IMAGE REGISTRATION ALOGORITHMS FOR SATELLITE AND REMOTE SENSING APPLICATION

A Thesis submitted to Gujarat Technological University

for the Award of

Doctor of Philosophy

in

Electronics & Communication Engineering

by

Manish I. Patel

119997111009

under supervision of

Dr. Vishvjit K. Thakar

GUJARAT TECHNOLOGICAL UNIVERSITY

AHMEDABAD

March-2017

IMAGE REGISTRATION ALOGORITHMS FOR SATELLITE AND REMOTE SENSING APPLICATION

A Thesis submitted to Gujarat Technological University

for the Award of

Doctor of Philosophy

in

Electronics & Communication Engineering

by

Manish I. Patel

119997111009

under supervision of

Dr. Vishvjit K. Thakar

GUJARAT TECHNOLOGICAL UNIVERSITY

AHMEDABAD

March-2017

DECLARATION

I declare that the thesis entitled "***Image Registration Algorithms for Satellite and Remote Sensing Application***" submitted by me for the degree of Doctor of Philosophy is the record of research work carried out by me during the period from Sept 2011 to March 2017 under the supervision of **Dr. Vishvjit K. Thakar** and this has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other University or other institution of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis. I shall be solely responsible for any plagiarism or other irregularities, if noticed in the thesis.

Signature of the Research Scholar: .... Date: 16-03-2017

Name of Research Scholar: **Manish I. Patel**

Place: Ahmedabad

CERTIFICATE

I certify that the work incorporated in the thesis ***"Image Registration Algorithms for Satellite and Remote Sensing Application"*** submitted by **Mr. Manish I. Patel** was carried out by the candidate under my supervision. To the best of my knowledge: (i) the candidate has not submitted the same research work to any other institution for any degree/ diploma, Associateship, Fellowship or other similar titles (ii) the thesis submitted is a record of original research work done by the Research Scholar during the period of study under my supervision, and (iii) the thesis represents independent research work on the part of the Research Scholar.

Signature of Supervisor: ....   Date: 16-03-2017

Name of Supervisor: **Dr. Vishvjit K. Thakar**

Place: Ahmedabad

ORIGINALITY REPORT CERTIFICATE

It is certified that PhD Thesis titled ***"Image Registration Algorithms for Satellite and Remote Sensing Application"*** by **Manish I. Patel** has been examined by us. We undertake the following:

a. Thesis has significant new work / knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.

b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results, or words of others have been presented as an Author own work.

c. There is no fabrication of data or results which have been compiled / analyzed.

d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

e. The thesis has been checked using **Turnitin** (copy of originality report attached) and found within limits as per GTU Plagiarism Policy and instructions issued from time to time.

Signature of the Research Scholar: ……………………… Date: 16-03-2017
Name of Research Scholar: **Manish I. Patel**
Place: Ahmedabad

Signature of Supervisor: ………………………………… Date: 16-03-2017
Name of Supervisor: **Dr. Vishvjit K. Thakar**
Place: Ahmedabad

Image Registration Algorithms for Satellite Images and Remote Sensing Applications

ORIGINALITY REPORT

%14	%14	%2	%1
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

13%

★ u.cs.biu.ac.il

Internet Source

EXCLUDE QUOTES	ON	EXCLUDE MATCHES	< 1%
EXCLUDE BIBLIOGRAPHY	ON		

PhD THESIS Non-Exclusive License to
GUJARAT TECHNOLOGICAL UNIVERSITY

In consideration of being a PhD Research Scholar at GTU and in the interests of the facilitation of research at GTU and elsewhere, I, **Manish I. Patel** having **119997111009** hereby grant a non-exclusive, royalty free and perpetual license to GTU on the following terms:

a) GTU is permitted to archive, reproduce and distribute my thesis, in whole or in part, and/or my abstract, in whole or in part (referred to collectively as the Work") anywhere in the world, for non-commercial purposes, in all forms of media;

b) GTU is permitted to authorize, sub-lease, sub-contract or procure any of the acts mentioned in paragraph (a);

c) GTU is authorized to submit the Work at any National / International Library, under the authority of their "Thesis Non-Exclusive License";

d) The Universal Copyright Notice (©) shall appear on all copies made under the authority of this license;

e) I undertake to submit my thesis, through my University, to any Library and Archives. Any abstract submitted with the thesis will be considered to form part of the thesis.

f) I represent that my thesis is my original work, does not infringe any rights of others, including privacy rights, and that I have the right to make the grant conferred by this non-exclusive license.

g) If third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required, I have obtained such permission from the copyright owners to do the acts mentioned in paragraph (a) above for the full term of copyright protection.

h) I retain copyright ownership and moral rights in my thesis, and may deal with the copyright in my thesis, in any way consistent with rights granted by me to my University in this non-exclusive license.

i) I further promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to my University in this non-exclusive license.

j) I am aware of and agree to accept the conditions and regulations of PhD including all policy matters related to authorship and plagiarism.

Signature of the Research Scholar: ……………………… Date: 16-03-2017
Name of Research Scholar: **Manish I. Patel**
Place: Ahmedabad

Signature of Supervisor: ………………………………… Date: 16-03-2017
Name of Supervisor: **Dr. Vishvjit K. Thakar**
Place: Ahmedabad

THESIS APPROVAL FORM

The viva-voce of the PhD Thesis submitted by **Mr. Manish I. Patel** (Enrollment No. **119997111009**) entitled **"Image Registration Algorithms for Satellite and Remote Sensing Application"** was conducted on 16th March, 2017 (Thursday) at Gujarat Technological University.

(Please tick any one of the following options)

☐ The performance of the candidate was satisfactory. We recommend that he/she be awarded the Ph.D. Degree.

☐ Any further modifications in research work recommended by the panel after 3 months from the date of first viva-voce upon request of the Supervisor after which viva-voce can be re-conducted by the same panel again.

(briefly specify the modification suggested by the panel)

☐ The performance of the candidate was unsatisfactory. We recommend that he/she should not be awarded the Ph.D. Degree.

(The panel must give justifications for rejecting the research work)

Name and Signature of Supervisor with Seal

1) External Examiner 1 Name and Signature

2) External Examiner 2 Name and Signature

3) External Examiner 3 Name and Signature

ABSTRACT

Image Registration is the process of overlaying images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It is a very important pre-processing step in the applications such as remote sensing, medical diagnosis, computer vision etc. where the final information is derived based on the comparison of the images. The reason for the increased significance of image registration for satellite images is that the remote sensing is currently moving towards operational use in many important applications, for societal benefits as well as scientific study. The satellite images are multi-temporal (taken at different dates), multisource (captured from multiple sensors), multi-spectral (captured at different frequency bands) or multimodal (obtained with different acquisition modalities). Image registration for remote sensing is also difficult due to the challenges such as large image size, having nonlinear variations in intensity level, atmospheric effects, noise, presence of clouds, occlusions etc.

Broadly there are two classes of approaches for image registration: area (or intensity) based methods and feature based methods. In the framework of area based methods, choice of similarity measure and search strategy play significant role. Image registration using mutual information as a similarity measure is investigated as it is best suited for multimodal images; but the computational complexity is challenging. An alternative approach to image registration is to transform the spatial information of image into another transform domain such Fourier transform and then to use the properties of the transform domain to estimate the registration parameters. Use of radon transform is investigated to estimate the registration parameters, which founds to be close to the actual parameters and robust to noise as well.

In feature based methods, basically there are four steps: feature detection, feature matching, estimation of registration parameters and re-sampling. Due to its advantages, Speeded up Robust Feature (SURF) is widely used in other image processing applications, including image registration also for the field other than remote sensing. SURF is explored which is found little in literature for satellite image registration mainly because of associated challenges. Direct use of SURF is not appropriate for many cases for registration of satellite image or remote sensing applications. Satellite image registration with varying intensity level is improved using Histogram of Oriented Gradient (HOG)

descriptor in SURF, as HOG descriptor is more illumination invariant which is a requirement for image registration of many satellite images such as multispectral images. Further, if the images are having occlusions such as clouds or shadows, they mislead image registration by giving incorrect matches. The features related to clouds or shadows are removed using support vector machine based classification, this results in reduction of incorrect matches. By simulation results it is shown that by both the modifications i.e. HOG as a descriptor and feature refinement step using classification just before matching step lead to improvement in correct match rate which ultimately improves the feature based image registration.

ACKNOWLEDGEMENT

I wish to express my sincere appreciation to those who have contributed to this thesis and supported me in one way or the other during this amazing journey.

First, I am extremely grateful to my supervisor, ***Dr. Vishvjit K. Thakar***, Professor and head of the Department of Electronics and Communication, A. D. Patel Institute of Technology, New Vallabh Vidhya Nagar, Gujarat for his guidance and all the useful discussions and brainstorming sessions, especially during the difficult conceptual development and decision stage. His deep insights helped me at various phases of my research. His invaluable suggestions and constructive criticisms from time to time enabled me to maintain my enthusiasm to complete my work successfully.

The completion of this work would not have been possible without, the Doctorate Progress Committee (DPC) members: ***Dr. Mehul S. Raval***, Associate Professor, Ahmedabad University, Ahmedabad and ***Dr. Tanish H. Zaveri,*** Associate Professor, Department of Electronics and Communication, Nirma University, Ahmedabad. I am thankful for their rigorous examinations and precious suggestions during my research. I am also obliged for the valuable suggestions from foreign co-supervisor ***Dr. Shishir Shah***, University of Houston, Texas.

My gratitude goes out to the assistance and support of ***Dr. Akshai Agrawal***, Ex. Vice Chancellor, ***Dr. Navin Sheth,*** Vice Chancellor, ***Shri J. C. Lilani***, Registrar, Data Entry Operator and other staff members of PhD Section, GTU.

At this stage, I would also like to acknowledge the support provided by my institute ***Sankalchand Patel College of Engineering, Visnagar***. I would also like to acknowledge my ***family members*** for their helping hands and patience as and when required.

Manish I. Patel

Table of Contents

List of Abbreviation

ABM	Area Based Method
CC	Cross Correlation
CMR	Correct Match Rate
FBM	Feature Based Method
FFT	Fast Fourier Transform
GCP	Global Control Points
GPS	Global Positing System
HOG	Histogram of Oriented Gradients
MI	Mutual Information
MLMI	Maximum Likelihood Mutual Information
NDVI	Normalised Difference Vegetation Index
PCA	Principle Component Analysis
RANSAC	Random Sample Consensus
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SIFT	Shift Invariant Feature Transform
SURF	Speeded-Up Robust Feature
SVM	Support Vector Machine

List of Figures

List of Tables

Chapter 1

Introduction

Image processing and analysis has become an integral part of many applications such as computer vision, robotics, instrumentation, medical, remote sensing etc. Depending on the complexity, there are three levels of image processing techniques. Low level of image processing techniques is logical or mathematical operators that perform simple image processing tasks. Medium level image processing combines the simple low level operators to perform feature extraction and pattern recognition functions. High level image processing uses combinations of medium level functions to perform some interpretation.

Many applications involve two or more than two images and it is required to combine those images. For example if growth rate of cancer is to be determined then its images are captured at appropriate time interval and by observing the changes, the growth rate of cancer can be estimated. Similarly if the patient is under treatment, then also by similar way the effect of treatment can be observed. In natural hazards, the effect analysis can be observed by combining the two images: before hazard and after hazard image.

There are many other examples under remote sensing applications such as weather forecasting, urban growth, vegetation, flood, fire, glacier monitoring, environment monitoring, creating super resolution images, Geographical Information System, image mosaic etc. where it is required to combine or fuse two images. In such applications it is not necessary that the available images are already geometrically aligned. So, in simple words, before fusion process it is required to align the images as a very important mandatory pre-processing step, termed as ‘image registration’.

1.1 Image Registration

Image registration is defined by various ways in literature. It is the process for aligning two images. As a simple example, as shown in Fig. 1.1, before fusion process image 2 is required to be rotated to align it to the first image 1.

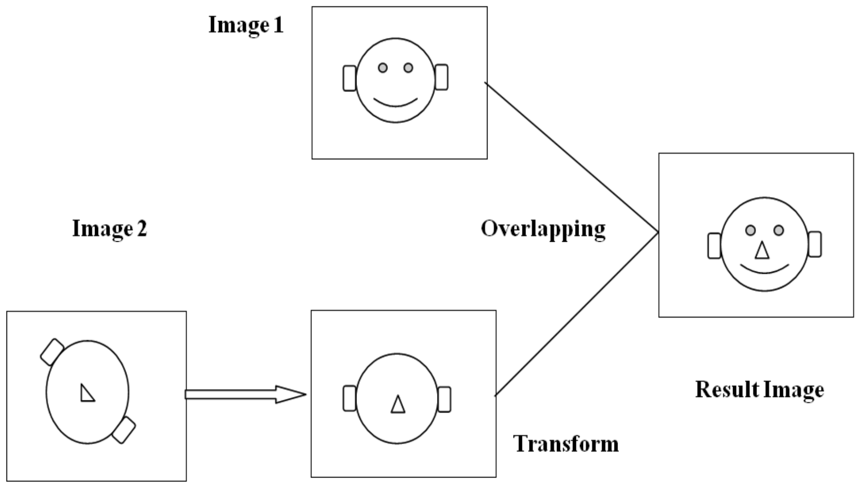


Figure 1.1 Simple concept of image registration

In (Barbara Zitová, 2003), image registration is defined as 'the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors'.

Before exploring the numerous approaches of image registration, some of the terminologies are introduced here:

Target (or reference or fixed): the image that is kept unchanged and is used as a basis for the warping.

Source (or sensed or moving): the image that is geometrically transformed to be aligned with the target image.

Transformation (or warping): the function used to modify the source towards the target image.

Registration is determining the spatial transformation that maps points in the target image to points in the source image. In (Brown, 1992), image registration is defined as a mapping between two images both spatially and with respect to intensity. If the images are defined as two dimensional arrays of a given size denoted by I_1, and I_2 where $I_1(x, y)$ and $I_2(x, y)$

each map to their respective intensity values, then the mapping between images can be expressed as:

$$I_2(x, y) = g\ (I_1(f(x, y)) \tag{1.1}$$

where f is a 2-D spatial coordinate transformation, i.e., $(x', y') = f(x, y)$ and g is 1-D intensity or radiometric transformation. Finding the spatial or geometric transformation is generally the key to any registration problem.

Basically image registration determines the geometrical transformation parameters of the transformation model, which can be applied to sensed image to align it with the reference image. In similarity transformation model, rotation, scale and translations are involved. Rigid transformation includes only translations and rotations, although in the literature, rigid transformations are sometimes allowed to include scaling. An affine transformation can furthermore include shearing. This type of transformation maps straight lines to straight lines and preserves the parallelism between lines. Another class consists of curved transformations, which allow the mapping of straight lines to curves. It is also known as elastic or deformable transformation. These transformations are shown in Fig. 1.2.

The complexity of image registration algorithms depend on the complexity of the geometrical transformation model or say degree of freedom, under consideration. For example, in the image registration with affine transformation is more complex than the image registration with similarity transformation as degree of freedom is larger in former case.

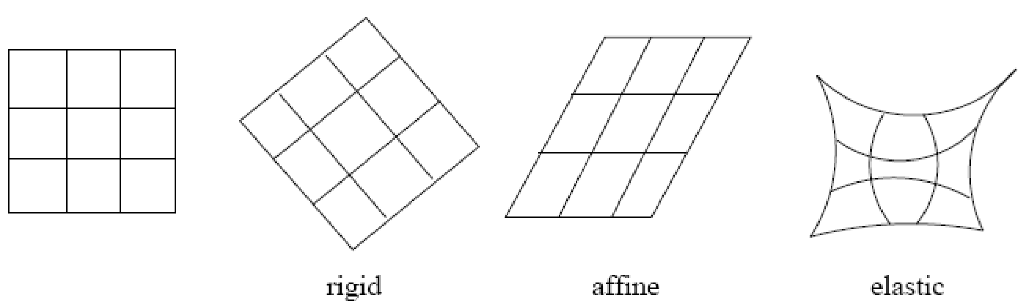


Figure 1.2 Transformation types

Mathematically, for example, 2-D rigid transformation can be represented by the matrix operation:

$$\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{1.2}$$

$$\mathbf{x}' = T(x|p) = T(x, y|t_x, t_y, \theta) \tag{1.3}$$

So in this case the main goal for image registration is to find the parameters t_x, t_y and θ that optimize some image similarity metric.

With reference to remote sensing, when images are captured by sensors mounted on satellite it can have the geometrical errors. These geometrical errors may be because of some reasons. Some of the reasons are:

- The rotation of the Earth during image capturing
- The finite scan rate of the sensors
- Variations in field of view of the sensors
- Curvature of the Earth
- Sensor non-idealities
- Variations in platform altitude and velocity

Correction to the Geometric Distortion

In general there are two techniques those can be used to correct various geometric distortions present in the satellite image datasets. One method is systematic correction as it is systematically based on some models. In this method the nature and magnitude of the sources of geometric distortions are modelled and these models are used to correct the geometric errors by establishing the correction formula. Systematic correction is effective when the distortions are well characterized such as curvature of the Earth and Earth rotation. The second method depends on the establishing mathematical relationships between the coordinates of pixels in an image and the corresponding coordinates of those points of the reference image. This is known as image registration. These relationships can be established to correct the geometrical errors irrespective of any knowledge of the image acquisition process or associated sensors. One more terminology 'geo-referencing' is quite confusing. In image registration if the reference image is a map then it becomes image to map registration also known as geo-referencing.

Most of the image registration methods can be best explained by the following four steps:

(1) Feature detection/extraction

In the first step of feature detection, specific and discriminating features such as point features, line features are either manually or automatically detected. These detected features are represented by the point representative called as feature points, interest points or control points. In addition to the points, a descriptor is used to describe the nature of the region surrounded by the detected points, which can be used for further processing.

(2) Feature matching

In feature matching step, the mapping or correspondence is found between the detected features of both the images. For that purpose, the feature descriptor and similarity function (also known as similarity measure or cost function) are used. Based on the nature of the similarity measure it is required to be maximized or minimized to obtain best matching.

(3) Transform model estimation

Geometrical transformation parameters such as rotation, translation etc. are estimated based on the correspondence found in the previous step. The complexity of the estimation depends on the number of parameters or the transform model under consideration.

(4) Image re-sampling and transformation

The sensed image is geometrically transformed based on the parameters estimated in the previous step, to align it with the reference image. Image pixel intensity values in non-integer coordinates are computed by an appropriate interpolation technique such as bilinear, nearest neighbour, bicubic etc.

These four steps (Barbara Zitová, 2003) are also shown in Fig. 1.3. All the image registration approaches differs in terms how they differ for the above mentioned steps. Further if any of the steps is interactive, it is said to be manual or semi manual image registration according to the amount of human intervention. If no human intervention is required then it is automatic image registration.

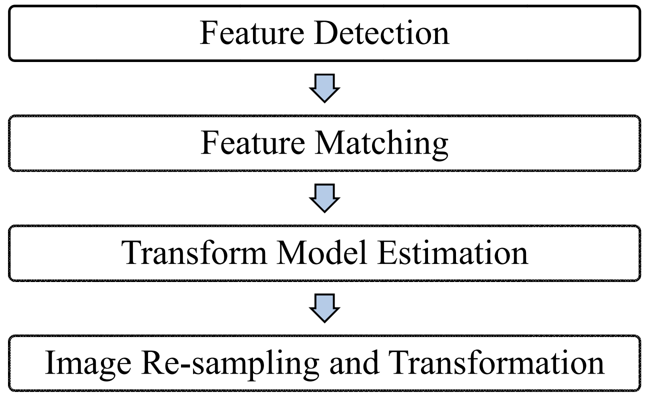


Figure 1.3 Steps for image registration

1.1.1 Classification of Image Registration Methods

The approaches for the image registration can be broadly classified as Area (also called as spatial or intensity or pixel) Based Methods (ABM) or Feature Based Methods (FBM). A third category can be considered as Transform domain Based Method (TBM), where some transformation is applied on the image and their properties are utilized to estimate the geometrical transformation parameters. It can also be considered under ABM as first it transforms the whole image into another transform domain. This classification is shown in Fig. 1.4.

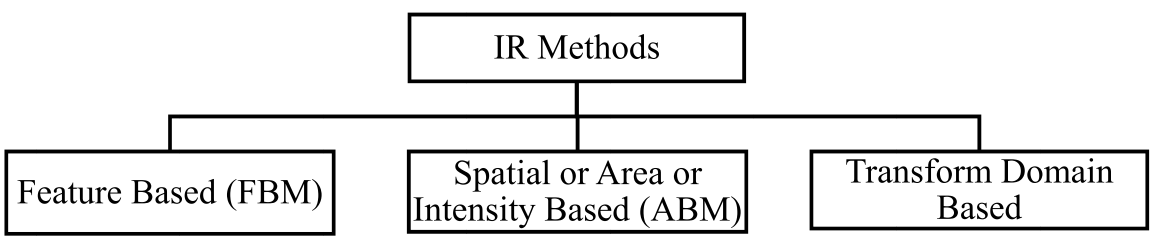


Figure 1.4 Classification of image registration approaches

In ABM, intensity of every pixel is used as feature, so ABM put emphasis on the feature matching step and its optimization step rather than on their detection as shown in Fig. 1.5. In ABM no features are detected so the first step of image registration shown in Fig. 1.3 is omitted, while in FBM salient features are first detected and only those are proceed further in the remaining steps.

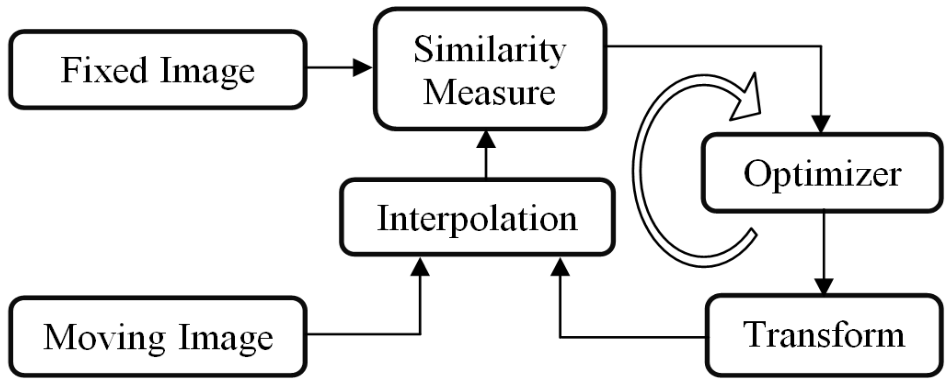


Figure 1.5 Generic flow of ABM for image registration approaches

1.2 Image Registration for Remote-sensing

Remote sensing can be defined as "the process by which information about an object or phenomenon is acquired from a remote place". In our case the objects are on the Earth and the remote place is a satellite or an aircraft. Satellite imaging is referred as "the use of sensors located on space-borne platforms to capture electromagnetic energy that is reflected or emitted from surface of the Earth". In this case the Sun is a source of energy so the sensor is termed as passive source. While in case of active sensors such as radar, they use their own source of energy to capture specific targets.

Sensor Characteristics

In remote sensing, generally different sensors are designed for different types of features to be observed which in turn depends on the requirements or applications. This defines the spatial, spectral, radiometric and temporal resolution of the sensors. The term resolution is "the smallest unit of granularity that can be measured by the sensor". The spatial resolution is "the area on the ground from which reflectance is obtained and integrated to compute the value assigned to each pixel". The spectral resolution is "the bandwidths utilized in the electromagnetic spectrum". The radiometric resolution is defines as "the number of bits which are used to record a given energy corresponding to a given wavelength". The temporal resolution is "the number of observations, defined by the orbit of the satellite and scanning of the sensor".

Most of the sensors used for remote sensing are multispectral i.e. utilizes several bands to capture the energy emitted or reflected from Earth features. Addition of the panchromatic imagery, which is of higher spatial resolution than that of multispectral imagery in the visible part of the spectrum, provides detailed information. In Landsat-4 and 5, the numbers of bands was increased from four to seven relative to Landsat-1 and 2, which added the bands from visible and thermal frequency range. The Landsat series was further extended by introducing Landsat-7 which contains additional panchromatic band.

Significance of Image Registration for Remote Sensing

Remote sensing is currently moving towards the operational uses in different applications for societal benefit and/or scientific study. This has increased the significance of image registration of satellite image database. The applications based on the monitoring the Earth surface over time are management of natural disasters, management of natural resources, preservation of environment, assessment of climate changes etc. Nowadays, due to multiple missions by many countries, there is an increasing availability of number of images with different characteristics. So a growing need emerges to process different remote sensing image for information extraction and fusion. This includes the fusion of newly acquired images with previous images captured with different modalities or geometric configuration or with cartographic data. So the satellite images or remote sensing images captured in this way can be multi-temporal (captured at different dates), multisource (captured by different sensors), multimode (obtained with different acquisition modalities), or multi-view/stereo images (taken from different viewpoints).

Image registration of such remote sensing data obtained from different satellites and airborne has become critical for several reasons. Image registration plays an important role in spatial and radiometric calibration of multi-temporal measurements for obtaining large, integrated datasets for long-term tracking of various phenomena. In another example, for change detection over time, it is required to register multi-sensor and multi-temporal images accurately. In earlier studies by (J. R. G. Townshend, 1992) and (Khorram, 1998) it is shown that even a small error in registration might have a large impact on the accuracy of global change measurements. It is best explained in terms of Normalized Difference Vegetation Index (NDVI). When looking at simulated data of MODIS at 250-m spatial resolution, a mis-registration error of a pixel may produce a 50% error in computation of the NDVI. For such applications, very accurate registration which is very close to pixel-to-pixel matching i.e. sub-pixel level of accuracy is required.

Image registration for remote sensing can be classified as follows:

- Multimodal registration

This enables the complementary information from different sensors. Some of application examples are agriculture and crop forecasting, water urban planning, mineral and oil exploration, cartography, flood monitoring, crop disease control, real estate tax monitoring and detection illegal crop. Here the images to be registered are captured by different sensors. In such application say for example, combinations of remote sensing and Geo Informatics System to help in critical decision making process.

- Multi-temporal registration

This can be used for detecting the changes from the data or images obtained from one or more sensors over a period of time or at different time. Cloud removal is another application of temporal registration, in which images of observations over several days are required to be integrated to create cloud-free data.

- Multi-view point registration

This integrates information from one or more moving platforms navigating together into three-dimensional models. Landmark navigation and planet exploration are examples of such applications.

- Multi-template registration

This is to find the correspondence of the small template in the base or reference image. It is useful for map updating and content based searching.

Automatic Vs Manual Image Registration for Remote Sensing

In the process of image registration it is required to locate and match similar regions in the given two images to be registered. If user performs one or more of these tasks visually using interactive software then it is called manual image registration. If these tasks are performed autonomously then it is called automatic image registration. In manual image registration the user selects the distinctive points from both the images which are normally also known as control points or tie points. These control points are manually matched and corresponded. These corresponded points are used to compute the parameters of the geometric transformation under consideration. The commercially available tools such as ENVI, Geomatica etc. are using such manual approach. But it has some drawbacks. It is repetitive, laborious, and time consuming so becomes prohibitive for large amount of data. In certain cases more efforts and care is required for control points. It may required to visit the site and label with the Global Positioning System (GPS) coordinates a set of robust ground features such as any well known place or some road intersections (Wang and Ellis, 2005b). Under this case the control points are also known as Ground Control Points (GCP). These geo-referenced control points are utilised for the process of image registration.

Systematic Correction Vs. Image Registration

While capturing the image by knowing some of the system related parameters such as type, orientation and shape of orbit, angle of view etc., it is possible to give approximate ground coordinates to the image. Such information is sometimes supplied in metadata associated with the image. This kind of correction is known as systematic correction, sometimes referred as navigation as it is based on navigation model with certain parameters. Systematic correction has systematic or random errors as the concerned parameters are having accuracy up to certain level only. Precision correction is required to correct these errors. At ground base station a precision correction is performed using the features or content of the image which is termed as image registration. Systematic correction is model based and image registration is feature or content based. Depending on the sensor and how it is old, the accuracy of systematic correction can be within a few pixels up to a few tens of pixels. Some of the models those use information from GPS (E1-Rabbany, 2002) are usually accurate within a few pixels. But in many applications such as change detection the desired accuracy is up to sub-pixel. So image registration is the necessary steps to refine the accuracy to the desired level.

The remote sensing images may be systematically corrected if the metadata is available. Then after, image registration is performed to have desired accuracy, where only small range of geometrical transformation parameters is required to consider as it is coarsely corrected using systematic correction. But if the images are not systematically corrected and no metadata is available then they have to be registered directly. Here more range of geometrical transformation parameters is required to consider.

Challenges in Image Registration for Satellite Images and Remote Sensing Applications

Many methods are developed for image registration in other fields such as medical. But there is no single, universal, stand-alone method for image registration in remote sensing field. This is mainly due to some specific challenges associated with the satellite images as summarized below (Eastman, 2010).

Remote sensing vs. other images:

Compared to medical images, the remote sensing images have certain variations which make its registration quite difficult.

(1) Variety in the types of the sensor data and the conditions of data acquisition: One technique that works well on the images which are captured by some particular sensors, in some particular atmospheric condition or some particular location may not necessarily work well for images which are captured by some other sensor or in different environmental condition or different location. In the field of remote sensing there is so diversity for the sensors in terms spatial, radiometric and spectral resolution as well as technology behind it. There is significant effect of atmospheric condition on image registration for satellite images because it is not in-door as it is for medical field.
(2) The size of the data: Compared to the other field the satellite images are of very large size say in case of Landsat image, typically it is in terms of 7000X7000 pixels. For processing of such large data, enough computational resources such as speed and memory are the mandatory requirements.
(3) The lack of well-distributed control points resulting in the difficulty to validate image registration methods in the remote sensing.

Atmospheric and cloud interactions:

Main atmospheric effects are scattering and absorption. The atmospheric effects on the data fidelity depend on the distance travelled by the radiation through the atmosphere and magnitude of the energy of the signal. Normally the remote sensors take the observations within the some specific ‘atmospheric windows’ that are defined by considering the range of wavelength having minimum absorption. Most of the atmospheric effects such as humidity and the concentration of atmospheric particles are corrected by the respective model. But yet the effects related to the temporal and local weather data acquisition are normally not included in the models so they can’t be systematically corrected. Recognition and removal of clouds is also an important pre-processing step as far as image registration is considered.

Multi-temporal effects:

There are several multi-temporal effects present in the images which make the image registration challenging. The effects may be natural or man-made. An example of natural effect is different lighting condition due the change in the Sun angle. The viewing angle of the sensor can change from pass to pass. Further with seasonal changes, such as weather changes, crop changes; the surface reflectance also varies. Over the urban, development occurs which is an example of man-made multi-temporal effect. In both the cases, natural or man-made multi-temporal effects, certain features of the images are drastically changed or not visible, which creates difficulties in image registration.

