8-5-2006

Automated Image Registration And Mosaicking For Multi-Sensor Images Acquired By A Miniature Unmanned Aerial Vehicle Platform

Adnan Orduyilmaz

AUTOMATED IMAGE REGISTRATION AND MOSAICKING FOR MULTI-SENSOR IMAGES ACQUIRED BY A MINIATURE UNMANNED AERIAL VEHICLE PLATFORM

By

Adnan Orduyilmaz

A Thesis
Submitted to the Faculty of
Mississippi State University
In Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Electrical Engineering
in the Department of Electrical and Computer Engineering

Mississippi State University

August 2006

AUTOMATED IMAGE REGISTRATION AND MOSAICKING FOR

MULTI-SENSOR IMAGES ACQUIRED BY A

MINIATURE UNMANNED AERIAL

VEHICLE PLATFORM

By

Adnan Orduyilmaz

Approved:

Dr. Jenny Q. Du Assistant Professor of Electrical and Computer Engineering (Director of Thesis)	Dr. Lori M. Bruce Associate Professor of Electrical and Computer Engineering (Co-Director of Thesis)
Dr. Nicholas H. Younan Professor of Electrical and Computer Engineering (Committee Member and Graduate Coordinator)	Dr. Roger King Associate Dean of College of Engineering

Name: Adnan Orduyilmaz

Date of Degree: August 5, 2006

Major Field: Electrical and Computer Engineering

Major Professor: Dr. Jenny Q. Du

Co-Major Professor: Dr. Lori M. Bruce

Title of Study: AUTOMATED IMAGE REGISTRATION AND MOSAICKING FOR MULTI-SENSOR IMAGES ACQUIRED BY A MINIATURE UNMANNED AERIAL VEHICLE PLATFORM.

Pages in Study: 67

Candidate for Degree of Master of Science in Electrical Engineering

Algorithms for automatic image registration and mosaicking are developed for a miniature Unmanned Aerial Vehicle platform. Three cameras acquire images in a single frame simultaneously at Green (550nm), Red (650 nm), and Near- Infrared (820nm), but with shifting and rotational misalignment. The area-based method is employed in the developed algorithms for control point detection. Since the three images to be registered have different spectral characteristics, region of interest determination and control point selection are the two key steps that ensure the quality of control points. Affine transformation is adopted for spatial transformation, followed by bilinear interpolation for image resampling. Mosaicking is conducted between adjacent frames after three-band co-registration. Manual evaluation confirms the effectiveness of the developed algorithms. The codes are converted into a software package, which is executable under the Microsoft Windows environment. The final products are color-infrared composite and normalized difference vegetation index images

DEDICATION

I would like to dedicate this thesis to my family and all my friends for support, motivation, and guidance during pursuing my degree.

ACKNOWLEDGEMENTS

I would like to thank both my supervisors, Drs. Jenny Q. Du and Lori M. Bruce, for their advice, support, and guidance during my research at Mississippi State University. I would also like to thank my committee member, Dr. Nicholas H. Younan, for the time and effort he spends to read and comment on my work.

I would also like to thank my labmate Ms. Nareenart Raksuntorn for the fruitful research discussions. I am grateful to Mr. Shangshu Cai and other colleages at the Signal Processing Research Applications Laboratory (SPIRAL) for the help in my research and the enjoyable working environment. Finally, I would like to thank my roommate, Tutku Karacolak, for the motivation in my academic life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

1.1 Overview

Image registration is the process of aligning two or more images, which are taken at different times, from different viewpoints, and by different or same sensors, such that the same pixels in these images correspond to the same location in the scene. Image registration has a variety of applications from remote sensing to medical image analysis. In remote sensing, image registration is an inevitable pre-processing step when fusing data from different resources to generate final image analysis products for different purposes, such as target detection, change detection, object classification, etc. The goal is to gain more information about the image scene to best support practical decision-making.

Many image registration tasks are accomplished manually, requiring expert knowledge of image analysts. During the last decades, image acquisition devices have developed rapidly. They capture a huge amount of images with great diversity. This development invokes the research on automatic image registration. But a single automatic image registration program cannot be applied to all applications, because specific requirements, sensor characteristics, and the nature of the imaged area can be different. The performance of a single automatic image registration program may not be well suited for all these different characteristics. The data type, features in the image scene,

registration accuracy, variations of the image, and noise characteristics of the sensed images are the factors that need to be considered in the development of automated image registration [1].

Image mosaicking is the act of combining two or more images with overlapping areas for an overview of a large image scene. The aim is to combine images with undistorted and smooth transition area so that it seems to be acquired from a single sensor. Radiometric normalization and blending processes can be employed for this purpose.

1.2 Application Area and Data

Unmanned Aerial Vehicles (UAVs) are of great importance in remote sensing. These small platforms are able to reach the environments that are difficult or even impossible for human beings to explore, such as chemical accident sites. Air-O-Space International (AOSI) L.L.C., a small business in Mississippi, has developed a miniature unmanned aerial vehicle (mini-UAV) system HL-UAV-10 with 8 ft wingspan and 22 lb gross weight. It has autopilot capability using global positioning system (GPS) navigation. Three charged coupled device (CCD) cameras are onboard this mini-UAV, simultaneously recording images at different bands (Green (550nm), Red (650 nm), and Near-Infrared (NIR) (820nm) bands). It usually operates at about 500 ft altitude to acquire images with 0.15-0.3 m spatial resolution. Images are taken contiguously about every two seconds during the mini-UAV's flight. The mini-UAV employed in this research is shown in Figure 1.1.

Figure 1.1 The mini-UAV used in this research

The digital images are multiplexed and converted to analog image frames. These frames are transmitted to the ground operation station via an S-band video transmitter. The ground control station allows high quality real-time digital image recording. The received analog frames are then reconverted to digital images and recorded onto a magnetic hard disk on a personal computer (PC). This system has been frequently used for different civilian purposes, such as agricultural and environmental monitoring.

After the acquisition operations are completed, the data is submitted to the laboratory for pre-processing, i.e., three band co-registration. This step is necessary because of the existence of misalignment between these cameras even after careful adjustment. Currently, trained personnel do this process manually before creating the commercial products, such as color-infrared (CIR) and normalized difference vegetation index (NDVI) images. The incurred time delay caused by the manual co-registration needs to be minimized for a competitive operation. In order to decrease this time delay, an

automated image registration algorithm is developed, which works in a near real-time fashion during the field operations instead of in the laboratory. In addition, image mosaic needs to be generated using the co-registered image frames for a larger view of the monitoring fields. The developed algorithms should be operated in the easily accessible Microsoft Windows ™ environment of a PC platform for a potential Commercial-Of-The-Shelf (COTS) software package.

1.3 The Challenges

Image registration and mosaicking techniques take advantage of the intensity similarity and/or distinctive features in two images. Our research is challenging because the existing techniques cannot be directly applied and can be even inapplicable in many studies.

The first challenge in this multi-sensor image registration research comes from the fact that the three images, which are acquired from different spectral sensors, visually appear different, because of different solar reflectance characteristics as shown in Figure 1.2.

NIR Band Red Band Green Band

Figure 1.2. The same scene from different bands

In Figure 1.3, the reflectance of four objects/materials in different spectral bands is given. If we focus on the Green, Red, and NIR bands, it can be easily seen that the reflectance of soil, concrete objects, and artificial buildings do not change significantly. For forests and other vegetation, the reflectance is fairly high in the Green band, low in the Red band, and there is a sharp peak in the NIR band. This special signature makes forests and other vegetation areas distinguishable from non-vegetation areas using these three bands.

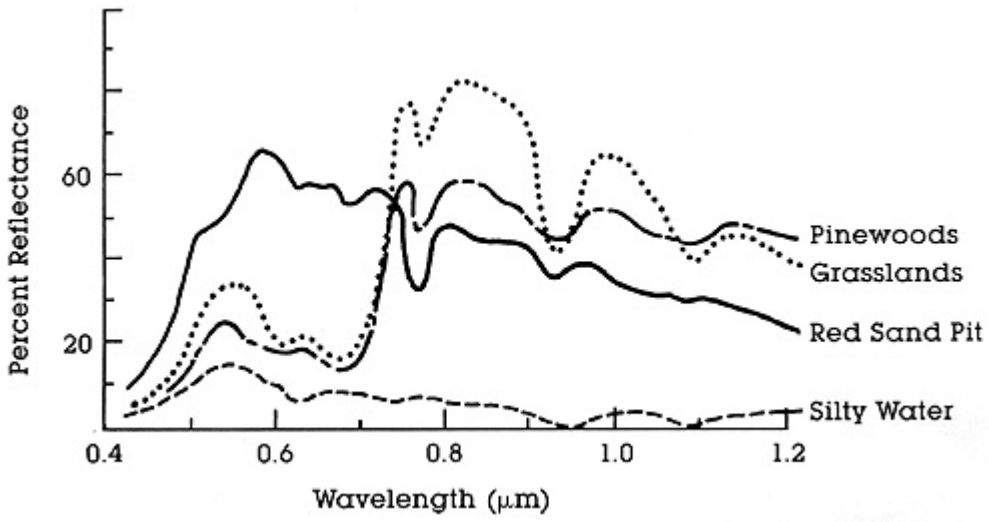


Figure 1.3 Spectral signatures of vegetation and non-vegetation areas [2]

The second challenge is induced when most images are about agricultural areas with no distinctive features being present. In feature-based image registration algorithms, distinctive features are needed for the comparison of two images [3], such as the cross sections of roads, building corners, sharp edges, and close loop boundaries. Unfortunately, in our case, it is difficult to get these reliable features for registration.

1.4 Summary

This thesis aims at developing an automatic image registration and mosaicking technique, which overcomes the difficulties when the multi-sensor system onboard AOSI's mini-UAV platform is being operated for agricultural fields. In order to meet this need, first, a preprocessing step of region of interest (ROI) determination is performed using an objective and automated criteria (e.g. entropy and correlation). Within the selected ROIs, control points (CPs) are detected, whose coordinates in two images are used for the calculation of transformation parameters for image registration. The CP selection is arguably the most important step in the registration process [4], keeping the good CPs for registration. An affine transform is used for image registration because of its simplicity. After transformation, pixel coordinates may be mapped onto fractional coordinates, so interpolation methods are used to determine the coordinates on the regular integer grid (e.g. nearest neighbor, bilinear interpolation). After the three band images in the same frame are co-registered, image mosaicking can be performed between two adjacent frames. The basic steps are similar to the registration process except that a larger rotational misalignment needs to be taken care of. The resulting images are used to create the final products, i.e., CIR and NDVI images.

This thesis is organized as follows. A literature survey of image registration and mosaicking methods is given in Chapter 2. Chapter 3 presents the detailed description of methodologies adopted in our algorithms. The experimental results and discussions are provided in Chapter 4. Finally, Chapter 5 summarizes this research and discusses the scope of future work.

CHAPTER II

LITERATURE REVIEW

2.1 Overview

The methods and implementation of image registration vary with the applications. Its primary applications can be categorized into four topics (Multimodal, Multitemporal, Multiview, and Template Analysis), which are described as follows [5].

Multimodal Analysis: In this category, images taken from different sensors are combined and integrated to have an image with better quality in terms of spectral and spatial resolutions. The resulting image has more detailed information about the image scene. This process is also called data fusion in remote sensing. For instance, the spatial resolution of a multispectral image can be improved by its fusion with a high-resolution panchromatic image [6]; Radar images and optical satellite images are combined to alleviate the effects of clouds and solar illumination variation [4], [7].

Multitemporal Analysis: In this category, registration is applied to the images acquired at different times. The environmental conditions such as clouds and solar illumination can vary at different times, and these effects need to be taken care of during the registration. Multitemporal analysis is very useful in monitoring land cover change [8].

Multiview Analysis: In this category, images to be registered are acquired from different viewpoints. In remote sensing, different viewpoints of the same image scene can be combined to have a better overview look in image mosaicking [9-11].

Template Analysis: In this category, the correspondence between newly sensed data and a previously developed template or dataset is evaluated [12]. In remote sensing, aerial or satellite images can be compared with maps. It is particularly useful in land cover/land use mapping.

In general, an automated image registration technique includes three main steps, which are given in Figure 2.1.

Figure 2.1 Main steps for image registration

Control Point Identification: The image to be registered is called *sensed image*, and the image that is compared with the sensed image is called *reference image*. A control point (CP) is the pixel whose coordinates are known in both images. There are two main methods for the CP detection: area-based and feature-based algorithms. In area-based methods, a small window of points in the reference image is statistically compared with windows of the same size in the sensed image. The comparison uses a similarity metric, which measures the similarity between two given windows [1]. In feature-based algorithms, the image is represented in a compact form by a set of features. The features

are invariant to the scaling, rotation, and gray level modification. The common features are edges, regions, lines, line endings, line intersections, or region centroid. In general, the CP selection is necessary to keep CPs with good quality from the detected CP pool.

Spatial Transformation: Once CPs are identified, the transformation parameters in a mapping function for registration can be determined. In order to define the mapping function, *a priori* information is needed about degradations. If there is no *a priori* information, mapping functions must be flexible so they can be suitable to all possible combinations of degradations.

Image Resampling: After the transformation, the registered image pixel coordinates are not integers anymore. The corresponding integer-valued pixel intensity is computed by an appropriate interpolation technique.

Each aforementioned step can be implemented using different methods. But the following properties are expected [13]:

1. The result of the registration should be accurate.
2. The registration scheme should be robust and reliable. It should work for different displacements and a variety of intensity levels.
3. First two constraints should be accomplished within a small amount of running time, and the registration should be as unsupervised as possible.

2.2 Control Point Identification

2.2.1 Area-Based Methods

In the area-based methods, a window of predefined size or the actual image itself is used for the similarity measurement [14]. The selected window scene should not be smooth, such as soil areas or lakes, because these areas are easy to mismatch with a smooth area at a different location in the other image. Consequently, the selection of the window scene is a very important issue, since this method is sensitive to the intensity changes. The widely used similarity metrics include correlation coefficient (CC) and mutual information (MI).

2.2.1.1 Correlation Coefficient

CC is a commonly used similarity metric in image registration [15-17]. It is calculated by taking a template window from the sensed image and comparing it with the other window from the reference image. The CC measure is formulated by [18]

$$CC = \frac{\sum_{i,j}\left(W(i,j)-E(W(i,j))\right)\left(B(i,j)-E(B(i,j))\right)}{\sqrt{\left[\sum_{i,j}\left(W(i,j)-E(W(i,j))\right)^2\right]\left[\sum_{i,j}\left(B(i,j)-E(B(i,j))\right)^2\right]}} \tag{2.1}$$

where B is the window in the reference image, W is the template window in the sensed image, $E(\cdot)$ denotes the mean operation, and (i,j) is the pixel coordinates in the windows. The best match occurs when the value of CC is maximized. The CC can be generally used when the rotation and scaling factors are small.

Since a simple correlation is not suitable for more complex transformations, new forms of the correlation are designed for geometrically distorted images. The CC was calculated over the geometrically transformed images, so it could handle more complex transformations [19]. Berthilsson applied this approach using the affine transform [20]. Simper used CC in a way of "divide and conquer" for the registration of images with perspective change and lens adjustment problems [21]. Kaneko *et al*. proposed a new version of the CC, which was called increment sign correlation for the partially occluded images [22-23].

