FEATURE POINT BASED IMAGE REGISTRATION BETWEEN SATELLITE IMAGERY AND AERIAL IMAGES OF AGRICULTURAL LAND

by

Mohsin Abbas

Supervised By

Dr. Sajid Saleem

Co-Supervised By

Dr. Moeenuddin Tariq

Submitted for partial fulfillment of the requirements of the degree of MSCS to the Faculty of Engineering and Computer Science

NATIONAL UNIVERSITY OF MODERN LANGUAGES, ISLAMABAD

JANUARY 2019

FEATURE POINT BASED IMAGE REGISTRATION BETWEEN SATELLITE IMAGERY AND AERIAL IMAGES OF AGRICULTURAL LAND

by

Mohsin Abbas

Supervised By

Dr. Sajid Saleem

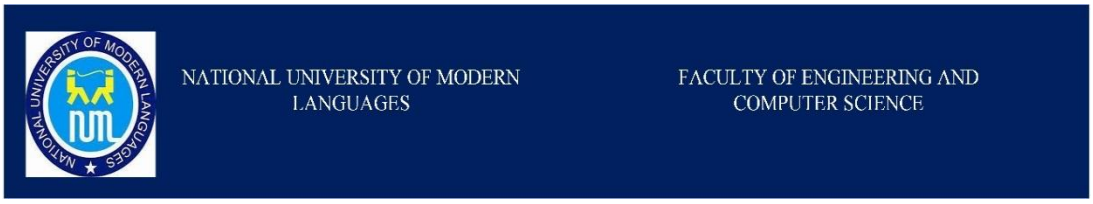
Co-Supervised By

Dr. Moeenuddin Tariq

Submitted for partial fulfillment of the requirements of the degree of MSCS to the Faculty of Engineering and Computer Science

NATIONAL UNIVERSITY OF MODERN LANGUAGES, ISLAMABAD

JANUARY 2019



THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with overall exam performance, and recommend the thesis to the Faculty of Engineering and Computer Sciences.

THESIS TITLE: FEATURE POINT BASED IMAGE REGISTRATION BETWEEN SATELLITE IMAGERY AND AERIAL IMAGES OF AGRICULTURAL LAND

Submitted By: Mohsin Abbas Registration #: 16/MS/CS/F16

Master of Science
Degree Name in Full

Computer Science
Name of Discipline

Dr. Sajid Saleem
Name of Research Supervisor
Signature: _____

Dr. Moeenuddin Tariq
Name of Co-Supervisor
Signature: _____

Dr. Muhammad Akbar
Name of Dean (FE&CS)
Signature: _____

Brig. Muhammad Ibrahim
Name of Director General (NUML)
Signature: _____

31st January, 2019
(Date)

CANDIDATE DECLARATION

I declare that this thesis entitled "*Feature Point Based Image Registration between Satellite Imagery and Aerial Images of Agricultural Land*" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : _____

Name : Mohsin Abbas

Date : January 31st, 2019

ABSTRACT

Image Registration is a process of geometrically aligning two images (reference and target) of the same scene, taken from different viewpoints, at different times or by different sensors. Image registration is used in a wide range of remote sensing applications. The rapid advancement in remote sensing sensors has drastically increase the use of Satellite Imagery (SI) and Unmanned Aerial Vehicle (UAV) images in different applications such as traffic monitoring, agriculture land analysis, early warning systems and damage assessment.

This thesis focuses on an agricultural application of SI and UAV images. The SI are low resolution images as they are captured from very high altitude, whereas, the UAV images are taken from low flying platform, have high resolution and relatively good quality. But the UAV images lack geo-referencing and cannot be used directly in remote sensing applications. This problem in literature is dealt with feature point based geo-registration between SI-UAV images. In case of agricultural SI-UAV images, the registration process is a challenging task. This is due to temporal nature of agricultural crops, which results in high textural and intensity differences between the SI-UAV images. Existing feature points such as Scale Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB), are not invariant to temporal, textural and intensity differences and under-perform in the image registration task. This thesis proposes a new method that combines the strength of Nearest Neighbor (NN) and Brute Force (BF) descriptor matching strategies to register SI-UAV images. The proposed method is named as NN-BF method.

In the proposed NN-BF method, the corresponding feature point matches are first identified between SI-UAV images of the training set with overlap error. Then the corresponding feature point matches are used with NN and BF based descriptor matching strategies to register the SI-UAV images of the test set. Experiments are performed on SI-UAV image dataset of agricultural land. The experimental results show that NN-BF improves the matching and precision scores of SIFT by 20.4% and 32% respectively. Whereas in case of SURF, the NN-BF method improves the matching and precision scores by 19.5% and 21.8% respectively.

Keywords: Agriculture land, Feature point detectors, Feature point descriptors, Image registration, Satellite imagery, UAV images.

DEDICATION

This thesis work is dedicated to my parents and my teachers throughout my education career who have not only loved me unconditionally but whose good examples have taught me to work hard for the things that I aspire to achieve.

ACKNOWLEGEMENT

First of all, I wish to express my gratitude and deep appreciation to Almighty Allah, who made this study possible and successful. This study would not be accomplished unless the honest espousal that was extended from several sources for which I would like to express my sincere thankfulness and gratitude. Yet, there were significant contributors for my attained success and I cannot forget their input, especially my research supervisors, Dr. Sajid Saleem and Dr. Moeenuddin Tariq, and my teacher Dr. Fazli Subhan, who did not leave any stone unturned to guide me during my research journey.

I would like to say a special thanks to my parents and family for their support throughout my educational career, specially my father Mr. Qamar Abbas Awan. I would like to thank my colleagues Miss. Sara Azeem and Mrs. Sarah Mazhar for their utmost moral support throughout my degree tenure. Specially, I would like to thank my colleagues Mr. Muhammad Ali Jamro for providing vital advice about my research area and also Miss. Falak Tahir for her guidance throughout this research. I would like to thank all my colleagues studied during this MS-CS degree.

I shall also acknowledge the extended assistance from the administrations of Department of Computer Sciences who supported me all through my research experience and simplified the challenges I faced. For all whom I did not mention but I shall not neglect their significant contribution, thanks for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BFDM	Brute Force Descriptor Matcher
BRIEF	Binary Robust Independent Elementary Features
BRISK	Binary Robust Invariant Scalable Keypoints
CP	Control Point
DOG	Difference of Gaussians
FAST	Features from Accelerated Segment Test
LHOPC	Local Histogram of Orientated Phase Congruency
MI	Mutual Information
MMI	Maximization of Mutual Information
NN	Nearest Neighbor
OE	Overlap Error
ORB	Oriented FAST and Rotated BRIEF
PE	Projection Error
RANSAC	Random Sample Consensus
RIFT	Radiation Invariant Feature Transform
SAR	Synthetic Aperture Radar
SI	Satellite Imagery
SIFT	Scale Invariant Feature Transform
SR	Scale Restriction
SURF	Speeded-Up Robust Features
UAV	Unmanned Aerial Vehicle

CHAPTER 1

INTRODUCTION

1.1 Overview

Image Registration is a process in which two images (reference and target) are combined together in order to obtain an optimized alignment between the images [1, 2]. The alignment is such that the reference image and the resulting image becomes geometrically aligned [3]. The images used in the registration might be taken from different viewpoints or/and at different times or/and by different sensors. In image processing, image registration is considered as an essential task for wide range of applications related to computer vision, remote sensing and medical image analysis [1, 2].

In medical image analysis, the image registration is used for disease confinement and also to detect tumor by registering different images of Computed Tomography, Photon Emission Tomography and Magnetic Resonance [1, 4]. In computer vision applications, it is used in object recognition, motion tracking, signature verification and stereo-mapping tasks [2, 5, 6, 7]. In case of remote sensing, it is used in environment monitoring, map updating, image mosaic, urban growth monitoring, change detection, and location identification [1, 2, 8, 9].

Recently, remote sensing sensors have experienced a quick advancement in data quantity, quality and attributes [10]. The earth observation satellites now observe the earth surface regularly. These sensors provide data that covers different portions and wavelengths of electromagnetic spectrum at different spatial and temporal resolutions [11]. With this advancement, an enormous increase in quality and accessibility of information has occurred, and has resulted in enormous increase in the use of Satellite Imagery (SI) and images taken from Unmanned Aerial Vehicle (UAV) in different applications of remote sensing [12, 13]. For instance, traffic monitoring [10], change detection [14], agricultural land analysis [15, 16, 17], urban

growth [18], early warning systems and damage assessment [19]. In particular, during natural disaster or any catastrophic changes, one of the major sources of information is remote sensing images [20]. Generally, images from different sensors are utilized in different applications of remote sensing depending upon the application need or data unavailability.

1.2 Image Registration Types

In general, the image registration can be divided into three categories, which are, multi-modal registration, template registration and temporal registration [1].

1.2.1 Multi-modal Registration

This registration category deals with registering same scene images obtained with different sensors such as microwave, infrared and visible spectrum. Viewpoint registration is also incorporated with multi-modal registration for images of remote sensing, as generally during scene acquisition remote sensing sensors of various modalities also possess different viewpoints. Furthermore, various sensor acquisition standards and geometries (such as optical and high resolution Synthetic Aperture Radar (SAR)) can also make this kind of registration complicated [21, 22].

1.2.2 Template Registration

This type of registration is used to locate a reference patch within an image. Generally, this type of registration is used for well-defined scenes such as airports, lakes, indoor and outdoor images etc. [23, 24]

1.2.3 Temporal Registration

Temporal registration is used to register same scene images captured at different times or under various changing conditions. There are immense applications of this type of registration especially in change detection, natural resource monitoring, urban growth monitoring, crop monitoring and precision agriculture [14, 25, 26].

1.3 Image Registration Techniques

The techniques used for image registration can be categorized into two major categories, which are area-based and feature-based techniques [2, 9].

1.3.1 Area-based Techniques

Area-based techniques are also known as intensity-based techniques. In these techniques, the registration of images is carried out largely on the basis of correlation

between pixel intensity values of two images because distinguishing data is provided through pixel intensities. This enables the procedure of registration to emphasize on matching of intensity map of images. The area-based techniques are susceptible to change in intensity values, noise distortions, illumination fluctuation, and change in sensors [18, 27]. The area/intensity-based registration technique carried out in five steps [1], which are, (i) Similarity metric, (ii) Search space and strategy, (iii) Transformation model estimation, (iv) Resampling and transformation of Image, (v) Registration quality assessment

1.3.2 Feature-based Techniques

These techniques are based on extraction of salient objects or features that represent landmarks, shapes, sharp edges, corners, line intersections or road junctions, that are invariant to image variations such as scale, rotation and affine. Significant regions (fields, lakes, forests), lines (coastlines, rivers, region boundaries, roads) or points (region corners, points on high curvature curves, line intersections) are considered as features. These features describe higher level of information and are appropriate for cases when differences in illumination are assumed or required a multi-sensor analysis. A transformation matrix based on spatial information of the features is obtained by matching the features of reference and target images and for image registration [3, 15, 26, 28, 29]. Feature-based registration techniques consist of four steps, (i) Feature detection, (ii) Feature matching, (iii) Transformation model and parameter estimation, (iv) Image warping

Both area/intensity and feature based technique are widely used in many applications. In medical image processing, the intensity based techniques are reported in [4, 30, 31]. In case of computer vision applications, the intensity based techniques are used but due to robust feature detection and description algorithms, they are replaced with feature based approaches [6, 32]. In remote sensing, feature based methods are preferred and used to register SAR images [29], infrared and visible color images [33], multi-spectral and multi-modal images [5, 32, 34, 35], and satellite images and aerial images [15, 26, 36, 37, 38].

The feature based techniques rely heavily on extracted features which should remain invariant to illumination, rotation, scale, photometric and projective changes. Many feature detection and description algorithms have been proposed over the course of time. Table 1.1 summarizes some well-known feature point algorithms. Some algorithms have both detector and descriptors parts while other have either detector or descriptor part.

Table 1.1: Summary of feature point detection and description algorithms

Method	Abbreviation	Detector Part	Descriptor Part
Scale Invariant Feature Transform [39]	SIFT	Yes	Yes
Speeded-Up Robust Features [40]	SURF	Yes	Yes
Binary Robust Invariant Scalable Keypoints [41]	BRISK	Yes	Yes
Oriented FAST and Rotated BRIEF [42]	ORB	Yes	Yes
Features from Accelerated Segment Test [43]	FAST	Yes	No
Binary Robust Independent Elementary Features [44]	BRIEF	No	Yes
DAISY [45]	DAISY	No	Yes
Local Histogram of Orientated Phase Congruency [46]	LHOPC	No	Yes
Radiation Invariant Feature Transform [47]	RIFT	No	Yes

1.4 Motivation

In remote sensing, the SI is considered as a key source of information and required frequently in order to monitor the earth surface [37]. Frequently obtaining the recent SI from agencies such as French Earth Observation Satellite System's SPOT, National Aeronautics and Space Administration (NASA) is an expensive exercise. However, there are some sources such as Google Earth^{TM} which provides SI free of cost but normally these images are outdated. In agricultural domain, the updated SI are frequently required for precision agriculture tasks like monitoring crops growth, health and yield estimation. Acquiring recent images from SPOT or NASA on regular basis is not an economical solution and also not affordable for farmers of developing countries like Pakistan. There is a need to investigate a cost effective solution which farmers can afford and use easily. In this regard, this thesis proposes a new method to update freely available SI with recently taken UAV images. The SI is updated through image registration process. The UAV images are high quality and high resolution images as compared to SI. In this thesis, Google Earth^{TM} images are used as SI, whereas the UAV images are acquired from low flying platform with ordinary RGB camera.

1.5 Problem Statement

SIs are low resolution and blurred images compared to UAV images. An example of SI and corresponding UAV image is shown in Figure 1.1. Both type of images depicts the same agriculture area, but the appearance of the area is different in SI and UAV image due to high temporal, textural and intensity differences. These changes are due to difference in altitude, sensors, acquisition time and date and also due to growth or changes in agricultural crops with the passage of time. In Figure 1.1, the left image is SI which is taken from Google Earth^{TM} while the right image is an UAV image of the same area.

(a) (b)

Figure 1.1: An example of agricultural land images. (a) Satellite image (b) corresponding UAV image

It is evident that both images possess high intensity and texture differences. Existing feature points such as Scale Invariant Feature Transform (SIFT), Speeded-up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB) and Binary Robust Invariant Scalable Keypoints (BRISK) are not invariant to such type of differences as they have been developed to overcome common types of transformations and deformations such as scale, rotation, affine and projective differences. These feature points underperform in SI-UAV image registration task. Therefore, an investigation is required to overcome such differences. Moreover, a robust feature matching strategy is also required for image registration. An investigation is also required whether using a machine learning approach the number of correct feature point matches can be increased for image registration between SI-UAV images.

1.6 Research Questions

i. Which feature point detection and description algorithm is best for registration between SI-UAV images of agricultural land?

ii. Can feature matching strategies based on machine learning techniques help in registration between SI-UAV images?

1.7 Contribution

The main contributions of this thesis are as follows:

i. A performance evaluation of various feature point detection and description algorithms to identify the best feature point for SI-UAV images registration.

ii. A new feature matching strategy based on nearest neighbor and brute force techniques for registering SI-UAV images of agricultural land.

1.8 Thesis Organization

The introduction has provided perceptions into importance, applications, problems and objectives of the research. Chapter 2 presents an overview of various image registration techniques and applications of image registration in different fields. Chapter 3 briefly describes various feature point detectors and descriptors algorithms used in this thesis. Also explains the importance of these feature point algorithms for image registration process. Chapter 4 presents the proposed method. Each step of proposed method is thoroughly explained. Chapter 5 presents experimental setup and results. It provides the results obtained with and without using the proposed method. And identify the difference between the performances of feature point algorithms. Finally, the thesis is concluded in Chapter 6 with future directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

In image processing, registration of two images is considered as essential task. It is extensively used in remote sensing, medical image analysis and 3-D image mosaics [1, 2]. In the last few decades, various image registration techniques have been developed. Selection of technique for an application depends on types of variations between the images [1]. Variations can be classified into three types.

i. Viewpoint, scale, rotation and affine variations, which are due to using different camera or different viewpoint for image acquisition.

ii. Intensity based variations, which are due to atmospheric and lighting conditions during the acquisition of images.

iii. Temporal variations, which results in change in the image features, for example, object location, scene change or growth of fields or forests.

Intensity based and feature based variations make registration process more complicated because establishing a match of a feature in other images is difficult. Type of variations in the images plays a major part in choosing the best feature space, search space and similarity measure to accomplish image registration. Causes of transformation and deformation and their competitive solutions are extremely important and should be considered for selecting or designing the method or technique for a specific application of image registration [1].

The image registration methods can be divided into two categories, i) area-based ii) feature-based [2]. Area-based techniques are more focused on matching of pixel intensities while feature-based approach focused on detection and matching of essential interest point named as feature points between the images. Major advantages,

contributions, drawbacks of these approaches and selection of appropriate method for image registration depending upon the problem and application is important [2].

Image registration can be multi-modal [21, 22], intra-sensor [8] or inter-sensor [36]. In case of multi-modal image registration, a method based on Maximum Likelihood is used [48]. This method uses intensity and the Control Points (CP). The likelihood function is maximized to approximate the parameters of the homography for registration. Similarly, to register intra and inter-sensor images, contour based approaches using the object boundaries and strong edges are used in [49].

2.2 Registration of Medical Images

In medical field, the contour based approaches are used for registering the nonlinear medical images [50]. The contour matching based method first extracts object boundary curves, then match the boundaries to estimate transformation for registration. In case of multi-modal image registration, the Mutual Information (MI) based approaches are used [4]. Maximization of MI criteria is used to determine the image intensities and their information redundancy or dependence in corresponding pixels [30, 31].

In [51] a computer program named as $elastix$ is developed by Klein el al., that contains stack of algorithms to address problems related to multi-modal image registration. A method based on histogram is presented in [52], which can be used for approximating and maximizing MI between multi-modal and multi-band images. For approximating MI, histogram estimation methods are utilized. To overcome the problem associated with histogram estimation the kernel density estimation is combined with non-uniform signal quantization. This approach has applications in several medical imaging [52]. Similarly, an intensity mapping in presented in [53] that uses MI along with coefficient of cross-correlation to deal with several registration issues occurs during the image registration process. MI is also used as similarity measure in a framework designed for image-geometry registration [54]. Graphic properties relative to illumination are utilized for this purpose.

A technique based on local phase [55] is presented in order to lower the computational cost of registering the multi-modal images. Fast Fourier Transform is utilized for an extraction method for detection of features based on local phase and on multiple scales. Mapping Complexity similarity measure is used by minimizing the complexity of intensity and spatial mapping [56]. This similarity measure is implemented with kernel matrix and eigenvectors. An enhanced Maximization of

Mutual Information (MMI) algorithm that uses combination of feature-based technique and spatial information for registering multi-modal airborne images is presented in [57]. Registration through MMI is robust to changes in brightness produced by multi-modal sensors. Efficiency is enhanced by identifying information-rich areas through classification using Harris corner filter (HCL). These areas are used as descriptors for feature matching for registration purpose. In [58], Loeckx el al., introduces a similarity measure for registering non-rigid images, which is an extension of MMI, named as Conditional Mutual Information.

