO-SIFT	0.62	0.75	0.71
PC-SC	0.59	0.63	0.66
FIRM Approach	0.42	0.51	0.61

#### *F. AWiFS Multi-Date Across-Track Co-Registration*

The AWiFS camera mounted on the Resourcesat spacecraft captures a broad region to generate a large area mosaic. Co-registration of multi-temporal AWiFS scenes is essential to align them to a common geographic coordinate system, ensuring the creation of a seamless mosaic and allowing for the assessment of geometric fidelity in overlap segments across multiple acquisition dates. Table-4.7 provides details of AWiFS image scenes, primarily covering Bangladesh. The AWiFS sensor provides data with a spatial resolution of 56 meters, and the corresponding reference is sampled at this particular resolution for registration. The main difficulty in co-registering AWiFS data arises from its broader swath, resulting in localized internal distortion. FIRM is utilized to align AWiFS images, modeling spatial alignment with the reference stack to produce a sub-pixel co-registered AWiFS data product.

Figure-4.16 presents the AWiFS False Color Composite (FCC) mosaic covering Bangladesh and the surrounding region. The enlarged tiles in Figure-4.16 highlight the geometric consistency in overlapping regions (illustrated as oval shapes) between adjacent AWiFS images. The consistency

of features along the edges indicates that multi-temporal images are accurately aligned across the swath, leading to the creation of mosaic data products at the country level.

Table-4.7. Details of AWiFS Image Scenes

S.No.	Path	Row	Scene/Swath (in km)	Date of Acquisition
1.	108	56	Full Scene/ 740 km	30-DEC-2018
2.	109	55		24-OCT-2018
3.	110	56		22-NOV- 2018

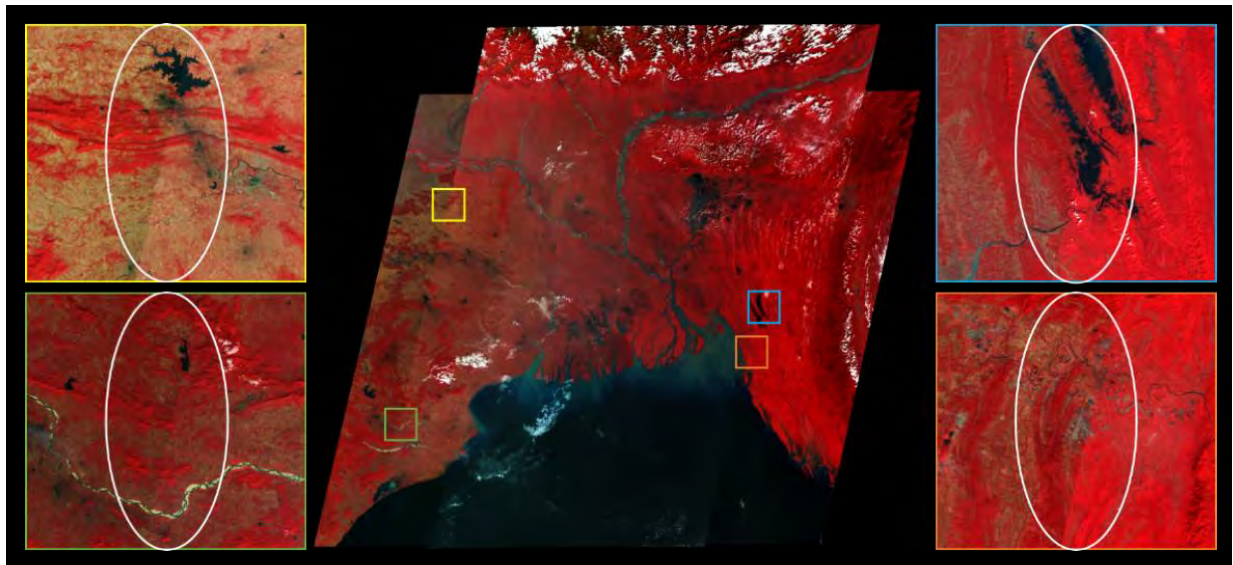


Figure-4.16. Geometric Consistency in Overlap Areas in the AWiFS Image Mosaic Covering Bangladesh and its Surroundings

#### 4.1.4. Summary

Every Resourcesat data piece possesses unique characteristics, with dynamic geometric errors attributed to steering, a broader swath, and spacecraft errors not systematically modeled. To address this, feature-based image registration is introduced, utilizing a reference layer to enhance geometric location accuracy and achieve sub-pixel co-registration accuracy in multi-temporal

Resourcesat scenes. For feature-based remote sensing image registration, it is essential to have a clear understanding of which feature points to detect and how to accurately describe them to establish correspondence. Minimizing false correspondences or outliers in the match set is crucial for determining effective transformation parameters. The paper provides a detailed description of the developed methodology, highlighting co-registration performance in multi-satellite images for harmonization and the quantitative evaluation of co-registered data products over various landscapes.

FIRM has demonstrated promising outcomes for enhancing the sub-pixel level accuracy of geometric location in Resourcesat multispectral remote sensing data. Simply relying on a technique for detecting features with an optimization algorithm is insufficient to address the registration challenges in Resourcesat multisensor data. The solution lies in integrating robust feature detection techniques strategically within the framework to attain the intended result. Employing Patch Affine ORB for coarse estimation and mode-guided SIFT for finer estimation, combined with the additional capability of motion smoothness constraint, enables the generation of precision co-registered data. FIRM has successfully automated sub-pixel co-registration in multi-temporal optical images with a decade-long difference in acquisition, covering diverse land features and terrains. FIRM acts as a versatile framework that can be customized for aligning medium-resolution remote sensing images through parameter adjustments. However, it's essential to assess FIRM's effectiveness on images with extremely high spatial resolution. On a broader scale, the georeferenced data products streamline the efforts of the remote sensing scientific community, allowing them to smoothly advance to the next stage of image analysis.

## 4.2. Multi-Satellite Scanty Feature Island Image Registration Methodology Development

### 4.2.1. Introduction

The section introduces a procedure for registering multi-satellite images in remote sensing scenes featuring sparse islands, referred to as MIRACLE. Figure-4.17 displays Resourcesat LISS-4 multi-spectral images, demonstrating their utilization in the automatic image registration task. The majority of these scenes comprise extensive oceanic expanses, featuring a scattering of small islands situated in Lakshadweep, India. As far as we are aware, this study represents the first attempt to develop a specialized multi-satellite image registration technique designed for scenes characterized by scattered islands. The inclusion of a data pre-processing stage aims to improve the image by utilizing anisotropic coherence localization for enhanced feature interpretation. Additionally, the multi-spectral image undergoes transformation through Principal Component Analysis (PCA). A technique for detecting features is created using nonlinear diffusion filtering, coupled with marginalization optimization for outlier removal guided by Mahalanobis distance. This process is designed to generate keypoint correspondences without errors. MIRACLE is subjected to quantitative assessment across diverse island scenes and is compared with state-of-the-art techniques. The evaluation of visual quality confirms that the proposed method, MIRACLE, achieves sub-pixel level spatial alignment accuracy in different island regions during the process of image co-registration.



Figure-4.17. Lakshadweep Islands, primarily located in the vast expanse of the deep ocean, as captured by the Indian Resourcesat LISS-4, along with the geographical coordinates (latitude and longitude) of the central point of the scene.

## 4.2.2. Methodology Developed

The initiation of the proposed image registration method, MIRACLE, involves a data pre-processing stage aimed at improving the input and reference images of scanty feature islands through anisotropic coherence locality enhancement. In addition to enhancement, Principal Component Analysis (PCA) is applied to the input multi-spectral image to maximize variance information, facilitating the detection of unique and consistent key points. The matching of the multi-satellite remote sensing image incorporates a non-linear diffusion filtering method, expedited by the Fast Explicit Diffusion (FED) approach. Despite this, the aligned control points are prone to outliers, requiring their refinement through the Mahalanobis Distance Guided Marginalization (MDGM) optimization process. The estimation of transformation parameters employs an affine transformation model, resulting in the generation of a data product registered across multiple satellites. The subsequent sections provide a thorough explanation of the algorithmic steps and processing workflow employed in MIRACLE.

### *A. Enhancement and Transformation through Anisotropic Coherence Locality*

Enhancing remote sensing images from multiple satellites, which primarily feature small, sparse islands, is essential to improve the demarcation of distinctive features. This enhancement of the image involves a critical role in extracting features accurately for the co-registration of island images. In this particular scenario, it is observed that there are flow-like formations around the coastlines in proximity to the islands, attributed to coral reefs. These structures require emphasis in the image to enhance feature matching effectively. Achieving this goal involves the use of a more advanced structure descriptor. Anisotropic coherence locality-enhanced diffusion filtering is utilized to accentuate coherent flow patterns in the vicinity of the coastal islands. Coherence enhancement can be accomplished using the structure tensor, which can be defined as follows.

$$J_p(\nabla\mu_\sigma) := K_p * (\nabla\mu_\sigma\nabla\mu_\sigma^T) \quad (4.11)$$

Here,  $K_p$  represents the Gaussian operator responsible for averaging orientation information across an integration scale denoted by  $p$ ,  $J_p$  is a semidefinite matrix,  $\sigma$  stands for the regularization parameter, and  $\mu$  signifies the eigenvalues.

In this context, the eigenvalues act as a measure for the average contrast along the eigen direction at a given scale. The difference between these eigenvalues measures local coherence, assisting in

the recognition of structures and areas that necessitate enhancement, as well as areas requires smoothing in the image (Weickert 1999).

In our case, where rectifying multi-spectral images in relation to a reference image is essential, we employ the PCA dimension reduction technique (Misra et al., 2022b). PCA calculates the primary component containing the highest variance information, resulting in a more perceptible and strengthened representation of feature targets around coastal islands compared to specific wavelength channels within the multi-spectral image dataset. Utilizing PCA for increased variability in features improves the matching performance in the multi-satellite image registration process between the input and reference images.

To simplify the image registration process, remote sensing images from multiple satellites are converted into a shared coordinate system. The geometric bounding corners of the input datasets are acquired, and the region of overlap is determined by performing a polygon intersection of the geo-bounding rectangles. As the input geo-referenced images exhibit diverse spatial resolutions, ranging from 5.0 meters to 15.0 meters, the extracted images undergo geometric transformation to a uniform projection system and are resampled to attain a consistent pixel resolution.

### *B. Employing Nonlinear Diffusion Filtering for Matching Features*

Methods for detecting features across multiple scales, like SIFT (Lowe, 1999), identify features by creating a Gaussian Kernel within a pyramid structure. Nevertheless, it has been noticed that the use of the Gaussian Filter can result in blurring, compromising the boundaries of objects and negatively impacting the accuracy of localization and distinctiveness in remote sensing images that contain sparse features.

To tackle this concern, the suggested algorithm proposes identifying and describing features within a non-linear space through the application of non-linear diffusion filtering (Alcantarilla & Solutions, 2011). To address this issue, the proposed algorithm recommends identifying and describing features within a non-linear space by employing non-linear diffusion filtering (Alcantarilla, Bartoli, and Davison, 2012). The traditional non-linear diffusion equation is represented at the pixel location  $(x, y)$  as follows:

$$\frac{dL}{dt} = \text{div}(c(x, y, t) \cdot \nabla L) \quad (4.12)$$

In this context,  $\text{div}$  and  $\nabla$  denote the divergence and gradient operators, respectively, with  $L$  representing the luminance of the image. The scale parameter,  $t$ , influences the complexity of the

image representations, with larger values yielding simpler representations. The conductivity function  $c$  is defined as:

$$c(x, y, t) = g(|\nabla L_\sigma(x, y, t)|) \quad (4.13)$$

In this context, the function  $\nabla L_\sigma$  corresponds to the gradient of an image  $L$ , which has undergone smoothing through a Gaussian filter.

Acceleration in nonlinear diffusion filtering becomes feasible through the use of Fast Explicit Diffusion (FED) (Alcantarilla et al., 2012). FED amalgamates the benefits of both explicit and semi-implicit schemes, thereby improving the efficiency of the computational process. Stability concerns linked to computationally demanding operations in the explicit scheme are resolved by performing multiple iterations to attain the desired scale level. The FED schemes draw inspiration from breaking down box filters into explicit schemes and are incorporated into a pyramidal framework, greatly accelerating the detection of features in non-linear spaces. Iterative box filters offer a dependable approximation of Gaussian kernels and are easy to implement. FED is configured to perform  $m$  cycles of  $n$  explicit diffusion steps, each with different step sizes  $t_j$ , derived from the factorization of the box filter.

$$t_j = \frac{t_{max}}{2\cos^2(\pi\frac{2j+1}{4n+2})} \quad (4.14)$$

Here,  $t_{max}$  represents the maximum step size that adheres to the stability requirement of the explicit scheme (Alcantarilla and Solutions, 2011).

### C. Guided by Mahalanobis Distance

In our method, Mahalanobis distance is employed to measure the distance between each element in the set of differences of matched points and the established distribution (Mahalanobis, 1936). This distribution is computed by finding the mean of the set of differences of matched points. The standard deviation is subsequently calculated to indicate the degree of deviation of each element in the set of differences of matched points from the mean distribution. A threshold, determined from the image, is used to direct the elimination of the initial set of outlier matched points, leading to the establishment of potential keypoint correspondences. The Mahalanobis distance in the horizontal direction ( $D_{mx}$ ) for a sample point  $P$  is calculated in the following manner.



$$D_{mx} = \frac{PD_x - M_x}{SD_x} \quad (4.15)$$

where,  $PD_x$  represents the difference in sample matched points for point  $P$  along the horizontal direction,

$M_x$  is the average of the differences in the set of matched points along the horizontal direction,  $SD_x$  denotes the deviation from the mean in the horizontal direction.

The Mahalanobis distance in the vertical direction  $D_{my}$  for a sample point  $P$  is calculated as follows.

$$D_{my} = \frac{PD_y - M_y}{SD_y} \quad (4.16)$$

where,  $PD_y$  denotes the sample matched point difference for point  $P$  in the vertical direction,

$M_y$  represents the mean of the matched point set differences in the vertical direction,

$SD_y$  signifies the standard deviation in the vertical direction.

The Mahalanobis distance estimate  $D_m$  for a sample point  $P$  is expressed as follows.

$$D_m = \sqrt{(D_{mx})^2 + (D_{my})^2} \quad (4.17)$$

The threshold distance limit, denoted as  $t$ , is determined empirically based on the set of differences in the matched points. If the computed Mahalanobis distance falls below the threshold distance limit  $t$ , the matched point is classified as an inlier and retained as a reliable, accurately matched point.

#### *D. Marginalization Optimization*

The Mahalanobis distance metric yields a smaller set of matched points, which serves as the input for the Marginalization procedure. This optimization method, devoid of a user-defined threshold, is introduced to enhance the accuracy of robust estimation significantly (Barath et al., 2021). In our methodology, we utilize the Marginalization Optimization Outlier Removal Technique to determine the final matched correspondences between the input multi-spectral remote sensing image and its associated reference image. The core steps of the Marginalization Optimization Algorithm are outlined as follows.

##### *i. Marginalization over $\sigma$*

Suppose the noise  $\sigma$  is considered a stochastic variable with a probability density function  $f(\sigma)$ . Let's establish a fresh quality metric for the model  $\theta$  marginalized over  $\sigma$  as stated:

$$Q_{MAGSAC}^*(\theta, P) = \frac{1}{\sigma_{max}} \int_0^{\sigma_{max}} \ln L(\theta, \frac{P}{\sigma}) d\sigma \quad (4.18)$$

In this expression, where  $Q$  represents the Quality Function for the Marginalization Optimization Model,  $P$  is the collection of data points,  $\theta$  denotes the parameters of the model, and  $\sigma$  represents the standard deviation of the noise.

### *ii. Model Fitting through $\sigma$ -Consensus*

As Marginalization Optimization does not have a predefined set of inliers that could be used to estimate the model derived from a minimal sample, it is recommended to utilize weighted least-square fitting. In this method, the weights are calculated based on the probabilities of points being inliers, as proposed by Barath et al. (2019). The probability of a point  $p \in P$  being an inlier, given the model  $\theta_\sigma$  with a residual function  $D$  and noise standard deviation  $\sigma$ , is expressed as:

$$L\left(\frac{p}{\theta_\sigma}\right) = 2C(P)\sigma^{-p}D^{p-1}(\theta_\sigma, p)\exp\left(\frac{-D^2(\theta_\sigma, p)}{2\sigma^2}\right) \quad (4.19)$$

### *iii. Termination Criteria*

The lack of an inlier set and, consequently, an initial estimate of the inlier ratio makes the standard termination criteria of RANSAC impractical. To determine  $k$  without depending on a particular noise standard deviation, a straightforward approach is to conduct marginalization similar to the model quality. The process is described as follows:

$$k^*(P, \theta) = \frac{1}{\sigma_{max}} \int_0^{\sigma_{max}} k(\theta, \sigma, P) d\sigma \quad (4.20)$$

Therefore, determining the necessary number of iterations for Marginalization Optimization is an iterative process that is continuously updated whenever a new optimal solution is identified, similar to the approach used in RANSAC. The effectiveness of Marginalization Optimization is evaluated in the context of matched feature points, even when the data is corrupted with numerous erroneous points. As shown in Figure-4.18, the Mahalanobis Distance Guided Marginalization Optimization trims the matched keypoints, generating precise keypoint correspondences that are essential for estimating transformation parameters.

Before the removal of outliers, the initial count of identified corresponding points among images captured by multiple satellites stands at 419. However, the keypoint set includes numerous spurious correspondences, as indicated by the red circles in Figure-4.18. Following the application of the proposed technique, the number correctly matched control points stands to 177. It is

noteworthy that approximately 42% of these are deemed reliable correspondences, represented by the orange-colored circles in Figure-4.18.

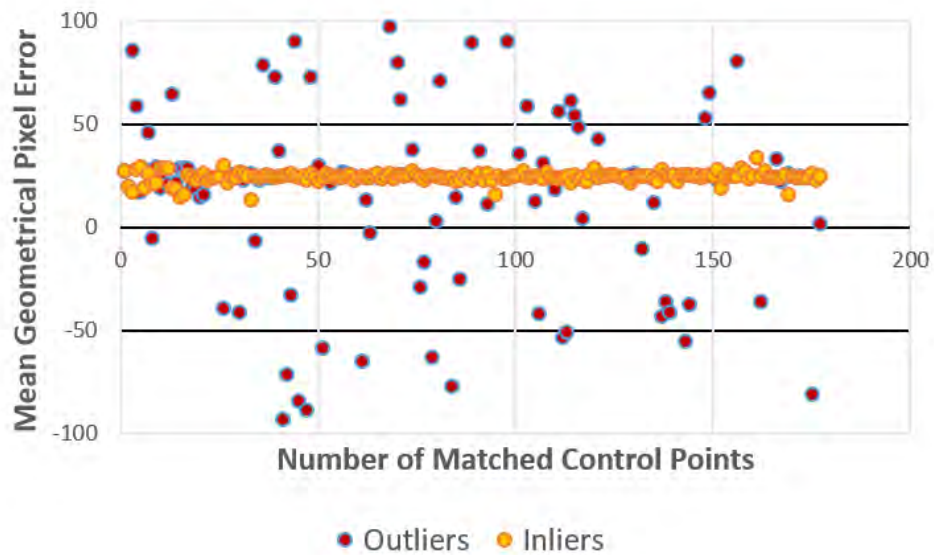


Figure-4.18. Outlier Removal process using Mahalanobis Distance Guided Marginalization Optimization

#### E. MIRACLE Processing Workflow

The process begins by applying anisotropic coherence-based locality enhancement and performing Principal Component Analysis (PCA) transformation on the input multi-spectral remote sensing image. The initial PCA component is created to improve the delineation of features in different spectral classes on the island, facilitating the identification of distinctive and stable characteristics. MIRACLE showcases resilience in addressing noise and occlusion found in the satellite image input by maximizing variance information derived from multi-spectral data for island image registration.

The Panchromatic (PAN) image, selected as the reference, offers superior geometric location accuracy and originates from a distinct satellite. The reference image is also improved through anisotropic coherence-based locality enhancement. The overlapping area between the multi-spectral and PAN images is extracted using the geographic information linked to each pixel. The reference image is automatically chosen from the directory of the satellite image database, guided by the geographic corner coordinates of the multi-spectral image.

The next step includes geometrically transforming the unaligned images from multiple satellites to a consistent map projection and pixel size, ensuring image registration. The MIRACLE

methodology employs nonlinear diffusion filtering as a technique for feature detection and description in multi-satellite images. The initial extraction of matched control points involves the application of the brute force matching descriptor with a K-Nearest Neighbor match measure.

Outlier elimination is achieved through the Mahalanobis Distance Guided Marginalization (MDGM) Optimization Parameter Estimation Method, leading to accurately identified keypoint correspondences. During the processing steps, adjustable parameters are chosen to guarantee the detection of an adequate number of feature points in scenes with limited features on the island. This helps establish putative keypoint correspondences, ultimately achieving sub-pixel image registration accuracy.

The affine transformation parameters are subsequently calculated using the pruned matched points, identifying relative variable image shifts. Following this, the original input multi-spectral remote sensing image is resampled to produce a co-registered data product from multiple satellites. Figure-4.19 visually illustrates the comprehensive workflow for multi-satellite image registration in scenes with limited features on the island, referred to as MIRACLE.

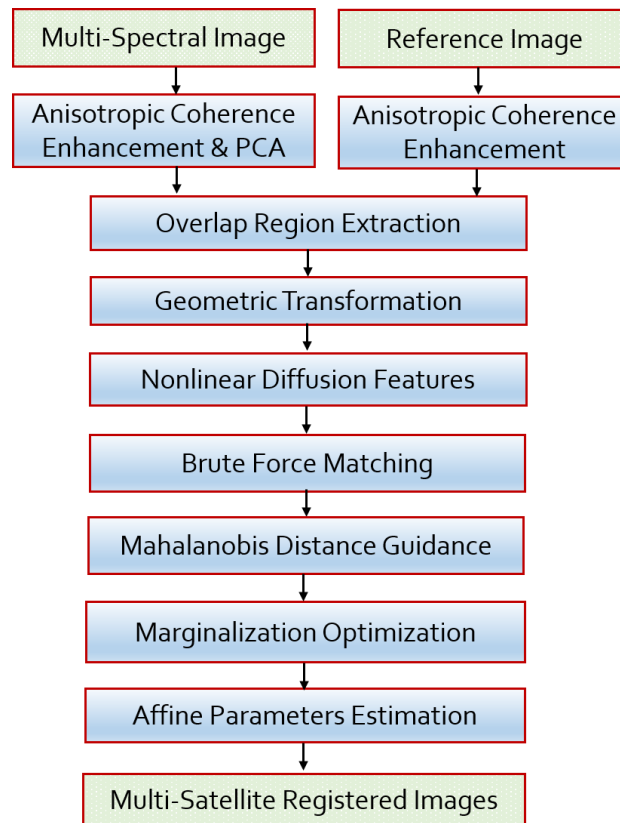


Figure-4.19. Workflow of the MIRACLE Island Image Registration process.

### 4.2.3. Experimental Results

The effectiveness of the MIRACLE image registration method is evaluated across various islands located in Lakshadweep, along the Tamil Nadu (TN) Coast, and the Gujarat Coast. The islands with sparse features are captured in multi-temporal imaging acquisitions using Resourcesat-2 LISS-4 multi-spectral images, which have a spatial resolution of 5.0 meters. Significantly, Resourcesat-2 demonstrates an increased magnitude of geo-location error attributed to uncertainties in spacecraft perturbation (Misra et al., 2021a). Even after reaching the end of its mission life, Resourcesat-2 consistently delivers radiometrically stable images suitable for a range of Earth observation applications.

The suggested approach employs Resourcesat-2 LISS-4 multi-spectral images as input, improving the accuracy of location data products using the MIRACLE procedure with respect to a standard reference satellite image source. Landsat-8 OLI PAN images, characterized by a spatial resolution channel of 15.0 meters and encompassing the wavelength spectrum of (0.5-0.68)  $\mu\text{m}$ , are selected as the reference. The radial RMSE geo-location error for Landsat-8 OLI is less than 12 meters for Level-1 terrain-corrected data products (Storey et al., 2014). Table-4.8 provides details of the satellites and sensors used in the image registration exercise.

Table-4.9 and Figure-4.20 provide information on the multi-satellite images and the recognized islands, along with the geographic coordinates of their corresponding center points. Notable islands such as Amini, Bitra, and Kiltan in the Lakshadweep region are considered, each with unique characteristics and cultural significance. The proposed approach, MIRACLE, is also applied to other prominent islands in the Lakshadweep region. Moreover, regions along the coast of Tamil Nadu and Gujarat, which include small islands characterized by limited land areas, are chosen to verify the applicability and effectiveness of the MIRACLE approach across diverse terrains. This validation aims to demonstrate the versatility of MIRACLE in addressing demanding tasks like mapping coral reefs and identifying shorelines through the production of data products registered at the sub-pixel level. Figure-4.20 visually represents images of the islands in Lakshadweep, Tamil Nadu (TN), and Gujarat.

Table-4.8. Details of Satellites and Sensors Utilized for the Evaluation of Island Image Registration

Satellite	Sensor	Swath (in km)	Spatial Resolution (in m)	Wavelength (in $\mu\text{m}$ )
Resourcesat-2	LISS-4	70	5.8	0.52-0.59
				0.62-0.68
				0.77-0.86
Landsat-8	OLI PAN	185	15.0	0.5-0.68

Table-4.9. Details of Multi-Satellite Images encompassing Lakshadweep Islands, Tamil Nadu Coast, and Gujarat Coast.

Islands (Lat., Lon.)	Reference Satellite/Sensor/ Path/Row/Image Date	Input Satellite/Sensor/ Path/Row/Image Date
Amini (11.1243,72.7239)	Landsat-8/OLI/ 147/52/12-Mar-2017	Resourcesat-2/LISS-4/ 096/066/21-Jan-2017
Bitra (11.5532,72.1611)	Landsat-8/OLI/ 147/52/28-OCT-2019	Resourcesat-2/LISS-4/ 095/065/19-Nov-2019
Kiltan (11.4826,73.0057)	Landsat-8/OLI/ 147/52/23-Jan-2017	Resourcesat-2/LISS-4/ 096/065/21-Jan-2017
Minicoy (8.2867, 73.0596)	Landsat-8/OLI/ 146/54/31-Dec-2016	Resourcesat-2/LISS-4/ 097/068/11-Aug-2016
Suhelipar (10.0808, 72.2805)	Landsat-8/OLI/ 147/53/25-12-2017	Resourcesat-2/LISS-4/ 095/066/09-Feb-2017
Chetlat (11.6929, 72.7124)	Landsat-8/OLI/ 147/52/13-Jan-2019	Resourcesat-2/LISS-4/ 096/065/11-Jan-2019
Baliyapaniyam (12.3625, 71.9027)	Landsat-8/OLI/ 147/52/23-Jan-2017	Resourcesat-2/LISS-4/ 095/065/16-Jan-2017
Kavaratti	Landsat-8/OLI/	Resourcesat-2/LISS-4/

<b>(10.5653, 72.6392)</b>	147/53/29-Apr-2017	096/066/21-Jan-2017
Tamil Nadu (TN) Coast-1 <b>(9.1849, 79.3213)</b>	Landsat-8/OLI 142/54/02-Apr-2018	Resourcesat-2/LISS-4/ 102/67/22-May-2018
Tamil Nadu (TN) Coast-2 <b>(8.5885, 78.1177)</b>	Landsat-8/OLI 143/54/09-Jan-2016	Resourcesat-2/LISS-4/ 101/68/28-Jan-2016
Gujarat Coast-1 <b>(22.2940, 69.6750)</b>	Landsat-8/OLI 150/44/20-Mar-2018	Resourcesat-2/LISS-4/ 090/056/23-Mar-2018
Gujarat Coast-2 <b>(22.4260, 70.1601)</b>	Landsat-8/OLI 150/44/04-Mar-2018	Resourcesat-2/LISS-4/ 091/056/04-Mar-2018

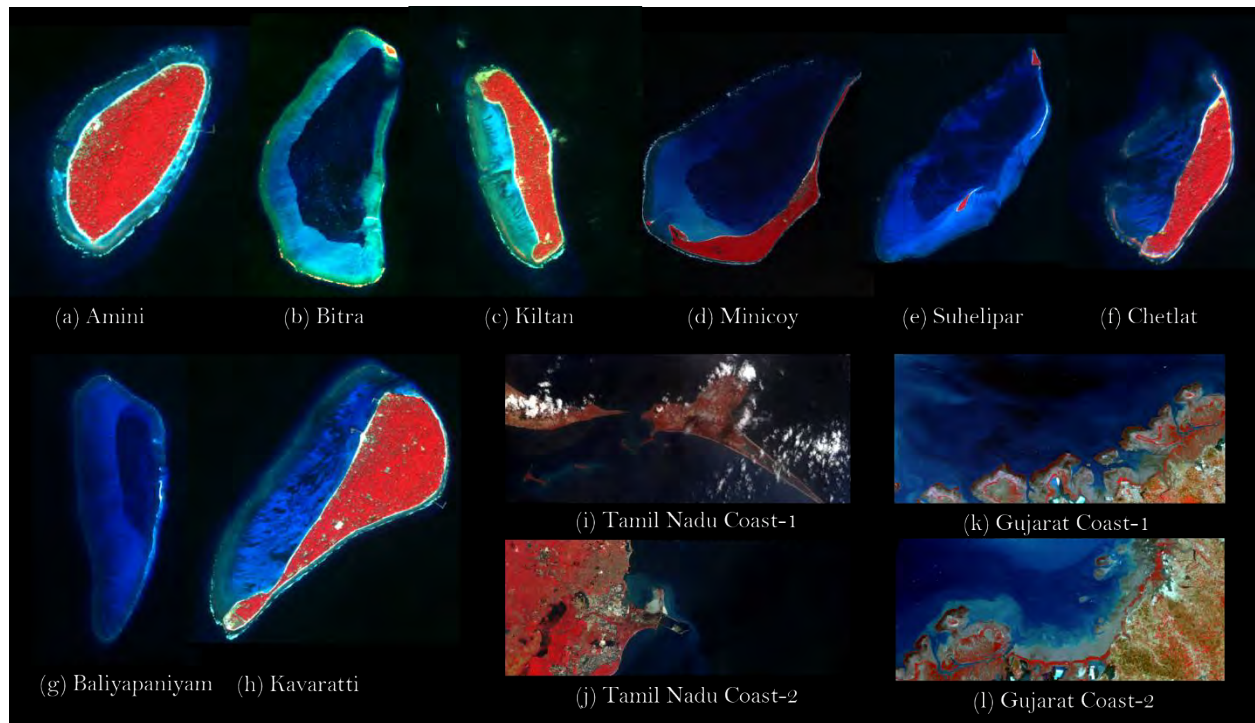


Figure-4.20. Evaluation of Image Registration across the Lakshadweep Islands, Tamil Nadu, and Gujarat Coast Regions.

#### A. Quantitative Evaluation and Comparative Analysis

The Root Mean Square Error (RMSE) is a suitable metric for assessing the performance of image co-registration in Resourcesat images. In our context, the definition of RMSE is as follows.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|^2} \quad (4.21)$$

In this context, where  $N$  is the total number of matched points,  $X_i$  denotes the  $(x_i, y_i)$  coordinates of the reference image from Landsat-8 OLI PAN, and  $\hat{X}_i$  represents the  $(\hat{x}_i, \hat{y}_i)$  coordinates of the Resourcesat LISS-4 multi-spectral image after co-registration.

Furthermore, the Correct Matching Ratio (CMR), as calculated by Zhang in 2020, serves as an effective metric for evaluating the outcome of image registration. CMR is defined in the following manner.

$$CMR = \frac{N_c}{N_T} \quad (4.22)$$

In this context,  $N_T$  represents the number of corresponding points that remain after the elimination of outliers, while  $N_c$  is the count of accurate matches obtained after eliminating the incorrect ones.

RMSE (measured in pixels) and CMR are calculated after applying the proposed image registration methodology, MIRACLE, to correct LISS-4 images. The process of image registration is also performed using alternative cutting-edge techniques for satellite image registration. These techniques incorporate feature detection methods such as Harris with Multi-Scale Histogram (MSH) as proposed by Guo et al. (2022) and Misra et al. (2012), along with SIFT combined with the robust statistical outlier removal method RANSAC, as suggested by Misra et al. (2021c). Furthermore, ORB (Luo et al., 2019) is utilized in conjunction with Motion Smoothness Constraint (MSC) (Bian et al., 2017) to produce refined sets of matched control points.

Furthermore, various recent methods for image registration, such as SIFT with Mode Guidance (Misra et al., 2021a), SURF with Steerable Filter Descriptor (SFD) (Ye et al., 2022a) coupled with M-Estimator Consensus (Wu et al., 2020), HarrisZ+ (Bellavia & Mishkin, 2022) with Local Self Similarity (LSS) descriptor (Ye & Shan, 2014), KAZE (Alcantarilla & Solutions, 2011) with MAGSAC (Barath & Matas, 2021), and Nonlinear SIFT (Yu et al., 2021) with Triangular Area Ratio (TAR) (Guo et al., 2022), have been selected for comparison with the proposed method. Recent techniques in remote sensing image registration based on deep learning, such as the Deep Neural Network (DNN) employing the Semantic Template (ST) matching method (Li et al., 2021) and the multi-scale framework involving unsupervised learning known as MU-NET (Ye et al., 2022b), are also included for evaluation and comparative analysis. All the image registration



methods selected for comparison have been developed within the last three years. Tables 4.10 and 4.11 present comparisons of RMSE and CMR for LISS-4 island scenes using various state-of-the-art image registration techniques.

It is clear that the MIRACLE method, as proposed, attained an RMSE of less than half a pixel, a level of performance not observed in any other image registration technique across diverse island scenes. The approach demonstrates superiority in terms of CMR, extracting the greatest number of correctly matched points for all the islands with limited features captured by Resourcesat LISS-4 and Landsat-8 images.

Table-4.10. Comparison of RMSE for Island Scenes employing Different Techniques for Remote Sensing Image Registration

Islands	ORB+MSC 2019)	SURF+SFD+ MSAC (2020)	SIFT+ RANSAC (2021)	SIFT+MG (2021)	KAZE+ MAGSAC (2021)	Harris+MSH (2022)	HarrisZ'+ LSS+MLE SAC (2022)	Nonlinear SIFT+TAR (2022)	MU- NET (2022)	DNN+ ST (2021)	MIRACLE
Amini	0.71	0.66	0.59	0.51	0.54	0.66	0.52	0.52	0.51	0.58	0.45
Bitra	0.74	0.75	0.67	0.58	0.66	0.81	0.67	0.59	0.57	0.63	0.47
Kiltan	0.63	0.61	0.52	0.52	0.55	0.69	0.51	0.58	0.65	0.61	0.43
Minicoy	0.79	0.75	0.78	0.56	0.58	0.87	0.61	0.60	0.60	0.67	0.49
Suhelipar	0.88	0.89	0.82	0.71	0.59	0.66	0.62	0.57	0.62	0.59	0.48
Chetlat	0.81	0.72	0.79	0.69	0.54	0.82	0.62	0.58	0.58	0.71	0.46
Baliyapaniyam	0.75	0.68	0.78	0.67	0.64	0.77	0.61	0.51	0.59	0.51	0.44
Kavaratti	0.78	0.81	0.75	0.80	0.69	0.64	0.78	0.62	0.67	0.63	0.49
TN Coast1	0.69	0.74	0.68	0.56	0.54	0.85	0.67	0.52	0.62	0.53	0.43
TN Coast2	0.78	0.65	0.77	0.68	0.63	0.85	0.77	0.54	0.55	0.58	0.48
Gujarat Coast1	0.53	0.66	0.72	0.72	0.55	0.64	0.58	0.68	0.52	0.61	0.41
Gujarat Coast2	0.78	0.83	0.72	0.81	0.52	0.62	0.68	0.51	0.61	0.70	0.46

Table-4.11. Comparison of CMR for Island Scenes employing Different Techniques for Remote Sensing Image Registration

Islands	ORB +MSC (2019)	SURF+SFD+ MSAC (2020)	SIFT + RANSAC (2021)	SIFT +MG (2021)	KAZE+ MAGSAC (2021)	Harris +MSH (2022)	HarrisZ'+ LSS+MLESAC (2022)	Nonlinear SIFT+TAR (2022)	MU-NET (2022)	DNN+ ST (2021)	MIRACLE
Amini	0.91	0.92	0.93	0.92	0.92	0.90	0.91	0.93	0.92	0.90	0.95
Bitra	0.85	0.84	0.82	0.85	0.81	0.83	0.83	0.85	0.84	0.86	0.88
Kiltan	0.81	0.82	0.91	0.88	0.87	0.87	0.89	0.89	0.92	0.91	0.94
Minicoy	0.92	0.88	0.89	0.79	0.89	0.91	0.81	0.82	0.83	0.88	0.90
Suhelipar	0.91	0.84	0.86	0.82	0.83	0.87	0.86	0.85	0.84	0.89	0.91
Chetlat	0.78	0.81	0.88	0.83	0.85	0.85	0.84	0.86	0.86	0.82	0.90
Baliyapaniyam	0.83	0.84	0.81	0.85	0.80	0.82	0.88	0.83	0.87	0.81	0.92
Kavaratti	0.83	0.86	0.81	0.82	0.88	0.84	0.85	0.83	0.84	0.89	0.91
TN Coast1	0.82	0.87	0.83	0.88	0.84	0.81	0.82	0.87	0.86	0.88	0.89
TN Coast2	0.88	0.90	0.86	0.87	0.91	0.82	0.83	0.91	0.92	0.89	0.93
Gujarat Coast1	0.83	0.87	0.82	0.89	0.82	0.86	0.81	0.83	0.93	0.91	0.95
Gujarat Coast2	0.80	0.81	0.82	0.82	0.84	0.78	0.86	0.82	0.84	0.81	0.90

Table-4.12. Processing Time (in seconds) for Image Registration Methods across Various Island Image Scenes

Islands	ORB + MSC (2019)	SURF+SFD+ MSAC (2020)	SIFT + RANSAC (2021)	SIFT +MG (2021)	KAZE+ MAGSAC (2021)	Harris +MSH (2022)	HarrisZ'+ LSS+MLESAC (2022)	Nonlinear SIFT+TAR (2022)	MU-NET (2022)	DNN+ ST (2021)	MIRACLE
Amini	183	193	199	201	187	189	192	205	211	215	184
Bitra	185	188	192	196	185	188	198	199	207	201	181
Kiltan	187	186	196	197	188	186	195	199	215	204	182
Minicoy	181	193	195	199	184	189	199	202	210	207	178
Suhelipar	170	189	201	200	174	199	202	204	203	209	169
Chetlat	165	192	197	199	168	194	197	206	206	216	163
Baliyapaniyam	178	185	193	198	181	183	193	208	211	215	173
Kavaratti	167	189	188	191	178	178	189	197	204	199	168
TN Coast1	202	211	216	222	205	206	213	234	244	264	201
TN Coast2	199	208	219	224	204	215	218	239	239	269	198
Gujarat Coast1	201	219	251	254	212	231	241	244	251	254	199
Gujarat Coast2	202	217	248	251	219	242	248	258	272	270	202

Table-4.13. Performance and Comparative Analysis of Outlier Removal Techniques

Islands	MSC (Recall, Precision)	MLESAC (Recall, Precision)	RANSAC (Recall, Precision)	MAGSAC (Recall, Precision)	Proposed MDGM (Recall, Precision)
Amini	(0.85, 0.88)	(0.88, 0.90)	(0.87, 0.89)	(0.90, 0.92)	(0.92, 0.94)
Bitra	(0.88, 0.91)	(0.87, 0.89)	(0.85, 0.86)	(0.92, 0.94)	(0.94, 0.97)
Kiltan	(0.87, 0.89)	(0.85, 0.86)	(0.82, 0.85)	(0.88, 0.90)	(0.91, 0.95)
Minicoy	(0.84, 0.87)	(0.88, 0.90)	(0.87, 0.89)	(0.87, 0.91)	(0.89, 0.92)
Suhelipar	(0.88, 0.90)	(0.86, 0.88)	(0.84, 0.87)	(0.89, 0.92)	(0.94, 0.96)
Chetlat	(0.92, 0.94)	(0.90, 0.92)	(0.88, 0.90)	(0.90, 0.93)	(0.91, 0.95)
Baliyapaniyam	(0.81, 0.83)	(0.84, 0.86)	(0.79, 0.81)	(0.84, 0.86)	(0.88, 0.92)
Kavaratti	(0.83, 0.85)	(0.86, 0.88)	(0.81, 0.84)	(0.93, 0.97)	(0.94, 0.97)
TN Coast1	(0.86, 0.88)	(0.84, 0.89)	(0.83, 0.86)	(0.86, 0.88)	(0.87, 0.90)
TN Coast2	(0.89, 0.91)	(0.87, 0.91)	(0.82, 0.85)	(0.91, 0.94)	(0.92, 0.96)
Gujarat Coast1	(0.83, 0.86)	(0.88, 0.90)	(0.86, 0.89)	(0.88, 0.92)	(0.89, 0.94)
Gujarat Coast2	(0.85, 0.88)	(0.78, 0.82)	(0.82, 0.87)	(0.90, 0.94)	(0.93, 0.95)

The computation of processing time is carried out to facilitate the comparison of different image registration techniques. Table-4.12 displays the processing time of distinct image registration techniques for island scenes, each having an image size of approximately 16000 x 16000 pixels. The processing time for island image registration using the proposed MIRACLE method is observed to be shorter when compared to other well-known image registration techniques.

To assess the effectiveness of the outlier removal stage, quality metrics such as *Recall* and *Precision* are calculated for the proposed optimization method, Mahalanobis Distance Guided Marginalization (MDGM), and compared to alternative state-of-the-art outlier removal techniques like RANSAC (Fischler & Bolles, 1981), MLESAC (Torr & Zisserman, 2000), MSC (Bian et al., 2017), and MAGSAC (Barath et al., 2019). *Recall* and *Precision* are defined as follows:

$$Recall = \frac{True\_Positives}{Total\_Positives} \quad (4.23)$$

$$Precision = \frac{True\_Positives}{True\_Positives + Wrong\_Positives} \quad (4.24)$$

In this context, where *True\_Positives* denote the count of accurately matched keypoints, *Wrong\_Positives* represent the count of incorrectly matched keypoints, and *Total\_Positives* is the initial total count of correctly matched keypoints. Table-4.13 presents the recall and precision efficacy across different techniques for outlier removal. The results clearly demonstrate that the proposed outlier removal technique, MDGM, achieves higher levels of recall and precision for all remote sensing images of the islands.

### *B. Visual Evaluation of Image Quality across Island Regions*

The visual evaluation of the registration accuracy achieved by the proposed image registration method, MIRACLE, is conducted across different regions at a 2X zoom level to confirm its effectiveness. Figures 4.21, 4.22, and 4.23 showcase the Resourcesat LISS-4 multi-spectral image in False Color Composite (FCC) juxtaposed with the Landsat-8 OLI PAN reference image, which has undergone horizontal and vertical shifts before correction. These figures unveil noticeable misalignment between the images in various island regions.

The co-registered output data product from multiple satellites illustrates the sub-pixel alignment accomplished through the proposed method, MIRACLE, affirming that the images are precisely overlaid. Checkerboard patterns are created between the registered image of the red wavelength channel from LISS-4 and the reference image from Landsat-8 OLI PAN. The interleaved tiles in the checkerboard pattern between the images from multiple satellites signify geometric consistency. The checkerboard pattern depicted in Figure-4.24 confirms that feature continuity is preserved across various geographic locations of islands, and the images are aligned at a sub-pixel level.

Besides visual interpretation, Ground Control Points (GCPs) are chosen manually at various locations on multiple islands, as outlined in Table-4.14. The pixel shifts at GCPs are calculated, indicating an average pixel shift of approximately 25 pixels horizontally and 15 pixels vertically between the Resourcesat LISS-4 image and the Landsat-8 OLI PAN image. Following correction with MIRACLE, the RMSE-based measurement for the output pixel shift is 0.3 pixels.

Additionally, MIRACLE is assessed for multi-temporal remote sensing imagery encompassing these islands or neighboring areas with varying spectral characteristics. The results suggest that MIRACLE attains sub-pixel co-registration accuracy in multi-date imaging acquisitions, showcasing consistency in an extensive dataset of remote sensing images.

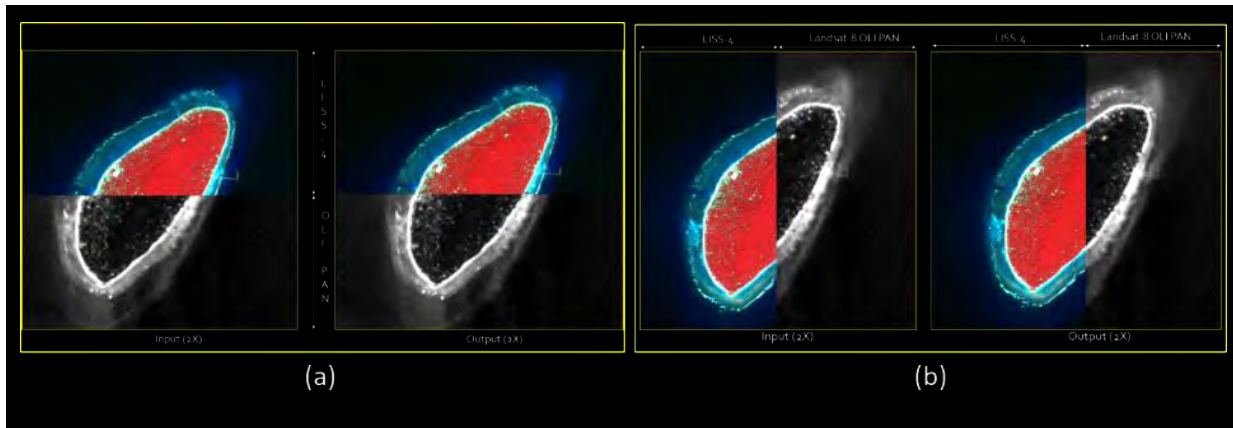


Figure-4.21. Quality Assessment of Co-registration of Amini Island: (a) Vertical adjustments prior to and following correction, (b) Horizontal shifts prior to and following correction

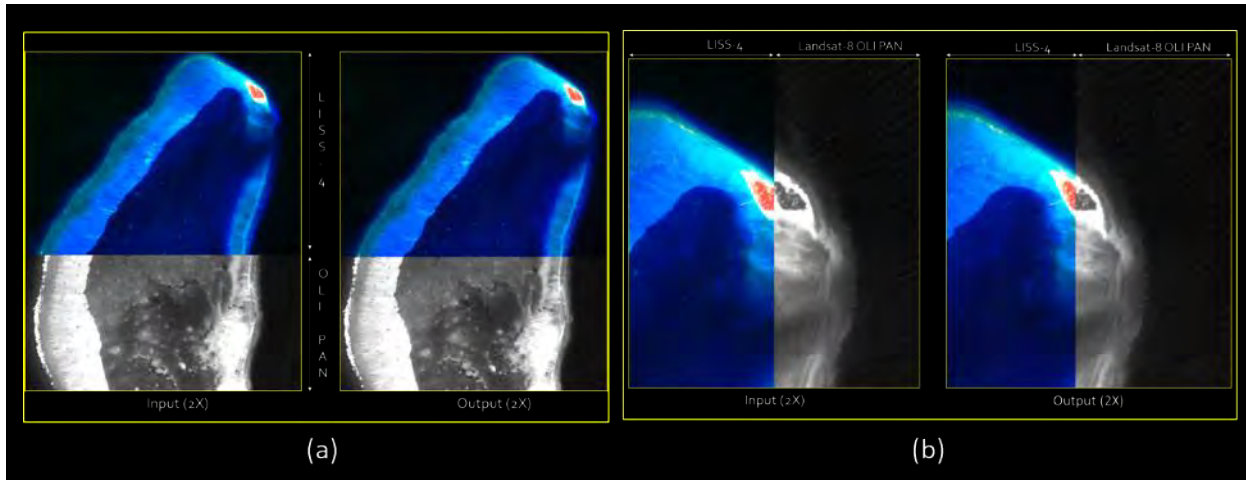


Figure-4.22. Quality Assessment of Co-Registration of Bitra Island (a) Vertical shifts prior to and following correction, (b) Horizontal shifts prior to and following correction

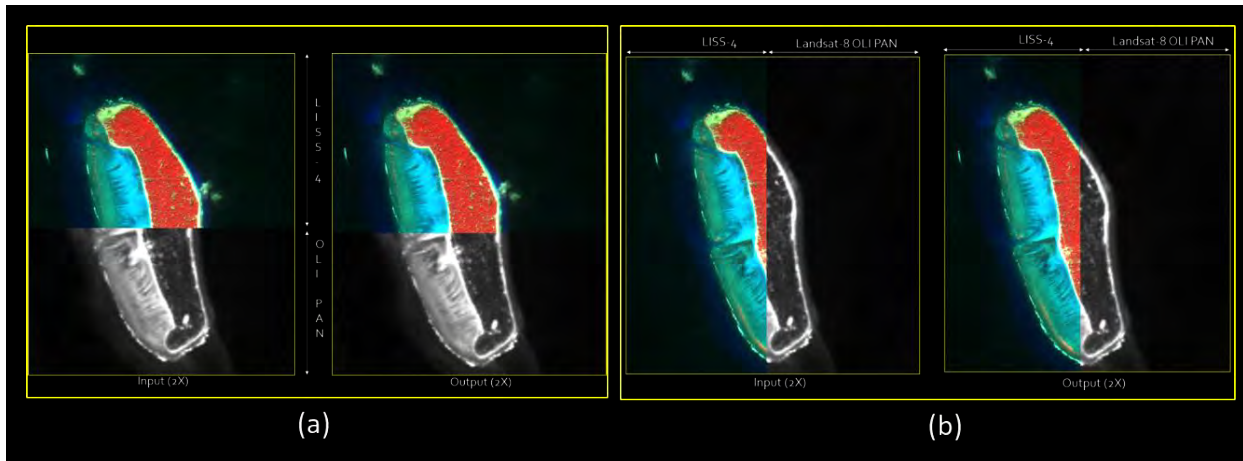


Figure-4.23. Quality Assessment of Co-Registration of Kiltan Island: (a) Vertical shifts prior to and following correction, (b) Horizontal shifts prior to and following correction

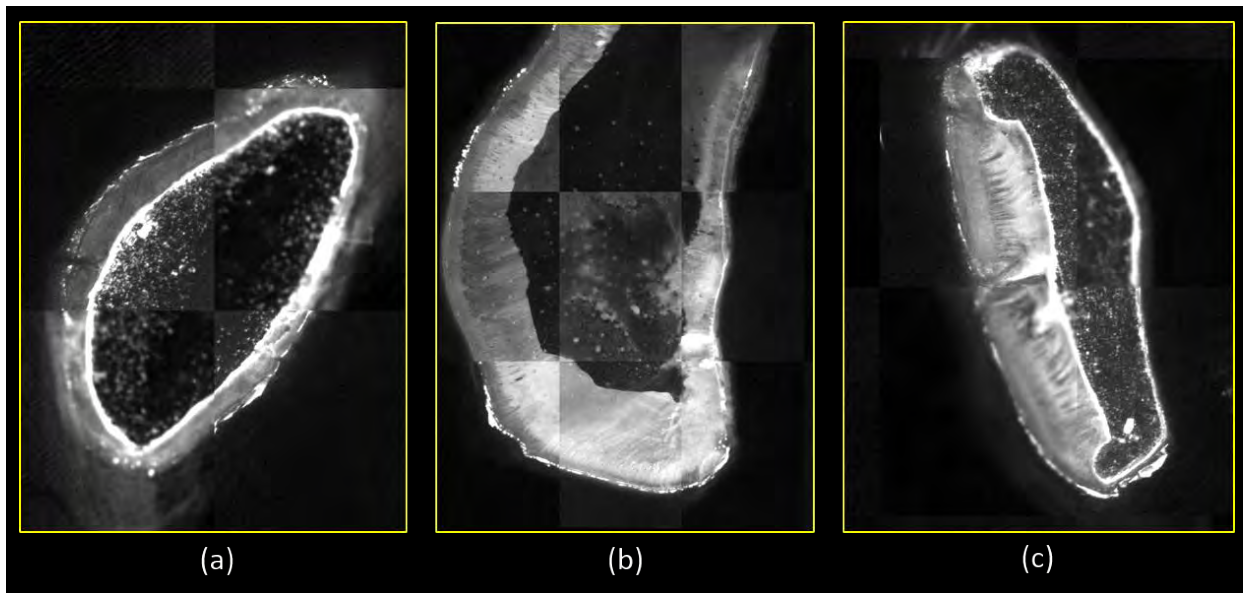



Figure-4.24. Evaluation of Checkerboard Pattern in Co-registration of images between the red channel registered image from Resourcesat LISS-4 and the Landsat-8 OLI PAN image (a) Amini, (b) Bitra, (c) Kiltan

Table-4.14. Shifts in Multi-Satellite Images prior to and following Correction at Ground Control Points (GCPs)

GCPs Marked over Amini and Kadmat Feature Islands	Input Vertical Shift (in pixels)	Input Horizontal Shift (in pixels)	Output Vertical Shift (in pixels)	Output Horizontal Shift (in pixels)
	15.95	23.53	-0.23	0.41
	16.07	24.16	0.22	0.35
	16.36	23.82	0.41	0.15
	16.21	23.95	-0.31	0.27
	14.93	24.83	0.28	0.45
	15.78	24.77	0.33	-0.24
	13.49	24.07	0.42	0.35
	15.08	23.59	-0.36	0.25
	15.64	25.47	-0.31	0.46
	16.1	24.32	0.41	0.37
	13.71	27.33	-0.21	0.29
	16.11	24.31	0.12	0.32
	11.67	28.56	0.26	0.31
	14.32	26.23	0.18	0.47

### C. Ablation Study

The MIRACLE methodology consists of several components, each playing a crucial role in every phase of the image registration process. Among these, the data pre-processing engine is essential and plays a central role in achieving sub-pixel image co-registration for island scenes. An experimental analysis is performed on various pairs of remote sensing images, denoted as  $MIRACLE^{NOENH}$ , wherein MIRACLE is implemented without the data pre-processing engine. This omits the calculation of PCA components and the enhancement of anisotropic coherence for the multi-spectral image. Moreover, the anisotropic coherence enhancement is not applied to the

reference image, constituting part of the data pre-processing stage. The guidance provided by Mahalanobis distance is a crucial aspect in generating initially pruned matched keypoints. MIRACLE<sup>NOMAH</sup> implements the entire proposed methodology without the use of Mahalanobis distance guidance for various remote sensing images of islands. Table-4.15 displays the RMSE performance of MIRACLE, MIRACLE<sup>NOENH</sup>, and MIRACLE<sup>NOMAH</sup> across all scenarios, clearly demonstrating that MIRACLE surpasses both MIRACLE<sup>NOENH</sup> and MIRACLE<sup>NOMAH</sup>. Table-4.16 presents the CMR estimates for all island scenes, indicating that the matching performance of MIRACLE is superior to both MIRACLE<sup>NOENH</sup> and MIRACLE<sup>NOMAH</sup>.

Table-4.15. RMSE Performance Comparison of MIRACLE<sup>NOENH</sup>, MIRACLE<sup>NOMAH</sup>, and MIRACLE Across Island Scenes

Islands	MIRACLE <sup>NOMAH</sup>	MIRACLE <sup>NOENH</sup>	MIRACLE
Amini	0.59	0.63	0.45
Bitra	0.57	0.51	0.47
Kiltan	0.63	0.58	0.43
Minicoy	0.68	0.57	0.49
Suhelipar	0.52	0.55	0.48
Chetlat	0.63	0.61	0.46
Baliyapaniyam	0.68	0.57	0.44
Kavaratti	0.55	0.74	0.49
TN Coast1	0.48	0.85	0.43
TN Coast2	0.57	0.65	0.48
Gujarat Coast1	0.77	0.67	0.41
Gujarat Coast2	0.62	0.59	0.46



Table-4.16. CMR Performance Comparison of MIRACLE<sup>NOENH</sup>, MIRACLE<sup>NOMAH</sup>, and  
MIRACLE Across Island Scenes

Islands	MIRACLE <sup>NOENH</sup>	MIRACLE <sup>NOMAH</sup>	MIRACLE
Amini	0.83	0.88	0.95
Bitra	0.81	0.87	0.88
Kiltan	0.88	0.83	0.94
Minicoy	0.87	0.88	0.90
Suhelipar	0.85	0.89	0.91
Chetlat	0.81	0.84	0.90
Baliyapaniyam	0.87	0.88	0.92
Kavaratti	0.84	0.82	0.91
TN Coast1	0.83	0.85	0.89
TN Coast2	0.81	0.81	0.93
Gujarat Coast1	0.87	0.87	0.95
Gujarat Coast2	0.89	0.88	0.90

Table-4.17. Comparison of Matched points after pruning and the corresponding matching time  
(in seconds) with and without the enhancement-based data pre-processing stage

Islands	MIRACLE (No. of Pruned Matched Points, Matching Time)	MIRACLE <sup>NOENH</sup> (No. of Pruned Matched Points, Matching Time)
Amini	(6188, 184)	(5988, 178)
Bitra	(5435, 184)	(5089, 181)
Kiltan	(6241, 182)	(5761, 175)
Minicoy	(6723, 178)	(6127, 169)
Suhelipar	(6823, 169)	(6327, 161)
Chetlat	(4633, 163)	(4208, 157)
Baliyapaniyam	(5874, 173)	(5342, 168)
Kavaratti	(6823, 168)	(6428, 162)
TN Coast1	(5165, 201)	(4783, 197)
TN Coast2	(6287, 198)	(6012, 191)
Gujarat Coast1	(5124, 199)	(4723, 192)
Gujarat Coast2	(6234, 202)	(6103, 194)

Furthermore, the impact of anisotropic coherence locality enhancement and the PCA algorithm in the proposed method, MIRACLE, is assessed by calculating pruned matched control points and matching time with or without this step. Table-4.17 presents the summary of pruned matched points and the corresponding matching time (in seconds) for both MIRACLE<sup>NOENH</sup> and MIRACLE. The findings reveal that the number of pruned matched points is higher for MIRACLE compared to MIRACLE<sup>NOENH</sup> across all remote sensing image scenes featuring islands with limited features. This suggests that the data pre-processing stage indeed enhances the matching between the input and reference satellite images. However, the matching time for MIRACLE is slightly longer than for MIRACLE<sup>NOENH</sup>, primarily due to a relatively larger number of pruned matched control points.

#### **4.2.4. Summary**

This section introduces a novel framework for multi-satellite image registration, termed as MIRACLE, designed specifically for remote sensing images with limited features. The objective is to align these images with a standard reference image to achieve the necessary geometric location accuracy. Enhancing both the input and reference images through anisotropic coherence localization improves feature interpretation for image registration purposes. Converting the input multi-spectral image using Principal Component Analysis (PCA) maximizes variance information, which is crucial for extracting stable feature points. The integration of nonlinear diffusion filtering with Mahalanobis Distance Guided Marginalization Outlier Removal forms a robust technique for generating potential matched point correspondences, facilitating accurate estimation of affine parameters.

MIRACLE is assessed using multi-temporal scenes from Resourcesat LISS-4, encompassing areas with islands characterized by limited features. It aligns these scenes with Landsat-8 OLI PAN reference images, achieving registration accuracy at the sub-pixel level. Quantitative measures are contrasted with feature-based image registration techniques, and MIRACLE consistently outperforms others across a range of image quality metrics. Visual evaluations, which include horizontal and vertical swipe, checkerboard pattern assessment, and pixel shifts at various Ground Control Points (GCPs), affirm the successful overlay of island regions. In the future, MIRACLE will be utilized to co-register remote sensing images with very high spatial resolution that capture deep ocean areas containing islands.

## **4.3. Band-to-Band Registration of Multispectral Images**

### **4.3.1. Introduction**

The section introduces an innovative approach to improving Band-to-Band Registration (BBR) by employing Co-occurrence Scale Space (CSS) for stable feature point extraction and description. It is noted that the use of conventional estimation methods results in a significant number of outliers in the matched control points, leading to increased BBR error. To address this issue, we propose a unique method called Spatial Confined RANSAC (SC-RANSAC) to eliminate outliers and enhance the estimation of transformation parameters for BBR.

To mitigate local distortion between multispectral bands, we employ a Segmented Affine Transformation (SAT) model. The proposed method is transformed into a processing pipeline and assessed using multispectral remote sensing images from the Indian Nano Satellite (INS), which exhibit band-to-band misregistration at the system level. The spatial misalignment results in significant spectral distortion and blurring in the False Color Composite (FCC) of INS multispectral images that combine Near Infrared (NIR), Red (R), and Green (G) wavelength channels.

Our methodology rectifies multiple band alignments at a sub-pixel level, and a visual quality check is conducted to evaluate BBR performance. Quantitative evaluations are performed in different regions by calculating sub-pixel band shifts in both horizontal and vertical directions. The developed methodology is compared with state-of-the-art image registration techniques. We evaluate multispectral remote sensing imagery with band-to-band misregistration errors across various spectral classes, computing image quality metrics for BBR assessment.

### **4.3.2. Methodology Developed**

#### *A. Feature Point Detection using Co-occurrence Scale Space (CSS)*

The Co-occurrence Filter (CoF) functions as a boundary-preserving filter, deviating from the conventional use of a Gaussian on range values for edge preservation by relying on a co-occurrence matrix (Jevnisek & Avidan, 2017). In CoF, pixels with frequent co-occurrences, typically found in textured regions of an image, are assigned higher weights, contributing to a smoother image through blending. Conversely, pixel values with infrequent co-occurrences receive lower weights, preserving the boundaries between textures. It is noteworthy that co-occurrence information not

only enhances edge preservation but also effectively addresses boundary-related aspects. The CoF is formally defined as follows.

$$Co_p = \frac{\sum_{q \in N(p)} G_{\sigma_s}(p,q) \cdot M(I_p I_q) \cdot I_q}{\sum_{q \in N(p)} G_{\sigma_s}(p,q) \cdot M(I_p I_q)} \quad (4.25)$$

Here,  $Co_p$  represents the output pixel image value at the pixel index  $p$ , while  $I_p$  and  $I_q$  denote the input pixel image values at pixel indices  $p$  and  $q$ , respectively.  $G_{\sigma_s}$  refers to the Gaussian filter, with  $\sigma_s$  being the user-specified parameter.

$M$  is a matrix of dimensions 256 by 256, and it can be expressed as follows:

$$M(a, b) = \frac{C(a,b)}{H(a)H(b)} \quad (4.26)$$

In this context,  $M(a, b)$  is derived from the co-occurrence matrix  $C(a, b)$ , where  $C(a, b)$  records the instances of the values  $a$  and  $b$  occurring together. The frequencies of the pixel values  $a$  and  $b$ , denoted as  $H(a)$  and  $H(b)$  respectively, can be calculated from the histogram of pixel values.

The proposed algorithm for feature detection begins by constructing scale space extrema through the use of the Co-occurrence Kernel. It has been observed that the Laplacian of Co-occurrence (LoC) offers stable features and a robust representation of scales, but it comes with a high computational cost. To address this, a Difference of Co-occurrence (DoC) pyramid is generated, serving as a close approximation to the Laplacian of Co-occurrence (LoC). Equation-4.27 establishes the relationship between the Co-occurrence Smooth Image ( $L$ ) and the Input Image ( $I$ ) at the point  $(x, y)$ .

$$L(x, y, \sigma) = Co(x, y, \sigma) * I(x, y) \quad (4.27)$$

In this context,  $Co(x, y, \sigma)$  represents the CoF as defined in Equation-4.25.

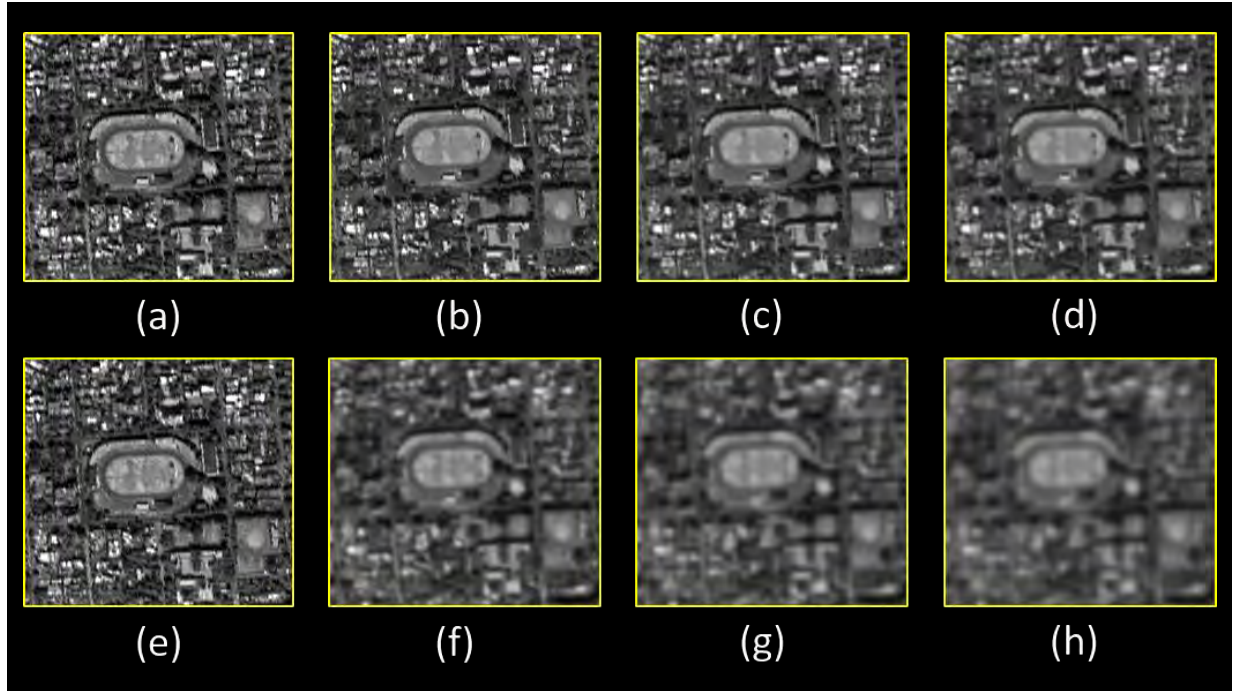


Figure-4.25. Scale Space Construction for Feature Detection and Description (a) Input Image for Co-occurrence Scale Space (CSS) (b) CSS at Layer-1 (c) CSS at Layer-2 (d) CSS at Layer-3 (e) Input Image for SIFT Scale Space (f) SIFT Scale Space at Layer-1 (g) SIFT Scale Space at Layer-2 (h) SIFT Scale Space at Layer-3

The subsequent phase involves the generation of a Difference of Co-occurrence (DoC) pyramid to detect minimum and maximum values across the immediate points, incorporating the scale. Figure-4.25 illustrates the creation of the DoC pyramid and the localization of extrema utilizing various scales. The DoC is computed by taking the difference between two neighboring scales separated by a constant multiplicative factor  $k$ . The formulation for DoC is expressed as follows.

$$DoC(x, y, \sigma) = (Co(x, y, k\sigma) - Co(x, y, \sigma)) * I(x, y) \quad (4.28)$$

Using equation (4.27), we can rewrite equation (4.28) in terms of smooth image  $L$  as:

$$DoC(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (4.29)$$

To achieve sub-pixel level localization of potential stable feature points and eliminate initial outliers, Taylor Series approximation is employed. In this context, the Taylor Expansion, considering up to quadratic terms, of the scale-space function  $DoC(Z)$  is shifted such that the origin aligns with the sample point.

$$DoC(Z) = DoC + \frac{\partial DoC^T}{\partial Z} Z + \frac{1}{2} Z^T \frac{\partial^2 DoC^T}{\partial Z^2} Z \quad (4.30)$$

In this expression,  $DoC$  and its derivatives are assessed at the sample point, with  $Z = (x, y, \sigma)$  representing the offset from this particular point. The extremum's location, denoted as  $Z$ , is determined by equating the derivative to zero.

$$\hat{Z} = -\frac{\partial^2 DoC^{-1}}{\partial Z^2} \frac{\partial DoC}{\partial Z} \quad (4.31)$$

The orientation of keypoints is determined by calculating the gradient of each blurred image. This orientation assignment ensures rotational invariance for the keypoints. The process primarily involves computing central derivatives, gradient magnitudes ( $m$ ), and direction ( $\theta$ ) of the smooth image ( $L$ ) at the keypoint's scale ( $x, y$ ). A weighted direction histogram is then generated within the keypoint's vicinity, structured as bins. Ultimately, the direction of the keypoint is identified by selecting the peak of the histogram.

$$m(x, y) = \sqrt{\begin{matrix} (L(x+1, y) - L(x-1, y))^2 + \\ (L(x, y+1) - L(x, y-1))^2 \end{matrix}} \quad (4.32)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (4.33)$$

The process of constructing descriptors begins with the sampling of points around the keypoint. The gradients and coordinates are then rotated according to the previously computed orientation, and the region is subdivided into smaller regions. A histogram is generated for each sub-region with specified bins. The descriptor for each keypoint is stored as an element vector, which is utilized for feature matching across multispectral bands. Figure 4.25 illustrates the construction of the Co-occurrence Scale Space (CSS) for feature detection and description. Additionally, for comparison purposes, a scale space is created using the SIFT operator in Figure-4.25. Notably, it is observed that SIFT, at higher scale space, fails to preserve sharp image boundary features, appearing almost blurred, and consequently, numerous distinct feature points are overlooked during detection. Conversely, the proposed co-occurrence scale space maintains a consistent representation across scale levels, demonstrating that even at higher scale space, prominent feature points remain detectable for the Band-to-Band Registration (BBR) task.

### B. Spatial Confined RANSAC (SC-RANSAC)

Random Sample Consensus (RANSAC) is an effective technique for removing outliers, generating potential solutions by employing the minimum required number of observations

(control points) to estimate the underlying model parameters (Fischler & Bolles, 1981). The RANSAC process unfolds in two phases: Hypothesis Generation, during which a set of matched keypoints is randomly selected, and Hypothesis Evaluation, where the chosen hypothesis is assessed to determine if it meets the criteria for inlier candidates whose model error is within a predefined threshold.

The selection of the number of iterations, denoted as  $N$ , is crucial to ensure that at least one set of randomly sampled data does not include an outlier. It's important to note that RANSAC utilizes a subset of the data rather than the entire dataset. If the chosen data consist entirely of inliers, they can contribute to a hypothesis that closely approximates the true model. This assumption dictates that the number of iterations  $N$  should be sufficiently high to ensure the inclusion of all inlier samples at least once.

$$N = \frac{\log \alpha}{\log(1-\gamma^m)} \quad (4.34)$$

In this context, where  $\alpha$  represents the failure probability,  $m$  denotes the number of matched keypoints required to generate a hypothesis, and  $\gamma$  is the probability of selecting an inlier—specifically, the ratio of inliers to the entire sample data (inliers ratio).

Nevertheless, the inliers ratio  $\gamma$  remains unknown in numerous practical scenarios and requires empirical determination by users. RANSAC transforms an estimation problem in the continuous domain into a selection problem in the discrete domain. The challenge becomes selecting the most appropriate pair from an extensive array of pairs. However, RANSAC algorithms may not consistently identify accurate matched pairs devoid of outliers in multispectral remote sensing images.

Spatial Confined RANSAC (SC-RANSAC) based estimation proves effective in distinguishing between true and false matches, ensuring accurate correspondence between the input and reference images. The spatial confinement aspect assumes that a small neighborhood around a true match is perceived as the same location in the other image. Conversely, the neighborhood around a false match is considered to represent geometrically distinct locations. If spatial consistency is maintained, neighboring pixels and features exhibit cohesive movement. SC-RANSAC is transformed into a statistical model that identifies and rejects false matches, thereby increasing the

number of reliable feature correspondences. In Figure-4.26, pruned matched points using RANSAC and SC-RANSAC are illustrated, highlighting that SC-RANSAC generates a greater number of reliable pruned feature matched points for transformation. SC-RANSAC is characterized by its simplicity, robustness, and superior performance compared to classical RANSAC keypoint optimization techniques in specific types of remote sensing images, particularly for achieving sub-pixel Band-to-Band Registration (BBR).

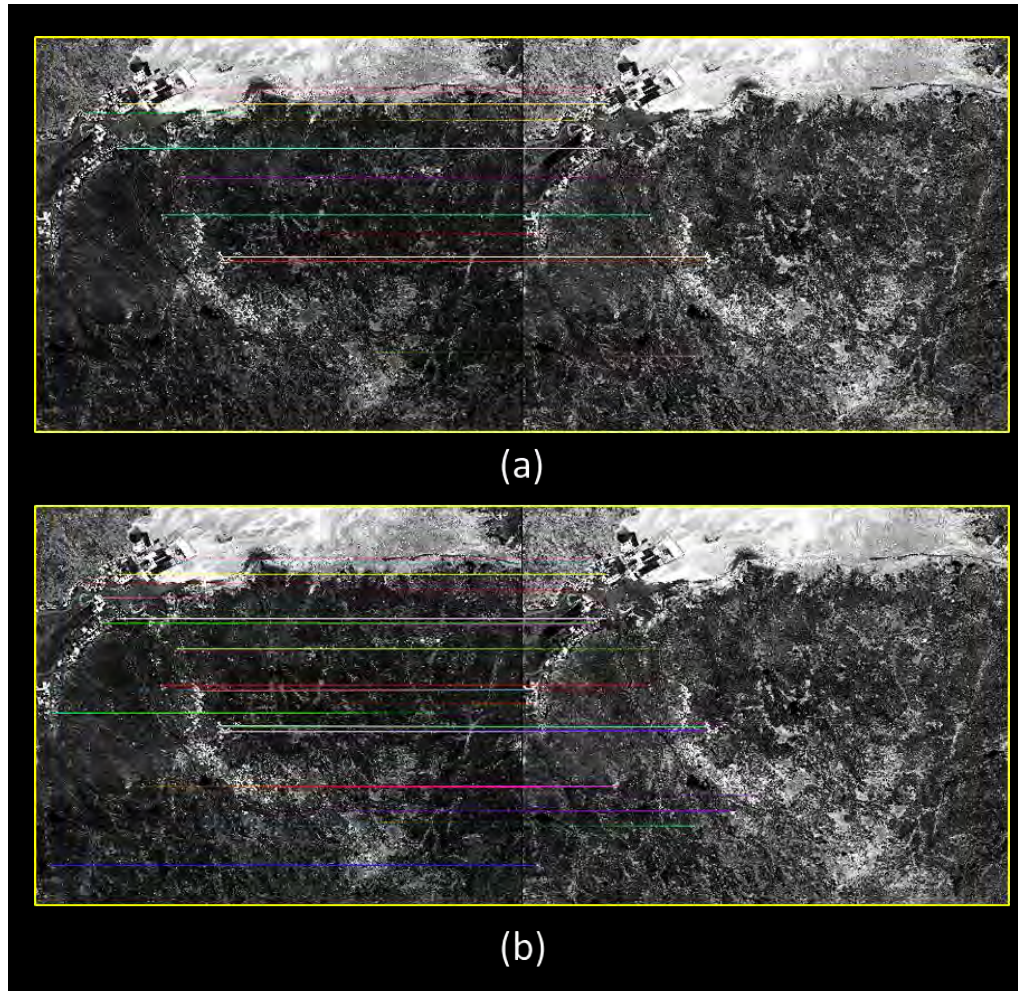


Figure-4.26. Pruned Matched Points between two Image Bands (a) RANSAC (b) SC-RANSAC

For the implementation of SC-RANSAC, the band images are divided into grid windows of equal spatial size. In our case, we have determined empirically that a grid window of 64\*64 pixels yields a satisfactory number of matched points for estimation. Within each grid cell, SC-RANSAC identifies the number of inlier feature-matched points and stores these matched points as pruned



matches for that specific grid window. This process is repeated for all grid windows, and inlier-matched feature points are extracted. SC-RANSAC proves to be a rapid, reliable, and efficient method for eliminating outlier matched points, generating potential keypoint correspondences for BBR transformation parameter estimation. The procedural steps of SC-RANSAC are outlined in Algorithm-4.4.