Terrain/relief effect:

Another source affecting the registration of remote sensing imagery is the topography or the terrain. Various terrain features can be represented by variations in image brightness, in different ways, depending on the angle of illumination. This means that depending on the slope of the geographic relief, the characteristics of the sensor and the satellite orbit, and the time of the day, terrain relief effects might appear very differently in the images to be registered. Large topographic variations can be corrected using a terrain model but small local effects remain present.

Multi-sensor (having different spatial and spectral resolutions):

When dealing with multiple sensors, with different geometries and various spatial, spectral, radiometric and temporal resolutions it is necessary to address the following image registration issues:

(1) Choice of geometric transformations that respond to various spatial resolutions and different scanning patterns

(2) Extraction of image features that are invariant to radiometric differences due to multispectral and multi-temporal resolution

1.3 Motivation

Based on the above mentioned challenges, the characteristics of satellite images can be summarized by the following:

- Multi-sensor
- Multi-modal
- Multi-temporal
- Multi-spectral
- Multi-resolution
- Large size
- Lack of known control points
- Nonlinear variation in illumination level
- Noise and clouds
- Atmospheric condition

Due to these characteristics of satellite images, an image registration algorithm used for medical field or computer vision field may not work satisfactory. Further, an image registration algorithm that addresses one of above challenge and work on specific dataset may not work satisfactory for the other challenges or other datasets. Due to this fact still there is no single stand alone algorithm that works for all kind of satellite images or remote sensing applications, which has motivated to work on it.

1.4 Research Objectives

The main research objectives are:

- To study and investigate area based methods and feature based methods of image registration for satellite images and remote-sensing applications

- Using signal/image processing techniques, suggest an approach to address some of the challenges for satellite image registration

- The approach should be automatic and take care of computational complexity for large sized satellite images
- To address the intensity level variation issue for satellite image registration, as it commonly occurs in case of multi-spectral, multi-modal, multi-temporal and multi-sensor cases.
- To address the issue of presence of clouds/shadows in satellite images for image registration, which are frequently present in many cases and affect the accuracy of the image registration.

1.5 Contribution

The original contribution of the thesis is in terms of modifications suggested in image registration algorithms for satellite images. Computational speed is investigated using mutual information as a similarity measure in ABM, while properties of radon transform are investigated for estimating image registration parameters; which shows ABM can be preferred only for fine image registration. In SURF based FBM algorithm, Correct Match Rate (CMR) is improved by using HOG as descriptor in case of satellite images with nonlinear variations in illumination level. To address cloud or other occlusion present in the satellite images, the use of SVM classification of features before feature matching step also improved CMR.

1.6 Organization of the Thesis

In rest of the thesis, chapter 2 covers related literature reviews which include various area based methods, Fourier transform based methods, feature based methods as well as state of the art in the field of image registration for satellite images and remote sensing. Chapter 3 is on investigation of ABM, where Mutual Information (MI) as a similarity measure is investigated. It is followed by radon transform based approach for image registration. Chapter 4 and chapter 5 are for FBM. In chapter 4, the work focuses on feature descriptor step, while chapter 5 is for the work on image matching step. The chapters 3 to 5 are also covering the related simulation results and discussion. Conclusion and future work is presented in chapter 6.

Chapter 2

Literature Review

2.1 Overview

This chapter represents the survey of diversified image registration approaches applied to the satellite images and remote-sensing field. The survey will cover some of the basic concept, overall framework, some selected algorithms etc. Though manual and semi-automatic approaches are still in use for remote sensing, here only automatic approaches are focused. Surveys on various articles on image registration are found in (Barbara Zitová, 2003), (Medha V. Wyawahare, 2009). General text books on image registration are (Goshtasby, 2005) and (Eastman, 2010). A specific survey on image registration for medical image is in (J. P. W. Pluim, 2003) and (J.B.Antoine Maintz, 1998).

For any image registration algorithm, the objective is to implement an accurate, automatic image registration algorithm. Depending on the satellite image datasets or remote sensing applications, the specific objectives may cover one or more of the following.

- Improving robustness and reliability i.e. the algorithms should work well under the case of noise, occlusion, some other variation or problem
- Refining the geometric transformation to better model the imaging process of an instrument and satellite
- Improving the accuracy of the transformation computed up to sub-pixel to satisfy the requirement of the application at hand
- Increasing the speed of image registration process for large satellite image datasets
- Managing multimodal image registration so that it can be applicable to the images with radiometric, scale and other differences that might be present across band, instrument or platform.

In general, majority of the researchers have contributed to one or more of the basic components of image registration approach by some modifications to the component/s or suggesting alternative method to the component/s. So knowledge of related basics, principles and their combinations can easily help for establishing a new image registration technique. According to the major surveys, image registration algorithm consists of the following main elements.

- Degrees of freedom and their range for the possible geometrical transformation between two images that will be considered i.e. search space
- Extracted features or feature space of information content
- Similarity (or dissimilarity) measure or metric that gives the merit of matching image features, by maximising (or minimising) it; this is also termed as cost function
- Optimization method used to find the optimal geometric transformation by maximising (or minimising) the cost function

As discussed in the previous chapter, various image registration algorithms are generally classified as ABM or FBM. In ABM, areas or regions of the original images are matched. These methods compute the differences of pixel values or use all pixel values to compute an intermediate full information representation such as Fourier coefficients. The methods in which some transformation such as Fourier is used it can be treated a class under ABM or a separate class of transform domain based method. In area based approach no point correspondence step is required; instead it matches whole area; so it is also termed as correspondence-less matching. In FBM on the other hand the original images are pre-processed to extract distinctive & highly informative features. These features only are used for further step of matching.

In some work combination of both ABM and FBM are also reported. Such two step approach of FBM followed by ABM may be appropriate when large distortions or displacement make feature point matching more robust. This is because local regions are warped so individually the pixels are aligned poorly but derived features, based on their nature, are comparatively more invariant to the distortions. After those large transformations are initially accounted for, ABM can be effectively performed.

2.2 Correlation Based Methods

Correlation-related methods directly compute a similarity measure for corresponding image regions by pixel-wise comparisons of intensity values. Also known as template matching, a region from one image is translated around the other to find alignment that optimizes the similarity measure. The similarity measure for the absolute difference of pixel intensities is given by

$$D[\Delta x, \Delta y] = \sum_{x}\sum_{y}|I_1(x, y) - I_2(x + \Delta x, y + \Delta y)| \tag{2.1}$$

where $\Delta x, \Delta y$ denote, respectively, the horizontal and vertical shifts in the sensed image and summation is carried out over all x, y locations of an image region. This sum of absolute difference (SAD) similarity measure is also known as L_1-norm or Manhattan norm. Another similarity measure sum of squared difference (SSD) known as L_2-norm or Euclidian distance can also be used alternatively.

Correlation based methods are widely accepted methods for image registration in various filed such as computer vision, medical field as well as remote sensing. Simultaneously it has many disadvantages which are addressed by some of the researchers for their practical applications. First and foremost problem is its computational complexity as it is a brute-force approach. It requires O (n^2m^2) number of operations, where m and n are the vertical and horizontal size of the image or image window. The problem of large computation of the brute-force approach becomes dominant when image registration is required with accuracy of sub-pixel level. So this is only useful in case of smaller size of image. Else improvement in computation speed is required. There are certain approaches to improve the speed. One of the approaches is computing the correlation coefficient in frequency domain sounds effective.

Some other approaches are partial computations, coarse-to-fine pyramid search, specialized parallel hardware, numerical optimization etc. In partial computation technique, similarity measure is computed at sampled locations in the search window (R. J. Althof, 1997), instead of computing it for all locations in the search window. Another technique known as Sequential Similarity Detection Algorithm (SSDA) accumulates the sum of absolute differences until the measure becomes large enough for the current alignment to provide the likely minimum value. The SSDA based image registration is performed on Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS) data in (Solberg, 2005). In coarse-to-fine pyramid search wavelet transform is used. In the technique the image is decimated or downscaled into a sequence of smaller lower resolution images of sizes that are decreasing sequence, normally of powers of two. Consequently, the actual translation in the original image is also reduced into smaller images, hence search window becomes smaller. The solution obtained to this smaller image can be extrapolated to the higher resolution image. Further the blurring and decimation can smooth the cost function (or similarity measure) which reduces the impact of local minima and noise.

Further the problem of higher computational requirement increases in case of complex geometric transformation instead of translation only. The addition of each geometrical transformation parameters such as rotation, scale etc. multiplies the size of the search space. If more complex models are required, with relatively low order the extra parameters can be incorporated into the least-squares formulation by a linearized approximation (P. Thevenaz, 1998).

In image registration process the correlation as a similarity measure is degraded by the existence of noise, occlusions, temporal changes, and radiometric differences in the multimodal, multi-temporal, multispectral or multi-sensor and other source that may affect the pixel intensity or creates variations in pixel intensity. This is a driving force for the development of the alternative approaches of image registration for satellite images. This includes variety of complex statistic based similarity measures and feature based methods. Some efforts are placed to make the correlation based approach more robust. Robust statistical measures, such as M-estimators, can be used to reduce the influence of outlying noise values. A version of normalized correlation, based on M-estimators, which is robust to occlusions and noise is presented in (Arya, 2007).

Some of the operational image registration systems are reviewed in (R. D. Eastman, 2007). They all deal only with translation (since it dominates in small regions). And to eliminate cloudy regions, cloud masking or threshold is used. The operational groups has reported some of the practical image registration issues, which are required to be addressed: effectiveness of normalized correlation in cross-band registration, adaption to thermal changes in satellite geometry and minor problems in orbit data, inadequate uniform sampling of control points across the image, and suitability of a specific ground location for correlation for different reasons. A ground location can't be suitable for correlation because the ground features are uniform and indistinct, because seasonal changes in temperature and vegetation cause the image to significantly vary, or the image may vary because of human activity.

In (Fujisada, 2005) the image registration system is described for the Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER), a 14-band multispectral imager launched in 1999 on the Terra (EOS-AM1) satellite. Registration was done to a database of about 300–600 CPs which were mapped onto topographic maps. The similarity measure used was normalized correlation with transformation limited to translation. Matches were rejected for correlation less than 0.7 or if clouds were detected.

The automatic orthoimage production system, ANDORRE is described in (Baillarin, 2005), using its algorithmic method; they called TARIFA (French acronym for Automatic Image Rectification and Fusion Processing) for the French satellite SPOT-5 launched in 2002. For matching, multi-resolution search strategy is used with the number of levels set to keep a 5×5 pixel size search window. Control points are automatically found, matched by correlation. Geometric outliers and control points with correlation coefficient below 0.80 are rejected.

2.3 Fourier Domain and other Transform Domain Based Methods

As shown in previous section, though various efforts are found in literature to improve correlation based method, computational complexity is need consideration. Very convincing alternative approach for fast computation of correlation as a similarity measure is frequency domain method, where a 2-D image has 2-D Fourier transform.

Depending on how pixel intensity values are varying, corresponding Fourier transform coefficients are generated. For lower variations in pixel intensity or smooth area of the image correspond to higher amplitude of lower frequency coefficient. Similarly, higher variations in pixel intensity in the image correspond to higher amplitude of larger frequency coefficient. Any transform domain based image registration methods are based on the premise that the representation in the transformed image will have capability to recover the geometric transformation easily. The core part of the Fourier domain methods is the Fourier's shift theorem. According to it, if there are two images f(x, y) and g(x, y) and the image g(x, y) is translated version of the image f(x, y) in spatial x-y domain, then the respective Fourier transform F(u,v) and G(u,v) in the frequency u-v domain are related by a phase shift that can recovered by efficient way. So, for image registration, the two images are first transformed to the frequency domain by computing their Fast Fourier Transform (FFT). Then after they are multiplied efficiently in the frequency domain, and the product is transformed back to spatial domain by computing inverse FFT; and by finding its peak the translation can be recovered. This way of computing correlation in frequency domain is equivalent to that of spatial domain but computational cost in frequency domain is now $O((n+m)^2\log(n+m))$ instead of the brute-force $O(n^2m^2)$ (Dewdney, 1978).

Though the relative computational efficiency of standard Fourier-based methods, compared to the brute-force computation of the correlation coefficient in the spatial domain, there are certain issues such as sensitivity to noise, geometric transformation beyond translation, displacement within pixel, large translation. Certain efforts are found to address them. (Chatterji, 1996) has extended the frequency domain approach so that rotation, translation and scale can be recovered within certain limit. In the method by (Qin-Sheng Chen, 1994), the shift-invariant Fourier transform and the scale invariant Mellin transform are combined and termed as 'Fourier-Mellin Transform'. The method is also called as symmetric phase-only matched filtering (SPOMF). One of the application of this invariant Fourier-Mellin transform to register Interferometric Synthetic Aperture Radar (In-SAR) images is found in (J., 2005), where larger signal to noise ratio is claimed than the standard correlation method. The authors in (Y. Keller, 2005) used the pseudo polar Fourier transform for computing rotation in a more robust and efficient manner.

The aliasing artifacts effects, which come from interpolation of rotated images, are addressed in (Harold S. Stone, 2003). This is required to increase to accuracy to subpixel level. The authors in (Wolberg, 2005) used the method applicable to large displacements. The authors in (Orchard, 2007) also addressed the recovery of large displacements and rotations, combining the Fourier transform and a gradient descent framework for an exhaustive global search. Robustness to multimodal radiometric differences in remote sensing and medical imagery is reported.

Basis functions of Fourier transform have infinite support and they do not have spatially localized frequency response. But basis functions of wavelet transform do have finite support and so they can localise the frequency response. Some of the wavelet basis functions Haar, Gabor, Daubechies, finite Walsh and Simoncelli wavelets have been used in image registration (H. Li, 1995), (Ping, 1997), (Zavorine, 2000); (Moigne, 2005); (Petrou, 2006). Fourier coefficients do have the information about the frequency of features but don't have any information exactly where it occur i.e. there is no spatial information. While wavelet transform provides combination of both frequency and spatial information. Though the wavelet transform don't have shift theorem similar to Fourier transform, so can't be used to recover the geometric transformation parameters. Instead it is used to decompose the image into higher-frequency, edge-like features, and lower-frequency features.

2.4 Mutual Information Based Approaches

In general, say for image registration in computer vision the required constraint of the images having equal intensities is valid. But this is not always valid for satellite images. For example in cross-band and cross-sensor image registration, the image intensities may be related by nonlinear, non-monotonic or even non-functional relationships. So to address this problem SAD, SSD and cross correlation do not work, so an appropriate alternative similarity measure is required. One class of such similarity measures is based on statistical relationships. In that the similarity measure is based on comparing local intensity distributions rather than individual pixel intensity values. Widely used such similarity measures are mutual information and its variants.

Mutual information based image registration is first introduced by (Collignon, 1995) and (Wells, 1995). Mutual information based methods are for the weak assumptions about the relationship between the pixel intensities in the two images and they do not require a monotonic or any functional model also. Mutual information is an information-theoretic quantity that measures the spread of the values in a joint histogram represented by a 2-D matrix. If two images match perfectly, then the joint histogram of their intensities should cluster around the diagonal of the matrix; otherwise, the values spread off the diagonal.

Many authors have worked to improve mutual information based image registration with reference to efficiency, accuracy, and robustness. In (Maes, 1999) ways of obtaining computational speed are presented by integrating multi-resolution pyramid approach with the gradient based approach. More efficiently an initial scale and translation is computed in

(Shams, 2007) using mutual information to gradient values. This is before refining the registration by a variation of Powell's numeric optimization.

The mutual information measure is also subject to scalloping artifacts that stem from interpolation errors, which may give false local optima. Thanks to the authors whose contribution can be treated as refinements to improve the computation of joint histogram, for example (mei, 2003), (Dowson N. a., 2006), (Liberata, 2008), (Dowson N. K., 2008), (Rajwade, 2009).

Mutual information as similarity measure has been widely used in medical imaging as it is having good potential for multimodal fusion. Some limited work or work on some limited datasets is found for mutual information based image registration with satellite images or remote sensing applications (Eastman, 2010). Still, research-oriented articles have demonstrated promising approaches on limited image datasets. The authors of (Kern, 2007) gave a thorough review of implementation details of mutual information, and developed a gradient descent version that was extensively tested on synthetically generated, multiband image pairs acquired by the Multispectral Thermal Imager (MTI) satellite.

The authors in (Cole-Rhodes, 2003) integrated MI with a wavelet pyramid, and used a stochastic gradient numeric optimization search approach to register multimodal, multiscale imagery acuired by IKONOS, Landsat ETM, MODIS, and Sea-viewing Wide Field-of-view Sensor (SeaWiFS). (mei, 2003) applied MI with different interpolations (nearest neighbour, linear, cubic, and partial volume) to register Landsat TM, Indian Remote Sensing Satellite (IRS), Panchromatic (PAN), and Radarsat Synthetic Aperture Radar (Radarsat/SAR) images over the San Francisco Bay Area, California (Partial volume interpolation was found to be most effective). In (Giro, 2004) systematically investigated the application of different similarity measures to multisensor satellite registration of SPOT and ERS-2 data, including correlation, correlation ratio, normalized standard deviation, MI and the related Kolmogorov measure. The authors in (Cariou, 2008) used MI with gradient descent to register airborne pushbroom sensor data to a reference orthoimage, computing a geometric transformation based on explicit airplane orientation parameters. Addressing the challenge of noisy SAR image is found in (Nies, 2008) by the application of mutual information.

2.5 Feature Based Methods

To overcome the main drawback of large computation cost for area based method, instead of every pixel, only specific features are used in feature based methods. The features can be a point, line, intersection of line, contour etc. The features should have the characteristics of distinctiveness and repeatability.

As shown in Fig. 1.3 of chapter 1, first of highly informative feature points are extracted from both images. The feature points are described by appropriate local image properties called as descriptor. Then the features are matched based on the descriptor. In literature the feature points are also termed as, control points, ground control points, and landmarks. Typically if the points are selected manually then they are termed as control points; if the points are based on some Geo informatics system or GPS data, then they are termed as ground control points. If the points are related to some well known place then termed as landmark points. While feature points are automatically extracted using some operator and for subsequent operation of matching of features its descriptor is used. For matching step appropriate similarity measure is required. The similarity measures used for the intensity based methods are different from that of the feature based methods. In area based method the similarity measure applied on spatial information, while in feature based methods the similarity measure is applied on the descriptor of the features. Individual matched pairs of features may have some incorrect matched pairs called as outliers; the correct matched pairs are called as inliers. Random Sample Consensus (RANSAC) (Bolles, 1981) is of the widely used algorithm for outlier removal.

The feature based method is said to be of low level features if they are of dense edge point of type. If smaller number of features having higher information content then the feature based method is said to be of high level features. In low level feature based methods comparatively it is easy to extract the features. But enough efforts are required in later stage of matching in large search space as well as to remove the outliers. The low level feature points are represented by their coordinates i.e. position; and they are described by some descriptor which is basically computed from the local neighbourhood.

Many articles are found in literature for the point pattern matching algorithm in the field of pattern recognition and image processing. The basic and simple similarity measures among unlabelled point are Hausdorff distance and its variants. The Hausdorff distance is a geometric measure.

In feature based methods if it is based on smaller, sparse sets of feature points then it require more effort to locate and describe the feature points. As a first step of features extraction, first of all it required to apply an interest operator to the images for finding distinctive points. The original images are converted into feature sets for computing an optimal match.

Syntactic interest operators look for distinctive regions with high-content information that can be localized in two directions simultaneously, such as corners or line intersections. They do so by computing local image properties (such as first and second image derivatives). Commonly used operators are (Harris, 1988) and (Forstner, 1987). In (Schmid C. M., 2000) systematically reviewed and evaluated a number of interest operators and concluded, based on the criterion they defined, that the Harris operator was the most repeatable and informative. In (Schmid M. K., 2004) extended the Harris detector to an affine invariant, scale space version (Harris-affine), that detects local maxima of the detector at multiple image scales. The authors in (Schmid M. K., 2005) reviewed six different affine invariant operators, including the Harris-affine, and concluded under what conditions each operator was most useful. In (Kelman, 2007) evaluated the three interest operators: Laplacian-of-Gaussian, Harris, and maximally stable extremal regions (MSERs), for the repeatability of their locations and orientations.

The topic is rich enough to support a book surveying interest and feature operators (Tuytelaars, 2008). However, much of the literature does not account for the complexity of remote sensing imagery. In (Hong S. H., 2006) emphasized the complexity of extracting GCPs in remote sensing data, as the extraction operator needs to account for the physical imaging models of the sensor, the satellite orbit, and the terrain's Digital Elevation Model (DEM) to accurately recover invariant feature points in SAR and optical images. Once feature points have been detected, a feature descriptor operator extracts local information to define a similarity measure for point-to-point matching. By reducing a neighbourhood to a smaller set of descriptors relatively invariant to geometric and radiometric transformations, the matching process can be made more efficient and robust. Local descriptors include moments of intensity or gradient information, local histograms of intensity of gradients, or geometric relationships between local edges.

Lowe defined the scale-invariant feature transform (SIFT) (Lowe D. , 1999) (Lowe D. G., 2004), which is a weighted, normalized histogram of local gradient edge directions invariant to minor affine transformations. The SIFT operator has been widely used in a number of registration applications. (Yang, 2006) used SIFT in an extension of the dual bootstrap iterative closest point (ICP) algorithm, originally developed for retinal image registration. Then the SIFT operator employed to obtain an extended version that performed successfully on the challenging image pairs that were experimented with. Following are the additional relevant papers: (Dufournaud, 2004), (Sukthankar, 2004), (Schmid M. K., 2005), (Schmid M. K., 2004). In remote sensing, there have been a number of applications of feature point detectors and descriptors. In (Grodecki, 2005) the authors have performed registration for redundant validation of IKONOS geometric calibration. They used two test sites with features points selected manually and automatically with the Forstner and Gulch interest operator of (Forstner, 1987). The work in (Bentoutou, 2005) registered SPOT and SAR images to subpixel accuracy, reported by small root mean square error (RMSE). They located feature points using an enhanced Harris detector, and then matched the points with affine invariant region descriptors with a subpixel interpolation of the similarity measure. The work in (Carrion, 2002) described the GEOREF system that the authors tested with Landsat imagery. In the GEOREF system, features extracted by the Forstner operator are located and matched at multiple levels using an image pyramid. An interesting twist is that once two feature points are matched, the match is refined in a local neighbourhood around the feature point by least-squares, area based method. Here, feature-based and area-based approaches are combined to improve the matching.

2.6 Geometric Transformations

If enough information regarding the image capturing model of the sensor, geometric distortion from satellite, atmospheric condition etc. is available then an appropriate transformation model can be selected, and accordingly suitable algorithm can be chosen (Solberg, 2005), otherwise an approximate geometric transformation model can be used. When an image registration algorithm is based on matching small image regions as chips or control points, empirical models of a simple geometric transformation as translation can be sufficiently accurate even without taking into account additional elements such as perspective. Thus, many image registration algorithms use empirical, low-order geometric models over small regions as chips or control points, and then use the control points to compute more accurate, parametric geometric models of higher order. Many algorithmic techniques, such as those based on numerical optimization or the Fourier transform, are primarily designed for the explicit recovery of the parameters of a low-order model, such as translation, rotation, and scale. Empirical models include:

- Rigid transformation with Rotation, scale, and translations i.e. four parameters (translation includes horizontal translation t_x and vertical translation t_y)
- Affine transformation of six parameters
- Projective transformation of eight parameters
- Homography, which consists of eight degrees of freedom
- Higher order 2D and 3D polynomial functions
- Rational polynomials

Similarity Transformation:

Similarity transformation represents global translational, rotational and scaling difference between two images defined by

$$X = S(x \cos\theta + y \sin\theta) + t_x \tag{2.2}$$

$$Y = S(-x \sin\theta + y \cos\theta) + t_y \tag{2.3}$$

Here (x,y) and (X,Y) are the coordinates of the points in both the images respectively; S, θ, t_x,t_y are scaling, rotational and translation differences between the images. In simple method the rotational difference between the images can be determined by angle between the lines connecting the points in the images. The scaling can be computed from the ratio of the distance the points in the image. After estimating scaling and rotational differences, solving equation (2.2) and (2.3), two unknowns t_x and t_y can be estimated.

Typically, the correspondences are noisy and inaccurate so more than two correspondences should be used and transformation parameters are estimated using least square or clustering method.

Rigidity Transformation:

If the similarity transformation doesn't have the scaling difference i.e. it has only translational and rotational differences then it is said to be rigidity transformation. So only θ, t_x and t_y are the geometric transformation to be estimated in case of rigid transformation.

Projective and Affine Transformation:

If image capturing is a projective process and if it has no associated nonlinearities, the relation between the two images can be modelled by projective transformation given by

$$X = \frac{ax + by + c}{dx + ey + 1} \tag{2.4}$$

$$Y = \frac{fx + gy + h}{dx + ey + 1} \tag{2.5}$$

From the correspondence, the projective transformation parameters a-h are estimated. At least four point correspondences are required under best conditions. If the scene is very far from the sensor, the projective transformation can be approximated as affine transformation given by

$$X = ax + by + c \tag{2.6}$$

$$Y = fx + gy + h \tag{2.7}$$

The further details are found in (Goshtasby, 2005). While the geometric transformations used for local matching are usually global and rigid, sometimes non-rigid or locally rigid transforms are important to account for distortions that vary across an image pair.

2.7 Re-sampling

Digital images are represented by values on discrete uniformly spaced rectangular grid. In image registration process if one image is to be transformed geometrically to match with the reference image or other image, the values have to be re-sampled to the new grid locations.

Such re-sampling is normally required at the end of finding the registration or geometric transformation parameters, which are applied on the sensed image or image to be registered with the reference image. In this re-sampling enough accuracy is necessary as it is for end product i.e. for registered image.

In addition to this, for the case of intensity based approach re-sampling is also required intermediately in all iterations. Here, even by compromising with the accuracy, faster re-sampling is required because it is for intermediate step to take decision related to the iterations and going to be deleted then after. Another issue related to the re-sampling is the re-sampling artifacts as it degrade the data as well as the accuracy of image registration.

Re-sampling is realised by reverse sampling of image values to re-map the sensed image to a new image.

If f (u, v) is the geometric transformation that maps the sensed image into the reference image, then the inverse transformation f^{-1} (x, y) is applied to map a pixel in the new sensed image to a sub-pixel location (u, v) between four neighbouring input pixels in the grid of the old sensed image. The values of these four pixels (or pixels in a larger surrounding neighbourhood) are then interpolated to compute the new pixel value. The basic interpolation methods are nearest neighbour, bilinear, bi-cubic and related details found in (Goshtasby, 2005).

2.8 Evaluation of Image Registration Algorithms

Performance of the image registration algorithm can be represented by several parameters such as accuracy, reliability, robustness, and computational complexity (Goshtasby, 2005). Accuracy is the difference between the true and estimated values. With reference to image registration accuracy can be expressed for the estimated registration parameters. Accuracy can be also referred to the mean or root mean squared distance between points in the reference image and corresponding points in the sensed image. Reliability is the number of times the algorithm succeeds in finding the correct answer with reference to the total number runs performed. For example if image registration is performed for n numbers of image pairs out of which m numbers of image pairs are registered correctly then the reliability of the algorithm is m/n. Robustness tells about the impact of variations in certain parameters on image registration. Robustness measures the degree of stability of the approach. It can be measured with respect to noise, illumination variations, occlusion, non-overlapping region etc. The computation complexity shows how much time is required to execute the algorithm.

The overall performance of the image registration depends on the performance of its components i.e. feature selection, feature correspondence etc. For example in feature correspondence the performance can be represented by true positive probability, which is termed as correct match rate in the community of image registration. Among the n numbers of features, correctly matched features are m then the correct match rate is given by m/n.

2.9 State of the Art in Image Registration for Remote Sensing

In (H. Goncalves, 2011), an automatic image registration method is proposed in which histogram-based image segmentation (HAIRIS) is proposed. Most of the methods used for segmentation are not giving enough performance specific to remote sensing applications. The paper has focused on the relaxation parameter which allows more efficient histogram based segmentation having in mind a posterior image registration procedure. Specific rigid body model is under consideration, so only rotation and translations are involved. The application is illustrated by applying the approach to the simulated rotation and translation on images. It is also applied to some multi-spectral, multi-sensor, multi-temporal datasets. The accuracy below 1 degree for rotation and sub-pixel level for translation is obtained.

In (Pattichis, 2007), the authors developed a frequency-domain model for the mutual information surface around the optimal parameters and used it to develop a robust gradient descent algorithm. The discrete derivatives of the mutual information surface tend to be extremely noisy. The problem persists even if the underlying probability density functions are estimated using kernel functions. Convergence from such noisy estimates cannot be guaranteed, except when the algorithm that is used for maximizing the mutual information surface is initialized in the neighbourhood of the actual maximum. So, in the work the authors found that the algorithm should be expected to converge, as long as the registration parameters are initialized to be within the correlation-length distance from the optimum.

In (S. Klein, 2010) (Klein, Staring, Murphy, Viergever, & Pluim, 2010), for medical field non-rigid image registration is performed using a variant of mutual information that is, conditional mutual information (cMI). For real clinical application cMI provides better results than gMI but at the cost of extra computational cost.

In (Reinartz, 2010), histogram based image registration method is used specific to the TerraSAR-X and Ikonos images acquired over urban areas. SAR images have different characteristics than normally used optical images; especially they are having speckle noise. It is relatively difficult to extract common features from SAR and optical images so instead of feature based method, intensity may be preferred. In intensity based method dominant components are similarity measure and optimization. The work is to obtain an automated mutual information based global (large shift) followed by a fine registration for the images acquired by TerraSAR-X and Ikonos sensors over dense urban area.

In (Zuliani, 2006) the work is related to detection of feature locations, where a novel, physically motivated curve/region descriptor suitable to establish image correspondence in a geometrically invariant fashion is introduced. The method is to estimate the image transformation parameters robustly in presence of large quantities of outliers.

Image registration using Cross Cumulative Residual Entropy (CCRE) as a similarity measure is investigated in (M. Hasan M. R., 2012) as it can accommodate images with varying contrast/brightness and faster convergence than the other information theory based similarity measures. Also, Parzen-window optimization is applied in the calculation of the gradients of the similarity measure directly. This enables the implementation of an optimization procedure directly based on partial volume interpolation.

In (Zhang Z. X., 2009), some of the area based and feature based algorithms are surveyed. It is followed by a novel approach for interest point matching for high resolution satellite images. In the survey of the methods it is summarized that the area based methods are having suitability only for images with little distortion, they can't deal with smooth areas and they have high computational complexity, while feature based methods are having effect of the existence of outliers, effect of noise on in matching and the feature descriptor must fulfil several conditions involving invariance, uniqueness, stability and independence. In the proposed approach, first 'super points' those points which have the greatest interest strength (i.e. which represent most prominent features), are extracted. Then a control network is constructed using these super points. Next, each remaining interest point is assigned a unique position with regard to the closest control point. And finally an iterative 'closest point' algorithm is applied to search for correspondences based on the position that has been assigned to each interest point.