2.3 Image Registration of Remote Sensing

Remote sensing sensors have experienced a quick advancement in recent times therefore, the use of satellite images in different kind of fields has achieved a colossal lift [10]. Now multi-sensor, multi-frequency and multi-temporal image data is available from Earth observation satellites thus, digital fusion of this image data is considered as beneficial tool in remote sensing applications. Image fusion based on pixels and Earth observation satellite data is presented in [11]. A hybrid method is proposed in [59] that combines intensity based and feature based methods for image fusion and registration. The hybrid approach enables the solution for overlapping problem occurs in intensity based methods. It uses edge based registration approach with coarse-to-fine multi-scale iterative refinement. The purpose of hybrid scheme is to balance out the limitations of individual approaches.

Mutual information is also used in remote sensing. Several algorithms have been developed using mutual information for registration of remote sensing images. One such approach is known as Generalized Partial Volume Estimation [60]. This approach is based on estimation of joint histogram for mutual information computation to accurately register multi-temporal remote sensing images. However complete accuracy is not possible but this approach can be used to an extent to get better results.

A framework of [61] is used for remote sensing image registration that uses idea of combining active contour with MI. Active contour model is used for segmentation of edges, and using these edges with MI for image registration. MI is used as similarity measure after removal of non-matched curves.

2.3.1 Feature Point based Image Registration

The feature based methods are commonly used for image registration due to their robustness towards image variations such as scale, rotation and illumination. Features based methods consists of four steps, i) feature detection, ii) feature matching,

iii) transformation parameter estimation, iv) transformation and resampling [2]. Various algorithm for feature point detection have been proposed [39, 40, 42, 43]. But finding the correspondence between features under scale, viewpoint and illumination changes is a challenging problem. To overcome this problem, feature points are detected in such a manner that they remain invariant to scale, rotation and other image variations. The invariance to illumination and geometry makes these features appropriate solution for finding the correspondences between the images. Furthermore, descriptors are computed by using the image patches around the feature points. By matching the descriptor, correspondence are established which are processed with RANSAC and epipolar constraint for outlier filters [62]. An insight detail is presented in [29] in order to review the capability of SIFT as a robust feature point for image registration in applications of geometric registration such as matching and registering SAR images.

An investigation is carried out in [5] regarding image registration of multi-spectral remote sensing images using the SIFT methods. The results indicate high false matching rate because spatial information is not observed. Therefore, to solve this issue, method is proposed that utilize neighborhood information to raise the matched CP points of SIFT. Transformation model known as Local Weighted Mean is combined with SIFT methods to enhance the correct matching rate [5]. A method to obtain robustness against nonlinear intensity differences is presented [63], in which histogram bins with 180° difference are combined to get shorter vector descriptors. This approach is named as Orientation Restricted SIFT, that offers better performance in multi-spectral satellite image registration process. In [32], Teke and Temizal explored the applicability of SURF method for registering multi-spectral satellite images. Adaption of Scale Restriction (SR) method to SURF enhance the performance of registration process.

A feature point based approach is proposed in [64] for registering visible and infrared images. The method uses trajectories of moving objects by detecting foreground and background in infrared and visible images. RANSAC [65] is used to match trajectory points and overlap of pixels of composite foreground images for the registration process. For infrared and visible spectrum images, a feature point descriptor is presented in [66]. Features are detected using SIFT-like detector and descriptors are computed through approach based on edge oriented histogram. Matching is carried out by searching nearest couple to the feature point. An investigation about performance of different conventional feature descriptors carried out in [33]. The comparison is done on infrared and visible spectrum images. A

method for image matching based on local features is also proposed [67], which are invariant to illumination and viewpoint. Iterative estimation of illumination and relative view relationship between the images are carried out in iterative way in order to reproduce the view of one image into another.

An image registration approach presented in [3] for remote sensing using feature matching strategy and histogram of Triangle Area Representation (TAR). This approach is based on transformation parameter estimation algorithm which act as replacement to RANSAC [65] and Progressive Sample Consensus (PROSAC) [68] methods as it can efficiently compute the higher accuracy consensus set. A homography estimation technique that combines the keypoint correspondences with appearance similarity is presented in [69]. A method named as Particle Swarm Optimization Sample Consensus Algorithm, which is able to deal with low correct matches rate and helps in finding more matches is also proposed in [70]. Fast Sample Consensus is introduced in [71], which is an enhanced version of RANSAC, to improve the accuracy of feature matching of remotely sensed images. Similarly, a technique named as Progressive Vector Field Consensus is also presented [72] to enhance correspondence between features of images. A robust algorithm based on feature point matching is introduced in [73], that is based on feature descriptors computed with K-Nearest Neighbors TAR (KNN-TAR). The descriptor enables the identification of outliers, and then local structure combined with global information is used to remove the outliers. Histogram of Collinear Gradient-Enhanced Coding (HCGEC) [34] and Log-Gabor Histogram Descriptor (LGHD) [74] descriptors are proposed that are robust for matching multi-spectral images.

For images relating to multi-spectral remote sensing, automatic registration could be demanding task because of differences in non-linear intensities between the images. To solve the issue a two-step registration method based on local descriptor is introduced in [75]. In first step, Scale Restricted SIFT is utilized for removing any differences related to rotation, translation and scale between images. In second step keypoints are detected through Harris corner method. Then modified local self-similarity descriptor is utilized to find bonding points. For registration piecewise linear model is utilized. Similarly, another two-step procedure is proposed in [27] for registering remotely sensed images automatically. SIFT method is used in first step and MMI is used in second step to achieve automatic registration of images. In [18], Liang et al., designed an approach that combines spatial information of detected features with intensities and MI, which is capable of automatically registering the remote sensing images. The combination of spatial information and MI is carried by introducing a

similarity metric name as Spatial and Mutual Information. Similarly for multi-modal remote sensing images, in [76] a method based on feature descriptor which is called as Histogram of Orientated Phase Congruency (HOPC) is used, that depends on images structural properties. Further a similarity metric is presented by introducing orientation in phase congruency model to get a modified model. This similarity metric utilizes the coefficients of normalized correlation to provide basis for automated registration method for remote sensing images [76].

A SIFT based approach is introduced in [28] for registration of remote sensing images. A new interpretation of gradient that includes magnitude and orientation is presented to surpass any intensity differences and then enhanced method which combines scale, orientation and location of keypoint is applied to improve correct correspondences during feature matching. Adaptive Binning SIFT (AB-SIFT), a feature descriptor is presented in [77] in order to get automatic image matching system for remotely sensed images. The primary concept of the method is computation of descriptor through adaptive binning technique. Extraction of region carried out using modified Hessian affine technique and then utilizing adaptive histogram quantization approach to compute the descriptor. Similarly, for remote sensing images another feature matching algorithm called as Locally Linear Transforming is proposed in [9] that is capable of managing rigid as well as non-rigid transformations when large number of outliers are present. The method approximate transformation and produce correct correspondences at the same time using maximum-likelihood structure. Another feature point matching approach for remote sensing images is presented in [35] that uses SIFT method to create a preliminary set of correspondences. Then inliers are increased using global structure constraint (shape context), and outliers are removed through local structure constraint i.e. TAR.

2.3.2 Registration of Satellite and Aerial Images

SI is key source of information in many remote sensing applications, that has to be updated regularly by acquiring new images with the passage of time [37]. For this purpose, a system is designed [8] for automatic registration of satellite images with remote sensing aerial images. The basic idea for designing this system is to overcome the problem of updating the remote sensing images continuously, that are acquired from separate sensors and/or at different intervals. The system is capable of managing geometrical distortions of remote sensing images occurs due to several transformations like rotation, scale and translation. Phase congruency model is the core for processes of CP detection and matching. The system is able to address several issues related to contrast or intensity during remote sensing image registration

with intensity invariant features and adaptive scheme of CP matching. At the end a modified RANSAC approach is used for removal of outliers and improving the accuracy [8]. Similarly, another automated system is introduced in [36], that tackles several issues like environmental and sensor noise, contrast and intensity differences between satellite images. Geman-McClure M-estimation approach is modified to get a new CP matching approach that uses complex feature representation. For improving localization of CP pair, CP pairs are iteratively refined and used for removal of outliers to estimate transformation for image registration.

The SIFT based method has been suggested for image registration in [5], similarly, in [78], a new method for automatic registration of images is proposed that combines SIFT method with the concept of image segmentation. Principal Component Analysis (PCA) and a robust process removes outliers and provides aid for designing the method. By combining different techniques, it enables registration of images having differences in translation, pixel size and rotation leading to subpixel accuracy. An approach based on deep learning is explored in [79] with the aim of improving the accuracy of optical satellite image geo-localization using SAR image data. A trained neural network, consisting of extraction and similarity measure stages, is used for achieving the goal. In [80], Zeng el al., presented a feature matching approach based on neighborhood geometry for improving the feature matching accuracy for images of remote sensing acquired through geostationary satellites. The approach comprises of three steps and using edge information as a key concept of the approach.

Gaussian mixture model along with the shape feature detector is used for image registration in [81]. Registering historical aerial image of an area with recent aerial image of same area helps in determining the changes occured during the period of time. For this regard, an approach is presented [26] that uses Time Invariant Line (TIL) features to register historic aerial images. Orientation is determined by area minimization between historical and recent images relative features. TIL are modified presentation of line features but without line-to-line correspondence. Geometric features depicted by lines are quite balanced which are useful for registering images of multi-temporal remote sensing. Therefore, a procedure is introduced [38] that uses line features for registration of multi-temporal planar segment aerial optical images. The main step of this procedure is to find the correspondence in images through lines. Furthermore, homography transformation is derived using corresponding lines parameters in images. Similarly, another registration method using composite deformable matching is presented in [82] that is based on edges and image entropy as features to register satellite images and UAV images. Edge features are used to

overcome the differences of illumination and sensors.

A method for effective stitching of aerial images is presented in [83] that uses ORB feature points to achieve accurate stitching. An approach for fast image registration on the basis of geometric invariants and features utilizing the ORB is put forward [84]. Features are extracted and described using slightly modified ORB approach. Then a distance constraint is applied for removing mismatches based on keypoints distribution. Geometric constraints are introduced to provide base for feature matching procedure.

For gathering geospatial data, a camera attached with UAV, now-a-days, is used as low-cost platform. Ortho-images are fundamental products and primary layer for databases of Geographic Information System, therefore, [19] exhibits that ortho-images accuracy can be improved through image registration. It also proposes a scheme for image registration that comprises Accelerated BRISK approach along with analysis of corresponding spatial CPs. Feature descriptors of images are compared to find the match and false matched features are removed using sorted ring to make matching process fast and accurate. In disaster situation, UAVs can also provide fast support for real time monitoring, therefore, in [85] methods are proposed for registering inter-spectral images using aerial thermal and visual images captured from UAVs with low-altitude.

A multi-viewpoint method [20] presented for registering remote sensing images comprising three preliminary steps, i) construction of mixture model using multiple features, ii) combining the Euclidean distance, shape context and SIFT distance to complete the feature mixture model, iii) introduction of geometric constraint for non-rigid transformation. The method is applied on images taken from Google Earth^{TM} and UAV and performance is examined.

2.4 Image Registration for Agricultural Applications

In [86], Wei et al., proposes a method for registering multi-temporal images and monitoring the change for agricultural terraces. This method is a four stage approach i) image filtering, ii) feature extraction, iii) point set registration, iv) feature points based registration and utilizes a small UAV to capture the aerial images of agricultural terraces. An automatic workflow is presented in [87] for processing high resolution satellite images for smallholder agricultural lands. This workflow is based on computer software with primary purpose of monitoring the development of crops and fields with passage of time by evaluating multi-sensor and multi-temporal images.

Updating the agricultural land images using satellite images or aerial images taken from UAV is vital for evaluating the changes due to new crops, crop yields, forest growth etc. In this regard, an approach for registering SI of agriculture land with aerial images is proposed by [17]. Farmland field junctions are used as features. By extracting boundary junctions and then computing junction descriptors by using structural and geometric properties of boundaries connected to junction. For calculating these properties, a rectangle was fitted to the fields. Junction descriptor contains patterns capable of matching and establishing junction correspondence. Another method is proposed in [37] for updating outdated agriculture land SI with high resolution UAV images. In this approach fields are described by measuring relative field boundaries to field boundary junction locations. Every field have separate description which make junction description distinct and helps in reliably finding the correct matches. Similarly, image matching between agricultural land images using different detector-descriptor pairs is presented in [15]. A comprehensive analysis on performance of feature detectors and descriptors is carried out that one may found as appropriate feature detector and descriptor for registering agricultural land images along with proposing a new feature descriptor named as Modified Normalized Gradient SIFT for achieving further ways and better performance.

In [46] LHOPC, a descriptor with advanced configurations is proposed by Ye et al. that uses an extended phase congruency feature. The purpose of developing LHOPC is to provide a robust solution for radiometric and geometric changes for remote sensing images. RIFT method is proposed in [47] by Li et al., to eliminate the problem of Nonlinear Radiation Distortions during the feature matching process. This method is used for multi-modal and remote sensing image matching.

In [88], Mikolajczyk and Schmid explored many feature descriptors based on interest regions on base of assessment criteria of same scene matching and recognition. A comprehensive and detailed overview is presented in [89], about feature detectors that are locally invariant and widely used in many applications, with a qualitative assessment of their robustness and deficiencies. In [90], Juan and Gwun examined three feature based methods for their application in different recognition systems. Similarly an investigation [91] is carried out for seven famous detector-descriptor combinations using different effects and variations. Similarly, the performances of computationally low-cost binary descriptors are examined in [92] and [93] provides suggestions for finding best combination of detector and descriptor as per requirement.

Table 2.1: Summary of image registration methods

Application	Method used	Comments
Medical Image Analysis [1, 2, 3, 30, 50]	Mutual information, active contour models	Mutual information based methods are mostly used for this field
Computer Vision and Object Tracking [1, 2, 3, 5, 7, 32, 55]	Active contours, mutual information, phase based, similarity metric, feature based	Mutual information based methods are used but with new trend, feature based are becoming popular choice
Remote Sensing Image Processing [1, 2, 3, 8, 9, 15]	Mutual information, active contours, feature based, similarity metric, deep learning	Feature based methods are frequently used for most of the applications.
Satellite and Aerial Images (Remotely Sensed) [15, 17, 20, 37, 87]	Feature based	Feature based methods are priority choice satellite and aerial images registration

2.5 Summary

A brief overview of image registration applications in different fields and method is presented and used to overcome challenges relating to different fields are summarized and depicted in Table 2.1 below.

The feature point based methods has attracted the attention of many researchers in the area of image registration. In the literature, feature point based methods have been found very useful and a reliable option for registration of SI-UAV images of agriculture land. Therefore, the proposed method is also based on feature points, and uses the feature points in a novel way to register SI-UAV images in the presence of high temporal textural and intensity differences. But feature based methods require the feature points that are computed in such way, so they remain invariant to any illumination, rotation, scale, photometric and projective changes between images. Many feature detection and description algorithms are introduced over the course of time and their performances are regularly evaluated and examined for different type of applications.

CHAPTER 3

AN OVERVIEW OF FEATURE POINT ALGORITHMS

3.1 Overview

Feature points are distinct location in an image that are invariant to change in intensity, viewpoint, scale and rotation. Many algorithms have been presented regarding how to extract and record such feature points. During the last decades, these features are used in various applications of image processing and computer vision. The well-known feature point detector and descriptor algorithms are SIFT [39], SURF [40], ORB [42] and BRISK [41], which are briefly explained in below sections.

3.2 Scale Invariant Feature Transform

Scale Invariant Feature Transform (SIFT) image features are invariant to rotation, scaling and intensity variations. SIFT works efficiently against affine and projective variations. The SIFT algorithm takes an input image and convert it into huge group of local feature vectors [94]. These feature vectors of image are invariant to any rotation, translation or scaling. For feature extraction, SIFT detector consist of four stages, i) Scale Space Extrema Detection, ii) Keypoint Localization iii) Orientation Assignment, iv) Keypoint Descriptor.

3.2.1 Scale Space Extrema Detection

In this stage image is scanned to identify those locations and scales that can become potential candidates for feature points which are orientation and scale invariant. This can be achieved effectively by scale space function. Furthermore, it has been shown that Gaussian function is the basis of this function. The scale space $L(x, y, \sigma)$ of an image is defined in Equation 3.1 as:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3.1}$$

where (x, y) represents feature locations, $G(x, y, \sigma)$ is Gaussian with variable-scale σ, $*$ is convolution operator and image $I(x, y)$ is given as input.

Stable keypoints in scale space can be distinguished using various approaches. One of them is Difference of Gaussians (DOG). Scale-space extrema, $D(x, y, \sigma)$ is located by computing the difference between a image and similar image but with k times scale factor. $D(x, y, \sigma)$ is then given in Equation 3.2 as by:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{3.2}$$

where $L(x, y, \sigma)$ denotes scale space of an image, and $L(x, y, k\sigma)$ is k times scale factor of same image scale space. In order to find the local maxima and minima of $D(x, y, \sigma)$, each point is compared to its eight neighbors in current scale and nine neighbors in the scale up and down. It can be extrema only if it's value is maximum in all of these neighbors or minimum in all of them [39].

3.2.2 Keypoint Localization

This stage refines the candidate keypoints by eliminating weak points. The points that are localized poorly along an edge or low contrast points are also rejected and eliminated. It can be achieved by computing the value of Laplacian for each point. The detection of extremum location, $\hat{x}$, is defined in Equation 3.3 by:

$$\hat{x} = -\frac{\partial^2 D}{\partial x^2}^{-1} \frac{\partial D}{\partial x} \tag{3.3}$$

where D is scale space extrema, x denotes feature location and ∂ represents partial derivatives. All extrema with the value of function at $\hat{x}$ less than specified threshold are discarded. This eliminates extrema having low contrast. Extrema that is localized poorly can be eliminated in different cases where it is noted that across the edge large principal curvature exists. But in perpendicular direction, there is a small curvature in Gaussian function difference. If the difference is less than the ratio of largest to smallest eigenvector at scale and the location of keypoint, from a 2×2 Hessian matrix, the keypoint is unaccepted [39].

3.2.3 Orientation Assignment

In this stage, an orientation is assigned to the keypoints based on local properties of image. This orientation is useful in representing the keypoint descriptor, that achieves invariance to rotation. Gaussian smoothed image L is selected by using the keypoints scale for computation of gradient magnitude, $m(x, y)$, and orientation $\theta(x, y)$ from Equation 3.4 and Equation 3.5:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{3.4}$$

$$\theta(x,y) = tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \tag{3.5}$$

where $L(x,y)$ represents Gaussian smoothed image at (x,y) location. Formation of an orientation histogram takes place with the help of the gradient orientations of sample keypoints. The $360°$ orientations range has been distributed using 36 bins of orientation histogram. The highest peak is identified in the histogram. The identified highest peak is considered as keypoint dominant orientation. Additional orientation to the keypoint can be assigned if there is any peak within 80% of identified highest peak for creating the orientated keypoint. Multiple orientations can be assigned to several keypoints [39].

3.2.4 Keypoint Descriptor

Image location, orientation and scale have been assigned to each keypoint in previous stages. The local gradient data used in those stages is used here to compute the keypoint descriptors. A set of 16 histograms is used by the SIFT keypoint descriptor algorithm, which are positioned in 4×4 array each containing eight orientation bins. Division of eight orientation bins are carried in such a way that one bin is nominated for each of the main orientation directions and one bin is nominated for each of directions in the mid-points. Thus resulting into a 4×4×8 = 128 elements size feature vector. The resulting feature vector are known as SIFT descriptors. They are specifically useful because of their distinctiveness, that enables the correct match for a feature to be found from a considerable sized feature database [39, 94].