A simple and fast version of CC was proposed by Barnea and Silverman, which was called *sequential similarity detection* [24]. The absolute difference between the two image pixels is used for the comparison. It accumulates the sum of absolute differences. A predefined threshold is applied to the accumulated value. If the accumulated value exceeds the threshold, the image pair is rejected.

The registration of multimodal images is comparatively difficult to achieve with the CC, because the intensity of the same object changes in different image with different modalities. Roche *et al.* proposed a new method for multimodal image registration [25]. In this method, the dependency of intensity is assigned to a function. Cain *et al.* proposed a projection-based registration for noisy images [26]. This method has a better performance compared to the classical CC.

Correlation-based image registration algorithms have a smooth similarity map with less variation. To overcome this problem, preprocessing can be applied to extract areas

with relatively large variation for the correlation comparison [27-28]. Pratt used filtering in a preprocessing step in order to have a sharp slope in the similarity map [29].

2.2.1.2 Mutual Information

Mutual information (MI) is another widely used similarity metric [30,31]. There are different forms of mutual information [32-34]. The most general definition is given as follows [35],

$$I(A,B) = H(A) - H(B) - H(A,B) \tag{2.2}$$

where $H(A)$ and $H(B)$ are the entropy of image A and image B, respectively, and $H(A,B)$ is their joint entropy. The formula of the *Shannon entropy* is given as [33]

$$H = \sum_i p_i \log \frac{1}{p_i} \tag{2.3}$$

where p_i is the probability of the i-th intensity. Eq. (2.2) can be represented as

$$I(A,B) = \sum_{a,b} p(a,b) \log \frac{p(a,b)}{p(a)p(b)} \tag{2.4}$$

where $p(a)$ and $p(b)$ are the distributions of images A and B, respectively, and $p(a,b)$ is their joint distribution.

MI measures the dependence between two images. The assumption is that if the images are aligned correctly, MI is maximized. Discussions on its properties in image registration are given in [36].

The window size affects the MI, because as the number of samples increases, the estimation of the probability distribution becomes more accurate. Studholme *et al.* proposed a normalized version of the MI (NMI) [37]. Maes *et al.* proposed another form of NMI, called entropy correlation coefficient (ECC) [38].

Viola and Wells presented one of the earliest applications of MI in the registration of magnetic resonance imaging (MRI) images and 3D object model matching [35]. Thevenaz and Unser developed different algorithms for each step of the MI image registration [39-41]. The maximization of MI was achieved by a Parzen window, Jeeves method [40], and Marquardt-Levenberg method [34]. Ritter *et al.* employed a hierarchical search method in order to quickly locate the maxima of the MI [42]. Studholme *et al.* used three different methods such as MI, joint entropy, and NMI, and compared the performance of these methods [37]. Maes *et al.* adopted Brent's method in the optimization of MI for the registration of MRI, computed tomography (CT), and positron emission tomography (PET) images [38]. The joint probability estimation method was evaluated by Likar and Pernus for the registration of muscle fibre images [43]. The comparison of the MI with other similarity metrics was given in [44]. Rangarajan *et al.* employed MI on the extracted feature sets instead of the original image intensities [45].

2.2.1.3 Fourier Domain Methods

The previous methods descended in 2.2.1.1 and 2.2.1.2 are based on image intensity and can be quite time-consuming. Intensity-based image registration techniques may lead to misregistration especially for noisy images or images acquired from different sensors. Under these circumstances, Fourier domain representations can be used.

Fourier shift theorem is a typical method, where the cross-power spectrum of the sensed and reference images is calculated and maxima location is selected as the CP [46]. This method works better for frequency dependent noise and non-uniform illumination. Also the computational cost is significantly decreased.

De Castro and Morandi developed the extension of phase correlation for the rotation cases [47]. Fourier-Mellin Transform [48], phase correlation [49], and cepstrum filter [50] are also proposed methods for registration, and satisfactory results were reported in [51]. 3-dimensional (3D) application of phase correlation was given in [52]. Anuta proposed to compute the correlation in frequency domain [53]. Since edges are used for the registration process, this method can be employed in multimodal registration.

2.2.1.4 Optimization Methods

The maximization of a similarity function can be represented as an optimization problem in which the parameters depend on the characteristic of geometric transformation. The dimensions of the optimization problem are related to the degrees of freedom of the transformation [14]. As the transformation becomes complex, the number of parameters increases, so the optimization gets sophisticated.

For the projective geometric transformation, a Gauss-Newton numerical minimization algorithm was employed in order to minimize the sum of squared differences [54]. In another application, the MI maximum is achieved by using the gradient descent optimization method [55]. For the variance minimization in the intensities of corresponding pixels, the Levenberg-Marquardt optimization method was used [56]. A

global optimization method for the robust registration of brain images was applied in [57].

2.2.2 Feature-Based Methods

There are two main steps in a feature-based algorithm. In the first step, salient features are extracted from both images. In the second step, these extracted features are matched in the feature space.

2.2.2.1 Feature Extraction

Salient features can be regions, lines, or points. These features should be significant and well-spread in both images. For the purpose of robustness, the number of features should be as large as possible. A change in image geometry, wavelength of the sensor, noise, and image scene should not affect these features. The first type of features is *region*. These features are extracted from the segmented areas, which are homogenous and different in terms of the texture from the background. The examples of such areas are close-boundary regions [58], lakes [59], buildings [60], forests [61], urban areas [62], and shadows [63]. The centroids of these areas are selected as CPs. The second type of features is *line*. Lines are sampled from object contours [64], coastal lines [65], or roads [66]. The centers or the end points are selected as CPs. Edge detection methods such as Canny detector [67], Laplacian of Gaussian [68], and region growing methods can be used for line extraction. The comparison of these methods can be found in [69]. *Point* is another type of feature. This includes line intersections [70], road crossings [62], high variance points [71], local extrema of wavelet transform [72], and corners [73,60,74].

Corners are difficult to detect and define, which were evaluated in [75-76]. In addition, Kitchen and Rosenfeld employed the second-order partial derivatives of an image to detect the corners [77]. Fortsner *et al.* employed the first-order derivative to detect the corners [78]. Trajkovic and Hedley developed an approach based on the fact that the corners have high intensity change in all directions [79].

2.2.2.2 Feature Matching

The correspondence between the features in the two images needs to be found. In order to avoid the misalignment, this correspondence should be well defined. In the feature domain, the most similar feature sets are matched for CP detection. Matching methods are based on spatial relations, invariant descriptors, relaxation, pyramids, or wavelets.

Spatial Relations: Spatial relations can be explored if the features are not very clear and there is a local distortion in the neighborhoods. Goshtasby *et al.* transformed the sensed image features such that after the transformation, these features are as close as possible to the features in the reference image [80]. Stockman *et al.* used abstract lines for the matching process [70]. Barrow *et al.* proposed chamfer matching for image registration, where the line features were matched such that the distance between them was minimized [81].

Invariant Descriptors: Different descriptors are assigned for each feature, and these descriptions should be invariant, unique, stable, and independent from each other [14]. The features with similar descriptions are selected as the CPs. To avoid mismatches, a threshold distance measure is applied, which varies with respect to the features. Sester

et al. described the forests by elongations parameter, compactness, number of holes, and several characteristics of minimum rectangle [61]. Zana *et al.* described each feature point by the angles between relevant intersection lines [82]. For the closed-boundary region features, chain code representation is preferred [83]. Suk *et al*. described the regions using polygons [84]. Moment-based invariant descriptions are often used for the region representation. Holm represented closed-boundary regions by their perimeter, area, compactness, moments, and moment invariants [59].

Relaxations: In this method, each feature in both images is labeled. The matching quality of feature pairs is iteratively calculated until a stable labeling is established [85]. Ranade and Rosenfeld proposed some interesting work in [86]. Wang *et al.* used the description for corners with the relaxation method [74]. Medioni and Nevatia used the line feature descriptions such as coordinates, orientation, and average contrast [87]. Other relaxation methods can be found in [88].

Pyramids and Wavelets: In order to decrease the computational cost, pyramids or wavelets can be employed. The matching estimation starts in the coarsest level of both sensed and reference images. More accurate parameters are calculated at finer resolutions. To avoid misalignment, a back-tracking or consistency check should be done. Wang and Chen extracted features at each level and calculated the parameters [89]. Sharma and Pavel used multiresolution Laplacian pyramid [90]. In addition to pyramids, wavelet decompositions are often used recently. Turcajova and Kautsky [91] and Le Moigne [92] used different wavelets in conjunction with the CC. Fonseca and Costa

detected the maxima from the low high (LH) and high low (HL) coefficients for feature matching [72].

2.2.3 Summary

Area-based methods do not include the salient points and distinctive features. Instead, intensity values of a window are compared with the other image using similarity metrics. In this type of methods, the images should have similar or dependent intensity values. This method generally can handle small rotation and translation misalignments. Computational cost is high, and pyramids or wavelets can be used to decrease the computational time.

Feature-based methods are used when the features of objects are distinctive. These methods are relatively more powerful for the registration of different types of images with distortions. In order to avoid misregistration, the features and their descriptors should be robust and invariant to transformation and distortion.

2.3 Spatial Transformation

After the CPs are identified, the transformation parameters in the mapping function are calculated using the correspondence of the CPs. According to the complexity and properties of the distortion, the mapping function type is selected. Based on the area where the transformation is applied, mapping functions can be divided into global and local functions. Based on the linearity, mapping functions can be divided into linear and nonlinear functions. The frequently used mapping functions are described below.

2.3.1 Affine Transform

The affine transform is linear and it can map a parallelogram to a square. The formula of the affine transform is given as [14]

$$u = a_0 + a_1 x + a_2 y$$
$$v = b_0 + b_1 x + b_2 y \tag{2.5}$$

where a_0 and b_0 are for shifting adjustment, and a_1, b_1, a_2, and b_2 are for rotational adjustment. Here, (x,y) and (u,v) are coordinates of CPs before and after registration. Theoretically, three independent CPs are necessary to solve the equation. In practice, the number of CPs is much more than the needed number for higher registration accuracy, and the least squares solution is used to estimate the parameters that can minimize the estimation error. The affine transform can be used for both local and global transformation.

2.3.2 Radial Basis Function

A radial basis function performs nonlinear mapping. The mapping function is given as [14]

$$u = a_0 + a_1 x + a_2 y + \sum_{i=1}^{N} c_i g(\mathbf{x}, \mathbf{x}_i)$$

$$v = b_0 + b_1 x + \mathrm{b}_2 y + \sum_{i=1}^{N} c_i g(\mathbf{x}, \mathbf{x}_i) \tag{2.6}$$

where a_0, a_1, a_2, b_0, b_1, b_2, and c_i are the parameters to be determined; g(·) is a radial basis function, which can be chosen as multiquadrics [93], reciprocal multiquadrics [93],

Gaussian, Wendland's functions [94], or thin-plate splines [14]; $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ is a vector of pixel coordinates before the registration and $\mathbf{x}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$ is the vector of pixel coordinates of the i-th CP for $1 \leq i \leq N$.

2.4 Image Resampling

After transformation the pixel coordinates are not integers. So interpolation needs to be applied, such as the nearest neighbor function, bilinear and bicubic functions, quadratic splines [95], cubic B-splines [96], Gaussians [97], and truncated sinc functions [98]. Different interpolation techniques are reviewed in [99-100].

Bilinear interpolation is commonly used because of its effectiveness and low computational cost. Cubic interpolation is used for high enlargements. Nearest neighbor is the simplest method, but it produces jagged edges and chunky artifacts.

2.5 Image Mosaicking

Image mosaicking is the process of obtaining images with a larger field of view by combining two or more overlapping images with some different areas. These images may have different radiometric, geometric, and quality properties.

The image mosaicking process is composed of two main steps. In the first step, the images are aligned using an appropriate registration method. In the second step, the registered image is post-processed in order to have a smooth and consistent mosaic. There are two main problems with the resulting registered image. The first problem is the

radiometric equalization of the image mosaics, if the images are taken from different types of sensors. The second problem is the blending of the two images. At the transition areas, there are some artifacts remaining. In order to have a smooth mosaic, those artifacts should be removed.

Radiometric normalization uses the overlapped area of the two images to find appropriate correction parameters. The correction parameters can be obtained using some objective measures [101-102]. The techniques frequently used in radiometric normalization include linear regression [103], pseudo-invariant features [104], bright and dark pixel sets [105], and the classification of land cover [101].

At the boundary, there may be some artifacts, distorting the smoothness and compactness of the mosaic. The removal of discontinuities and artifacts to make the resulting mosaic a compact image is called blending process. Weighted average of the transition or boundary areas is a basic method for this purpose. There are some other complicated methods to blend the subimages correctly. For instance, a linear ramp function was applied to equalize the intensity values at transition areas [106]; the histogram was calculated in the overlapping area to find the intensity difference between two subimages [107]; an iterative algorithm was adopted to minimize the error between subimages and increase the smoothness of the image [108]; and multiresolution spline functions was proposed for blending in [109].

CHAPTER III

METHODOLOGIES

3.1 Overview

In this thesis, automatic registration and mosaicking algorithms for the images acquired by three-band sensors mounted on AOSI's mini-UAV are developed. The flowchart shown in Figure 3.1 summarizes all the steps. First, the region of interest (ROI) is determined for control point (CP) detection; CPs with good quality are selected to determine the mapping functions; after a three-band single frame coregistration, the consecutive frames are mosaicked to have a large overview of the scene.

3.2 Control Point Identification

3.2.1 Region of Interest Selection

Since images are taken from different sensors with different solar reflectance variation, CPs cannot be found directly using an area-based method. For example, a forest area is very bright in the NIR band, but the corresponding area in the Red band is very dim. Consequently, the similarity between these two areas is very low in terms of intensity values. Since images are taken from the agricultural areas, the soil and grass fields are dominant in the images. The windows should be chosen from the region with relatively large variation to avoid misalignment. To select the distinctive areas, entropy is calculated, which can be used to represent the variation within an image.

