
UCGE Reports
Number 20190

Department of Geomatics Engineering

# Semi-Automatic Registration of Multi-Source Satellite Imagery with Varying Geometric Resolutions

**(URL: http://www.geomatics.ucalgary.ca/links/GradTheses.html)**

**by**

**Rami Al-Ruzouq**

**May 2004**

UNIVERSITY OF CALGARY

# Semi-Automatic Registration of Multi-Source Satellite Imagery with Varying Geometric Resolutions

by

Rami Al-Ruzouq

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

MAY, 2004

## ABSTRACT

Image registration concerns the problem of how to combine data and information from multiple sensors in order to achieve improved accuracy and better inferences about the environment than could be attained through the use of a single sensor. Registration of imagery from multiple sources is essential for a variety of applications in remote sensing, medical diagnosis, computer vision, and pattern recognition. In general, an image registration methodology must deal with four issues. First, a decision has to be made regarding the choice of primitives for the registration procedure. The second issue concerns establishing the registration transformation function that mathematically relates images to be registered. Then, a similarity measure should be devised to ensure the correspondence of conjugate primitives. Finally, a matching strategy has to be designed and implemented as a controlling framework that utilizes the primitives, the similarity measure, and the transformation function to solve the registration problem. The Modified Iterated Hough Transform (MIHT) is used as the matching strategy for automatically deriving an estimate of the parameters involved in the transformation function as well as the correspondence between conjugate primitives. The MIHT procedure follows an optimal sequence for parameter estimation. This sequence takes into account the contribution of linear features with different orientations at various locations within the imagery towards the estimation of the transformation parameters in question.

Accurate co-registration of multi-sensor datasets captured at different times is a prerequisite step for a reliable change detection procedure. Once the registration problem has been solved, the suggested methodology proceeds by detecting changes between the

registered images. Derived edges from the registered images are used as the basis for change detection. Edges are utilized because they are invariant regardless of possible radiometric differences between the images in question. Experimental results using real data proved the feasibility and robustness of the suggested approach.

## ACKNOWLEDGEMENTS

I wish to thank my adviser, Professor Ayman Habib, for his tenacious dedication and support offered during my graduate studies, research, and dissertation work. His unflagging enthusiasm and encouragement propelled my work while his intellectual vigor greatly inspired the methodology, approaches, and critical notions adopted in my research.

My parents and family always gave me boundless love and encouragement to follow whatever path I chose. Their continual support gave me the strength, motivation, and willingness to continue my research. My gratitude is beyond words.

I wish to express my deep appreciation to Dr. Naser El-Sheimy, Dr. Caterina Valeo, Dr. Nigel Waters, and Dr. James Bethel for carefully reading and providing comments concerning various aspects of this research.

I appreciate the help I received from my colleagues in the digital photogrammetry group. My special gratitude goes to my friend Mwafag Ghanma for his kind encouragement and support.

The author gratefully acknowledges Dr. Kyungok Kim from the Korean Electronics and Telecommunications Research Institute (ETRI) for providing the satellite imagery, which were used in the experimental results section. Discussions with Dr. Kim and her team members were invaluable for this work.

## DEDICATION

*To my beloved parents*

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF SYMBOLS

| | |
|---|---|
| $\rho$, $\theta$ | Polar Coordinates Representing a Line Segment |
| $\delta\kappa$ | Non-Orthogonality Angle |
| $\kappa$ | Rotation Angle |
| $s_x$ | Scale Factor along x-Axis |
| $s_y$ | Scale Factor along y-Axis |
| $\hat{\sigma}_o^2$ | Variance Component |
| $x_T$ , $y_T$ | Shifts in x and y Direction |
| $\vec{b}$ | Image Base (Vector between Two Perspective Centers of the Stereopair) |
| $\vec{p}_l, \vec{p}_r$ | Vectors from the Perspective Center to a Conjugate Point in the Left and Right Images Respectively |
| R | Rotation Matrix |
| $s$ | Scale Factor |
| $x$ , $y$ | Image Coordinates in the Reference Image |
| X, Y, Z | Ground Point Coordinates |
| $x'$ , $y'$ | Image Coordinates in the Input Image |
| $X_0$, $Y_0$, $Z_0$ ω, φ, κ | Exterior Orientation Parameters ($X_0$, $Y_0$ and $Z_0$ Represent the Position of Perspective Center with Respect to Ground Coordinate System, where ω, φ and κ Represent the Rotation Angles between the Ground and Image Coordinate Systems) |

| | |
|---|---|
| $x_p$, $y_p$, c | Interior Orientation Parameters (Calibrated Principal Point Position and Principal Distance of The Camera with Respect to Image Coordinate System) |

## LIST OF ABBREVIATIONS

| | |
|---|---|
| 2-D | Two Dimensional |
| 3-D | Three Dimensional |
| CCD | Charge Coupled Device |
| CVA | Change Vector Analysis |
| DEM | Digital Elevation Models |
| DLT | Direct Linear Transformation |
| EOP | Exterior Orientation Parameters |
| FOV | Field of View |
| GCP | Ground Control Point |
| GPS | Global Positioning System |
| INS | Inertial Navigation System |
| IOP | Interior Orientation Parameters |
| LOG | Laplacian-of-Gaussian |
| MIHT | Modified Iterated Hough Transform |
| MMS | Mobile Mapping Systems |
| MSS | Multi-Spectral Scanner |
| TM | Thematic Mapper |
| NDVI | Normalized Difference Vegetation Index |
| NIR | Near Infrared Red |
| PCA | Principal Component Analysis |
| RFM | Rational Function Models |
| RMSE | Root Mean Square Error |

| | |
|---|---|
| VIS | Visible |

# CHAPTER 1

# INTRODUCTION

## 1.1 Problem Definition

Image registration concerns the problem of how to combine data and information from multiple sensors in order to achieve improved accuracy and better inferences about the environment than could be attained through the use of a single sensor data. In some applications, such as interactive remote sensing and medical imaging, image registration is the final goal; in other applications, it is a prerequisite for accomplishing high-level tasks such as sensor fusion, surface reconstruction, and object recognition. With the flux of high resolution scenes captured by space-borne platforms (e.g., LANDSAT-7, IKONOS, QUICKBIRD, ORBVIEW, EROS-A1, and SPOT-5), there is an increasing need for a robust registration technique that can tolerate varying geometric resolutions of the available scenes.

Automatic and even manual registration of imagery remains challenging for several reasons. First, images are usually acquired using different sensor types, each having its inherent noise. Furthermore, radiometric as well as geometric properties of the same object in the involved imagery might differ as a result of changes in the sensor view point, imaging methodology, imaging conditions (e.g., atmospheric changes, cloud coverage, and shadows), and spectral sensitivity of the implemented imaging systems (e.g., panchromatic, multi- and hyper-spectral imaging systems). Finally, the registration

process can be complicated by changes in object space caused by movements, deformations, and urban development between the epochs of capture associated with the involved images.

Although a vast body of research has dealt with automatic image registration, we still do not have a methodology that meets the current challenges posed by image registration. This research will investigate and develop a semi-automatic, accurate, and robust registration paradigm that can cope with those challenges.

## 1.2 Motivation

In recent years, there has been an enormous increase in the volume of remotely sensed images being acquired by an ever-growing number of earth observation satellites. This surge in use mandates the development of accurate and robust registration procedures that can handle imagery with varying geometric and radiometric properties. Moreover, the need to develop a registration methodology is motivated by the fact that its application areas span the following fields (Brown, 1992)**:**

- Remotely sensed data processing for military and civilian applications in agriculture, geology, oceanography, oil, mineral exploration, pollution control, urban expansion monitoring, forestry, and target location and identification.

- Medical image analysis for diagnosis purposes such as tumor detection and disease localization. Image registration can be also useful for biomedical applications such as

classification of microscopic images of blood cells, cervical smears, and chromosomes.

- Computer vision and pattern recognition applications such as segmentation, object recognition, shape reconstruction, motion tracking, stereo-mapping, and character recognition.

## 1.3 Scope of the Research

This research aims at developing a registration methodology for handling imagery with varying geometric and radiometric properties. This thesis describes in detail the essential components and the suggested implementation of an effective image registration methodology, which includes selecting appropriate primitives, transformation function, similarity measure, and matching strategy.

The first stage in this study includes an investigation into the most appropriate primitives that can be used for image-to-image registration. Several primitives (e.g. points, linear features, and areal features) can be used in the registration process (Fonseca and Manjunath, 1996). Linear features, and more specifically straight line features, will be the main focus of this study. The rationale behind using straight lines instead of points and areal features will be discussed. Moreover, a comparative study will be done of the performance of point and linear features as the registration primitives.

Investigation of the most appropriate registration transformation functions is the second stage. Within this stage, simplified (i.e., approximate) as well as rigorous (i.e., based on

the geometric characteristics of the imaging sensor) transformation functions will be analyzed (Habib and Morgan, 2002). A comprehensive analysis will be conducted to evaluate the validity of the transformation function and the gained improvement in the registration quality, as contrasted with the increased complexity of the registration methodology.

The third stage concerns the development of a similarity measure, which mathematically describes the coincidence of conjugate elements after the application of the registration transformation function. A similarity measure incorporates the attributes of the registration primitives to derive the necessary constraints that can be used to estimate the parameters of the transformation function relating the images to be registered.

Automating the solution to the registration problem requires the establishment of a controlling framework that utilizes the primitives, similarity measure, and transformation function. This framework is usually referred to as the matching strategy and is the final stage. In this research, the Modified Iterated Hough Transform (MIHT) first proposed by Habib et al (2001a, 2001b) is used as the matching strategy. Such a methodology is attractive since it allows for simultaneous matching and parameter estimation. MIHT has been successfully implemented in several photogrammetric operations such as automatic single photo resection and relative orientation (Habib et al., 2001a, 2001b; Habib and Kelley 2001a, 2001b).

In general, the key contributions of this thesis are as follows:

- Utilizing straight-line segments in image-to-image registration as a remedy for expected geometric differences between multi-source satellite images. This is an important means of increasing the robustness of the registration procedure to cope with the flux of newly available high-resolution satellite scenes (e.g., IKONOS, LANDSAT-7, SPOT-5, EROS-A1, KOMPSAT-II, QUICKBIRD, and ORBVIEW).

- Showing that line segments are superior to point primitives in identifying the registration primitives in multi-resolution satellite imagery.

- Introducing a new mathematical model representing the similarity measure, which describes the necessary constraints for ensuring the correspondence of conjugate primitives. The similarity measure has been developed in light of the fact that the end points of conjugate line segments are not identical. This is a critical consideration; because of varying geometric and radiometric properties of the respective imaging systems and different imaging conditions, the manual or automatic extraction methodology might not reliably identify corresponding points in the reference and input images.

- Establishing an automatic matching strategy (MIHT) that utilizes the introduced similarity measure together with the transformation function to establish the correspondence between the extracted primitives and simultaneously solve for the parameters involved in the registration transformation function. This method would allow for investigating and evaluating the appropriateness of the selected registration transformation function. Previous research has rarely considered this issue. The

MIHT approach does not assume a one-to-one correspondence between the primitives, nor does it require approximate registration of the involved scenes.

- Deriving the optimal sequence for parameter estimation within the MIHT procedure. This process takes into account the contribution of linear features with different orientations at various locations within the imagery to estimate the transformation parameters in question. An optimal sequence is derived for 2-D similarity and affine transformation functions. The derivation has been performed by analyzing the deviations from the similarity measure constraints associated with line segments with different orientations at various regions within the imagery as a result of incremental changes in the transformation parameters.

- Detecting changes between the registered images. After the registration problem has been solved, the suggested methodology derives edges from the registered images and uses those edges as the basis for change detection. Edges are utilized because they are invariant regardless of possible radiometric differences between the images in question.

## 1.4 Definition of Terms

This section provides the definitions of some frequently used terms.

**Reference image and input image:** One of two given images is called a reference image; and the other, an input image. Image registration attempts to find the transforms in the input image relative to the reference image. The reference image is assumed to be of

good quality as well as high resolution, and more representative of the object space**.** A reference image would have no cloud, good contrast and negligible geometrical distortions. The input image, on the other hand, is relatively lower in resolution than the reference image. Some coverage such as fog or clouds may be present, along with the geometrical distortions.

**Conjugate or corresponding features:** This term refers to the primitives (i.e., points, linear features, and areal features) on different images, which represent the same object space feature. In the case of image registration, conjugate or corresponding features are the matched (coupled pairs) primitives in the reference and input images. Although semantically the term "corresponding" is more general, and "conjugate" refers to two images only, these terms are used interchangeably.

**Geometric resolution or resolving power:** These interchangeable terms refer to the smallest visible separation between similar objects that can be clearly reproduced by a remote sensing system. Geometric resolution is usually expressed as the maximum number of line pairs per unit length.

**Residual:** The residual is the linear distance between a fixed reference point (ground control point) and the point that gives the best fit, as determined through the application of the transformation function to the observed data.

**Root Mean Square Error (RMSE):** One determines the RMSE by calculating the deviations of points from their true position, averaging the squares of such deviations, and then taking the square root of the average.

**Accuracy:** Accuracy refers to the quality of the nearness to the truth if one assumes no biases in the measurement procedure. Accuracy represents the relationship of a set of features to a defined reference system and is expressed as the RMSE of a set of derived points.

## 1.5 Thesis Outline

The balance of this thesis is organized as follows. Chapter 2 contains a comprehensive review of the image registration paradigm elements which include selecting appropriate primitives, transformation function, similarity measure, and matching strategy. This chapter then explains various existing techniques and studies for image registration and their limitations. Finally, the importance of accurate image registration for reliable change detection techniques is investigated and a review of the change detection algorithms is presented.

Chapter 3 is dedicated to selecting the most appropriate primitive, transformation function, similarity measure, and matching strategy for incorporation into the suggested image registration paradigm to ensure accurate, and robust image-to-image registration. In this chapter, the motivation for using linear features in photogrammetric applications and in particular automatic image registration is explained. This explanation is followed by an investigation of the most appropriate registration transformation functions; detailed analysis of rigorous and approximate models is provided and the validity of such models for different applications is discussed. The calculation of the similarity measure, which mathematically describes the coincidence of conjugate line segments after application of

the registration transformation function, is the next step. Finally, the Hough Transform is introduced, and the development of the MIHT algorithm, which is a foundation of this work, is described.

Within the proposed image registration methodology, the MIHT procedure follows an optimal sequence for parameter estimation. This will be the main issue of Chapter 4 where the optimal sequence of the transformation functions (2-D similarity and affine transformation) parameters are established.

Chapter 5 deals with the change detection algorithm. Once the registration problem has been solved, the suggested methodology detects changes between the registered imagery. The suggested change detection algorithm depends on geometrical properties of the images rather than on radiometric properties. This algorithm is explained in Chapter 5.

Chapter 6 describes the experiments carried out to demonstrate the feasibility and robustness of the proposed algorithm for image registration and change detection.