**Algorithm-4.4: SC-RANSAC**

1. For each grid window  $W_i$  in the input image patch, where  $i$  represents the index of the input window.
2. Randomly select a minimum number of matched points  $N_{mp}$  within  $W_i$ .
3. Estimate the model parameters for the grid window  $W_i$  using the selected  $N_{mp}$ .
4. Identify inlier matched points  $IN_{mp}$  from  $N_{mp}$  that conform to the model within a specified tolerance.
5. If the ratio of  $IN_{mp}$  to  $N_{mp}$  exceeds the inlier ratio  $I_R$ , store  $IN_{mp}$  as inlier matched points for the window  $W_i$ .
6. Otherwise, repeat steps 2 through 5 a maximum of  $N$  times, as estimated from equation-4.38.
7. Accumulate total inlier matched control points  $IN_t$  from all grid windows  $W_i$ .
8. Estimate transformation parameters  $T_p$  for BBR using the collected  $IN_t$ .

*C. Segmented Affine Transformation (SAT) and Processing Workflow*

The refined set of control points obtained through SC-RANSAC is utilized to calculate transformation parameters for Band-to-Band Registration (BBR) of multispectral images. We introduce the Segmented Affine Transformation (SAT) model, which computes affine values for distinct segments, aiming to minimize local distortion among the bands. The image bands are divided into three sections, and individual affine parameters are derived based on the control points within each specific segment. These SAT parameters are then applied to each segment, and the input multispectral band is resampled relative to the reference multispectral band. The SAT model is defined as follows.

$$SAT(k) = \{t_{0k}, t_{1k}, t_{2k}, t_{3k}, t_{4k}, t_{5k}\} \tag{4.35}$$

In this context, where  $k$  represents the segment number, and  $t_{ik}$  denotes the affine parameters.

The proposed methodology has been transformed into a comprehensive processing workflow, as illustrated in Figure-4.27, designed for the Band-to-Band Registration (BBR) of multispectral images. The input bands of the multispectral image are fed into the workflow for BBR. In this setup, a specific multispectral band is chosen as the reference, while the remaining bands serve as inputs requiring registration relative to the chosen reference band. Typically, the reference multispectral band is selected to be positioned in the middle of the electromagnetic spectrum among the available spectral bands for the instrument. In our depicted scenario in Figure-4.27, Band-3 has been designated as the reference, and all other input bands are registered in reference to Band-3.

The initial step involves detecting and describing distinctive feature points using Co-occurrence Scale Space (CSS). The subsequent removal of outlier-matched control points is carried out through the Spatial Confined RANSAC (SC-RANSAC) model. The pruned matched control points are then utilized for estimating the parameters of the Segmented Affine Transformation (SAT) model, ensuring consistent BBR across the entire image. This proposed workflow achieves the co-registration of all spectral bands within the multispectral image with sub-pixel level accuracy.

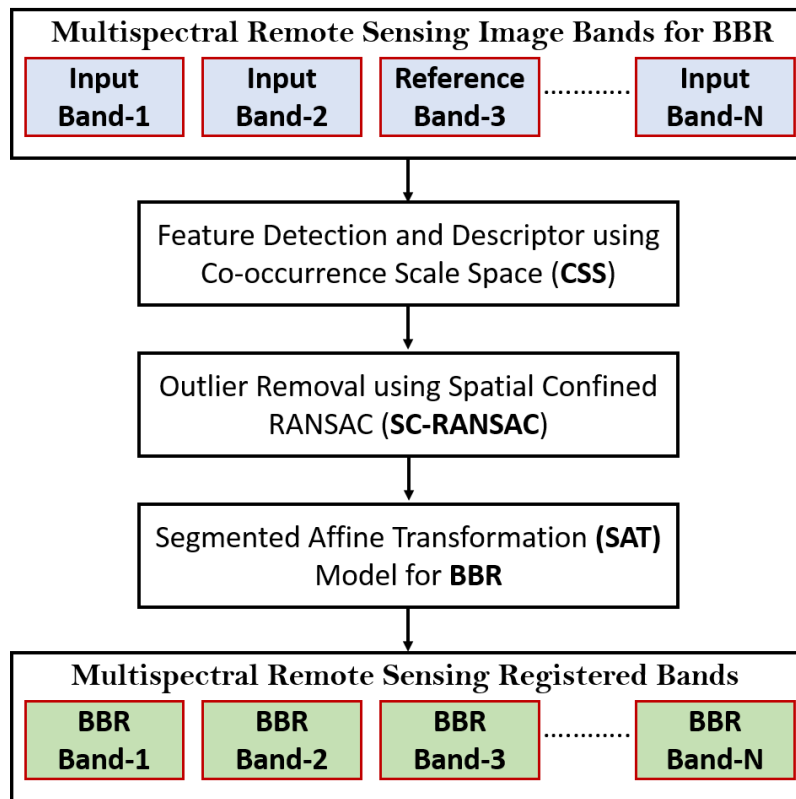


Figure-4.27. Proposed Workflow for Band-to-Band Registration (BBR)

### 4.3.3. Experimental Results

The proposed Band-to-Band Registration (BBR) methodology is assessed using INS-2B Nano-MX multispectral images captured over the Indian terrain. Nano-MX is equipped with four Visible-Near Infrared (VNIR) bands, offering a radiometric resolution of 10 bits, a spatial resolution of 30.0 meters, and covering a swath of 116 km during image acquisition. The potential applications of Nano-MX data products span a wide range, including land use/land cover mapping, forest area assessment, crop estimation, water quality index determination, and other spaceborne studies.

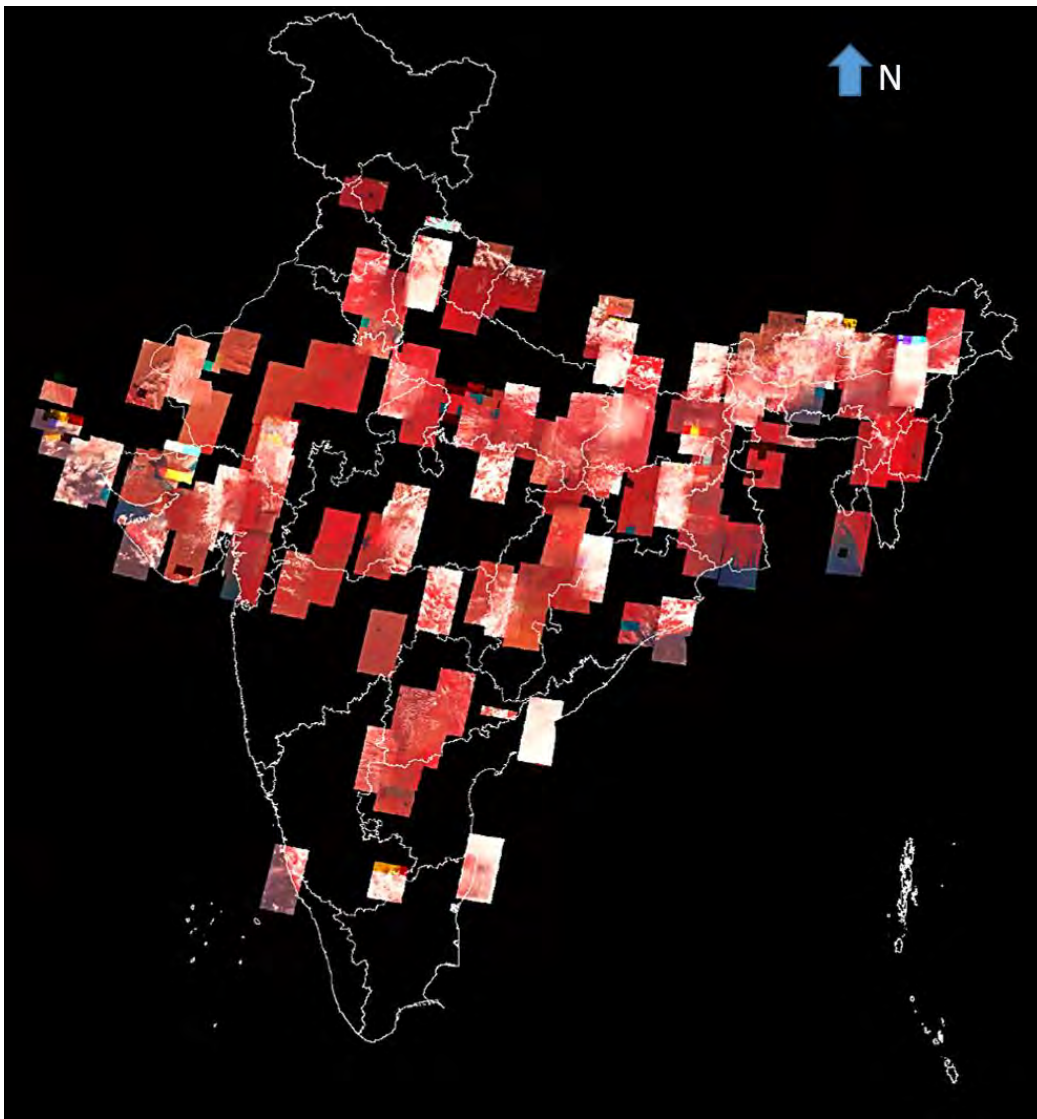


Figure-4.28. INS-2B Nano-MX Images (No of Image Scenes =220) utilized for BBR evaluation, blended over the boundary of Indian States

In our BBR experiment, approximately 220 Nano-MX images, encompassing diverse landscapes, are considered for both visual and quantitative evaluation. Figure-4.28 displays Nano-MX False Color Composite (FCC) images superimposed on an Indian state boundary shape file, serving as the basis for BBR assessment. The Nano-MX multispectral images exhibit various spectral classes such as vegetation, water, snow, and soil, allowing for the evaluation and comparison of BBR performance against state-of-the-art image registration techniques.

Additionally, clouds are present in the optical multispectral images, and the BBR performance in the presence of clouds is also assessed to verify the robustness of the proposed method.

#### *A. Visual Analysis of BBR*

The Nano-MX multispectral bands span from the blue to near-infrared wavelength range, and a noticeable misregistration is observed among these spectral bands. Band-to-Band misregistration in Nano-MX arises due to the effective focal length mismatch between the spectral bands, uncertainties in relative attitude values, and uncertainties in the knowledge of interior parameters (look angles/distortion) of different spectral bands. Figure-4.29 illustrates the BBR performance across various feature targets, including a water body, sparse urban region, and vegetation.

In the multispectral input False Color Composite (FCC) image, it is evident that the bands are not spatially aligned, leading to spectral distortion. The output multispectral image in Figure-4.29, after BBR correction using the proposed method, achieves spatial alignment at the sub-pixel level. The BBR output proves to be a spatially and spectrally consistent remote sensing image. Figure-4.30 showcases the BBR performance over hilly terrain dominated by snow and glacier regions. For studying glacier dynamics, such as ice flow and glacier surface velocity using multispectral remote sensing imagery, it is essential to register the wavelength bands. The input image in Figure-4.30 is not suitable for remote sensing image interpretation and analysis. The BBR correction output in Figure-4.30 presents a spatially aligned multispectral image, ready for advanced snow and glacier studies.

To assess the robustness of the proposed method, BBR performance is verified in an image dominated by cloudy pixels (Figure-4.31). The input composite images exhibit different color patterns around the cloudy pixels due to band-to-band misregistration. The output BBR-corrected image aligns the cloudy pixels at the sub-pixel level, appearing as pure white spectra in the False Color Composite image. Other prominent feature targets, such as water bodies and vegetation around the cloudy region, are also spatially superimposed.



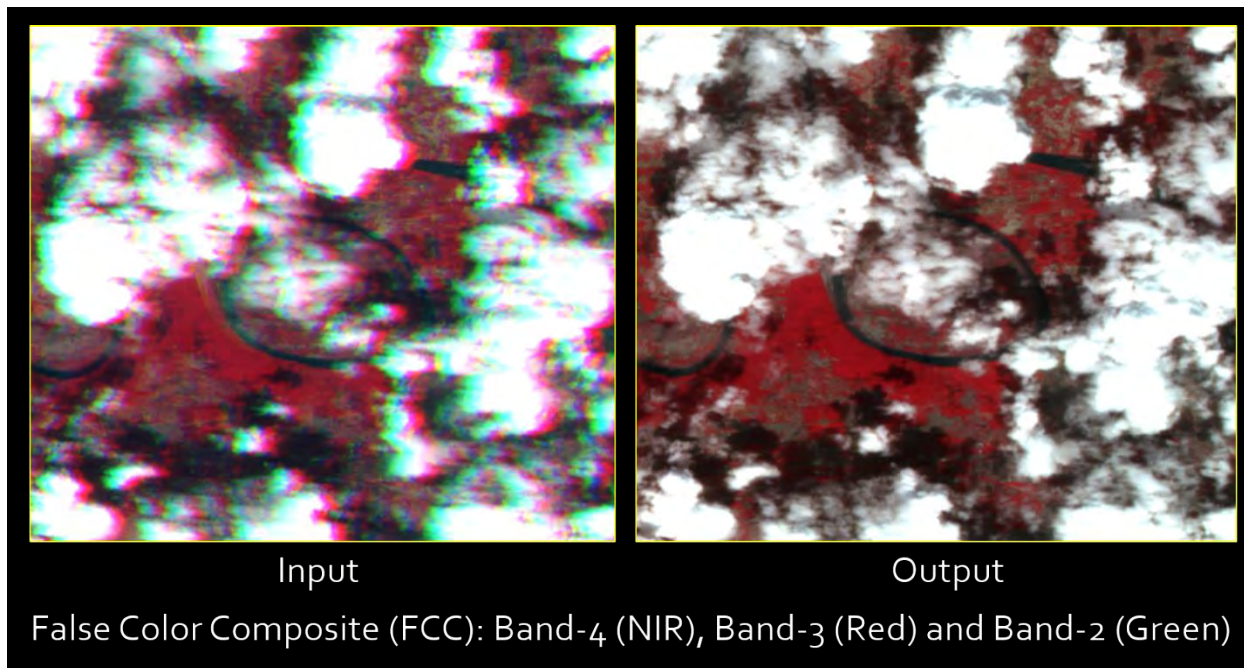


Figure-4.31. BBR Performance over Cloudy Region with a 4X Zoom.

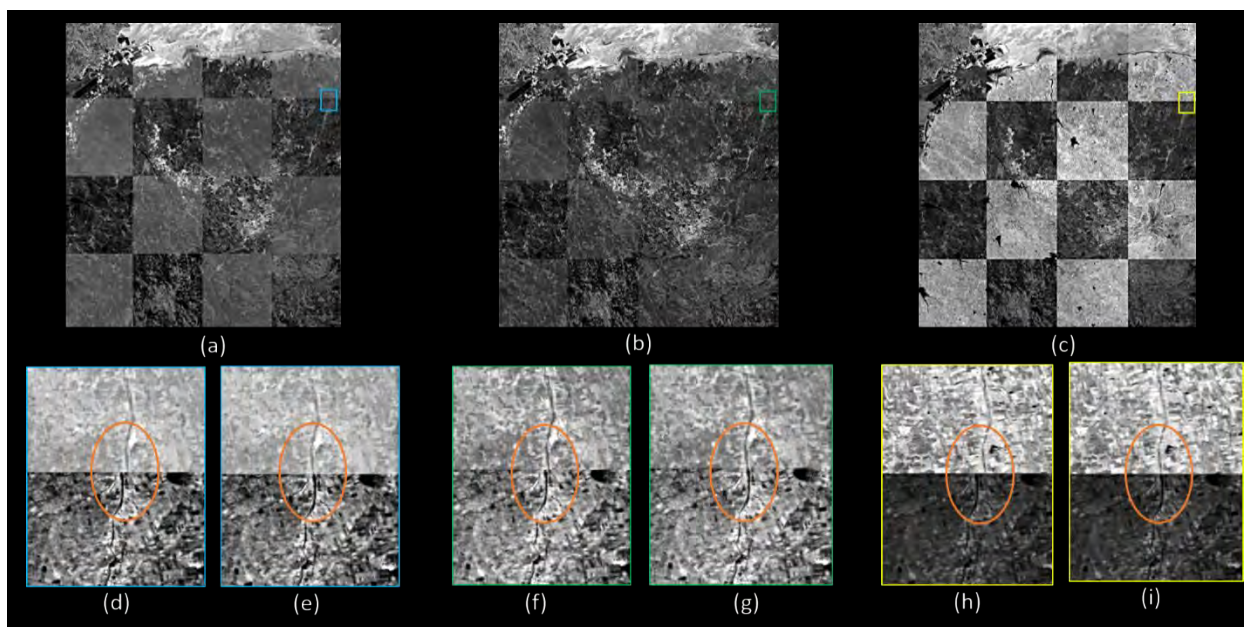


Figure-4.32. Checkerboard Pattern based Visual BBR Evaluation (a) Band-1 w.r.t. Band-3 (b) Band-2 w.r.t. Band-3 (c) Band-4 w.r.t. Band-3 (d) Band-1 w.r.t. Band-3 before BBR Correction (e) Band-1 w.r.t. Band-3 after BBR Correction (f) Band-2 w.r.t. Band-3 before BBR Correction (g) Band-2 w.r.t. Band-3 after BBR Correction (h) Band-4 w.r.t. Band-3 before BBR Correction (i) Band-4 w.r.t. Band-3 after BBR Correction

Additionally, checkerboard pattern images are generated between input multispectral bands (Band-1, Band-2, Band-4) concerning the reference red spectral band (Band-3) to further evaluate BBR performance (Figure-4.32). The checkerboard pattern images illustrate the discontinuity of a river stream before BBR correction, highlighted within an orange oval region. After BBR correction using the proposed method, the river stream is aligned between the spectral band pairs, demonstrating the effectiveness of the correction in achieving spatial alignment.

#### *B. Band Misregistration Effect on NDVI Geophysical Parameter*

An INS-2B Nano-MX multispectral image False Color Composite (FCC) is created around a water body, both with and without Band-to-Band Registration (BBR) correction. In the Nano-MX FCC before BBR correction, depicted in Figure-4.33 (a), dense vegetation is noticeable along the left-side perimeter of the water body, highlighted by a white oval shape in the image. Post BBR correction, in the Nano-MX FCC image shown in Figure-4.33 (b), the water body is observed without the dense vegetation. To validate the BBR correction outcome, the same region is extracted from a Landsat-9 satellite image captured around the same time frame, as illustrated in Figure-4.33 (c). The Landsat-9 OLI sensor captures remote sensing multispectral images at a spatial resolution of 30.0 m, similar to the spatial resolution of the INS-2B Nano-MX sensor, and serves as a reference for a comparative evaluation of the band misregistration effect. Table-4.18 provides details of the satellite images used for band misregistration evaluation.

It is observed that the Landsat-9 FCC exhibits a similar appearance to the Nano-MX FCC after BBR correction, with no dense vegetation around the water body, as highlighted by a white oval in the images. All three images (Nano-MX FCC before BBR, Nano-MX FCC after BBR, and Landsat-9 FCC) are utilized to calculate the Normalized Difference Vegetation Index (NDVI) geophysical parameters [48]. Figure-4.33 (d) displays the NDVI map corresponding to Nano-MX FCC before BBR, indicating a high vegetation index around the left-side perimeter of the water body. Conversely, the NDVI maps from Nano-MX FCC after BBR and Landsat-9 FCC exhibit a similar NDVI trend, indicating significantly less vegetation in the left region of the water body.

The range of NDVI measurements from Nano-MX FCC after BBR and Landsat-9 FCC differs slightly due to the distinct wavelength ranges of the two satellites' Near-Infrared (NIR) and Red spectral bands. Additionally, the imaging times of the two satellites differ by more than two hours (as shown in Table-4.18), leading to variations in the range of estimated NDVI values.

Table-4.18. Satellite Image Details used for analysis of the impact of Band Misregistration

Satellite	Sensor	Date and Time of Imaging	NIR & Red Spectral Resolution for NDVI (in $\mu\text{m}$ )	Spatial Resolution (in m)	Geographical Region
INS-2B	Nano-MX	20-Dec-2022, 07:23:33 UTC	NIR: 0.77-0.86 Red: 0.62-0.68	30.0	Gujarat, India
Landsat-9	OLI	22-Dec-2022, 05:33:12 UTC	NIR: 0.85-0.88 Red: 0.64-0.67	30.0	

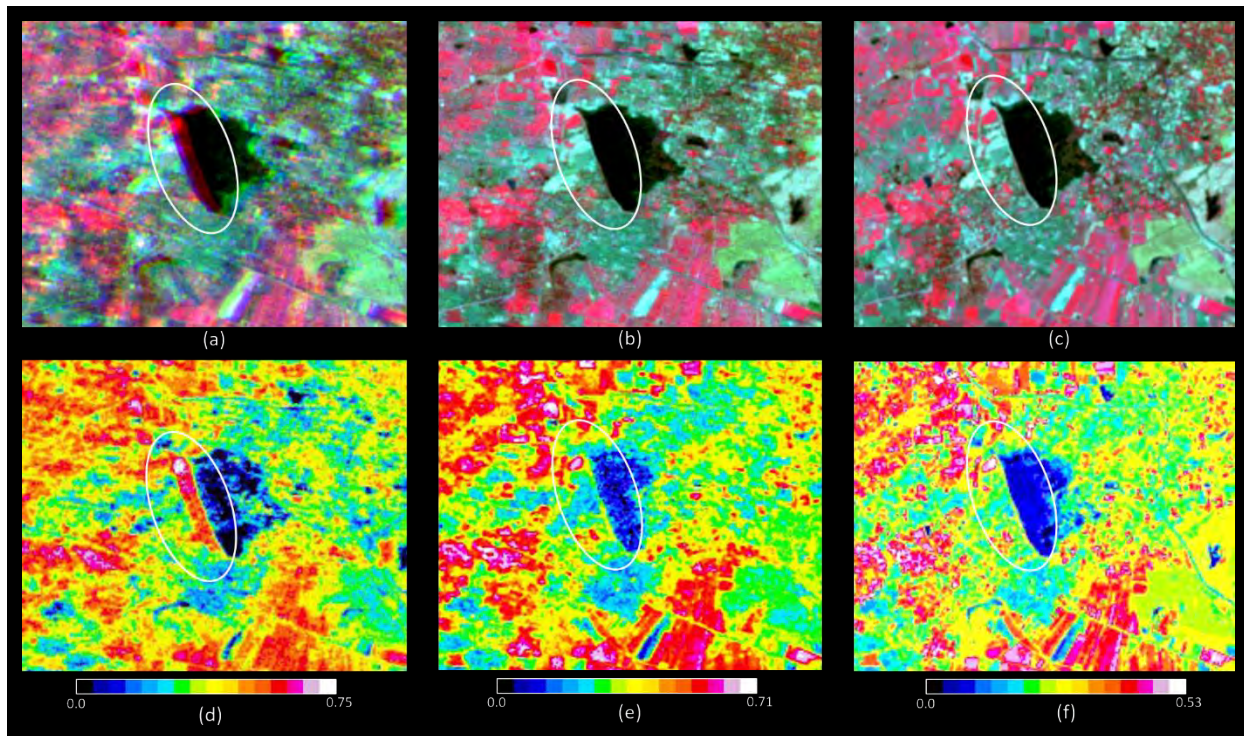


Figure-4.33. Band Misregistration Effect on NDVI (a) Nano-MX FCC before BBR (b) Nano-MX FCC after BBR (c) Landsat-9 FCC (d) NDVI Map from Nano-MX FCC before BBR (e) NDVI Map from Nano-MX FCC after BBR (f) NDVI Map from Landsat-9 FCC



### C. Quantitative Evaluation of BBR

Nano-MX multispectral images are segmented into three regions, R1 (left), R2 (middle), and R3 (right), for BBR estimation. The across track (horizontal) direction and along track (vertical) direction pixel shifts are estimated before and after BBR correction using the proposed methodology. Table-4.19 shows BBR estimation before correction at system level. It is observed that maximum pixel shift of around 10 pixels is observed between extreme bands before correction. Table-4.20 shows misregistration estimate after BBR correction. It is clearly found that the average band misalignment after correction is less than 0.15 pixel. The proposed BBR technique is able to achieve sub-pixel spatial alignment between the multispectral bands.

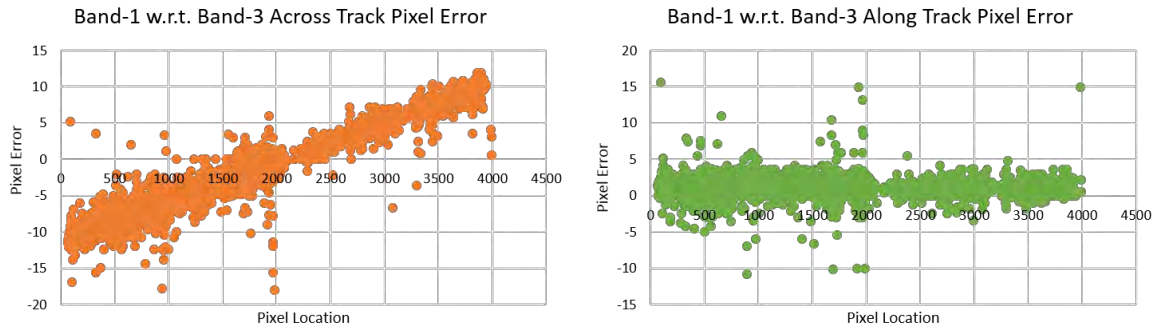
Table-4.19. Estimation of Band-to-Band Misregistration before Correction

<b>BBR Estimation for Band Combination</b>	<b>R1 (Across, Along) (in pixels)</b>	<b>R2 (Across, Along) (in pixels)</b>	<b>R3 (Across, Along) (in pixels)</b>
Band-1 wrt. Band-3	(-7.19,0.79)	(0.19, 0.89)	(7.48, 0.95)
Band-2 wrt. Band-3	(-4.48, 0.94)	(-0.15, 0.92)	(4.26, 0.82)
Band-4 wrt. Band-3	(-9.71, 0.37)	(0.01, 0.68)	(10.12, 0.82)

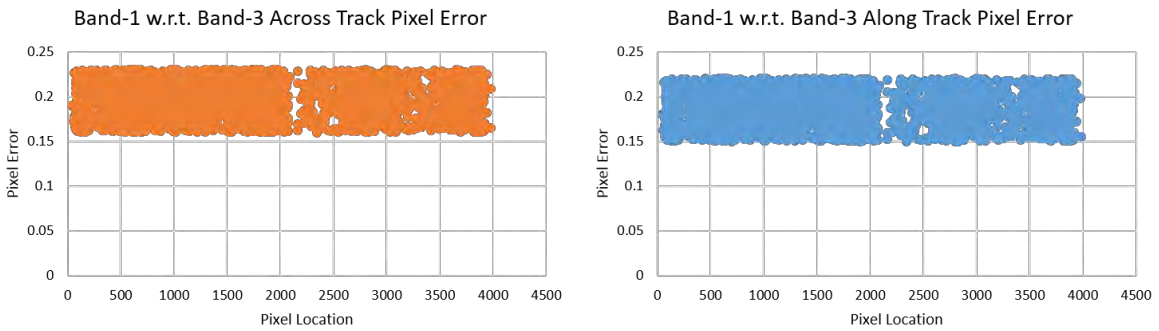
Table-4.20. Estimation of Band-to-Band Misregistration Estimation after Correction

<b>BBR Estimation for Band Combination</b>	<b>R1 (Across, Along) (in pixels)</b>	<b>R2 (Across, Along) (in pixels)</b>	<b>R3 (Across, Along) (in pixels)</b>
Band-1 wrt. Band-3	(0.07,0.01)	(-0.03, -0.02)	(0.04, -0.03)
Band-2 wrt. Band-3	(0.03,0.01)	(-0.04,0.02)	(0.02, 0.01)
Band-4 wrt. Band-3	(0.15, 0.07)	(0.12, 0.07)	(0.01, 0.01)

The Band-to-Band Misregistration before and after correction is visualized by plotting the pixel error in along-track and across-track pixel locations. The proposed methodology is executed before and after correction to extract pruned matched feature points between the band combinations such as Band-1 w.r.t. Band-3 (shown in Figure-4.34), Band-2 w.r.t. Band-3 (shown in Figure-4.35), and Band-4 w.r.t. Band-3 (shown in Figure-4.36). It is found that before BBR correction, the pixel error is at a higher margin in the across-track direction for all the band combinations. The developed algorithm corrects the misregistration across-track and along-track directions and achieves sub-pixel registration accuracy for all the band combinations.

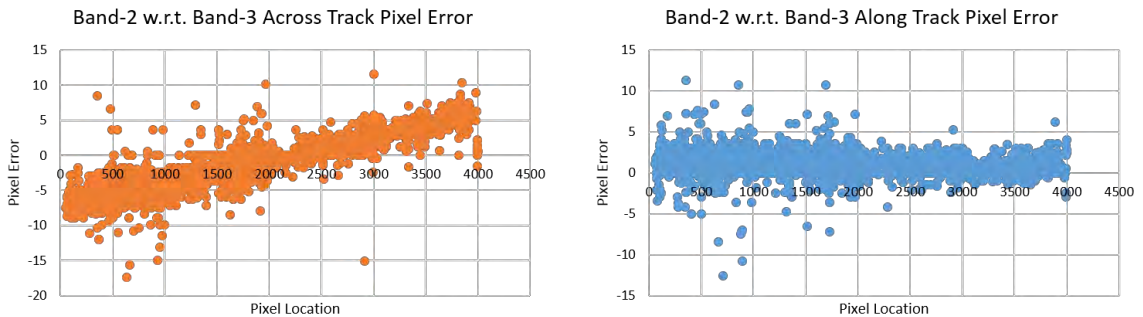


Before BBR Correction

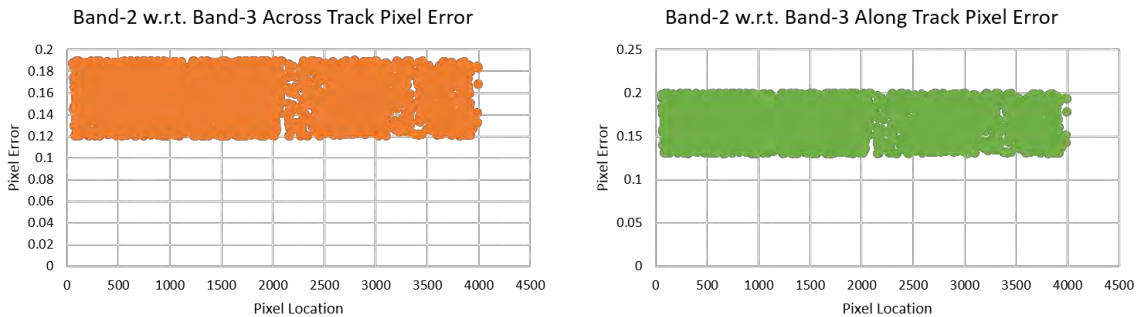


After BBR Correction

Figure-4.34. Band-1 wrt. Band-3 BBR Performance before and after correction



Before BBR Correction



After BBR Correction

Figure-4.35. Band-2 wrt. Band-3 BBR Performance before and after correction



Figure-4.36. Band-4 wrt. Band-2 BBR Performance before and after correction

#### D. Comparison with State-of-the-Art Techniques

To assess its efficacy, the proposed method is juxtaposed with contemporary remote sensing image registration techniques, and the Band-to-Band Registration (BBR) performance is scrutinized using image quality metrics. Control matched points in BBR image bands, produced by various BBR techniques, are manually identified for computation. Root Mean Square Error (RMSE) emerges as a suitable metric to quantify the spatial alignment among the multispectral bands. The RMSE, employed for BBR quantification of each band in relation to the reference, is defined as:

$$RMSE = \sqrt{\frac{1}{K} \sum_{i=1}^K \|X_i - \hat{X}_i\|^2} \quad (4.36)$$

In this context, where  $K$  is the number of matched control points for BBR evaluation,  $X_i$  denotes the coordinates  $(x_i, y_i)$  of the reference multispectral band, and  $\hat{X}_i$  represents the coordinates  $(\hat{x}_i, \hat{y}_i)$  of the output registered multispectral band.

To verify the accuracy of the matched control points, the Correct Matching Ratio (CMR) metric is also calculated (Zhang, 2020). CMR serves as a robust measure to assess the Band-to-Band Registration (BBR) output and is defined as follows:

$$CMR = \frac{K_c}{K_T} \quad (4.37)$$

In this context,  $K_T$  represents the number of matched correspondences after the removal of outliers, and  $K_c$  is the number of precisely correct matches obtained after eliminating the false ones.

Furthermore, the Number of Correct Matches (NCM), which represents the count of accurate correspondences in the image pairs, is also computed. Root Mean Square Error (RMSE), Correct Matching Ratio (CMR), and NCM are evaluated after applying Band-to-Band Registration (BBR) correction to Nano-MX multispectral images using state-of-the-art image registration techniques. The comparison includes PSO-SIFT (Ma et al., 2016), SURF with M-estimator sample consensus (SURF+M-SAC) (Wu et al., 2020), ORB with Motion Smoothness Constraint (ORB+MSC) (Misra et al., 2021), KAZE with MAGSAC Sample Consensus (KAZE+MAGSAC) (Alcantarilla et al., 2012, Barath et al., 2019), Nonlinear SIFT (Yu et al., 2021) with Triangular Area Ratio (NSIFT+TAR) (Guo et al., 2022), and Co-occurrence Filter Space Matching (CoFSM) (Yao et al., 2022) with MSC (Bian et al., 2017). Additionally, recent deep learning-based remote sensing image registration techniques, such as Deep Neural Network with Semantic Template Matching (DNN+ST) (Li et al., 2021) and multi-scale deep learning framework with unsupervised learning (MU-NET) (Ye et al., 2022), are considered for comparison.

Twenty Nano-MX multispectral images covering diverse landscapes are selected for the comparison and evaluation. The average RMSE and CMR are calculated for each band with respect to the reference band of Nano-MX. Table-4.21 presents the RMSE comparison of BBR Nano-MX multispectral images using eight prominent image registration techniques. It is evident that the RMSE for BBR estimation using the proposed method is superior to all contemporary image registration techniques, and the average RMSE for multiple Nano-MX multispectral images using the proposed BBR approach is 0.20 pixels. Table-4.22 shows the CMR computed for all output registered bands with respect to reference bands using identified image registration methods. The CMR values indicate that the matching ratio of the proposed method is better than state-of-the-art remote sensing image registration techniques. Table-4.23 displays the NCM computation, revealing that the number of correct correspondences between the bands with reference Band-3 is highest in the proposed method for all image pairs.

The computation time for BBR methods is also assessed for different spectral band combinations. The same system configuration is considered to measure the computation time (in seconds), with the Nano-MX image size being around 4096\*4096 pixels. Table-4.24 illustrates the BBR computation time using prominent remote sensing image registration methods. It is observed that NSIFT+TAR takes more compute time than other eminent BBR techniques, while ORB+MSC requires the least BBR compute time for Nano-MX images. The proposed method exhibits relatively less computation time than recent robust BBR methods.

Table-4.21. RMSE for the Comparison of BBR

<b>BBR Estimation</b>	<b>PSO-SIFT</b>	<b>SURF+MSAC</b>	<b>ORB+MSC</b>	<b>KAZE+MAGSAC</b>	<b>NSIFT+TAR</b>	<b>CoFSM+MSC</b>	<b>DNN+ST</b>	<b>MU-NET</b>	<b>Proposed Method</b>
Band-1 w.r.t. Band-3	0.48	0.41	0.67	0.56	0.34	0.46	0.41	0.35	<b>0.21</b>
Band-2 w.r.t. Band-3	0.32	0.27	0.38	0.41	0.22	0.35	0.31	0.40	<b>0.15</b>
Band-4 w.r.t. Band-3	0.44	0.51	0.35	0.52	0.25	0.48	0.42	0.43	<b>0.23</b>

Table-4.22. Computation of CMR for the Comparison of BBR

<b>CMR Estimation</b>	<b>PSO-SIFT</b>	<b>SURF+MSAC</b>	<b>ORB+MSC</b>	<b>KAZE+MAGSAC</b>	<b>NSIFT+TAR</b>	<b>CoFSM+MSC</b>	<b>DNN+ST</b>	<b>MU-NET</b>	<b>Proposed Method</b>
Band-1 w.r.t. Band-3	0.91	0.92	0.86	0.90	0.90	0.91	0.92	0.90	<b>0.94</b>
Band-2 w.r.t. Band-3	0.82	0.87	0.83	0.91	0.92	0.87	0.89	0.91	<b>0.93</b>
Band-4 w.r.t. Band-3	0.84	0.87	0.85	0.88	0.89	0.85	0.87	0.89	<b>0.91</b>

Table-4.23. NCM computation for BBR Comparison

<b>CMR Estimation</b>	<b>PSO-SIFT</b>	<b>SURF+ MSAC</b>	<b>ORB+ MSC</b>	<b>KAZE+ MAGSAC</b>	<b>NSIFT+ TAR</b>	<b>CoFSM+ MSC</b>	<b>DNN+ ST</b>	<b>MU-NET</b>	<b>Proposed Method</b>
Band-1 w.r.t. Band-3	191	256	278	183	262	290	262	298	<b>320</b>
Band-2 w.r.t. Band-3	237	337	367	202	297	320	359	384	<b>438</b>
Band-4 w.r.t. Band-3	169	198	208	148	189	217	229	259	<b>272</b>

Table-4.24. Computation Time (in seconds) for Comparison of BBR

<b>Compute Time</b>	<b>PSO-SIFT</b>	<b>SURF+ MSAC</b>	<b>ORB+ MSC</b>	<b>KAZE+ MAGSAC</b>	<b>NSIFT+ TAR</b>	<b>CoFSM+ MSC</b>	<b>DNN+ ST</b>	<b>MU-NET</b>	<b>Proposed Method</b>
Band-1 w.r.t. Band-3	25	21	12	28	26	21	20	22	<b>19</b>
Band-2 w.r.t. Band-3	32	28	11	35	33	26	24	26	<b>16</b>
Band-4 w.r.t. Band-3	22	19	9	26	25	18	15	17	<b>12</b>

### E. Ablation Study

The proposed Band-to-Band Registration (BBR) method incorporates several crucial components to achieve the desired registration accuracy among spectral bands. The detection of outliers, exemplified by SC-RANSAC, plays a critical role in refining key point correspondences and generating potential feature match points. Additionally, Segmented Affine Transformation (SAT) addresses local geometric distortion across various image segments, ensuring consistent registration of spectral bands. Ablation study evaluations are conducted considering different combinations such as CSS with Affine Transformation (CSS+AT), CSS+SAT, CSS+RANSAC+SAT, and our proposed technique CSS+SC-RANSAC+SAT.

Image registration quality parameters, including Root Mean Square Error (RMSE) and Correct Matching Ratio (CMR), are derived using Nano-MX registered image pairs, and corresponding averages are computed. Table-4.25 presents the RMSE and CMR comparison for various BBR component combinations. Notably, our proposed method, CSS+SC-RANSAC+SAT, yields the lowest RMSE values for all band image pairs, indicating superior spatial alignment among the spectral bands. Moreover, the CMR value for the spectral band pairs is highest for our proposed method compared to other BBR component combinations.

Table-4.25. RMSE among different combinations of BBR Components

<b>BBR Component Combinations</b>	<b>CSS+ AT (RMSE, CMR)</b>	<b>CSS+ SAT (RMSE, CMR)</b>	<b>CSS+ RANSAC+ AT (RMSE, CMR)</b>	<b>CSS+ SC-RANSAC+ SAT (Ours) (RMSE, CMR)</b>
Band-1 w.r.t. Band-3	(1.03, 0.82)	(0.84,0.84)	(0.56,0.88)	<b>(0.21,0.94)</b>
Band-2 w.r.t. Band-3	(0.78,0.88)	(0.66,0.86)	(0.51,0.89)	<b>(0.15, 0.93)</b>
Band-4 w.r.t. Band-3	(0.96,0.79)	(0.74,0.81)	(0.61,0.85)	<b>(0.23, 0.91)</b>

#### 4.3.4. Summary

BBR stands as a pivotal preprocessing step in the realm of remote sensing image analysis, facilitating precise data integration and analysis across diverse applications. The selection of a registration method hinges on the specific characteristics of the images and the desired level of accuracy. This article delves into the intricacies of a novel BBR methodology designed for sub-pixel spatial alignment. The proposed approach excels in detecting and characterizing unique feature points for matching using Co-occurrence Scale Space (CSS), outperforming established feature detection techniques. The Spatial Confined Random Sample Consensus (SC-RANSAC) outlier removal model meticulously prunes matched control points, ensuring the faithful estimation of the Segmented Affine Transformation (SAT) model. SAT adeptly addresses residual BBR, managing local distortions and aligning multispectral bands with precision.

Experiments span diverse feature targets, including water bodies, vegetation, urban regions, and snow/glacier feature classes, validating the consistency and robustness of the developed BBR technique. Visual assessments and quantified pixel shifts relative to the reference band reveal seamless cohesion among multispectral bands. It's important to note that the proposed method's performance on hyperspectral remote sensing imagery warrants evaluation, with potential parameter tuning to achieve the desired sub-pixel registration accuracy. The methodology achieved a Root Mean Square Error (RMSE) of less than 0.25 pixels and a Correct Matching Ratio (CMR) exceeding 0.91 across geographically distant Nano-MX multispectral images. Comparative analysis with state-of-the-art image registration techniques, employing image quality metrics, underscores the superiority of the proposed method.



## **4.4. Direct Feature Extraction and Co-Registration of Bayer Pattern Raw Planetary Images**

### **4.4.1. Introduction**

The section presents a novel method for extracting features directly from MCC Bayer pattern raw images. Upon receiving MCC raw data on the ground, it undergoes a sequence of radiometric and geometric processing procedures to produce an MCC color image. Demosaicking and denoising of MCC images represent a time-consuming stage in this procedure. The image processing pipeline that has been developed is complex and requires significant computational resources (Misra et al., 2019b).

In attempting to derive gradient features directly from MCC Bayer Pattern raw images, we utilized edge detectors grounded in the Color Difference Constancy (CDC) assumption (Zhou et al., 2021). The primary tools employed for directly extracting gradients from MCC Bayer Pattern raw images are the Central Difference Edge Operator and the Sobel Edge Detector (Kanopoulos et al., 1988). We devised a method for co-registering images from MCC raw data by employing the Mode-Mean Combo Patch Filler and Gradient Intensity-induced Scale Invariant Feature Transform (GI-SIFT), as proposed by Lowe in 1999, for matching features. To enhance this, we utilized Feature Similarity Score-guided Random Sample Consensus (FSS-RANSAC) algorithm to eliminate the outliers (Misra et al., 2024a).

The experimental assessment shows that the circumference outlines of crater boundaries, Martian mons, and valleys on the Mars surface are nearly identical when derived from the Gradient Magnitude Map (GMM) extracted from MCC Bayer Pattern raw images and MCC demosaic images. We subjected the results obtained from the Gradient Magnitude Map (GMM) to a quantitative evaluation against the Mars Digital Image Model (MDIM) 2.1 reference (Archinal et al., 2004) to demonstrate the effectiveness and reliability of our proposed method. Image co-registration experiments were performed across different Martian regions, demonstrating that our proposed method for direct image co-registration achieves a relative geometric accuracy at a sub-pixel level. Upon comparison, it is evident that co-registration estimates obtained directly from MCC Bayer Pattern raw images closely match those derived from MCC demosaic images.

## 4.4.2. Methodology Developed

### A. Extracting Features from MCC Bayer Pattern Raw Images

The primary objective of directly employing the MCC Bayer Pattern raw image is to fetch information for the gradient intensity map. The computation of gradients is determined by the derivatives of image pixels with respect to the directions in the horizontal and vertical planes. The chosen gradient operator for the MCC Bayer Pattern raw image needs to meet the following conditions:

- 1) Showing a gradient intensity value close to zero in regions with uniform characteristics.
- 2) Offering an elevated estimate of differences on inclined surfaces.
- 3) Generating a non-zero measurement at the start or conclusion of a slope.

To meet these requirements, it is crucial for an edge detector to have alternating zero and non-zero coefficients, ensuring compliance with the Color Difference Constancy (CDC) assumption. The demosaicking procedure generally depend on Color Difference Constancy (CDC) to produce a complete color image, as it provides superior performance in terms of Peak Signal-to-Noise Ratio (PSNR). In our method, we extract information from the MCC Bayer image using gradient operators that adhere to the Color Difference Constancy (CDC) assumption. In particular, we have opted for the Central Difference Operator and Sobel Edge Filter, illustrated in Figure-4.37, to extract notable gradient features from the MCC Bayer Pattern raw image.

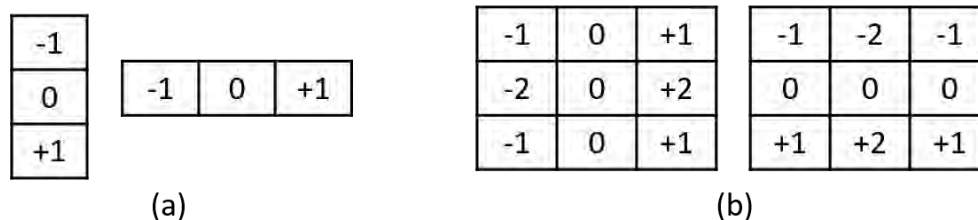


Figure-4.37. Edge Operators: (a) Central Difference Operator (b) Sobel Operator

The main concept involves directly applying a gradient filter to the MCC raw Bayer Pattern image. Both edge operators depicted in Figure-4.37 conform to the Color Difference Constancy (CDC), which is a vital factor for producing the gradient intensity map from raw MCC data. The definition of the central difference gradient operator is as follows:

$$D_x(x, y) = I(x + 1, y) - I(x - 1, y) \quad (4.38)$$

$$D_y(x, y) = I(x, y + 1) - I(x, y - 1) \quad (4.39)$$

Here,  $D_x$  and  $D_y$  denote the image gradients in the horizontal and vertical directions respectively, and  $I(x, y)$  denotes the intensity value at the pixel coordinates  $(x, y)$  of the MCC Bayer pattern image.

Rather than carrying out comprehensive planetary image data processing steps along with demosaicking processes on MCC images, we directly apply the filter template shown in Figure-4.37 (a) to the MCC Bayer Pattern image. This method, meeting the conditions mentioned earlier, enables the direct extraction of gradient information. Figure-4.38 displays both the traditional and the proposed pipelines for feature extraction. The traditional image processing pipeline for planetary remote sensing is computationally demanding, requiring a sequence of image correction steps. The first stage involves Photo Response Non-Uniformity (PRNU) correction, where the detector response is normalized using coefficients determined from the radiometric calibration carried out during pre-launch at the lab setup (Kinch et al., 2020; Qiao et al., 2021). Subsequently, the PRNU-corrected Bayer Pattern raw image from MCC undergoes a demosaicking process to reconstruct and generate a natural color image of the Martian surface (Kiku et al., 2016). The demosaicked image further go through the process of denoising with restoration to generate enhanced planetary image. In addition, the color balancing process (Mathew et al., 2017) also take place to show a closer representation of the Martian landscape.

In contrast, our research is motivated by the goal of bypassing the resource-intensive planetary image processing pipeline for the extraction of features. Our proposed approach involves the direct extraction of features from raw Bayer pattern planetary images, making a meaningful contribution to diverse planetary science exploration. As illustrated in Figure-4.38(b), our method simplifies the processing sequence by directly extracting potential features from MCC Bayer Pattern raw images. The methodology we have devised depends on a straightforward yet sturdy framework for extracting features and co-registering raw Bayer pattern planetary remote sensing images. For a comprehensive understanding of the process, Figure-4.39 presents an illustrative diagram of a typical MCC RGGB Bayer Pattern scheme. It highlights the unambiguous extraction of features from MCC raw data using a gradient operator compliant with the Color Difference Constancy (CDC) at MCC pixel locations.

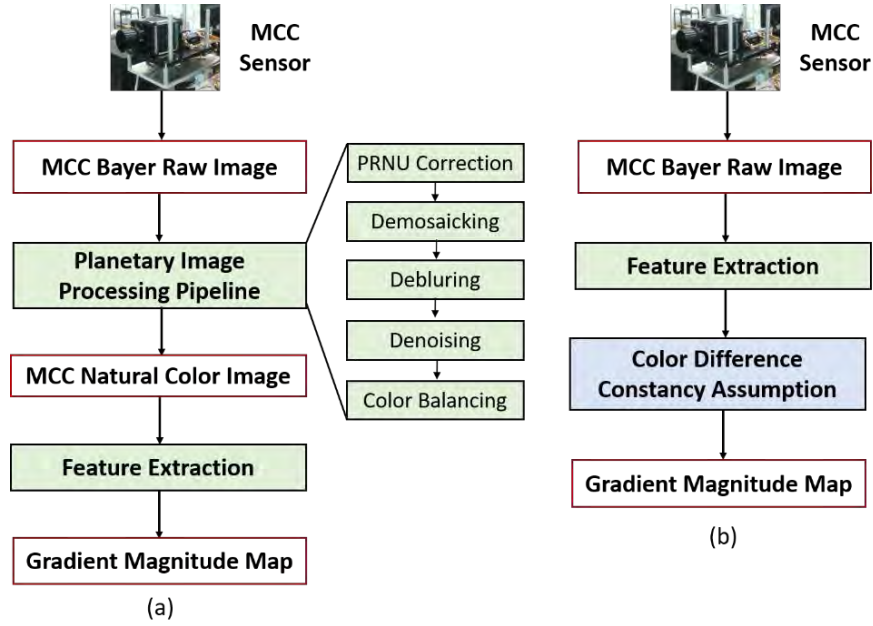


Figure-4.38. Feature Extraction Pipelines (a) Traditional Feature Extraction Pipeline (b) Novel Pipeline for the Direct Extraction of Features from MCC Bayer Pattern Raw Image

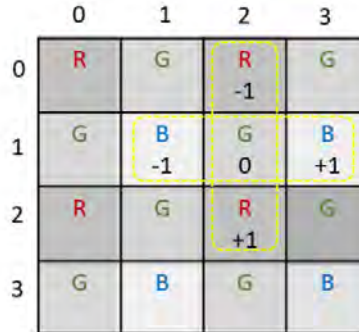


Figure-4.39. Extraction of Gradient Magnitude from MCC Raw Bayer Pattern Image

The two MCC pixels used as inputs for coefficients 1 and -1 in the convolution templates belong to the identical spectral band. This ensures that differences are consistently computed on pixels within the same spectral channel. The formula for calculating the gradient magnitude at the position (1, 2) in Figure-4.39 is as follows:

$$D_{x(1,2)}^B = I_{(1,3)}^B - I_{(1,1)}^B \quad (4.40)$$

$$D_{y(1,2)}^R = I_{(2,2)}^R - I_{(0,2)}^R \quad (4.41)$$

Here,  $D^B$  and  $D^R$  represent the gradients of the blue and red spectral channels, respectively, while  $I_{(x,y)}^B$  and  $I_{(x,y)}^R$  are the intensity measurements of the blue and red spectral channels at the pixel coordinates  $(x, y)$ .

It is noted that if pixel locations  $(x_1, y_1)$  and  $(x_2, y_2)$  are situated in close proximity, then

$$I^G(x_1, y_1) - I^G(x_2, y_2) \approx I^R(x_1, y_1) - I^R(x_2, y_2) \quad (4.42)$$

$$I^G(x_1, y_1) - I^G(x_2, y_2) \approx I^B(x_1, y_1) - I^B(x_2, y_2) \quad (4.43)$$

Utilizing equations (4.40) and (4.41) in conjunction with equations (4.42) and (4.43), we can infer that

$$D \approx D^R \approx D^G \approx D^B \quad (4.44)$$

This implies that the gradients of planetary multispectral imagery can be calculated using any of the three spectral wavelength channels if the Color Difference Constancy (CDC) assumption is maintained. Hence, in cases where two color pixel samples are not present at image location, the gradient for that location in both the horizontal and vertical directions can be directly calculated from the MCC Bayer Pattern raw image using the nearest color spectral values, as detailed below.

$$D_{x(1,2)} = D_{x(1,2)}^B = I_{(1,3)}^B - I_{(1,1)}^B \quad (4.45)$$

$$D_{y(1,2)} = D_{y(1,2)}^R = I_{(2,2)}^R - I_{(0,2)}^R \quad (4.46)$$

The aforementioned idea can be extended to include other gradient detectors that are defined by alternating zero and non-zero coefficients. These detectors can then be utilized to directly extract significant feature targets from Bayer Pattern images. The Sobel edge detector, illustrated in Figure-4.37 (b), is applied to MCC Bayer Pattern raw data, and the calculation for the gradient estimate at a pixel location is as follows:

$$\begin{aligned} D_{x(1,2)}^S &= I_{(0,3)}^G + 2 * I_{(1,3)}^B + I_{(2,3)}^G - I_{(0,1)}^G - 2 * I_{(1,1)}^B - I_{(2,1)}^G \\ &= (I_{(0,3)}^G - I_{(0,1)}^G) + 2 * (I_{(1,3)}^B - I_{(1,1)}^B) + (I_{(2,3)}^G - I_{(2,1)}^G), \end{aligned} \quad (4.47)$$

$$\begin{aligned} D_{y(1,2)}^S &= I_{(2,1)}^G + 2 * I_{(2,2)}^R + I_{(2,3)}^G - I_{(0,1)}^G - 2 * I_{(0,2)}^R - I_{(0,3)}^G \\ &= (I_{(2,1)}^G - I_{(0,1)}^G) + 2 * (I_{(2,2)}^R - I_{(0,2)}^R) + (I_{(2,3)}^G - I_{(0,3)}^G). \end{aligned} \quad (4.48)$$

where  $D_{x(1,2)}^S$  and  $D_{y(1,2)}^S$  are values of the gradient using the Sobel edge detector at the pixel location (1, 2) in both horizontal and vertical directions,  $I_{(x,y)}^{spectra}$  is the intensity value of the spectrum (either Red (R), Green (G), or Blue (B) sample) at the pixel coordinates  $(x, y)$ .

The Sobel Edge Operator upholds the CDC assumption, enabling the direct extraction of features from MCC raw Bayer Pattern images. According to the CDC assumption, equations (4.47) and (4.48) can be reformulated as follows:

$$\begin{aligned}
 D_{x(1,2)}^S &\approx (I_{(0,3)}^{\hat{R}} - I_{(0,1)}^{\hat{R}}) + 2 * (I_{(1,3)}^{\hat{R}} - I_{(1,1)}^{\hat{R}}) + (I_{(2,3)}^{\hat{R}} - I_{(2,1)}^{\hat{R}}) \\
 &\approx (I_{(0,3)}^{\hat{G}} - I_{(0,1)}^{\hat{G}}) + 2 * (I_{(1,3)}^{\hat{G}} - I_{(1,1)}^{\hat{G}}) + (I_{(2,3)}^{\hat{G}} - I_{(2,1)}^{\hat{G}}) \\
 &\approx (I_{(0,3)}^{\hat{B}} - I_{(0,1)}^{\hat{B}}) + 2 * (I_{(1,3)}^{\hat{B}} - I_{(1,1)}^{\hat{B}}) + (I_{(2,3)}^{\hat{B}} - I_{(2,1)}^{\hat{B}}), \quad (4.49)
 \end{aligned}$$

$$\begin{aligned}
 D_{y(1,2)}^S &\approx (I_{(2,1)}^{\hat{R}} - I_{(0,1)}^{\hat{R}}) + 2 * (I_{(2,2)}^{\hat{R}} - I_{(0,2)}^{\hat{R}}) + (I_{(2,3)}^{\hat{R}} - I_{(0,3)}^{\hat{R}}) \\
 &\approx (I_{(2,1)}^{\hat{G}} - I_{(0,1)}^{\hat{G}}) + 2 * (I_{(2,2)}^{\hat{G}} - I_{(0,2)}^{\hat{G}}) + (I_{(2,3)}^{\hat{G}} - I_{(0,3)}^{\hat{G}}) \\
 &\approx (I_{(2,1)}^{\hat{B}} - I_{(0,1)}^{\hat{B}}) + 2 * (I_{(2,2)}^{\hat{B}} - I_{(0,2)}^{\hat{B}}) + (I_{(2,3)}^{\hat{B}} - I_{(0,3)}^{\hat{B}}). \quad (4.50)
 \end{aligned}$$

Here,  $\hat{R}$ ,  $\hat{G}$ , and  $\hat{B}$  represent the absent color components at the specific MCC pixel coordinates.

The pixel indices (1, 2) represent the MCC pixel location used for computing the gradient map. The collection of MCC pixel locations (1, 2) used for feature extraction computation is depicted with rounded corner rectangles having yellow-colored dashed boundaries, encompassing two regions in Figure-4.39 utilized for gradient magnitude extraction. As a result, it is clear that the Sobel edge detector can be utilized to directly extract features from MCC Bayer Pattern raw images.

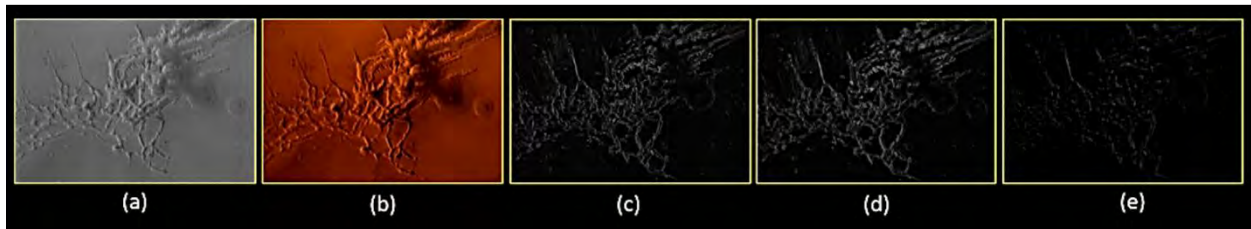


Figure-4.40. (a) Image in MCC Bayer Pattern (b) Image Obtained from MCC Demosaicking (c) Gradient Magnitude Map Derived from Bayer Pattern (d) Gradient Magnitude Map Obtained from Demosaicking (e) Image of Gradient Differences

It is observed that planetary remote sensing images capturing the Martian surface typically display limited texture information, posing a challenge for extracting crucial features. In flat regions, the Sobel Edge Operator provides a minimal difference estimate between two identical

spectral pixels. On the other hand, regions with high texture, such as crater boundaries and terrain slopes, demonstrate substantial differences between two identical spectral pixels. This leads to the extraction of sharp edge information in the gradient derived from the MCC Bayer Pattern image. Figure-4.40 illustrates the Gradient Magnitude Map (GMM) obtained from both the MCC Bayer Pattern raw image and the MCC demosaic image through the application of the Sobel Edge Operator. The features extracted from the MCC Bayer data correspond to the gradient features extracted from the MCC demosaic image. In certain cases, the Gradient Magnitude Map (GMM) from the Bayer Pattern might appear slightly subdued visually, given the presence of only a single spectral sample value at each pixel location. However, the comprehensive Gradient Magnitude Map (GMM) from MCC Bayer Pattern still offers the essential abstraction of image information, which is crucial for tasks such as object recognition and direct planetary image co-registration.

The gradient difference image shown in Figure-4.40 (e) demonstrates minimal deviation between the Gradient Magnitude Map (GMM) from MCC Bayer Pattern and the GMM from MCC demosaic. The highlighted area belongs to Valles Marineris, the largest canyon system in our solar system. The GMM from the MCC Bayer Pattern raw image accurately outlines the perimeter of this areo-morphic structure, providing insights into the prominent ramps or slopes in that areo-morphological region. Although the Bayer Pattern Color Filter Array is mainly utilized in planetary remote sensing sensors, the developed methodology is primarily assessed using such types of remote sensing images. Nevertheless, the methodology is broadly applicable to color filter array patterns like RYYB and CYYM which adhere to CDC. These patterns are also employed in payload camera systems developed to sense remotely planetary surfaces. Therefore, the method devised for direct feature extraction and co-registration is scalable.

### *B. Planetary Image Feature Matching using GI-SIFT*

Co-registering images with the SIFT starts by building a scale space through the use of a Gaussian Kernel (Goncalves et al., 2011; Zhong, Liu & Chen, 2015). Yet, directly employing Gaussian filtering on the MCC Bayer Pattern raw image might disturb the mosaic structure of the CFA, potentially resulting in the loss of vital spectral information crucial for feature matching. To tackle this problem, we suggest a Mode-Mean-based Patch Filler that directly recovers missing color samples from the CFA without undergoing intricate interpolation or demosaicking procedures.

The term "Mode" refers to the color sample with the highest frequency of occurrences within a patch. If the highest occurrence of a mode value in a patch surpasses a predefined sample score limit, the missing color samples are filled with the estimated mode value. If the sample score is under a predefined limit, we calculate the mean of the patch using the available color spectra, and the missing values are subsequently filled with that calculated mean. In our methodology, we choose a 5x5 patch to calculate missing components at every MCC Bayer Pattern pixel location, and based on empirical observations, this particular patch window is determined to contain adequate single spectral samples for recovering missing colors.

Next, red, green, and blue pixel samples are separated, and the mode-mean for the selected patch is calculated. Based on empirical findings, it is established that a sample score limit ( $l$ ) of three yields optimal results for the patch filler. The Bayer intensity image and the Gradient Magnitude Map (GMM) are created using the filled samples at each MCC image pixel, serving as a direct basis for matching the features and spatial alignment of planetary Bayer Pattern images. The computation of missing spectra value  $S_m$  in a patch  $p$  is defined as:

$$(S_m, p) = \{\mu, l \leq 3 \\ \Omega, l > 3\} \quad (4.51)$$

where  $\mu$  represents mean of the patch,  $\Omega$  represents the mode of the patch.

To tackle variations in radiometric intensity within MCC Bayer Pattern raw images, we introduce Gradient Intensity-based SIFT (GI-SIFT) for feature matching. GI-SIFT establishes a connection, as detailed in Equation-4.52 and Equation-4.53, between the smoothed image  $L$  and the input Bayer intensity raw image  $I$  at a pixel location  $(x, y)$  by applying a Gaussian kernel across various scales. This method proficiently handles image noise and demonstrates resilient performance even in imaging conditions with a low Signal-to-Noise Ratio (SNR).

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (4.52)$$

$$\text{where } G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma^2} \quad (4.53)$$

The intensity image undergoes Gaussian filtering at the scale space level to create a Gaussian pyramid, and the gradient magnitude is directly extracted from the MCC intensity image using the Sobel Edge operator. The Gradient Magnitude of the MCC Intensity image  $D_\sigma^1$  is defined as:

$$D_\sigma^1 = \sqrt{(D_{\sigma,x}^1)^2 + (D_{\sigma,y}^1)^2} \quad (4.54)$$



Here,  $\sigma$  represents the Gaussian scale space, and  $D_{\sigma,x}^1$  and  $D_{\sigma,y}^1$  are the gradient of the MCC intensity image in both the horizontal and vertical directions., respectively, with scale  $\sigma$ .

Next, the gradient orientation  $O_{\sigma}^2$  and gradient magnitude  $D_{\sigma}^2$  expressed as:

$$O_{\sigma}^2 = \arctan \left( \frac{D_{\sigma,y}^2}{D_{\sigma,x}^2} \right) \quad (4.55)$$

$$D_{\sigma}^2 = \sqrt{(D_{\sigma,x}^2)^2 + (D_{\sigma,y}^2)^2} \quad (4.56)$$

where  $D_{\sigma,x}^2$  and  $D_{\sigma,y}^2$  are the gradient magnitude derivatives of the MCC intensity image.

Here,  $D_{\sigma,x}^2$  and  $D_{\sigma,y}^2$  represent the derivatives of the gradient magnitude of the MCC intensity image.

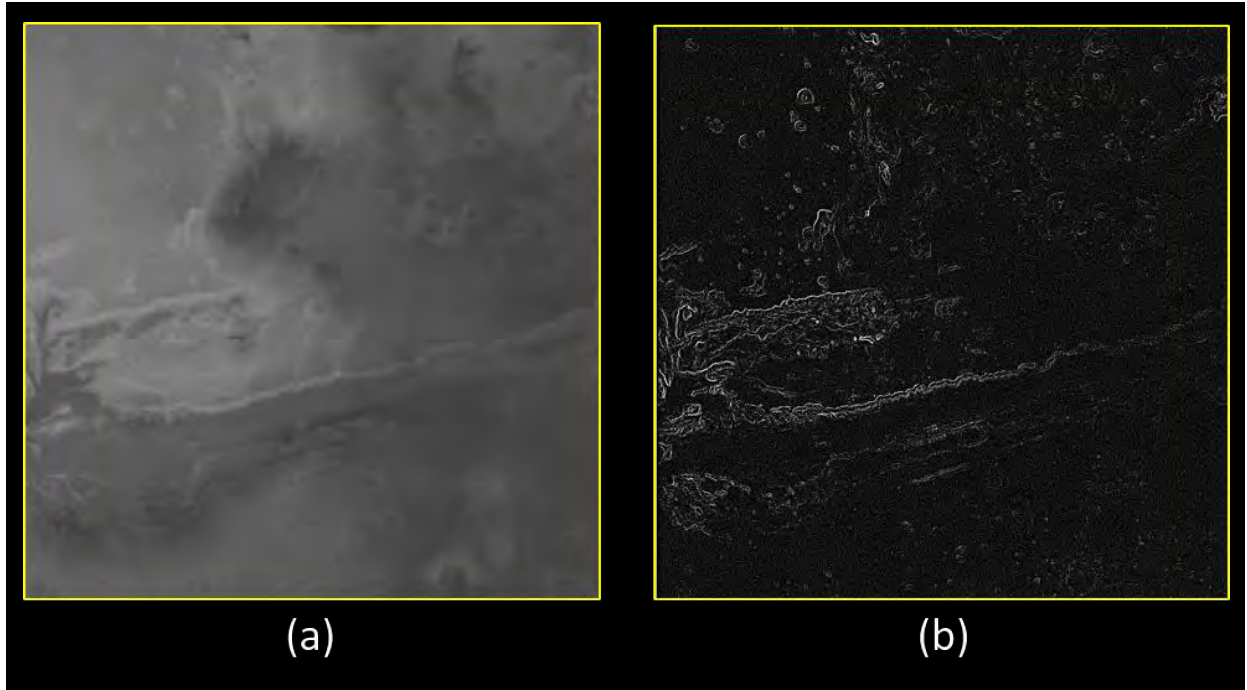


Figure-4.41. (a) Bayer pattern MCC Intensity Raw Image (b) Gradient Magnitude Map (GMM) obtained directly from the MCC Bayer Intensity Raw Image.

The gradients are computed through the Sobel edge operator and are utilized as inputs for assignment of orientation and extraction of descriptor in the process of matching features between the MCC Bayer Pattern raw image and the Mars image reference. Figure-4.41 illustrates the outputs directly produced from the MCC Bayer Pattern raw image, acting as inputs for feature matching based on GI-SIFT. The MCC Bayer Pattern raw image is utilized for reconstructing the intensity image using the suggested patch filler technique, leading to a more precise depiction of

MCC Bayer Pattern raw data. Subsequently, Gradient Magnitude Map is created from the MCC Bayer intensity raw image.

### *C. FSS-RANSAC Outlier Removal Method*

While GI-SIFT feature matching offers advantages, it faces difficulties in dealing with numerous falsely matched keypoints in planetary images, leading to errors in the ultimate estimation of transformation parameters. The Random Sample Consensus (RANSAC) technique proves to be a robust approach for eliminating these outliers. RANSAC produces potential solutions by utilizing the minimum required number of observations (control points) to estimate the underlying model parameters (Misra et al., 2021a; Misra et al., 2022a). The RANSAC procedure consists of two stages: the "Hypothesis Generation" phase, in which it randomly chooses a set of matched keypoints, and the "Hypothesis Evaluation" phase, during which it examines whether the selected hypothesis meets the specified inlier candidate criteria, with the parameter model error falling within a predefined threshold. The RANSAC algorithm proceeds through the following steps:

1. Select a set of points at random, ensuring it meets the minimum required number to establish the model parameters.
2. Determine the model parameters by solving for them using the selected points.
3. Determine the count of points within the entire set that fall within a predefined tolerance.
4. If the ratio of inliers to the total points in the set exceeds a predefined threshold, recalibrate the model parameters using all the identified inliers and conclude the process.
5. Alternatively, repeat steps 1 to 4 for a maximum of  $N$  iterations if the condition is not met.

The iteration count, represented by  $N$ , is set sufficiently high such that at least one of the randomly chosen samples sets does not contain an outlier. RANSAC processes sample data portion, not the entire dataset. If the data portion selected comprises only inliers, it can offer a hypothesis that is close to the actual truth. This assumption implies that the specified number of iterations,  $N$ , should be sufficient to include all samples of inliers at least once, as explained below.

$$N = \frac{\log \alpha}{\log(1-\gamma^m)} \quad (4.57)$$

In this context,  $\alpha$  denotes the failure probability,  $m$  signifies the quantity of matched keypoints required to formulate a hypothesis, and  $\gamma$  represents the likelihood of choosing an inlier. Specifically,  $\gamma$  is defined as the proportion of inliers to the complete set of sample data, often referred to as the inliers ratio.

Yet, in numerous practical situations, the inliers ratio  $\gamma$  is frequently unspecified and requires users to establish it. The difficulty now lies in choosing the most appropriate pair from an extensive array of pairs. Although certain methods in the literature aim to enhance the efficiency of RANSAC (Wu et al., 2014; Zheng et al., 2021), the improved RANSAC algorithms do not consistently detect precise matched pairs devoid of outliers in planetary remote sensing images.

In order to improve RANSAC performance, we introduce a new approach called FSS-RANSAC, which involves computing the Feature Similarity Score (FSS) for the matched keypoints and preemptively eliminating significant outliers. The Coefficient of Variation ( $CoV$ ) is calculated by defining a local window around the keypoint in both the input MCC Bayer intensity image and the reference image, as detailed below:

$$CoV = \{(\beta/\mu) * 100, w\} \quad (4.58)$$

Here,  $\beta$  represents the standard deviation,  $\mu$  is the mean, and  $w$  stands for the window size.

$$FSS_p = CoV_{ip} - CoV_{rp} \quad (4.59)$$

Here,  $FSS_p$  denotes the FSS for the matched pixel  $p$ ,  $CoV_{ip}$  represents the  $CoV$  at the pixel location  $ip$  in the input MCC image, and  $CoV_{rp}$  signifies the  $CoV$  at the pixel location  $rp$  in the reference image.

#### *D. Workflow for Direct Image Co-Registration*

The devised approach is depicted as a data processing pipeline, wherein the MCC Bayer Pattern raw image serves as input. The pipeline produces the MCC intensity raw image by employing the Mode-Mean Combo Patch Filler, and sharp edges are identified through the Sobel Edge detector. The suggested procedure employs the Sobel operator to directly extract features from raw Bayer pattern planetary images. This novel method for extracting crucial characteristics from raw Bayer pattern planetary data has not been explored within the planetary science community. It can be employed in advanced image processing tasks, including multi-temporal image co-registration.

Furthermore, the direct extraction of planetary features can make valuable contributions to diverse studies in planetary space exploration, such as crater counting and analysis of areomorphological features.

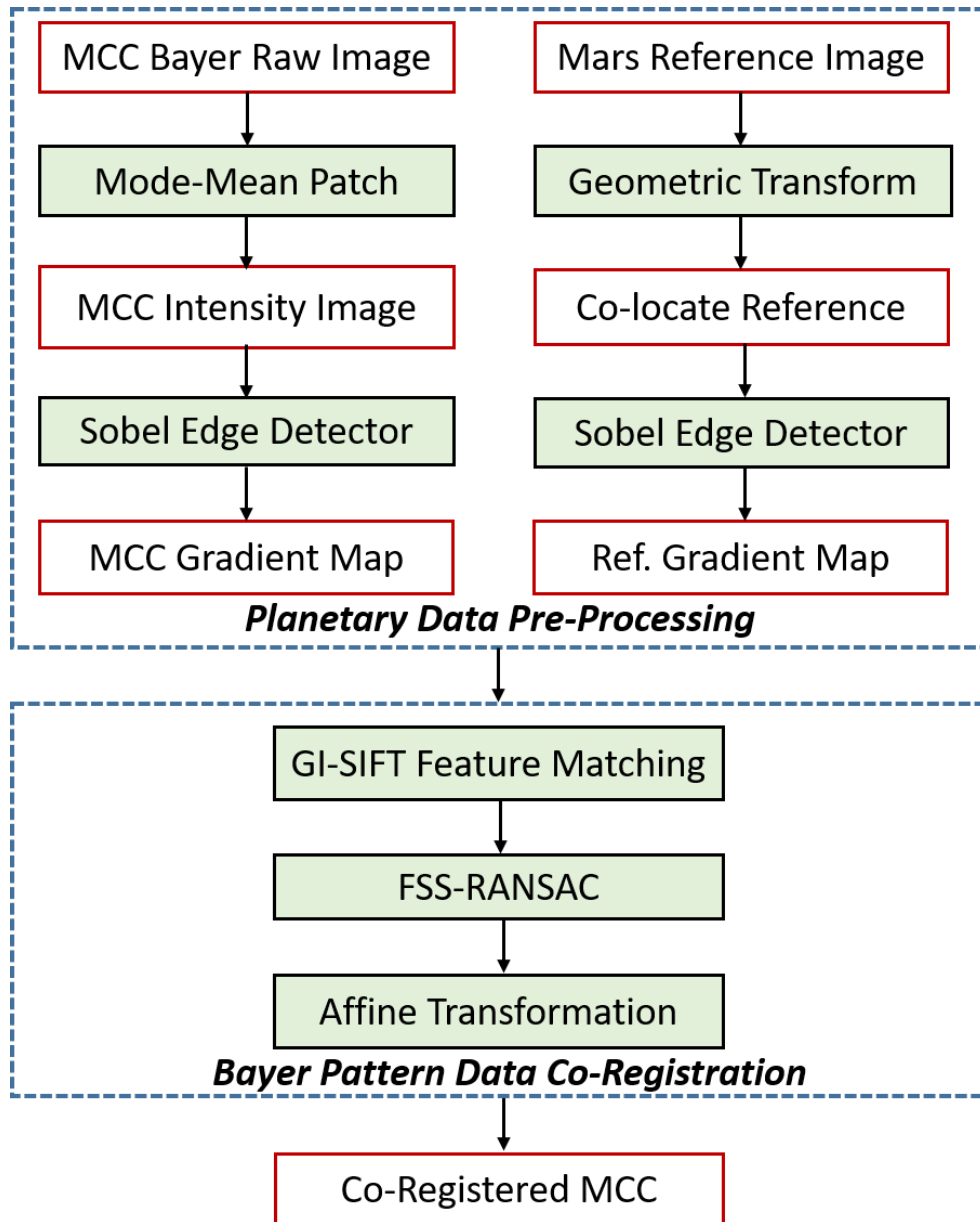


Figure-4.42. Workflow for Direct Co-Registration of Bayer Pattern Raw Images

At the same time, the MDIM 2.1 Mars reference image is integrated into the workflow for image co-registration. The reference image undergoes geometric transformation, followed by cropping and resampling, to align with the spatial resolution of MCC. Following this, the Gradient

Magnitude Map (GMM) for the reference is produced. Stable matched keypoints are extracted using GI-SIFT-based feature matching. Empirically, it is established that a window size of 5 yields an optimal estimate for the Coefficient of Variation ( $CoV$ ). A threshold of 0.8 for the Feature Similarity Score ( $FSS$ ) is defined to extract potential correspondences between keypoints. The suggested FSS-RANSAC technique efficiently eliminates outliers. The refined matched points are employed for estimating the affine transformation parameters (Misra et al., 2021b), and the MCC Bayer Pattern raw image is resampled using a cubic convolution kernel to ensure spatial alignment with the Mars planetary reference. Figure-4.42 depicts the devised workflow for direct image co-registration of MCC Bayer pattern raw images. The assumption of CDC is valid for the majority of CFA pattern images utilized in the design of remote sensing cameras. Hence, it is anticipated that the method devised for feature extraction and co-registration will operate with precision and dependability.