3.3 Speeded-Up Robust Features

Due to high time consumption of SIFT, an algorithm is proposed inspired from same steps and principles but with a different phenomenon to provide faster results known as Speeded-Up Robust Features (SURF) algorithm [40]. Like SIFT, SURF is also able to compute the interest points as well as descriptors that are invariant to rotation and scale. Because of rotation and scale invariant nature of SURF descriptors and less computation time, it is widely used instead of other methods for feature extraction [91]. Simple Hessian-matrix approximation is used for detecting the interest point. Using integral images, the time of computation can be reduced significantly [40].

3.3.1 Hessian Matrix based Interest Points

Hessian matrix is the basis of SURF detector. Blob-like structures are identified at locations with maximum determinant. In a given image I, a Hessian matrix $H(x,\sigma)$

at a point $x = (x, y)$, is defined in Equation 3.6 as follows:

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \tag{3.6}$$

where $L_{xx}(x,\sigma)$ is convolution of second derivative of Gaussian (g) of standard deviation σ with the given image I and it is defined in Equation 3.7 and Equation 3.8 as:

$$L_{xx}(x,\sigma) = I(x) * \frac{\partial^2}{\partial x^2} g(\sigma) \tag{3.7}$$

$$L_{xy}(x,\sigma) = I(x) * \frac{\partial^2}{\partial xy} g(\sigma) \tag{3.8}$$

where $I(x)$ is image at point $x = (x, y)$, σ represents scale and $\frac{\partial^2}{\partial x^2} g(\sigma)$ and $\frac{\partial^2}{\partial xy} g(\sigma)$ are second derivative of Gaussian for L_{xx} and L_{xy}, respectively. The characteristic aspect of performance of SURF is non-maximal-suppression of the hessian matrices determinants [91].

3.3.2 Integral Images

To calculate the convolutions is costly task therefore, by the using integral images this process is speeded-up and approximated. Integral Images enables fast calculation of box type convolution filters. It is defined as an image $I(x)$ where at location $x = (x,y)^T$, it represents the sum of all pixels between origin and location x within a rectangular area. It is described in Equation 3.9 as follows:

$$I(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x,y) \tag{3.9}$$

where $I(x,y)$ denotes image at (x, y) location.

The second derivative Gaussian kernel used for the Hessian matrix (H) have to be distinctive and cropped. A 9×9 kernel is then applied. These kernels are approximated by SURF algorithm with box filters. Through this it is possible to compute the approximated convolution for arbitrarily sized kernel effectively using the integral image.

$$Det(H) = D_{xx} D_{yy} - (w D_{xy})^2 \tag{3.10}$$

where the discrete and approximated kernels are denoted as D_{xy} for $L_{xy}(x,\sigma)$ and D_{yy} for $L_{yy}(x,\sigma)$. The kernel refers to $\sigma = 1.2$ pixels and they are the lowest scale that can be handled by SURF algorithm. Weight w is used to calculate the determinant of hessian matrix using the approximated kernels. The weight w is sensitive to scale.

Theoretically it can be constant at 0.9 [40]. The integral images enable calculation of responses with arbitrary large kernels by the SURF algorithm.

3.3.3 Scale Space Representation

To detect features across scale several octaves and levels are to be considered. Therefore, scale space is divided into octaves and levels. The integral images make the convolution with regular Gaussian kernel having the large filter size. It becomes computationally fast and less expensive, and also there is no need to downscale the image.

3.3.4 Interest Point Localization

Interest point is localized in the image and over scales by applying non-maximum suppression in $3\times3\times3$ neighborhood. The maxima of Hessian matrix determinant is introduced with the help of method described in [62] in scale and image space.

3.3.5 Descriptor

The descriptor of SURF algorithm is based on Haar wavelet responses. With the integral images they are efficiently and easily computed. The purpose of descriptor is to provide description of the feature that is unique and robust. On the basis of surrounding area of feature, a descriptor can be generated. Like SIFT, determining the orientation is important need in SURF because rotational invariance can be achieved by giving each feature a unique orientation. SURF descriptor divides keypoint neighborhood region or interest area into 4×4 sub-regions. Then a 2D Haar wavelet response is calculated for every sub-region. Four values are contributed by each response of sub-region to the descriptor. Therefore, $4 \times 4 \times 4 = 64$-dimensional feature description is described for each keypoint for all sub-regions [40]. SURF has been proved to work faster than SIFT on occasions. But in some scenarios like intensity and viewpoint change SIFT outperformed the SURF methods [91].

3.4 Oriented FAST and Rotated BRIEF

In number of applications, SIFT [39] have been proved successful and efficient. But its computational cost is significantly high which leads to search for an alternative method which can perform with lower computation cost. SURF [40] is an option for replacing SIFT but yet in some situations it cannot match the performances of SIFT. Therefore, a method with effective computational cost is presented having similar performance ability to SIFT as a replacement. This method is known as Oriented FAST and Rotated BRIEF (ORB) [42] that is based on FAST detector [43] and BRIEF

descriptor [44]. Purpose of using FAST detector is because it finds decent corner keypoints efficiently and is widely used in real-time system applications. BRIEF descriptor is simple regarding its construction. It is a bit string description, considering a smoothed image patch defined by simple binary intensity tests between pixels of image.

3.4.1 Detection

FAST detector is used for detecting the keypoints because of computationally low cost but it cannot support the orientation facility. To overcome the problem a modified version of FAST detector is employed. Firstly, from an image FAST points are detected. For measure of cornerness which lacks in FAST, after detecting keypoints, Harris corner measure is applied for the sake of sorting them. Based on threshold value, top N points are selected and then sort them according to Harris measure. By using intensity centroid [95], ORB computes the local orientation, that is a weighted averaging of pixel magnitudes in the local patch also using first-order moments of patch in order to achieve rotation invariant features. Between the centroid and feature location, the vector is known as orientation. This method looks unstable but it is competitive with orientation assignment used in SIFT [92].

3.4.2 Description

BRIEF is a bit string descriptor that provides description of smoothed image patch defined by binary intensity test set. Given a smoothed image patch p and a binary test τ can be defined in Equation 3.11 as:

$$\tau(p; x, y) = \begin{cases} 1 & p(x) < p(y) \\ 0 & p(x) \geq p(y) \end{cases} \tag{3.11}$$

where $p(x)$ is intensity of image patch p at point $x = (x, y)$. A vector contains n binary tests is defined as a feature which is given in Equation 3.12 as:

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i) \tag{3.12}$$

where p is image patch and $f_n(p)$ is feature of p, (x_i, y_i) denotes feature location.

For test distribution, a Gaussian distribution in all directions of patch center is used. BRIEF descriptor performance in many aspects like blur, robustness to lighting and distortions is similar to SIFT but its matching performance falls of significantly in situation of in-plane rotation [42]. To solve the problem of rotation invariant, an

efficient method is proposed that uses keypoints orientations to steer BRIEF. Given a set of feature at location (x_i, y_i) having n binary tests, a $2\times n$ matrix is defined in Equation 3.13 as:

$$S = \begin{pmatrix} x_1, \ldots\ldots, x_n \\ y_1, \ldots\ldots, y_n \end{pmatrix} \tag{3.13}$$

where S is $2\times n$ matrix and (x_i, y_i) represents feature location. Patch orientation θ, corresponding rotation matrix R_θ is utilized to construct S_θ, a "steered" version of S in Equation 3.14 as:

$$S_\theta = R_\theta S \tag{3.14}$$

where R_θ is rotation matrix and S_θ is steered form of matrix S. Therefore, steered BRIEF operator will be defined in Equation 3.15 as:

$$g_n(p, \theta) := f_n(p) | (x_i, y_i) \in S_\theta \tag{3.15}$$

where $g_n(p, \theta)$ represent steered BRIEF operator, $f_n(p)$ is feature of intensity **p** at location (x_i, y_i) and S_θ is steered matrix with patch orientation θ.

The key to computation of descriptor is the utilizing the set of correct points S_θ, which can be achieved by consistency of keypoint orientation across views [42]. In ORB the sampling pattern utilize 256 pairwise intensity comparisons. This is established through machine learning, descriptor's variance maximization and under various orientation changes minimization of correlation [92].

3.5 Binary Robust Invariant Scalable Keypoints

Local feature and their vector description provided by SIFT, SURF and likewise methods proved to be successful in many real-time applications but with expensive computational cost. This challenge leads to alternative solutions like binary descriptors and BRISK is one of them. Binary Robust Invariant Scalable Keypoints (BRISK) [41] is a method that provides scale invariance as well as rotation invariance. BRISK includes stages of detection of features, computation of descriptors and matching of keypoints.

3.5.1 Detection

To detect the regions of interest or location of features in the image while keeping the focus on efficient computational cost BRISK uses Adaptive and Generic Corner Detection based on the Accelerated Segment Test (AGAST) [96] method. This is an extension of FAST by improving speed with having similar detection performance provides the base for detection of feature locations. BRISK detects keypoint by

searching maxima in image plane as well as scale space in order to make it invariant to scale. It performs non-maxima suppression and compute around all scales.

3.5.2 Description

For description of features, BRISK uses symmetric patterns and avoid the unsystematic or learned patterns of BRIEF and ORB. The pattern defined the sample points placed in circles concentric around the keypoint. Each sample point represents the Gaussian blurring of its neighboring pixels. This blurring uses standard deviation which is increased with distance between points on circles. This pattern looks identical to DAISY descriptor [45] but BRISK differentiate with it because DAISY was specifically used for dense matching and holds extra information for description of keypoint [41]. The sampling-point pairs are then defined into two subset pairs named as short-distance pairing and long-distance pairing. Orientation is found by taking sum of computed local gradients between short-distance pairs and long-distance pairs. Thus the keypoint is rotated according to orientation to make BRISK invariant to rotation.

The descriptor is constructed by performing comparisons of intensity values of all short-distance pairs. For each pair, if the first point in pair value is greater than the second point in pair value then the bit string is represented by "1", otherwise it is represented by "0". Each of bit string obtained for point pair combined to provide 512-bit length descriptor vector. It can be defined in Equation 3.16 as:

$$b = \begin{cases} 1 & I({p_j}^{\alpha}, \sigma_j) > I({p_i}^{\alpha}, \sigma_i) \\ 0 & otherwise \end{cases} \tag{3.16}$$

where b represents a bit string, $({p_i}^{\alpha}, {p_j}^{\alpha})$ represents points in short-distance pair S with α as rotation orientation. BRISK utilizes Hamming distance rather than Euclidean distance for matching purpose because of its fast execution time. But yet, BRISK needs more computation and more storage space than either BRIEF or ORB [92].

3.6 DAISY

DAISY descriptor [45] is used for dense computation and matching. DAISY descriptors have also shown better performance in patch correspondence problems. A major advantage of DAISY descriptors is that while constructing descriptor the spatial binning of oriented derivatives is characteristic of different resolutions. But its limitation is the need of additional memory storage [41]. DAISY descriptor works by firstly computing the orientation maps from an input image, each for quantized direction. After that, each orientation map is convolved at various times with Gaussian kernel. To control region size, different values of Gaussian kernels are used thus

obtaining convolved orientation maps for various region sizes. Primary objective is to lower the computational cost and efficient implementation of convolutions using separable Gaussian filters. Moreover, it leads to low cost computation of various sizes orientation maps because various successive convolutions with smaller kernels can achieve large Gaussian kernel convolutions.

It is noted that in convolved orientation maps, every pixel location carries a value which is a weighted sum of gradient norms calculated over a neighborhood. Unlike other methods like SIFT etc., DAISY uses Gaussian kernel. DAISY comprises of vector at every pixel location, that contains convolved orientation maps values situated on concentric circles pivoted on location, and where the measure of Gaussian smoothing is corresponding to the radii of the circles. The vectors are then normalized to unit form. Normalization in each histogram is achieved independently not as a whole, to be able to portray pixels close to occlusions correctly. If whole descriptor was normalized, then the same point descriptors that is near an occlusion when imaged from different viewpoints would be quite different [45]. DAISY descriptor gives better results with DOG based detector [93], and it is evident that SIFT is also based on DOG technique as mentioned in earlier section, therefore in experimental setup combination of SIFT detector and DAISY descriptor will be used.

3.7 Local Histogram of Orientated Phase Congruency

LHOPC is proposed in [46] that uses an extended phase congruency feature [76] to compute the descriptor with advanced configurations. The purpose of developing LHOPC is to provide a robust solution for radiometric and geometric changes for remote sensing images. Phase congruency is a model used for feature detection which is impartial of signal magnitude. That makes it robust to changes like contrast and illumination mainly against gradient information. This conventional phase congruency model only considers magnitude of feature (i.e. gradient magnitude). But it cannot give consideration to information of feature orientation (i.e. gradient orientation) that reflects the feature variation directions. Therefore, conventional model cannot be used for construction of features which are robust and locally invariant. Thus, extension of the model of phase congruency is presented to assemble the orientation representation for building the feature descriptor [76]. The extended model returns both feature orientations and magnitudes that are mentioned respectively as orientation and magnitude of phase congruency. LHOPC use these parameters for descriptor building relative to DAISY-style spatial arrangement. Because it is using the benefits of phase congruency as well as DAISY therefore, this descriptor is expected to robust against radiometric and geometric changes [46].

To compute LHOPC descriptor for a keypoint, dominant orientation is assigned to a keypoint. This can be achieved by using orientation histogram created from phase congruency magnitude and orientation. The orientation histogram is distributed along $360°$ by each bin of $10°$ making total of 36 bins. Phase congruency magnitudes is used to make entries in histogram and after that highest peak is identified in histogram. The identified highest peak is considered as keypoint dominant orientation. Additional orientation to the keypoint are assigned if there is any peak within 80% of identified highest peak. This process is similar to orientation assignment of SIFT [39]. Around a keypoint a local region is formed and DAISY-style spatial arrangement is designed for representation of descriptor relative to dominant orientation. Using phase congruency orientations and magnitudes, descriptor is constructed relative to DAISY-style arrangement because it is proved to be more robust to illumination and geometric differences [46]. Combination of SIFT detector and LHOPC descriptor will be used in experimental setup.

3.8 Radiation Invariant Feature Transform

RIFT is proposed in [47] to eliminate the problem of Nonlinear Radiation Distortions (NRD) during the feature matching process. This method is used for multi-modal and remote sensing image matching. Traditional feature descriptors normally use gradient arrangement or image intensity to compute feature vectors but gradient and intensity are sensitive to NRD. For task of multi-modal image matching, these descriptors performance wise are not good. Phase congruency measure appeared to be robust against NRD. Therefore, initially using phase congruency map, a 216-dimensional feature vector is constructed based on distribution histogram approach. Discouraging results leads to presentation of Maximum Index Map (MIM) measure for better description of feature.

The log-Gabor convolution sequence is used to compute MIM and then distribution histogram approach is utilized for feature description that is similar to SIFT. An image patch focused on feature is selected for every feature point and weights are assigned to each pixel using Gaussian function with standard deviation. Local patch is then divided into 6×6 sub-grids and distribution histogram is constructed with a bin for each sub-grids. All histograms are concatenated resulting into a feature vector. Finally, by normalizing the feature vector results in invariance to illumination changes.

This procedure can be suitable considering no rotation changes, otherwise this process will not be useful. To make it rotation invariant further analyzing and processing is required. By going through further analysis that by computing feature

point orientation, rotation difference is removed between image patch, thus making it invariant to rotation. This method is specifically constructed for NRD challenges and has been proved to be resistant to NRD [47]. Like LHOPC, RIFT method also possess similarity to SIFT method therefore, SIFT detector and RIFT descriptor combination will be used in experimental setup. Table 3.1, lists the feature point detector-descriptor pairs used in experimental setup.

Table 3.1: List of feature point detector-descriptor pairs used in experimental setup and results

S. No	Detector	Descriptor	Descriptor Size
1	SIFT	SIFT	128
2	SURF	SURF	64
3	BRISK	BRISK	64
4	ORB	ORB	32
5	SIFT	DAISY	200
6	SIFT	LHOPC	200
7	SIFT	RIFT	216

3.9 Summary

By going through all discussion of feature detectors and descriptors, it is summarized that for experimental setup, several feature point detector-descriptor pairs are used in this thesis in order to obtain the results and performance comparison. As described in Table 1.1, SIFT, SURF, BRISK and ORB have both detector and descriptor part. But DAISY, LHOPC and RIFT only have descriptor part, therefore, their detector part is provided by SIFT detector in experimental setup (see Table 3.1).

CHAPTER 4

METHODOLOGY

4.1 Overview

The methodology of the thesis is based on experimental and quantitative approach for image registration of SI-UAV images of agricultural land. The proposed method is based on Nearest Neighbor (NN) and Brute Force (BF) descriptor matching strategy therefore, it is named as NN-BF method. The block diagram for the NN-BF method is shown in Figure 4.1. Each block is briefly described below.

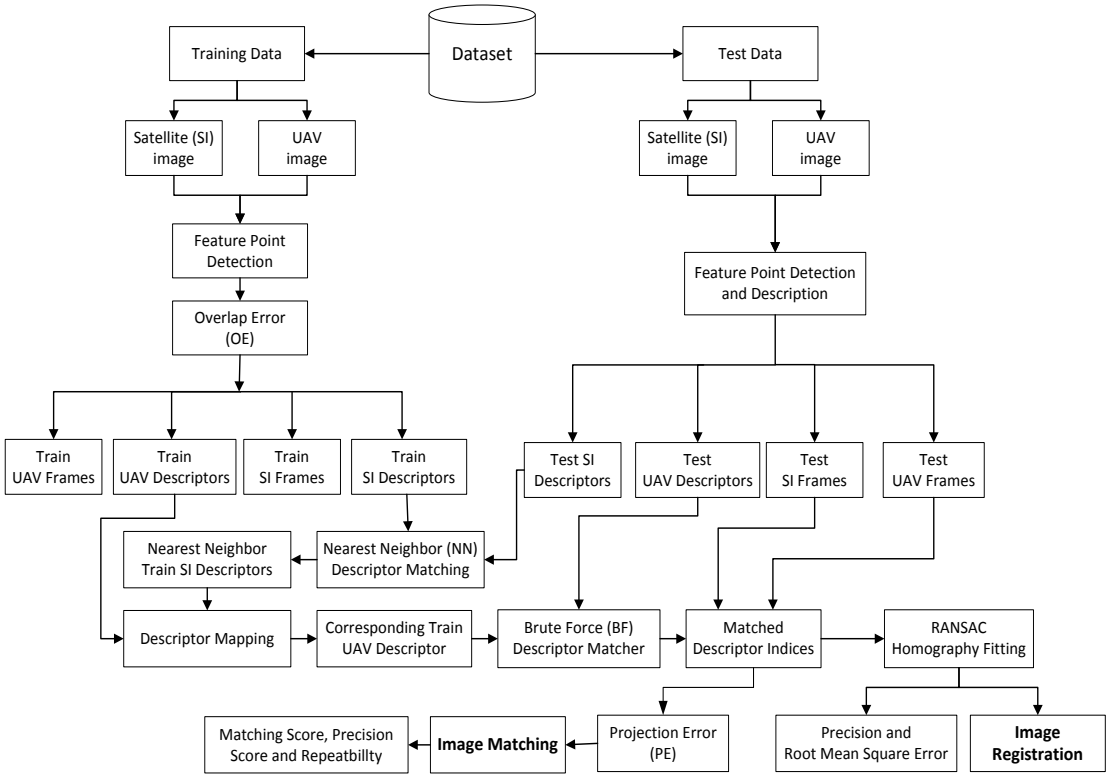


Figure 4.1: Block diagram for proposed NN-BF method for registration of SI-UAV images

4.2 Agricultural Land Image Dataset

The dataset consists of eight SI-UAV image pairs as shown in Figure 4.2 and Figure 4.3. Each pair consists of an SI image taken from Google Earth^{TM} and the corresponding UAV image captured with low flying UAV with ordinary RGB camera. The UAV (aerial) images were taken near Lahore. Each UAV image is geotagged, which helps in finding its corresponding location in the Google Earth. Around that location, an image patch is cropped which completely show the UAV image contents. This process is repeated for all the UAV's images of training and test set.