In Chapter 7, conclusions are drawn and relevant suggestions for future research are given.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Introduction

The recent development of new sensors has created a need for data processing techniques that can fuse observations from a variety of different sensors. Image registration aims at geometrically aligning two or more images so that corresponding pixels or their derivatives (edges, corner points, etc.)**,** representing the same underlying structure in the object space, may be integrated or fused. Registration of images captured by different types of sensors under different conditions is a challenging problem. The difficulty comes from the fact that these images have varying radiometric and geometric resolutions and properties.

In general, an image registration methodology must deal with four issues. First, a decision has to be made regarding the choice of primitives for the registration procedure. Then, a similarity measure should be devised to ensure the correspondence of conjugate primitives. The third issue concerns establishing the registration transformation function that mathematically relates the images under consideration. Finally, a matching strategy has to be designed and implemented as a controlling framework that utilizes the primitives, the similarity measure, and the transformation function to solve the registration problem.

This chapter reviews the basic components of the image registration process and then analyzes the various techniques that are currently used for image-to-image registration and change detection. The chapter begins with an overview of different alternatives that can be used as registration primitives (Section 2.2); this is followed by a description of different transformation functions and their characteristics considering various imaging systems used to capture the images (Section 2.3). Then, Section 2.4 reviews different criteria for the similarity measure. An overview of the most common image registration algorithms and their associated advantages and limitations is given in Section 2.5. Finally, Section 2.6 describes various procedures that have been developed for change detection purposes.

## 2.2 Primitives

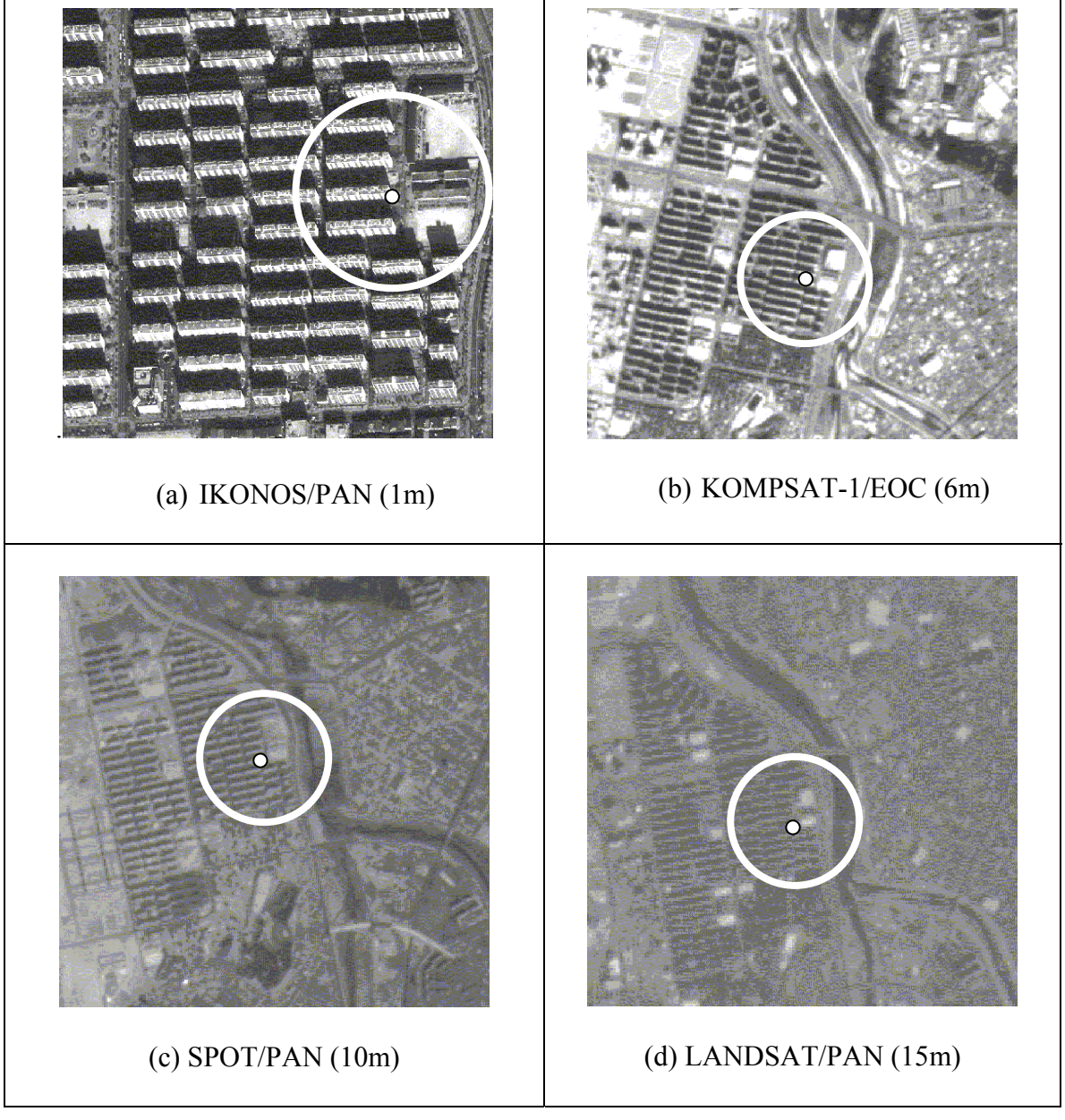
Registration primitives encompass the domain in which information is extracted from input imagery for the registration process. Hence, to carry out the registration process, the appropriate primitives must be chosen. The three fundamental and most commonly used spatial domain features are points, lines and homogenous/areal regions, Figure 2.1. Candidate features include lakes, rivers, cost-lines, roads or similar dominant man-made or natural structures. Each of these features will be assigned one or more point locations (e.g. centroid of area, line endings, etc.) to be used as the registration primitive (Fonseca and Manjunath, 1996).


**Figure 2.1:** Registration Primitives Alternatives

Traditional procedures for manually registering an image pair require interactive selection of tie points in each image. The points are then used to determine the parameters of a registration transformation function, which is subsequently used to resample one of the images into the reference frame associated with the other image. However, such a procedure, which relies on manual identification of conjugate points, can lead to inaccurate results and is slow to execute, especially if large numbers of images with varying geometric and radiometric properties need to be registered. One could even argue that manual registration of such imagery using points would be very difficult. For example, a visual inspection of the imagery in Figure 2.2 shows that manual identification of conjugate points is extremely difficult, if not impossible.

Automation of the registration procedure requires the replacement of manual tie point selection with automatic algorithms for locating corresponding points in both images (Brown, 1992). Points can be automatically extracted using an interest operator (Förstner

and Gulch, 1987; Moravec, 1977). Then, extracted points can be automatically matched through consideration of the radiometric properties of the surrounding pixels or the geometric distribution of the whole set of selected points across the entire image (Boardman et al., 1996).



(a) IKONOS/PAN (1m)

(b) KOMPSAT-1/EOC (6m)

(c) SPOT/PAN (10m)

(d) LANDSAT/PAN (15m)

**Figure 2.2:** Scenes with Varying Geometric and Radiometric Properties

Automatic extraction of points based on the radiometric information results in different sets of points from each image due to varying radiometric properties of the involved imagery. This situation extends to the problem of finding conjugate points (matching) where point extraction algorithms are not likely to be able to identify the same point. In other words, for multi-source imagery with varying geometric and radiometric resolutions, the texture and gray levels at the location of conjugate points are unlikely to be similar. Therefore, automatically or manually extracted points are difficult to match and are not suitable primitives for registration.

Consequently, linear and areal features are other promising alternative primitives that are more suited for multi-source image registration; since the geometric distribution of the pixels making up the feature, rather than their radiometric attributes, can be used in the matching. Linear features can be extracted through the use of derivative-based edge detectors (Pratt, 1991) or line extraction algorithms such as Hough transform (Hough, 1962). On the other hand, areal features (patches) can be extracted using classification or segmentation algorithms (Gonzalez and Woods, 1992).

Areal primitives (e.g., lakes, oceans, and homogeneous regions), might not be always available especially in the case of satellite scenes over urban areas. Moreover, registration procedures based on areal primitives use the centers of gravity of these features as the registration primitives. The estimated centers of gravity are susceptible to potential errors associated with the identified boundaries of these patches. Linear features are more appropriate than areal features in terms of availability in nature, complexity of extraction algorithms, and existence of geometric constraints. Areal features can be represented as a

sequence of linear features along the boundaries. For change detection applications, linear features can be broken into smaller subsets, which can be individually matched. However, dividing an areal feature into smaller subsets is not a trivial task.

## 2.3 Transformation Functions

The second issue in a registration procedure involves the establishment of the transformation function that mathematically describes the mapping between imagery in question. In other words, given a pair of images, reference and input images, the transformation function attempts properly to overlay these images. The functions, used to align two images, may be global or local. A global transformation is given by a single set of equations, which optimally registers all the pixels in the two images. Local transformations map the images depending on the spatial location; this results in several sets of equations for one map. Local transformations are usually more accurate but also more computationally demanding (Fonseca and Manjunath, 1996).

For different imaging systems, geometric distortions vary considerably with different factors such as the platform (airborne versus satellite), the sensor (LANDSAT versus IKONOS), the total field of view, and the scanning trajectory. To overcome the problem of geometric distortions, several types of transformation functions have been considered, starting with the two dimensional conformal transformation in Equation 2.1, also known as *2-D similarity*. This transformation is sufficient to match two images with rigid-body distortion (Brown, 1992) where the true shape is retained. This is a four-parameter

transformation that includes two translations in x- and y-directions, one scale and one rotation.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x_T \\ y_T \end{bmatrix} + s \begin{bmatrix} \cos\kappa & \sin\kappa \\ -\sin\kappa & \cos\kappa \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.1}$$

where

$s$ : Scale factor

$x_T$ & $y_T$ : Shifts in x and y direction

$\kappa$ : Rotation angle

$x$ & $y$ : Image coordinates in the reference image

$x'$ & $y'$ : Corresponding image coordinates in the input image.

At least two tie points are required to solve for the parameters of the 2-D similarity transformation. However, using only the minimum number of tie points is unwise, since it allows no room for monitoring observation errors. An increase in accuracy of the results is accomplishable only through the use of many well distributed tie points across the images.

The *affine transformation* in Equation 2.2 is frequently used to obtain a mapping between two coordinate systems. There are two more parameters than in the 2-D similarity transformation. Additional allowance is made for two different scale factors, one in the x-direction and the other in the y-direction, and there is a nonorthogonality correction between the x and y axes. In general, the more parameters included in the transformation function, the greater the ability to compensate for possible distortions between the two involved images.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x_T \\ y_T \end{bmatrix} + \begin{bmatrix} s_x \cos\kappa & s_y \sin(\kappa + \delta\kappa) \\ -s_x \sin\kappa & s_Y \cos(\kappa + \delta\kappa) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.2}$$

where

$s_x$ : Scale factor along x-axis

$s_y$ : Scale factor along y-axis

$\delta\kappa$ : Non-orthogonality angle

$x_T$ & $y_T$ : Shifts in x and y direction

$\kappa$ : Rotation angle

$x$ & $y$ : Image coordinates in the reference image

$x'$& $y'$ : Corresponding image coordinates in the input image.

At least three tie points are required to solve for the parameters of the affine transformation. This transformation carries parallel lines into parallel lines, preserves collinearity (i.e., all points lying on a line before transformation still lie on a line after transformation) and ratios of distances (i.e., the midpoint of a line segment remains the midpoint after transformation). The affine transformation will not preserve orthogonality.

A transformation that maps lines to lines, and does not necessarily preserve parallelism, is the *projective transformation* (Equation 2.3), also known as eight-parameter transformation. It is the appropriate transformation to use when the transformation takes place between two planes.

$$x' = \frac{a_0 + a_1 x + a_2 y}{a_3 x + b_3 y + 1}$$

$$y' = \frac{b_0 + b_1 x + b_2 y}{a_3 x + b_3 y + 1} \qquad (2.3)$$

where

$x$ & $y$ : Image coordinates in the reference image

$x'$& $y'$ : Corresponding image coordinates in the input image.

In fact, if $a_3$ and $b_3$ in Equation 2.3 are equal to zero, these equations become the affine transformation. With eight unknown parameters, this transformation requires a minimum of four tie points.

Although the previous three transformation types are the most commonly used, higher order polynomials and surface splines have been proposed to overcome the problems of significant geometric distortions (Flusser, 1992; Goshtasby, 1988; Goshtasby et al., 1986). For example, the second order polynomial transformation in Equation 2.4 includes twelve parameters and can be expressed as follows:

$$x' = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2$$

$$y' = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 xy + b_5 y^2 \qquad (2.4)$$

where

x & y : Image coordinates in the reference image

x′& y′ : Corresponding image coordinates in the input image.

Prior research did not investigate the validity of the registration transformation function. Simplified and sometimes incorrect registration transformation functions are assumed in earlier studies. Moreover, only direct image-to-image transformations were considered. Further investigation of the object-to-image transformation was not taken as a preliminary step, so that the sensor geometry and complexity of the object space could be explored to determine a valid image-to-image transformation.

## 2.4 Similarity Measure

The next step in the registration paradigm is the selection of the similarity measure, which describes the necessary constraints for ensuring the correspondence of conjugate primitives. The similarity measure formulation depends on the selected registration primitives and their respective attributes. The division of similarity measure criteria falls into two broad categories: radiometric and geometric similarity measures.

### 2.4.1 Radiometric Similarity Measure

Radiometric similarity measure describes the degree of similarity between the gray level distribution functions at the vicinity of the selected primitives. Small windows composed of gray values serve as matching primitives where the center pixel of the window can be used for the definition of the location of a point to be matched. For point primitives, radiometric similarity measure plays an important role in both interest point extraction and matching criteria. Interest points are image locations where the interest operator computes a high variance value. In other words, a certain pixel within the image

corresponds to an interesting point if there is a significant difference between the gray-value at this pixel and its neighbouring pixels. On the other hand, point matching compares the gray level distribution of a small sub image patch with its counterpart in the other image (Schenk, 1999). Figure 2.3 introduces a frequently used terminology. The template is an image patch, which usually remains fixed in the input image. The search window refers to the search space within which image patches (Matching window) in the reference image are compared with the template. Cross-correlation and least squares matching are the best known criteria for similarity measure. These criteria represent a quantitative measure of how well conjugate points corresponds to each other.

