Automatic Registration of Multi-Modal Airborne Imagery

by

Xiaofeng Fan

B.S. University of Science & Technology of China, 1997
M.S. University of Science & Technology of China, 2001

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology

May. 1, 2011

Signature of the Author ______________________________

Accepted by ______________________________
Coordinator, Ph.D. Degree Program Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Xiaofeng Fan
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation required for the
Ph.D. degree in Imaging Science

____
Dr. Harvey Rhody, Dissertation Advisor

____
Dr. Eli Saber

____
Dr. Carl Salvaggio

____
Dr. Robert Teese

____
Date

DISSERTATION RELEASE PERMISSION

ROCHESTER INSTITUTE OF TECHNOLOGY

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

Title of Dissertation:

Automatic Registration of Multi-Modal Airborne Imagery

I, Xiaofeng Fan, hereby grant permission to Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature ______________________________

Date

Automatic Registration of Multi-Modal Airborne Imagery

by

Xiaofeng Fan

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment of the requirements
for the Doctor of Philosophy Degree
at the Rochester Institute of Technology

Abstract

This dissertation presents a novel technique based on Maximization of Mutual Information (MMI) and multi-resolution to design an algorithm for automatic registration of multi-sensor images captured by various airborne cameras. In contrast to conventional methods that extract and employ feature points, MMI-based algorithms utilize the mutual information found between two given images to compute the registration parameters. These, in turn, are then utilized to perform multi-sensor registration for remote sensing images. The results indicate that the proposed algorithms are very effective in registering infrared images taken at three different wavelengths with a high resolution visual image of a given scene. The MMI technique has proven to be very robust with images acquired with the Wild Airborne Sensor Program (WASP) multi-sensor instrument. This dissertation also shows how wavelet based techniques can be used in a multi-resolution analysis framework to significantly increase computational efficiency for images captured at different resolutions.

The fundamental result of this thesis is the technique of using features in the images to enhance the robustness, accuracy and speed of MMI registration. This is done by using features to focus MMI on places that are rich in information. The new algorithm smoothly integrates with MMI and avoids any need for feature-matching, and then the applications of such extensions are studied. The first extension is the registration of cartographic maps and image datum, which is very important for map updating and change detection. This is a difficult problem because map features such as roads and buildings may be mis-located and features extracted from images may not correspond to map features. Nonetheless, it is possible to obtain a general global registration of maps and images by applying statistical techniques to map and image features. To solve the map-to-image registration problem this research extends the MMI technique through a focus-of-attention mechanism that forces MMI to utilize correspondences that have a high probability of being information rich.

The gradient-based parameter search and exhaustive parameter search methods are also compared. Both qualitative and quantitative analysis are used to assess the registration accuracy.

Another difficult application is the fusion of the LIDAR elevation or intensity data with imagery. Such applications are even more challenging when automated registrations algorithms are needed. To improve the registration robustness, a salient area extraction algorithm is developed to overcome the distortion in the airborne and satellite images from different sensors. This extension combines the SIFT and Harris feature detection algorithms with MMI and the Harris corner label map to address difficult multi-modal registration problems through a combination of selection and focus-of-attention mechanisms together with mutual information. This two-step approach overcomes the above problems and provides a good initialization for the final step of the registration process. Experimental results are provided that demonstrate a variety of mapping applications including multi-modal IR imagery, map and image registration and image and LIDAR registration.

Acknowledgements

First and foremost I offer my sincerest gratitude to my advisor, Dr Harvey Rhody, who supported me throughout my research with his patience and knowledge. I attributes the level of my Ph.D degree to his inspiration and effort. Without him my research and this thesis would not have been completed or written. I also wishes to express my gratitude to my co-adviser, Dr. Eli Saber, who was abundantly helpful and offered invaluable assistance, support and guidance.

Deepest gratitude are also due to the members of the graduate committee, Dr. Carl Salvaggio and Dr. Robert Teese, for all their advice, encouragement. Without whose knowledge and assistance this dissertation would not have been successful.

I appreciate the countless help from Image Algorithm and System (LIAS) group. Special thanks go to Dr. John Kerekes and Dr. Jan van Aardt, who has offered me much advice and insight throughout my research. I also thank the LIAS lab staff, especially Mr. Donald McKeown. In my daily work I have been blessed with a friendly and cheerful group. Mr. Jason Faulring and Mr. Robert Krzaczek help me to solve the software and equipment issue. Without their help, my experiments would not have been successful. I also acknowledge the USA National Geospatial Intelligence Agency for their financial support for this project.

I am also indebted to many faculties and admin staffs of the Chester F. Carlson Center of Imaging Science, especially Ms. Susan Chan, Ms. Cindy Schultz, Ms. Joyce French and Mr. Joe Pow. Thanks also go to other staffs for all student assistant and help.

The Chester F. Carlson Center of Imaging Science has provided the support and equipment I have needed to produce and complete my Research.

Last but not the least, I would like to thank my family, especially my wife and my parents for supporting and encouraging me to pursue this degree. Without their encouragement, I would not have finished the degree.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

During the last twenty years with the development of high-speed digital image processing, there are more and more applications in remote sensing, computer vision and medical imaging. Those applications bring forward a requirement for accurate and efficient image registration which is a necessary pre-treatment for processes that combine multiple images. Image fusion enables the extraction of the information about scene objects that can't be gotten from any single image. In remote sensing, image registration is applied for multi-spectral classification, monitoring the weather and integrating the information from the geographic information systems (GIS) [1–3]; In medical imaging, registration helps doctors combine computer tomography (CT) and nuclear magnetic resonance (NMR) images to achieve better analysis of the disease area and treatment verification [4–6]; In cartography, registration is used for map updating and environment monitoring [7].

Image registration is the process of overlaying two or more images of the same scene taken at different times, under different lighting conditions, from different viewpoints and/or by different sensors [8]. These overlapping images can detect the difference between the two images (moving object detection [9,10]), stack information in 3D medical imaging

[4, 11] [12] and reveal the relationship between the two images [13]. This combination of the raw data from various sources is the critical step for many modern analysis methods (PCA [14], ICA [15] and multi-spectral classifications etc. [16]).

The core problem for image registration is to project the raw images into a common coordinate system to enable image analysis that is based on multiple views of the same scene.

1.1 A classification of image registration

Image registration falls into four categories based on acquisition [8]:

- Different location (Multi-view) registration

  The raw data for registration are the images of the same region taken from different viewpoints. The aim is to mosaic the images [17–19] or do stereo image construction [20–23].

- Different time (Multi-temporal) registration

  The raw data for registration are the images of the same region taken at different times or under different conditions [24, 25]. This registration task is to evaluate the change that happened in the target region [26, 27]. This technology is widely used in the environment monitoring and surveillance systems [28–30].

- Different sensor or multi-model

  The raw data for registration are the images of the same region taken from different sensors. This kind of image registration has received more and more attention recently especially in medical image processing [31–33] and remote sensing [34–36]. In medical image, the multi-modal image registration can overlap the photos from

X-ray, CT or MRI and give doctor detailed information about the patient's disease [37–39]. In remote sensing this technology is useful in combining panchromatic and hyper-spectral images [1, 40].

- Scene to model registration

  This registration is not between real images but a registration between the real image and a model of certain scene objects [38]. The model is based on previous knowledge: in medical imaging, the model could be a digital representation of anatomy [41]; while in remote sensing the model could be the GIS data or a map [42, 43]. This registration is one of the most difficult applications but very valuable for medical diagnosis, specimen classification and map updating [44].

A general block diagram of image registration processing is shown in Figure 1.1.

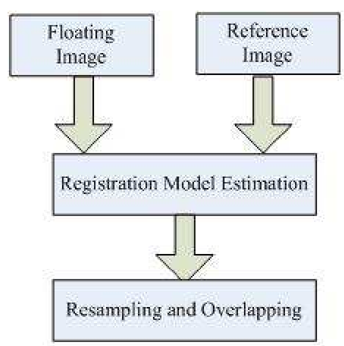


Figure 1.1: Blockchart for the image registration

The above introduction has shown the complexity of the image registration: it includes various images and the various degradations. This diversity makes it impossible to design a universal method working on all registration applications. There is a trade-off between the design for specific conditions and the design for general applications. To meet the different requirements, various registration algorithms have been designed. Currently the majority of the registration algorithms could be divided into the following three categories:

- Feature-based registration

  Control points (CP) are extracted from the image and used to estimate the parameters of the transformation by minimizing a measure of the residual matching error. Most current registration algorithms belongs to this kind of algorithm [45].

- Region-based or area-based registration

  This registration method uses a measure of similarity of two images. The reference image acts as template and the transformation parameters are adjusted to reach the maximum similarity.

- Statistic-based registration

  This registration is achieved by measuring common information between the reference and target images. It doesn't directly use the intensity values but is based instead on statistical information hidden in the probability distributions. This is a critical difference with the region-based registration algorithm.

All three types of registration have real applications in practice. In chapter 2 detailed summary about current image registration algorithms will be presented.

1.2 Objective

The object of this research is to construct an automatic (or semi-automatic) multi-modal registration algorithm for generic kinds of applications. The direct application of this research is for the WASP system, which is composed of three infra-red cameras and one high-resolution visual camera. It will be shown that the algorithm is able to address such challenging application as image matching to maps, GIS data and LIDAR derived images. These applications cover all four registration categories: multi-view (images from serial frames of one flight line); multi-temporal (images between two flight lines); multi-model (registration between the WASP bands); and image to model registration (airborne images with map or GIS data).

The external and internal distortion introduced by the geometric deformation, radiant blur and the variance among the airborne images from different sensors leads to challenging requirements for algorithm robustness. In these applications the registration system should not only be capable of handling the modal difference but also the noise and distortion, which may be hard to predict or estimate. In some extreme cases, manual processing is necessary, but the goal is to reduce human intervention as much as possible.

Due to the large volume of raw data, remote sensing applications require high efficiency as well as automatic or semi-automatic operation to minimize human workload. An important goal of this research is to improve the efficiency and robustness of registration algorithms to facilitate automated processing with the challenging kinds of images described above. The basic requirements for the proposed registration system are:

- The registration algorithm must be suitable for different registration applications, especially for the multi-modal image registration.
- The registration algorithm must be robust to noise, geometric distortion and minor lens radial distortion.
- The registration algorithm should be efficient enough to deal with large datum. Human intervention should be limited by enabling the automatic (or semi-automatic) processing.
- The algorithm structure must be extensible to fit various applications. The components of the algorithm should be flexible to meet the different requirements.

1.3 Organization

The focus of this thesis is providing an alternative solution that meets the above requirements. The proposed algorithm has the following features:

- The algorithm is based on MMI methods, which are able to register multi-modal images.

- The extensions to the MMI algorithm improve the robustness to noise, and reduce effects of geometric distortion and minor lens radial distortion.

- The deploying of pyramid coarse-to-fine methods and classification of the images with Harris corner labeling instead of pixel intensity reduces processing time by significant factor.

- The framework of the algorithm is modular. Each block of the algorithm is for different applications.

The details of the proposed algorithm will be described below. Chapter 2 introduces the background; Chapter 3 describes the MMI application to WASP remote sensing image registration and discusses the search algorithm for registration. Chapter 4 introduces the Harris corner detector and the corresponding feature enhanced MMI (FE-MMI) algorithm; Chapter 5 discusses the robustness and performance accuracy of the MMI and FE-MMI registration algorithms, and demonstrates registration between a label image and a real image; Chapter 6 analyzes the performance via a sample set; Chapter 7 covers the SIFT improved registration algorithm by localizing the best part for registration; Chapter 8 discusses experimental in real applications and analyzes the algorithm performance; Chapter 9 summarizes the thesis and describes the future work.

Chapter 2

Background

In this chapter we will review the current popular image registration algorithms. The interest in automatic registration is associated with the rapid development of image acquisition devices and the corresponding rapid growth in the size of image archives. According to the database of the Institute of Science Information (ISI), more than 1500 papers were published on image registration [8] and many registration methods were proposed for the diversity of images and variety of applications in computer vision, medical imaging and remote sensing. As mentioned in the first chapter, those algorithms generally can fall into three categories. The following section is the brief summary of registration algorithms.

2.1 The Classification of Registration Algorithms

2.1.1 Feature-based Algorithm

This registration algorithm is based on the extraction of salient features from the floating and reference images. Salient features could be regions [46, 47], lines [48–50] or corner points [51, 52]. They should be distinct, spread throughout the image and efficiently detectable in both images. The corresponding registration algorithm is based on the assumption that features are stable and fixed at positions in both images. The majority of the feature-based algorithm include the following four steps [8]:

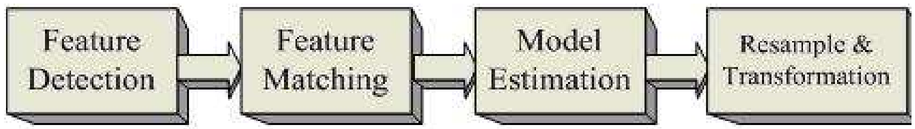


Figure 2.1: The Blockchart for the Feature-based Algorithm

- Feature detection: This step extracts the feature points by various filters or descriptors. The feature sets could be region boundaries [53], road or bank curves [54], building edges or intersection corners [13]. It has to be decided what kind of features are easy to detect, and what physical characteristic that both the sensed and reference images have in common. The feature sets in the floating and reference images should be comparable with respect to their invariance and position accuracy. A common problem is that the images do not exactly cover the same region because of view-point motion or object occlusion. The number of common elements of the feature sets must be sufficiently high to overcome changes in image geometry, lens radial distortion and additive noise, and the feature points should cover the whole image to overcome object occlusion or other unexpected changes. Good feature descriptors should be invariant to the assumed degradations and noise, and should be unique enough for the future processing [5, 55, 56].

- Feature matching: This step builds the relationship between the feature point sets from the reference and floating images. The establishment of the correspondence is based on the various feature descriptors and similarity measures [26]. The matching step is highly dependent on which invariant property is associated with the extracted feature points. One common error-causing factor is incorrect feature detection or image degradation [57]. Another is physical dissimilarity due to the different image conditions or different spectral sensitivities of the sensor. The matching step should be able to distinguish different features in the presence of image distortion and degradation as well as being stable so as not to be influenced by feature variations and additive noise [27, 58].

- Model estimation: This step decides the type and parameters that are needed for the mapping function. The parameters are computed from the feature pairs from the correspondences built in feature matching. An appropriate mapping model should be robust with respect to the expected distortion and noise [10, 59]. In addition, the accuracy of the feature detection, the reliability of the feature matching and the required approximation error have to be considered. We will introduce the registration model being used in the following section.

- Resample and Transformation: After the mapping function and corresponding parameters have been found, the floating image is transformed to the coordinate system of the reference image. Resampling and interpolation are necessary to get the values of non-integer coordinates. The choice of the appropriate resampling technique depends on the accuracy requirements. In many cases nearest-neighbor or bilinear interpolation are sufficient. On the other hand, the bicubic interpolation [60], quadratic interpolation [61] and spline interpolation [62,63] are designed for the more precise requirements. There is the trade-off between the accuracy and the computation complexity [64, 65]. That is a common step for all registration algorithms and is not emphasis of this thesis.

In summary, the feature-based methods are often used in applications where the images contain enough details that are distinctive and easy to detect. These methods have been widely applied in remote sensing (towns, rivers, forests and etc.), computer vision (objects, facilities and etc.) and cartography (roads, buildings and etc.). However, if the image is not rich in detail, or the detail is unsharp and hard to detect, one may resort to area-based algorithms [66]. The area-based registration algorithms also have wide application in medical imaging, which often lacks distinctive objects [67,68]. This is the leading reason for area-based methods, which can work well on single modalities even if the features are unreliable. The applicability of area-based and feature-based methods for images with different contrast and sharpness are discussed in [69]. On the other hand, both feature-based and area-based algorithm face big challenges for multi-model image or image-to-model registration due to the different sensors' characteristics (such as MRI, CT, X-ray in medical imagery or hyper-spectral image in remote sensing etc.), which is the motivation for the development of the statistic-based algorithm.

2.1.2 Area-based Algorithms

Area-based methods, sometimes called correlation-like methods or template matching, do not need the step of feature detection. Instead of extracting the salient points or areas, these methods merge the feature detection and matching processes and then estimate the parameters. The correspondences are calculated using a predefined-size window or even the whole image. The classical area-based methods, like cross-correlation or the Fourier transform, are based on matching of image intensities [70, 71].

- Correlation-like methods

  The commonly used area-based registration method is the normalized cross-correlation

(CC). The definition of CC is:

$$CC(i,j) = \frac{\sum_W (W - E(W))(I_{(i,j)} - E(I_{(i,j)}))}{\sqrt{\sum_W (W - E(W))^2}\sqrt{\sum_{(i,j)} (I_{(i,j)} - E(I_{(i,j)}))^2}} \tag{2.1}$$

where W is the reference image and I is the floating image. Although classical CC-based registration can exactly align translated images only, it can also work well with the slight rotation or scaling application.

To meet the requirement of more complicated geometrical deformation, one may extend the typical CC algorithm to the generalized CC version. One computes the CC for each assumed geometric transformation of the floating and reference image window [72] instead of the translation only transform. Berthilsson extended the CC algorithm to affinely deformed images in this manner [73] and Simper proposed use of a divide and conquer system with the CC technique for registering images that differ by perspective changes as well as changes due to the lens imperfections [74]. Because the CC algorithm uses image patches or even the whole image to calculate the optimization, it's more robust to the partial occlusion and additive noise. However, the computing-consumption grows rapidly with the predefined window size and the number of independent transform parameters.

Another simple CC-like algorithm is sequential similarity detection algorithm (SSDA) [75]. This method uses a sequential search approach to save computation. It accumulates the sum of absolute difference of the image intensity and applies a threshold criterion. If the difference is higher than the threshold the corresponding parameter will be rejected. In [76], the sum of squared difference similarity measure was used for iterative estimation of perspective deformation.

- Fourier Methods:

  Another kind of region-based registration algorithm is built on Fourier transform. This algorithm is prefered for applications that have a limitations on computing time. The algorithm has good performance on the images corrupted by frequency-dependent noise. The Fourier methods exploit the Fourier representation of the image in the frequency domain. The spatial shift is proportional to the phase change [71]. It computes the cross-power spectrum of the floating and reference images and searches for the peak value to find the shift. [24].

  $$\frac{F(f)F(g)^*}{|F(f)F(g)^*|} = e^{2\pi i(ux_0+vy_0)} \tag{2.2}$$

  where f and g represent the floating and reference images.

  FFT registration algorithm shows robust performance on the frequency dependent noise and non-uniformities caused by the illumination disturbances. Computation time is saved, especially for the large image.

  Phase correlation [77] is a type of FFT registration while a change of image scale is also present, the images can be registered using the combination of polar-log mapping of the spectral magnitude and phase correlation by Fourier-Mellin transformation [78, 79]. The applications of the extended algorithm in remote sensing and medical imaging are described in [78]. Wolberg [76] presented a method to register affinely distorted images by means of phase correlation and log-polar mapping.

  The area-based method uses a metric to measure the similarity of the images without extracting features explicitly and the above study demonstrated the effective and robust performance. However, there are three main drawbacks of the area-based methods: one is the flatness of the similarity measure maxima; another is the computation complexity, with Fourier-based methods showing a big improvement in computing time compared with the correlation-like methods [77] [79]; the third drawback is the difficulty in handling modality-dependent intensity variations [80]. The leading algorithm for multi-modal image registration is the following statistic-based algorithm [81].

2.1.3 Statistic-based algorithms

In contrast with typical region-based algorithms, such as the correlation-like algorithm or Fourier algorithms, statistic-based algorithms achieve registration based on the statistical information contained in the pixels between the floating image and reference image. In some literature the statistic-based algorithm is classified into the region-based algorithm. These algorithms randomly sample the images and extract the feature from the sample pixels, (the most direct feature is the intensity, but it's not only limited to the pixel intensity, we will discuss that at following chapter) then build feature pairs that are determined by a choice of registration parameters. The statistical information is then calculated over the samples for each choice of parameters [33]. The common statistic-based algorithm measure the similarity by calculating the mutual information between the two images. Greater mutual information usually means a better registration result. [82, 83].

The statistic-based algorithms avoid the step of feature point matching, which is the key problem with the use of feature-based algorithm for multi-modal images. In comparison with the other region-based algorithms, the registration is not directly relative to the pixel character. This improves the algorithms' robustness to the image variation from different sensors. The statistic-based algorithm has shown the powerful capability to handle multi-modal image registration in medical imaging processing [84, 85] and remote sensing [86–88].

The above section introduced the techniques that often are used for registration. The following section will briefly cover another critical factor for successful registration: the registration model. A good registration model will make the floating image overlap the reference image as closing as possible. It not only includes the mathematical rule to project the floating image to the reference image, but also the consideration of the transform complexity. For example, it can tell us how many feature pairs are needed to solve the linear function for the model's parameters. Usually the good estimation of the transformation model is based on prior knowledge.

2.2 Registration Model Estimation and Transformation

The registration model is the critical step to achieve registration. Generally the registration model estimation is composed of two steps: the first step is the transformation function choice: This step decides which registration model should be used and what parameters need to be estimated. The appropriate mapping function must correspond well with the assumed geometric deformation of the floating image, to the method of image acquisition and to the required accuracy of the registration; the second step is the estimation of the appropriate parameter values for the transform function.

There are a lot of transformation models and they have success in different applications. According to the region used for the registration, those algorithms can be divided into global mapping models and local mapping models. In global models, the mapping functions use Control Points (CPs for feature-based algorithm) [45, 89, 90] or even the whole images (area-based algorithm) [46,91,92] to estimate the mapping function parameter set and apply that mapping function into the entire image; on the other hand, the local model treats the image as jigsaw puzzle and finds an individual mapping function for each patch. The local model works well for the complicated image registration and is widely used in medical imaging [93]. However, the local mapping function technique needs pre-processing and tessellation of the image, usually by triangulation, ands need an extra step to mosaic and blend the borders of different patches [94–96].

2.2.1 Linear model mapping

The global model being most frequently used is the perspective model. The planar projective linear transformation in our real world could be defined as non-singular three by three matrices:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.3}$$

The transformation matrix has nine elements and eight degree of freedom. That eight degrees of freedom can represent the translation, rotation and perspective distortion in homogeneous coordinates. The task of the registration is to measure the similarity of invariant features and then extract those coefficients from the two images [97–99].

The major considerations for registration are efficiency and accuracy. Those are determined by the number of degrees of freedom (d.o.f) of the transformation matrix H in Equation 2.3. The transformation can be described by four levels related to the freedom factors:

- The most general transformation is projective, where the H matrix has 8 d.o.f. In this kind of transformation, the concurrency, collinearity, order of contact (intersection, tangency, inflection, etc.), and cross ratio are invariant.
- If the vanish point keeps on the infinite plane, the transformation reduced to affine, which has 6 d.o.f. Under this transformation, the parallelism, ratio of areas, ratio of

lengths on parallel lines (e.g midpoints), linear combinations of vectors (centroids) are invariant factors for registration. The transformation matrix can be represented as:

$$\begin{bmatrix} h_{11} & h_{12} & t_x \\ h_{21} & h_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Where h_{11}, h_{12}, h_{21} and h_{22} represent the skew of the image; t_x and t_y contain the shift information.

- Further, if the transformation can keep the ratios of lengths and angles invariant,the matrix is a 4 d.o.f similarity matrix. The transformation matrix is then:

$$\begin{bmatrix} s\cos\theta & s\sin\theta & t_x \\ -s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

  The skew in the affine matrix is reduced to a zoom parameter s and rotation angle θ. The four d.o.f are zoom, rotation, translation on X and translation on Y.

- At last, if the zoom is restricted, the transformation is Euclidean, the simplest transformation. The transformation matrix is 3 d.o.f.

The registration processing is to find the combination of those parameters that meet the acceptance threshold while maintaining efficiency. The more d.o.fs that are used, the more computation that is needed. For some applications and imaging conditions some parameters will be known or fixed. According to the properties of WASP airborne system, the similarity transformation is used in most application in this thesis.

2.2.2 Non-linear model mapping

Another estimation method for the projective transformation is based on inhomogeneous coordinates. By selecting a section of the image corresponding to a planar section of the world, the matching pairs between the image plane (x, y) and the world plane (x', y') are built and the transformation model are estimated as [100–102] :

$$x' = \frac{x'_1}{x'_3} = \frac{h_{11}x+h_{12}y+h_{13}}{h_{31}x+h_{32}y+h_{33}} \quad y' = \frac{x'_2}{x'_3} = \frac{h_{21}x+h_{22}y+h_{23}}{h_{31}x+h_{32}y+h_{33}} \tag{2.4}$$

Each point correspondence can generate two equations for the elements of H, then four point correspondences can build eight linear equations of H, which is enough to solve for H matrix.

$$\begin{aligned} x'(h_{11}x + h_{12}y + h_{13}) &= h_{31}x + h_{32}y + h_{33} \\ y'(h_{21}x + h_{22}y + h_{23}) &= h_{31}x + h_{32}y + h_{33} \end{aligned} \tag{2.5}$$

The projective deformation is introduced by a pinhole camera because the optical axis is not always perpendicular to the scene [8]. The only restriction for the points to build the correspondence is that no three points are collinear [21]. Another non-linear model is using the second or high-order polynomial function to solve the slight deformation in the image 2.6 [44, 103, 104]. Because the number of available matching pairs for registration are more than four, the optimization method can use the error minimization method to estimate the registration parameters.

$$\begin{aligned} x' &= a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 + a_5xy + \cdots \\ y' &= b_0 + b_1x + b_2y + b_3x^2 + b_4y^2 + b_5xy + \cdots \end{aligned} \tag{2.6}$$

2.3 Summary of Registration Algorithms

In summary, Table 2.1 provides a comparison for image registration algorithms.

Table 2.1: The comparison of Image Registration Algorithm

	Multi-modal image	Robust to noise and deformation	Efficiency	Local mapping support
Feature-based algorithm	Low	Low	High	Medium
Area-based Algorithm	Low	High	Low	Low
Fourier-based Algorithm	Medium	Medium	Low	Low
MI-based Algorithm	High	High	Low	Low
Local Mapping Model	Low	Medium	Low	High

The above discussion shows the capabilities of different registration algorithms. Multi-modal image registration is necessary for WASP system, which feature-based and area-based methods are not dependable for. MMI-based algorithms are the best choice for WASP remote sensed-image registration due to the good performance on the multi-modal image registration. The first step in this research is to apply the MMI techniques to the images from the three WASP IR sensors. These images exhibit extreme conditions and enable us to understand the limitation of the standard MMI registration methods.

A potential problem related to conventional MMI-based registration algorithm is the performance of robustness. Meanwhile, efficiency limits its use in real-time applications. A previous research combines the gradient feature to build high dimension vector to increase the capture range [105]. However, the extension of high-order mutual information deteriorate the efficiency performance.

This paper address the efficiency issue by combining a multi-resolution algorithm and various search algorithms to speed the processing. This approach has achieved a reduction in processing time of about 90% for WASP image registration.

In addition to efficiency improvement, this research improves the robustness performance of MMI-based algorithms. The new extension could work for map-to-image registration of map updating application, which the conventional MMI-based algorithm always failed.

This research combines MMI-algorithm with spatial features to gain higher robustness and reliability. The spatial feature detector is applied on map and airborne images and the area rich in structure information, such as roads and buildings, are utilized by MMI algorithm. The robustness and accuracy of MI-based algorithm are highly improved.

Another extension for MMI algorithm is the fusion of the LIDAR elevation or intensity data with imagery, which is an extremely difficult registration application. The robustness of registration algorithm is highly improved by the use of salient areas, which can be extracted by the SIFT algorithm. The use of salient region pairs extends the local mapping capability of the MMI algorithm, overcomes the negative effects of high distortion or high noise and improves the registration algorithm robustness. Experimental results are provided in the following chapters that demonstrate a variety of mapping applications including multi-modal IR imagery, map and image registration and image and LIDAR registration.

Chapter 3

Proposed Mutual Information Algorithm

This chapter discusses the application and improvement of MMI-based image registration algorithms for remote sensing image. Satellites, aircraft, and other remote sensing systems continuously capture a tremendous volume of data using a variety of sensors and modalities. These include low and high resolution still/video cameras in the visual spectrum, multi-spectral cameras using a variety of infra-red wavelengths, radar systems, and LIDAR. The sensors typically run at various capture rates, possess significant differences in resolution capabilities, and are not cross-registered spatially.

Data fusion from different sensors can provide better knowledge about the target and background but this also brings a big challenge for spatial registration of the images. The vast disparity in capture rates, resolution and spectral capabilities coupled with the lack of automatic and effective registration algorithms places the majority of the burden on human observers to analyze and exploit the data. Analysis by human observers is quite laborious, prone to errors and oversights, expensive, and time consuming, if not infeasible, given the continuously rising volume of data. Hence, the need for sophisticated systems capable of automatically registering multi-modal images, and exploiting their inherent information for a variety of applications (e.g. scene modeling, material analysis, etc.).

Conventional feature matching techniques often require manual intervention to improve the accuracy and reduce the computational complexity of the search space. Area-based methods [106] that utilize correlation [107] have been used to register single modality images, but are often ineffective in handling multi-modal images due to the scene-dependent deference in intensity in images from different sensors. They are also computationally expensive in cases where translations, rotations and scale variations are allowed.

Building the model between the different sensors has proven difficult because this need to build a good model for each sensor. A possible approach is to design a registration algorithm that is independent of pixel intensity. Viola introduced a new registration technique based on the Maximization of Mutual Information (MMI) between images [108]. This algorithm builds the registration on the statistical information associated with pixel intensity between sets of pixels [68] [37]. Its effectiveness and superiority to correlation based methods have been demonstrated on medical imagery [86], [109]. This advantage is already used to process remote sensing image [86, 88].

3.1 MMI Algorithm

The information contained in the image is a pixel-wise measurement that relies on the image's pixel intensity distribution. Let A be a discrete random variable with values a_i, $i = 1, 2, \ldots, m$ and probability $P_A(a_i)$. The entropy of A is defined as:

$$E[A] = -\sum_{i=1}^{m} P_A(a_i) \log_2 P_A(a_i) \tag{3.1}$$

Let B be a second random variable with values b_j, $j = 1, 2, \ldots, n$. Its entropy $H(B)$ is calculated in the same manner. The joint entropy $E(A, B)$ is:

$$E(A,B) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P_{AB}(a_i, b_j) \log_2 P_{AB}(a_i, b_j) \tag{3.2}$$

If A and B are statistically independent then $P_{AB}(a_i, b_j) = P_B(a_i)P_B(b_j)$ for all (i,j), and it follows immediately that:

$$E(A,B) = E(A) + E(B) \tag{3.3}$$

It is readily shown that maximum entropy is reached with statistically independent random variables. The mutual information is the difference between the maximum possible and actual joint entropy. That is:

$$I(A,B) = E(A) + E(B) - E(A,B) \tag{3.4}$$

The mutual information can be scaled by the summation value of $E(A)$ and $E(B)$:

$$R(A,B) = \frac{I(A,B)}{E(A) + E(B)} \tag{3.5}$$

In this application, the random variables A and B correspond to the brightness values of the two images that are to be registered. Let X =[x_1,x_2,...,x_n] represent a set of pixel locations in image A and let A(x_i) denote the brightness value at position x_i. The samples can be used to construct a histogram $h_A(A(X))$ and to estimate entropy E(A). These can use the same approach to estimate E(B). However, as will be seen later, these entropies are never actually needed in the matching algorithm.

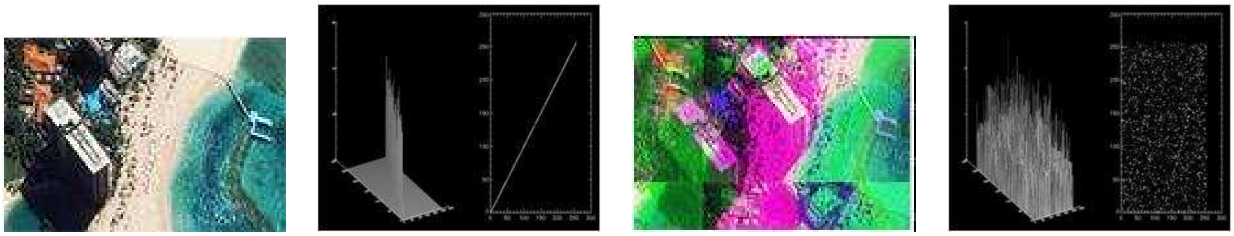
Let T be a spatial transform and let $Y = T(X)$ be a set of locations in image B.

Then $B(Y) = B(T(X))$ is a set of samples of B and $[A(X), B(T(X))]$ is a set of sample pairs from images A and B associated with the transform T. The goal is to find the transform that provides the best spatial match of the images. Conventional MMI searches for the transform that provides the maximum value of the mutual information between the samples sets. An estimate $\hat{\mathrm{I}}(\mathrm{A}, \mathrm{B})$ of the mutual information can be computed by finding the histogram $h_{AB}(A(X), B(T(X)))$, which, after normalization, is an approximation to P_{AB}. The transformation can be estimated by $\hat{T} = \arg\max_T \hat{I}(A(x), B(T(x)))$ or $\hat{T} = \arg\max_T \hat{R}(A(x), B(T(x)))$.