In (Zhang G. H., 2007), area based matching and feature based matching are combined for image processing of high resolution satellite images. Cross-correlation matching and relaxation based image matching techniques are used. Two pairs of datasets of panchromatic images of IKONOS and multispectral images of Quickbird, are used. In the first step wavelet multi-resolution property is used to produce pyramid images.

In (David, 2010), an Automated Inter-sensor/Inter-band Satellite Image Registration-AISIR is proposed for registering the images acquired from the different sensors and at different frequency bands by addressing some of the associated issues. In AISIR a novel control point matching and matching function estimation schemes based on modified

Geman-McClue M-estimation scheme is used to improve robustness to outliers. It also used an iterative refinement based on the modified Geman-McClure objective function to improve the localization accuracy of control point pairs.

For the case of multi-sensor image registration, intensity based approach is used in (Uma, 2014). In the work Normalized Mutual Information (NMI) based rigid image registration performed using genetic algorithm and Particle Swarm Optimization (PSO) algorithm. Genetic algorithm is random based technique and it is complex to implement. PSO algorithm converges prematurely and weak local search ability is the main drawback of PSO. So use of hybrid GA-PSO algorithm is investigated.

In (Clausi, 2007), the authors have presented an automatic registration system ARRSI-Automatic Registration of Remote Sensing Images. The system is designed specifically to address the problems associated with registration of remotely sensed images captured from different sensors. It employed technique based on phase-congruency model in the control point's detection to address global and local contrast and illumination conditions that may affect the accuracy of the detected control points. An advancement of the RANSAC algorithm called Maximum Distance Sample Consensus (MDSAC) is also introduced in ARRSI to remove the outliers which reduces RMSE in less number of iterations than RANSAC. The system works by first selecting random subsets of control points, then after a tentative geometric transformation is computed. If the transformation consistently extends to a significant portion of the full set, then it is accepted as correct. Under the rotation condition the performance is degraded.

In (Bentoutou, 2005), control points are detected by improved version of Harris corner detector at regions where the gradient magnitude is high. It reduces the time complexity and results in lower RMSE. But the registration fails if the images don't contain distinctive features such as mountain scene.

In (S. Chen, 2011), registration parameters can be automatically tuned so that both fusion and registration can be optimised simultaneously using maximum likelihood approach. Algorithm converges in 17 to 42 iterations. But the approach is not as good as mutual information based method when the SNR is less than 10 dB. This is due to the fact that the mutual information based approach is having exhaustive search strategy which guarantees a global optimization.

One of the common strategies to non-rigid image registration is hierarchical subdivision which decomposes a non-rigid matching problem to local rigid transformation. While using mutual information in such decomposed or small images it loses its statistical consistency. So Information theoretical measures are proposed in (P, 2008), to identify the concerned problematic regions to overcome the problem of mutual information; this improves robustness, accuracy as well reduces the computational cost by efficient stopping criteria. In the proposed hybrid approach, two similarity measure mutual information and cross correlation are used to get advantage of both i.e. robustness of cross correlation and multi-modal use of mutual information.

Chapter 3

Area Based Methods for Image Registration

3.1 Overview

Choice of ABM or FBM is critical for image registration. It depends on the application as well as content and type of the image. In general, if salient features are available in the images then FBM can be preferred. It should be followed by accurate matching technique otherwise the image registration is degraded. If salient features are not available and/or can't be easily extracted from the images such plain area then it is difficult to use FBM. Further if the images are having regular patterns in some of the region then there are more chances to have incorrect matches in FBM. ABM is well adopted for medical field but the selection of similarity measure (also known as cost function) is crucial and plays important role. In image registration for varying illumination level such as multi-modal, multi-spectral cases, MI is reported as a good candidate for similarity measure (R. K. Gambhir, 2013).

In some of the image registration methods (Chatterji, 1996), (Sejdi, 2011) (R. Matungka, 2008), (Martucci, 2001), (Varshney, 2009) particular transformation is applied on whole image and then the properties of the transform domain are used to find the registration parameters. Such approach can be categorized as a separate category of say Transform domain Based Method (TBM) or it can be also be categorized under ABM as the first step is to transform the whole image (instead of some features only) domain or spatial domain to transform domain.

This chapter covers study of ABM for image registration. In this chapter section 3.2 covers use of MI as a similarity measure for image registration and section 3.3 covers the use of radon transform for image registration.

3.2 Mutual Information as a Similarity Measure

Different similarity measures are reported for the image registration application such as Sum of Squared Difference (SSD), Sum of Absolute Difference (SAD), Cross Correlation (CC), Normalized CC (NCC), Mutual Information (MI), Normalized MI (NMI) etc. If the images are having linear relationship between the image intensities then SAD, SSD and CC works well. For images with nonlinear or even non functional relationship between their intensities such as multi-modal or multi-sensor images SAD, SSD and CC are not effective, they leads to inaccurate or even failure of image registration. But mutual information can be used (R. K. Gambhir, 2013).

Mutual information was first introduced in information theory by Shannon in 1948. According to it mutual information can be used to measure the statistical dependence between the image intensities of corresponding pixels in two images. The definition of mutual information can be presented in various ways (Sabuncu, 2006). For two images A and B, if the concerned probabilities and entropies are given then mutual information can be defined by various ways:

$$I(X;Y) = \sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} = \sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) log \frac{p(x|y)}{p(x)} \quad (3.1)$$

$$I(X;Y) = H(X) - H(X|Y) \quad (3.2)$$

$$I(X;Y) = H(Y) - H(Y|X) \quad (3.3)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \quad (3.4)$$

$$I(X;X) = H(X) - H(X|X) = H(X) \quad (3.5)$$

Here I is mutual information, H is entropy and p is probability.

Mutual information is found maximum when the images are geometrically aligned. Image registration using mutual information was introduced by two independent groups: Collignoon and Maes et al. (F. Maes, 1997) and viola (Wells, 1995). Then after, mutual information is widely accepted as a similarity measure in image registration. For medical images, MI is widely used as similarity measure for image registration (J. P. W. Pluim, 2003), (S. Klein, 2010). For medical images MI become de facto similarity measure for image registration. For satellite images however, MI is used in some recent work only (Eastman, 2010). For satellite images various similarity measures are compared in (R. K. Gambhir, 2013), which shows superiority of MI.

Though mutual information is found to be more suitable candidate as similarity measure, a measure concern is its computational complexity. Further, it becomes worse for the images of large in size such as satellite images. One of the techniques to estimate mutual information is Maximum Likelihood Mutual Information (MLMI) (T. Suzuki, 2009). The technique is to estimate mutual information between two random one dimensional variables.

In this section of the chapter, mutual information is selected as a similarity measure and using it image registration is performed. To observe the effectiveness of MLMI estimation method, the mutual information is computed by the standard histogram based method also. At the end, computation time required to perform image registration is observed, using both the approaches of computing mutual information that is, MLMI and histogram. Computation time of image registration based on MLMI found to be less than that of histogram approach.

3.2.1 Maximum Likelihood Based Mutual Information

Maximum Likelihood Estimation

Suppose it is required to measure the true value of some quantity (x_T), for which the repeated measurements of this quantity are $\{x_1, x_2, \dots x_n\}$. The standard way to estimate x_T from the measurements is to calculate the mean value and set $x_T = m_x$. The maximum likelihood method provides a quite similar but general method for estimating parameters of interest from data.

Let, parameter space $\boldsymbol{\Theta}$,

$\boldsymbol{X_1, \dots, X_n}$: i.i.d.(independent identically distributed) observations

$\boldsymbol{\theta}$**:** an unknown parameter

$\boldsymbol{\Theta}$: The set of all possible values of θ

Joint pdf or pmf of $X_1, \dots, X_n$

$$f(x_1, \dots \dots x_n|\theta) = f(x_1|\theta)\, f(x_2|\theta) \dots \dots . f(x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

Likelihood Function of $\boldsymbol{\theta}$, For observed $\boldsymbol{\chi_1, \dots, \chi_n}$:

$$L(\theta|x_1, \dots \dots x_n) = \prod_{i=1}^{n} f(x_i|\theta)$$

The joint p.d.f. or p.m.f. is a function of $\chi_1,\ldots,\chi_n$ for given $\boldsymbol{\theta}$ while likelihood function is a function of $\boldsymbol{\theta}$ for given $\chi_1,\ldots,\chi_n$.

Log-likelihood function is defined as

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta)$$

Setting the derivative of L (θ) equal to zero and solving for θ, will give maximum value of likelihood function

$$\frac{d[\ln L(\theta)]}{d\theta} = 0$$

MLMI

Some of the methods for the estimation of mutual information are surveyed and compared in (Walters-Williams, 2009), (Verleysen, 2012). Accordingly, the estimation methods can be classified as parametric and nonparametric estimation. Parametric estimators are Bayesian estimator, edgeworth estimator, maximum likelihood estimator and least square estimator. Nonparametric estimators are histogram based estimator, adaptive partitioning of XY plane, kernel density estimator, B-spline estimator, nearest neighbour estimator and wavelet density estimator. For example in kernel density estimator the densities $p_{xy}(x, y), p_x(x)$ and $p_y(y)$ are estimated separately and using them mutual information is estimated, where the division may magnify the error.

In (T. Suzuki, 2009), new method for the estimation of mutual information based on maximum likelihood is proposed which is called as MLMI. This MLMI method has several advantages such as it does not involve separate density estimation and directly models the density ratio,

$$w(x, y) = p_{xy}(x, y) \,/\, p_{xy}(x)\, p_y(y) \tag{3.6}$$

Thus it is a single-shot procedure without division by estimated quantities and therefore the estimation error is not further expanded. Therefore the unique global optimal solution can be obtained efficiently. In (T. Suzuki, 2009), MLMI method is compared with kernel density estimator, k-nearest neighbor, edgeworth expansion; and it is shown that MLMI is better as compared to those methods. Additionally, using MLMI method, mutual information is estimated between two data vectors for various cases of their relationship such as linear dependence, independence, nonlinear dependence with correlation, and nonlinear dependence without correlation.

3.2.2 Image Registration Using MLMI

Use of MLMI method is investigated for images. MLMI method can be used for images also because the image can be represented by 2-D vector. And this estimated mutual information can be used as a similarity measure in the image registration process. In (Cochoff, 2002) and (S. Chen, 2011), maximum likelihood approach is used to obtain image registration parameter itself. In case of (Cochoff, 2002) it is not for satellite images. In case of (S. Chen, 2011), computation time is very large. In our approach maximum likelihood approach is not to find the registration parameter but to find MI, which requires relatively lower computation time.

For the image registration process it is focused only on mutual information as a similarity measure and corresponding computation time only.

To perform image registration process, one test image say reference image is taken. Second rotated image say sensed image is synthesized by applying small known arbitrary rotation. In image registration process, this sensed image is required to be aligned with the reference image. For that the sensed image is rotated in step within predefined range of angle. For every step of rotation mutual information between reference image and sensed image is found. The maximum mutual information is found from all mutual information values. The angle corresponding to this maximum mutual information value is the angle of rotation. So the sensed image is de-rotated by the same angle to align it with the reference image. In this way image registration is performed.

3.2.3 Results and Discussion

In the experiment, mutual information is estimated using two different methods-histogram based and MLMI. It is repeated for different images with different size which includes satellite images also. All simulations are carried out in MATLAB, on Pentium Dual-Core CPU with 2 GHz and 2 GB RAM. Simulation results of image registration process using MLMI based mutual information is shown in Fig. 3.1, while using histogram based mutual information is shown in Fig. 3.2. For both methods and for different images, the processing time is observed. It is summarized in Table 3.1. It shows the processing time is less for image registration using MLMI based mutual information as compared to the image registration using histogram based mutual information.

Figure 3.1 in each image dataset, first image is original, second image is generated by applying known rotation and third image is de-rotated by MLMI method after finding the rotation

Figure 3.2 in each image dataset, first image is original, second image is generated by applying known rotation and third image is de-rotated by histogram based MI method after finding the rotation

Comparison with the other result is difficult due to wide selection of image registration components (or modules), techniques and geometric transformation under consideration. Accuracy is also important but it highly depends on interpolation and optimization techniques used in image registration. And hence comparison of accuracy is also difficult.

Here the focus is only on computation time of MI as a similarity measure as it is very important in image registration process for large images such as satellite images.

3.3 Transform Domain Based Image Registration

As discussed in chapter 1, in general the approaches for image registration are classified as area or intensity based method and feature based method. It is found in some of the scientific studies that the problem sometime becomes simper in another transform domain. This is also applicable to image registration. So a third category of image registration approaches can be considered as transform domain based image registration. Choice of method highly depends on the specific application and the image contents. Application of Fourier transform to estimate registration parameters are found in (Chatterji, 1996) and improved in (Varshney, 2009), (Martucci, 2001).

In this section after case study of Fourier transform based approach, rotation and translation invariant properties of radon transform are used to find the amount of rotation and translation required to align the images, i.e. to perform image registration. Simulation results are shown for different images, with different amount of rotation and translations, to show the accuracy and reliability of the method. Again noise level is also varied, to observe the robustness of the method to noise. The required average computation time is in seconds, depending on the size of images.

Table 3.1 Computation Time for Image Registration using MI as Similarity Measure

Images	Rotation applied to sensed image (degree)	Steps for Rotation (degree)	Computation Time for MLMI (second)	Computation Time for histogram based mutual information (second)
image-1 1024X1024	-4	-5 to 5	355	610
image-2 512X512	2	-3 to 3	77	144
image-3 2091X2018	2	-3 to 3	474	730

Images	Rotation applied to sensed image (degree)	Steps for Rotation (degree)	Computation Time for MLMI (second)	Computation Time for histogram based mutual information (second)
image-4 512X512	-2	-3 to 3	109	141
image-5 512X512	-4	-9 to 9	167	478
image-5 512X512	-6	-8:8	94.2	176.5
image-5 512X512	7	-9:9	105.7	226

3.3.1 Fourier Transform Based Image Registration

To overcome the problem of high computation time of area based method (i.e. ABM), many algorithms are proposed using transform domain. Image registration using FFT is proposed by (Chatterji, 1996). Improvements in Fourier transform based image registration algorithm are found in (Martucci, 2001) and (Varshney, 2009). Image registration using fractional Fourier transform is found in (W. Pan, 2009). In (Varshney, 2009) image registration using Fourier transform is shown to be better than mutual information based method. Some image registration algorithms FFT based, contour based, wavelet based, Harris based are compared in (R.M. Ezzeldeen, 2010). Image registration algorithms based on DCT, Haar, DWT transform and correlation are compared in (H B kekre, 2012). Image registration using hough transform and phase correlation is evaluated in (B Summar, 2005).

To investigate the scope of research in transform domain based image registration for satellite images, FFT based image registration (Chatterji, 1996) is studied for learning purpose. It follows the following steps:

(1) Read image I_1

(2) Read image I_2

(3) Take FFT of I_1 and shift it to center on zero frequency
(4) Take FFT of I_2 and shift it to center on zero frequency
(5) Perform convolution between the magnitude of step (3) with high pass filter
(6) Perform convolution between the magnitude of step (4) with high pass filter
(7) Perform transformation of (5) into log polar space
(8) Perform transformation of (6) into log polar space
(9) Take FFT of (7)
(10) Take FFT of (8)
(11) Compute phase correlation between (9) and (10)
(12) In (11) find the location of peak of the phase correlation
(13) Compute angle (360/Y-size of image)*y from (12)
(14) Rotate the image from (2) by an angle obtained in (13)
(15) Rotate the image from (2) by (180 + an angle obtained in (13))
(16) Compute FFT of (14)
(17) Compute FFT of (15)
(18) Compute phase correlation of (3) and (16)
(19) Compute phase correlation of (3) and (17)
(20) Find the location (x,y) in (18) of the peak of the phase correlation
(21) Find the location (x,y) in (19) of the peak of the phase correlation
(22) If phase peak in (20)>phase peak in (21), (y,x) from (20) is the translation. Else (y,x) from (21) is the translation and also if the angle from (13) < 180, add 180 to it else subtract 180 from it.

The steps can be also understood by Fig. 3.3. The corresponding simulation result is shown in Fig. 3.4 and Fig. 3.5. This shows the execution time is extremely low but the accuracy is also poor. The estimated angle of rotation is 2.8125 degree instead of the actual angle of 3 degree.

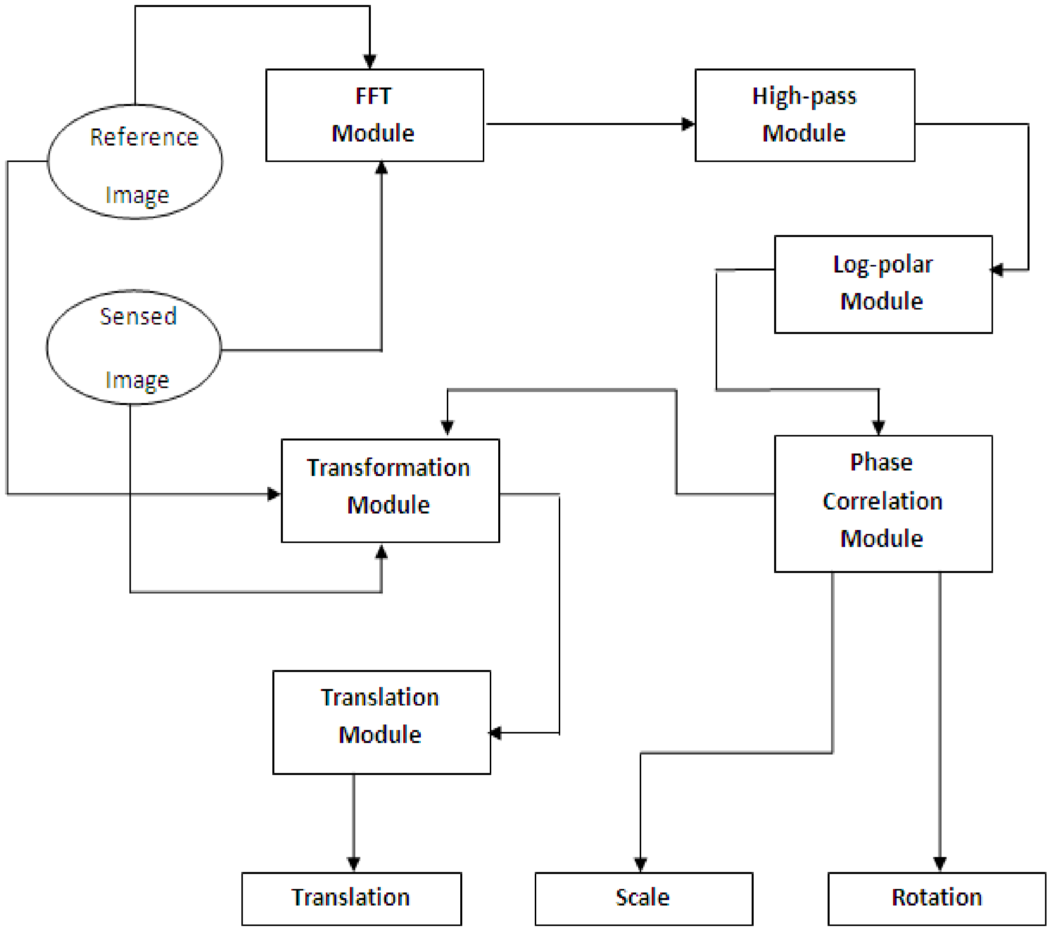


Figure 3.3 R, S, T parameter extraction using FFT

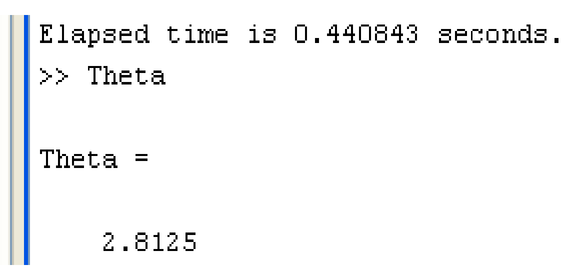


Figure 3.4 Simulation result on command window of MATLAB for FFT based image registration: original image, 3° rotated image, aligned image and execution time

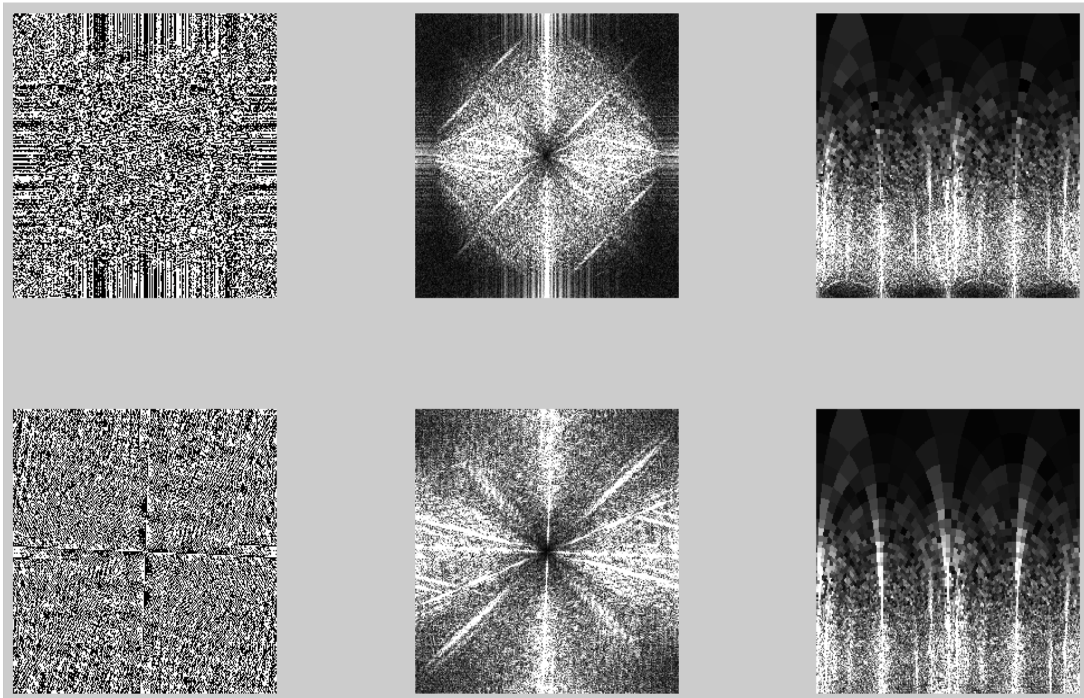


Figure 3.5 Simulation result for FFT based image registration: FFT, high-pass filtered and log-polar transform for both images

3.3.2 Radon Transform and Its Properties

The Radon transform computes the projections of an image matrix along specified directions as shown in Fig. 3.6. Mathematically, for a 2-D function f (x, y), the radon transform is defined as

$$R_f(r,\theta) = \int f(x,y)\,\delta\big(r - xcos(\theta) + ysin(\theta)\big)dxdy \qquad (3.7)$$

$$= F^R(r,\theta)$$

$$x, y, r \in R \; and \; \theta \in [0,\pi]$$

Equation (3.7) can be interpreted as the line integral of function f (x, y) over the line

$$r = (xcos(\theta) + ysin(\theta)) \qquad (3.8)$$

The lines are defined by the perpendicular distance from the origin, r, and the angle that r makes with the horizontal axis, θ.

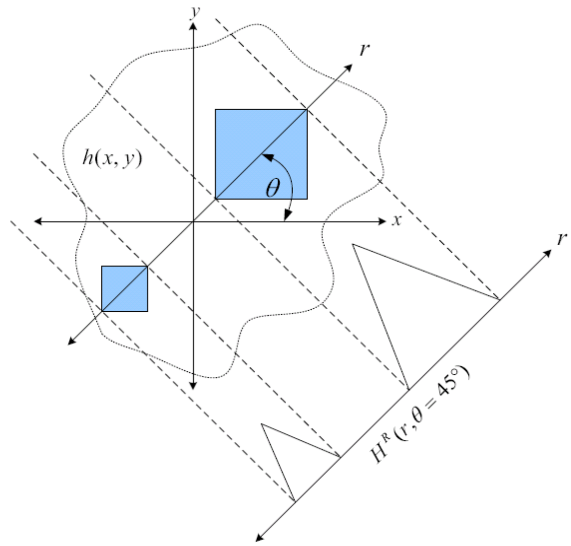


Figure 3.6 Radon transform - projection

As explained in (Deans, 2000), any translation in spatial domain leads to translation in the r direction of Radon domain. The amount of the translation varies with the theta dimension. The scaling of the original image along both axes results in the scaling along the r axis in the Radon domain. The value of the transform is also scaled. The rotation in spatial domain leads to circular translation along the theta axis in the Radon domain.

$$R\{f_\varphi(x,y)\} = F^R(r, \theta - \varphi) \tag{3.9}$$

$$R\{f_T(x,y)\} = F^R(r - T_x cos\theta + T_y sin\theta, \theta) \tag{3.10}$$

$for\ \theta = 0\ degree,$

$$R\{f_T(x,y)\} = F^R(r - T_x, \theta) \tag{3.11}$$

$for\ \theta = 90\ degree,$

$$R\{f_T(x,y)\} = F^R(r + T_y, \theta) \tag{3.12}$$

Table 3.2 Properties of Radon Transform

Image	**Input image f**	**Radon Transform F^R**
Original	$f(x, y)$	$F^R(r, \theta)$
Translated	$f(x - x_o, y - y_o)$	$F^R(r - x_o\cos\theta - y_o\sin\theta, \theta)$
Scaled	$f(ax, ay)$	$\frac{1}{\lvert a\rvert} F^R(ar, \theta)$
Rotated	$f(r, \emptyset + \theta_0)$	$g(r, (\emptyset + \theta_0\ , \mathrm{mod}\ 2\pi))$

The properties of the Radon transform are summarized in Table 3.2. The Fig. 3.7-3.11 explain these properties. Fig. 3.7 is estimate amount of rotation. Fig. 3.8 is to observe the effect of different amount of vertical translation t_x in the radon domain; then they are superimposed in Fig. 3.9. Similarly, Fig. 3.10 is to observe the effect of different amount of horizontal translation t_y in the radon domain; then they are superimposed in Fig. 3.11.