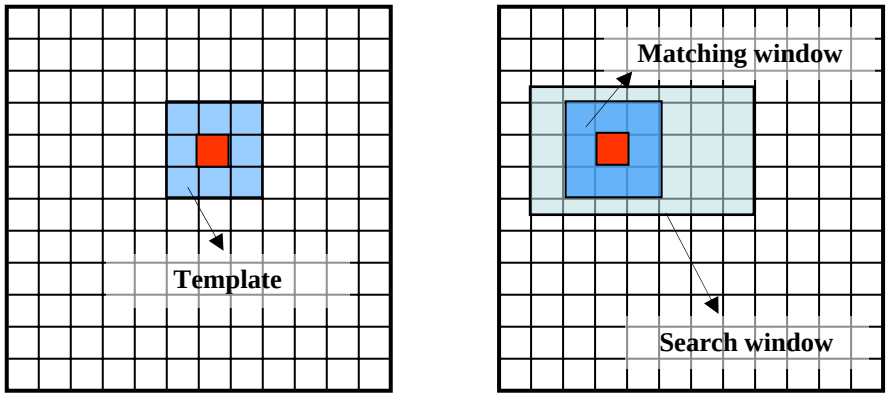



**Figure 2.3:** Concept of Area Based Matching

In cross-correlation, the idea is to measure the similarity between the template and the matching window. The template window moves over the search window where a correlation coefficient between the template and matching window is calculated at the center of each window (point primitive). The position of the conjugate points is given by the position of the maximum correlation coefficient. The correlation coefficient can take

values that range from -1 to +1 where zero indicates no similarity at all, -1 indicates an inverse similarity and +1 indicates a perfect match (the highest possible similarity).

In least square matching, the idea is to minimize the gray level differences between the template and the matching window, where the position and shape of the matching window are parameters to be estimated in the adjustment process. The position and shape of the matching window are changed until the gray level differences between the template and the matching window reach a minimum (Schenk, 1999). Least square matching is sensitive to the approximate values needed during the adjustment process where it might diverge if bad approximations are selected.

Although cross-correlation and least square matching are very successful in certain situations, these methods suffer from a number of limitations. The images in question must be radiometrically very similar, preferably imaged by the same sensor. However, gray level characteristics of the images can vary from sensor to sensor; hence, correlation measures become unreliable (Fonseca and Manjunath, 1996). Moreover, applying cross-correlation requires two images with the same geometric resolution, and this is not the case with existing satellite images (i.e., IKONOS (1m), SPOT (10m), LANDSAT (30m), etc.). Furthermore, the success of these methods depends on external influences, such as illumination and atmospheric conditions. For these reasons, the radiometric similarity measure, in general, is not suitable for images with varying geometric and radiometric resolutions.

Primitives other than points can also be extracted on the basis of radiometric properties, such as edge detection algorithms for lines and segmentation and classification

procedures for areal primitives. However, finding matching primitives by depending on radiometric attributes is difficult if not impossible. Some criteria of similarity measures depend on radiometric properties such as the average brightness for regions, the difference in gray value or gray value variance between two adjacent regions; however, these attributes are ineffective unless geometric attributes (e.g., area of regions and length of lines) are incorporated. Moreover, these criteria are invalid for dealing with images with various radiometric resolutions where conjugate regions appear with different gray values.

### 2.4.2 Geometric Similarity Measure

A geometric similarity measure mathematically describes the fact that conjugate primitives should coincide with each other after application of the proper registration transformation function. Geometric similarity measure depends on the selected registration primitive (e.g., points, linear features, areal regions) as well as the registration transformation function (e.g., 2-D similarity or affine transformation).

Feature-based matching techniques do not use the gray levels themselves as the description of the images; rather, they employ an abstract image representation derived through a feature extraction algorithm. Feature-based matching employs conjugate features appearing in both images. One important requirement is that the features be robust against changes in sensor geometry, wavelength and noise characteristics (Fonseca and Manjunath, 1996). Moreover, features should be distinct with respect to their neighborhoods, invariant with respect to geometric and radiometric influences, and stable with respect to noise (Förstner 1986).

Each feature is characterized by a set of local or global attributes. An example of a local attribute would be the image coordinates of a point where the distance between the conjugate points after application of the transformation function is used as a similarity measure. Further examples of local attributes are the edge orientation and strength (gradient across the edge) for edge elements, the length and curvature of lines, and the area of regions. Global features are usually composed of different local features. Besides the attributes of the local features, relations between these local features are introduced to characterize global features. These relations can be geometric or topologic. Geometric relations include the angle between two adjacent polygon sides and the minimum distance between two edges. An example of a topologic relation is the notion that one feature is contained in another. Matching with global features is also referred to as relational matching.

The $\Psi$-$S$ curve is a good example of geometric similarity measures. It is a functional representation of a line where the arc length $s$ is the parameter of the tangent $\Psi$. Straight lines in the spatial domain correspond to horizontal straight lines in the $\Psi$-$S$ domain (the derivative is constant), see Figure 2.4 (a). Note that a rotation in the spatial domain amounts to vertical shift in the $\Psi$-$S$ representation, see Figure 2.4 (b) (Schenk, 1999). Circles are represented as straight lines with slope proportional to the curvature. The rationale for using the $\Psi$-$S$ representation for feature matching is that the representation is invariant with respect to the edge position in the image and the rotation between two images results in a simple shift. Another motivation for using this representation is its stability to extract distinct shape features such as change in curvature.

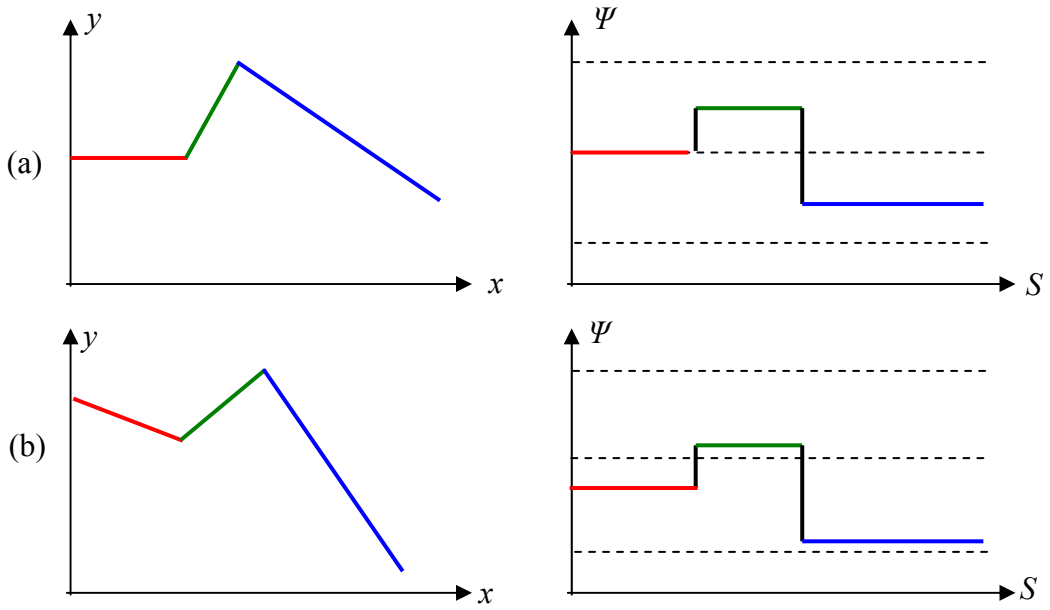



**Figure 2.4:** (a) Example of Line in the x,y Domain and its Representation in the Ψ-S Domain, (b) Rotated Version of the Line in Figure 2.4 (a).

In conclusion, similarity measures based on geometric attributes are more suitable than those based on radiometric attributes. Features in general are more invariant with respect to geometric and radiometric influences. Moreover, geometric constraints can be established between features to ensure the correspondence between conjugate primitives.

## 2.5 Matching Strategy

In photogrammetry and remote sensing, matching can be defined as the establishment the correspondence between various data sets. The matching problem is also referred to as the correspondence problem. The datasets involved in matching might include images, maps and GIS data. Image matching is an important step and prerequisite for many

applications, such as image registration, image orientation, DEM generation, ortho-rectification, data fusion and relative orientation.

Matching strategy refers to the concept or overall scheme of the solution of the matching problem (Schenk, 1999). It encompasses the selected primitives, transformation functions and similarity measures for automatically solving the registration problem.

## 2.6 Registration Algorithms

A previous survey of registration techniques (Fonseca and Manjunath, 1996) reviewed registration techniques developed for many different types of applications and data. An earlier survey by Brown (1992) was far more wide-ranging and compared numerous different applications of image registration, including remote sensing, computer vision and medical imaging. The following survey outlines relevant research attempts in chronological order.

Goshtasby et al. (1986) used closed boundary regions primitives to register LANDSAT MSS and simulated TM data. Regions were extracted as matching primitives through the use of image segmentation. Centers of gravity of closed boundary regions were taken as tie points and correspondence was established between the tie points. These points were used to establish an approximation of the scale factor, rotation and translation required to register the images. Through registration of the images, corresponding patches could be identified and their edges refined so that they became optimally similar. The centroids of the patches were then re-determined to an anticipated higher degree of accuracy. The co-ordinates of these centroids were then used to determine the parameters of a higher order transformation function (polynomial transformation). Sub-pixel accuracy was achieved.

Goshtasby (1988) used point primitives to register LANDSAT MSS and LANDSAT-TM scenes. The centers of gravity of fourteen corresponding regions were used as control points to register the images. Two surface splines, which represent the x and y components of a transformation function, were used to register the two images. Results showed the root mean square error of 0.63 pixels, compared to 1.9 pixels when a polynomial and least squares technique was used. The surface fitting approach takes into consideration the local geometric distortion between the images. However, the surface fitting approach proved to have a high computational cost and relied on the number of correctly identified tie points to solve the equations.

Flusser (1992) used point primitives to establish image-to-image registration. The method relied on a number of correctly identified tie points being selected in advance. Flusser's paper addressed the problem of selecting a valid and appropriate transformation function between the images. Images of 3D-scenes had different viewing angles, or taken by different sensors, had local geometric distortions which prevented global polynomial transformation functions from achieving accurate registration. More accurate results were obtained using surface spline transformation functions. However, the computational efforts to solve the surface spline functions were too great and intensified quickly with increasing numbers of tie points and larger images. To solve the problem, an adaptive algorithm that split the image up into smaller tiles and used a much simpler equation to represent the surface spline function was used where sub-pixel accuracy was achieved.

Flusser and Suk (1994) used closed boundary regions as matching primitives to register SPOT and LANDSAT-TM images of size 512 × 512 pixels. To extract closed boundary

regions, a Sobel mask was used first to detect edges in various directions; then, the image was binarized where closed-boundary regions were found. Only the regions having perimeters between ten and one hundred pixels were taken into account. Correspondence between the regions was then established in two stages. The first stage depended on the local information of the regions to find three pairs of the most likely correspondence regions. The second stage used the center of gravity of the regions to find the parameters of the assumed transformation function. Root mean square error was used to assess the accuracy of the registration where most errors were less than one pixel.

Abbasi-Dezfouli and Freeman (1994) used areal features as matching primitives to register SPOT stereo pairs of size 500 × 500 pixels. The primitives were extracted by searching the image for patches of uniform colour and were then matched using several criteria that corresponded to some feature's attributes such as area, dimensions of bounding rectangle, perimeter, linearity, concavity and relative geometry. The aim was to register the images fully automatically and determine the terrain height across the region of interest. After the patches had been matched, tie points were generated by matching significant points on the boundaries of corresponding patches with each other through the use of a correlation method. No mention was made about the order of the transformation function or the matching strategy.

Li et al. (1995) presented a contour-based approach using region boundaries and other strong edges as matching primitives to register LANDSAT and SPOT images of size 512 × 512 pixels. Images were convolved with a Laplacian-of-Gaussian (LOG) operator and the edges were detected at the zero crossing points. For every closed contour, five shape

attributes were compared: the perimeter, the longest and shortest distances from the boundary to the centroid, and the first and second invariant moments. An invariant moment is a function of centroid coordinates and the length of the contour. A pair of contour lines was accepted as matching a candidate if it passed a certain threshold. For open contours, measures of curvature at certain points within certain thresholds were used to find the matched entities. The relationship between corresponding points was assumed to be an affine transformation. Results showed that the contour-matching algorithm was quite robust and reliable as long as corresponding contours were available. The registration scheme would fail if sufficient contour information could not be extracted.

Dowman et al. (1996) used points to register two SPOT images of size 1024 × 1024 pixels. The matching primitives were extracted from the images using the Förstner interest operator (Förstner, 1986). Large number of corresponding points was automatically found. The interest operator was applied first at the top level of the pyramid and the matching was applied at that level where matched points were used to define the initial transformation parameters. The transformation passed down to the next level where extraction and matching were applied again. The transformation from the previous level of the pyramid was used to predict the position of the point in the reference and input images. Cross-correlation was used to find conjugate points. An approximate initial affine transformation was used to register the images. Results showed an RMS error of 2 pixels (approximately 20m on the ground).

Boardman et al. (1996) used point primitives to register two SPOT images. In this work, the Förstner operator was used to determine interest points in an image. Transformation function and accuracy of the results was not mentioned. The suggested system components were initial image registration, image smoothing, image sub-sampling, areas of interest, interest point extraction, feature point matching and intensity matching. The procedure required an initial image registration to kick off the hierarchical and iterative matching procedure. The suggested system achieved sub-pixel accuracy in ideal situations and accuracy of two pixels in areas of great relief change, with much fewer matchable features.

Fonseca and Costa (1997) presented an automatic registration algorithm that used points to register SPOT and LANDSAT-TM images of size 512 × 512 pixels. Point features were detected from the gray level information content of the images and their local wavelet transform modulus maxima. A correlation coefficient was used as a similarity measure. An affine transformation was used to model the deformation between the two images. Because the registration procedure used the gray level information content of the images in the matching process, it was only adequate for registering images of the same sensor with similar spectral bands.

Morgado and Dowman (1997) used areal features as the matching primitives to register aerial photographs to a map. Initial registration was performed through the matching of areal features on the basis of attributes such as size, shape and perimeter length. Registration was then refined based on a dynamic programming technique. Tie points were generated from the pixels that made up the edges of the patches. The image was

registered to the map using an affine transformation, the parameters of which were determined from the tie points using a least squares technique. An RMS error of approximately 7m (two pixels) in the x-y plane was obtained. As in other studies, areal and edge features were utilized to extract tie points such as edge endpoints for use in matching where point-to-point correspondence was required.

Hsieh et al. (1997) introduced an edge-based approach for image registration. Their approach applied a wavelet transform to extract a number of points as the basis for registration. Each selected point was an edge point whose edge response was the maximum within a neighborhood. The rotation angle corresponding to the maximum peak of the histogram was used to compensate for the difference between two target images. On the basis of the rotation angle, an initial matching could be performed. Then, cross-correlations were used to establish and find matching entities for use in the final registration process. 2-D similarity transformation was used to model the geometric transformation between two images.

Dare and Dowman (2001) used points and areal features as registration primitives to automatically register SAR and SPOT sub-images of size 512 × 512 pixels. Multiple feature extraction and matching algorithms were incorporated to identify common features from which accurate tie points could be derived. An affine transformation was used to model the geometric distortion present in the images. The image registration model described in this work was based on three steps. Initial alignment with the use of manually selected tie points was followed by approximate registration using patch matching. Finally, an edge extraction algorithm was applied to provide a much larger

number of tie points with more extensive spatial distribution than the number and distribution found in the patch matching algorithm. The results showed that multiple feature extraction algorithm increased the number of identified common tie points usable for more accurate and robust image registration. The method relied on a number of correctly identified tie points resulting from different matching procedures. Therefore, mismatch points would lead to significant change in the estimated transformation parameters of the suggested transformation function.