For a fixed set X of samples, $E(A(X))$ will be a constant, and will therefore not be of interest in searching for a maximum. Although the sample set will vary with the choice of T, it is reasonable to assume that the statistics of the image samples $B(Y)$ will not change in a systematic or significant manner during the search. If this turns out to be untrue then it is a simple matter to calculate $h_B(B(Y))$ and include the estimate of $E(Y)$ in the calculation. One can generally omit it in favor of reducing the number of calculations. If one assumes that $E(B(Y))$ is constant, then maximizing $I(A(X), B(T(X)))$ is equivalent to minimizing $E(A(X), B(T(X)))$ over the range of parameters in the transformation T.

The following is a simple illustration for the mutual information between the images in two bands. Figure 3.1(a) is a common RGB three-band image. The Figure 3.1(b) is the joint histogram. The joint histogram distribution is very sharp. However, shifting the green band ruins the RGB picture, shown in the Figure 3.1(c). The corresponding joint histogram is changed to Figure 3.1(d), the distribution is very random like noise. The entropy in Figure 3.1(b) is much lower than the entropy in Figure 3.1(d).

The above estimates are based entirely upon the histogram, which, in turn, is determined by the set X and the mapping T. Therefore, one is led to consider ways to choose a good sample set. It's important to consider which qualities of the sample set are im-



(a) (b) (c) (d)

Figure 3.1: The illustration of mutual information analysis between the images from different bands: (a) and (b) are original RGB image and the joint histogram between the red band and green band, the joint entropy $E(A, B)$ is 9.916; (c) and (d) represent the ruined RGB images and corresponding changed joint histogram, the joint entropy $E(A, B)$ is 14.898.

portant. Let T^* be the transformation that maximizes I and let $\hat{T}$ be the transformation that maximizes $\hat{I}$. Since the goal is to find T^*, one wants to choose the samples such that $\hat{T} = T^*$ and ensure that the value of $\hat{I}$ is sensitive to variations in T. We are not actually concerned with making an accurate estimate of the mutual information. The primary requirement is to have the information peak produced by the estimate have the same parameter-space location as the true peak. These considerations lead us to select locations for the set X that fall on or near edges and corners in image A. This will provide a histogram that is skewed away from the true probability distribution, but it will tend to increase the sensitivity to parameter variation. This topic will be discussed in the following chapter.

3.2 Efficiency Improvement

Remote sensing applications tend to produce large quantities of raw data and to put forward challenging requirements for the computation efficiency. The efficiency is an important factor that one has to consider when designing a registration algorithm. The conventional MMI algorithm is computation-intensive which is a limitation for use in remote sensing. Two steps in the algorithm require the most computation: the first is the search for the best parameter values for registration and the second is the calculation of the mutual information of data sets. To speed the parameter search, exhaustive and gradient search methods are investigated. The wavelet pyramid is built to achieve coarse-to-fine analysis, which can reduce both the computation for mutual information and parameter searching.

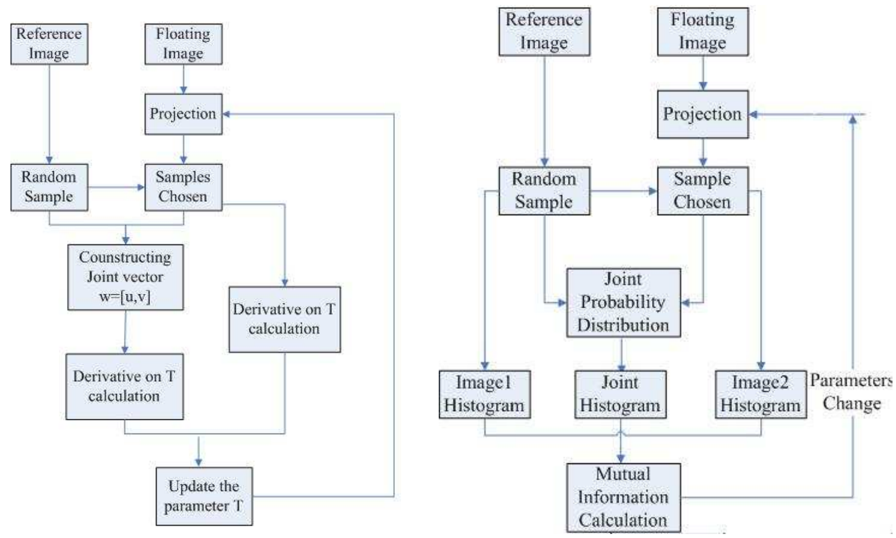
3.2.1 Search Method

In this section, exhaustive search and gradient search methods are tested. The efficiency improvement of the gradient search algorithm is discussed and analyzed for the WASP system. The Figure 3.2 is the flowchart for the two search methods. The samples are randomly chosen from the reference image, and the corresponding positions in floating image is decided by the transformation. Then the joint histogram is measured for each transformation parameters.

The search process can be structured as either exhaustive or gradient search. Exhaustive search tries all of the values of all parameters while gradient search tries to find a search path to the optimum set of parameter sets.

Exhaustive search

The most straightforward parameter search method is exhaustive search. The parameters construct the space and each parameter can be looked at as a dimension in the space. The exhaustive parameter search has to cover the entire space with resolution that depends on the desired resolution, or bin size. A small bin size provides more accurate resolution; however, it requires more computation time.



(a) (b)

Figure 3.2: (a) The flowchart for the exhaustive search; (b) the flowchart for the gradient search.

Gradient search

To reduce the computation required in the exhaustive search, the alternative method of gradient search is used. Estimating the probability density function by Parzen window, the estimation for derivative of the mutual information can be expressed as [108]:

$$
\begin{aligned}
&\frac{\widetilde{dI}}{dT} = \frac{1}{N_B} \sum_{x_i \in B} \sum_{x_j \in A} (v_i - v_j)^T [W_u(v_i, v_j)\psi_v^{-1} - W_{uv}(w_i, w_j)\psi_{vv}^{-1}] \frac{d}{dT}(v_i - v_j) \\
&W_v(u_i, v_j) = \frac{g_{\psi_v}(v_i - v_j)}{\sum\limits_{x_k \in A} g_{\psi_v}(v_i - v_k)} \quad W_{uv}(w_i, w_j) = \frac{g_{\psi_{uv}}(w_i - w_j)}{\sum\limits_{x_k \in A} g_{\psi_{uv}}(w_i - w_k)}
\end{aligned}
\tag{3.6}
$$

u is the reference image while v is the floating image that needs to be registered, w is joint intensity $[u, v]$. Given different transformation models, the derivative for each model is different. In the following example, the use of gradient search with translation, rotation and scaling is illustrated, which together comprise the elements of affine transformations.

The new parameters vector on the gradient-search path is required by gradient feedback $T' = T - \lambda \frac{dI}{dT}$, which is iterated till the maximal I is located. λ is the search step size.

- Translation

  The corresponding transformation matrix is defined as: $T = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix}$ So the derivative is:

$$\begin{aligned} &\tfrac{d}{dT} v(x', y') = \tfrac{d}{dT} v(x + t_x, y + t_y) \\ &= \left[\ \tfrac{d}{dt_x} v(x + t_x, y + t_y) \quad \tfrac{d}{dt_y} v(x + t_x, y + t_y) \ \right] \\ &= \left[\ \tfrac{dv}{dx'} \times \tfrac{dx'}{dt_x} \quad \tfrac{dv}{dy'} \times \tfrac{dy'}{dt_y} \ \right] = \left[\ \tfrac{dv}{dx'} \quad \tfrac{dv}{dy'} \ \right] \end{aligned} \tag{3.7}$$

- Rotation

  The corresponding transformation matrix is defined as: $T = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

  So the derivative can be expressed as:

$$\begin{aligned} &\tfrac{d}{dT} v(x', y') = \tfrac{d}{d\theta} v(\cos\theta \times x - \sin\theta \times y, \sin\theta \times x + \cos\theta \times y) \\ &= \tfrac{dv}{dx'} \times \tfrac{dx'}{d\theta} + \tfrac{dv}{dy'} \times \tfrac{dy'}{d\theta} \\ &= -\tfrac{dv}{dx'} (\sin\theta \times x + \cos\theta \times y) + \tfrac{dv}{dy'} (\cos\theta \times x - \sin\theta \times y) \end{aligned} \tag{3.8}$$

- Scale, rotation & translation

  The scale matrix is : $T = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix}$, the derivative on the scale s is:

$$\begin{aligned}&\tfrac{d}{dT}v(x',y') = \tfrac{d}{ds}v(s\times x, s\times y) = \\ &\tfrac{dv}{dx'}\times\tfrac{dx'}{ds} + \tfrac{dv}{dy'}\times\tfrac{dy'}{ds} = -\tfrac{dv}{dx'}\times s + \tfrac{dv}{dy'}\times s\end{aligned} \tag{3.9}$$

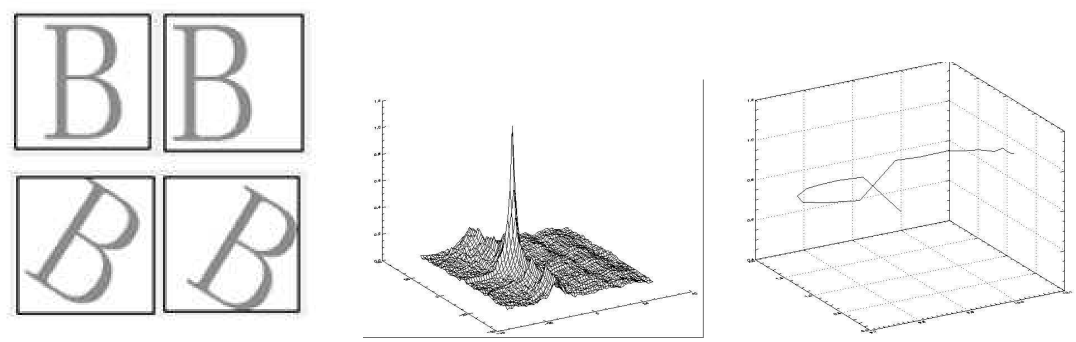
The above steps show the derivative for translation, rotation and scale parameters separately. But in real application the conditions can be combined:

$$d = d_S d_R d_T = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \tag{3.10}$$

In this complicated case the derivative for the transformation parameter is:

$$\frac{dv}{dT} = \begin{bmatrix} \frac{dv}{ds} \\ \frac{dv}{d\theta} \\ \frac{dv}{dt_x} \\ \frac{dv}{dt_y} \end{bmatrix} = \begin{bmatrix} \frac{dv}{dx'}(\cos\theta\times x - \sin\theta\times y) + \frac{dv}{dy'}(\sin\theta\times x + \cos\theta\times y) \\ -\frac{dv}{dx'}(\sin\theta\times x + \cos\theta\times y)s + \frac{dv}{dy'}(\cos\theta\times x - \sin\theta\times y)s \\ \frac{dv}{dx'} \\ \frac{dv}{dy'} \end{bmatrix} \tag{3.11}$$

The following is a simple example for the comparison of exhaustive and gradient search. The sample images are shown in Figure 3.3(a). Figure 3.3(b) displays the parameter search surface created by the exhaustive search. There are only translation between the model image (top left image) and the floating image(top right image), then the position of the peak indicates the optimum X and Y coordinates for registration. The Figure 3.3(c) shows the gradient search track to the best parameter set of translation and rotation (bottom right image).



(a) (b) (c)

Figure 3.3: a) Sample Example (128 × 128) b)Exhaustive Searching Surface c) Gradient searching curve

The following table shows the comparison of exhaustive searching and gradient searching:

Table 3.1: The simple comparison of time and computation consumption

	Translation		Translation & Rotation	
	Points searched	Time Consumed	Points searched	Time Consumed (s)
Gradient Search	18	4.84s	90	18.48s
Exhaustive Search	3721	6.72s	26047	48.094s

From the experiment results (shown in Table 3.1), it can be seen that the gradient search and exhaustive search both can reach the correct registration result. Because of smaller number of points required for search, the gradient algorithm needs fewer steps than exhaustive search. However, the gradient search need extra computation on calculating the gradient to get the new search point.

3.2.2 Wavelet pyramid

Pyramid processing is a type of multi-scale signal processing developed for the computer vision, image processing and signal processing, in which a signal or an image is subject to repeated smoothing and subsampling [110] [111]. The coarse-to-fine procedure divides the parameter search into two parts: the first part is the search in the higher layer to reach the coarse parameter estimate; The second is the search around the coarse parameter estimate in the following lower layer to reach finer results. The corresponding computation is only relative to the search windows size and pyramid height.

With the coarse-to-fine procedure, the image size is reduced by the pyramid height. The required computation can be adjusted by system capability instead of the input data. The bigger the original image size is, the higher pyramid could be built therefore the more computation could be saved.

Another efficiency improvement is from reduction of the sample points for the mutual information calculation. The biggest part of the mutual information calculation is done between the small images on the high layer of the pyramid, where the dimensions are much smaller than the original size. Fewer sample points are needed to compute the mutual information. If the sample ratio is fixed, then the sample points in the 3rd layer is only 1/16 and in the 2nd layer only 1/4 compared to the registration on original image. Therefore the computation and corresponding hardware burden are reduced.

The block diagram of the proposed wavelet-based MMI image registration algorithm is shown in the following figure:

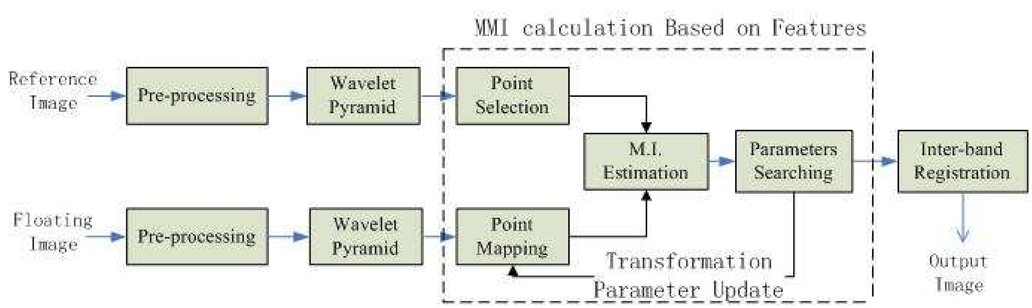


Figure 3.4: Block Diagram of the wavelet based MMI registration algorithm

3.3 Experiments & Results

In this section the performance of MMI algorithm with multi-resolution will be demonstrated. The first part will test the MMI algorithm performance on multi-modal image registration, then the efficiency will be compared in the following part.

3.3.1 MMI Algorithm Registration on WASP Image

The images utilized are captured by the WASP system which employs three IR cameras and a high-resolution visible camera. The specifications of the system are shown in the table 3.2. (*http://lias.cis.rit.edu/projects/wasp*)

Table 3.2: WASP camera description

	SWIR	MWIR	LWIR	VNIR
Bandwidth	0.9-1.7 μm	3-5 μm	8-9.2 μm	0.4-0.9 μm RGB
Resolution	640×510	640×510	640×510	2048×2048

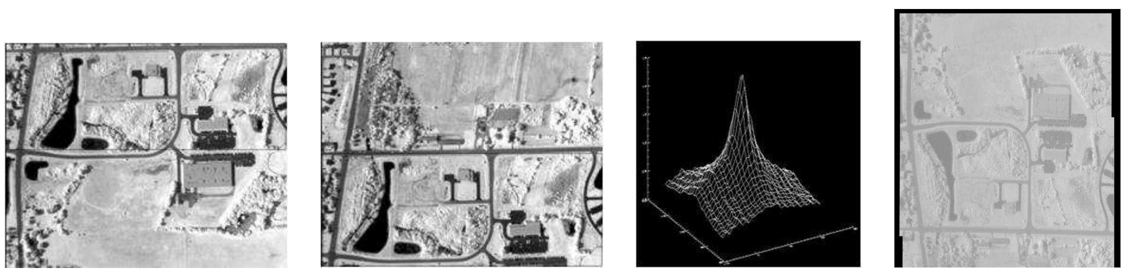
Figure 3.5: WASP's infrared cameras cover three different ranges of the infrared spectrum

The images from the different IR bands provide spectral information that can be used for fire detection as well as for other kinds of environmental data gathering. The images are difficult to register spatially because of the differences in intensity variations. The short-wave band depends mainly on reflected light while the long-wave band depends upon the thermal values across the scene, and the medium-wave band is a combination.

The different pixel intensity variations with temperature emissivity and reflectivity provide a challenge for registration by conventional feature-based and correlation algorithms.

Intra-band Image Registration

This section tests the MMI-based algorithm's performance on the registration of images from same band. The images shown in Figure 3.6(a) and Figure 3.6(b) represent the short-wave IR band of serial frames taken 4 seconds apart from the WASP system. Figure 3.6(c) shows the mutual information distribution along the parameter space and the registration result is shown in the Figure 3.6(d).



(a) (b) (c) (d)

Figure 3.6: The registration for the short-wave band WASP image: (a) and (b) are serial short wave band image; (c) is the search surface: X and Y coordinates are the shift on X and Y directions and the mutual information is shown the value on Z; (d) represent the registration result.

Section 3.1 describe how the mutual information can be measured by a joint histogram. The left side of Figure 3.7 shows the 2D joint probability and the corresponding compactness is displayed on the right. Each row represents one parameter set. Observe that parameter set #2 (the optimum set) demonstrates that the 2D joint probability is the most compact and has the highest mutual information.

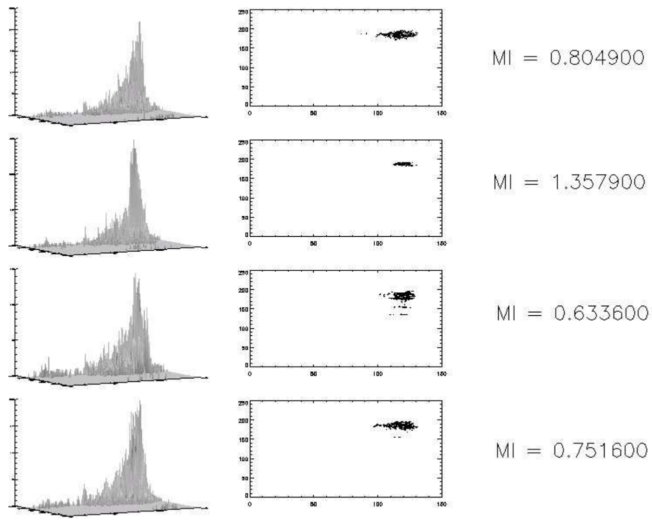


Figure 3.7: The 2D Joint Probability for Different Parameter Sets. Each row represents the 2D joint histogram (left) and contour projection (medium) for a parameter sets. The right column is the mutual information for each parameter. The smaller the joint distribution area, the more joint probability is shown, which is the higher and compacter distribution.

Inter-band Image Registration

This section will demonstrate the effectiveness of the proposed algorithm in registering multi-modal imagery acquired by the WASP system using the three IR bands in Table 3.2 above. The images from different IR bands provide spectral information that can be used for fire detection as well as for other kinds of environmental data gathering. The images are difficult to register spatially because of the differences in intensity variations. The short-wave band depends mainly on reflected light while the long-wave band depends upon the thermal values across the scene, and the medium-wave band is a combination. The different pixel intensity variations with temperature and reflectivity provide a challenge for registration by conventional feature-based and correlation algorithms. The short-wave, medium-wave and long-wave IR images are shown in Figure 3.8(a), (b) and (c) respectively.

Due to the sensors' resolution difference, the scale is an important factor for successful registration. The addition of the scale parameter increases the dimensionality of the search space, which increases the computation cost.

To find the best scale value, the original images are interpolated for each scale. The peak value of the mutual information on translation and rotation are searched, then the one with highest mutual information are founded to decide the best parameter set.

The results computed by our algorithm are shown in Figure 3.8(d) and displayed in Table 3.3, where the short-wave band was utilized as the reference one. The table shows the scale for medium-wave image is 1.13 while 1.12 scale for long-wave band image (short-wave band image is reference image and the size is fixed). Interpolation processing is required to build the 'bigger' medium-wave and long-wave band image. Then translation parameters are searched for the scaled image.

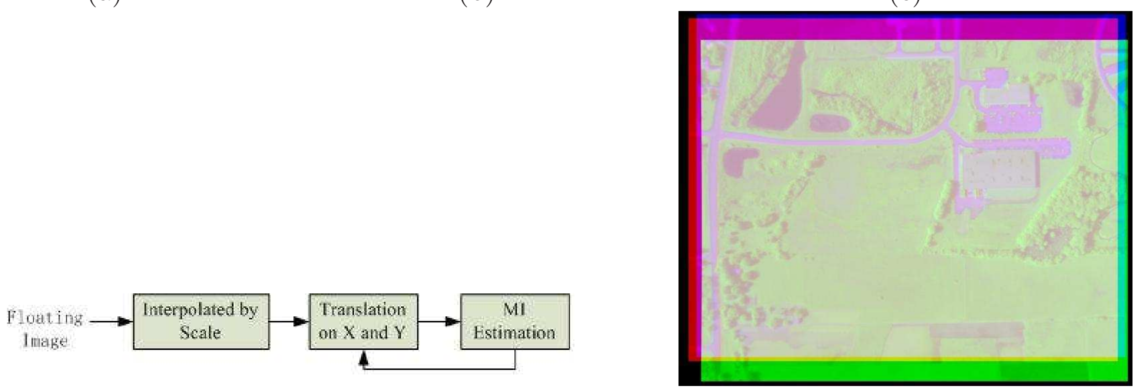
Table 3.3: Inter-band Registration Results

	Scale	Shift on X	Shift on Y
Medium-wave band	1.13	19.0	32.0
Long-wave band	1.12	6.0	36.0

The visible band (see Table 3.2) is 2048×2048 as compared to the IR bands which are 640×512 in resolution. This makes registering the visible band to the IR bands much more difficult. Figure 3.9(b) shows the corresponding visible band for the IR images shown in Figure 3.8. The results of our registration algorithm for the visible and short wave band are shown in Figure 3.9(c). With the visible high resolution image, one can reach the super-resolution for IR band, which is very valuable for the information extraction and analysis.

The above experiments shows the validity of MMI-based algorithm on intra-band image and inter-band image registration. The figure 3.6 (d), 3.8 (d) and 3.9 (c) show the overlapping images. The registration results match the requirement in accuracy. Next

(a) (b) (c)



(d) (e)

Figure 3.8: Inter-band Registration Example: (a), (b) and (c) are short-wave band image, medium-wave band image and long-wave band image from left to right successively; (d) is the simple flowchart for parameter search; (e) is the semi-color registration result.

(a) (b) (c)

Figure 3.9: IR and Visible Band Registration

section will study how to speed the processing and improve the efficiency.

3.3.2 Efficiency Improvement

The efficiency improvement of our algorithm will be demonstrated in this section. First the coarse-to-fine algorithm is analyzed and then the exhaustive search and gradient search are compared.

Pyramid Algorithm

In this section the performance of coarse-to-fine search will be tested on WASP single band images. The images in Figure 3.10(a) and Figure 3.10(b) which represent the short-wave IR band images. Figure 3.10(c) shows the first 2 levels of the wavelet decomposition, while Figure 3.10(d) displays the parameter search from “coarse” to “fine”. The registered images are overlaid (See Figure 3.10(e)) to indicate the degree of accuracy obtained by the proposed approach. Noted in this example, only translation is considered as the registration parameter.

This example shows the pyramid and compare the computation for search w/o coarse-to-fine processing. The images of size M x N are decomposed by using a wavelet-based pyramid [112]. The number of layers in the pyramid is defined as L. The registration is then performed on the ’coarse’ set of images and refined as one navigates the wavelet pyramid from the ’coarse’ to the ’fine’ layer. The initial search window is selected as M/4 by N/4. The parameters computed at each layer are passed onto the higher resolution layer thereby providing an accurate initialization for the registration process at the new layer. The search window size is set to decide the search area in new layer.

In this experiment, the window size was selected as K1 = K2 = 16. $S1$ represents the number of search iterations directly at the highest resolution layer while $S2$ encompasses

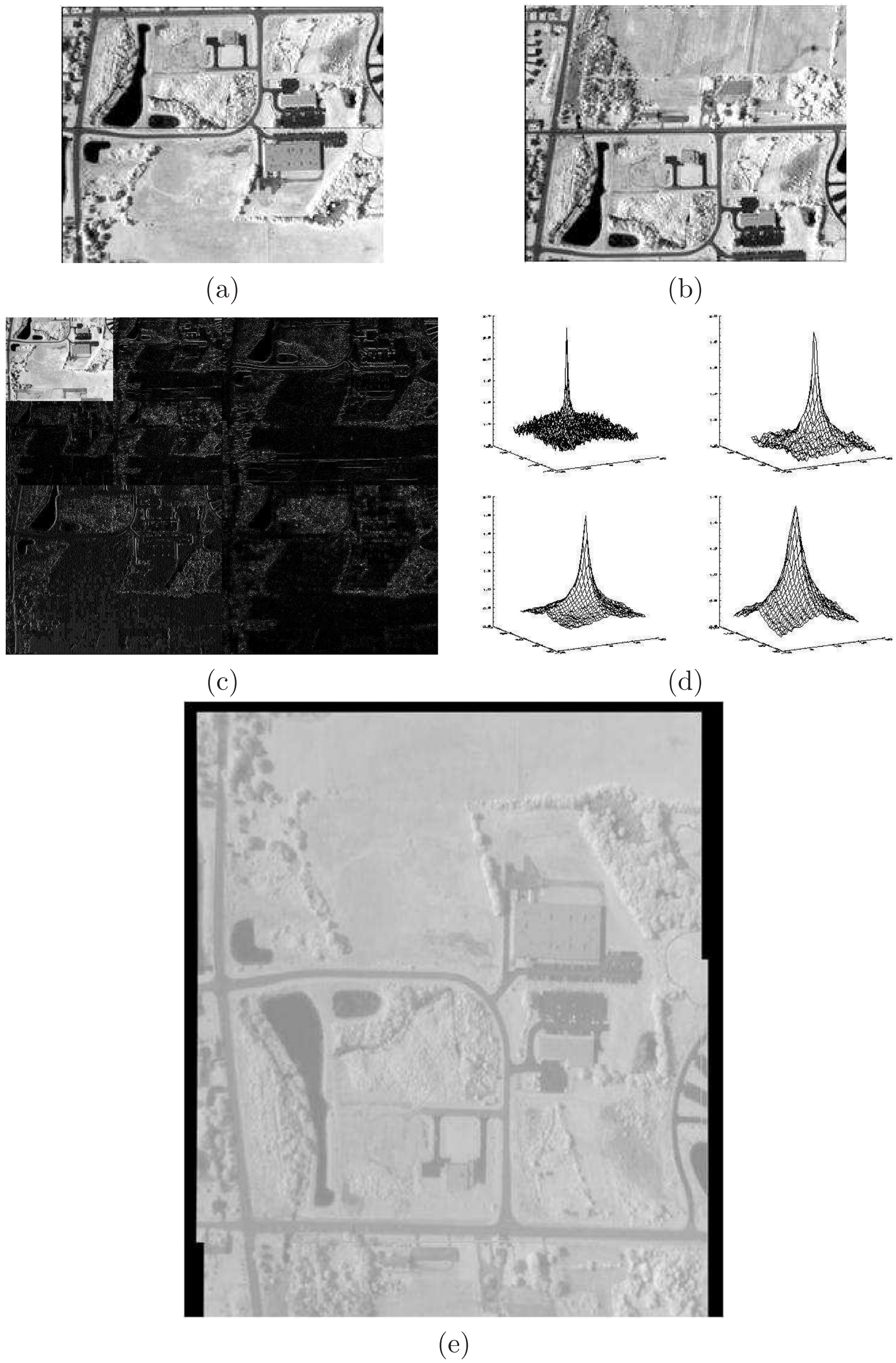


Figure 3.10: Coarse to fine Registration Example: (a) and (b) represent two sequential frames, (c) wavelet decomposition of frame shown in Fig 2a, (d) MMI parameter search spaces, the coarse-to-fine search space are shown from the top left image to the bottom right image, (e) Overlaid registered images.

the number of iterations through the wavelet pyramid as described above. Hence, the computational reduction for an image of size $M = 640$ by $N = 510$ (the size of WASP IR images), with a pyramid decomposition of $L = 3$ is:

$$\eta = \frac{S_2}{S_1} = \frac{1787}{20400} = 8.75\%$$

where

$$S_1 = \frac{M}{4} \cdot \frac{N}{4} = \frac{640 \times 510}{16} = 20400$$

and

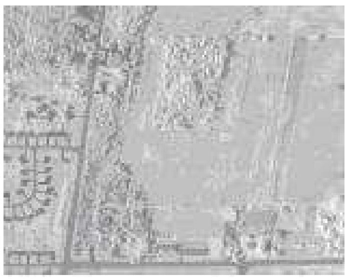
$$S_2 = \frac{M}{4 \times 2^{(L-1)}} \cdot \frac{N}{4 \times 2^{(L-1)}} + (L-1) \cdot (K_1 \cdot K_2) = 1787$$

In this sample, about 92% the search iterations are saved.

Exhaustive Search vs. Gradient Search

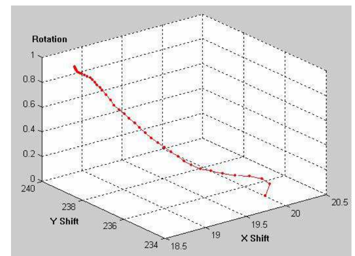
This section compares the gradient search and exhaustive search performance in the WASP images. This sample is more complicated because the transformation parameter set includes both translation and rotation. The experiment is illustrated in the Figure 3.11.

Table 3.4 shows the registration result with exhaustive search and gradient search. Rotation increases the search complexity which in turn increases number of operations especially for exhaustive search. The Figure 3.11(d) shows the search surface along the rotation angle in exhaustive search: each search surface is for a special rotation angle and the X, Y coordinates represent the shift on X and Y directions; the step size of the angle search is 0.5. The search surface shows that registration has the best result under the rotation angle $\Theta = 1.0$. The Figure 3.11(c) shows the gradient search points. The Figure 3.11(c) shows the conversion step of gradient search. In this example the gradient search saves about 40% time relative to exhaustive search.

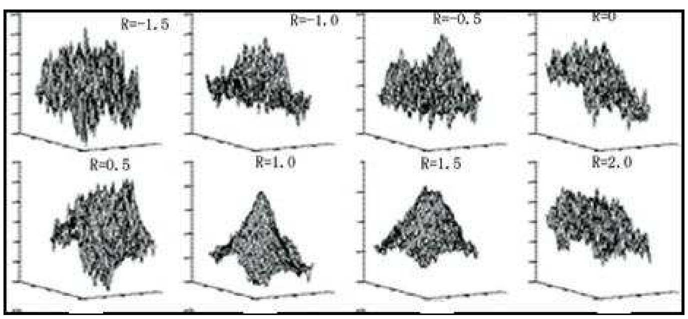


(a)

(b)



(c)



(d)

(e)

Figure 3.11: WASP Image Registration Example: (a) and (b) represent two sequential frames, (c) Gradient search curve, (d) Exhaustive searching space on different rotation, (e) Overlaid registered images

Table 3.4: Experiment Result for Gradient & Exhaustive Search

	X Axis (pix)	Y Axis (pix)	Rotation (degree)	Time Consumed
Gradient Search	18.84	239.7	0.88	72.422s
Exhaustive Search	20	235	1.0	115.45s

This examples shows the efficiency improvement of gradient search on exhaustive search. However, the shortcoming is robustness. The registration of multi-modal images is a difficult task. Usually the 'bumpy' search surfaces make the gradient search easy to be trapped into local minimum, which will cause the registration fail. Taking those factors into account, this thesis uses the exhaustive search in most cases.

3.4 Summary

This chapter described a wavelet based MMI algorithm for effectively and efficiently registering multi-modal images. The algorithm is capable of handling a wide variety of translations and practical scale variations by searching the parameter space in order to maximize mutual information. Two different search algorithms are discussed. The results indicate that both search algorithms are feasible for registering IR and visual images and robust at multiple resolutions. Compared to exhaustive search, the gradient search method is more efficient. However, the gradient search relies more on the starting point, appropriate step interval and the image quality. The use of wavelet pyramid provides a significant savings in computational efficiency while maintaining equivalent accuracy. The results indicate that the MMI method is feasible for registering IR and visual images and robust at multiple resolutions.

Chapter 4

Proposed Feature Enhanced MMI Algorithm

Although MMI is a powerful method for multi-modal image registration, there are still some problems for infrared image registration:

- The robustness of the algorithm is questionable. Experiments show it often fails in the registration of infrared images due to the a resulting flat search surface (Figure 4.1 (c)).

(a) (b) (c)

Figure 4.1: An example of flat search surface

- Usually remote sensing applications produce a huge volume of raw data from airborne or satellite images. The corresponding efficiency requirement is difficult to meet with MMI due to it's computational intensity.

The above two problems become obstacles to applying the conventional MMI algorithm to remote sensing registration. The technique described in this chapter encompasses a set of related algorithms for multi-modal image registration to solve those problems. It includes two basic steps: the first responds to spatial features and the second step uses them to efficiently compute the MMI registration. The enhanced algorithm flowchart is shown in Figure 4.2.