#### **4.4.3. Experimental Results**

The evaluation of the direct feature extraction approach covers diverse regions of the Martian surface captured by the Indian MCC instrument on the MOM spacecraft. The MCC, with its ability to capture surface features of Mars at various resolutions, is proficient in observing dynamic phenomena like dust storms and cloud patterns. It has the capacity to capture an image of the entire Mars disc at around 4.0 km spatial resolution and offer a more detailed view of specific areo-morphological structures at spatial resolutions better than 15.0 meters. The distinctive imaging capability of MCC enables us to perform experiments using datasets characterized by varying spatial resolutions.

The specifics of the MCC Bayer Pattern raw data employed in our experiments are delineated in Table-4.26. The MCC raw data complies with the Planetary Data System (PDS) standard and is available for download at no cost through the MOM LTA web browse portal (MOM, 2022) for the scientific user community. The surface features of Mars include a variety of areo-morphic structures such as Mons, craters, canyon systems, fractures, and grabens. We choose datasets that showcase a range of Mars surface structures for our direct feature extraction experiments, with the goal of generalizing the performance of the proposed method. Additionally, the datasets are selected at different spatial resolutions to acquire insights into the performance of feature extraction across diverse scales.

Table-4.27 delineates the parameter settings employed in the experiments for the proposed method, aiding in the assessment of feature extraction performance. The experimental configuration is meticulously crafted to guarantee that the direct feature extraction method produces optimal results under varying conditions, including different regions, scales, and imaging conditions. The evaluation of image co-registration performance involves using MCC Bayer Pattern raw images with diverse resolutions and comparing them with their demosaic counterparts.

Table-4.26. Details of MCC Bayer Pattern Raw Images

<b>Region Covered</b>	<b>PDS File Name</b>	<b>Imaging Time (YYYY-MM-DD, HH:MM:SS)</b>	<b>Spatial Resolution (in m)</b>
Arisa Mons	MCC_MEB_20141227T020400953_D_D32.IMG	2014-12-27, 02:04:00	453
Gale Crater	MCC_MEB_20150117T213420018_D_D32.IMG	2015-01-17, 21:34:20	468
Pital Crater	MCC_MEB_20150423T072306117_D_D32.IMG	2015-04-23,07:23:06.11	44
Tharsis Tholus	MCC_MEB_20150903T221933652_D_D32.IMG	2015-09-03,22:19:33.65	292

Table-4.27. Experimental Parameters Configuration for the Proposed Method

<b>S.No.</b>	<b>Experimental Parameters</b>	<b>Specified Value</b>
1.	Sigma	4.0
2.	Pre-defined Threshold	0.75
3.	FSS Threshold	0.8
4.	Window Size	5.0

*A. Comparison of Visual Quality in Gradient Magnitude Maps (GMM)*

Gradient information is extracted from MCC Bayer Pattern images, and Gradient Magnitude Maps (GMM) are directly created from the raw data. Demosaicking is applied to the MCC Bayer raw image to generate a full RGB color image, followed by the extraction of the Gradient Magnitude Map (GMM) from the entire color spectrum. Figures 4.43 to 4.48 depict the Gradient Magnitude Maps (GMM) derived from MCC Bayer Pattern raw images in comparison to the

GMM from MCC demosaic images, facilitating visual evaluation. In our experiment, we consider a variety of areo-morphological structures, including different spatial resolutions and physically distant locations on Mars.

In Figure-4.43, the Arisa Mons outer ring is directly extracted from the MCC Bayer raw image, demonstrating similarity with the Gradient Magnitude Map (GMM) generated from the MCC demosaic image. Figure-4.44 illustrates the precise tracking of the circumference of Gale Crater in the Gradient Magnitude Map (GMM) derived from MCC Bayer data, a significant landing site for NASA's Curiosity rover. Pital Crater, featured in Figure-4.45, is an impact crater captured by MCC at a relatively higher spatial resolution. The Pital crater wall is accurately extracted from the GMM in the MCC Bayer image, disclosing a diameter of around 40 kilometers. Tharsis Tholus, a shield volcano portrayed in Figure-4.46, exhibits a dome-shaped areo-morphic feature on Mars, and the direct Gradient Magnitude Map (GMM) illustrates the area covered by the volcano on the Martian surface.

The visual outcomes from the experiments confirm that the proposed method undeniably extracts features from MCC Bayer pattern raw planetary images. The developed methodology accurately captures prominent areomorphological features such as craters, Mons, and Valles on the Martian planetary surface, regardless of the specific topographical feature targets. Moreover, the feature map of gradient magnitude from the raw Bayer planetary image closely mirrors the features extracted from the demosaic color planetary image.

The introduced method is applied to diverse types of remote sensing images, as exemplified in the experimental section using planetary remote images from the Viking Orbiter Mission (Snyder, 1977) and the Mars Global Surveyor (MGS) Mission (Albee et al., 1998). To evaluate its efficacy, a simulated Bayer pattern image is created from color multi-spectral images. Figure-4.47 showcases a Bayer pattern image of the entire Mars disc alongside a color multispectral image captured by the Viking Orbiter. Figures 4.47(c) and 4.47(d) illustrate the respective gradient magnitude maps derived from the Bayer pattern image and the multispectral color image. The identical procedure is implemented for MGS images, and the resulting gradient magnitude maps are exhibited in Figure-4.48. Significantly, the extraction of features from Bayer pattern images closely mirrors the features found in multispectral color images from both the Viking Orbiter and MGS.

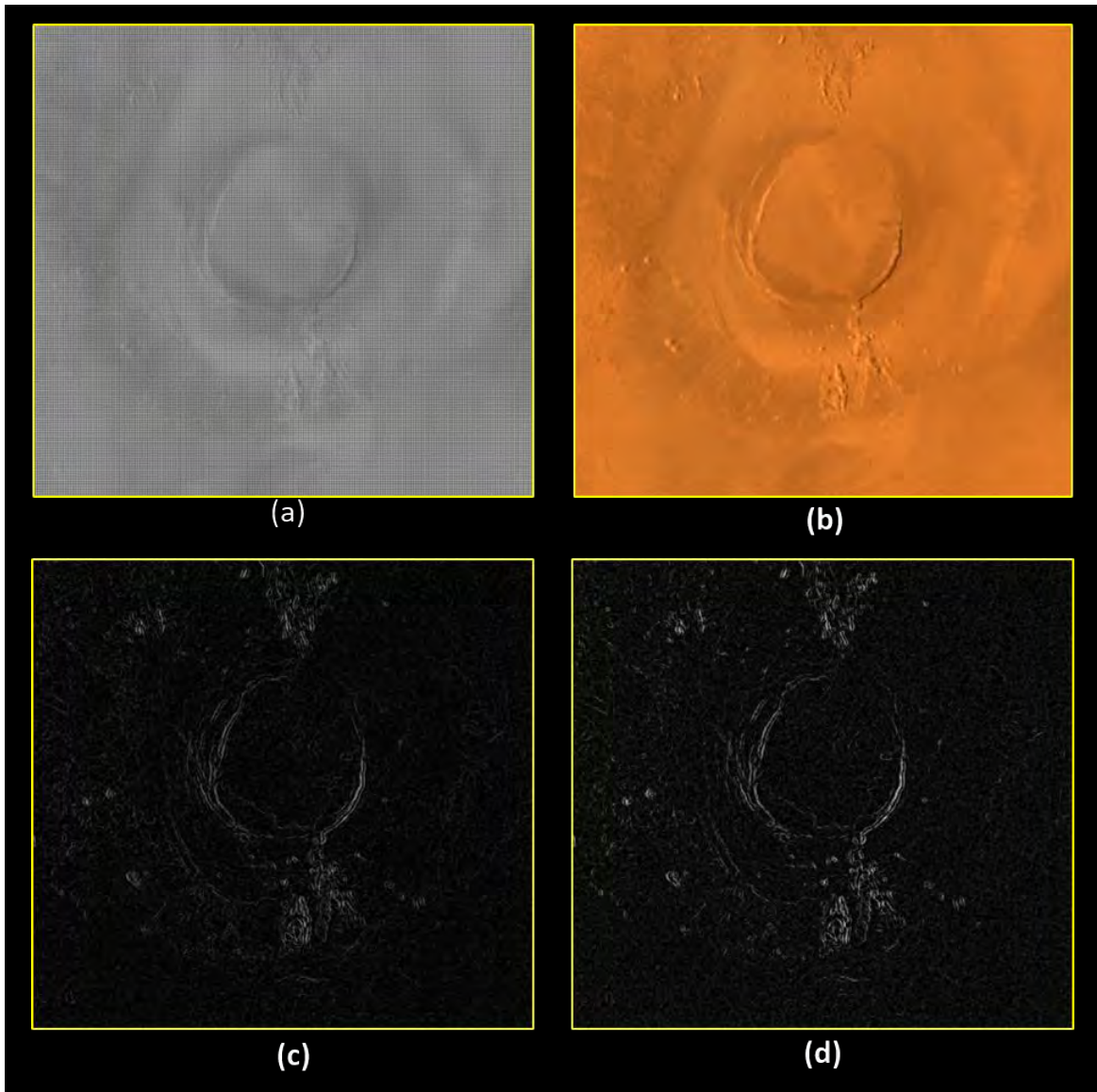


Figure-4.43. Extraction of Features Directly at Arisa Mons (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image



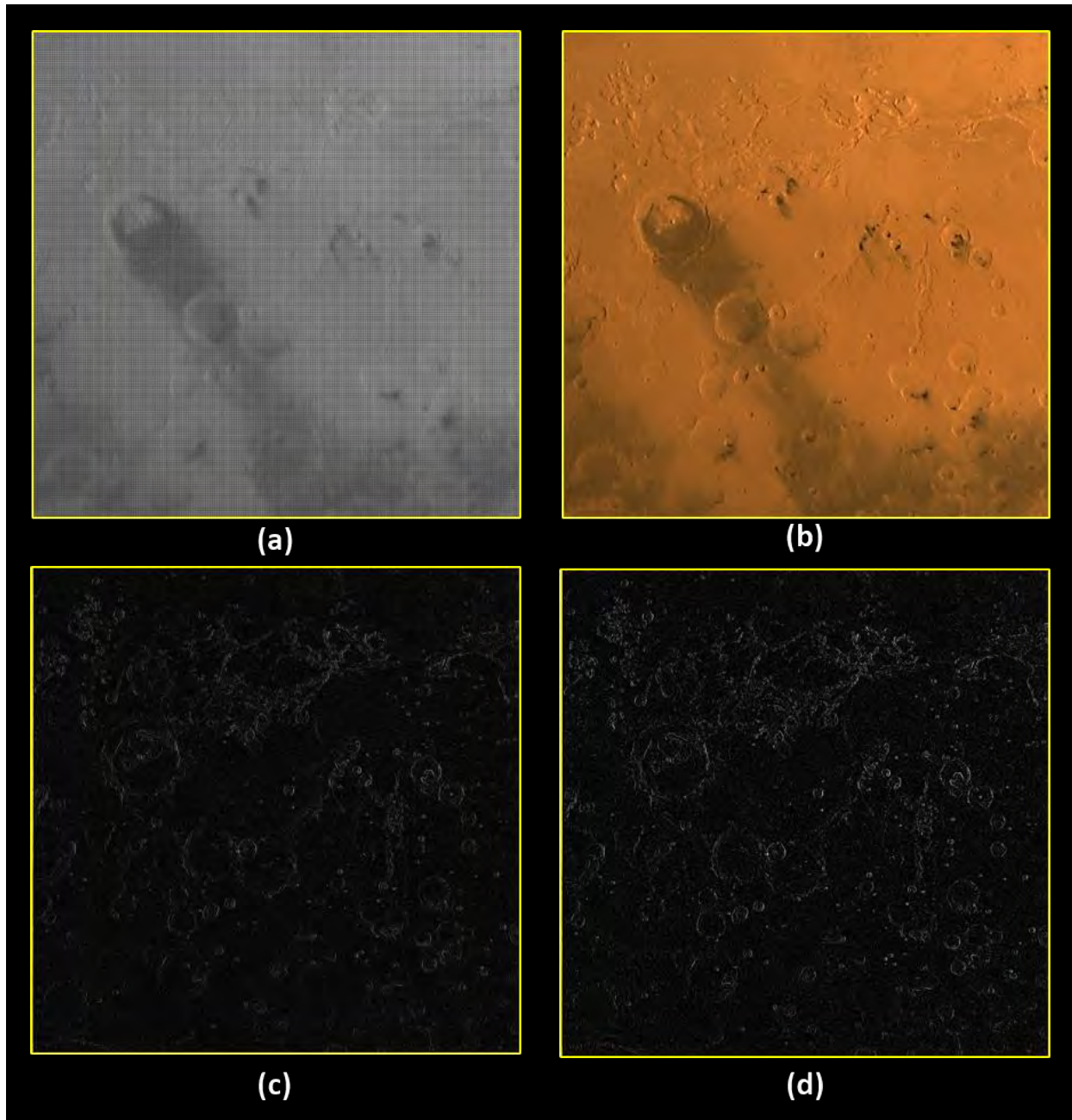


Figure-4.44. Extraction of Features Directly in the Vicinity of the Gale Crater Region (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image

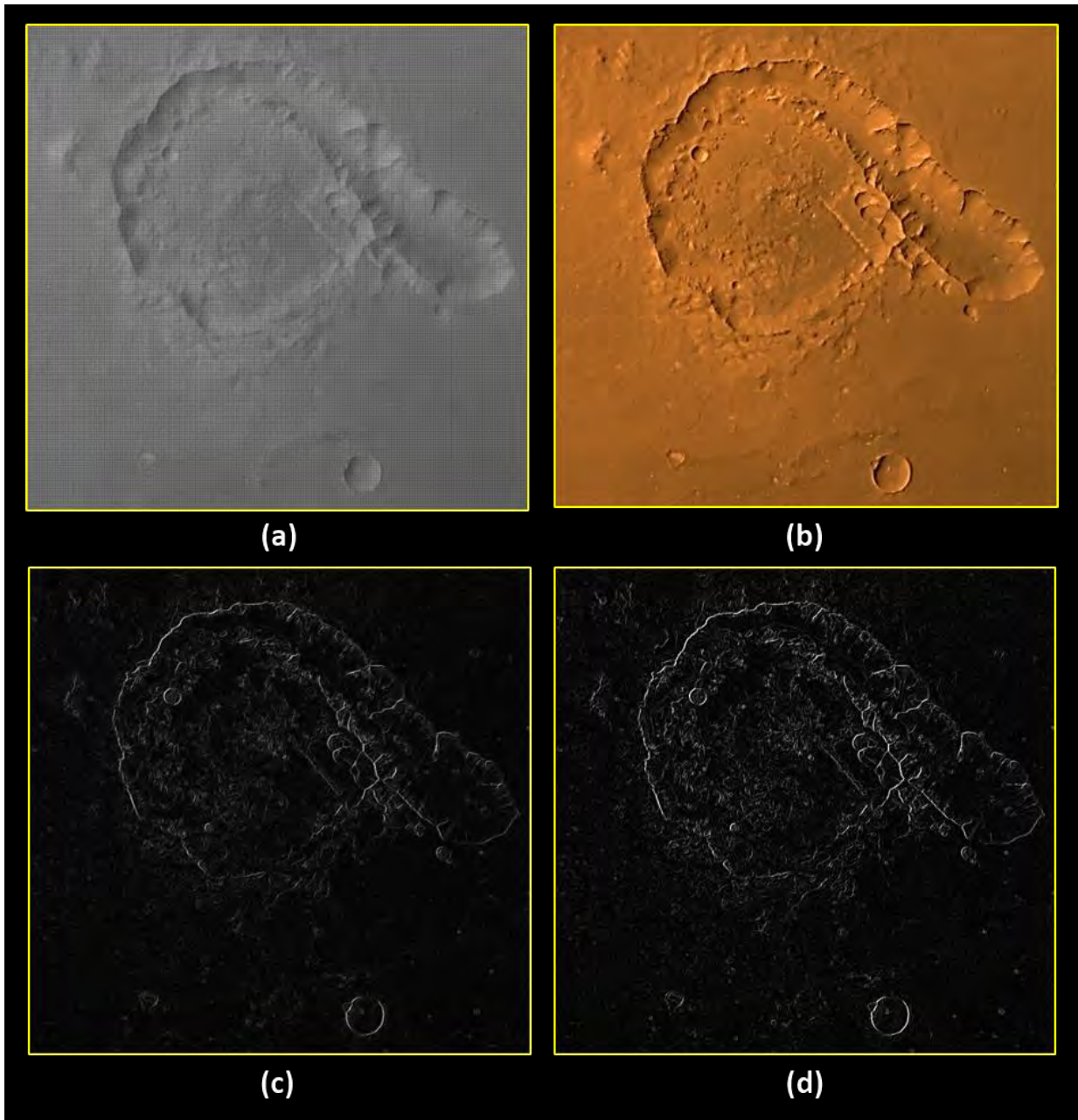


Figure-4.45. Extraction of Features Directly at the Pital Crater (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image

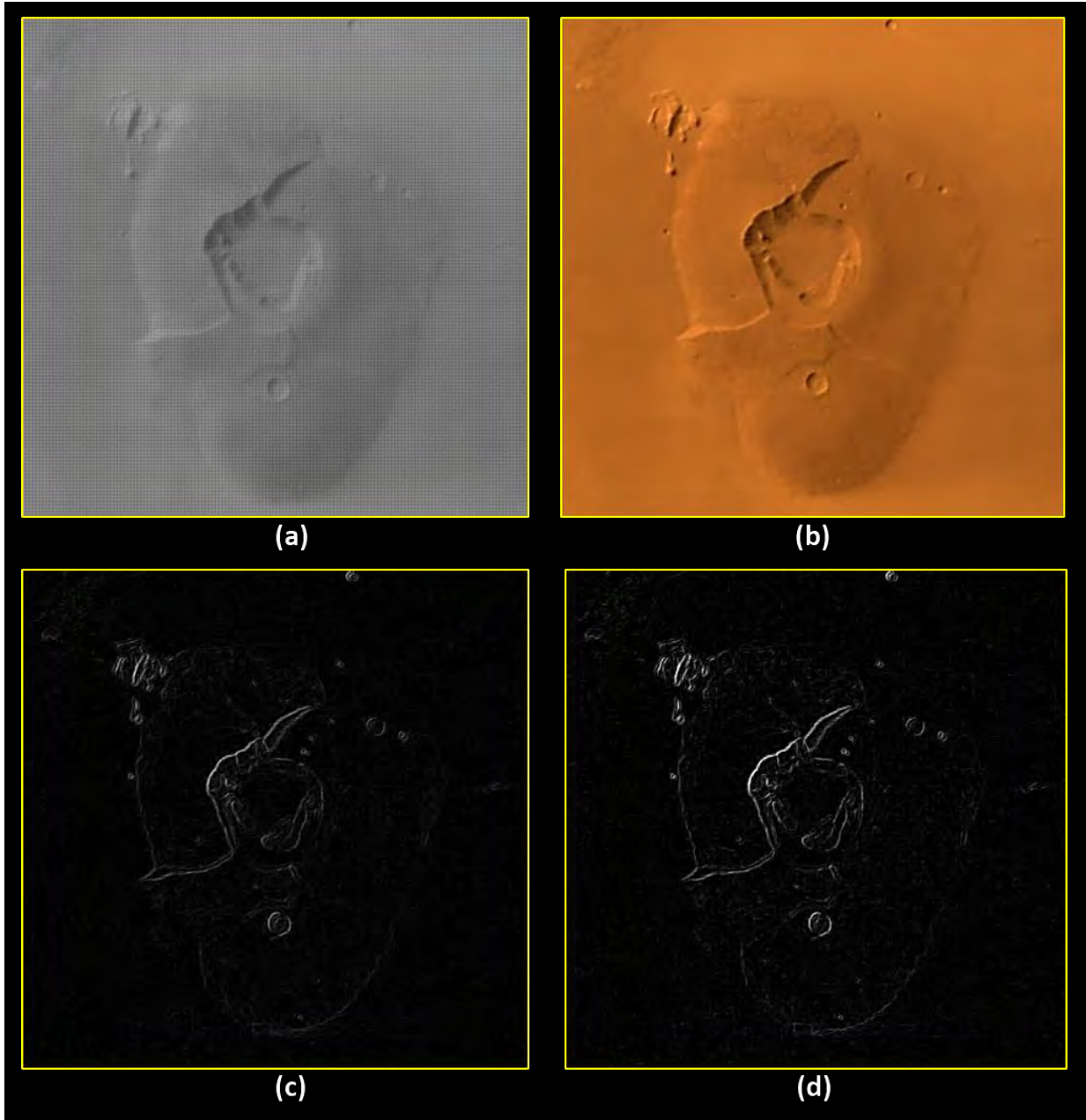


Figure-4.46. Extraction of Features Directly at Tharsis Tholus (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image

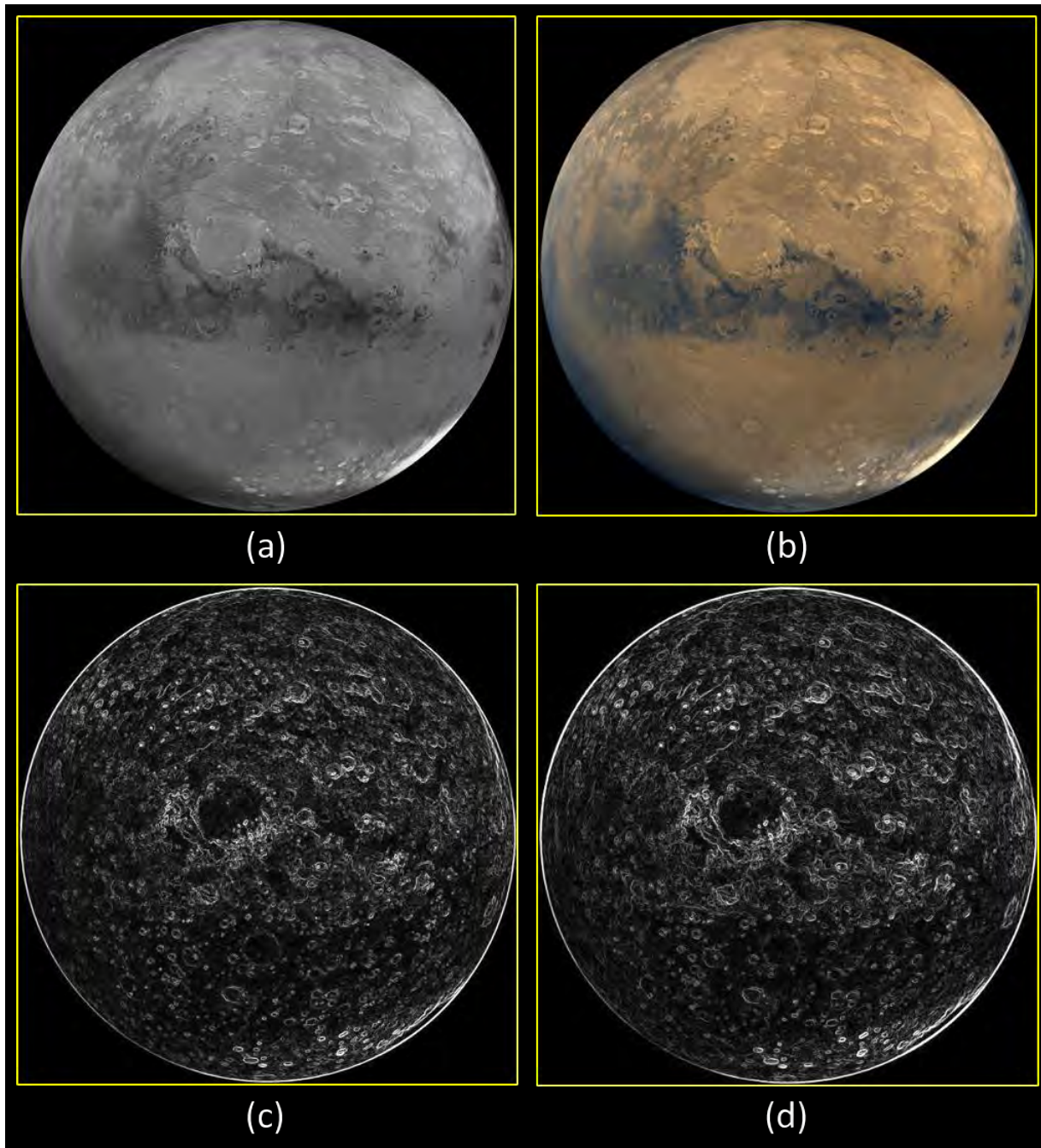


Figure-4.47. Extraction of Features Directly from a Full Disc Image of Mars Captured by the Viking Orbiter Mission (a) Raw Bayer Pattern Image (b) Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from Demosaicked Image

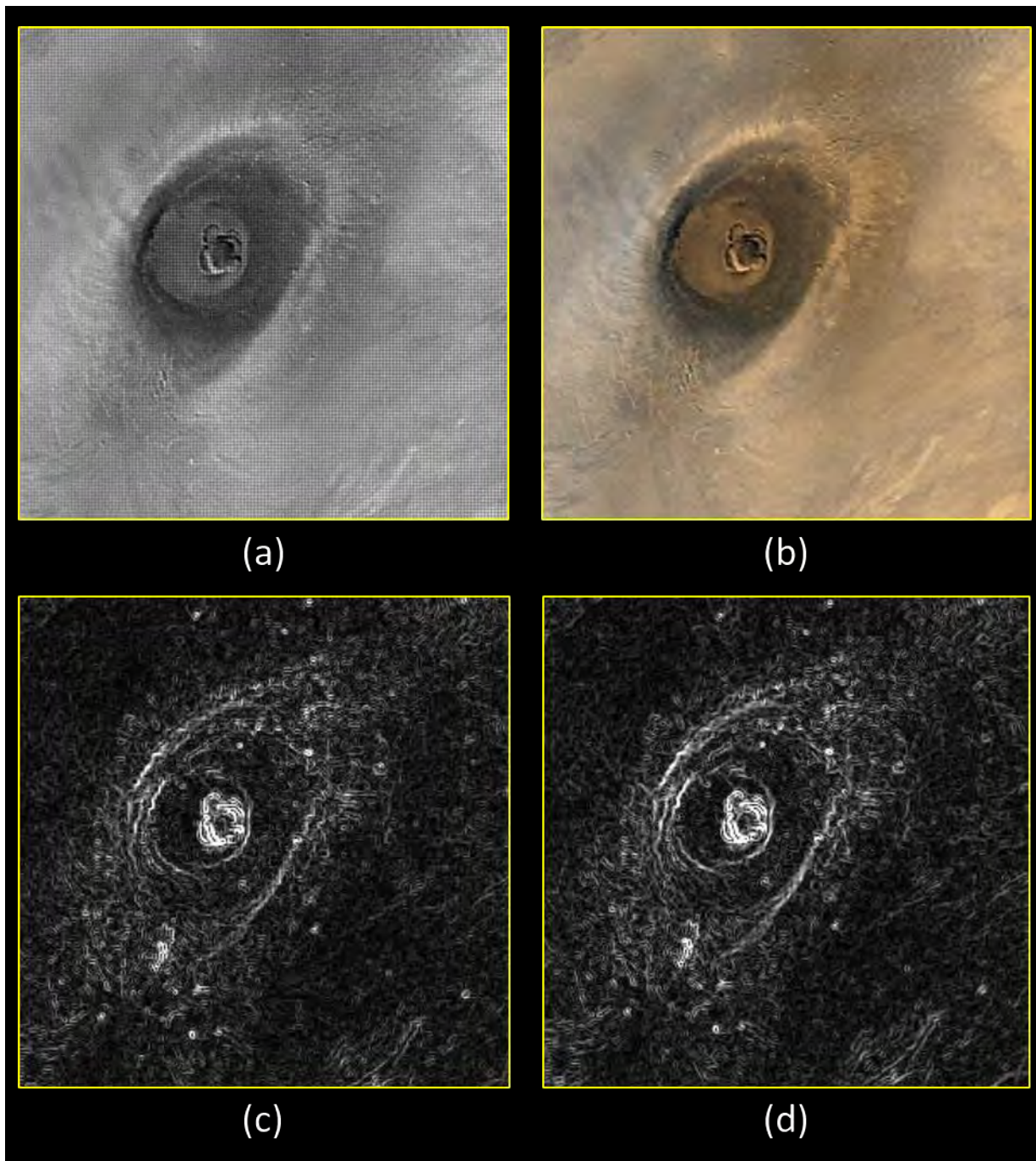


Figure-4.48. Extraction of Features Directly from an Image of Mars Captured by the Mars Global Surveyor Mission (a) MCC Raw Bayer Pattern Image (b) MCC Demosaicked Image (c) Gradient Magnitude Map (GMM) derived from MCC Bayer Pattern (d) Gradient Magnitude Map (GMM) obtained from MCC Demosaicked Image

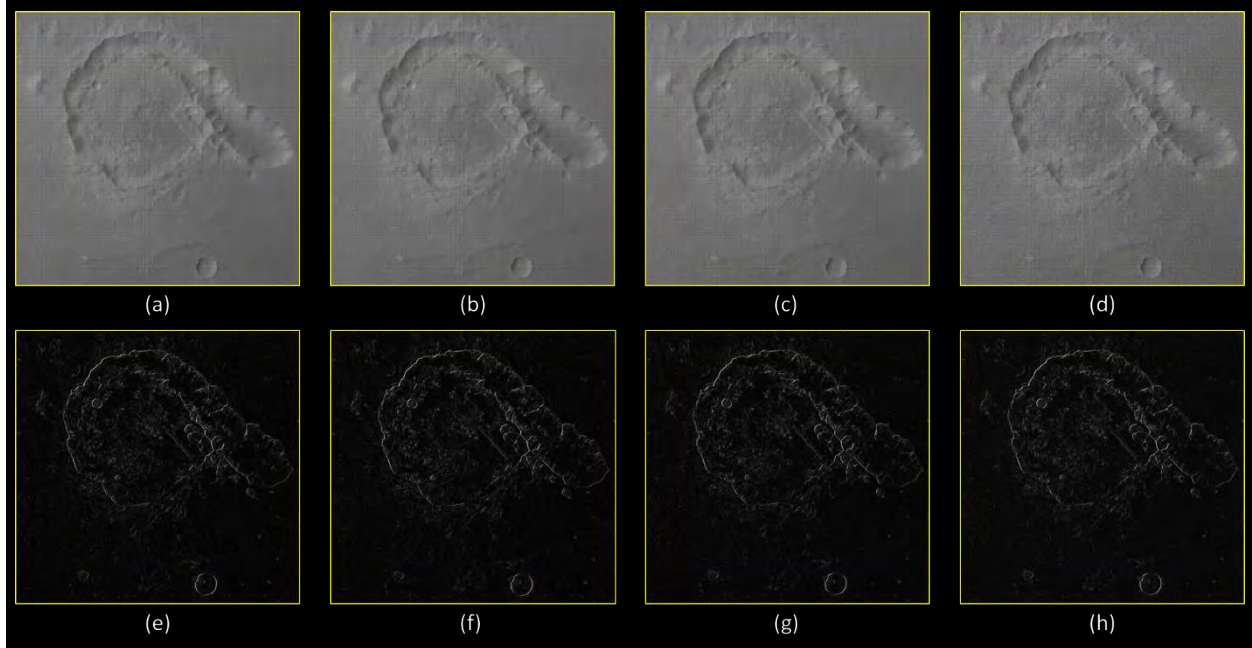


Figure-4.49. Evaluating feature extraction performance on the Pital Crater under different levels of Gaussian Noise (a) Raw MCC Bayer Pattern Image with no noise. (b) Image in its raw form with the addition of Gaussian Noise  $\sigma = 2.0$  (c) Image in its raw form with the addition of Gaussian Noise  $\sigma = 5.0$  (d) Image in its raw form with the addition of Gaussian Noise  $\sigma = 10.0$  (e) Image with extracted features and no noise (f) Image with extracted features and the presence of noise  $\sigma = 2.0$  (g) Image with extracted features and the presence of noise  $\sigma = 5.0$  (h) Image with extracted features and the presence of noise  $\sigma = 10.0$

Additionally, the effectiveness of the proposed method is assessed under various noise conditions. To replicate this scenario, different intensities of Gaussian noise are introduced to MCC raw Bayer pattern images. Subsequently, feature extraction is directly conducted from the noisy Bayer pattern planetary images. Figure-4.49 depicts an instance by showcasing an image of the Pital crater situated on the surface of Mars. From a visual standpoint, it is evident that features are successfully extracted from MCC Bayer pattern images, even when Gaussian noise is present.

### *B. Quantitative Evaluation and Performance in Co-Registration*

The quantitative assessment of Gradient Magnitude Maps (GMM) entails employing diverse image quality metrics. The Gradient Magnitude Similarity Deviation (GMSD) (Xue et al., 2013) is utilized as a widely used metric to measure the similarity of gradient magnitudes. It signifies the

difference between the direct gradient map extracted from MCC Bayer pattern raw images and the reference gradient map. A reduced GMSD value indicates a more accurate match between the gradient maps. The Mean Structural Similarity Index Map (MSSIM) (Wang et al., 2004) serves as a conventional full-reference image quality metric, evaluating luminance, contrast, and structure alignment between two images. Peak Signal to Noise Ratio (PSNR), which varies inversely with Mean Square Error (MSE), offers a consistent benchmark for assessing image quality. Moreover, metrics such as Canberra Distance (CD) and Sum of Absolute Distance (SAD) unveil the gradient map performance of Bayer and Demosaic images in comparison to reference data.

Table-4.28 displays these image quality metrics along with their descriptions and interpretations, employed for assessing the performance of direct feature extraction. Figure-4.50 exhibits the references of the extracted Region of Interest (ROI) from Mars Digital Image Model (MDIM) 2.1, illustrating three recognized Martian terrains—Arisa Mons, Gale Crater, and Pital crater. MDIM 2.1, acting as a global base map of Mars with a spatial resolution of 231 meters, functions as a reference image with enhanced spatial resolution to ensure precise image registration. The Region of Interest (ROI) is extracted from MDIM 2.1, subjected to geometric transformation to match the MCC image resolution through the application of a standard cubic resampling kernel, resulting in the creation of a seamlessly corresponding reference image. The reference gradient image derived from the transformed MDIM 2.1 image is generated through the application of a Sobel edge detector. This serves as a benchmark gradient map for evaluating Gradient Magnitude Maps (GMM) from MCC Bayer Pattern raw images and GMM from MCC demosaic images.

Table-4.28 offers a quantitative comparison of the extracted Gradient Magnitude Maps (GMM) from MCC Bayer pattern images and MCC Demosaic images in three specific regions of Mars—Arisa Mons, Gale Crater, and Tharsis Tholus. In various Martian landscapes, evaluations of GMSD, MSSIM, PSNR, CD, and SAD indicate that the metric values of Gradient Magnitude Maps (GMM) from MCC Bayer Pattern images closely correspond to those from MCC Demosaic images.

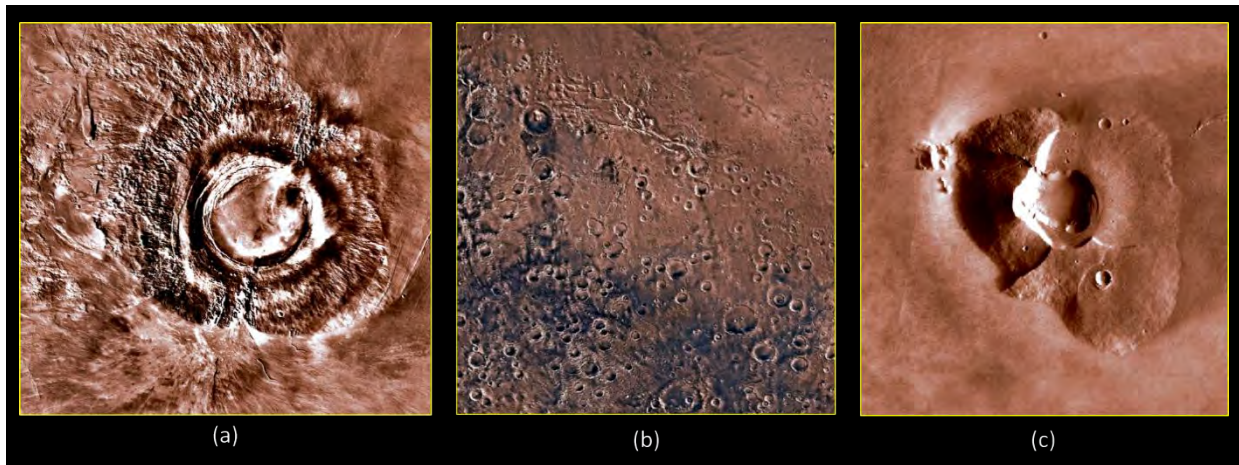


Figure-4.50. Transformed Region of Interest (ROI) reference extracted from MDIM 2.1 (a) Arisa Mons (b) Gale Crater (c) Tharsis Tholus

Table-4.28. Comparative Analysis of Quantitative Gradient Maps Between Raw image using the MCC Bayer pattern and Image demosaicked with MCC

<b>Gradient Map</b>	<b>Region Covered</b>	<b>MSSIM</b> ↑	<b>GMSD</b> ↓	<b>CD</b> ↓	<b>PSNR</b> ↑	<b>SAD</b> ↓
MCC Bayer Pattern Image	Arisa Mons	0.842	0.152	0.968	22.923	0.803
	Gale Crater	0.782	0.138	0.941	27.731	0.692
	Tharsis Tholus	0.812	0.115	0.929	29.324	0.812
MCC Demosaic Image	Arisa Mons	0.851	0.145	0.961	22.952	0.782
	Gale Crater	0.786	0.131	0.936	27.781	0.678
	Tharsis Tholus	0.818	0.102	0.925	29.335	0.793

The image quality metrics, such as GMSD, MSSIM, PSNR, CD, and SAD, are computed at different levels of Gaussian noise to evaluate the method's effectiveness in the presence of noise. Table-4.29 demonstrates the performance of feature extraction under varying levels of Gaussian noise. The features extracted from the MCC demosaic image are utilized as a benchmark for calculating quality metrics at different noise levels. The findings suggest that features of the Pital



crater are consistently and distinctly extracted, exhibiting minimal decline in performance even when subjected to different noise levels.

Table-4.29. Performance of Feature Extraction Across Various Levels of Noise

<b>Pital Crater Feature Extracted Image at different noise level</b>	<b>MSSIM</b> ↑	<b>GMSD</b> ↓	<b>CD</b> ↓	<b>PSNR</b> ↑	<b>SAD</b> ↓
No Gaussian Noise	0.912	0.141	0.916	28.923	0.731
Gaussian Noise $\sigma=2.0$	0.902	0.143	0.921	28.742	0.739
Gaussian Noise $\sigma=5.0$	0.894	0.145	0.928	28.553	0.744
Gaussian Noise $\sigma=10.0$	0.890	0.148	0.951	28.422	0.765

The MCC Bayer Pattern images taken above Arisa Mons, Gale Crater, and Tharsis Tholus undergo the process of co-registration with the MDIM 2.1 transformed Region of Interest (ROI), as depicted in Figure-4.50. This is achieved using the suggested GI-SIFT feature matching technique. Concurrently, the MCC demosaic images over these Martian locations are matched and aligned through the utilization of the GI-SIFT technique. The MCC demosaic image is generated through a process that includes end-to-end radiometric correction. This process involves PRNU correction, deblurring, denoising, color balancing procedures, and demosaicking. Following the application of the FSS-RANSAC outlier removal method, accurately matched points are identified in both the MCC Bayer Image and the MCC Demosaic Image. The RMSE (Chai & Draxler, 2014) is subsequently calculated by using the correspondences established between co-registered keypoints of the MCC image and the MDIM 2.1 reference through the GI-SIFT and FSS-RANSAC techniques. The definition of RMSE is as follows:

$$RMSE = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|^2\right)} \quad (4.64)$$

where, N represents the number of accurately matched points,  $X_i$  denotes the estimated coordinates  $(x_i, y_i)$  in the MCC image, and  $\hat{X}_i$  represents the actual coordinates  $(\hat{x}_i, \hat{y}_i)$  in the image reference of Mars.

Another valuable metric for assessing the effectiveness of image registration is the Correct Matching Ratio (CMR) (Zhang, 2020). Its definition is provided below.

$$CMR = \frac{N_c}{N_T} \quad (4.65)$$

In this context,  $N_T$  represents the count of matched correspondences post outlier elimination, while  $N_c$  signifies the number of precisely correct matches achieved after the removal of false matches.

Table-4.30. Comparison of RMSE and CMR Accuracy for Image Co-Registration

Image Details	Martian Region	CMR ↑	RMSE (in pixel) ↓
MCC Bayer Pattern Image	Arisa Mons	0.923	0.452
	Gale Crater	0.956	0.338
	Tharsis Tholus	0.911	0.515
MCC Demosaic Image	Arisa Mons	0.928	0.312
	Gale Crater	0.961	0.268
	Tharsis Tholus	0.931	0.435

Table-4.30 displays the RMSE values acquired from various Martian landscapes, utilizing both MCC Bayer Pattern raw images and MCC demosaic images. The precision of co-registration achieved by directly employing MCC Bayer Pattern raw images closely aligns with that attained using MCC demosaic images. The mean root mean square error (RMSE) obtained using MCC Bayer Pattern raw images consistently remains within the 0.5 pixel range across various Martian areas. Additionally, the table provides a comparison of CMR accuracy between MCC Bayer Pattern images and MCC demosaic images. The CMR metric, computed using the proposed method on MCC Bayer pattern images, exhibits a high degree of similarity to CMR values obtained from MCC demosaic images, indicating comparable matching features' performance across a variety of Martian terrains.

To evaluate the resilience of the suggested approach in extracting features directly from raw Bayer pattern planetary images and aligning morphological structures, the processing workflow is carried out iteratively. Table-4.31 showcases the RMSE results for the co-registration of images related to the Arisa Mons feature on the surface of Mars. The outcomes presented in Table-4.31

suggest that, even when the proposed approach randomly chooses points, the co-registration performance maintains consistency across numerous iterations.

The investigation also takes into account multi-temporal imaging acquisitions covering Arisa Mons and Gale Crater, obtained under varying lighting conditions attributable to the Mars Orbiter Mission spacecraft's highly elliptical orbit. Table-4.32 demonstrates the RMSE results for multi-temporal imaging acquisitions featuring intricate backgrounds and negligible local variations. The results presented in Table-4.32 indicate that the performance of image co-registration using the suggested method remains uniform across multi-temporal images taken under different solar illumination conditions.

Table-4.31. Co-Registration Performance of Images for Arisa Mons Across Iterative Executions of the Proposed Method

Execution Number	RMSE (in pixel) ↓
Run-1	0.452
Run-2	0.456
Run-3	0.451
Run-4	0.454

Table-4.32. Co-Registration Performance of Multi-temporal Images in the Presence of Complex Backgrounds and Diverse Illumination Conditions

Region Covered	RMSE (in pixel) ↓
Gale Crater Multi-temporal Image-1	0.338
Gale Crater Multi-temporal Image-2	0.381
Gale Crater Multi-temporal Image-3	0.318
Arisa Mons Multi-temporal Image-1	0.452
Arisa Mons Multi-temporal Image-2	0.406
Arisa Mons Multi-temporal Image-3	0.482

Figure-4.51 illustrates horizontal image swipes designed to assess the registration accuracy between the MCC Bayer Pattern intensity image (displayed on the left) and the MDIM 2.1 reference map (shown on the right) across different categories of Mars craters. A visual examination validates that the gradient derived from the MCC Bayer Pattern intensity image

effectively aligns various feature targets on Mars with co-registration accuracy at a sub-pixel level, in comparison to the MDIM 2.1 reference.

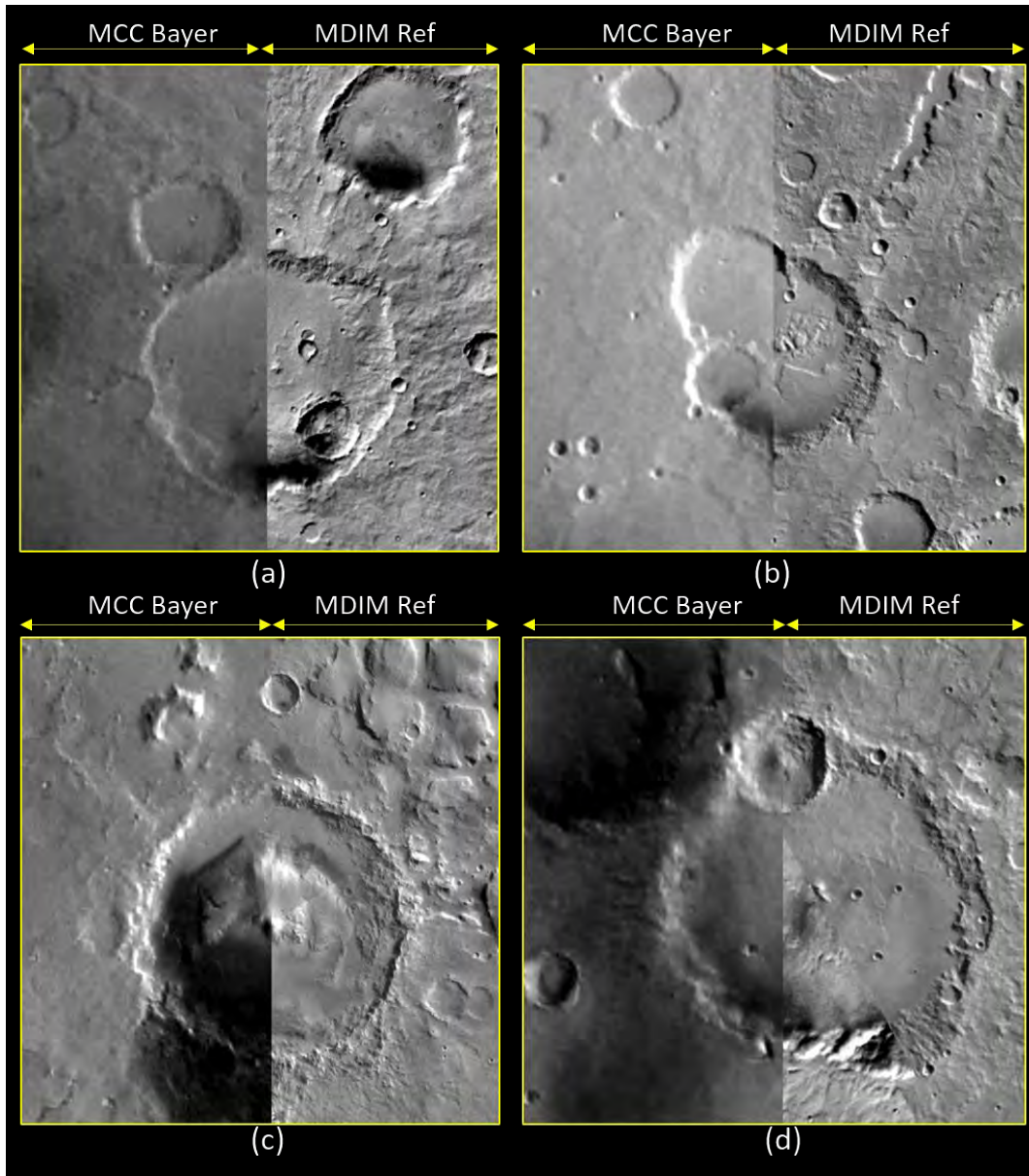


Figure-4.51. Visual Assessment of Sub-Pixel Alignment between MCC Bayer and MDIM Reference for Different Mars Craters (a) Crater-01 (b) Crater-02 (c) Crater-03 (d) Crater-04

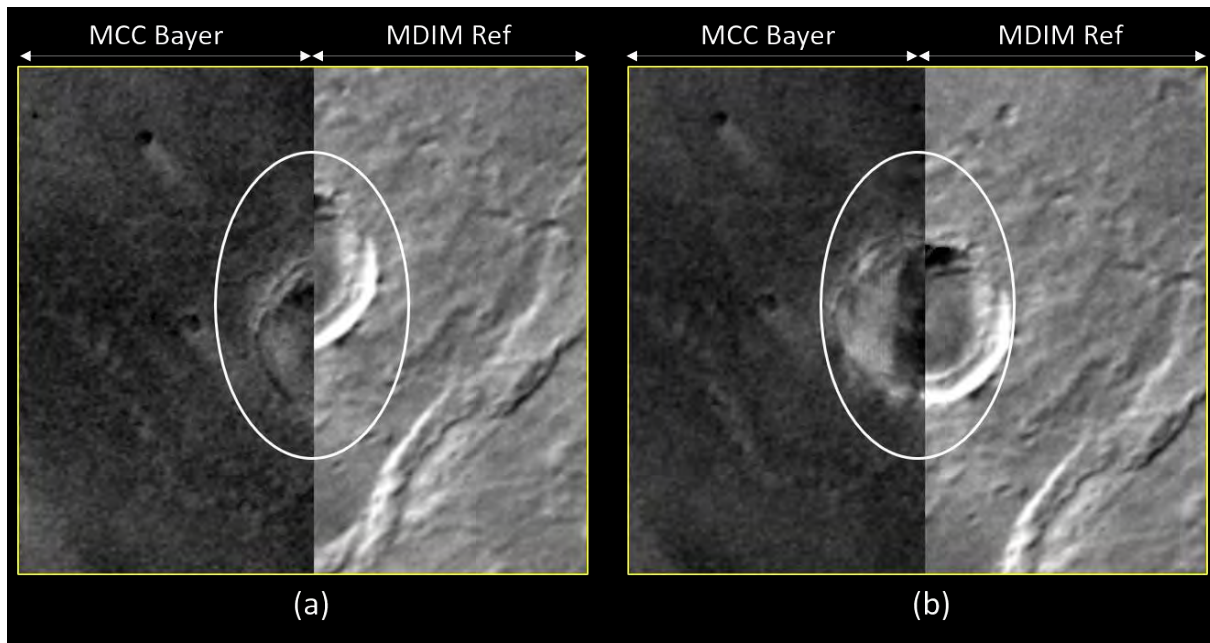


Figure-4.52. Results of Co-Registration for Planetary Images Across a Martian Crater (a) Prior to Registration (b) Post Registration

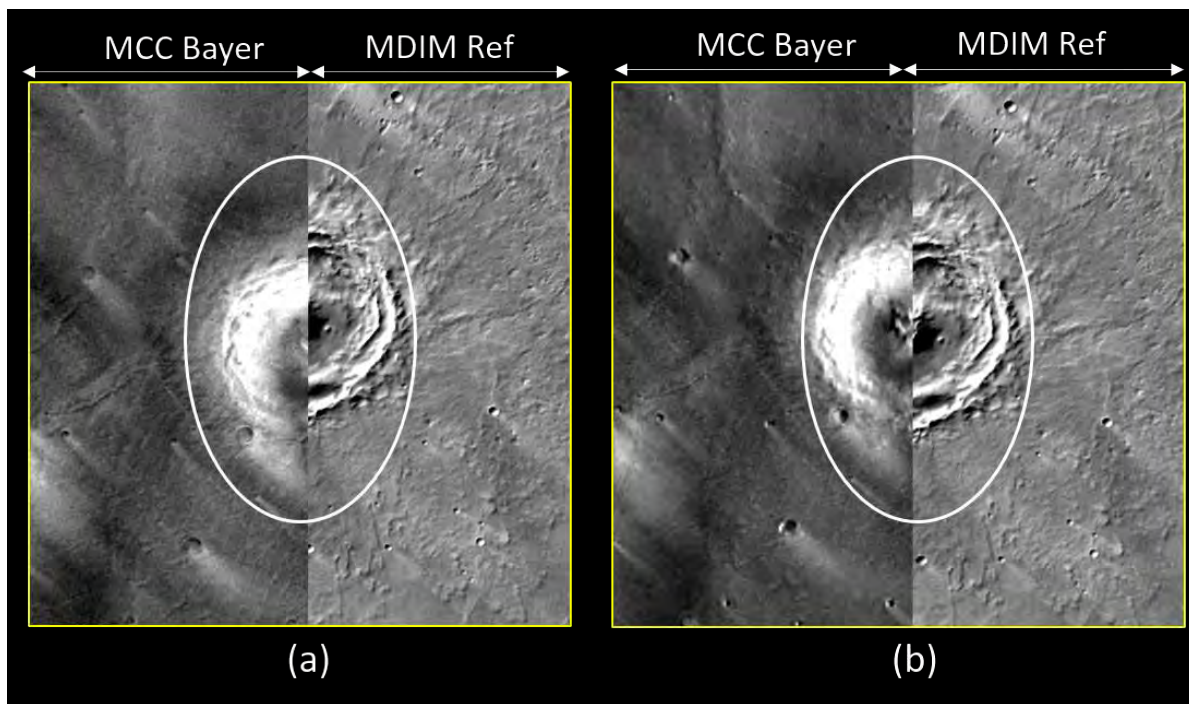


Figure-4.53. Results of Co-Registration for Planetary Images Across Kinkora Crater (a) Prior to Registration (b) Post Registration

The evaluation of the proposed method is extended by horizontally aligning images of various Martian feature terrains, employing both MCC Bayer and MDIM reference images. In Figure-4.52(a), the MCC Bayer Pattern raw image is presented pre-registration, offering a visual representation of the misalignment observed in the crater region. After the co-registration of planetary images using the developed methodology, the MCC Bayer image is overlaid onto the MDIM reference, as depicted in Figure-4.52(b). The effectiveness of the methodology is also tested at Kinkora Crater, a comparatively larger crater, as illustrated in Figure-4.53, showcasing the sub-pixel level alignment achieved post-registration in comparison to the MDIM reference image.

Figures 4.54, 4.55, and 4.56 showcase the effectiveness of FSS-RANSAC across Martian terrain. The matched points, refined through the FSS-RANSAC outlier removal algorithm, demonstrate uniformity across diverse regions of the Martian landscape. In the figures, lines shaded in a dark blue hue depict potential keypoint correspondences identified by FSS-RANSAC during the estimation of the final model parameters.

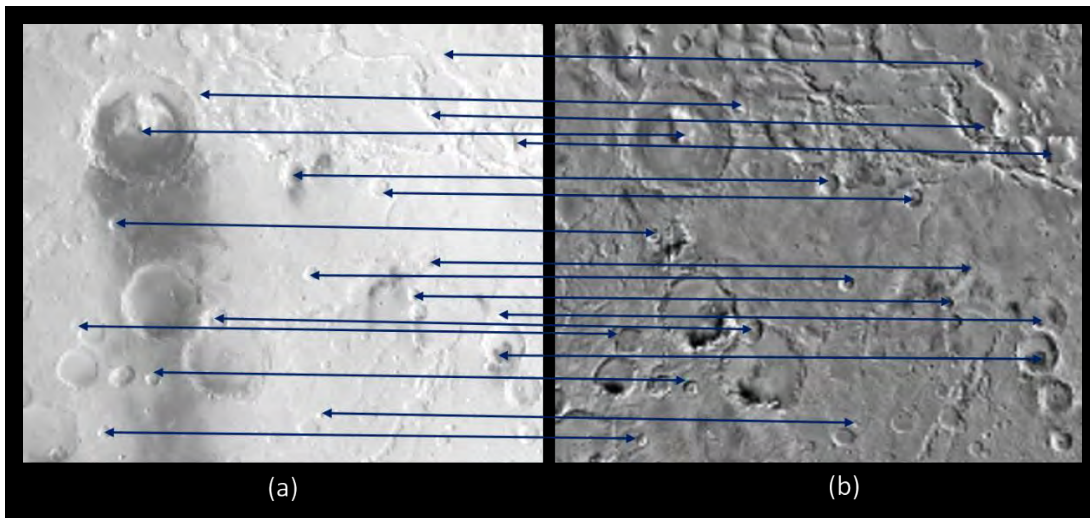


Figure-4.54. Performance of Outlier Removal Utilizing FSS-RANSAC Near Gale Crater (a) MCC Bayer Intensity Image (b) MDIM 2.1 Reference Image

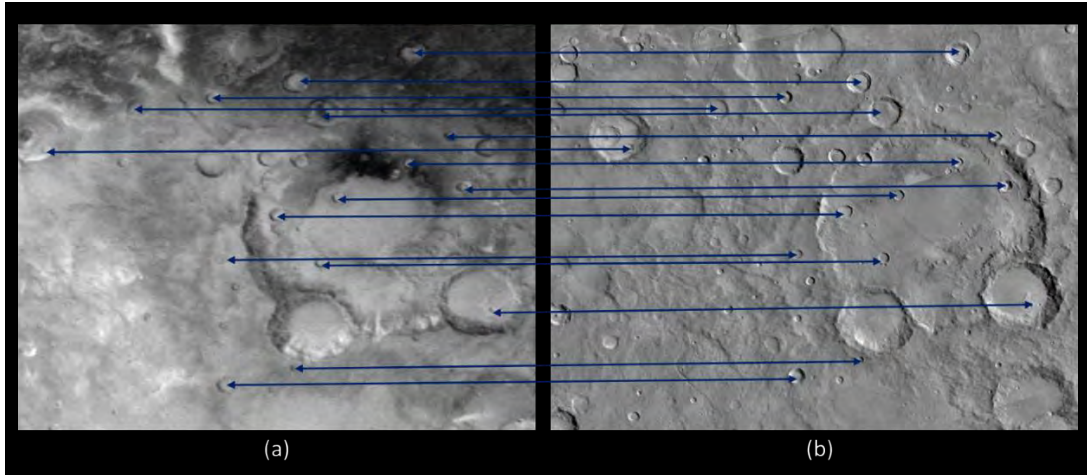


Figure-4.55. Performance of Outlier Removal Utilizing FSS-RANSAC Near Martian Craters (a) MCC Bayer Intensity Image (b) MDIM 2.1 Reference Image

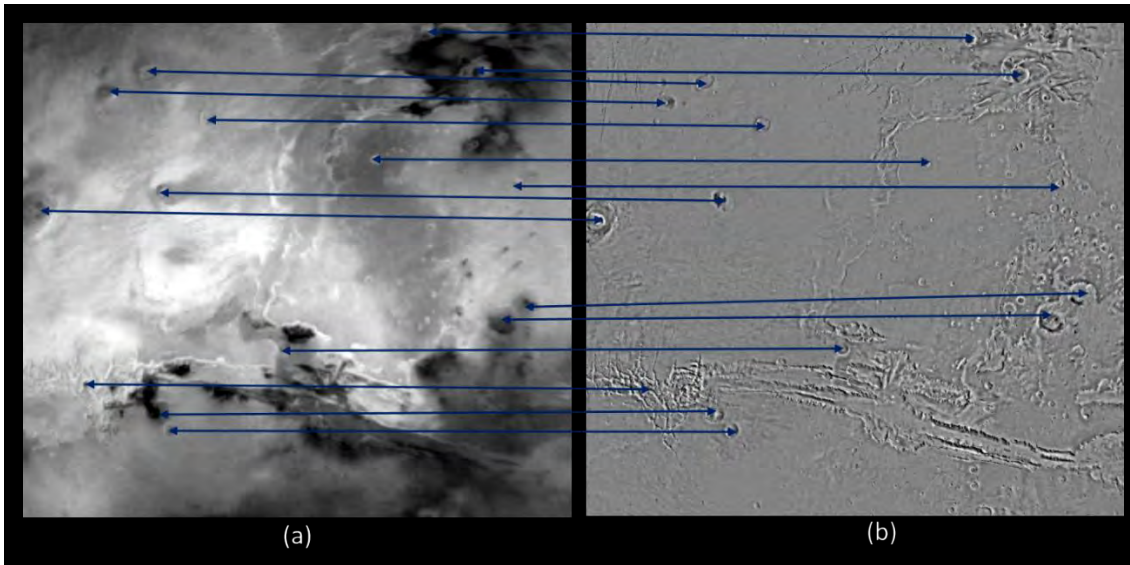


Figure-4.56. Performance of Outlier Removal Utilizing FSS-RANSAC Near Valles Marineris (a) MCC Bayer Intensity Image (b) MDIM 2.1 Reference Image

The goal for co-registration accuracy in planetary remote sensing images is set at less than 0.5 pixels. The suggested approach showcases proficient feature extraction, achieving image alignment at a sub-pixel precision and successfully meeting the established target. The assessment across diverse feature points on the Martian landscape reveals an average RMSE metric of 0.435 pixels, staying within the specified target threshold.

### *C. Comparison with State-of-the-Art Feature based Techniques*

The proposed image registration method is compared with eminent feature-based image registration techniques across diverse Mars surface regions. Conducting a comparison with state-of-the-art techniques enables a more in-depth quantitative evaluation, particularly since other registration methods also achieve sub-pixel accuracy and may exhibit visual similarities. This quantitative analysis provides valuable insights into the superior registration accuracy of the proposed method in comparison to other well-known registration techniques across various areomorphological structures on the Martian surface.

The pre-processing stage for planetary data encompasses the extraction of features from MCC raw Bayer pattern images, preparing abstract information for the co-registration of Bayer pattern data. The compared methodologies follow identical data pre-processing procedures to ensure a fair and consistent comparison. The proposed method is contrasted with Harris with RANSAC (Misra et al., 2012a), SIFT with RANSAC (Misra et al., 2021b; Hernández & Ortiz, 2020), ORB with MSC (Misra et al., 2021a), SURF with MLESAC (Misra et al., 2022a), and KAZE with MAGSAC (Alcantarilla et al., 2012; Misra et al., 2023a). Furthermore, a deep learning-based image registration technique is incorporated for comparison and assessment, involving a Deep Neural Network (DNN) with Semantic matching (Li et al., 2021), and a multi-scale framework employing the unsupervised learning method MU-NET (Ye et al., 2022) for planetary image registration. The assessment involves calculating the evaluation metrics RMSE (Chai & Draxler, 2014) and CMR (Zhang, 2020) to evaluate the image registration methods.

The proposed method exhibits superiority over other established and recent feature-based image co-registration techniques, as evidenced by improvements in both RMSE and CMR. The examination indicates that, in all three assessed Mars regions designated for image registration evaluation, the proposed method attains the lowest RMSE, as outlined in Table-4.33. Moreover, the CMR, which signifies the highest accuracy in image registration for Bayer pattern raw planetary images taken over the Martian surface, is significantly superior in the proposed method, as depicted in Table-4.34.



Table-4.33. Comparative Assessment of RMSE Utilizing Different Image Registration Techniques on Bayer Pattern Raw Planetary Images

S.No.	Mars Region Covered	SIFT + RANSAC	Harris + RANSAC	SURF + MLESAC	ORB + MSC	KAZE+ MAGSAC	MU-NET	Semantic DNN	Proposed Method
1.	Arisa Mons	0.576	0.786	0.584	0.662	0.612	0.526	0.598	0.452
2.	Gale Crater	0.674	0.581	0.498	0.563	0.553	0.491	0.488	0.338
3.	Tharsis Tholus	0.611	0.855	0.675	0.681	0.674	0.575	0.591	0.515

Table-4.34. Comparative Assessment of CMR Utilizing Different Image Registration Techniques on Bayer Pattern Raw Planetary Images

S.No.	Mars Region Covered	SIFT + RANSAC	Harris + RANSAC	SURF + MLESAC	ORB + MSC	KAZE+ MAGSAC	MU-NET	Semantic DNN	Proposed Method
1.	Arisa Mons	0.901	0.865	0.889	0.893	0.895	0.891	0.902	0.923
2.	Gale Crater	0.912	0.898	0.934	0.906	0.913	0.925	0.899	0.956
3.	Tharsis Tholus	0.862	0.861	0.857	0.856	0.862	0.857	0.882	0.911

In evaluating the feature matching performance for planetary image co-registration, the Number of Correct Matches (NCM) is calculated. NCM is determined for multi-temporal MCC images covering craters of diverse sizes using various image co-registration methods. The dimensions of the planetary image are 1817 pixels in width and 886 pixels in height. Table-4.35 provides NCM values, clearly demonstrating that the proposed method produces the highest number of correct feature matching points among multi-temporal planetary images. Figure-4.57 visually depicts the feature matching performance of various well-known image co-registration techniques, emphasizing that the proposed method consistently outperforms state-of-the-art counterparts in feature matching.

Table-4.35. Comparative Assessment of NCM Utilizing Different Image Registration Techniques on Multi-Temporal Bayer Pattern Raw Planetary Images

S.No.	Image Co-Registration Techniques	Number of Correct Matches (NCM)
1.	SIFT + RANSAC	40
2.	Harris + RANSAC	50
3.	SURF + MLESAC	80
4.	ORB + MSC	85
5.	KAZE + MAGSAC	45
6.	MU-NET	85
7.	Semantic DNN	65
8.	Proposed Method	120

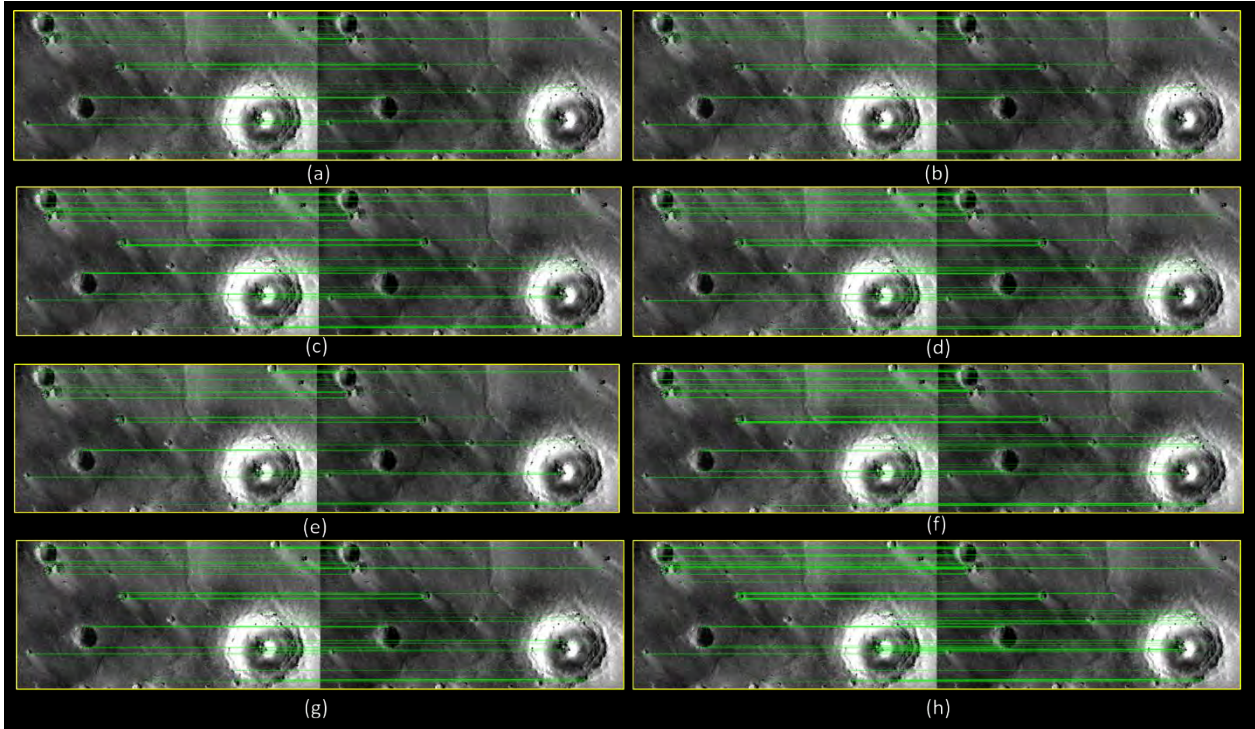


Figure-4.57. Performance in Matching Features Between Multi-temporal Planetary Images (a) Harris+RANSAC (b) SIFT+RANSAC (c) ORB+MSC (d) SURF+MLESAC (e) KAZE+MAGSAC (f) Semantic DNN (g) MU-NET (h) Proposed Method

#### 4.4.4. Summary

MCC Bayer Pattern raw images are employed for the direct extraction of information related to diverse areomorphological features on the surface of Mars. By capitalizing on the CDC assumption, the GMM is created, outlining crater boundaries and mons circumferences directly from the MCC raw data. Applying the Central Difference Gradient Operator and Sobel Edge Operator in accordance with CDC, these operations are executed across MCC raw images covering various Martian terrains. The visual comparisons indicate minimal discrepancy between Gradient Magnitude Maps (GMMs) derived from MCC Bayer Pattern raw images and those obtained from MCC demosaic images. The proposed method undergoes additional evaluation using planetary remote sensing images from the Viking Orbiter and Mars Global Surveyor, showcasing consistent feature extraction performance. The quantitative metrics calculated for both iterations of GMMs confirm the resilience of the developed direct feature extraction method.

The use of GI-SIFT feature matching, combined with the FSS-RANSAC outlier removal technique, effectively aligns MCC Bayer Pattern raw images with the standard Mars reference, achieving spatial alignment at a sub-pixel level. Both subjective visual quality assessment and quantitative evaluation underscore the efficacy of the proposed methodology in the co-registration of planetary images. Nevertheless, it's crucial to emphasize that the method is specifically designed for extracting features directly from planetary images formatted in Bayer pattern color schemes. The suggested technique is found to be appropriate for planetary remote sensing images captured under diverse lighting conditions. The assessment with multi-temporal MCC images, captured under distinct solar illumination conditions owing to the highly elliptical orbit of the Mars Orbiter Mission (MOM) spacecraft, illustrates the method's ability to consistently extract features across varied imaging scenarios. The established workflow for the direct extraction of features and image co-registration with MCC raw images offers fundamental characteristics applicable to tasks such as crater counting, identification of morphological structures, change detection, and subsequent image processing operations with minimal impediment.

## **5. NOVEL COUNTRY-LEVEL IMAGE MOSAICKING SYSTEM USING OPTICAL REMOTE SENSING IMAGERY**

### **5.1. Introduction**

The section introduces an innovative approach to create image mosaics utilizing the Resourcesat remote sensing dataset over India. Our research is driven by the goal of constructing a comprehensive image mosaicking framework capable of producing mosaic data products at the country level. The primary challenge lies in automatically managing the radiometry and the spatial characteristics of images captured through remote sensing, particularly in the overlapping regions, to generate a seamlessly stitched mosaic covering a large area. This article addresses key issues such as the alignment of extensive image strips and the creation of an automated blending technique for images (Misra et al., 2022c).

The established systematic process manages the preparation of input datasets and conducts the geo-registration of image strips using the Scale Invariant Feature Transform (SIFT) (Lowe, 1999), in conjunction with the Mode-Biased Random Sample Consensus (MB-RANSAC) outlier removal approach. The automated procedure for stitching images together utilizes a statistical approach based on regions, incorporating a joint optimization strategy across multiple frames. This technique creates a mosaic data product at the national level by combining co-registered image strips.

The experiments utilize Resourcesat LISS-3 and AWiFS imagery to evaluate the effectiveness of strip co-registration in diverse landscapes across the Indian terrain. The statistics for co-registration and comparisons of reflectance are carried out both prior to and following correction to quantitatively evaluate the effectiveness of the proposed methodology. Visual evaluations of the performance in overlap regions are conducted at various targets of interest to confirm consistency in both geometric and radiometric aspects.

### **5.2. Methodology Developed**

The system for creating image mosaics utilizes input consists of geo-referenced remote sensing image strips that have undergone system-level correction, and they are prepared for the processing of mosaic data. An image strip comprises multiple scenes captured at different times by the remote

sensing instrument, with its extension in the north-south or east-west direction depending on the instrument's imaging scheme. The developed mosaic procedure primarily entails meticulous radiometric and geometric data processing during the data processing stage to generate mosaic images.

A notable difficulty in creating remote sensing image mosaics involves achieving strip co-registration to guarantee geometric precision in the areas of overlap. Another significant challenge in constructing extensive mosaics lies in the automatic mosaicking of image strips within the geographic coordinate space using only one image resampling operation in the processing operation. The suggested methodology for mosaic data processing encompasses taking in input strips, re-projecting them into the mosaic space, implementing an innovative image spatial alignment process, and automatically stitching images to create a continuous and cohesive mosaic data output at the country level.

In the output phase, the system produces mosaic data on a band-by-band basis, assuming the essential metadata details are accessible, rendering the mosaic data product readily applicable for the scientific community. Figure-5.1 presents a diagrammatic representation illustrating the proposed integrated system for image mosaicking on a country-wide scale.

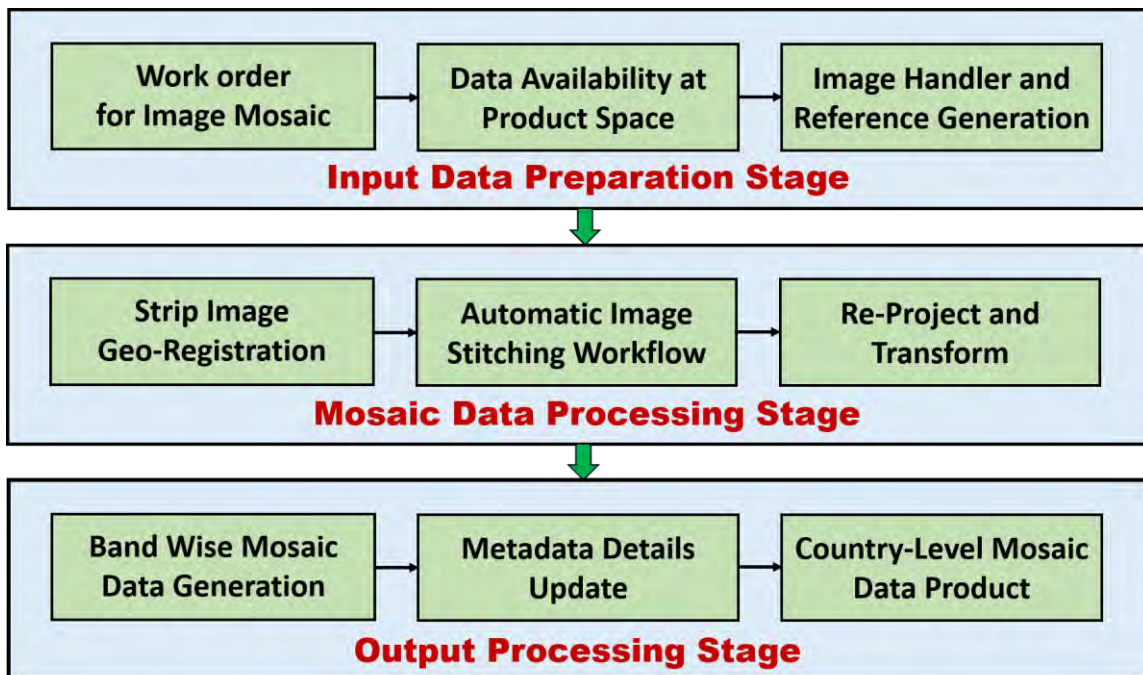


Figure-5.1. Proposed System for Integrated Image Mosaicking at the Country Level

### A. Satellite Data Preparation Stage

The initial phase of mosaic processing involves checking the request for work order, which must specify the image strips intended for stitching to achieve country-level coverage. It is imperative to ensure the availability of strip data for ingestion. The data preparation procedure assesses the handling of image strips, emphasizing adherence to the conventional geospatial data format. Moreover, a corresponding reference image is created to assist in strip co-registration.

The input image strip and its associated reference image are reprojected into the complete map projection space of India (Misra et al., 2012c). The projection selected for mosaic data is Lambert Conformal Conic (LCC), recognized for its appropriateness in mapping extensive geographical regions (Snyder, 1982). Table-5.1 furnishes the parameters of map projection for the complete India mosaic and the neighboring regions surrounding the Indian subcontinent. Geometric referencing and Ground control points for geographic registration of image strips are established using reference images from Resourcesat-2A LISS-3 (Misra et al., 2018). The extraction of the reference image occurs and subjected to geometric transformation to align the data with the coordinate system of the input image strip. Ensuring the geometric accuracy spanning the strips is accomplished by geo-registering the imagery with reference from Resourcesat-2A LISS-3.