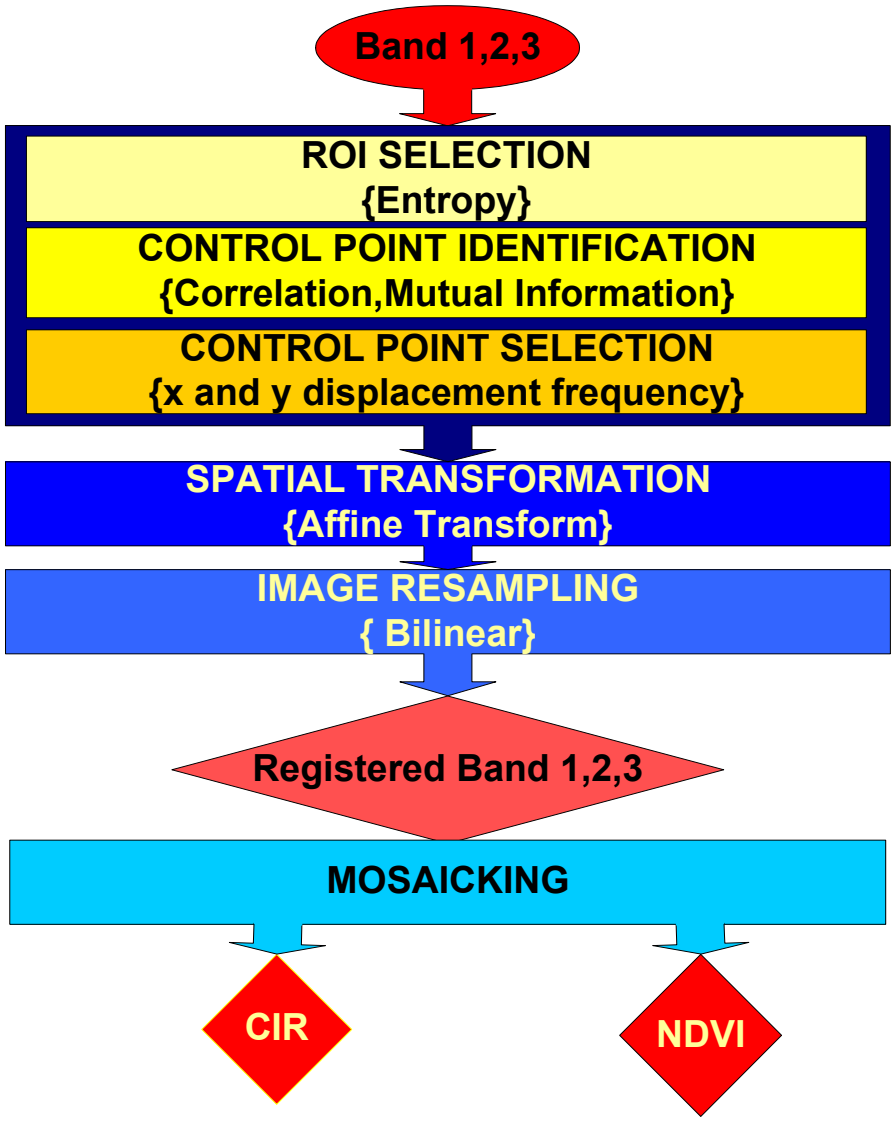


Figure 3.1 Overall block diagram of the registration and mosaicking system

An NIR image and its corresponding entropy map are given in Figure 3.2. We can see that the soil and grass areas have lower entropy, and buildings, trees, and roads have higher entropy.

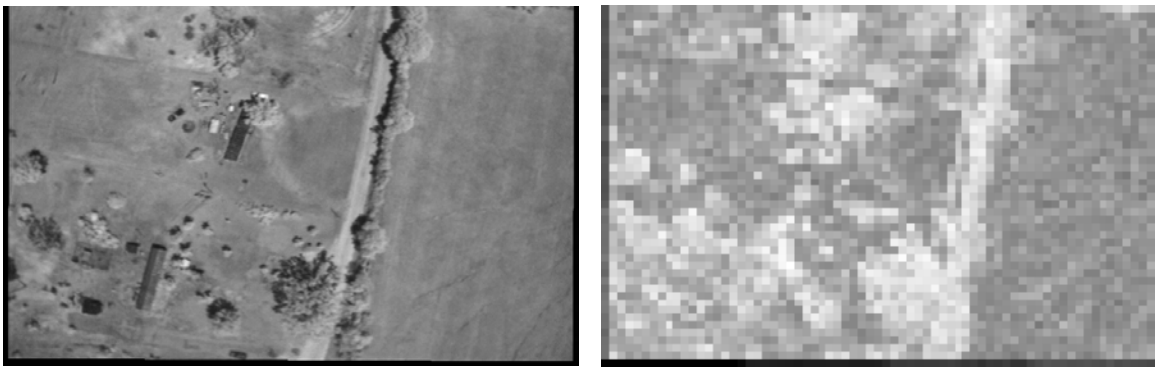


Figure 3.2 NIR band image and corresponding entropy map

3.2.2 Control Point Detection

To find similar areas of the sensed and reference images, a template window is selected at the ROI centers. In our experiments, 51 × 51 is a good choice for speed and resultant CP quality. Each window of the sensed image is compared to the corresponding window in the reference image as shown in Figure 3.3.

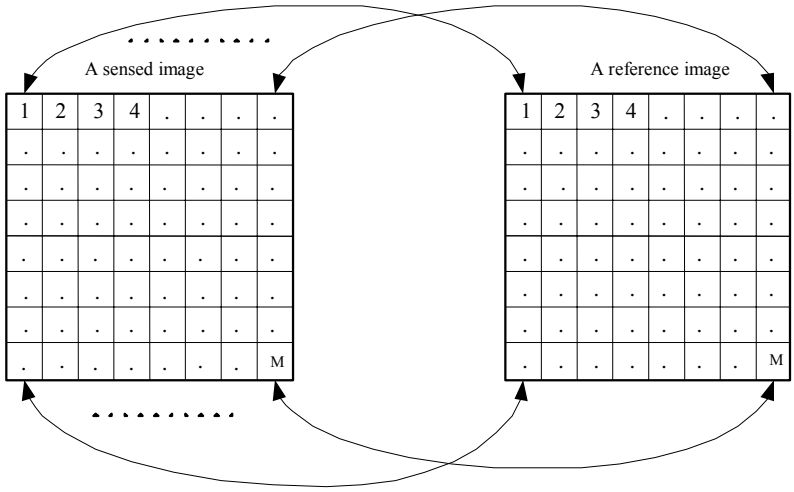


Figure 3.3 Similarity comparison

If the similarity is maximal or above a threshold, we consider these two areas are similar, and the central pixels will be selected as CPs. An example of the similar areas after comparison is given in Figure 3.4.

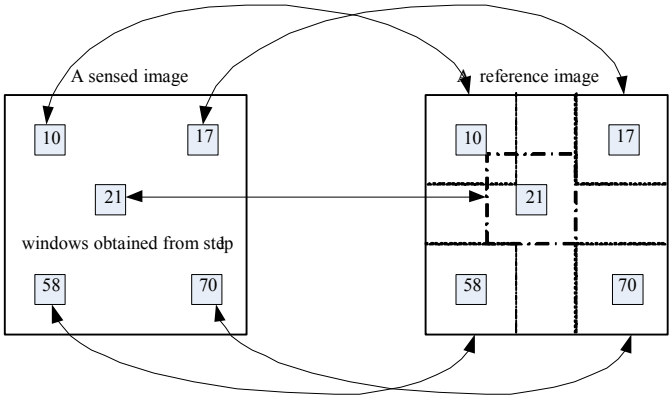


Figure 3.4 An example of the similar areas after comparison

The correlation coefficient (CC) in Eq. (2.1) and mutual information (MI) in Eq. (2.4) are the similarity metrics used in this thesis. If two areas are very close, the CC will be close to 1. In Figure 3.5, the comparison of the template window from sensed image to the subimages of the reference image is illustrated.

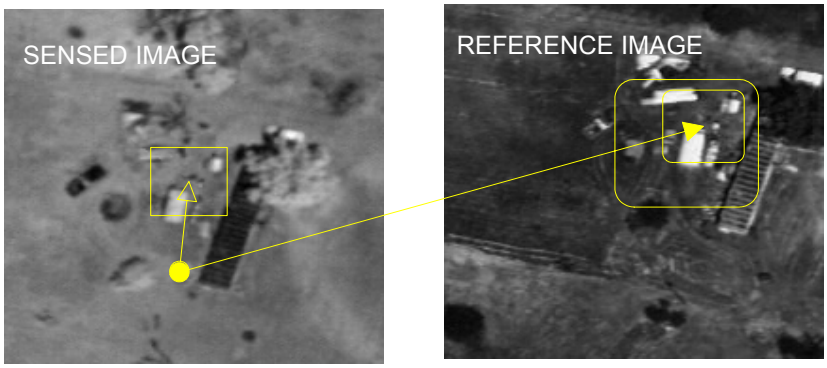
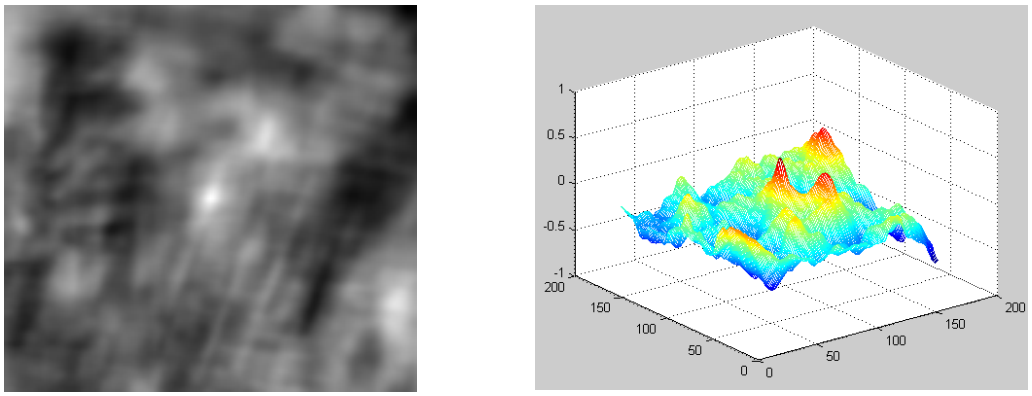


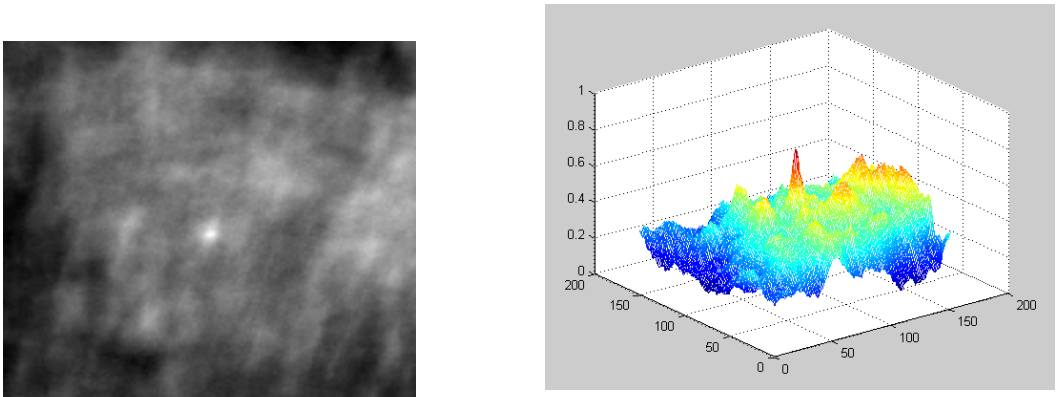
Figure 3.5 Illustration of the comparison process

In Figure 3.6, the CC map is given. The brighter areas in the CC map represent high correlation and the darker areas represent low correlation. Also a 3-D mesh is given to have a better illustration. In Figure 3.7, the resulting MI map of the same image is given, where a high MI is represented with a bright pixel in the 2-D MI map. In the 3-D MI map, sharper peaks are present which prevent false alarms.



2-D CC Map 3-D CC Map

Figure 3.6 The CC map



2-D MI Map 3-D MI Map

Figure 3.7 The MI map

3.2.3 Control Point Selection

CPs will then be selected from the candidate CPs. As illustrated in Figure 3.8, the selection includes the following steps:

1. Compute the *x*-directional displacement and *y*-directional displacement of each pair of potential CPs.
2. Calculate the occurrence frequencies of all *x* and *y*-directional displacements.
3. Detect the center point that have the largest occurrence frequencies in *x* and *y* directions.
4. Keep CPs around the center point within the threshold values.

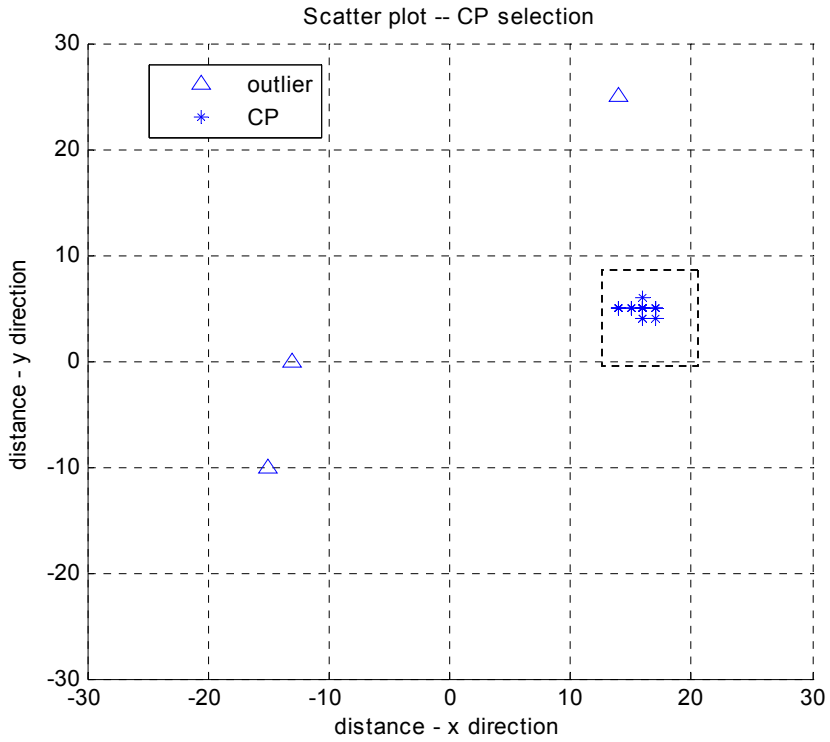


Figure 3.8 The illustration of the control point selection

3.3 Spatial Transformation

After the CPs are identified, the transformation parameters can be determined using their coordinates. According to the complexity and properties of the distortion, the mapping function type is selected. In our case, the mini-UAV takes the three bands of the same frame simultaneously at a specific altitude. So the scaling factor does not need to be included. After the adjustment of sensor location is made, there is small rotational and translation misalignment between the images. So the simple linear affine transform in Eq. (2.5) is appropriate for our case.

Theoretically, three independent CPs are sufficient to solve the equation. However, more CPs are needed for a better registration. The least squares solution is employed to estimate the parameters with the least error. The matrix form of the affine transform is

$$\mathbf{B} = \mathbf{A}\,\mathbf{R} \tag{3.1}$$

where **R** is the matrix including the pixel coordinates of the CPs in the sensed image, **A** is the affine transformation matrix, and **B** is the matrix with the CP pixel coordinates in the reference images. The transformation parameters can be estimated as

$$\hat{\mathbf{A}} = \mathbf{B}(\mathbf{R}^{\mathrm{T}}(\mathbf{R}\mathbf{R}^{\mathrm{T}})^{-1}) \tag{3.2}$$

where $\mathbf{R}^{\mathrm{T}}$ is the transposed matrix of **R**, $\hat{\mathbf{A}}$ is the estimate of the affine transformation matrix, and superscript –1 indicates a matrix inverse.

3.3 Image Resampling

After transformation, pixel coordinates are not integers any more. If using rounding, one or more pixels can be transformed to the same pixel, and some pixels are not assigned in the registered image leaving them as dark pixels as shown in Figure 3.9.