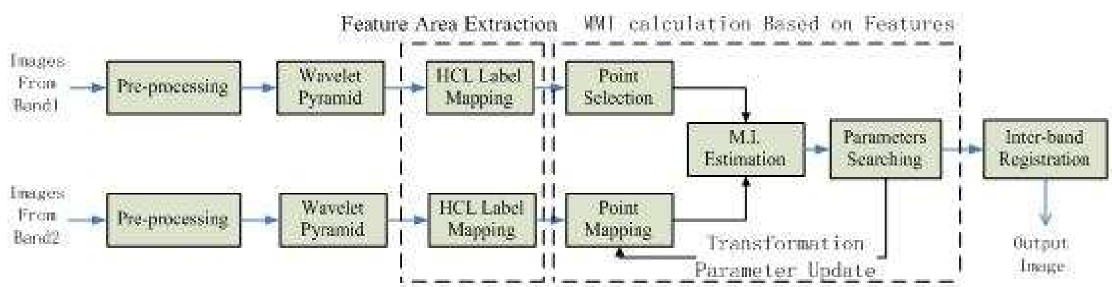


Figure 4.2: Flowchart for the HCL-Based MI registration

Extended from the flowchart in 3.4 on the page 31, the new algorithm adds a feature-extracting module. The image pixels are classified and labeled in this step. The following processing is based on the label instead of original pixel intensity.

In this implementation, the Harris corner detector was used to extract the feature information because it is fast and translation and rotation invariant. However, it can be replaced by other feature detectors, if that is a designer's preference. The wavelet pyramid algorithm is integrated to achieve multi-resolution for computation efficiency. To further improve the algorithm efficiency, the Harris Corner Label (HCL) map can be used for MMI calculation in place of the original images. This will be described in detail later.

The robustness of the MMI algorithm for multi-modal image registration is based on the statistical relationships between corresponding sets of pixels. Some spatial features, such as edges and corners, are rich in information and good candidates for point sets for statistical analysis. The key point of the proposed algorithm is to utilize feature detection as a focus-of-attention mechanism to improve the robustness and efficiency of MMI.

The Harris filter conveniently provides a label map with three levels: [0,1,2] which are associated with background (no corner or edge), edge, and corner, respectively, at each pixel location. Similar classification can be done with a variety of detectors and it is likely that other detectors would give similar results. The three-level label map can be used in place of an image for the purpose of matching. With suitable Harris threshold settings, it tends to provide a less noisy parameter search surface than is available with the original images.

4.1 Harris Corner Detector

The Harris corner detector is a popular interest-point detector due to its strong invariance to translation, rotation and illumination variation with relative tolerance of image noise [113]. It utilizes the average gradient c(x,y) computed over a small region w as follows:

$$c(x, y) = \sum_{w} [S(x_i, y_i) - S(x_i + \Delta x, y_i + \Delta y)]^2 \tag{4.1}$$

where (x, y) is the location and $S(x, y)$ represents the intensity of the corresponding pixel. This can be expressed in matrix form as:

$$c(x, y) \approx \mathbf{d}^T \mathbf{M}(x, y) \mathbf{d} \tag{4.2}$$

where $\mathbf{d} = [\Delta x, \Delta y]^T$.

$$\mathbf{M} = \begin{bmatrix} \sum_w (S_x(x_i, y_i))^2 & \sum_w S_x(x_i, y_i) I_y(x_i, y_i) \\ \sum_w S_x(x_i, y_i) I_y(x_i, y_i) & \sum_w (S_y(x_i, y_i))^2 \end{bmatrix}$$

$S_x(x_i, y_i) = S(x_i, y_i) - S(x_i + \Delta x, y_i)$ and $S_y(x_i, y_i) = S(x_i, y_i) - S(x_i, y_i + \Delta y)$ are gradient measure. Based on the relationship between the eigenvalues of the matrix M(x,y), the pixels is classified into the three categories [0,1,2]: (Shown in Figure 4.3)

- If λ_1 and λ_2 are both are small, then the local area is relatively smooth.
- If one eigenvalue is large and the other is small it means the local auto-correlation function is ridge shaped and the local shift only happens in one direction which indicates an edge.
- If both eigenvalues are large, so the local auto-correlation function is sharply peaked, this indicates a corner.

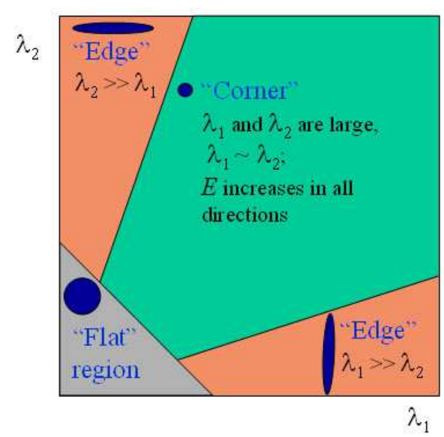
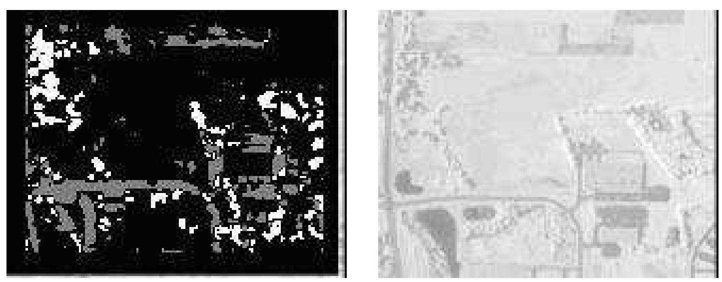


Figure 4.3: Illustration for Harris Corner Label Categories

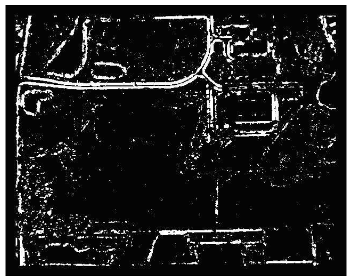
The feature-enhanced MMI (FE-MMI) algorithm uses the label information to identify pixels located on edges and corners and gives them priority in choosing the sample locations. This has the effect of incorporating spatial information into the selection of data for the MMI calculations while avoiding the difficulty of maintaining an explicit description of spatial data. The labels can either be used to guide the sampling of the image data or sampling of a filtered image. It was found that the label image itself can be used as a surrogate for the original image in the registration process. This reduces the range of histogram values and the sensitivity to image noise.



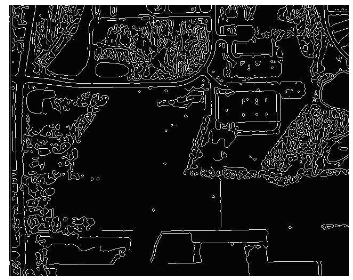
(a) (b)

Figure 4.4: The WASP Airborne Image and the Corresponding Harris Corner Label Mapping

The Harris corner detector is a good choice as a focus-of-attention device for this approach to MMI registration. Other focus-of-attention methods were tried such as typical edge detection, but it was found that the HCL map provides more robust performance with MMI because the Harris detector provides a richer set of pixels that can be used in the MMI computation. This can be seen in Figure 4.5 which compares a HCL map with a conventional Canny edge map. The edge map is so sparse that the MMI algorithm has problems with pull-in range (Figure 4.5 (b)). Furthermore, the edge position itself is somewhat sensitive to relocation between images and maps. The Harris detector can be set to provide a broader edge and this reduces that problem. The Harris corner detector was found to be less sensitive to feature relocation across the image and maps. Another important advantage is the threshold can be adjusted for different applications.



(a)



(b)

Figure 4.5: Comparison between the HCL Mapping and Edge Mapping

4.2 Harris corner label directed MMI

This section will discuss the relationships between the traditional MMI method and the FE-MMI algorithm and illustrate the reason that FE-MMI algorithm improves robustness. Suppose that S^r and S^f are the intensity domains for the reference and floating images respectively, and the H^r and H^f are their corresponding Harris Corner Label (HCL) map images. The use of spatial information can improve the conventional MMI convergence range and accuracy [105]. The high dimension mutual information can be constructed by integrating the Harris corner label:

$$MI(S^f, H^f, S^r, H^r) = \sum_{S^f, H^f, S^r, H^r} p(S^f, H^f, S^r, H^r) \log_2 \frac{p(S^f, H^f, S^r, H^r)}{p(S^f, H^f) p(S^r, H^r)} \tag{4.3}$$

The probability $p(S^r, H^r, S^f, H^f)$ over the sampling set can be approximated by the 4D joint histogram $h(S^r, H^r, S^f, H^f)$, and the probability distribution $p(S^r, H^r, S^f, H^f)$

can be approximated as:

$$p(S^f, H^f, S^r, H^r) = \frac{h(S^f, H^f, S^r, H^r)}{\sum\limits_{S^f, H^f, S^r, H^r} h(S^f, H^f, S^r, H^r)} \tag{4.4}$$

the similar approximation can be achieved on $p(S^f, H^f)$ and $p(S^r, H^r)$:

$$p(S^f, H^f) = \frac{h(S^f,H^f)}{\sum\limits_{I^f,H^f} h(S^f,H^f)} \quad p(S^r, H^r) = \frac{h(S^r,H^r)}{\sum\limits_{I^r,H^r} h(S^r,H^r)}$$

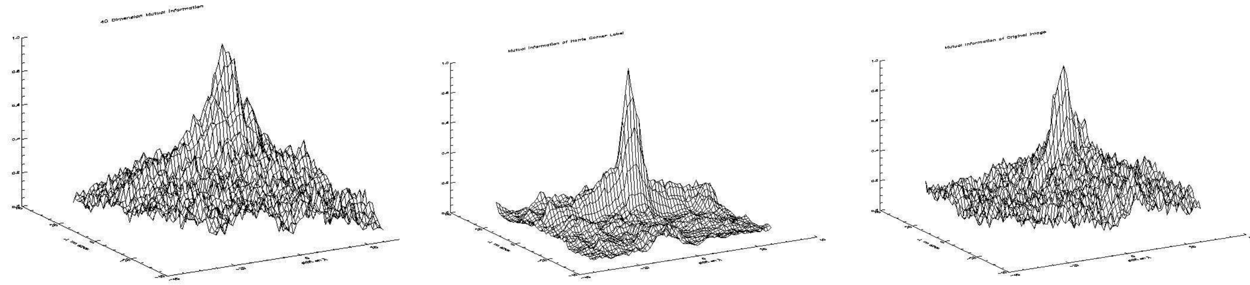
The above equation provides good estimation for the 4D mutual information. However, one disadvantage of the 4D mutual information is huge computation requirements. To speed the calculation, we express the 4D mutual information as:

$$\begin{aligned} &MI(S^f, H^f, S^r, H^r) = \sum_{S^f, H^f, S^r, H^r} p(S^f, H^f, S^r, H^r) \log_2 \frac{p(S^f,H^f,S^r,H^r)}{p(S^f,H^f)p(S^r,H^r)} \\ &= E(H^f) + E(S^f|H^f) + E(H^r) + E(S^r|H^r) - \{E(H^f, H^r) + E(S^f, S^r|H^f, H^r)\} \\ &= MI(H^f, H^r) + O(S^f, H^f, S^r, H^r) \quad \text{where} \\ &O(S^f, H^f, S^r, H^r) = \quad E(S^f|H^f) + E(S^r|H^r) - E(S^f, S^r|H^f, H^r) \end{aligned} \tag{4.5}$$

The derivation shows the 4D mutual information is composed of two items: the mutual information between the Harris corner label mappings and the remaining terms. The experiments show the MI of the Harris corner label is dominant (Figure 4.6), and that the HCL maps provide effective image registration with MMI. The sampling is done using direction from the HCL labels.

The use of the HCL maps for registration reduces the range of required histogram bins that are required, which can speed the MI calculation significantly.

To compare the sensitivity of different scaled mutual information, one method is to



(a) (b) (c)

Figure 4.6: The Illustration for Equation 13, x and y axis are the shift parameters: (a) Normalized 4D high dimensional mutual information; (b) Normalized Mutual information between the Harris Corner Label Maps; (c) Normalized Conditional Entropy of pixel intensity.

use the kurtosis to measure the sharpness of each scaled search surface. The sharper the search surface is, the higher the kurtosis [114], which represents more confidence in the registration result. Table 4.1 shows the mutual information kurtosis of Harris corner label

Table 4.1: The comparison of Mutual Information

	$MI(S^f, H^f, S^r, H^r)$	$MI(S^f, S^r)$	$MI(H^f, H^r)$
Kurtosis	7.756	6.646	15.33

is about double of that of the original pixel intensity, that means the mutual information between the Harris corner label maps is more effective for registration, which can be shown as the sharper peak in the search surface (Figure 4.6(b)).

Edge descriptors are the important composition for human visual registration and the FE-MMI algorithm simulates that procedure. Note that to increase the capture range, the Harris corner detector can give us both edge and corner candidate pixels instead of only the corner pixels. For conventional feature-based algorithm, increasing candidate points leads to more search time for matching pairs and reduces the efficiency drastically. But, for the FE-MMI algorithm, it can improve registration's smoothness and makes the algorithm more robust with very small efficiency reduction.

4.3 Scene and Map Registration

The purpose for the FE-MMI algorithm is the registration between regular scene images, including different wavelengths. The effectiveness of registration with HCL maps leads to the idea of a way to register images and graphic maps, which is a necessary step in image-to-map conflation and very valuable for updating and correcting maps and for the use of map information in image understanding systems. Automated tools for image-to-map registration are therefore of substantial interest for a variety of civilian and military applications. This problem has certain peculiarities that make it more challenging: (a) map features such as roads and buildings may be mis-located; (b) features extracted from images may not correspond to map features; (c) maps contain symbolic information that is not a part of the imagery; (d) maps and images have very different backgrounds, textures and colorations and these have little or no correlation. It was found that the FE-MMI technique can address these difficult problems through the use of a focus-of-attention mechanism based on conventional feature detection that is used to select pixels with high mutual information content in the image to build the MMI calculation with graphic maps.

The realization that a MMI calculation on the label maps was robust and efficient led to experiments with mapping of images to actual cartographic maps. The block diagram of the FE-MMI registration for scene and label image is shown in Figure 4.7. It is robust and reasonably efficient, depending on the extent of the required parameter search. Results below show that this can be an effective technique and a separate method for cartographic maps.

The following discussion presents some experiments involving image-map registration. The reference image is a map and the floating image is an airborne image. The interior differences between image and map details such as road width, building shape and labels will produce mis-located features and present serious problems for conventional feature-based

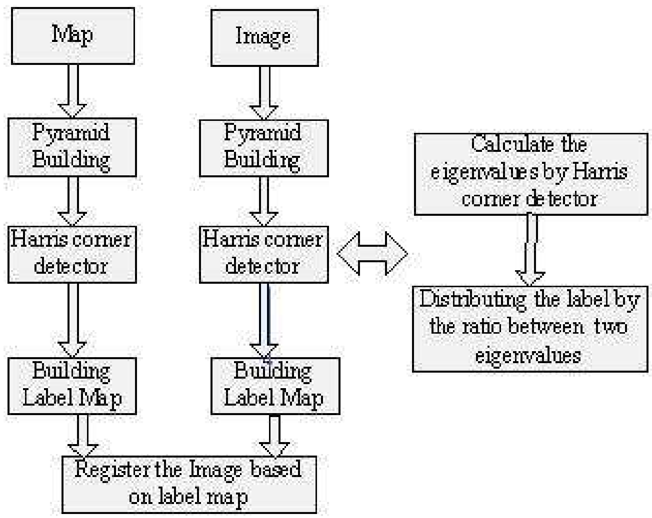


Figure 4.7: Flowchart for the Scene and Map MI registration

algorithms. Conventional cross-correlation methods fail because the statistical correlation between the pixel values for images and maps lack consistency. These differences are shown in Figures 4.8(a) and (b) on the page 52.

Although there are many differences, it is also obvious that the map and image share the same structure, which is the key requirement for using the Harris corner detector to extract the spatial information (Figure 4.8(c) & (d)). The results computed by the FE-MMI algorithm are shown in Figure 4.8(f)

Another experiment was the registration of the satellite image to a map of the RIT campus. The map is from the RIT website (*http://facilities.rit.edu/campus/maps/general/general.pdf*) and the images are from Google satellite image.

The experimental results show that the proposed HCL enhanced MMI algorithm provide an alternative efficient and effective way for map-image registration.

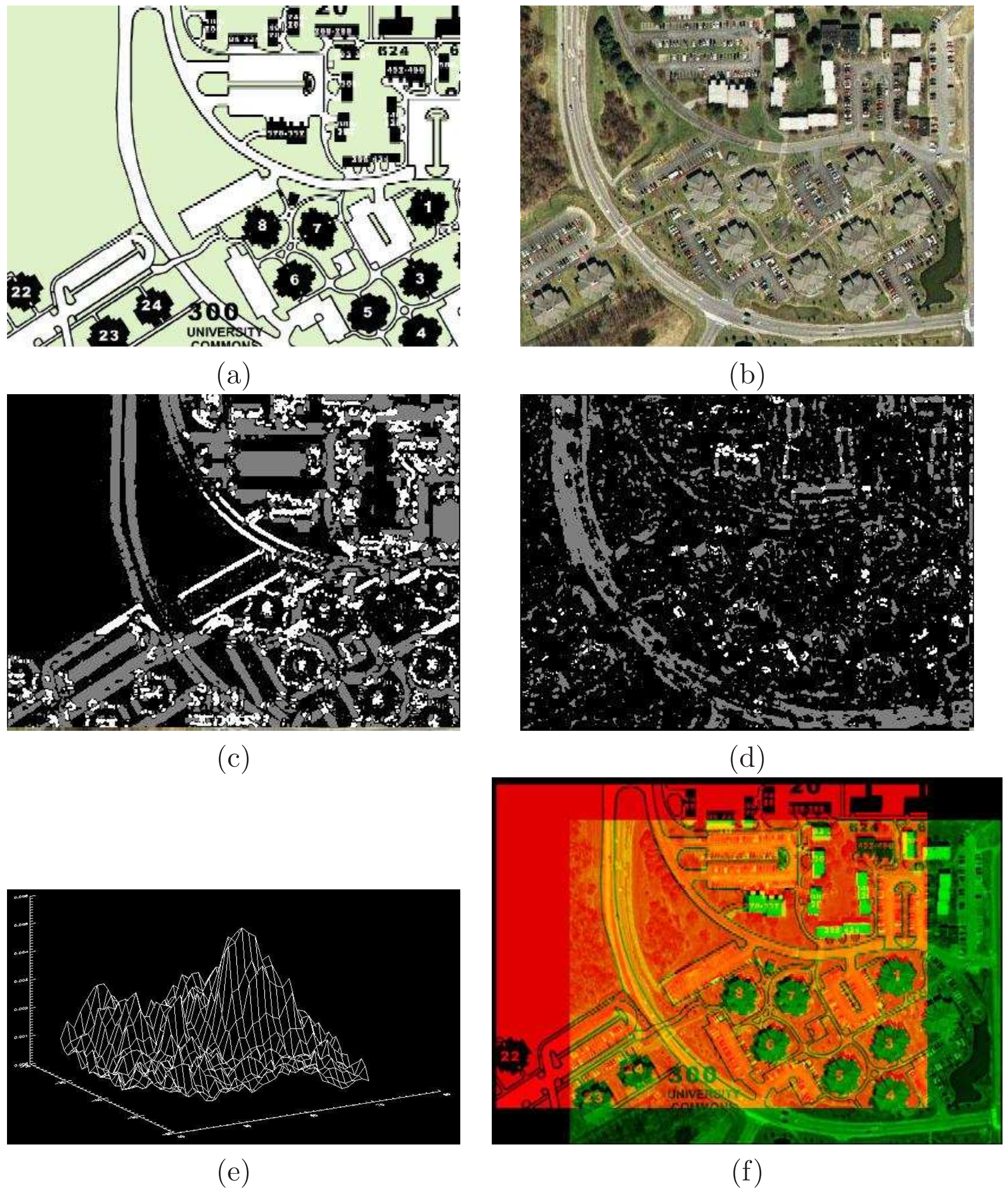
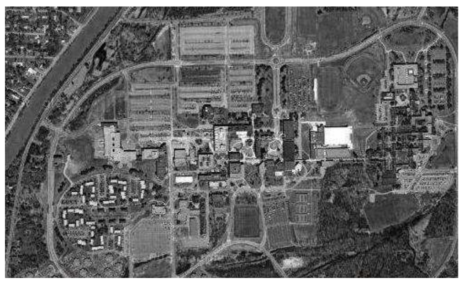
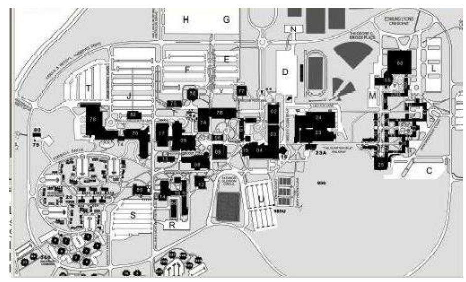


Figure 4.8: Map-Image Registration Example: (a) and (b) represent map and image; (c) and (d) is the Harris Corner Label Mapping for (a) and (b); (e) is the FE- MMI parameter search function ; (f) Overlaid registered images



(a)



(b)

(c)

Figure 4.9: RIT Campus Registration Result

4.4 Summary

This chapter described a novel feature-enhanced MMI algorithm to improve the registration effectively. The results indicate that the techniques are useful for registering maps and images at multiple resolutions. The main assumption is that the maps and images include enough similar structure information for registration. If the Harris corner detector cannot extract enough information-rich data from the image, or the spatial information from images cannot match that in the map, the accuracy and stability will decline.

Because of the lack of the ground truth data or matching control points between the maps and images, it is difficult to measure the registration accuracy. The next chapter will discuss the robustness and accuracy evaluation of the registration result.

Chapter 5

Performance Analysis of Proposed Algorithms

The previous chapter introduced the feature enhanced MMI algorithm and its application to registration between real images and scene vs. cartographic maps. Experimental results show the improvement of the robustness of the algorithm. This chapter will compare FE-MMI with conventional MMI algorithms and analyze their robustness in detail. Another important consideration is the accuracy of the registration results. To measure the accuracy performance both qualitative and a quantitative analysis for the FE-MMI algorithm are presented.

With the improvement of robustness, the FE-MMI algorithm will be tested for the registration of graphic images. It is common for graphic images to contain features that do not appear in images, such as labels and symbols, shapes of things like buildings that do not actually match the images, and features such as roads that are somewhat mislocated. These variations present very difficult problems for accuracy analysis. It's impossible to build a complete accurate sample pair for accuracy analysis, see Figures 5.6 and 5.8.

Although there is a large difference in brightness, coloration and texture, it is obvious that graphic images and scenes share the same general structure. Under this condition it's reasonable to use qualitative analysis to measure accuracy.

5.1 Robust Performance Analysis

This section will describe an experiment that challenges the robustness of the registration algorithm and presents a case where the conventional MMI algorithm fails but the FE-MMI algorithm succeeds. An experiment will then be described that utilizes real-world multi-modal imagery gathered with the four-camera WASP sensor. Finally, an analysis of the relative computational effort of the FE-MMI and traditional MMI algorithms will be discussed.

5.1.1 Experiment to Test Robustness

This experiment compares the performance of the conventional MMI registration algorithm and the feature enhanced MMI algorithm on a pair of synthetic images at different noise levels to simulate data from two different noisy sensors with different spectral responses (Figure 5.1(a)-(b)). The image in Figure 5.1(c) has a constant background, and the object is covered with spatially independent random pixel noise with a normal distribution. The image in Figure 5.1(d) has a reversal in contrast of the structure of Figure 5.1(c). Both the conventional MMI and the FE-MMI algorithms have no difficulty registering the original images, Figure 5.1(e)-(f). However, the conventional MMI algorithm fails to register Figures 5.1(c)-(d). Note that when they are aligned, every pixel is random in one image and constant in the other, making every pixel in one image statistically independent of its counterpart in the other image. Any misalignment leads to overlaps of like regions. Overlaps of random pixel areas maintain pair-wise independence, causing the entropy to be high for all shift positions. Conventional MMI is expected to fail this test and it does [115]. Figure 5.1(g) shows the MMI surface as a function of horizontal and vertical shift that is achieved with Figure 5.1(c)-(d).FE-MMI registration has good performance on both the noisy images as well as on the noiseless pair. Although the pixel intensity inside the region is random, there is useful information at the object edges which the algorithm is able to extract. Focusing on these pixels enables MMI to detect misalignments. From Figure 5.1(h), it could be seen that even though the noise causes a "bump" in the FE-MMI search surface, the peak value is still very clear and easy to locate.

It is expected that the performance of both algorithms will deteriorate as more noise is added. To test this, 100 registration runs were conducted with each algorithm at several noise levels. Table 5.1 shows the percent of successful registrations for each algorithm with a range of noise levels. In each pair, the standard deviation of white noise varied from 1% to 9%. The performance of the traditional MMI algorithm falls off rapidly as noise is added, while the decline of FE-MMI algorithm is much slower because the edge information is more stable than the individual pixel intensity. This is shown in Figure 5.2

Table 5.1: Correct ratio for FE-MMI and MMI algorithm on synthetic image registration

Noise Level	Rate for FE-MMI	Rate for MMI
0 (no noise)	100%	100%
1%	100%	56%
2%	93%	38%
3%	95%	31%
4%	96%	27%
5%	95%	15%
6%	89%	9%
7%	83%	9%
8%	89%	5%
9%	79%	7%

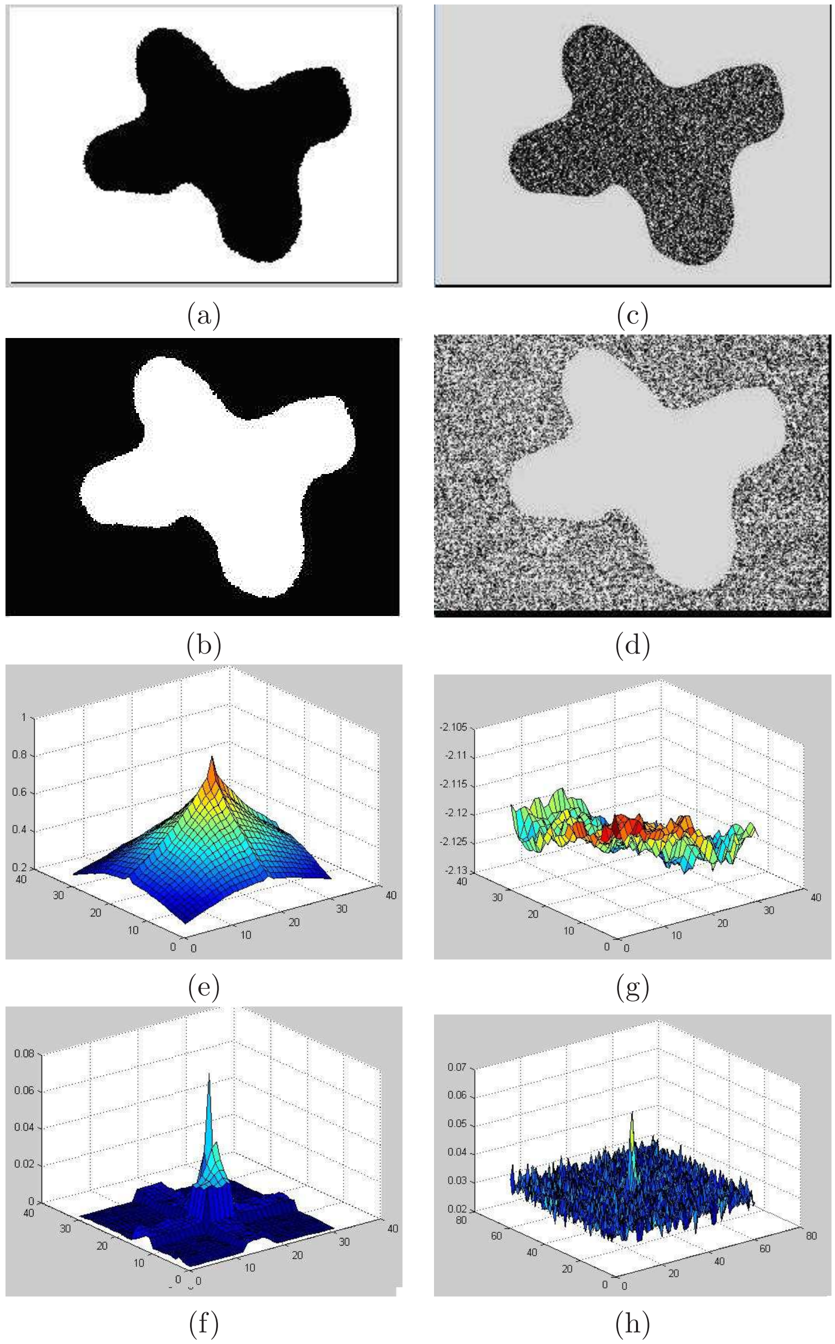


Figure 5.1: Noised image registration Example: (a) and (b) are original image; (c) and (d) represent two noised images, (e) and (f) show the search surface for for two images without noise level; (g) and (h) shows the searching surface for conventional MMI and FE-MMI registration algorithm.

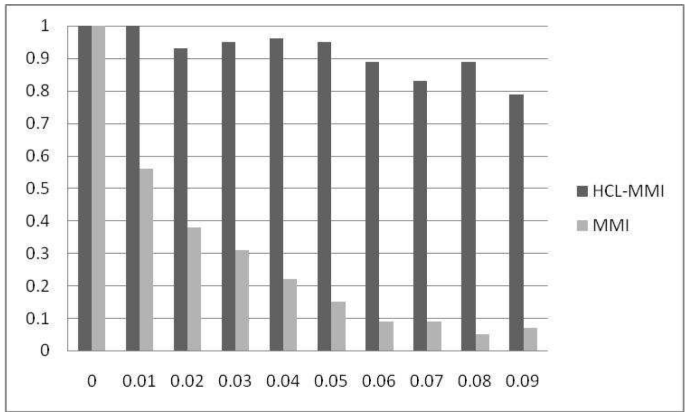


Figure 5.2: Bar Graph for Correct ratio of Table 5.1

5.1.2 Experiments on IR images

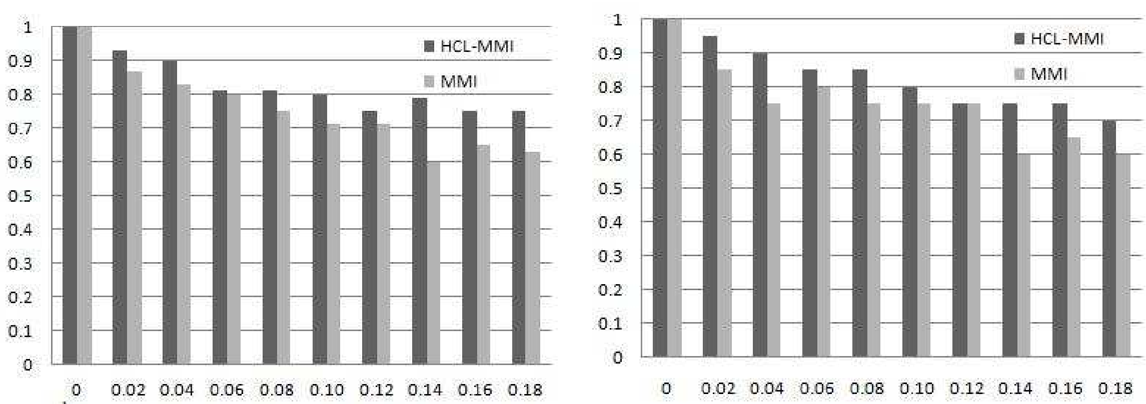
This section will analyze the robustness improvement in the registration of the images of real scenes using synthetic IR images that were generated using DIRSIG [116, 117]. Simulated images provide accurate ground truth data which is usually lacking in the real airborne images. This enables the evaluation of algorithm performance.

Synthetic images such as those shown in Figure 5.3 were used. These images simulate 640×510 IR images of WASP system (Chapter 3.3.1 on page 32). The simulated image hold 10 time the number of pixels on each dimension to enable the evaluation of accuracy at a sub-pixel level.

The floating images were rotated by an angle in the range of [-10, 10] degrees and translated from [50, 150] WASP pixels in x and y directions in this experiment. The registration algorithms were run 100 times each for various levels of noise that was added to each image.

Both translation and rotation error were calculated relative to the ground truth. The registration is a success if both translation on x and y axes and rotation error are below 2 pixels of simulated image and 2 degrees, respectively. The results are shown in the following chart and table:

Figure 5.3: The Synthetic Noised IR images



(a) (b)

Figure 5.4: Bar Graph for Correct ratio vs. noise level for FE-MMI and MMI algorithm on WASP image registration: (a) Intra-band registration with SW band; (b) Inter-band registration between SW & MW band

The result shows clearly that the FE-MMI is more robust than the conventional MMI. Although the success rate for both algorithms is reduced with increasing noise, the FE-MMI algorithm outperformed the traditional MMI algorithm at all noise levels. The registration accuracy can be measured by the means and standard deviations of rotation

Noise Level	Rate for FE-MMI (i)	Rate for MMI (i)	Rate for FE-MMI (ii)	Rate for MMI (ii)
0 (no noise)	100%	100%	100%	100%
2%	95%	85%	93%	87%
4%	90%	75%	90%	83%
6%	85%	80%	81%	80%
8%	85%	75%	81%	75%
10%	80%	75%	80%	71%
12%	75%	75%	75%	71%
14%	75%	60%	79%	60%
16%	75%	65%	75%	65%
18%	70%	60%	75%	63%

Table 5.2: Correct ratio for FE-MMI and MMI algorithm on WASP image registration: i): the left two columns are the correct ratio for FE-MMI and MMI algorithm on shortwave band image registration; ii) the right two columns are the correct ratio for FE-MMI and MMI algorithm on image registration

error and shift error. The Table 5.3 shows comparison results (the better is shown in bold face). In most cases, the registration accuracy and stability are improved. Both algorithms generally achieved sub-pixel accuracy.