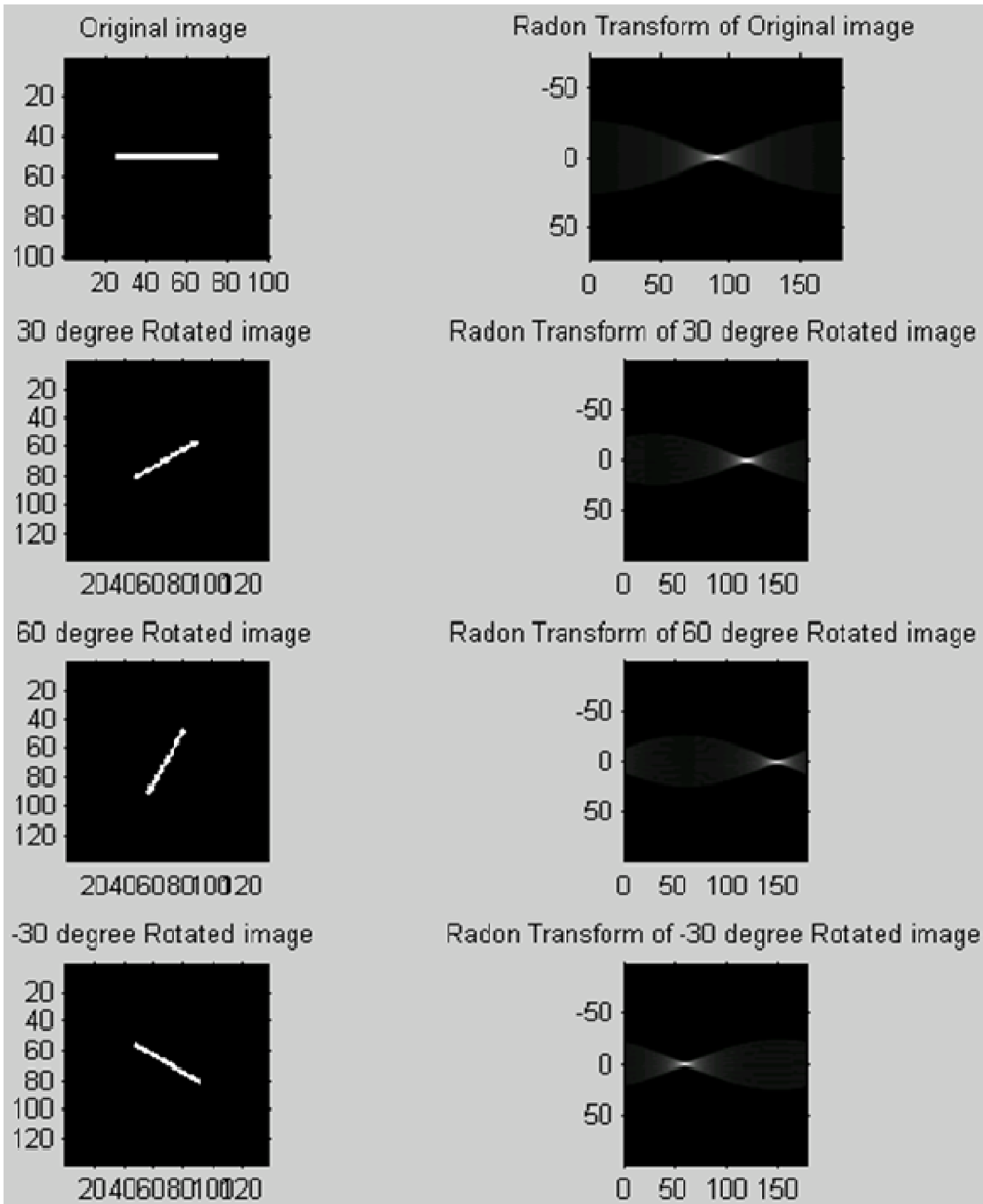


Figure 3.7 Effect of rotation is observed in radon domain

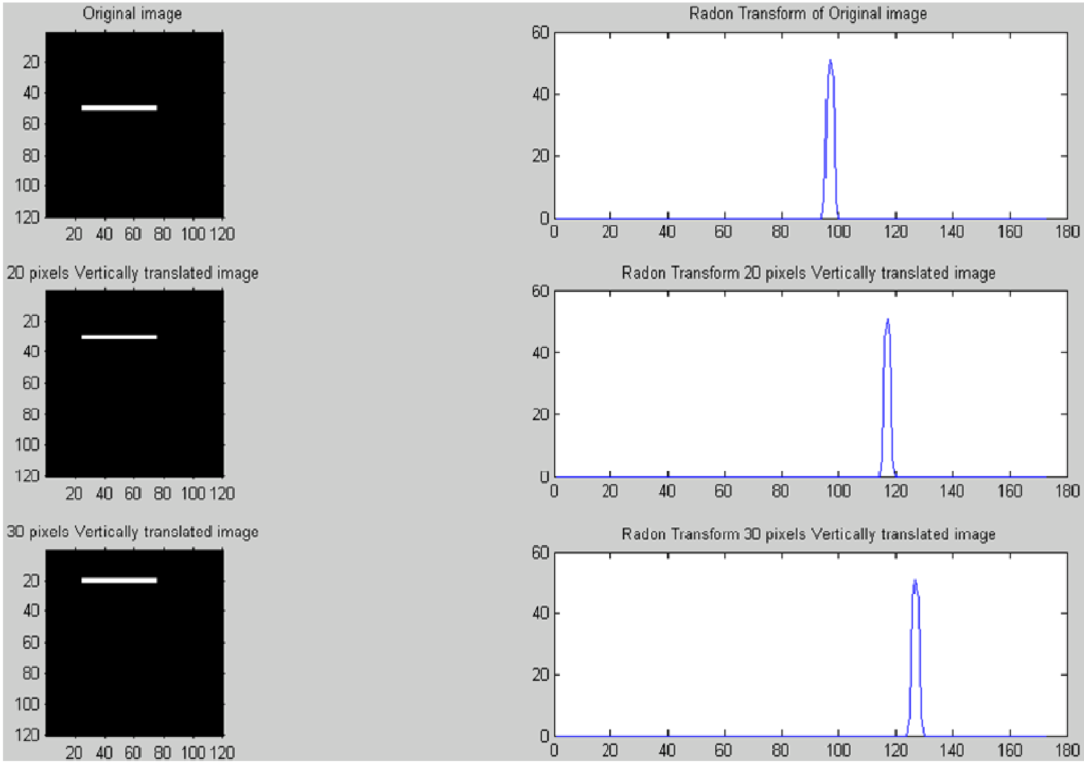


Figure 3.8 Effect of vertical translation is observed in radon domain

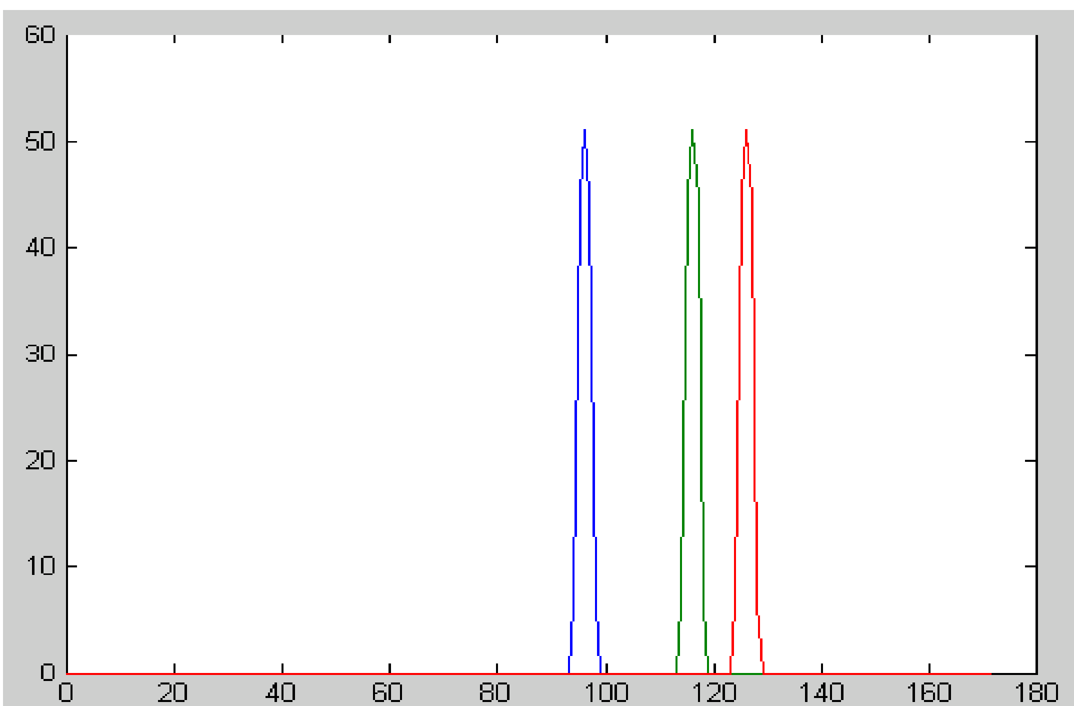


Figure 3.9 Horizontal distances can be used to estimate the amount of vertical translation

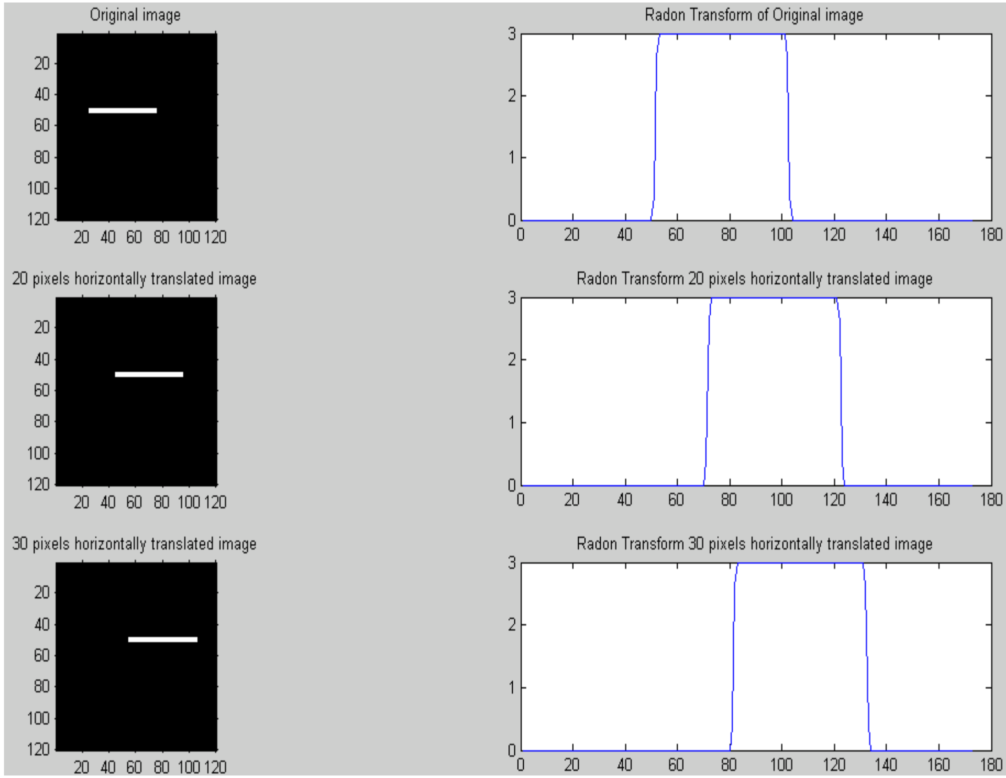


Figure 3.10 Effect of horizontal translation is observed in radon domain

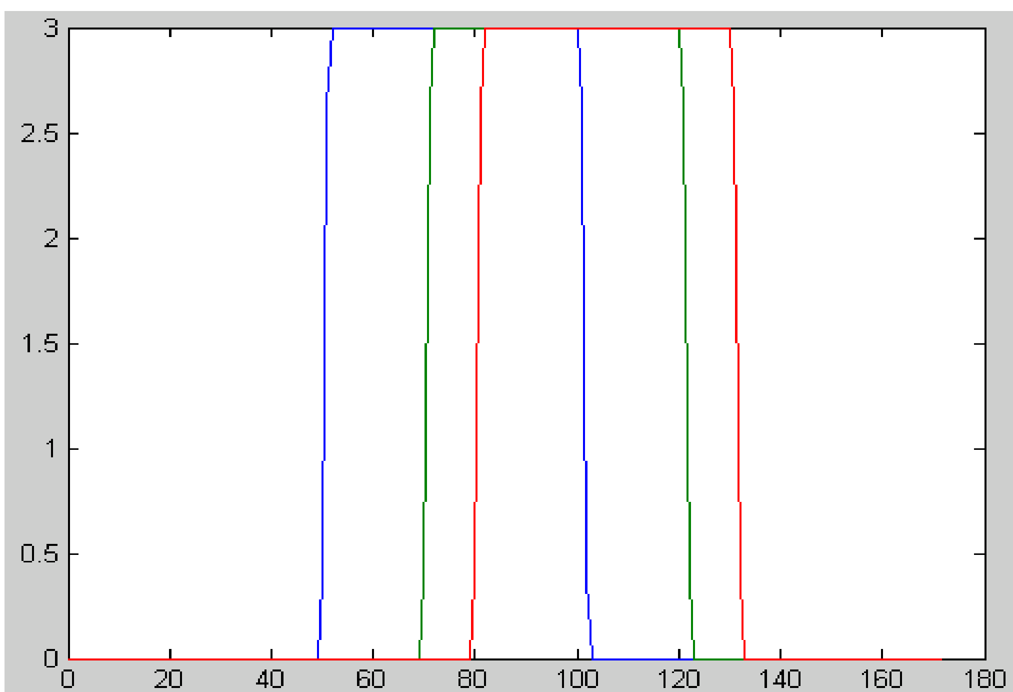


Figure 3.11 Horizontal distances can be used to estimate the amount of horizontal translation

3.3.3 Image Registration Using Radon Transform

Radon transform has been used in many applications. Rotational invariance property of Radon transform is used to estimate the orientation of printed text in (Soltanian-Zadeh, 2005). Radon transform is used for motion blur estimation application in (Krahmer, 2006). In (Pradip M. Patil, 2012) radon transform is used in robust shoeprint matching algorithm. In (Y. C. Chen, 2013), properties of radon transform are used for clustering application. In (Shehan Fernanado, 2011) radon transform is used for image registration application.

Four different images are selected with different resolution and size. Different amount of rotations and translations have been applied on these images. Correspondingly different images with θ degree rotated, t_x pixels translated in the x direction and t_y pixels translated in the y direction have been synthesized. Between the synthesized images (geometrically transformed) and original image, the radon transform is operated and using its properties θ, t_x and t_y are estimated to perform image registration.

All simulations are carried out in MATLAB, on Pentium Dual-core CPU with 2 GHz and 2 GB RAM

3.3.4 Results and Discussions

Using the properties discussed in previous section, the θ, t_x and t_y have been estimated. Among the other images, for two images i.e. image-1 and image-2, Fig. 3.12-3.16 describes the approach by their respective intermediate images for the first image. Similarly Fig. 3.17-3.21 describes the approach by their respective the intermediate images for the second image. The known actual values and the values estimated from simulation are summarized in Table 3.3. In most of the cases the estimated parameters values are very close to the actual values. This shows the accuracy and reliability of the approach. Computation time is also comparable as far as image registration is considered. The Table 3.3 is without noise. To observe the effect of noise on the performance, same steps have been repeated after adding various amount of noise in the images. Table 3.4 shows the corresponding results

Figure 3.12 Original image-1 of 489X300

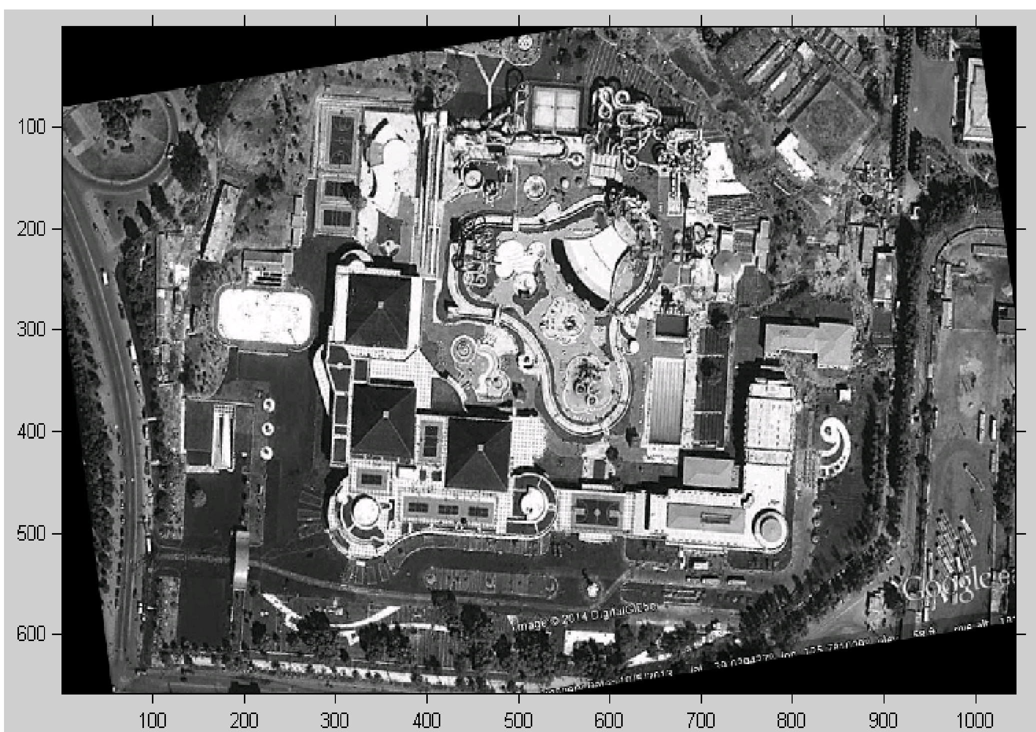


Figure 3.13 Rotated and translated image-1

Figure 3.14 De-rotated image-1 after estimating angle

Figure 3.15 De-rotated and de-translated image-1

Figure 3.16 For image-1, both images (i.e. original and geometrically corrected) have combined after geometric correction with the average pixel intensity value

Figure 3.17 Original airport image-2 of 579X481

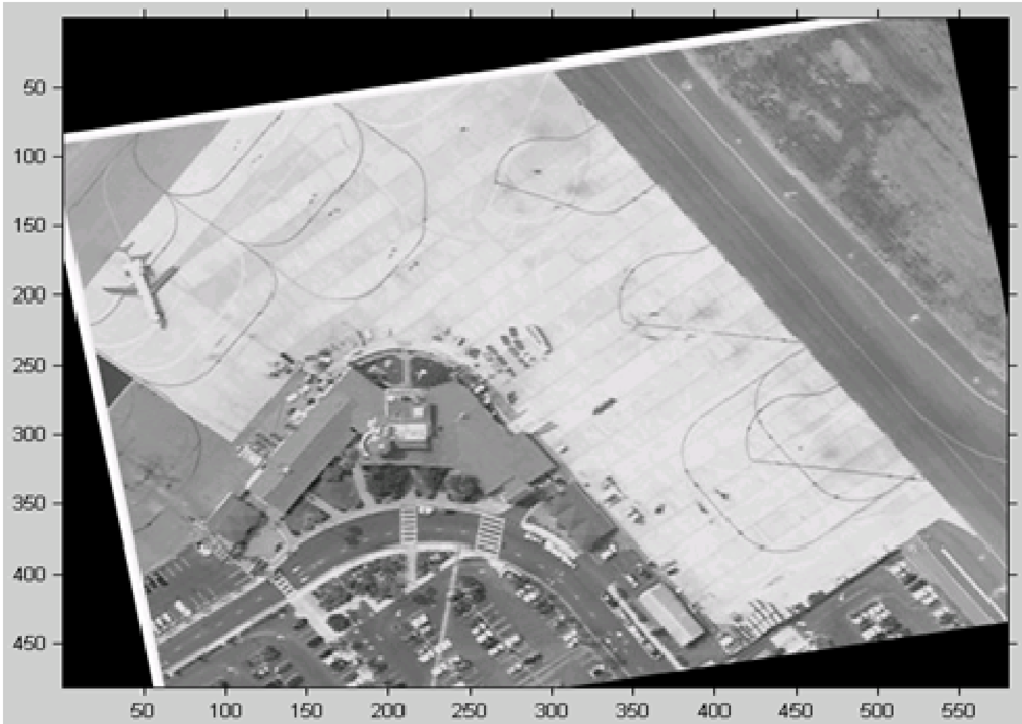


Figure 3.18 Rotated and translated image-2

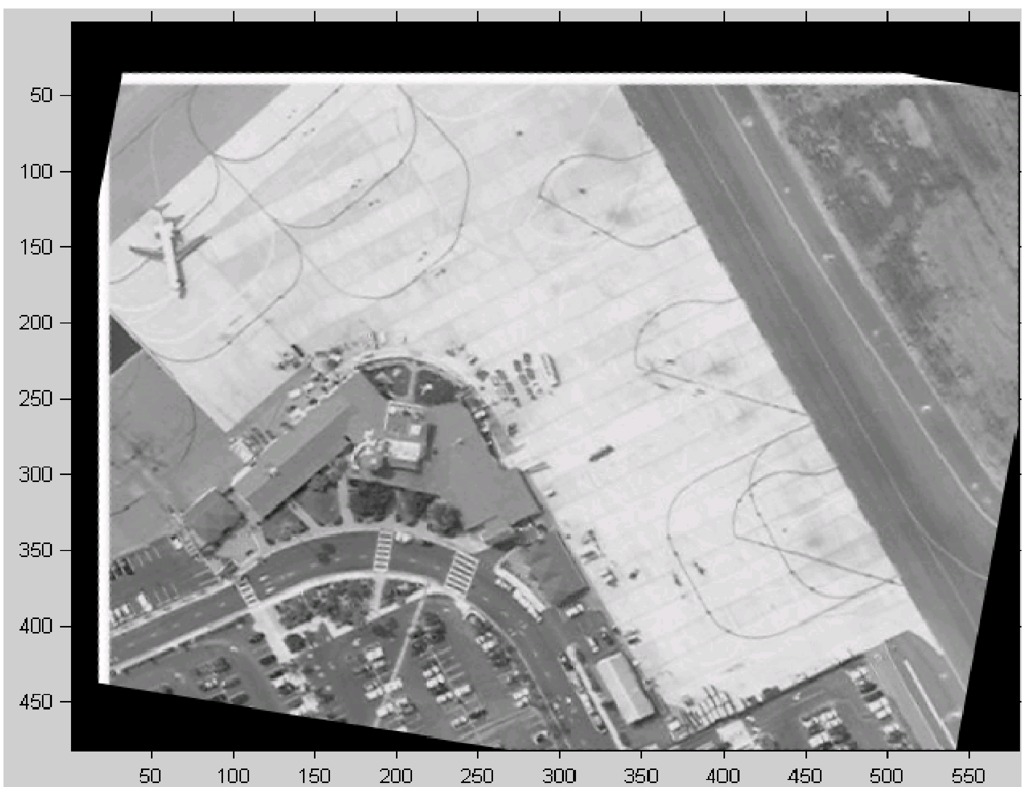


Figure 3.19 de-rotated image-2 after estimating angle

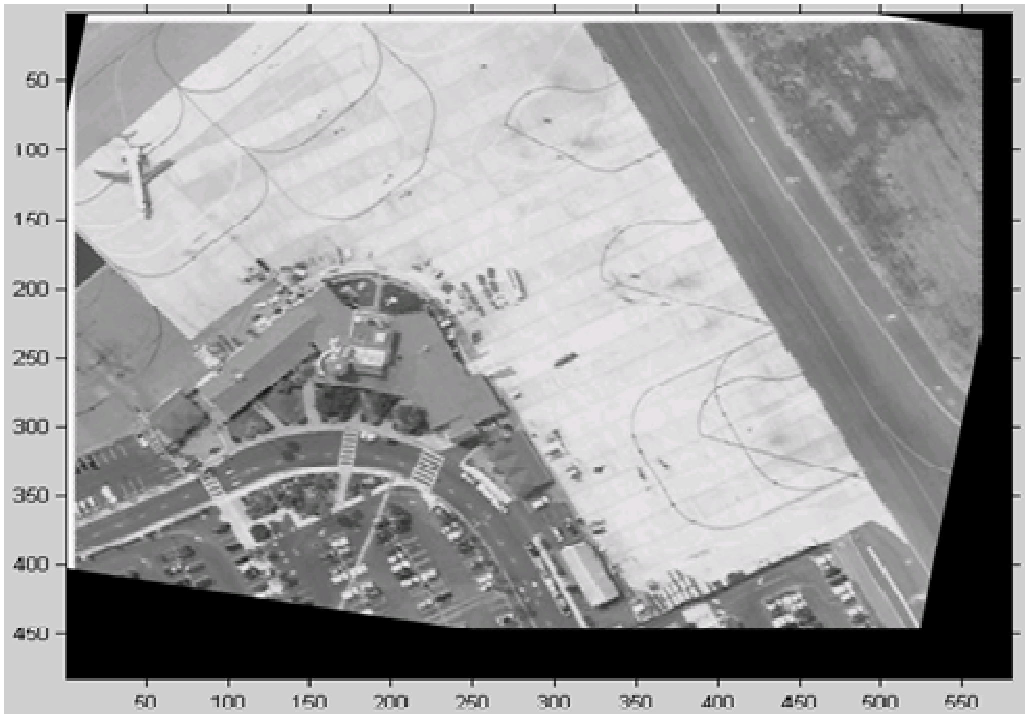


Figure 3.20 de-rotated and de-translated image-2

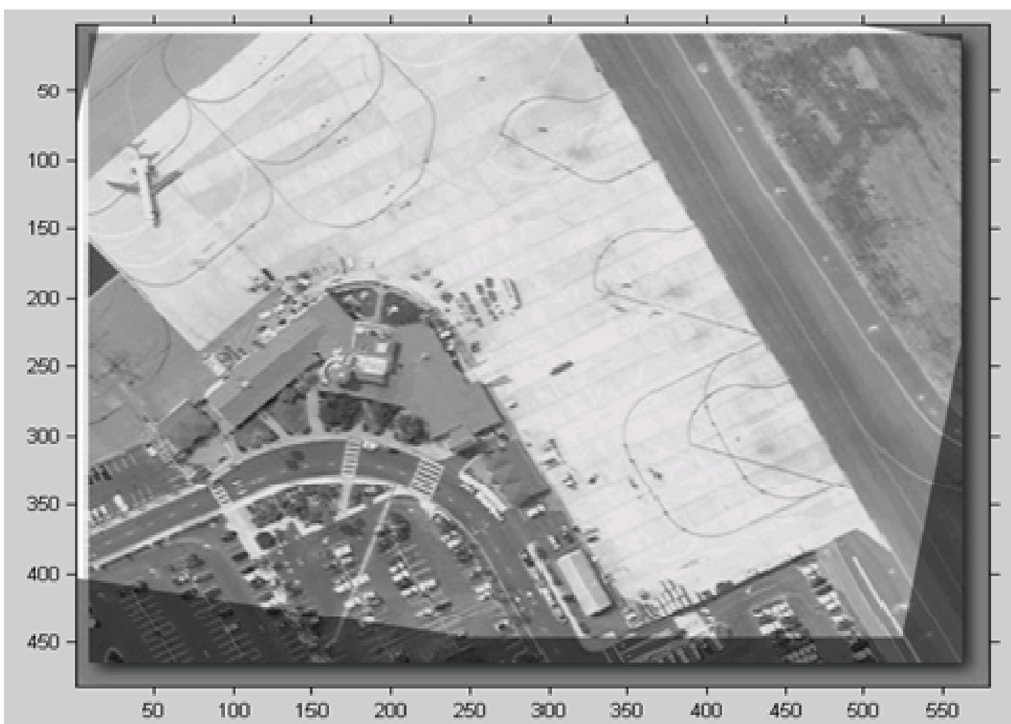


Figure 3.21 for image-2, both images (i.e. original and geometrically corrected) have combined after geometric correction with the average pixel intensity value

3.4 Summary

Image registration is performed using mutual information as a similarity measure. Two methods are used for the estimation of mutual information. Image registration is performed on various images and computation time is observed. It is observed that image registration which uses MLMI technique for mutual information estimation requires less computation time than the image registration which uses histogram based mutual information estimation. In this paper during the image registration process rotation is performed in step and accordingly maximum mutual information is found. This can be done by some optimization algorithm which maximizes the mutual information.

Table 3.3 Actual and Estimated θ, t_x and t_y Parameters for various images with Its Computation Time

Images	Actual Parameters			Estimated Parameters			Computation time (Second)
	Θ (degree)	t_x (pixel)	t_y (pixel)	θ' (degree)	t'_x (pixel)	t'_y pixel)	
image-1 489X300 125kb	0	20	35	0.5	20	27	4.3
	20	5	17	18.5	7	15	4.24
	10	35	17	10.5	34	17	4.38
	25	14	30	23.5	15	19	4.2
image-2 579X481 43.7kb	0	20	35	0.5	20	35	8.35
	20	5	17	20	4	18	7.78
	10	35	17	10	35	18	7.91
	25	14	30	25	14	31	7.73
image-3 3264X2448 467kb	0	20	35	0.5	16	37	223.52
	20	5	17	20.5	6	20	208.32
	10	35	17	10.5	13	18	214.3
	25	14	30	25.5	15	33	206.63
image-4 256X256 65kb	0	20	35	0.5	20	35	2.42
	20	5	17	18	5	16	2.20
	10	35	17	9.5	34	17	2.25
	25	14	30	24	13	29	2.19

For images which have been geometrically misaligned in terms of rotation and translation, the amount of rotation and translations are estimated using the properties of radon transform. Using these parameters, the images can be aligned or registered. For the different combination of rotation, translation and images, the average error in the estimation is small. This shows the accuracy and reliability of the approach. Further the effect of noise is relatively low which shows robustness of the approach to noise.

Table 3.4 Actual and Estimated θ, t_x and t_y Parameters For various images with gaussion noise

Images	Actual parameters			Estimated Parameters with Gaussian noise of variance 0.001				Estimated Parameters with Gaussian noise of variance 0.005			
	θ (degree)	t_x (pixel)	t_y (pixel)	θ' (degree)	t'_x (pixel)	t'_y (pixel)	Computation time (Second)	θ' (degree)	t'_x (pixel)	t'_y (pixel)	Computation time (Second)
image-1 489X300 125kb	0	20	35	0.5	20	27	4.6	0.5	20	26	4.5
	20	5	17	18	7	15	4,3	18	7	14	4.3
	10	35	17	10.5	34	17	4.27	10.5	34	17	4.29
	25	14	30	23.5	15	19	4.13	23.5	15	18	4.12
image-2 579X481 43.7kb	0	20	35	0.5	20	35	8.4	0.5	20	35	8.35
	20	5	17	20	4	18	7.8	20	4	18	7.98
	10	35	17	10	35	18	8.15	10	35	18	8.09
	25	14	30	25	14	31	8.19	25	14	31	8.1

CHAPTER 4

Feature Based Method: HOG as Feature Descriptor

4.1 Overview

Keeping in mind the large computational load of ABM, comparatively FBM is preferred for satellite images and remote-sensing applications. In FBM, salient features of the images such as points, lines, edges etc. are detected and corresponded to find the required geometric transformation parameters. Relatively, this is faster and works well if salient features are available in the images, as only those features are required to proceed further.

As discussed chapter 1, the steps in any FBM for image registration: feature extraction, feature matching (using descriptor of the extracted features), geometric transformation estimation and re-sampling. In image registration, depending on application, selection of feature extraction, its descriptor and matching method play important roles. The error in any of the steps of image registration is propagated in the next steps; accordingly the accuracy of image registration is reduced. Accuracy of consequent processing say, change detection depends on accuracy of image registration.

For multi-modal, multi-sensor, multi-spectral satellite images one of the challenges for image registration is varying illumination level according to the sensor characteristics, which reduces CMR. In this chapter the challenge is addressed by using Histogram of Oriented Gradient (HOG) along with Speeded-Up Robust Feature (SURF). It is shown that illumination variation gives some incorrect matches with SURF only which degrades image registration. Incorrect matches are reduced by using HOG as descriptor in SURF. Supporting simulation results for satellite images are presented which show the improvement in the correct matching rate.

4.2 Feature Extraction using SURF

During the last decade, Scale Invariant Feature Transform (SIFT) (Lowe D. G., 2004) and SURF (Herbert Bay, 2008) have been widely used for point feature extraction. As per survey in (Vaithiyanathan, 2016), thousands of papers have cited the SIFT method in their work, year wise statistics are shown in Fig. 4.1. Then after, in the paper specifically variants of SIFT based image registration approaches are also compared. SIFT has different variants such as PCA-SIFT, GSIFT, CSIFT, ASIFT, SURF. The variants are some modification/s in one or more steps of SIFT with regard to some specific application or objective. For example in (Sukthankar, 2004) a well established technique of dimensionality reduction Principle Component Analysis (PCA) is used to reduce the dimension of the 128-D SIFT descriptor to a lower dimensionality, which results in reduction in matching time and hence increase in overall speed. Some of the variants are also compared in (Wu J., 2013) in which SURF is shown to be fastest; and this is one of the requirements for large satellite image processing.

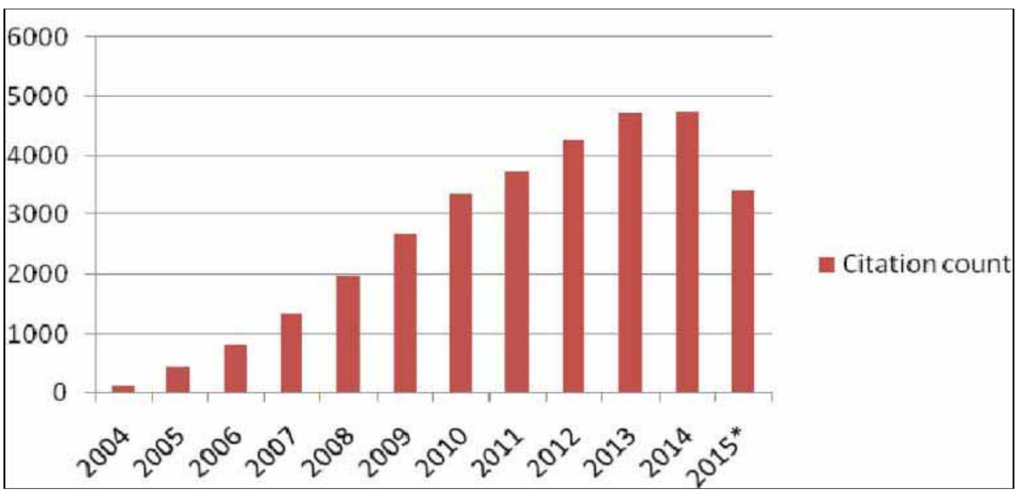


Figure 4.1 SIFT Based Techniques: Citation count,*upto mid of 2015 (Vaithiyanathan, 2016)

SURF is basically derived from SIFT with some improvements which are obtained by involving integral image, Haar wavelet response and approximation of hessian matrix. The most important characteristics of SURF is three times faster speed than SIFT, simultaneously it has good performance of repeatability, distinctiveness and robustness.