Each pair depicts the same agricultural land but possesses high temporal, textural and intensity differences which give the same agricultural land a different appearance in aerial view. The dataset is further divided into two disjoint sets; a training set and a test set.

4.2.1 Training Set

The training set comprises of two SI-UAV image pairs as shown in Figure 4.2. First column of the figure shows the SIs, and the second column shows the corresponding UAV images. The third column show manually registered SI-UAV image, which shows the physical location of the UAV image inside its corresponding SI.

(a)

(b)

Figure 4.2: Agriculture land images used as training set, first column shows SIs, second column shows corresponding UAV images and last column shows manually registered SI-UAV images.

4.2.2 Test Set

Test set consists of six SI-UAV pairs as shown Figure 4.3. The idea behind the training and test set is to learn the corresponding feature point matches between the SI-UAV images and then use these matches in the feature matching step to register the SI-UAV images of the test set. SI-UAV image pairs for training and test set and randomly selected. No certain criteria is used for selecting the image pairs for training and test set.

4.3 Manual Registration

The process of manual registration and homography estimation is carried out with $cpselect()$ function of MATLAB. For this purpose, the control points are selected manually in SI and their corresponding points are identified manually in UAV image. After the selection of control points, RANSAC is applied on control points to estimate the homography matrix.

4.4 Feature Points

The SIFT algorithm [39] is used in the proposed NN-BF method for the detection and description of feature points (see Section 3.2). The selection of SIFT is based on the experimental results presented in the next chapter where SIFT demonstrates better image matching and registration performances than SURF, ORB and BRISK (see Section 5.3.2). The feature points detected by SIFT have two parts: (i) Frames that contain information about pixel location *(x,y)*, scale, orientation of each detected feature point (ii) Descriptor, which is 128 element vector representing the distribution of image gradients around the pixel location of each feature point.

4.5 Homography

Homography (H) is a transformation matrix that is known in advance between each SI-UAV image pair of the training and test sets. When target image is aligned with reference image, parameters of mapping are estimated which is known as transformation estimation. Transformation matrix is a 3×3 matrix. Transformation is divided into two classes, affine and projective. Affine transformation is a special case of projective transformation. Following matrix represents both transformations:

$$\begin{bmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ c_1 & c_2 & 1 \end{bmatrix}$$

(a) Pair 1

(b) Pair 2

(c) Pair 3

(d) Pair 4

(e) Pair 5

(f) Pair 6

Figure 4.3: Pairs of agriculture land images used as a test set, first column shows SIs, second column shows corresponding UAV images and last column shows manually registered SI-UAV images.

Affine transformation has 3 variations, Translation, Scale and Rotation. The following vector elements of this matrix

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

are used as scale and rotation matrix used in presence of scale and rotation.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

is translation matrix used for moving the points.

$$\begin{bmatrix} c_1 & c_2 \end{bmatrix}$$

is projection matrix. Projective transformations are frequently used to register images that are out of alignment. 2D Projective transformation is can be given as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ E & F & 1 \end{bmatrix}$$

where E and F are vanishing points. When E and F are equal to zero, the transformation becomes an affine.

The homography is used in multiple ways in this thesis. First it is used to show manually registered images in Figure 4.2 and Figure 4.3 and to depict the physical location of all UAV images in their corresponding SIs. Secondly, it is used with Overlap Error in the training step to identify corresponding features points between SI-UAV image pairs of the training set. Thirdly, it is used with Projection Error in the testing phase to compute ground truth data for image matching on SI-UAV image pairs of the test set. Finally, it is used in the quantitative analysis (Precision and Root Mean Square Error) for the proposed NN-BF method and also for comparison with the homography estimated by RANSAC in proposed NN-BF method for registration of SI-UAV image pairs of the test set. The given (i.e, ground truth) homography is denoted by *H* and the homography estimated by RANSAC with *K*.

4.6 Overlap Error

OE [88] measures how well two regions (i.e, image patches) around the feature points of reference (SI) and target (UAV) images correspond under a known homography H. It is computed with Equation 4.1 as:

$$OE = 1 - \frac{n \cap (H^T m H)}{n \cup (H^T m H)} \tag{4.1}$$

where n and m are regions around the feature point of reference and target images, respectively. $\cap$ and $\cup$ represents union and intersections between regions n and m.

n and m regions are considered as corresponding regions, if OE between them is less than or equal to a threshold. In the proposed NN-BF method, thresholds less than or equal to 15% and 25% is used to find the corresponding feature points between SI-UAV image pairs of the training set. Figure 4.4 illustrates OE between n and m, which are shown as green and red circles.

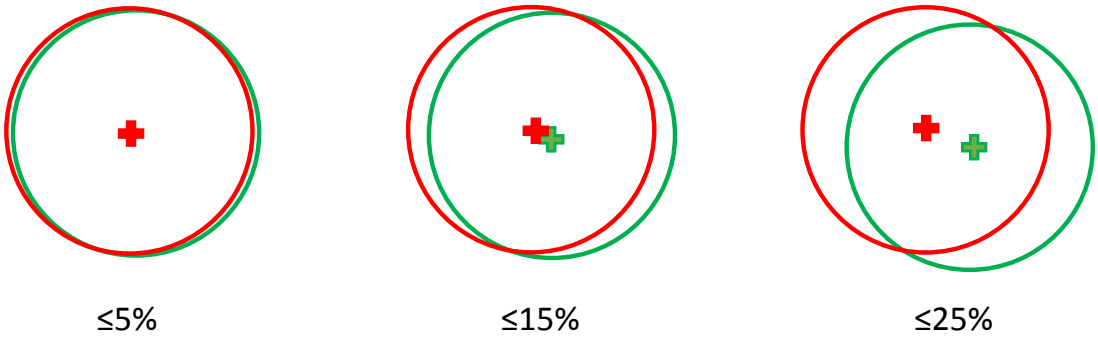


Figure 4.4: Three different cases are shown to illustrate Overlap Error. In each case, overlap error is shown below the regions and it is due to difference in size (scale) of the regions around the feature points, which are shown as '+'. The regions of reference (SI) image are shown in green circles, whereas the regions projected from target (UAV) image onto reference image is shown in red color circles.

4.7 Train Descriptors

These are the SIFT descriptors, which are computed for the corresponding feature points which are identified with OE between the SI-UAV image pairs of the training set, as explained above. The descriptor of corresponding feature points of SI is referred to as Train SI descriptors and in case of UAV they are called Train UAV

descriptors.

4.8 Test Descriptors

These are the SIFT descriptors, which are computed for SIFT feature points on each SI-UAV image pair of the test set. The descriptors of SI image in case of test set are referred to as Test SI descriptors and whereas the descriptors of test UAV image are called Test UAV descriptors.

4.9 Descriptor Matching

The proposed method is based on NN and BF descriptor matching strategies, therefore, descriptor matching is carried out in two ways: i) Nearest neighbor descriptor matching, ii) Brute force descriptor matching

4.9.1 Nearest Neighbor Descriptor Matching

Nearest neighbor matching takes a descriptor from one feature vector and find its best match in other vector using distance ratio. Lower the distance ratio, better will be the feature match and vice-versa. In this thesis, this descriptor matching is between the Train SI descriptors and the Test descriptors of SI of the test set. For this purpose, the SI-UAV image pairs of the test set is processed one by one for image registration. The nearest neighbor descriptor matching is one to one descriptor matching. For each Test SI descriptor, a nearest neighbor descriptor match is searched in Train SI descriptor. Having computed the nearest neighbor descriptors, the Test SI descriptors are replaced with their nearest neighbor Train SI descriptors. The nearest neighbor Train SI descriptors are further replaced with their corresponding Train UAV descriptors (which were identified with OE between SI-UAV image pairs of the train set during the training phase). The reason for this replacement is that direct descriptor matching between the Test SI descriptors and Test UAV descriptors is not possible due to high temporal and textual differences. By replacing the Test SI descriptors with their nearest neighbor Train SI descriptors, and then in turn replacing the nearest neighbor Train SI descriptors with their corresponding Train UAV descriptors leaves a descriptor matching between the Test UAV descriptors and the corresponding Train UAV descriptors. This descriptor matching is not difficult because the training and test UAV images were acquired on same date, time, illumination condition and using the same camera.

4.9.2 Brute Force Descriptor Matching

The Brute Force Descriptor Matcher (BFDM) is an approach which takes a descriptor of a feature from reference data vector and find its closest neighbor using

some distance threshold, by matching it with all feature descriptors of target data vector. In this way the reference data vector features would be integrated with all combinations of target data vector features. OpenCV implementation of Brute Force Descriptor Matcher (BFDM) [1] is used in this thesis for descriptor matching between the Test UAV descriptors and the corresponding Train UAV descriptors. This matching is also one to one descriptor matching. Please note that the descriptor part of the Test SI SIFT is replaced and the frame part of the SIFT is retained, which is used by RANSAC to estimate homography for registration of SI-UAV images of the test set. The frame part carries information about the pixel location of the Test SIFT feature points of SI-UAV test image pairs.

4.10 Random Sample Consensus

After the Brute Force (BF) descriptor matching, the matched feature points of test SI-UAV are processed with RANSAC [65] algorithm. The RANSAC estimates the homography and removes outliers. It also helps in the computation of the number of correct matches, false matches and the precision score to evaluate the performance of proposed NN-BF method.

4.11 Image Warping

Having computed the homography, the test UAV image of each SI-UAV image pair of the test set is geometrically aligned and registered with its corresponding test SI. The image warping is used in the visual inspection step to analyse the results obtained with the proposed method.

4.12 Root Mean Square Error

Root Mean Square Error (RMSE) [75] is computed using Equation 4.2 as follows:

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N} residual_n^2}{N}} \tag{4.2}$$

where $residual = \sqrt{(x_r - K(x_t))^2 + (y_r - K(y_t))^2}$, (x_r, y_r) and (x_t, y_t) are pixel coordinates of reference (SI) and target (UAV) images. K is a homography estimated by RANSAC between SI-UAV image pairs of the test set. N represents total number of pixel coordinates.

[1] https://docs.opencv.org/3.4/d3/da1/classcv_1_1BFMatcher.html

4.13 Summary

This chapter presents an overview of proposed NN-BF method. The proposed NN-BF method comprises of agriculture land image dataset divided into two disjoint sets. On the basis of experimental results, SIFT algorithm is used in proposed NN-BF method with OE. Transformation matrix known as Homography is used in four different ways in the proposed method. The proposed NN-BF method use NN and BF descriptor matching strategies to match the train and test descriptors of SI-UAV images. Finally, proposed NN-BF method uses RANSAC algorithm for removal of outliers and mismatches.

CHAPTER 5

EXPERIMENTAL SETUP AND RESULTS

5.1 Overview

This chapter presents experimental setup and results. The experimental results are obtained on SI-UAV image pairs of the test set as shown in Figure 4.3. The experimental setup section describes the evaluation benchmark and results obtained through experiments are provided in experimental results section.

5.2 Experimental Setup

The experimental setup of proposed NN-BF method consists of projection error and evaluation criteria which describes the benchmark and performance metric for evaluation of feature point algorithms used in the experiments on SI-UAV images.

5.2.1 Projection Error

Projection Error (PE) is Euclidian distance between the features point of SI and the projected feature points of UAV image. The projected feature points are obtained by projecting the UAV feature points onto SI with the ground truth homography 'H'. PE is computed in Equation 5.1 as follows:

$$PE = \|Hx_R - x_T\| \tag{5.1}$$

where $x_R = (x_r, y_r)$ and $x_T = (x_t, y_t)$ are pixel locations of feature points of SI (reference) and UAV (target) images, respectively. PE less than or equal to $0.5, 1.0, 1.5, 2.0, 2.5, 3.0(pixels)$ is used in this thesis. Please note, PE is only used in the testing phase to compute number of corresponding feature points, correct descriptor matches, false descriptor matches, repeatability score and matching score for image matching on SI-UAV image pairs of the test set. This is because the homography *'H'* is considered to be known between the images for image matching [88, 40, 97]. The PE has no role in image registration. The correct and false matches in case of image registration is identified with RANSAC. Projection error is illustrated in 5.1.

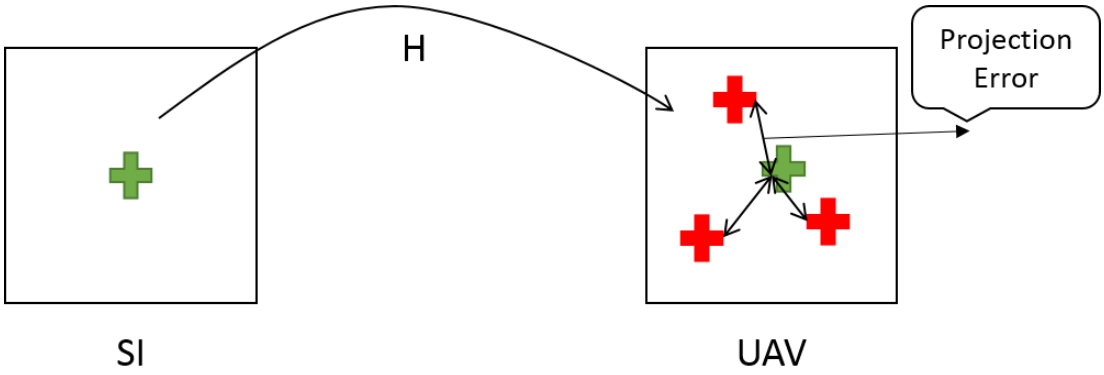


Figure 5.1: Illustration of Projection Error

5.2.2 Evaluation Criteria

This thesis uses the following performance metrics [88]. These metrics are computed for image matching using different PE thresholds.

i. Repeatability metric is used to evaluate the performance of feature point detectors. It is defined in Equation 5.2 as:

$$repeatability = 100 \times \frac{\#\ corresponding\ feature\ points}{minFeature} \tag{5.2}$$

where $minFeature$ are minimum number of feature points in pair of images i.e SI-UAV and $\#correspondingfeaturepoints$ denotes the number of corresponding features between SI-UAV images.

ii. Matching score is used for performance evaluation of feature point descriptors and defined in Equation 5.3 as:

$$matchScore = 100 \times \frac{\#\ correct\ matches}{minFeature} \tag{5.3}$$

where $\#correctmatches$ denotes the correct number of matched feature points.

iii. Precision score is a ratio between correct number of matches and total number of matches (correct and false matches) and defined in Equation 5.4 as:

$$precision = 100 \times \frac{\#\ correct\ matches}{\#\ correct\ matches + \#\ false\ matches} \tag{5.4}$$

where $\#correctmatches$ are correctly matched and $\#falsematches$ are

outliers or mismatched feature points. Precision score is also used to evaluate the performance of feature point descriptors.

5.3 Experimental Results

The experimental results are divided into three parts (i) Image Matching to evaluate the performance of different feature points on SI-UAV image pairs (ii) Image Matching with the proposed NN-BF method (iii) Image Registration with the proposed NN-BF method.

5.3.1 Image Matching

Image matching on SI-UAV images are carried out with the help of protocols described in [88] and [46]. Image matching is widely used as a test problem to evaluate the performance of feature points [15, 40, 97]. Image matching is also used to evaluate the performance of the proposed NN-BF method. Additionally, the type of feature point used in the proposed NN-BF method is SIFT, which is then replaced with other types of feature points such as SURF, ORB, BRISK etc., to check how the performance of the proposed method vary with respect to types of feature points. The performance metric used for evaluation are repeatability score (%), matching score (%) and precision score (%). These performance measures are computed for different PE thresholds as described in [46] and Section 5.2.1.

Table 5.1 shows the comparison of feature points based on number of correspondence and Table 5.2 depicts the comparison based on repeatability score (%). The detector used are SIFT, SURF, ORB, and BRISK. The correspondence and repeatability score are obtained with PE $\leq$ 2.5 pixels. The results show that ORB on average demonstrates the best number of correspondence and repeatability score and outperforms all other detectors. The best performance of ORB is due to a fact that it detects feature points that are closely packed and having very small inter feature distances. The last column represents the mean value which is computed over all the SI-UAV image pairs of test set.

Table 5.1: Number of correspondence obtained by feature points on Test SI-UAV image pairs with PE $\leq$ 2.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	772	589	727	686	862	802	740.0
SURF	1326	936	893	1070	1120	1094	1073.0
ORB	**4797**	**6174**	**5662**	**5633**	**6028**	**5235**	**5588.0**
BRISK	341	263	299	345	360	336	324.0

Figure 5.2 show the avearge number of correspondence and Figure 5.3 show the average repeatability score with respect to increasing PE threshold values. The figure shows the average value computed over all the six SI-UAV image pairs of the test set. It can be seen that ORB outperforms all detectors. SURF achieves the second best performance followed by SIFT and BRISK. It is evident from the figure that repeatability score increases with the increase in the PE values.