5.2 Registration Accuracy Analysis

Another important consideration is the accuracy of the registration results. To measure the accuracy performance both a qualitative and a quantitative analysis for the enhanced MMI algorithm are presented.

5.2.1 Accuracy Analysis Method

The registration accuracy can be evaluated by qualitative and quantitative analysis. Qualitative analysis measures the sharpness of the peak value of the mutual information as the parameters are varied. This method is especially useful for multi-modal images that lack

Table 5.3: Means and standard deviations of errors for the success result of FE-MMI and MMI algorithm on WASP image registration. (*Bold black represents the better results.*)

i: Result from MMI Algorithm

Noise level	Rotation	X	Y
0 (no noise)	0.05 ± 0.23	0.75 ± 0.23	0.5 ± 0.26
2%	0.1 ± 0.09	0.3 ± 0.75	$\mathbf{0.3 \pm 0.75}$
4%	0.7 ± 0.16	0.35 ± 0.76	0.35 ± 0.36
6%	$\mathbf{0 \pm 0}$	0.50 ± 0.76	0.55 ± 0.76
8%	0.15 ± 0.73	0.50 ± 0.76	0.35 ± 0.36
10%	0.05 ± 0.05	0.55 ± 0.7	0.5 ± 0.76
12%	$\mathbf{0.05 \pm 0.05}$	0.65 ± 0.33	$\mathbf{0.3 \pm 0.75}$
14%	0.15 ± 0.13	0.65 ± 0.76	0.75 ± 0.30
16%	$\mathbf{0 \pm 0.05}$	$\mathbf{0.65 \pm 0.73}$	$\mathbf{0.55 \pm 0.75}$
18%	0.13 ± 0.15	$\mathbf{0.35 \pm 0.76}$	0.6 ± 0.75

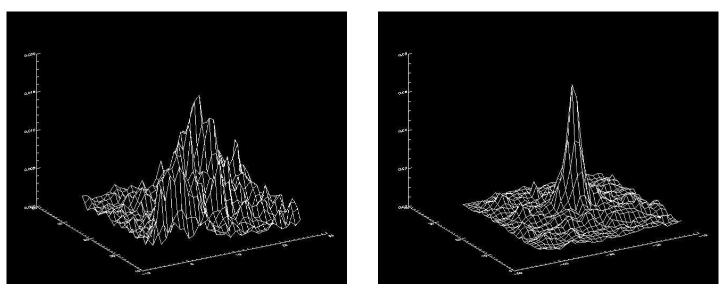
ii: Result from HCL-MMI Algorithm

Noise level	Rotation	X	Y
0 (no noise)	$\mathbf{0 \pm 0}$	$\mathbf{0.3 \pm 0.23}$	$\mathbf{0.3 \pm 0.27}$
2%	$\mathbf{0 \pm 0}$	$\mathbf{0.25 \pm 19}$	0.35 ± 0.73
4%	$\mathbf{0 \pm 0}$	$\mathbf{0.25 \pm 0.19}$	$\mathbf{0.35 \pm 0.26}$
6%	0.15 ± 0.13	$\mathbf{0.35 \pm 0.73}$	$\mathbf{0.35 \pm 0.36}$
8%	$\mathbf{0 \pm 0}$	$\mathbf{0.35 \pm 0.7}$	$\mathbf{0.35 \pm 0.23}$
10%	$\mathbf{0 \pm 0}$	$\mathbf{0.50 \pm 0.73}$	$\mathbf{0.35 \pm 0.73}$
12%	0.1 ± 0.09	$\mathbf{0.6 \pm 0.75}$	0.3 ± 0.75
14%	$\mathbf{0.1 \pm 0.09}$	$\mathbf{0.35 \pm 0.77}$	$\mathbf{0.3 \pm 0.75}$
16%	0.1 ± 0.09	0.65 ± 0.73	0.6 ± 0.35
18%	$\mathbf{0 \pm 0}$	0.5 ± 0.13	$\mathbf{0.3 \pm 0.75}$

accurate ground truth data. On the other hand, for maps or GIS images with high spatial accuracy, one can take advantage of the available location data to provide a quantitative analysis. It is also possible to test the algorithm with a variety of simulated imagery for which ground truth is automatically available to provide a quantitative assessment of performance.

Qualitative analysis

A straightforward measure of the registration sensitivity to parameter variation is the sharpness of the peak value in the scaled search surface over the transformation parameters. A measure of the search surface's sharpness is provided by the kurtosis of the distribution, where larger kurtosis represents a sharper peak. Experiments show that the kurtosis provides a good measure of sharpness for the non-Gaussian distributions, especially for the Generalized Gaussian distribution [118].



(a) (b)

Figure 5.5: The sharpness of the search surface with low and high kurtosis: (a) search surface with kurtosis = 2.96; (b) search surface with kurtosis = 17.55.

The definition of kurtosis is:

$$k = \frac{n \sum_{i=1}^{n} (x_i - \bar{x})^4}{(\sum_{i=1}^{n} (x_i - \bar{x})^2)^2}$$

Experiments use kurtosis to measure whether the data are peaked or flat relative to the normal distribution. If the kurtosis is high the peak tends to be sharp, whereas low kurtosis indicates a distribution with a flat top instead of distinct peak. This measurement works well and more reliably than simple variance.

Quantitative analysis

Quantitative analysis of registration accuracy is often difficult because of the lack of accurate ground truth data, especially from the multi-modal imagery. When it is available, a GIS image can provide accurate ground truth data. The registration error can be calculated by comparing points in the GIS image to the corresponding points in the transformed images. In these experiments the ground truth points and corresponding image were picked manually, but this could potentially be replaced by an automatic algorithm. A mean squared geometrical error can be measured as:

$$e = \sqrt{\frac{1}{N}\sum_{N} ||G - T(I)||^2}$$

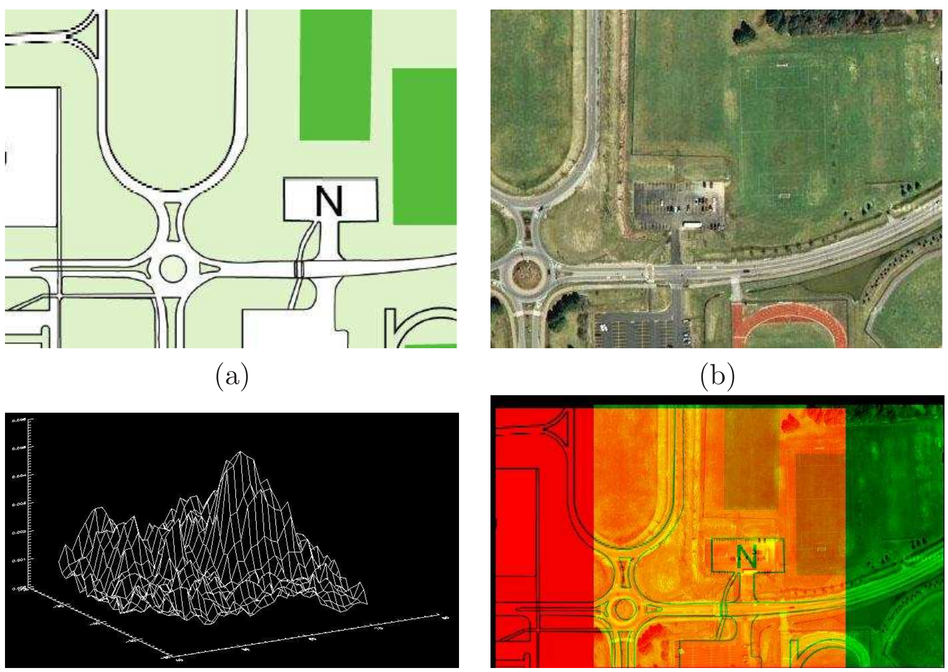
Where G is the sample set in reference image (GIS image), I is the set of corresponding points in the floating image and T is the geometrical transformation. A geometrical registration error can be estimated by the Euclidian distance between sample sets from two images.

5.2.2 Accuracy Analysis Example

This section shows two examples of registration between sensor images and graphic images. The sensor image is an airborne remote sensing image while the graphic image is a map in the first experiment and a GIS model in the second experiment. The matching algorithm for these experiments was found to be consistent, accurate and robust.

Map vs. Image Registration

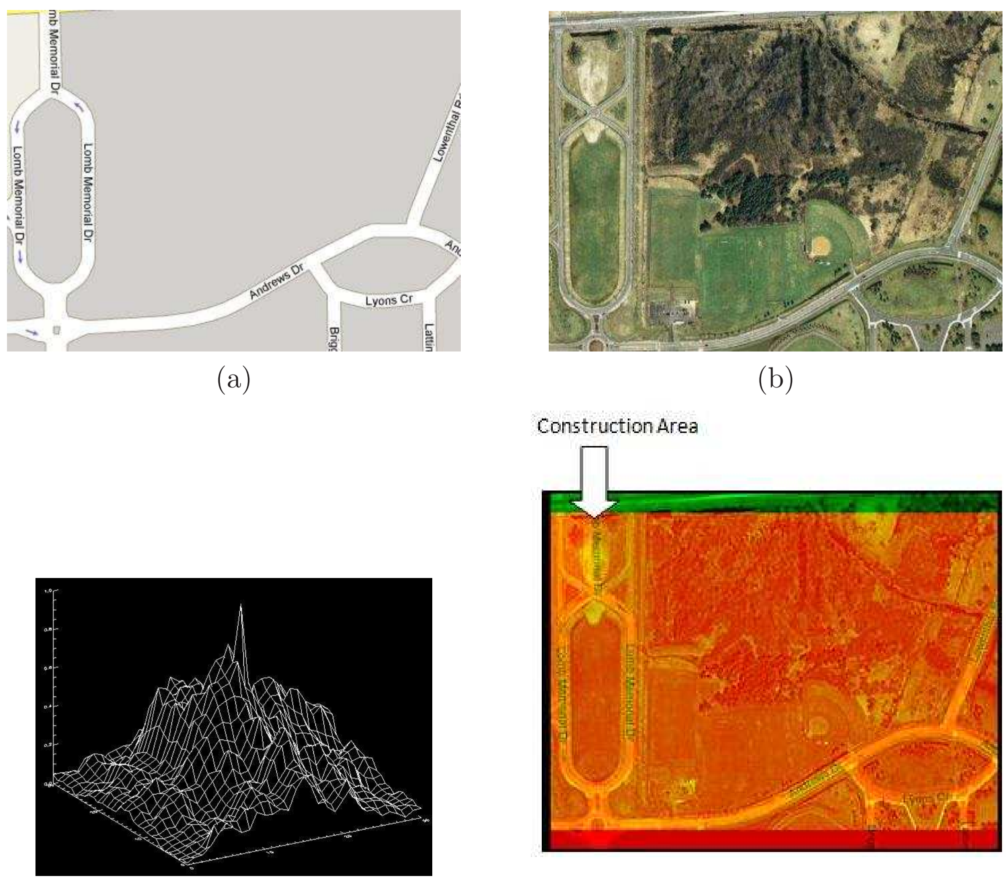
Image-to-map registration can be used for change detection and map revision. The fact that changes are expected means that the registration method must not fail when a portion of the data is in disagreement. To illustrate this application, a map from Google Maps and an image that was taken perhaps a year later were registered using the enhanced MMI method. The results are shown in Figure 5.7. The change in a road position that was made before the image was taken is evident. A road-following algorithm would be readily able to pick up this change between the registered image and map.



(a) (b)

(c) (d)

Figure 5.6: Map-to-Image Registration Example: (a) and (b) represent map and image; (c) is the HCL- MMI parameter search function ; (d) Overlaid registered images



(a) (b)

(c) (d)

Figure 5.7: Map update Example: (a) and (b) represent map and image; (c) is the FE-MMI parameter search function ; (d) Overlaid registered images

The sensitivity of the enhanced MMI algorithm for the two examples above is quantified by the variance and kurtosis values.

Table 5.4: The kurtosis of search surface

	Variance	Kurtosis
Example 1	0.0471	2.624
Example 2	0.0211	3.761

The variance and kurtosis analysis show that the sensitivity of the calculation is higher in the second example, which matches the visual observation: the search surface of example 2 has the sharper peak. The sharp peak in Figure 5.7(c) compared to the flatter surface in Figure 5.6(c) is indicated by the respective kurtosis values. It is also notable that both surfaces have a rather rough quality, which indicates that caution must be used in how the parameter search is implemented. Gradient-based techniques failed on these examples but exhaustive search implemented through a resolution pyramid succeeded.

The main limitation for this algorithm is the need for the maps and images to include enough similar structure information to permit a statistical measure of correspondence. If there is not enough detail in the map or if the feature detector cannot extract enough information-rich data from the image then accuracy and stability both decline.

GIS vs. Image Registration

Another valuable application of this algorithm is automated image-to-GIS registration. An image can be formed from GIS data and then registered by the FE-MMI technique. An automated registration tool can be used to locate in an image objects that are described by models, for example. An example is shown in Figure 5.8, where an image of a group of buildings was formed from GIS data and registered to a remote sensing image. Note that this was done on a small portion of a much larger image. This technique could be extended to provide a search method for known shapes within an image, and could provide a statistics-based target finding or recognition algorithm.

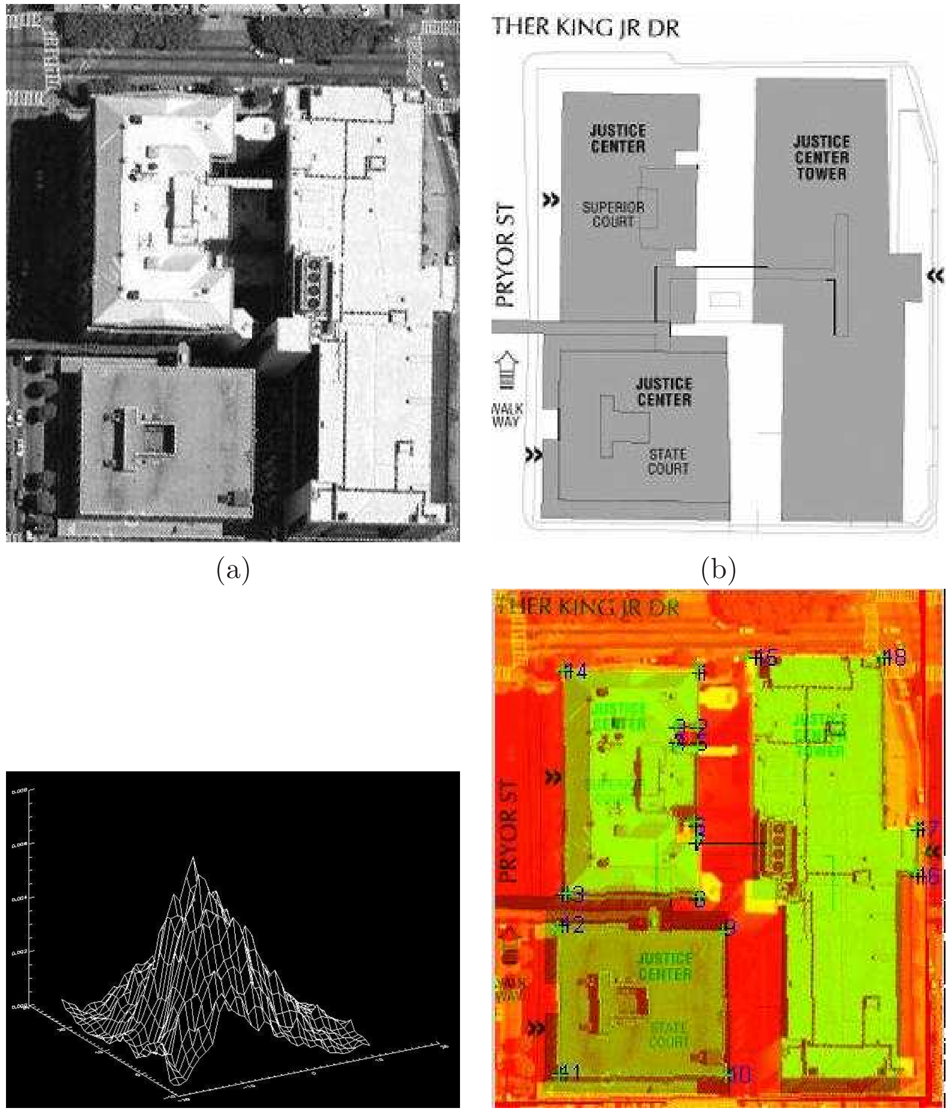
To evaluate the quality of the registration for this example, a number of matching points on prominent features were manually extracted from the GIS image and scene image and the GIS image was set as fixed image and the real image as the floating image. The transformation was applied and the registration error was measured using Euclidean distance.

Table 5.5: Registration Results

Registration Result	Rotation	X	Y	Mean Error
Feature point based algorithm	-1.15	-10.5	-5.10	3.64 pix
FE-MMI based algorithm	-1.5	-10	-8	3.18 pix

To compare the results presented with other registration algorithms, the prominent point sets were randomly divided into two parts: the first part was for registration and the second was for accuracy testing. The optimization method to find the registration parameters were set to minimize the Euclidian distance between the first part of feature sets from the GIS image and transformed floating image. Then the registration parameter was applied to the second part of feature sets to measure the registration accuracy. To minimize the random noise effect, both registration algorithms were run 100 times. The mean value of registration parameters and registration error are in Table 5.5 on page 68.

The quantitative analysis in Table 5.5 shows the difference of the results from two algorithms are a half degree difference on rotation and and about two pixels difference on shifting. It is doubtful that the objects in these images can be placed with more accuracy. Automatic processing makes it possible to handle large images efficiently and effectively, and will enable a high-speed registration workflow.



(a) (b) (c) (d)

Figure 5.8: Image-to-GIS registration example: (a) An airborne image from region of Atlanta, GA showing buildings and roads; (b) GIS map taken from Google Earth; (c) The FE-MMI search surface, which shows a sharp peak at the parameter values that provide the best transformation; (d) A composite image produced by overlaying the registered map and image.

5.3 Computational Effort

This section compares computation effort between conventional MMI and FE-MMI. Conventional MMI registration uses a set of points X that are distributed randomly over the image. The samples are used to construct the joint histogram of data values between the images. The registration process looks for a transformation that minimizes the joint entropy of the histogram constructed from $A(X)$ and $B(T(X))$. The computational effort depends upon the number of samples N, the histogram bins K and the number of steps S used in the parameter search process. Since the computational effort is in direct proportion to the number of samples and the number of bins required in the joint histogram, the computational cost to the samples and histogram bins are $O(N)$ and $O(K^2)$. The total computation needed is:

$$O(T) = O(N) \cdot O(K^2) \cdot O(S)$$

The FE-MMI algorithm improves the efficiency three ways. (i) In conventional MMI the samples are taken uniformly across the images. The sample grid must be small enough to ensure the capture of a sufficient number of points with high information value. This algorithm reduces the computation requirement by selecting the information-rich pixels that then reduces the sample size. In the following experiment, the number of samples per image can be reduced by 80% without loss of registration accuracy. (ii) The label maps, rather than the original imagery, can also be used in the mutual information calculation. In the HCL maps, the pixels on background, edges and corners are denoted by classification values 0, 1, 2. This removes most of the background variation but preserves essential information about the location of the edges and corners. The reduction of the number of histogram bins from 256 (for typical byte-scale normalized images) to 3 enables the 2D histogram of sample values to be reduced from 256^2 to 3^2. This can save substantial memory and computing time. (iii) A pyramid approach can be used to enable a hierarchical search for the optimum parameter values. An approximate registration can be accomplished by using larger steps in parameter values with lower-resolution images. The value for maximum entropy at low-resolution was used as the starting point for a refined search at the next higher resolution. The Harr wavelet [119] was used to build a pyramid and have found that the pyramid algorithm leads to about 90% reduction in the volume of search region without accuracy loss comparing with the exhaustive search [120].

Taken together, the three steps described above can reduce the computing time by 90%. For the intra-band registration, the rotation and shift, and scale plus shift for inter-band registration were measured. The running time for a typical registration was reduced from 2 minutes to about 10 seconds. The following table is the comparison of time consumption: (The running environment is P4 2.0GHz 1G RAM, IDL 6.0)

Table 5.6: Time Consuming for HCL-MMI and MMI algorithm on WASP image registration

	Shortwave band *vs.* Shortwave band	Shortwave band *vs.* Middle-wave band	Shortwave band *vs.* Long-wave band
HCL-MMI	8.78(s)	10.79 (s)	10.86(s)
MMI	103(s)	117(s)	115(s)

This improvement is significant and may enable the proposed algorithm to be suitable for serial image registration in near real-time applications.

5.4 Summary

This chapter analyze the efficiency, accuracy and robustness of the registration algorithms. The experiments show efficiency is improved by using features to select information-rich areas, by using the classification label image to reduce the joint histogram calculation, and using pyramid techniques to achieve coarse-to-fine analysis. The reduction in computing time is not done at the expense of registration accuracy.

Chapter 6

Validation Analysis of Sample Sets

The previous chapter introduced the application of conventional MMI algorithm on WASP images and the extension, FE-HCL algorithm. The performance in the registration application is analyzed. The improvement of the FE-MMI is through the focusing of the MMI process on information-rich pixels. However, one problem for FE-MMI still exists: the registration results are not always stable. This instability will make the registration results very unreliable in some extreme conditions. The reason for unstable performance is random sample selection. All pixels are treated the same and samples are randomly selected from the whole image. This chapter will provided a more effective method based on the focus-of-attention features to construct the sample sets for MMI registration. Comparing with the random selection, fewer samples are needed without any loss of the registration performance.

6.1 The Analysis of Sample Set Validation

Good sample sets should improve the the robustness, accuracy and efficiency performance of the registration algorithm. In chapter 4, the Harris corner detector is used to divide the pixels into the information-rich area and flat region, and this classification is effective in improving the MMI registration algorithm as shown in Chapter 5. It is straightforward to bring forward this question: Can the Harris corner detector help to select the samples? Or in other words, if only the pixels rich-information are selected, is the sample set better than random selection?

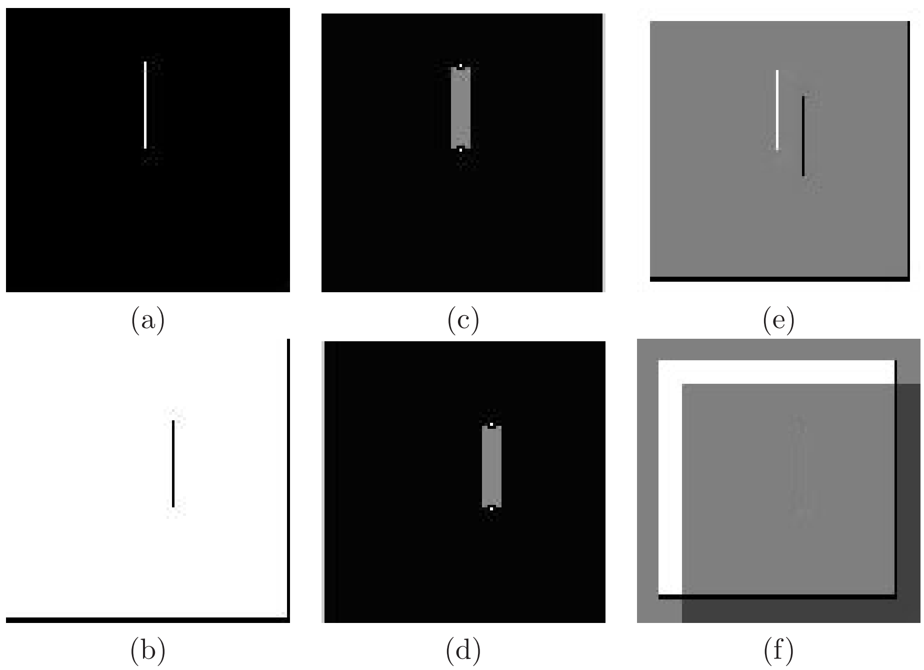


Figure 6.1: Experiments to test special sample set: (a) and (b) are original images; (c) and (d) are the Harris corner label map, the gray part is the information-rich area and the black part is the flat region; (e) and (f) are the well-aligned and misaligned registration results.

A simple experiment is designed to test this idea: the Figure 6.1 (a) and (b) are the images of a line in solid background and the reversed foreground and background. Figure 6.1(b) is shifted down-left from the Figure 6.1(a). The Figure 6.1(c) and (d) show the information-rich area and the flat region by the classification of Harris corner detector. A special sample set only from the information-rich area was built and tested by registration.

As Figure 6.2 shows, the mutual information of this special sample makes FE-MMI algorithm fail on this registration. The search surface is very flat (Figure 6.2). The Table 6.3 shows the mutual information reaches a maximum using both well-aligned (Figure 6.1(f)) and misaligned images (Figure 6.1 (e)). The reason is that all pairs fall into the [1,0] when mis-aligned and [1,1] when well-aligned, however, the joint entropy stays the same because it is the summation of the total distribution of the pairs without distinguishing the single pair probability. To build an effective sample set needs a complicated analysis.

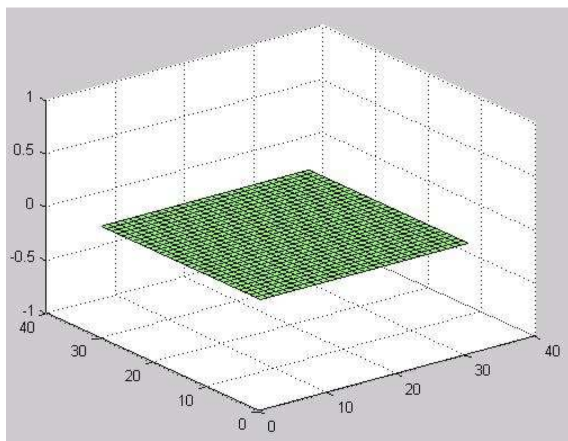


Figure 6.2: The search surface of the registration between sample set X and Y

Table 6.1: The joint histogram of well-aligned and mis-aligned between sample set X and Y

	[0,0]	[0,1]	[1,0]	[1,1]	$H(X,Y)$
Well Aligned	0	0	142	0	0
Mis Aligned	0	0	0	142	0

The question is simplified for Harris Corner Label map. The Harris corner label includes three values [0,1,2] and value [1,2] could be merged to represent the information-rich region. The sample set X from Figure 6.1 (a) consist of two subsets: u (0 value points) and v (1 value points), while the corresponding subsets Y in Figure 6.1(b)is u' and v'.

The mutual information is expressed in Chapter 3 as:

$$I(A(X), B(Y)) = H(A(X)) + H(B(T(X))) - H(A(X), B(T(X))) \tag{6.1}$$

With the two-value sample set, the last term could be represented as:

$$H(A(X), B(Y)) = -[P(u, u') \log P(u, u') + P(u, v') \log P(u, v') + P(v, u') \log P(v, u') + P(v, v') \log P(v, v'$$

where $P(u, v')$ could be calculated by $P(u) \cdot P(v'|u)$ (Bayes theorem), the $H(A(X), B(Y))$ could be expressed as:

$$\begin{aligned} &H(A(X), B(Y)) = -[P(u) \cdot \log P(u) + P(v) \cdot \log P(v) + \\ &P(u) \cdot (P(u'|u) \cdot \log P(u'|u) + P(v'|u) \cdot \log P(v'|u)) + \\ &P(v) \cdot (P(u'|v) \cdot \log P(u'|v) + P(v'|v) \cdot \log P(v'|v))] \end{aligned}$$

The conditional entropy is defined as:

$$\begin{aligned} H(Y|u) &= P(u'|u) \cdot \log P(u'|u) + P(v'|u) \cdot \log P(v'|u) \\ H(Y|v) &= P(u'|v) \cdot \log P(u'|v) + P(v'|v) \cdot \log P(v'|v) \end{aligned}$$

The final expression of $H(A(X), B(Y))$ is:

$$H(A(X), B(Y)) = H(A(X)) + P(u) \cdot H(Y|u) + P(v) \cdot H(Y|v) \tag{6.2}$$

Combining equations 6.1 and 6.2, the mutual information of two images is:

$$I(A(X), B(Y)) = H(B(Y)) - [P(u) \cdot H(Y|u) + P(v) \cdot H(Y|v)] \tag{6.3}$$

Equation 6.4 demonstrates mutual information is positive proportion to the entropy in subsets Y and the conditional entropy has negative effect. The good sample sets should make the mutual information detect the subtle parameter change promptly, which requires mutual information has enough space to reflect the change of the parameter set. The following is the detailed analysis of the results of applying equation 6.4 to the two value sample sets.

Table 6.2: The relationship between sample set X and Y

$P(u)$	$u \to u'$	$u \to v'$
α	β	$1-\beta$
$P(v)$	$v \to u'$	$v \to v'$
$1-\alpha$	γ	$1-\gamma$

α is the ratio of u in sample X, while β is the ratio u in X projected to u' and γ is the ratio v in X being projected to v'. From the Table 6.2, the composition of sample Y may be derived easily:

Table 6.3: The composition of the sample set Y

u'	v'
$\alpha\beta + (1-\alpha)(1-\gamma)$	$\alpha(1-\beta) + (1-\alpha)\gamma$

With the Table 6.3, the three terms of equation 6.4 may be calculated as:

$$H(Y) = -[\{\alpha\beta + (1-\alpha)(1-\gamma)\} \cdot \log\{\alpha\beta + (1-\alpha)(1-\gamma)\}+$$
$$\{\alpha(1-\beta) + (1-\alpha)\gamma\} \cdot \log\{\alpha(1-\beta) + (1-\alpha)\gamma\}]$$

$$H(Y|u) = -[\beta \log \beta + (1-\beta)\log(1-\beta)]$$

$$H(Y|v) = -[\gamma \log \gamma + (1-\gamma)\log(1-\gamma)]$$

First some extreme condition will be discussed:

- $\beta = \gamma = 0$ or $\beta = \gamma = 1$. The corresponding $H(Y|u)$ and $H(Y|v)$ are equal to 0. Under this condition, the mutual information is:

$$I = -[\alpha \log \alpha + (1 - \alpha) \log(1 - \alpha)]$$

The mutual information I reaches the maximal value $\log 2$ when $\alpha = \frac{1}{2}$. This result is instructive to the experiment discussed in chapter 5 section 5.1.1.

- $\beta = \gamma$. $H(Y|u)$ is equal to $H(Y|v)$ under this condition and can be treated as the constant. The mutual information $I \propto H(Y)$. It is known in statistic theory that the $H(Y)$ get maximized when u' and v' take half to half. From the table 6.3, $P(u') = P(v')$.

$$\alpha\beta + 1 - \alpha - \gamma + \alpha\gamma = \alpha - \alpha\beta + \gamma - \alpha\gamma \tag{6.4}$$

$\alpha = \frac{1}{2}$ where $\beta = \gamma$. In this case, joint entropy is:

$$H(A(X), B(Y)) = \log 2 + \frac{1}{2} \cdot (H(Y|u) + H(Y|v))$$

- The general case will be discussed here. The α should maximize $H(Y)$ without any special knowledge about β and γ. Continuing the discussion in equation 6.4, the best α value is:

$$\alpha = \frac{1}{2} \frac{2\gamma - 1}{\beta + \gamma - 1}$$

The above shows how to choose a two-value sample set effectively in theory. Unfortunately, in practice it's typically not possible to use this method, since one rarely knows β and γ accurately. Fortunately the difference between β and γ is small if the sample sets are big enough. In most application it could be assumed as the same and the error of the approximation could be ignored. This analysis is easily extended to multi-value sample sets.

6.2 Experimental Results

Previous section analyzed the effective method to construct the sample set. In brief the good sample sets should have the maximal entropy. This section will present the application in registration experiments. Let's come back the experiment in section 6.1 first and compare the registration result.

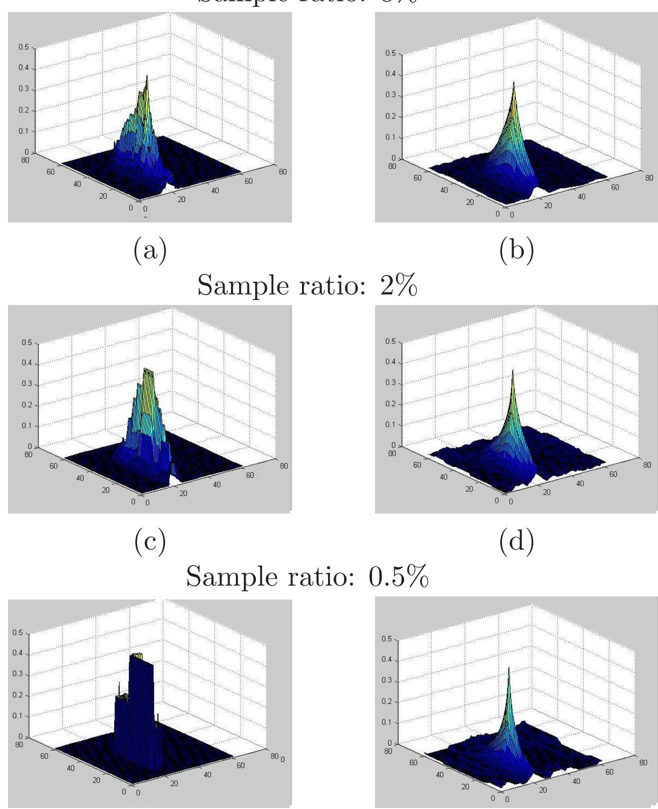
The original images for registration are the Figure 6.1 (a) and (b) on page 74. The information-rich area takes about 8% of the whole image (Shown in Figure 6.1 (c) and (d)). The left column and right column of Figure 6.2 are the comparison for different sample sets and each row is of different sample ratio. Both sample sets work well at 8% sample ratio. When the sample ratio is below 2%, the registration result from the sample set randomly selected becomes unstable and registration fails. However, the registration performance of the optimized sample set remains stable even at the sample ratio 0.5%.