Seedahmed and Martucci (2002) used points as the registration primitive to register two sub-images of SPOT scenes of size 1024 × 1024 pixels taken at different times (1987 and 1991). Point features were extracted through the use of a Moravec operator. The suggested approach assumed that the two SPOT images could be aligned by a 2-D similarity transformation. This paper introduced an automatic registration procedure largely based on the Modified Iterated Hough Transform (MIHT) strategy (Habib et al., 2001a, b). The results showed that sub-pixel accuracy in the final registration parameters was achieved. The suggested approach significantly differed from the registration strategies described above, as it simultaneously determined the correspondences between the involved primitives and solved for the parameters of the registration transformation function. However, this work started with the extraction of point primitives that could not be reliably extracted from imagery with different geometric and radiometric properties. Moreover, the registration transformation function was not investigated and was assumed without valid justification.

Table 2.1 summarizes the work described above in terms of the primitives and transformation functions used.

**Table 2.1:** Summary of Current Registration Techniques

| Authors | Registration primitives | Transformation functions |
|---|---|---|
| Goshtasby et al. (1986) | Regions | Polynomial function |
| Goshtasby (1988) | Regions | Surface splines |
| Flusser (1992) | Points | Surface Spline |
| Flusser and Suk (1994) | Regions | Affine |
| Abbasi-Dezfouli and Freeman (1994) | Regions | Affine |
| Li (1995) | Points and Regions | Affine |
| Dowman et al. (1996) | Points | Affine |
| Boardman et al. (1996) | Points | 2-D Similarity |
| Fonseca and Costa (1997) | Points | Affine |
| Morgado and Dowman (1997) | Regions | Affine |
| Hsieh et al. (1997) | Points | 2-D Similarity |
| Dare and Dowman (2001) | Points and Regions | Affine |
| Seedahmed and Martucci (2002) | Points | 2-D Similarity |

Although a vast body of research has dealt with image registration, methodologies that can meet the current challenges posed by image registration are not available yet. The following is a summary of the drawbacks of the methodologies suggested in current literature:

- Extracted points from multi-source imagery with varying radiometric and geometric properties would be difficult to match. Moreover, for this imagery, point extraction algorithms likely would not be able to identify the same point. One could even argue that manual registration of such imagery using points would be extremely difficult, Figure 2.2.

- Areal primitives might not be always available, especially in the case of high resolution satellite scenes over urban areas.

- Registration procedures based on areal primitives use the centers of gravity of these features as the registration primitives. The estimated centers of gravity are susceptible to potential errors associated with the identified boundaries of these patches.

- Developed similarity measures for matching those primitives are empirical and sometimes subjective. Also, the involved imagery has to be approximately aligned or registered prior to the automatic registration procedure to avoid ambiguities in the matching of the involved primitives.

- The appropriate registration transformation function is not investigated, i.e., simplified and sometimes invalid registration transformation function is assumed. Moreover, previous methods draw out the results based on small patches of the test fields (e.g., 500 × 500 pixels). This may be considered a local transformation because such patches do not necessarily represent the whole scene.

As mentioned before, change detection is one of the most important applications of image registration where accurate image registration is required for reliable and effective change detection (Singh, 1989; Townshend et al.1992; Lillesand and Kiefer 2000; and Li et al., 2002). The next section discusses the importance of change detection and reviews different algorithms and techniques that have been mentioned in the literature.

## 2.7 Change Detection Algorithms

Change detection is the process of identifying differences in the state of objects or phenomena through observation at different times. Change detection is an important process for monitoring and managing natural resources, urban development, environmental changes and disaster assessments. Recent advances in satellite imagery, in terms of improved spatial and temporal resolutions, allow for efficient identification of change patterns and the prediction of areas of growth. Change detection analysis might involve multi-spectral, multi-source, and multi-resolution images that have been captured at different times. The reliability of the change detection process is strongly affected by environmental factors such as atmospheric effects, illumination conditions, lake level, winds or soil moisture. Moreover, seasonal changes such as differences in land cover must be considered (Lillesand and Kiefer 2000).

Traditional change detection studies are based on visual comparison of temporal datasets (such as satellite scenes, aerial images, maps, etc.). However, the huge flux of imagery that is being captured by an ever-increasing number of earth observing satellites necessitates the development of automatic, reliable, and fast change detection techniques. Such techniques are essential to reduce the high cost associated with spatial data updating activities.

Several change detection methods have been developed and reported in the literature (Singh, 1989; Fung, 1990; Coppin and Bauer, 1994; Dowman, 1998; Sohl, 1999; Mas, 1999; Bruzzone and Prieto, 2000; Cho, 2000; Li et al., 2002 ; Palandro et al., 2003; Li and Narayanan, 2003, and Townshend et al., 1992). Basically, two main solutions for the

change detection problem have been proposed: the supervised and unsupervised approaches. The former is based on supervised classification methods, which require the availability of multi-temporal ground truth in order to derive a suitable training set for the learning process of the classifier. The latter performs change detection by directly comparing the two images under consideration, without relying on any additional information (Bruzzone and Prieto, 2000).

In supervised classification, data from two images is separately classified; thus, the problem of normalizing such data for atmospheric and sensor differences between two different times is minimized (Singh, 1989). The supervised approach exhibits some advantages over the unsupervised, mainly the capability to recognize the kinds of land cover transition that have occurred, robustness to different atmospheric and light conditions at the two acquisition times, and the ability to process multi-sensor/multi-source images (Bruzzone and Serpico, 1997). A major drawback of the supervised classification is that the generation of an appropriate multi-temporal ground truth is usually a difficult and expensive task; in addition greater computational and labelling efforts are required. On the other hand, unsupervised classification is used mainly to create “difference images”. It involves image differencing, image ratio, vegetation index differencing, image regressions, change vector analysis (CVA), and principal component analysis (PCA). Changes are then identified through analysis (e.g., thresholding) of the difference image.

*Image differencing* is the most widely used technique for unsupervised change detection (Singh, 1989). In this technique, two or more images taken at different times are

subtracted, pixel by pixel, where changes are detected based on the difference of the gray values. The differences in areas of no change will be close to zero while areas of change will have relatively large differences. A threshold must be decided upon to set the boundary between changed and unchanged areas.

*Image ratioing* finds the ratio in the gray value where images are compared on a pixel by pixel basis. The ratio in areas of no change will be very close to one while in areas of change the ratio would be significantly greater or less than one. Ratioing tends to normalize the data for changes such as sun angle and shadow (Lillesand and Kiefer, 2000). However, the problem lies in selecting the threshold values in order to separate the areas of change from those of no change.

*Vegetation index differencing* technique uses the Normalized Difference Vegetation Index (NDVI) as the basis of the change detection algorithm instead of making direct use of spectral radiance values. If one band is in the visible region (VIS) and another band is in the near infrared (NIR), then the NDVI is (NIR - VIS)/(NIR + VIS). The technique provides a crude estimate of vegetation health and a way of monitoring changes in vegetation over time. Simply*, in vegetation index differencing technique,* the NDVI is calculated for both dates and then subtracted. This tends to enhance the random or coherent noise that is not correlated in different bands. Nelson (1982, 1983) reported that *vegetation index differencing* is a more accurate technique than image differencing and ratioing to describe forest canopy changes. Banner and Lynham (1981) used the *vegetation index differencing* to delineate forest clear-cuts. They compared the results

with the supervised classification approach using multi-spectral MSS Band 5 images. The results showed that the *vegetation index differencing* method was less accurate.

*Image regression* assumes a linear relationship for changes between pixels and applies regression techniques to predict unknown information. In this method, a thresholding technique is also used to detect the area of changes. This technique accounts for differences in the mean and variance of pixel values so that the effect of different atmospheric conditions and sun angles tends to be reduced (Jenson, 1983).

*The Change Vector Analysis* algorithm can be considered as an extension of image differencing. Two spectral variables are plotted for Date 1 and Date 2 of a given pixel. The vector connecting the two datasets describes the magnitude and direction of spectral changes between two images. A threshold on the magnitude can be established as the basis for determining the areas of changes, and the direction of the spectral changes vector often relates to the type of the change (Lillesand and Kiefer 2000). CVA requires accurate geometric registration and radiometric normalization (Johnson and Kasischke, 1998).

*The Principal component analysis technique is* used to compress all of the information contained in an original n-bands dataset into fewer than n bands (components), where linear combinations of the original dataset are derived to form new band images containing all the information of all bands in each image. Then, several uncorrelated principal components are used in image differencing to find areas of changes. This

technique is often difficult to interpret and it is not easy to identify the specific nature of the changes involved (Lillesand and Kiefer 2000).

Based on the previous techniques of change detection, several procedures have been proposed by researchers. Fung (1990) assesses the information content and accuracy of LANDSAT TM digital images for change detection purposes. Two sub-areas of LANDSAT-5 TM images of size 500 × 700 dated August 3, 1985 and July 21, 1986 were used for the analysis. Both images were radiometrically calibrated and converted to reflectance values to alleviate any differences in solar elevation. To register geometrically the images, polynomial functions were generated by using twenty nine ground control points. The RMSE was reported to be below 0.2 pixels. On the basis of the different combinations of the available special bands of LANDSAT, change detection techniques of image differencing and principal component analysis were applied to generate twelve change images. To delineate change from no-change areas, each of the twelve images was thresholded through the selection of threshold values at ± N standard deviations from the mean. The results show that images associated with changes in the near-infrared reflectance or greenness detect changes between vegetation and non-vegetation areas. Images related to changes in the visible reflectance are able to detect changes due to rural-to-urban land conversion. Moreover, researchers found that mid-infrared do not provide additional information about land-cover changes.

Townshend et al. (1992) studied the impact of misregistration of images on the detection of changes in land cover using LANDSAT MSS images. They focused attention on simulated images of the Normalized Difference Vegetation Index (NDVI) of two of the

spatial resolution of the planed Moderate Resolution Imaging Spectrometer (MODIS). The results indicate that in the absence of real changes in the object space, the consequences of misregistration were clearly highlighted even for sub-pixel misregistration. Also, the results showed that for four out of seven tested areas an error equivalent to 50% of the actual changes has been introduced. Registration accuracy of 0.2 pixels is required in order to reduce the error to 10% of the actual changes.

Coppin and Bauer (1994) developed a digital procedure to optimize the information content of multi-spectral LANDSAT TM images for forest cover change detection. Images from three different years (1984, 1986, and 1990) were radiometrically calibrated and geometrically rectified being before subjected to two change detection algorithms, image differencing and principal component analysis. Vegetation indexes were calculated from band reflectance values and used in the change detection algorithms. It was found that changes in brightness and greenness identified the most important forest canopy change features. The researchers noted that post-classification for change images would eliminate small but real features of interest. They also reported that spatially accurate forest cover monitoring required the precise registration of the multi-date imagery.

Dowman (1998) developed procedures for change detection using aerial, SPOT and SAR images. The first procedure uses the information present in a database to identify corresponding objects in an image and then compare the detailed geometry as given by the database and image. The second procedure compares two images pixel by pixel after normalizing and noise reduction. As has been mentioned, accurate image registration is essential for the validity of change detection output.

Mas (1999) used LANDSAT Multi-Spectral Scanner (MSS) to test six change detection procedures for detecting areas of change in a coastal zone. 2045×1687 pixel sub images were extracted from February 15, 1974 and April 29, 1992 LANDSAT MSS scenes. The change detection techniques considered were image differencing, vegetation index differencing, principal component analysis, unsupervised classification, supervised classification and a combination of supervised classification and image enhancement. The accuracy of the results was evaluated by comparison with aerial photographs. Thresholds were applied to the change images to isolate the pixels with no change. Supervised classification comparison was found to be the most accurate procedure and had the added advantage of indicating the nature of changes. Due to differences in soil moisture and vegetation between the two scenes, poor performance was obtained by image enhancement procedures. It was found that methods based on classification were less sensitive to radiometric variations.

Cho (2000) used LANDSAT TM and SPOT panchromatic images to detect environmental changes on an area of Mountain Moscow, Idaho. LANDSAT TM images were captured on July 8, 1990 and July 20, 1991, while a SPOT image was captured on August 14, 1992. Images were registered with RMSE of 0.286 pixels where TM images were resampled from 30 m to 10 m spatial resolution. The image enchantment process then was performed to improve the quality of the images. Supervised classification was performed using the maximum likelihood classifier method. Pixel-by-pixel comparison was used to detect changes in the study area. Overall accuracy was reported to be higher than 90%.

Palandro et al. (2003) used two aerial photographs acquired on February 5, 1981 and Januarry 17, 1992 and 4 m spatial resolution IKONOS imagery acquired on October 15, 2000 for a change detection application. Aerial photographs were scanned and resampled to 2 m pixel size and then geo-corrected using IKONOS image as a reference image. Supervised classification was used to divide each image into four classes. The training pixels were chosen according to the database available for each year. Results showed the IKONOS imagery to be a good source of information when used in conjunction with historic color aerial photographs. The authors did not release any information about the size of involved imagery, accuracy of registration between the images, radiometric variations between the images, quality of the output results, and the types of changes that occurred. Moreover, the change detection procedure was not explained.

Li and Narayanan (2003) presented a shape-based approach to detect changes of lakes in the Nebraska region using supervised classification. Thirty-six four-band LANDSAT MSS images of size 256×256 pixels from 1981 to 1987 and 10 six-band TM of size 768×768 from 1992 to 1997 were employed. All the images were radiometrically and geometrically rectified. As a first stage, each image was classified and lakes represented by polygons were retrieved. Then, shape similarity measures based on the Euclidian distances between centroids of two conjugate polygons were conducted. Results showed that the centroids of the studied lakes did not change although the boundary shapes of some lakes experience strong variations

Table 2.2 summarizes the work described above in terms of the images as well as the methods used for change detection.

**Table 2.2:** Summary of Change Detection Techniques

| Authors | Images | Change detection Method |
|---|---|---|
| Fung (1990) | ▪ LANDSAT-TM | ▪ Image differencing<br>▪ Principal component analysis |
| Townshend et al. (1992) | ▪ LANDSAT-MSS | ▪ Vegetation index differencing |
| Coppin and Bauer (1994) | ▪ LANDSAT-TM | ▪ Vegetation index differencing |
| Dowman (1998) | ▪ Aerial<br>▪ SPOT<br>▪ SAR | ▪ Image differencing |
| Mas (1999) | ▪ LANDSAT-MSS | ▪ Image differencing<br>▪ Vegetation index differencing<br>▪ Principal component analysis |
| Cho (2000) | ▪ LANDSAT-TM<br>▪ SPOT | ▪ Image differencing |
| Palandro et al. (2003) | ▪ Aerial<br>▪ IKONOS | ▪ Image differencing |
| Li and Narayanan (2003) | ▪ LANDSAT-MSS | ▪ Image differencing |

In summary, the following issues have to be considered for change detection techniques:

- Image differencing methods assume that differences between radiometric values are due to changes in the object space. Intact, these differences could be a result of other factors, such as different atmospheric conditions, different illumination conditions, changes in soil moisture and changes of sunlight angle. Several solutions were suggested to overcome such a problem. Basically, these solutions depend on image enhancement and radiometric corrections that tend to reduce radiometric differences between images under consideration. Ingram et al., (1981) used a normalization

procedure, in which a selected value for the mean and standard deviation was input along the images to be normalized.