Table 5.2: Repeatability scores (%) obtained by feature points on Test SI-UAV image pairs with PE $\leq$ 2.5 pixels

Method	**Pair 1**	**Pair 2**	**Pair 3**	**Pair 4**	**Pair 5**	**Pair 6**	**Mean**
SIFT	41.5	41.4	46.2	34.6	46.6	37.4	41.3
SURF	36.8	59.4	51.7	47.6	54.8	45.1	49.2
ORB	**128.6**	**184.0**	**160.1**	**152.1**	**167.5**	**140.6**	**155.5**
BRISK	25.8	32.3	31.6	30.8	36.1	28.7	30.9

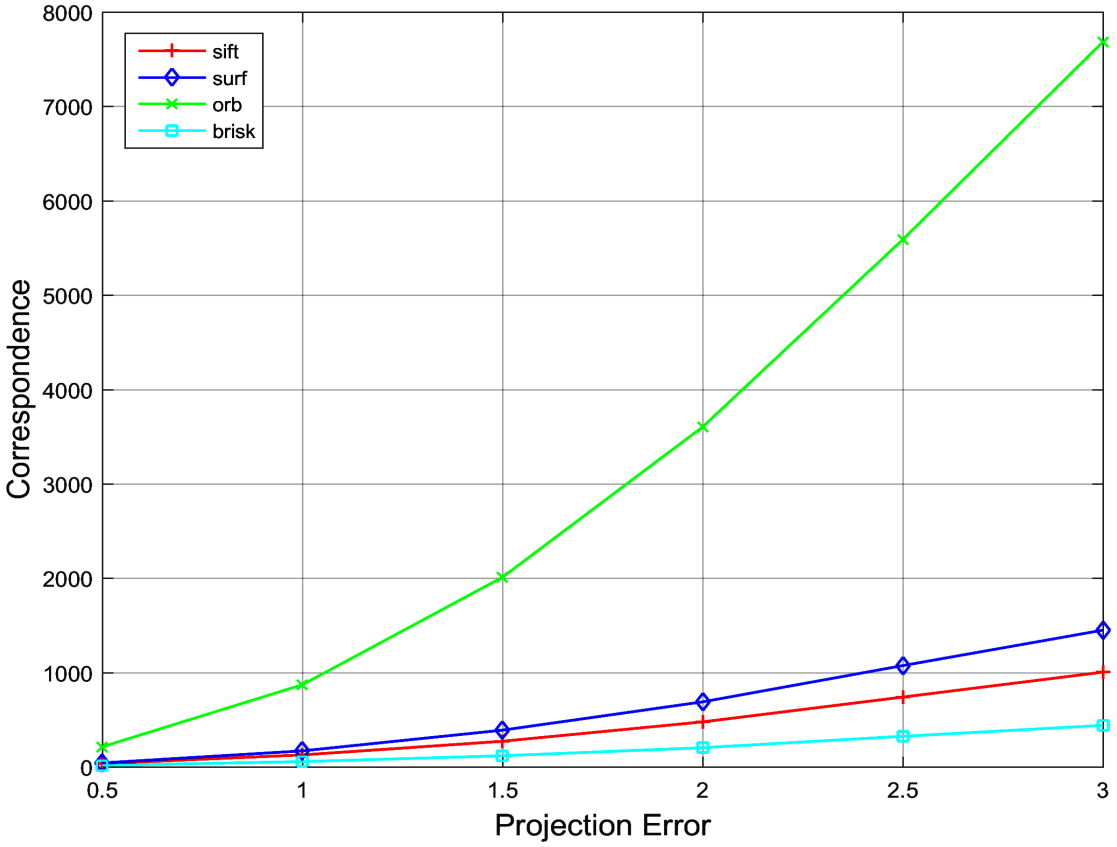


Figure 5.2: Comparison of feature points based on average number of correspondence for image matching between the SI-UAV image pairs of test set

Table 5.3 shows a comparison of feature points based on matching score (%). PE $\leq$ 2.5 pixels is used. DAISY, LHOPC and RIFT have only descriptor part (see Table 1.1). Their detector part is provided by SIFT detector, means SIFT+LHOPC denotes that SIFT is detector and LHOPC is descriptor.

The matching results show that all feature points demonstrate extremely low matching scores. SIFT demonstrate zero matching score due to high textural, intensity and temporal differences between test SI-UAV images. SURF achieves on average the best performance but it is also too low i.e., 0.05%. Figure 5.4 shows the average matching score (%) obtained by feature points for different PEs.

Table 5.4 shows the average precision score (%) based comparison between the feature points with PE $\leq$ 2.5 pixels whereas Figure 5.5 shows the precision score obtained with different PEs.

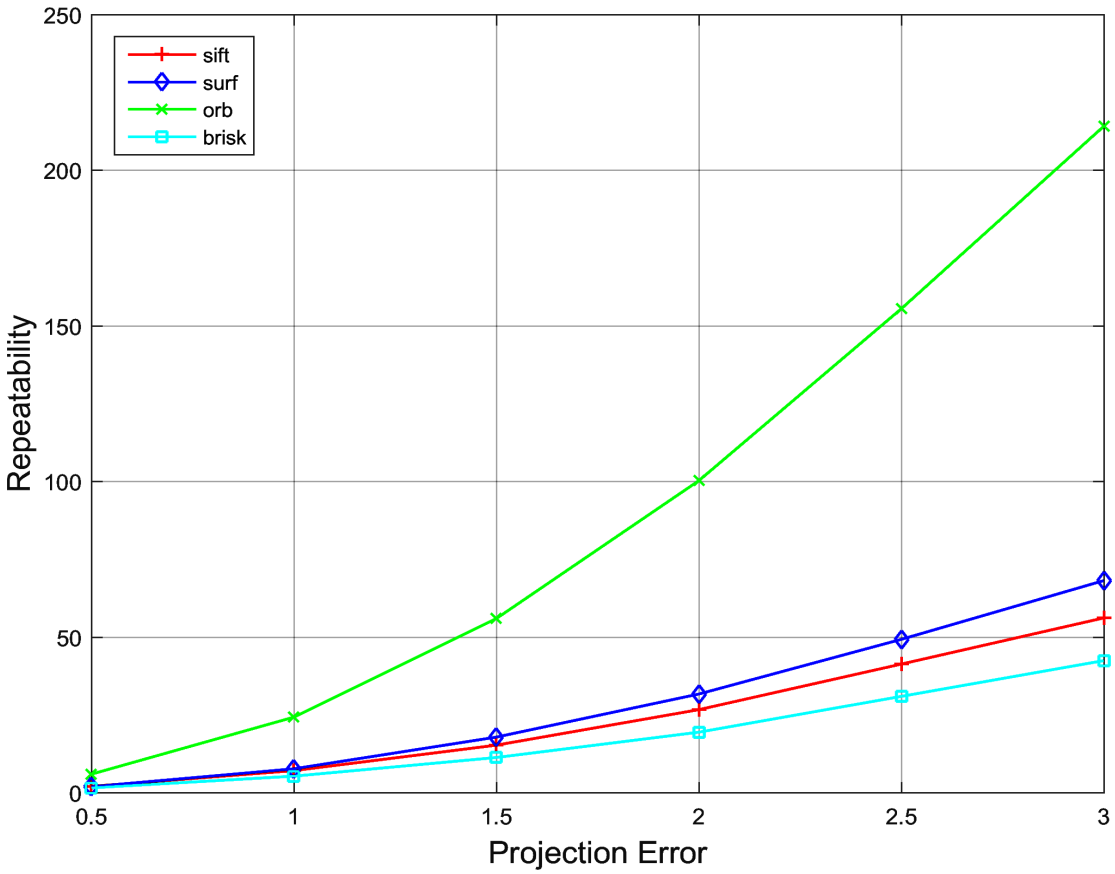


Figure 5.3: Comparison of feature points based on average number of correspondence for image matching between the SI-UAV image pairs of test set

Table 5.3: Matching score (%) obtained by feature points on Test SI-UAV image pairs with PE $\leq$ 2.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SURF	**0.11**	**0.06**	0.00	0.04	0.00	**0.08**	**0.05**
ORB	0.03	0.03	0.03	0.03	0.00	0.00	0.02
BRISK	0.00	0.00	0.11	0.00	0.00	**0.08**	0.03
DAISY	0.00	0.00	0.00	**0.05**	0.00	0.00	0.01
LHOPC	0.00	0.00	0.00	0.00	**0.11**	0.00	0.02
RIFT	0.00	0.00	**0.13**	0.00	0.05	0.00	0.03

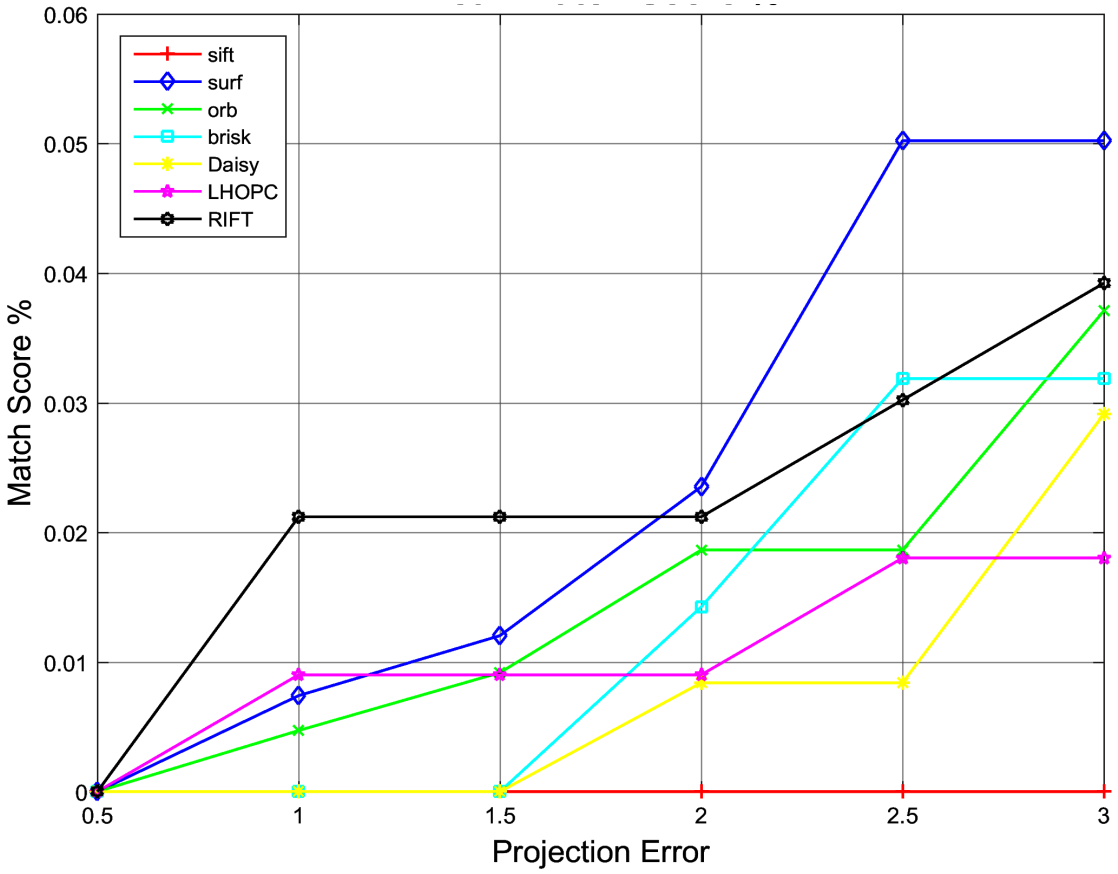


Figure 5.4: Comparison of feature points based on average matching score (%) for image matching between the SI-UAV image pairs of test set

Table 5.4: Precision score (%) obtained by feature points on Test SI-UAV image pairs with PE $\leq$ 2.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SURF	**0.21**	**0.09**	0.00	0.08	0.00	0.15	**0.09**
ORB	0.06	0.08	0.07	0.07	0.00	0.00	0.05
BRISK	0.00	0.00	0.26	0.00	0.00	**0.16**	0.07
DAISY	0.00	0.00	0.00	**0.13**	0.00	0.00	0.02
LHOPC	0.00	0.00	0.00	0.00	**0.30**	0.00	0.05
RIFT	0.00	0.00	**0.30**	0.00	0.13	0.00	0.07

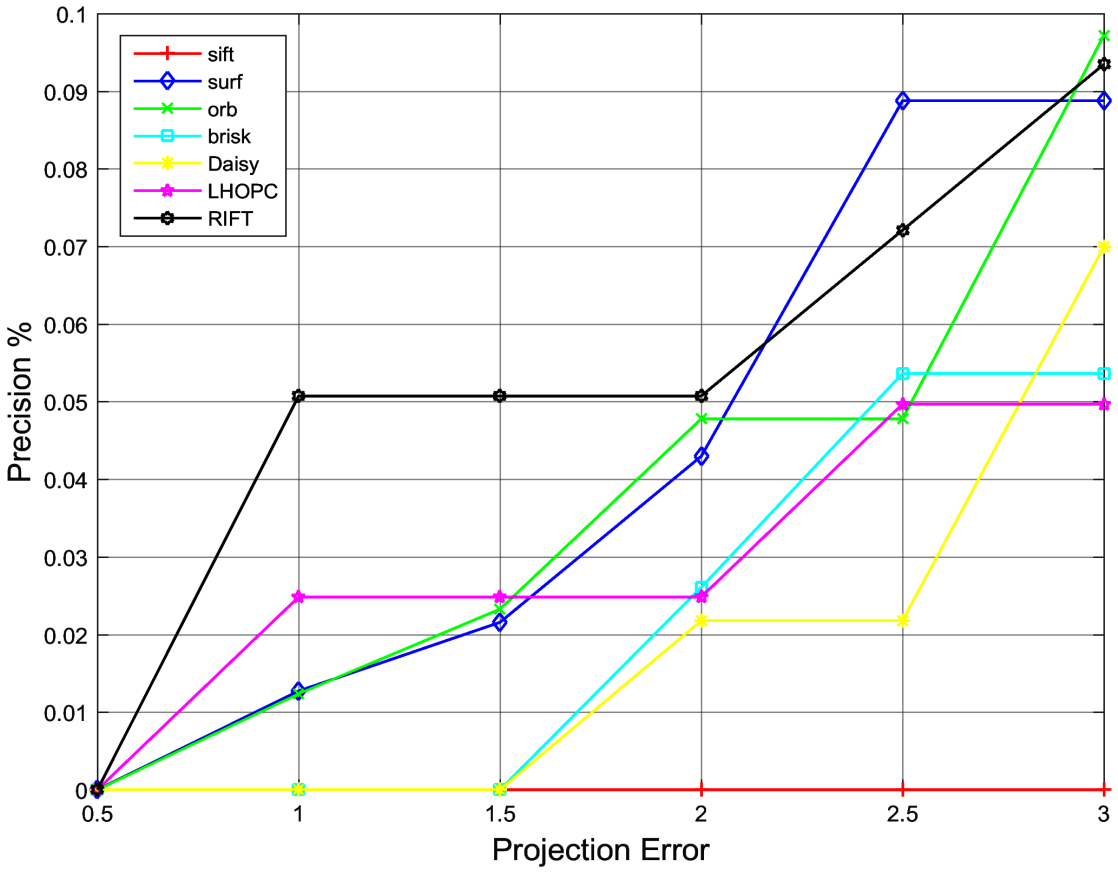


Figure 5.5: Comparison of feature points based on average precision score (%) for image matching between the SI-UAV image pairs of test set

5.3.2 Image Matching with Proposed NN-BF Method

This section presents the image matching results by using the proposed NN-BF method. Last section shows that feature points under perform in image to matching task on test SI-UAV images. This section shows that the proposed method significantly improves the performance of feature points. The experimental results shown in this section are obtained by using OE $\leq$ 15% and OE $\leq$ 25% in the proposed method (see Section 4.6).

Table 5.5 shows the image matching results based on matching score (%) obtained by using different feature points in the proposed method with OE $\leq$ 15%. The PE used is less than or equal to 1.5 pixels. It can be seen that SIFT obtains on average a matching score of 20.4% and outperforms all other feature points, SURF obtains the second best performance followed by ORB and BRISK. DAISY, LHOPC and RIFT do not perform well compared to SIFT and SURF.

Table 5.5: Matching score (%) obtained with proposed NN-BF method, OE $\leq$ 15% and PE $\leq$ 1.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	**34.5**	**19.6**	17.6	**24.5**	0.0	26.3	**20.4**
SURF	30.6	7.0	**25.0**	15.4	1.3	**37.5**	19.5
ORB	18.5	4.6	6.8	16.2	**2.8**	13.6	10.4
BRISK	8.5	2.4	6.4	7.8	0.0	8.8	5.7
DAISY	9.4	4.0	5.2	4.7	0.0	5.3	4.8
LHOPC	9.9	3.4	2.5	2.2	0.9	6.4	4.2
RIFT	13.1	2.2	1.1	5.0	0.0	2.2	3.9

Table 5.6 shows the image matching results based on precision score (%) obtained by different feature points in the proposed NN-BF method with OE $\leq$ 15%. SIFT obtains on average a precision score of 32% and outperforms all other feature points, ORB obtains the second best performance followed by SURF and BRISK. DAISY, LHOPC and RIFT demonstrates low precision score (%).

Figure 5.6 shows the average matching score (%) obtained by feature points with proposed method and using OE $\leq$ 15% values and different PE values. Figure 5.7 shows the average precision score (%) obtained by feature points with proposed method and using OE $\leq$ 15% values and different PE values.

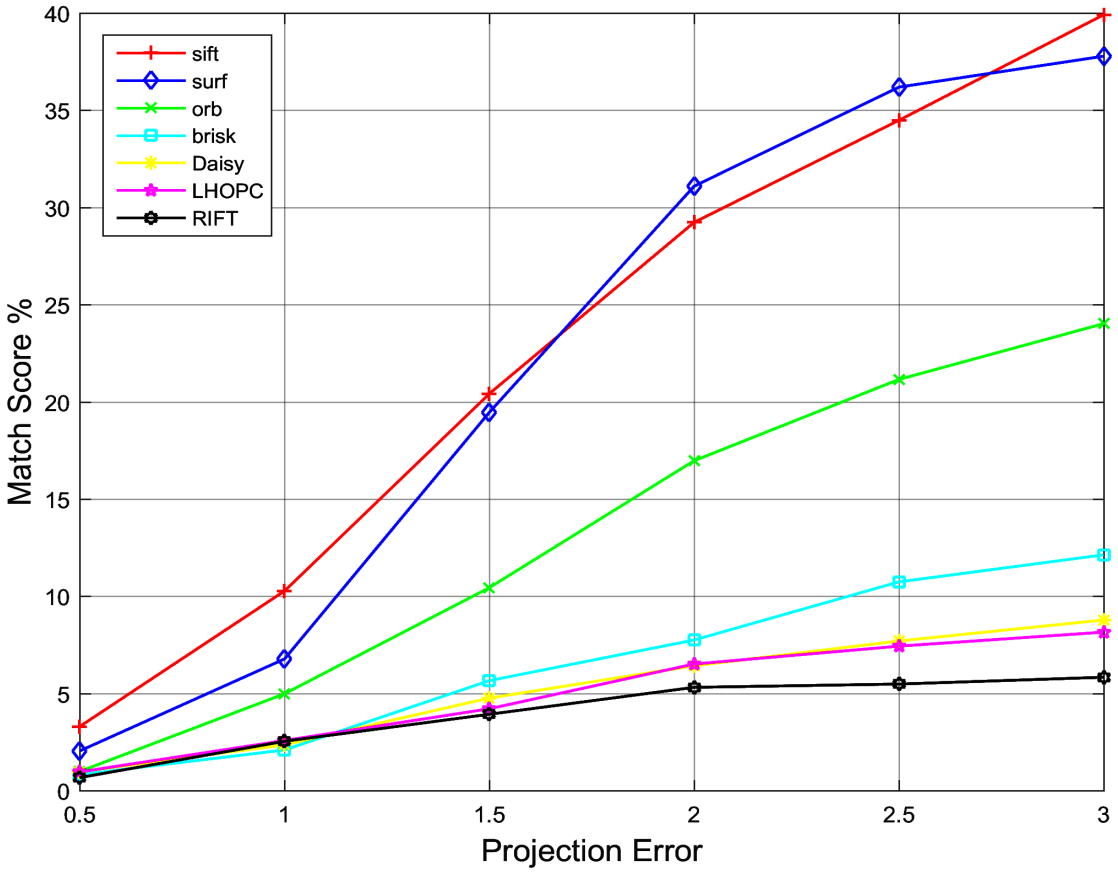


Figure 5.6: Average matching score (%) obtained with proposed NN-BF method using OE $\leq$ 15%

Table 5.6: Precision score (%) obtained with proposed NN-BF method, OE $\leq$ 15% and PE $\leq$ 1.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	**52.7**	**32.1**	26.1	**37.5**	0.0	**43.5**	**32.0**
SURF	35.8	9.1	**26.9**	17.7	1.4	40.0	21.8
ORB	52.6	9.5	16.7	36.0	**6.2**	35.3	26.0
BRISK	9.5	2.8	7.5	8.9	0.0	8.8	6.3
DAISY	17.2	7.1	9.9	8.3	0.0	9.5	8.7
DAISY	16.2	5.4	4.8	4.3	1.4	10.9	7.2
RIFT	21.3	3.8	2.3	9.3	0.0	4.4	6.8