Table-5.1. Parameters of Map Projection for the Indian Terrain

<b>S. No.</b>	<b>Map Projection Parameters</b>	<b>Value</b>
1.	Projection Name	<b>LCC</b>
2.	Datum/Ellipsoid	<b>WGS 1984</b>
3.	Latitude of Origin (in degrees)	<b>22.0</b>
4.	Longitude of Origin (in degrees)	<b>84.0</b>
5.	Standard Parallel-1 (in degrees)	<b>30.0</b>
6.	Standard Parallel-2 (in degrees)	<b>14.0</b>

The processing begins with the extraction of a shared region from the reference image of India, readying it for the subsequent co-registration of image strips. The products of the image strips are conventional geo-referenced products, and they might display local internal distortion within a few

pixels along the strip. These strips undergo the process of geo-registration with the reference through the image registration process specially tailored for imagery captured through optical remote sensing sensors, as per the proposed methodology. The aligned strips achieve sub-pixel registration with one another, leading to the continuity of features within the common region. The primary objective of the developed image registration workflow is to guarantee the production of a stitched dataset that maintains geometric consistency. The images are transformed by resampling to achieve the specified ground sample distance, and the projection is carried out according to the map projection parameters specified in Table-5.1.

### B. Precise geometric alignment between images

To attain sub-pixel matching of image features, a tiled Scale Invariant Feature Transform (SIFT) model incorporating Mode Biased RANSAC (MB-RANSAC) is utilized. The SIFT algorithm initiates by constructing scale space extrema through the application of a Gaussian Kernel ( $G$ ). Equation-5.1 demonstrates the connection between the Smooth Image ( $L$ ) and the Input Image ( $I$ ) at the coordinate point  $(x, y)$ .

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5.1)$$

$$\text{where } G(x, y, \sigma) = 1/2\pi\sigma^2(e^{-(x^2+y^2)/2\sigma^2}) \quad (5.2)$$

It has been observed that the Laplacian of Gaussian (LoG) provides consistent features and a resilient comprehension of scales, but its computational cost is significant. Therefore, a pyramid of the Difference of Gaussian (DoG) is generated and regarded as an estimate of the Laplacian of Gaussian (LoG). The subsequent step involves scanning each DoG image to identify the minimum and maximum values across neighboring points, including scale considerations.

One of the advantages of the Scale Invariant Feature Transform (SIFT) lies in the correlation of extracted keypoints with a descriptor comprising 128 elements (Misra et al., 2021a). This vector descriptor is instrumental in matching features between the input imagery and its related reference data, leading to the creation of dependable matched keypoints. To achieve accurate geometric correspondence between images, the geo-referenced strip image is matched with its corresponding reference strip using the SIFT algorithm.

In our methodology, we employ an approach to feature matching based on interleaved tiles designed to obtain a sufficient number of stable keypoints while reducing the execution time for

substantial image dimensions. A tile dimension of 256\*256 pixels is selected, demonstrating its optimality in attaining precise matching in strips of remote sensing images.

It has been noted that even after matching points, a considerable number of outliers persist, significantly impacting the estimation of transformation parameters. To address this, we propose a hierarchical outlier detection approach. At the first level, obvious erroneous matched points are removed, and at the second level, an effective model estimation is computed from the optimal subset of potential keypoint correspondences.

The process starts with the calculation of sets of differences are calculated in both the side-to-side and up-and-down orientations for every identified keypoint. A mode-biased approach is utilized to eliminate outliers, wherein the mode is obtained from the collection of difference image points, and a threshold limit is established from the image to identify the genuine corresponding keypoints (Kupfer et al., 2014). Experimental results suggest that establishing a standard deviation of six around the calculated mode estimate includes all the targeted corresponding points in lengthy strip remote sensing images. The keypoints pruned with a mode-biased approach are utilized as input for estimating model parameters through the Random Sample Consensus (RANSAC) algorithm.

RANSAC functions in two stages: the "Hypothesis Generation" phase, where it selects a random set of corresponding keypoints, and the "Hypothesis Evaluation" phase, where it assesses if the chosen hypothesis produces the required inlier candidate with a model error parameter within a pre-defined limit. The RANSAC method is utilized in various scenarios within remote sensing image processing, including mosaicking lunar images and matching planetary images for spatial enhancement using unpaired image frames (Yang et al., 2020 ; Misra et al., 2021b). The RANSAC algorithm comprises the following steps.

- i. Randomly choose the minimum required number of points to establish the parameters of the model.
- ii. Determine the model parameters using the selected points.
- iii. Identify the number of matched points from the entire set that align with a pre-established tolerance.
- iv. If the ratio of inliers to the overall point count exceeds a pre-established limit, reassess the parameters of the model and bring the process to a conclusion.
- v. If not, repeat steps i through iv, up to a maximum of  $N$  times.



The count of iterations, represented by  $N$ , is chosen to be sufficiently large to guarantee that at least one of the randomly chosen sets of matched samples does not contain an outlier. RANSAC employs a subset of the data rather than the entire dataset. In cases where the selected data consist entirely of inliers, they can represent a hypothesis that closely approximates the truth. This assumption necessitates that  $N$  iterations are sufficient to include all samples of inliers at least once.

$$N = \frac{\log \alpha}{\log(1-\gamma^m)} \quad (5.3)$$

where  $\alpha$  denotes the failure probability,  $m$  represents the count of corresponding keypoints needed to form a hypothesis, and  $\gamma$  is the likelihood of choosing an inlier (inliers ratio).

In the final processing stage, a segmented affine transformation model is employed to calculate affine values for distinct image segments within a strip. This procedure entails updating the geometric grid for each segment to alleviate local distortion, as outlined by Misra et al. (2019). An essential phase in the data processing includes transforming the geo-registered dataset to the map projection selected for the Indian terrain (as indicated in Table-5.1). The detailed process for fine image geo-registration is depicted in Figure-5.2. The representation of the affine transformation for a segment is given by:

$$aff(k) = \{a_{0k}, a_{1k}, a_{2k}, a_{3k}, a_{4k}, a_{5k}\} \quad (5.4)$$

where  $k$  denotes the segment number, and  $a_{ik}$  represents the affine parameters.

The adjusted geographic coordinates ( $lat_{cor}, lon_{cor}$ ) for the image point  $(x, y)$ , calculated using the input geographic coordinates  $(lat, lon)$  and the affine parameters  $a_{ik}$  for segment number  $k$ , are:

$$\begin{bmatrix} lat_{cor} \\ lon_{cor} \end{bmatrix} = \begin{bmatrix} a_{0k} & a_{1k} \\ a_{2k} & a_{3k} \end{bmatrix} \begin{bmatrix} lat \\ lon \end{bmatrix} + \begin{bmatrix} a_{4k} \\ a_{5k} \end{bmatrix} \quad (5.5)$$

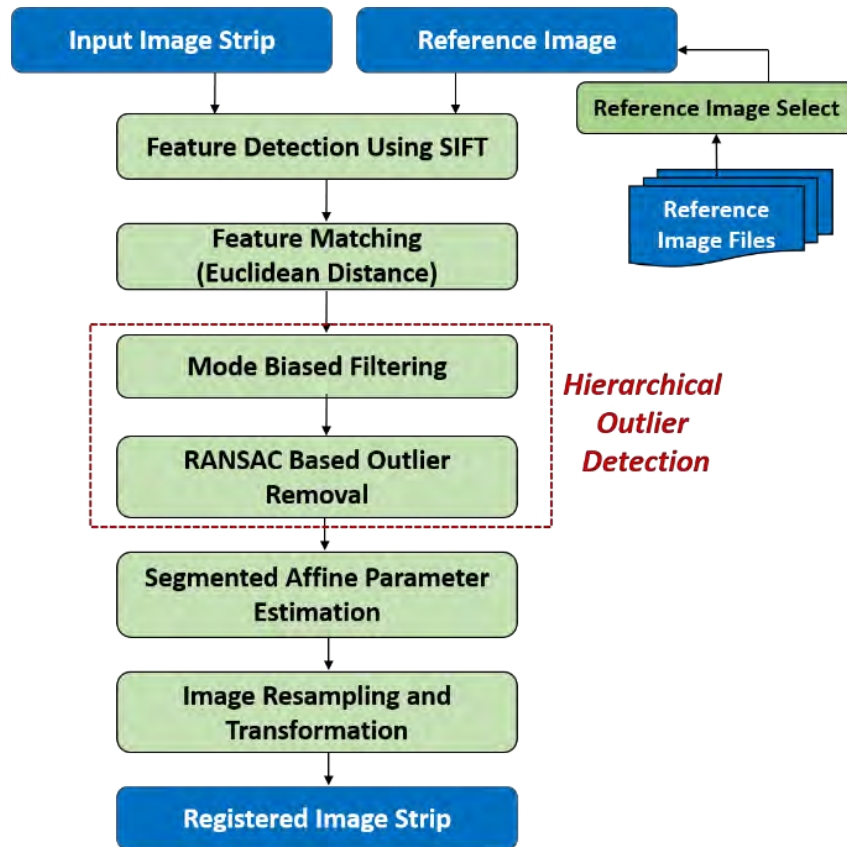


Figure-5.2. Workflow for Fine Geo-Registration of Image Strips

### C. Automated Image Stitching

In the mosaicking stage, the integration of information from overlapping pixels facilitates the seamless merging of co-registered image datasets (Edwards et al., 2011). The unified image mosaic system consolidates several co-registered image strips for the purpose of mosaicking. The automated geometric mosaic seamlessly stitches together all strips, resulting in the formation of a mosaic at the country level. In this case, radiometric normalization is excluded to reduce spectral distortion in the reflectance data of the mosaic. This is vital for purposes like geographic mapping, analyzing scientific images and making decisions (Li et al., 2019).

In the initial stage, the overlapping region among neighboring remote sensing images is identified through the intersection of polygons in spatial location space (Misra et al., 2019b). Typically, the optimal seamline, determined by the minimal intensity difference of pixels between adjacent images, should exhibit negligible geometric differences (Duplaquet, 1998). In our method, the seamline is directly detected within the image by extracting equivalent geometric

information from the overlapping region. The suggested attention technique is proficient in handling medium-resolution remote sensing imagery, where straight geometric lines are derived for seamline determination and subsequently optimized using a multi-frame joint strategy (Li et al., 2015; Li et al., 2016).

The process of blending images handles radiometric inconsistencies around the seamline, ensuring smooth transitions from one co-registered image strip to an adjacent one (Levin et al., 2004). Weighting factors are calculated by normalizing based on statistics derived from the overlap region among the images. These factors are subsequently utilized on the following image to produce a cohesive mosaic without visible seams. The blending algorithm proposed concurrently merges all co-registered images into an output map projection space. The ultimate mosaic is formed by computing the weighted average intensities of overlapping image strips.

The comprehensive image stitching approach includes the single resampling of geo-registered strip data products used as input. Figure-5.3 illustrates the developed workflow for the automatic stitching of images captured through optical remote sensing. Apart from scaled radiance image stitching, the system has the capability to stitch strips with Top-of-Atmosphere (TOA) and Surface Reflectance (SR) data to produce a reflectance mosaic data product at the country level. The suggested workflow for creating image mosaics is customized to encompass nations in the vicinity of the Indian terrain.

The mosaic system produces stitched data on a band-by-band basis, encapsulated in the Geo Tagged Image File Format (Geo-TIFF). Metadata details associated with the mosaic data products are revised for user reference. For user convenience, a defined buffer for the Indian subcontinent is set up, and the stitched data is displayed within the designated bounding corner frame. The ultimate output is placed within the designated product directory location, ready for further utilization in applications in the scientific field. The steps for the suggested system for combining images into a mosaic are delineated in Algorithm-5.1. The mosaic system created is adaptable and robust, capable of stitching together remote sensing images from other satellites, encompassing a variety of geographical regions.

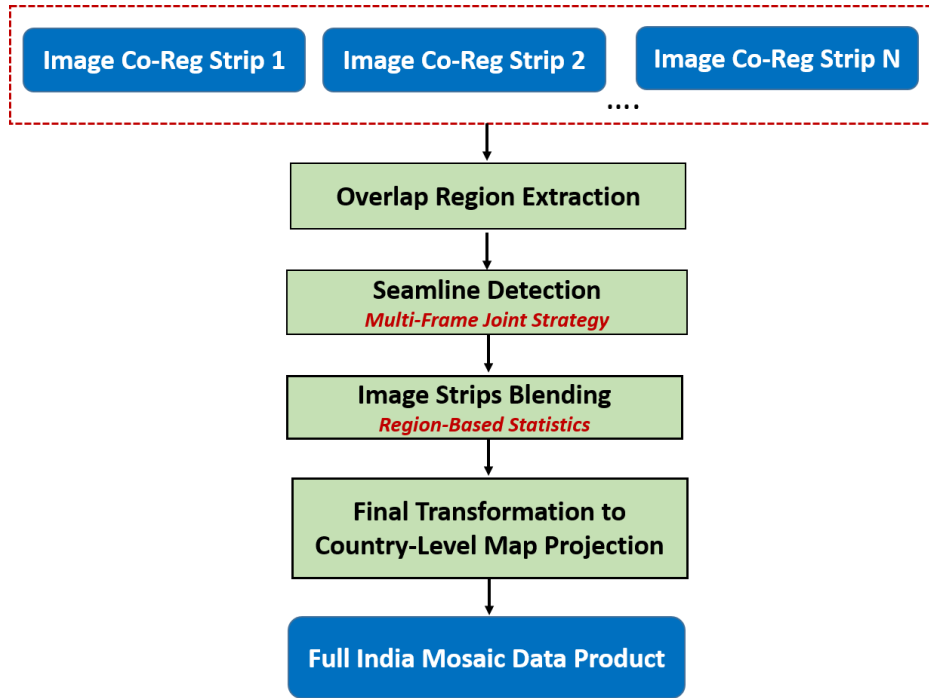


Figure-5.3. Workflow for Automatic Image Stitching

#### Algorithm-5.1: Integrated Mosaicking for Country-Level Images

**Input:** Geo-located input image strips from remote sensing  $I_1, I_2, \dots, I_n$ , where  $n$  represents the number of image strips for the mosaic.

1. Check image strip handling for all  $n$  image strips.
2. Extract reference strips  $(R_1, R_2, \dots, R_n)$ .
3. Transform reference strips  $(\overline{R_1}, \overline{R_2}, \dots, \overline{R_n})$ .
4. Register strips  $(I_1, I_2, \dots, I_n)$  using SIFT+ MB-RANSAC.
5. Estimate segmented affine parameters  $(a_{k1}, a_{k2}, a_{k3}, a_{k4}, a_{k5}, a_{k6})$ , where  $k$  is the number of segments.
6. Generate geo-registered strips  $(GRI_1, GRI_2, \dots, GRI_n)$ .
7. Detect optimal seamlines  $(SD_1, SD_2, \dots, SD_n)$ .
8. Compute region-based statistics  $(RS_1, RS_2, \dots, RS_n)$ .
9. Perform automatic image stitching to create the mosaic image

$$I_s = \prod_{i=1}^n \{\text{stitch}(GRI_i, SD_i, RS_i)\}$$

10. Re-project and transform to the country-level projection space.
11. Format the product in Geo-TIFF and update metadata.

**Output:** Mosaic data product at the country level  $I_M$

## 5.3. Experimental Results

### *A. Performance in Geo-Registering Strip Images*

To improve the assessment of the image co-registration method, intentional translation and affine errors are introduced to the image strip from Resourcesat LISS-3. The process described in above methodology section for image geo-correspondence is subsequently utilized to co-register this extensive image data. Geometric SIFT matching in conjunction with MB-RANSAC pruning, is employed in a congruence mode to achieve sub-pixel spatial correspondence. The errors intentionally introduced into the image strip are detailed in Table-5.2. These deliberately introduced errors result in a visibly varying shift in the image strip, as evident in Table-5.2.

Figure-5.4 illustrates a multi-temporal mosaic of LISS-3 images covering the expanse from Gujarat in the west to the easternmost part of India. In the mosaic created using LISS-3 images, the regions of overlap between two adjacent strips are emphasized and represented by yellow, green, and blue boxes denoting Regions of Interest (RoI) in Figure-5.4. For a more comprehensive examination of the overlap regions encompassing diverse feature targets, please refer to Figure-5.5, Figure-5.6, and Figure-5.7 in their full-resolution versions. Within the identified overlap regions, two neighboring image strips are seamlessly joined using the suggested methodology, both prior to and following geo-registration procedure.

In the overlapping region, there is a blurred appearance attributed to geometric misalignment between the strips prior to correction. Nevertheless, the correction in co-registration of image data leads to uniform geometric accuracy throughout the strips, seamlessly integrating the regions of overlap into the feature space.

Table-5.2. Introduced Errors in Input Image Strips/Scenes


	Type of Errors Induced	Error Values (in pixels)
 <p data-bbox="277 625 626 688">Error Induced in Input Image Strip</p>	Translation Error	4.2 (in horizontal direction) 9.6 (in vertical direction)
	Affine Error	aff_0 = 0.999900, aff_1 = 0.000004, aff_2 = -0.000004, aff_3 = 1.000176 aff_4 = -0.083587, aff_5 = 0.083861



Figure-5.4. Large Area Mosaic of Multi-Temporal LISS-3 and Identification of RoIs (Yellow, Green, and Blue) for Assessing Overlap Region Performance of Remote Sensing Image Strips

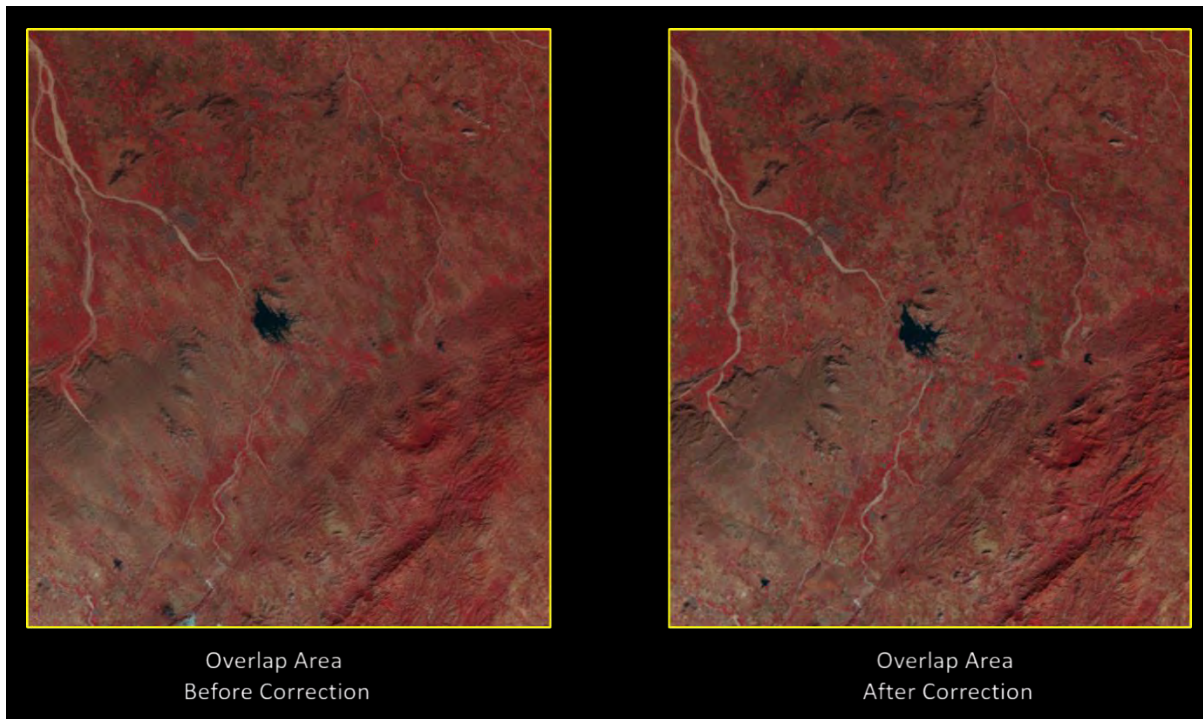


Figure-5.5. Full-Resolution View of Overlapping Region within the Yellow ROI Box (Sparse Vegetation)

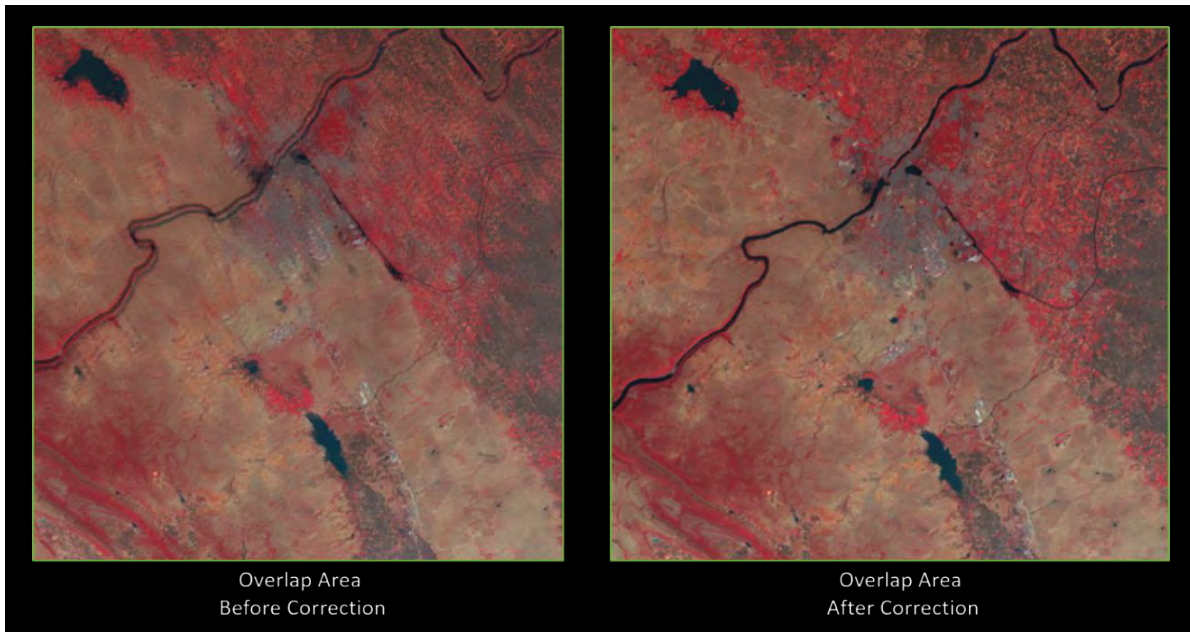


Figure-5.6. Full-Resolution View of Overlapping Region within the Green ROI Box (Urban/Water Bodies)

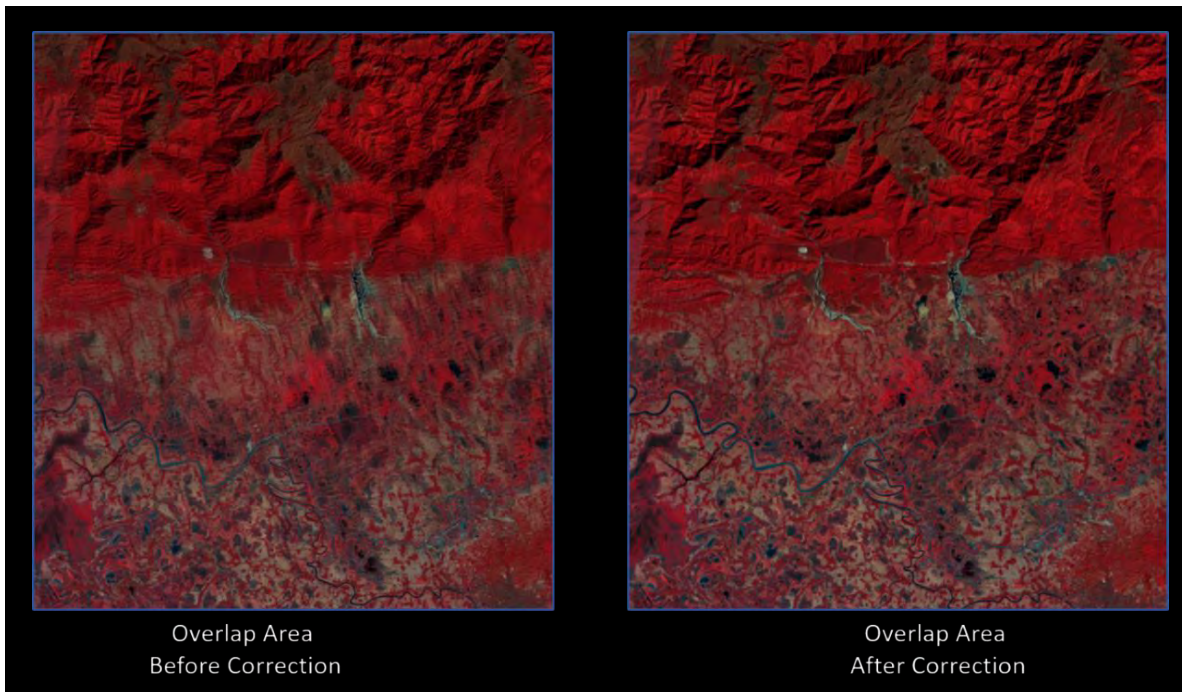


Figure-5.7. Full-resolution perspective of the overlap region within the Blue ROI Box (Hilly Region).

The Root Mean Square Error (RMSE) is calculated for various regions of interest, denoted by the yellow, green, and blue Regions of Interest (RoI) boxes in Figure-5.4, both before and after correction, to quantify the changes. The presented approach, SIFT + MB-RANSAC, is compared with SIFT + RANSAC (Misra et al., 2021b), demonstrating that the Root Mean Square Error (RMSE) achieved is superior for all three Regions of Interest (ROI) boxes when employing the proposed SIFT + MB-RANSAC image registration method shown in Table-5.3. The mean RMSE following correction using SIFT + MB-RANSAC is found to be 0.36 pixels. In our context, RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|^2} \quad (5.6)$$

where  $N$  is the total number of matched points,  $X_i$  represents the coordinates  $(x_i, y_i)$  of the reference image, and  $\hat{X}_i$  represents the coordinates of the input/output strip image.

Table-5.3. Root Mean Square Error (RMSE) in Pixel Values Prior to and Following Correction with SIFT + RANSAC and SIFT + MB-RANSAC

Region of Interest (ROI)	RMSE (in pixels) before Correction	RMSE (in pixels) after Correction using SIFT+RANSAC [36]	RMSE (in pixels) after Correction using SIFT + MB-RANSAC
Yellow ROI	8.95	0.54	<b>0.35</b>
Green ROI	8.56	0.62	<b>0.31</b>
Blue ROI	9.81	0.69	<b>0.42</b>

### B. Assessment of Reflectance in Country-Level Mosaic and Overlap Regions

Following the co-registration of image strips/scenes, we involve multiple strips in the automated image stitching process. The LISS-3 sensor on the Resourcesat satellite, which includes approximately 350 scenes, provides comprehensive coverage of India. A solitary LISS-3 strip includes numerous scenes captured on the same date, extending from top to bottom and covering a more extensive area in contrast to a single scene. In order to encompass the entire Indian landmass, thirty LISS-3 strips are geo-registered in relation to their respective reference strips, preparing the strip datasets for image mosaicking.

The process, depicted in Figure-5.2 and Figure-5.3, entails integrating multiple strips, performing strip co-registration, extracting overlap regions, identifying seamlines, estimating



blending parameters, and ultimately stitching the images together to automatically generate a comprehensive mosaic data product for India. Considering the scale of the data and the increased computational complexity, conventional step-by-step computation techniques fail to fulfill the timing specifications for mosaicking extensive image areas. To improve the efficiency of the integrated image mosaicking system in creating mosaic data products at the country level, ample parallelism is introduced, and the process is integrated into the created software library.

The image mosaicking system software is written in C/C++ and run in Linux Operating system. The hardware configuration of system used for testing comprise of Intel(R) Xeon(R) CPU E5-2967 v3 @ 2.60GHz CPU, dual socket with 14 cores per CPU and have 256 GB RAM. The comprehensive mosaic data processing from start to finish is completed within a 30-minute timeframe for the Indian landmass. In Figure-5.8(a), the complete India mosaic derived from LISS-3 sensor data is showcased, and a False Color Composite (FCC) is produced using strip datasets from the Green, Red, and Near Infrared wavelength channels. The bounding corners of the country mosaic product are standardized to ensure a consistent image size across varied acquisition scenarios. Figure-5.8(b) offers a closer view of the Gujarat region within the India Mosaic, highlighting the preservation of both radiometric and geometric consistencies in the mosaic data.

The image mosaicking system is a versatile procedure capable of merging reflectance strips to produce a mosaic of reflectance data product at the country level. An experiment was carried out to annually create mosaics of Bangladesh's landscape using reflectance images from Resourcesat-2A LISS-3. The mosaic covering the entire country across several years enables the monitoring of diverse environmental dynamics that undergo changes over time. Figure-5.9 illustrates the timeline of Bangladesh LISS-3 reflectance mosaics. The visuals offer a natural color mosaic of the land terrain in Bangladesh by combining Short Wave Infrared (SWIR), Near Infrared (NIR), and Red (R) spectral bands. The faint seams in the overlapping region of the mosaicked image arise from reflectance images captured at different seasons in multiple time periods. Utilizing regional statistics for blending guarantees a seamlessly smooth overlap region to the best possible extent.

Nevertheless, additional radiometric normalization alters the original physical measurement quantity, namely surface reflectance, and diminishes the scientific significance of conducting data analysis with mosaic data products at the country level. Figure-5.10 showcases the effectiveness of the suggested image blending at the full resolution level. The use of mean statistics for blending exposes radiometric inconsistency in the common region, as depicted in Figure-5.10(a). To achieve

a seamless mosaic outcome, radiometric balancing is implemented using region-based statistics to align the intensities of the images captured at different time points as closely as possible. Figure-5.10(b) illustrates the effectiveness of the suggested image mosaicking procedure, which minimizes variations along the seamline and seamlessly integrates the image strips.

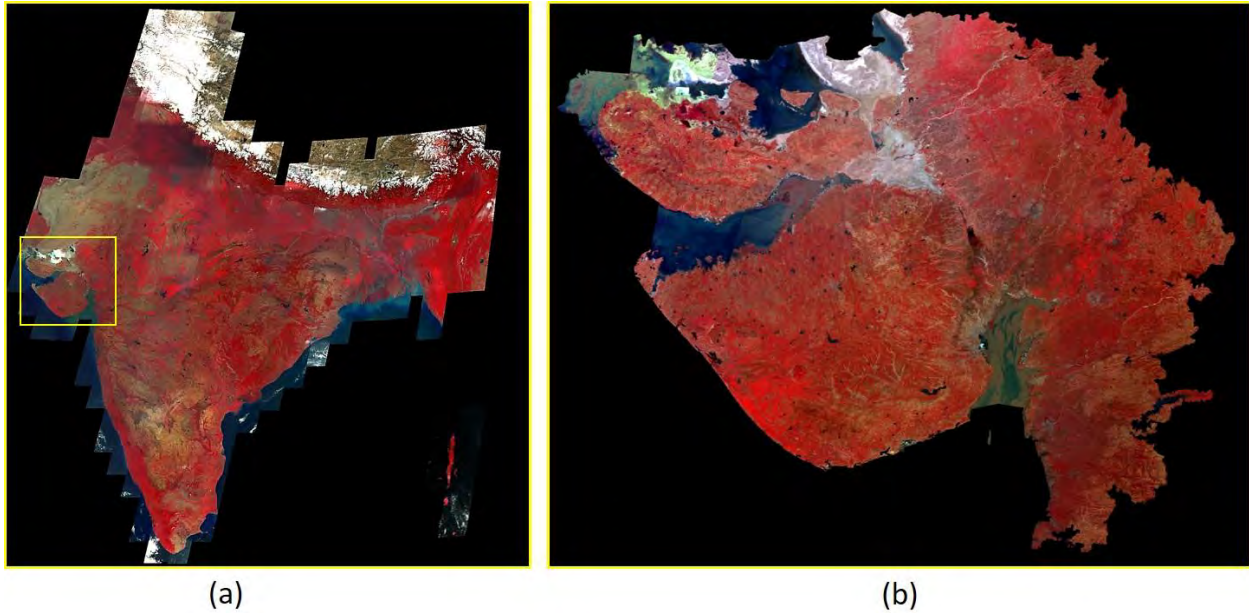


Figure-5.8. (a) Complete Mosaic of India (Image Layer: RS-2A LISS-3 False Color Composite (FCC) utilizing Near Infrared, Red, and Green Wavelength Channels) (b) Enlarged View of Gujarat Region Demonstrating a Seamless Mosaic.

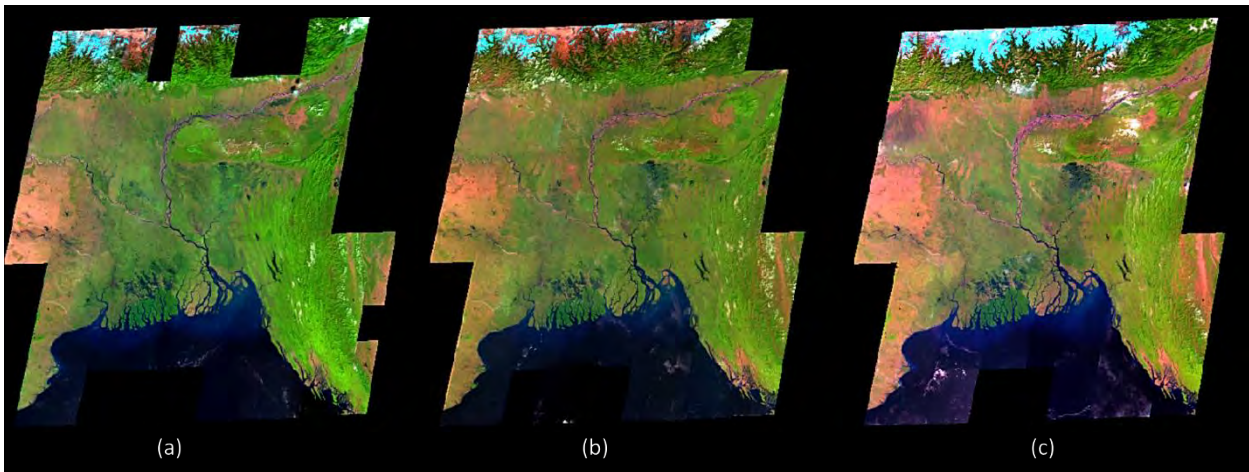


Figure-5.9. Reflectance Image Mosaic of Bangladesh at the Country Level for Various Years (Band Combination: SWIR, Near Infrared, and Red) (a) 2017 (b) 2018 (c) 2019 utilizing Resourcesat-2A LISS-3 Data

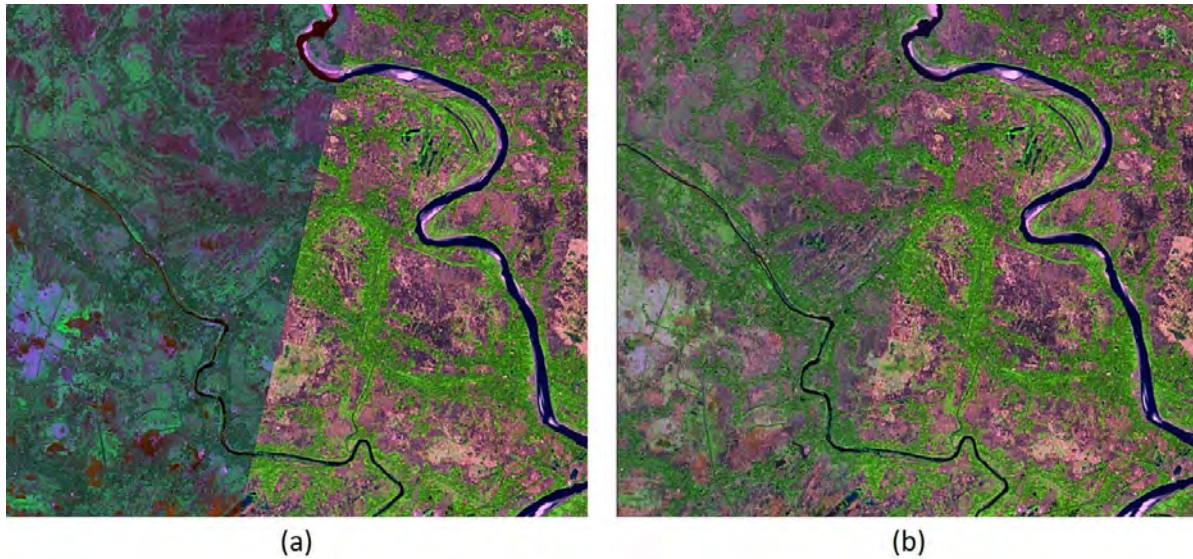


Figure-5.10. Image Blending Comparison (a) Image Blending using Mean Statistics in the Overlapping Area (b) Image blending using region-based statistics as proposed

The evaluation includes examining the alteration in reflectance between the reflectance of individual strips and the reflectance of the mosaic in the overlapping areas between two reflectance strips and spanning different feature targets. Surface reflectance data is utilized to mitigate the impact of sporadic atmospheric scattering and absorption. Radiative Transfer Modeling (RTM) using 6S (Vermote et al., 1997) effectively conducts atmospheric correction of remote sensing images employing a ML technique (Rusia et al., 2021). In Figure-5.11, spectral classes are emphasized for the evaluation of overlap performance. A compact Region of Interest (RoI) within a yellow tile box is manually selected to capture a distinct signature of an individual class. Table-5.4 displays the initial reflectance value of the strip and the predicted reflectance mosaic at different wavelength channels.

The spectral reflectance deviation is slightly more pronounced in urban and vegetation classes in contrast to the spectral class associated with water, attributed to mixed spectral signatures in urban and vegetation pixel regions. The mean deviation in spectral reflectance within mosaic reflectance values concerning the original surface reflectance measurement at various feature targets is noted to be below 0.127%. The comparison suggests that the mosaic reflectance measurement displays minimal deviation when contrasted with the input reflectance values, making the mosaic data readily applicable for large-scale earth observation applications. The Spectral Reflectance Deviation (*SRD*) is expressed as follows.

$$SRD(\%) = \left| \left( \frac{p_s - p_m}{p_s} \right) \right| * 100 \quad (5.7)$$

where  $p_s$  represents the strip reflectance measurement, and  $p_m$  represents the mosaic reflectance measurement.



Figure-5.111. Spectral Classes in the Overlapping Region for Comparing Mosaic Reflectance  
(Image Layer displayed in False Color Composite)

Table-5.4. Comparison of Strip and Mosaic Reflectance Measurements in the Overlapping  
Region

Spectral Class – Water			
Spectral Bands	Strip Reflectance	Mosaic Reflectance	Spectra Reflectance Deviation (%)
Green	0.1784	0.1783	0.056
Red	0.1638	0.1637	0.061
Near Infrared	0.1425	0.1424	0.070
Spectral Class – Vegetation			
Spectral Bands	Strip Reflectance	Mosaic Reflectance	
Green	0.1782	0.1779	0.168
Red	0.1596	0.1592	0.250
Near Infrared	0.3305	0.3304	0.030
Spectral Class – Urban			
Spectral Bands	Strip Reflectance	Mosaic Reflectance	
Green	0.2105	0.2103	0.095
Red	0.2101	0.2109	0.380
Near Infrared	0.2564	0.2563	0.039

### C. India Mosaic and Geo-Physical Product from AWiFS over a 5-Day Period

The planned integrated processing system for mosaics is specifically created to handle geo-referenced strips from Resourcesat AWiFS as its input. The process of preparing input data includes the identification of AWiFS strips captured over a span of five consecutive Julian days

starting from the commencement of the year. Subsequently, these strips undergo co-registration and stitching during the mosaic processing phase, resulting in the creation of mosaic products at the country level that encompass the entirety of the country's landmass. Figure-5.12 depicts the structure of the AWiFS mosaic covering the terrain of India. Information layers, including geo-physical parameters like NDVI, are derived from the reflectance data of the mosaic. The extracted data is then employed for monitoring natural resources and diverse operational space-based applications.

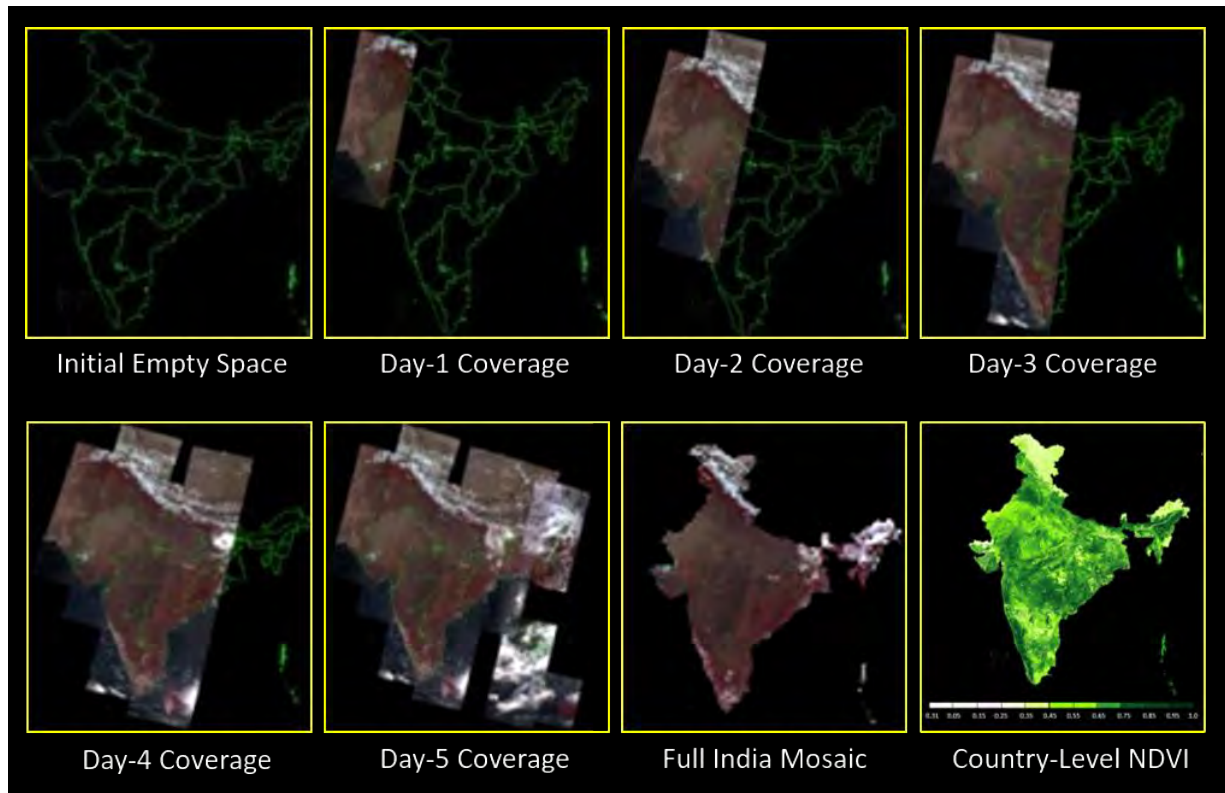


Figure-5.12. Schema of AWiFS Country-Level Mosaic

#### 5.4. Summary

The proposed integrated image mosaicking system is skilled in handling remote sensing images captured at various time intervals. It demonstrates excellence in sub-pixel-level co-registration of strips and effectively stitches multiple strips together, resulting in the creation of a comprehensive mosaic data product at the country level. Improving the correspondence between images entails aligning the input image strip with a reference database through the use of the SIFT and MB-RANSAC outlier removal technique. Utilizing a multi-frame joint approach and region-based

statistics in geographic space, automatic image stitching proves effective in producing mosaic data that is nearly seamless. There is a minimal deviation noted in the reflectance values of the mosaic across spectral classes within the overlapping region.

Experimental assessment reveals that, following correction, there is no visible blurring in the overlapping area between two strips, and the Root Mean Square Error (RMSE) estimate for strip co-registration falls within a single pixel. The developed mosaicking technique demonstrates versatility and is specifically tailored to manage optical remote sensing images captured at different times across various geographical locations. This ensures the generation of mosaic images that are both spatially continuous and of high quality. In the future, there is a possibility of improvement by integrating deep learning models into our integrated image mosaicking processing workflow to address various tasks related to remote sensing image processing. Furthermore, there are intentions to expand the developed technique for mosaicking high spatial resolution images to encompass extensive geographical areas in the upcoming period.

## 6. SPECTRA PRESERVING REMOTE SENSING IMAGE FUSION FRAMEWORK

### 6.1. Spectra Preserving Pan-Sharpener Technique

#### 6.1.1. Introduction

Our endeavor is driven by the aspiration to create a well-balanced and efficient image fusion method. This approach aims to integrate the capabilities of deep learning framework with a spatial injection relies on a robust statistical framework rooted in the principles of physics in the Variational Optimization (VO) space. The goal is not only to enhance spatial intricacies while simultaneously reducing radiometric distortion in diverse surface features sensed by low-resolution multispectral imaging camera.

Our proposed method stands out by focusing on the adaptive configuration of regularization parameters based on surface characteristics. This adaptive approach is crucial for achieving optimal fusion performance and represents an under explored area in Bayesian-based pan-sharpening methods. In this section, we present a novel image fusion technique named SPRINT, which leverages the integration of comprehensive deep edges for spatial improvement (Xie & Tu, 2015), along with terrain guidance through the Minnaert function using Digital Elevation Model (DEM) and solar angles, within a Bayesian probabilistic model (Bernardo & Smith, 2009) to ensure spectral fidelity.

The SPRINT method involves using a distinct medoid aggregate intensity, combined with a denoised panchromatic image, to extract dependable sharp edges for focused spatial attention. The denoising of the panchromatic image is achieved through the application of a Bilateral Filter (Elad, 2002) to mitigate noise in regions with uniform characteristics. Our data pre-processing system utilizes the Segmented Affine Scale Invariant Feature Transform (SA-SIFT) method proposed by Lowe in 1999 for achieving multi-sensor image co-registration with sub-pixel accuracy.

Observations indicate that the SPRINT processing workflow effectively enriches spatial information from the panchromatic image while preserving the spectral characteristics of the multispectral image. We evaluate the performance of the SPRINT fusion method using datasets from the fusion pairs of Indian Cartosat-1 (Baltsavias et al., 2007) and Resourcesat-2/2A (Pandya et al., 2013), along with IKONOS (Dial et al., 2003). Comparative assessments are made against

cutting-edge methods for image fusion. Additionally, the SPRINT fused reflectance image generated from Landsat (Roy et al., 2014) and Resourcesat closest imaging time acquisition, demonstrating close consensus with actual reflectance measurements (Vermote et al., 2016).

To evaluate the output of the SPRINT fusion, diverse feature targets including urban landscapes, various vegetation types, water bodies, and hilly terrain regions are chosen. Qualitative and quantitative evaluations are conducted, including measuring Normalized Difference Vegetation Index (NDVI) (Pettorelli et al., 2005) values at different classes. The results indicate minor deviations compared to reference values.

### **6.1.2. Methodology Developed**

#### *A. Pre-Processing Data with SA-SIFT*

The panchromatic and multispectral input images must undergo a geometric transformation to ensure they share the identical spatial resolutions and are in a shared map projection system. Achieving image fusion necessitates aligning image co-registration at a sub-pixel level for multi-sensor datasets. In our approach, we utilized a feature matching technique based on Scale Invariant Feature Transform (SIFT), combined with a segmented affine transform known as SA-SIFT (Misra et al., 2019a, Misra et al., 2021a). This methodology addresses adjusts the images for relative internal distortion and overlays them with sub-pixel precision.

Figure-2 illustrates the performance of the pair of panchromatic and multispectral images both prior to and following correction. The multispectral image is displayed in False Color Composite (FCC) by merging channels of Near-Infrared (NIR), Red (R), and Green (G) wavelengths. Significantly, the misalignment of solar panels, indicative of permanent manmade structures, becomes apparent when horizontally aligning the panchromatic and multispectral images. The error in geo-alignment has a notable impact on image fusion because of a compromised spatial-spectral relationship model.

Co-registration at the sub-pixel level emerges as a crucial preliminary data processing stage, preparing the image pair for the fusion task. Following image co-registration, precise sub-pixel spatial alignment of the panel structures is systematically accomplished at each pixel, as illustrated on the right side of Figure-6.1. To emphasize the differences, a yellow oval is superimposed on the image set prior to and following the co-registration process. The left side vividly displays feature misalignment, while the right side showcases feature alignment.



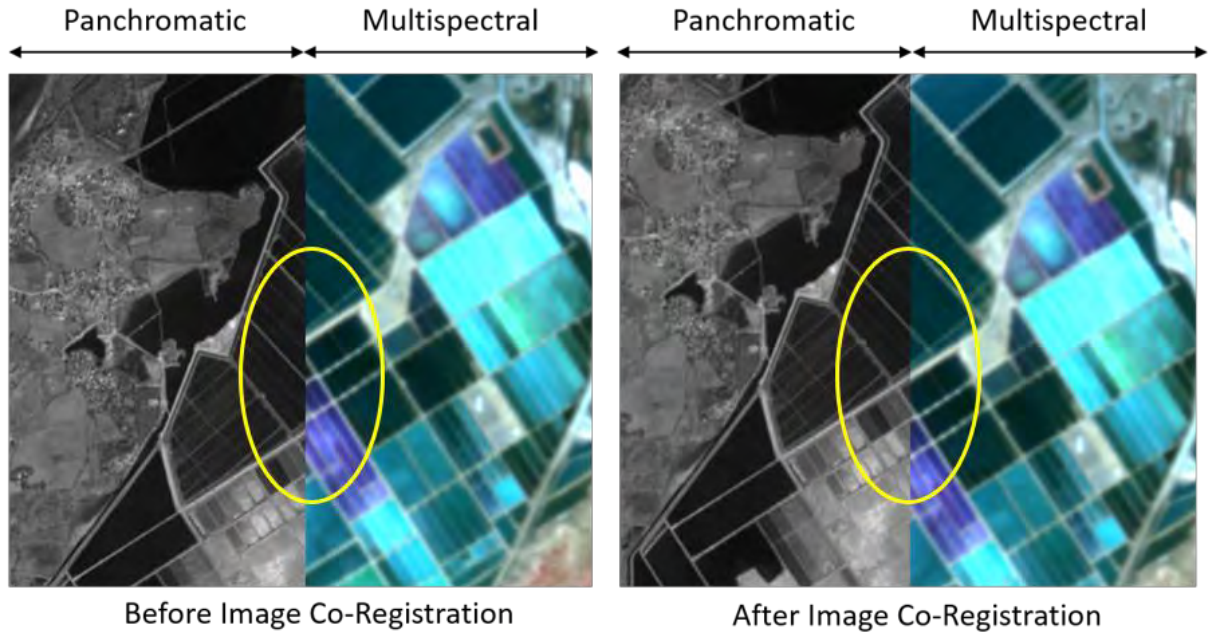


Figure-6.1. Performance of co-registering panchromatic and multispectral images.

#### B. Calculation of Medoid Intensity Matching

Medoids are objects that serve as representatives within an information index, chosen for their minimal dissimilarities to each multispectral pixel element within a cluster. They find particular utility when neither a mean nor a centroid accurately captures the essence of multivariate datasets. In the realm of remote sensing, medoids hold potential for generating composite geophysical parameters through remote sensing observations over a period of time (Toumisto, 2016).

In our methodology, we have devised an approach to derive the intensity component from multispectral images through the utilization of medoids. This method proves to be a more reliable measure of the intensity brightness value for the particular pixel location being examined (Misra et al., 2021b). The computation of medoid intensity from multispectral data is defined as follows:

$$m_k = \underset{\mu_k}{\operatorname{argmin}} \left( \sum_{i=1}^n \|x_{ik} - \mu_k\| \right) \quad (6.1)$$

Let  $m_k$  represent the medoid intensity value at a specific pixel location  $k$ ,  $\mu_k$  represent the average intensity value at a specific pixel location  $k$ ,  $n$  represent the number of spectral bands, and  $x_{ik}$  denote the intensity value of pixel at location  $k$  for spectral band  $i$ .

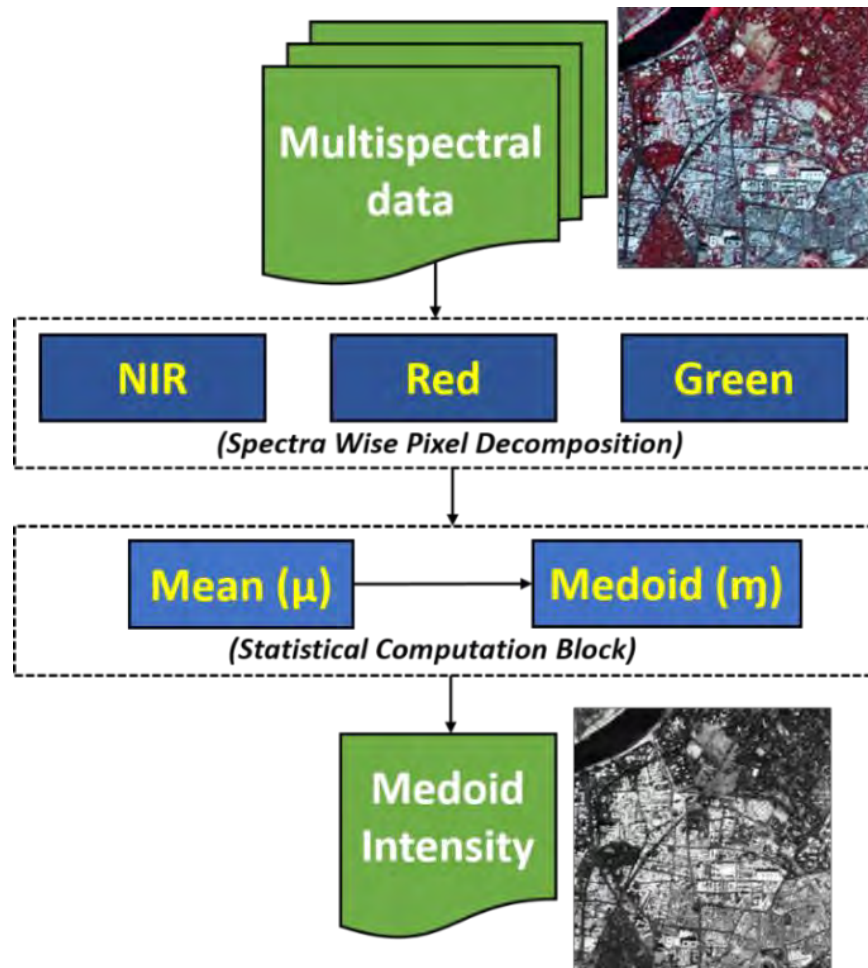


Figure-6.2. Computation of medoids from a multispectral image.

Figure-6.2 illustrates the steps involved in medoid computation visually. The multispectral information in a band-sequential format is separated into distinct wavelength channels to facilitate pixel decomposition based on the spectrum. In our context, the multispectral image includes elements from the Near-Infrared (NIR), Red (R), and Green (G) wavelength channels. These components are subsequently transformed into the Red, Green, and Blue color space model to enable False Color Composite (FCC) visualization.

The average intensity is computed for every pixel location coordinate, and it is then employed to determine the medoid value on a pixel-wise basis using Equation (6.1). The intensity image based on medoid values undergoes histogram matching with the panchromatic image, ensuring the selection of the most significant and accurate values for rendering the intensity component image. This method stands out for its superiority in providing a clear and precise value for each pixel.

### C. Panchromatic Denoising through Bilateral Filtering

The bilateral filter is a non-linear method used to maintain edges and reduce noise in uniform areas of images (Zhang & Allebach, 2008). This approach mainly links neighboring pixels by considering their spatial characteristics and photometric similarity. It assigns pixel values that adhere to closeness in both spatial and photometric domains. The bilateral filter functions on the intensity of individual pixels, substituting it with the weighted average intensity values of nearby pixels. In our application, the weight is calculated based on a Gaussian distribution, effectively maintaining sharp edges in the panchromatic image while reducing noise in uniform regions. The bilateral filter (BF) is defined by the following equation:

$$BF(I_p) = 1/W_p \sum_{I_q \in t} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(\|I_p - I_q\|) \quad (6.2)$$

Here,  $W_p$  represents the normalization factor,  $p$  denotes the current pixel position in the image,  $q$  signifies the pixel position that affects the given pixel  $p$ ,  $I_p$  represents the value of intensity at the pixel location  $p$ ,  $I_q$  represents the value of intensity at the pixel location  $q$ ,  $t$  is the central point of the window surrounding the pixel position  $p$ ,  $G_{\sigma_s}$  stands for the spatial Gaussian kernel for smoothing differences in pixel coordinates, and  $G_{\sigma_r}$  represents the Gaussian kernel in the range for smoothing variations in image intensity values.

The normalization factor  $W_p$  is calculated based on both the spatial proximity (from the spatial kernel) and the difference in intensity (from the range kernel). As depicted in Figure-6.3, noise is efficiently mitigated in uniform regions, while preserving the edge and texture details in the panchromatic image. In the zoom view of 4X, specific area demonstrates the noise reduction achieved through the bilateral filter technique in the input panchromatic image. Table-6.1 showcases the Signal-to-Noise Ratio (SNR) (Atkinson et al., 2005) and the estimated noise level  $\sigma$  for both the input and output panchromatic images (Chen et al., 2015). A higher Signal-to-Noise Ratio (SNR) and lower levels of noise estimation distinctly suggest that the output from the bilateral filter more effectively preserves the signal information.

Table 6.1. Comparison of SNR (Signal-to-Noise Ratio) and Noise Level

Image Type	Noise Estimation Level $\sigma$	SNR
Original Image	4.8251	2.3132
Filter Image	4.6252	2.4415

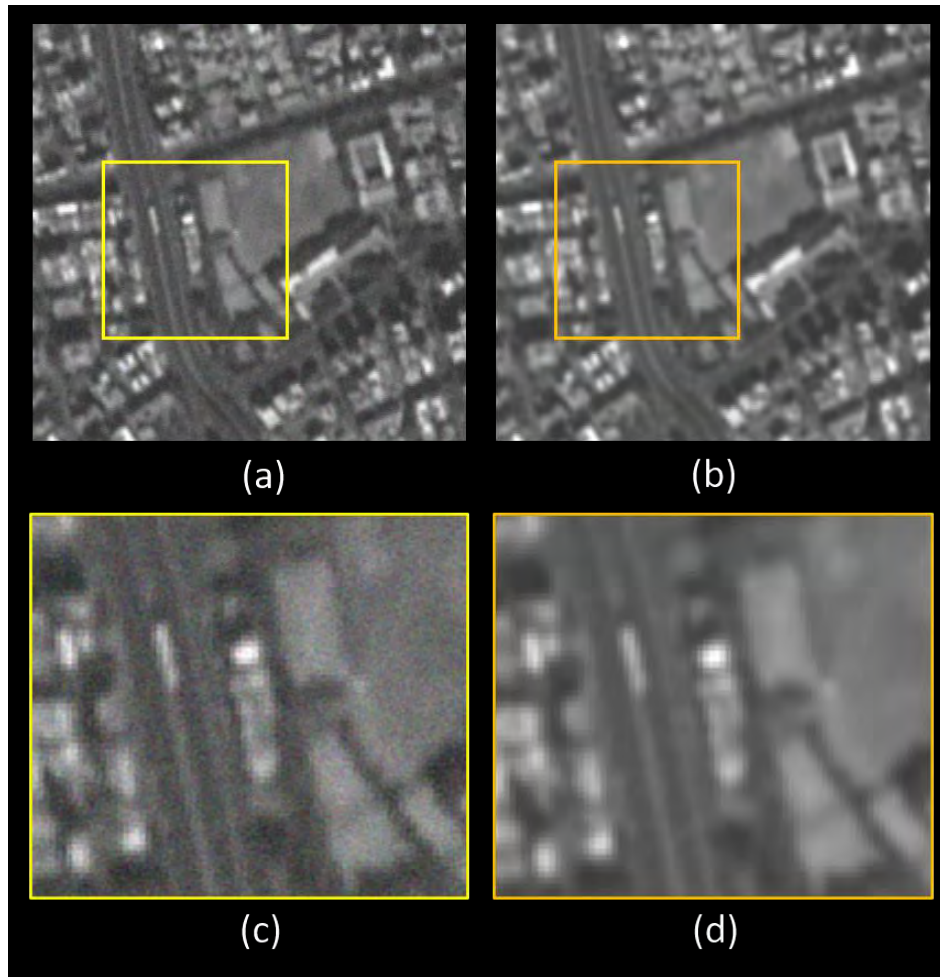


Figure-6.3. Performance of the Bilateral Filter (a) Original Panchromatic Image (b) Output from the Bilateral Filter (c) Zoomed-in View (4X) of the Original Panchromatic Region (d) Zoomed-in View (4X) of the Denoised Region after Applying Bilateral Filter

#### *D. Holistic Nested Edge Detection (HNED) for Remote Sensing Images*

The extraction of edge information is vital as it guides focus towards significant regions within an image (Torre & Poggio, 1986). Detecting edges in satellite images is a vital characteristic with applications in remote sensing tasks like registration of images, detection of changes, and merging of images. In the process of pan-sharpening, the edges recognized in the panchromatic image are pivotal in the spatial attention mechanism, directing focus to inject specific details in a controlled manner.

HNED (Xie & Tu, 2015) is a deep convolutional network designed based on the architecture of VGG16 (Simonyan & Zisserman, 2014). HNED excels in reliably detecting edges in an image due to its holistic structure. The goal of HNED is to train and predict edges in a manner where the

emphasis is on the generated and iteratively refined edge maps, treated as side outputs in an image-to-image fashion. The effectiveness of HNED has been demonstrated on natural images, surpassing classical edge detectors like the Canny operator (Canny, 1986).

Each side output layer  $l_{side}$  is linked with a classifier, and to tackle the imbalance between edge and non-edge pixels, a class balancing weight  $\delta$  is introduced. The cross-entropy function with class balancing can be expressed as follows.

$$l_{side}^{(m)}(W, w^{(m)}) = -\delta \sum_{j \in Y^+} \log Pr(y_j = 1 | X; W, w^{(m)}) - (1 - \delta) \sum_{j \in Y^-} \log Pr(y_j = 0 | X; W, w^{(m)}) \quad (6.3)$$

In this context,  $X$  represents the original image,  $Y$  is the edge map,  $j$  denotes the pixel location,  $m$  signifies the side output layer,  $Pr$  is calculated using the sigmoid activation function at the pixel level  $j$ ,  $W$  encompasses all conventional network parameters, and  $w^{(m)}$  represents the weights of the layer  $m$ .

For the direct fusion of predictions from the side output layers, a fusion layer with weighting is introduced, and the loss function at the fusion layer  $L_{fuse}$  can be expressed as:

$$L_{fuse}(W, w, h) = Dist(Y, Y_{fuse}) \quad (6.4)$$

where  $Y_{fuse} = \sigma(\sum_{m=1}^M h_m A_{side})$ ,  $h_m$  is the fusion weight and  $A_{side}$  are the activations originating from the side output of the layer  $m$ ,  $Dist$  is the the distance between the fused predictions and the ground truth label map

The minimization of the loss is achieved through the standard approach employing backpropagation stochastic gradient technique, which can be articulated as:

$$(W, w, h) = argmin(L_{side}(W, w) + L_{fuse}(W, w, h)) \quad (6.5)$$

During the testing phase, when provided with an image  $X$ , we can acquire predictions of edge maps from both the side output layers and the fusion layer with weighting.

$$(Y_{fuse}, Y_{side}^1, \dots, Y_{side}^M) = CNN \quad (6.6)$$

The ultimate consolidated result, denoted as  $Y_{HED}$ , can be achieved through additional aggregation of these generated edge maps. The formulation of  $Y_{HED}$  is as follows:

$$Y_{HED} = Average(Y_{fuse}, Y_{side}^1, \dots, Y_{side}^M) \quad (6.7)$$

HNED is not extensively utilized in the realm of remote sensing images because the pre-trained model is trained on BSDS datasets (Martin et al., 2001). The training datasets are created by extracting patches from samples of panchromatic images that encompass a range of spectral

signatures. To tailor HNED for remote sensing applications, we initially performed semantic segmentation on denoised panchromatic patches. The criterion set for semantic segmentation is that if any of the eight pixels surrounding the pixel in question differs from it, it is labeled as an edge pixel point; otherwise, it is classified as a non-edge point. The automated segmentation process involves a notable inclusion of non-edge pixels in the segmented output. This is later rectified through manual annotation by human intervention to produce the ultimate ground truth map.

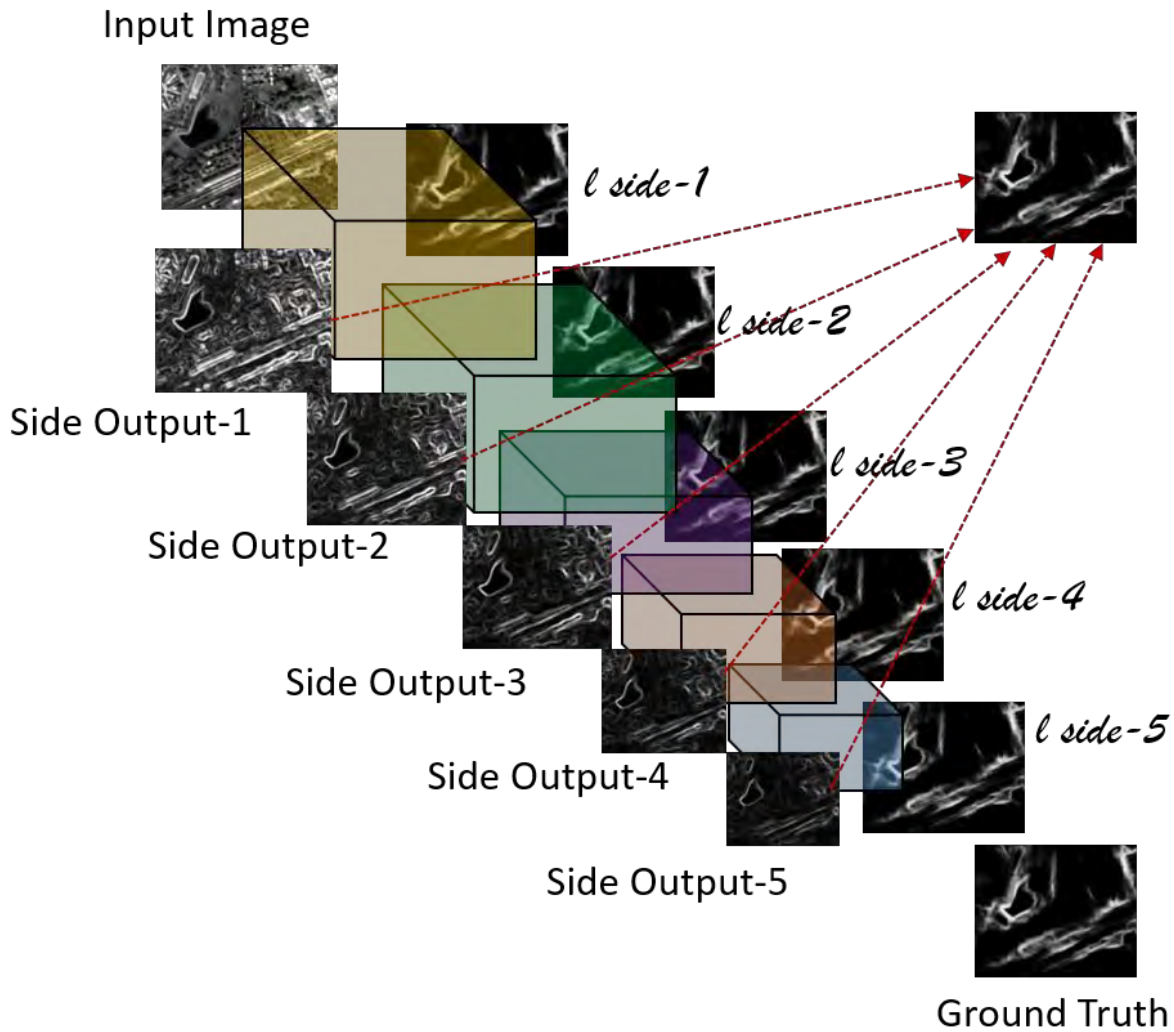


Figure- 6.4. Architecture of the The Holistic Nested Edge Detection Network implemented on an image with medoid intensity matching.

In the training stage, the pre-trained model is enhanced with features that are specific to remote sensing imagery. The versatility of the HNED framework is demonstrated in its ability to identify crisp and prominent edges in a variety of high-resolution remote sensing images. The precision of accurate edge information directly influences fusion performance and requires particular attention in tasks related to image merging. HNED serves as a resilient multi-scale deep feature learning network for precise edge detection, overcoming the drawbacks of traditional methods where the probability of detecting false positive edge information is elevated.

Figure-6.4 illustrates the application of applying the HNED network to a satellite image. Within the HNED framework, the leading edge map captures finer spatial details, while the trailing edge map encapsulates improved semantic information. The resulting edge map is an amalgamation of both intricate spatial details and enriched semantic information. The deep edge map, identified from the image with matched medoid intensity, provides information about regions where extra spatial attention is required during the fusion of multisensor remote sensing images.

#### *E. Optimized Weight Determination using the Minnaert Function*

The Minnaert function was originally employed to characterize the non-Lambertian surface properties of the moon (Minnaert, 1941). Recovering the geometry and radiance properties of a non-Lambertian surface poses challenges, yet it is crucial for modeling the terrain's topography. The function calculates the radiance factor of the surface based on the phase angle and geometric location. Minnaert introduces a parameter  $k$ , denoted as the Minnaert coefficient, to ascertain the surface characteristics of the celestial body. The Radiance Factor  $RF$  is expressed as:

$$RF = \frac{I}{\pi F} = A_m \mu_o^k \mu^{k-1} \quad (6.8)$$

$$\mu_o = \cos\varphi \cos(\alpha - \omega) \quad (6.9)$$

$$\mu = \cos\varphi \cos\omega \quad (6.10)$$

where  $A_m$  represents the Minnaert Albedo,

$k$  is the Minnaert Constant,

$I$  denotes Scattered Radiance,

$\pi F$  represents Incidence Radiance,

$\varphi$  stands for Geographic Latitude,

$\omega$  represents Geographic Longitude,

$\alpha$  is the Phase Angle, and

$k$  signifies the Minnaert Parameter.

To assess planetary body surface roughness, the semi-empirical Equation (6.8) can be rephrased using the Minnaert coefficient parameter to determine the normalized radiance at every pixel in the image.

$$\rho_H = \rho_T \frac{(\cos \lambda)^{-k+1}}{(\cos \beta)^k} \quad (6.11)$$

where  $\rho_H$  represents the radiance of a horizontal surface,

$\rho_T$  denotes the radiance of an inclined terrain,

$\beta$  stands for the solar illumination angle,

$\lambda$  represents the terrain slope, and

$k$  is the Minnaert Parameter.

The local solar illumination angle  $\beta$  can be computed as follows:

$$\cos \beta = \cos \lambda \cos \Phi + \sin \lambda \sin \Phi \cos(\Theta_a - \Theta_b) \quad (6.12)$$

where  $\lambda$  represents the terrain slope,

$\Phi$  is the solar zenith angle,

$\Theta_a$  is the solar azimuth, and  $\Theta_b$  is the topographic azimuth (aspect angle).

The various angles used in the computation of the solar illumination angle are visually presented in Figure-6.5 (a) to enhance representation and analysis. Figure-6.5 (b) shows an example of a Digital Elevation Model (DEM) used for computing terrain angles. In Figure-6.5 (c), the surface normal vector establishes the solar illumination angle at the local surface, playing a vital role in shaping surface characteristics and profiles.

The Minnaert model depends on Digital Elevation Model (DEM) data to obtain topographic information. DEM records the elevation details of a particular geographic area, supplying essential information for a range of Earth observation applications. Techniques for modeling illumination conditions require a corresponding DEM with a spatial resolution equivalent to that of the remote sensing image. A DEM is fundamentally necessary for calculating terrain slope and aspect angle to determine the incident angle (depicted as  $\beta$  in Figure-6.5 (a)), expressed as the angle between the ground's normal vector and the rays of the sun. The calculation of the surface normal involves using terrain angles as an input parameter, aiding in determining solar illumination angles at every pixel location in the image. In our scenario, we employed the Carto DEM (Muralikrishnan et al., 2013), which comes with a pixel spacing of 10 meters.



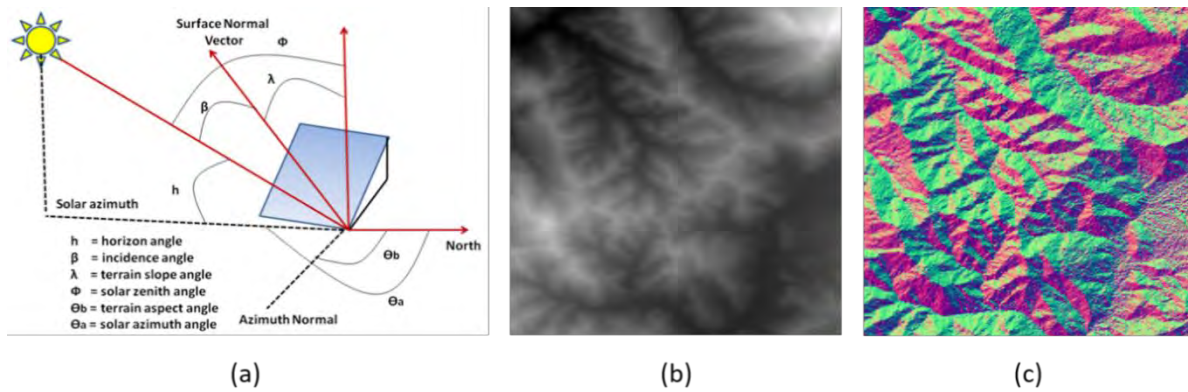


Figure-6.5. (a) Angles used in the calculations for solar illumination, (b) Digital Elevation Model (DEM), and (c) Surface normal derived from the DEM

The term  $k$  effectively characterizes non-Lambertian behavior. When  $k = 1$ , it signifies a surface that is a perfect Lambertian reflector. The computation of the  $k$  value involves linearizing Equation (6.12) by taking the logarithm.

$$\log(\rho_T \cos \lambda) = \log \rho_H + k \log(\cos \lambda \cos \beta) \quad (6.13)$$

The  $k$  value is calculated from Equation (6.13) through a simple linear regression process. The  $k$  value depends on the characteristics of the surface, the topographic factor, and the wavelength. In areas where incident angles are significant, the  $k$  value can avoid division by small values, enhancing the spatial-spectral impact in the resulting fused image. The  $k$  value ranges from 0 to 1, with elevated  $k$  values for flat areas and lower  $k$  values for steeper slopes. A reduced  $k$  value of emphasizes topographically steep regions, where slopes are perceived akin to edges in the image. For accurate identification of the surface type in image fusion, the Minnaert parameter is calculated on a pixel-by-pixel basis by establishing the correlation between Minnaert coefficients and topographic slopes.

The SPRINT fusion framework utilizes the Minnaert parameter  $k$  calculated at every pixel position in the panchromatic/multispectral image pair, with the denoised panchromatic image serving as input (Ge et al., 2008). This parameter assists in forecasting the per-pixel weight of an image based on its roughness, directing the image fusion process to enhance spatial injection in accordance with the spectral characteristics. In our case,  $k$  serves as a regularization parameter that directly evaluates the spatial-spectral relationship between the panchromatic and multispectral images. As a result, greater emphasis is placed on edge and high-texture information, with less weight assigned to flat terrain in the panchromatic data.

$$w_i \propto \frac{1}{k_i} \quad (6.14)$$

The weight  $w_i$  at each pixel position  $i$  is inversely related to the Minnaert parameter  $k_i$  determined at that particular pixel location, obtained from the denoised panchromatic image.