In general there are four steps in SURF: keypoint detection (also known as interest point or feature point), orientation assignment (optional step), local descriptor and keypoint matching (using its descriptor). The keypoint detector is based on the determinant of hessian. SIFT approximated laplacian of Gaussian using difference of Gaussian while SURF approximated it using box filter.

The work in this chapter is related to the descriptor. In SURF, descriptor is generated in 20s square region around the keypoint where s is the concerned scale. The region is divided into 4X4 square sub-regions. For each sub-region, Haar wavelet response in horizontal direction d_x and in vertical direction d_y are computed from 5X5 sample points. Finally, the responses and their absolute values are summed for each sub-region and accordingly 4-D descriptor vector (Σd_x, Σd_y, $\Sigma|d_x|$, $\Sigma|d_y|$) is formed. By combining this 4-D descriptor vector for all 4X4 sub-regions, results in a descriptor vector of length 64. Various stages of the SURF approach can also be explained in the Fig. 4.2 (a) (Elsalamony, 2015), while 4-D descriptor vectors for three different kinds image details are shown in Fig. 4.2 (b).

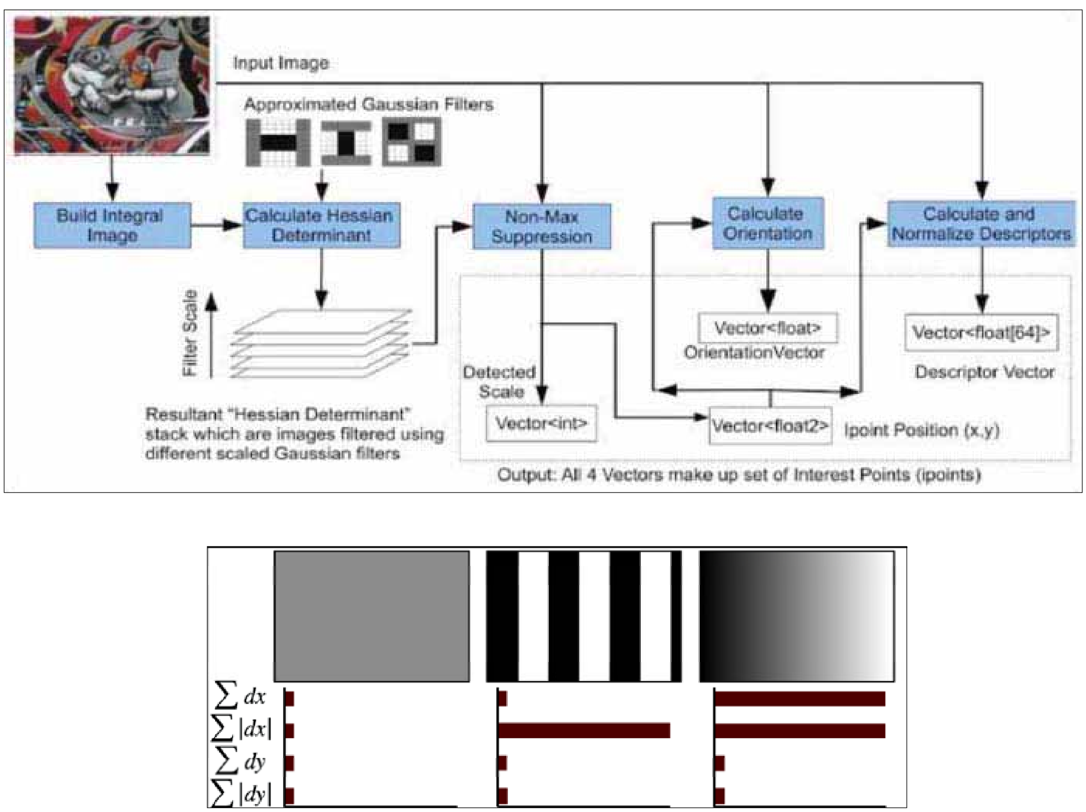


Figure 4.2 (a) Block diagram representation of SURF (b) Haar based response used to prepare 4-D descriptor vectors for three different kinds of image details

4.3 Feature Description using HOG

In SURF, Haar response based descriptor is used. Some alternatives for feature descriptors are compared in (Schmid M. K., 2005). In (Triggs, 2005), Histogram of Oriented Gradient (HOG) is used as feature descriptor for human detection.

The essential thought behind the HOG is that the local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The implementation of these descriptors can be achieved by dividing the image into small connected regions, called cells, and for each cell compiling a histogram of gradient directions for the pixel within the cell. The combination of these histograms then represents the descriptor. For improved performance the local histogram can be contrast-normalized by calculating a measure of the intensity across a larger partially overlapping region of the image. This normalization results in better invariance to changes in illumination or shadowing.

In (Triggs, 2005) for gradient computation, filtering is performed using the kernels, $D_x = [-1\ 0\ 1]$ and $D_y = [1\ 0\ -1]^t$. The steps for computing HOG descriptor is shown in Fig. 4.3. In (Triggs, 2005), various parameters are analyzed in detail before finalizing their values. The comparisons with other descriptor on standard dataset MIT and INRIA are shown in Fig. 4.4. Nine number of histogram channel is also suggested in their experiments. Fig. 4.5 shows increase in the numbers of orientation bins increases performance significantly up to about 9 bins. Effect of variations in cell size and block size is also presented in Fig. 4.6.

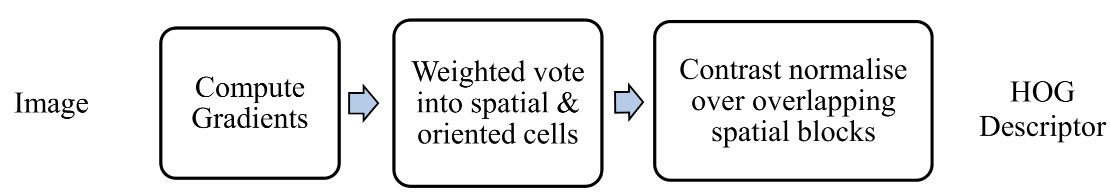


Figure 4.3 Steps for HOG descriptor

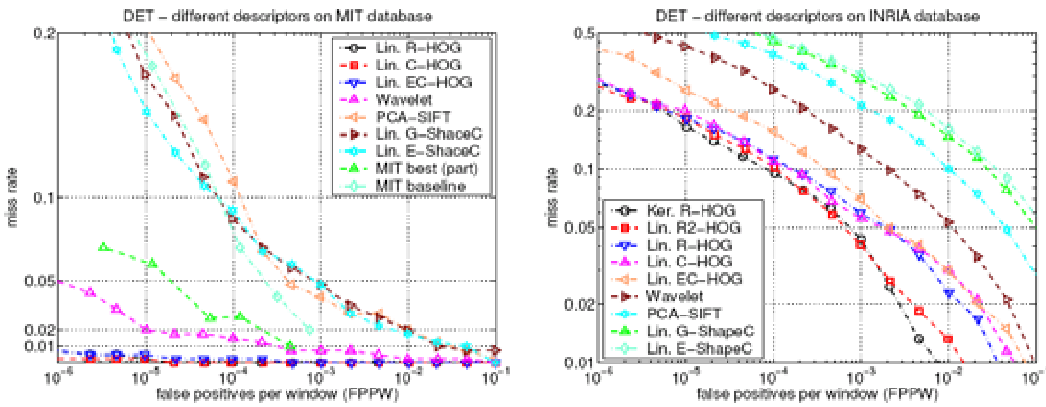


Figure 4.4 Performances of selected detectors on MIT and INRIA data sets (Triggs, 2005)

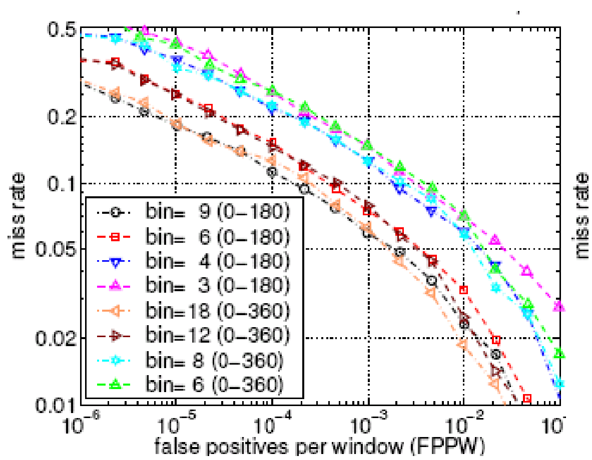


Figure 4.5 Effect of variations in number of bins (Triggs, 2005)

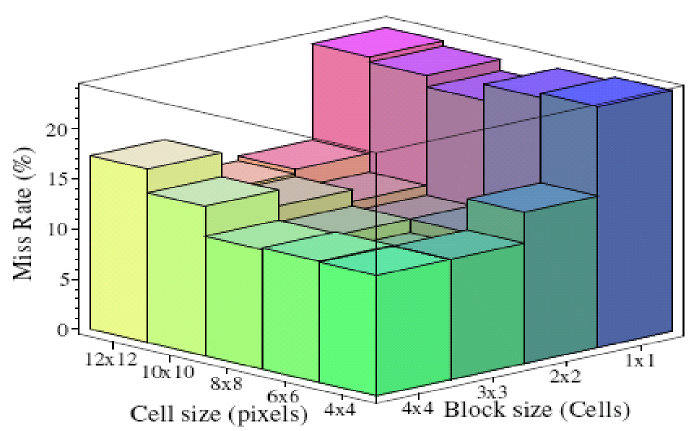


Figure 4.6 Effect of variations in block size and cell size (Triggs, 2005)

4.4 HOG as a Descriptor in SURF

Because of its nature and as claimed by the authors, HOG is illumination invariant. This is useful requirement for image registration of satellite images with varying illumination level. To compare descriptor of SURF and HOG descriptor, around the same point two image patches are selected from the two images having different illumination level as shown in Fig. 4.7. For both the image patches, Haar based SURF descriptor vectors are plotted in Fig. 4.8 while HOG descriptor vectors are plotted in Fig. 4.9.This shows HOG descriptor is more illumination invariant compared to the descriptor of SURF.

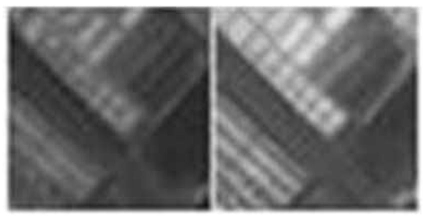


Figure 4.7 Small low resolution image patches of size 41X41 pixel with different illumination level (increased in size for proper visual display purpose only)

Here the idea is to use HOG as feature descriptor for SURF point features to address the illumination variation present between two satellite images. Such illumination variation may occur in certain cases such as multi-spectral images, multi-sensor satellite images.

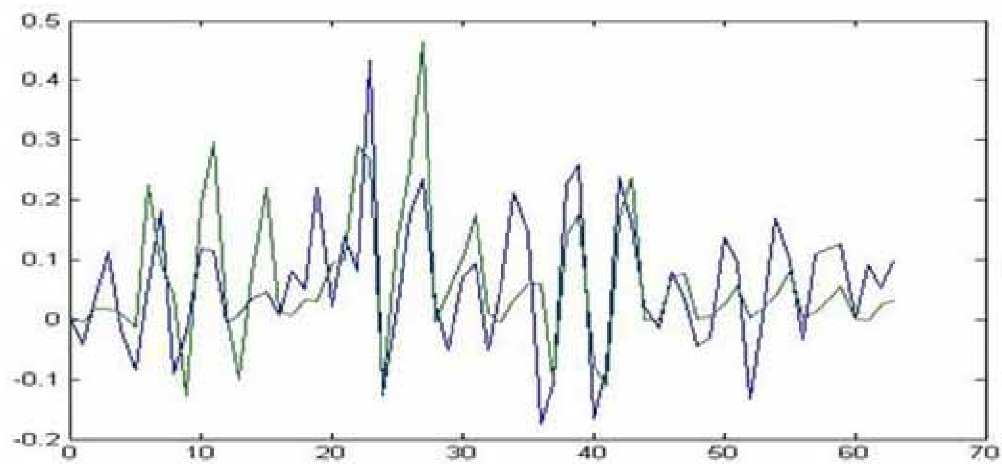


Figure 4.8 SURF descriptor vectors of two image patches

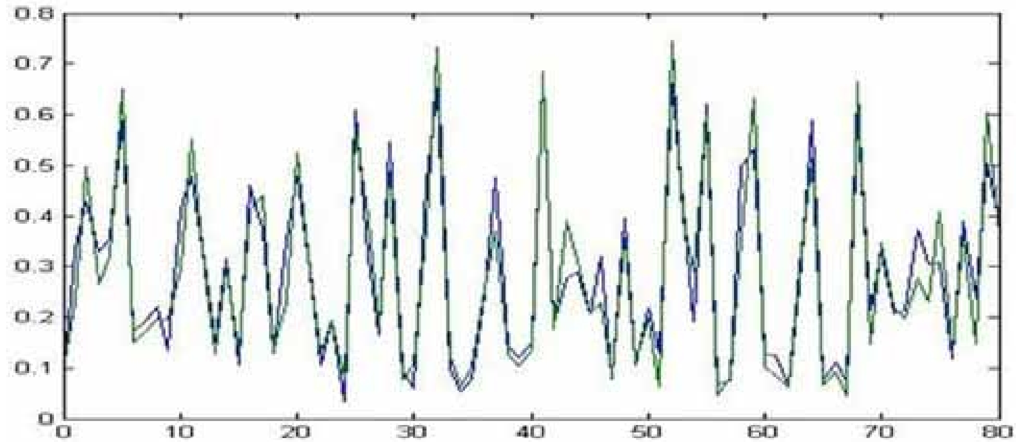


Figure 4.9 HOG descriptor vectors of two image patches

4.5 Image Registration Using HOG Descriptor

SIFT is used in (A. Sedaghat, 2011), (Wang, 2012), (C. Huo, 2012), (Gong, 2014), (Li, 2009) for image registration. In (Gong, 2014), for satellite images coarse image registration is performed using SIFT to get its advantage of robustness and then fine image registration is performed using mutual information to get its advantage of accuracy. Similarly in (Lee, 2010), coarse image registration is performed using SURF and fine image registration is performed using Harris corner detector. However such strategy of coarse-to-fine image registration requires re-sampling process two times, so corresponding errors are added.

For the work in this chapter, SURF point features are used. Satellite images may be multi sensor, multi-spectral, multi-resolution or multi-temporal; they are typically large in size. Due to these characteristics of satellite images, conventional image registration algorithms used for computer vision or medical images may face some problems. SURF is also giving some incorrect matches, and hence improved in (Wang Kai, 2012), (Lin, 2014), (Temizel, 2010), (Xiaopeng, 2014), (Lee, 2010) for satellite image registration.

In (Wang Kai, 2012) the normalized SURF algorithm extracts more accurate matching points than the original SURF algorithm; however the stability and robustness of the normalized SURF method still needs further study. In (Lin, 2014) feature points are extracted using SUSAN algorithm and they are described using SURF algorithm, where marginal improvement is found but results are not shown for challenging satellite images.

In (Temizel, 2010) performance of SURF for registration of high resolution satellite images captured at different bands is evaluated and then Scale Restriction (SR) method, which has been already proposed for SIFT, is adapted to SURF. In (Xiaopeng, 2014), SURF descriptor is modified according to the gradient reversals. This improves the Correct Match Rate (CMR) for multi-modal images but at the cost of reduced CMR for mono-modal images.

In this approach, first optional step is to remove intensity difference between two images using their mean values. This step can't remove it completely as intensity levels are not necessarily related linearly in case of satellite images such as multi-spectral and multi-sensor images. Further as it operates at globally, the local details may change sometimes changes significantly and satellite images are rich of low level details.

Keypoints are extracted using SURF. Around every extracted keypoints, a 41X41 image patch as used in (Schmid M. K., 2005) is selected. For these image patches, corresponding HOG feature descriptors are computed. Number of bins selected in HOG, will decide the size of descriptor. The feature descriptor vectors are matched using Euclidean distance. The steps for proposed approach are shown in Fig. 4.10.

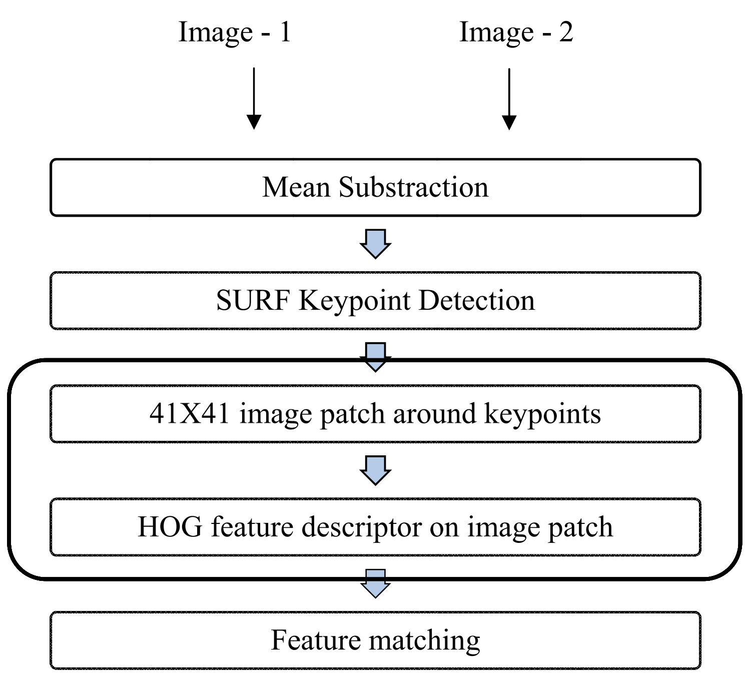


Figure 4.10 Steps for proposed approach

4.6 Datasets

Among the other datasets, two datasets of size 400X300 obtained from **(Gong, 2014)**. Two more multi-spectral satellite image datasets from LISS-III sensor of size 1000X1000 obtained from Bhuvan portal of NRSC, ISRO with details shown in table 4.1. All these four datasets are shown in Fig. 4.11, which show large illumination variation.

Table 4.1 Datasets obtained from BHUVAN portal of ISRO

near the bay of Kutch	
Name	L3-NF42E12-091-056-18dec11
Resolution	24m
Name of the Satellite	Resourcesat-1
Sensor	LISS-III

Part of Ahmedabad city and Gandhinagar city	
Name	L3_SAT_8B_v1_72.5E23N_F43A12 _01nov08
Resolution	24m
Name of the Satellite	Resourcesat-1
Sensor	LISS-III

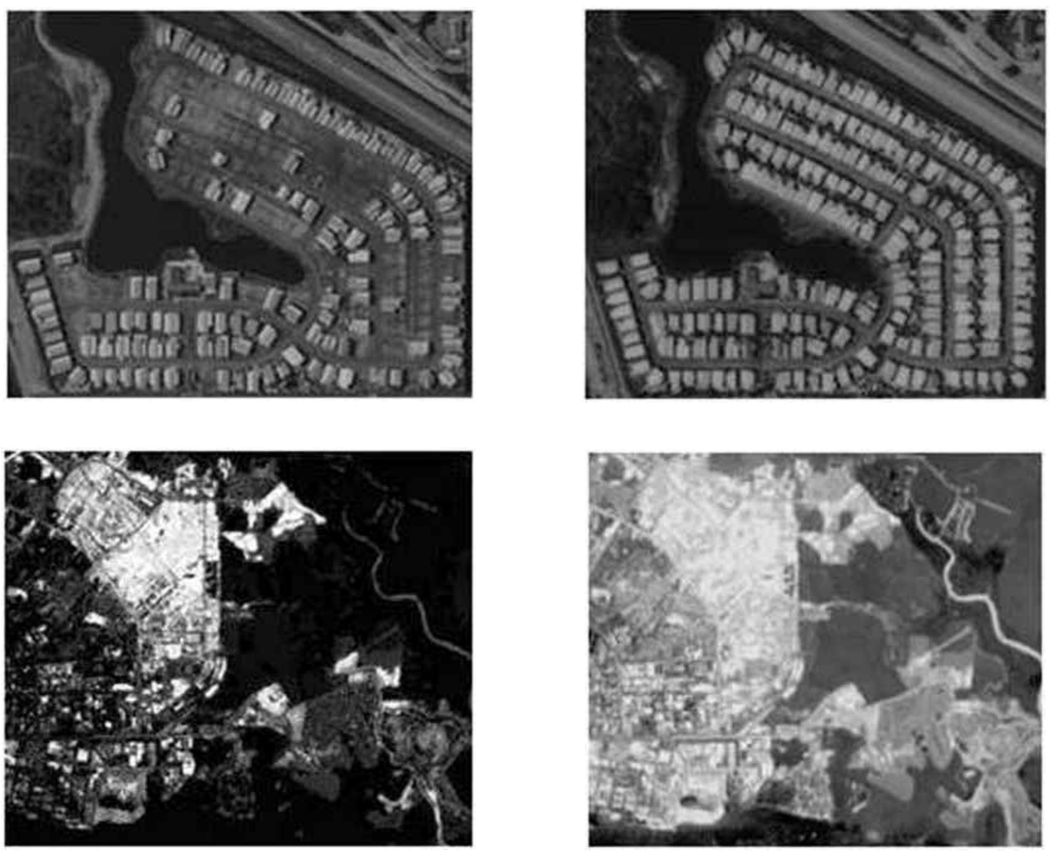
4.7 Results and Discussion

For performance parameter CMR is used. This is with the consideration that the matched point is within five pixels of the actual matched point, which is the same approach used in (Schmid M. K., 2005). Further correct@N (Xiaopeng, 2014) is also used, because in image registration first few best matched features are used to estimate the registration parameters, based on the geometric transformation under consideration like rigid, affine etc. In our analysis N=20 is taken, same as in (Xiaopeng, 2014). Implementation is in MATLAB, on Pentium Dual-Core CPU with 2 GHz and 2 GB RAM.

Best 20 matched feature points for Dataset-1 are shown in Fig. 4.12 for two approaches: Approach-A using SURF with its Haar based descriptor of 64 size called SURF-64, and Approach-B using HOG descriptor (with number of bins=9 i.e. descriptor size of 81 as it is suggested as best value in (Triggs, 2005)) with SURF called HOG-81. The Fig. 4.12 shows, for Approach-A, 10 matches are correct out of best 20 matches i.e. correct@N is 10, while for Approach-B, 15 matches are correct out of best 20 matches i.e. correct@N is 15. Similar observation for dataset-3 is shown in Fig. 4.13.

Further analysis was also carried out for different bin size in HOG. Comparable results are found in case of seven numbers of bins i.e. HOG descriptor size is 63. For comparison purpose this is also included as approach-C called HOG-63.

For all four datasets, Fig. 4.14 shows the CMR while Fig. 4.15 shows the correct@N (N=20). This shows improved performance of using HOG as descriptor for satellite images. Computation time is not significantly changed as shown in table 4.2.



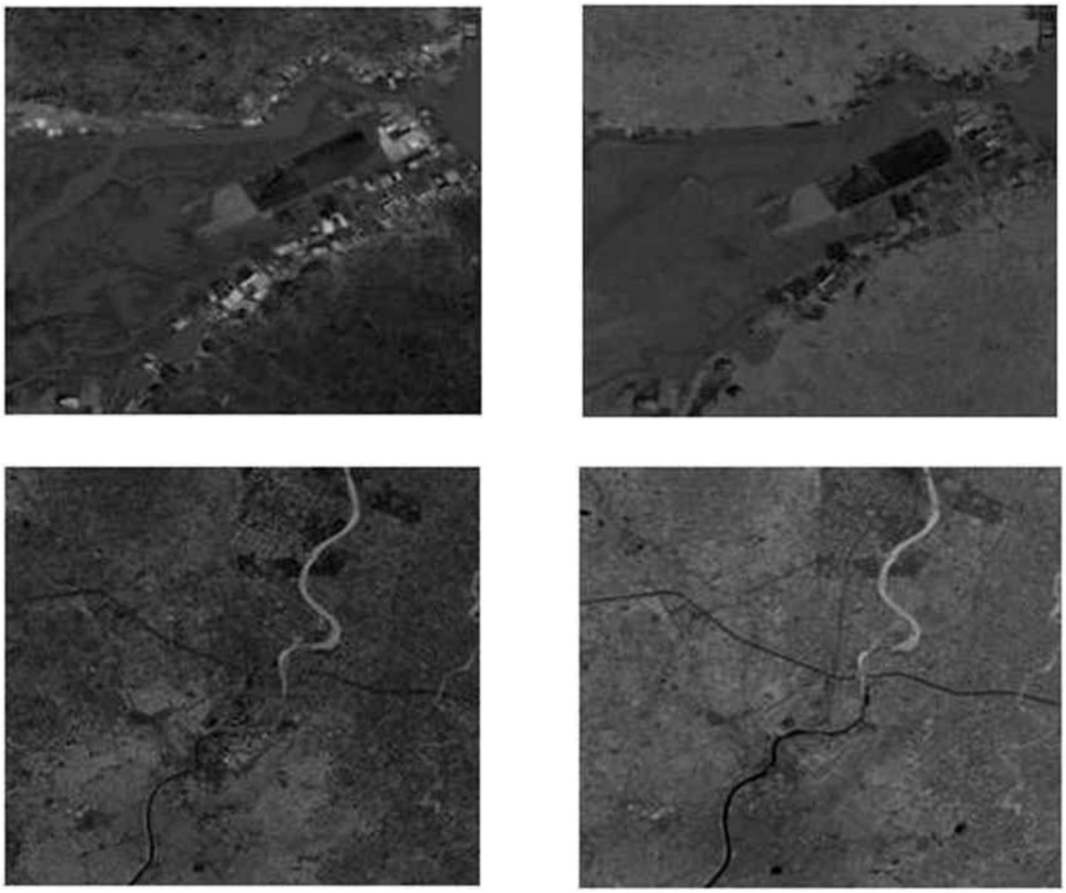
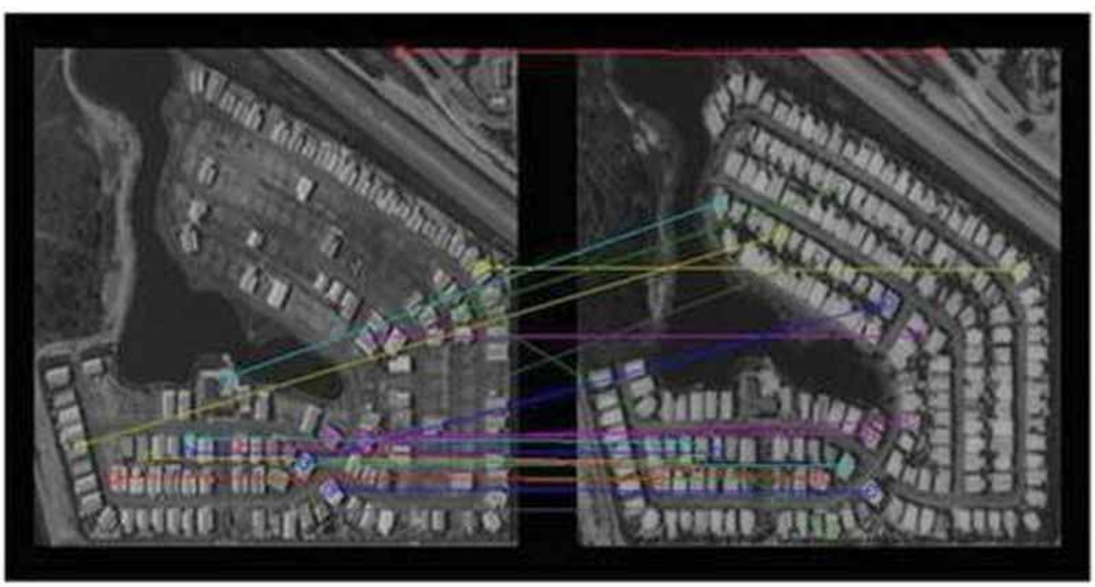


Figure 4.11 Dataset-1, dataset-2, dataset-3 (near bay of Kutch) and dataset-4 (near Ahmedabad city)