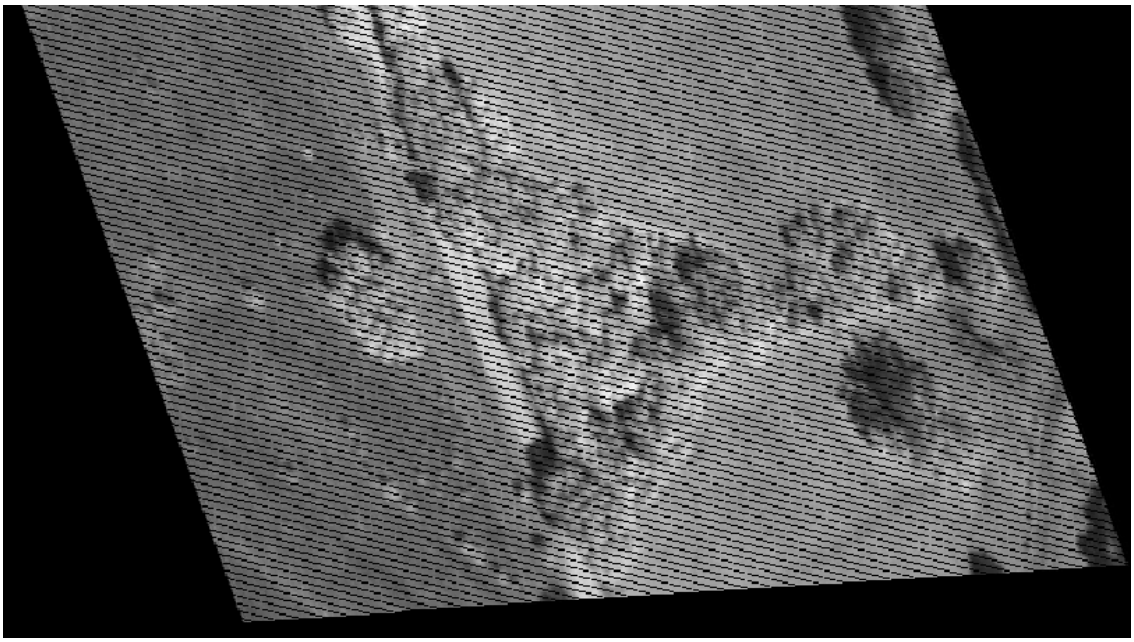


Figure 3.9 The illustration of the blank pixels after transformation

An interpolation method needs to be used. Bilinear interpolation is a good choice due to its simplicity and effectiveness. It determines the value of a new pixel based on a weighted average of four pixels in its nearest neighborhood. An illustration of bilinear interpolation is given in Figure 3.10.

The bilinear interpolation can be represented as

$$U(x,y) = \alpha x + \beta y + \gamma xy + \delta \tag{3.3}$$

where U is the intensity of the pixel at the integer coordinates (x,y), and α, β, γ, and δ are the constants to be determined. To solve Eq. 3.3, four equations are needed which can be constructed from the four neighbors. The computation cost is relatively smaller than other interpolation methods, and generally the result is satisfying. So bilinear interpolation is used as the default method.

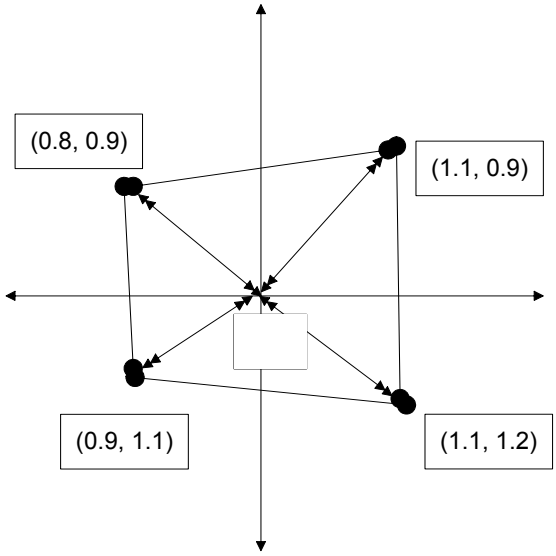


Figure 3.10 Bilinear Interpolation

3.5 Image Mosaicking

Image Mosaicking includes as an image registration process. Post-processing such as radiometric normalization is not required in this case because images to be mosaicked are taken from the same sensor during the same time period. In the registration process for image Mosaicking, slight differences include:

1. The images to be registered, which have been coregistered within each frame, are taken in the same band in consecutive frames.

2. In the registration step, there was a slight translation, so the searching area is selected as three or four times of the window size. However, in image mosaicking the searching area is the entire overlapped area, so the resultant computation cost is higher than image registration.
3. In the multiband image registration, there is a small rotation, which can be handled using the affine transform. However, in the image mosaicking, the images have larger rotation as the UAV flying. So the rotation angle should be detected with a pre-processing step before CP detection.
4. Since the distinct features of the forest, roads, and buildings are well represented in the NIR band, the reference image is selected as the NIR band. After the NIR bands are mosaicked, the Red and Green bands can be mosaicked accordingly.

Considering the differences explained above, the mosaicking algorithm has an additional step – rotation angle detection, described in the following section.

3.5.1 Rotation Angle Detection

The major difficulty in image mosaicking is the large rotation between the consecutive image frames. The window content at the same location is changing at each degree, which leads to the difficulties in similarity comparison. The rotation should be estimated and adjusted before the comparison of windows.

With pre-introducing the rotation, the comparison of windows is still conducted using CC and MI. The resulting rotation angle versus CC and MI are given in Figure 3.11. The

rotation angle corresponding to the maximum coefficient is selected for each window. The rotation angle with maximum occurrence frequency is considered as the actual rotation angle. From Figure 3.11, we can conclude that MI has a sharp peak and CC has a more smooth curve. So the rotation angle can be more efficiently detected using MI.

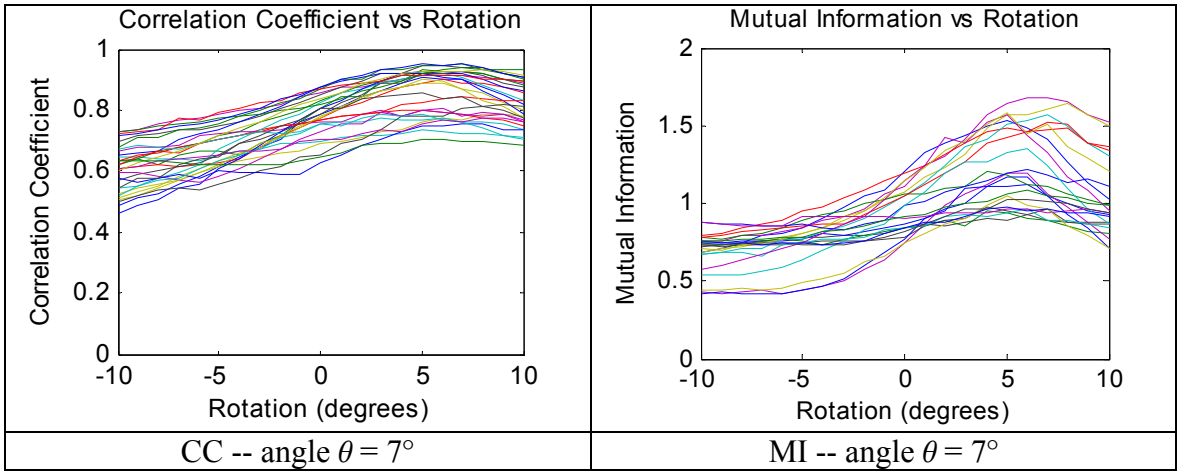


Figure 3.11 The rotation detection using MI and CC for the angle $\theta = 7°$

After the rotation angle is detected, the NIR image is first rotated back using this angle, and the remaining steps in the registration algorithm is the same as in the previous sections.

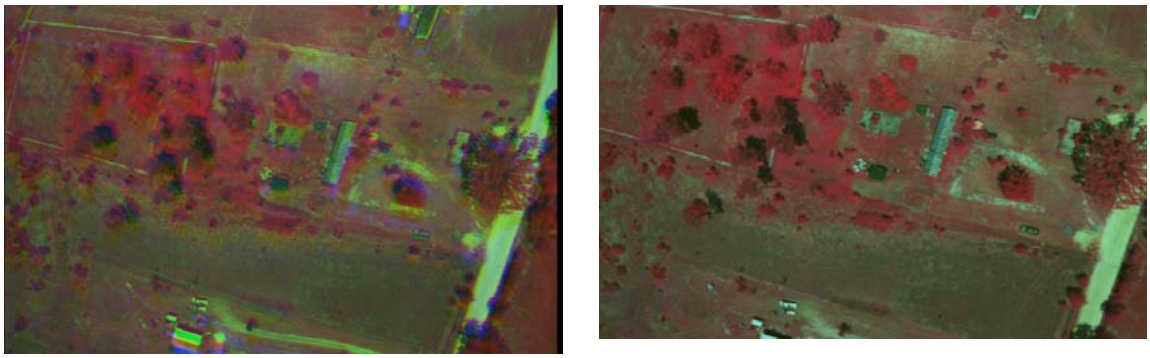
Since the images are taken within a very close amount of time, the radiometric difference between the overlap areas in the reference and sensed images is assumed to be very small. At the transition areas close to the boundary of the images, the intensity values are weighted on average to have a smooth and compact mosaic.

3.6 Final Product Generation

After the multiband registration and multiframe mosaicking, the data is now available for use. There are two widely used products in agricultural studies: pseudo color infrared (CIR) composite image and normalized difference vegetation index (NDVI) images.

3.6.1 CIR Color Composite Images

CIR images are simply generated by using the NIR band as the Red component, the Red band as the Green component, and the Green band as the Blue component. This representation is shown in Figure 3.12.



CIR – Before registration CIR – After registration

Figure 3.12 CIR images before and after registration

CIR images reveal very useful information about the image scene, particularly about vegetation density and distribution. As shown in Figure 3.12, the vegetation areas are dominant in agricultural image scenes, so the most part of the images are red. Healthiness of the vegetation is varying with respect to the red color density. The roads and concrete objects, like buildings, are generally represented in light blue or gray. The soil has a brown color. In Figure 3.12, the colors after registration are more accurate, but the colors

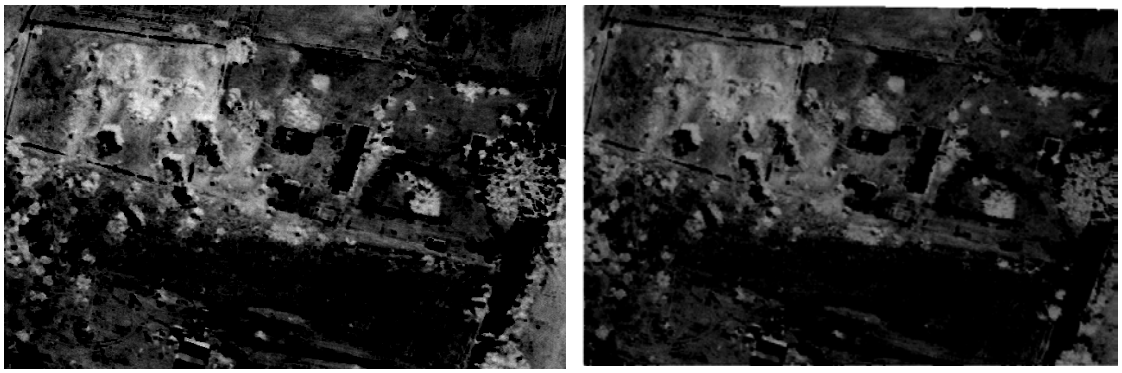
before registration have some distinct problems. For instance, the colors around buildings are purple.

3.6.2 NDVI Images

NDVI is defined as

$$NDVI = \frac{NIR - RED}{NIR + RED}. \tag{3.4}$$

NDVI images can be used for the detection of the healthiness of vegetation areas, where the bright pixels correspond to the healthy vegetation areas, and the buildings are represented as low intensity. An example is given in Figure 3.13, where the NDVI image after registration has higher contrast and bright areas (such as small trees) are well concentrated.



Before Registration After Registration

Figure 3.13 NDVI images before and after registration

3.6.3 Executable Program

The automatic image registration and mosaicking algorithms are developed in MATLAB environment. Using a MATLAB compiler, the written codes are converted into a software package, which is executable under the Windows, Linux, or Unix PC platforms without the requirement of MATLAB or other special software package.

CHAPTER IV

RESULTS AND DISCUSSIONS

4.1 Database

The images collected at four sites – Lake Columbus, Stennis Space Center, Greenwood, and Oswalt in Mississippi in 2005, were used in this research.

4.2 Image Registration

Two examples of the original and registered images in three bands and resulting CIR and NDVI images are given in Figures 4.1-4.2. The images were acquired from Oswalt. In the CIR and NDVI images, the effect of the registration can be easily recognized. As the accuracy of the alignment increases, registered images have more compact and clear CIR and NDVI images. The misregistration causes blurry CIR images and some artificial artifacts in NDVI images. In Figure 4.2, the colors around the panel were not clear in CIR image, but after the registration the colors were more compact. There was not much difference between MI and CC, because both methods selected enough number of good CPs. However, the small house at the low part of the image in Figure 4.1 was not well registered using CC, but that house was better registered using MI. The registration error was increased for objects far away from the image center. Overall, both MI and CC performed successful registration.