#### *F. Maintaining Spectral Integrity via Bayesian Probabilistic Fusion*

Developing a well-balanced and resilient image fusion method that can incorporate spatial details in a controlled manner without causing substantial distortion in the spectral signature of multispectral images poses a formidable challenge. Existing methods, such as Component Substitution (CS) based approaches like adaptive IHS with a guided filter (Yang et al., 2016), achieve spatial goals but experience spectral distortions in the resultant merged data product. Conversely, techniques based on Multi-Resolution Analysis (MRA) such as SFIM (Liu, 2000) maintain spectral information but do not fully achieve the intended spatial enhancement. While Deep Convolutional Neural Networks (CNN) have the potential to generate improved fusion output, they require specific tailoring for the datasets in remote sensing being modeled.

To address this challenge, our proposed methodology introduces a Minnaert-guided Bayesian probabilistic approach. The Bayesian theoretical model addresses the issue of image fusion within a well-defined probabilistic framework (Bernardo & Smith, 2009). The objective of Bayesian fusion with spectrum preservation, incorporating Minnaert guidance and spatial attention from the deep edge map, is to seamlessly merge panchromatic and multispectral images, guaranteeing adaptability, resilience, and efficiency.

The formulation of Bayesian probabilistic fusion is grounded in the variables of interest, represented as the vector, which cannot be directly observed. These variables are connected to observable variables  $Y$  through a functional model:

$$Y = g(Z) + E \quad (6.15)$$

where  $g(Z)$  is a function set,

$E$  is a vector of random errors that are statistically independent of  $Z$

In accordance with probability calculus, the Probability Density Function (PDF) of vector  $Z$ , given the observed variables, can be expressed as:

$$f\left(\frac{z}{y}\right) \propto f_z(z)f_E(y - g(z)) \quad (6.16)$$

$f_z(\cdot)$  is prior PDF of  $Z$ ,  $f_E(\cdot)$  is prior PDF of random error  $E$

In our context,  $Z$  represents the merged pixel achieved by combining the multispectral pixel with the high spatial resolution panchromatic pixel. Additionally,  $Y$  comprises a collection of observed variables and can be presented as:

$$Y = (Y_S, Y_P, Y_H) \quad (6.17)$$

$Y_S$  represents the pixel in the multispectral band of the image with low spatial resolution,

$Y_P$  represents the pixel in the panchromatic band of the image with high spatial resolution,

$Y_H$  represents the pixel representing holistic nested deep edge information from the medoid intensity-matched image.

The error term  $E$ , representing a vector of errors originating from the panchromatic, multispectral, and holistic nested edge images, can be formulated as

$$E = (E_S, E_P, E_H) \quad (6.18)$$

Under zero mean condition,  $Y_S = g_S(Z) + E_S$ ,  $Y_P = g_P(Z) + E_P$  and  $Y_H = g_H(Z) + E_H$ . Assuming independence among the error terms, Equation (6.16) can be expressed as:

$$f\left(\frac{Z}{Y_S, Y_P, Y_H}\right) \propto f_Z(Z) * f_{E_S}(y_S - g_S(Z)) * f_{E_P}(y_P - g_P(Z)) * f_{E_H}(y_H - g_H(Z)) \quad (6.19)$$

Each observed variable  $Y_S, Y_P, Y_H$  contains quality of information regarding the variable  $Z$ , and it can be assigned a weight to generate

$$f\left(\frac{Z}{Y_S, Y_P, Y_H}\right) \propto f_Z(Z) * f_{E_S}(y_S - g_S(Z))^{2(1-w)} * f_{E_P}(y_P - g_P(Z))^{2w} * f_{E_H}(y_H - g_H(Z))^w \quad (6.20)$$

In this context, the parameter  $w$  acts as the weight that adjusts spatial resolution details, potentially sacrificing spectral resolution information. The holistic nested edge pixel is also assigned a weight to accentuate edge pixels derived from the medoid intensity-matched image. Spatial attention is incorporated into the SPRINT workflow through the identification of deep edges using HNED (Xie & Tu, 2015). It has been noted that HNED precisely identifies significant edges in the medoid intensity-matched image while neglecting flat regions.

The weight is calculated by deciphering a Minnaert coefficient image on a pixel-wise basis, as outlined in subsection *E*. The Minnaert function generates coefficients depending on the terrain type, and the weight value adaptively derived at each pixel forms the core of our Bayesian fusion approach. The weight value, influenced by the Minnaert coefficient at each pixel position, accurately signifies surface properties, such as steep slopes or uniform regions. The adaptable weight, utilized in the Bayesian probabilistic model alongside the spatial attention of the holistic

deep edge map, has the potential to improve and precisely retain spectral characteristics in the fused image.

The devised approach has been translated into a resilient data processing pipeline termed as SPRINT. The initial panchromatic (PAN) image undergoes bilateral filtering to produce a denoised PAN. The multispectral image is aligned with the PAN through co-registration, setting the stage for its passage through the SPRINT network. The multispectral image is resampled using the nearest neighbor method for spectral consistency with optimal spatial injection, resulting in a multispectral interpolated image. The interpolated multispectral image acts as the input to the Bayesian probabilistic model, in conjunction with the PAN and multispectral image pair.

The intensity component is produced through medoid aggregation and histogram matching with the denoised PAN, yielding a matched intensity image. The salient edges are identified from the derived matched intensity component using HNED. The Minnaert function calculates the surface roughness parameter by relying on the Digital Elevation Model (DEM) and solar angles. The Minnaert parameter is determined at each pixel, directing the Bayesian model to merge PAN and multispectral images in a regulated manner based on the feature target.

The Bayesian probabilistic fusion method, focused on preserving spectra, integrates essential inputs, along with the holistic deep edge map for spatial attention and the estimated Minnaert parameter value at each pixel in the image. This procedure anticipates the creation of a fused image with optimal spatial and spectral resolution. The innovative SPRINT framework for combining PAN with multispectral images is illustrated in Figure-6.6, and the procedural steps for generating the fused image are delineated in Algorithm-6.1.

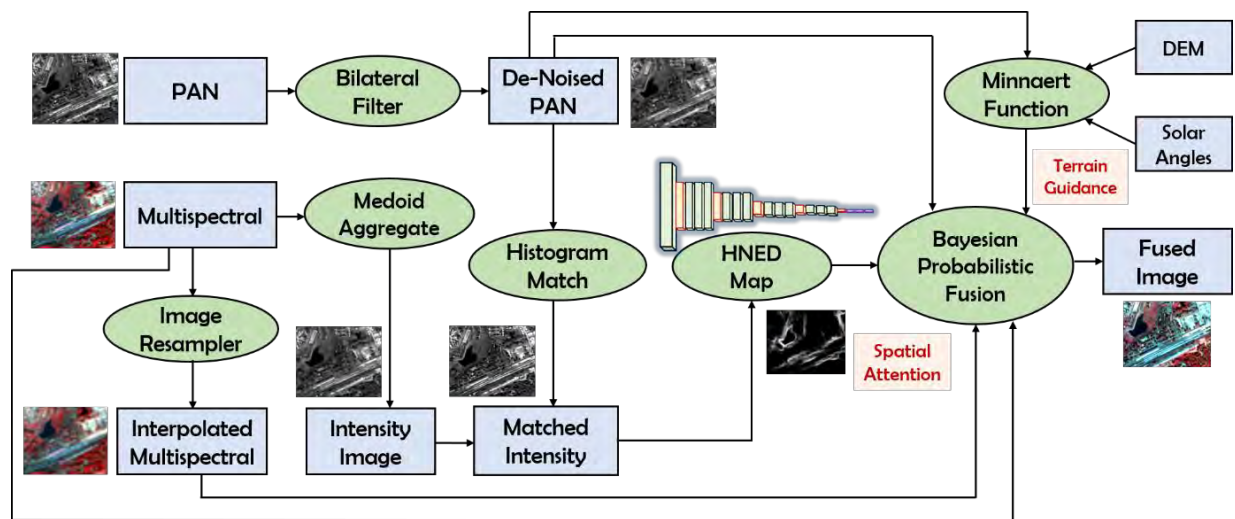


Figure-6.6. Processing Workflow of SPRINT

### Algorithm 6.1: SPRINT

**Input:** Low Resolution multispectral image  $I_{MS}$  and corresponding PAN image  $I_p$

**Output:** Fused Image  $I_F$

**Begin**

- i.  $I_{MS}$  geo-aligned with respect to  $I_p$ .  
$$I_{GMS} = SA\_SIFT\_REG(I_{MS}, I_p)$$
- ii.  $I_p$  de-noised using Bilateral Filter.  
$$I_{DN} = BF(I_p)$$
- iii. Interpolated  $I_{GMS}$  for spectral consistency.  
$$I_{\hat{G}MS} = resample\_nn(I_{GMS})$$
- iv. Medoid Intensity generation.  
$$I_{MI} = medoid(I_{MS}, I_{DN})$$
- v. Deep edge generation.  
$$I_{DE} = HNED(I_{MI})$$
- vi. DEM Extraction.  
$$I_{dem} = extract(I_{DN}, DEM, corners)$$
- vii. Minnaert Function for normalized radiance.  
$$I_{norm} = minnaert(I_{DN}, I_{dem}, S_{angles})$$
- viii. Pixel wise Minnaert Parameter.  
$$k = regress(I_{norm}, I_{DN}, \lambda, \beta)$$
- ix. Spectra Preserving Fusion.  
$$I_F = fusion(I_{DN}, I_{GMS}, I_{\hat{G}MS}, I_{DE}, k)$$

**End**

### 6.1.3. Experimental Results

The evaluation of the SPRINT methodology involves the analysis of various earth observation satellite datasets, such as Resourcesat, Cartosat, Landsat, and IKONOS remote sensing images. The LISS-4 sensors on-board Resourcesat comprise three multispectral wavelength channels (Near-Infrared, Red, and Green), encompassing a 70 km swath. They provide a spatial resolution of 5.0 meters at nadir and possess a radiometric resolution of 10 bits. LISS-3 consists of four multispectral wavelength channels, with the inclusion of an extra Short-Wave Infrared (SWIR) band. Each band possesses a spatial resolution of 24 meters, equivalent radiometric resolution to LISS-4, but encompasses a broader swath of 140 kilometers.

The input for pan-sharpening in the SPRINT network comprises either the Cartosat Panchromatic (PAN) image with a spatial resolution of 2.5 meters, the Landsat OLI-PAN band with a 15-meter spatial resolution, or the nearest imaging acquisition of Resourcesat multispectral image. IKONOS supplies a simultaneous PAN and multispectral image of the identical area,

creating an optimal pair for the process of pan-sharpening. The datasets employed for fusion consist of physical measurements represented in scaled radiance, often referred to as digital number measurements. These satellite pictures include cityscapes, thick greenery, bodies of water, human-made constructions, and hilly landscapes, revealing varied spectral characteristics and enabling a thorough assessment of SPRINT effectiveness. The details of the image fusion pairs used in our experiments are provided in Table-6.2.

Table-6.2. Metadata details for the panchromatic/multispectral image pairs utilized in the fusion evaluation

Pair No.	Regions Covered	Satellite/Sensor	Spatial Resolution
Fusion Pair No. 1	Jabalpur, India (Urban)	Resourcesat-2/L4 Cartosat-1/PAN	5.0 m 2.5 m
Fusion Pair No. 2	Indore, India (Stadium)	IKONOS/Multispectral IKONOS/PAN	4.0 m 1.0 m
Fusion Pair No. 3	Gandhinagar, India (Convention Centre)	Resourcesat-2A/L4 Cartosat-1/PAN	5.0 m 2.5 m
Fusion Pair No. 4	Mithapur, India (Solar Power Plant)	Resourcesat-2/L4 Cartosat-1/PAN	5.0 m 2.5 m
Fusion Pair No. 5	Tehri Dam, India (Hilly Terrain)	Resourcesat-2/L4 Cartosat-1/PAN	5.0 m 2.5 m
Fusion Pair No. 6	Kolkata, India (Urban)	Resourcesat-2A/L3 Landsat-8/OLI-PAN	24.0 m 15.0 m
Fusion Pair No. 7	Sundarbans, India (Dense Vegetation)	Resourecsat-2A/L3 Landsat-8/OLI-PAN	24.0 m 15.0 m

#### A. Visual Quality Evaluation

The amalgamation of the Cartosat-1 PAN image and the Resourcesat-2 LISS-4 multispectral image serves as a reliable benchmark for evaluating the quality of the fused data output. The several remote sensing image fusion methods are examined and contrasted with SPRINT to assess visual quality. The seven image fusion methods, including ACS (Aiazzi et al., 2007), AIHS (Yang et al., 2016), APCA (Shah et al., 2008), SFIM (Liu, 2000), HQB (Wang et al., 2018), PNN (Masi et al., 2016), and Pan-GAN (Ma et al., 2020) ranging from CS/MRA/VO techniques to sophisticated deep learning models, have been chosen for assessment alongside our approach, SPRINT. The majority of these techniques have their source code publicly accessible.

Visual assessments in our experiments reveal that ACS, AIHS, and APCA enhance spatial details in the combined image, but frequently introduce distortions in spectral characteristics is observed across different areas (refer to Figure-6.7, Figure-6.8, Figure-6.9, and Figure-6.10).

Spectral fidelity is crucial for studying specific regions of interest, and distortions can impact the scientific significance of the ultimate merged data output. Pan-sharpening methods based on MRA, such as SFIM, maintain spectral information but do not incorporate spatial details. HQB, PNN, and Pan-GAN effectively maintain spatial-spectral integration but have limitations in adjusting spectral behavior precisely according to the multispectral image, particularly in regions with mixed spectral targets.

SPRINT, functioning as a resilient image fusion framework, attains the best possible spatial and spectral resolution through a controlled and optimized merging of PAN data with multispectral images. Figures 6.7, 6.9, and 6.10 illustrate that the spectral characteristics of SPRINT are in harmony with Resourcesat's multispectral image, and significant spatial details are improved through the incorporation of Cartosat PAN data. Experiments in image fusion are also carried out using diverse remote sensing image pairs obtained from the IKONOS satellite. In Figure-6.8, an enlarged perspective of a stadium structure is presented, produced using various pan-sharpening techniques, showcasing the superior spatial enhancement of SPRINT with minimal spectral distortion.

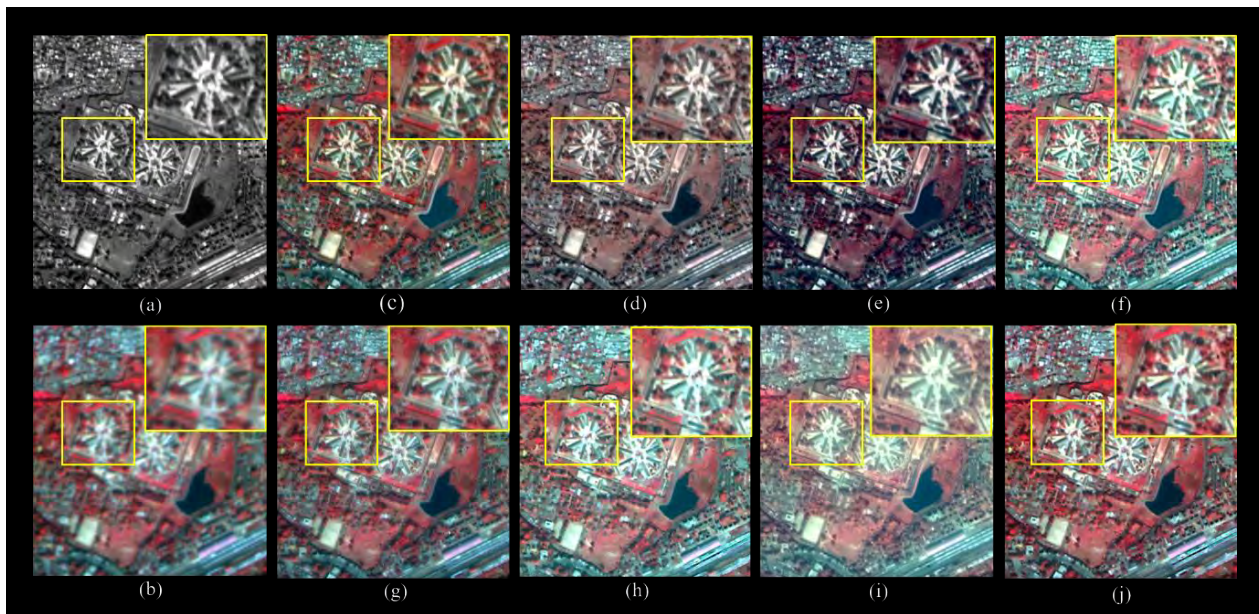


Figure-6.7. Visual Evaluation of Urban Landscape Fusion Pair No. 1 [(a) Cartosat-1 PAN (b) Resourcesat LISS-4 Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]

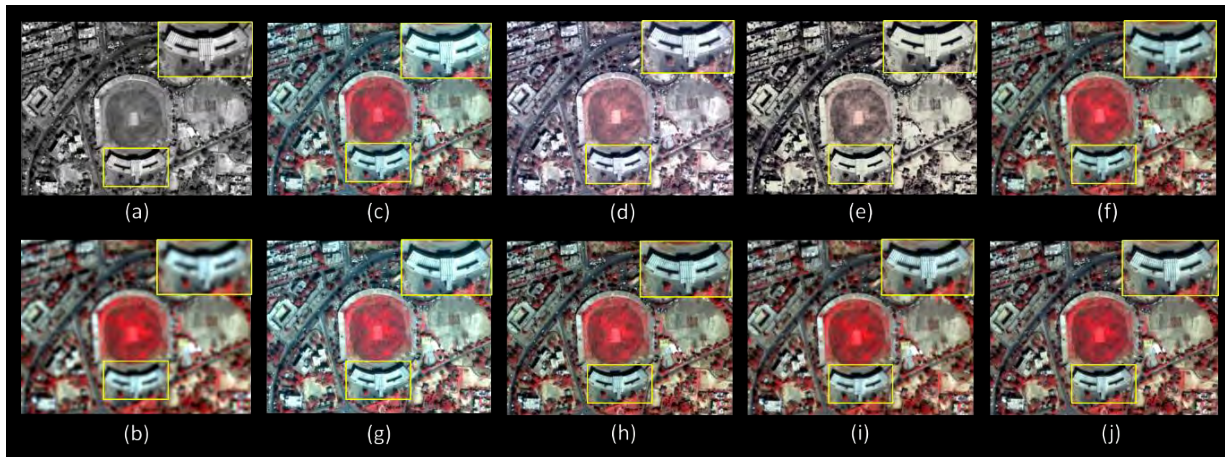


Figure-6.8. Visual Evaluation of Stadium Fusion Pair No. 2 [(a) IKONOS PAN (b) IKONOS Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]

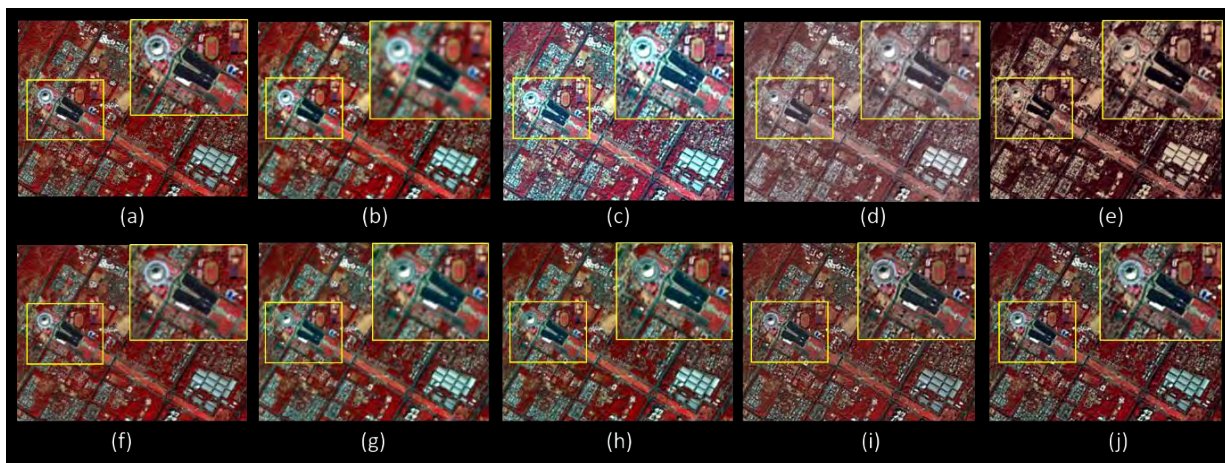


Figure-6.9. Visual Evaluation of Convention Centre Fusion Pair No. 3 [(a) Ground Truth (b) Degraded Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]

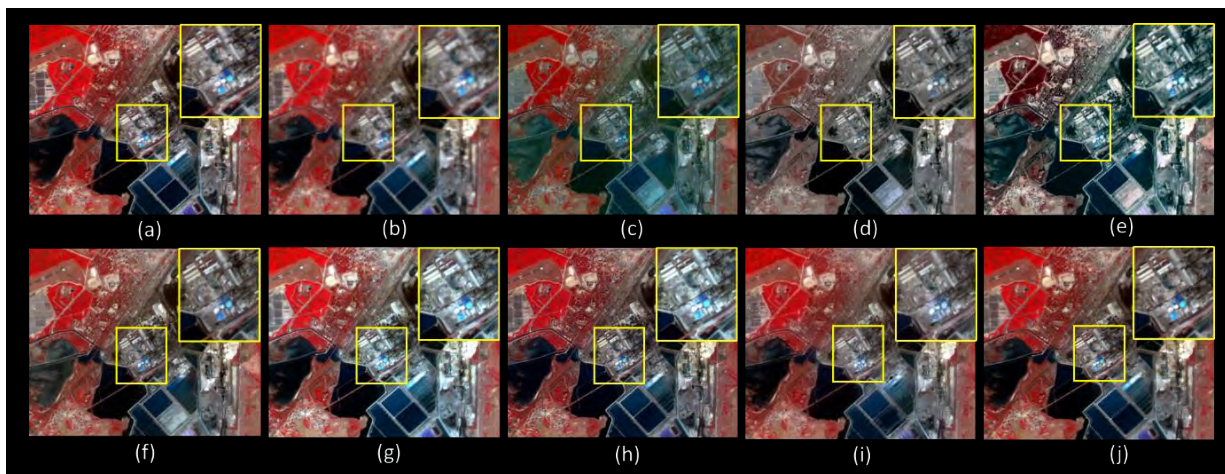


Figure-6.10. Visual Evaluation of Fusion Pair No. 4 featuring Solar Power Plant Structure [(a) Ground Truth (b) Degraded Multispectral (c) ACS (d) AIHS+GF (e) APCA (f) SFIM (g) HQB (h) PNN (i) Pan-GAN (j) SPRINT]



## B. Quantitative Assessment of Fusion Metrics

Various metrics are employed to assess the image fusion methods, including Structural Similarity (SSIM) (Wang et al., 2004), Correlation Coefficient (CC) (Taylor, 1990), Root Mean Square Error (RMSE) (Chai & Draxler, 2014), Spectral Angle Mapper (SAM) (Yuhua et al., 1992), Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) (Wald et al., 2000), and Universal Image Quality Index (UIQI) (Wang & Bovik, 2002). In the case of urban landscape (Fusion Pair No. 1), structural similarity is measured using SSIM, comparing the Cartosat PAN data with Resourcesat-2 multispectral images serving as the reference for evaluating spectral quality metrics. Table-6.3 displays the metrics computed through different image fusion techniques across the urban landscape. While AIHS exhibits a slight advantage in SSIM, SPRINT excels in other metrics specifically associated with spectral quality. Evaluation metrics indicate SPRINT's superiority over other image fusion techniques.

Similar assessments are conducted for the stadium structure (Fusion Pair No. 2), convention centre (Fusion Pair No. 3), and solar power plant structure (Fusion Pair No. 4), demonstrating SPRINT's consistent superiority in preserving spectral profiles while enhancing spatial details and tabulated in Table-6.4, Table-6.5 and Table-6.6. Reference ground truth multispectral images are generated for solar power plant structures, and SPRINT's robustness is evident in the evaluation metrics. Table-6.6 showcases the metrics for solar power plant images, with Pan-GAN showing marginal improvement in SSIM but SPRINT proving to be a more resilient approach to merging panchromatic and multispectral images. Bold font values in the tables indicate the optimal value for each metric across various fusion methods.

To provide a visual comparison plots are created for the image fusion techniques, illustrating the average values of each image quality metric over several images (Figure-6.11). These plots reveal that SPRINT outperforms all other image fusion techniques in terms of performance.

The no-reference image quality metrics provide valuable insights into essential image properties, indicating achieved image enhancement without notable spectral distortion. In our experiment, where a reference or ground truth image is unavailable for Fusion Pair No. 1 and 2, evaluating the performance of various fusion methods is more effectively carried out using no-reference image assessment metrics like spectral distortion indices  $D_s$  and  $D_\lambda$  (Yang et al., 2018). QNR, an additional no-reference image index comprised of  $D_s$  and  $D_\lambda$ , quantifies the luminance

and contrast of remote sensing images (Yang et al., 2018). Table-6.7 offers an impartial assessment of the outcomes achieved using different image fusion techniques.

In the urban landscape scenario of Fusion Pair No. 1, HQB (Wang et al., 2018) produces a superior outcome, with SPRINT ranking as the second-best performer. The fused images of the stadium taken by IKONOS PAN/Multispectral are also assessed using no-reference image quality metrics, demonstrating that SPRINT surpasses other fusion methods. Furthermore, the processing time is calculated by executing fusion methods on the identical computing system. The time required by various methods is indicated in the final column of Table-6.7, assuming a standard image scene size for a fusion pair of approximately 6000\*6000 pixels. The suggested approach, SPRINT, exhibits comparatively lower complexity when compared to modern state-of-the-art image fusion techniques in the field of variable observation and deep learning.

Table-6.3. Evaluation of Image Fusion Metrics in Urban Landscape (Fusion Pair No. 1)

Fusion Methods	SSIM ↑	CC ↑	RMSE ↓	SAM ↓	ERGAS ↓	UIQI ↑
ACS	0.7112	0.3288	3.0812	8.3426	2.1125	0.0542
AIHS	<b>0.7851</b>	0.3511	3.0631	9.7812	2.0832	0.1964
APCA	0.7551	0.4134	4.0388	6.9123	2.0416	0.1846
SFIM	0.7224	0.5811	3.0312	6.2311	1.0404	0.2122
HQB	0.7143	0.7155	4.0315	5.8322	2.0356	0.2315
PNN	0.7212	0.7882	4.0287	5.4112	1.0362	0.2711
Pan-GAN	0.7422	0.7942	3.0251	4.8112	1.0351	0.2822
SPRINT	0.7848	<b>0.8024</b>	<b>3.0242</b>	<b>4.2311</b>	<b>1.0339</b>	<b>0.3068</b>

Table-6.4. Evaluation of Image Fusion Metrics for Stadium Structure (Fusion Pair No. 2)

Fusion Methods	SSIM ↑	CC ↑	RMSE ↓	SAM ↓	ERGAS ↓	UIQI ↑
ACS	0.7218	0.4214	3.1323	6.8921	2.5981	0.0752
AIHS	0.7323	0.4416	3.0891	6.7812	2.4512	0.0781
APCA	0.7456	0.4654	2.9871	6.6721	2.3125	0.0891
SFIM	0.7215	0.6545	2.6521	6.4982	2.3081	0.0912
HQB	0.7536	0.6851	2.4332	6.2412	2.1281	0.1143
PNN	0.7891	0.7296	2.1215	5.7641	2.0465	0.1421
Pan-GAN	0.8012	0.7893	1.9476	4.9481	1.8819	0.1681
SPRINT	<b>0.8028</b>	<b>0.7936</b>	<b>1.6715</b>	<b>4.7188</b>	<b>1.6142</b>	<b>0.1983</b>

Table-6.5. Evaluation of Image Fusion Metrics for Convention Centre (Fusion Pair No. 3)

Fusion Methods	SSIM ↑	CC ↑	RMSE ↓	SAM ↓	ERGAS ↓	UIQI ↑
ACS	0.8121	0.5311	4.4322	8.9821	3.1823	0.0943
AIHS	0.8291	0.5432	4.3125	8.6123	3.2145	0.0987
APCA	0.8312	0.5621	3.9321	8.3123	3.1131	0.0995
SFIM	0.8119	0.5685	3.8713	7.9846	2.9832	0.1045
HQB	0.8432	0.6753	3.4521	7.7634	2.8832	0.1087
PNN	0.8481	0.7541	2.9345	6.9865	2.7912	0.1096
Pan-GAN	0.8561	0.7784	2.5632	5.8121	2.6312	0.1154
SPRINT	<b>0.8621</b>	<b>0.7891</b>	<b>2.1654</b>	<b>5.6825</b>	<b>2.3418</b>	<b>0.1269</b>

Table-6.6. Image Fusion Metrics Assessment over Solar Power Plant Structure (Fusion Pair No. 4)

Fusion Methods	SSIM ↑	CC ↑	RMSE ↓	SAM ↓	ERGAS ↓	UIQI ↑
ACS	0.7008	0.3138	2.0925	7.9812	2.1251	0.0712
AIHS	0.7143	0.3412	2.0723	7.1231	1.1162	0.0812
APCA	0.7412	0.3614	1.0561	6.8213	1.0952	0.1004
SFIM	0.7325	0.4245	1.0425	6.4124	1.0854	0.1153
HQB	0.7362	0.7223	1.4371	6.2355	2.0476	0.1421
PNN	0.7327	0.7282	1.0315	5.8123	2.0342	0.1764
Pan-GAN	<b>0.8121</b>	0.7561	1.0282	4.9324	1.0326	0.2261
SPRINT	0.8115	<b>0.7936</b>	<b>1.0195</b>	<b>4.5458</b>	<b>1.0295</b>	<b>0.2659</b>

The no-reference image quality metrics offer us insights into essential properties of image that signify achieved image improvement with minimal spectral distortion. In our experiment, there is no reference or ground truth image available for Fusion Pair No. 1 and 2, making it more suitable to assess the different fusion techniques performance using no-reference image assessment metrics like spectral distortion indices  $D_s$  and  $D_\lambda$  (Yang et al., 2018). QNR is an additional no-reference image index, comprised of  $D_s$  and  $D_\lambda$ , designed to gauge the luminance and contrast of remote sensing images (Yang et al., 2018). Table-6.7 presents an impartial assessment of the outcomes achieved through various image fusion approach. The Fusion Pair No. 1 represents urban landscape demonstrates a superior outcome with HQB (Wang et al., 2018), and the second-best result is achieved with our proposed approach, SPRINT. The fused images of the stadium taken by IKONOS PAN/Multispectral are also evaluated using no-reference image quality metrics, revealing that fusion performance of SPRINT's is superior. Additionally, the processing time is calculated by executing fusion methods on the same computing system. The time required by

various methods is indicated in the final column of Table-6.7, taking into account a standard fusion pair image scene size of approximately 6000\*6000 pixels. The complexity of the suggested method, SPRINT, is relatively modest when contrasted with the current state-of-the-art image fusion techniques in the domain of VO and deep learning.

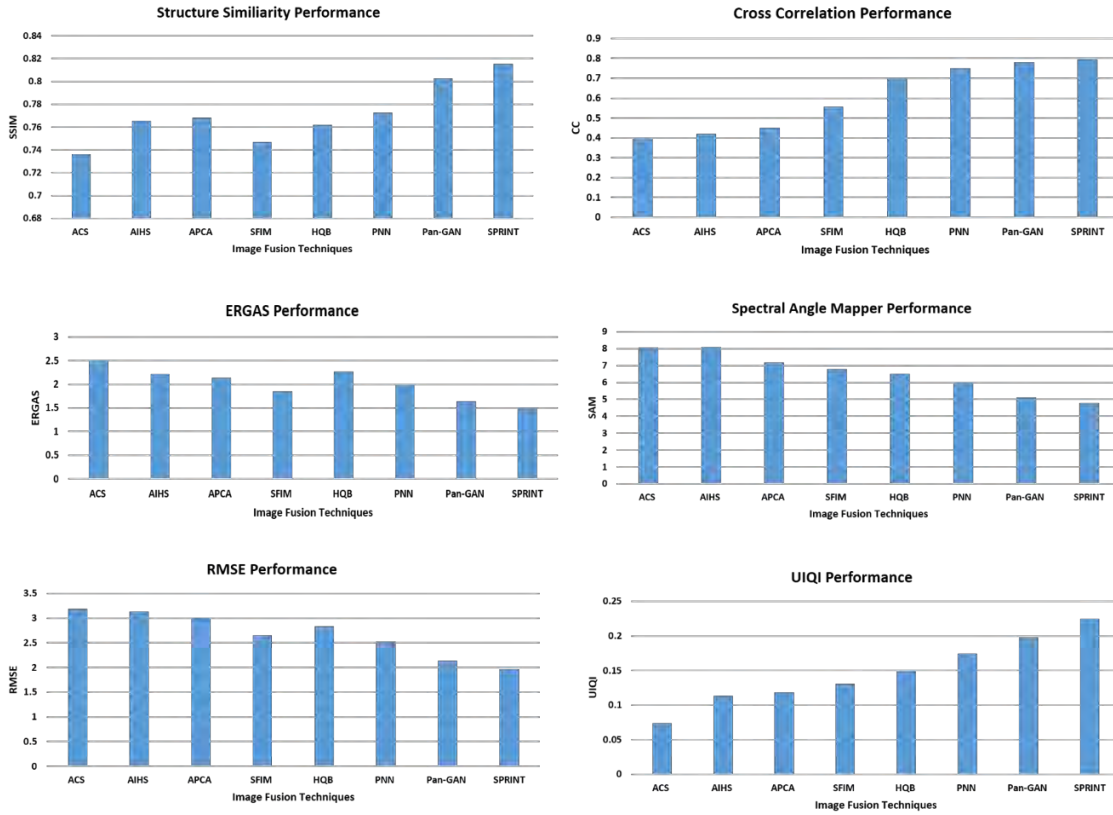


Figure-6.11. Charts depicting image quality metrics for the comparison of image fusion Table-6.7. Assessment of Fused Image Quality without Reference

Fusion Methods	Fusion Pair No. 1 (Urban)			Fusion Pair No. 2 (Stadium)			Runtime(s)
	$D_\lambda \downarrow$	$D_s \downarrow$	QNR $\uparrow$	$D_\lambda \downarrow$	$D_s \downarrow$	QNR $\uparrow$	
ACS	0.0821	0.1923	0.7413	0.0654	0.1787	0.7675	<b>12.2443</b>
AIHS	0.0801	0.1916	0.7436	0.0612	0.1776	0.7720	13.5365
APCA	0.0721	0.1903	0.7513	0.0606	0.1753	0.7747	18.3256
SFIM	0.0613	0.1892	0.7610	0.0655	0.1655	0.7798	17.1265
HQB	<b>0.0388</b>	<b>0.1564</b>	<b>0.8108</b>	0.0548	0.1594	0.7945	26.3264
PNN	0.0581	0.1645	0.7869	0.0521	0.1483	0.8073	19.4366
Pan-GAN	0.0413	0.1621	0.8032	0.0483	0.1444	0.8142	21.3442
SPRINT	0.0407	0.1589	0.8068	<b>0.0469</b>	<b>0.1421</b>	<b>0.8176</b>	19.1354

### C. Evaluation of SPRINT Performance over Hills and NDVI Comparison

The performance of the SPRINT workflow is evaluated using a scene pair in hilly terrain (Fusion Pair No. 5) that includes a variety of land classes. The assessment is conducted using ortho-rectified tiles from Cartosat-1 and images from Resourcesat-2A LISS-4. A reference ground truth image is created by magnifying the original LISS-4 multispectral image. In the experiment, the LISS-4 image is reduced to a resolution of 10.0 meters, while the Cartosat-1 panchromatic image is downsampled to 5.0 meters. The results indicate that the SPRINT output exhibits spatial enhancement while preserving spectral signatures. The radiometric and geometric properties of the SPRINT fused image align with those of the reference image. To enhance the evaluation, the Normalized Difference Vegetation Index (NDVI) is calculated using the original reference and SPRINT fused radiance images, employing the Near-Infrared (NIR) and Red (R) wavelength channels. NDVI is formulated as:

$$NDVI = \frac{NIR - R}{NIR + R} \quad (6.21)$$

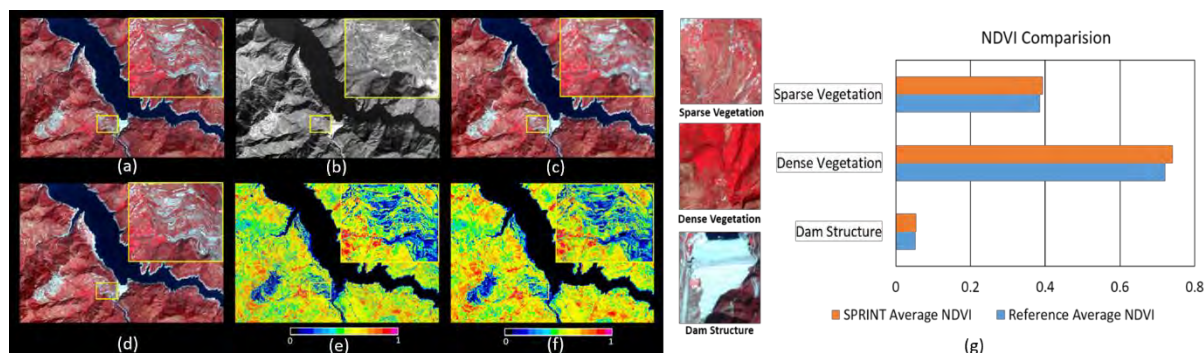


Figure 6.12. Performance of SPRINT over hilly terrain for Fusion Pair No. 5 (a) Ground Truth Reference, (b) Cartosat-1 Degraded Ortho Rectified Tile, (c) LISS-4 Degraded Multispectral, (d) SPRINT, (e) Ground Truth NDVI, (f) SPRINT NDVI, and (g) Comparison of NDVI between SPRINT and the reference across various classes

The NDVI obtained from the SPRINT fused image closely correspond to the authentic NDVI measurements. Figure-6.12 visually demonstrates the effectiveness of SPRINT fusion and NDVI in a hilly terrain region. For qualitative evaluation, mean NDVI values are calculated for various classes, such as sparse vegetation, dense vegetation, and a dam structure. Class values are manually chosen to guarantee that pixel values within each group do not include contributions from mixed spectral signatures. The dense vegetation class exhibits the maximum deviation, whereas the dam structure shows the minimum deviation. In general, the deviation in NDVI values is determined to

be negligible, suggesting that the SPRINT fused image maintains spectral properties with spatial enhancement. Figure-6.12 (g) showcases different multispectral classes and a juxtaposition of average NDVI values between reference measurements and NDVI estimates obtained through SPRINT.

#### *D. Surface Reflectance Profile Derived by SPRINT*

To conduct this distinctive analysis, Surface reflectance images are generated using the fused radiance images produced by SPRINT. The efficiency of the SPRINT reflectance image is subsequently assessed by comparing it to the original LISS-3 reflectance image across various features like urban areas, vegetation, and water bodies. Surface reflectance data products are crucial for mitigating the impact of intermittent atmospheric scattering and absorption. Atmospheric correction is carried out using Radiative Transfer Modelling (RTM) with 6S (Vermote et al., 1997), ensuring computational efficiency in this intensive operation (Rusia et al., 2021). The details of the Landsat and Resourcesat image pair for this experiment (Fusion Pair No. 6 and 7) are specified in Table-6.2. The SPRINT fused reflectance image is produced with a spatial resolution of 15.0 meters.

Figure-6.13 offers a qualitative evaluation of the SPRINT fused reflectance image, contrasting it with the corresponding original interpolated reflectance from Resourcesat-2A LISS-3 in urban, water, and vegetation regions. The fused image displays enhanced delineation of roads and structures in urban areas when compared to their original counterparts. A comparable enhancement is observed in vegetation areas and narrow water channels.

Comparing spectral reflectance across water, vegetation, and urban regions in the fused product generated by SPRINT, as shown in Figure-6.14, indicates a strong resemblance to the original LISS-3 multispectral reflectance measurements across all wavelength channels. For comparing spectral reflectance profiles, small subsets of each feature target are selected, cropped, and mean reflectance values are calculated. The observed average deviation in radiometric accuracy is below 1%, rendering the SPRINT fused image suitable for scientific data analysis in space-borne applications.

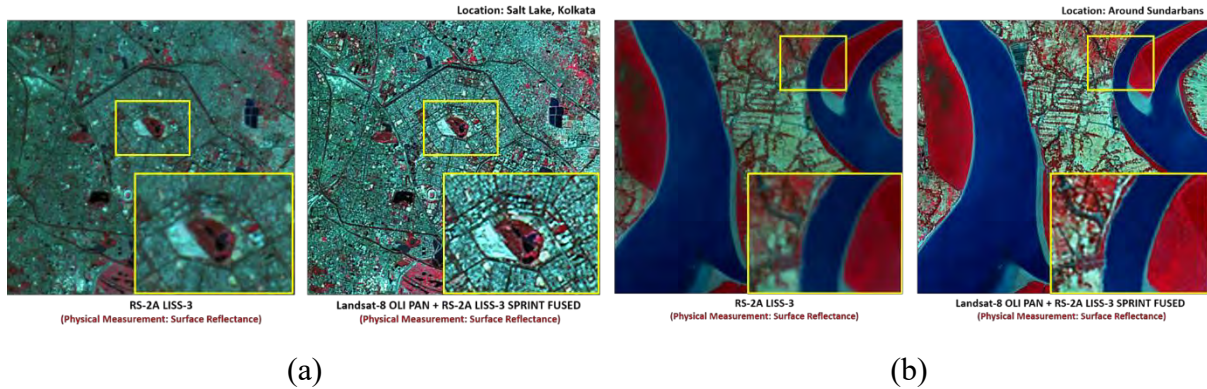


Figure 6.13. (a) Visual comparison of SPRINT reflectance fusion over an urban area (Fusion Pair No. 6).  
 (b) Visual comparison of SPRINT reflectance fusion over water and vegetation (Fusion Pair No. 7).

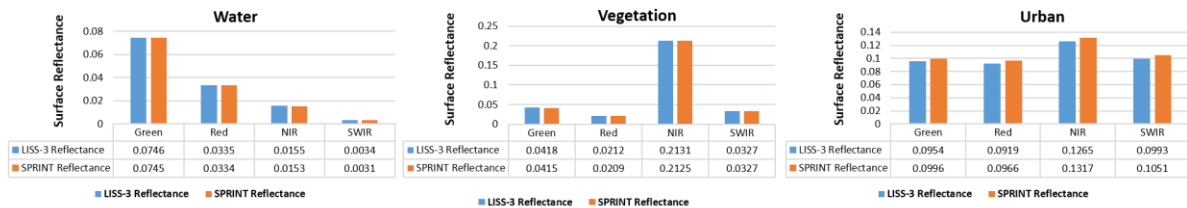


Figure-6.14. Comparison of spectral reflectance profiles between the original interpolated reflectance from LISS-3 and SPRINT-generated reflectance across various spectral classes

### 6.1.4. Summary

We introduced SPRINT, a proficient fusion technique designed for integrating panchromatic data with multispectral image stacks. The data pre-processing engine integrated into SPRINT prepares multi-sensor data for the fusion process. HNEF amplifies significant gradient magnitude features, while the Minnaert-based control mechanism discerns terrain topography. These elements are input into a Bayesian probabilistic model to produce a fused data output. The medoid intensity component, utilizing a bilateral filter, proficiently mitigates noise and acts as an ideal representative of the multispectral/panchromatic image pair.

SPRINT is assessed across diverse feature targets, and its visual performance is scrutinized with fusion techniques based on Component Substitution (CS), Multi-Resolution Analysis (MRA), Variable Optimization (VO), and deep learning. The image fusion quality metrics consistently indicate that SPRINT excels in combining multispectral data with detailed panchromatic information. The minimal deviation observed between SPRINT fused NDVI values and the actual

NDVI values across various classes suggests that SPRINT retains spectral information, rendering the data analysis suitable for scientific applications.

Experimental findings indicate that the surface reflectance product derived from SPRINT offers more informative content with improved spatial resolution while preserving spectral consistency. The fused reflectance profile generated by SPRINT corresponds to the spectral signature trend across various wavelength channels and multiple classes, displaying a closer agreement with the actual reflectance measurements. SPRINT acts as a versatile image fusion framework suitable for producing optimal fused data from a variety of satellite remote sensing images. In future endeavors, our goal is to automate the adjustment of fusion parameters and further enhance SPRINT for tasks related to hyperspectral pan-sharpening.



## 6.2. Infrared-Visible Image Fusion

### 6.2.1. Introduction

The research work introduces an innovative approach to fuse thermal infrared and multispectral visible images. In this method, spatial attributes are derived from thermal broadband image, whereas multispectral visible imagery is utilized to acquire spectral information. The amalgamation of high-resolution thermal infrared images from various satellites improves pixel information, providing valuable insights into landscapes. The fused product combining infrared and visible data proves advantageous for monitoring changes in vegetation health, evaluating urban heat islands, and investigating the effects of climate change.

The high-resolution thermal data with multispectral visible image combination assists in discerning different materials or objects. This differentiation is especially valuable in applications such as military reconnaissance and land cover classification. The suggested approach, referred to as Co-Occurrence with Laplacian Intensity Modulation (CLIM), is utilized for extracting boundaries from thermal images and incorporating them into improved multispectral remote sensing images (Misra et al., 2023c).

CLIM improves the thermal infrared dataset by applying a Laplacian of Gaussian Filter to enhance specific features and utilizes a Co-occurrence Image Filter to refine the image. The enhancement of the multispectral visible image is achieved by employing the IHS color space transformation and adaptive contrast enhancement. The sharp and smooth components ratio extracts essential edge information, and this information is modulated with the nearest acquisition multispectral image to produce the fused thermal-visible image.

The Indian Nano Satellite Technology Demonstrator (INS-2TD) is spotlighted as an advanced broadband thermal infrared sensor system specifically crafted for Nano satellite platforms. The Landsat-8 satellite, featuring a Thermal Infrared camera that spans the 10–12.5 micron band, is also acknowledged for its applicability in various remote sensing scenarios.

The effectiveness of the suggested method is assessed using the most recent captures from INS-2TD, thermal infrared images from Landsat-8, and the most proximate imaging acquisition from MODIS multispectral visible images. The fusion based on the CLIM method is visually evaluated across various spectral regions and compared with advanced remote sensing image fusion

techniques. The metrics for image quality are calculated to quantitatively assess the output of fusion obtained by combining thermal infrared images with multispectral visible images.

## 6.2.2. Methodology Developed

### A. Co-occurrence Smoothing Filter

The Co-occurrence Filter (CoF) is a filter designed to preserve boundaries by relying on a co-occurrence matrix instead of using a Gaussian on the range values to retain edges, as described by Jevnisek and Avidan in 2017. In CoF, pixels with frequently co-occurring values, typically found in texture-rich regions of an image, are assigned higher weights, contributing to a smoother image by blending them together. Conversely, pixels with infrequent co-occurrences receive lower weights, thereby preserving the boundaries between textures. It has been observed that co-occurrence information not only enhances edges but also proves effective in handling boundaries. The definition of CoF is as follows.

$$O_p = \frac{\sum_{q \in N(p)} G_{\sigma_s}(p, q) \cdot M(I_p I_q) \cdot I_q}{\sum_{q \in N(p)} G_{\sigma_s}(p, q) \cdot M(I_p I_q)} \quad (6.22)$$

In this context,  $O_p$  represents the pixel value of the output image at the specified pixel index  $p$ , while  $I_p$  and  $I_q$  denote the pixel values of the input image at the specified pixel indices  $p$  and  $q$ , respectively.  $G_{\sigma_s}$  refers to the Gaussian filter, with  $\sigma_s$  representing the parameter specified by the user.

$M$  is the 256\*256 matrix that can be written as:

$$M(a, b) = \frac{C(a, b)}{H(a)H(b)} \quad (6.23)$$

In this context,  $C(a, b)$  is derived from the co-occurrence matrix  $(a, b)$ , where  $(a, b)$  records the co-occurrence count of values  $a$  and  $b$ .  $H(a)$  and  $H(b)$  represent the frequencies of the values of the pixels  $a$  and  $b$ , respectively, and these frequencies can be calculated from the histogram of pixel values.

CoF is employed in the processing of broadband thermal infrared remote sensing images. The  $\sigma_s$  parameter plays a crucial role in regulating the extent of smoothing. In our specific instance, with a window size of 11\*11, we have opted for  $\sigma_s = 2\sqrt{11} + 1$ . As illustrated in Figure-6.15, the thermal image undergoes smoothing, yet the filtered output effectively preserves the boundaries.

It is evident from the image that co-occurrence information plays a key role in maintaining the contours of tributaries flowing into larger rivers or seas.

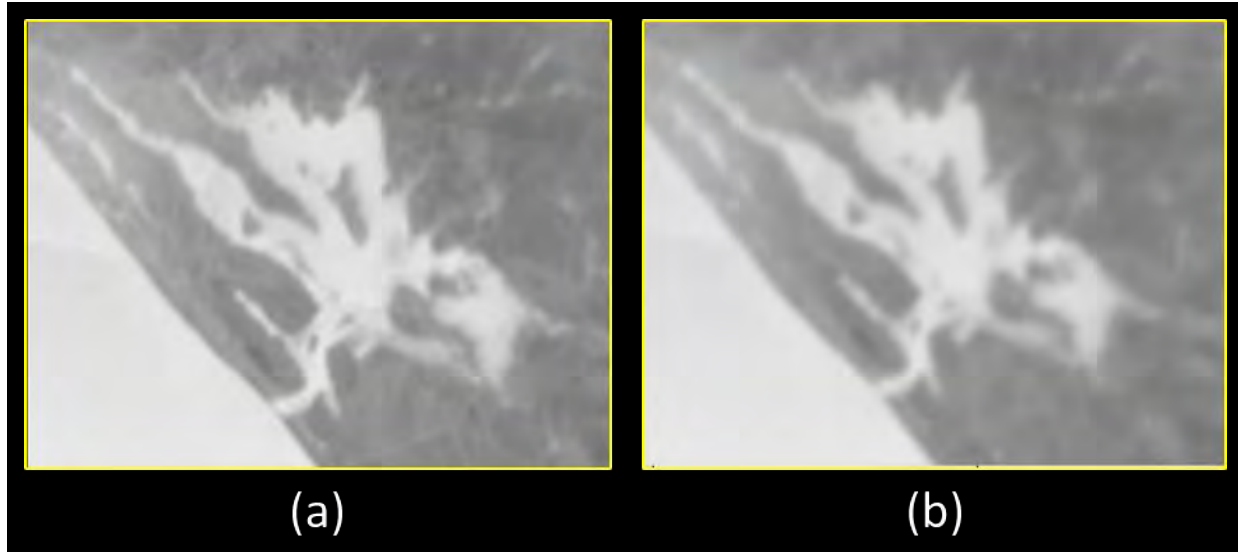


Figure-6.15. (a) Original Broadband Thermal Image (b) Output of Co-occurrence Filter

#### B. Enhancement of Broadband Thermal Images Using Laplacian of Gaussian

Laplacian Filters, recognized as derivative filters, are utilized to identify areas of rapid changes, such as edges (Fu et al., 2021). As derivative filters are highly sensitive to noise, it is crucial to initially smooth the image with a Gaussian filter before implementing the Laplacian. The Laplacian of Gaussian (LoG) functions as a high-pass filter, proficiently identifying prominent sharp edges in the broadband thermal image. The connection between the high-pass image  $I_H$  and the thermal input image  $I_T$  at a specific pixel location  $(x, y)$  is articulated as follows.

$$I_H = \log(x, y, \sigma) * I_T \quad (6.24)$$

$$\text{where, } \log(x, y, \sigma) = \frac{-1}{\pi\sigma^4} \left[ 1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (6.25)$$

The image sharpening, based on Laplacian of Gaussian ( $LoG$ ), is implemented on the Broadband Thermal Image to improve and incorporating edge information by combining it with the original thermal infrared image. In Equation-6.24 and Equation-6.25,  $\sigma$  represents the standard deviation of the Gaussian kernel. As depicted in Figure-6.16, the sharpened image now more clearly delineates the river track and texture regions.

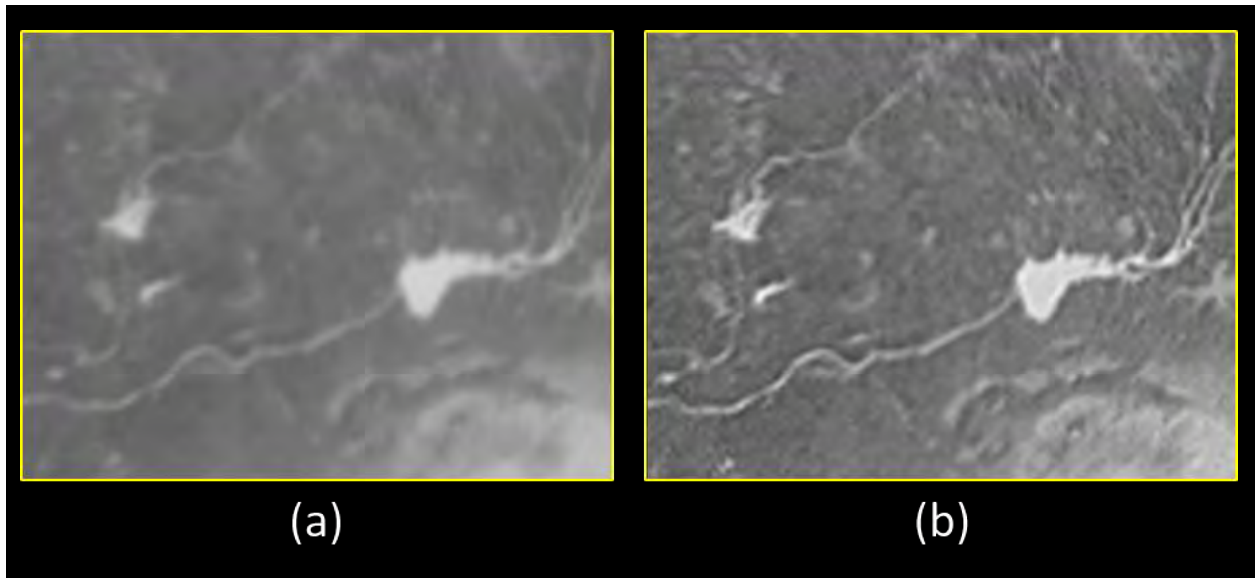


Figure-6.16. (a) Original Thermal Infrared Image (b) Sharpened Thermal Infrared Image using *LoG*

### C. Visible Image Enhancement Using IHS Transform and CLAHE

The multispectral visible image possesses a less detailed spatial resolution and necessitates enhancement to refine its texture characteristics. In our approach, we convert the visible image from the RGB color space to the IHS color space (Schetsellar et al., 1998) using the conventional IHS transformation method (Chu & Zhu, 2008). The intensity component is subsequently enhanced through Contrast Limited Adaptive Histogram Equalization (CLAHE) to better emphasize local features, thereby elevating the overall quality of the visible image. Image enhancement based on CLAHE restricts the amplification of contrast stretch according to a predefined histogram limit (Misra et al., 2021a). A crucial factor in this procedure is the size of the window around a pixel, for which the histogram is equalized.

Figure-6.17 displays the initial multispectral visible remote sensing image alongside the resulting improved multispectral visible image achieved with different window sizes, denoted as  $w$ . It is noted that establishing the window size  $w$  as 256 for CLAHE enhancement of the multispectral visible image yields the most favorable visual appearance. The inverted transformation from IHS to RGB color space is applied to the improved intensity image, in conjunction with the hue and saturation components, to generate the enhanced multispectral visible image.

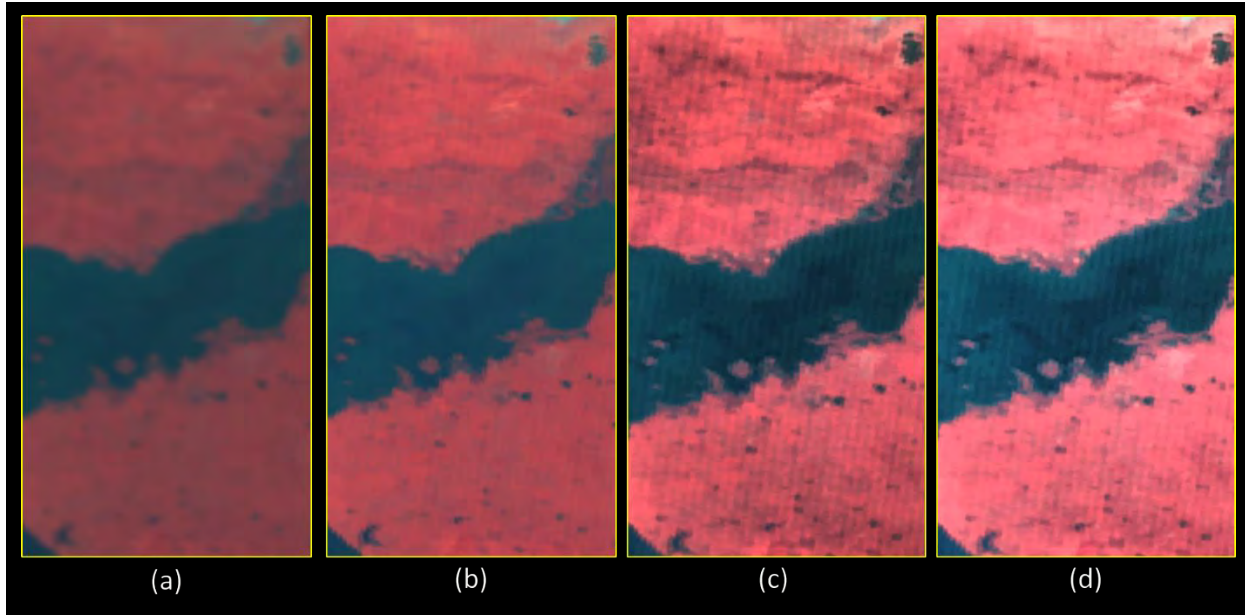


Figure-6.17. Comparative Analysis of CLAHE for enhancing multispectral visible images (a) Input Image (b) Enhanced with  $w = 64$  (c) Enhanced with  $w = 256$  (d) Enhanced with  $w = 512$

#### D. Algorithm and Workflow for CLIM Image Fusion

The main goal of our suggested infrared-visible remote sensing image fusion technique, CLIM, is to boost the spatial resolution of multispectral images by incorporating spatial information derived from thermal radiance. At the same time, it is essential to carefully maintain the spectral characteristics of the multispectral visible image while integrating the fine details of high spatial resolution from the thermal channel. Generally, discernible features in a daytime thermal image frequently exhibit similarities to those in multispectral reflective spectral bands. The theoretical basis of the CLIM method revolves around extracting structural information from thermal infrared images while preserving the spectral characteristics of multispectral remote sensing images.

The CLIM approach includes determining the ratio between the improved thermal image derived from the Laplacian of Gaussian (LoG) filter and the Co-occurrence Filter (CoF) smoothing filter. This ratio is formulated to detect significant boundary information within the broadband thermal image. The ratio accentuates topographic variable characteristics within a high spatial resolution broadband thermal pixel, portraying different slope and edge features in an image.

To incorporate this information into the intensity of the visible band, the spatial information ratio from the thermal data is adjusted by combining it with the nearest simultaneously improved

multispectral visible remote sensing image. The procedural steps of CLIM are delineated in Algorithm-6.2.

**Algorithm-6.2: Image Fusion using CLIM**

Inputs:  $I_T$  = Remote Sensing Image in the Thermal Infrared Spectrum,

$I_{MX}$  = Remote Sensing Image in the Multispectral Visible Spectrum.

High-pass filtering using the Laplacian of Gaussian:  $I_{HPF} = \log(I_T)$

Co-occurrence Low Pass Filtering:  $I_{LPF} = CoF(I_T)$

Thermal Image Enhancement:  $I_{SH} = I_T + I_{HPF}$

IHS Transform and Intensity Generation:  $I_H = IHS(I_{MX})$

CLAHE Enhancement:  $I_{CLAHE} = clahe(I_H)$

Inverse IHS to RGB Transform:  $I_{EMX} = inverse(I_{CLAHE})$

CLIM Image Fusion:  $I_{CLIM} = \frac{I_{SH}}{I_{LPF}} * I_{EMX}$

Output:  $I_{CLIM}$  = Image Fusion Using the CLIM Method for Infrared and Visible Bands

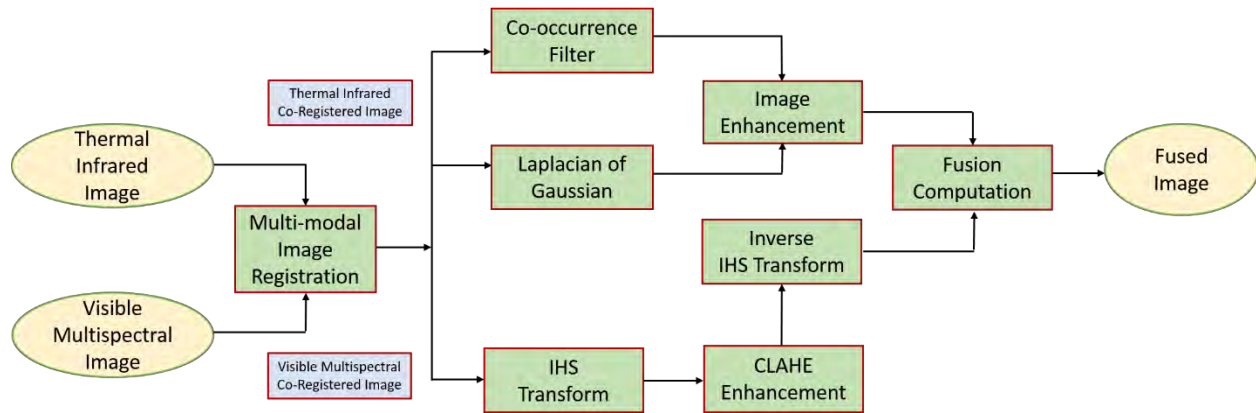


Figure-6.18. CLIM Workflow for Infrared-Visible Fusion

The CLIM algorithm is transformed into a processing workflow aimed at generating a fused image from thermal infrared and multispectral visible images. The multi-modal images are subjected to co-registration to achieve precision in registration at the sub-pixel level (Misra et al., 2021b). The co-registered thermal image is subsequently fed into the CoF and LoG computation for the enhancement of the thermal infrared image. At the same time, the co-registered multispectral visible image undergoes a transformation into the IHS color space. The intensity component is improved through CLAHE, followed by an inverted IHS transformation, to produce an enhanced multispectral remote sensing image.

The proposed CLIM framework includes essential modulation operations and convolutions with the improved multispectral image to create the fused image. Figure-6.18 depicts the processing workflow of the CLIM methodology.

### 6.2.3. Experimental Results

The evaluation of the proposed CLIM methodology is conducted using the INS-2TD broadband thermal infrared image and the MODIS multispectral visible image acquired on the same day. Furthermore, the methodology is validated using Landsat-8 thermal infrared bands along with MODIS multispectral visible channels. Access to INS-2TD data is available (or can be requested) from the Indian Space Science Data Center (ISSDC) (ISSDC, 2023). The datasets from Landsat-8 and MODIS can be accessed online through the Earth Explorer web portal (Earth Explorer, 2023) and the MODIS Level-1 land products web portal (MODIS, 2023). The study area includes sections of Gujarat, India, encompassing both land and coastal regions.

Additionally, we make use of TNO multi-band image collection datasets, incorporating visible, near-infrared, and long-wave infrared images, to contribute to the overall applicability of the CLIM methodology (Toet, 2017). Table-6.8 furnishes information regarding the data employed in our image fusion experiments. The datasets for training and testing encompass diverse remote sensing feature targets, including dams, refineries, ship-docking zones, vegetation, and water bodies. The dataset's spatial resolution is detailed in Table-6.8.

The experimental software environment operates on the Linux operating system, and the computer programs are coded in C and C++. The qualitative and quantitative fusion outcomes are derived through post-registered pairs of infrared and visible images. The efficacy of the multi-modal image registration, which is a component within the CLIM framework, is depicted in Figure-6.19. In Figure-6.19(a), it is evident from the image that the island is not properly aligned during the pre-registration phase of the multi-modal image pair. Nevertheless, as depicted in Figure-6.19(b), following image registration, the infrared and visible image pair is overlaid, accurately aligning the island. This alignment is emphasized with a white oval area in the region of interest in the image swipes.

Table-6.8. Specifications of Thermal and Visible Remote Sensing Images for Fusion Evaluation

S. No.	Satellite/Sensor Type	Wavelength (in $\mu\text{m}$ )	Spatial Resolution (in m)
1.	Landsat-8/Thermal Infrared	10.6-11.9	100
2.	INS-2TD/Thermal Infrared	7-12	175
3.	MODIS/Multispectral Visible	0.459-0.479 0.545-0.565 1.230-1.250	500
4.	TNO Multi-Band Image Collection	0.3-0.7 (Visible) 8-12 (Long Wave Infrared)	NA

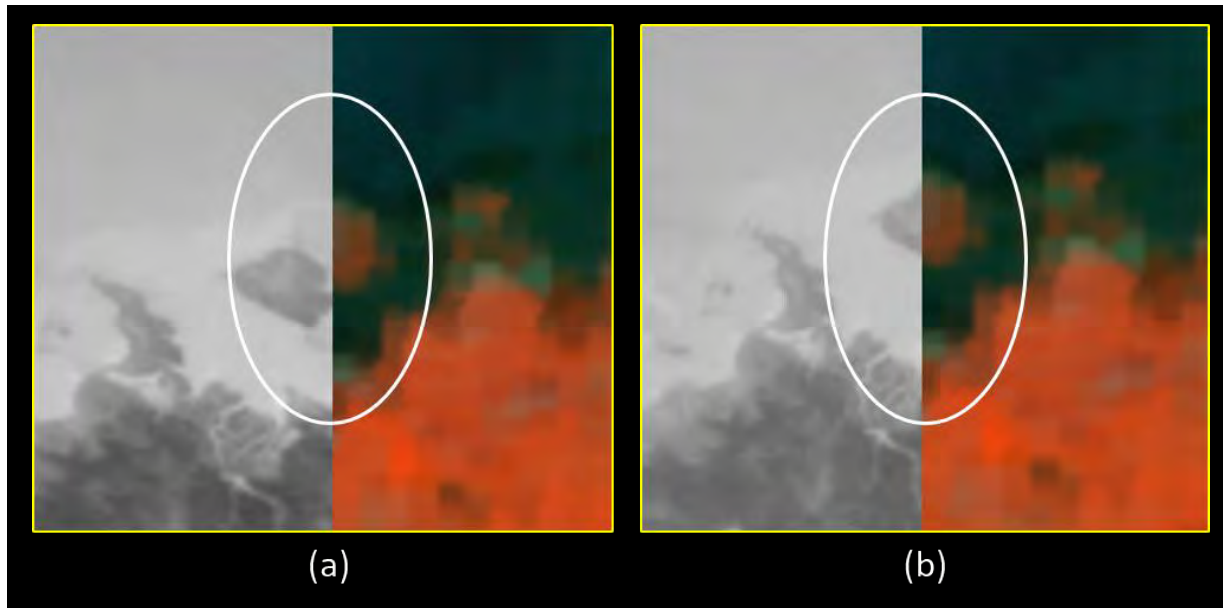


Figure-6.19. Registration Performance Comparison for Multi-Modal Images of Infrared and Visible Channels (a) Before Registration (b) After Registration

A. Visual Assessment and Comparison of Fusion Results

The image fusion process utilizing CLIM is executed on both thermal and visible image datasets, leading to the generation of a fused output. In Figures 6.20(a) and 6.20(b), we witness the INS-2TD thermal infrared image with relatively high spatial resolution and the MODIS visible image, both resampled to a spatial resolution of 175 meters to ensure consistent pixel size before the fusion process. Figure 6.20(c) displays the fused image created through CLIM, integrating spatial details from the thermal band while preserving the spectral characteristics of the MODIS visible image. All images are enlarged by a factor of two to focus on specific region targets. The Sasoi dam and Refinery complex are indicated by green and blue square boxes, respectively, in Figure 6.20. Importantly, these structures are now clearly visible in the fused image. The branches



of the dam are distinctly outlined, a feature that is difficult to discern in the MODIS visible image. The refinery complex near the dam displays an elevated brightness temperature captured by the thermal image, seamlessly integrated in a controlled manner through CLIM-based image fusion. The amalgamation of characteristics from thermal and visible spectral channels produces a more comprehensive image, providing detailed insights into the feature target compared to their individual counterparts.

Additionally, the evaluation of CLIM involves the utilization of Landsat-8 thermal infrared images in combination with MODIS multispectral visible images. Figure 6.21 presents the Landsat-8 infrared image, MODIS visible image, and the fused image obtained through the CLIM method. Linear stretching is applied to all images to improve visualization and facilitate image interpretation. The CLIM fused image clearly emphasizes ship docking zones while maintaining the spectral information from the multispectral visible channels in the ultimate output.

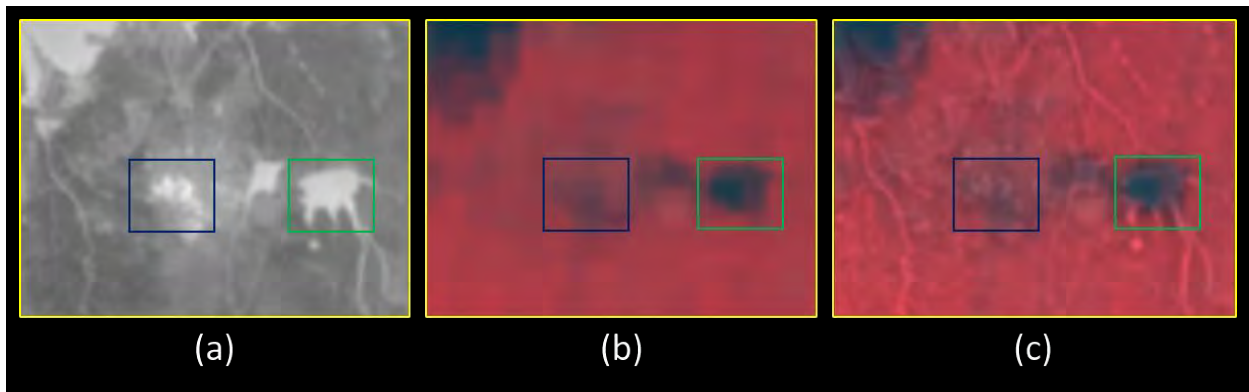


Figure-6.20. Region of Sasoi Dam and Refinery Complex (a) INS-2TD Broadband Thermal Image (b) Resampled MODIS Visible Multispectral Image (c) CLIM Fused Image

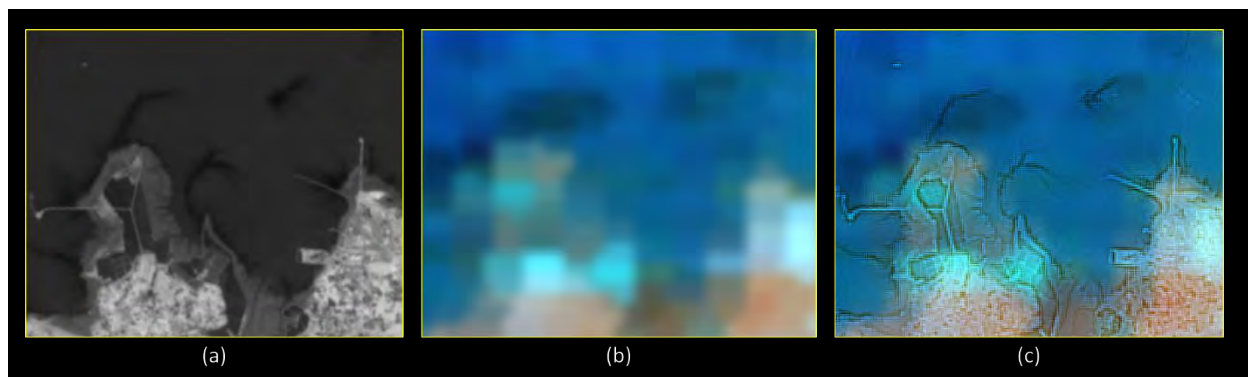


Figure-6.21. Areas with Ship Docking Zones (a) Thermal Infrared Image from Landsat-8 (b) Resampled Multispectral Visible Image from MODIS (c) CLIM Fused Image

Various methods for fusing remote sensing images are utilized to combine thermal infrared images with multispectral visible images. Figure-6.22 showcases the effectiveness of various image fusion techniques for visual quality assessment. Fusion techniques like BT (Misra et al., 2012b), IIHS (Tiwari et al., 2021), PCA (Wu et al., 2018), SOE (Liu et al., 2022), and FGIF (Guo et al., 2022) designed for Infrared-Visible fusion are employed on multi-modal image pairs, resulting in fusion outputs. The region marked with a yellow square box is zoomed in Figure-6.22 and Figure-6.23 to improve the visualization of the image fusion outcomes. As depicted in Figure-6.22, all Component Substitution (CS)-based methods display notable spectral distortion, affecting the interpretation of remote sensing images for various Earth observation applications. SOE (Liu et al., 2022) incorporates spatial details into the multispectral image with comparatively minimal spectral distortion. The Fast Guided Image Filtering (FGIF) method (Guo et al., 2022) showcases enhanced spatial enhancement in the fusion of thermal and visible images. Nevertheless, FGIF also leads to irregular radiometric properties in comparison to the multispectral image.

On the contrary, the fused output produced by CLIM captures pertinent spatial details from the thermal broadband channel and merges them with the multispectral visible image, creating the most representative image possible. The spectral characteristics of the water and land regions in the CLIM fused image closely match those of the multispectral visible image. Moreover, spatial improvement is noticeable at the edges and prominent feature areas.

The assessment of infrared-visible image fusion is carried out on an image pair comprising Landsat-8 thermal infrared and MODIS multispectral visible images, fused using diverse image fusion strategies. Figure-6.23 illustrates a comparison of fusion outcomes, and the images are stretched to facilitate improved visual assessment. The outcomes reveal that BT (Misra et al., 2012b) enhances the spatial aspect of the multispectral visible image using thermal infrared spatial information but results in a loss of spectral information. IIHS (Tiwari et al., 2021) maintains spectral details to a certain extent, but it falls short in terms of spatial enhancement. PCA (Wu et al., 2018) and SOE (Liu et al., 2022) exhibit slight enhancement, introducing spatial information while indicating the preservation of spectral characteristics. FGIF (Guo et al., 2022) combines the infrared-visible image, exposing texture information along with spatial details from the Landsat-8 infrared channel. In comparison, our proposed method, CLIM, produces an optimal output where both spatial and spectral information are integrated in a controlled manner. The CLIM fused image

distinctly showcases river tributaries, and its spectral information closely resembles that of the original MODIS multispectral visible image.

The effectiveness of CLIM fusion is also assessed in the TNO image dataset, encompassing long-wave infrared (LWIR) and visible images. In Figure-6.24, an individual in front of a house is detected in the LWIR image but remains unnoticed in the corresponding visible image. The figure illustrates the performance of CLIM fusion in comparison to other well-known image fusion techniques within the TNO multi-band image collection. The CLIM fused image yields the most accurate output, correctly identifying the person in front of the house while preserving the characteristics of the visible image in the fused result. This clearly indicates the applicability of the proposed fusion algorithm beyond remote sensing images in a broader context.