Next search surfaces of the sample sets with different entropy are compared under a fixed sample ratio. Figure 6.1 (a) and (b) are still used as original images. The size of sample sets are limited to 1% of the whole image. The registration on small sample set is more sensitive to the entropy change. The ratio between the pixels from the information-rich area and flat region is varied from 10% to 90%.

Figure 6.4 shows the registration results of sample sets with different entropy under the 1% sample ratio. Experiment shows that both too few and too many pixels from information-rich area or flat region harm the registration stability. The information-rich area and flat region can be regarded as a kind of classification. The registration is very

Sample set randomly selected   Optimized sample set

Sample ratio: 8%



(a) (b)

Sample ratio: 2%

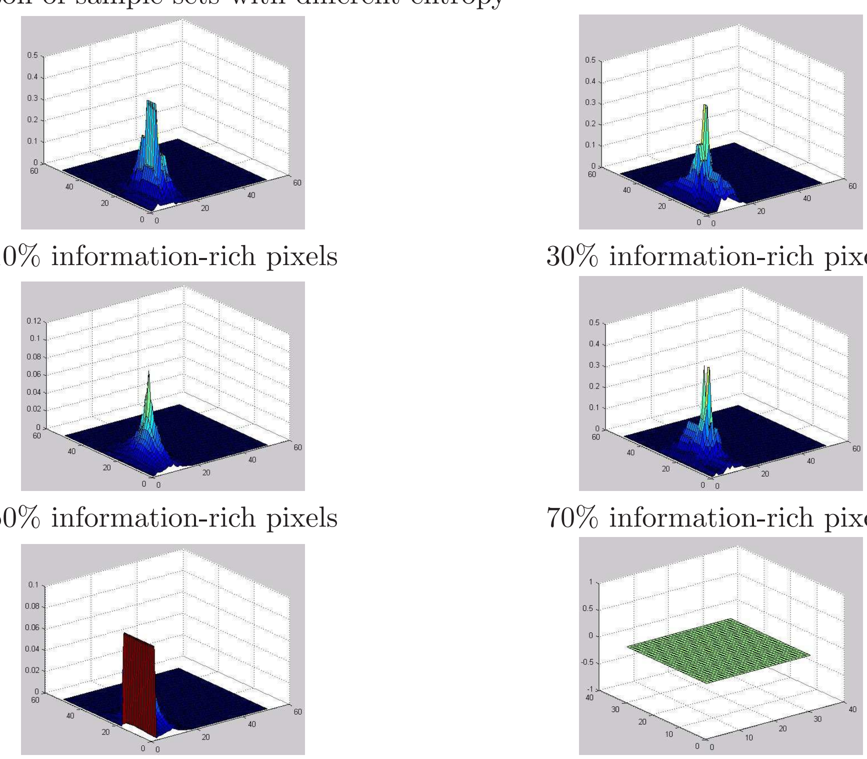
(c) (d)

Sample ratio: 0.5%

(e) (f)

Figure 6.3: The comparison of the registration result of Figure 6.1(a) and (b) by different sample sets. The Figure (a) (c) and (e) in left column are the search surfaces of the sample set randomly selected while the figure (b) (d) and (f) in right column are the search surfaces of the optimized sample set. Each row represents different sample ratio: (a) and (b): sample ratio is the 8%; (c) and (d): sample ratio is the 2%; (e) and (f) sample ratio is the 0.5%.

The comparison of sample sets with different entropy



10% information-rich pixels

30% information-rich pixels

50% information-rich pixels

70% information-rich pixels

90% information-rich pixels

100% information-rich pixels

Figure 6.4: The comparison of the registration result of Figure 6.1(a) and (b) by sample sets with different entropy.

unstable and fails at 90% samples from one category. The registration results are better with more balance between the sample pixels from two categories. This result is valuable for constructing effective sample sets because it is very rare for the information-rich area and the flat region to each take half in a real application.

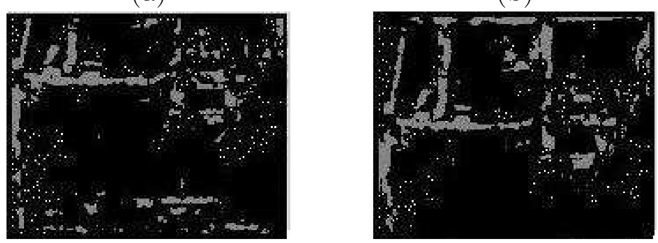
The following experiment tests the performance on WASP image. The Figure 6.5(a) and (b) are the WASP short-wave images and the Figure 6.5(c) and (d) are the Harris corner label images. Almost 12% of the points are in the information-rich area. In previous experiment, the left column is the search surface of the randomly selected sample set and the right is for the optimized sample set.

The Figure 6.5 (e)-(j) represent the comparison of registration result of two sample sets in different sample ratio. The experiment result shows the performance of the optimized sample set is better than the random one. The registration could be achieved with fewer sample points, which is an effective method to improve efficiency.

6.3 Summary

The traditional MMI algorithm select sample sets randomly from the original image. It is easy to show cases in which the registration result is unstable. This chapter describes a method to optimize the sample set. With the classification by Harris corner detector, the sample set is optimized by maximizing the entropy of the Harris corner label. The experiment shows the registration stability is improved with fewer sample points.

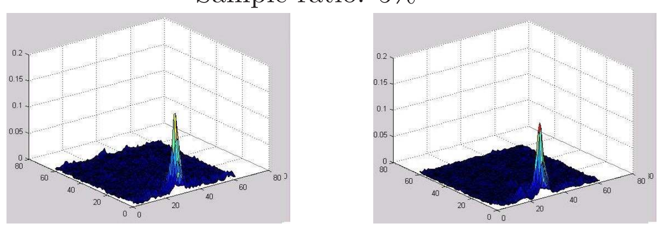
(a) (b)



(c) (d)

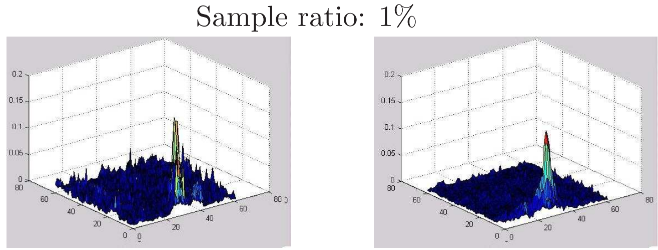
Sample set randomly selected Optimized sample set

Sample ratio: 5%



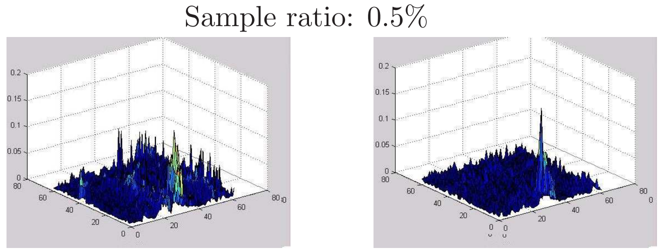
(e) (f)

Sample ratio: 1%



(g) (h)

Sample ratio: 0.5%



(i) (j)

Figure 6.5: The comparison of the registration result of the WASP short wave band by different sample sets: (a) and (b) are the original images; (c) and (d) are the corresponding Harris Corner label map. The left column are the search surfaces of the sample set randomly selected while the right column are the search surfaces of the optimized sample set. Each row represents different sample ratio: (e) and (f): sample ratio is the 5%; (g) and (h): sample ratio is the 1%; (i) and (j) sample ratio is the 0.5%.

Chapter 7

Registration Area Selection

In previous chapters, the success of enhanced mutual-information techniques in registering multi-modal remote sensing imagery was reported [120]. The enhancement is provided by the use of Harris feature detection [113] to locate places that are likely to have high information content, and then these locations are used in the MMI process. This improved the speed and robustness of the algorithm for a variety of airborne image registration tasks. In [121], it was shown that the Harris feature map, which is an image with three brightness values, could be used as an element in MMI image registration. The result is a robust registration tool with reduced computing time.

However, one problem for the algorithm is that the area for mutual information calculation is based on the whole image. In most conditions this can provide enough information for registration. However, in the application of remote sensing image registration, the following condition may arise:

- The area of overlap between the images may be small and in unknown locations. Samples taken from non-overlap regions can dominate the statistical analysis and lead to registration failure.

- Different parts of the images provide better registration information due to noise, illumination conditions and sensor characteristics. Some areas like forest and grass are highly textured but statistically uniform, while other areas that include buildings or roads are more useful. Areas in each image that have good potential for registration are called salient areas.
- The distribution of the noise, illumination conditions and variance caused by sensor characters are not constant in the whole image. Some parts are more similar than other parts and registration based on similar areas has potential improving the registration.

These problems could be relieved by an initialization procedure. A common condition is that the image quality is poor for registration in remote sensing applications because of extreme illumination changes. A reasonable assumption is that registration will be improved by only using the overlap or salient areas that include more local object information or less noise. To use this strategy, image segmentation can be done after initialization procedure.

This chapter will further extend the approach of using feature detection to bootstrap the MMI algorithm to achieve robust and efficient performance in the registration of imagery with diverse characteristics. By combining the SIFT algorithm [122] with an MMI analysis to find regions in the two images that are likely candidates for matching, the salient areas in each image can be identified and then each possible pairing between the images can be analyzed to find salient area matches. Harris feature detection with MMI on the candidate regions can then be used to determine the parameters of the transform for matching. This two-step approach overcomes the above problems and provides a good initialization for the final step of the registration process.

The following sections will introduce the extension in detail. Section 7.2 discusses salient area extraction by the combination of SIFT and local entropy. Section 7.3 discusses experimental tests. The results are analyzed in Section 7.4.

7.1 Salient Region Extraction

A salient region is a small part of an image that is noticeable because it has a high density of features. Spatial information such as shape, size and structure can be measured to find the coarse matches of similar regions and then registration can be refined by additional processing. It has been found that a multi-scale approach to region matching to be effective because it facilitates comparisons between images from modalities with different spatial resolutions. In this algorithm the scale invariant feature transform (SIFT) was used to identify salient regions.

7.1.1 Scale-Invariant Feature Transform

The scale-invariant feature transform (SIFT) was used by David Lowe in 1999 [123] to detect local features in images. The key point of the SIFT algorithm is it makes use of the local extrema in the intensity domain, which is difficult to obtain from edge detectors or corner detectors. Such extrema points are the points of interest because they can provide important information about regions. These regions are the potential salient regions for our registration.

The SIFT algorithm convolves the image with Gaussian filters at different scales, and then takes the difference of successive Gaussian-blurred images. The maxima/minima of the Different of Gaussian (DOG) that occur at multiple scales are the extracted keypoints. The DOG of the image could be defined as:

$$D(x, y, \sigma) = L(x, y, k_i\sigma) - L(x, y, k_j\sigma) \tag{7.1}$$

where $L(x, y, k\sigma)$ is the Laplacian of Gaussian, which is the original image $I(x, y)$ convolved with the Gaussian blur $G(x, y, k\sigma)$ at scale $k\sigma$:

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

The DOG image comes from the different Gaussian-blurred images at scale $k_i\sigma$ and $k_j\sigma$. In practice the Gaussian blur functions are grouped by octave, and each octave corresponds to doubling the value of σ. The k value is selected so a fixed number of convolved images per octave are obtained. The difference-of-Gaussian (DOG) is calculated from adjacent Gaussian-blurred images per octave.

The local extrema points detection are taken with respect to both space and scale. This is done by comparing each pixel in the DOG image with it's eight neighbors at the same scale and nine corresponding pixels in each of the neighboring scales. Thus, the pixel will be the compared with the nearest 26 neighbors in scale-space volume. If the pixel value is a maximum or minimum value among all compared pixels, it is a candidate key point. Figure 7.1 illustrate SIFT points extraction.

7.1.2 SIFT Application in salient Region Extraction

The SIFT algorithm locates points of interest for further processing. These points may signal the presence of objects or parts of objects in the image domain. A characteristic of SIFT feature points is scale, shift and rotation invariance.

The SIFT algorithm can be used to find candidate locations that are rich in information. This is done by shifting a rectangular region over the SIFT image and counting the SIFT points in the rectangle at each position. Regions with the highest count are regions of interest. The entropy in the high-interest regions is then calculated, and the regions

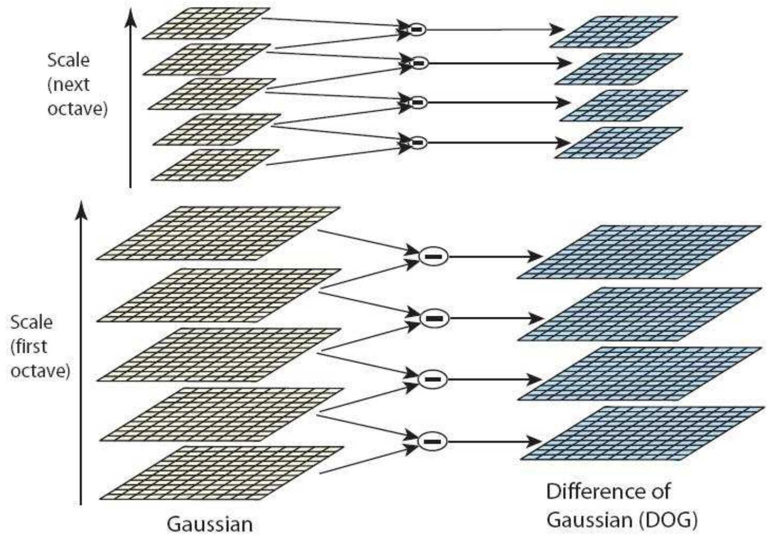


Figure 7.1: The illustration for SIFT algorithm

with the highest entropy are selected. These salient regions are then used in the next stage of processing. It is the combination of SIFT point density and high entropy that identifies salient regions. More SIFT feature points and higher entropy increase the possibility that a region contains valuable spatial information for the MMI registration algorithm. The SIFT process finds stable features, such as the buildings, cross-sections and landmarks, etc., and high entropy indicates that the region has enough intensity variation for a good MMI calculation.

After salient regions being extracted, then the next step is building salient region pair between images. These matching pair in this algorithm is built for the region. The number of salient regions is much less than the number of point candidates used in traditional feature-based algorithm. This property limit the time consumption to estimate the parameters effectively, which is main disadvantage of traditional feature-matching algorithm.

An example of salient region location is shown in Figure 7.4. A short-wave IR image is shown in (a). The SIFT point extrema are highlighted in (b). A Harris corner label image is shown in (c) and a SIFT point density map is shown in (d). The entropy is calculated in (c) for the regions with high point density in (d). The salient regions are locations where the entropy, weighted by the point density, exceeds a threshold. These are shown in (e). The HCL image is used in the entropy calculation because it is faster and consistent with the HCL image use in the following registration process. The entropy calculation could as well be done on the original image.

7.2 Image Registration by Salient Region

There are three stages to the registration process for a pair of images. The first stage identifies salient regions in each image. This is done independently on the two images. The second stage finds matching salient regions between the two images. This provides a rough match and constrains the search space for the parameters of the match transform. The third stage refines the match by applying label-enhanced MMI to the paired salient regions and, optionally, to the overlap regions of the full images. The first stage processing has been described in above section 7.1. It is simply applied to both images.

7.2.1 Stage 2: Salient Region Matching

The second stage requires a metric for the match of salient regions. It was found that an effective comparison can be made by computing the MMI surface between each possible pair of salient regions and then evaluating the kurtosis of the MMI surface. This will be illustrated with an example.

The two infrared images shown in Figure 7.3 (a-b) are successive images from the WASP airborne sensor. These images have both high and low detail regions. The first stage of the registration process produced the salient regions that are outlined in (c-d). Note that the some regions are likely to be found in both images, but they may be shifted,

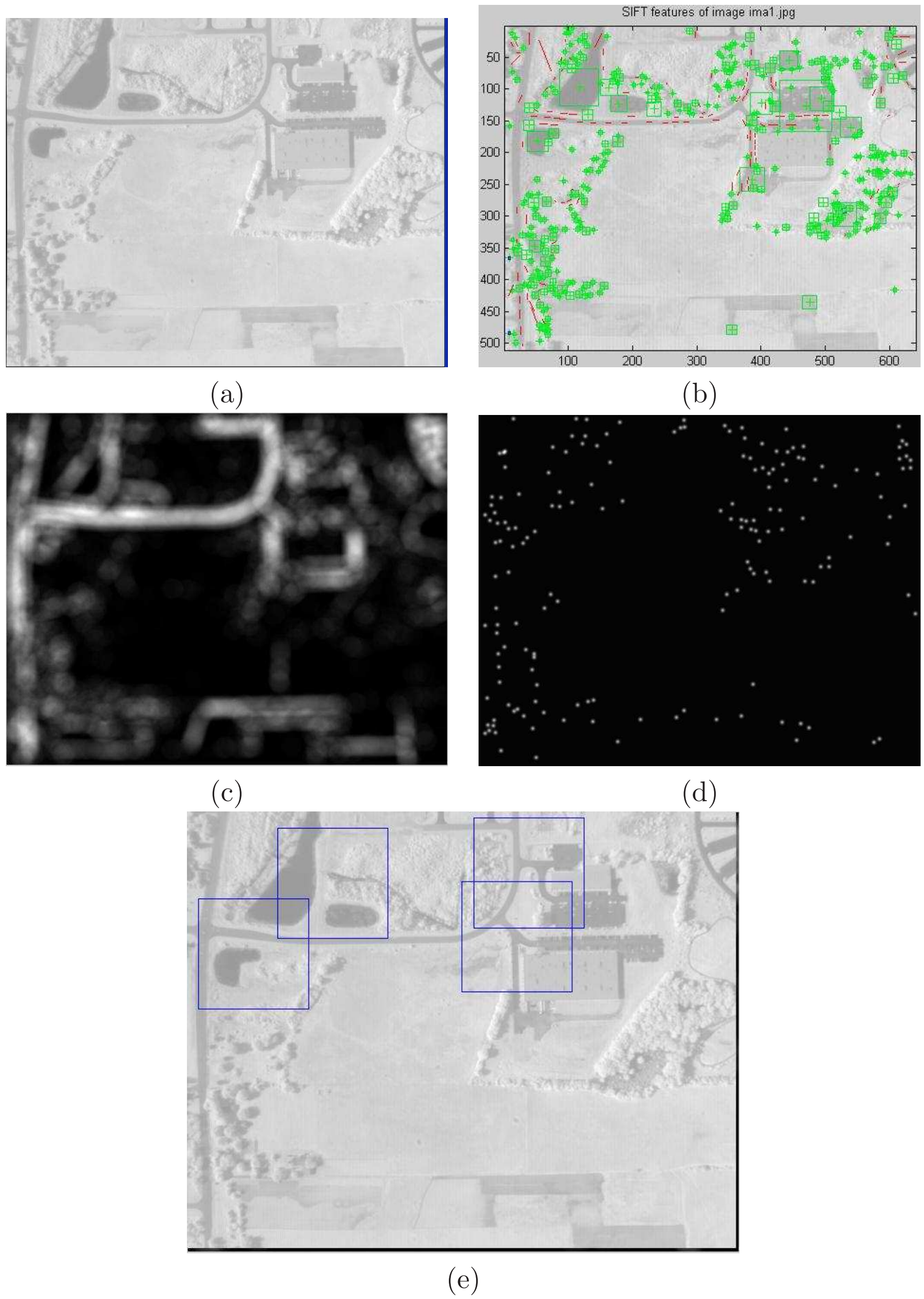


Figure 7.2: The salient region extraction: (a) An airborne image from WASP short wave band; (b) The SIFT points; (c) The distribution of local entropy; (d)The SIFT points distribution ; (e) The extracted best salient region.

and other regions are found in only one of the images.

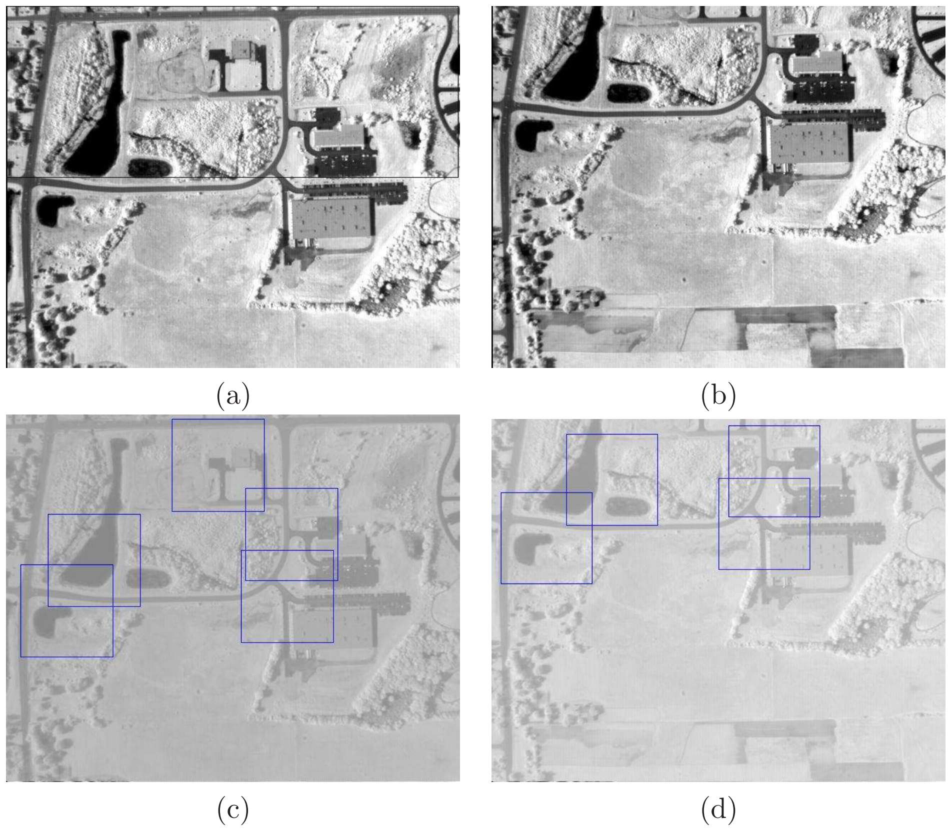


Figure 7.3: Experiment on WASP short wave band image: a) and (b) An airborne image from WASP short wave band; (c) and (d) The extracted salient region.

The MMI algorithm is then run on each pair of salient regions. This produces a MMI surface map for each pair, from which the kurtosis can be calculated. Five salient regions were found in Figure 7.3(c) and four in 7.3(d). The kurtosis for each pair is shown in Figure 7.5. The matching salient regions have a much higher kurtosis, as indicated by the highlighted values.

The MMI process also produces estimates of the transformation parameters between the matched regions, from which an initial estimate of the transformation can be produced.

If consistent parameter estimates are produced by two or more pairs then confidence in good matches is increased. If there were many match options, a RANSAC approach [124] could be used to find a good set of initial transform parameters.

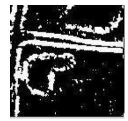
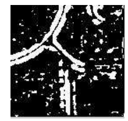
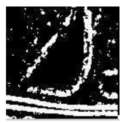
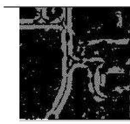
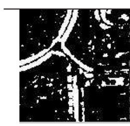
				
	**6.15**	2.37	1.21	4.57
	4.85	**37.34**	6.16	4.33
	2.12	5.43	**32.87**	3.74
	5.28	4.66	4.13	**27.56**
	3.32	3.58	2.94	4.11

Figure 7.4: The kurtosis of the registration between the salient region pairs.

Table 7.1: Image registration parameters from the salient region pairs.

	Shift on X (Pixels)	Shift on Y (Pixels)
Match Pair1	8.0	105
Match Pair2	9.0	103
Match Pair3	10	105
Match Pair4	11	103

The MMI search surfaces for the matched pairs from Figure 7.4 are shown in Figure 7.5. The kurtosis of each search surface corresponds to the highlighted value in Figure 7.4,

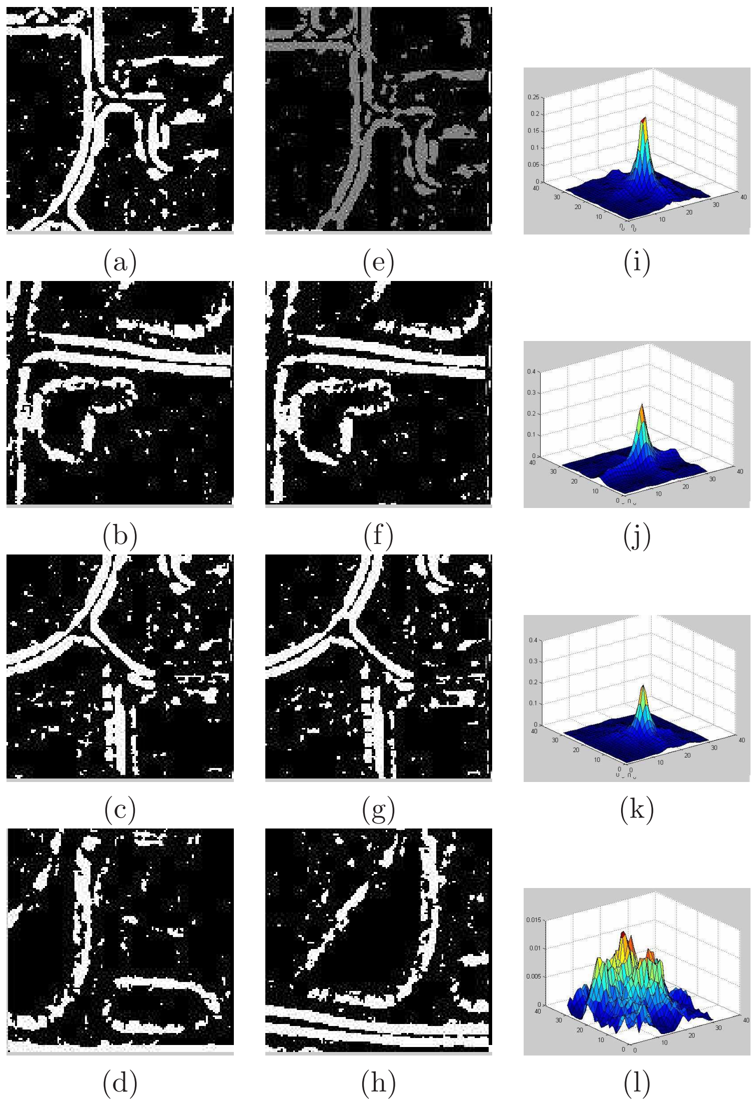


Figure 7.5: The registration pairs between the salient region: (a)-(d) The salient region in the Figure 7.3(a); (e)-(h) The corresponding pair in the Figure 7.3(b), (i)-(l) The corresponding search surface for each pair.

and it is seen that surfaces with a more pronounced peak have higher values for kurtosis.

7.2.2 Stage 2: Transform Parameters Estimation

If one or more salient region matches are obtained then the registration process can proceed to the third stage. The approximate transform parameters produced by the second stage can be used to identify the overlap regions of the images. The search region for refined transform parameters can be confined to a small volume in parameter space and the search can be carried out using a variety of MMI techniques. One option is to simply use conventional MMI between the images in the overlap region. However, it is more efficient and robust to use feature-enhanced MMI based on the Harris corner detector.The floating image can be either the original or the HCL map. These techniques were described in [120].

7.3 Experimental Results

This section will describe two experiments for registration of imagery. The first experiment uses infrared images from two different bands and the second uses a LIDAR image produced from intensity and elevation data with a cartographic map. The first experiment is challenging because the features appear quite differently in the two bands. The second is very challenging because both images are essentially derived products from other data to which arbitrary coloring and markup has been applied.

7.3.1 Registration of Multi-band IR

Two images produced by the WASP airborne sensing system are shown in Figure 7.6 (a) & (b). It is evident that the short wave image shows buildings or roads in more detail while the medium wave image is more sensitive to the temperature variation. The first stage of processing has identified the salient regions shown in Figure 7.6 (c) & (d). While the regions are not identical, it can be seen that there is at least on matching pair.

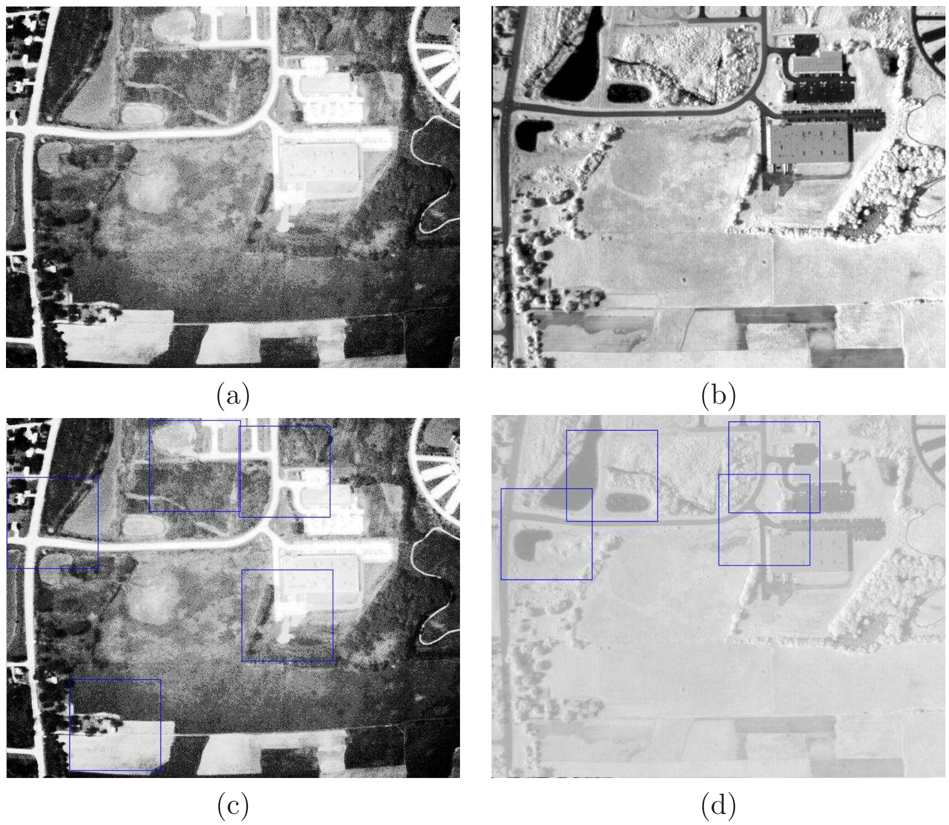


Figure 7.6: Experiment on WASP short wave band image and media wave band image: (a) and (b) An airborne image from WASP medium wave band and short wave band; (c) and (d) The extracted salient region.

The search for matching regions was carried out by the second stage of processing. Since there are four salient regions in each image, there are sixteen pairs to be tested. The test results are shown in Figure 7.7, where the kurtosis test finds one pair match. and the Figure 7.8 is the best match pair from the Figure 7.6 (c) & (d)

Based on the results from above matching pair, the two corresponding region from the middle-wave band image can be extracted, then the registration on those two pairs can be

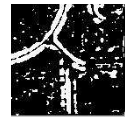
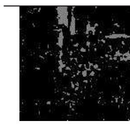
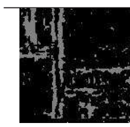
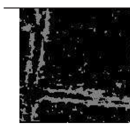
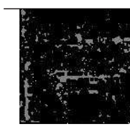
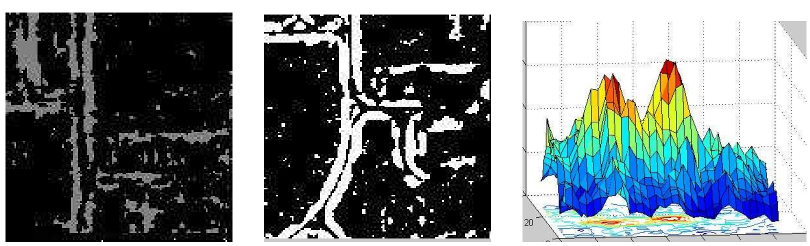
				
	1.33	1.37	3.21	3.75
	3.58	3.45	1.64	3.44
	1.12	**17.64**	2.78	2.47
	2.58	3.46	0.75	2.65
	4.23	3.44	4.29	1.14

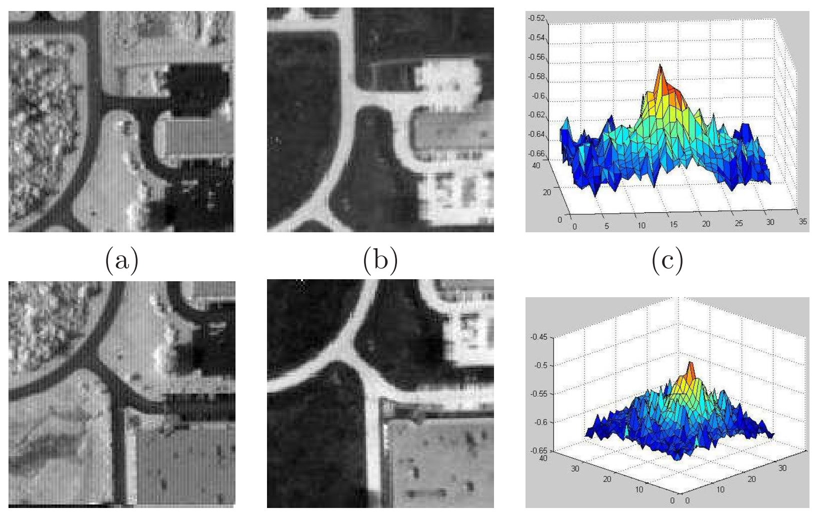
Figure 7.7: The kurtosis value of the registration between the salient region pairs.