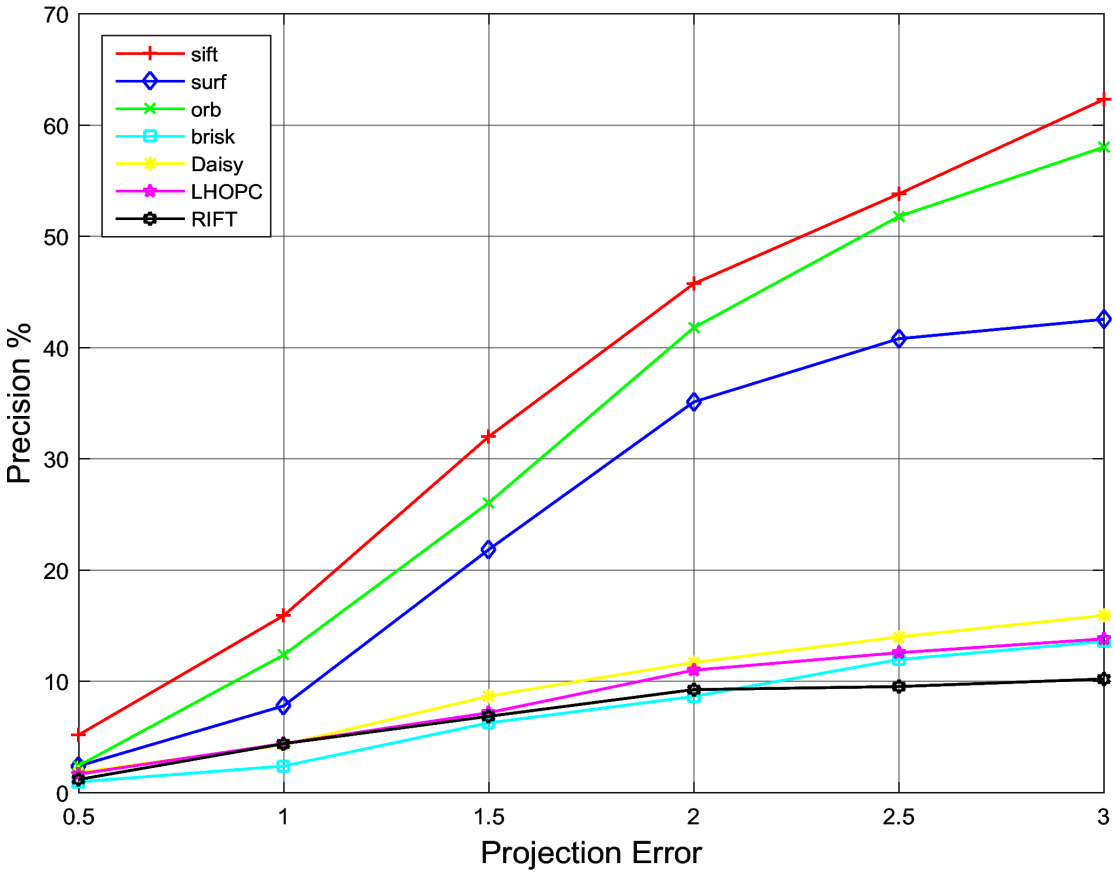


Figure 5.7: Average precision score (%) obtained with proposed NN-BF method using OE $\leq$ 15%

Table 5.7 shows the image matching results based on matching score (%) obtained by using different feature points in the proposed NN-BF method with OE $\leq$ 25%. It is evident that SIFT obtains on average a matching score of 29.1% and outperforms all other feature points, SURF obtains the second best performance followed by ORB. The average matching score (%) results shown in Figure 5.8 are obtained with proposed methods and OE $\leq$ 25%.

Table 5.8 shows the image matching results based on precision score (%) obtained by different feature points in the proposed method with OE $\leq$ 25%. SIFT obtains on average a precision score of 46.5% and outperforms all other feature points, ORB obtains the second best performance (43.0%) followed by SURF (29.8%). Figure 5.9 shows the average precision score (%) obtained by feature points with proposed method and using OE $\leq$ 25% values and different PE values.

Table 5.7: Matching score (%) obtained with proposed NN-BF method, OE $\leq$ 25% and PE $\leq$ 1.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	**37.7**	**27.4**	**33.3**	**34.2**	4.0	37.7	**29.1**
SURF	29.4	13.7	25.0	26.6	**5.0**	**38.0**	23.0
ORB	13.5	7.7	13.4	16.8	4.4	18.6	12.4
BRISK	17.9	7.8	6.0	2.4	3.1	7.4	7.5
DAISY	11.5	6.1	6.6	4.9	0.6	5.9	5.9
LHOPC	16.1	7.1	3.6	2.6	0.8	3.9	5.7
RIFT	19.4	9.2	2.4	4.0	0.6	0.0	5.9

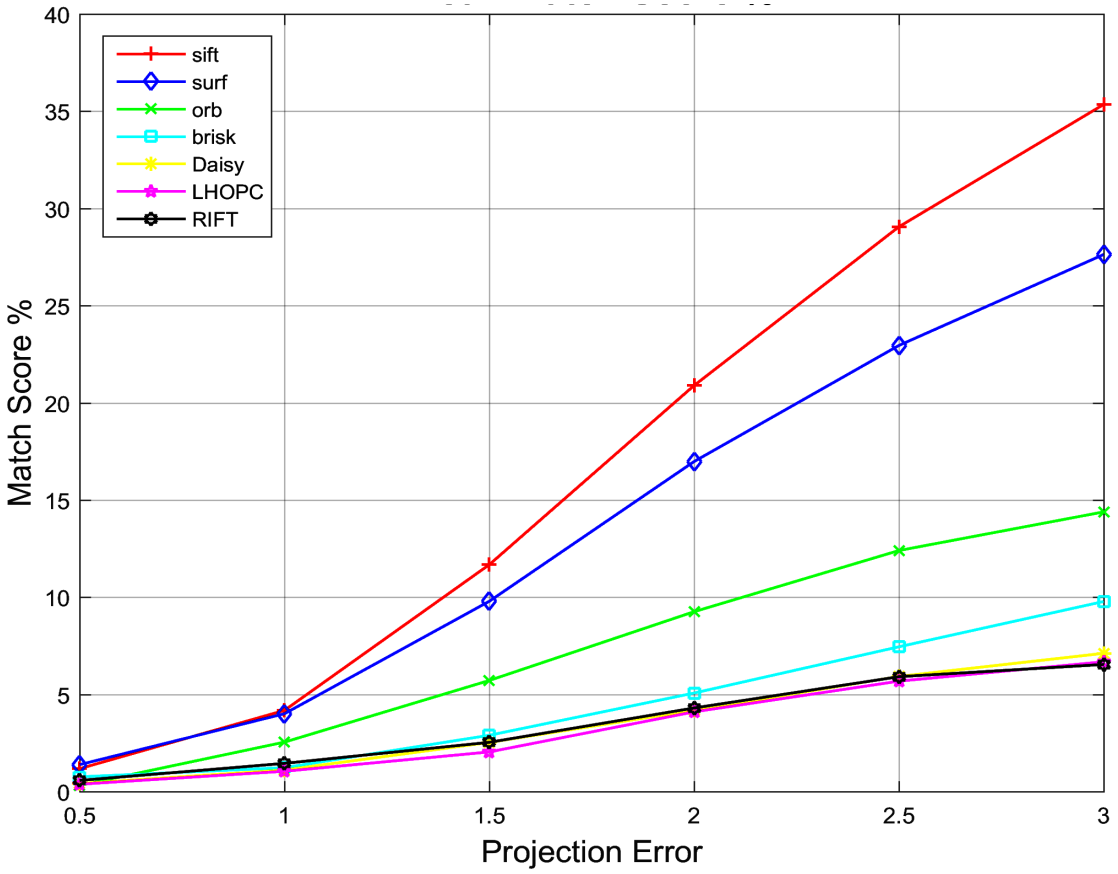


Figure 5.8: Average matching score (%) obtained with proposed NN-BF method using OE $\leq$ 25%

Table 5.8: Precision score (%) obtained with proposed NN-BF method, OE $\leq$ 25% and PE $\leq$ 1.5 pixels

Method	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Mean
SIFT	58.4	**39.0**	**52.2**	55.9	6.2	**67.4**	**46.5**
SURF	41.6	19.7	33.3	32.4	6.1	45.8	29.8
ORB	**67.2**	23.9	39.7	**58.0**	**12.1**	57.1	43.0
BRISK	24.1	8.5	7.6	4.8	3.7	8.4	9.5
DAISY	25.3	13.1	15.8	10.8	1.3	12.9	13.2
DAISY	31.4	14.6	10.0	6.8	1.9	9.6	12.4
RIFT	38.5	20.7	6.8	10.6	1.2	0.0	13.0

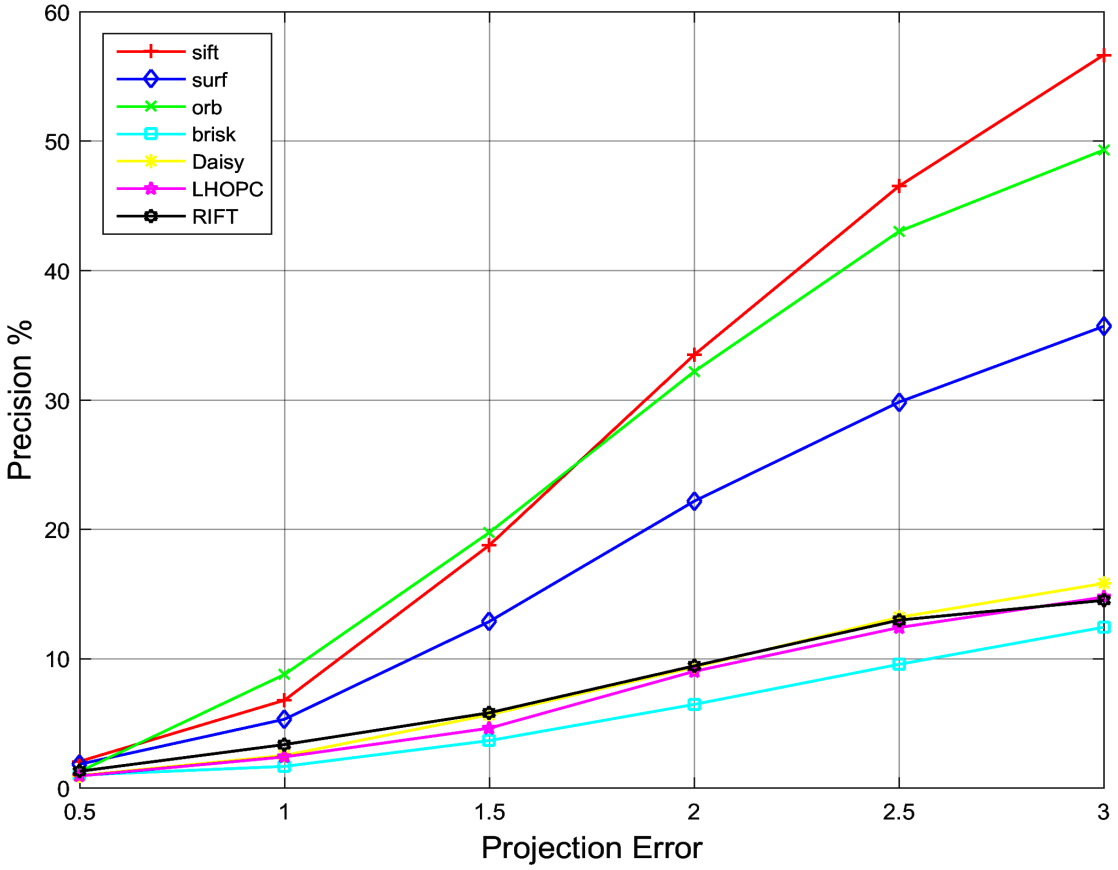


Figure 5.9: Average precision score (%) obtained with proposed NN-BF method using OE $\leq$ 25%

Table 5.9 and Table 5.10 summaries the result of image matching and provides a comparison of results using and without using the proposed NN-BF method. OE

equal to or less than 15% and 25% is used. It can be seen that SIFT outperforms all other feature points, that's why we use SIFT as feature points in the proposed method. The comparison also shows that results obtained with proposed NN-BF method are much better than the results obtained without the proposed method.

Table 5.9: Matching score (%) obtained with and without proposed NN-BF method using OE $\leq$ 15% and 25%

Method	Overlap Error $\leq$ 15%		Overlap Error $\leq$ 25%	
	Without Proposed Method	With Proposed Method	Without Proposed Method	With Proposed Method
SIFT	0.00	20.42	0.00	11.68
SURF	0.01	19.46	0.05	9.79
ORB	0.01	10.44	0.02	6.14
BRISK	0.00	5.66	0.03	3.97
DAISY	0.00	4.76	0.01	3.82
LHOPC	0.01	4.21	0.02	3.64
RIFT	0.02	3.93	0.03	3.83

Table 5.10: Precision score (%) obtained with and without proposed NN-BF method using OE $\leq$ 15% and 25%

Method	Overlap Error $\leq$ 15%		Overlap Error $\leq$ 25%	
	Without Proposed Method	With Proposed Method	Without Proposed Method	With Proposed Method
SIFT	0.00	31.97	0.00	19.51
SURF	0.02	21.82	0.09	12.83
ORB	0.02	26.03	0.05	20.00
BRISK	0.00	6.25	0.07	4.57
DAISY	0.00	8.67	0.02	6.18
LHOPC	0.02	7.17	0.05	5.39
RIFT	0.05	6.84	0.07	6.19

5.4 Image Registration with Proposed NN-BF Method

This section presents the image registration results using the proposed NN-BF method. Different feature points are used in the proposed method such as SIFT, SURF, ORB. The image registration results are demonstrated:

i. Quantitative Analysis

ii. Visual Inspection

5.4.1 Quantitative Analysis

The quantitative analysis is based on average correct matches, false matches, precision score and RMSE. These scores are obtained with the RANSAC algorithm. The results are based on OE $\leq$ 15% and 25% in the proposed NN-BF method. The best three feature points identified in the previous section are only used in the quantitative analysis which are SIFT, SURF and ORB. Table 5.11 and 5.12 shows the results obtained using proposed NN-BF method with OE $\leq$ 15% and 25% respectively.

Table 5.11: RANSAC based image registration results obtained with proposed method and OE$\leq$15%

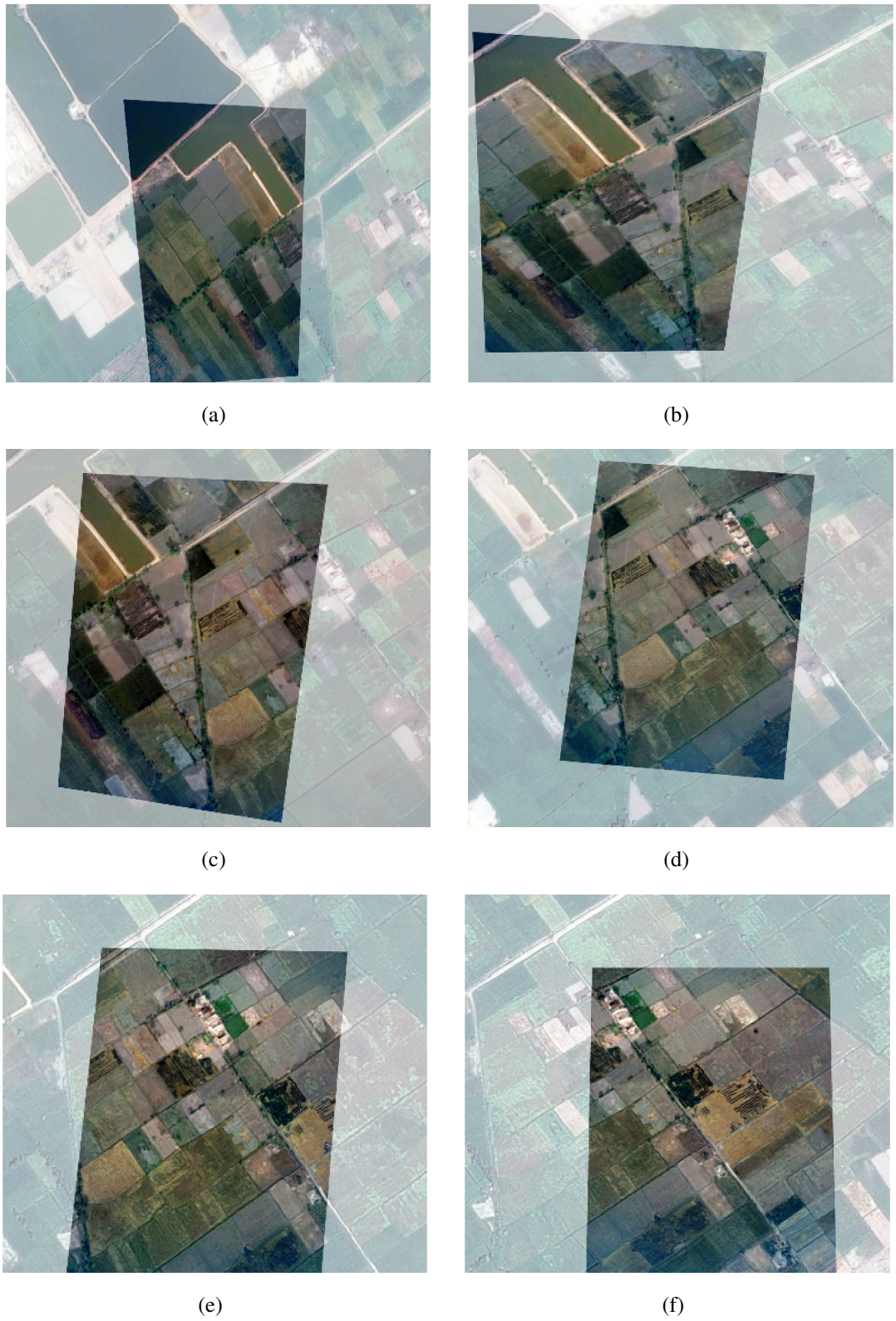
Method	**Avg. Correct**	**Avg. False**	**Precision (%)**	**RMSE**
SIFT	72.79	27.21	72.83	1.3497
SURF	59.51	40.49	66.82	1.5497
ORB	50.25	49.75	53.92	1.6869

Table 5.12: RANSAC based image registration results obtained with proposed method and OE$\leq$25%

Method	**Avg. Correct**	**Avg. False**	**Precision (%)**	**RMSE**
SIFT	50.95	49.05	51.88	1.3754
SURF	53.97	46.03	59.57	1.5987
ORB	34.12	65.88	40.87	1.8718

5.4.2 Visual Inspection

In this section the visual inspection based results are shown. In each result the UAV image is shown geometrically aligned and superimposed on corresponding SI. The superimposed images are obtained by using the homography estimated by RANSAC algorithm in the proposed NN-BF method. The farmland boundaries can be used as features to check whether the superimposed UAV image is perfectly aligned with the SI or not. Figure 5.10 and 5.11 shows SIFT based image registration results using proposed method and OE $\leq$ 15% and 25% respectively. Figure 5.12 and 5.13 shows SURF based image registration results using proposed method and OE $\leq$ 15% and 25% respectively. Figure 5.14 and 5.15 shows ORB based image registration results using proposed method and OE $\leq$ 15% and 25% respectively.