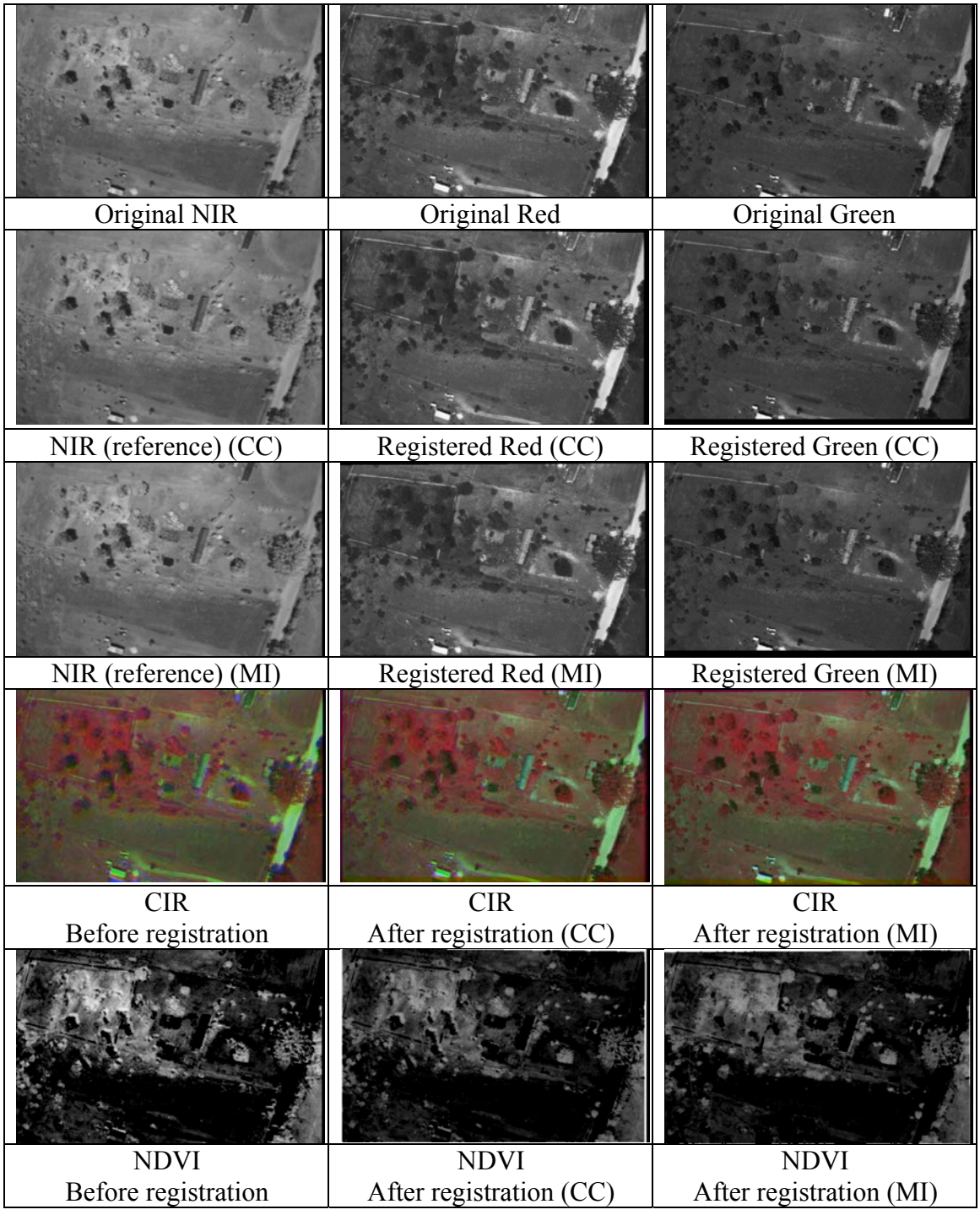


Figure 4.1 An example of image registration using CC and MI

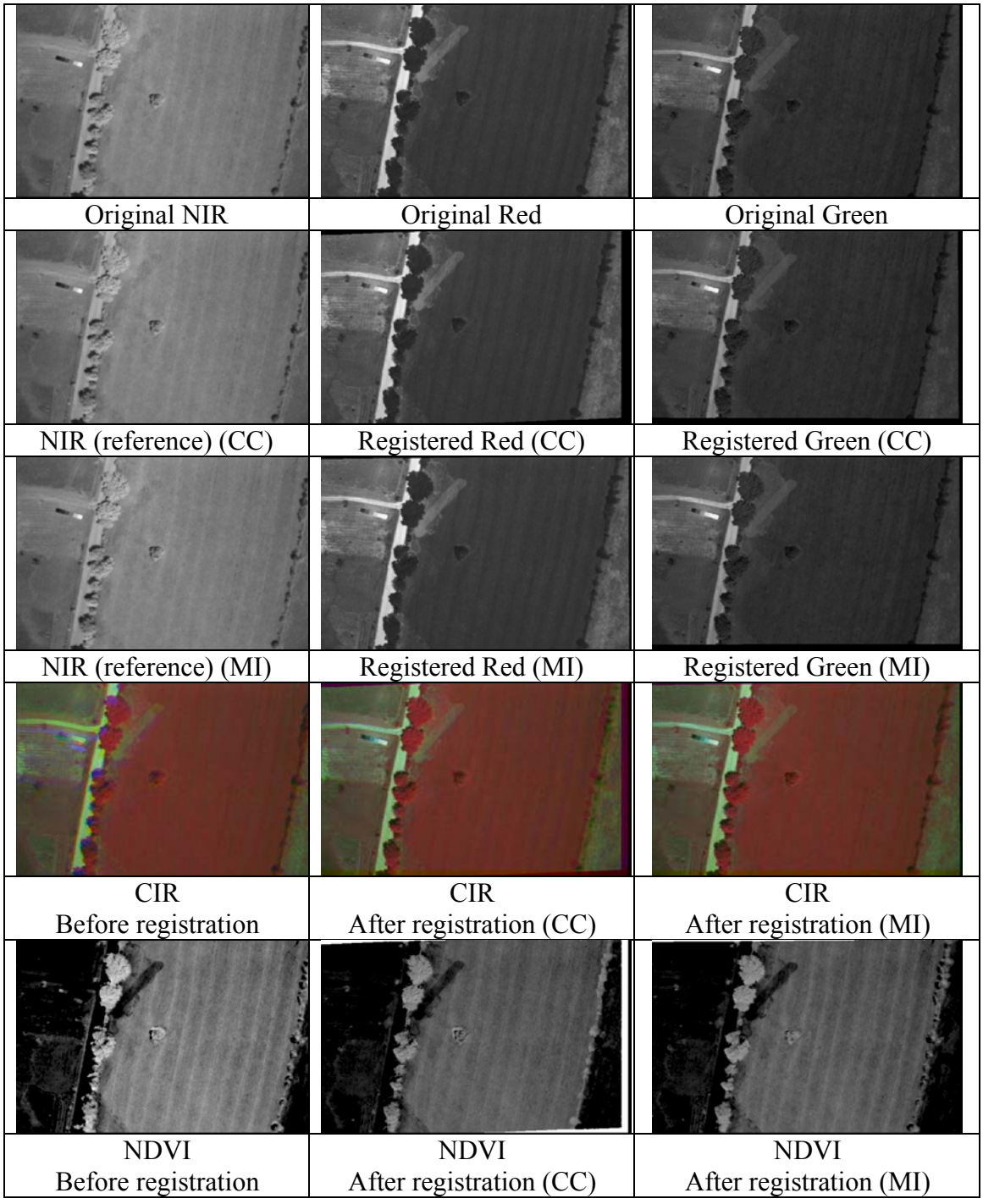
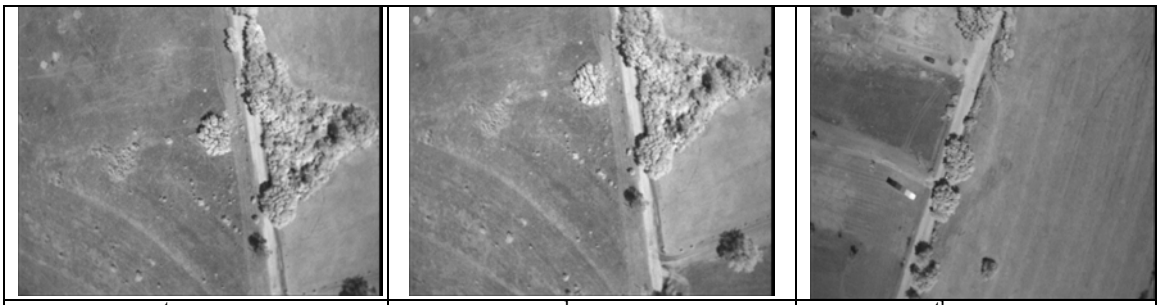


Figure 4.2 Another example of image registration using CC and MI

4.3 Image Mosaicking

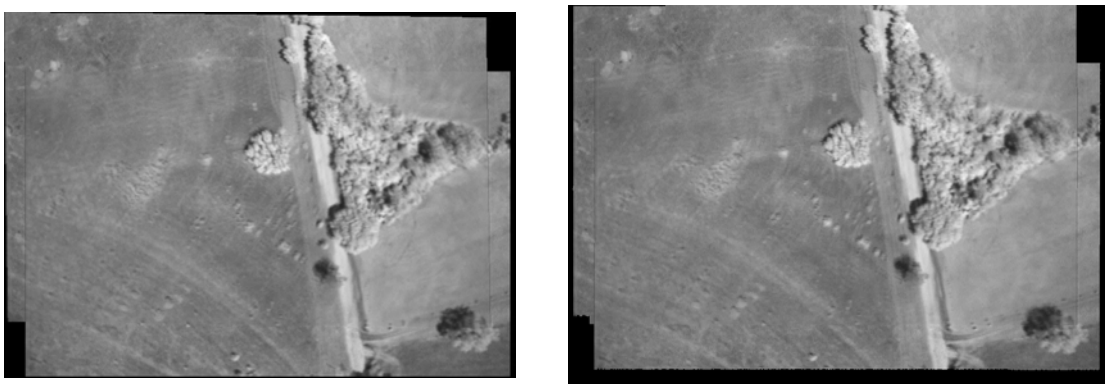
The images used for image mosaicking are shown in Figure 4.3. Ten consecutive frames were mosaicked. Only the NIR bands are shown here, because the NIR band is selected as the reference image for the mosaicking. The Red and Green bands were mosaicked using the parameters found for the NIR image.



1st frame	2nd frame	8th frame

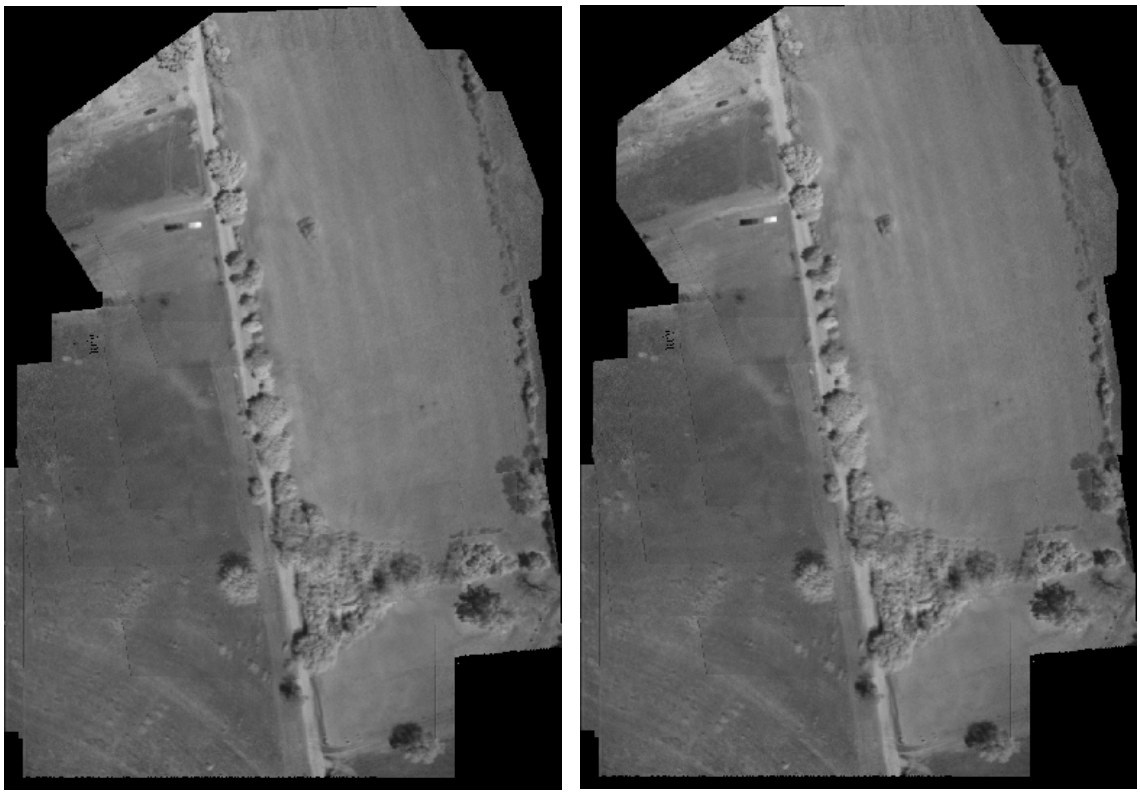
Figure 4.3 Three image frames used in the mosaicking

The resulting mosaics using the first two consecutive frames and eight consecutive frames together are given in Figures 4.4 and 4.5, respectively. The corresponding CIR and NDVI images are given in Figures 4.6-4.9.