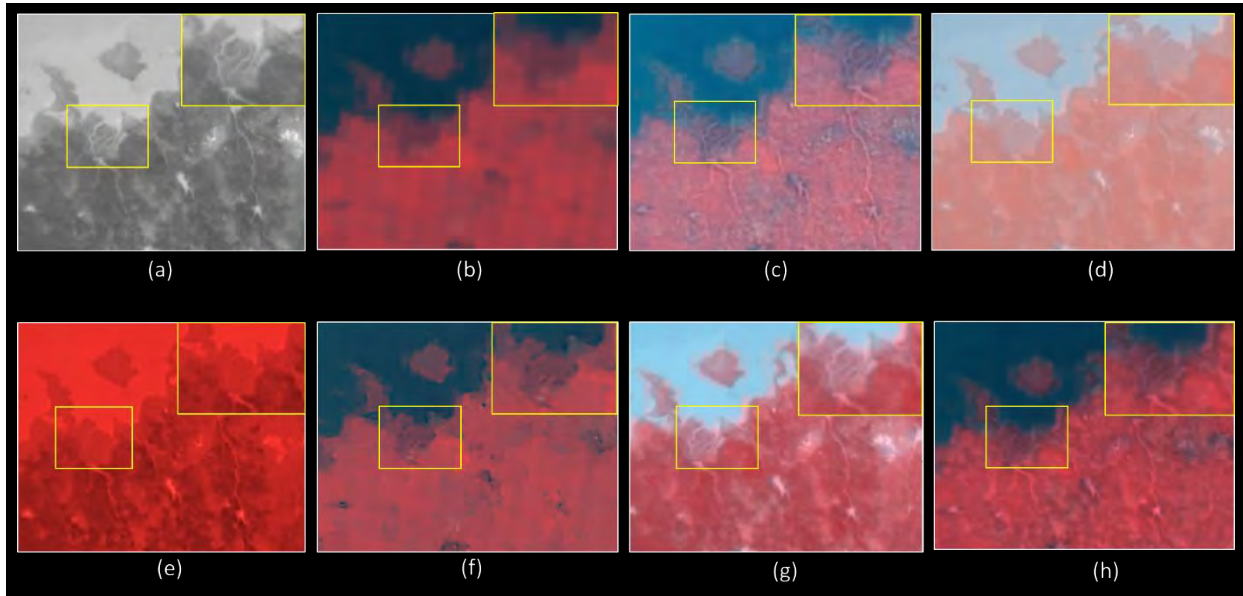


Figure-6.22. Comparison of Fusion Results for INS-2TD Broadband Thermal Infrared and MODIS Visible Images across Different Feature Targets (a) Thermal Infrared Image (b) Resampled Multispectral Image (c) BT (d) IIHS (e) PCA (f) SOE (g) FGIF (h) CLIM

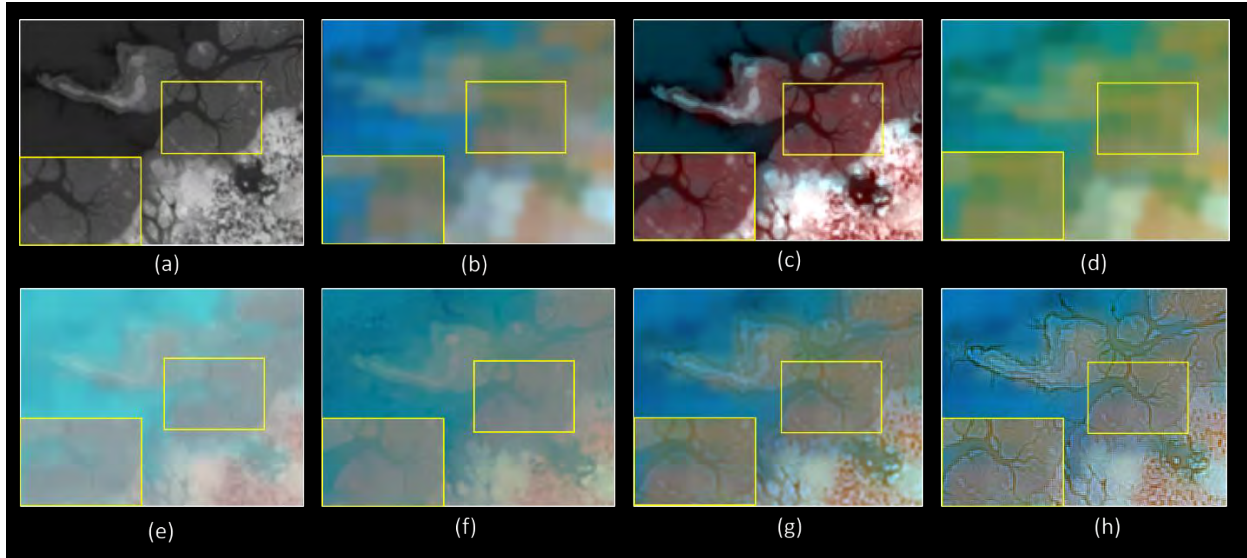


Figure-6.23. Comparison of Fusion Outcomes for Landsat-8 Thermal Infrared and MODIS Visible Images across Different Feature Targets (a) Thermal Infrared Image (b) Resampled Multispectral Image (c) BT (d) IIHS (e) PCA (f) SOE (g) FGIF (h) CLIM

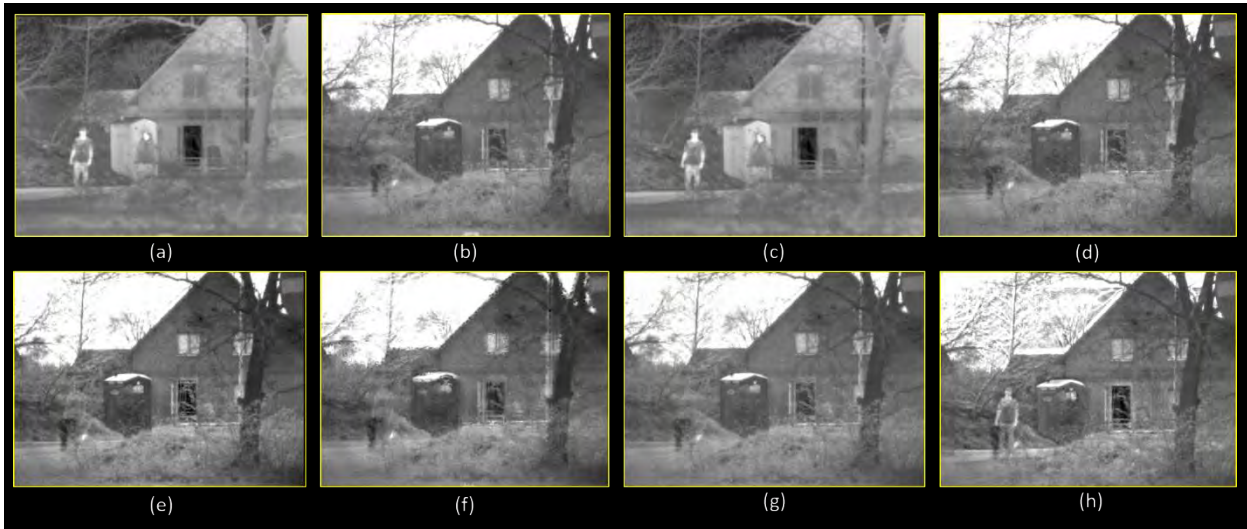


Figure-6.24. Assessment of Fusion on TNO Dataset (a) Infrared Image (b) Visible Image (c) BT (d) IIHS (e) PCA (f) SOE (g) FGIF (h) CLIM

### B. Quantitative Evaluation

The assessment of the fusion between INS-2TD thermal infrared and MODIS visible images involves evaluating merging performance using different techniques, which includes the analysis of the quality metrics based on reference points. To measure the structural information present in

the fused data, the SSIM (Structural Similarity Index) metric (Wang et al., 2004) is calculated, utilizing a thermal infrared image with a comparatively high spatial resolution as the reference basis. SSIM stands as a widely recognized quality metric employed for assessing structural distortion between two images. Spectral quality metrics, including Correlation Coefficient (CC) (Zhang et al., 2008), Root Mean Square Error (RMSE) (Misra et al., 2012a), Spectral Angle Mapper (SAM) (Shahdoosti & Ghassemian, 2014), Relative Dimensionless Global Error in Synthesis (ERGAS) (Alparone et al., 2007), and Universal Image Quality Index (UIQI) (Wang & Bovik, 2002), predominantly assess the spectral distortion arising from the image fusion process. RMSE is utilized to examine the disparity between the reference and fused images by directly calculating the variation in pixel values. CC assesses the resemblance of spectral characteristics between the reference and fused images. SAM calculates the spectral angle by examining the pixel vectors of both the reference and fused images. ERGAS assesses the quality of the fused image by considering the normalized average error for each band of the processed image. UIQI quantifies the extent of transformation of pertinent data from the reference image to the fused image. In this instance, the reference for assessing spectral image quality parameters is selected as the resampled MODIS visible image. Table-6.9 presents the calculated quality metrics for the reference image for each method of image fusion. Furthermore, contemporary deep learning techniques like VIF-Net (Hou et al., 2020) and CMFA-Net (Ding et al., 2021) are taken into account for a quantitative comparison. FGIF transform exhibits slightly superior performance compared to others in SSIM. Nevertheless, the spatial performance of CLIM is observed to be nearly comparable to FGIF, and the SSIM metric value for CLIM suggests a close correspondence. CLIM demonstrates a distinct superiority over advanced image fusion techniques, especially concerning radiometry and spectral image quality parameters.

The assessment of fusion quality, relying on reference images, is also performed for the fusion outcomes of Landsat-8 thermal infrared and MODIS multispectral visible images employing diverse techniques. Table-6.10 provides the quality metrics and measurements acquired for all methods of infrared-visible image fusion. Despite the superior performance of deep learning fusion methods like VIF-Net and CMFA-Net, the fusion quality metrics suggest a deficiency in preserving the spectral properties of the multispectral visible image. On the contrary, the suggested method, CLIM, exhibits superior metrics in both spatial and spectral image quality parameters when compared to alternative image fusion methods. Image fusion metrics based on reference are

calculated for the TNO image dataset using various fusion techniques. As indicated in Table-6.11, CLIM's fusion performance exceeds that of other fusion methods, even when applied to the TNO multi-band collection.

Table-6.9. Assessment of Reference Quality Parameters for Thermal-Visible Image Fusion in the Pair of Thermal Infrared from INS-2TD and Multispectral Visible Images from MODIS

<b>Fusion Methods</b>	<b>SSIM</b> ↑	<b>CC</b> ↑	<b>RMSE</b> ↓	<b>SAM</b> ↓	<b>ERGAS</b> ↓	<b>UIQI</b> ↑
BT [9]	0.8112	0.3288	6.8426	10.3146	5.3415	0.0421
IIHS [10]	0.6120	0.3541	8.6453	10.2813	5.2412	0.1532
PCA [11]	0.7001	0.3962	12.3244	9.9471	5.1675	0.1832
SOE [16]	0.6224	0.6814	7.0312	8.3521	4.1436	0.2531
FGIF [18]	<b>0.8313</b>	0.7455	6.2313	7.5213	3.1524	0.2735
VIF-Net [26]	0.8123	0.7564	6.0133	7.7144	3.2341	0.2876
CMFA-Net [27]	0.8032	0.7432	6.2342	7.8233	3.3421	<b>0.2985</b>
CLIM	0.8252	<b>0.7771</b>	<b>5.1432</b>	<b>7.4552</b>	<b>2.1357</b>	0.2891

Table-6.10. Evaluation of Reference Quality Parameters for Thermal-Visible Image Fusion in the Pair of Thermal Infrared Images from Landsat-8 and Multispectral Visible Images from MODIS

<b>Fusion Methods</b>	<b>SSIM</b> ↑	<b>CC</b> ↑	<b>RMSE</b> ↓	<b>SAM</b> ↓	<b>ERGAS</b> ↓	<b>UIQI</b> ↑
BT [9]	0.8231	0.3144	14.6456	12.5421	6.4244	0.0981
IIHS [10]	0.5135	0.5537	11.4253	11.3233	5.2132	0.1234
PCA [11]	0.6742	0.5973	10.5241	10.4321	4.4211	0.1462
SOE [16]	0.6942	0.7014	8.5121	9.4211	3.3211	0.2721
FGIF [18]	0.8321	0.7572	7.5531	7.5321	3.1743	0.2864
VIF-Net [26]	0.8098	0.7223	7.3433	7.6543	3.5612	0.2944
CMFA-Net [27]	0.7983	0.7234	7.4543	7.8722	4.0643	0.2783
CLIM	<b>0.8423</b>	<b>0.7984</b>	<b>4.3146</b>	<b>6.3163</b>	<b>2.5711</b>	<b>0.3033</b>

Table-6.11. Assessment of Reference Quality Parameters for Thermal-Visible Image Fusion in the dataset containing Long-Wave Infrared (LWIR) and Visible Images from TNO

<b>Fusion Methods</b>	<b>SSIM</b> ↑	<b>CC</b> ↑	<b>RMSE</b> ↓	<b>SAM</b> ↓	<b>ERGAS</b> ↓	<b>UIQI</b> ↑
BT [9]	0.8362	0.3523	12.4253	12.5672	6.6134	0.1143
IIHS [10]	0.6624	0.4976	11.7345	10.9324	6.2325	0.1256
PCA [11]	0.6824	0.5532	9.7234	9.4511	5.3463	0.1353
SOE [16]	0.7452	0.6723	8.2134	8.6234	4.4513	0.2987
FGIF [18]	0.8411	0.7361	7.6532	7.4366	3.7712	0.3867
VIF-Net [26]	0.8523	0.7543	6.9812	7.2311	3.3244	0.4023
CMFA-Net [27]	0.8123	0.7124	7.1243	8.0421	3.6531	0.3234
<b>CLIM</b>	<b>0.8625</b>	<b>0.7845</b>	<b>4.5123</b>	<b>7.1352</b>	<b>3.1642</b>	<b>0.4726</b>

The image quality metrics without reference provide an impartial assessment of fusion outputs from different techniques in the thermal-visible domain. The evaluation of CLIM-based image fusion performance, relative to other fusion techniques, is conducted using spectral distortion indices  $D_s$  and  $D_\lambda$  (Misra et al., 2023b). Moreover, we calculate the QNR (Quality with No Reference) image index, consisting of  $D_s$  and  $D_\lambda$ , to assess the enhancement in image quality achieved through multispectral radiometric consistency (Alparone et al., 2008). Table-6.12 displays the values of no-reference image metrics for different image fusion techniques applied to the INS-2TD Infrared-MODIS Visible Image combination. All three image quality parameters confirm that CLIM surpasses its counterparts in effectively combining high spatial resolution thermal radiance details with multispectral visible images. The no-reference quality parameters also suggest that the CLIM fused image displays minimal spectral distortion.

Similarly, no-reference image quality metrics are calculated for fused images of Landsat-8 Thermal and MODIS Visible using different techniques. Table-6.13 presents the no-reference metrics of fusion methods for the purpose of comparison. Consistently, CLIM demonstrates superior performance compared to other state-of-the-art techniques in all three no-reference image quality metrics.

Table-6.12. Evaluation of No-Reference Image Quality Metrics for Fusion of INS-2TD  
Infrared-MODIS Visible Images

Fusion Methods	$D_s$ ↓	$D_\lambda$ ↓	QNR ↑
BT [9]	0.193	0.378	0.525
IIHS [10]	0.183	0.251	0.511
PCA [11]	0.112	0.170	0.572
SOE [16]	0.087	0.193	0.568
FGIF [18]	0.078	0.147	0.634
VIF-Net [26]	0.077	0.145	0.604
CMFA-Net [27]	0.075	0.163	0.512
CLIM	<b>0.066</b>	<b>0.133</b>	<b>0.638</b>

Table-6.13. Assessment of No-Reference Image Quality Metrics for Fusion of Landsat-8  
Infrared-MODIS Visible Images

Fusion Methods	$D_s$ ↓	$D_\lambda$ ↓	QNR ↑
BT [9]	0.152	0.238	0.455
IIHS [10]	0.145	0.242	0.412
PCA [11]	0.142	0.215	0.545
SOE [16]	0.123	0.182	0.563
FGIF [18]	0.116	0.157	0.585
VIF-Net [26]	0.102	0.145	0.598
CMFA-Net [27]	0.121	0.190	0.559
CLIM	<b>0.098</b>	<b>0.123</b>	<b>0.723</b>

The processing time for diverse image fusion methods is computed across various datasets, and Table-6.14 showcases the time needed to generate fused images from image pairs. It is observed that the processing time for the Landsat-8/MODIS visible image pair is greater compared to other pairs, attributed to its larger image size. The TNO dataset, being relatively small, exhibits a shorter processing time. The proposed CLIM method's processing time is not the minimum, but it surpasses that of the majority of other image fusion techniques.



Table-6.14. Processing Time (in seconds) of Image Fusion Methods for Infrared-Visible Image Pairs

Image Pair	BT [9]	IIHS [10]	PCA [11]	SOE [16]	FGIF [18]	VIF-Net [26]	CMFA-Net [27]	CLIM
Landsat-8 TIR/MODIS Visible	32	46	53	67	62	68	72	42
INS-2TD TIR/MODIS Visible	16	26	35	41	44	48	52	24
TNO Dataset	3	4	8	14	18	23	26	6

### C. Ablation Study

The CLIM method consists of various components, each holding importance at distinct phases of the infrared-visible image fusion process. Among these, enhancing thermal images is an important step in image processing, where features are more clearly outlined. The CoF acts as a filter that preserves boundaries and plays a vital role in extracting abstract information from thermal infrared images. We conduct two experimental analyses, denoted as  $CLIM^{NOENH}$  with thermal image enhancement turned off, and  $CLIM^{NOCOF}$  with the co-occurrence filter step disabled, in the proposed fusion method CLIM. Table-6.15 showcases the fusion performance of  $CLIM^{NOENH}$ ,  $CLIM^{NOCOF}$ , and CLIM for the Landsat-8 Thermal Infrared (TIR) and MODIS Visible image pair. The evaluation metrics suggest that the fused output from CLIM surpasses both  $CLIM^{NOENH}$  and  $CLIM^{NOCOF}$ .

Table-6.15. Comparison of Fusion Performance for Landsat-8/MODIS Visible Image Pair among CLIM<sup>NOENH</sup>, CLIM<sup>NOCOF</sup>, and CLIM

Fusion Quality Metrics	CLIM <sup>NOENH</sup>	CLIM <sup>NOCOF</sup>	CLIM
SSIM ↑	0.7673	0.7532	0.8423
CC ↑	0.6893	0.6561	0.7984
RMSE ↓	5.2311	5.9823	4.3146
SAM ↓	7.4225	7.8761	6.3163
ERGAS ↓	3.8745	3.9087	2.5711
UIQI ↑	0.2987	0.2899	0.3033

#### 6.2.4. Summary and Avenues for Future Research

CLIM emerges as a robust and powerful image fusion technique with the capability to seamlessly combine relatively high spatial resolution broadband thermal infrared images with low spatial resolution multispectral visible remote sensing images. The proposed methodology, CLIM, effectively integrates the results of the Co-occurrence Smooth Filter with a LoG sharpened image in an optimized manner, successfully extracting clear boundary features from the thermal image. The enhancement of the multispectral visible image is achieved through the IHS color space transform and CLAHE enhancement, ensuring the optimal injection of texture. The resulting fused output from CLIM represents a comprehensive image that integrates spatial details from the thermal channel and spectral information from visible channels. CLIM offers detailed information, significantly improving the delineation of distinct feature targets for the interpretation of remote sensing images. By employing both visual examination and quantitative assessment against contemporary image fusion techniques, CLIM showcases its superiority in tasks involving the fusion of multi-modal images. In the future, exploring the fine-tuning of CLIM algorithm parameters for various classes of multi-modal image pairs, such as Synthetic Aperture Radar (SAR) and optical remote sensing image fusion, holds promise.

## **6.3. Multi-Modal Image fusion and processing using Venus remote sensing images**

### **6.3.1. Introduction**

The section introduces an innovative end-to-end framework for processing visible images of Venus. The primary goal is to create a data processing pipeline for Venus's visible images, marking the initial effort in this field. Addressing the key challenges in the Venus visible image processing workflow, the framework involves denoising and enhancing the visible image, sub-pixel level co-registration of the enhanced visible planetary image with radar topographic data, and fusion of the multi-modal registered datasets to generate a representative Venus image. The proposed image processing workflow is designed to be adaptable for other Venus visible images (Misra et al., 2024b).

A radiometric correction technique is developed for denoising the raw image using the Tri-State median filter (Chen et al., 1999). The denoised image undergoes enhancement using CLAHE (Reza et al., 2004) to improve visual image quality. The visible enhanced image is co-registered with Magellan radar topographic data utilizing SA-SIFT (Lowe, 2004) and Motion Smoothness Constraint (MSC) (Bian et al., 2017; Misra et al., 2021a) for outlier removal. The Venus visible enhanced co-registered image is fused with radar topographic data using a novel image fusion algorithm employing guided filters (He et al., 2012). The quality of the processed visible image is assessed at each step of the proposed pipeline.

SNR and Noise Estimation Level are calculated for the denoised image and raw visible image (Atkinson et al., 2005). Additionally, various image quality metrics (Rubel et al., 2022) are employed for comprehensive comparisons of planetary image processing results. The image co-registration performance is visually assessed by overlaying multi-modal images and quantitatively evaluated at different regions of Venus. The final fused Venus image is utilized for planetary feature identification, fuzzy image segmentation to gain insights into diverse surface features, and is validated against publicly available annotated radar image maps of Venus.

### 6.3.2. Methodology Developed

#### A. Image Denoising using Tri-State Median (TSM) Filter

Planetary images commonly exhibit various types of noise issues (Plebani et al., 2022). The presence of diverse statistical noise in sparse image patches necessitates the use of robust image denoising techniques (Yuan & He, 2021). Venus's visible planetary remote sensing images, in particular, often suffer from impulse noise and streak noise, as illustrated in Figure-6.25 (a). To comprehensively study the characteristics of the planetary surface, it is crucial to suppress these noises while preserving the planetary features and their contextual information. In our approach, the Tri-State Median (TSM) filter is selected to address streak and impulse noise. The TSM filter, a non-linear filter, has been identified for its ability to retain essential image feature details while effectively suppressing streak and impulse noise in planetary remote sensing images (Ramachandran & Kishorebabu, 2019). The TSM filter integrates the standard median (SM) filter and the center-weighted median (CWM) filter into a noise detection framework to assess whether a pixel is corrupted before applying unconditional filtering (George et al., 2018). The TSM filter for an image pixel  $X_{ij}$  is defined as follows:

$$TSM_{ij} = \begin{cases} X_{ij} & T \geq d_1 \\ CWM_{ij} & d_2 \leq T < d_1 \\ SM_{ij} & T < d_2 \end{cases} \quad (6.26)$$

where  $T$  is the threshold,  $d_1 = |X_{ij} - SM_{ij}|$  and  $d_2 = |X_{ij} - CWM_{ij}|$

Furthermore, the dark background noise is calculated directly from the image and subtracted to eliminate any remaining deviation noise in the Venus' visible image. Figure-6.25 (b) displays the TSM output with background subtraction, showcasing the effective denoising of the Venus' visible image and enhanced visibility of planetary surface features.

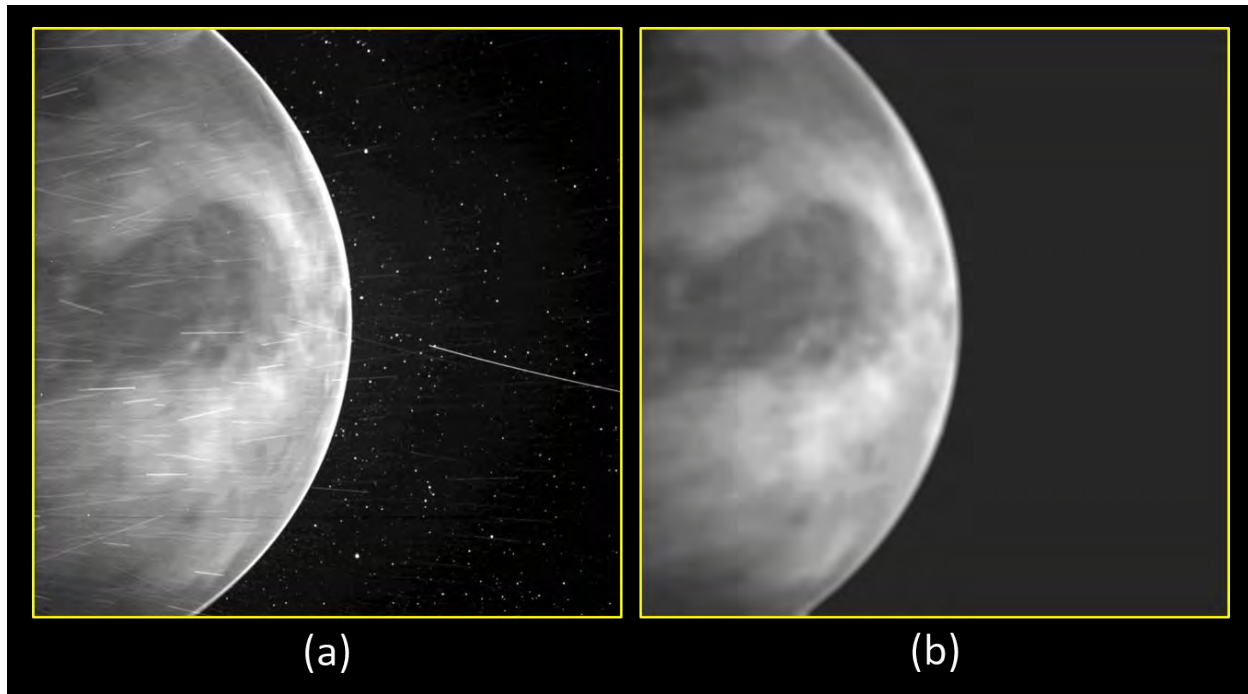


Figure-6.25. Performance of Venus' Visible Image Denoising (a) Raw Venus' Visible Image (b) Denoised Venus' Visible Image

#### B. *Image Enhancement using CLAHE*

The denoised image of Venus appears somewhat blurry, posing a challenge to enhance it, especially under low-light imaging conditions, for more effective analysis of planetary surface features (Samarasinha & Larson, 2014; Hao & Xu, 2021). To address this, we apply the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique [50]. CLAHE proves effective in enhancing local features and sharpening the denoised Venus visible image, thereby improving perceptibility. The image enhancement through CLAHE is designed to restrict the amplification of contrast stretch based on a predefined histogram limit (Chang et al., 2018). An essential parameter in this process is the window size around a pixel for which the histogram is equalized. For CLAHE enhancement of the Venus denoised image, a window size of 256 is empirically determined to provide optimal visual appearance. Figure-6.26 illustrates the performance of image enhancement on Venus's denoised visible image, showcasing the noticeable improvement in contrast for local features at various locations.

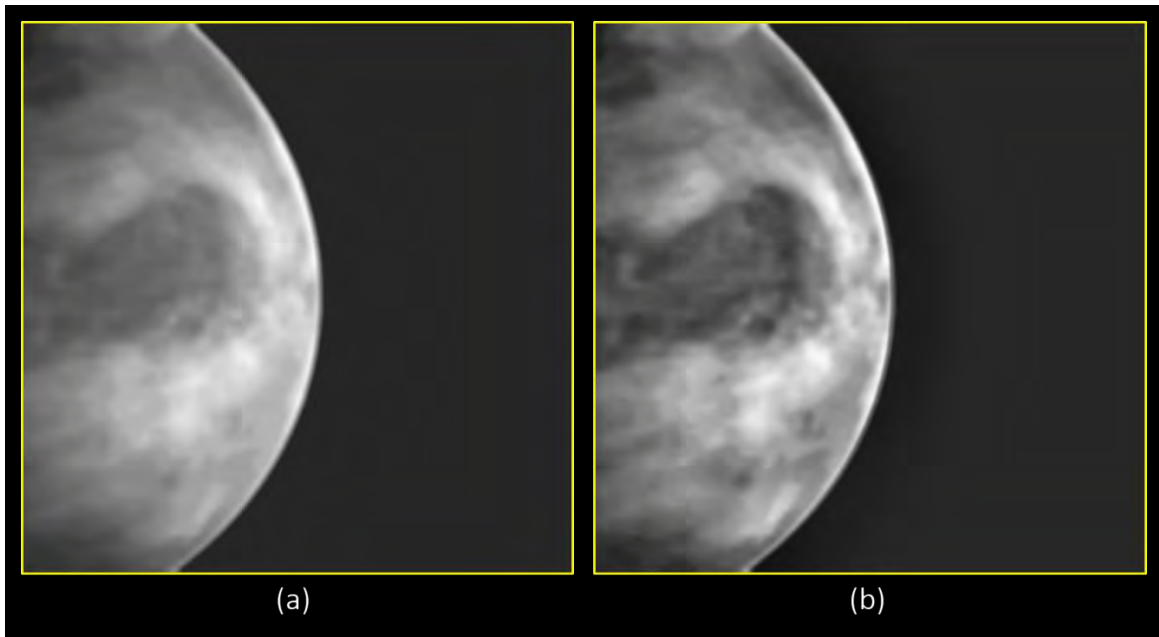


Figure-6.26. Visual Assessment of Venus' Image Enhancement (a) Denoised Venus' Image (b) Enhanced Venus' Image

*C. Registration of multi-modal images using SA-SIFT and MSC*

Multi-modal datasets in remote sensing provide complementary information that, when merged, can produce a more comprehensive and representative image. In the case of Venus, the visible image offers unique texture information, while the Magellan mission maps the planet in the microwave region of the electromagnetic spectrum. To effectively utilize both data sources, it is essential to co-register the multi-modal Visible-Radar image pair at a sub-pixel level. Figures 6.27(a) and 6.27(b) depict the multi-modal image pair of the planet Venus for the image registration task. Achieving sub-pixel spatial alignment between the visible and radar images is challenging due to the significant radiometric differences. For precise feature matching between the multi-modal image pair, a simulated image is generated from Venus' radar image using an inverse look-up table, ensuring similarity in planetary features between the visible and radar remote sensing images. Figure-6.27(c) illustrates the simulated image created for the multi-modal image co-registration task.

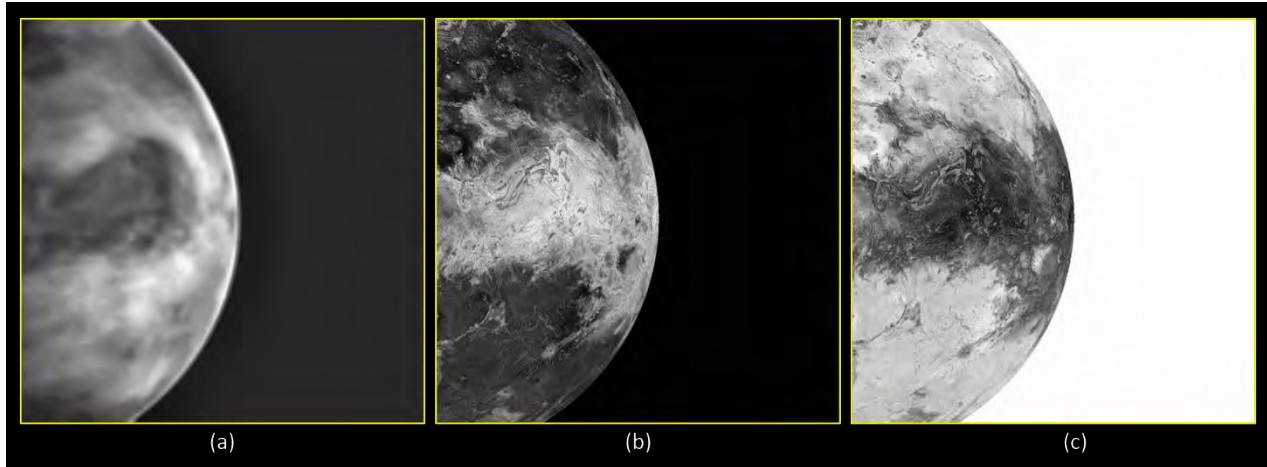


Figure-6.27. Multi-modal Pair of Venus' Images (a) Denoised and Enhanced Visible Image (b) Radar Image from Magellan Mission (c) Simulated Venus' Image from Radar Data for Image Registration

The registration of multi-modal planetary images in our scenario employs the Segmented Affine Scale Invariant Feature Transform (SA-SIFT) technique. The simulated Venus' radar image serves as the reference, and the Venus' visible image is rectified in relation to this reference. The SIFT algorithm begins by constructing scale space extrema based on a Gaussian Kernel in the visible-radar image pair (Misra et al., 2021a).

It is observed that SA-SIFT-derived matched feature points exhibit a considerable number of false matches. This is primarily due to inaccuracies in the estimation parameters of the transformation caused by erroneous matches. Motion smoothness-based estimation effectively discerns between true and false matches, contributing to the establishment of accurate correspondences between the input and reference images (Yu et al., 2018). The notion behind motion smoothness is that a small neighborhood surrounding a true match is expected to appear in the same location in another image. Conversely, the neighborhood of a false match is assumed to be at geometrically distinct locations. Smooth motion implies a collective movement of neighboring pixels and features (Zheng et al., 2018). The MSC is formulated as a statistical model that identifies and rejects false matches in multi-modal planetary remote sensing images. MSC is characterized by its simplicity, robustness, and superior performance compared to techniques like RANSAC (Choi et al., 1997) and other matched keypoint optimization methods, particularly in certain types of remote sensing image pairs such as Venus' visible and radar planetary remote sensing images.

Following the extraction of pruned matched control points with SA-SIFT, a segmented affine transformation model is employed to calculate affine values for different image segments within the planetary dataset. This process aims to enhance the geometric grid for each segment, mitigating local distortions. The parameters of the segmented affine transform are then utilized to update the geometric grid for each planetary image segment, enabling the resampling of the input Venus' visible image. This results in a sub-pixel co-registered output. The affine transform for a Venus' visible image segment  $k$  is expressed as:

$$aff(k) = \{a_{0k}, a_{1k}, a_{2k}, a_{3k}, a_{4k}, a_{5k}\} \quad (6.27)$$

where  $k$  represents the segment number and  $a_{ik}$  denotes the affine parameters.

#### D. Fusion with Radar Topographic Data

The enhanced and co-registered Venus' visible image is now ready for fusion with the corresponding Venus' radar topographic data of the same region. The fusion process involves the generation of a smooth pixel from the higher resolution Venus' radar image using the Guidance Image Filter (GIF) (Caraffa et al., 2015). GIF is a local linear filter controlled by two parameters: regularization ( $\epsilon$ ) and window radius ( $r$ ) (He et al., 2012; Misra et al., 2022b). The enhanced Venus' visible image serves as the guidance image ( $G$ ) to smooth the corresponding Venus' radar image. The guidance image  $G$  is used as the input for GIF (Liu et al., 2016; Prema & Arivazhagan, 2022), and the smoothed radar image becomes the input for the subsequent step in the proposed image fusion algorithm. The output  $f$  at a pixel  $i$  is defined as follows:

$$f_i = \sum_j w_{ij}(G)s_j \quad (6.28)$$

where  $w_{ij}$  is the weight between pixel  $j$  and  $i$ ,  $s_j$  is the value of the  $j$ th input pixel,  $j$  denotes the pixel index in the window  $W$ , and  $i$  is the center of the window.

$$w_{ij}(G) = \frac{1}{|W|^2} \sum_{(i,j) \in W} \left[ 1 + \frac{(G_i - \mu_w)(G_j - \mu_w)}{\sigma_w^2 + \epsilon} \right] \quad (6.29)$$

$W$  is the square area with a radius of  $r$ , where  $\mu_w$  and  $\sigma_w$  are the mean and variance of the guidance image  $G$  in window  $W$ , and  $|W|$  is the number of pixels in  $W$ .

In the guided filter, we presume that  $f$  is a linear transformation of  $G$  in a window  $w_k$  centered at a pixel  $k$ .

$$f_i = cG_i + d \quad (6.30)$$

Here,  $c$  and  $d$  are considered constant linear coefficients within the window radius.



The Component Ratio (*C. R.*) is calculated by comparing Venus' radar image with the output of the guidance filter. This computation eliminates the spectral and topographical contrast of the high-resolution Venus' radar image while preserving sharp edges for integration with the lower resolution Venus' visible image. Equations 6.31 and 6.32 outline the key steps of the proposed fusion algorithm, involving the computation of *C. R.* and the blending of *C. R.* with the co-registered Venus' visible image to produce the fused image.

$$C.R. = I_{Radar}/I_{GIF} \quad (6.31)$$

where,  $I_{Radar}$  represents the Venus' Radar Image,

$I_{GIF}$  represents the Guidance Image Filter Output.

$$I_{fusion} = C.R.* I_{Visible} \quad (6.32)$$

where,  $I_{fusion}$  = Fused Image,  $I_{Visible}$  = Venus' Visible Image

The aforementioned fusion process for planetary remote sensing images is implemented on the Venus' multi-modal image pair. Figure-6.28 displays the co-registered Venus' visible enhanced image and the resulting fused image. The fused output enhances the delineation of various morphological features on the planet, providing comprehensive details about its surface characteristics. In summary, the fused image successfully preserves the spectral characteristics post-fusion.

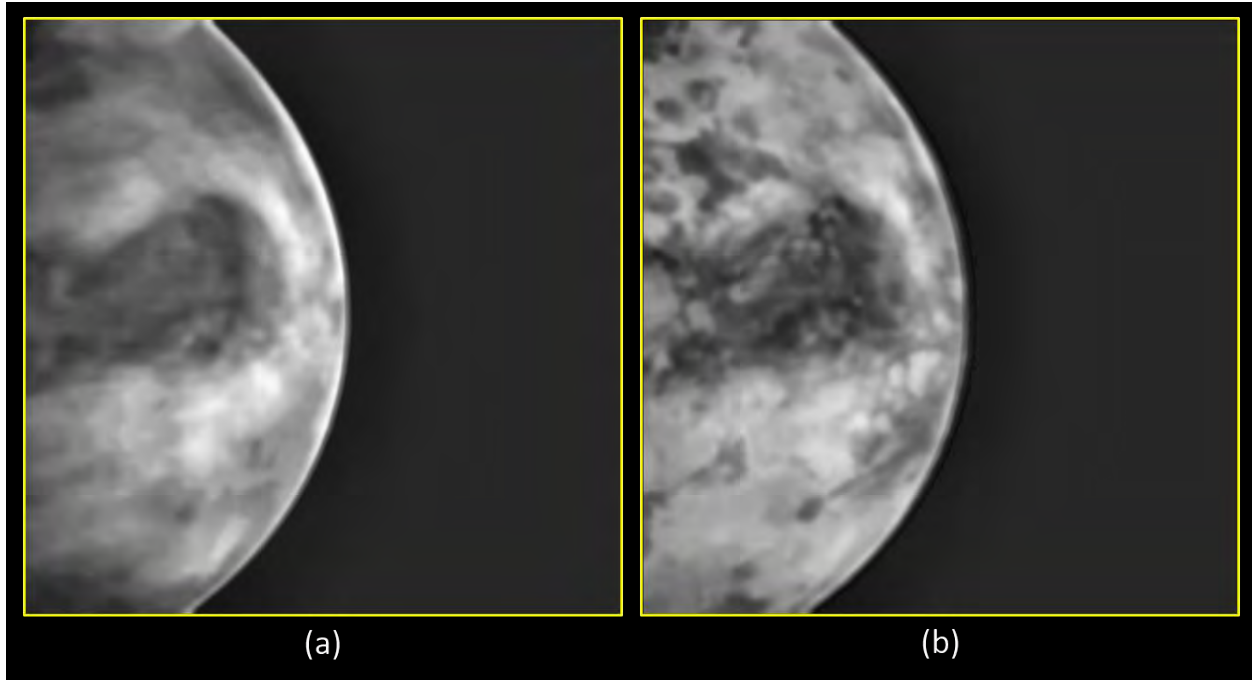


Figure-6.28. Venus' Image Fusion Results (a) Enhanced Co-Registered Visible Image (b) Fused Image

#### *E. Venus' Image Processing Workflow*

The proposed methodology has been transformed into a planetary data processing pipeline, as depicted in Figure-6.29. This neoteric workflow begins with taking Venus' visible raw image as input, followed by the necessary radiometric processing to de-noise and enhance the raw planetary image. The subsequent step involves spatially aligning the visible and radar images over Venus through the proposed SA-SIFT feature matching with MSC outlier removal technique. The resulting multi-modal image pair is then processed for fusion using the innovative algorithm with GIF. The resulting Venus' fused image proves effective in delineating surface morphological features for planetary studies and scientific analysis.

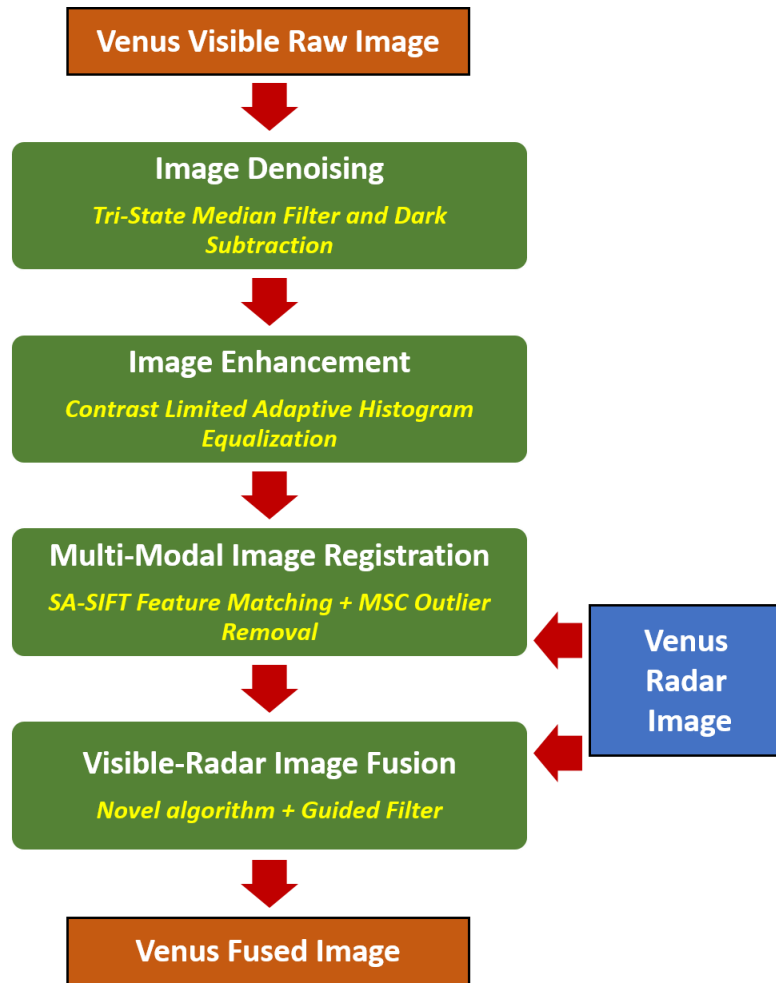


Figure-6.29. Proposed Advanced Workflow for Venus' Image Processing

### 6.3.3. Venus Processed Image Quality Assessment

#### A. SNR and Noise Estimation Evaluation

The denoised image of Venus effectively eliminates undesirable streak noise and impulse noise inherent in the raw image. In Figure-6.7 (a), a preview of the denoised output for Venus' visible image is presented, with a specific region of interest highlighted in a yellow-colored box for subsequent qualitative and quantitative assessment. This yellow box is then cropped and zoomed to a 4X level for both the raw and denoised images to facilitate clearer data visualization. As depicted in Figure-6.30 (b), various types of noise dominate the raw image, and this noise is notably diminished in the denoised output, as illustrated in Figure-6.30 (c). The denoising process effectively suppresses noise in homogeneous areas while preserving edge and texture information in Venus' visible image.

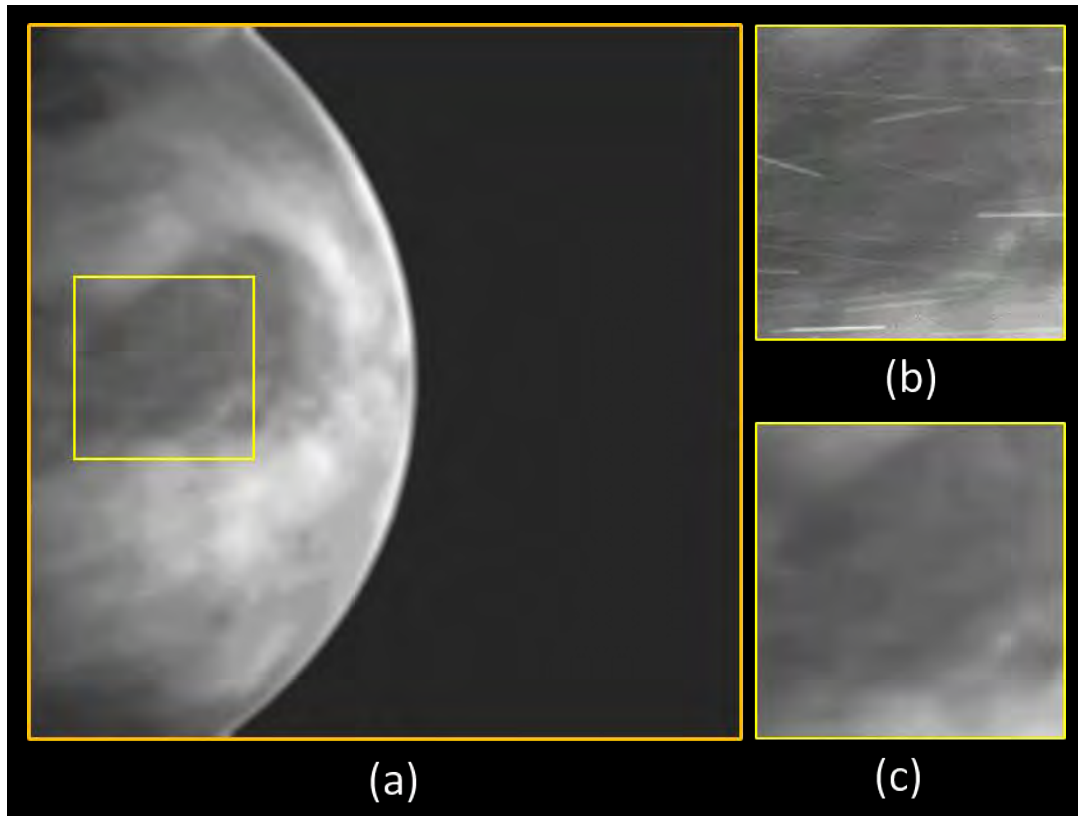


Figure-6.30. Quality Evaluation of Venus' Visible Image Region of Interest (a) Preview of Denoised Image (b) 4X Zoomed Raw Visible Image (c) 4X Zoomed Denoised Visible Image

Furthermore, the SNR and the noise estimation level  $\sigma$  are computed for both Venus' visible raw image and the corresponding denoised image (Chen et al., 2015). Table-6.16 provides the SNR and  $\sigma$  values for both the raw and visible images. The results of this measurement highlight an enhancement in SNR, accompanied by a reduction in the noise estimation level for the denoised image.

Table-6.16. Comparison of SNR and Noise Level

Venus Visible Image Type	Noise Level $\sigma$ ↓	SNR ↑
Raw Image	4.9953	4.3076
Denoised Image	4.5361	5.7930

### B. No-Reference Image Quality Assessment

As a visible reference image is unavailable for the corresponding region on Venus, the Blind/Reference-less Image Spatial Quality Evaluator (BRISQUE) (Mittal et al., 2012) score is computed for both denoised and enhanced images. Additionally, entropy, as an image quality metric, is calculated to signify the amount of information content gained in the image after enhancement (Liu et al., 2014). The region overlapping with Venus is extracted from both denoised and enhanced datasets, encompassing diverse planetary morphological features. Figure-6.31 displays the Venus denoised-enhanced image pair for BRISQUE evaluation. BRISQUE employs scene statistics to quantify distortion, providing a comprehensive measure of image quality, where a lower score indicates better image quality. Table-6.17 presents the BRISQUE scores, suggesting that the Venus visible enhanced image exhibits superior image quality compared to the denoised counterpart.

Table-6.17. Performance of Venus' Visible Image Quality without Reference

S. No.	Venus' Visible Image	BRISQUE Model Score ↓	Entropy ↑
1.	Denoised Image	73.8398	6.9067
2.	Enhanced Image	67.4371	7.1718

### C. Venus' Visible Image Co-Registration Accuracy

The qualitative assessment of planetary image registration involves swiping the Venus visible co-registered image across various regions of interest on the Venus radar image, as illustrated in Figure-6.32. Different locations on Venus are cropped to create multiple regions of interest for geometric evaluation. The zoomed 2X multi-modal image view allows a closer examination of the achieved registration accuracy. The results clearly demonstrate that Venus' visible image aligns with sub-pixel level spatial accuracy at various morphological features on the planet when compared to the high-resolution Venus radar image reference.

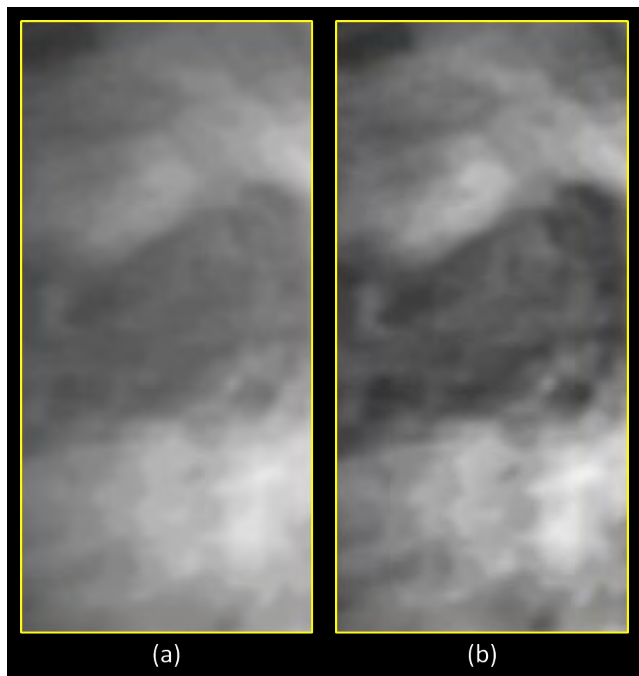


Figure-6.31. Venus' Visible Cropped Image Pair for Entropy and BRISQUE Evaluation (a) Denoised Image (b) Enhanced Image

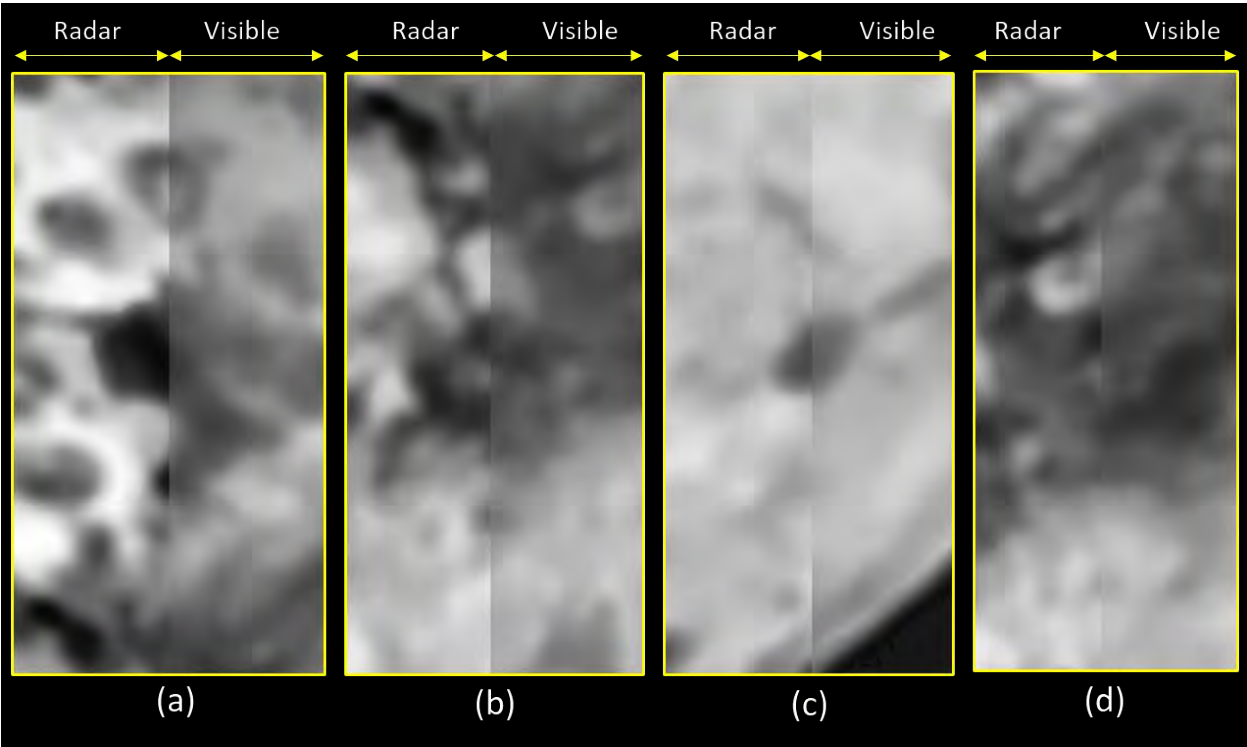


Figure-6.32. Horizontal Swipe Visual Assessment of Venus' Image Co-Registration at Different Regions of Interest (RoI) (a) RoI-1 Swipe (b) RoI-2 Swipe (c) RoI-3 Swipe (d) RoI-4 Swipe

To assess the effectiveness of the planetary multi-modal image registration technique utilizing SA-SIFT and MSC outlier removal, the Venus visible enhanced registered images are once again correlated with SA-SIFT+MSC using the simulated Venus radar image as a reference. The RMSE is then calculated using the pruned control matched points obtained from the developed framework. RMSE is commonly employed as a suitable metric for evaluating the performance of a planetary image co-registration model (Misra et al., 2012a). The registration accuracy is evaluated at four Regions of Interest (RoIs) indicated in Figure-6.32. Table-6.18 presents the output pixel shift in both directions and the RMSE computed at RoIs for Venus' visible enhanced co-registered image concerning the simulated Venus radar reference frame. The RMSE computation is expressed as follows:

$$RMSE = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|^2\right)}, \quad (6.33)$$

where N denotes the total number of matched points,

$X_i$  represents the  $(x_i, y_i)$  coordinates in the Venus' simulated radar image (Reference),

$\hat{X}_i$  is the  $(\hat{x}_i, \hat{y}_i)$  estimated coordinates based on the final transformation model of the output Venus' visible enhanced registered image.

Table-6.18. Evaluation of Registration Accuracy for Venus' Visible Co-Registered Image with Respect to Venus' Radar Image Reference

S. No.	Region of Interest (RoI)	Shift in Vertical Direction (in pixel)	Shift in Horizontal Direction (in pixel)	RMSE (in pixel)
1.	RoI-1	0.45	0.32	0.37
2.	RoI-2	0.46	0.41	0.42
3.	RoI-2	0.42	0.36	0.39
4.	RoI-4	0.48	0.46	0.45

The planetary image registration technique proposed in this study is compared against state-of-the-art feature-based image registration methods, including Harris + MLESAC, SIFT + RANSAC, ORB + MG, SURF + MLESAC, and KAZE + RANSAC. The evaluation is based on the computation of RMSE for RoIs and listed in Table-6.19. The results indicate the superiority of our proposed method over other feature-based image registration approaches.

Table-6.19. Comparative RMSE Analysis of Planetary Image Registration Methods

S. No.	Region of Interest (RoI)	ORB + MG	Harris + MLESAC	SURF + MLESAC	SIFT + RANSAC	KAZE + RANSAC	Proposed Method
1.	RoI-1	0.88	0.81	0.69	0.65	0.82	0.37
2.	RoI-2	0.74	0.89	0.74	0.55	0.69	0.42
3.	RoI-2	0.71	0.73	0.64	0.68	0.66	0.39
4.	RoI-4	0.56	0.76	0.66	0.51	0.78	0.45

*D. Quantitative Assessment of Fusion and Identification of Surface Features*

In the absence of a reference image for Venus’ fusion results, an evaluation of the proposed fusion method is conducted through a no-reference image assessment. This involves estimating the entropy of the image (Chibani & Houacine, 2002), employing BRISQUE (Fernandez-Beltran et al., 2017), a no-reference image index that gauges scene statistics to quantify the potential naturalness of planetary remote sensing images. Additionally, the structural similarity (SSIM) metric (Wang et al., 2004) is computed, measured with respect to the Magellan radar topography image. The results of these quantitative fusion quality parameters for both Venus’ visible enhanced image and the fused image are presented in Table-6.20. The findings indicate that the fused image outperforms in various fusion image quality parameters.

Table-6.20. Evaluation of Venus’ Image Fusion Quality

S. No.	Venus Visible Image Type	BRISQUE ↓	Entropy ↑	SSIM ↑
1.	Enhanced Image	67.4371	7.1718	0.6923
2.	Fused Image	65.6324	7.2995	0.9048

The Venus’ fused image serves as a valuable tool for the analysis of surface features, and for feature identification, a comparison is made between the processed visible image and Venus’ radar nomenclature map (Batson et al., 1994). Figure-6.33 presents the Venus’ fused image annotated map alongside Venus’ radar image nomenclature map. Notably, Aphrodite Terra, an elevated large



area region situated near the equator of Venus (Basilevsky et al., 2012), appears as a dark patch in the fused image compared to other morphological features. This observation suggests that the relatively low temperature in the region is indicative of its higher elevation. Ovda Regio, a Venusian crustal plateau located in the western highland region of Aphrodite Terra, is also mapped in the fused image. Additionally, Tahmina Planitia, a low plain region of Venus below Aphrodite Terra, appears brighter in the fused image, indicating a relatively higher temperature even during nighttime. Lastly, Unelanuhi Dorsa, a region above Ovda Regio, exhibits low intensity measurements in the fused image, signifying a cooler region compared to Tahmina Planitia.

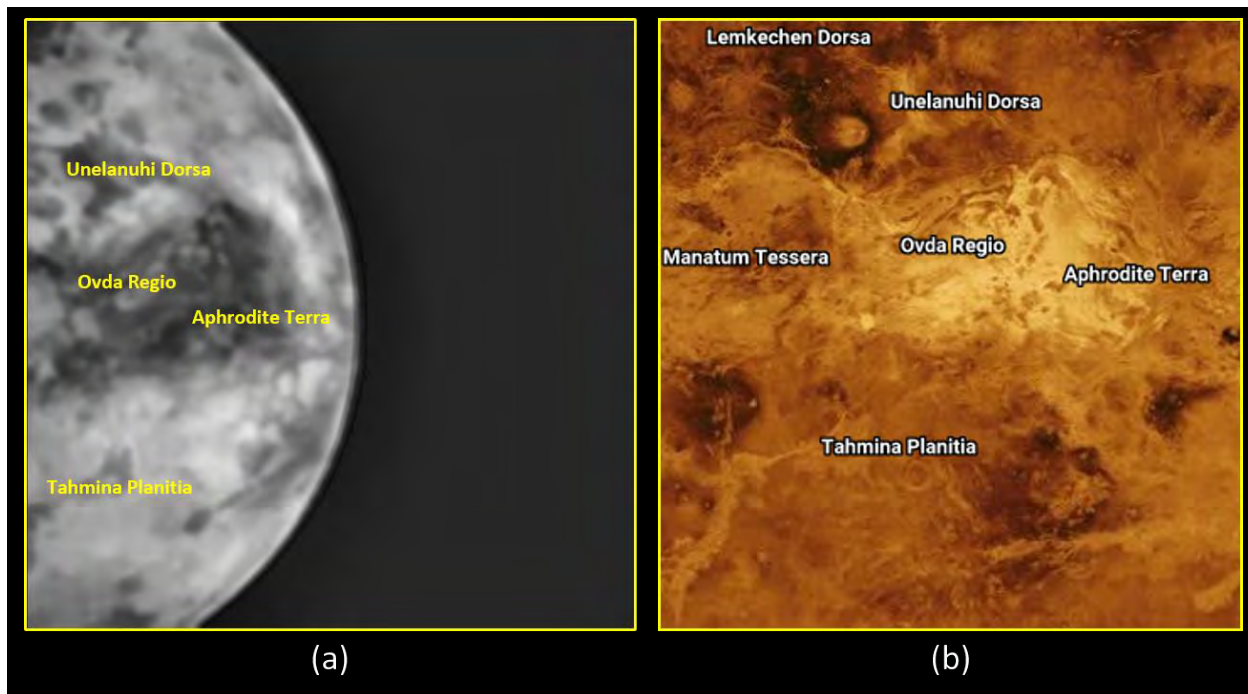


Figure-6.33. Validation of Venus' Fused Annotated Image with Venus' Radar Nomenclature Map

#### *E. Fuzzy Image Segmentation Analysis*

A segment of the Venus' fused image is isolated and subjected to fuzzy logic for enhanced evaluation of surface features (Seidpisheh & Bamdadi, 2023). Unlike conventional segmentation processes that rely on thresholding estimates, our proposed method computes the association of membership for each pixel (Bonnet et al., 2002). In the planetary image space, probabilistic relaxation is employed, and the segmented image is generated through defuzzification (Huet & Philipp, 1998). For the fuzzy planetary image segmentation task, 15 classes are chosen, and the number of iterations is set to 10. Additionally, a color-coded map is created from the segmented

image using a look-up table, representing different classes with distinct colors to enhance the scientific visualization of planetary surface features (Moreland, 2009). Figure-6.34 illustrates the Venus' fused image, the fuzzy segmented image, and the corresponding color-coded map. The various shades of blue clearly indicate that the Aphrodite Terra region of Venus experiences relatively low temperatures during nighttime. The bottom-right interior region of Tahmina Planitia is predominantly filled in pink, with a white area at the extreme right end, as depicted in Figure-6.34 (c). The identified white region is confirmed to be the hottest area in the Venus' color-coded segmented map. The segmented output has been cross-validated and is found to align with the morphological information available about Venus' surface features. The granular information extraction about Venus' surface is facilitated by the generation of the processed image, which distinctly delineates the various morphological features.

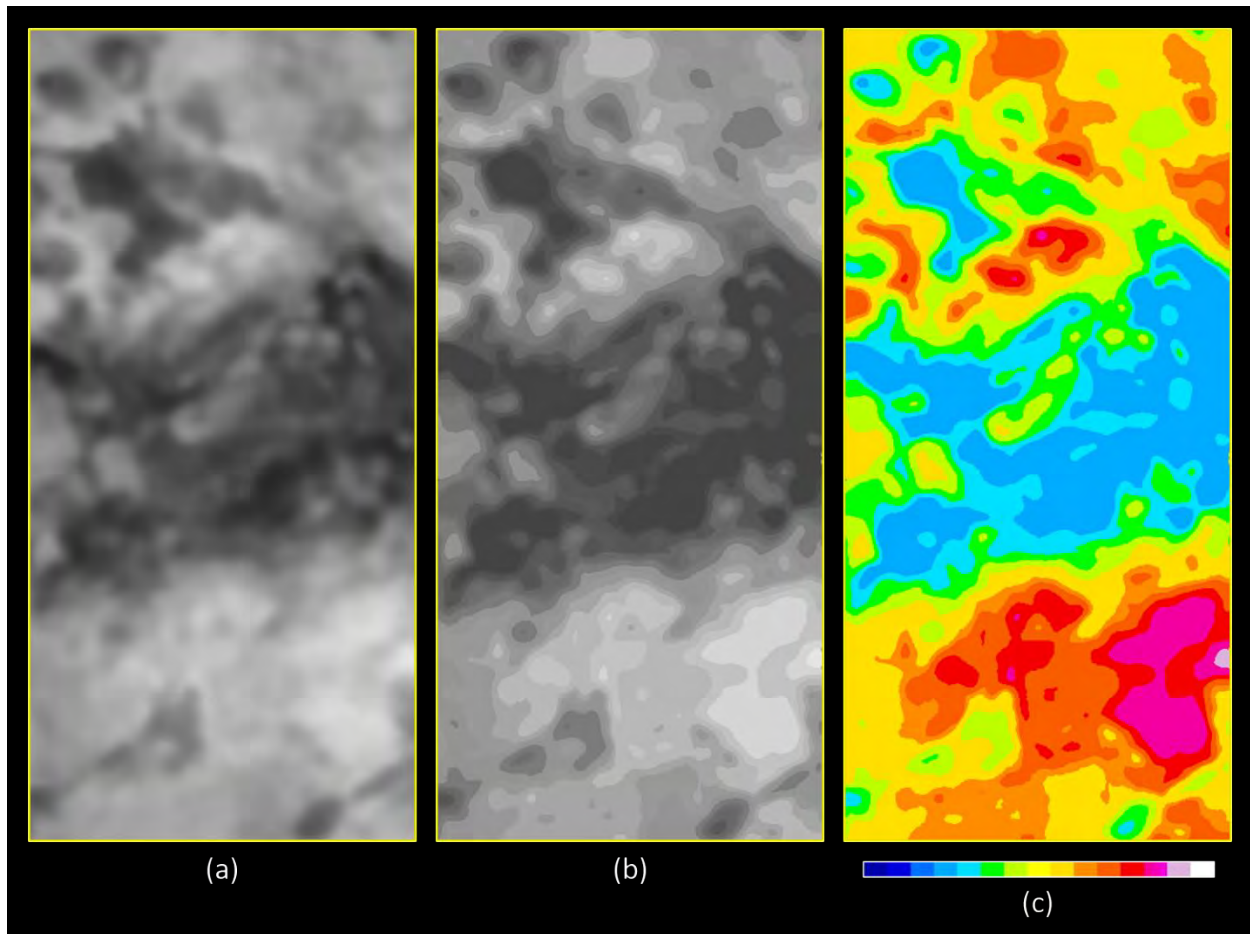


Figure-6.34. Fuzzy Segmentation on Venus' Visible Image (a) Venus' Fused Image (b) Fuzzy Segmented Image (c) Color-coded Fuzzy Segmented Map

### **6.3.4. Summary and Future Research Prospects**

We have introduced an innovative end-to-end workflow for processing Venus' visible raw images, aiming to generate the most enhanced and fused images possible. Employing the Tri-State median filter with background subtraction effectively eliminates streak and impulse noise in the raw image. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied for enhancing the image, improving the contrast of various morphological features on the planet's surface. Our developed multi-modal image registration technique, utilizing SA-SIFT and MSC outlier removal, ensures sub-pixel co-registration accuracy. The resulting multi-modal co-registered image pair is fused using our proposed novel algorithm, incorporating a guided filter for optimal spatial and spectral resolution in the planetary fused image. Each stage of the processing pipeline is assessed using diverse image quality metrics, demonstrating that the final fused image accurately identifies various morphological features on Venus. The output is validated against Venus' radar nomenclature map. Additionally, the fuzzy color-coded segmented map provides further scientific insights into Venus' planetary surface characteristics. Future research endeavors could involve applying this novel image processing workflow to the next available Venus' visible raw image, potentially captured by instruments onboard upcoming Venus orbiter missions conducted by various space agencies globally.

## **7. ENHANCED PROCEDURE FOR MULTI-DECADAL CHANGE DETECTION USING PLANETARY REMOTE SENSING IMAGES**

### **7.1. Introduction**

The chapter emphasizes the development of a methodology for automatically detecting changes utilizing temporal remote sensing imagery. Detecting historical changes requires innovative data processing techniques capable of accurately identifying alterations on the surface, irrespective of target features, morphological formations, geographical positions, or celestial entities. Achieving change detection in an operational mode, covering the entire process from preprocessing data to creating maps indicating the detected changes, is a challenging task. Automatically identifying changes in a region of interest using multi-temporal satellite imagery requires substantial computational processing and the development of a systematic methodology.

The generic workflow developed is capable of handling heterogeneous remote sensing data on Earth and Mars, one of our neighboring dynamic planets. The bi-temporal datasets considered have a time difference of more than a decade in imaging. The output is a change detection map that automatically deduces surface changes resulting from various geological processes. The experimental outcomes demonstrate the visual evaluation of changes identified in dual-temporal images using maps detecting change intensity and color composite images focusing on various features. The quantitative evaluation of change detection includes assessing overall accuracy and kappa coefficients (Misra et al., 2022b).

### **7.2. Proposed Method**

The commencement of data processing starts by ingesting the latest geometrically corrected data for a particular area, along with corresponding historical remote sensing data with a spatial resolution at a similar scale. Digital image processing techniques are utilized throughout each step of processing to ultimately extract change data from remote sensing images spanning multiple decades. The methodology devised includes the successive execution of the following processing stages.

### A. Geometric Transformation of Unpaired Bi-Temporal Images

Unpaired datasets undergo geometric transformation, where the overlapping region is extracted and mapped onto a shared coordinate system for change detection purposes. The geographical bounding corners are derived from geometrically corrected data stored in the GeoTIFF (Geo-Tagged Image File Format) (Ritter et al., 2000). The geographic corner points can be utilized to create a bounding rectangle. The overlapping region is identified by performing a polygon intersection on the geo-bounding rectangles obtained from dual-temporal images. The calculation of the geo-bounded image rectangle for the input temporal images  $I_{(x,y)}^{(1)}$  and  $I_{(x,y)}^{(2)}$  is completed, and the overlapping area is identified using the provided equations below.

$$Rect_{(1)} = o_{bound} \left( I_{(x,y)}^{(1)} \right) \quad (7.1)$$

$$Rect_{(2)} = o_{bound} \left( I_{(x,y)}^{(2)} \right) \quad (7.2)$$

where  $Rect_{(1)}$  and  $Rect_{(2)}$  represent the geographical boundaries of temporal image rectangles  $I_{(x,y)}^{(1)} \wedge I_{(x,y)}^{(2)}$  respectively,

Here,  $x$  denotes the pixel position in the horizontal direction, and  $y$  denotes the pixel position in the vertical direction.

$$Inter_{Rect} = poly_{Intersect} (Rect_{(1)}, Rect_{(2)}) \quad (7.3)$$

where  $Inter_{Rect}$  represents the intersected geo-rectangle between  $I_{(x,y)}^{(1)} \wedge I_{(x,y)}^{(2)}$ .

$$OI_{(x,y)}^{(1)} = overlap_{extract} \left( I_{(x,y)}^{(1)}, Inter_{Rect} \right) \quad (7.4)$$

$$OI_{(x,y)}^{(2)} = overlap_{extract} \left( I_{(x,y)}^{(2)}, Inter_{Rect} \right) \quad (7.5)$$

where  $OI_{(x,y)}^{(1)}$  and  $OI_{(x,y)}^{(2)}$  denote the overlap image regions extracted from  $I_{(x,y)}^{(1)} \wedge I_{(x,y)}^{(2)}$ , respectively.

Images from multiple dates are obtained with varying spatial resolutions. On a global scale, extensive area coverage images are utilized, originating from distinct perspective views. These images undergo rectification through a geometric correction process, which incorporates a map projection step. To align the legacy image that has been extracted with the geometry of the most recent image, both images are subjected to the same projection system. Furthermore, they are resized to a consistent pixel size employing the standard cubic convolution resampling method (Parkar et al., 1983). Below are the equations utilized for the geometric transformation.

$$GI_{(x,y)}^{(1)} = geo_{transform} \left( OI_{(x,y)}^{(1)} \right) \quad (7.6)$$

$$GI_{(x,y)}^{(2)} = geo_{transform} \left( OI_{(x,y)}^{(2)} \right) \quad (7.7)$$

where  $GI_{(x,y)}^{(1)}$  and  $GI_{(x,y)}^{(2)}$  are geometrically transformed images of  $I_{(x,y)}^{(1)}$  and  $I_{(x,y)}^{(2)}$  respectively.

where  $GI_{(x,y)}^{(1)}$  and  $GI_{(x,y)}^{(2)}$  represent the geometrically transformed images of  $I_{(x,y)}^{(1)}$  and  $I_{(x,y)}^{(2)}$  respectively.

### B. Mode Improved SIFT (M-SIFT) Image Registration

The SIFT algorithm commences by constructing scale space extrema through the utilization of a Gaussian Kernel (Lowe 1999; Kupfer et al., 2014). While the Laplacian of Gaussian (LoG) is known to offer stable features and an excellent understanding of scales, its computational cost is too high. Consequently, a Difference of Gaussian (DoG) pyramid is generated as a computationally more efficient approximation of the LoG.

The procedure for building the descriptor starts by sampling points in the vicinity of the matched point. Afterwards, the gradients and coordinates are rotated according to the previously identified orientation, and the region is subdivided into sub-regions. A histogram is generated for each sub-region using predefined bins. The descriptor for each keypoint is stored as a vector element and utilized for feature matching in dual-temporal images. Any unmatched points that still exist as outliers are discarded using technique based on seeking the mode (Misra et al., 2021).

To enhance the SIFT-matched keypoints, the mode scale is identified. Although SIFT naturally handles its own scale and orientation differences, we employ the modes derived from the geo-spatial shifts between matched feature points, as they are frequently more accurate. Subsequently, the matched points are filtered based on user-defined thresholds, which are determined from the image and the mode obtained from geo-spatial shifts in both directions.

The term "mode" refers to the highest frequency of a geo-spatial shift value observed between dual-temporal images. It is computed separately in each direction to aid in the removal of matched points. The determination of mode values is conducted in the following manner:

$$mode_x = mode_{comp}(\Delta_{x1}, \Delta_{x2}, \dots, \Delta_{xn}) \quad (7.8)$$

$$mode_y = mode_{comp}(\Delta_{y1}, \Delta_{y2}, \dots, \Delta_{yn}) \quad (7.9)$$

$mode_x$  represents the mode estimate in the horizontal direction,  $mode_y$  signifies the mode estimate in the vertical direction,  $\Delta x$  denotes the spatial displacement between dual-temporal images along the horizontal direction.,  $\Delta y$  denotes the spatial displacement between dual-temporal

images along the vertical direction, and  $n$  corresponds to the quantity of corresponding feature points.

Matched feature points, represented as  $FP_{match}$ , experience pruning using modes from both directions and predefined thresholds  $Th$ . The computation of pruned matched feature points, denoted as  $Prune\_FP_{match}$ , is carried out as follows:

$$Prune\_FP_{match} = mode_{filter}(FP_{match}, mode_x, mode_y, Th) \quad (7.10)$$

An affine transformation model is built using the ultimately identified potential matching points, and the slave image input undergoes resampling to produce a co-registered data output (Misra et al., 2019a; Xin et al., 2018).

The efficiency of M-SIFT registration is assessed using a set of twenty historical images obtained from IRS-1C/1D LISS-3 captured between 2000 and 2002. RS-2A LISS-3 images captured between 2017 and 2020, notable for their enhanced geometric precision, serve as a benchmark for correcting IRS-1C/1D LISS-3 images. The precision of image co-registration is assessed through both qualitative and quantitative methods across diverse feature targets. Table-7.1 provides information on the IRS-1C/1D LISS-3 images included in the study, along with their respective Root Mean Square Error (RMSE) values (Chai & Draxler, 2014) in relation to the RS-2A LISS-3 reference.

In this study, the M-SIFT approach is contrasted with a registration technique that combines RANSAC (Fischler & Bolles, 1981) and SIFT, referred to as R-SIFT. Root Mean Square Error (RMSE) values are calculated for each IRS scene using both R-SIFT and M-SIFT, and the outcomes are presented in Table-7.1. Significantly, M-SIFT demonstrates superior performance compared to R-SIFT in terms of RMSE values. In a specific situation within mountainous terrain, the RMSE value of R-SIFT exceeds one pixel. The mean RMSE achieved with M-SIFT for LISS-3 scenes is approximately 0.42 pixels, while the average RMSE with R-SIFT is 0.59 pixels, indicating a comparatively higher value.