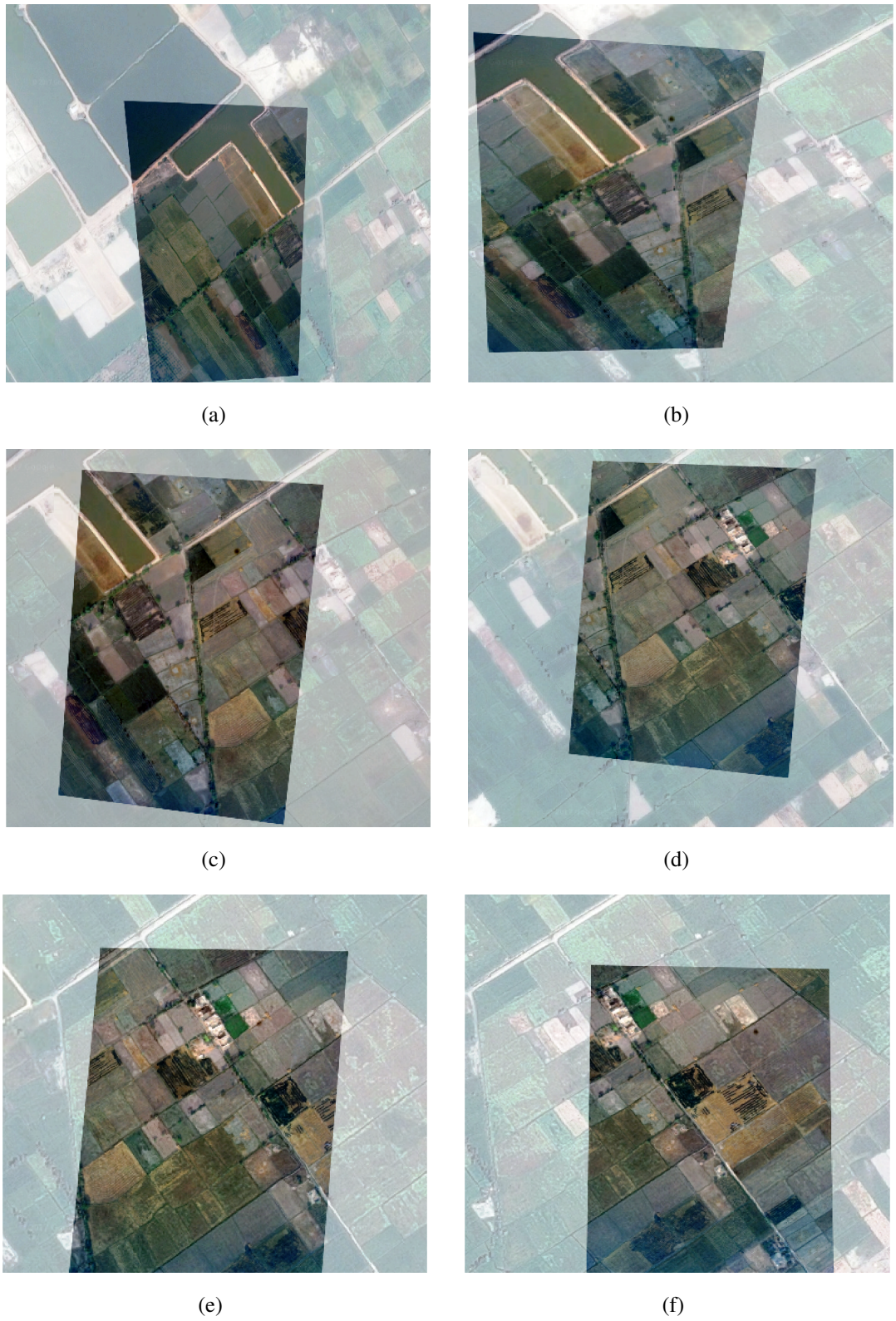


(a) (b)

(c) (d)

(e) (f)

Figure 5.10: Registered images obtained by using SIFT features in the proposed NN-BF method with OE $\leq$ 15%

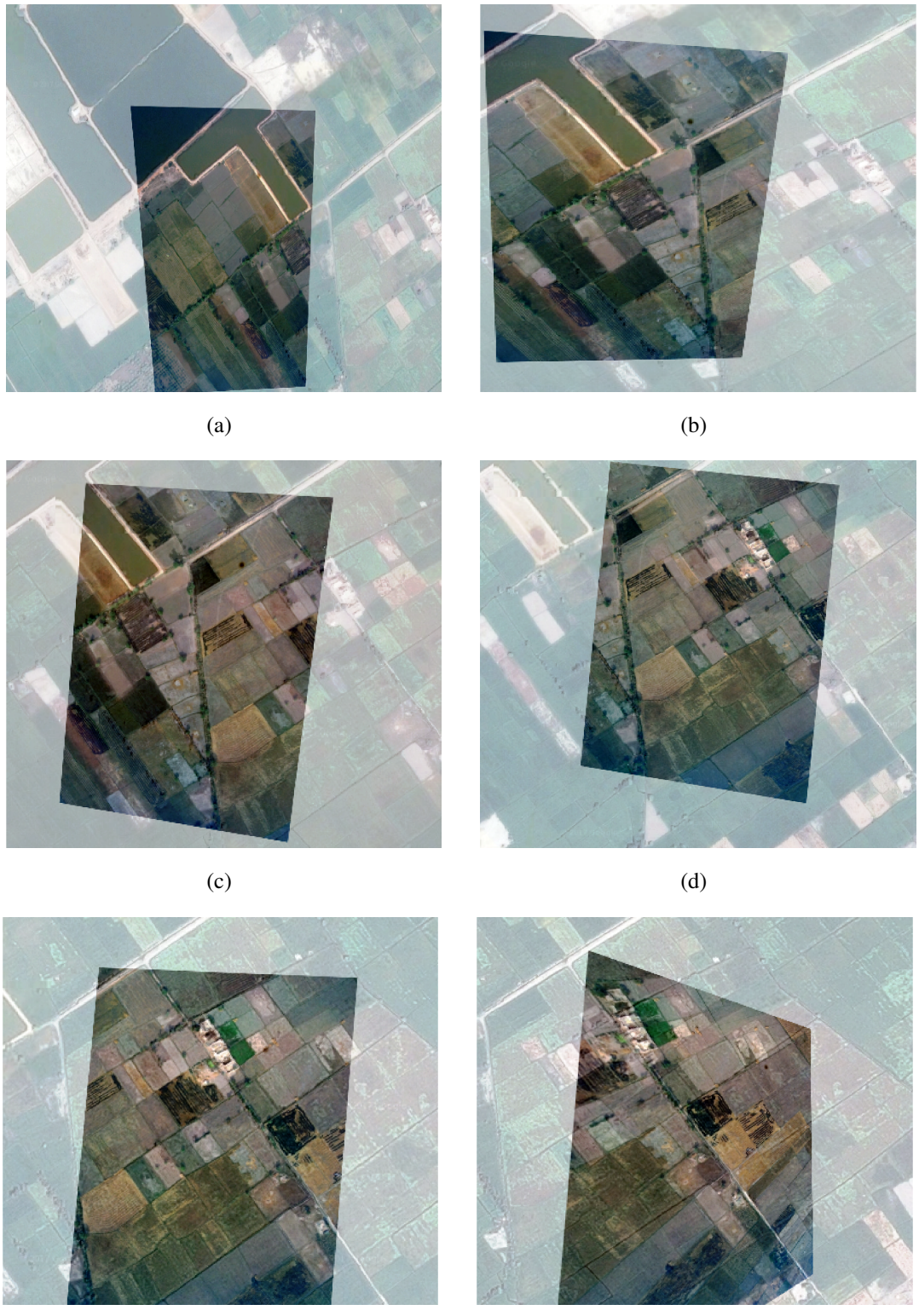


(a) (b)

(c) (d)

(e) (f)

Figure 5.11: Registered images obtained by using SIFT features in the proposed NN-BF method with OE $\leq$ 25%

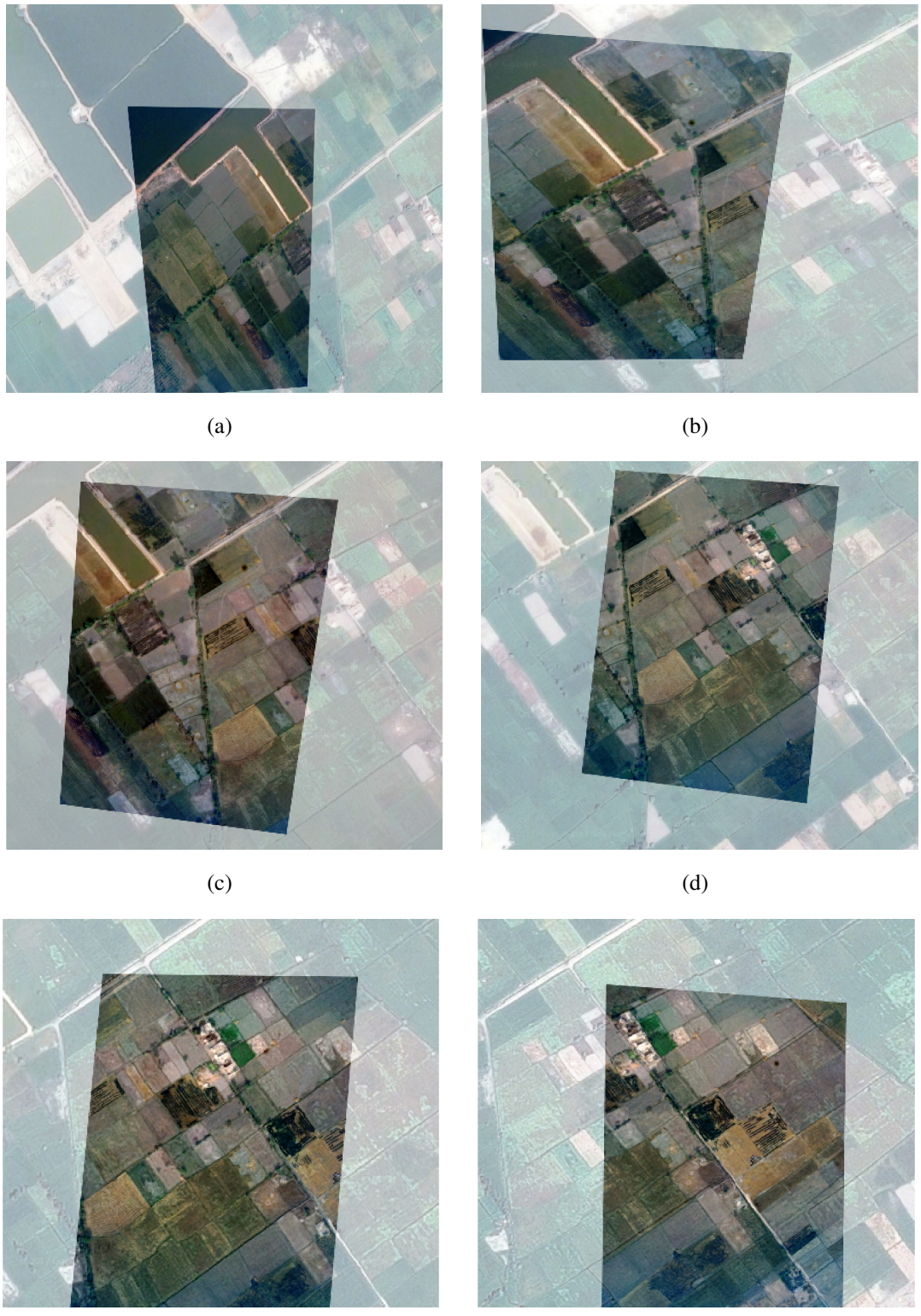


(a) (b)

(c) (d)

(e) (f)

Figure 5.12: Registered images obtained by using SURF features in the proposed NN-BF method with OE $\leq$ 15%

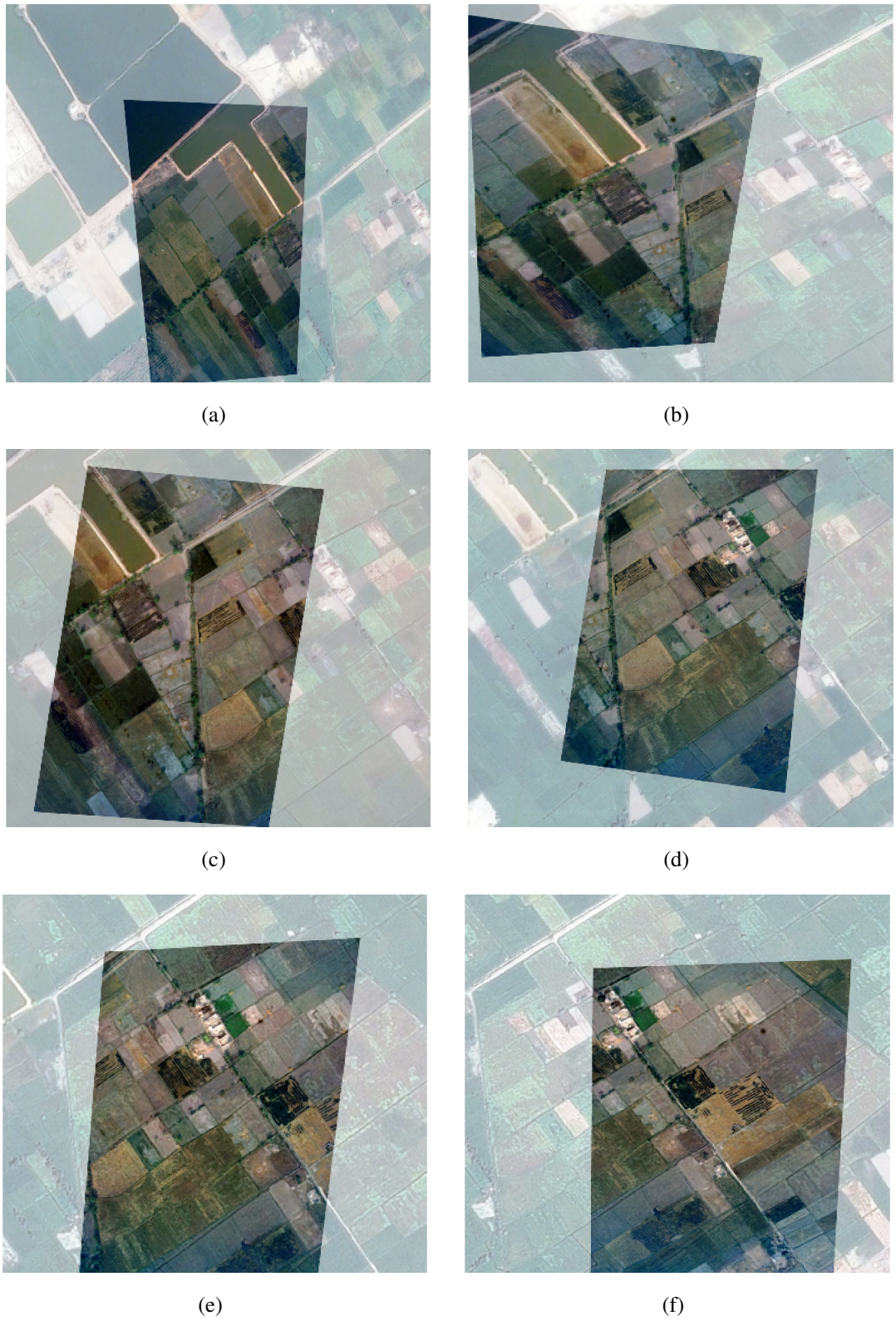


(a) (b)

(c) (d)

(e) (f)

Figure 5.13: Registered images obtained by using SURF features in the proposed NN-BF method with OE $\leq$ 25%

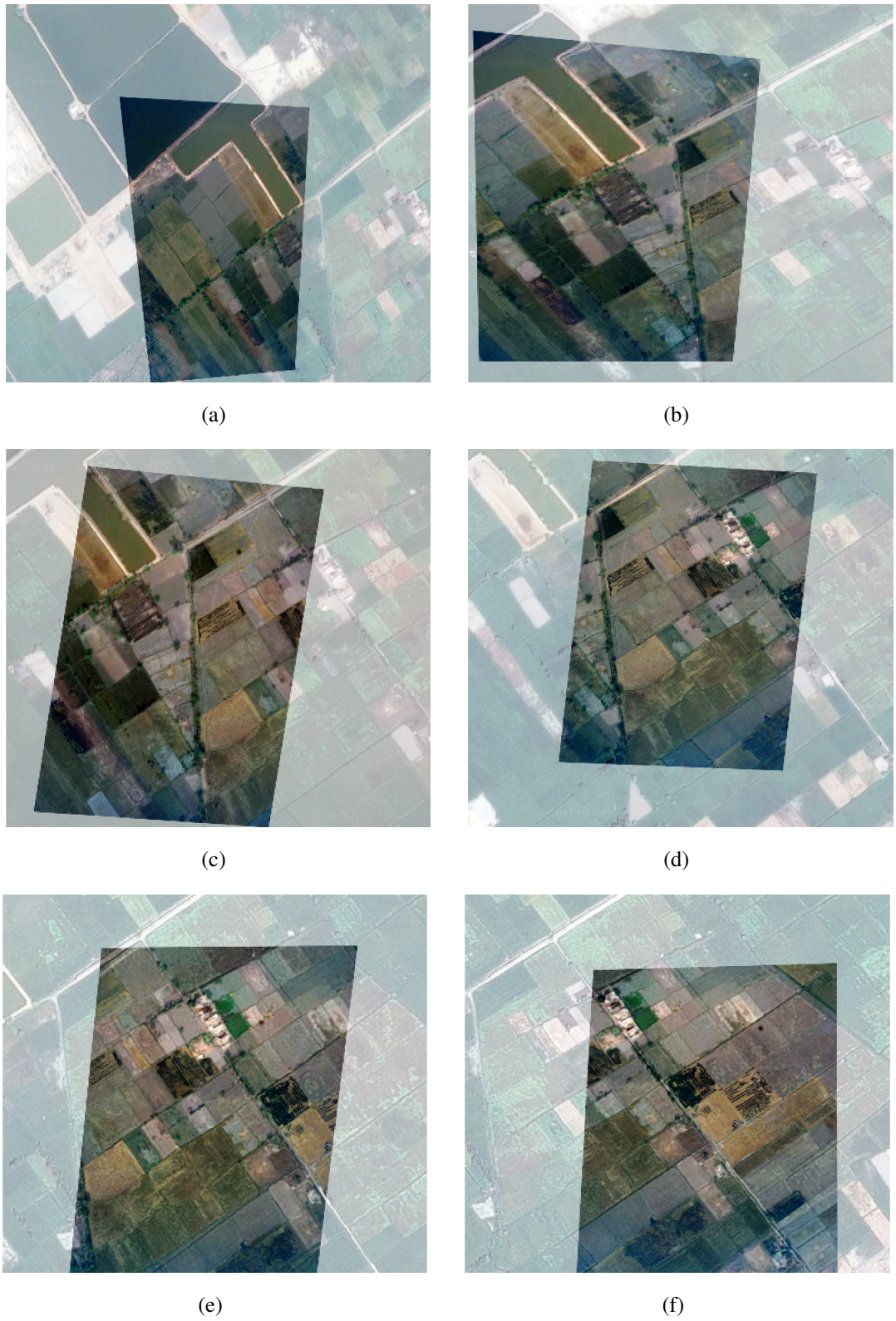


(a) (b)

(c) (d)

(e) (f)

Figure 5.14: Registered images obtained by using ORB features in the proposed NN-BF method with OE $\leq$ 15%



(a) (b)

(c) (d)

(e) (f)

Figure 5.15: Registered images obtained by using ORB features in the proposed NN-BF method with OE $\leq$ 25%

5.5 Summary

On the basis of the results demonstrated above, it can be seen that SIFT outperforms all other feature points by achieving the best average number of correct matches, precision score (%) and low RMSE. SURF obtains the second best performance followed by ORB. SIFT is capable of handling the image rotation, affine transformations, intensity, and viewpoint changes during the features matching. While SI and UAV images of agriculture land possess high temporal and textural differences. Using training phase and strengths of NN and BF matching strategies, SIFT outperforms other feature point algorithms. Best results are obtained using OE $\leq$ 15%. SIFT achieves 72.79% average correct matches compared to SURF (59.51%) and ORB (50.25%). SIFT also demonstrate on average 6% and 18.91% better precision (%) compared to SURF and ORB. However, in case of OE $\leq$ 25%, SURF has better average number of correct matches and precision (%) than SIFT and ORB. But overall scores are low compared to scores of OE $\leq$ 15%.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Overview

The regularly developing significance of SI in applications such as border monitoring, traffic studies, agriculture studies, damage assessment in case of disasters and many more, calls for persistent research work and innovations in the remote sensing area. In general, images from different sensors are utilized in different applications of remote sensing depends upon the application requirements. Therefore, utilization of techniques related to remote sensing for different applications has achieved a colossal lift. Feature-based techniques are widely used for processing of satellite images. Similarly, for data, relating to agricultural studies, regularly updated imagery is required to monitor the changes occur in agricultural areas during the course of time. This chapter contains a summary, conclusions and buildup of a point of view for the future research.