- Most of these methods require a decision as to where to place the threshold boundaries in order to separate the areas of changes from those of no changes (Singh, 1989). In fact, classical techniques perform thresholding based on empirical strategies or manual trial and error procedures, which significantly affect the reliability and the accuracy of the final change detection results (Li et al., 2002).

- In general, classification methods require two or more bands for the classification process. These are not always available, especially in the case of aerial images, which represent an important source of historical information needed for change detection purposes.

- Image differencing techniques are sensitive to misregistration between the reference and input images (Singh, 1989; Townshend et al., 1992; Li et al., 2002; Buruzzone 2003). Studies pointed out that the accuracy of the image registration process is the key factor that controls the validity and reliability of the change detection outcome.

Traditional approaches to change detection have failed. They are based on differencing intensity images, and the illumination of the scene is not under control in many applications. To overcome this problem, this thesis proposes features invariant to changes of the illumination conditions.

# CHAPTER 3

# IMAGE REGISTRATION METHODOLOGY

## 3.1 Introduction

In this chapter, the basic concepts, the mathematical model, and the methodology of the proposed image registration paradigm are introduced. The elements of the paradigm have been established and chosen in a way that allows them to handle multi-source imagery with varying geometric and radiometric properties.

Section 3.2 is a discussion of the rationale behind adopting linear features, and in particular straight lines, as the preferred registration primitives. In Section 3.3, rigorous as well as approximate transformation functions are analyzed in search of the most appropriate transformation functions. Then, a geometric similarity measure based on straight lines is mathematically derived (Section 3.4). Finally, a matching strategy based on Modified Iterated Hough Transform is introduced. This is used as a framework utilizing straight line primitives, a similarity measure, and a transformation function to calculate the parameters of the transformation function and simultaneously establish the correspondence between conjugate lines (Section 3.5).

## 3.2 Linear Features

As mentioned before, the selection of the most appropriate primitives, which encompass the domain in which information is extracted from input imagery for the registration process, is the first step of the image registration paradigm. In contrast to point primitives, linear features have a set of appealing properties, especially in the case of multi-resolution images. These properties include the following facts:

- Compared to distinct points, linear features have higher semantics, which can be useful for subsequent processes (such as DEM generation, map compilation, change detection, and object recognition).
- It is easier to automatically extract linear features from multi-resolution imagery rather than distinct points (Kubik, 1991). This is attributable to the nature of linear features, since they represent discontinuities in the gray value function in one direction. On the other hand, point features represent discontinuity in all directions. Even if the extraction process is done manually, the identification of conjugate linear features in multi-resolution imagery is much easier than the identification of conjugate distinct points.
- Images of a man-made environment are rich with linear features.
- Geometric constraints are more likely to exist among linear features. This can lead to a simple and robust registration procedure.
- Linear features in multi-resolution imagery can be extracted with sub-pixel accuracy across the direction of the edge.

- Linear features allow for the incorporation of areal features through the use of their boundaries. Moreover, linear features are easier to use in change detection applications than are areal features. The superiority of linear features stems from the possibility of dividing them into smaller subsets. On the other hand, breaking areal features into smaller subsets is not a trivial task.
- Terrestrial Mobile Mapping Systems (MMS) can economically provide accurate and current object space linear features in real time.
- Linear features increase the redundancy and improve the robustness and geometric strength of various photogrammetric adjustment activities.
- Point correspondence on matched linear features is not necessary, so the use of such features allows more flexibility than the use of points or areal features.

Linear features can be represented either by an analytical function (e.g., straight lines, conic sections, or parametric functions) or by a free form shape. In this research, straight-line segments have been chosen as the registration primitives for the following reasons:

- Man-made environments are rich with straight lines.
- Straight lines are easier to detect in multi-resolution imagery, and the correspondence problem between conjugate features in the input imagery becomes easier to solve.
- Straight-line parameters can be obtained with sub-pixel accuracy.
- It is straightforward to develop mathematical constraints (similarity measures) describing the correspondence of conjugate straight-line segments.
- Free-form linear features can be represented with sufficient accuracy as a sequence of straight-line segments (polylines).

After selecting straight-line segments as the registration primitives, one must decide how to represent them. In this thesis, the line segments will be represented by their end points. This representation is chosen since it will have no singularity (i.e., it is capable of representing all line segments in 2-D space). The end points defining corresponding line segments in the imagery need not be conjugate, Figure 3.1.

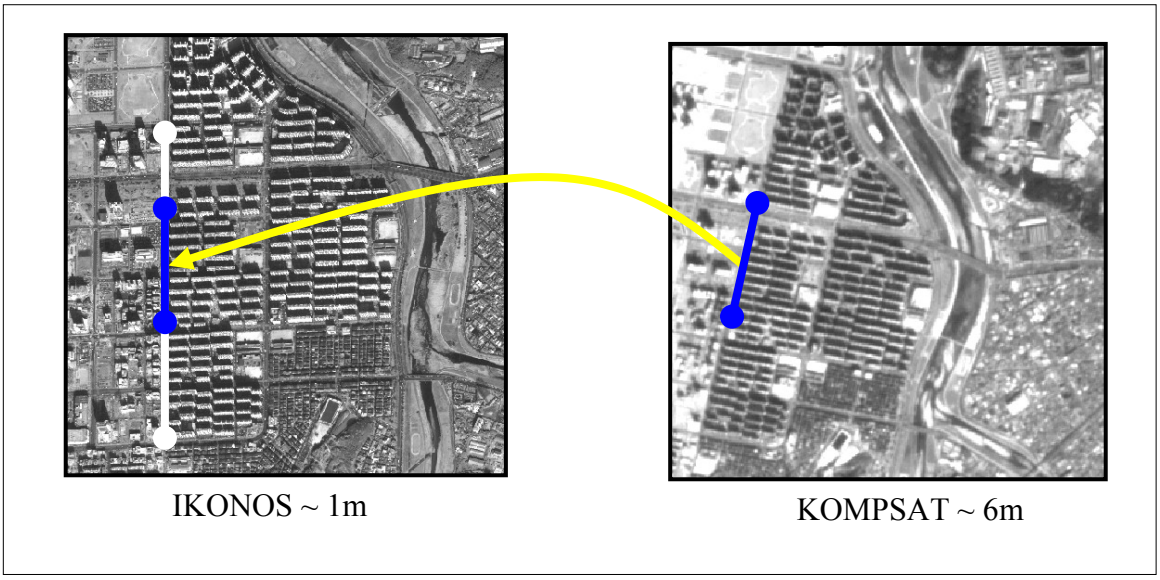



**Figure 3.1:** Conjugate Straight Lines without Corresponding End Points

It has to be mentioned that manual digitization was adopted in this research since the main objective is focused on image-to-image registration through the use of straight lines, not the extraction method. Automatic extraction of straight lines was beyond the objective of this study and will be investigated in future work.

Once straight lines are digitized and adopted as the most suitable primitive to be used in the registration process, the next step is to select a valid and proper transformation function that can faithfully represent the transformation between the input and the reference images.

### 3.3 Registration Transformation Functions

At this stage, one should establish a registration transformation function that properly aligns the images relative to each other. Given a pair of images, reference and input images, the registration process attempts to find the relative transformation between them. The type of spatial transformation needed to properly overlay the input and reference images is one of the most fundamental and difficult tasks in any image registration technique. Images involved in the registration process might have been taken from different viewpoints with the use of different imaging technologies. The registration transformation function must suit multi-resolution and multi-spectral images that could be captured under different circumstances.

Throughout this thesis, $(x, y)$ denotes the coordinates of a point in the reference image, $(x', y')$ is used for the coordinates of the conjugate point in the input image, and $(X, Y, Z)$ represents the ground coordinates of the corresponding object point. In the next subsections, available options for establishing the transformation models between such images will be investigated. For the various alternatives, object to image space transformation for both the reference and input images is discussed first. Then, the transformation function is analyzed to derive the mapping function relating the reference and input images, Figure 3.2.

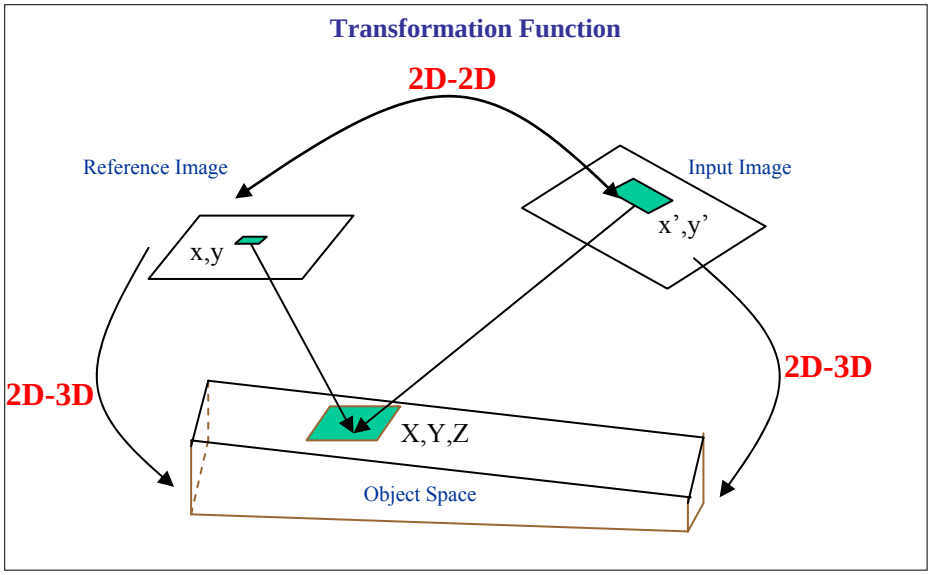



**Figure 3.2:** Object to Image (3D-2D) Transformation and Image-to-Image (2D-2D) Transformation

Different transformation models can be utilized to handle object to image transformation. Such models fall mainly into two broad categories: rigorous and approximate models. The following subsections describe these models in some detail.

### 3.3.1 Perspective Projection: Rigorous Models

A great challenge for digital photogrammetry is the investigation of different types of sensors such as frame cameras, push broom scanners, panoramic scanners, radar and laser scanners (Ackermann, 1995). In addition to the many different types of imaging sensors, supportive position and attitude sensors (GPS/INS) are also available. The integration of these different types of sensors constitutes the real progress in digital photogrammetry (Habib and Beshah, 1998).

Despite the current progress in the development of frame cameras, digital frame systems that can replace film-based cameras are still not available. A Charged Coupled Device (CCD) camera with a resolution comparable to the classical frame cameras must be in the order of 20,000 × 20,000 sensing elements (Schenk, 1999), and is not available yet. The chief advantage of digital frame cameras over the classical film-based technology is the instant availability of images for further processing and analysis. This feature makes digital frame cameras ideal for real time photogrammetry (Habib and Beshah, 1998).

Today, digital systems using line scanning geometry are the only imaging sensors that can compete with film-based photos in terms of acquired area and image resolution. Digital frame cameras capture images through a single exposure of a two-dimensional CCD array; linear array scanners capture scenes with large ground coverage and high geometric and radiometric resolutions through multiple exposures of a few scan lines along the focal plane. Successive coverage of different areas on the ground is achieved either through the motion of the imaging platform (push-broom scanners) or the motion of the sensor relative to the imaging platform (panoramic scanners). Depending on the sensor type, viewing angles, system altitude, and scan trajectory, different mathematical models have been devised to describe the relation between the image and object space as well as the relation between stereo-images.

The image formation process can be described by a central (perspective) projection in which the projection rays from the object to the image space pass through a single point, the perspective center. For frame cameras, there are three models that can be used to rigorously describe the mathematical relationship between corresponding image and

ground coordinates, collinearity equations, Direct Linear Transformation (DLT), and projective transformation.

The main objective of the collinearity equations (Equation 3.1) is to define the mathematical relationship between corresponding image and object space coordinates in perspective views, where the image coordinates of a point are expressed as a function of the Interior Orientation Parameters (IOP), the Exterior Orientation Parameters (EOP), and the ground coordinates of the corresponding object point, Figure 3.3 (Kraus, 1997). The IOP defines the sensor or camera characteristics required for the reconstruction of the image space bundle of rays from corresponding image points. The IOP can be obtained from the system manufacturer or a calibration procedure (Habib et al, 2001b). The EOP establishes the position and orientation of the bundle of rays with respect to the object space coordinates (Mikhail and Bethel, 2001). EOP can be directly determined through the use of GPS/INS or indirectly estimated through the use of ground control points in a bundle adjustment procedure (Habib et al, 2000; Habib et al, 2001b)

The concept of the collinearity equations stems from the fact that image point, object point, and the perspective center are collinear.

$$x = x_p - c\frac{r_{11}\cdot(X-X_o)+r_{21}\cdot(Y-Y_o)+r_{31}\cdot(Z-Z_o)}{r_{13}\cdot(X-X_o)+r_{23}\cdot(Y-Y_o)+r_{33}\cdot(Z-Z_o)}$$

$$y = y_p - c\frac{r_{12}\cdot(X-X_o)+r_{22}\cdot(Y-Y_o)+r_{32}\cdot(Z-Z_o)}{r_{13}\cdot(X-X_o)+r_{23}\cdot(Y-Y_o)+r_{33}\cdot(Z-Z_o)} \qquad (3.1)$$

where

x, y: Image point coordinates corresponding to object point (X, Y, Z)

X, Y, Z: Corresponding ground point coordinates

$x_p$, $y_p$, c: Interior orientation parameters (calibrated principal point position and principal distance of the camera with respect to image coordinate system)

$X_0$, $Y_0$, $Z_0$: ω, φ, κ Exterior orientation parameters ($X_0$, $Y_0$ and $Z_0$ represent the position of perspective center with respect to ground coordinate system, where ω, φ and κ represent the rotation angles between the ground and image coordinate systems).