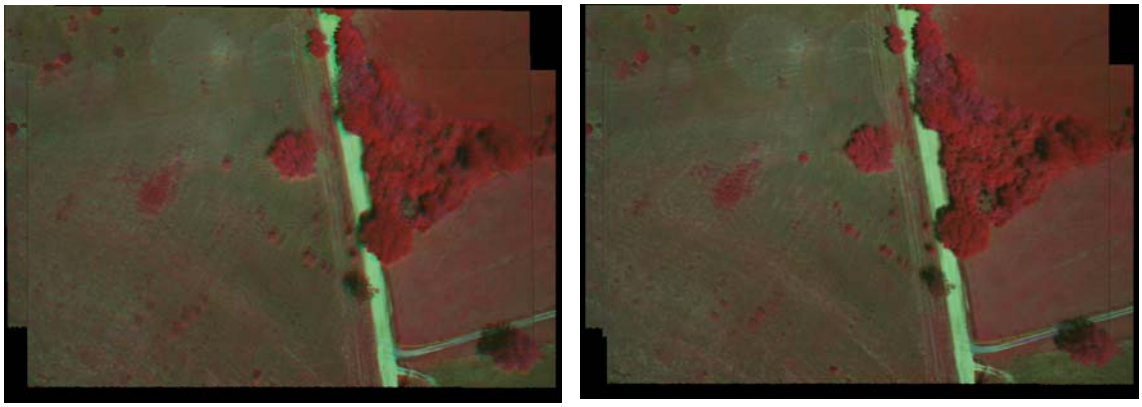
Mosaic 1 – CC Mosaic 1 – MI

Figure 4.4 Mosaic 1 using CC and MI



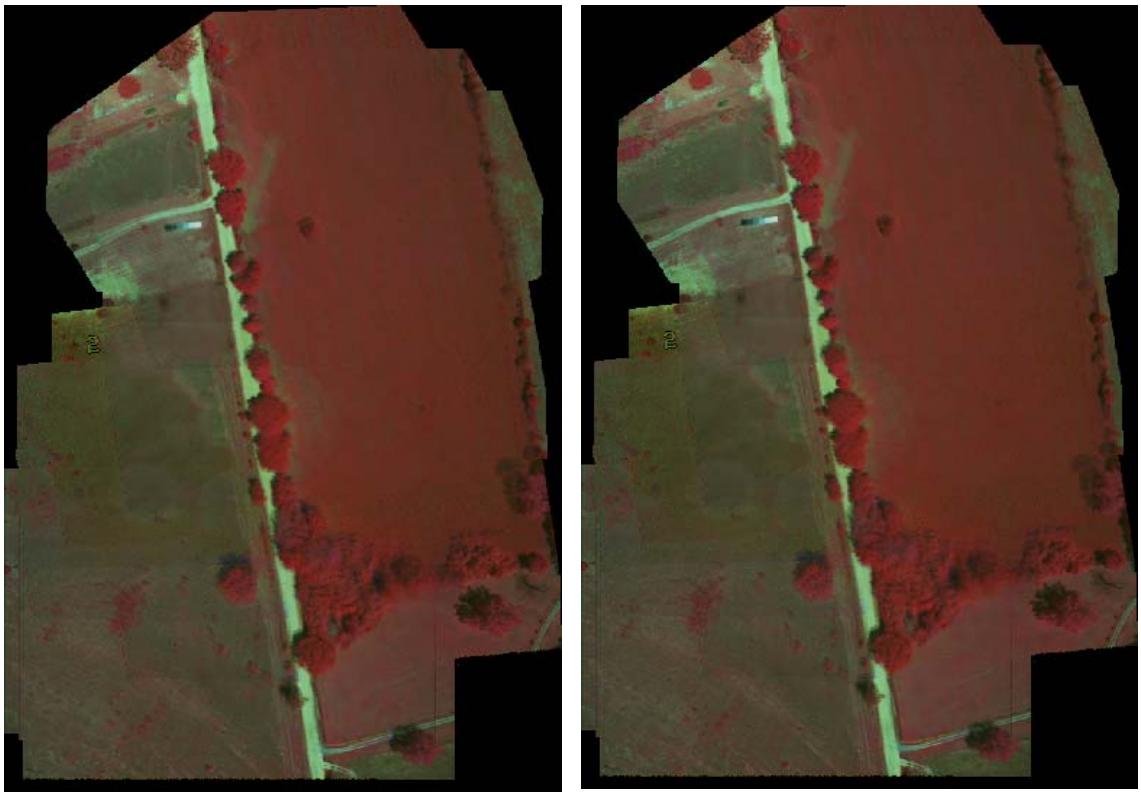
Mosaic 8 – CC Mosaic 8 – MI

Figure 4.5 Mosaic 8 using CC and MI



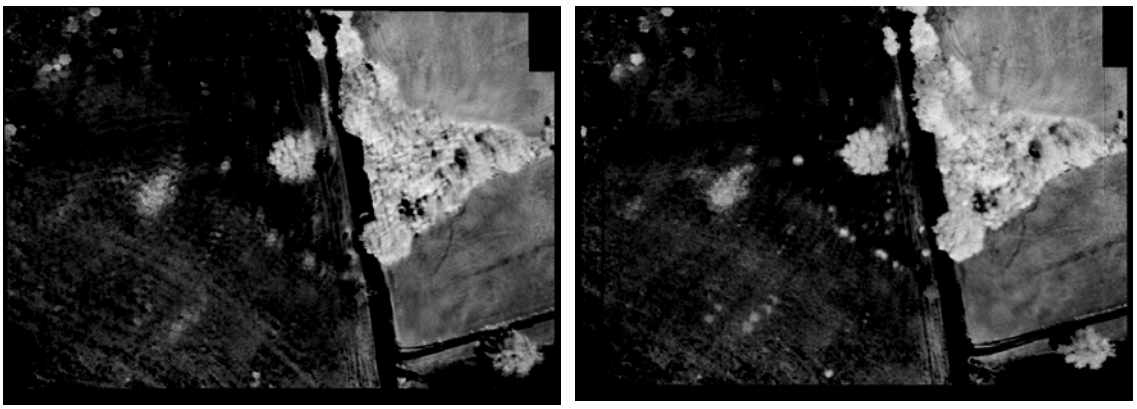
CIR Mosaic 1 – CC CIR Mosaic 1 – MI

Figure 4.6 Mosaic 1 CIR images using CC and MI



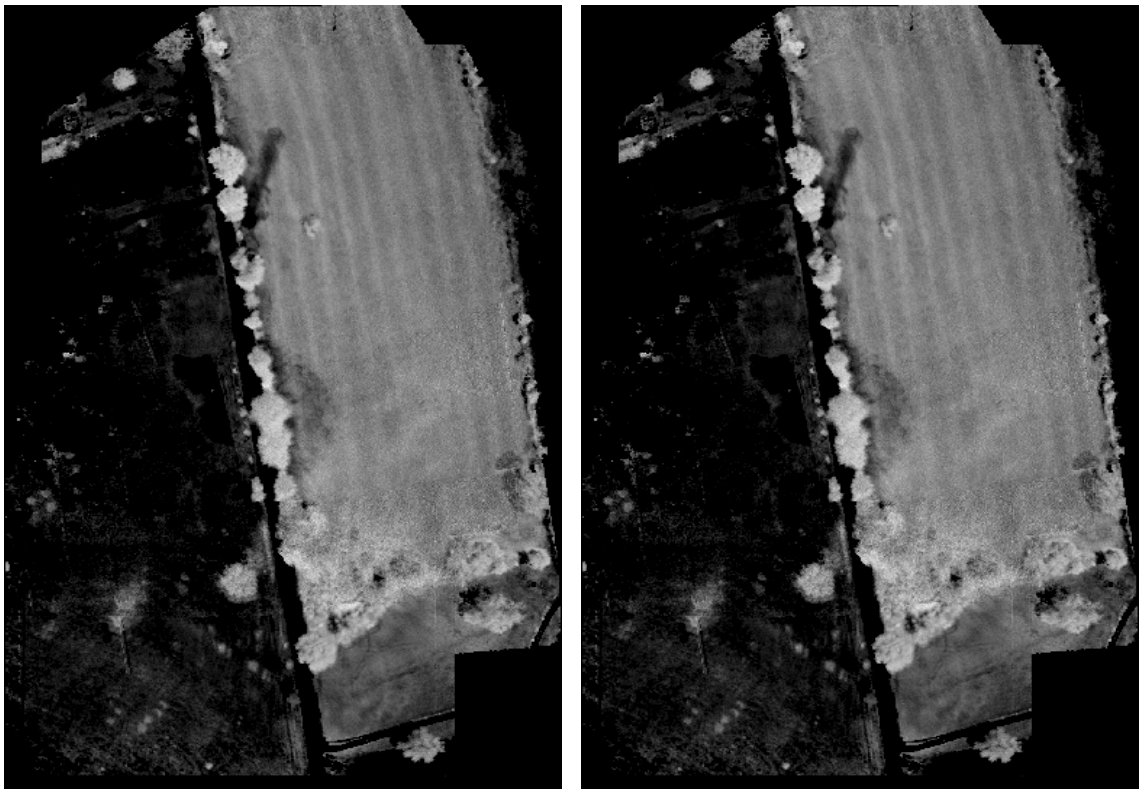
CIR Mosaic 8 – CC CIR Mosaic 8 – MI

Figure 4.7 Mosaic 8 CIR images using CC and MI



NDVI Mosaic 1 – CC NDVI Mosaic 1 – MI

Figure 4.8 Mosaic 1 NDVI images using CC and MI



NDVI Mosaic 8 – CC NDVI Mosaic 8 – MI

Figure 4.9 Mosaic 8 NDVI images using CC and MI

The difference between the MI and CC can be distinguished in Figure 4.8. The NDVI image using CC was blurry, if the small trees are considered at the center of the image. However, NDVI image using MI was clearer.

4.4 Evaluation of the Results

4.4.1 Manual Evaluation for Registration

4.4.1.1 Evaluation Process

Before the flights, 6 panels with the size of 3×3 meters were put in the image scenes.

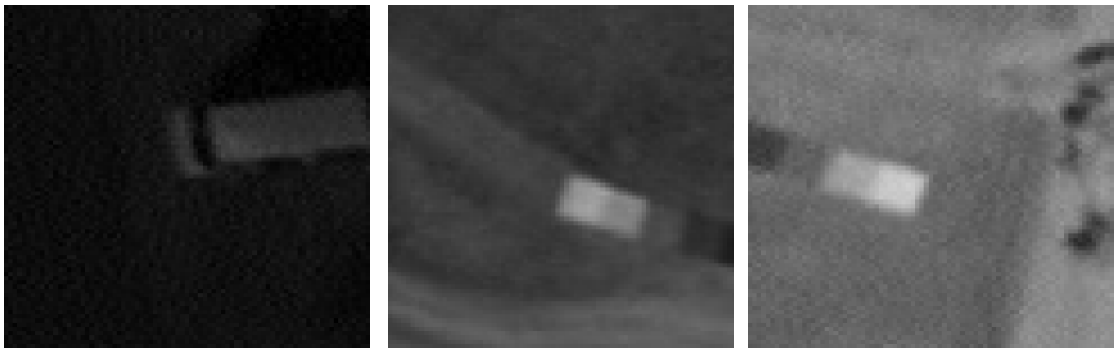
The first panel is white, and the last is black. The others have the gray level tones. The corners of these panels are distinctive, so these corners can be used for evaluation. The samples of the images with panels are shown in Figure 4.10.

Stennis Space Center Greenwood Oswalt

Figure 4.10 The images with panels in Stennis Space Center, Greenwood, and Oswalt

In the evaluation process, totally 8 frames with panels were used: 3 about Stennis Space Center, 2 about Greenwood, and 3 about Oswalt. Four graduate students and two faculty members participated in the evaluation. Each frame has 6 images composed of three bands before and after registration. The evaluation process was employed for both MI and CC methods. Since the spatial resolution of the images is not high enough to locate the corner exactly, there will be a user-dependent error. Examples about the corners are given in Figure 4.11.

Participants were given zoomed images as shown in Figure 4.11. Participants were instructed to select a corner by simply clicking on the corner pixel on the image for each frame. They selected the corners for three bands before and after registration. This process was repeated for each corner. When they clicked the corner, coordinates of the corners were saved by the program. The same program was employed for all the participants.



Stennis Space Center Greenwood Oswalt

Figure 4.11 The zoomed images around the panel corners in Figure 4.10

4.4.1.2 Registration Accuracy

The selected corner locations for NIR, Red, and Green bands are presented in Figure 4.12, where the blue points represent the pixel locations of the corners in NIR band, red points represents Red band, and green points represents Green band. Typically, one cluster corresponds to one corner. Before registration, the corners were misaligned with small rotation and translation, so the red, green, and blue clusters for the same corner were far away from each other. After the registration, the red, green, and blue clusters for the same corner were very close. Ideally, the corner coordinates should be the same for the bands after registration.

The quantified misalignment between the 3 bands in 8 frames is given in Tables 4.1-4.3. The first three image sets were taken from Oswalt; image set 4 and 5 were from Greenwood; and last three images were taken from Stennis Space Center. The Euclidian distance was calculated between each point and its cluster center. Here, μ represents the mean of the misalignment for 4 corners determined by 6 users, *std* represents the standard deviation, and the unit is in pixels. We can see that the misalignment was fairly high

before the registration, and it was very small after registration. Comparing the results from MI with CC, the accuracy using MI was slightly higher than CC.

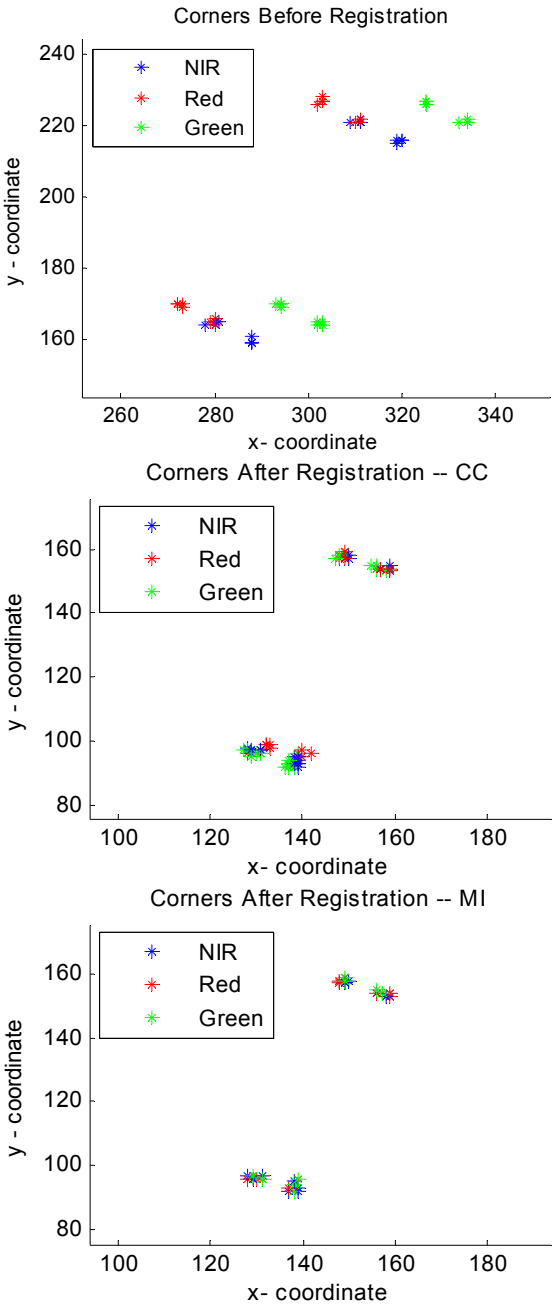


Figure 4.12 Manual evaluation for Oswalt image set using MI and CC

Table 4.1 The misalignment between NIR and Red band

Image Set	Before		After - MI		After – CC	
	μ	*std*	μ	*std*	μ	*Std*
1	9.019776	0.97032	0.4375	0.314576	1	0.889757
2	9.9159	0.413218	0.625	0.433013	1.3125	0.239357
3	9.676206	0.590657	0.3125	0.239357	1.25	0.456435
4	10.07202	0.918386	0.5	0.204124	0.4375	0.426956
5	5.479581	0.221177	0.625	0.829156	0.625	0.322749
6	4.096143	0.880142	0.5625	0.426956	1.625	0.661438
7	5.583762	0.685904	0.75	0.408248	0.6875	0.314576
8	5.761751	0.651004	0.625	0.595119	0.5	0.408248

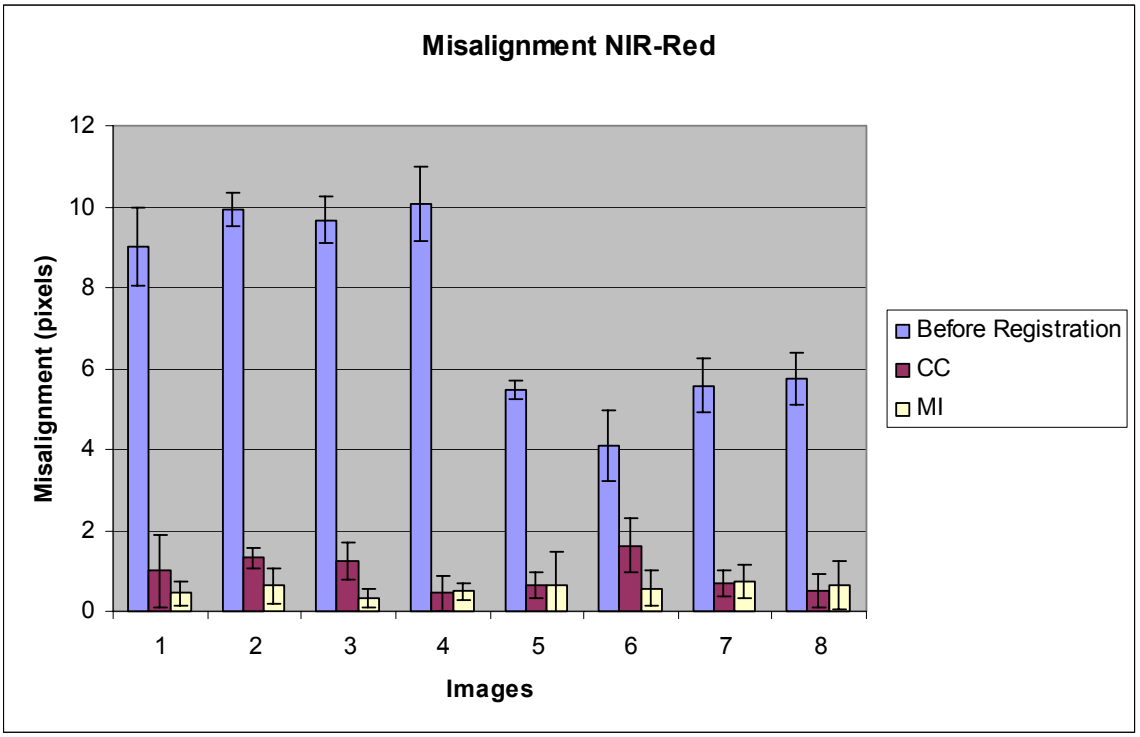


Figure 4.13 The misalignment between NIR and Red bands

Table 4.2 The misalignment between Red and Green band

Image Set	Before		After – MI		After – CC	
	μ	*std*	μ	*std*	μ	*Std*
1	22.84188	0.430907	0.478553	0.44119	2.904333	2.052186
2	22.58105	0.561803	1.172345	0.409239	0.818689	0.596935
3	22.25278	0.708408	0.915062	0.303122	0.941942	0.258833
4	15.39892	0.762159	0.813534	0.897246	0.478143	0.087595
5	14.20028	0.621466	1.040475	0.278175	0.717419	0.25248
6	13.13568	0.693013	1.75023	0.776592	0.732998	0.341982
7	14.38619	0.972493	1.405955	0.568086	0.781085	0.352194
8	14.99833	0.580894	0.702665	0.392661	0.639754	0.316416