(a) (b) (c)

Figure 7.8: Salient region registration on short wave band image and long medium band image: (a) Salient region in short wave band; (b)Salient region in medium wave band; (c) Registration search surface. (Kurtosis = 17.64).

completed (shown in the Figure 7.9) and individual parameter set can be obtained from each pair. All registration parameter sets were averaged and the registration error was reduced in theory. This method has special benefit for the applications without ground-truth data.



(a) (b) (c)

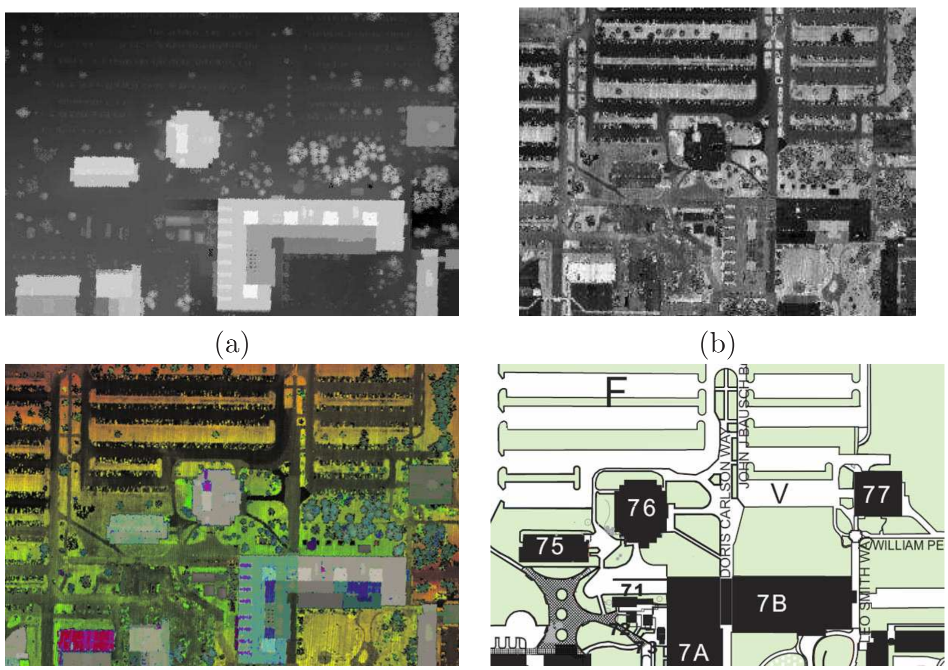
(d) (e) (f)

Figure 7.9: The registration pairs between the short wave band image and medium wave band image: a() and (d) The salient region in the Figure 7.6(d); (b) & (e) The corresponding region in medium-wave band image; (c) is the search surface of (a) and (b); (f) is the the search surface of (d) and (e).

7.3.2 LIDAR to Map Registration

This experiment is the registration of an image derived from LIDAR point cloud data and a cartographic map. The results demonstrate a method that can be extended to the task of fusing LIDAR data with imagery from various sources. The map image was chosen to demonstrate the method on a challenging problem. It has application in its own right as a method for relating archival maps to LIDAR collections for the purpose of map updating.

The LIDAR image was constructed from the point cloud by combining elevation (Figure 7.16 (a)) and intensity data (b). The result is a composite rasterized LIDAR image (CRLI) (c). We seek to register this image to a map such as that shown in (d). We can use any of the three LIDAR image products in the registration algorithm.

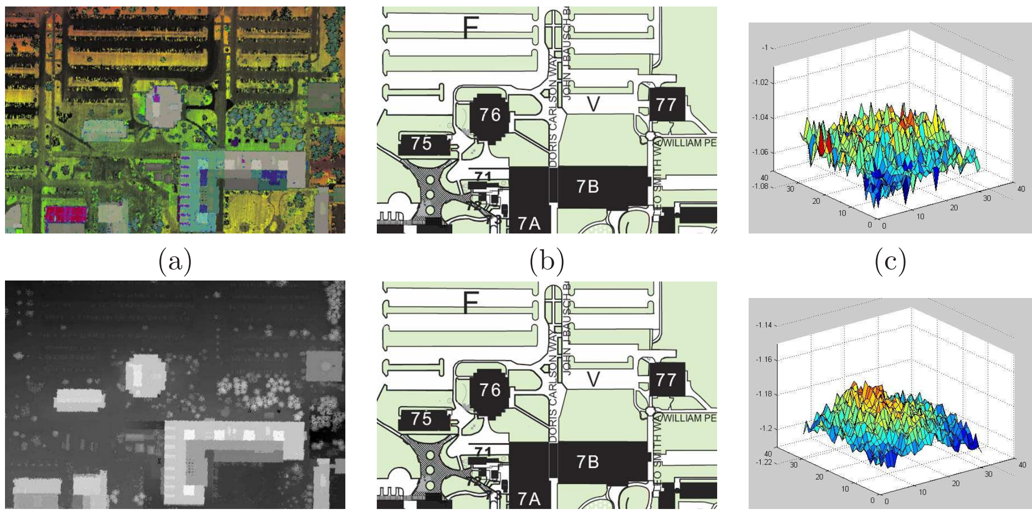


(a) (b)

(c) (d)

Figure 7.10: (a) Range image; (b)Intensity image; (c) CLRI image; (d) Campus map.

Maps are diagrams that show roads, buildings and symbolic information. The idea is to register these and similar images using the FE-MMI algorithm as well as with conventional MMI. All of the attempts were unsuccessful because of the large variation in texture and the incidental labeling in the map data. While it would be possible to construct special filters that use custom techniques on such images, the goal is to find a more general-purpose approach. The top row of Figure 11 shows an example of registration results using MMI for the map image with the CLRI image. The MMI search surface is very noisy without any indication of a peak. Use of the elevation with a map image using conventional MMI is shown in the bottom row. There is a slight indication of a registration peak, but it is very noisy and broad.



(d) (e) (f)

Figure 7.11: MMI registration result

Next attempt was to register the LIDAR and map images using the salient region method. The Figure 7.12 (c)&(d) shows the salient regions extracted with the SIFT algorithm. The two images do not have identical salient regions, but there is one that the two images share. The preliminary registration is achieved using this pair.

The initial registration result provided by matching salient regions can be used as the start point for stage 3. It was found that the best results were obtained using the elevation image shown in Figure 7.12. The search surface was now focused in the parameter space provided by the initial registration and is shown in Figure 7.14(c)) .

7.3.3 Performance Analysis

Assuming that the first two stages of the process provide a good starting point for the final registration step, one can expect to have a better chance for a good result. The quality of

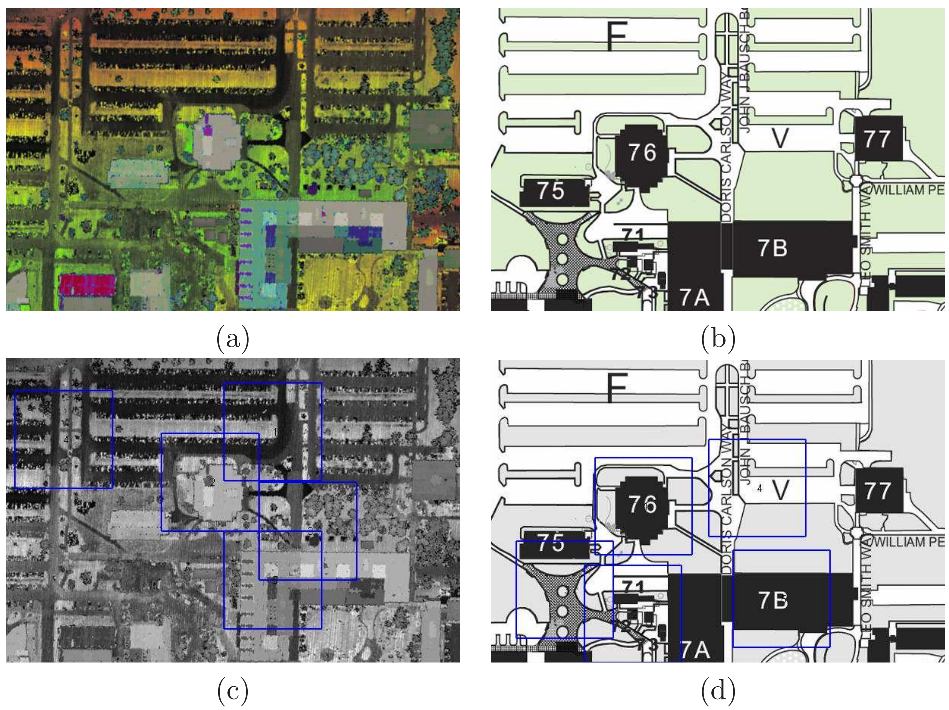


Figure 7.12: Registration on LADAR image and Map: a) Ladar image; b)Campus map; (c) and (d) The extracted salient region.

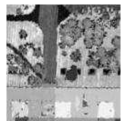
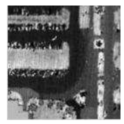
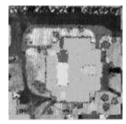
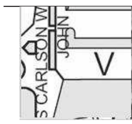
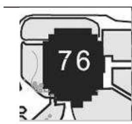
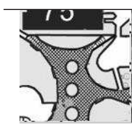
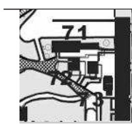
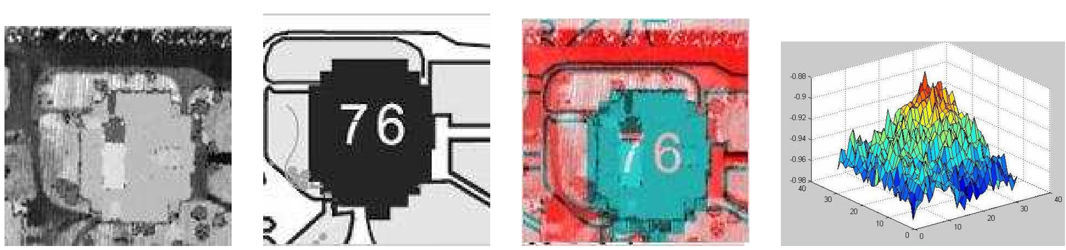
					
	0.13	2.20	1.23	1.67	2.79
	0.63	3.74	1.41	2.84	2.82
	4.51	0.92	**8.89**	3.78	3.57
	1.49	4.67	1.76	3.91	1.29
	3.83	1.82	5.91	2.59	4.75

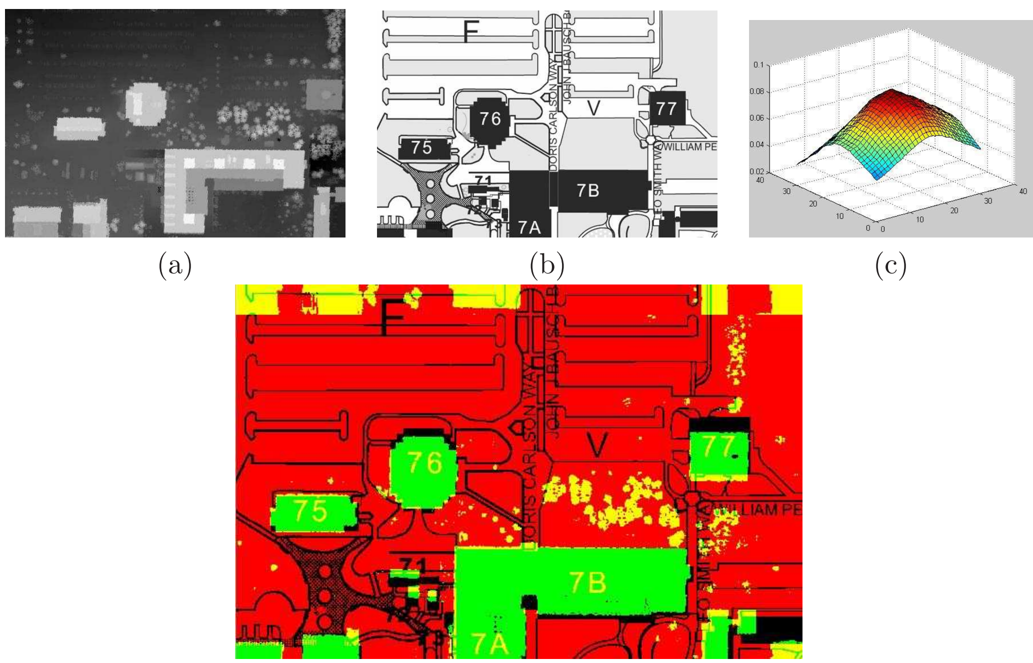
Figure 7.13: The kurtosis value of the registration between the salient region pairs.



(a) (b) (c) (d)

Figure 7.14: Salient region registration on short wave band image and long medium band image: (a) Salient region in map; (b)Salient region in Ladar image; (c) Registration result; (d) Registration search surface.

the final result will depend upon the characteristics of the algorithm used in stage 3. In this paper we are emphasizing the use of combined SIFT and MMI algorithms in stages 1 and 2 to provide a good starting point for the final registration and are not concentrating on refinement of the final step. This is an area where a number of algorithms can be used. In this section a conventional MMI algorithm and a feature enhanced MMI algorithm were used for the final registration step. One can compare the results to those obtained by manual registration as an indication of the performance of the system.



(a) (b) (c)

(d)

Figure 7.15: The refined registration result: (a) the height information in the LIDAR image Figure 7.12(a); (b) The campus map; (c) The corresponding search surface for each pair; (d) The overlapping image.

From the overlapping image (Figure 7.15(d)), it can be seen the alignment of the building 7A and 7B is reasonable, while there is the some registration error around the building 76. Although we don't have ground truth data, in this registration case we can select a building corner to make the quantitative error estimation. The error of building 77 area are in the Table 7.2. A mean squared geometrical error can be computed as:

$$e = \sqrt{\frac{1}{N}\sum_{N} ||G - T(I)||^2}$$

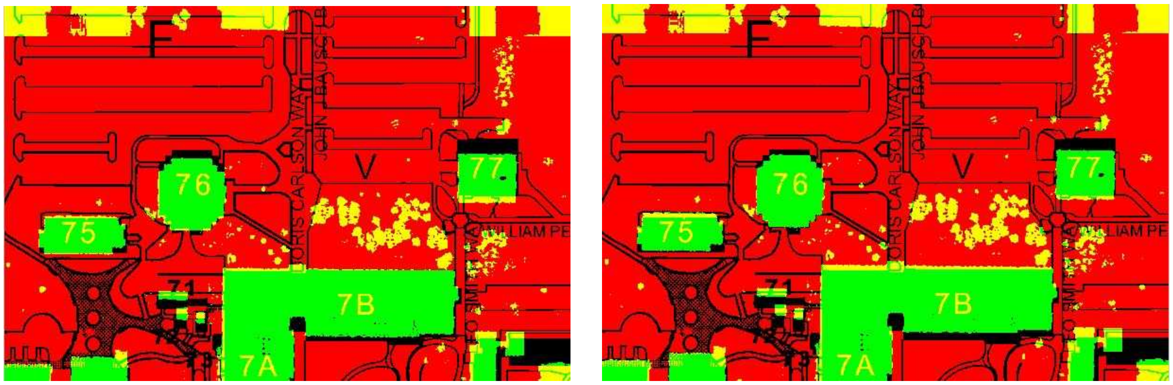
Where N is the number of sample points. The average geometrical error is about 9.4 pixels. A visual comparison can be seen in Figure 7.16.

Now the problem of quality evaluation of the registration is examined. Because of the lack of ground truth data, one alternative measurement is the comparison with manual registration. The feature pairs was selected manually and randomly divided into two parts: the first part is for registration and the second is for accuracy testing. The optimization method is to find the registration parameter set to minimize the Euclidian error, and then apply the registration parameters to the second part of the feature sets to measure the registration accuracy. A total of 100 repetitions were run for both registration algorithms, using different points selected randomly on each run, to get the average value. The average registration error was 7.3 pixels, which is a little better than the SIFT enhanced MMI algorithm. Comparing the registration results in Figure 7.16, both results were not perfect due to the difference in the image and map.

Table 7.2: The comparison of image registration result.

	Rotation	Shift on X (Pixels)	Shift on Y (Pixels)
Automatical Registration	0	37	-51
Manual Registration	-0.20	41	-45

The SIFT algorithm is the method to extract the salient region in this research. The proposed region extraction algorithm provide an alternative way to improve MMI/FE-MMI registration algorithm. Note that the algorithm is very flexible and one can choose proper method to make the full use of the information to fit different applications.



(a) (b)

Figure 7.16: The comparison of registration result: (a) The registration result from SIFT-MMI algorithm; (d) The manual registration by feature points.

7.4 Summary

This chapter described a novel salient region extraction for FE-MMI algorithm that combine the SIFT points and the local region entropy selection. The algorithm enables the FE-MMI algorithm to be applied on salient region, then the SIFT points density can be used to estimate the spatial information stability. Then information/entropy threshold of local region is measured to guaranty there is enough information for MMI registration, which can estimate the confidence for registration. The FE-MMI registration can be readily followed by salient region step to find the parameters.

The main limitation is in the assumption that at least one salient region will overlap and the registration pair can be built. In worst case the whole image can be used as the salient image. In these applications the algorithm can fall back to the conventional method.

Chapter 8

Experiments and Analysis

The previous chapters introduced the MMI-based registration algorithm and improvements for multi-modal image registration. This section presents experiments that compose conventional and improved MMI registration algorithms in real applications. Some parts of the experiments have been already used to demonstrate the algorithm in the previous chapters. In this section the complete experiments will be described.

8.1 Experiment I: Conventional MMI on WASP Images Registration

The first experiment is to test the conventional MMI-based registration algorithm on WASP multi-band images. The experiment contains nine WASP frames and each frame includes the short, medium and long wave band images 8.1. The registration tests both inter-band and intra-band parts. Part of results were shown in chapter 3 to demonstrate the flowchart and describe the MMI method. Note that Figure 8.1 shows that the short-wave images rotated 180^o relative to the medium-wave and long-wave band image due to the camera positions in the WASP system. This kind of error must be corrected in the pre-processing procedure.

Inter-band registration is used to overlap the different bands in each frame. Table 8.2 is the inter-band registration results. The short-wave band acts as reference image and the medium-wave and long-wave bands are registered to it. The parameters of scale and shift on X and Y are measured. The experiment shows the registration results from each frame are almost fixed because they depend primarily on the relative position and orientation of the cameras. The mean value of the results could be used as the internal parameters for WASP camera system to speed the registration.

The short-wave band is most sensitive to the grass/forest area. The short wave band is put into the green band because it contains more texture detail. The medium and long wave band are more sensitive to the thermal distribution and they tend to highlight buildings and roads. The image in Figure 8.2 are in pseudo-color and have been constructed by nine WASP frames.

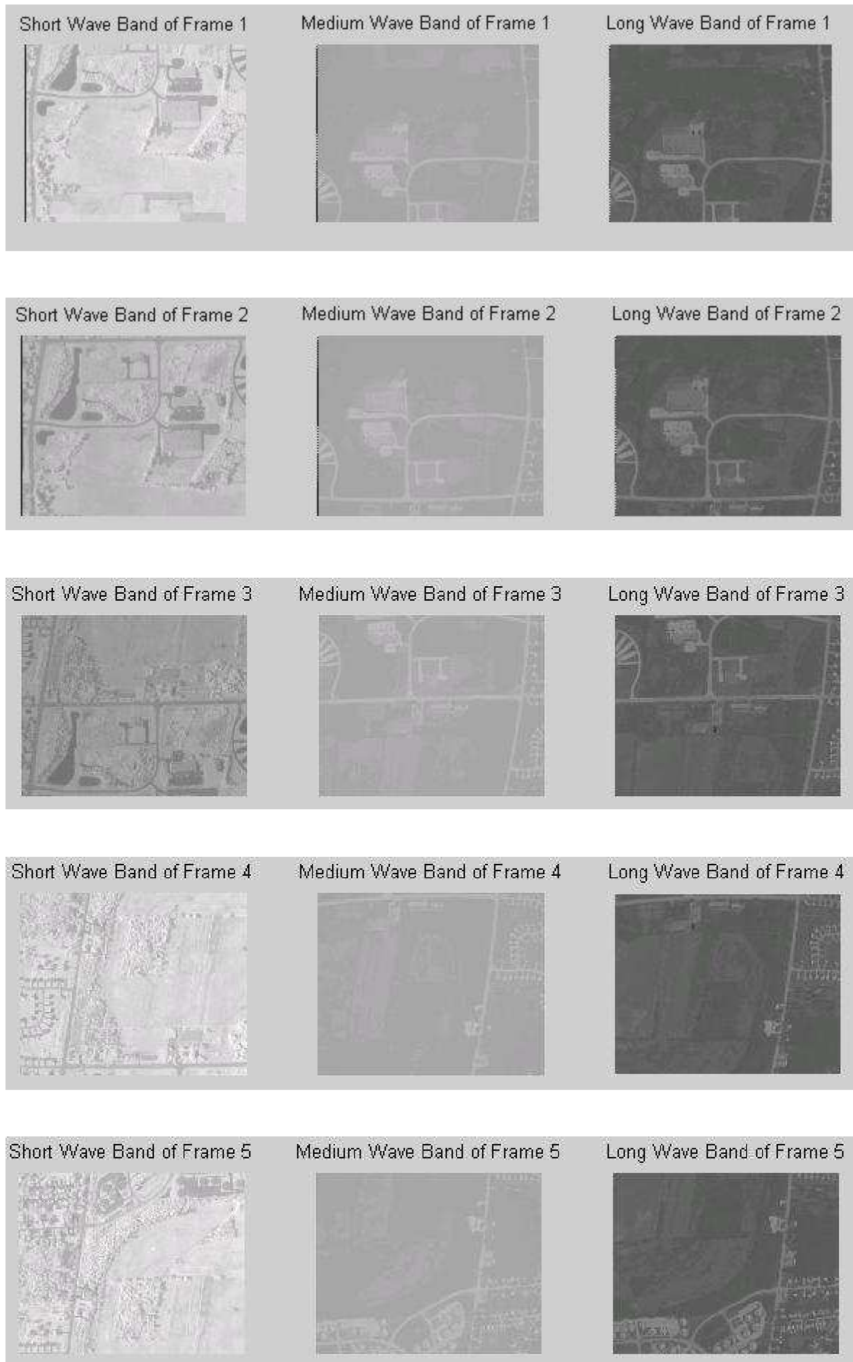
Another test of registration is intra-band image registration between the short-wave band image from different frames.

This experiment shows MMI algorithm works on intra-band and inter-band registration in most condition. Combined results from inter-band and intra-band registration can be used to build a mosaic image. This is shown in Figure 8.2.

8.2 Experiment II: FE-MMI algorithm on wild fire image registration

The last section demonstrates that the conventional MMI algorithm is a powerful method for multi-modal image registration. The MMI algorithm worked well with most WASP

WASP Frames:


Short Wave Band of Frame 1
Medium Wave Band of Frame 1
Long Wave Band of Frame 1
Short Wave Band of Frame 2
Medium Wave Band of Frame 2
Long Wave Band of Frame 2
Short Wave Band of Frame 3
Medium Wave Band of Frame 3
Long Wave Band of Frame 3
Short Wave Band of Frame 4
Medium Wave Band of Frame 4
Long Wave Band of Frame 4
Short Wave Band of Frame 5
Medium Wave Band of Frame 5
Long Wave Band of Frame 5

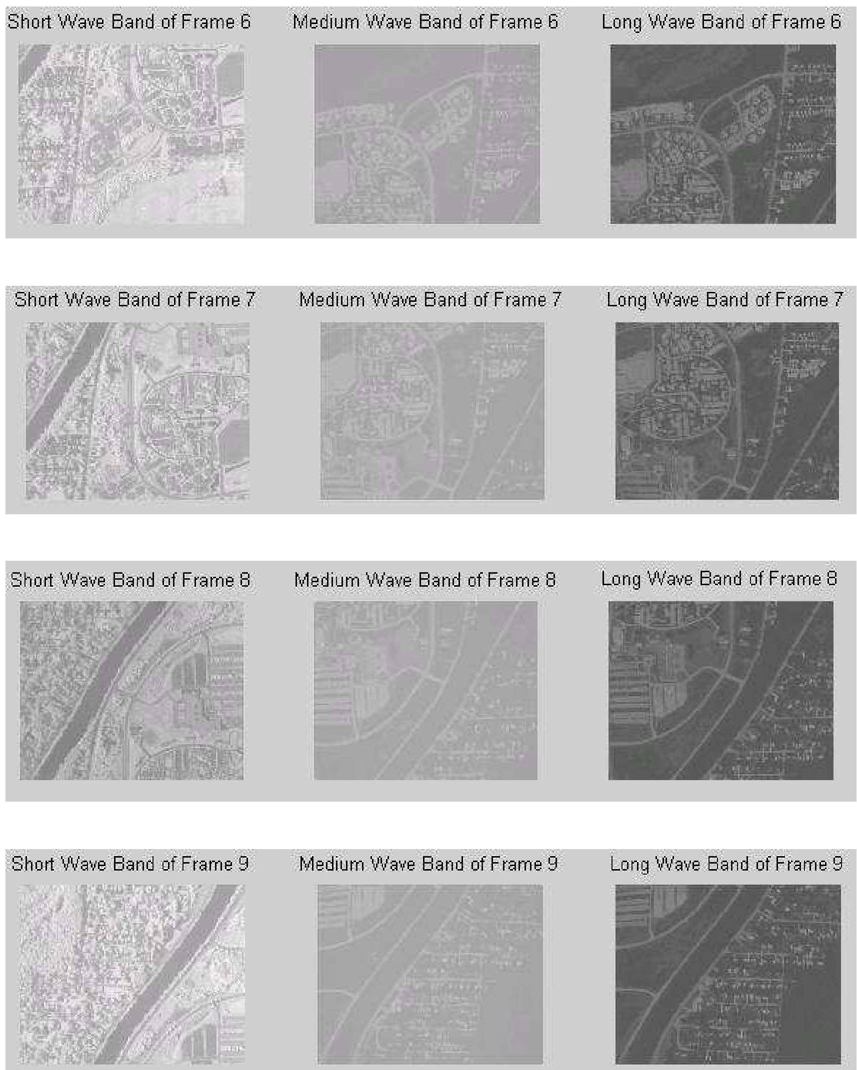
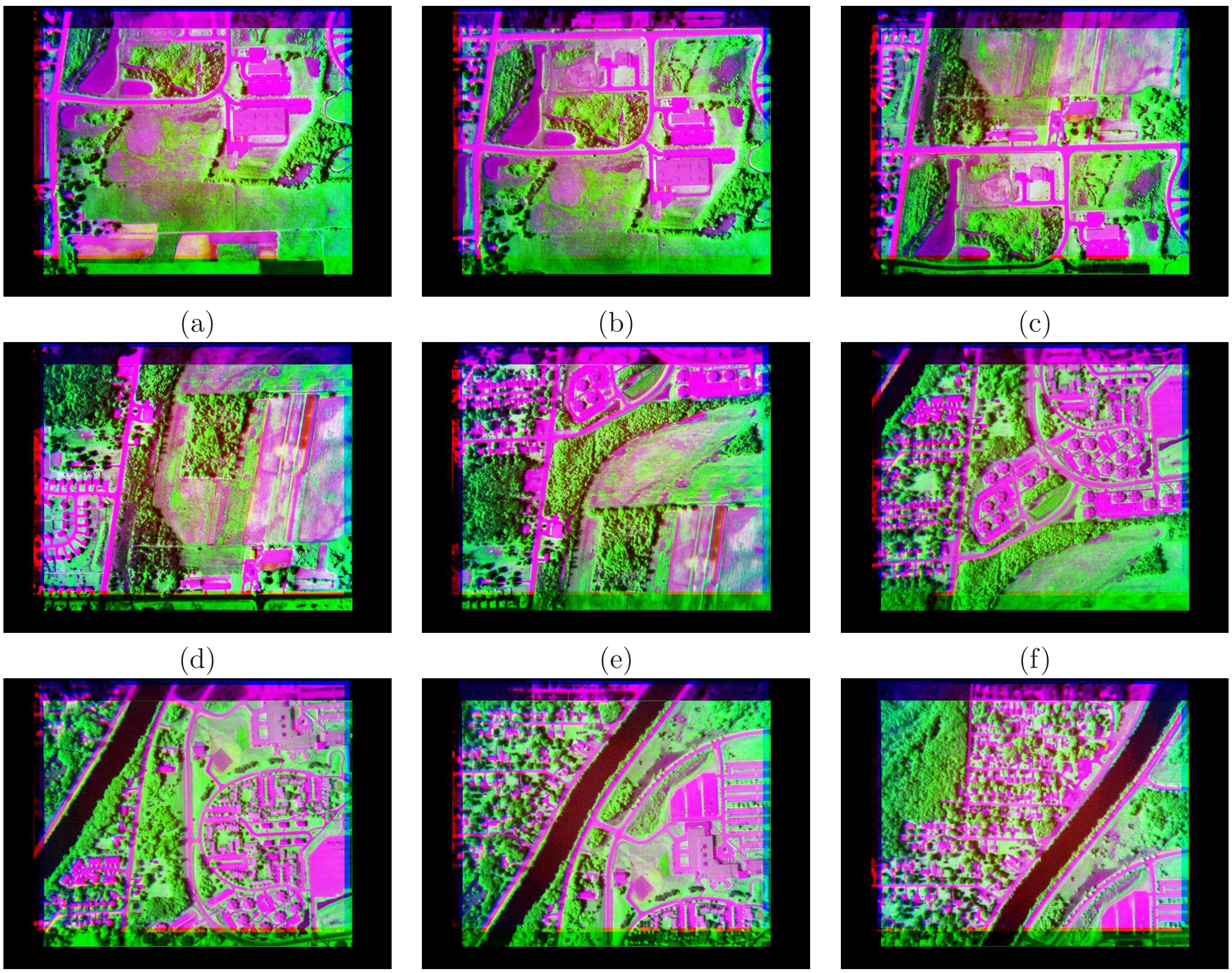


Figure 8.1: The comparison of the WASP frames. It includes three bands: short-wave, medium-wave and long-wave band.



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 8.2: WASP Inter-band registration Experiment: a) - i) are overlapping three-band images for each WASP frame.

Table 8.1: Inter-band registration results of WASP frames. (*High-light black represents mean values*).

i: Result of short-wave band vs. medium-wave band

Frame	Scale	X	Y
1	1.12	18	33
2	1.12	16	33
3	1.12	16	32
4	1.12	18	33
5	1.12	17	34
6	1.12	16	32
7	1.12	17	31
8	1.12	17	32
9	1.12	17	33
Mean	**1.12**	**17**	**33**

ii: Result of short-wave band vs. long-wave band

Frame	Scale	X	Y
1	1.12	5	40
2	1.12	5	40
3	1.12	5	39
4	1.12	4	41
5	1.12	5	40
6	1.12	5	40
7	1.12	6	40
8	1.12	5	40
9	1.12	5	40
Mean	**1.12**	**5**	**40**

Table 8.2: Intra-band registration results for short-wave band.

Frame	Rotation	X	Y
1 vs. 2	0	-10	-104
2 vs. 3	0	0	-112
3 vs. 4	0	-82	-121
4 vs. 5	-1.5	21	-235
5 vs. 6	-0.5	29	-249
6 vs. 7	0.5	7	-255
7 vs. 8	1.0	-30	-239
8 vs. 9	1.0	-30	-240

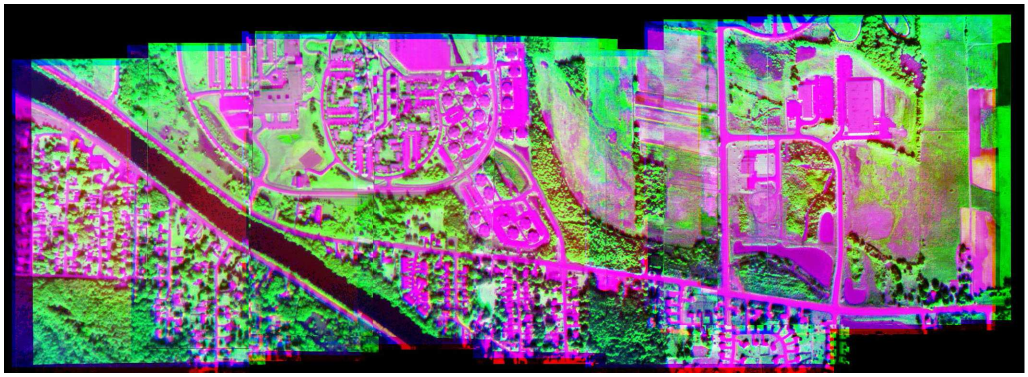
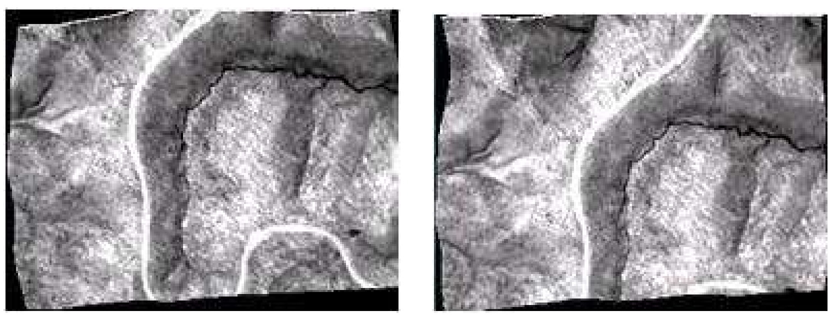


Figure 8.3: The mosaic image of the WASP frames

image registration tasks. However, in some experiments, the image difference are subtle especially in pixel intensity distribution (Shown in Figure 8.4).