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \Rightarrow \text{Rotation Matrix}$$

$$= \begin{bmatrix} \cos\varphi\cos\kappa & -\cos\varphi\sin\kappa & \sin\varphi \\ \cos\omega\sin\kappa + \sin\omega\sin\varphi\cos\kappa & \cos\omega\cos\kappa - \sin\omega\sin\varphi\sin\kappa & -\sin\omega\cos\varphi \\ \sin\omega\sin\kappa - \cos\omega\sin\varphi\cos\kappa & \sin\omega\cos\kappa + \cos\omega\sin\varphi\sin\kappa & \cos\omega\cos\varphi \end{bmatrix}$$

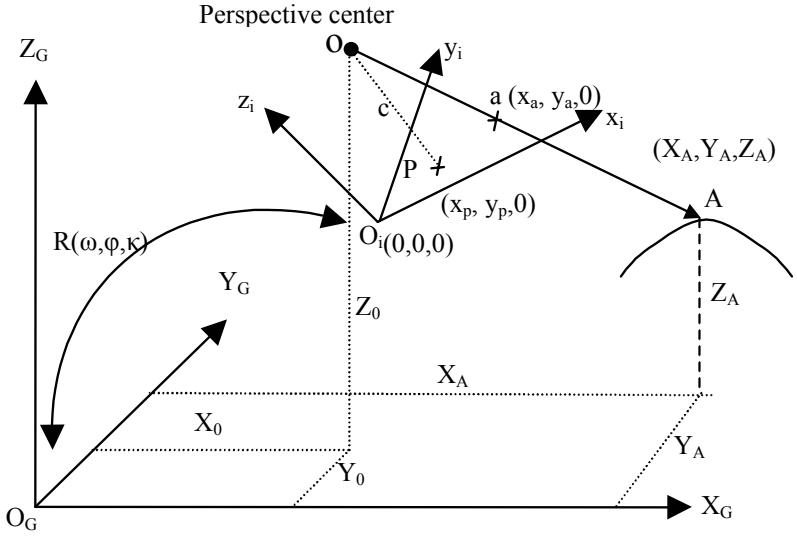



**Figure 3.3:** Relationship between Image and Object Coordinate System

DLT in Equation 3.2 is a simpler model in which the relationship between image and ground coordinates is formulated through the use of eleven parameters that encompass the interior and exterior orientation parameters (Abdel-Aziz and Karara, 1971). The wide popularity of the DLT is due to the linear formulation of the relationship between image and object coordinates. The DLT model requires well distributed 3-D object space control points to estimate the full set of its parameters. In this model, IOP and EOP are not explicitly needed.

$$x = \frac{A_1 + A_2 X + A_3 Y + A_4 Z}{1 + A_9 X + A_{10} Y + A_{11} Z}$$
$$y = \frac{A_5 + A_6 X + A_7 Y + A_8 Z}{1 + A_9 X + A_{10} Y + A_{11} Z} \quad (3.2)$$

where

x, y : Image point coordinates corresponding to object point coordinates (X, Y, Z)

$A_1, \ldots, A_{11}$ : Direct linear transformation parameters.

Finally, projective transformation (Equation 3.3), which involves eight parameters, assumes a planar object space. Projective transformation can be used for high altitude photography over flat terrain. At least four planimetric ground control points are needed to solve for the eight parameters involved in the projective transformation. As in the case of the DLT, the IOP and EOP are not explicitly involved in the projective transformation.

$$x = \frac{A_1 + A_2 X + A_3 Y}{1 + A_7 X + A_8 Y}$$
$$y = \frac{A_4 + A_5 X + A_6 Y}{1 + A_7 X + A_8 Y} \tag{3.3}$$

where

x, y : Image point coordinates corresponding to object point coordinates (X, Y, Z)

$A_1, \ldots, A_3$ : Projective transformation parameters.

The collinearity model used for frame imagery can be modified so as to be valid for linear array scanners (Habib and Beshah, 1998). In the case of linear array scanners, each image line is the result of a perspective projection in the CCD line direction and has its own EOP. The collinearity equations for linear array scanners are as follows:

$$x_t = x_p - c\frac{r_{11}^t \cdot (X - X_o^t) + r_{21}^t \cdot (Y - Y_o^t) + r_{31}^t \cdot (Z - Z_o^t)}{r_{13}^t \cdot (X - X_o^t) + r_{23} \cdot (Y - Y_o^t) + r_{33}^t \cdot (Z - Z_o^t)}$$
$$y_t = y_p - c\frac{r_{12}^t \cdot (X - X_o^t) + r_{22}^t \cdot (Y - Y_o^t) + r_{32}^t \cdot (Z - Z_o^t)}{r_{13}^t \cdot (X - X_o^t) + r_{23}^t \cdot (Y - Y_o^t) + r_{33}^t \cdot (Z - Z_o^t)} \tag{3.4}$$

where

$x_t$, $y_t$: Image point coordinates corresponding to object point (X, Y, Z) at time t

X, Y, Z: Corresponding ground point coordinates

$x_p$, $y_p$, c: Interior orientation parameters (calibrated principal point position and principal distance of the camera with respect to image coordinate system)

$r_{11}^t, r_{22}^t, \ldots, r_{33}^t$: Elements of rotation matrix $R^t$, which are function of $\omega^t$, $\varphi^t$ and $\kappa^t$ at time t

$X_o^t, Y_o^t, Z_o^t$: The position of the perspective center at time of capturing the scene line under consideration.

Once perspective transformations (3D to 2D) have been discussed, the mathematical relation between stereo-pair (2D to 2D) has to be described. Before details of the mathematical model are discussed, explanation of the idea of the epipolar geometry must be clarified. Figure 3.4 shows the concept of epipolar geometry of a stereopair. The perspective centers associated with the images of a stereopair and a single point on the ground define a plane known as the epipolar plane; epipolar lines are defined by the intersection of the epipolar plane with the focal planes of the images.

The relationship between conjugate points in stereo-pair (2D to 2D) captured according to perspective projection can be described by the co-planarity condition. This condition mathematically describes the fact that conjugate points in the reference and input images belong to the corresponding epipolar plane (Habib and Kelley, 2001b). The coplanarity constraint can be defined by constraining the normal to the epipolar plane to be perpendicular to the base vector (Equation 3.5). This condition is defined as follows.

$$(\vec{p}_l \times \vec{p}_r) \bullet \vec{b} = 0 \tag{3.5}$$

where

$\vec{b}$: The vector between the two perspective centers of the stereopair, referred to as the image base

$$\vec{b} = Pc_2 - Pc_1 = \begin{bmatrix} X_{or} - X_{ol} \\ Y_{or} - Y_{ol} \\ Z_{or} - Z_{ol} \end{bmatrix}$$

$\vec{P}_l$, $\vec{P}_r$ : The vectors from the perspective center to conjugate points in the left and right images, respectively.

$$\vec{P}_l = R(\omega,\theta,\kappa)_l \begin{bmatrix} x_l - x_p \\ y_l - y_p \\ -c \end{bmatrix} \qquad \vec{P}_r = R(\omega,\theta,\kappa)_r \begin{bmatrix} x_r - x_p \\ y_r - y_p \\ -c \end{bmatrix}$$

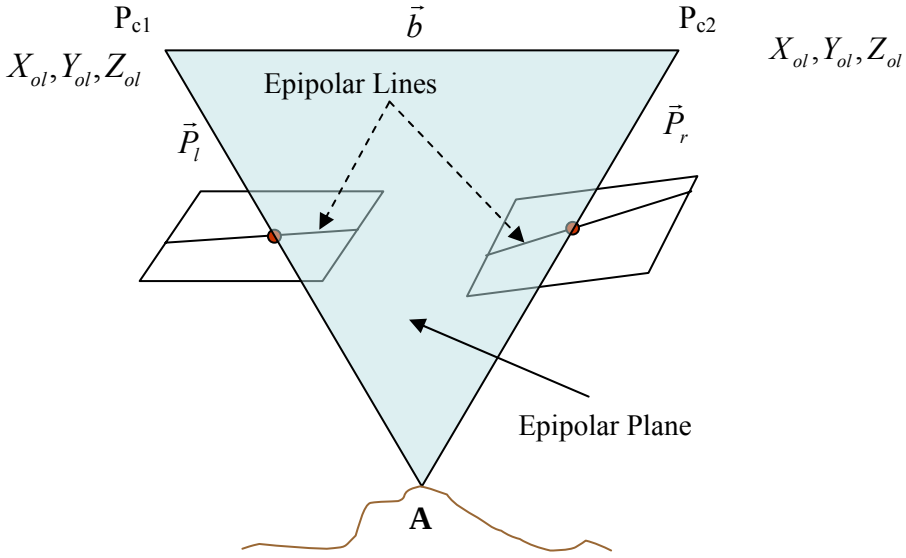



**Figure 3.4:** Epipolar Geometry of a Stereo-pair

Equation 3.5 can be re-written as follows:

$$f(x_l, y_l, x_r, y_r, EOP_L, EOP_r, IOP) = 0 \tag{3.6}$$

Assuming the availability of the EOP and IOP, Equation 3.6 formulates a mathematical relationship between the stereo-pair. This model relates a point in the left image

$(x_l, y_l)$ with a line in the right image (epipolar line). Therefore, there is no point-to-point mathematical relation. Instead any given point in the input image belongs to a line in the reference image (epipolar line) Figure 3.5. Point-to-point relation can be established if and only if a digital elevation model (DEM) of the object space is available. This is not practical since the DEM is an end product.

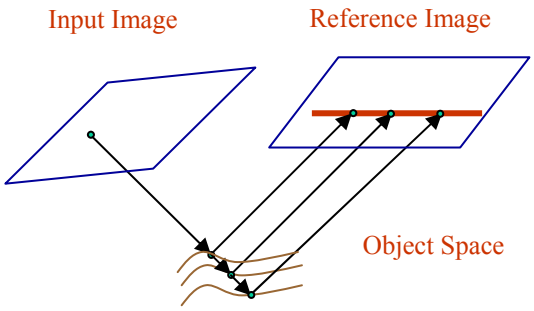



**Figure 3.5:** Rigorous Mathematical Relationship between Conjugate Points in Stereo-Images

### 3.3.2 Perspective Projection: Approximate Models

Work with rigorous models usually results in accurate mathematical representation of object to image relation. However, these models suffer from several difficulties. First, the interior orientation parameters (IOP) and exterior orientation parameters (EOP) for sensors (e.g. IKONOS) might not be always available for security reasons. Second, images provided by linear CCD array sensors consist of lines scanned independently at different instants of time with different EOP. The use of rigorous models for such images requires six parameters for each scan line, and this results in a huge set of parameters to be considered. For example 36000 parameters are involved in SPOT with 6000 scan lines. Third, rigorous modelling requires complete understanding of the nature and

operation principles of the imaging system used to capture the images. For all these reasons, there has been an increasing trend within the photogrammetric community towards using approximate models to describe the mathematical relationship between image and object space points for scenes captured by high altitude line cameras with a narrow angular field of view (e.g., IKONOS, SPOT, LANDAST, EROS-A1, QUICKBIRD, and ORBVIEW) (Figure 3.6). The main advantage of approximate models is their capability to extract reliable and accurate 3D information from stereo satellite images without explicit reference to either camera model or satellite ephemeris information (Fraser et al., 2001).

Among these models, Rational Function Models (RFM) are gaining popularity since they can handle any type of imagery without the need for a comprehensive understanding of the operational principles of the imaging system (Tao and Hu, 2001). RFM are fractional polynomial functions that express the image coordinates as a function of object space coordinates (Equation 3.7). RFM have been extensively used in processing satellite scenes in the absence of the rigorous sensor model (e.g., IKONOS scenes). However, the use of RFM would not allow for the development of a closed form transformation function between the coordinates of conjugate points in the reference and input images.

$$x = \frac{p_1(X,Y,Z)}{P_2(X,Y,Z)} \qquad y = \frac{p_3(X,Y,Z)}{P_4(X,Y,Z)} \tag{3.7}$$

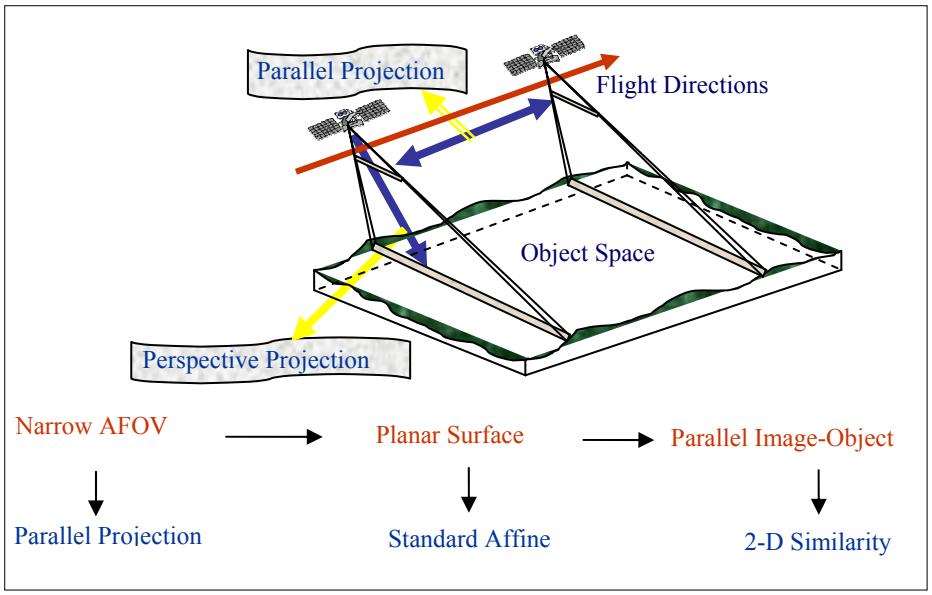



**Figure 3.6:** Approximate Models

For scenes captured by high altitude line cameras with narrow angular field of view (e.g., IKONOS, SPOT, and LANDAST), parallel projection approximates the mathematical relationship between image and object space coordinates (Habib and Morgan, 2002). Image to object space coordinate transformation using parallel projection involves eight parameters, Equation 3.8.

$$x = A_1X + A_2Y + A_3Z + A_4$$
$$y = A_5X + A_6Y + A_7Z + A_8 \tag{3.8}$$

As in the case of rational functions, parallel projection would not allow for the establishment of point-to-point correspondence between the input and reference image without the knowledge of height information (Z coordinate), Figure 3.5.

For relatively planar object space (i.e., when height variation within the object space is very small compared to the flying height), the parallel projection can be simplified to an

affine transformation involving six parameters. In other words, corresponding image (either the reference or the input image) and planimetric object coordinates are related through a six-parameter affine transformation. Due to the transitive property of an affine transformation, the relationship between corresponding points in the input and reference images can be represented by an affine transformation as well (Figure 3.7). For situations where the image is almost parallel to the object space, the affine transformation function can be further approximated by a 2-D similarity transformation. Once again, since similarity transformation is transitive, coordinates of conjugate points in the reference and input images can be related to each other through a 2-D similarity transformation.

In summary, one concludes that for scenes captured by high altitude imaging satellites with narrow angular field of view of a relatively flat terrain, the mathematical relationship between the coordinates of conjugate points in the reference and input images can be described by an affine transformation. Moreover, for scenes that are almost parallel to the object space, such transformations can be further simplified to a 2-D similarity transformation involving only four parameters. Since this thesis focuses on the registration of multi-resolution satellite imagery (e.g., IKONOS, KOMPSAT, SPOT, and LANDAST), affine and 2-D similarity transformation functions will be used to establish the mathematical relationship between conjugate elements of the involved image pair.

After discussing the choice of the most appropriate registration primitives as well as the transformation function between the reference and input images, one can proceed to the third issue of the registration paradigm, namely, the similarity measure.