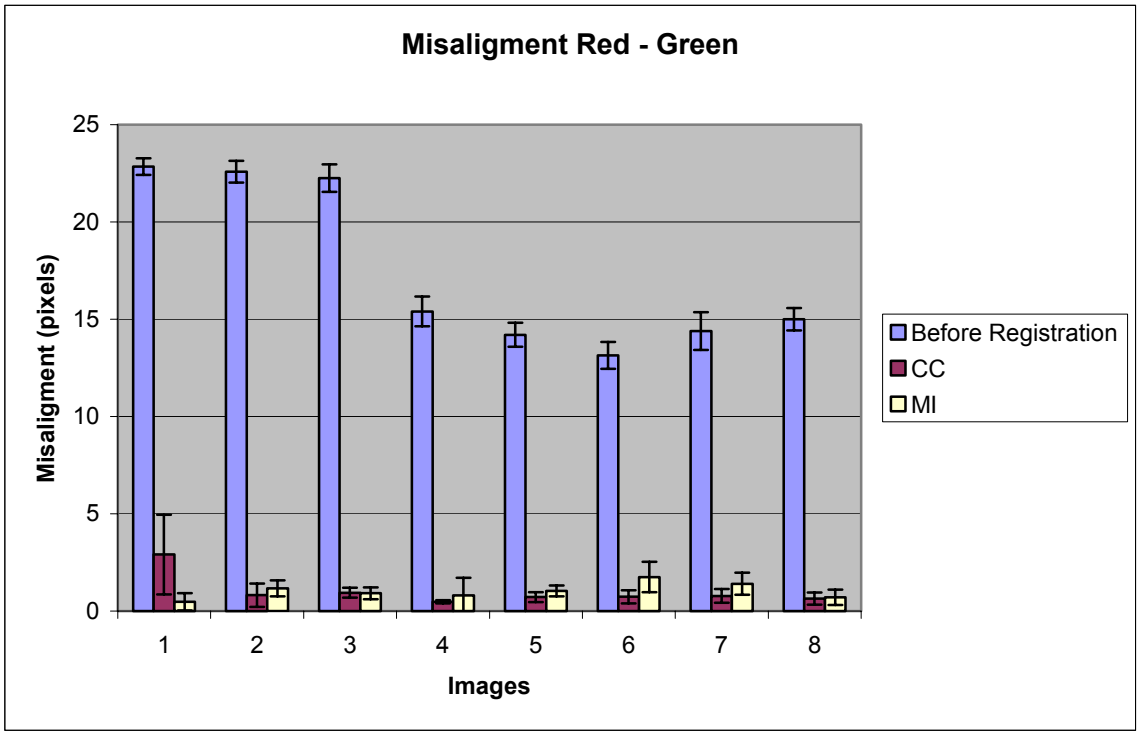


Figure 4.14 The misalignment between Red and Green bands

Table 4.3 The misalignment between NIR and Green band

Image Set	Before		After – MI		After – CC	
	μ	*std*	μ	*Std*	μ	*Std*
1	17.04342	0.741543	0.790062	0.206729	1.468698	0.558033
2	15.81644	0.495166	0.668725	0.302064	1.215448	0.509327
3	15.16733	0.200277	0.86574	0.722144	0.838388	0.382283
4	16.1422	0.770237	1.344653	0.600038	1.172082	0.325267
5	12.85259	0.643826	0.902948	0.282098	0.632651	0.697976
6	9.73757	1.025741	2.221667	1.1999	1.331696	0.568285
7	9.059672	1.080416	1.111932	0.64341	1.094056	0.418197
8	9.364144	0.2894	1.021354	0.360484	0.728143	0.383724

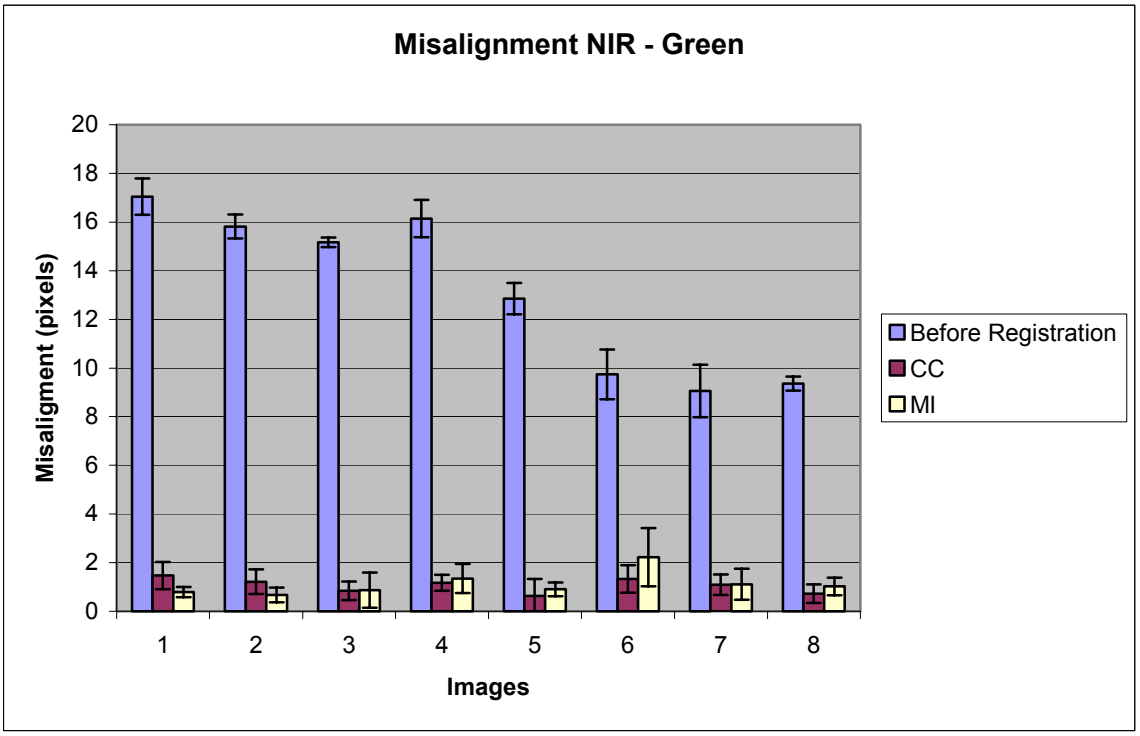


Figure 4.15 The misalignment between NIR and Green bands

4.4.1.3 Comparison between Correlation Coefficient and Mutual Information

The duration of algorithm execution on a 2.8 GHz and 2Gb Ram Dell PC is given in Table 4.4, where the same 8 frames were used as in Section 4.4.1.1. The computational cost of the MI is much higher compared to CC. Here, MI and CC used the same window size, and MI took every four gray levels as one bin for histogram computation.

Table 4.4 The registration time for image sets using MI and CC

Image Set	Time (minutes)	
	MI	CC
1	6.356483	0.70335
1	6.356483	0.70335
2	6.114133	1.00445
3	5.760617	0.996267
4	4.4696	0.901356
5	4.4015	0.882267
6	4.059167	0.947017
7	4.961633	0.7696
8	4.25445	1.00595

Table 4.5 The number of the selected control points using MI and CC

Image Sets	# of CP			
	Green-NIR		Red-NIR	
	MI	CC	MI	CC
1	27	24	30	21
2	30	28	28	29
3	30	29	29	26
4	30	25	28	23
5	30	28	29	24
6	22	17	24	20
7	23	26	24	21
8	24	29	20	21

The number of selected CPs is listed in Table 4.5, where generally MI results in more good CPs. As explained in Chapter 3, the similarity map of the CC is smooth compared to the MI; after the CP selection, the number of CPs that can be used is smaller. In other words, MI can detect CPs with better quality, which makes the following CP selection easier. A large number of good CPs means higher registration accuracy.

4.4.2 Manual Evaluation for Mosaicking

4.4.2.1 Mosaicking Accuracy

The same evaluation was also applied to image mosaics for the mosaicking data. The corner locations for a single image and previous mosaic were determined before and after the registration. The result is shown in Figure 4.16, where the Oswalt image set with panels were used for evaluation. Blue points represent the corner locations in a single image to be mosaicked, and red points represent the corner locations in the previous mosaic. Before mosaicking, the misalignment was larger than the registration of the three bands, and the rotational misalignment was obvious. After registration, the corners were matched successfully with very small misalignment. From Figure 4.16, it can be seen that MI produces more compact mosaics, because the misalignment for MI is smaller than with the CC method.

4.4.2.2 Comparison between Correlation Coefficient and Mutual Information

More comparison between MI and CC was conducted. In Table 4.6, the misalignment values between a single image and its previous mosaic image are given before and after registration using CC and MI. Obviously, CC yields lower registration accuracy.

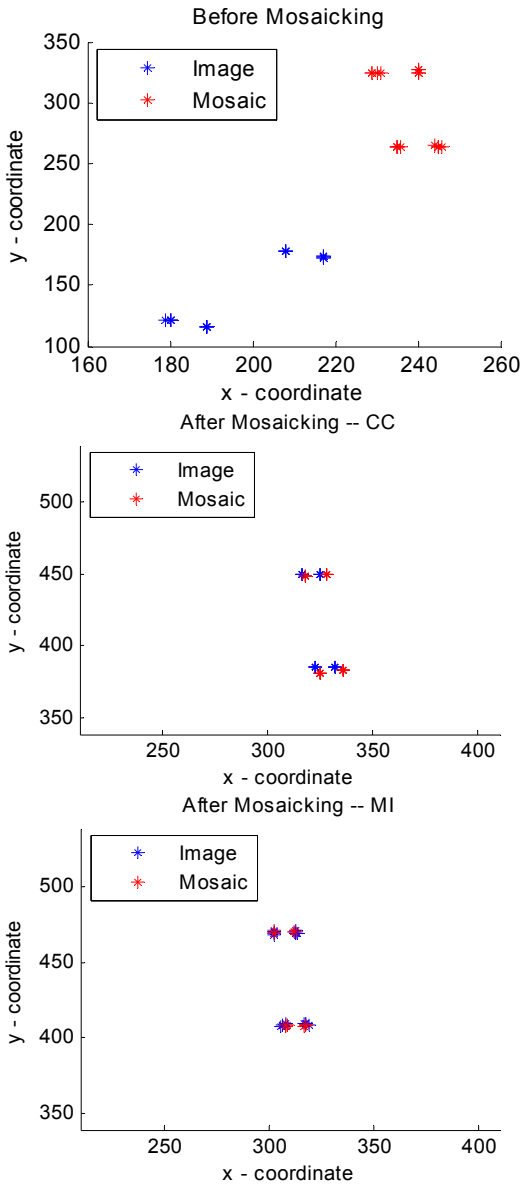


Figure 4.16 Manual evaluation for Oswalt image set using MI and CC for Mosaicking

Table 4.6 The misalignment before and after the mosaicking using MI and CC

Image Set	Before		After -- MI		After -- CC	
	μ	*std*	μ	*std*	μ	*std*
1	153.4979	4.092079	0.642836	0.446111	3.643831	1.634522
2	102.4206	3.989842	1.082004	0.480303	3.524256	2.072115

As shown in Table 4.7, the computational time is relatively higher for MI than CC, because the searching area for CPs is the overlapped area of the images.

Table 4.7 Computational time for mosaicking using MI and CC

Image Set	Time (minutes)	
	MI	CC
1	65.29773	8.796817
2	30.22078	3.548783

The numbers of the selected CPs are given in Table 4.8. The number of the selected CPs using MI is larger than the CC. As the number of the selected CPs increases, the accuracy of the registration increases. More CPs are distributed over the image scene, so the registration error decreases.

Table 4.8 The number of the selected control points using MI and CC

Image Set	# of CPs	
	MI	CC
1	15	12
2	21	14

CHAPTER V

CONCLUSION AND FUTURE WORK

5.1 Conclusion

Algorithms for automatic image registration and mosaicking are developed for a mini-UAV platform. After their installation into the ground control station with the capability of real-time image recording, the near real-time decision-making support is achievable with the final commercial products, such as CIR and NDVI images, for agricultural, forestry, and environmental studies.

The specialties of the acquired images include: most image scenes are about vegetation areas and agricultural crop fields without distinctive features; the three cameras take images in a single frame simultaneously with small shifting and rotational misalignment; every two adjacent frames taken at similar altitudes have overlapping areas and relatively large rotational misalignment. Algorithms are developed based on these image characteristics. The area-based method is employed, which is applicable when no prominent feature details are present in image scenes. The correlation coefficient (CC) and mutual information (MI) are the two metrics used for similarity comparison. Since the image scenes are taken from the vegetation areas and farms, the texture features of areas are similar to each other. Without any pre-processing, the similarity measures may lead to large registration error. So the region of interest (ROI) determination is important. Because the registration is applied to three different bands in

a single frame but small misalignment, the ROI is selected by the calculation of the entropy of non-overlapping blocks and control point detection is spatially confined within the corresponding ROIs. Mosaicking is for the same bands in adjacent frames but with large misalignment, so the ROI is determined as regions with relatively large intensity variations. Control point identification is another key step where CC or MI is used for control point detection, followed by control point selection to eliminate outliers and to ensure the quality of final control points. Affine transformation without the scaling factor is employed for the spatial transformation, and bilinear interpolation is used for spatial sampling for the transformed images. In image mosaicking, pre-introducing rotation is the major contribution, which makes the area-based method feasible when the rotational misalignment cannot be ignored.

Manual evaluation confirms the effectiveness of the developed algorithms. In particular, MI is demonstrated as a better similarity metric, which can produce more control points with good quality thereby achieving higher registration accuracy, but it is computationally more expensive.