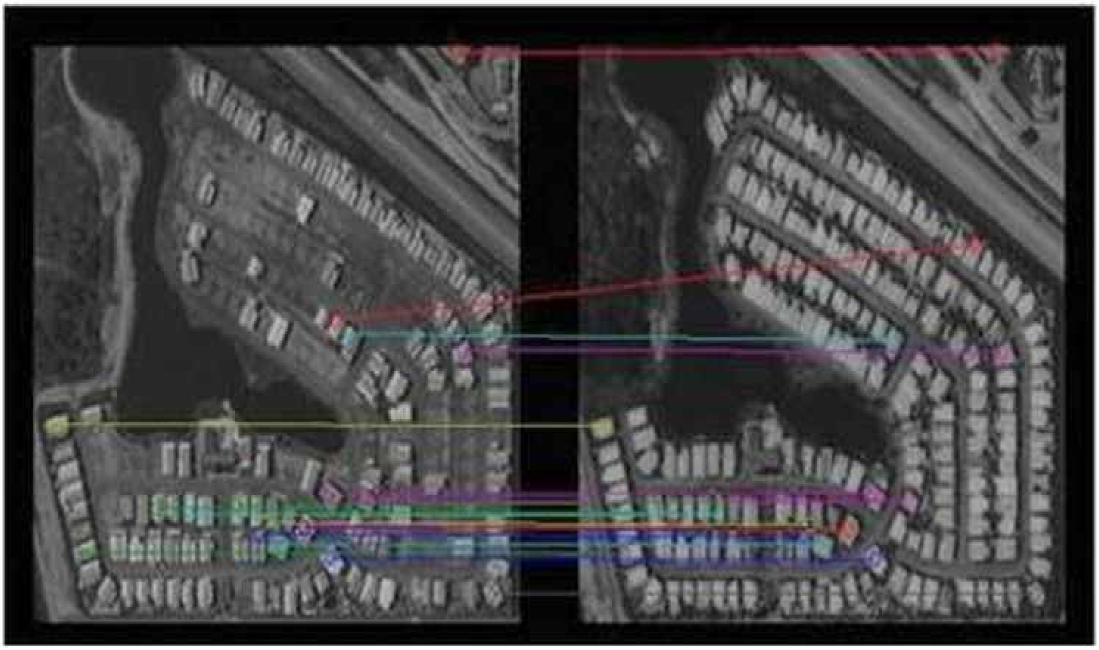
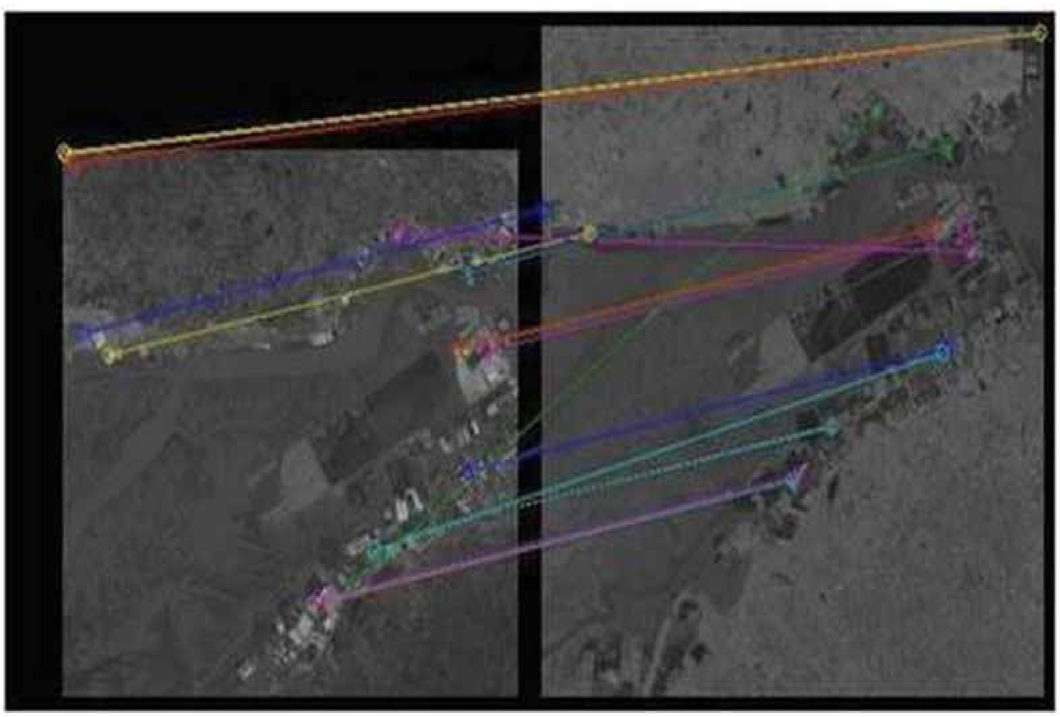


Figure 4.12 Matched point features for dataset-1 using (a) Approach-A and (b) Approach-B



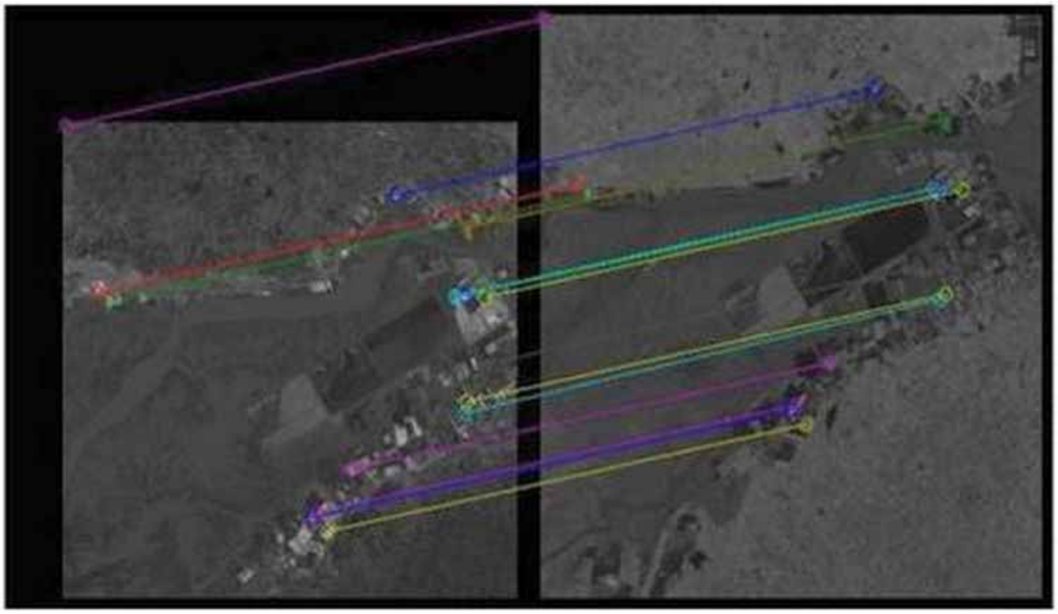


Figure 4.13 Matched point features for dataset-3 using (a) Approach-A and (b) Approach-B

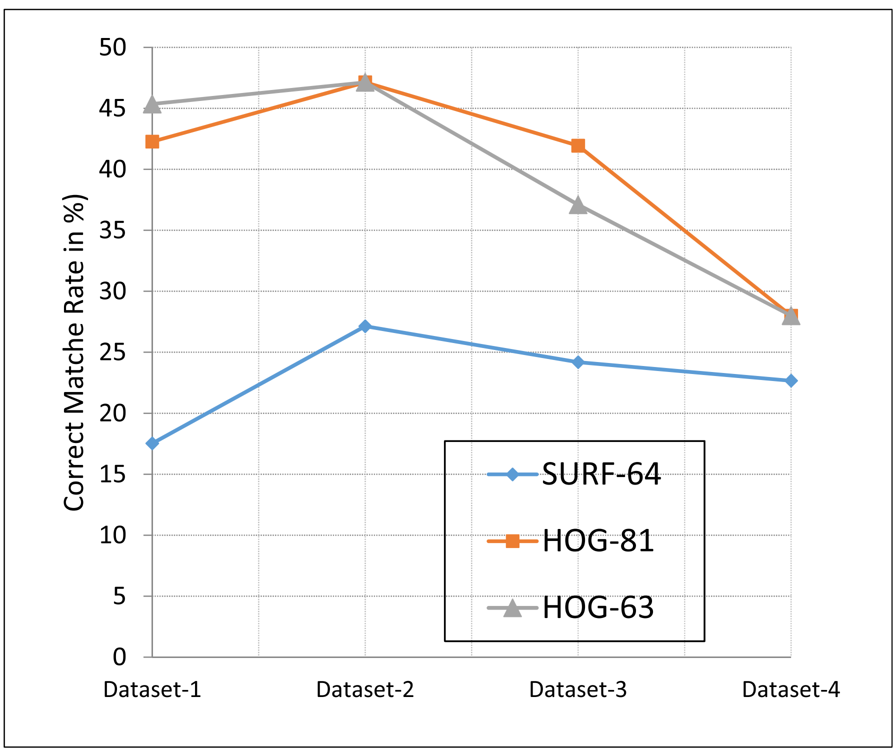


Figure 4.14 CMR for all four datasets

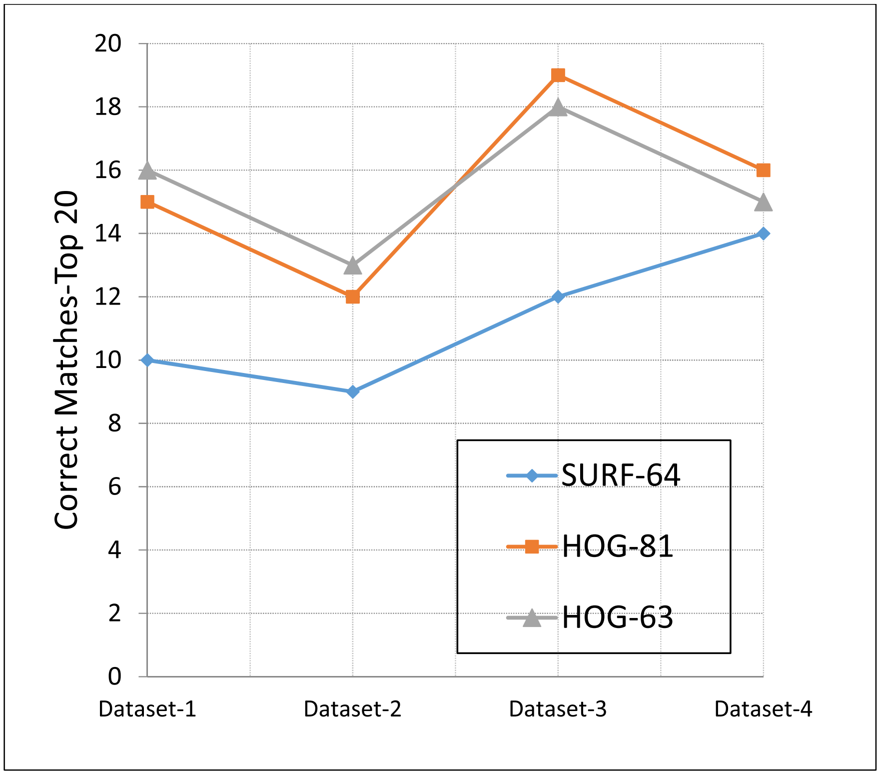


Figure 4.15 Correct@N, N=20 for all four datasets

Table 4.2. Computation Time for HOG descriptor based approach

	Computation Time in second			
Approach	Dataset-1	Dataset-2	Dataset-3	Dataset-4
SURF-64	3.33	2.63	13.80	15.28
HOG-81	3.18	2.37	13.52	16.19
HOG-63	3.14	2.32	13.62	15.30

For further analysis, images of a same scene captured by various sensors such as SPOT, Landsat, IRS and air photo are used, which are having illumination variation as shown in Fig. 4.16. From these images with different combination, six image pairs are formed. For these six image pairs observed CMR using the SURF-64 and HOG-81 approaches are compared in Fig. 4.17. This shows the performance improvement in case of multi-sensor satellite images.

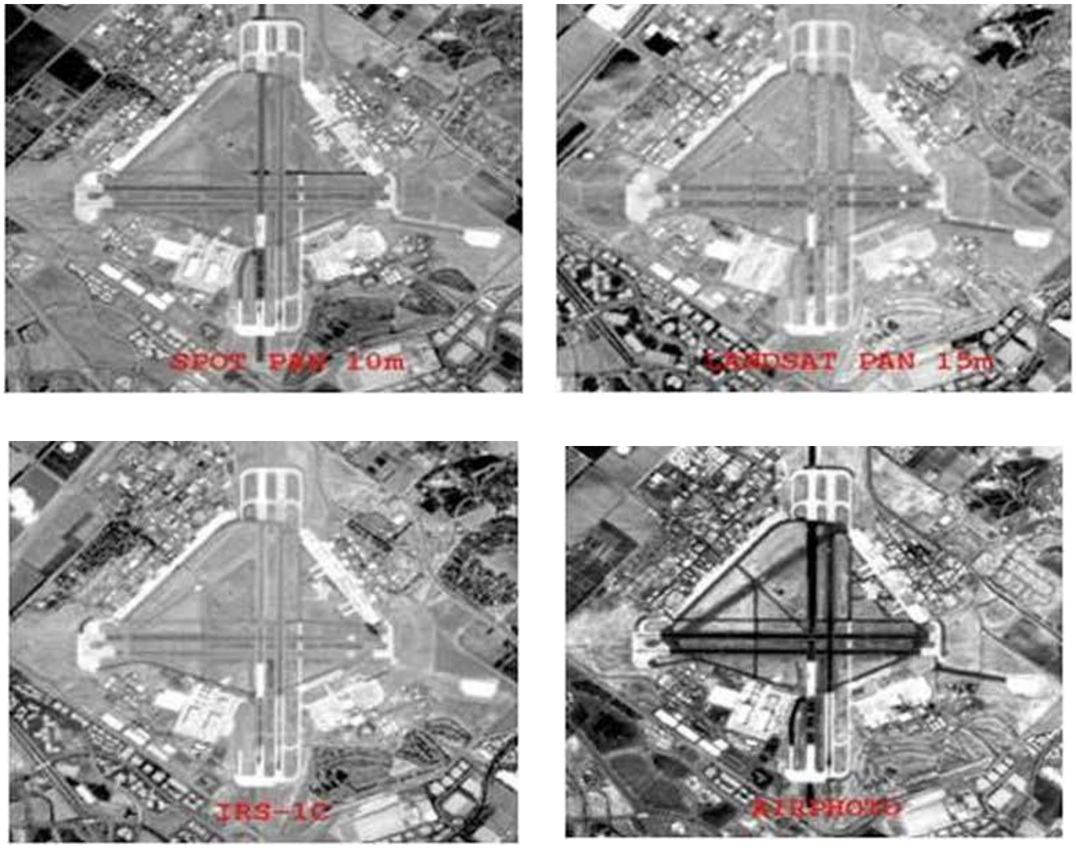


Figure 4.16 Multi-sensor image dataset for the same scene: SPOT, Landsat, IRS and Air photo

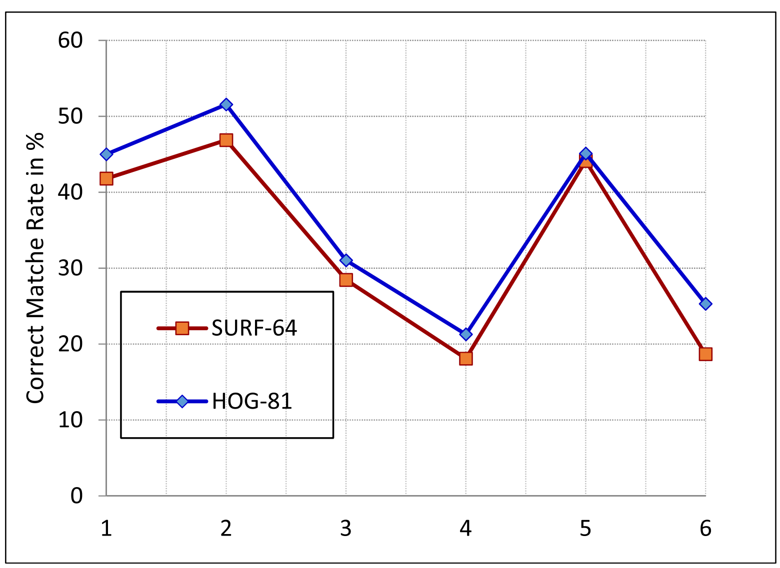


Figure 4.17 CMR for six multi-sensor image pair datasets

To observe the effect of improved CMR on image registration, first of all incorrect matches i.e. outliers are removed from the initial few matches using RANSAC approach. Then after, predefined numbers of correct matches are used to find the registration or transformation parameters. If ground truth is not available then it is difficult to compare the results. But for the known registration parameters the results can be compared. During various experiments, improved accuracy is observed in estimating registration parameters for the suggested approach. This improvement in image registration is due to the improvement in CMR. For example the dataset-3 was translated in x-direction and y-direction by 100 pixels and 200 pixels respectively. The SURF-64 approach has estimated it as 100.90 and 200.6 pixels while HOG-81 approach has estimated it as 100.70 and 200.51. Similar improvement is also observed with small rotation as well.

4.8 Summary

Compared to SURF with Haar response based descriptor, SURF with HOG based descriptor is found more appropriate in case of images with illumination variation such as multi-spectral and multi-sensor satellite images. Good improvement is observed in CMR and correct@N value for HOG based SURF. For small rotation performance in terms of correct@N is comparable but large rotation degrades it.

CHAPTER 5

Feature Based Method: Feature Refinement using SVM Classification

5.1 Overview

Among the main steps of feature based methods of image registration, the previous chapter is on feature descriptor step, while this chapter is on feature matching step. In this work CMR is improved for the case of satellite images having some occluding objects in the scene such as clouds or shadows. For example if clouds are present in the images then the features corresponding to them are not effective and they may lead to incorrect matches. In case of multi-temporal images if the clouds are shifted in the test image with respect to the reference image and if the corresponding features are correctly matched then it results in incorrect estimation of registration parameters; because the shifting in clouds is not actual geometrical shift in between the two images. Such circumstances can also be applicable to the images captured at different day time having varied amount of shadow or having any other varied occluding region. So it is better if the feature points related to say such cloudy areas are removed before the matching step. This will reduce the false matches and so improves the CMR. In this work it is achieved by support vector machine based classification of the features.

Related Work

Such idea of refinement of features before matching step to improve CMR is also found in (C. Huo, 2012), where coarse to fine strategy for image registration is used. After coarse image registration, the images are divided in blocks and block wise features are matched as explained in Fig. 5.1. Features of respective blocks say local features only are taking part in matching step. This reduces the possibilities of matching with the similar features which may exist in other part of the image and are away from the block. This improves the CMR.

The drawback of such coarse-to-fine approach is that it is required to perform re-sampling twice so corresponding errors is accumulated.

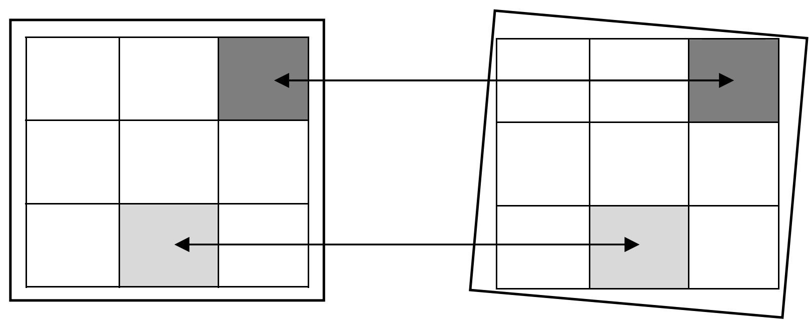


Figure 5.1 concept of local block level matching in fine image registration step of coarse-to-fine image registration strategy

Somewhat similar approach using block level processing is also used in (Choi, 2013). This increased the number of features and improved the feature distribution quality. In these both approaches the refinement of features can be thought in spatial representation. In our approach the refinement of features is targeted to some specific types of features related to say occluding objects only; so the refinement is content based.

In (Hong G. , 2007) as future scope of work it is suggested to refine the control points i.e. extracted feature points automatically to improve the robustness. In the work it was required to remove some of features associated with trees or building manually. So the authors have also suggested use some method to remove those kind of specific features automatically rather than manually.

In (Eikvil, Holden, & Huseby, 2009) a system for co-registration of time series which depend on learning based theory is described. During the training phase the system learns to recognize regions in an image suited for registration. It allows the intelligent selection of an appropriate registration algorithm for each region in the image. Simultaneously, unsuited regions for registration can be discarded. This approach is region based and whole image is classified prior to the registration. In our approach, only the extracted point features are classified, prior to the feature matching step.

In (Shen, 2010), combination of SIFT and image classification is used for remote sensing image registration. First using maximum likelihood, images are classified according to the spectral features of different objects approach. Then keypoints belongs to the same classes are matched so this reduces likelihood of incorrect matches so CMR is improved. Compared to this approach, in our approach, faster SURF is used; the classification is not on whole image but on extracted features only and classification is done using SVM.

Most of the remote sensing or satellite images contain more or less clouds distributed in the image. For instance, Landsat scenes are globally estimated to be 35% cloud covered (Roy, 2008). Clouds are required to be detected, analyzed and/or classified for meteorological study. In (A. Taravat, 2015) an effort has been made to investigate machine learning methods for automatic cloud detection in the image. The pixel values have been used as inputs and output is classified pixels in terms of cloud coverage or others cloud free pixels.

But as far as image registration is concerned there is a critical impact on the accuracy and results if the clouds are present in the scene. The features associated with clouds may result in incorrect matches as the clouds are not static and they don't give stable features. Further for example if the clouds are shifted and features from them are correctly matched then it results in incorrect estimation of translation registration parameters because the cloud shift and the translation registration parameters are not related at all.

In (Luis Gomez-Chova, 2012) a tool is provided for masking of clouds in time series from Earth observation satellites. It is to minimize the errors such as land-cover changes and co-registration; and so maximizing the specific changes in multi-temporal images. The method enables automatic cloud detection in multi-spectral time series. Image registration of such cloudy images can be performed after cloud removal. In our approach a separate cloud removal is not required but it is taken care in the feature refinement step before feature matching step. Cloud removal works on whole image while feature refinement works only on extracted features.

Though SIFT is widely used in many applications including image registration and object recognition, directly applying SIFT to register satellite images results in high false rate. So, in (M. Hasan X. J., 2010), spatial information is used for feature refinement. For a feature points a list of their neighbouring feature points on each image separately within a window of W pixels and then instead of matching all the feature points in the target image with all the points in the reference image, only the points of the respective list are matched. Supporting results for band to band registration is presented. Similar spatial information based approach so called Spatial Consistent Matching (SCM) is also used in (Fan, 2013). One disadvantage of the approach is that the utilized low distortion constraint which restricts its direct use for registering images with large distortion.

5.2 Support Vector Machine

Support Vectors Machine (SVM) has recently shown its ability in pattern recognition and classification (Chapelle, 1998). Some of the examples of such applications are text categorization, hand-written character recognition, image classification and bio-informatics. In literature SVM favourably to more advanced classifiers like neural networks, discriminate analysis and decision trees. Performance is found better than the others. Main properties of SVM classifier are: distribution-free classification approach and training step is reduced to a convex optimization problem. Intuitively, given a set of points which belong to either of two classes, a linear SVM finds the hyperplane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyperplane. According to (Chapelle Olivier, 2003), this hyperplane minimizes the risk of misclassifying examples of the test set.

Optimal Separating Machines

Let $(x_i, y_i)_{1 \le i \le N}$ be a set of training example, $x_i \in R^n$ and belongs to class labeled by $y_i \in \{-1,1\}$. The aim is to carry out the equation of an hyperplane which divides the set of examples such that all the points with the same label are on the same side of the hyperplane. This means find w and b such that

$$y_i(w.x_i + b) > 0, i = 1, \dots., N \tag{5.1}$$

If there exists a hyperplane satisfying the eq. 5.1, the set is said to be linearly separable. In this case, it is always possible to rescale w and b such that

$$\min_{1 \le i \le N} y_i(w.x_i + b) \ \ge\ 1, \qquad i = 1, \dots., N \tag{5.2}$$

i.e. such that the closest point to the hyperplane has a distance of $1/\|w\|$. Then, eq. (5.1) becomes

$$y_i(w.x_i + b) \geq 1 \tag{5.3}$$

Among the separating hyperplanes, the one for which the distance to the closest point is maximal is called optimal separating hyperplane (OSH). Since the distance to the closest point is $1/\|w\|$, finding the OSH amounts to solve the following problem:

$$\text{minimize } \frac{1}{2}w.w \tag{5.4}$$

under constraint (5.3)

The quantity $2/\|w\|$ is called the margin and thus the OSH is separating hyperplane which maximizes the margin. The margin can be seen as a measure of difficulty of the problem: the smaller the margin is the more difficult the problem is, while the larger margin is, the better the generalization is expected to be as shown in Fig. 5.2.

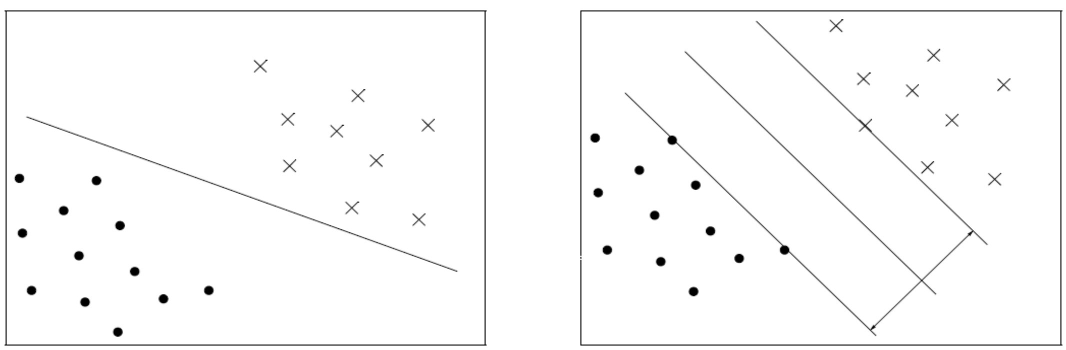


Figure 5.2 Both the hyperplane separate correctly the training examples, but the OSH on the right side has a larger margin and expected to give better generalization

Since$\|w\|^2$ is convex, minimizing it under linear constraints of eq. (5.3) can be achieved by the use of Langrange function

$$L(w,b,\alpha) = \frac{1}{2}w.w \; - \; \sum_{i=1}^{N} \alpha_i[y_i(w.x_i + b) - 1] \tag{5.4}$$

To find this point, one has to minimize this function over w and b and to maximize it over Lagrange multiplier $\alpha_i \; \geq 0$

If the data are not separable, the problem of finding OSH becomes meaningless. To allow the possibility of examples violating (5.3), one can introduce slack variables. This leads to the selection appropriate kernels.

5.3 Feature Classification using SVM

To improve CMR in FBM for image registration, a feature refinement step to target occluding objects such as clouds or shadows is added before feature matching step. The algorithmic steps are shown in Fig. 5.3. For point feature extraction SURF is used. For the every feature point, a 41X41 patch around them are obtained. The mean and variance values of the patches are used to characterize the patches to call them as say 'cloudy' or 'non-cloudy'. Typically these values are larger for cloudy and smaller for non-cloudy, but difficult to separate. SVM classification is best in the category of such two class classifier. So using SVM the feature patches are classified and accordingly the point features without clouds can be taken as say class-1 (i.e. 'non-cloudy') and with clouds can be taken as say class-2 (i.e. 'cloudy'). There is more likelihood of incorrect matches with the class-2 cloudy patches. So in the matching step, they are just discarded. The point features associated with class-1 i.e. 'non-cloudy' are only allowed in matching step.

In case of images with shadows of different amount, the patches without shadows can be classified as class-1 and patches with shadow can be classified as class-2. Then after point features concerned with class-1 only are allowed to take part in the matching step.

Data-sets

Among the other datasets, three image pairs are shown in Fig. 5.4. The first image pair is air photos captured by me from the window of airplane having noise and clouds. The 700X600 size only from the original image is selected here. Second image pair is also having clouds covering small area of image. It is clearly observed that in both the image pairs the clouds are shifted with respect to each other. Third image pair doesn't have the clouds but significant amount of change is present. Original second and third image pairs are of thousands by thousands size but can be resized to take care of computation time.

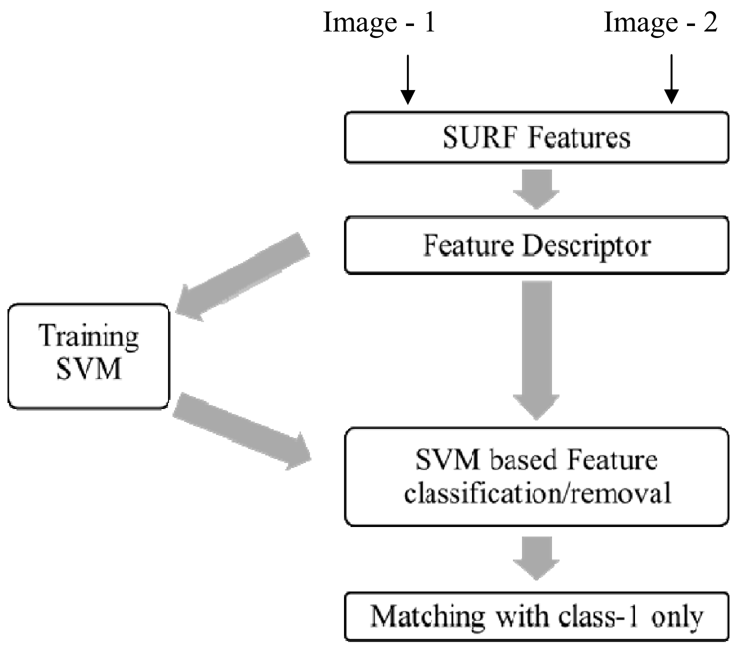
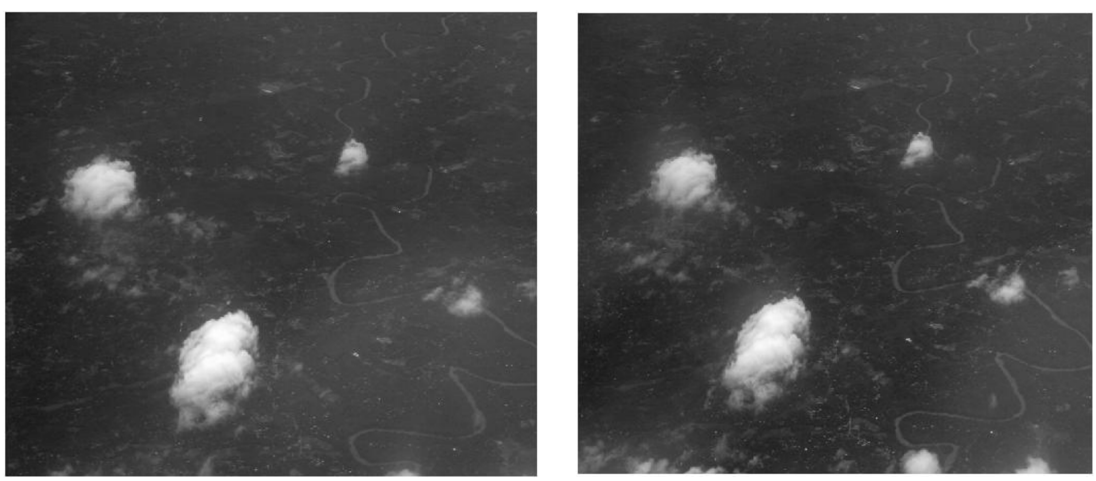


Figure 5.3 Feature refinement step in the steps of FBM of image registration



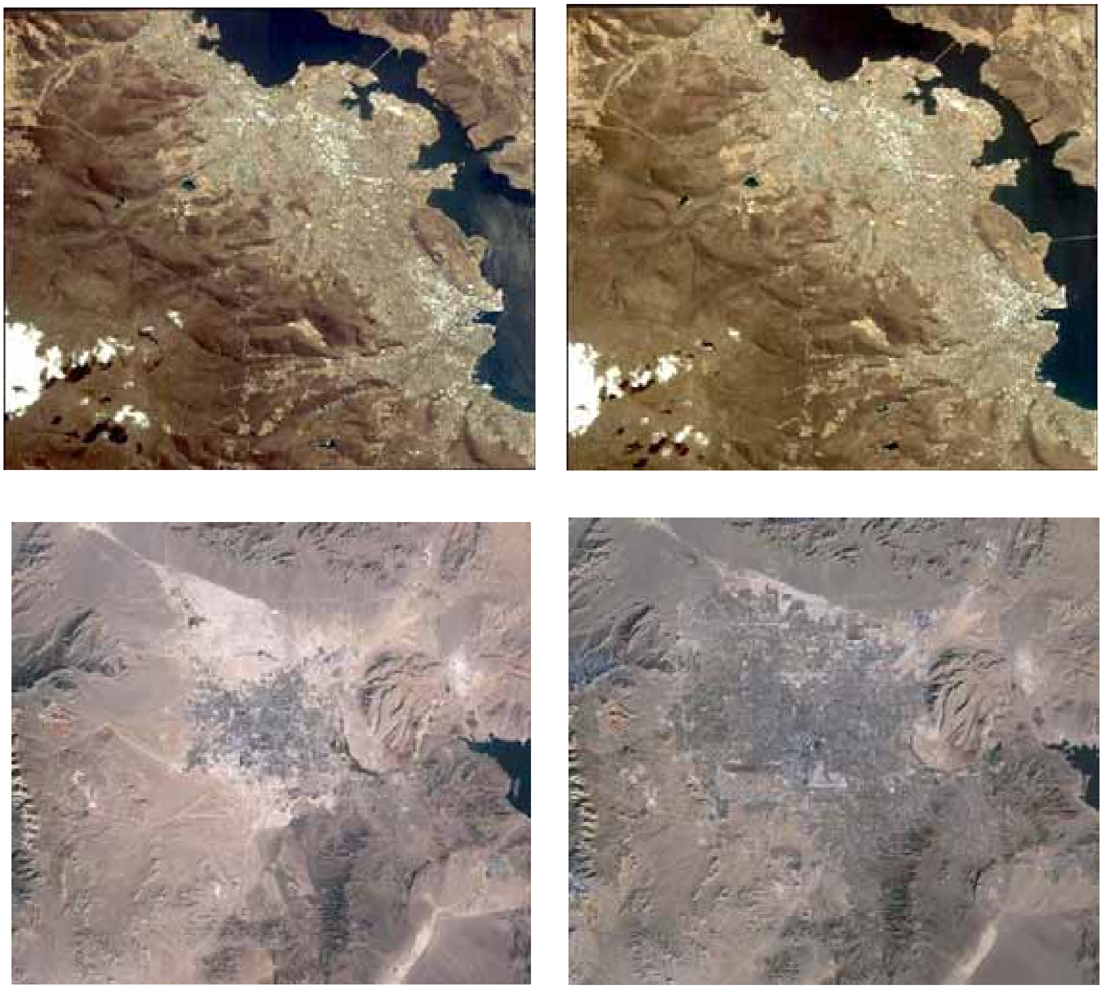


Figure 5.4 Three image pairs used for the analysis of refinement step

5.4 Training of SVM Classifier

The performance of the suggested modification highly depends on the training of the SVM for the classification. For this purpose image patches around the extracted points are used. Some examples of such patches are shown in Fig. 5.5. Among them, very carefully chosen some 'cloudy' and 'non-cloudy' image patches are used for training purpose. So for one image pair, among the extracted features, a fraction of them can be used for training purpose for removal of cloudy features. Then after the same image pair can be tested under various parameters to analyze the approach. Here threshold of SURF is varied and CMR is observed with and without feature refinement step. For other occlusion effect, say removal of varying shadows, it is required to train the SVM as per the patches with and without shadows.