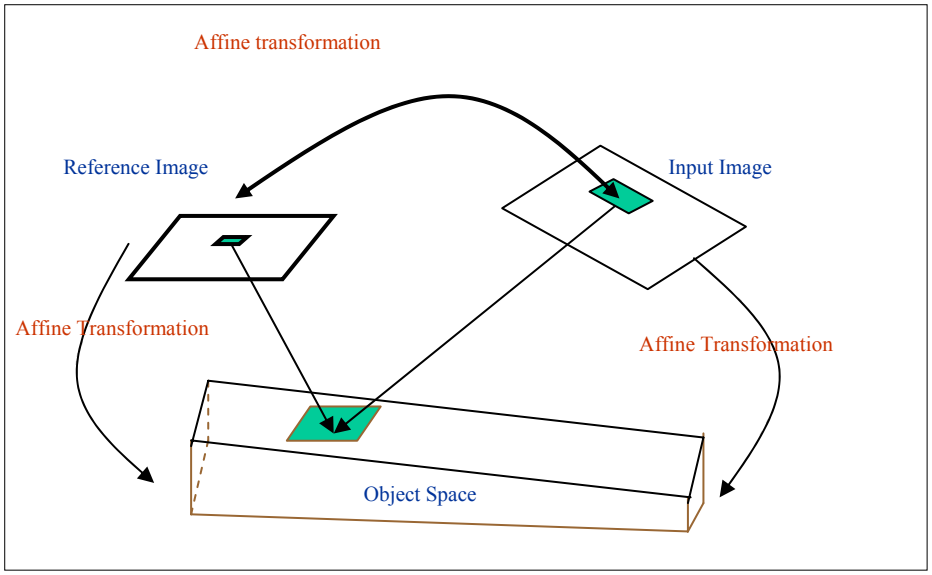



**Figure 3.7:** Transitive Property of Affine Transformation

## 3.4 Similarity Measure

The similarity measure mathematically describes the coincidence of conjugate line segments after application of the registration transformation function. The similarity measure incorporates the attributes of the registration primitives to derive the necessary constraints that can be used to estimate the parameters of the transformation function relating the reference and input images. In other words, having two datasets, which represent the registration primitives (straight-line segments) manually or automatically extracted from the input and reference images, one should derive the necessary constraints to describe the coincidence of conjugate primitives after applying the appropriate registration transformation function.

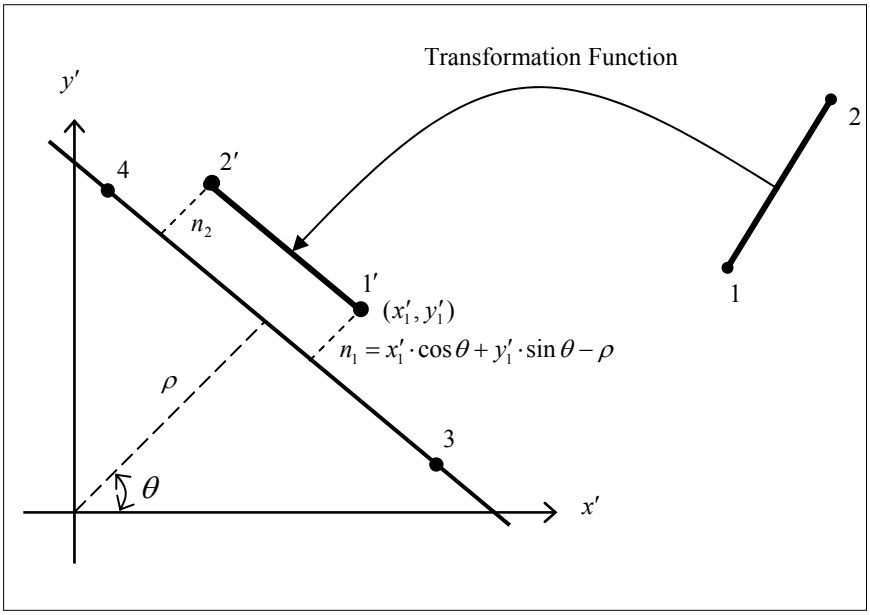



**Figure 3.8:** Similarity Measure Using Straight Line Segments

In Figure 3.8, it is assumed that the line segment (1-2) in the reference image corresponds to the line segment (3-4) in the input image. As mentioned earlier, the end points of the two segments need not be conjugate. The similarity measure should describe mathematically the fact that the line segment (1-2) should coincide with the corresponding line segment (3-4) after application of the transformation function relating the reference and input images. Such a measure can be derived by forcing the normal distances between the end points of a line segment in the reference image, after applying the transformation function, and the corresponding line segment in the input image to be zero (i.e., $n_1 = n_2 = 0$, Figure 3.8). Equation 3.9 mathematically describes such a constraint for one of the end points of the line segment in the reference image.

$$x_1' \cdot \cos\theta + y_1' \cdot \sin\theta - \rho = 0 \tag{3.9}$$

where

$(\rho, \theta)$: are the polar coordinates representing the line segment 3-4 in the input image

$(x_1', y_1')$: are the transformed coordinates of point 1 in the reference image after applying the registration transformation function.

2-D similarity and affine registration transformation functions, represented by Equations 2.1 and 2.2 respectively, can be used to describe the mathematical relationship between $(x_1, y_1)$ and $(x_1', y_1')$. Re-parameterization of Equations 2.1 and 2.2 results in Equations 3.10 and 3.11, respectively.

$$\begin{bmatrix} x_1' \\ y_1' \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \tag{3.10}$$

where

$$a_1 = s\cos\kappa \qquad b_1 = s\sin\kappa$$

$$\begin{bmatrix} x_1' \\ y_1' \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \tag{3.11}$$

where

$$a_0 = x_T \qquad a_1 = s_x\cos\kappa \qquad a_2 = s_y\sin(\kappa + \delta\kappa)$$

$$b_0 = y_T \qquad b_1 = s_x\sin\kappa \qquad b_2 = s_y\cos(\kappa + \delta\kappa)$$

One pair of conjugate line segments would yield two constraints of the form in Equation 3.9. Using a given set of "n" corresponding line segments, one can incorporate the resulting "2n" constraints in a least squares adjustment procedure to solve for the parameters of the registration transformation function ($a_0$, $b_0$, $a_1$, and $b_1$ for 2-D similarity transformation or $a_0$, $b_0$, $a_1$, $a_2$, $b_1$, and $b_2$ for affine transformation).

## 3.5 Matching Strategy

After establishing the registration primitives, transformation function, and similarity measure, one should focus on how to establish the correspondence between conjugate primitives. Corresponding primitives in the reference and input images can be manually identified. However, the large amount of data and the need for fast registration methods mandate the automation of the process of identifying conjugate primitives. Therefore, a matching strategy has to be developed to manipulate the registration primitives, the transformation function, and the similarity measure to automatically establish the correspondence between conjugate primitives. In this thesis, the Modified Iterated Hough Transform (MIHT) is used as the matching strategy. Such a methodology is attractive since it allows for simultaneous matching and parameter estimation. Moreover, it does not require a complete correspondence between the primitives in the reference and input images. MIHT has been successfully implemented in several photogrammetric operations such as automatic single photo resection and automatic relative orientation (Habib et al., 2001a; Habib and Kelley 2001a, 2001b).

Hough (1962) introduced a method of determining parameters by way of a voting scheme (Appendix A). The basic principle of his approach was to switch the roles of parameters and spatial variables. The Hough transform can be used to estimate the parameters of a mathematical model relating conjugate entities of two datasets after some modification. The modified Hough transform assumes the availability of two datasets (Figure 3.9) where the attributes of conjugate primitives are related to each other through a mathematical function (similarity measure incorporating the appropriate transformation function). The approach starts by making all possible matching hypotheses between the primitives in the datasets under consideration. For each hypothesis, the similarity measure constraints are formulated where the parameters of the transformation function can be estimated simultaneously or sequentially, depending on the number of hypothesized matches simultaneously considered.

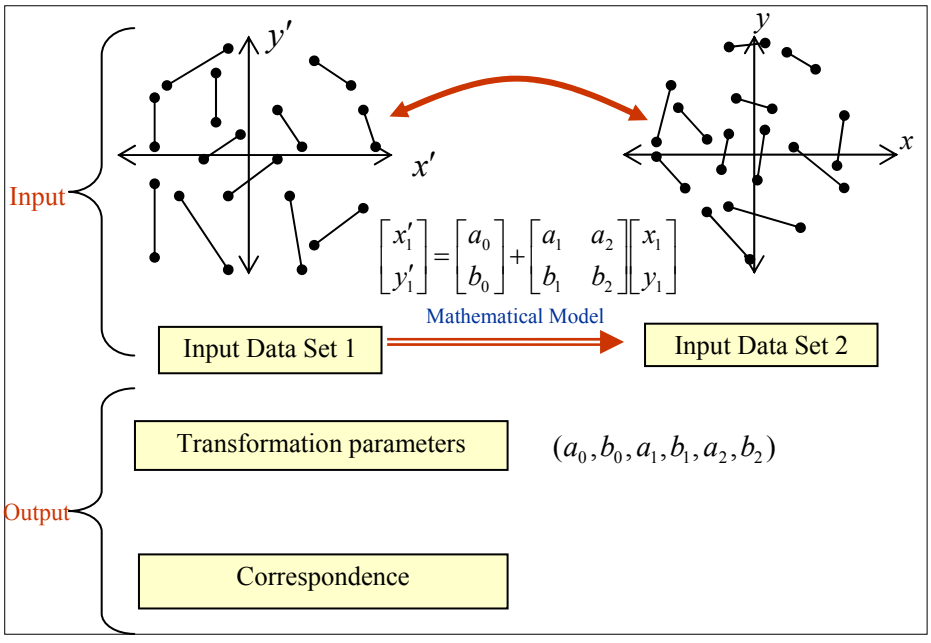



**Figure 3.9:** Summary of the Input and Output of the Modified Hough Transform

All possible entity matches are evaluated, and the results (parameter estimations) are stored in an accumulator array, which is a discrete tessellation of the range of expected numerical values for the parameters under consideration. Within the considered correspondences, correct matching hypotheses would produce the same parameters, which will manifest themselves as a distinct peak in the accumulator array. Moreover, matching hypotheses that contributed to the peak can be tracked to establish the correspondence between conjugate primitives in the involved datasets.

The number of parameters being simultaneously solved determines the dimension of the accumulator array. In order to solve "*n*" parameters simultaneously, one must utilize the number of hypothesized entity matches needed to generate the required *n* equations. However, this approach is not practical. Evaluating all permutations of entities leads to combinatorial explosion. For example, if there are *x* entities in dataset one and y entities in dataset two, solving *n* parameters simultaneously would lead to $\frac{x!}{(x-n)!} \times \frac{y!}{(y-n)!}$ combinations. In addition, the memory requirements of an *n* dimensional accumulator array create another problem.

An alternative is to solve for each parameter sequentially in an iterative manner, updating the approximations at each step (MIHT). Consequently, the accumulator array becomes one-dimensional and the memory problem disappears. Also, if there are *x* elements in dataset one and y elements in dataset two, the total number of evaluated entity matches becomes *xy*, and this reduces the computational complexity of the problem. After each iteration, the approximations are updated and the cell size of the accumulator array can be

reduced. In this manner, the parameters can be estimated with high accuracy (Habib and Kelley, 2001a).

The convergence rate towards the correct parameters depends on the independence of the parameters and the non-linearity of the transformation function. The effect of the non-linearity is similar to least squares adjustment of non-linear models. Highly non-linear models have slower convergence rates and would require more iterations. On the other hand, the independence of the parameters is more crucial. Since we are sequentially solving for the involved parameters, the quality of the estimated parameters at any time depends on the quality of other parameters (assumed to be correct). Therefore, if the parameters are independent, then we can partition the given data into subsets. The partitioning should be done in such a way that each set is only affected by a single parameter. This partitioning will lead to a faster convergence rate (Habib and Kelley, 2001a). Partitioning depends on the mathematical models under consideration. The partitioning based on affine and 2-D similarity transformation functions will be discussed in Chapter 4.

The implementation of the MIHT strategy for automatic image registration can be summarized as follows:

- An accumulator array is formed for the parameters involved in the registration transformation function (e.g., 2-D similarity or affine). The accumulator array is a discrete tessellation of the range of expected parameters solutions. The dimension of this array depends on the number of parameters to be simultaneously solved for, which is related to the number of entity pairings simultaneously considered as well as

the number of constraints provided by a single matching hypothesis. In this research, the parameters are sequentially estimated one by one and a one-dimensional accumulator array is always used (Chapter 4).

- Approximations are assumed for the parameters which are not yet determined. The cell size of the accumulator array depends on the quality of the initial values; poor approximations will require larger cell sizes.
- All possible matches between individual registration primitives within the reference and input images are evaluated. The accumulator array is incremented at the location of the resulting solution from each matching hypothesis.
- After all possible matches have been considered, the peak in the accumulator array will indicate the correct solution of the parameter in question. Only one peak is expected for a given accumulator array, Figure 3.10.

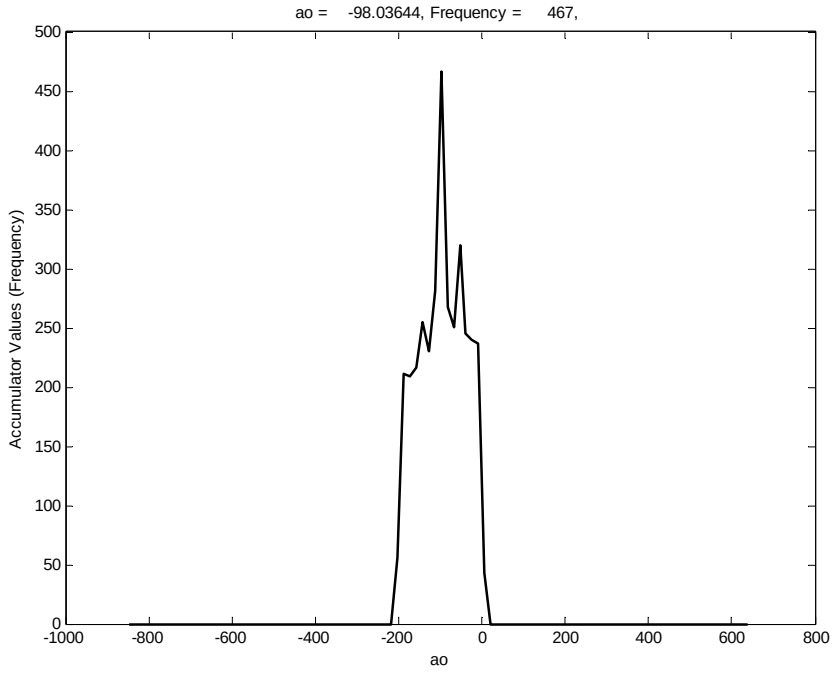



**Figure 3.10:** Example of an Accumulator Array for $a_0$ Parameter

- After each parameter is determined in a sequential manner, the approximations are updated. For the next iteration, the accumulator array cell size is decreased to reflect the improvement in the quality of the parameters. Then, the above two steps are repeated until convergence is achieved (i.e., the estimated parameters do not significantly change from one iteration to the next).