6.2 Conclusion

In this thesis, a new method for registration of SI-UAV images of agricultural land is presented. The proposed method is named as NN-BF method as it is based on NN and BF descriptor matching strategies. The SI-UAV images possess high temporal, textural and intensity differences and affect the performance of feature points. The proposed method overcomes this problem by computing feature point descriptor matches in a novel way.

Experiments are performed on SI-UAV image dataset of agricultural land. The experiments are divided into three parts: image matching without proposed method, image matching with proposed method and image registration with proposed method. The experimental results of image matching show that ORB detector obtains the best repeatability score followed by SURF and SIFT. But the comparison of descriptors on SI-UAV images show that SURF obtains the best matching and precision scores.

But these scores are too low, i.e., 0.05% (matching) and 0.09% (precision). By using the proposed NN-BF method, the matching and precision scores of these descriptors are significantly increased. The experimental results show that proposed method increase the matching score and precision of SIFT up to 20.4% and 32%, respectively. Followed by SURF, which achieves 19.5% (matching) and 21.8% (precision). Without the proposed method both SIFT and SURF were unable to perform well on SI-UAV images. The experimental results for image registration show that with the help of proposed NN-BF method SIFT demonstrates the best performance by achieving 6.01% and 18.91% better precision scores than SURF and ORB, respectively. The proposed method also enhance the root mean square error of SIFT, which is 0.2 and 0.34 pixels' lower than SURF and ORB, respectively.

6.3 Future work

More work can be done in future in order to find robust methods and techniques and even using more feature points that can helps in improving the performance of registration process. Developing new feature point detector-descriptor algorithm that are invariant to temporal, photo-metric and textural changes between SI and UAV of agriculture land may result in better and improved performance. In other case, different combinations of existing feature point algorithms can be used, such as, SIFT detector and SURF descriptor or ORB detector and SIFT descriptor. Similarly, applying other machine learning techniques such as Support Vector Machine (SVM) etc., instead of NN and BF and also applying deep learning methods may prove to be effective and beneficial for registration of SI-UAV images of agricultural land.

REFERENCES

[1] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[2] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[3] Z. Song, S. Zhou, and J. Guan, "A novel image registration algorithm for remote sensing under affine transformation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4895–4912, 2014.

[4] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.

[5] M. Hasan, X. Jia, A. Robles-Kelly, J. Zhou, and M. R. Pickering, "Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints," in *IEEE International Geoscience and Remote Sensing Symposium*, pp. 1011–1014, 2010.

[6] C. Liu, J. Yuen, and A. Torralba, "SIFT flow: Dense correspondence across scenes and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 978–994, 2011.

[7] M. E. Linger and A. A. Goshtasby, "Aerial image registration for tracking," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2137–2145, 2015.

[8] A. Wong and D. A. Clausi, "ARRSI: Automatic registration of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1483–1493, 2007.

[9] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6469–6481, 2015.

[10] C. Toth and G. Jóźków, "Remote sensing platforms and sensors: A survey," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 22–36,

2016.

[11] C. Pohl and J. L. V. Genderen, “Review article multisensor image fusion in remote sensing: Concepts, methods and applications,” *International Journal of Remote Sensing*, vol. 19, no. 5, pp. 823–854, 1998.

[12] L. Tang and G. Shao, “Drone remote sensing for forestry research and practices,” *Journal of Forestry Research*, vol. 26, no. 4, pp. 791–797, 2015.

[13] G. Pajares, “Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs),” *Photogrammetric Engineering & Remote Sensing*, vol. 81, no. 4, pp. 281–330, 2015.

[14] M. I. Patel, V. K. Thakar, and S. K. Shah, “Image registration of satellite images with varying illumination level using HOG descriptor based SURF,” *Procedia Computer Science*, vol. 93, pp. 382–388, 2016.

[15] S. Saleem, A. Bais, and R. Sablatnig, “Towards feature points based image matching between satellite imagery and aerial photographs of agriculture land,” *Computers and Electronics in Agriculture*, vol. 126, pp. 12–20, 2016.

[16] Q. Du, N. Raksuntorn, A. Orduyilmaz, and L. M. Bruce, “Automatic registration and mosaicking for airborne multispectral image sequences,” *Photogrammetric Engineering & Remote Sensing*, vol. 74, no. 2, pp. 169–181, 2008.

[17] S. Saleem, A. Bais, and Y. M. Khawaja, “Registering aerial photographs of farmland with satellite imagery,” in *IEEE International Conference on Image Processing*, pp. 945–948, 2010.

[18] J. Liang, X. Liu, K. Huang, X. Li, D. Wang, and X. Wang, “Automatic registration of multisensor images using an integrated spatial and mutual information (SMI) metric,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 603–615, 2014.

[19] C.-H. Tsai and Y.-C. Lin, “An accelerated image matching technique for UAV orthoimage registration,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 128, pp. 130–145, 2017.

[20] K. Yang, A. Pan, Y. Yang, S. Zhang, S. Ong, and H. Tang, “Remote sensing image registration using multiple image features,” *Remote Sensing*, vol. 9, no. 6, pp. 581.1–581.21, 2017.

[21] T. Butz and J.-P. Thiran, “Affine registration with feature space mutual information,” in *Medical Image Computing and Computer-Assisted Intervention*, vol. 2208, pp. 549–556, Springer Berlin Heidelberg, 2001.

[22] X. Shen, L. Xu, Q. Zhang, and J. Jia, "Multi-modal and multi-spectral registration for natural images," in *European Conference on Computer Vision*, vol. 8692 LNCS, pp. 309–324, Springer International Publishing, 2014.

[23] L. Ding, A. Goshtasby, and M. Satter, "Volume image registration by template matching," *Image and Vision Computing*, vol. 19, no. 12, pp. 821–832, 2001.

[24] C. Lee, J. Oh, C. Hong, and J. Youn, "Automated generation of a digital elevation model over steep terrain in antarctica from high-resolution satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1186–1194, 2015.

[25] S. Nussbaum and I. Niemeyer, "Automated extraction of change information from multispectral satellite imagery," *Esarda Bulletin*, vol. 36, pp. 19–25, 2007.

[26] S. Nagarajan and T. Schenk, "Feature-based registration of historical aerial images by area minimization," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 15–23, 2016.

[27] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4328–4338, 2014.

[28] W. Ma, Z. Wen, Y. Wu, L. Jiao, M. Gong, Y. Zheng, and L. Liu, "Remote sensing image registration with modified SIFT and enhanced feature matching," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 1, pp. 3–7, 2017.

[29] P. Schwind, S. Suri, P. Reinartz, and A. Siebert, "Applicability of the SIFT operator to geometric SAR image registration," *International Journal of Remote Sensing*, vol. 31, no. 8, pp. 1959–1980, 2010.

[30] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[31] B. Wei, Z. Zhao, and X. Peng, "Spatial information based medical image registration using mutual information," *International Symposium on Networking and Network Security*, vol. 1, pp. 174–177, 2010.

[32] M. Teke and A. Temizel, "Multi-spectral satellite image registration using scale-restricted SURF," in *IEEE International Conference on Pattern Recognition*, pp. 2310–2313, 2010.

[33] P. Ricaurte, C. Chilán, C. Aguilera-Carrasco, B. Vintimilla, and A. Sappa, "Feature point descriptors: Infrared and visible spectra," *Sensors*, vol. 14, no. 2, pp. 3690–3701, 2014.

[34] Y. Qin, Z. Cao, W. Zhuo, and Z. Yu, "Robust key point descriptor for multi-spectral image matching," *Journal of Systems Engineering and Electronics*, vol. 25, no. 4, pp. 681–687, 2014.

[35] J. Jiang and X. Shi, "A robust point-matching algorithm based on integrated spatial structure constraint for remote sensing image registration," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 11, pp. 1716–1720, 2016.

[36] A. Wong and D. A. Clausi, "AISIR: Automated inter-sensor/inter-band satellite image registration using robust complex wavelet feature representations," *Pattern Recognition Letters*, vol. 31, no. 10, pp. 1160–1167, 2010.

[37] S. Saleem, A. Bais, and Y. M. Khawaja, "Updating farmland satellite imagery using high resolution aerial images," in *IEEE Frontiers of Information Technology*, pp. 133–136, 2011.

[38] C. Zhao and A. A. Goshtasby, "Registration of multitemporal aerial optical images using line features," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 117, pp. 149–160, 2016.

[39] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[40] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[41] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *IEEE International Conference on Computer Vision*, pp. 2548–2555, 2011.

[42] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *IEEE International Conference on Computer Vision*, pp. 2564–2571, 2011.

[43] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

[44] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust

independent elementary features," in *European Conference on Computer Vision*, pp. 778–792, Springer Berlin Heidelberg, 2010.

[45] E. Tola, V. Lepetit, and P. Fua, "DAISY: An efficient dense descriptor applied to wide-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 815–830, 2010.

[46] Y. Ye, J. Shan, S. Hao, L. Bruzzone, and Y. Qin, "A local phase based invariant feature for remote sensing image matching," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 142, pp. 205–221, 2018.

[47] J. Li, Q. Hu, and M. Ai, "Rift: Multi-modal image matching based on radiation-invariant feature transform," *arXiv preprint arXiv:1804.09493*, pp. 1–15, 2018.

[48] W. Li and H. Leung, "A maximum likelihood approach for image registration using control point and intensity," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1115–1127, 2004.

[49] H. Li, B. Manjunath, and S. Mitra, "A contour-based approach to multisensor image registration," *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 320–334, 1995.

[50] C. Davatzikos, J. Prince, and R. Bryan, "Image registration based on boundary mapping," *IEEE Transactions on Medical Imaging*, vol. 15, no. 1, pp. 112–115, 1996.

[51] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim, "elastix: A toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196–205, 2010.

[52] C. S. R. Fransens and L. V. Gool, "Multimodal and multiband image registration using mutual information," *European Space Agency, (Special Publication)*, no. 553, pp. 41–49, 2004.

[53] G. S. A. Andronache, M. von Siebenthal and P. Cattin, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Medical Image Analysis*, vol. 12, no. 1, pp. 3–15, 2008.

[54] M. Corsini, M. Dellepiane, F. Ponchio, and R. Scopigno, "Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties," *Computer Graphics Forum*, vol. 28, no. 7, pp. 1755–1764, 2009.

[55] A. Wong and P. Fieguth, "Fast phase-based registration of multimodal image data," *Signal Processing*, vol. 89, no. 5, pp. 724–737, 2009.

[56] A. Myronenko and X. Song, "Image registration by minimization of mapping complexity," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2009 IEEE, pp. 17–24, 2009.

[57] X. Fan, H. Rhody, and E. Saber, "A spatial-feature-enhanced MMI algorithm for multimodal airborne image registration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 6, pp. 2580–2589, 2010.

[58] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid image registration using conditional mutual information," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 19–29, 2010.

[59] Z. Zhang and R. S. Blum, "A hybrid image registration technique for a digital camera image fusion application," *Information Fusion*, vol. 2, no. 2, pp. 135–149, 2001.

[60] H.-M. Chen, P. Varshney, and M. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 11, pp. 2445–2454, 2003.

[61] Y. Yang and X. Gao, "Remote sensing image registration via active contour model," *AEU - International Journal of Electronics and Communications*, vol. 63, no. 4, pp. 227–234, 2009.

[62] M. Brown and D. Lowe, "Invariant features from interest point groups," in *British Machine Vision Conference*, pp. 23.1–23.10, 2002.

[63] M. F. Vural, Y. Yardimci, and A. Temlzel, "Registration of multispectral satellite images with orientation-restricted SIFT," in *IEEE International Geoscience and Remote Sensing Symposium*, pp. III–243–III–246, 2009.

[64] G. Bilodeau, A. Torabi, and F. Morin, "Visible and infrared image registration using trajectories and composite foreground images," *Image and Vision Computing*, vol. 29, no. 1, pp. 41–50, 2011.

[65] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[66] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, and R. Toledo, "Multispectral image feature points," *Sensors*, vol. 12, no. 9, pp. 12661–12672, 2012.

[67] Y. Yu, K. Huang, W. Chen, and T. Tan, "A novel algorithm for view and illumination invariant image matching," *IEEE Transactions on Image*

Processing, vol. 21, no. 1, pp. 229–240, 2012.

[68] O. Chum and J. Matas, "Matching with PROSAC — progressive sample consensus," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 220–226, 2005.

[69] Q. Yan, Y. Xu, X. Yang, and T. Nguyen, "HEASK: Robust homography estimation based on appearance similarity and keypoint correspondences," *Pattern Recognition*, vol. 47, no. 1, pp. 368–387, 2014.

[70] Y. Wu, Q. Miao, W. Ma, M. Gong, and S. Wang, "PSOSAC: Particle swarm optimization sample consensus algorithm for remote sensing image registration," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 242–246, 2018.

[71] Y. Wu, W. Ma, M. Gong, L. Su, and L. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 1, pp. 43–47, 2015.

[72] J. Ma, Y. Ma, J. Zhao, and J. Tian, "Image feature matching via progressive vector field consensus," *IEEE Signal Processing Letters*, vol. 22, no. 6, pp. 767–771, 2015.

[73] K. Zhang, X. Li, and J. Zhang, "A robust point-matching algorithm for remote sensing image registration," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 2, pp. 469–473, 2014.

[74] C. A. Aguilera, A. D. Sappa, and R. Toledo, "LGHD: A feature descriptor for matching across non-linear intensity variations," in *IEEE International Conference on Image Processing (ICIP)*, vol. 2015-Decem, pp. 178–181, 2015.

[75] Y. Ye and J. Shan, "A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 90, pp. 83–95, 2014.

[76] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2941–2958, 2017.

[77] A. Sedaghat and H. Ebadi, "Remote sensing image matching based on adaptive binning SIFT descriptor," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 10, pp. 5283–5293, 2015.

[78] H. Goncalves, L. Corte-Real, and J. A. Goncalves, "Automatic image

registration through image segmentation and SIFT," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2589–2600, 2011.

[79] N. Merkle, W. Luo, S. Auer, R. Müller, and R. Urtasun, "Exploiting deep matching and SAR data for the geo-localization accuracy improvement of optical satellite images," *Remote Sensing*, vol. 9, no. 6, pp. 586.1–586.18, 2017.

[80] D. Zeng, T. Zhang, R. Fang, W. Shen, and Q. Tian, "Neighborhood geometry based feature matching for geostationary satellite remote sensing image," *Neurocomputing*, vol. 236, pp. 65–72, 2017.

[81] C. Wu, Y. Wang, and H. R. Karimi, "A robust aerial image registration method using gaussian mixture models," *Neurocomputing*, vol. 144, pp. 546–552, 2014.

[82] B. Fan, Y. Du, L. Zhu, and Y. Tang, "The registration of UAV down-looking aerial images to satellite images with image entropy and edges," in *Intelligent Robotics and Applications*, pp. 609–617, Springer Berlin Heidelberg, 2010.

[83] G. Wang, Z. Zhai, B. Xu, and Y. Cheng, "A parallel method for aerial image stitching using ORB feature points," in *IEEE International Conference on Computer and Information Science*, pp. 769–773, 2017.

[84] Y. Lu, K. Gao, T. Zhang, and T. Xu, "A novel image registration approach via combining local features and geometric invariants," *PLOS ONE*, vol. 13, no. 1, pp. 1–18, 2018.

[85] S. Yahyanejad and B. Rinner, "A fast and mobile system for registration of low-altitude visual and thermal aerial images using multiple small-scale UAVs," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 104, pp. 189–202, 2015.

[86] Z. Wei, Y. Han, M. Li, K. Yang, Y. Yang, Y. Luo, and S.-H. Ong, "A small UAV based multi-temporal image registration for dynamic agricultural terrace monitoring," *Remote Sensing*, vol. 9, no. 9, pp. 904.1–904.19, 2017.

[87] D. Stratoulias, V. Tolpekin, R. de By, R. Zurita-Milla, V. Retsios, W. Bijker, M. Hasan, and E. Vermote, "A workflow for automated satellite image processing: from raw VHSR data to object-based spectral information for smallholder agriculture," *Remote Sensing*, vol. 9, no. 10, pp. 1048.1–1048.20, 2017.

[88] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[89] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: A survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2007.

[90] L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," *International Journal of Image Processing*, vol. 3, no. 4, pp. 143–152, 2009.

[91] Ş. Işık, "A comparative evaluation of well-known feature detectors and descriptors," *International Journal of Applied Mathematics, Electronics and Computers*, vol. 3, no. 1, pp. 1–6, 2014.

[92] J. Heinly, E. Dunn, and J.-M. Frahm, "Comparative evaluation of binary features," in *European Conference on Computer Vision*, vol. 7573 LNCS, pp. 759–773, Springer Berlin Heidelberg, 2012.

[93] A. L. Dahl, H. Aanæs, and K. S. Pedersen, "Finding the best feature detector-descriptor combination," in *IEEE International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, no. April 2016, pp. 318–325, 2011.

[94] D. Lowe, "Object recognition from local scale-invariant features," in *IEEE International Conference on Computer Vision*, vol. 2, pp. 1150–1157 vol. 2, 1999.

[95] P. L. Rosin, "Measuring corner properties," *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 291–307, 1999.

[96] E. Mair, G. D. Hager, D. Burschka, M. Suppa, and G. Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test," in *European Conference on Computer Vision*, vol. 6312 LNCS, pp. 183–196, Springer Berlin Heidelberg, 2010.

[97] S. Saleem and R. Sablatnig, "A robust SIFT descriptor for multispectral images," *IEEE Signal Processing Letters*, vol. 21, no. 4, pp. 400–403, 2014.