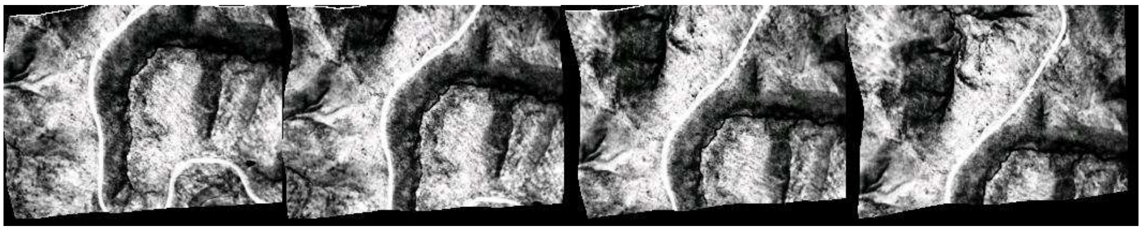
(a) (b)

Figure 8.4: (a) & (b) are two serial images.

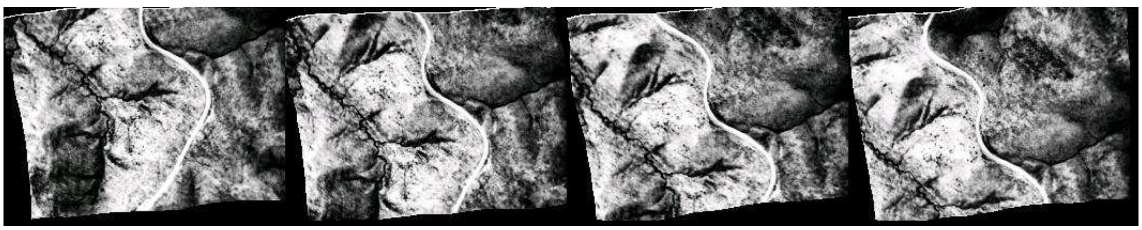
The performance of conventional MMI registration algorithm will decline because the mutual information between the sample sets couldn't reflect that change. Experiment shows that the ratio of accurate registration of MMI algorithm falls below 50%. Some registration failures of MMI algorithm are shown in Figure 8.6 on page 114.

To improve the registration, the research in this thesis extends the traditional MMI algorithm in two ways: First, extracting the spatial features by using the Harris corner detector; Second, achieving the MMI-based registration on the Harris corner label. The

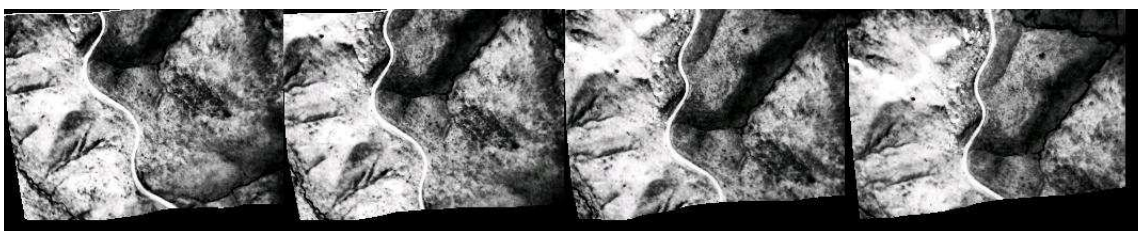
Wild Fire Frames:



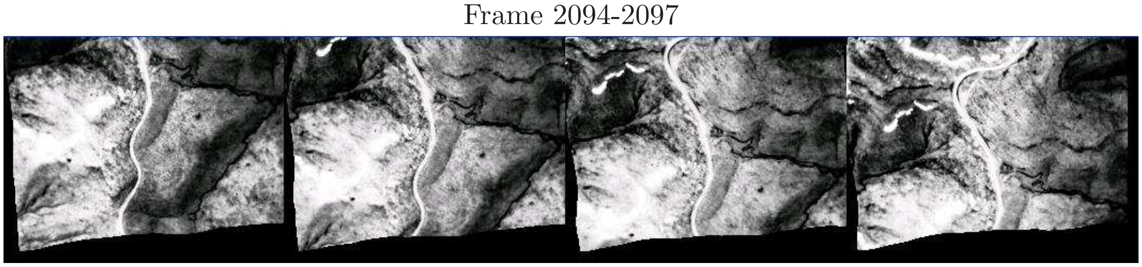
Frame 2086-2089



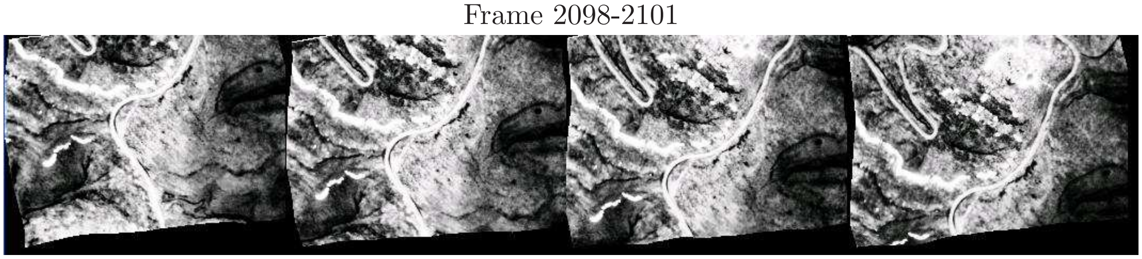
Frame 2090-2093



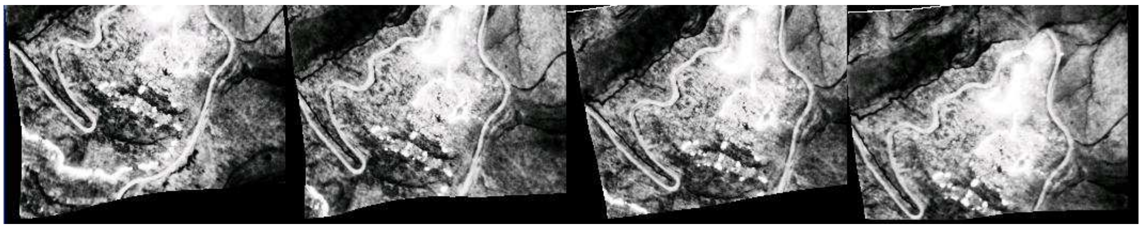
Frame 2094-2097



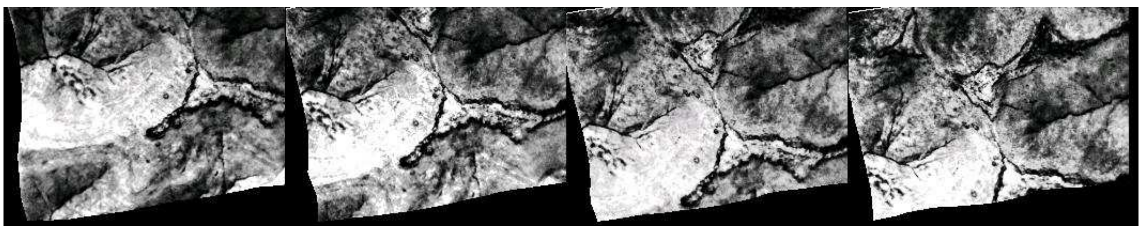
Frame 2098-2101



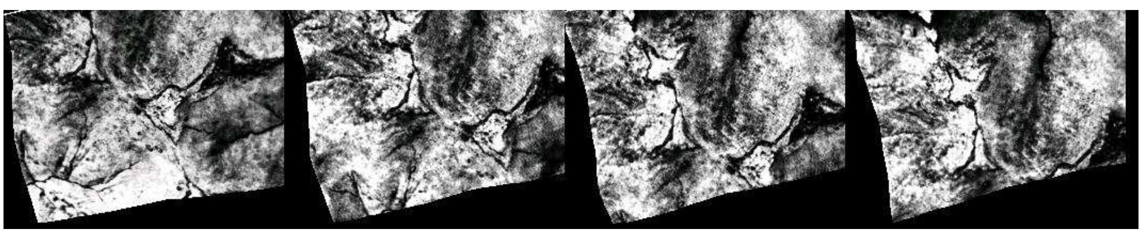
Frame 2102-2105



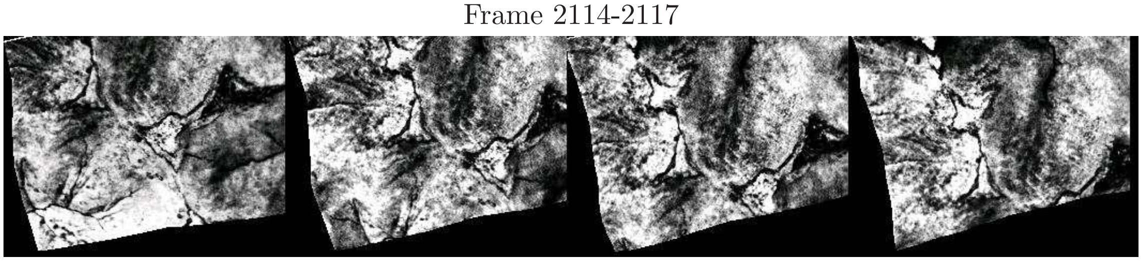
Frame 2106-2109



Frame 2110-2113

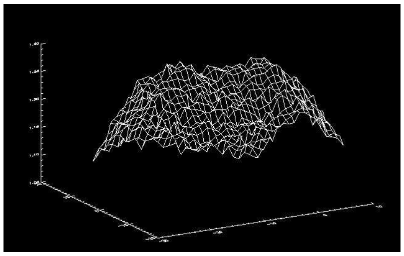


Frame 2114-2117

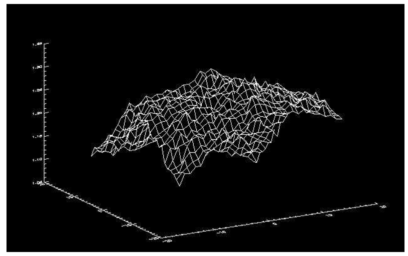


Frame 2118-2121

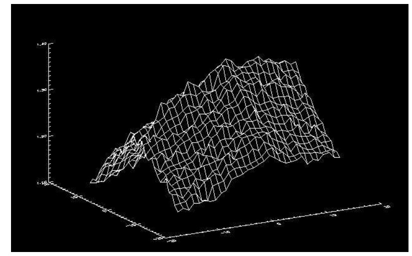
Figure 8.5: The wild fire frames. *Noted that the contrast is adjusted for presentation.*



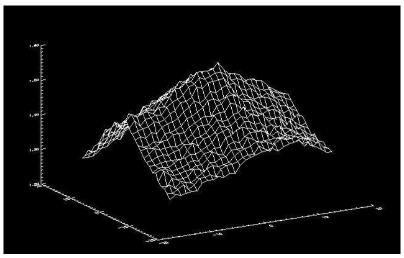
Frame 2086 vs. 2087



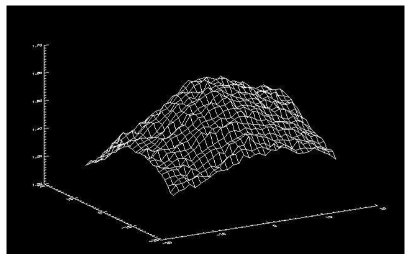
Frame 2087 vs. 2088



Frame 2090 vs. 2091



Frame 2094 vs. 2095



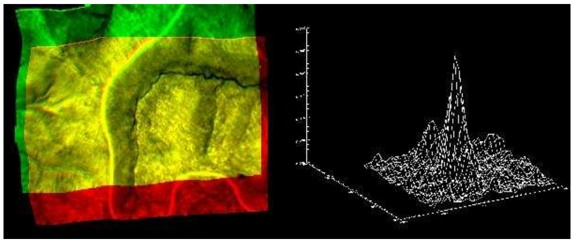
Frame 2099 vs. 2100

Figure 8.6: The registration failures of conventional MMI algorithm on wild fire image registration.

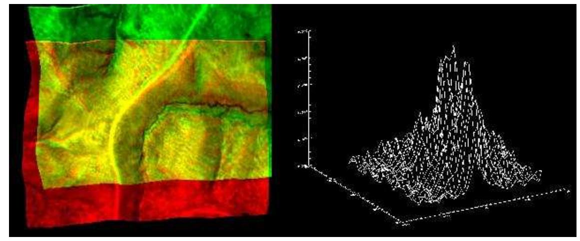
spatial features share more similarity and the mutual information registration is more robust. (discussed in page 49 in Chapter 4.

Another important enhancement is the optimization of the sample sets (discussed in Chapter 6). The sample sets is not randomly selected and the sample set is optimized by maximizing the entropy. Then the region of rich-in-information could be used fully and the stability of registration is highly improved.

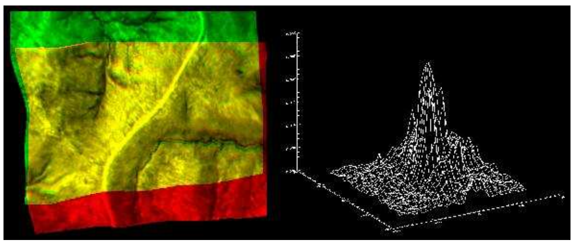
This experiment uses kurtosis to measure the registration quality (Chapter 5). The 80% kurtosis value is above 10. Those registration are achieved with high quality. Table 8.3 shows only two examples of registration kurtosis for the frame 2108 vs. 2109 and the frame 2113 vs. 2114 below 5, which is the threshold for registration acceptance. The ratio of correct registration is about 94%. This experiment result demonstrate the improvement of FE-MMI algorithm over conventional MMI algorithm.



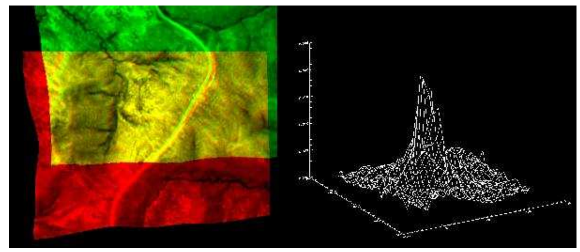
Frame 2086 vs. 2087



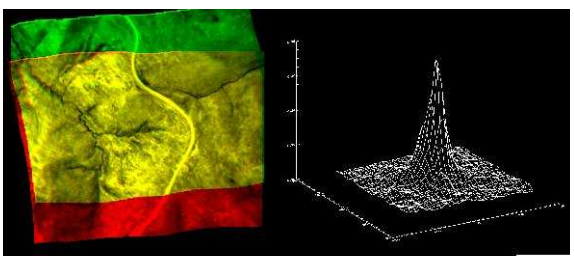
Frame 2087 vs. 2088



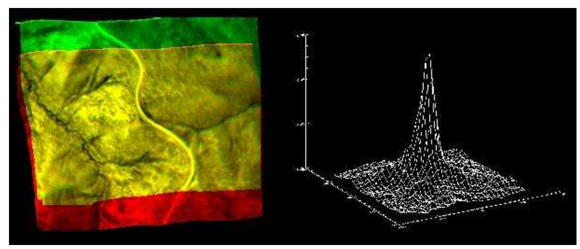
Frame 2088 vs. 2089



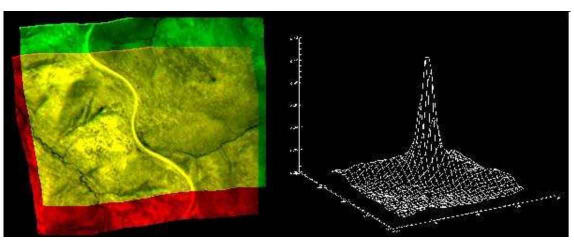
Frame 2089 vs. 2090



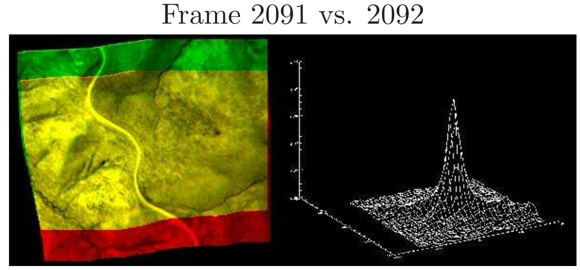
Frame 2090 vs. 2091



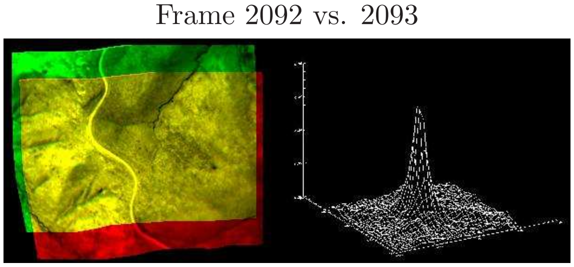
Frame 2091 vs. 2092



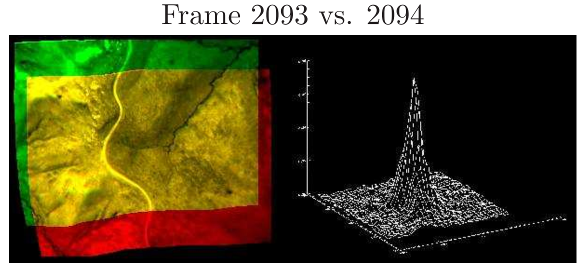
Frame 2092 vs. 2093



Frame 2093 vs. 2094



Frame 2094 vs. 2095



Frame 2095 vs. 2096

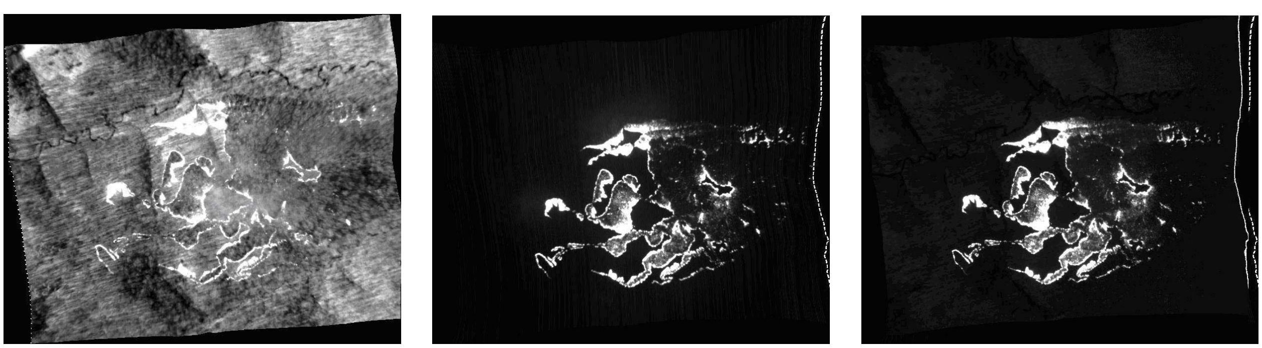
Figure 8.7: The performance of FE-MMI on wild fire image registration: the left is the overlapping image and the right is the search surface

Table 8.3: The registration results for wild fire images.

Frame	Rotation	X	Y	Kurtosis
2086 vs. 2087	-2	-9	35	41.17
2087 vs. 2088	1.5	3	38	7.825
2088 vs. 2089	-2.5	-4	32	12.25
2089 vs. 2090	2.5	8	74	9.71
2090 vs. 2091	-2.5	2	44	34.39
2091 vs. 2092	2.5	-2	28	27.52
2092 vs. 2093	0	8	31	32.38
2093 vs. 2094	-2.5	-3	33	35.15
2094 vs. 2095	0	-7	29	39.82
2095 vs. 2096	0	-14	32	31.93
2096 vs. 2097	1	-15	24	36.43
2097 vs. 2098	0	2	27	40.20
2098 vs. 2099	0	-2	36	24.078
2099 vs. 2100	0	4	32	36.82
2100 vs. 2101	0	-7	33	37.97
2101 vs. 2102	2.5	0	32	34.69
2102 vs. 2103	-1.5	5	34	40.97
2103 vs. 2104	-2.5	1	23	39.02
2104 vs. 2105	1.5	-4	31	40.02
2105 vs. 2106	2.5	-1	18	41.16
2106 vs. 2107	-2	8	41	38.86
2107 vs. 2108	2.5	-24	13	6.28
2108 vs. 2109	-2.5	23	31	4.58
2109 vs. 2110	-2.5	-26	30	26.97
2110 vs. 2111	2.5	-2	25	24.12
2111 vs. 2112	-2.5	-7	34	41.67
2112 vs. 2113	0	-44	17	39.46
2113 vs. 2114	2	10	29	2.77
2114 vs. 2115	2.5	30	24	7.11
2115 vs. 2116	0	0	34	32.79
2116 vs. 2117	-1.5	2	31	39.5
2117 vs. 2118	1	-26	20	49.04
2118 vs. 2119	1.5	-43	26	6.66
2119 vs. 2120	1.5	15	34	27.56
2120 vs. 2121	0	-20	19	37.09

8.3 Experiment III: SIFT enhanced MMI on wild fire Image registration

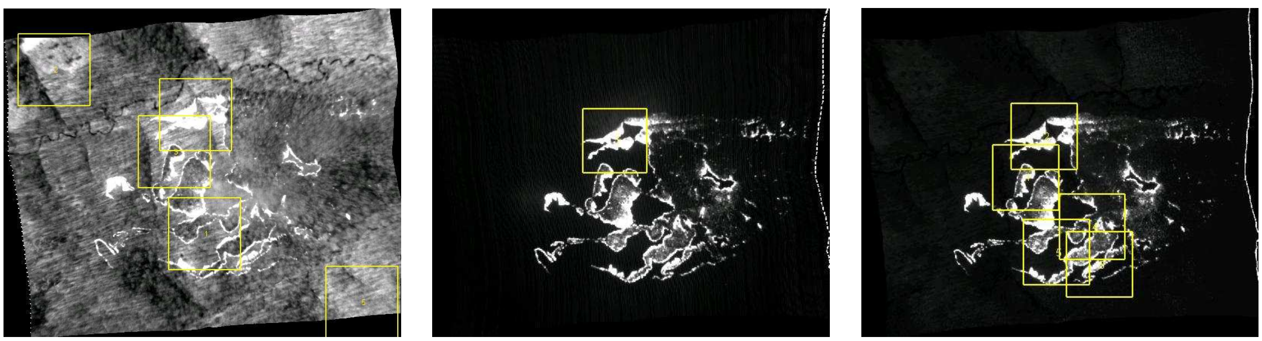
The difference of pixel intensity in different band is the inspiration for us to choose the MMI-based registration algorithm. As described before, due to different sensitivity, each sensor responds to scene elements in different ways, which affects the histogram distribution and degrades the conventional MMI performance. On the other hand, the noise will increase the fake corner points and augment the failure possibility of FE-MMI registration algorithm. For example, the Figure 8.8 shows three bands for wild fire images. The short wave band is brighter and catches more terrain detail; the medium band is more sensitive to thermal variation and the part of fire is very clear; the long wave band is very sensitive to thermal variation either, but still catches some terrain information in the dark areas. These variations will blur the statistical relationship between the pixel intensity, which is the base for the MMI-based registration. Meanwhile, the information-rich areas are different due to the sensor difference. Some regions will not match a counterpart in another band and will act like noise to harm the registration.



Short wave band   Medium wave band   Long wave band

Figure 8.8: The example of the various sensitivity of different sensors

Under such difficult conditions, the method can be improved by relating regions with common characteristics. A salient region detection and matching method is needed to find the matching pair. It has been found that Scale-invariant Feature Transform (SIFT) can be used to identify potential salient regions. A further region selection is provided by the region with enough spatial entropy. Then the registration is tested for each pair. The pairs for all salient regions that pass above tests are selected. The transform parameters for these pairs are used to initiate the registration for whole image.



Short wave band Medium wave band Long wave band

Figure 8.9: The salient region of the three bands

The Figure 8.10 shows the HCL maps of the best salient region pairs and corresponding search surface for each pair. The Table 8.4 shows the kurtosis value for each pair. All three search surfaces are capable of finding the parameters although the kurtosis of short-wave band and mid-wave band is much lower than other two. The Table shows the translation on X and Y directions for each pair.

Table 8.4: The kurtosis of the salient region pairs

	Short vs. Medium	Medium vs. Long	Short vs. Long
Kurtosis	5.23	14.72	25.57

The best match parameters for whole image can be calculated with the center of salient regions and the registration parameters for each pair of salient region. In this example, the long wave band image is chosen as reference image because the registration of long-

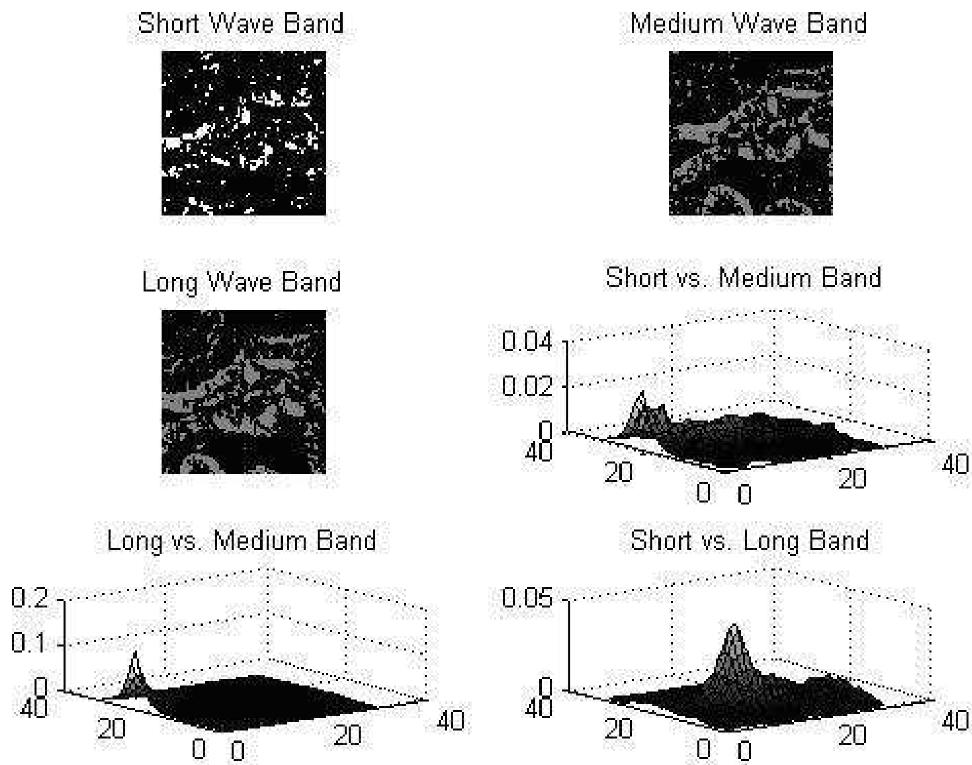
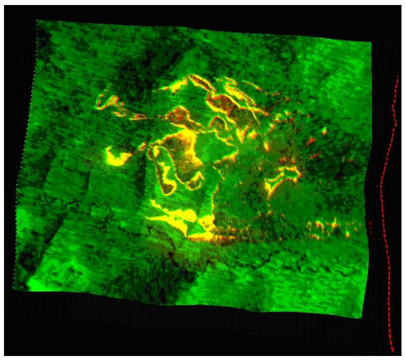


Figure 8.10: The salient region pair with best registration

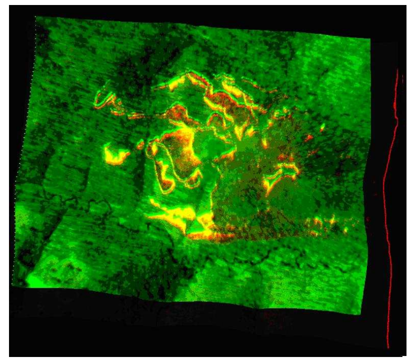
Table 8.5: The translation on X & Y coordinators of each pair

	Short vs. Medium	Medium vs. Long	Short vs. Long
Shift on X	-15	13	-3
Shift on Y	7	-6	0

wave band image with short-wave and medium-wave band images both get better kurtosis value. The Table 8.6 shows the final parameters for three bands and the corresponding registration result for each band is in Figure 8.11.



Short wave band vs. Medium wave band



Short wave band vs. Long wave band

Figure 8.11: The registration result for between two bands

Table 8.6: The translation parameters for registration

	Medium vs. Long	Short vs. Long
Shift on X	-9	3
Shift on Y	-16	33

Most of the fire regions overlapped. But, due to the sensors' differences, the alignment of the border isn't perfect, which make the images a little fuzzy. That's more obvious in overlapping image of the short vs. long wave band image. The final registration result is shown in Figure 8.12.

8.4 Experiment IV: The Registration Between a LWIR Image and Scene.

The most difficult registration task is the registration between the LWIR image and the scene image. LWIR image is more sensitive to thermal variation than the visible image.

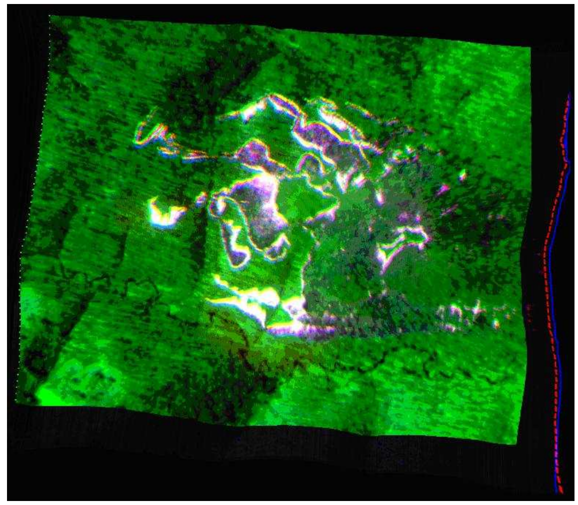


Figure 8.12: The overlapping image of three bands

The registration is valuable for assessing wildfire damage to forest and structures. The potential problems that can cause mis-registration are:

- The fire line is the good feature region for the long wave band image registration. However, it will cause disturbance of the registration because there is the lack of the corresponding information in the base image.
- Parts of rivers and roads are covered by the tree overhang and this increases the difficulty of registering the images with those features.
- The base image is full of forest texture which increases the risk of extracting invalid feature areas which decreases the robustness of the registration algorithm.

The above is a big challenge for automatic registration algorithms, and all automatic registration algorithms failed in these extreme cases in our experiments. To achieve the correct registration, manual pre-processing may be used to highlight such regions. The manual pre-pressing includes two steps: first we build the mask to eliminate the fire area.

This step is to get rid of the effect of the change caused by wild fire; the second step is to build the mask to help to locate valid information in the base image. This is in the spirit of using focus-of-attention to guide the MMI algorithm, albeit in the case with manual help.

This experiment shows that registration can be achieved by highlighting with a mask (Figure 8.13g)). Although in our experiment the mask was built manually, it could be generated automatically with the GIS data information.

8.5 Experiment V: The counterexample of FE-MMI algorithm.

The previous discussion and experiment shows the general improvement of FE-MMI algorithm on the traditional MMI algorithm. But under some special condition, MMI algorithm has better performance. The Figure 8.14 (a) and (b) are medium-wave and long-wave band images. The images are first to be corrected for radial distortion. Note that, there is a lot of vertical streaks in the medium-wave band, which leads to fake edge/corner areas. This will cause the HCL-MMI algorithm fail on registration.