Figure-7.1 depicts the registration accuracy in urban areas, vegetation, road networks, and mountainous regions. The left portion of the horizontal image swipe displays a multispectral False Color Composite (FCC) image from IRS-1C/1D LISS-3, while the right side features RS-2A LISS-3, presented in grayscale to emphasize only the red spectral band for improved visual assessment. It's observed that enduring features, like road networks and ridges in hilly terrain, align with high precision at a sub-pixel level of image co-registration accuracy.

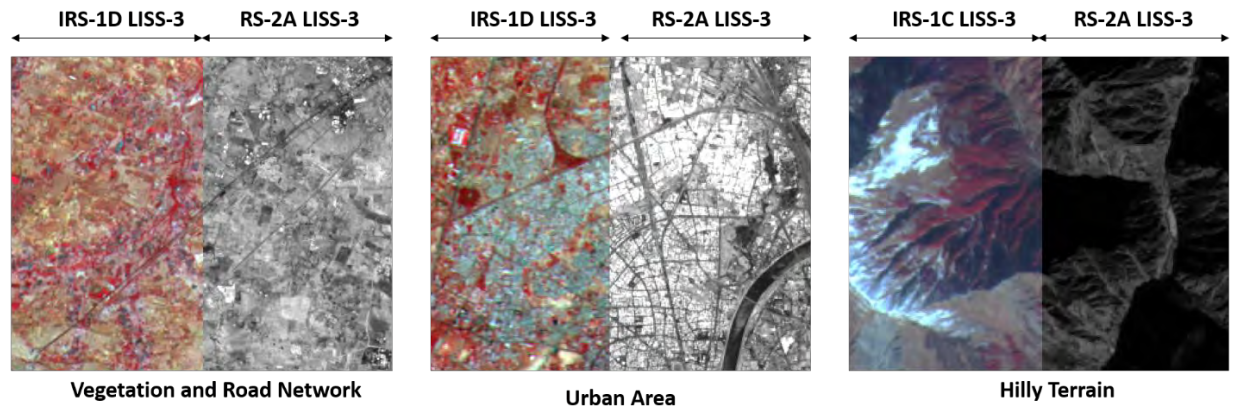


Figure-7.1. Image Co-Registration Performance of M-SIFT Across Various Features

Table-7.1. Details of IRS-1C/1D Scenes and Corresponding RMSE in Comparison to RS-2A Reference

S.No.	Satellite/Sensor	Path/Row	Date of Imaging	RMSE (in pixels) using R-SIFT	RMSE (in pixels) using M-SIFT
1.	IRS-1C/LISS-3	100/060	10 Feb. 2001	0.49	0.41
2.	IRS-1D/LISS-3	100/060	05 Dec. 2000	0.51	0.45
3.	IRS-1C/LISS-3	096/061	03 Jan. 2000	0.83	0.42
4.	IRS-1D/LISS-3	096/061	01 May. 2002	0.45	0.32
5.	IRS-1C/LISS-3	100/061	20 Dec. 2002	0.39	0.38
6.	IRS-1D/LISS-3	100/061	05 Dec. 2000	0.46	0.33
7.	IRS-1C/LISS-3	093/055	02 Oct. 2000	0.55	0.45
8.	IRS-1D/LISS-3	093/055	01 Nov. 2002	0.51	0.47
9.	IRS-1D/LISS-3	097/050	03 Apr. 2002	0.87	0.44
10.	IRS-1C/LISS-3	097/050	03 Apr. 2002	0.63	0.49
11.	IRS-1C/LISS-3	097/049	04 Sep. 2000	1.08	0.51
12.	IRS-1D/LISS-3	097/049	02 Feb. 2001	0.46	0.42
13.	IRS-1C/LISS-3	096/062	15 Mar. 2000	0.38	0.34
14.	IRS-1D/LISS-3	096/062	02 Jan. 2000	0.42	0.36
15.	IRS-1C/LISS-3	093/056	14 Nov. 2001	0.50	0.43
16.	IRS-1D/LISS-3	093/056	01 Dec. 2000	0.59	0.48
17.	IRS-1C/LISS-3	092/052	07 Jan. 2000	0.89	0.32
18.	IRS-1D/LISS-3	092/052	02 Feb. 2002	0.39	0.38
19.	IRS-1D/LISS-3	092/051	02 Feb. 2002	0.56	0.45
20.	IRS-1C/LISS-3	092/051	03 Sep. 2000	0.88	0.47



The calculation of the RMSE value is performed as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|^2} \quad (7.11)$$

where  $N$  represents the total number of keypoint correspondences,

$X_i$  denotes the coordinates  $(x_i, y_i)$  of the RS-2A LISS-3 reference image,

and  $\hat{X}_i$  represents the coordinates  $\hat{x}_i, \hat{y}_i$  of the IRS-1C/1D LISS-3 input image.

### C. Guided Image Filter Enhanced Multivariate Alteration Detection (GIF-MAD)

Examining physical occurrences in remote sensing images usually entails the analysis of numerous pixels within a scene. Identifying alterations in a particular area is achieved using supervised techniques, wherein the spectral attributes of pixels are examined, and change categories are assigned labels. Yet, the process of supervised change detection is frequently labor-intensive and demands manual involvement. Hence, it is more advantageous to directly infer changes from the image data and extract essential information about modifications in a geographical area.

In this context, Multivariate Alteration Detection (MAD), employing canonical variates, provides an accurate evaluation of the magnitude and spatial distribution of changes autonomously. MAD achieves this by identifying spatially coherent patterns of notable changes within an image sequence, while maintaining the spatial context of neighboring pixels. Furthermore, MAD remains unaffected by offset or gain settings of measuring devices, as well as by radiometric and atmospheric correction techniques that exhibit a linear relationship with brightness counts. Hence, when utilizing MAD for change detection in time-series remote sensing images, the need for radiometric normalization preprocessing becomes unnecessary (Canty et al., 2004; Nielsen et al., 1998). The MAD transformation is defined as:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \rightarrow \begin{bmatrix} a_p^T X - b_p^T Y \\ \vdots \\ a_1^T X - b_1^T Y \end{bmatrix} \quad (7.12)$$

Here,  $a_i$  and  $b_i$  denote the characteristic coefficients derived from standard canonical correlation analysis, where  $X$  and  $Y$  are vectors, and  $p$  is the vectors dimension.

The MAD transformation involves obtaining variates by subtracting the corresponding canonical variates in reverse order. The dispersion matrix  $D$  for the MAD variates is depicted below:

$$D\{a^T X - b^T Y\} = 2(M - R) \quad (7.13)$$

In this context,  $M$  represents the  $p \times p$  identity matrix, while  $R$  denotes the  $p \times p$  matrix containing the arranged canonical correlations along the diagonal and zeros elsewhere (Hotelling, 1992).

Incorporating the Guided Image Filter (GIF) enhances the MAD variates' change detection map to a superior level. GIF is a local linear filter mainly controlled by two parameters: regularization, denoted by  $\varepsilon$ , and window radius, represented by  $r$  (He et al., 2012). Principal Component Analysis (PCA) transformation is employed to detect changes between multispectral images captured over multiple decades. The first component of PCA effectively delineates changes in features, serving as a guiding image  $G$  to enhance the MAD variates. The guidance image  $G$  acts as input for the GIF, improving the change detection capability of MAD variates to a higher level. The output  $f$  of the filter at a pixel  $i$  is defined as follows:

$$f_i = \sum_j w_{ij}(G)s_j \quad (7.14)$$

In this context,  $w_{ij}$  denotes the weight between pixel  $j$  and  $i$ , and  $s_j$  represents the  $j$ th input pixel value. In this context,  $j$  denotes the pixel index within the window  $W$ , while  $i$  denotes the center of the window.

$$w_{ij}(G) = \frac{1}{|W|^2} \sum_{(i,j) \in W} \left[ 1 + \frac{(G_i - \mu_w)(G_j - \mu_w)}{\sigma_w^2 + \varepsilon} \right] \quad (7.15)$$

Here  $W$  denotes the square region with a radius of  $r$ , while  $\mu_w$  and  $\sigma_w$  denotes the average and variability of the guidance image  $G$  within the window  $W$ . The term  $|W|$  represents the number of pixels in  $W$ .

In the guided filter, we assume that  $f$  is a linear conversion of  $G$  within a window  $w_k$  centered on a pixel  $k$ .

$$f_i = cG_i + d \quad (7.16)$$

Here,  $c$  and  $d$  represent linear coefficients that are considered to remain consistent within the window radius.

The MAD variates are provided as input  $m_i$  to the GIF filter. GIF works to trace and smooth out undesirable changes, which act as noise in the MAD components.

$$f_i = m_i - n \quad (7.17)$$

Here,  $m_i$  represents the MAD input variates, and  $n$  denotes the noise manifested as undesired fluctuations in the MAD variates.

The output of the filter  $f_i$  is influenced by the guidance image, which is obtained from the PCA transformation. This procedure aids in maintaining sharp edges indicative of significant alterations

while also smoothing areas devoid of change. GIF-MAD demonstrates effectiveness in integrating PCA with MAD variates, thereby improving the resulting change detection image. Figure-7.2 illustrates the workflow of GIF-MAD for change detection using bi-temporal images. GIF-MAD offers a statistically robust approach for detecting spatially coherent patterns of significant changes in bi-temporal images. As a result, GIF-MAD serves as an effective method for detecting historical changes using multi-decadal time-lapse remote sensing data, such as IRS-1C and RS-2A images.

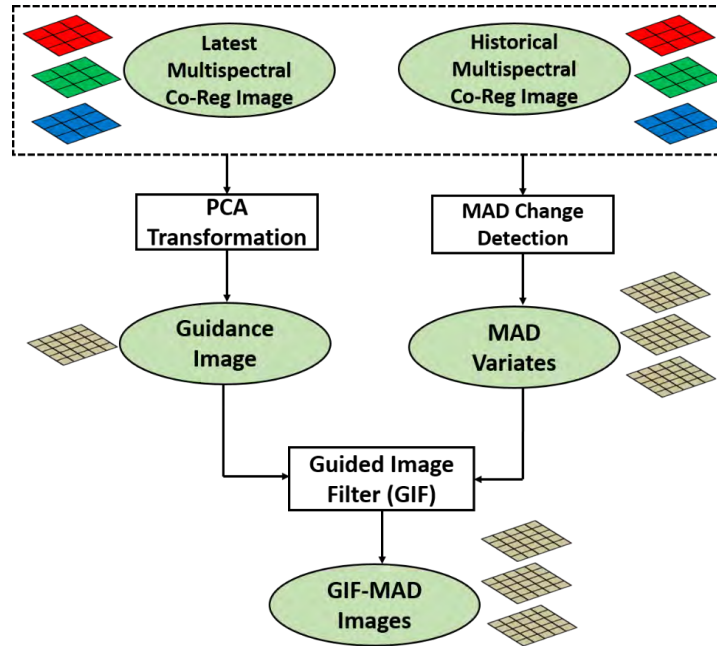


Figure-7.2. Workflow for change detection using GIF-MAD

#### D. Performance Evaluation and Comparison of GIF-MAD Simulation

To assess GIF-MAD in comparison to other change detection methods, a small-scale simulation is conducted. The bi-temporal images utilized in the simulation include IRS-1C LISS-3 and RS-2A LISS-3, both with a spatial resolution of 24 meters, and taken approximately two decades apart. The reference data is directly extracted from the images. Both sets of bi-temporal images are supplemented with zero intensity values within a central region surrounding the patches, effectively designating this central area as a no-change region for simulation and performance evaluation purposes.

In the experiment, a small patch of size 1000 by 1000 pixels from IRS-1C LISS-3, along with its corresponding counterpart from RS-2A LISS-3, is supplemented with zero values across all spectral bands. This patch is positioned geometrically in the center of a 1050 by 1050 pixel background. The disparity between the two 1050 by 1050 scenes is assessed by computing the

Local Mean Difference (LMD), PCA, MAD, and GIF-MAD. Table-7.2 showcases the alterations identified in the region devoid of changes, as highlighted by standardized values derived from various change detection methods. The findings indicate that GIF-MAD surpasses other conventional change detection methods based on image transformation in the area where no change is observed.

Table-7.2. Change Detected in No Change Region for IRS Data

<b>Channels/Components</b>	<b>1</b>	<b>2</b>	<b>3</b>
LMD	0	0	0
PCA	0.21	0.39	0.54
MAD	0.22	0.28	0.31
GIF-MAD	<b>0.15</b>	<b>0.18</b>	<b>0.19</b>

Figure-7.3 illustrates the qualitative evaluation of all the previously mentioned change detection methods. Figure-7.3 presents change intensity detection maps produced utilizing various techniques for image transformation, visually depicting significant changes occurring across the landscape. Areas with higher brightness levels in the change intensity detection maps suggest potential locations of changes. The change detection maps based on color composites provide an easily understandable interpretation of changes utilizing a vivid color scheme.

LMD provides a thorough evaluation of change at individual points, its accuracy is significantly influenced by the absolute calibration accuracy and temporal consistency of the satellite instrument. PCA surpasses LMD in performance, but it is susceptible to linear scaling and affine transformations of the input data sets, potentially leading to a loss of spatial context information in the change detection output. MAD, employing canonical variates, excels in maintaining the spatial context of data in change detection results compared to PCA and LMD. However, MAD has a propensity to categorize unchanged pixels as changed, resulting in increased false positive rates, particularly in urban areas.

GIF-MAD combines PCA and MAD synergistically to enhance change detection results, maintaining edges while smoothing non-changing areas, thus surpassing existing methods. The change detection maps indicate that GIF-MAD efficiently detects changes in the eastern vicinity of Ahmedabad airport, demonstrating significant urbanization expansion in contrast to data from two decades earlier.

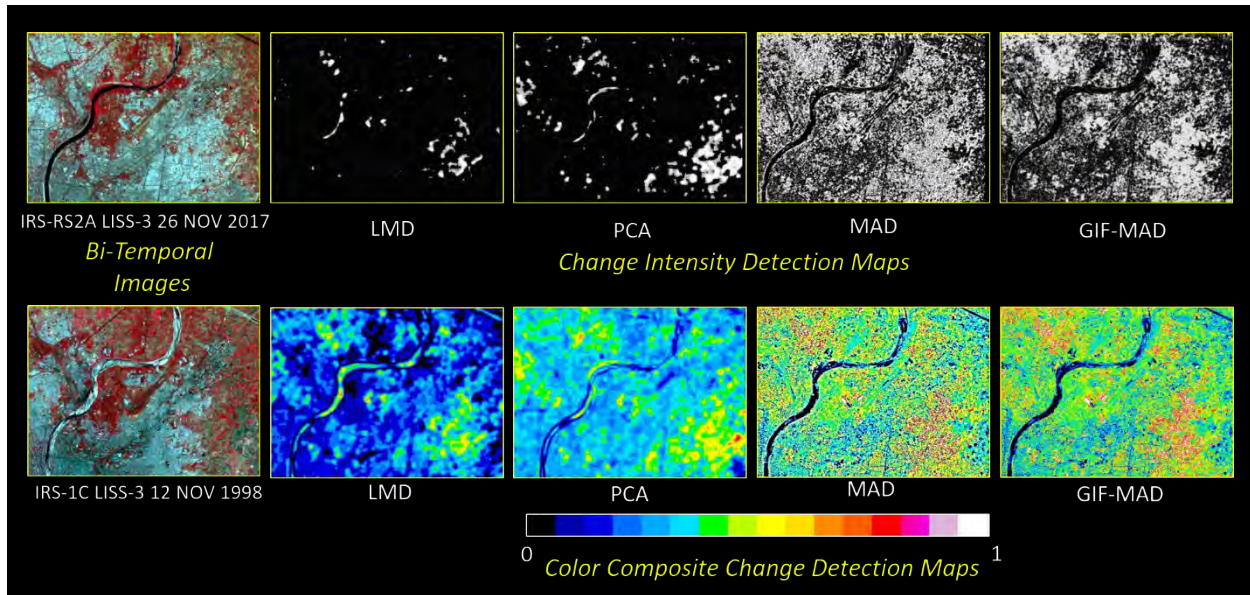


Figure-7.3. Visual Quality Evaluation of Various Change Detection Techniques

#### E. Development and Workflow of the EPOCH Algorithm

The outlined data processing procedures have been structured into a processing pipeline named EPOCH. The pipeline receives geometrically corrected data products as input for Earth observation data, the overlapping area is extracted from the historical data archive and geometrically transformed to match the same projection system and pixel dimensions as the most recent acquisition data input. For Mars processing, the workflow is initiated by ISRO's MCC most recent acquisition data, while the Viking color mosaic (Mars Viking, 2021) serves as the legacy data. The multi-temporal images after transformation undergo co-registration through the Mode Improved SIFT technique. Afterward, the change detection module based on GIF-MAD produces a change detection map for data analysis.

The Mars Digital Image Model (MDIM 2.1), featuring a resolution of 231 meters, serves as a comprehensive base map for Mars, improving absolute accuracy by integrating globally distributed ground control points (Archinal et al. 2004; Misra et al., 2019b). MDIM 2.1 acts as a reference in the Mars data workflow, and the pertinent region is extracted to validate the final change detection map (Misra et al., 2021c). Prepared reference images, created using Resourcesat data over Indian landscape (Misra et al., 2018), are utilized as a standard for validating the identified changes in Earth observation data.

The software chain is crafted in C++ and functions on the Linux operating system. The EPOCH process has been assessed on a system equipped with an Intel Xeon CPU E5-2697 v4 @ 2.30GHz

processor boasting 18 cores and 128 GB of Random Access Memory (RAM). The processing time for EPOCH is observed to be less than a minute for bi-temporal remote sensing images sized 6000 by 6000 pixels. The algorithmic steps of the developed methodology are outlined in Algorithm-7.1. The EPOCH data processing workflow includes data pre-processing and the change detection stage, illustrated in Figure-7.4.

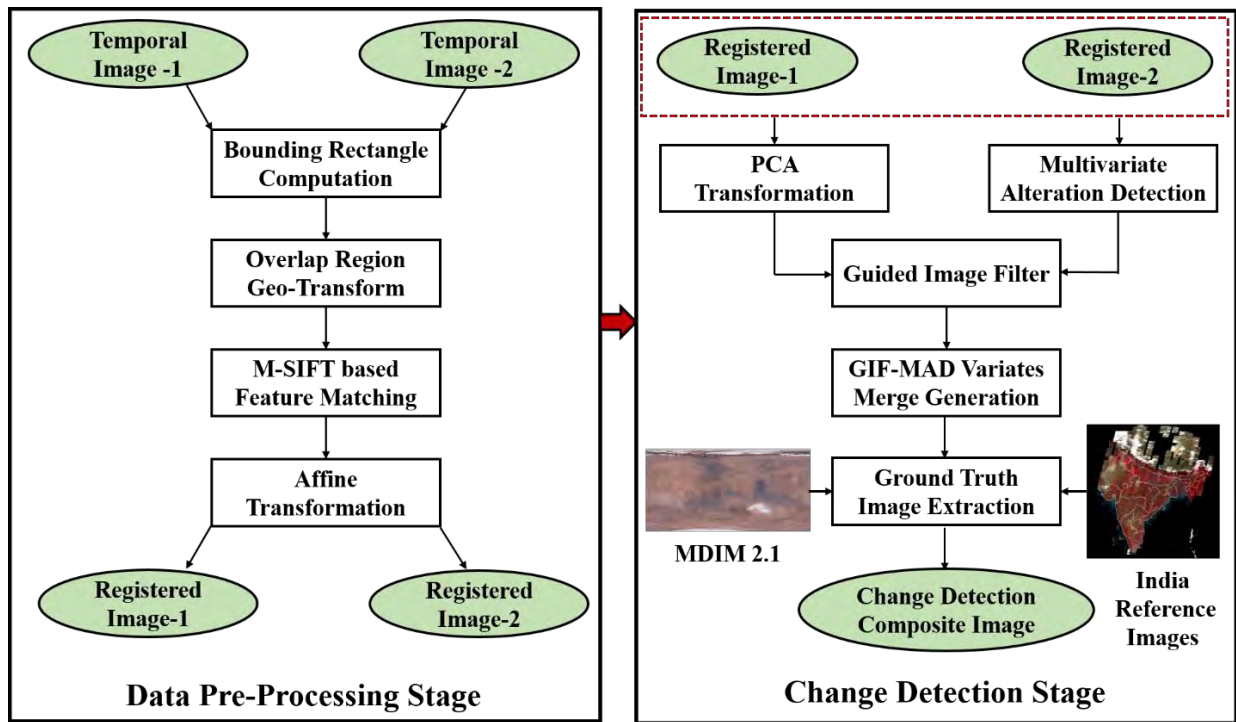


Figure-7.4. EPOCH Processing Framework

### **Algorithm-7.1**

1. Calculate geographic bounding rectangles from bi-temporal images  $I_{(x,y)}^{(1)} \wedge I_{(x,y)}^{(2)}$
2. Utilize polygon intersection to extract the overlapping region between the image pair.
3. Geometrically segment the bi-temporal images based on the intersected polygon.
4. Geometrically transform the overlapping images to align them onto the same projection system and dimensions.
5. Utilize SIFT-based invariant feature detection and matching between the transformed images.
6. Employ Mode Seeking-based pruning to ascertain the ultimate putative matches.
7. Conduct an Affine Transformation of the older image relative to the recently acquired image.
8. Implement Multivariate Alteration Detection (MAD) on the registered image pair.
9. Carry out Principal Component Analysis (PCA) transformation on the registered image pair.
10. Extract the first component of PCA, which acts as a guiding image.
11. Improve MAD variates through PCA using Guided Image Filter (GIF-MAD).
12. Merge the GIF-MAD variates to produce a color composite change detection map.

### **7.3. Experimental Results**

The effectiveness of the proposed framework has been evaluated through testing the EPOCH algorithm across diverse landscapes, encompassing various features including vegetation, urban zones, and artificial constructions. Studies in Earth observation utilized IRS and Landsat data, showcasing major changes in the change detection color legend image. In the detection of surface changes on Mars over several decades, recent data from the Mars Color Camera (MCC) and the Viking image repository were utilized. Figure-7.5 depicts the chosen study regions on both Earth and Mars, denoted by yellow square boxes, where the proposed methodology is implemented for historical change detection. Table-7.3 presents the specifics of the metadata utilized in the experiment.

Table-7.3. Metadata Specifications of the Remote Sensing Datasets Employed in the Experiment

Bi-Temporal Image ID	Planet	Satellite	Sensor	Date of Acquisition	Region Covered
Bi-Temporal Images No. 1	Earth	IRS-1C	LISS-3	12 Nov. 1998	Ahmedabad, India (Urban)
		RS-2A	LISS-3	26 Nov. 2017	
Bi-Temporal Images No. 2	Earth	IRS-P6	LISS-4	29 Dec. 2005	Shillong, India (Man-made structures)
		RS-2A	LISS-4	27 Nov. 2018	
Bi-Temporal Images No. 3	Earth	IRS-1A	LISS-2	29 Nov. 1992	Gandhinagar, India (Vegetation)
		RS-2A	LISS-3	26 Nov. 2017	
Bi-Temporal Images No. 4	Earth	Landsat-5	TM	10 Apr. 1985	Sioux Falls, United States (Urban Growth)
		Landsat-8	OLI	29 Apr. 2021	
Bi-Temporal Images No. 5	Mars	Viking Orbiter	Vidicon Camera	Aug. 1980	Elysium and Amenthes Quadrangles
		MOM	MCC	Jan./Apr./Jul.2016	

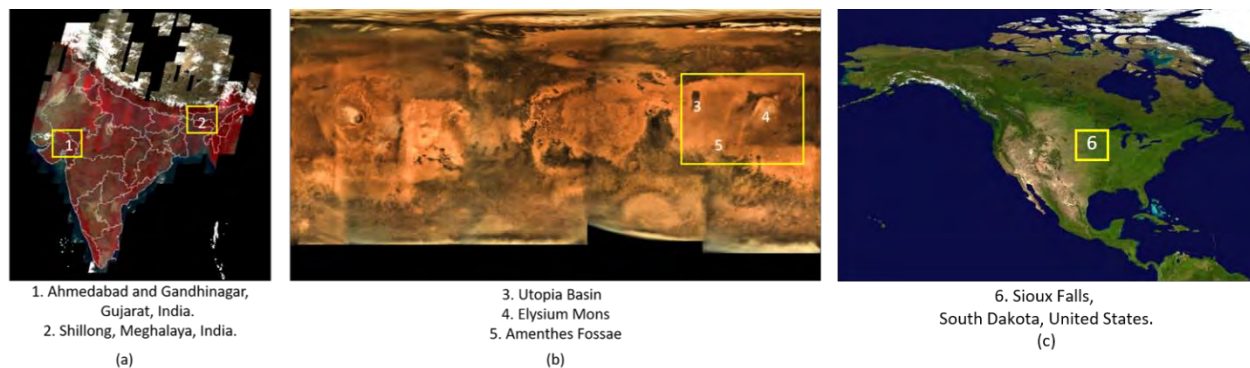


Figure-7.5. (a) Study Regions (Highlighted in Yellow Boxes) depicted on the Indian Resourcesat LISS-3 Mosaic (Earth) (b) Study Regions (Highlighted in Yellow Boxes) presented on ISRO's MCC Global Mosaic (Mars) (c) Study Region (Highlighted in Yellow Box) displayed on the Blue Marble Image of the United States (Earth)



*A. Performance of Surface Change Detection on Earth*

The bi-temporal images undergo processing through the pipeline to automatically produce a change detection image. Figure-7.6 portrays the evolution of Ahmedabad airport and its vicinity utilizing historical IRS-1C LISS-3 data, which predates the current RS-2A LISS-3 data by two decades. The existence of the new international terminal near Ahmedabad airport is noticeable in the RS-2A data but not apparent in the older IRS-1C data. The change detection image generated by EPOCH is shown in white, highlighting the most notable modifications between the bi-temporal multispectral images.



Figure-7.6. Expansion of Ahmedabad Airport depicted through Bi-Temporal Images No. 1

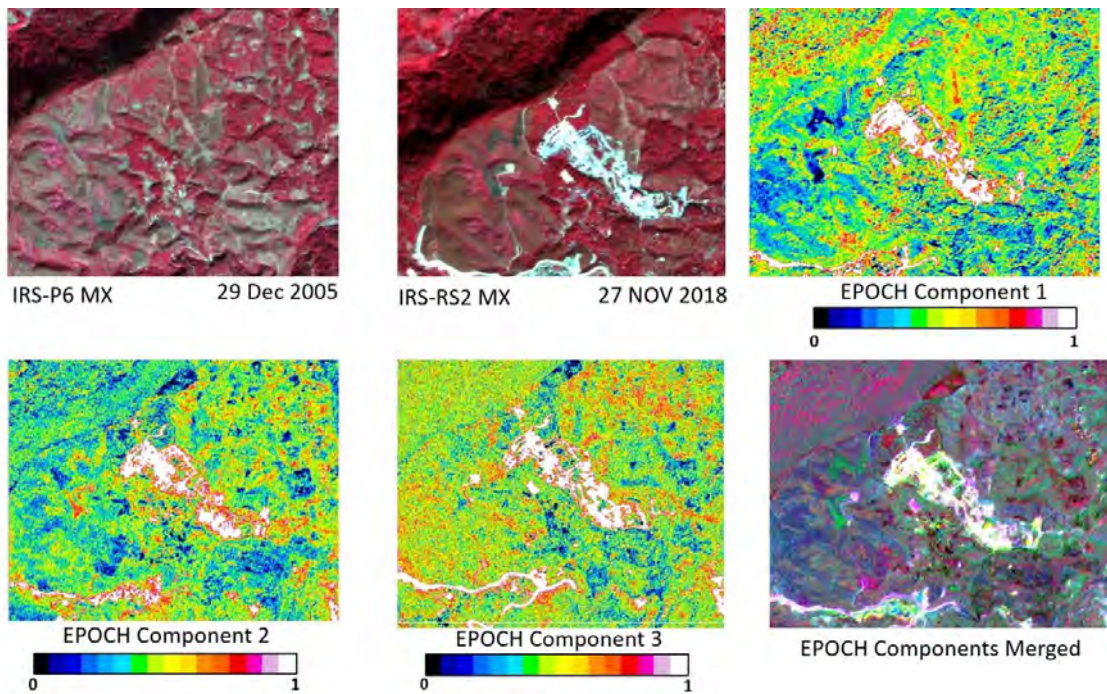


Figure-7.7. Identification of man-made features around Umsawli, India, depicted in Bi-Temporal Images No. 2

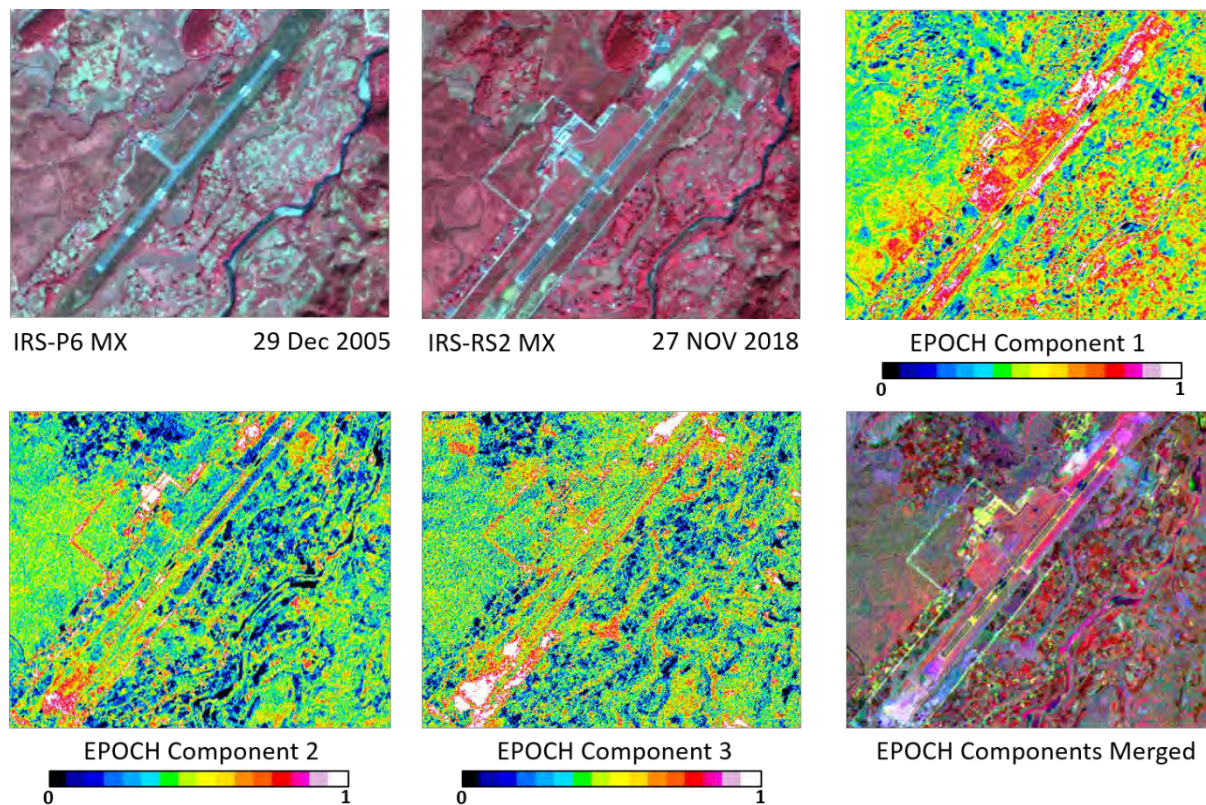


Figure-7.8. Change detection observed over Shillong Airport, India, depicted in Bi-Temporal Images No. 2

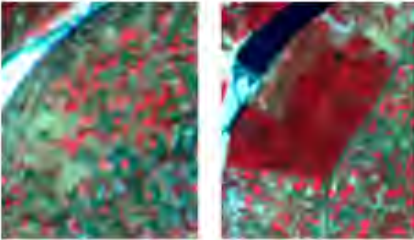
Human-made structures across hilly terrain are effectively recognized, precisely capturing the identified changes, as demonstrated in Figure-7.7. The highlighted area corresponds to the new site of IIM Shillong in Umsawli, India. Figure-7.8 illustrates the identified alterations at Shillong Airport. Each component of EPOCH is showcased individually, and in the color composite change detection map, pink and white hues indicate areas of maximum change between the bi-temporal images.

To examine alterations in vegetation within the Gandhinagar region, the historical path-oriented data from IRS-1A LISS-2 is aligned to the north and resampled to match the spatial resolution of 24 meters in RS-2A LISS-3. Geometric registration is carried out between IRS-1A LISS-2 and the most recent RS-2A LISS-3 acquisition for change detection investigations. The vegetation change image generated via the automated processing workflow based on EPOCH is produced. NDVI is computed for both IRS-1A and RS-2A, and the discrepancy in NDVI is analyzed to identify changes in vegetation.

The comparison between EPOCH and NDVI difference is depicted in Figure-7.9, demonstrating that EPOCH effectively identifies changes in greenery (vegetation cover) within the study area. Pixels with NDVI difference values greater than 0.6 are categorized as dense vegetation. Ground truth data is manually derived from the LISS-3 reference image covering the site, outlining the mapping of healthy changed vegetation pixels. Table-7.4 showcases the number of changed vegetation pixels across multi-decadal IRS images employing different change detection methods, such as CVA, LMD, MAD, PCA, SFA, NDVI difference, PCANet, and EPOCH. To optimize PCANet hyperparameters, we have 8 filters per layer, using patch sizes of 7x7. The block size for histogram calculation is 8\*8 and have 2 pyramid levels for effective change detection. The comparison is conducted within an area of interest marked by the yellow square box in Figure-7.9.

While MAD surpasses PCA, it tends to identify unaltered pixels as changed vegetation indicators. Conversely, SFA and PCANet demonstrate superior capability in capturing vegetation changes within the region of interest, displaying improved detection performance. Nonetheless, there remains a relatively high degree of variability compared to the ground truth data. EPOCH is observed to closely approximate ground truth measurements, exhibiting minimal deviation in comparison to alternative change detection techniques. In general, Gandhinagar seems to be more verdant, particularly around the Sabarmati River, in contrast to its state 25 years ago.

Table-7.4. Alterations in Healthy Vegetation within a Specific Area near Gandhinagar, India

Region of Interest	Change Detection Techniques	Change in Vegetation Pixels (numbers)
 <p data-bbox="282 1688 456 1709">IRS-1A LISS-2 (1992)</p> <p data-bbox="500 1688 673 1709">RS-2A LISS-3 (2017)</p>	LMD	512
	CVA	553
	PCA	543
	MAD	598
	NDVI Difference	565
	SFA	569
	PCANet	572
	EPOCH	<b>578</b>
	Ground Truth	<b>585</b>

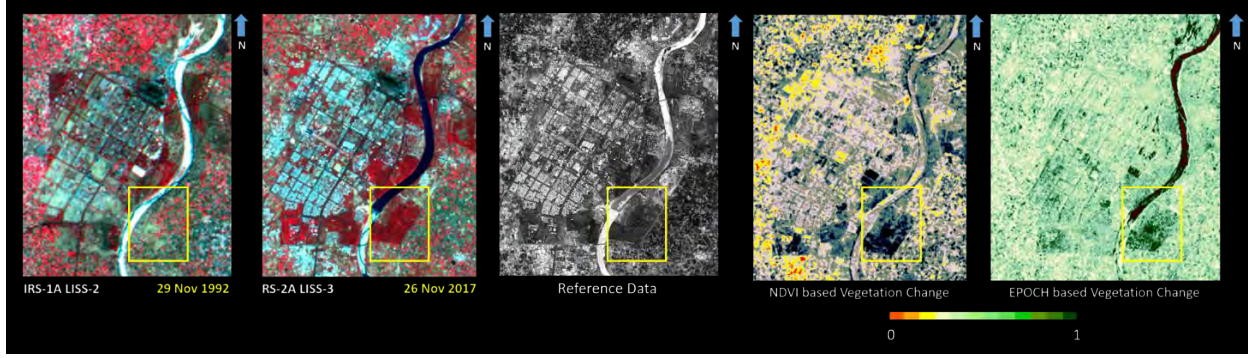


Figure-7.9. Detection of greenery in Gandhinagar using IRS data (Bi-Temporal Images No. 3)

The quantitative evaluation of the developed methodology involves detecting keypoints in the ground truth image using the SIFT feature detection technique. In this study, the RS-2A LISS-3 reference image over Indian land terrain serves as the basis for the ground truth. These reference images are the most recent ortho-rectified data acquisitions, largely devoid of clouds, and possess an absolute location accuracy of up to 12 meters. The change layer is validated by dispersing SIFT keypoints across regions where changes have been detected, as well as across features where no changes are evident. The selected points are categorized into change and no-change groups, and accuracy is assessed using a standard confusion matrix (Foody, 2002). The corresponding error matrices are illustrated in Table-7.5, 7.6, and 7.7. The overall accuracy is computed by summing up the correctly classified values and dividing by the total number of SIFT points (Story & Congalton, 1986). The kappa coefficient, which assesses the agreement between classification and truth values (Stehman, 1996; Fung & LeDrew, 1988), is also calculated using the formulas provided below.

$$\text{Overall Accuracy} = \frac{\sum_{i=1}^n m_i}{N}, \quad (7.18)$$

$$\text{Kappa} = \frac{N \sum_{i=1}^n m_i - \sum_{i=1}^n \text{predicted}_i \text{truth}_i}{N^2 - \sum_{i=1}^n \text{predicted}_i \text{truth}_i}, \quad (7.19)$$

where  $N$  is the total count of classified pixels,  $i$  represents the class number,  $n$  is the number of classes,  $m_i$  is the number of correctly classified values belonging to class  $i$ ,  $\text{predicted}_i$  is the total count of predicted values attributed to the class  $i$ , and  $\text{truth}_i$  is the total count of true values attributed to the class  $i$ .

Table-7.5. Error Matrix for Change/No-Change Classes in Bi-Temporal Images No. 1

Classification				
		Change	No Change	Sum
Reference	Change	575	40	615
	No Change	60	525	585
	Sum	635	565	1200
Overall Accuracy	0.916			
Kappa	0.846			

Table-7.6. Error Matrix for Change/No-Change Classes in Bi-Temporal Images No. 2

Classification				
		Change	No Change	Sum
Reference	Change	655	35	690
	No Change	54	620	674
	Sum	709	655	1364
Overall Accuracy	0.934			
Kappa	0.869			

Table-7.7. Error Matrix for Change/No-Change Classes in Bi-Temporal Images No. 3

Classification				
		Change	No Change	Sum
Reference	Change	797	80	877
	No Change	64	653	717
	Sum	861	733	1594
Overall Accuracy	0.909			
Kappa	0.817			

The mean overall accuracy attained for bi-temporal Earth observation images is 91.9%. The corresponding Kappa value calculated for the identical set of image pairs is 0.844, signifying the precise identification of changes in remote sensing images spanning multiple decades by EPOCH. It is noteworthy that precise detection of changes is observed across diverse land features, encompassing urban areas, vegetation, and man-made structures.

Different techniques yield optimal change detection results for Bi-Temporal Images No. 3 by configuring suitable threshold values. Table-7.8 provides a contrast of change detection outcomes in regards to overall accuracy, Kappa value, and F1 score. The F1 score, computed from precision and recall, can be formulated as:

$$F1 = 2 * \left( \frac{Precision * Recall}{Precision + Recall} \right) \quad (7.20)$$

It is clear that LMD faces challenges in accurately detecting alterations in remote sensing images over multiple decades. Although NDVI difference, SFA, and PCANet offer reasonably precise estimations of altered pixels in vegetation areas, their accuracy declines in varied feature targets. Conversely, EPOCH consistently surpasses other change detection methods across all evaluated accuracy metrics.

Table-7.8. Comparison of Change Detection Results for Bi-Temporal Images No. 3

S.No	Change Detection Methods	Overall Accuracy	Kappa	F1 Score
1.	LMD	0.864	0.754	0.812
2.	CVA	0.882	0.771	0.842
3.	PCA	0.855	0.785	0.862
4.	MAD	0.861	0.789	0.885
5.	NDVI Difference	0.893	0.802	0.902
6.	SFA	0.887	0.791	0.893
7.	PCANet	0.874	0.809	0.904
8.	EPOCH	<b>0.909</b>	<b>0.817</b>	<b>0.916</b>

EPOCH is evaluated using Landsat series data covering Sioux Falls, South Dakota, in the United States. The bi-temporal images acquired by Landsat-5 TM and Landsat-8 OLI are fed into

the EPOCH processing workflow. In Figure-7.10, the results of change detection illustrate urban growth in the southern part of Sioux Falls, where pink and white colors indicate regions of significant change in the detection map. The Global Land Survey (GLS) datasets serve as a reference, providing terrain-corrected, consistent, and coordinated remote sensing data. Table-7.9 displays the error matrix for classifications of change and no-change using multitemporal Landsat images. The overall accuracy and Kappa coefficients validate the dependability of the EPOCH method in identifying changes, underscoring its capacity to produce precise change data across various geographical areas for subsequent scientific analysis.

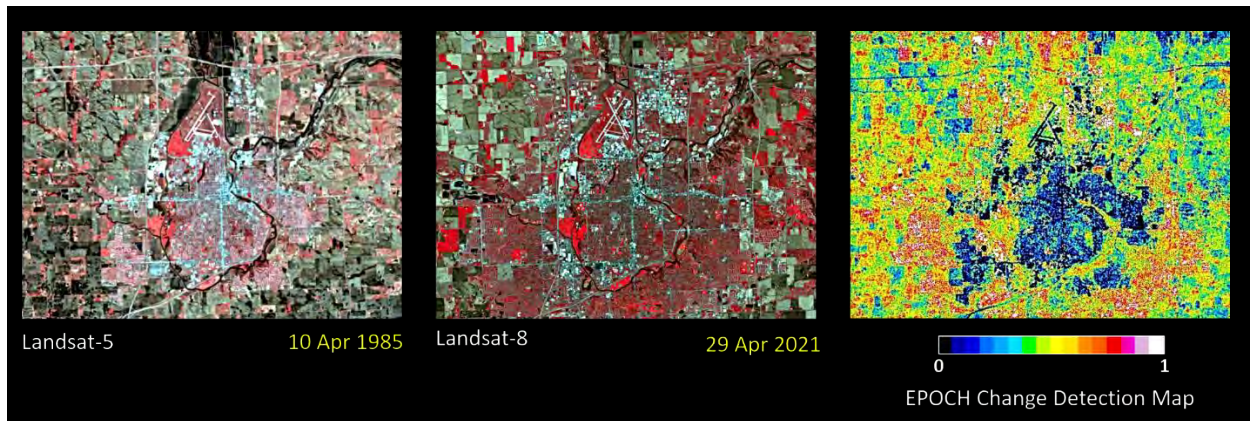


Figure-7.10. Urban expansion in Sioux Falls, United States, as observed from Landsat images (Bi-Temporal Image No. 4)

Table-7.9. Error matrix for change/no-change classifications in Bi-Temporal Images No. 4

		Classification		
		Change	No Change	Sum
Reference	Change	672	44	716
	No Change	56	611	667
	Sum	728	655	1383
Overall Accuracy				0.927
Kappa				0.855

### B. Performance Evaluation of Change Detection on the Surface of Mars

The EPOCH methodology was implemented and evaluated within the Elysium (Zimelman & Leshin, 1987) and Amenthes (Craddock & Maxwell, 1990) Quadrangles on Mars, covering an extensive surface area with features including Mons, impact craters, dust devils, and clouds. For

this region, the MCC dataset was utilized, offering data at different spatial resolutions. The Viking image from three decades ago underwent processing using the EPOCH workflow, yielding a modified Viking image and a color composite change detection map. Figure-7.11 displays these elements, highlighting significant changes, denoted by the pink color, occurring in impact basins and Mons near Amenthes Fossae.

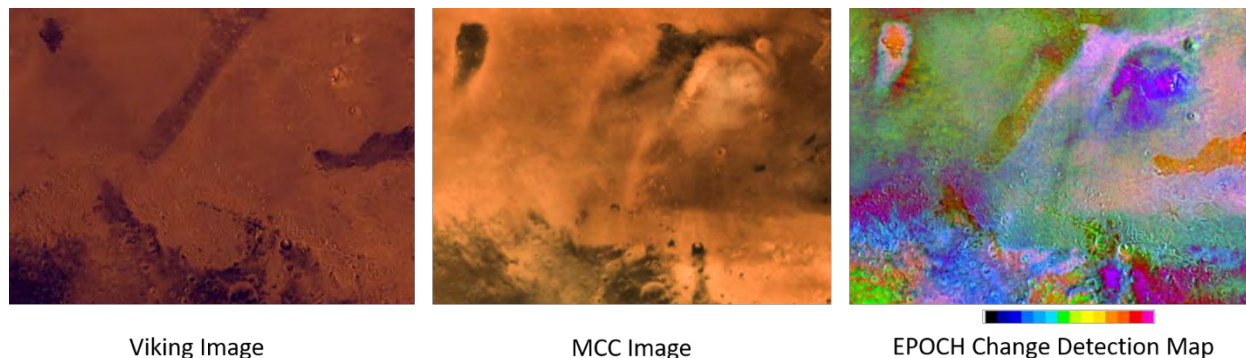


Figure-7.11. Change Detection across Mars Quadrangles (Bi-Temporal Images No. 5)

Elysium Mons, a prominent volcano on Mars situated in the eastern hemisphere, rises approximately 12.6 kilometers (41,338 feet) above its base and 14.1 kilometers (46,259 feet) above the Martian datum. This makes it the third highest Martian mountain in terms of relief and the fourth tallest in elevation. On April 25, 2016, MCC captured images of Elysium Mons and its surroundings, revealing clouds above the towering volcano (depicted in Figure-7.12). The automated change detection images generated by MCC and Viking exhibit significant differences over Elysium Mons, which can be attributed to cloud cover and the appearance of dust devils. Figure-7.12 depicts the EPOCH variates color composite change detection map over Elysium Mons, employing light pink and dark pink colors hues to emphasize areas with the most significant surface changes. The change detection map also delineates significant changes, evident as brighter regions surrounding the Elysium Mons volcano.



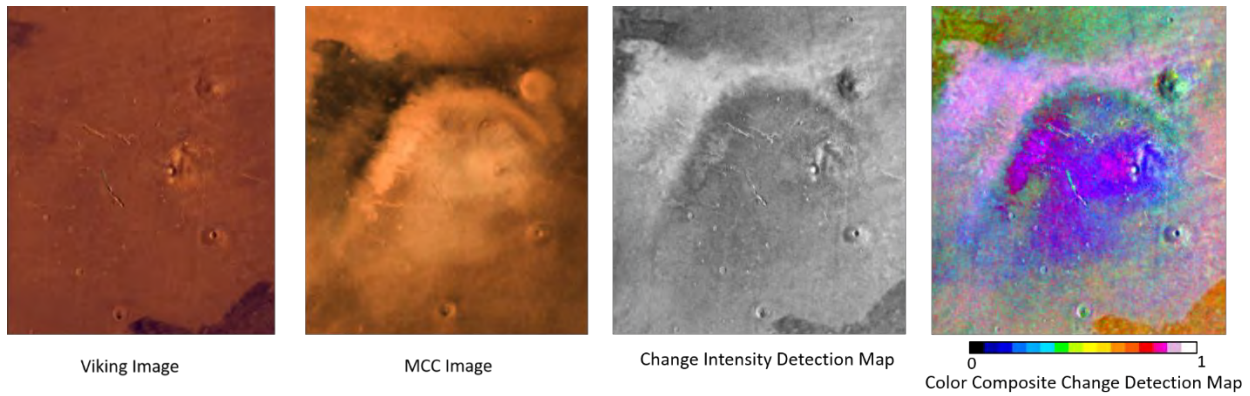


Figure-7.12. Change Detection around Elysium Mons (Bi-Temporal Images No. 5)

The Utopia basin represents the most prominent easily identifiable impact structure in the northern hemisphere of Mars, displaying a unique large-scale impact characteristic. Following the coverage of this region by MCC and subsequent comparison with older Viking images, a significant alteration is observed in this area. The expansion of the lower portion of the Utopia basin is precisely outlined, with the change manifesting as a bright area in the change intensity detection map. Additionally, Figure-7.13 illustrates the color composite change detection image generated by EPOCH covering the Utopia basin, offering insights into the magnitude of changes surrounding this geographical feature on Mars.

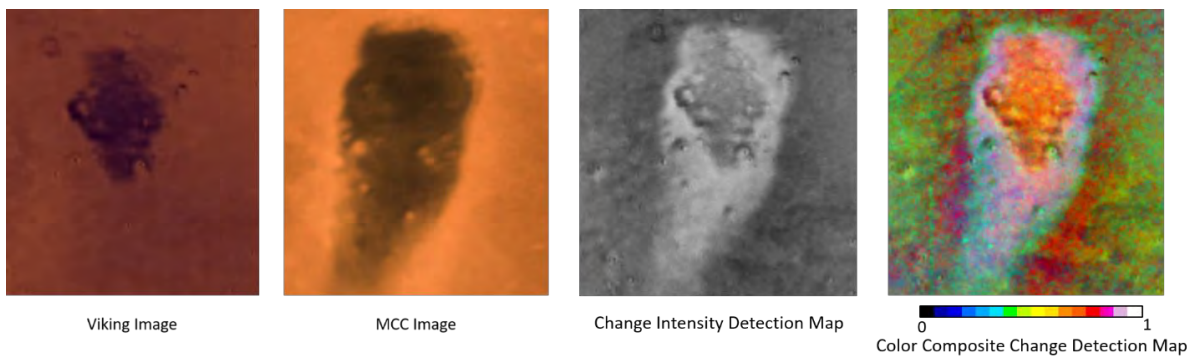


Figure-7.13. Change Detection over Utopia Basin (Bi-Temporal Images No. 5)

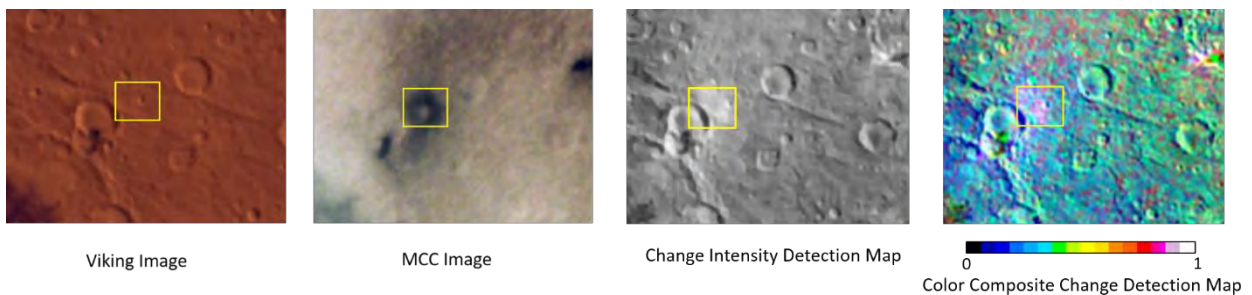


Figure-7.14. Fresh Impact Crater Formation near Amenthes Fossae (Bi-Temporal Images No. 5)

The latest MCC image acquisition has unveiled the emergence of a fresh impact crater near Amenthes Fossae, a component of a network of troughs in the Amenthes quadrangle on Mars. This newly created crater is not visible in the archived Viking image. The change intensity detection map, exhibiting a bright-toned area, and the change detection image using color composites in light pink (as shown in Figure-7.14), clearly emphasize the newly detected impact crater (enclosed in the yellow box) along with other modified geological phenomena.

Table-7.10. Error matrix for change/no-change classifications of Bi-Temporal Images No. 5.

		Classification		
		Change	No Change	Sum
Reference	Change	361	29	390
	No Change	45	365	410
	Sum	365	435	800
Overall Accuracy				0.907
Kappa				0.830

The quantitative assessment of Mars Bi-Temporal Images No. 5 involves identifying stable keypoints using SIFT feature detection in the reference image. MDIM 2.1 serves as the reference ground truth for this evaluation, characterized by outstanding geometric accuracy and radiometric attributes. SIFT detects eight hundred stable matched points in the reference image, serving as the foundation for accuracy verification. The distribution of keypoints across change and no-change regions is even, facilitating classification into change and no-change categories. The associated error matrix, shown in Table-7.10, denotes an overall accuracy of 90.7% and a kappa value of 0.830. Despite the bi-temporal images of Mars being acquired over a period of more than 30 years, a relatively high level of accuracy is attained.

#### 7.4. Summary

Images captured at different times with a gap of over a decade offer a comprehensive record of possible surface alterations. The emphasis of change detection data products are located in various areas of interest for temporal analysis. The EPOCH change detection processing pipeline, as detailed in the paper, has shown encouraging outcomes in automatically detecting alterations across geographically separated areas. The developed approach is evaluated on various feature targets to identify changes spanning multiple decades, employing Landsat and IRS data on Earth.

The alignment of time-lapse remote sensing images utilizes SIFT feature matching, which produces potential feature matches using the mode geo-shift procedure. GIF-MAD provides a statistically sound approach to identify spatially consistent patterns of significant changes in sequences of remote sensing images. EPOCH demonstrates efficacy in monitoring urban expansion, alterations in healthy vegetation regions, and enables the analysis of land cover changes in an operational capacity. The accuracy of change detection is assessed both quantitatively and qualitatively by employing change detection color composite images generated through EPOCH. EPOCH, a versatile processing workflow, not only evaluates surface alterations on Earth but also demonstrates promise in detecting changes on the Martian surface. The fusion of ISRO's MCC imagery with archival NASA's Viking orbiter data reveals substantial alterations spanning over a decade in various morphological aspects of Mars. The change intensity detection map offers a more profound understanding of surface alterations on Mars. The vast MCC image repository can be leveraged for conducting change detection analyses across various Martian regions, facilitating the monitoring of dynamic phenomena such as dust devils/streaks, clouds, and the emergence of new impact craters. Moving forward, EPOCH will undergo testing using multi-temporal hyperspectral remote sensing data to observe changes at a more detailed spectral resolution.

## 8. CONCLUSIONS

Feature-based remote sensing image registration is a critical process in the field of remote sensing and geospatial analysis. It involves aligning images from different sensors or time periods to create accurate, composite images for various applications. Feature-based registration methods have demonstrated high accuracy and precision in aligning remote sensing images. They are particularly effective when distinct, recognizable features are present in the images, such as landmarks, roads, or natural features. Feature-based methods are robust in handling variations in scale, rotation, translation, and even some degree of perspective changes. This robustness makes them suitable for a wide range of remote sensing applications. To handle scanty features in remote sensing images such as island scenes, it is needed to perform data pre-processing to enhance the image for visibility and recognize the distinctness of the features. This may include image enhancement, noise reduction, and contrast adjustment. In addition, it is observed that matched feature points suffer from lot of outliers which ultimately increase the error in estimation of transformation parameters for image co-registration. Robust Outlier removal techniques generate putative keypoint correspondences and helps to achieve sub-pixel image co-registration accuracy between multi-temporal and multi-sensor remote sensing images. Some feature-based registration techniques can be computationally demanding, especially when dealing with large datasets or high-resolution images. However, advances in hardware and algorithms have improved the computational efficiency of these methods. Automation is a key consideration in remote sensing, and feature-based registration methods can be automated to handle large scale data co-registration task and implement batch processing pipelines to handle multiple images efficiently.

Remote sensing image fusion is a vital process in the field of remote sensing and geospatial analysis, aimed at combining information from multiple sensors or data sources to create a more comprehensive and informative image. Image fusion enhances the quality and content of remote sensing data by integrating information from different sensors, resolutions, and spectral bands. This results in images that provide a more complete and detailed representation of the planet's surface. Fused images facilitate more accurate and meaningful analysis of various land cover and land use applications, such as agriculture, forestry, urban planning, and disaster monitoring. They enable better discrimination of objects and features within the imagery. Fusion techniques can be applied to combine data from different remote sensing platforms, including visible, infrared and

radar sensors. This multisource data integration can lead to a richer understanding of complex environmental phenomena. Despite its advantages, remote sensing image fusion can face challenges related to radiometric and geometric differences between the source images, as well as the selection of appropriate fusion techniques and parameters. Before fusion, preprocess the source images to ensure that they are radiometrically and geometrically compatible. The sub-pixel image co-registration of remote sensing image pair need to be ensured before image fusion task. The main goal of our remote sensing image fusion techniques development is to preserve the spectral characteristics of multispectral visible images while improve the spatial resolution from high resolution remote sensing image.

Remote sensing change detection is a crucial process in the field of remote sensing and geospatial analysis, aimed at identifying and quantifying changes in the Earth's surface over time using satellite or aerial imagery. Change detection enables the identification of various environmental and land-use changes, including urban expansion, deforestation, agricultural shifts, and natural disasters. These insights are essential for decision-making in areas such as urban planning, forestry management, disaster response, and environmental monitoring. Integrating data from multiple sensors and time periods enhances the accuracy and reliability of change detection. Different sensors provide complementary information, allowing for a more comprehensive understanding of changes. The choice of image acquisition time intervals is critical for effective change detection. Shorter intervals may capture rapid changes but may also introduce noise, while longer intervals may miss subtle or gradual changes. Various change detection algorithms are available, including pixel-based, object-based, and hybrid approaches. The choice of method should be based on the specific application, data characteristics, and desired level of detail. Change detection can be challenging due to factors such as sensor calibration, atmospheric conditions, radiometric and geometric differences between images, and seasonal variations. Addressing these challenges is crucial for accurate results. We have proposed an enhanced change detection framework that take cares from data pre-processing stage that involves geometric transformation, image co-registration, region of interest extraction to change detection stage that generates change detection map to monitor and capture changes and trends over time. This is especially important for environmental and land-use monitoring for policy level decision making.

## SPECIFIC CONTRIBUTIONS

The contributions of my research work are listed as follows.

- i. Developed a multistage feature based remote sensing image registration technique to handle large geometrical location error at gross stage and can perform perfect adjustment at finer stage to meet sub-pixel co-registration accuracy.
- ii. Multi-satellite island image registration technique using anisotropic coherence locality enhancement and synergistic use of nonlinear diffusion filtering with Mahalanobis distance guided marginalization optimization technique to determine correctly matched points.
- iii. Developed a novel Band-to-Band Registration (BBR) technique for multispectral images using Co-occurrence Scale Space (CSS) based feature detection/description and Spatial Confined RANSAC outlier removal model along with segmented affine transformation model to superimpose spectral bands over each other.
- iv. Direct Feature extraction from raw Bayer pattern planetary remote sensing images, which is not yet explored in planetary science community. In addition, a novel image co-registration methodology is developed to co-register planetary images at raw level and compared the proposed method with state-of-the-art image registration techniques.
- v. Designed novel spectra preserving Bayesian image fusion network using holistic nested deep edges for spatial attention and minnaert function for terrain guidance to combine panchromatic and multispectral remote sensing images.
- vi. The fusion of infrared and visible images is achieved through the innovative combination of Co-occurrence with Laplacian Intensity modulation and advanced color space transformation. This method, to our knowledge, marks the pioneering approach in creating a novel Infrared-Sharpening technique. It effectively merges a thermal infrared broadband image with a multispectral visible remote sensing image, where the former has relatively high spatial resolution and the latter possesses low spatial resolution.
- vii. Designed a Venus visible image processing workflow that includes de-noising and enhancement of the Venus' visible image, co-registration of visible planetary enhanced

image with radar topographic data at sub-pixel level, and fusion of the multi-modal registered datasets to generate best possible representative Venus' image.

- viii. An improved method for detecting changes in remote sensing data involves processing multi-decadal time-lapse optical images to handle data pre-processing and produce a change detection map. This map facilitates direct interpretation of significant changes occurring in the region. The approach is adaptable for detecting surface alterations using various satellite images on Earth and even extends to analyzing changes on Mars.

## **FUTURE SCOPE OF RESEARCH WORK**

The future prospects for feature-based image registration are promising, with numerous avenues for research and development. Firstly, there is a need to explore and enhance feature detection algorithms to efficiently identify distinctive features across various types of images, including those with low texture or repetitive patterns. Additionally, investigating the application of deep learning techniques, such as CNNs or GANs, for feature extraction holds potential to improve registration accuracy by automatically extracting discriminative features from images. The challenges such as occlusions, viewpoint changes, and variations in illumination can be addressed by developing novel feature descriptors and registration algorithms with robust outlier removal model that are resilient to such conditions. Moreover, enhancing the efficiency and scalability of feature-based registration methods for real-time processing of large-scale datasets is crucial, particularly in applications like surveillance and disaster management. Incorporating semantic information into registration methods can lead to more meaningful alignment of images based on scene semantics, while estimating and quantifying uncertainty in the registration process is essential for remote sensing applications. Adaptive registration frameworks and self-supervised learning approaches can further improve registration performance by dynamically adjusting parameters based on image content or learning without explicit supervision.

As for remote sensing image fusion, there is a need to develop advanced fusion algorithms capable of integrating information from multiple sensors and modalities effectively. Semantic-driven fusion methods can prioritize important features or regions in the scene, while dynamic fusion techniques can adaptively adjust parameters based on scene and sensor characteristics. Uncertainty estimation in fused images is crucial for decision-making and risk assessment, and methods to quantify and propagate uncertainty need to be developed. Temporal fusion techniques can enable the creation of temporally consistent fused products, while fusion methods tailored for hyperspectral and LiDAR data can provide valuable complementary information for various applications. The integration of artificial intelligence techniques into fusion processes can automate feature extraction and quality assessment tasks, while tailored fusion algorithms for specific application domains can address unique requirements and challenges. The benchmarks standardization and evaluation protocols are essential for comparing and validating fusion algorithms across different applications and datasets.



In the realm of remote sensing change detection, there is a need to develop algorithms leveraging machine learning and deep learning techniques for automatic detection and classification of changes in imagery. Semantic change analysis and spatio-temporal analysis techniques can provide deeper insights into changes occurring in the environment, while methods for change detection in high-resolution imagery and urban areas can capture fine-scale changes and urban development dynamics. Automated change detection workflows need to be extended to process large volumes of remote sensing data efficiently and generate actionable insights for end-users. This involves integrating change detection algorithms with data preprocessing, feature extraction, and visualization tools to facilitate decision-making and monitoring applications across various domains.

## LIST OF REFERENCES

- Abaspur Kazerouni, I., Dooly, G., & Toal, D. (2020). Underwater image enhancement and mosaicking system based on a-kaze feature matching. *Journal of Marine Science and Engineering*, 8(6), 449.
- Abd El-Samie, F. E., Hadhoud, M. M., & El-Khamy, S. E. (2012). *Image super-resolution and applications*. CRC press.
- Aiazzi, B., Alparone, L., Baronti, S., Garzelli, A., & Selva, M. (2006). MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogrammetric Engineering & Remote Sensing*, 72(5), 591-596.
- Aiazzi, B., Baronti, S., & Selva, M. (2007). Improving component substitution pansharpening through multivariate regression of MS + Pan data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), 3230-3239.
- Albee, A. L., Palluconi, F. D., & Arvidson, R. E. (1998). Mars global surveyor mission: overview and status. *Science*, 279(5357), 1671-1672.
- Alcantarilla, P. F., & Solutions, T. (2011). Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Patt. Anal. Mach. Intell*, 34(7), 1281-1298.
- Alcantarilla, P. F., Bartoli, A., & Davison, A. J. (2012). KAZE features. In *Computer Vision—ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI 12* (pp. 214-227). Springer Berlin Heidelberg.
- Alparone, L., Aiazzi, B., Baronti, S., Garzelli, A., Nencini, F., & Selva, M. (2008). Multispectral and panchromatic data fusion assessment without reference. *Photogrammetric Engineering & Remote Sensing*, 74(2), 193-200.
- Alparone, L., Wald, L., Chanussot, J., Thomas, C., Gamba, P., & Bruce, L. M. (2007). Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), 3012-3021.
- Anzid, H., Le Goic, G., Bekkari, A., Mansouri, A., & Mammass, D. (2017, May). An automatic filtering algorithm for SURF-based registration of remote sensing images. In *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)* (pp. 1-7). IEEE.
- Archinal, B. A., Lee, E. M., Kirk, R. L., Duxbury, T. C., Sucharski, R. M., Cook, D. A., & Barrett, J. M. (2004). A new Mars digital image model (MDIM 2.1) control network. *International Archives of Photogrammetry and Remote Sensing*, 35, B4.
- Arunan, S., & Satish, R. (2015). Mars Orbiter Mission spacecraft and its challenges. *Current Science*, 1061-1069.
- Arya, A. S., Rajasekhar, R. P., Singh, R. B., Sur, K., Chauhan, P., Sarkar, S. S., ... & Kiran Kumar, A. S. (2015, March). Mars color camera onboard mars orbiter mission: Initial observations and results. In *46th Annual Lunar and Planetary Science Conference* (No. 1832, p. 2123).
- Arya, A. S., Sarkar, S. S., Srinivas, A. R., Moorthi, S. M., Patel, V. D., Singh, R. B., ... & Kumar, A. K. (2015). Mars Colour Camera: The payload characterization/calibration and data analysis from Earth imaging phase. *Current Science*, 109(1076), 1076-1086.

- Asokan, A., & Anitha, J. J. E. S. I. (2019). Change detection techniques for remote sensing applications: A survey. *Earth Science Informatics*, *12*, 143-160.
- Atkinson, P. M., Sargent, I. M., Foody, G. M., & Williams, J. (2005). Interpreting image-based methods for estimating the signal-to-noise ratio. *International Journal of Remote Sensing*, *26*(22), 5099-5115.
- Avola, D., Foresti, G. L., Martinel, N., Micheloni, C., Pannone, D., & Piciarelli, C. (2017). Real-time incremental and geo-referenced mosaicking by small-scale UAVs. In *Image Analysis and Processing-ICIAP 2017: 19th International Conference, Catania, Italy, September 11-15, 2017, Proceedings, Part I 19* (pp. 694-705). Springer International Publishing.
- Azarang, A., & Kehtarnavaz, N. (2020). Image fusion in remote sensing by multi-objective deep learning. *International Journal of Remote Sensing*, *41*(24), 9507-9524.
- Azarang, A., Manoochehri, H. E., & Kehtarnavaz, N. (2019). Convolutional autoencoder-based multispectral image fusion. *IEEE access*, *7*, 35673-35683.
- Ballester, C., Caselles, V., Igual, L., Verdera, J., & Rougé, B. (2006). A variational model for P+ XS image fusion. *International Journal of Computer Vision*, *69*, 43-58.
- Baltsavias, E., Kocaman, S., Akca, D., & Wolff, K. (2007). Geometric and radiometric investigations of Cartosat-1 data. In *ISPRS Workshop "High Resolution Earth Imaging for Geospatial Information"* (Vol. 36). ISPRS.
- Barath, D., Matas, J., & Noskova, J. (2019). MAGSAC: marginalizing sample consensus. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10197-10205).
- Barath, D., Noskova, J., & Matas, J. (2021). Marginalizing sample consensus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(11), 8420-8432.
- Barker, J. L., & Seiferth, J. C. (1996, May). Landsat thematic mapper band-to-band registration. In *IGARSS'96. 1996 International Geoscience and Remote Sensing Symposium* (Vol. 3, pp. 1600-1602). IEEE.
- Basilevsky, A. T., Shalygin, E. V., Titov, D. V., Markiewicz, W. J., Scholten, F., Roatsch, T., ... & Michalik, H. (2012). Geologic interpretation of the near-infrared images of the surface taken by the Venus Monitoring Camera, Venus Express. *Icarus*, *217*(2), 434-450.
- Batson, R. M., Kirk, R. L., Edwards, K., & Morgan, H. F. (1994). Venus cartography. *Journal of Geophysical Research: Planets*, *99*(E10), 21173-21181.
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. In *Computer Vision—ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9* (pp. 404-417). Springer Berlin Heidelberg.
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, *110*(3), 346-359.
- Bayer, B. (1976). Color imaging array. *United States Patent, no. 3971065*.
- Bellavia, F., & Mishkin, D. (2022). HarrisZ+: Harris corner selection for next-gen image matching pipelines. *Pattern Recognition Letters*, *158*, 141-147.

- Bentoutou, Y., Taleb, N., Kpalma, K., & Ronsin, J. (2005). An automatic image registration for applications in remote sensing. *IEEE transactions on geoscience and remote sensing*, 43(9), 2127-2137.
- Benzenati, T., Kallel, A., & Kessentini, Y. (2020). Two stages pan-sharpening details injection approach based on very deep residual networks. *IEEE Transactions on Geoscience and Remote Sensing*, 59(6), 4984-4992.
- Bernardo, J. M., & Smith, A. F. (2009). *Bayesian theory* (Vol. 405). John Wiley & Sons.
- Bian, J., Lin, W. Y., Matsushita, Y., Yeung, S. K., Nguyen, T. D., & Cheng, M. M. (2017). Gms: Grid-based motion statistics for fast, ultra-robust feature correspondence. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4181-4190).
- Bolles, R. C., & Fischler, M. A. (1981, August). A RANSAC-based approach to model fitting and its application to finding cylinders in range data. In *IJCAI* (Vol. 1981, pp. 637-643).
- Bonnet, N., Cutrona, J., & Herbin, M. (2002). A 'no-threshold' histogram-based image segmentation method. *Pattern Recognition*, 35(10), 2319-2322.
- Bouchiha, R., & Besbes, K. (2013). Automatic remote-sensing image registration using SURF. *International Journal of Computer Theory and Engineering*, 5(1), 88-92.
- Boukerch, I., Farhi, N., Karoui, M. S., Djerriri, K., & Mahmoudi, R. (2018, July). A Dense Vector Matching Approach for Band to Band Registration of Alsat-2 Images. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3401-3404). IEEE.
- Brook, A., & Ben-Dor, E. (2011). Automatic registration of airborne and spaceborne images by topology map matching with SURF processor algorithm. *Remote Sensing*, 3(1), 65-82.
- Brown, L. G. (1992). A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4), 325-376.
- Brown, M., & Lowe, D. G. (2007). Automatic panoramic image stitching using invariant features. *International journal of computer vision*, 74, 59-73.
- Bruzzo, L., & Prieto, D. F. (2000). Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote sensing*, 38(3), 1171-1182.
- Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). Brief: Binary robust independent elementary features. In *Computer Vision—ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11* (pp. 778-792). Springer Berlin Heidelberg.
- Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing*. Guilford press.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 679-698.
- Canty, M. J., Nielsen, A. A., & Schmidt, M. (2004). Automatic radiometric normalization of multitemporal satellite imagery. *Remote sensing of Environment*, 91(3-4), 441-451.
- Caraffa, L., Tarel, J. P., & Charbonnier, P. (2015). The guided bilateral filter: When the joint/cross bilateral filter becomes robust. *IEEE Transactions on Image Processing*, 24(4), 1199-1208.

- Carper, W., Lillesand, T., & Kiefer, R. (1990). The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogrammetric Engineering and remote sensing*, 56(4), 459-467.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific model development discussions*, 7(1), 1525-1534.
- Chang, H. H., Wu, G. L., & Chiang, M. H. (2019). Remote sensing image registration based on modified SIFT and feature slope grouping. *IEEE Geoscience and Remote Sensing Letters*, 16(9), 1363-1367.
- Chang, Y., Jung, C., Ke, P., Song, H., & Hwang, J. (2018). Automatic contrast-limited adaptive histogram equalization with dual gamma correction. *Ieee Access*, 6, 11782-11792.
- Chavez, P., Sides, S. C., & Anderson, J. A. (1991). Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic. *Photogrammetric Engineering and remote sensing*, 57(3), 295-303.
- Chen, C., Li, Y., Liu, W., & Huang, J. (2015). SIRF: Simultaneous satellite image registration and fusion in a unified framework. *IEEE Transactions on Image Processing*, 24(11), 4213-4224.
- Chen, G., Zhu, F., & Ann Heng, P. (2015). An efficient statistical method for image noise level estimation. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 477-485).
- Chen, G., Zhu, F., & Ann Heng, P. (2015). An efficient statistical method for image noise level estimation. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 477-485).
- Chen, H., & Shi, Z. (2020). A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sensing*, 12(10), 1662.
- Chen, J., Luo, L., Wang, S., & Wu, H. (2018, July). Automatic panoramic UAV image mosaic using ORB features and robust transformation estimation. In *2018 37th Chinese Control Conference (CCC)* (pp. 4265-4270). IEEE.
- Chen, S., Zhong, S., Xue, B., Li, X., Zhao, L., & Chang, C. I. (2020). Iterative scale-invariant feature transform for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4), 3244-3265.
- Chen, T., & Liu, Y. (2021). A Quick Band-to-Band Mis-Registration Detection Method for Sentinel-2 MSI Images. *Remote Sensing*, 13(17), 3351.
- Chen, T., Ma, K. K., & Chen, L. H. (1999). Tri-state median filter for image denoising. *IEEE Transactions on Image processing*, 8(12), 1834-1838.
- Cheng, C., Sun, C., Sun, Y., & Zhu, J. (2022). StyleFuse: An unsupervised network based on style loss function for infrared and visible image fusion. *Signal Processing: Image Communication*, 106, 116722.
- Cheng, L., Li, M., Liu, Y., Cai, W., Chen, Y., & Yang, K. (2012). Remote sensing image matching by integrating affine invariant feature extraction and RANSAC. *Computers & Electrical Engineering*, 38(4), 1023-1032.
- Cheng, M. L., & Matsuoka, M. (2021). An Efficient and Precise Remote Sensing Optical Image Matching Technique Using Binary-Based Feature Points. *Sensors*, 21(18), 6035.
- Chevrier, V., & Mathé, P. E. (2007). Mineralogy and evolution of the surface of Mars: a review. *Planetary and Space Science*, 55(3), 289-314.

- Chibani, Y., & Houacine, A. (2002). The joint use of IHS transform and redundant wavelet decomposition for fusing multispectral and panchromatic images. *International Journal of Remote Sensing*, 23(18), 3821-3833.
- Chicarro, A., Martin, P., & Trautner, R. (2004). The Mars Express mission: an overview. In: *Mars Express: the scientific payload*. Ed. by Andrew Wilson, scientific coordination: Agustin Chicarro. ESA SP-1240, Noordwijk, Netherlands: ESA Publications Division, ISBN 92-9092-556-6, 2004, p. 3-13, 1240, 3-13.
- Choi, J., Yu, K., & Kim, Y. (2010). A new adaptive component-substitution-based satellite image fusion by using partial replacement. *IEEE transactions on geoscience and remote sensing*, 49(1), 295-309.
- Choi, S., Kim, T., & Yu, W. (1997). Performance evaluation of RANSAC family. *Journal of Computer Vision*, 24(3), 271-300.
- Choi, T., Shao, X., & Cao, C. (2019, July). NOAA-20 Visible Infrared Imaging Radiometer Suite (VIIRS) on-orbit band-to-band registration estimation for reflective solar band (RSB) using scheduled lunar collections. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 9059-9062). IEEE.
- Chu, H., & Zhu, W. (2008). Fusion of IKONOS satellite imagery using IHS transform and local variation. *IEEE Geoscience and Remote Sensing Letters*, 5(4), 653-657.
- Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J. C., ... & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote sensing of environment*, 219, 145-161.
- Colleu, T., Shen, J. K., Matuszewski, B., Shark, L. K., & Cariou, C. (2006, September). Feature-based deformable image registration with ransac based search correspondence. In *AECRIS'06-Atlantic Europe Conference on Remote Imaging and Spectroscopy* (pp. 57-64).
- Craddock, R. A., & Maxwell, T. A. (1990). Resurfacing of the Martian highlands in the Amenthes and Tyrrhena region. *Journal of Geophysical Research: Solid Earth*, 95(B9), 14265-14278.
- DadrasJavan, F., Samadzadegan, F., & Fathollahi, F. (2018). Spectral and spatial quality assessment of IHS and wavelet based pan-sharpening techniques for high resolution satellite imagery. *Advances in image and video processing*, 6(2), 01.
- Davis, L. S. (1975). A survey of edge detection techniques. *Computer graphics and image processing*, 4(3), 248-270.
- De Bethune, S., Muller, F., & Binard, M. (1998, September). Adaptive intensity matching filters: a new tool for multi-resolution data fusion. In *AGARD CONFERENCE PROCEEDINGS AGARD CP* (pp. 28-28). AGARD.
- De Leeuw, A. J., Veugen, L. M. M., & Van Stokkom, H. T. C. (1988). Geometric correction of remotely-sensed imagery using ground control points and orthogonal polynomials. *International journal of remote sensing*, 9(10-11), 1751-1759.
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, 6(2), 182-197.
- Dechoz, C., Poulain, V., Massera, S., Languille, F., Greslou, D., De Lussy, F., ... & Trémas, T. (2015, October). Sentinel 2 global reference image. In *Image and signal processing for remote sensing XXI* (Vol. 9643, pp. 94-107). SPIE.