Figure 5.5 41X41 Image patches for image pair-2 and image pair-3 used for SVM training purpose

Further, SVM classification is model based or kernel based method. In general there are no straight rules for the selection of kernel methods to be used. So it is difficult to decide which kernel is to be used. So, further analysis can be carried out to compare various kernel methods, for our specific objective that is removal of cloudy image patches.

5.5 Results and Discussion

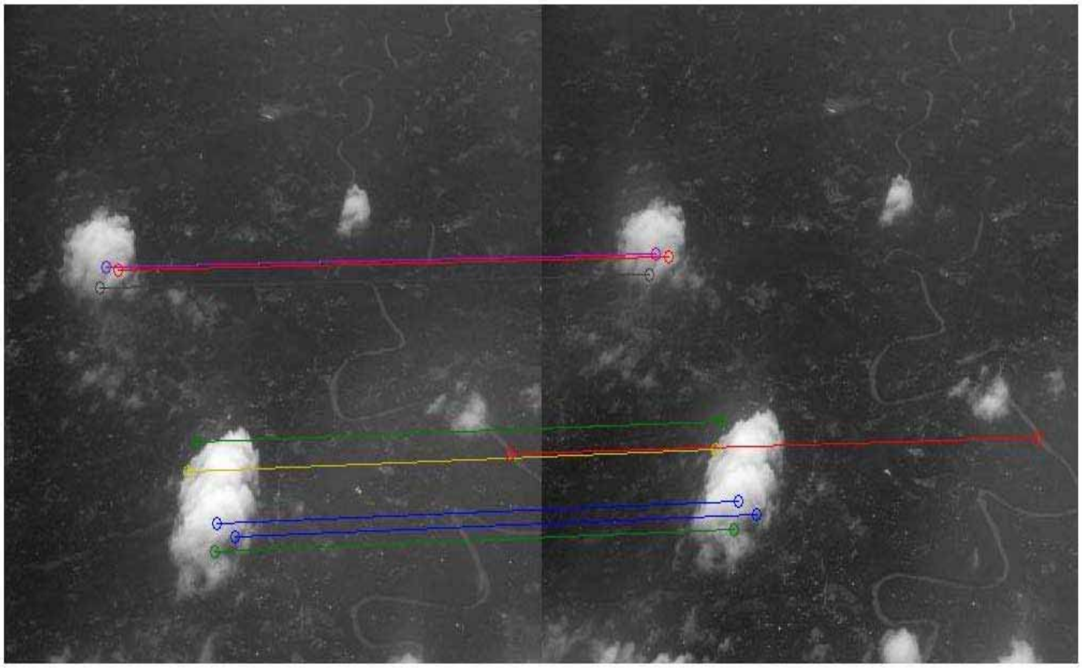
Though the approach is applied to other images also, results related to three image pairs shown in Fig. 5.3 are presented and discussed here. First image pair is used to see the visual effect of the suggested SVM based image refinement approach applied to cloudy features. Second pair is used for comparison of various kernels for SVM classification in terms of CMR. Third image pair is also used for CMR analysis but with the target of refinement is changes present in images instead of shadow or cloud.

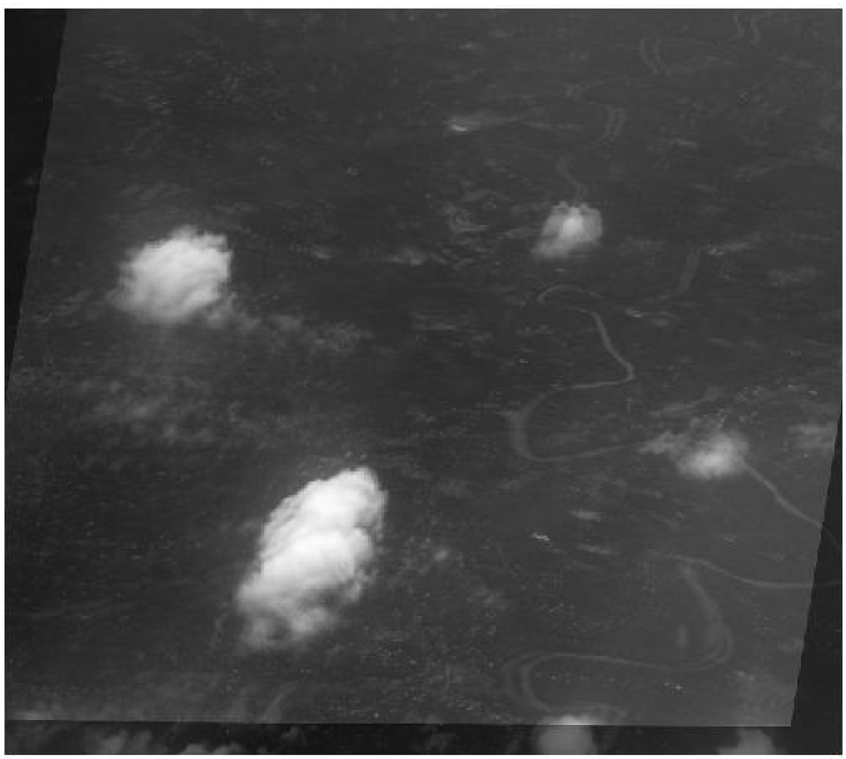
For the first image pair, as the ground truth are not available, after image registration both the images are fused image using average of their pixel values. The registration is based on best ten matched feature points only.

For the analysis on first image pair only best to matches are shown. Fig. 5.6 (a)-(c) are with SURF only i.e. without feature refinement. In Fig. 5.6 (a) if we look carefully then in the clouds are shifted, and the majority of the matched points belong to the cloudy scene. So this introduces the registration error. The registration error can be observed in Fig. 5.6 (b), a rectangle part of which has been zoomed and shown in Fig. 5.6 (c).

Fig. 5.7 (a)-(d) is for the case of with refinement of cloudy features using SVM classification. In Fig. 5.7 (a) if it is observed carefully then the majority of the matched points belong to the non-cloudy scene and they are taking part in the estimation of registration parameters.

The registration error can be observed in Fig. 5.7 (b), a rectangle part of which has zoomed and shown in Fig. 5.7 (c). This is low compared to the case of without feature refinement. Polynomial kernel based SVM classified features are shown in Fig. 5.7(d).





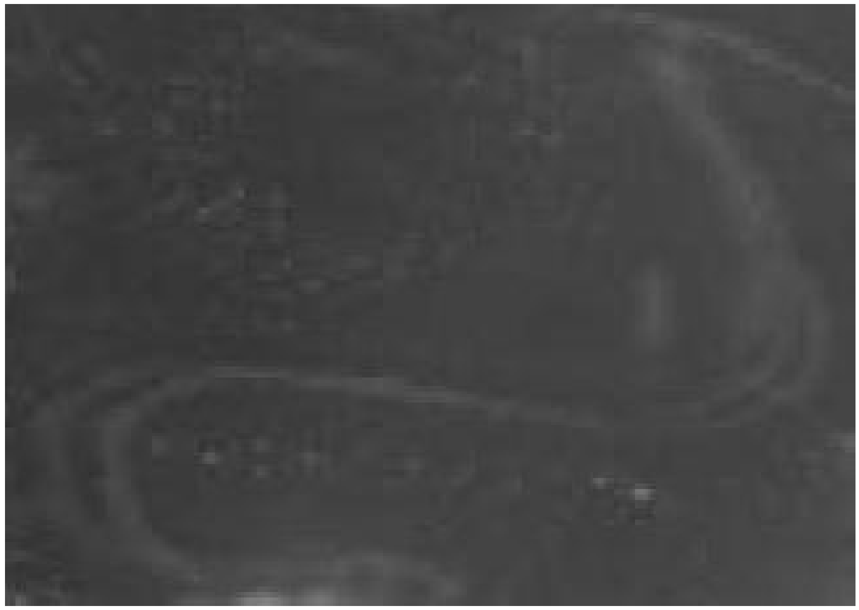
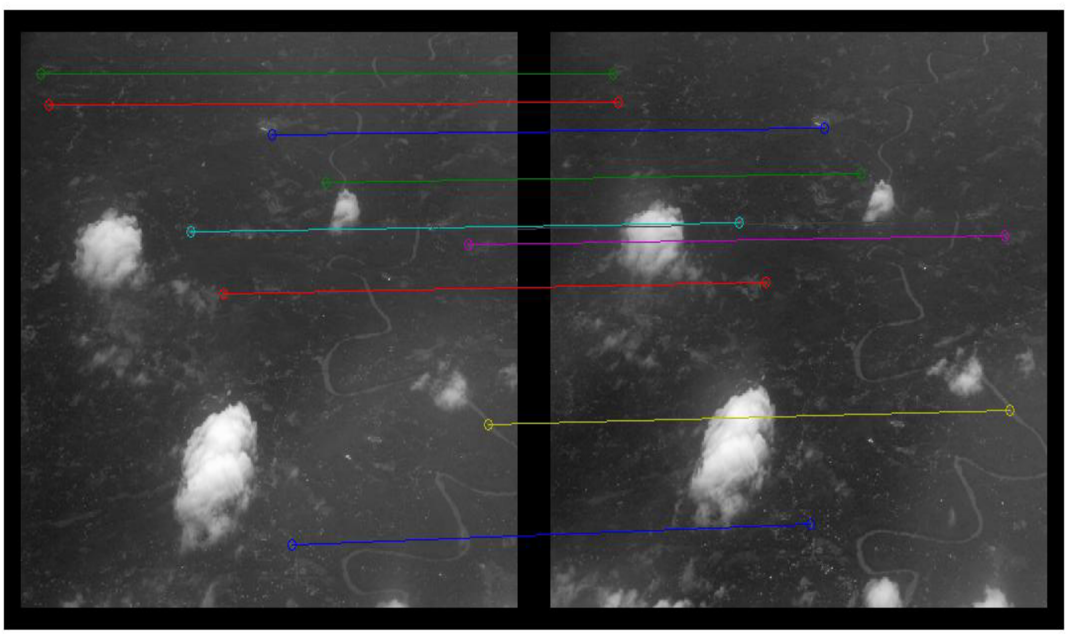
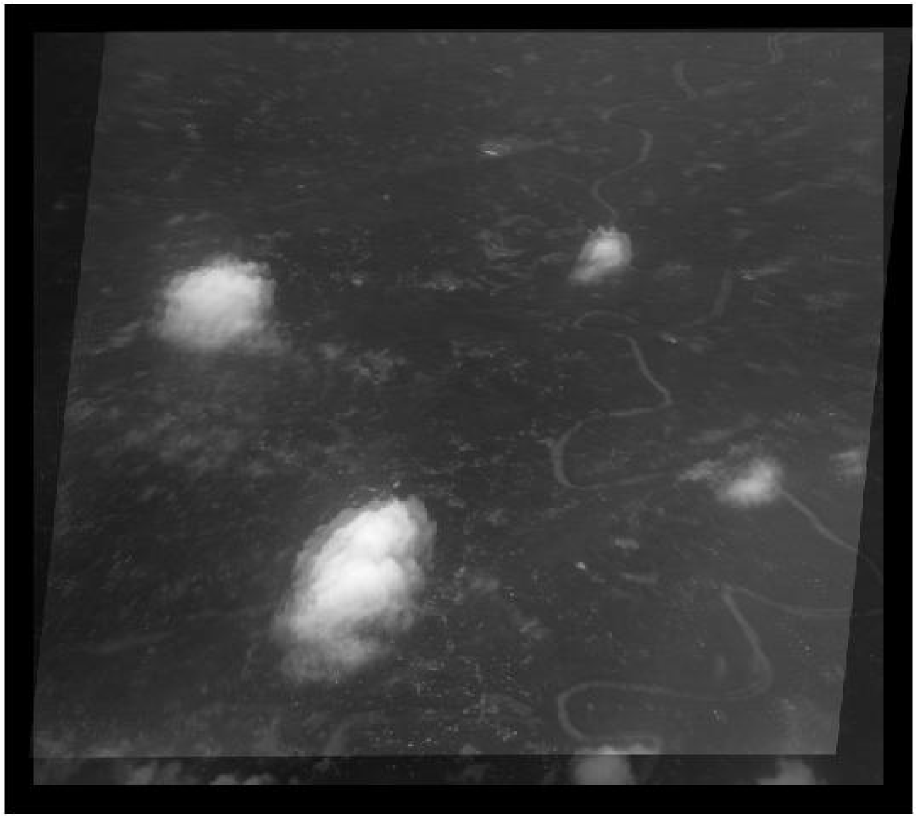
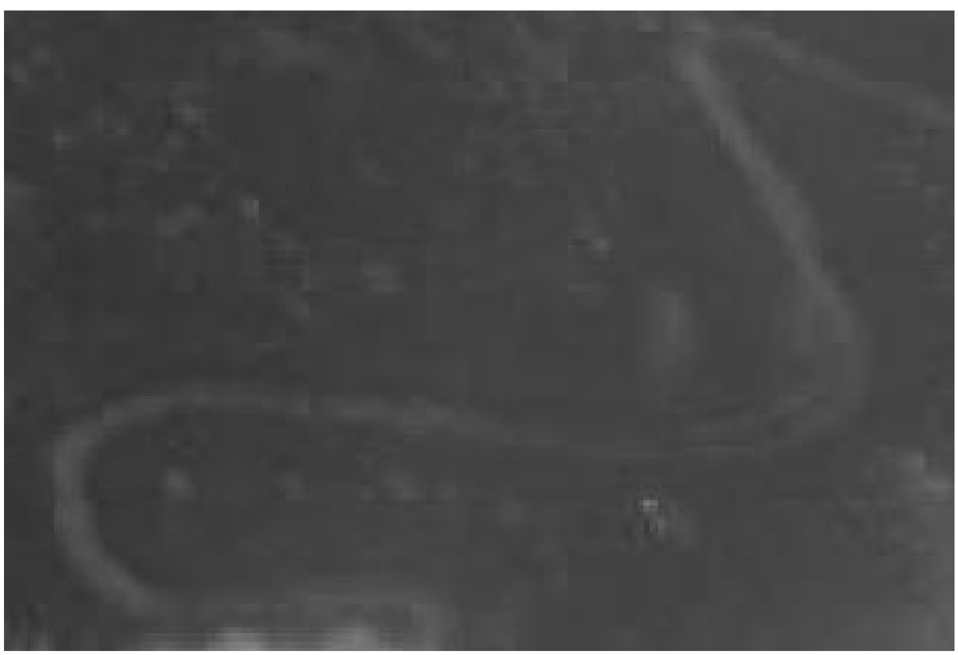


Figure 5.6 Results using SURF only for first image pair (a) correctly matches points (b) fused image (c) zooming of rectangle section of the fused image







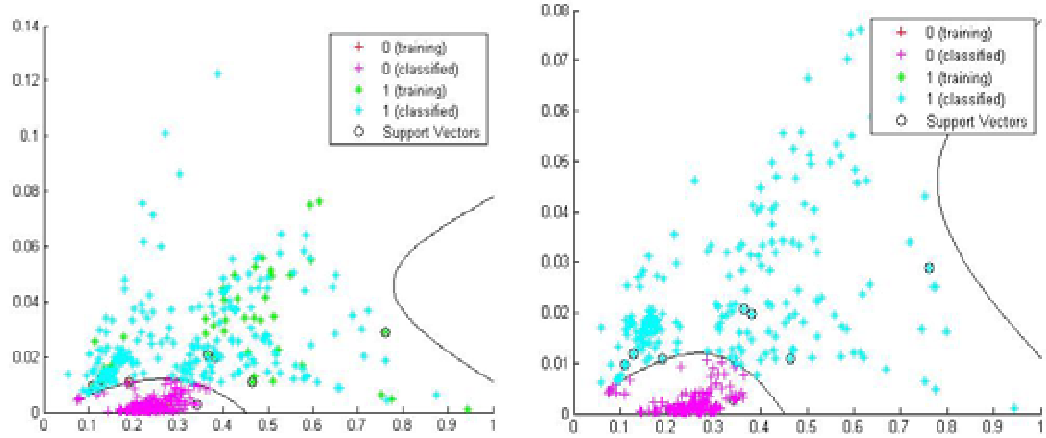
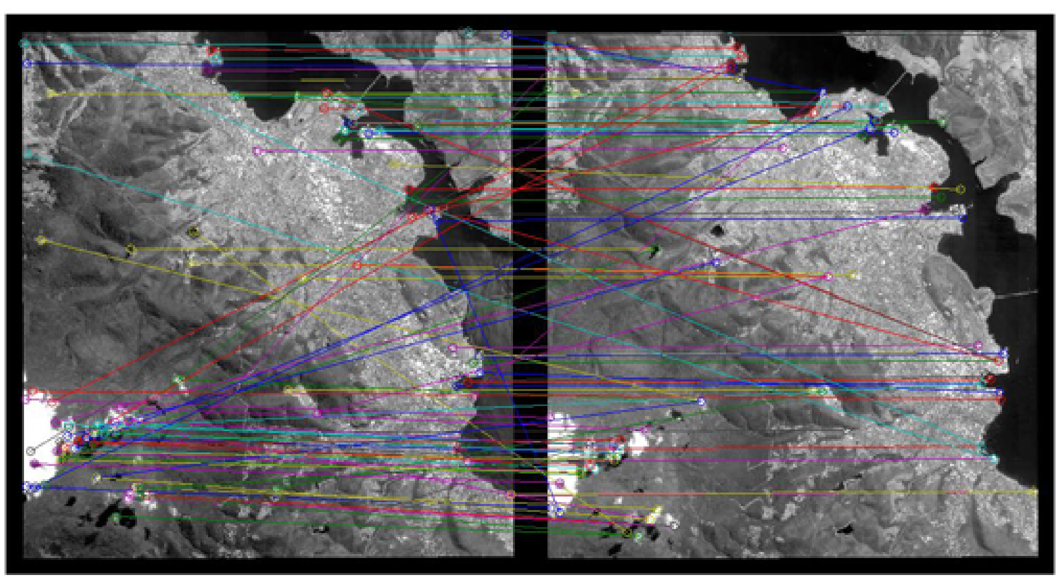


Figure 5.7 Results using SURF with SVM feature refinement for first image pair (a) correctly matches points (b) fused image (c) zooming of rectangle section of the fused image (d) SVM classified features

Now the results related to the second image pair is presented here for the CMR analysis including comparison of various kernel based SVM classification. Fig. 5.8 (a) is for all matched features without feature refinement. It clearly shows some of the matched feature points correspond to the cloudy area, which are in fact shifted and they are matched correctly, may results in incorrect estimation of registration parameters. Fig. 5.8 (b) is with the refinement steps for the cloudy features. It shows that the numbers of feature points corresponding to the cloudy area have been reduced; as they have been successfully rectified using SVM based classification.



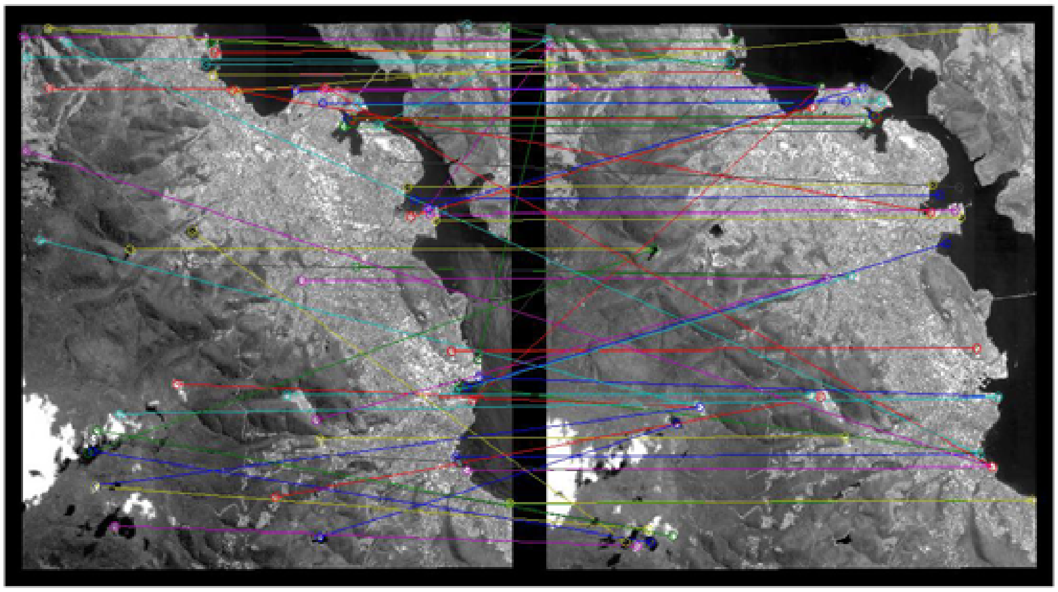
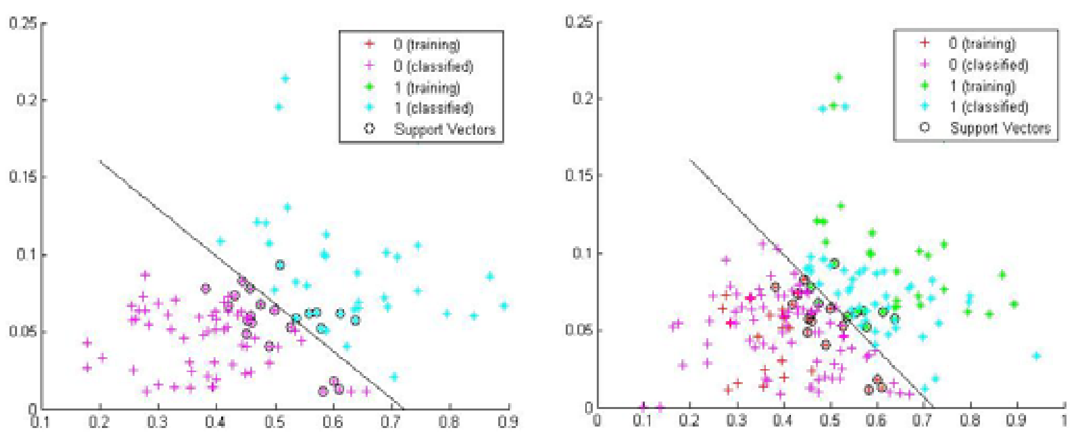


Figure 5.8 All matched points for second image pair (a) without feature refinement (b) with feature refinement

Further in Fig 5.9 (a)-(d), the classified features of both the images using different kernels: linear, radial basis function, polynomial and quadratic of SVM classification are shown. It is difficult to compare these classifications directly. So, corresponding CMR values are listed in Table 5.1 and plotted in Fig. 5.10. It is observed that the best kernel is polynomial. So in further analysis polynomial kernel based SVM is used.



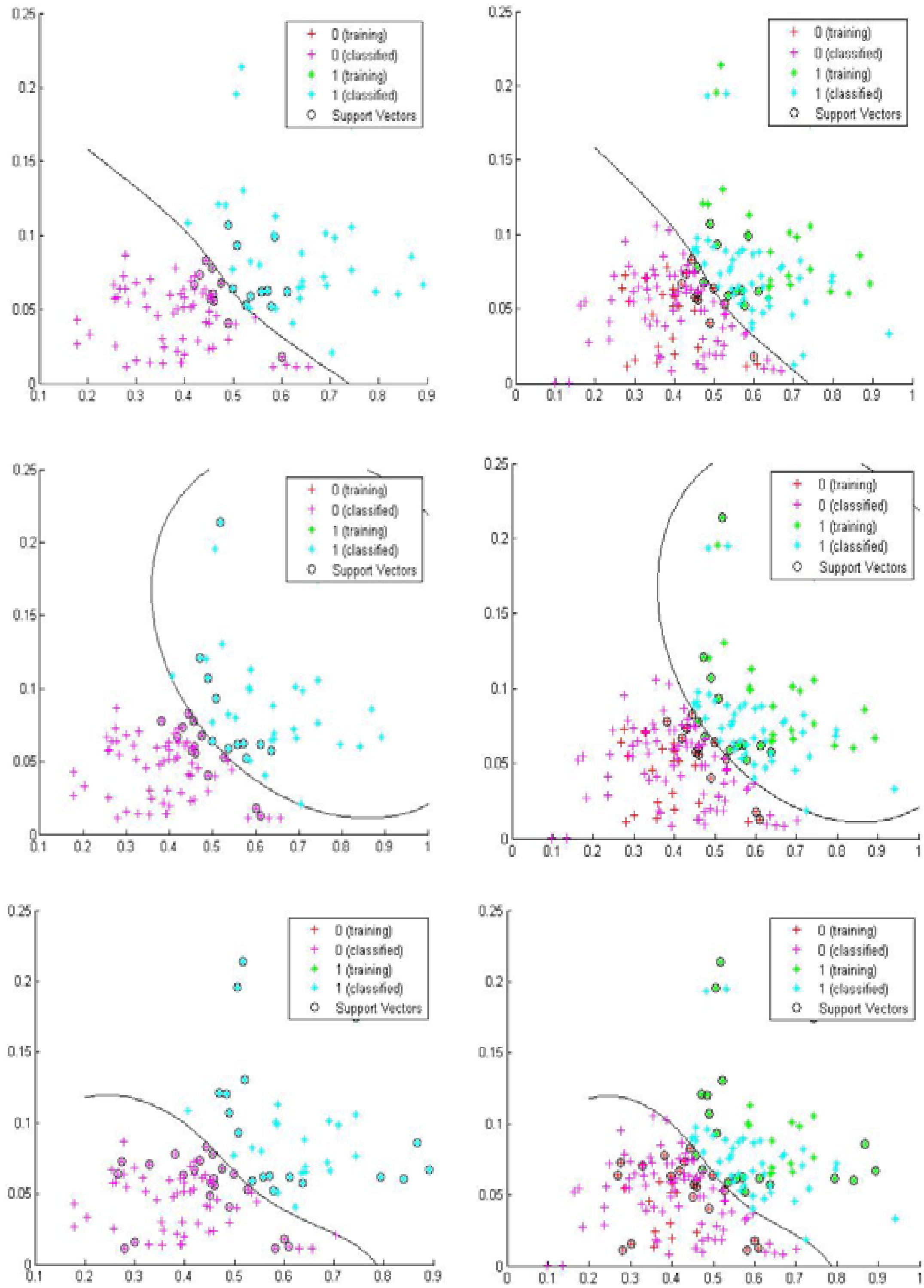


Figure 5.9 SVM classified features for various kernels (a) linear (b) radial basis function (c) polynomial (d) quadratic

Table 5.1 CMR values with different kernels

Kernel	CMR in %	
	Threshold=0.005	Threshold=0.0009
Linear	43.00	33.33
Radial Basis Function	42.46	33.33
Polynomial	44.93	37.50
Quadratic	43.66	35.30

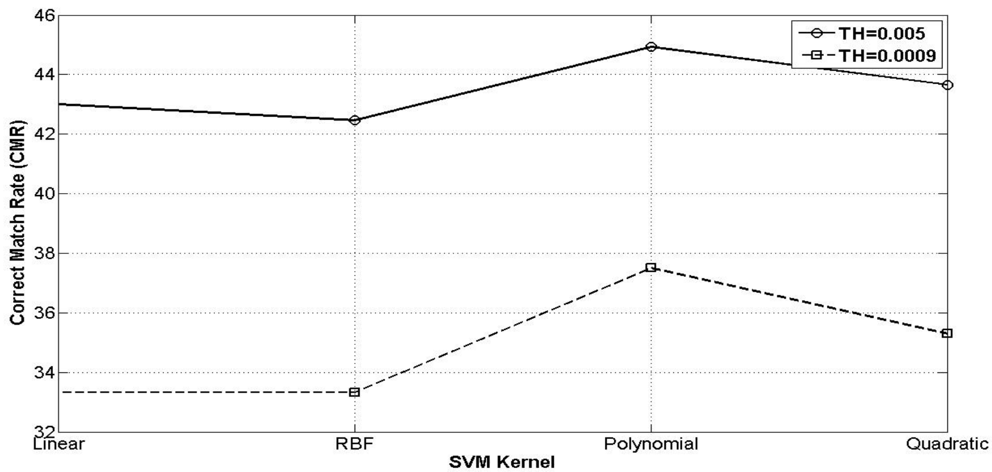


Figure 5.10 Graph of CMR with refinement using SVM classification with various kernels

In the third image pair of TM, the one image is captured in 2009 and another image is captured in 1999. So, drastic changes in the scene can be observed. If some of the feature points associated with the changes in first image then there is no correct match present in the second (here older) image. So there is likelihood of matching of such features points related to the changes, to some other similar features of the second image. This leads to incorrect matches and hence reduces CMR. Based on the previous analysis of some of the SVM kernels, here polynomial kernel is used for this third image pair. By changing the threshold of SURF, CMR is computed for both the cases, i.e. without and with the feature refinement step using SVM classification. The corresponding CMR values are listed in Table 5.2. It shows in every observation, compared to SURF only, the CMR is improved if the SVM based refinement of feature points is performed before matching step. These values are shown in the graph of Fig. 5.11.