Conventional MMI algorithm works better under such conditions. The mid-wave band image (Figure 8.14 (b)) is set as reference while long-wave band (Figure 8.14 (b) green) floating image. The rotation is 1 degree. The shift in X orientation is 24 pixels (to right) and 14 pixels down in Y direction. The registration result is shown in RGB color image, in which the mid-wave image is put into Red band while long-wave image being in Green band. The yellow regions represent the overlapping area while the green areas and red spots mean the unmatched areas (Shown by the brown arrows). The parameter search surface is shown in Figure 8.15(b). Due to the internal inter-band difference, the MMI

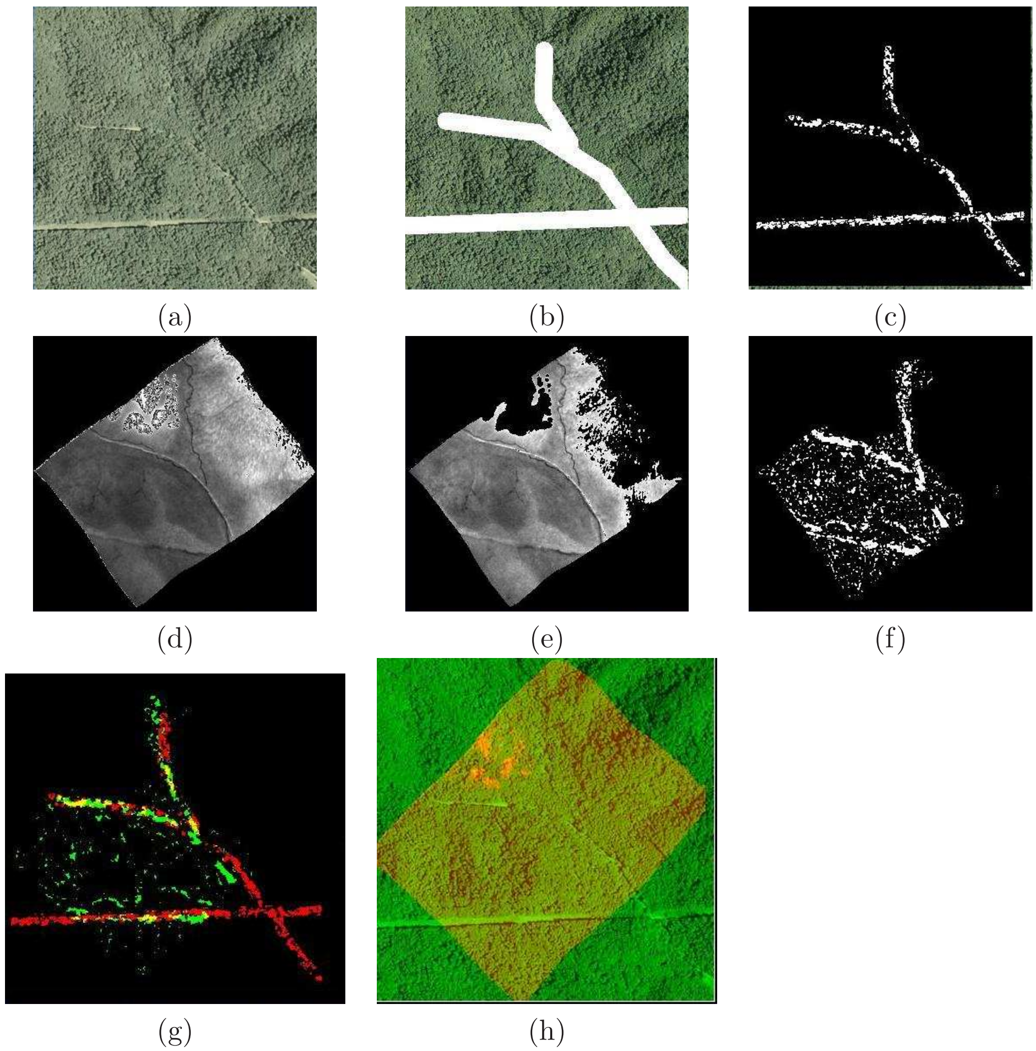
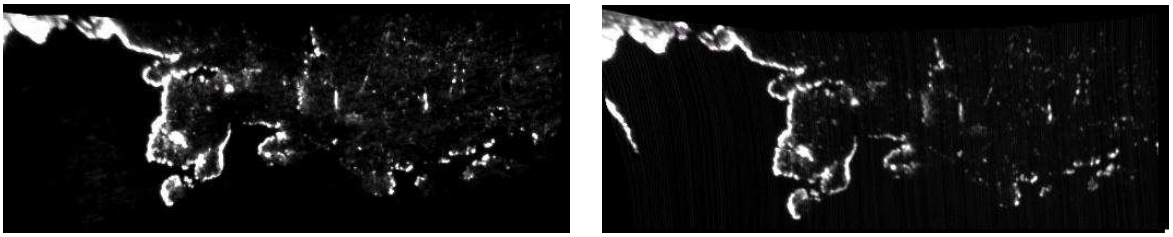


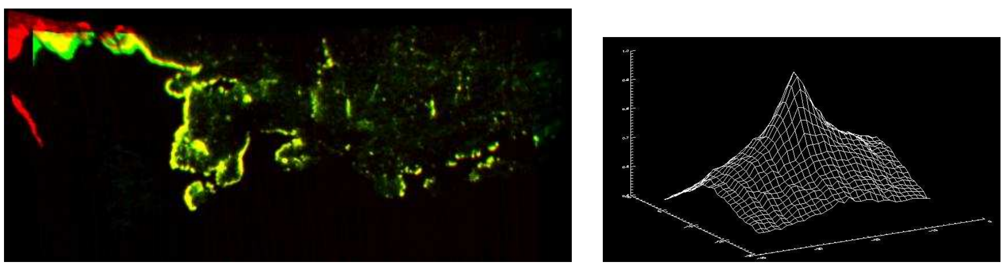
Figure 8.13: LWIR image registration: (a) & (b) represent original base image and mask; c) is corresponding Harris corner label map we used for registration; (d) & (e) are the LWIR image and mask; (f) is the Harris corner label map of (e); (g) is the registration result of (c) and (f); (h) is the overlap of base image and LWIR image by the registration parameter from (g)



Long wave band Medium wave band

Figure 8.14: The long-wave and medium-wave band of wild fire

registration result is acceptable.



Overlapping image MMI search surface

Figure 8.15: The long-wave and medium-wave band of wild fire. Notice the vertical streaks in the mid-wave band.

8.6 Summary

This chapter discusses the advantages and shortcomings of the conventional MMI, FE-MMI and SIFT improved MMI registration algorithms for real remote sensing applications. The MMI algorithm is a powerful tool for multi-band registration. FE-MMI is a good method if the Harris corner detector can extract information-rich areas correctly. SIFT enhanced MMI improves the registration by selecting regions with similar characteristics with sufficient detail.

There is no conflict among the three algorithms and one can be converted to the other easily. If the pixel intensity is used as a feature, FE-MMI is converted to conventional MMI. SIFT enhanced MMI is in the middle of the feature-based algorithm and statistical-based algorithm. The window-size is a critical factor for SIFT enhancement. If the window size is reduced to one pixel, the algorithm becomes the common feature-based algorithm; on the other hand, the algorithm will be conventional MMI if the whole image is the salient region. The choice can be decided to match the application and achieved the parameter adjustment.

Chapter 9

Conclusions & Future Research

This dissertation provides a solution for multi-modal airborne image registration. The MMI algorithm was originally designed for medical image registration. It developed and introduced a similarity measurement based on statistical information (probability) make it useful for multi-modal image registration because it can successfully avoid the negative effects of inter-modality intensity variation. The experimental results demonstrates that the MMI-based algorithm works on the registration of the infra-red images from different bands in WASP system.

However, the experiments also show that performance robustness, accuracy and efficiency need to be improved. False registration cases are common in flat search regions, and the registration error is often more than one or two pixels. The slow speed of MMI is another limitation. Solving those weaknesses is the motivation for our future research.

The previous versions of the MMI algorithm treat each point equally and select the samples randomly. However, from our visual experience, the points located on the edges or corners should play a more important role in registration. Those points contain more spatial information and should be emphasized for registration. The algorithm is improved this by focus-of-attention mechanism that selects rich-information locations for processing. Images can be divided into three categories with Harris corner detector.

There are best sample sets be constructed based on these classes. Furthermore it was found that synthetic images formed by the classification labels can be used effectively for the MMI registration, which leads to a significant increase in efficiency.

Another important extension is conversion of the MMI algorithm from the global based to local based. Conventional MMI chooses the sample set randomly across the whole image. The samples from non-overlap areas or high noise areas will impair the statistical relationship of between the two sample sets. That negative effect is eliminated or relieved by building the registration on extracted salient regions. Usually more than one salient region is extracted from each image, and the parameters from the best match pair are chosen as the starting point for further refinement. This initial registration can then be improved by a finer MMI search over a limited parameter space.

9.1 Conclusions

To sum up above discussion, we can draw the conclusion:

- MMI algorithm is able to register many examples of multi-modal image. The statistical relationship is a powerful similarity measurement;
- If mutual information for registration is changed from the pixel intensity to the spatial information (structure distribution) produced by classification with Harris corner detector, there is a substantial increase of robustness and accuracy;
- The measurement of the mutual information based on spatial features makes scene-and-graphic registration possible;

- The sample sets can be optimized by including the most information (maximizing the entropy), which can improve the performance in the same sample ratio, or reduce the sample ratio without the loss of accuracy;
- The salient region extraction converts the global region based algorithm to local region based. Such conversion is valuable for the registration when there are occlusions, noisy areas or low-overlap images;
- The calculation time can be greatly reduced by the applying the pyramid coarse-to-fine method and using classification images extracted from Harris corner detection instead of original pixel intensity. This is valuable for processing large data sets;
- The algorithm framework is modular. Each step could be applied to specific application requirements.

9.2 Limitation and Future Research

The main limitation of the proposed algorithm is in the assumption that the images include enough structure information for registration. If the Harris corner detector cannot identify enough information-rich pixels in the image, accuracy and stability will decline significantly. In this case, this technique may be extended with other feature detectors, which may be of interest in some applications, or "fall back" to the conventional MMI approach.

To keep the the generality of the algorithm, the common Harris Corner detector is chosen as classification filter in our research. However, the algorithm performance could be improved by application-specific filter design.

The algorithm can be adapted to a "pipeline" structure, which would be useful in applications such as real-time processing with moving platforms. The pipeline structure was not investigated here, but would be a natural extension in such applications.

Chapter 10

Publications

- MMI-based algorithm application on WASP image

  X. Fan, H. Rhody and E. Saber, 'Automatic Registration of Multi-Sensor Airborne Imagery', presented at the AIPR, Washington DC, 2005 [125].

- The study of search method

  X. Fan, H. Rhody and E. Saber, "A Comparison of Exhaustive Search vs. Gradient Search for Automatic Imagery Registration Based on MMI", presented at the WNYIP Workshop IEEE, Rochester, NY, 2006 [120].

- The extension of feature enhanced MMI algorithm

  X. Fan, H. Rhody and E. Saber, "A Harris Corner Label Enhanced MMI Algorithm for Multi-Model Airborne Image Registration", 2nd International Conference on Computer Vision Theory and Applications, Barcelona, 8-11 March, 2007 [126].

  X. Fan, H. Rhody and E. Saber, "An Algorithm for Automated Registration of Maps and Images Based on Feature Detection and Mutual Information", Electronic Image, SPIE/ISIT, San Jose, 27-30 Jan. 2008 [127].

- The analysis of registration performance

  X. Fan, H. Rhody and E. Saber, "A novel feature enhanced mmi based registration algorithm for automated maps and images", in 2008 IEEE International Geoscience & Remote Sensing Symposium, (Boston, MA), July 2008 [121].

  X. Fan, H. Rhody and E. Saber, "A spatial feature enhanced MMI Algorithm for multi-modal wild-fire image registration", presented at the AIPR, Washington DC, 2008 [128].

- The sift improved FE-HCL algorithm

X. Fan, H. Rhody, "The Novel Improvement for the HCL-MMI Multi-Modal Image Registration by SIFT Algorithm", Accepted and will be presented in IEEE Western NY image processing workshop at Sept. 27, 2009 [129].

- Journal Paper

  X. Fan, H. Rhody and E. Saber, "A spatial feature enhanced MMI algorithm for multi-modal airborne registration" IEEE, Transaction in Geo-science and remote sensing, submitted at 2008 and accepted at 2009 [130].

  S.R. Lach, J.P. Kerekes, and X. Fan, "Fusion of Multiple Image Types for the Creation of Radiometrically-Accurate Synthetic Scenes," Journal of Applied Remote Sensing, vol. 3, 033501, DOI: 10.1117/1.3075896, January 2009 [131].

Bibliography

[1] L. Fonseca and B. Manjunath, “Registration techniques for multisensor remotely sensed imagery,” *Photogrammetric Engineering and Remote Sensing*, vol. 62, pp. 1049–1056, 1996.

[2] E. Gulch, “Results of test on image matching of ISPRS WG,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 46, pp. 1–18, 1991.

[3] J. le Moigne, “First evaluation of automatic image registration methods,” in *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS 98*, pp. 315–317, 1998.

[4] D. Hill, P. Batchelor, M. Holden, and D. Hawkes, “Medical image registration,” *Physics in Medicine and Biology*, vol. 46, pp. R1–R45, 2001.

[5] S. Banerjee, D. Mukherjee, and D. Majumdar, “Point landmarks for registration of CT and NMR images,” in *Pattern Recognition Letters 16*, pp. 1033–1042, 1995.

[6] J. Maintz and M. Viergever, “A survey of medical image registration,” *Medical Image Analysis*, vol. 2, pp. 1–36, 1998.

[7] D. Fogel, “Image rectification with radial basis functions: Application to rs/gis data integration,” in *Proceedings of the Third International Conference on Integrating GIS and Environmental Modelling*, p. 19, 1996.

[8] B. Zitova and J. Flusser, “Image registration methods: a survey,” *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

[9] G. Stockman, S. Kopstein, and S. Benett, “Matching images to models for registration and object detection via clustering,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 4, pp. 229–241, 1982.

[10] L. Dreschler and H. Nagel, “Volumetric model and 3-d trajectory of a moving car derived from monocular tv-frame sequence of a street scene,” in *Proceedings of the Interantional Joint Conference on Artificial Intelligence, Vancouver, Canada*, pp. 692–697, 1981.

[11] H. Lester and S. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition*, vol. 32, pp. 129–149.

[12] B. Likar and F. Pernus, "Automatic extraction of corresponding points for the registration of medical images," *Medical Physics*, vol. 26, pp. 1678–1686, 1999.

[13] H. Li, B. Manjunath, and S. Mitra, "A contour-based approach to multisensor image registration," *IEEE Transactions on Image Processing*, vol. 4, pp. 320–334, 1995.

[14] I. Jolliffe, *Principal Component Analysis.* Springer Series in Statistics, second ed., 2002.

[15] P. Comon, "Computer vision, graphics and image processing," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[16] J. Maintz, P. van den Elsen, and M. Viergever, "Evaluation on ridge seeking operators for multimodality medical image matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 353–365, 1996.

[17] X. Dai and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 2351–2362.

[18] S. Abdelsayed, D. Ionescu, and D. Goodenough, "Matching and registration method for remote sensing images," in *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS 95, Florence, Italy*, pp. 1029–1031, 1995.

[19] J. Djamdji, A. Bajaoui, and R. Maniere, "Geometrical registration of images: the multiresolution approach," *Photogrammetric Engineering and Remote Sensing*, vol. 53, pp. 645–653, 1993.

[20] W. Ali and F. Cohen, "Registering coronal histological 2-d sections of a rat brain with coronal sections of a 3-d brain atlas using geometric curve invariants and b-spline representation," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 957–966, 1998.

[21] M. Davis, A. Khotanzad, D. Flaming, and S. Harms, "A physicsbased coordinate transformation for 3d image matching," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 317–328, 1997.

[22] J. Matas and O. Chum, "Local affine frames for widebaseline stereo," in *16th International Conference on Pattern Recognition ICPR 2002*, vol. 4, pp. 363–366, 2002.

[23] D. Bhattacharya and S. Sinha, "Invariance of stereo images via theory of complex moments," *Pattern Recognition*, vol. 30, pp. 1373–1386, 1997.

[24] P. Anuta, "Partial registration of multispectral and multitemporal digital imagery using fast fourier transform," *IEEE Transactions on Geoscience Electronics*, vol. 8, pp. 353–368, 1970.

[25] P. Besl and N. McKay, "A method for registration of 3d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, p. 239C254, 1992.

[26] J. Cheng and T. Huang, "Image registration by matching relational structures," *Pattern Recognition*, vol. 17, pp. 149–159, 1984.

[27] F. Cheng, "Point pattern matching algorithm invariant to geometrical transformation and distortion," in *Pattern Recognition Letters 17*, pp. 1429–1435, 1996.

[28] L. Fonseca and M. Costa, "Automatic registration of satellite images," in *Proceedings of the Brazilian Symposium on Computer Graphic and Image Processing, Brazil*, pp. 219–226, 1997.

[29] B. Ghaffary and A. Sawchuk, "A survey of new techniques for image registration and mapping," in *Proceedings of the SPIE: Applications of Digital Image Processing 432*, pp. 222–239, 1983.

[30] A. Goshtasby, "Registration of images with geometric distortions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, pp. 60–64, 1988.

[31] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *In Proceedings Visualization in Biomedical Computing VBC96, 1131, Springer Lecture Notes in Computer Science, Hamburg, Germany*, pp. 267–276, 1996.

[32] M. Audette and F. Ferrie, "An algorithmic overview ofsurface registration techniques for medical imaging," *Medical image Analysis*, no. 4, pp. 201–217, 2000.

[33] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 187–198, 1997.

[34] R. Sharma and M. Pavel, "Multisensor image registration," in *Proceedings of the Society for Information Display XXVIII*, pp. 951–954, 1997.

[35] X. Dai and S. Khorram, "Development of a feature-based approach to automated image registration for multitemporal and multisensor remotely sensed imagery," in *Pattern Recognition Letters 18*, pp. 243–245, 1997.

[36] C. Shekhar, V. Govindu, and R. Chellapa, "Multisensor image registration by feature consensus," *Pattern Recognition*, vol. 32, pp. 39–52.

[37] S. Nelson, M. Day, P. Buffone, L. Wald, T. Budinger, R. Hawkins, W. Dillon, S. Huhn, M. Prados, S. Chang, and D. Vigneron, "Alignment of volume mri and high resolution f-18 flurodeoxyglucose pet images for evaluation of patients with brain tumors," *Journal of Computed Assisted Tomography*, vol. 21, pp. 183–191, 1997.

[38] T. Lehmann, "A two stage algorithm for model-based registration of medical images," in *Proceedings of the Interantional Conference on Pattern Recognition ICPR 98, Brisbane, Australia*, pp. 344–352, 1998.

[39] J. Liu, B. Vemuri, and J. Marroquin, "Local frequency representations for robust multimodal image registration," *IEEE Transaction on Medical Imaging*, vol. 21, pp. 462–469, 2002.

[40] J. le Moigne, "Parallel registratin of multi-sensor remotely sensed imagery using wavelet coefficients," in *Proceedings of the SPIE: Wavelet Applications, Orlando, Florida, 2242*, p. 432C443, 1994.

[41] K. Rohr, H. Stiehl, R. Sprengel, W. Beil, T. Buzug, J. Weese, and M. Kuhn, "Nonrigid registration of medical images based on anatomical point landmarks and approximating thin-plate splines," in *Proceedings of the Aacheren Workshop: Bildverarbeiterung fur die Medizin, Aachen, Germany*, pp. 41–46, 1996.

[42] M. Roux, "Automatic registration of spot images and digitized maps," in *Proceedings of the IEEE International Conference on Image Processing ICIP 96, Lausanne, Switzerland*, pp. 625–628, 1996.

[43] V. Govindu, C. Shekhar, and R. Chellapa, "Using geometric properties for correspondence-less image alignment," in *Proceedings of the International Conference on Pattern Recognition ICPR98, Brisbane, Australia*, pp. 37–41, 1998.

[44] D. Shin, J. Pollard, and J. Muller, "Accurate geometric correction of atsr images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 997–1006, 1997.

[45] N. Vujovic and D. Brzakovic, "Establishing the correspondence between control points in pairs of mammographic images," *IEEE Transactions on Image Processing*, vol. 6, pp. 1388–1399, 1997.

[46] A. Goshtasby, G. Stockman, and C. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 24, pp. 390–399, 1986.

[47] M. Holm, "Towards automatic rectification of satellite images using feature based matching," in *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS 91*, pp. 2439–2442, 1991.

[48] Y. Hsieh, D. McKeown, and F. Perlant, "Performance evaluation of scene registration and stereo matching for cartographic feature extraction," in *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, pp. 214–237, 1992.

[49] S. Moss and E. Hancock, "Multiple line-template matching with em algorithm," in *International Geoscience and Remote Sensing Symposium IGARSS97, Singapore*, pp. 1283–1292, 1997.

[50] A. Vasileisky, B. Zhukov, and M. Berger, "Automated image coregistration based on linear feature recognition," in *Proceedings of the Second Conference Fusion of Earth Data, Sophia Antipolis*, pp. 59–66, 1998.

[51] D. Lavine, B. Lambird, and L. Kanal, "Recognition of spatial point patterns," *Pattern Recognition*, vol. 16, pp. 289–295, 1983.

[52] J. Ton and A. Jain, "Registering landsat images by point matching," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, pp. 642–651, 1989.

[53] M. Sester, H. Hild, and D. Fritsch, "Definition of ground control features for image registration using gis data," in *Proceedings of the Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels*, p. 7, 1998.

[54] S. Li, J. Kittler, and M. Petrou, "Matching and recognition of road networks from aerial images," in *Proceedings of the Second European Conference on Computer Vision ECCV'92*, pp. 857–861, 1992.

[55] J. Flusser and T. Suk, "Degraded image analysis: an invariant approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 590–603, 1998.

[56] J. Flusser and B. Zitova, "Combined invariants to linear filtering and rotation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, pp. 1123–1136.

[57] L. Ding, A. Goshtasby, and M. Satter, "Volume image registration by template matching," *Image and Vision Computing*, vol. 19, pp. 821–832, 2001.

[58] S. Chang, F. Cheng, W. Hsu, and G. Wu, "Fast algorithm for point pattern matching: Invariant to translations, rotations and scale changes," *Pattern Recognition*, vol. 30, pp. 311–320, 1997.

[59] R. Huseby, O. Halck, and R. Solberg, "A model-based approach for geometrical correction of optical satellite images," in *Proceedings of the International Geoscience Remote Sensing Symposium IGARSS 99*, (Hamburg, Germany), pp. 330–332, 1999.

[60] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 29, pp. 1153–1160, 1981.

[61] N. Dodgson, "Quadratic interpolation for image resampling," *IEEE Transactions on Image Processing*, vol. 6, pp. 1322–1326, 1997.

[62] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 26, pp. 508–517, 1978.

[63] T. Lehmann, C. Gonner, and K. Spitzer, "Addendum: B-spline interpolation in medical image processing," *IEEE Transaction on Medical Imaging*, vol. 20, pp. 660–665, 2001.

[64] T. Lehmann, C. G. onner, and K. Spitzer, "Survey: interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 1049–1075, 1999.

[65] G. Grevera and J. Udupa, "An objective comparison of 3d image interpolation methods," *IEEE Transactions an Medical Imaging*, vol. 17, pp. 642–652, 1998.

[66] P. Hellier and C. Barillot, "Coupling dense and landmark-based approaches for non rigid registration," *IRISA research report, PI 1368:30, France*, 2000.

[67] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI98), Lecture Notes in Computer Science*, vol. 1496, (Cambridge, MA), p. 1115C1124, 1998.

[68] N. Ritter and R. Owens, "Registration of stereo and temporal images of the retina," *IEEE Trans. on Medical Imaging*, vol. 18, pp. 404–418, 1999.

[69] B. Rezaie and M. Srinath, "Algorithms for fast image registration," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, pp. 716–728, 1984.

[70] W. Pratt, *Digital Image Processing.* New York: Wiley, 2nd ed., 1991.

[71] R. Bracewell, *The Fourier Transform and Its Applications.* New York: McGraw-Hill, 1965.

[72] H. Hanaizumi and S. Fujimura, "An automated method for registration of satellite remote sensing images," in *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS 93, Tokyo, Japan*, pp. 1348–1350, 1993.

[73] R. Berthilsson, "Affine correlation," in *Proceedings of the International Conference on Pattern Recognition ICPR98, Brisbane, Australia*, pp. 1458–1461, 1998.

[74] A. Simper, "Correcting general band-to-band misregistrations," in *Proceedings of the IEEE International Conference on Image Processing ICIP 96, Lausanne, Switzerland*, pp. 597–600, 1996.

[75] D. Barnea and H. Silverman, "A class of algorithms for fast digital image registration," *IEEE Transactions on Computing*, vol. 21, pp. 179–186, 1972.

[76] G. Wolberg and S. Zokai, "Image registration for perspective deformation recovery," in *SPIE Conference on Automatic Target Recognition X, Orlando, Florida, USA*, p. p. 12, 2000.

[77] E. Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 700–703, 1987.

[78] Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of fourier-mellin transform for image registration and recognition," *IEEE Transactions on Pattern Analysis and Machine Intellingence*, vol. 16, pp. 1156–1168, 1994.

[79] B. Reddy and B. Chatterji, "An fft-based technique for translation, rotation and scale-invariant image registration," *IEEE Transactions on Image Processing*, vol. 5, pp. 1266–1271, 1996.

[80] F. Zana and J. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and hough transform," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 419–428.

[81] P. Thvenaz and M. Unser, "Spline pyramids for inter-modal image registration using mutual information," in *Proceedings of SPIE: Wavelet Applications in Signal and Image Processing, San Diego, CA*, pp. 236–247, 1997.

[82] P. Thevenaz and M. Unser, "An efficient mutual information optimizer for multiresolution image registration," in *Proceedings of the IEEE International Conference on Image Processing ICIP 98, Chicago, I*, pp. 833–837, 1998.

[83] B. Likar and F. Pernus, "A hierarchical approach to elastic registration based on mutual information," in *Image and Vision Computing*, vol. 19, pp. 33–44.

[84] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognition*, vol. 32, pp. 71–86.

[85] A. Rangarajan, H. Chui, and J. Duncan, “Rigid point feature registration using mutual information,” *Medical Image Analysis*, vol. 4, pp. 1–17, 1999.

[86] H.-M. Chen, V. P.K., and A. M.K., “Performance of mutual information similarity measure for registration of multitemporal remote sensing images,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, pp. 2445 – 2454, November 2003.

[87] A. Cole-Rhodes, K. Johnson, and J. LeMoigne, “Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient,” *IEEE Transactions on Image Processing*, vol. 12, pp. 1495–1511, December 2003.

[88] C. C. and C. K., “Automatic georeferencing of airborne pushbroom scanner images with missing ancillary data using mutual information,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 15, pp. 1290 – 1300, 2008.

[89] W. Wang and Y. Chen, “Image registration by control points pairing using the invariant properties of line segments,” *Pattern Recognition Letters*, vol. 18, pp. 269–281, 1997.

[90] E. Guest, E. Berry, R. Baldock, M. Fidrich, and M. Smith, “Robust point correspondence applied to two- and three-dimensional image registration,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 165–179, 2001.

[91] G. Vanderbru and A. Rosenfeld, “Two stage template matching,” *IEEE Transactions on Computers*, vol. 26, pp. 384–393, 1977.

[92] H. Foroosh, J. Zerubia, and M. Berthod, “Extension of phase correlation to subpixel registration,” *IEEE Transactions on Image Processing*, vol. 11, pp. 188–200, 2002.

[93] Goshtasby, “Image registration by local approximation methods,” *Image and Vision Computing*, vol. 6, pp. 255–261, 1988.

[94] T. Peli, “An algorithm for recognition and localization of rotated and scaled objects,” *Proceedings of the IEEE*, vol. 69, pp. 483–485.

[95] A. Goshtasby, “Piecewise linear mapping functions for image registration,” *Pattern Recognition*, vol. 19, pp. 459–466, 1987.

[96] B. Zitova, J. Flusser, and F. Sroubek, “Application of image processing for the conservation of the medieval mosaic,” in *Proceedings of the IEEE International Conference on Image Processing ICIP02 Rochester, MI*, vol. 21, pp. 993–996, 2002.

[97] J. Flusser and T. Suk, “Pattern recognition by affine moment invariants,” *Pattern Recognition*, vol. 26, pp. 167–174, 1993.

[98] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, pp. 143–156, 2001.

[99] Y. Bentoutou, N. Taleb, M. C. E. Mezouar, M. Taleb, and J. Jetto, "An invariant approach for image registration in digital subtraction angiography," *Pattern Recognition*, vol. 35, pp. 2853–2865, 1996.

[100] T. Suk and J. Flusser, "Vertex-based features for recognition of projectively deformed polygons," *Pattern Recognition*, vol. 29, pp. 361–367, 1996.

[101] T. Suk and J. Flusser, "Point-based projective invariants," *Pattern Recognition*, vol. 33, pp. 251–261, 2000.

[102] J. Fitzpatrik and J. West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 917–927, 2001.

[103] J. Fitzpatrik, J. West, and C. M. Jr., "Predicting error in rigidbody point-based registration," *IEEE Trasnactions on Medical Imaging*, vol. 17, pp. 694–702, 1998.

[104] O. Thepaut, K. Kpalma, and J. Ronsin, "Automatic registration of ers and spot multisensor images in a data fusion context," in *Forest Ecology and Management 128*, pp. 93–100, 2000.

[105] R. Gan and A. C. Chung, "Multi-dimensional mutual information based robust image registration using maximum distance-gradient-magnitude," in *IPMI* (I. G. E. Christensen and M. Sonka, eds.), vol. LNCS 3565, (Berlin), pp. 210–221, Springer-Verlag, 2005.

[106] L.G.Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, pp. 326–376, 1992.

[107] W. Pratt, "Correlation techniques of image registration," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 10, pp. 353–358, 1974.

[108] P. Viola, *Alignment by Maximization of Mutual Information.* PhD thesis, Massachusetts Institute of Technology, 1995.

[109] B. Guo, G. S.R., D. R.I., and N. J.D.B., "Band selection for hyperspectral image classification using mutual information," *Geoscience and Remote Sensing Letters, IEEE*, vol. 3, pp. 522 – 526, October 2006.

[110] B. P.J., "Fast filter transforms for image processing," *Computer Vision, Graphics and Image Processing*, vol. 16, pp. 20–51, 1981.

[111] Burt, Peter, Adelson, and Ted, "The laplacian pyramid as a compact image code," *IEEE Trans. Communications*, vol. 9, p. 532C540, 1983.

[112] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation.," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, 1989.

[113] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151, 1988.

[114] R. C.S., C. Sacchi, A. Teschioni, and S. Giulini, "Higher-order-statistics-based sharpness evaluation of a generalizedgaussian pdf model in impulsive noisy environments," in *Statistical Signal and Array Processing, 1998. Proceedings., Ninth IEEE SP Workshop*, 1998.

[115] W. t. Wisniewski and R. A. Schowengerdt, "Information in the joint aggregate pixel distribution of two images," in *Visual Information Processing XIV*, vol. SPIE2005, 2005.

[116] J. R. Schott, C. Salvaggio, S. D. Brown, and R. A. Rose, "Incorporation of texture in multispectral synthetic image generation tools," in *Proceedings of the SPIE, SPIE AeroSense, Targets and Backgrounds: Characterization and Representation, Models and Generators*, vol. 2469, (Orlando, Florida, United States of America), pp. 189–196, SPIE, April 1995.

[117] M. K. Jakubowski, D. R. Pogorzala, T. J. Hattenberger, S. D. Brown, and J. R. Schott, "Synthetic data generation of high-resolution hyperspectral data using dirsig," in *Imaging Spectrometry XII*, vol. 6661, (San Diego, California, United States), SPIE, August 2007.

[118] J. Richiardi, K. Kryszczuk, and A. Drygajlo, "Quality measures in unimodal and multimodal biometric verification," in *Proc. 15th European Signal Processing Conf. (EUSIPCO)*, (Poznan, Poland), 2007.

[119] M. Ehlers and D. Fogel, "An introduction to wavelets," in *Academic Press, San Diego*, 1992.

[120] X. Fan, H. Rhody, and E. Saber, "A comparison of exhaustive search vs. gradient search for automatic imagery registration based on MMI," in *2006 Western New York Image Processing Workshop*, (Rochester, NY), September 2006.

[121] X. Fan, H. Rhody, and E. Saber, "A novel feature enhanced MMI based registration algorithm for automated maps and images," in *2008 IEEE International Geoscience and Remote Sensing Symposium*, (Boston, MA), July 2008.

[122] Lowe and D. G, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision 60 (2)*, pp. 91–110, 2004.

[123] Lowe and D. G, “Object recognition from local scale-invariant features,” *Proceedings of the International Conference on Computer Vision 2*, pp. 1150–1157, 1999.

[124] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Comm. Of the ACM*, vol. 24, p. 381C395.

[125] X. Fan, H. Rhody, and E. Saber, “Automatic registration of multi-sensor airborne imagery,” in *Applied Imagery Pattern Recognition (AIPR) annual workshop*, vol. AIPR2005, (Washington, DC), IEEE, October 2005.

[126] X. Fan, H. Rhody, and E. Saber, “A Harris Corner Label enhanced MMI algorithm for multi-model airborne image registration,” in *Second International Conference on Computer Vision Theory and Applications*, (Barcelona, Spain), March 2007.

[127] X. Fan, H. Rhody, and E. Saber, “An algorithm for automated registration of maps and images based on feature detection and mutual information,” in *Electronic Image, SPIE/ISIT*, (San Jose, CA), Janurary 2008.

[128] X. Fan, H. Rhody, and E. Saber, “A spatial feature enhanced MMI algorithm for multi-modal wild-fire image registration,” in *Applied Imagery Pattern Recognition (AIPR) annual workshop*, vol. AIPR2008, (Washington, DC), IEEE, October 2008.

[129] X. Fan and H. Rhody, “The novel improvement for the HCL-MMI multi-modal image registration by SIFT algorithm,” in *2009 Western New York Image Processing Workshop*, (Rochester, NY), September 2009.

[130] X. Fan, H. Rhody, and E. Saber, “A spatial feature enhanced MMI algorithm for multi-modal airborne registration,” *IEEE Transaction in Geo-science and remote sensing Accepted*, 2009.

[131] S. R. Lach, J. P. Kerekes, and X. Fan, “Fusion of multiple image types for the creation of radiometrically-accurate synthetic scenes,” *Journal of Applied Remote Sensing, Vol. 3*, pp. 120–140, 2009.