- Demchev, D., Volkov, V., Kazakov, E., Alcantarilla, P. F., Sandven, S., & Khmeleva, V. (2017). Sea ice drift tracking from sequential SAR images using accelerated-KAZE features. *IEEE Transactions on Geoscience and Remote Sensing*, 55(9), 5174-5184.
- Deng, J. S., Wang, K., Deng, Y. H., & Qi, G. J. (2008). PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, 29(16), 4823-4838.
- Derpanis, K. G. (2004). The harris corner detector. *York University*, 2, 1-2.
- Derpanis, K. G. (2010). Overview of the RANSAC Algorithm. *Image Rochester NY*, 4(1), 2-3.
- Desachy, J., Begni, G., Boissin, B., & Perbos, J. (1985). Investigation of Landsat-4 Thematic Mapper line-to-line and band-to-band registration and relative detector calibration. *Photogrammetric Engineering and Remote Sensing*, 51(9), 1291-1298.
- Dial, G., Bowen, H., Gerlach, F., Grodecki, J., & Oleszczuk, R. (2003). IKONOS satellite, imagery, and products. *Remote sensing of Environment*, 88(1-2), 23-36.
- Dibs, H., Mansor, S., Ahmad, N., & Pradhan, B. (2015). Band-to-band registration model for near-equatorial Earth observation satellite images with the use of automatic control point extraction. *International Journal of Remote Sensing*, 36(8), 2184-2200.
- Ding, L., & Goshtasby, A. (2001). On the Canny edge detector. *Pattern recognition*, 34(3), 721-725.
- Ding, Z., Li, H., Zhou, D., Li, H., Liu, Y., & Hou, R. (2021). CMFA\_Net: A cross-modal feature aggregation network for infrared-visible image fusion. *Infrared Physics & Technology*, 118, 103905.
- Du, B., Ru, L., Wu, C., & Zhang, L. (2019). Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12), 9976-9992.
- Du, J., Li, W., Lu, K., & Xiao, B. (2016). An overview of multi-modal medical image fusion. *Neurocomputing*, 215, 3-20.
- Duplaquet, M. L. (1998, July). Building large image mosaics with invisible seam lines. In *Visual information processing VII* (Vol. 3387, pp. 369-377). SPIE.
- Durgam, U. K., Paul, S., & Pati, U. C. (2016, March). SURF based matching for SAR image registration. In *2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (pp. 1-5). IEEE.
- Dwyer, J. L., Roy, D. P., Sauer, B., Jenkerson, C. B., Zhang, H. K., & Lymburner, L. (2018). Analysis ready data: enabling analysis of the Landsat archive. *Remote Sensing*, 10(9), 1363.
- Edwards, C. S., Nowicki, K. J., Christensen, P. R., Hill, J., Gorelick, N., & Murray, K. (2011). Mosaicking of global planetary image datasets: 1. Techniques and data processing for Thermal Emission Imaging System (THEMIS) multi-spectral data. *Journal of Geophysical Research: Planets*, 116(E10).
- Elad, M. (2002). On the origin of the bilateral filter and ways to improve it. *IEEE Transactions on image processing*, 11(10), 1141-1151.
- Erives, H., & Fitzgerald, G. J. (2005). Automated registration of hyperspectral images for precision agriculture. *Computers and Electronics in Agriculture*, 47(2), 103-119.

- Fanqing, M., & Fucheng, Y. (2013, December). A tracking algorithm based on ORB. In *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)* (pp. 1187-1190). IEEE.
- Fasbender, D., Radoux, J., & Bogaert, P. (2008). Bayesian data fusion for adaptable image pansharpener. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(6), 1847-1857.
- Fernandez-Beltran, R., Latorre-Carmona, P., & Pla, F. (2017). Single-frame super-resolution in remote sensing: A practical overview. *International journal of remote sensing*, *38*(1), 314-354.
- Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381-395.
- Flood, N. (2013). Seasonal composite Landsat TM/ETM+ images using the medoid (a multi-dimensional median). *Remote Sensing*, *5*(12), 6481-6500.
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment*, *80*(1), 185-201.
- Fotouhi, M., Hekmatian, H., Kashani-Nezhad, M. A., & Kasaei, S. (2019). SC-RANSAC: spatial consistency on RANSAC. *Multimedia Tools and Applications*, *78*, 9429-9461.
- Fu, J., Zhao, J., & Li, F. (2021). Infrared sea-sky line detection utilizing self-adaptive laplacian of gaussian filter and visual-saliency-based probabilistic Hough transform. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1-5.
- Fung, T., & LeDrew, E. (1987). Application of principal components analysis to change detection. *Photogrammetric engineering and remote sensing*, *53*(12), 1649-1658.
- Fung, T., & LeDrew, E. (1988). For change detection using various accuracy. *Photogramm. Eng. Remote Sens*, *54*(10), 1449-1454.
- Gao, C., Li, W., Tao, R., & Du, Q. (2022). MS-HLMO: multiscale histogram of local main orientation for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-14.
- Gao, F., Dong, J., Li, B., & Xu, Q. (2016). Automatic change detection in synthetic aperture radar images based on PCANet. *IEEE geoscience and remote sensing letters*, *13*(12), 1792-1796.
- Garzelli, A., Nencini, F., & Capobianco, L. (2007). Optimal MMSE pan sharpening of very high resolution multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(1), 228-236.
- Gastineau, A., Aujol, J. F., Berthoumieu, Y., & Germain, C. (2021). Generative adversarial network for pansharpener with spectral and spatial discriminators. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-11.
- Ge, H., Lu, D., He, S., Xu, A., Zhou, G., & Du, H. (2008). Pixel-based Minnaert correction method for reducing topographic effects on a Landsat 7 ETM+ image. *Photogrammetric Engineering & Remote Sensing*, *74*(11), 1343-1350.
- George, G., Oommen, R. M., Shelly, S., Philipose, S. S., & Varghese, A. M. (2018, March). A survey on various median filtering techniques for removal of impulse noise from digital image. In *2018 Conference on Emerging Devices and Smart Systems (ICEDSS)* (pp. 235-238). IEEE.



- Ghassemian, H. (2016). A review of remote sensing image fusion methods. *Information Fusion*, 32, 75-89.
- Gillespie, A. R., Kahle, A. B., & Walker, R. E. (1987). Color enhancement of highly correlated images. II. Channel ratio and “chromaticity” transformation techniques. *Remote Sensing of Environment*, 22(3), 343-365.
- Goncalves, H., Corte-Real, L., & Goncalves, J. A. (2011). Automatic image registration through image segmentation and SIFT. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7), 2589-2600.
- Gong, M., Zhao, S., Jiao, L., Tian, D., & Wang, S. (2013). A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE transactions on geoscience and remote sensing*, 52(7), 4328-4338.
- González-Audícana, M., Saleta, J. L., Catalán, R. G., & García, R. (2004). Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition. *IEEE Transactions on Geoscience and Remote sensing*, 42(6), 1291-1299.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, 18-27.
- Goudarzi, M. A., & Landry, R. J. (2017). Assessing horizontal positional accuracy of Google Earth imagery in the city of Montreal, Canada. *Geodesy and cartography*, 43(2), 56-65.
- Guiqin, Y. A. N. G., Chang, X., & Jiang, Z. (2019, August). A fast aerial images mosaic method based on ORB feature and homography matrix. In *2019 International Conference on Computer, Information and Telecommunication Systems (CITS)* (pp. 1-5). IEEE.
- Guizar-Sicairos, M., Thurman, S. T., & Fienup, J. R. (2008). Efficient subpixel image registration algorithms. *Optics letters*, 33(2), 156-158.
- Gungor, O. (2008). *Multi sensor multi resolution image fusion* (Doctoral dissertation, Purdue University).
- Guo, H., Cheng, C., & Yang, Y. (2010, July). An automated registration of RS images based on SURF and piecewise linear transformation. In *2010 The 2nd Conference on Environmental Science and Information Application Technology* (Vol. 3, pp. 133-136). IEEE.
- Guo, H., Xu, H., Wei, Y., Shen, Y., & Rui, X. (2022). Outlier removal and feature point pairs optimization for piecewise linear transformation in the co-registration of very high-resolution optical remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 193, 299-313.
- Guo, L. J., & Moore, J. M. (1998). Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery. *International Journal of Remote Sensing*, 19(13), 2477-2491.
- Guo, M., Li, J., Sheng, C., Xu, J., & Wu, L. (2017). A review of wetland remote sensing. *Sensors*, 17(4), 777.
- Guo, Z., Yu, X., & Du, Q. (2022). Infrared and visible image fusion based on saliency and fast guided filtering. *Infrared Physics & Technology*, 123, 104178.
- Gurunlu, B., & Ozturk, S. (2022). Efficient approach for block-based copy-move forgery detection. In *Smart Trends in Computing and Communications: Proceedings of SmartCom 2021* (pp. 167-174). Springer Singapore.

- Hao, S., Han, X., Zhang, Y., & Xu, L. (2021). Low-light enhancement based on an improved simplified Retinex model via fast illumination map refinement. *Pattern Analysis and Applications*, 24, 321-332.
- Harinath, N., Mahadevan, V., & Sarma, K. S. (2008). Resourcesat-1 mission planning, analysis and operations— Outline of key components. *International journal of applied earth observation and geoinformation*, 10(2), 124-132.
- Harris, C., & Stephens, M. (1988, August). A combined corner and edge detector. In *Alvey vision conference* (Vol. 15, No. 50, pp. 10-5244).
- Hashimoto, T. (1999). Effect of interpolation on NDVI derived from the images with band-to-band registration errors. *Journal of the Japan society of Photogrammetry and Remote Sensing*, 38(2), 61-65.
- Hassanpour, M., Dadras Javan, F., & Azizi, A. (2019). Band to band registration of multi-spectral aerial imagery— relief displacement and miss-registration error. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 467-474.
- He, K., Sun, J., & Tang, X. (2012). Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, 35(6), 1397-1409.
- Hernández, C. A. V., & Ortiz, F. A. P. (2020). A real-time map merging strategy for robust collaborative reconstruction of unknown environments. *Expert Systems with Applications*, 145, 113109.
- Honkavaara, E., Rosnell, T., Oliveira, R., & Tommaselli, A. (2017). Band registration of tuneable frame format hyperspectral UAV imagers in complex scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 96-109.
- Hosseini-Nejad, Z., & Nasri, M. (2017). An adaptive image registration method based on SIFT features and RANSAC transform. *Computers & Electrical Engineering*, 62, 524-537.
- Hosseini-Nejad, Z., Agahi, H., & Mahmoodzadeh, A. (2021). Image matching based on the adaptive redundant keypoint elimination method in the SIFT algorithm. *Pattern Analysis and Applications*, 24, 669-683.
- Hotelling, H. (1992). Relations between two sets of variates. *Proceeding of Breakthroughs in Statistics*, Springer, New York, NY. , 162-190.
- <https://earthexplorer.usgs.gov/> (Accessed on 17-April-2023)
- <https://ladsweb.modaps.eosdis.nasa.gov/> (Accessed on 17-April-2023)
- <https://www.issdc.gov.in> (Accessed on 17-April-2023)
- Huang, L., & Li, Z. (2010). Feature-based image registration using the shape context. *International Journal of Remote Sensing*, 31(8), 2169-2177.
- Huang, W., Fei, X., Feng, J., Wang, H., Liu, Y., & Huang, Y. (2020). Pan-sharpening via multi-scale and multiple deep neural networks. *Signal Processing: Image Communication*, 85, 115850.
- Huet, F., & Philipp, S. (1998). Fusion of images interpreted by a new fuzzy classifier. *Pattern Analysis and Applications*, 1, 231-247.

- Huo, C., Pan, C., Huo, L., & Zhou, Z. (2011). Multilevel SIFT matching for large-size VHR image registration. *IEEE Geoscience and Remote Sensing Letters*, 9(2), 171-175.
- Inglada, J. (2007). Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. *ISPRS journal of photogrammetry and remote sensing*, 62(3), 236-248.
- IRS-1A Spacecraft. (2021) [Accessed 2023 May 16]. <https://www.isro.gov.in/Spacecraft/irs-1a>.
- Jevnisek, R. J., & Avidan, S. (2017). Co-occurrence filter. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3184-3192).
- Jhan, J. P., Rau, J. Y., & Huang, C. Y. (2016). Band-to-band registration and ortho-rectification of multilens/multispectral imagery: A case study of MiniMCA-12 acquired by a fixed-wing UAS. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 66-77.
- Jhan, J. P., & Rau, J. Y. (2021). A generalized tool for accurate and efficient image registration of UAV multi-lens multispectral cameras by N-SURF matching. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 6353-6362.
- Jiang, M., Shen, H., Li, J., Yuan, Q., & Zhang, L. (2020). A differential information residual convolutional neural network for pansharpening. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163, 257-271.
- Jianya, G., Haigang, S., Guorui, M., & Qiming, Z. (2008). A review of multi-temporal remote sensing data change detection algorithms. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7), 757-762.
- Johnson, R. D., & Kasischke, E. S. (1998). Change vector analysis: A technique for the multispectral monitoring of land cover and condition. *International journal of remote sensing*, 19(3), 411-426.
- Kalantar, B., Ueda, N., Al-Najjar, H. A. H., Moayedi, H., Halin, A. A., & Mansor, S. (2019). UAV and lidar image registration: A SURF-based approach for ground control points selection. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 413-418.
- Kanopoulos, N., Vasanthavada, N., & Baker, R. L. (1988). Design of an image edge detection filter using the Sobel operator. *IEEE Journal of solid-state circuits*, 23(2), 358-367.
- Kasturirangan, K., Aravamudan, R., Deekshatulu, B. L., Joseph, G., & Chandrasekhar, M. G. (1996). Indian remote sensing satellite (IRS)-1C—The beginning of a new era. *Current science*, 70(7), 495-500.
- Kaur, H., Koundal, D., & Kadyan, V. (2021). Image fusion techniques: a survey. *Archives of computational methods in Engineering*, 28, 4425-4447.
- Khelifi, L., & Mignotte, M. (2020). Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis. *Ieee Access*, 8, 126385-126400.
- Kiku, D., Monno, Y., Tanaka, M., & Okutomi, M. (2016). Beyond color difference: Residual interpolation for color image demosaicking. *IEEE Transactions on Image Processing*, 25(3), 1288-1300.

- Kim, T., & Im, Y. J. (2003). Automatic satellite image registration by combination of matching and random sample consensus. *IEEE transactions on geoscience and remote sensing*, 41(5), 1111-1117.
- Kim, W., Jung, S., Moon, Y., & Mangum, S. C. (2020). Morphological band registration of multispectral cameras for water quality analysis with unmanned aerial vehicle. *Remote Sensing*, 12(12), 2024.
- Kinch, K. M., Madsen, M. B., Bell, J. F., Maki, J. N., Bailey, Z. J., Hayes, A. G., ... & Winhold, A. (2020). Radiometric calibration targets for the Mastcam-Z camera on the Mars 2020 rover mission. *Space science reviews*, 216, 1-51.
- Kiran Kumar, A. S., & Samudraiah, D. R. M. (2021). Remembering IRS-1C Imaging Payloads: A Giant Leap in Indian Space Imaging System Evolution. *Journal of the Indian Society of Remote Sensing*, 49, 23-40.
- Kittler, J. (1983). On the accuracy of the Sobel edge detector. *Image and Vision Computing*, 1(1), 37-42.
- Klein, H. P., Lederberg, J., Rich, A., Horowitz, N. H., Oyama, V. I., & Levin, G. V. (1976). The Viking Mission search for life on Mars. *Nature*, 262(5563), 24-27.
- Kumar, R., Singh, N., Kumar, A., Desai, J. H., & Kumaran, R. (2021, December). High Performance camera electronics for Infrared payloads. In *2021 IEEE 18th India Council International Conference (INDICON)* (pp. 1-5). IEEE.
- Kupfer, B., Netanyahu, N. S., & Shimshoni, I. (2014). An efficient SIFT-based mode-seeking algorithm for sub-pixel registration of remotely sensed images. *IEEE Geoscience and Remote Sensing Letters*, 12(2), 379-383.
- Le Moigne, J., Netanyahu, N. S., & Eastman, R. D. (Eds.). (2011). *Image registration for remote sensing*. Cambridge University Press.
- Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In *2011 International conference on computer vision* (pp. 2548-2555). Ieee.
- Levin, A., Zomet, A., Peleg, S., & Weiss, Y. (2004). Seamless image stitching in the gradient domain. In *Computer Vision-ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part IV 8* (pp. 377-389). Springer Berlin Heidelberg.
- Li, A., & Zhang, B. (2012, May). Registration algorithm of panchromatic and multi-spectral images based on sift and ransac. In *2012 Symposium on Photonics and Optoelectronics* (pp. 1-4). IEEE.
- Li, B., Ming, D., Yan, W., Sun, X., Tian, T., & Tian, J. (2014). Image matching based on two-column histogram hashing and improved RANSAC. *IEEE Geoscience and Remote Sensing Letters*, 11(8), 1433-1437.
- Li, C., Guo, B., Guo, X., & Zhi, Y. (2018, August). Real-time UAV imagery stitching based on grid-based motion statistics. In *Journal of Physics: Conference Series* (Vol. 1069, No. 1, p. 012163). IOP Publishing.
- Li, H., Qin, J., Xiang, X., Pan, L., Ma, W., & Xiong, N. N. (2018). An efficient image matching algorithm based on adaptive threshold and RANSAC. *Ieee Access*, 6, 66963-66971.
- Li, H., Yu, Y., & Du, W. (2022, March). Research on rapid image registration based on feature point detection. In *Eighth Symposium on Novel Photoelectronic Detection Technology and Applications* (Vol. 12169, pp. 964-969). SPIE.

- Li, L., Han, L., Ding, M., Cao, H., & Hu, H. (2021). A deep learning semantic template matching framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 181, 205-217.
- Li, L., Yao, J., Lu, X., Tu, J., & Shan, J. (2016). Optimal seamline detection for multiple image mosaicking via graph cuts. *ISPRS Journal of Photogrammetry and Remote Sensing*, 113, 1-16.
- Li, Q., Wang, G., Liu, J., & Chen, S. (2009). Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, 6(2), 287-291.
- Li, X., & Yeh, A. G. O. (1998). Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta. *International Journal of Remote Sensing*, 19(8), 1501-1518.
- Li, X., Feng, R., Guan, X., Shen, H., & Zhang, L. (2019). Remote sensing image mosaicking: Achievements and challenges. *IEEE Geoscience and Remote Sensing Magazine*, 7(4), 8-22.
- Li, X., Hui, N., Shen, H., Fu, Y., & Zhang, L. (2015). A robust mosaicking procedure for high spatial resolution remote sensing images. *ISPRS journal of photogrammetry and remote sensing*, 109, 108-125.
- Liao, F., Chen, Y., Chen, Y., & Lu, Y. (2020, September). SAR image registration based on optimized Ransac algorithm with mixed feature extraction. In *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1153-1156). IEEE.
- Liu, H., & Xiao, G. F. (2020). Remote sensing image registration based on improved KAZE and BRIEF descriptor. *International Journal of Automation and Computing*, 17(4), 588-598.
- Liu, J. G. (2000). Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details. *International Journal of Remote Sensing*, 21(18), 3461-3472.
- Liu, J., & Liang, S. (2016). Pan-sharpening using a guided filter. *International Journal of Remote Sensing*, 37(8), 1777-1800.
- Liu, L., Liu, B., Huang, H., & Bovik, A. C. (2014). No-reference image quality assessment based on spatial and spectral entropies. *Signal processing: Image communication*, 29(8), 856-863.
- Liu, Q., Zhou, H., Xu, Q., Liu, X., & Wang, Y. (2020). PSGAN: A generative adversarial network for remote sensing image pan-sharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 59(12), 10227-10242.
- Liu, X. J., Yang, J., & Shen, H. B. (2008). Automatic image registration by local descriptors in remote sensing. *Optical Engineering*, 47(8), 087206-087206.
- Liu, Y., Dong, L., & Xu, W. (2022). Infrared and visible image fusion via salient object extraction and low-light region enhancement. *Infrared Physics & Technology*, 124, 104223.
- Liu, Y., Lan, C., Li, C., Mo, F., & Wang, H. (2016). S-AKAZE: An effective point-based method for image matching. *Optik*, 127(14), 5670-5681.
- Liu, Y., Lan, C., Yao, F., Li, L., & Li, C. (2016, May). Oblique remote sensing image matching based on improved AKAZE algorithm. In *2016 Sixth International Conference on Information Science and Technology (ICIST)* (pp. 448-454). IEEE.

- Liu, Y., Zhang, H., Guo, H., & Xiong, N. N. (2018). A fast-brisk feature detector with depth information. *Sensors*, *18*(11), 3908.
- Liu, Y., Zhou, D., Nie, R., Ding, Z., Guo, Y., Ruan, X., ... & Hou, R. (2022). TSE\_Fuse: Two stage enhancement method using attention mechanism and feature-linking model for infrared and visible image fusion. *Digital Signal Processing*, *123*, 103387.
- Liu, Y., He, W., & Zhang, H. (2023). GLoCNet: Robust Feature Matching with Global-local Consistency Network for Remote Sensing Image Registration. *IEEE Transactions on Geoscience and Remote Sensing*.
- Loveland, T. R., & Dwyer, J. L. (2012). Landsat: Building a strong future. *Remote Sensing of Environment*, *122*, 22-29.
- Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1150-1157). Ieee.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*, 91-110.
- Lu, D., Mausel, P., Brondizio, E., & Moran, E. (2004). Change detection techniques. *International journal of remote sensing*, *25*(12), 2365-2401.
- Lu, Y., Gao, K., Zhang, T., & Xu, T. (2018). A novel image registration approach via combining local features and geometric invariants. *PloS one*, *13*(1), e0190383.
- Lukac, R., & Plataniotis, K. N. (2005). Universal demosaicking for imaging pipelines with an RGB color filter array. *Pattern Recognition*, *38*(11), 2208-2212.
- Luo, C., Yang, W., Huang, P., & Zhou, J. (2019, June). Overview of image matching based on ORB algorithm. In *Journal of Physics: Conference Series* (Vol. 1237, No. 3, p. 032020). IOP Publishing.
- Ma, C., Hu, X., Xiao, J., Du, H., & Zhang, G. (2020). Improved ORB algorithm using three-patch method and local gray difference. *Sensors*, *20*(4), 975.
- Ma, D., & Lai, H. C. (2019). Remote sensing image matching based improved ORB in NSCT domain. *Journal of the Indian Society of Remote Sensing*, *47*, 801-807.
- Ma, J., Yu, W., Chen, C., Liang, P., Guo, X., & Jiang, J. (2020). Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. *Information Fusion*, *62*, 110-120.
- Ma, W., Wen, Z., Wu, Y., Jiao, L., Gong, M., Zheng, Y., & Liu, L. (2016). Remote sensing image registration with modified SIFT and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, *14*(1), 3-7.
- Ma, W., Wu, Y., Liu, S., Su, Q., & Zhong, Y. (2018). Remote sensing image registration based on phase congruency feature detection and spatial constraint matching. *IEEE Access*, *6*, 77554-77567.
- Ma, W., Zhang, J., Wu, Y., Jiao, L., Zhu, H., & Zhao, W. (2019). A novel two-step registration method for remote sensing images based on deep and local features. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(7), 4834-4843.

- Mahalanobis, P. C. (2018). On the generalized distance in statistics. *Sankhyā: The Indian Journal of Statistics, Series A (2008-), 80*, S1-S7.
- Maini, D. S., & Aggarwal, A. K. (2018). Camera position estimation using 2D image dataset. *Int. J. Innov. Eng. Technol., 10*, 199-203.
- Maintz, J. A., & Viergever, M. A. (1998). A survey of medical image registration. *Medical image analysis, 2*(1), 1-36.
- Mair, E., Hager, G. D., Burschka, D., Suppa, M., & Hirzinger, G. (2010). Adaptive and generic corner detection based on the accelerated segment test. In *Computer Vision—ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part II 11* (pp. 183-196). Springer Berlin Heidelberg.
- Malin, M. C., & Edgett, K. S. (2001). Mars global surveyor Mars orbiter camera: interplanetary cruise through primary mission. *Journal of Geophysical Research: Planets, 106*(E10), 23429-23570.
- Markham, B. L., Storey, J. C., Williams, D. L., & Irons, J. R. (2004). Landsat sensor performance: history and current status. *IEEE transactions on geoscience and remote sensing, 42*(12), 2691-2694.
- Mars Viking. (2021). [accessed 2021 Aug 03]. [https://astrogeology.usgs.gov/search/map/Mars/Viking/Color/Mars\\_Viking\\_ClrMosaic\\_global\\_925m](https://astrogeology.usgs.gov/search/map/Mars/Viking/Color/Mars_Viking_ClrMosaic_global_925m).
- Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001, July). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001* (Vol. 2, pp. 416-423). IEEE.
- Masi, G., Cozzolino, D., Verdoliva, L., & Scarpa, G. (2016). Pansharpening by convolutional neural networks. *Remote Sensing, 8*(7), 594.
- Mason, A., Rioux, J., Clarke, S. E., Costa, A., Schmidt, M., Keough, V., ... & Beyea, S. (2019). Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of MR images. *IEEE transactions on medical imaging, 39*(4), 1064-1072.
- Mathew, K., Arya, A. S., Seth, H., Moorthi, S. M., & Babu, P. N. (2017). Correction of Mars Colour Camera images for identification of spectral classes. *Current Science, 1158-1164*.
- Mentzer, N., Mahr, J., Paya-Vaya, G., & Blume, H. (2019). Online stereo camera calibration for automotive vision based on HW-accelerated A-KAZE-feature extraction. *Journal of Systems Architecture, 97*, 335-348.
- Metternicht, G. (1999). Change detection assessment using fuzzy sets and remotely sensed data: an application of topographic map revision. *ISPRS Journal of Photogrammetry and Remote Sensing, 54*(4), 221-233.
- Minnaert, M. (1941). The reciprocity principle in lunar photometry. *Astrophysical Journal, vol. 93, p. 403-410 (1941).*, 93, 403-410.
- Misra, I., Moorthi, S. M., Dhar, D., & Ramakrishnan, R. (2012a). An automatic satellite image registration technique based on Harris corner detection and Random Sample Consensus (RANSAC) outlier rejection model. In *2012 1st International Conference on Recent Advances in Information Technology (RAIT)* (pp. 68-73). IEEE.

- Misra, I., Gambhir, R. K., Moorthi, S. M., Dhar, D., & Ramakrishnan, R. (2012b). An efficient algorithm for automatic fusion of RISAT-1 SAR data and Resourcesat-2 optical images. In *2012 4th International Conference on Intelligent Human Computer Interaction (IHCI)* (pp. 1-6). IEEE.
- Misra, I., Moorthi, S. M., Sharma, V., Sangar, G., & Dhar, D. (2012c). Full India NDVI map generation using Resourcesat-2 AWiFS data. *Ahmedabad: Space Applications Centre, ISRO*.
- Misra, I., Moorthi, S. M., & Dhar, D. (2015). Minnaert Topographic Correction of Mars Color Camera Images from Mars Orbiter Mission. In *2015 International Conference on Information Technology (ICIT)* (pp. 76-81). IEEE.
- Misra, I., Sharma, V., Moorthi, S. M., & Dhar, D. (2018). An automated approach for reference image generation using multi-tier Resourcesat-2A data over Indian land terrain. In *38th INCA International Congress*.
- Misra, I., Sharma, V., Moorthi, S. M., & Dhar, D. (2019a). An approach for generation of multi temporal co-registered optical remote sensing images from Resourcesat-2/2A sensors. *J. Geomatics*, 13(1), 173-178.
- Misra, I., Moorthi, S. M., & Dhar, D. (2019b). Techniques developed for large area Mars image mosaic using ISRO's Mars Color Camera (MCC) data. *Journal of Geomatics*, 13(1).
- Misra, I., Rohil, M. K., Moorthi, S. M., & Dhar, D. (2021a). FIRM: framework for image registration using multistage feature detection and mode-guided motion smoothness keypoint optimization. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-12.
- Misra, I., Rout, L., Arya, S., Bhateja, Y., Moorthi, S. M., & Dhar, D. (2021b). Phobos image enhancement using unpaired multi-frame acquisitions from Indian Mars Color Camera. *Planetary and Space Science*, 201, 105215.
- Misra, I., Rohil, M. K., Manthira Moorthi, S., & Dhar, D. (2021c). Mars surface multi-decadal change detection using ISRO's Mars color camera (MCC) and viking orbiter images. In *Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Prayagraj, India, December 4-6, 2020, Revised Selected Papers, Part III 5* (pp. 25-33). Springer Singapore.
- Misra, I., Rohil, M. K., Manthira Moorthi, S., & Dhar, D. (2022a). Feature based remote sensing image registration techniques: a comprehensive and comparative review. *International Journal of Remote Sensing*, 43(12), 4477-4516.
- Misra, I., Rohil, M. K., Subbiah, M. M., & Dhar, D. (2022b). EPOCH: enhanced procedure for operational change detection using historical invariant features and PCA guided multivariate statistical technique. *Geocarto International*, 37(25), 9369-9391.
- Misra, I., Rohil, M. K., Moorthi, S. M., & Dhar, D. (2022c). A novel country-level integrated image mosaic system using optical remote sensing imagery. *Earth Science Informatics*, 15(4), 2181-2193.
- Misra, I., Rohil, M. K., Manthira Moorthi, S., & Dhar, D. (2023a). MIRACLE: multi-satellite Island image registration using anisotropic coherence locality enhanced nonlinear diffusion and Mahalanobis distance guided marginalization. *International Journal of Remote Sensing*, 44(12), 3753-3776.
- Misra, I., Rohil, M. K., Moorthi, S. M., & Dhar, D. (2023b). SPRINT: Spectra Preserving Radiance Image Fusion Technique using holistic deep edge spatial attention and Minnaert guided Bayesian probabilistic model. *Signal Processing: Image Communication*, 113, 116920.



- Misra, I., Rohil, M. K., Moorthi, S. M., & Dhar, D. (2023c). CLIM: Co-occurrence with Laplacian Intensity Modulation and Enhanced Color Space Transform for Infrared-Visible Image Fusion. *Infrared Physics & Technology*, 135, 104951.
- Misra, I., Rohil, M. K., Moorthi, S. M., & Dhar, D. (2024a). Direct feature extraction and image co-registration of morphological structure from Bayer pattern raw planetary images. *Expert Systems with Applications*, 238, 122274.
- Misra, I., Rohil, M. K., Moorthi, S., & Dhar, D. (2024b). A novel Venus' visible image processing neoteric workflow for improved planetary surface feature analysis. *Pattern Analysis and Applications*, 27(2), 31.
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on image processing*, 21(12), 4695-4708.
- Moghimi, A., Celik, T., Mohammadzadeh, A., & Kusetogullari, H. (2021b). Comparison of keypoint detectors and descriptors for relative radiometric normalization of bitemporal remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 4063-4073.
- Moghimi, A., Sarmadian, A., Mohammadzadeh, A., Celik, T., Amani, M., & Kusetogullari, H. (2021a). Distortion robust relative radiometric normalization of multitemporal and multisensor remote sensing images using image features. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-20.
- MOM Web Browse Portal (Accessed on 26 July 2022). "<https://mrbrowse.issdc.gov.in/MOMLTA/>"
- Montgomery, H., Che, N., & Bowser, J. (2000, July). Determination of MODIS band-to-band co-registration on-orbit using the SRCA. In *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120)* (Vol. 5, pp. 2203-2205). IEEE.
- Moorthi, S. M., Arya, A. S., Mathew, K., Singh, R. P., Chauhan, P., Sarkar, S. S., ... & Kiran Kumar, A. S. (2015, March). Mars orbiter mission: science data products and archive pipeline. In *46th Annual Lunar and Planetary Science Conference* (No. 1832, p. 1317).
- Moorthi, S. M., Kayal, R., Krishnan, R. R., & Srivastava, P. K. (2008). RESOURCESAT-1 LISS-4 MX bands onground co-registration by in-flight calibration and attitude refinement. *International journal of applied earth observation and geoinformation*, 10(2), 140-146.
- Moorthi, S. M., Oza, M. P., Misra, I., Gambhir, R. K., Darji, N. P., Sharma, S., ... & Parihar, J. S. (2014). FASALSoft—An ISRO software framework for crop production forecast using remote sensing data analysis. *Journal of Geomatics*, 8(1), 28.
- Moreland, K. (2009, November). Diverging color maps for scientific visualization. In *International symposium on visual computing* (pp. 92-103). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Mou, L., Bruzzone, L., & Zhu, X. X. (2018). Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2), 924-935.
- Mumby, P. J., Skirving, W., Strong, A. E., Hardy, J. T., LeDrew, E. F., Hochberg, E. J., ... & David, L. T. (2004). Remote sensing of coral reefs and their physical environment. *Marine pollution bulletin*, 48(3-4), 219-228.

- Munyati, C. (2004). Use of principal component analysis (PCA) of remote sensing images in wetland change detection on the Kafue Flats, Zambia. *Geocarto International*, 19(3), 11-22.
- Muralikrishnan, S., Pillai, A., Narender, B., Reddy, S., Venkataraman, V. R., & Dadhwal, V. K. (2013). Validation of Indian national DEM from Cartosat-1 data. *Journal of the Indian Society of Remote Sensing*, 41, 1-13.
- Mutch, T. A., Arvidson, R. E., Head III, J. W., Jones, K. L., & Saunders, R. S. (1976). The geology of Mars. *Princeton*.
- Nain, J., & Mueller, J. (2019, October). Improving band to band registration accuracy of SEVIRI level 1.5 products. In *Image and Signal Processing for Remote Sensing XXV* (Vol. 11155, pp. 28-35). SPIE.
- Najafi, E., & Ivanov, S. V. (2021, October). Image Registration by Affine-ORB on SAR and UAV Images to Find Object Geolocation. In *2021 International Conference on Engineering and Emerging Technologies (ICEET)* (pp. 1-6). IEEE.
- NASA Visible Earth. 2020 [Online]. Available: <https://visibleearth.nasa.gov/images/147190/explorer-base-map>
- Nevatia, R., & Babu, K. R. (1980). Linear feature extraction and description. *Computer graphics and image processing*, 13(3), 257-269.
- Nielsen, A. A., Conradsen, K., & Simpson, J. J. (1998). Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, 64(1), 1-19.
- NRSC, H. (2011). RESOURCESAT-2 Data User's Handbook. *NRSC: SDAPSA: NDC: DEC11-364*.
- Oliveira, F. P., & Tavares, J. M. R. (2014). Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2), 73-93.
- Ordóñez, Á., Argüello, F., & Heras, D. B. (2018). Alignment of hyperspectral images using KAZE features. *Remote Sensing*, 10(5), 756.
- Ordóñez, Á., Argüello, F., Heras, D. B., & Demir, B. (2020). GPU-accelerated registration of hyperspectral images using KAZE features. *The Journal of Supercomputing*, 76(12), 9478-9492.
- Ordóñez, Á., Heras, D. B., & Argüello, F. (2019, July). SURF-based registration for hyperspectral images. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 63-66). IEEE.
- Oza, M. P., Pandya, M. R., & Rajak, D. R. (2008). Evaluation and use of Resourcesat-I data for agricultural applications. *International Journal of Applied Earth Observation and Geoinformation*, 10(2), 194-205.
- Ozcelik, F., Alganci, U., Sertel, E., & Unal, G. (2020). Rethinking CNN-based pansharpening: Guided colorization of panchromatic images via GANs. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4), 3486-3501.
- Pan, H., Tian, J., Wang, T., Wang, J., Liu, C., & Yang, L. (2022). Band-to-Band Registration of FY-1C/D Visible-IR Scanning Radiometer High-Resolution Picture Transmission Data. *Remote Sensing*, 14(2), 411.
- Pan, J., Zhu, Y., Wang, M., & Zhang, B. X. (2011, August). Parallel band-to-band registration for HJ-1A1B CCD images using openMP. In *2011 International Symposium on Image and Data Fusion* (pp. 1-4). IEEE.

- Pandya, M. R., Murali, K. R., & Kirankumar, A. S. (2013). Quantification and comparison of spectral characteristics of sensors on board Resourcesat-1 and Resourcesat-2 satellites. *Remote sensing letters*, 4(3), 306-314.
- Panigrahy, S., Murthy, T. V. R., Patel, J. G., & Singh, T. S. (2012). Wetlands of India: inventory and assessment at 1: 50,000 scale using geospatial techniques. *Current science*, 852-856.
- Parihar, J. S., & Oza, M. P. (2006, December). FASAL: an integrated approach for crop assessment and production forecasting. In *Agriculture and hydrology applications of remote sensing* (Vol. 6411, p. 641101). SPIE.
- Parker, J. A., Kenyon, R. V., & Troxel, D. E. (1983). Comparison of interpolating methods for image resampling. *IEEE Transactions on medical imaging*, 2(1), 31-39.
- Paul, S., & Pati, U. C. (2016). Remote sensing optical image registration using modified uniform robust SIFT. *IEEE Geoscience and Remote Sensing Letters*, 13(9), 1300-1304.
- Pearlman, A., Montanaro, M., Efremova, B., McCorkel, J., Wenny, B., Lunsford, A., & Reuter, D. (2020). Prelaunch radiometric calibration and uncertainty analysis of landsat thermal infrared sensor 2. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4), 2715-2726.
- Peng, J., Liu, L., Wang, J., Zhang, E., Zhu, X., Zhang, Y., ... & Jiao, L. (2020). PSMD-Net: A novel pan-sharpening method based on a multiscale dense network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(6), 4957-4971.
- Peng, X., Zhong, R., Li, Z., & Li, Q. (2020). Optical remote sensing image change detection based on attention mechanism and image difference. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9), 7296-7307.
- Pettengill, G. H., Ford, P. G., Johnson, W. T., Raney, R. K., & Soderblom, L. A. (1991). Magellan: Radar performance and data products. *Science*, 252(5003), 260-265.
- Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9), 503-510.
- Pizarro, D., & Bartoli, A. (2012), "Feature-based deformable surface detection with self-occlusion reasoning," *International Journal of Computer Vision*, 97, 54-70.
- Plebani, E., Ehlmann, B. L., Leask, E. K., Fox, V. K., & Dundar, M. M. (2022). A machine learning toolkit for CRISM image analysis. *Icarus*, 376, 114849.
- Pohl, C., & Van Genderen, J. L. (1998). Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International journal of remote sensing*, 19(5), 823-854.
- Pourfard, M., Hosseinian, T., Saeidi, R., Motamedi, S. A., Abdollahifard, M. J., Mansoori, R., & Safabakhsh, R. (2021). KAZE-SAR: SAR image registration using KAZE detector and modified SURF descriptor for tackling speckle noise. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-12.
- Prema, G., & Arivazhagan, S. (2022). Infrared and visible image fusion via multi-scale multi-layer rolling guidance filter. *Pattern Analysis and Applications*, 25(4), 933-950.

- Pu, R., Gong, P., Tian, Y., Miao, X., Carruthers, R. I., & Anderson, G. L. (2008). Using classification and NDVI differencing methods for monitoring sparse vegetation coverage: a case study of saltcedar in Nevada, USA. *International Journal of Remote Sensing*, 29(14), 3987-4011.
- Pulighe, G., Baiocchi, V., & Lupia, F. (2016). Horizontal accuracy assessment of very high resolution Google Earth images in the city of Rome, Italy. *International Journal of Digital Earth*, 9(4), 342-362.
- Qiao, L., Wang, M., Jin, Z., & Mao, D. (2021). A comprehensive non-uniformity correction method for EMCCD. *Scientific Reports*, 11(1), 23947.
- Quan, D., Wei, H., Wang, S., Gu, Y., Hou, B., & Jiao, L. (2023). A Novel Coarse-to-Fine Deep Learning Registration Framework for Multi-Modal Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*.
- Radhadevi, P. V., Solanki, S. S., Akilan, A., Jyothi, M. V., & Nagasubramanian, V. (2016). Performance assessment and geometric calibration of resourcesat-2. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, 237-244.
- Radhadevi, P. V., Solanki, S. S., Jyothi, M. V., Nagasubramanian, V., & Varadan, G. (2009). Automated co-registration of images from multiple bands of Liss-4 camera. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1), 17-26.
- Ramachandran, V., & Kishorebabu, V. (2019). A tri-state filter for the removal of salt and pepper noise in mammogram images. *Journal of medical systems*, 43(2), 40.
- Ranchin, T., & Wald, L. (2000). Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogrammetric engineering and remote sensing*, 66(1), 49-61.
- Ranganath, R., & Singh, R. P. (2010). The evolution of the earth observation system in India. *Journal of the Indian Institute of Science*, 90(4), 471-488.
- Redzuwan, R., Radzi, N. A. M., Din, N. M., & Mustafa, I. S. (2015, October). Affine versus projective transformation for SIFT and RANSAC image matching methods. In *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)* (pp. 447-451). IEEE.
- Resourcesat Mission Details (2011) [Online]. Available: <https://www.isro.gov.in/Spacecraft/resourcesat-2>
- Resourcesat-2A Spacecraft. (2021). [accessed 2023 May 16], <https://www.isro.gov.in/Spacecraft/resourcesat-2a>.
- Reuter, D. C., Richardson, C. M., Pellerano, F. A., Irons, J. R., Allen, R. G., Anderson, M., ... & Thome, K. J. (2015). The Thermal Infrared Sensor (TIRS) on Landsat 8: Design overview and pre-launch characterization. *Remote Sensing*, 7(1), 1135-1153.
- Reza, A. M. (2004). Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, 38, 35-44.
- Ridd, M. K., & Liu, J. (1998). A comparison of four algorithms for change detection in an urban environment. *Remote sensing of environment*, 63(2), 95-100.
- Ritter, N., Ruth, M., Grissom, B. B., Galang, G., Haller, J., Stephenson, G., ... & Lundh, F. (2000). Geotiff format specification geotiff revision 1.0. *SPOT Image Corp*, 1.

- Rosten, E., & Drummond, T. (2006). Machine learning for high-speed corner detection. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9* (pp. 430-443). Springer Berlin Heidelberg.
- Rosten, E., Porter, R., & Drummond, T. (2008). Faster and better: A machine learning approach to corner detection. *IEEE transactions on pattern analysis and machine intelligence*, 32(1), 105-119.
- Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., ... & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment*, 145, 154-172.
- Roy, S., Moorthi, S. M., Dhar, D., & Sarkar, S. S. (2014, November). Comparative analysis of demosaicking techniques for Mars Colour Camera data. In *2014 International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 862-867). IEEE.
- Rubel, A., Ieremeiev, O., Lukin, V., Fastowicz, J., & Okarma, K. (2022). Combined No-Reference Image Quality Metrics for Visual Quality Assessment Optimized for Remote Sensing Images. *Applied Sciences*, 12(4), 1986.
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011, November). ORB: An efficient alternative to SIFT or SURF. In *2011 International conference on computer vision* (pp. 2564-2571). Ieee.
- Rusia, P., Bhateja, Y., Misra, I., Moorthi, S. M., & Dhar, D. (2021). An efficient machine learning approach for atmospheric correction. *Journal of the Indian Society of Remote Sensing*, 49, 2539-2548.
- Safy, M., Shi, G., & Amein, A. S. (2013). Semi-automatic image registration using harris corner detection and random sample consensus (RANSAC).
- Salehpour, M., & Behrad, A. (2017). Hierarchical approach for synthetic aperture radar and optical image coregistration using local and global geometric relationship of invariant features. *Journal of Applied Remote Sensing*, 11(1), 015002-015002.
- Salomonson, V. V., Barnes, W. L., Maymon, P. W., Montgomery, H. E., & Ostrow, H. (1989). MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Transactions on Geoscience and Remote Sensing*, 27(2), 145-153.
- Samarasinha, N. H., & Larson, S. M. (2014). Image enhancement techniques for quantitative investigations of morphological features in cometary comae: A comparative study. *Icarus*, 239, 168-185.
- Saunders, R. S., Spear, A. J., Allin, P. C., Austin, R. S., Berman, A. L., Chandler, R. C., ... & Wall, S. D. (1992). Magellan mission summary. *Journal of Geophysical Research: Planets*, 97(E8), 13067-13090.
- Saxena, S., Rabha, A., Tahlani, P., & Ray, S. S. (2021). Crop situation in India, before, during and after COVID-19 lockdown, as seen from the satellite data of resourcesat-2 AWiFS. *Journal of the Indian Society of Remote Sensing*, 49(2), 365-376.
- Schetselaar, E. M. (1998). Fusion by the IHS transform: Should we use cylindrical or spherical coordinates?. *International Journal of Remote Sensing*, 19(4), 759-765.
- Schwind, P., & d'Angelo, P. (2015). Evaluating the applicability of BRISK for the geometric registration of remote sensing images. *Remote Sensing Letters*, 6(9), 677-686.

- Schwind, P., Schneider, M., & Müller, R. (2014). Improving hypslex sensor co-registration accuracy using brisk and sensor-model based ransac. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, 371-376.
- Schwind, P., Suri, S., Reinartz, P., & Siebert, A. (2010). Applicability of the SIFT operator to geometric SAR image registration. *International Journal of Remote Sensing*, 31(8), 1959-1980.
- Seal, A., Garcia-Pedrero, A., Bhattacharjee, D., Nasipuri, M., Lillo-Saavedra, M., Menasalvas, E., & Gonzalo-Martin, C. (2020). Multi-scale RoIs selection for classifying multi-spectral images. *Multidimensional Systems and Signal Processing*, 31, 745-769.
- Sedaghat, A., & Ebadi, H. (2015). Remote sensing image matching based on adaptive binning SIFT descriptor. *IEEE transactions on geoscience and remote sensing*, 53(10), 5283-5293.
- Sedaghat, A., & Mohammadi, N. (2019). High-resolution image registration based on improved SURF detector and localized GTM. *International Journal of Remote Sensing*, 40(7), 2576-2601.
- Seidpisheh, M., & Bamdadi, R. (2023). Fuzzy and non-fuzzy k-quantile clustering for high-variance data. *Pattern Analysis and Applications*, 26(2), 517-528.
- Seong, H., Choi, H., Son, H., & Kim, C. (2018). Image-based 3D building reconstruction using A-KAZE feature extraction algorithm. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 35, pp. 1-5). IAARC Publications.
- Sezer, A., & Altan, A. (2021a). Detection of solder paste defects with an optimization-based deep learning model using image processing techniques. *Soldering & Surface Mount Technology*, 33(5), 291-298.
- Sezer, A., & Altan, A. (2021b). Optimization of deep learning model parameters in classification of solder paste defects. In *2021 3rd international congress on human-computer interaction, optimization and robotic applications (HORA)* (pp. 1-6). IEEE.
- Shafique, A., Cao, G., Khan, Z., Asad, M., & Aslam, M. (2022). Deep learning-based change detection in remote sensing images: A review. *Remote Sensing*, 14(4), 871.
- Shah, V. P., Younan, N. H., & King, R. L. (2008). An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. *IEEE transactions on geoscience and remote sensing*, 46(5), 1323-1335.
- Shahbazi, M., & Cortes, C. (2019). Seamless Co-Registration of Images from Multi-Sensor Multispectral Cameras. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 315-322.
- Shahdoosti, H. R., & Ghassemanian, H. (2014). Fusion of MS and PAN images preserving spectral quality. *IEEE Geoscience and Remote Sensing Letters*, 12(3), 611-615.
- Shang, Y., Qin, J., & Cao, G. (2020). SAR and Optical Remote Sensing Image Registration Based on an Improved Point Feature. In *Big Data Analytics for Cyber-Physical System in Smart City: BDCPS 2019, 28-29 December 2019, Shenyang, China* (pp. 1967-1975). Springer Singapore.
- Shao, Z., & Cai, J. (2018). Remote sensing image fusion with deep convolutional neural network. *IEEE journal of selected topics in applied earth observations and remote sensing*, 11(5), 1656-1669.

- Sharif, S. M. A., Naqvi, R. A., Ali, F., & Biswas, M. (2023). DarkDeblur: Learning single-shot image deblurring in low-light condition. *Expert Systems with Applications*, 222, 119739.
- Shi, Q., Liu, Y., & Xu, Y. (2021). Image feature point matching method based on improved BRISK. *International Journal of Wireless and Mobile Computing*, 20(2), 132-138.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, A. (1989). Review article digital change detection techniques using remotely-sensed data. *International journal of remote sensing*, 10(6), 989-1003.
- Skakun, S., Roger, J. C., Vermote, E. F., Masek, J. G., & Justice, C. O. (2017). Automatic sub-pixel co-registration of Landsat-8 Operational Land Imager and Sentinel-2A Multi-Spectral Instrument images using phase correlation and machine learning based mapping. *International Journal of Digital Earth*, 10(12), 1253-1269.
- Snyder, C. W. (1977). The missions of the Viking orbiters. *Journal of Geophysical Research*, 82(28), 3971-3983.
- Snyder, J. P. (1982). *Map projections used by the US Geological Survey* (No. 1532). US Government Printing Office.
- Song, Z. L., & Zhang, J. (2010). Remote sensing image registration based on retrofitted SURF algorithm and trajectories generated from Lissajous figures. *IEEE Geoscience and Remote Sensing Letters*, 7(3), 491-495.
- Spoto, F., Martimort, P., & Drusch, M. (2012, April). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. In *First Sentinel-2 Preparatory Symposium* (Vol. 707, p. 2).
- Stehman, S. (1996). Estimating the kappa coefficient and its variance under stratified random sampling. *Photogrammetric engineering and remote sensing*, 62(4), 401-407.
- Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 operational land imager on-orbit geometric calibration and performance. *Remote sensing*, 6(11), 11127-11152.
- Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing*, 52(3), 397-399.
- Svedhem, H., Titov, D. V., McCoy, D., Lebreton, J. P., Barabash, S., Bertaux, J. L., ... & Coradini, M. (2007). Venus Express—the first European mission to Venus. *Planetary and Space Science*, 55(12), 1636-1652.
- Tan, B., Woodcock, C. E., Hu, J., Zhang, P., Ozdogan, M., Huang, D., & Myneni, R. B. (2006). The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. *Remote Sensing of Environment*, 105(2), 98-114.
- Tan, X., Yang, J., & Deng, X. (2018, April). Filtering method of star control points for geometric correction of remote sensing image based on RANSAC algorithm. In *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)* (Vol. 10615, pp. 1216-1221). SPIE.
- Tao, Y., Muller, J. P., & Poole, W. (2016). Automated localisation of Mars rovers using co-registered HiRISE-CTX-HRSC orthorectified images and wide baseline Navcam orthorectified mosaics. *Icarus*, 280, 139-157.

- Tarsha-Kurdi, F., Landes, T., & Grussenmeyer, P. (2007, September). Hough-transform and extended ransac algorithms for automatic detection of 3d building roof planes from lidar data. In *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007* (Vol. 36, pp. 407-412).
- Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35-39.
- Teke, M., & Temizel, A. (2010, August). Multi-spectral satellite image registration using scale-restricted SURF. In *2010 20th International Conference on Pattern Recognition* (pp. 2310-2313). IEEE.
- Tengfeng, W. (2018, May). Seamless stitching of panoramic image based on multiple homography matrix. In *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* (pp. 2403-2407). IEEE.
- Thomas, C., Ranchin, T., Wald, L., & Chanussot, J. (2008). Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5), 1301-1312.
- Tian, F., & Shi, P. (2014, October). Image Mosaic using ORB descriptor and improved blending algorithm. In *2014 7th International Congress on Image and Signal Processing* (pp. 693-698). IEEE.
- Tilton, J. C., Lin, G., & Tan, B. (2016). Measurement of the band-to-band registration of the SNPP VIIRS imaging system from on-orbit data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3), 1056-1067.
- Tilton, J. C., Wolfe, R. E., Lin, G., & Dellomo, J. J. (2019). On-orbit measurement of the effective focal length and Band-to-Band registration of satellite-borne whiskbroom imaging sensors. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(11), 4622-4633.
- Tiwari, M. M., Misra, I., Moorthi, S. M., & Dhar, D. (2021, December). An Improved IHS Image Fusion Algorithm using Medoid Intensity Match and Bilateral Filter. In *2021 IEEE International India Geoscience and Remote Sensing Symposium (InGARSS)* (pp. 500-503). IEEE.
- Toet, A. (2017). The TNO multiband image data collection. *Data in brief*, 15, 249-251.
- Torr, P. H., & Zisserman, A. (2000). MLESAC: A new robust estimator with application to estimating image geometry. *Computer vision and image understanding*, 78(1), 138-156.
- Torre, V., & Poggio, T. A. (1986). On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), 147-163.
- Tu, T. M., Su, S. C., Shyu, H. C., & Huang, P. S. (2001). A new look at IHS-like image fusion methods. *Information fusion*, 2(3), 177-186.
- Tuomisto, H. (2016). Influence of compositing criterion and data availability on pixel-based Landsat TM/ETM+ image compositing over Amazonian forests. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3), 857-867.



- Vázquez-Jiménez, R., Romero-Calcerrada, R., Novillo, C. J., Ramos-Bernal, R. N., & Arrogante-Funes, P. (2017). Applying the chi-square transformation and automatic secant thresholding to Landsat imagery as unsupervised change detection methods. *Journal of Applied Remote Sensing*, *11*(1), 016016-016016.
- Vermote, E. F., Tanré, D., Deuze, J. L., Herman, M., & Morcette, J. J. (1997). Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE transactions on geoscience and remote sensing*, *35*(3), 675-686.
- Vermote, E. F., Tanré, D., Deuze, J. L., Herman, M., & Morcette, J. J. (1997). Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE transactions on geoscience and remote sensing*, *35*(3), 675-686.
- Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, *185*, 46-56.
- Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G. A., ... & Wald, L. (2014). A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(5), 2565-2586.
- Vivone, G., Dalla Mura, M., Garzelli, A., Restaino, R., Scarpa, G., Ulfarsson, M. O., ... & Chanussot, J. (2020). A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods. *IEEE Geoscience and Remote Sensing Magazine*, *9*(1), 53-81.
- Vivone, G., Restaino, R., & Chanussot, J. (2017). A regression-based high-pass modulation pansharpening approach. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(2), 984-996.
- Wald, L. (2000, January). Quality of high resolution synthesised images: Is there a simple criterion?. In *Third conference " Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"* (pp. 99-103). SEE/URISCA.
- Wang, G., Zhai, Z., Xu, B., & Cheng, Y. (2017, May). A parallel method for aerial image stitching using ORB feature points. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)* (pp. 769-773). IEEE.
- Wang, M., Yang, B., Hu, F., & Zang, X. (2014). On-orbit geometric calibration model and its applications for high-resolution optical satellite imagery. *Remote sensing*, *6*(5), 4391-4408.
- Wang, Q., Zhang, X., Chen, G., Dai, F., Gong, Y., & Zhu, K. (2018). Change detection based on Faster R-CNN for high-resolution remote sensing images. *Remote sensing letters*, *9*(10), 923-932.
- Wang, S., Quan, D., Liang, X., Ning, M., Guo, Y., & Jiao, L. (2018). A deep learning framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 148-164.
- Wang, T., Fang, F., Li, F., & Zhang, G. (2018). High-quality Bayesian pansharpening. *IEEE Transactions on Image Processing*, *28*(1), 227-239.
- Wang, Y., Lai, H., Ma, H., Jia, Z., & Wang, L. (2020). A novel harris feature detection-based registration for remote sensing image. *Journal of the Indian Society of Remote Sensing*, *48*, 1245-1252.
- Wang, Z., & Bovik, A. C. (2002). A universal image quality index. *IEEE signal processing letters*, *9*(3), 81-84.

- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600-612.
- Wang, Z., Xiong, X., & Li, Y. (2015). Improved band-to-band registration characterization for VIIRS reflective solar bands based on lunar observations. *Remote Sensing*, 8(1), 27.
- Wei, Y., Yuan, Q., Shen, H., & Zhang, L. (2017). Boosting the accuracy of multispectral image pansharpening by learning a deep residual network. *IEEE Geoscience and Remote Sensing Letters*, 14(10), 1795-1799.
- Weickert, J. (1999). Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17(3-4), 201-212.
- Welstead, S. T. (1999). *Fractal and wavelet image compression techniques* (Vol. 40). Spie Press.
- Wood, B. E., Hess, P., Lustig-Yaeger, J., Gallagher, B., Korwan, D., Rich, N., ... & Raouafi, N. E. (2022). Parker solar probe imaging of the night side of Venus. *Geophysical Research Letters*, 49(3), e2021GL096302.
- Wu, C., Du, B., & Zhang, L. (2013). Slow feature analysis for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5), 2858-2874.
- Wu, S., Zeng, W., & Chen, H. (2020). A sub-pixel image registration algorithm based on SURF and M-estimator sample consensus. *Pattern Recognition Letters*, 140, 261-266.
- Wu, Y., Ma, W., Gong, M., Su, L., & Jiao, L. (2014). A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geoscience and Remote Sensing Letters*, 12(1), 43-47.
- Wu, Z., Huang, Y., & Zhang, K. (2018). Remote sensing image fusion method based on PCA and curvelet transform. *Journal of the Indian Society of Remote Sensing*, 46, 687-695.
- Xiang, Y., Wang, F., & You, H. (2018). OS-SIFT: A robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6), 3078-3090.
- Xiao, J., Suab, S. A., Chen, X., Singh, C. K., Singh, D., Aggarwal, A. K., ... & Avtar, R. (2023). Enhancing assessment of corn growth performance using unmanned aerial vehicles (UAVs) and deep learning. *Measurement*, 214, 112764.
- Xie, S., & Tu, Z. (2015). Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 1395-1403).
- Xie, T., Dai, K., Wang, K., Li, R., & Zhao, L. (2024). Deepmatcher: a deep transformer-based network for robust and accurate local feature matching. *Expert Systems with Applications*, 237, 121361.
- Xie, Y., Xiong, X., Qu, J. J., Che, N., & Summers, M. E. (2011). Impact analysis of MODIS band-to-band registration on its measurements and science data products. *International Journal of Remote Sensing*, 32(16), 4431-4444.
- Xie, Y., Xiong, X., Qu, J. J., Che, N., & Wang, L. (2007, October). Sensitivity analysis of MODIS band-to-band registration characterization and its impact on the science data products. In *Remote Sensing and Modeling of Ecosystems for Sustainability IV* (Vol. 6679, pp. 64-72). SPIE.

- Xin, X., Liu, B., Di, K., Jia, M., & Oberst, J. (2018). High-precision co-registration of orbiter imagery and digital elevation model constrained by both geometric and photometric information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *144*, 28-37.
- Xing, Y., Wang, M., Yang, S., & Jiao, L. (2018). Pan-sharpening via deep metric learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 165-183.
- Xiong, X., Sun, J., Angal, A., Xie, Y., Choi, T., & Wang, Z. (2011, September). Results of MODIS band-to-band registration characterization using on-orbit lunar observations. In *Earth Observing Systems XVI* (Vol. 8153, pp. 583-591). SPIE.
- Xu, C., Sui, H., Li, H., & Liu, J. (2015). An automatic optical and SAR image registration method with iterative level set segmentation and SIFT. *International Journal of Remote Sensing*, *36*(15), 3997-4017.
- Xue, W., Zhang, L., Mou, X., & Bovik, A. C. (2013). Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE transactions on image processing*, *23*(2), 684-695.
- Yan, L., Roy, D. P., Li, Z., Zhang, H. K., & Huang, H. (2018). Sentinel-2A multi-temporal misregistration characterization and an orbit-based sub-pixel registration methodology. *Remote sensing of environment*, *215*, 495-506.
- Yan, L., Roy, D. P., Zhang, H., Li, J., & Huang, H. (2016). An automated approach for sub-pixel registration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery. *Remote Sensing*, *8*(6), 520.
- Yan, W., She, H., & Yuan, Z. (2014). Robust registration of remote sensing image based on SURF and KCCA. *Journal of the Indian Society of Remote Sensing*, *42*, 291-299.
- Yang, J., & Li, J. (2018, February). An improved feature extraction algorithm based on KAZE for multi-spectral image. In *MIPPR 2017: Automatic Target Recognition and Navigation* (Vol. 10608, pp. 175-180). SPIE.
- Yang, J., Fu, X., Hu, Y., Huang, Y., Ding, X., & Paisley, J. (2017). PanNet: A deep network architecture for pan-sharpening. In *Proceedings of the IEEE international conference on computer vision* (pp. 5449-5457).
- Yang, K., Fleig, A. J., Wolfe, R. E., & Nishihama, M. (2000, July). MODIS band-to-band registration. In *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120)* (Vol. 2, pp. 887-889). IEEE.
- Yang, K., Pan, A., Yang, Y., Zhang, S., Ong, S. H., & Tang, H. (2017). Remote sensing image registration using multiple image features. *Remote Sensing*, *9*(6), 581.
- Yang, M., Jiao, L., Liu, F., Hou, B., & Yang, S. (2019). Transferred deep learning-based change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(9), 6960-6973.
- Yang, Y., Wan, W., Huang, S., Yuan, F., Yang, S., & Que, Y. (2016). Remote sensing image fusion based on adaptive IHS and multiscale guided filter. *IEEE Access*, *4*, 4573-4582.

- Yang, Y., Wu, L., Huang, S., Tang, Y., & Wan, W. (2018). Pansharpening for multiband images with adaptive spectral-intensity modulation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(9), 3196-3208.
- Yang, Z., Kang, Z., & Yang, J. (2020). A semiautomatic registration method for Chang'E-1 IIM imagery based on globally geo-reference LROC-WAC mosaic imagery. *IEEE Geoscience and Remote Sensing Letters*, 18(3), 543-547.
- Yao, G., Cui, J., Deng, K., & Zhang, L. (2017). Robust Harris corner matching based on the quasi-homography transform and self-adaptive window for wide-baseline stereo images. *IEEE Transactions on Geoscience and Remote Sensing*, 56(1), 559-574.
- Yao, Y., Zhang, Y., Wan, Y., Liu, X., Yan, X., & Li, J. (2022). Multi-modal remote sensing image matching considering co-occurrence filter. *IEEE Transactions on Image Processing*, 31, 2584-2597.
- Ye, F., Guo, Y., & Zhuang, P. (2019). Pan-sharpening via a gradient-based deep network prior. *Signal Processing: Image Communication*, 74, 322-331.
- Ye, Y., & Shan, J. (2014). A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences. *ISPRS Journal of Photogrammetry and Remote Sensing*, 90, 83-95.
- Ye, Y., Tang, T., Zhu, B., Yang, C., Li, B., & Hao, S. (2022). A multiscale framework with unsupervised learning for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-15.
- Ye, Y., Zhu, B., Tang, T., Yang, C., Xu, Q., & Zhang, G. (2022a). A robust multimodal remote sensing image registration method and system using steerable filters with first-and second-order gradients. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188, 331-350.
- Ye, Y., Yang, C., Gong, G., Yang, P., Quan, D., & Li, J. (2024). Robust optical and SAR image matching using attention-enhanced structural features. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1-12.
- Yilmaz, C. S., Yilmaz, V., Gungor, O., & Shan, J. (2019). Metaheuristic pansharpening based on symbiotic organisms search optimization. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158, 167-187.
- Yilmaz, V. (2021). A Non-Dominated Sorting Genetic Algorithm-II-based approach to optimize the spectral and spatial quality of component substitution-based pansharpened images. *Concurrency and Computation: Practice and Experience*, 33(5), e6030.
- Yokoya, N., Yairi, T., & Iwasaki, A. (2011). Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2), 528-537.
- Yu, M., Yang, H., Deng, K., & Yuan, K. (2018). Registrating oblique images by integrating affine and scale-invariant features. *International Journal of Remote Sensing*, 39(10), 3386-3405.
- Yu, Q., Ni, D., Jiang, Y., Yan, Y., An, J., & Sun, T. (2021). Universal SAR and optical image registration via a novel SIFT framework based on nonlinear diffusion and a polar spatial-frequency descriptor. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171, 1-17.

- Yu, Q., Ni, D., Jiang, Y., Yan, Y., An, J., & Sun, T. (2021). Universal SAR and optical image registration via a novel SIFT framework based on nonlinear diffusion and a polar spatial-frequency descriptor. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171, 1-17.
- Yuan, D., & Elvidge, C. D. (1996). Comparison of relative radiometric normalization techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(3), 117-126.
- Yuan, J., & He, J. (2021). Blocking sparse method for image denoising. *Pattern Analysis and Applications*, 24(3), 1125-1133.
- Yuan, Q., Wei, Y., Meng, X., Shen, H., & Zhang, L. (2018). A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(3), 978-989.
- Yuhas, R. H., Goetz, A. F., & Boardman, J. W. (1992, June). Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*.
- Yusuf, Y., Sri Sumantyo, J. T., & Kuze, H. (2013). Spectral information analysis of image fusion data for remote sensing applications. *Geocarto international*, 28(4), 291-310.
- Zhang, B., & Allebach, J. P. (2008). Adaptive bilateral filter for sharpness enhancement and noise removal. *IEEE transactions on Image Processing*, 17(5), 664-678.
- Zhang, H. K., & Huang, B. (2015). A new look at image fusion methods from a Bayesian perspective. *Remote sensing*, 7(6), 6828-6861.
- Zhang, M., & Shi, W. (2020). A feature difference convolutional neural network-based change detection method. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10), 7232-7246.
- Zhang, S., Li, S., Zhang, B., & Peng, M. (2020). Integration of optimal spatial distributed tie-points in RANSAC-based image registration. *European Journal of Remote Sensing*, 53(1), 67-80.
- Zhang, T., Zhao, R., & Chen, Z. (2020). Application of migration image registration algorithm based on improved SURF in remote sensing image mosaic. *IEEE Access*, 8, 163637-163645.
- Zhang, W. (2020). Combination of SIFT and Canny edge detection for registration between SAR and optical images. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
- Zhang, X., Wang, Y., & Liu, H. (2021). Robust optical and SAR image registration based on OS-SIFT and cascaded sample consensus. *IEEE Geoscience and Remote Sensing Letters*, 19, 1-5.
- Zhang, Y. (2008). Methods for image fusion quality assessment-a review, comparison and analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(PART B7), 1101-1109.
- Zhang, Y., & Hong, G. (2005). An IHS and wavelet integrated approach to improve pan-sharpening visual quality of natural colour IKONOS and QuickBird images. *Information fusion*, 6(3), 225-234.

- Zhang, Y., & Li, X. (2019, May). Bilateral Grid Statistics Combined with BRISK for Robust Matching. In *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)* (pp. 263-269). IEEE.
- Zhang, Y., De Backer, S., & Scheunders, P. (2009). Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(11), 3834-3843.
- Zhang, Y., Liu, C., Sun, M., & Ou, Y. (2019). Pan-sharpening using an efficient bidirectional pyramid network. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(8), 5549-5563.
- Zhang, Y., Zhao, P., Ma, Y., & Fan, X. (2021). Multi-focus image fusion with joint guided image filtering. *Signal Processing: Image Communication*, *92*, 116128.
- Zhao, H., Zhou, P., Zhang, Y., Wang, Z., & Shi, S. (2013). Development of a dual-path system for band-to-band registration of an acousto-optic tunable filter-based imaging spectrometer. *Optics Letters*, *38*(20), 4120-4123.
- Zheng, J., Peng, W., Wang, Y., & Zhai, B. (2021). Accelerated RANSAC for accurate image registration in aerial video surveillance. *IEEE Access*, *9*, 36775-36790.
- Zheng, Q., & Chellappa, R. (1993). A computational vision approach to image registration. *IEEE Transactions on Image Processing*, *2*(3), 311-326.
- Zheng, Z., Ma, Y., Zheng, H., Ju, J., & Lin, M. (2018). UGC: Real-time, ultra-robust feature correspondence via unilateral grid-based clustering. *IEEE Access*, *6*, 55501-55508.
- Zhong, S. H., Liu, Y., & Chen, Q. C. (2015). Visual orientation inhomogeneity based scale-invariant feature transform. *Expert Systems with Applications*, *42*(13), 5658-5667.
- Zhou, J., Kwan, C., & Budavari, B. (2016). Hyperspectral image super-resolution: A hybrid color mapping approach. *Journal of Applied Remote Sensing*, *10*(3), 035024-035024.
- Zhou, W., Zhang, L., Gao, S., & Lou, X. (2021). Gradient-based feature extraction from raw bayer pattern images. *IEEE Transactions on Image Processing*, *30*, 5122-5137.
- Zhu, J. T., Gong, C. F., Zhao, M. X., Wang, L., & Luo, Y. (2020). Image mosaic algorithm based on PCA-ORB feature matching. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 83-89.
- Zhu, Y., Wang, M., Zhu, Q., & Pan, J. (2014). Detection and compensation of Band-to-band registration error for multi-spectral imagery caused by satellite jitter. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*, 69-76.
- Zimbelman, J. R., & Leshin, L. A. (1987). A geologic evaluation of thermal properties for the Elysium and Aeolis Quadrangles of Mars. *Journal of Geophysical Research: Solid Earth*, *92*(B4), E588-E596.
- Zitova, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and vision computing*, *21*(11), 977-1000.
- Zong, J., Smyth, M. M., & Jovanovic, V. M. (1996, November). MISR band-to-band registration. In *Multispectral Imaging for Terrestrial Applications* (Vol. 2818, pp. 138-146). SPIE.

# LIST OF PUBLICATIONS

## A. Published Peer Review Journal Publications

1. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "A novel Venus' visible image processing neoteric workflow for improved planetary surface feature analysis." *Pattern Analysis and Applications* 27, no. 2 (2024): 31.
2. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "Direct feature extraction and image co-registration of morphological structure from Bayer pattern raw planetary images." *Expert Systems with Applications* 238 (2024): 122274.
3. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "CLIM: Co-occurrence with Laplacian Intensity Modulation and Enhanced Color Space Transform for Infrared-Visible Image Fusion." *Infrared Physics & Technology* 135 (2023): 104951.
4. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "MIRACLE: multi-satellite Island image registration using anisotropic coherence locality enhanced nonlinear diffusion and Mahalanobis distance guided marginalization." *International Journal of Remote Sensing* 44, no. 12 (2023): 3753-3776.
5. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "SPRINT: Spectra Preserving Radiance Image Fusion Technique using holistic deep edge spatial attention and Minnaert guided Bayesian probabilistic model." *Signal Processing: Image Communication* 113 (2023): 116920.
6. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "A novel country-level integrated image mosaic system using optical remote sensing imagery." *Earth Science Informatics* 15, no. 4 (2022): 2181-2193.
7. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "Feature based remote sensing image registration techniques: A comprehensive and comparative review." *International Journal of Remote Sensing* 43, no. 12 (2022): 4477-4516.

8. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "EPOCH: enhanced procedure for operational change detection using historical invariant features and PCA guided multivariate statistical technique." *Geocarto International* 37, no. 25 (2022): 9369-9391.
9. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "FIRM: framework for image registration using multistage feature detection and mode-guided motion smoothness keypoint optimization." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1-12.

## **B. Published Peer Review Conference Publications**

1. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. Himalayan Glacier and the Possibility of Past Ice Streams on Martian Outflow Channels using Remote Sensing Imagery. In *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, pp. 6160-6163. IEEE, 2024.*
2. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. Monitoring Riparian Vegetation Change Dynamics using Kullback-Leibler Divergence: A Case Study in the Vicinity of Sabarmati River, Gandhinagar, India. In *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, pp. 4731-4734. IEEE, 2024.*
3. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. "Mars surface multi-decadal change detection using ISRO's Mars color camera (MCC) and viking orbiter images." In *Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Prayagraj, India, December 4-6, 2020, Revised Selected Papers, Part III 5*, pp. 25-33. Springer Singapore, 2021.

## **C. Manuscript Communicated**

1. Misra, Indranil, Mukesh Kumar Rohil, S. Manthira Moorthi, and Debajyoti Dhar. Enhanced Multispectral Band-to-Band Registration using Co-occurrence Scale Space and Spatial Confined RANSAC Guided Segmented Affine Transformation. Communicated to *IEEE Transactions on Image Processing*.



## BRIEF BIOGRAPHY OF THE CANDIDATE



**Indranil Misra** received Bachelor of Engineering (Information Technology; Gold Medalist) from Madhav Institute of Technology and Science (MITS), Gwalior, India and Master of Science (Software System) from Birla Institute of Technology and Science (BITS), Pilani, India in 2007 and 2013 respectively. Currently pursuing PhD under Aspirant Scheme in Department of Computer Science and Information Systems, BITS, Pilani, Pilani Campus.

Presently working as Scientist/Engineer ‘SF’ at Space Applications Centre (SAC), Indian Space Research Organisation (ISRO), Ahmedabad. Worked for more than 15 years in diverse range of remote sensing missions and generate enhanced data products using innovative processing techniques. Responsible for algorithm design and software solutions development for satellite image processing and computer vision across multiple remote sensing projects. Contributed mainly in remote sensing data processing techniques development for geometric calibration, multi-temporal image registration, multi-sensor image fusion, change detection and topographic correction models for improving the radiometry of remote sensing images.

Mr. Misra received ISRO Young Scientist Award in 2015 for contribution in Indian Space technology. He also received the Prof. Kakani N Rao Young Achiever Award from Indian Society of Geomatics (ISG) in 2022. Got Best Paper Award in national symposium organized by the Indian Society of Remote Sensing (ISRS) and Indian Society of Geomatics (ISG) in year 2018. Attended Space Studies Program 2023 organised by International Space University as an ISRO Participant at Brazil. He has published more than 44 articles in peer reviewed journals and national/international conferences. He is a member of IEEE and life member of ISRS and ISG professional societies.

## BRIEF BIOGRAPHY OF THE SUPERVISOR



**Mukesh Kumar Rohil** received Bachelor of Engineering (Honours) in Civil Engineering and Master of Science (Honours) in Physics under Dual Degree Program in 1993, Master of Engineering in Systems & Information in 1995, and Ph.D. in Computer Science and Engineering in 2004, all from the Birla Institute of Technology and Science, Pilani (BITS Pilani), India.

He is working as Professor with the Department of Computer Science and Information Systems, BITS Pilani, India. Previously, he has worked there as Assistant Lecturer from June 1995 to December 1998, as Lecturer from December 1999 to May 2006, as Assistant Professor from June 2006 to January 2013, and as Associate Professor from February 2013 to February 2023. His areas of research interest include Computer Vision, Pattern Recognition, Artificial Intelligence, Geographic Information Systems, Remote Sensing, Machine Learning, Multimedia, Extended Reality, Human-Computer Interaction and Natural Language Processing.

Dr. Rohil is a Senior Member of IEEE, a Life Member of the Indian Society of Remote Sensing (ISRS), and a Life Member of the Indian Society for Technical Education (ISTE).



**Dr. S Manthira Moorthi** holds a PhD in Computational Physics and joined Signal and Image Processing Area (SIPA), Space Applications Centre (SAC), Indian Space Research Organisation (ISRO), Ahmedabad in 1991 as a scientist.

Presently he is working as a Scientist/Engineer 'G', Group Director, Optical Data Processing Group, SIPA, SAC, ISRO, Ahmedabad. He has been contributing to Remote Sensing Data processing domain in the development of methods, algorithms and software systems for image pre and post processing systems for Earth and planetary Observations. His areas of interests are developing models for satellite image rectification, registration, restoration, classification, super resolution, data fusion, AI/ML/DL based image processing solutions and space science data management.

Dr. Moorthi was involved in Chandrayaan-1, Mars Orbiter Mission, & Chandrayaan-3 ISRO Missions data processing aspects and received ISRO team excellence awards for Chandrayaan-1 and Mars Orbiter Mission being project team member and mission operations.