This result can be explained by another viewpoint that is by maintaining the numbers of detected features constant. From the Table 5.2, the numbers of matched features in case of threshold 0.0008 and without feature rectification are near about the same number of matched features in case of threshold 0.0005 and with feature refinement. If these two are compared based on equal matched points then we can observe that the refinement of features give about 5% improvement in CMR.

Table 5.2 CMR values without and with refinement using SVM classification with different threshold

TH	SURF only (without feature refinement)			SURF+SVM (with feature refinement)		
	No. of matched points	No. of correctly matched points	CMR (%)	No. of matched points	No. of correctly matched points	CMR (%)
0.0005	1563	574	36.72	**948**	**383**	**40.40**
0.0008	**949**	**333**	**35.00**	500	202	40.40
0.0009	822	283	34.43	413	164	39.71
0.001	711	241	33.90	340	135	39.70
0.002	201	60	29.85	76	29	38.16

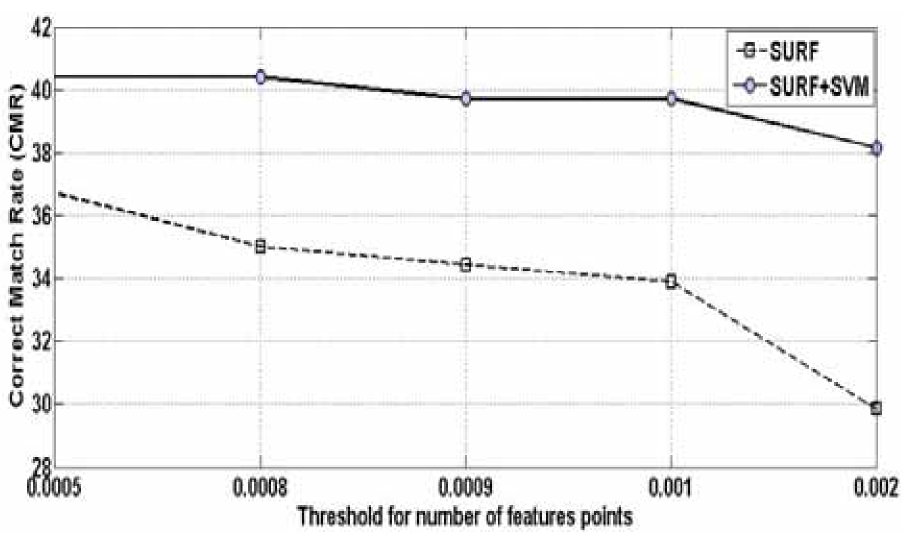


Figure 5.11 Graph of CMR without and with refinement using SVM classification with different threshold

5.6 Summary

Similar to some other work, this work is also for the improvement in one the steps of feature based image registration that is matching step. A feature refinement step is added before the feature matching step. The feature refinement step can be targeted to the features related to some specific occlusions in the image such as clouds and shadows. Presence of such features gives some incorrect matches and misleads the registration process. Due to the addition of feature refinement step before matching step removes such features so reduces incorrect matches. This is achieved by SVM classification. This results in improved correct match rate, which is presented by related graphs. Kernel methods are also compared with reference to correct match rate and polynomial kernel is found to be better. This approach of refinement of irrelevant features can be generalised also. For example in multi-temporal image registration, due to the presence of changes, they may contribute for incorrect matches. But the features corresponding to those changes can be refined before matching step, which also improves correct match rate.

CHAPTER 6

Conclusion and Future Work

6.1 Conclusion

In the study various methods for image registration for satellite images are investigated. The satellite images or remote sensing images are multi-view, multimodal, multi-sensor, and multi-temporal; they also have noise, occlusions, clouds, nonlinear illumination and other atmospheric effects. Based on the study, various approaches for image registration can be categorised as intensity based methods, feature based methods, transform domain based methods and combined or hybrid methods. But no single method is available which can work for all kinds of satellite images. For any algorithm, it works for the specific types of satellite images and/or it addresses the specific challenge/s. In this study the initial work is on intensity based methods followed by work related to feature based methods.

All the steps of image registration are assumed without any human intervention i.e. automatic image registration is under consideration.

In the intensity based method, mutual information is used as similarity measure. But to address its computational complexity, maximum likelihood based mutual information method is used, which results in reduction in computation time. Still computation is large especially for large satellite images, so the approach is suitable only in the case of if the initial solution is close to the final solution i.e. geometrical transformation should be within certain limit.

For a transform domain based approach, Fourier transform based approach is studied followed by use of radon transform for image registration. In such transform domain based approach, the spatial information is transformed to another domain and the properties of the domains are used to find the registration parameters. In the work based on radon transform rotation, horizontal translation and vertical translation parameters are estimated which are found close to the known applied parameters. Further robustness to noise is also observed in the approach. In the approach the estimation is single shot instead of iterative as found in intensity based method; so faster but still the accuracy depends on the resolution of the transform domain; which if increased the computation time also increases for large satellite images.

In general both MLMI based and radon based approaches have the problem of considerable computational time, can't be reduced significantly because the approaches basically work on every pixel of the image. Due to this inherent limitation of area based method there is little scope of work or improvement; preferred only if the initial solution is close to the final solution.

In further work based on feature based methods, the feature descriptor and feature matching steps are investigated. In one of the known feature extraction method that is SURF, its descriptor is replaced by HOG descriptor. This is to address the nonlinear illumination variation exists in the satellite images. Depending on the illumination variation exists between the two images, improvement in correct match rate in the range of 5-10% is found on the BHUVAN dataset of multi-temporal, multispectral images as well as on some datasets used in few other works. In similar manner, correct match rate with the top twenty matches only is also improved. This can also be used as one of the performance parameter as few numbers of correct matched points are required to compute the image registration parameters depending on the geometrical transformation or degree of freedom under consideration.

In the matching step, the correct match rate is improved by performing feature refinement step. This refinement step is targeted to clouds. The presence of clouds misleads the registration process and may give inaccurate registration or even failure of image registration. So based on the patch around the extracted features, they are classified as cloudy and non-cloudy features using SVM classifier. Then only the refined i.e. non-cloudy features are allowed to perform the feature matching step. This change in the algorithm is also investigated on the image having significant changes in the scene. The changes in the scene play similar role as cloud. So the features corresponding to the changes have been refined using SVM classification technique. This strategy has improved the correct match rate depending on the data. The SVM classification is kernel based so selection of kernel is critical and no standard way to know the suitable kernel for some specific application. So, comparative analysis of various kernels of SVM for correct match rate is carried out. It is found that the polynomial kernel is giving better result in terms of correct match rate. This approach of refining the feature points before the feature matching step can be generalized and extended to any other kind of occlusions present in image which degrades the registration process. One such example is shadow. Depending on day time at which the images captured, in both the images have different amount of shadows i.e. shorter or longer shadow. The features associated with them can also mislead the registration process. So by similar manner those can be removed before matching step.

6.2 Future Work

From the study here there are some possible future enhancement for the image registration algorithms for satellite images and remote sensing applications:

- Comparatively little work and improvement is found to address image registration of the challenging radar images. One of the main issues with the radar images is the speckle noise which creates difficulties in similarity measure, feature description, and feature matching steps of the image registration.
- One specific algorithm will not work for all community of satellite images. Generally, any of the component/s of the image registration framework are focused by researchers and they have contributed to improve the image registration algorithm with reference to applications or satellite datasets at hand. So a system can be prepared in which such algorithms are available in library and based on the input image characteristics, the components can be selected by the expert. Further, instead of the expert, first an intelligent system can be prepared which decides the components to be used based on the required applications or characteristics of the datasets at hand.
- Some standard datasets are available for many image processing and analysis application including image registration in computer vision and medical field. But no such standard satellite or remote sensing image datasets are available for image registration. The reason is diversity in the related applications, sensors and airborne mission. So, difficult to prepare some dataset of satellite images. Even though some effort may be placed to prepare a set of satellite images having some specific characteristics in general. Further there is a difficulty in comparing various

algorithms as most of them are applied on the images available to them. So preparing a satellite image datasets for image registration may also helpful to compare various algorithms on the same datasets i.e. evaluation of image registration algorithm on the general satellite datasets makes it more convincing.

- For the feature refinement SVM is used for classification. Alternatives for the classification can also be analysed for the comparative analysis purpose. Further the training can be made more systematic for the satellite image datasets.

List of References

A. Sedaghat, M. M. (2011). Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing, vol. 49*(11), pp. 4516-4527.

A. Taravat, F. D. (2015). Neural Networks and Support Vector Machine Algorithms for Automatic Cloud Classification of Whole-Sky Ground-Based Images. *IEEE Geoscience and Remote Sensing Letters, vol. 12*(3), pp. 666-670.

Althof, R. J. (1997). A Rapid and Automatic Image Registration Algorithm. *IEEE Transactions on Medical Imaging*, pp. 308–316.

Arya, K. V. (2007). Image Registration Using Robust M-estimators. *Pattern Recogn. Lett., vol. 28*(15), 1957-1968.

B Summar, O. C. (2005). Image registration using Hough transform and Phase correlation in Wavelet domain. *ICCAS.* Gyeonggi-Do, Korea.

Baillarin, S. B. (2005). Using a Three Dimensional Spatial Database to Orthorectify Automatically Remote Sensing Images. *Workshop on Service and Application of Spatial Data Infrastructure*, (pp. 89–93).

Barbara Zitová, J. F. (2003). Image registration methods: a survey. *Image and Vision Computing , vol.21*(11), pp. 977-1000.

Bentoutou, Y. T. (2005). An Automatic Image Registration for Applications in Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, pp. 2127-2137.

Bolles, M. A. (1981). Random sample consensus: A paradigm formodel fitting with applications to image analysis and automated cartography. *Communications of the ACM*, pp. 381–395.

Brown, L. G. (1992). A Survey of Image Registration Techniques. *ACM Computer Survey, 24*(4), 325-376.

C. Huo, C. P. (2012). Multilevel SIFT Matching for Large-Size VHR Image Registration. *IEEE Geoscience and Remote Sensing Letters, vol. 9*(2), pp. 171-175.

C. Serief, M. B. (2008). Elastic registration of remote-sensing images based on the nonsubsampled contourlet transform. *16th European Signal Processing Conference*, (pp. 1-5.).

Cariou, C. a. (2008). Automatic georeferencing of airborne pushbroom scanner images with missing ancillary data using mutual information. *IEEE Transactions on Geoscience and Remote Sensing, vol. 46*(5), pp. 1290–1300.

Carrion, D. G. (2002). GEOREF: A software for improving the use of remote sensing images in environmental applications. *Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society*, (pp. 360–365.). Lugano, Switzerland.

Chapelle Olivier, V. V. (2003, September). Model Selection for Support Vector Machines. *Neurocomputing*, pp. 221-249.

Chapelle, O. (1998). Support Vector Machines for Image Classification. pp. 1-31.

Chatterji, B. S. (1996). An FFT-Based Technique for Translation, Rotation and Scale-Invariant Image Registration. *IEEE Transcations on Image Processing*, pp. 1266-1271.

Choi, I. H. (2013). Accurate Registration using Adaptive Block Processing for Multi-spectral Images. *IEEE Transactions on Circuits and Systems for Video Technology, vol. 23*(9), pp. 1491-1501.

Clausi, A. W. (2007). ARRSI: Automatic Registration of Remote-Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing, vol. 45*(5), pp. 1483-1493.

Cochoff, Y.-M. Z. (2002). Likelihood maximization approach to image registration. *IEEE Transactions on Image Processing, vol. 11*(12), pp. 1417-1426.

Cole-Rhodes, A. J. (2003). Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Transactions on Image Processing, vol. 12*(12), pp. 1495–1511.

Collignon, A. M. (1995). Automated multimodality image registration based on information theory. *Proceedings of the 14th International Conference on Information Processing in Medical Imaging*, (pp. 263–274).

David, W. A. (2010). AISIR: Automated Inter-sensor/Inter-band Satellite Image Registration Using Robust Complex Wavelet Feature Representations. *Pattern Recognition Letters, vol. 31*(10), pp. 1160-1167.

Deans, S. R. (2000). Radon and Abel Transform. CRC Press.

Dewdney, A. K. (1978). Analysis of a steepest-descent image-matching algorithm. *Pattern Recognition, 10*(1), pp. 31-39.

Dowson, N. a. (2006). A unifying framework for mutual information methods for use in non-linear optimisation. *Proceedings of the Ninth European Conference on Computer Vision*, (pp. 365–378).

Dowson, N. K. (2008). Estimating the joint statistics of images using nonparametric windows with application to registration using mutual informatio. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30*(10), pp. 1841–1857.

Dufournaud, Y. S. (2004). Image matching with scale adjustment. *Computer Vision and Image Understanding, vol. 93*(2), pp. 175–194.

E. Abraham, S. M. (2013). HOG descriptor based registration (A new image registration technique). *15th International Conference on Advanced Computing Technologies (ICACT)*, (pp. 1-4).

Eastman, J. L. (2010). *Image Registration for Remote Sensing.* Cambridge University Press.

Eikvil, L., Holden, M., & Huseby, R. B. (2009). Adaptive registration of remote sensing images using supervised learning. *Photogrammetric Engineering and Remote Sensing, vol. 75*(11), pp. 1297-1306.

Elsalamony, H. A. (2015). Comparing Proposed Signature with SURF in Object Detection Process. *IETE Journal of Research*, pp. 466-474.

Ervin Sejdic, I. D. (2011). Fractional Fourier transform as a signal processing tool: An overview of recent developments. *Signal Processing*, pp. 1351-1369.

F. Maes, A. C. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging, vol. 16*(2), pp. 187-198.

Fan, B. H. (2013). Registration of Optical and SAR Satellite Images by Exploring the Spatial Relationship of the Improved SIFT. *IEEE Geoscience and Remote Sensing Letters, vol. 10*(4), pp. 657-661.

Flora Dellinger, J. D. (2015). SAR-SIFT : A SIFT-Like Algorithm for SAR Images. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, pp. 453-466.

Forstner, W. a. (1987). A fast operator for detection and precise location of distinct points, corners, and centers of circular features. *Proceedings of the ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data*, (pp. 281–385). Interlaken, Switzerland.

Fujisada, A. I. (2005). ASTER geometric performance. *IEEE Transactions on Geoscience and Remote Sensing, vol. 43*(12), pp. 2700-2706.

Giro, I. J. (2004). On the possibility of automatic multisensor image registration. *IEEE Transactions on Geoscience and Remote Sensing, vol. 41*(10), pp. 2104–2120.

Gong, M. Z. (2014). 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*(7), pp. 4328-4338.

Goshtasby, A. A. (2005). *2-D and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications.* Wiley-Interscience.

Grodecki, J. a. (2005). *IKONOS geometric calibrations, Technical report, Space Imaging.*

H B kekre, D. T. (2012). 2D Satellite Image Registration Using Transform Based and Correlation Based Methods. *International Journal of Advanced Computer Science and Applications, vol. 3* (5).

H. Goncalves, J. A.-R. (2011). HAIRIS: A Method for Automatic Image Registration Through Histogram-Based Image Segmentation. *IEEE Transactions on Image Processing, vol. 20*(3), pp. 776-789.

H. Li, B. M. (1995). Multisensor Image Fusion Using the Wavelet Transform. *Graphical Models and Image Processing, vol. 57*(3), pp. 235-245.

Harold S. Stone, B. T. (2003). Analysis of Image Registration Noise due to Rotationally Dependent Aliasing. *Journal of Visual Communication and Image Representation, vol. 14*(2), pp. 114-135.

Harris, C. a. (1988). A combined corner and edge detector. *Proceedings of the Fourth Alvey Vision Conference*, (pp. 147–151). Manchester.

Herbert Bay, A. E. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, pp. 346-359.

Hong, G. (2007). *Image Fusion, Image Registration, and Radiometric Normalization for High Resolution Image Processing.* Ph.D. dissertation, University of New Brunswick, Department of Geodesy and Geomatics Engineering, New Brunswick, Canada.

Hong, S. H. (2006). Extraction of ground control points (GCPs) from synthetic aperture radar images and SRTM DEM. *International Journal of Remote Sensing, vol. 27*(18), pp. 3813–3829.

J. P. W. Pluim, J. B. (2003). Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging, vol. 22*(8), pp. 986-1004.

J. R. G. Townshend, C. O. (1992). The impact of misregistration on change detection. *IEEE Transactions on Geoscience and Remote Sensing, vol. 30* (5), pp. 1054-1060.

J., M. A. (2005). InSAR Image Co-registration using the Fourier-Mellin Transform. *International Journal of Remote Sensing, vol. 26*(13), pp. 2865–2876.

J.B.Antoine Maintz, M. A. (1998). A survey of medical image registration. *Medical Image Analysis, vol. 2*(1), pp. 1-36.

Kelman, A. S. (2007). Keypoint descriptors for matching across multiple image modalities and non-linear intensity variations. *Proceedings of the IEEE Computer Vision and Pattern Recognition Workshop on Image Registration and Fusion*, (pp. 1–7). Minneapolis.

Kern, J. a. (2007). Robust multispectral image registration using mutual-information models. *IEEE Transactions on Geoscience and Remote Sensing, vol. 45*(5), pp. 1494–1505.

Khorram, X. D. (1998). The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing, vol. 36*(5), pp. 1566-1577.

Krahmer, F. (2006). Blind image deconvolution: Motion blur estimation. *IMA Preprints Series*.

Lee, S. R. (2010). A Coarse-to-Fine Approach for Remote-Sensing Image Registration Based on a Local Method. *International Journal of Smart Sensing and Intelligent Systems, vol. 3*(4), pp. 690-702.

Li, Q. W. (2009). Robust Scale-Invariant Feature Matching for Remote Sensing Image Registration. *IEEE Geoscience and Remote Sensing Letters, vol. 6*(2), pp. 287-291.

Liberata, C. M. (2008). Image registration using non-linear diffusion. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, (pp. 220–223). Boston.

Lin, Z. H. (2014). Image registration research based on SUSAN-SURF algorithm. *26th Chinese Control and Decision Conference (2014 CCDC)*, (pp. 5292-5296). Changsha.

Lowe, D. (1999). Object recognition from local scale-invariant features. *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 1150–1157). Corfu, Greece.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, vol. 20*(2), pp. 91–110.

Luis Gomez-Chova, J. A.-V.-M.-V. (2012). CLOUD-SCREENING FROM MULTISPECTRAL SATELLITE. *1st EARSeL Workshop on Temporal Analysis of Satellite Images*, (pp. 81-86). Mykonos, Greece.

M. Hasan, M. R. (2012). Robust Automatic Registration of Multimodal Satellite Images Using CCRE With Partial Volume Interpolation. *IEEE Transactions on Geoscience and Remote Sensing, vol. 50*(10), pp. 4050-4061.

M. Hasan, X. J. (2010). Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints. *IEEE International Geoscience and Remote Sensing Symposium*, (pp. 1011-1014). Honolulu.

M.M. El-gayar, H. S. (2013). A comparative study of image low level feature extraction algorithms. *Egyptian Informatics Journal, vol. 14*(2), pp. 175-181.

Maes, F. V. (1999). Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis, vol. 3*(4), pp. 272–286.

Mahmudul Hasan, X. J.-K. (2010). Multi-spectral Remote Sensing Image Registration via Spatial Relationship Analysis on SIFT Keypoints. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1011-1014.

Martucci, S. H. (2001). A Fast Direct Fourier-Based Algorithm for Subpixel Registration of Images. *IEEE Transactions on Geoscience and Remote Sensing, vol. 39*(10), pp. 2235-2243.

Medha V. Wyawahare, D. P. (2009). Image Registration Techniques : An overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, pp. 11-28.

mei, C. H. (2003). Mutual information-based image registration for remote sensing data. *International Journal of Remote Sensing*, pp. 3701-3706.

Moigne, I. Z. (2005). Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery. *IEEE Transactions on Image Processing, vol. 14*(6), pp. 770-782.

Nies, H. L. (2008). Image registration of TerraSAR-X data using different information measures. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, (pp. 419–422). Boston.

Orchard, J. (2007). Efficient Least Squares Multimodal Registration With a Globally Exhaustive Alignment Search. *IEEE Transactions on Image Processing, vol. 16*(10), pp. 2526-2534.

P, A. A. (2008). Non-rigid registration of multi-modal images using both mutual information and cross-correlation. *Medical Image Analysis, vol. 12*(1), pp. 3-15.

P. Thevenaz, U. E. (1998). A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing, vol. 7*(1), pp. 27-41.

Pattichis, J. P. (2007). Robust Multispectral Image Registration Using Mutual-Information Models. *IEEE Transactions on Geoscience and Remote Sensing, vol. 45*(5), pp. 1494-1505.

Petrou, G. L. (2006). Image registration using the Walsh transform. *IEEE Transactions on Image Processing, vol. 15*(8), pp. 2343-2357.

Ping, H. J. (1997). Image Registration Using a New Edge-Based Approach. *Computer Vision Image Understanding, vol. 67*(2), pp. 112-130.

Pradip M. Patil, M. P. (2012). Investigation of Shoeprints Using Radon Transform With Reduced Computational Complecity. *Journal of Pattern Recognition Research, 7*(1).

Qin-Sheng Chen, M. D. (1994). Symmetric Phase-only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16*(11), pp. 1156-1168.

R. D. Eastman, J. L. (2007). Research issues in image registration for remote sensing. *IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 1-8).

R. J. Althof, M. G. (1997). A rapid and automatic image registration algorithm with subpixel accuracy. *IEEE Transactions on Medical Imaging*, pp. 308-316.

R. K. Gambhir, S. M. (2013). Comparative study of intensity based cost functions for automated satellite image registration. *3rd IEEE International Advance Computing Conference (IACC)*, (pp. pp. 1240-1245). Ghaziabad.

R. Matungka, Y. F. (2008). Image registration using Adaptive Polar Transform. *15th IEEE International Conference on Image Processing*, (pp. 2416-2419). San Diego, CA.

R.M. Ezzeldeen, H. R.-W. (2010). Comparative study for image registration techniques of remote sensing images. *The Egyptian Journal of Remote Sensing and Space Science, 13*(1), pp. 31-36.

Rajwade, A. B. (2009). Probability density estimation using isocontours and isosurfaces: Applications to information-theoretic image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31*(3), 475–491.

Reinartz, S. S. (2010). Mutual-Information-Based Registration of TerraSAR-X and Ikonos Imagery in Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing, vol. 48*(2), pp. 939-949.

Roy, J. J. (2008). The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sensing of Environment, vol. 112*(3), pp. 1196–1211.

S. Chen, Q. G. (2011). A Maximum Likelihood Approach to Joint Image Registration and Fusion. *IEEE Transactions on Image Processing, vol. 20*(5), pp. 1363-1372.

S. Klein, M. S. (2010). elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging, vol. 29*(1), pp. 196-205.

Sabuncu, M. R. (2006). *Entropy-based Image Registration.* NJ, USA: Princeton University.

Schmid, C. M. (2000). Evaluation of interest point detectors. *International Journal of Computer Vision, vol. 37*(2), pp. 151–172.

Schmid, M. K. (2004). Scale & affine invariant interest point detectors. *International Journal of Computer Vision, vol. 60*(1), pp. 63–86.

Schmid, M. K. (2005). A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27*(10), pp. 1615–1630.

Sejdi, E. a. (2011). Fractional Fourier Transform As a Signal Processing Tool: An Overview of Recent Developments. *Signal Processing, vol. 91*(6), 1351-1369.

Shams, R. S. (2007). Gradient intensity: A new mutual information based registration method. *Proceedings of the IEEE Computer Vision and Pattern Recognition Workshop on Image Registration and Fusion*, (pp. 3249–3256).

Shehan Fernanado, T. C. (2011). Rotation Invariant Image Registration with Radon Transform. *SAITM Research Symposium on Engineering Advancements.*

Shen, Z. Z. (2010). Automatic remote sensing image registration based on SIFT descriptor and image classification. *18th International Conference on Geoinformatics .*

Solberg, R. B. (2005). A model-based approach for geometrical correction of optical satellite images. *International Journal of Remote Sensing, vol. 26*(15), pp. 3205-3223.

Soltanian-Zadeh, K. J.-K. (2005). Radon transform orientation estimation for rotation invariant texture analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27*(6), pp. 1004-1008.

Sukthankar, R. K. (2004). PCA-SIFT: A more distinctive representation for local image descriptors. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 506–513). Washington, DC.

T. Suzuki, M. S. (2009). Mutual information approximation via maximum likelihood estimation of density ratio. *IEEE International Symposium on Information Theory*, (pp. 463-467). Seoul.

Temizel, M. T. (2010). Multi-spectral Satellite Image Registration Using Scale-Restricted SURF. *20th International Conference on Pattern Recognition*, (pp. 2310-2313). Istanbul.

Triggs, N. D. (2005). Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, (pp. 886-893). San Diego, CA, USA.

Tuytelaars, T. a. (2008). *Local Invariant Feature Detectors: A Survey.* Hanover: MA:Now Publishers, Inc.

Uma, G. (2014). *Intensity based image registration of satellite images using evolutionary techniques, MTech thesis.*

Vaithiyanathan, K. D. (2016). Image Registration Techniques Based on the Scale Invariant Feature Transform. *IETE Technical Review*, pp. 22-29.

Varshney, X. M. (2009). A Subspace Method for Fourier-Based Image Registration. *IEEE Geoscience and Remote Sensing Letters, vol. 6*(3), pp. 491-494.

Verleysen, D. G. (2012). A Cmparison of multivariate mutual information estimators for feature selection. *Proceedings 1st International Conference on Pattern Recognition Applications and Methods*, (pp. 176-185). Belgium.

W. Pan, K. Q. (2009). An Adaptable-Multilayer Fractional Fourier Transform Approach for Image Registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31*(3), pp. 400-414.

Walters-Williams, J. (2009). Estimation of Mutual Information: A Survey. *Proceedings 4th International Conference:Rough Sets and Knowledge Technology* (pp. pp. 389-396). Gold Coast, Australia: Springer Berlin Heidelberg.

Wang Kai, C. B. (2012). Multi-source Remote Sensing Image Registration Based on Normalized SURF Algorithm. *International Conference on Computer Science and Electronics Engineering*, (pp. 373 - 377). Hangzhou.

Wang, S. Y. (2012). BFSIFT: A Novel Method to Find Feature Matches for SAR Image Registration. *IEEE Geoscience and Remote Sensing Letters, vol. 9*(4), pp. 649-653.

Wells, P. V. (1995). Alignment by maximization of mutual information. *Proceedings of IEEE International Conference on Computer Vision*, (pp. 16-23).

Wolberg, S. Z. (2005). Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations. *IEEE Transactions on Image Processing, vol. 14*(10), pp. 1422-1434.

Wu J., C. Z. (2013). A Comparative Study of SIFT and its Variants. *Measurement Science Review, vol. 13*(3), pp. 122-131.

Xiaopeng, Z. D. (2014). Rapid Multimodality Registration Based on MM-SURF. *Neurocomputing*, pp. 87-97.

Y. C. Chen, C. S. (2013). In-Plane Rotation and Scale Invariant Clustering using Dictionaries. *IEEE Transactions on Image Processing, 22*(6), 2166-2180.

Y. Keller, Y. S. (2005). The Angular Difference Function and its Application to Iage Registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27*(6), pp. 969-976.

Yang, G. S. (2006). Automatic robust image registration system: Initialization, estimation, and decision. In Proceedings of the. *Fourth IEEE International Conference on Computer Vision Systems*, (pp. 23–31). New York, NY.

Zavorine, J. L. (2000). Use of wavelets for image registration. *SPIE*, pp. 99-108.

Zhang, G. H. (2007). Combination of feature-based and area-based image registration technique for high resolution remote sensing image. *IEEE International Geoscience and Remote Sensing Symposium*, (pp. 377-380). Barcelona.

Zhang, Z. X. (2009). A Novel Interest-Point-Matching Algorithm for High-Resolution Satellite Images. *IEEE Transactions on Geoscience and Remote Sensing, vol. 47*(12), pp. 4189-4200.

Zuliani, M. (2006). *Computational methods for automatic image registration.* PhD Thesis, UNIVERSITY of CALIFORNIA, Santa Barbara.

List of Publications

Papers

- Manish I. Patel, V. K. Thakar, "Application of Radon Transform for Fast Image Registration", 2nd IEEE International Conference on Advanced Computing and Communication Systems, Coimbatore,5-7 Jan. 2015
- Manish I. Patel, V. K. Thakar, "Speed Improvement in Image Registration using Maximum Likelihood based Mutual Information", 2nd IEEE International Conference on Advanced Computing and Communication Systems, Coimbatore,5-7 Jan. 2015
- Manish I. Patel, V. K. Thakar, Shishir Shah, "Image Registration of Satellite Images with Varying Illumination Level using HOG Descriptor based SURF", 6th International Conference on Advances in Computing & Communications 2016. (Paper is published in Elsevier Procedia Computer Science, Volume 93, 2016, Pages 382-388, ISSN 1877-0509 and the publication is made available on sciencedirect.com)
- Manish I. Patel, "Image Registration with Reduced Dimensionality of SURF Descriptor using PCA", International Conference on Technology and Management, Visnagar 17-18, February, 2017 (Second Student-Best-Paper Award)

Miscellaneous

- Presented work in special session on PhD Forum During National Seminar on Computer Vision and Image Processing-2014 (NaSCoVIP-2014) an IEEE Gujarat section event jointly organized by SCET, Surat and SVNIT, Surat during 19-20, September, 2014
- Attended QIP short term course - on "Advanced Techniques for Satellite Image Analysis" at IIT, Bombay, during 10-14 February, 2014