An executable version of the software is developed for Intel PC platforms. The program does not require any additional software like MATLAB and C++. The enhancement in data quality results in more reliable data analysis products by implementing automated image registration and mosaicking software.

5.2. Future Work

The future work will be in two directions:

1. *Automatic assessment*: Currently, image registration and mosaicking are evaluated

by human inspection, or by using the manually selected control points. An automatic approach needs to be developed to avoid human-dependent errors.

2. *Image transformation*: Affine transform currently is used, which is linear. In order to more powerfully take care of the geometric distortion induced when the camera focal plane is not horizontal to the earth surface, nonlinear image mapping needs to be used.

REFERENCES

[1] L. M. G. Fonseca and B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 62, no. 9, pp. 1049-1056, 1996.

[2] NASA, "Electromagnetic Spectrum: Spectral Signatures", http://rst.gsfc.nasa.gov/Intro/Part2_5.html.

[3] 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, no. 5, pp. 2351-2362, 1999.

[4] T. Hong and R. A. Schowengerdt, "A robust technique for precise registration of radar and optical satellite images," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 5, pp. 585-593, 2005.

[5] L. Brown, "A survey of image registration techniques," *Association for Computational Machinery Computational Survey*, vol. 24, no. 4, pp. 325-376, 1992.

[6] M. Irani and S. Peleg, "Improving resolution by image registration," *Computer Vision, Graphics and Image Processing,* vol. 53, no. 3, pp. 213-239, 1991.

[7] G. A. Lampropoulos, J. Chan, J. Secker, Y. Li, and A. Jouan, "Automatic registration of Electro-optical and SAR images," *Proceedings of IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 219-226, 2003.

[8] X. Dai, S. Khorram, and H. Cheshire, "Automated image registration for change detection from thematic mapped imagery," *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, vol. 111**,** pp. 1609-1611, 1996.

[9] H. Nicholas, "New methods for dynamic mosaicking," *IEEE Transactions on Image Processing,* vol. 10, no. 8, pp. 1239-1251, 2001.

[10] M. S. Su, W. L. Hwang, and K. Y. Cheng, "Analysis on multiresolution mosaic images," *IEEE Transactions on Image Processing,* vol. 13, no. 7, pp. 952-959, 2004.

[11] Y. Du, J. Cihlar, J. Beaubien, and Rasim Latifovic, "Radiometric normalization, compositing, and quality control for satellite high resolution image mosaics over large areas," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, no. 3, pp. 2351-2362, 2001.

[12] Y. Amit and A. Kong, "Graphical templates for model registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 225-236, 1996.

[13] Y. Bentoutou, N. Taleb, K. Kpalma, and J. Ronsin, "An automatic image registration for applications in remote sensing," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, no. 9, pp. 2127-2137, 2005.

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

[15] B. K. Ghaffray and A. A. Sawchuk, "A survey of new techniques for image registration and mapping," *Proceedings of the SPIE: Applications of Digital Image Processing,* vol. 432, pp. 222-239, 1983.

[16] C. M. Lau, T. Adali, and Y. Wang, "Coregistration of PET/MR brain images by multi-feature correlation matching," *Proceedings of the 15th Fifteenth Southern Biomedical Engineering Conference*, 1996.

[17] G. Q. Maguire, M. E. Noz, E. M. Lee, and J. H. Schimpf, " Correlation methods for tomographic images using two and three dimensional techniques," *Proceedings of the 9th Conference Information Processing Medical Imaging,* pp. 266-279, 1985.

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

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

[20] R. Berthilsson, "Affine correlation," *Proceedings of the International Conference of Pattern Recognition*, Brishhbane, Australia, pp. 1458-1461, 1998.

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

[22] S. Kaneko, I. Murase, and S. Igarashi, "Robust image registration by increment sign correlation," *Pattern Recognition,* vol. 35, pp. 2223-2234, 2002.

[23] S. Kaneko, Y. Satoh, and S. Igarashi, "Using selective correlation coefficient for robust image registration," *Pattern Recognition,* vol. 36, pp. 1165-1173, 2003.

[24] D. I. Barnea and H. F. Silverman, "A class of algorithms for fat digital image registration," *IEEE Transactions on Computer,* vol. 21, no. 2, pp.179-186, 1972.

[25] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ration as a new similarity measure for multimodal image registration," *Proceedings of the 1st International Conference on Medical Image Image Computing and Computer-Assisted Intervention Lecture Notes in Computer Science,* Cambridge, USA, vol. 1496, pp. 1115-1124, 1998.

[26] S. C. Cain, M. M. Hayat, and E. E. Armstrong, "Projection based image registration in the presence of fixed-pattern noise," *IEEE Transactions on Image Processing,* vol. 10, pp. 1860-1872, 2001.

[27] P. E. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transforms," *IEEE Transactions on Geoscience Electronics,* vol. 8, pp. 353-368, 1970.

[28] P. Van Wie and M. Stein, "A Landsat digital image rectification system," *IEEE Transactions on Geoscience Electronics,* vol. 15, pp. 130-136, 1977.

[29] W. K. Pratt, "Correlation techniques of image registration," *IEEE Transactions on Aerospace and Electronic System,* vol. 10, pp. 353-358, 1974.

[30] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, *Medical Image Registration,* Boca Raton, FL, 2001.

[31] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in Medicine and Biology,* vol. 46, no. 3, pp. R1-R45, 2001.

[32] J. Aczel and Z. Daroczy, *On the Measures of Information and Their Characterizations,* New York, Academic, 1975.

[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* New York, Wiley, 1991.

[34] I. Vajda, *Theory of Statistical Inference and Information,* Dordrecht, The Netherlands, Kluwer, 1989.

[35] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision,* vol. 24, pp. 137-154, 1997.

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

[37] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3-D medical image alignment," *Pattern Recognition*, vol. 32, pp. 71-86, 1999.

[38] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging,* vol. 16, no. 2, pp. 197-198, 1997.

[39] P. Thenevaz and M. Unser, "An efficient mutual information optimizer for multiresolution image registration," *Proceedings of the IEEE International Conference on Image Processing,* Chicago, IL, pp. 833-837, 1998.

[40] P. Thenevaz and M. Unser, "A pyramid approach to sub-pixel image fusion based on mutual information," *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, pp. 265-268, 1996.

[41] P. Thenevaz and M. Unser, "Spline pyramids for inter-modal image registration using mutual information," *Proceedings of the SPIE: Wavelet Application in Signal and Image Processing,* San Diego, CA, vol. 3169, pp. 236-247, 1997.

[42] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. Van Saarloos, "Registration of stereo and temporal images of the retina," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 404-418, 1999.

[43] B. Likar and F. Pernus, "A hierarchical approach to elastic registration based on mutual information," *Image and Vision Computing,* vol. 19, pp. 33-34, 2001.

[44] A. Roche, G. Malandain, and N. Ayache, "Unifying maximum likelihood approaches in medical image registration," *International Journal of Imaging Systems and Technology,* vol. 11, pp. 71-80, 2000.

[45] A. Rangarajan, H. Chui, and J. S. Duncan, "Rigid point feature registration using mutual information," *Medical Image Analysis,* vol. 4, pp. 1-17, 1999.

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

[47] E. De 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.

[48] 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 Intelligence,* vol. 16, pp. 1156-1168, 1994.

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

[50] T. M. Lehman, "A two stage algorithm for model-based registration of medical images," *Proceedings of the International Conference on Pattern Recognition,* Brisbane, Australia, pp. 344-352, 1998.

[51] A. V. Cideciyan, "Registration of ocular fundus images," *IEEE Engineering in Medicine and Biology Magazine,* vol. 14, pp. 52-58, 1995.

[52] L. Lucchese, G. Doretto, and G. M. Cortelazzo, "A frequency domain technique for range data registration," *IEEE Transactions on Medical Imaging,* vol. 24, pp. 1468-1484, 2002.

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

[54] 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, 1999.

[55] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision,* vol. 24, pp. 137-154, 1997.

[56] H. S. Sawhney and R. Kumar, "True multi-image alignment and its application to mosaicking and lens distortion correction," *IEEE Transactions on Pattern analysis and Machine Intelligence,* vol. 21, pp. 235-243, 1999.

[57] M. Jenkinson and S. Smith, "A global optimization method for robust affine registration of brain images," *Medical Image Analysis,* vol. 5, pp. 2223-2234, 2002.

[58] A. Goshtasby, G. C. Stockman, and C.V. 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.

[59] M. Holm, "Towards automatic rectification of satellite images using feature based matching," *Proceedings of the International Geoscience and Remote Sensing Symposium*, Espoo, Finland, pp. 2439-2442, 1991.

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

[61] M. Sester, H. Hild, and D. Fritsch, "Definition of ground control features for image registration using GIS data," *Proceedings of the Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels*, Columbus, Ohio, pp. 1-7, 1998.

[62] M. Roux, "Automatic registration of SPOT images and digitized maps," *Proceedings of the IEEE International Conference on Image Processing,* Lausanne, Switzerland, pp. 625-628, 1996.

[63] P. A. Brivio, A. D. Ventura, A. Rampini, and R. Schettini, "Automatic selection of control points from shadow structures," *International Journal of Remote Sensing*, vol. 13, pp. 1853-1860, 1992.

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

[65] H. Maitre and Y. Wu, "Improving dynamic programming to solve image registration," *Pattern Recognition,* vol. 20, pp. 443-462, 1987.

[66] S. Z. Li, J. Kittler, and M. Petrou, "Matching and recognition of road networks from aerial images," *Proceedings of the 2nd European Conference on Computer Vision*, St Margherita, Italy, pp. 857-861, 1992.

[67] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679-698, 1986.

[68] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London,* vol. 207, pp. 187-217, 1980.

[69] J. B. A. Maintz, P. A. Elsen, and M.A. Viergever, "Comparison of edge-based and ridge-based registration of CT and MR brain images," *Medical Image Analysis*, vol. 1, pp. 151-161, 1996.

[70] 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.

[71] M. Ehlers, "Region-based matching for image registration in remote sensing databases," *Proceedings of the International Geoscience and Remote Sensing Symposium*, Espoo, Finland, pp. 2231-2234, 1991.

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

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

[74] C.Y. Wang, H. Sun, S. Yadas, and A. Rosenfeld, "Some experiments in relaxation image matching using corner features," *Pattern Recognition,* vol. 16, pp. 167-182, 1983.

[75] K. Rohr, "Localization properties of direct corner detectors," *Journal of Mathematical Imaging and Vision*, vol. 4, pp. 139-150, 1994.

[76] Z. Zheng, H. Wang, and E. K. Teoh, "Analysis of gray level corner detection," *Pattern Recognition Letters*, vol. 20, pp. 149-162, 1999.

[77] L. Kitchen and A. Rosenfeld, "Gray-level corner detection," *Pattern Recognition Letters*, vol. 1, pp. 95-102, 1982.

[78] W. Forstner and E. Gulch, "A fast operator for detection of precise location of distinct points, corners and centers of circular features," *Proceedings of the ISPRS Workshop on Fast Processing of Photogrammetric Data,* Interlaken, Switzerland, pp. 281-305, 1986.

[79] M. Trajkovic and M. Hedley, "Fast corner detection," *Image and Vision Computing*, vol. 16, pp. 75-87, 1998.

[80] A. Goshtasby and G. C. Stockman, "Point pattern matching using convex hull edges," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, pp. 631-637, 1985.

[81] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric correspondence and chamfer matching: two new techniques for image matching," *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, Cambridge, Massachusetts, pp. 659-663, 1977.

[82] F. Zana and J. C. 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, 1999.

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

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

[85] R. S. Mitra and N. N. Murthy, "Elastic maximal matching," *Pattern Recognition*, vol. 24, pp. 747-753, 1991.

[86] S. Ranade and A. Rosenfeld, "Point pattern matching by relaxation," *Pattern Recognition*, vol. 12, pp. 269-275, 1980.

[87] G. Medioni and R. Nevatia, "Matching images using linear features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 675-685, 1984.

[88] K. E. Price, "Relaxation matching techniques – a comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 617-623, 1985.

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

[90] R. K. Sharma and M. Pavel, "Multisensor image registration," *Proceedings of the Society for Information Display,* vol. 28, pp. 951-954, 1997.

[91] R. Turcajova and J. Kautsky, "A hierarchical multiresolution technique for image registration," *Proceedings of the SPIE Mathematical Imaging: Wavelet Applications in Signal and Image Processing,* Colorado, vol. 2825, pp. 686-696, 1996.

[92] J. Le Moigne, "Parallel registration of multi-sensor remotely sensed imagery using wavelet coefficients," *Proceedings of the SPIE: Wavelet Applications,* Orlando, Florida, vol. 2242, pp. 432-443, 1994.

[93] M. Ehlers and D. N. Fogel, "High-precision geometric correction of airborne remote sensing revisited: the multiquadric interpolation," *Proceedings of SPIE: Image and Signal Processing for Remote Sensing*, vol. 2314, pp. 814-824, 1994.

[94] M. Fornefett, K. Rohr, and H. S. Stiehl, "Radial basis functions with compact support for elastic registration of medical images," *Image and Vision Computing,* vol. 19, pp. 87-96, 2001.

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

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

[97] C. R. Appledorn, "A new approach to the interpolation of sampled data," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 369-376, 1996.

[98] P. Thevenaz, T. Blu, and M. Unser, "Image interpolation and resampling," *Handbook of Medical Image Processing*, Academic Press, New York, 2003.

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

[100] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 1049-1075, 1999.

[101] B. Guindon, "Assessing the radiometric fidelity of high resolution satellite image mosaics," *J. Photogramm. Remote Sensing*, vol. 52, pp. 229-243, 1997.

[102] B. Guindon, "Utilization of Landsat pathfinder data criterion of large area mosaic," *ACSM/ASPRS Annual Convention and Exposition Technical Papers,* Charlotte, NC, vol. 2, pp. 144-153, 1995.

[103] J. R. Jenson, "Urban/Suburban land use analysis," *Manual of remote Sensing,* 2nd ed., American Society of Photogrammetry, vol. 2, pp. 1571-1666, 1983.

[104] J. R. Schott, C. Salvaggio, and W. J. Volchock, "Radiometric scene normalization using pseudo invariant features," *Remote Sensing of Environment*, vol. 26, pp. 1-16, 1988.

[105] F. G. Hall, D. E. Strebel, J. E. Nickeson, and S. J. Goetz, "Radiometric rectification: toward a common radiometric response among multidate, multisensor images," *Remote Sensing of Environment*, vol. 35, pp. 11-27, 1991.

[106] D. L. Milgram, "Computer methods for creating photomosaics," *IEEE Transactions on Computers*, vol. 24, pp. 1113-1119, 1975.

[107] D. L. Milgram, "Adaptive techniques for photomosaicking," *IEEE Transactions on Computers*, vol. 26, pp. 1175-1180, 1977.

[108] S. Peleg, "Elimination of seams from photomosaics," *Computer Graphics and Image Processing*, vol. 16, pp. 90-94, 1981.

[109] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *Association for Computational Machinery Transactions Graphics,* vol. 2, pp. 217-236, 1983.