- By tracking the hypothesized matches that contributed towards the peak in the last iteration, one can determine the correspondence between conjugate primitives. These matches are then used in a simultaneous least squares adjustment to derive a stochastic estimate of the involved parameters in the registration transformation function. Figure 3.11 briefly explains the various steps in MIHT using 2-D similarity transformation.

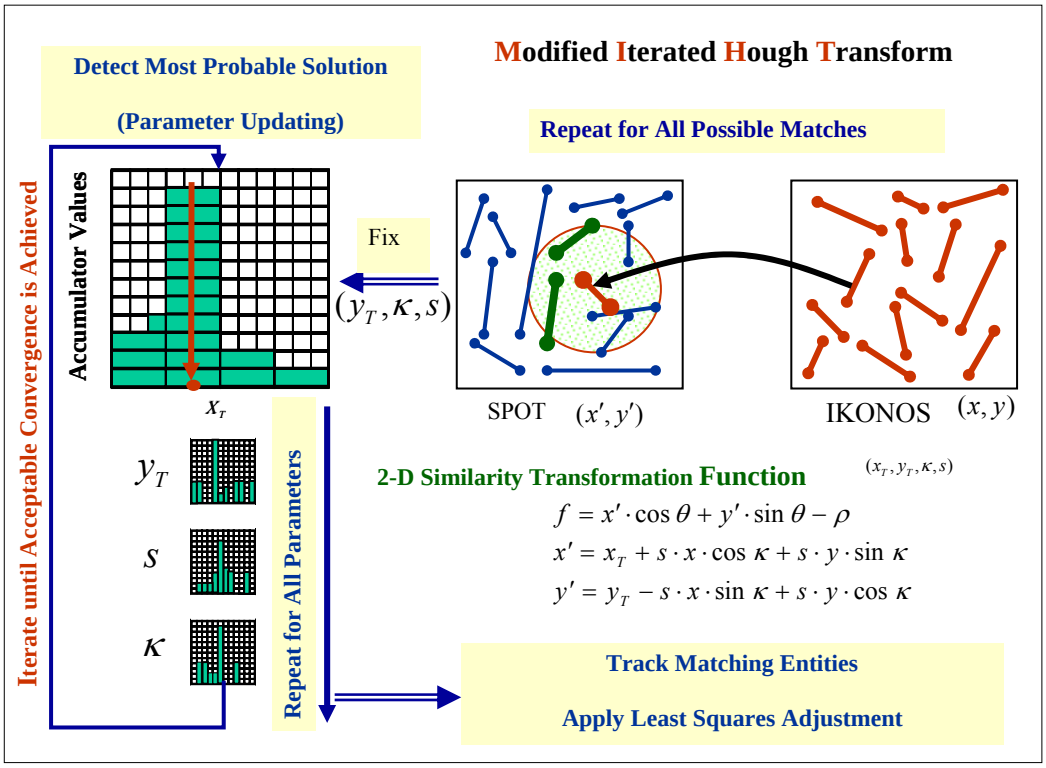



**Figure 3.11:** MIHT Implementation Using 2-D Similarity Transformation

In addition to providing simultaneous estimation of the parameters of the registration transformation function and the correspondence between conjugate primitives, the MIHT strategy will help in verifying the validity of the selected transformation function between the reference and input images. The MIHT is expected to converge if and only if the registration transformation function is appropriate (on the assumption that enough conjugate primitives exist in the involved datasets).

In summary, the automatic matching strategy (MIHT) utilizes the introduced similarity measure together with the transformation function to establish the correspondence between the extracted primitives while simultaneously solving for the parameters involved in the registration transformation function. In this approach, we assume no knowledge of conjugate entities and do not require complete correspondence between those entities Moreover, the suggested approach allows for investigating and evaluating the appropriateness of the selected registration transformation function; this has rarely been considered in previous research.

# CHAPTER 4

# OPTIMAL SEQUENCE FOR PARAMETER ESTIMATION

## 4.1 Introduction

MIHT sequentially derives an estimate of the parameters of the registration transformation function. Therefore, a decision has to be made concerning the optimum sequence for parameter estimation that guarantees fast and robust convergence to the correct solution. Through analysis of the similarity measure, it can be established that linear features with different orientations at various regions of the image are influenced differently by the involved parameters. In other words, some parameters have low influence on specific line segments at some regions while having larger influence on others. Thus, certain regions in the images under consideration would be useful for the estimation of certain parameters if they have large influence at that region, while other parameters have minor or no influence at the same region.

The following subsections deal with how to determine the optimal sequence for parameter estimation when affine and 2-D similarity registration transformation functions are used.

## 4.2 Affine Transformation

The assumption is that the mathematical model between the coordinate systems shown in Figure 4.1 can be represented by 2-D affine transformation.

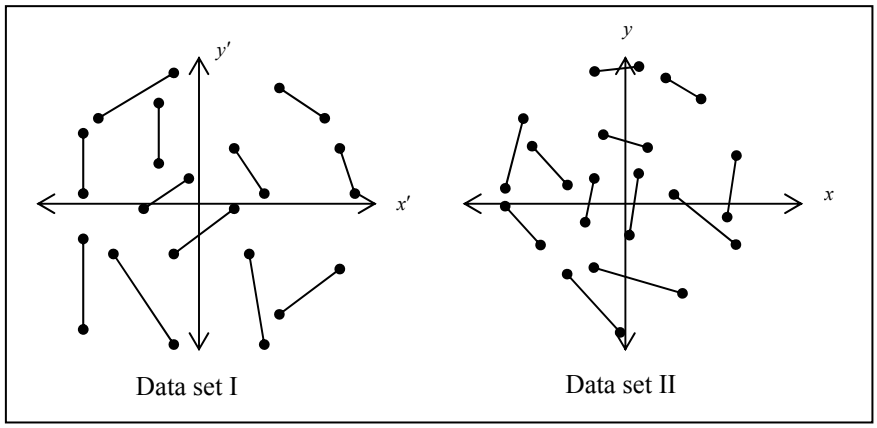



**Figure 4.1:** Example of Linear Features Extracted From Input and Reference Images

With reference to Figure 4.2., affine transformation function can be expressed as:

$$x' = x_T + s_x \cdot s \cdot \cos\kappa + s_y \cdot y \cdot \sin(\kappa + \delta\kappa)$$
$$y' = y_T - s_x \cdot x \cdot \sin\kappa + s_Y \cdot y \cdot \cos(\kappa + \delta\kappa) \quad (4.1a)$$

in matrix form as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x_T \\ y_T \end{bmatrix} + \begin{bmatrix} s_x \cos\kappa & s_y \sin(\kappa + \delta\kappa) \\ -s_x \sin\kappa & s_Y \cos(\kappa + \delta\kappa) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.1b)$$

where

$(x, y)$ : Coordinates of a point in the reference image.

$(x', y')$ : Coordinates of the conjugate point in the input image.

$x_T$ : Shift in x-axis

$y_T$ : Shift in y-axis

$\kappa$ : Rotation angle

$s_x$ : Scale factor along x-axis

$s_y$ : Scale factor along y-axis

$\delta\kappa$ : Non-orthogonality angle.

The previous equation can be rewritten as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{4.1c}$$

where

$a_0 = x_T$ $\quad a_1 = s_x \cos\kappa$ $\quad a_2 = s_y \sin(\kappa + \delta\kappa)$

$b_0 = y_T$ $\quad b_1 = s_x \sin\kappa$ $\quad b_2 = s_y \cos(\kappa + \delta\kappa)$

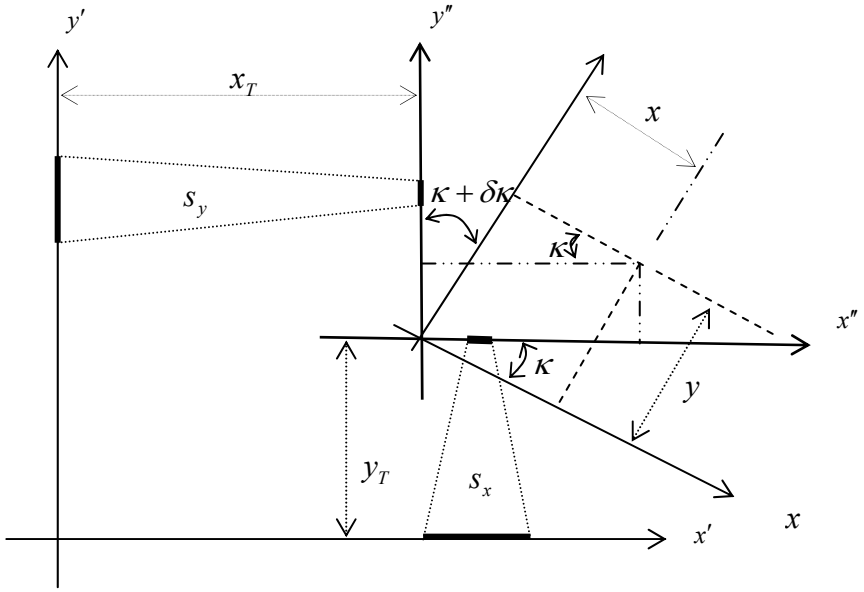



**Figure 4.2:** Affine Transformation Parameters

The similarity measure in Figure 3.8, which ensures that a line segment in the reference image is conjugate to another line segment in the input image, is expressed by two constraints involving the normal distance between conjugate lines as shown in Equation 4.2.

$$f = x' \cos\theta + y' \sin\theta - \rho = 0 \tag{4.2}$$

where

$$x' = a_0 + a_1\ x + a_2\ y$$

$$y' = b_0 + b_1\ x + b_2\ y$$

To establish the optimal sequence for parameter estimation, one has to determine how the normal distance, between the transformed end points of the line segment in the reference image and the corresponding segment in the input image, will change as a result of incremental changes in the parameters of the transformation function. The magnitude of these changes should be evaluated for line segments with different orientations at various locations in the involved imagery. Large changes resulting from an incremental change in a single parameter, using specific line segments at certain locations, indicate the appropriateness of these segments in this area for the computation of that parameter. For that purpose, the images to be registered are divided into five regions labeled from 1 to 5 (Figure 4.3). One can assume that points in Region 5 have very small *x*-, y-coordinate values ($x_5 \approx y_5 \approx 0$). Also, points in Regions 1 and 3 have smaller *x*-coordinate values

when compared to their *y*-coordinates ($|x_{1,3}| < |y_{1,3}|$), while points in Regions 2 and 4 have smaller *y*-coordinates values when compared to their *x*-coordinates ($|y_{2,4}| < |x_{2,4}|$).

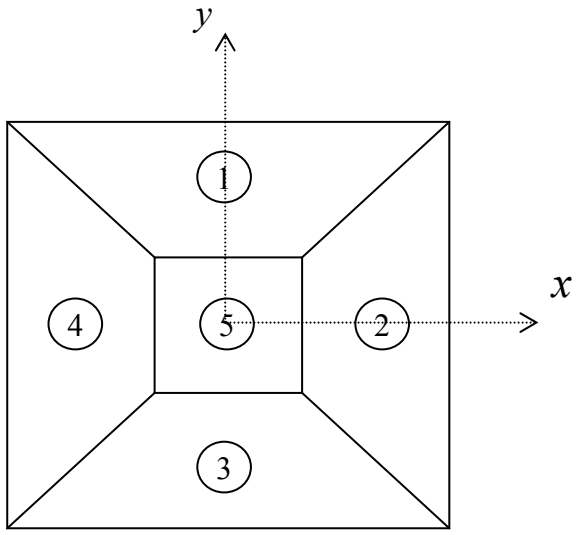



**Figure 4.3:** Image Partitioning for Sequential Estimation of the Transformation Parameters

To illustrate the optimal sequence concept, let us evaluate the effect of incremental changes in $(a_0, b_0, a_1, b_1, a_2, b_2)$, as expressed by the respective partial derivative, on the normal distance described by the similarity measure, as follows:

*Partial derivative with respect to* $(a_0)$:

$$\frac{\partial f}{\partial a_0} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial a_0} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial a_0}$$

$$\frac{\partial f}{\partial a_0} = \cos\theta \tag{4.3}$$

A closer look at Equation 4.3 reveals the following facts:

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to no change in the normal distance ($\partial f / \partial a_0 \approx 0$). Therefore, horizontal line segments are not useful for $a_0$ estimation.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to a change in the normal distance ($\partial f / \partial a_0 \approx 1$). Therefore, vertical line segments are useful for $a_0$ estimation.

*Partial derivative with respect to* ($b_0$):

$$\frac{\partial f}{\partial b_0} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial b_0} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial b_0}$$

$$\frac{\partial f}{\partial b_0} = \sin\theta \tag{4.4}$$

A closer look at Equation 4.4 reveals the following facts:

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to a change in the normal distance ($\partial f / \partial b_0 \approx 1$). Therefore, horizontal line segments are useful for $b_0$ estimation.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to no change in the normal distance ($\partial f / \partial b_0 \approx 0$). Therefore, vertical line segments are not useful for $b_0$ estimation.

*Partial derivative with respect to* ($a_1$):

$$\frac{\partial f}{\partial a_1} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial a_1} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial a_1}$$

$$\frac{\partial f}{\partial a_1} = x \cos\theta \tag{4.5}$$

A closer look at Equation 4.5 reveals the following facts:

- Line segments in the reference image with relatively small *x*-coordinates ($x \approx 0$) would lead to no change in the normal distance ($\partial f / \partial a_1 \approx 0$). Therefore, line segments that are close to the y-axis are not useful for $a_1$ estimation, regardless of their orientation.

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to no change in the normal distance ($\partial f / \partial a_1 \approx 0$). Therefore, horizontal line segments are not useful for $a_1$ estimation, regardless of their location in the images to be registered.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to a change in the normal distance that is proportional to their *x*-coordinate in the reference image ($\partial f / \partial a_1 \approx x$). Therefore, vertical line segments that are far from the y-axis are useful for $a_1$ estimation.

*Partial derivative with respect to* ($b_1$):

$$\frac{\partial f}{\partial b_1} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial b_1} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial b_1}$$

$$\frac{\partial f}{\partial b_1} = x \sin\theta \tag{4.6}$$

A closer look at Equation 4.6 reveals the following facts:

- Line segments in the reference image with relatively small *x*-coordinates ($x \approx 0$) would lead to no change in the normal distance ($\partial f / \partial b_1 \approx 0$). Therefore, line segments that are close to the y-axis are not useful for $b_1$ estimation regardless of their orientation.

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to a change in the normal distance that is proportional to their *x*-coordinate in the reference image ($\partial f / \partial b_1 \approx x$). Therefore, horizontal line segments that are far from the y-axis are useful for $b_1$ estimation.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to no change in the normal distance ($\partial f / \partial b_1 \approx 0$). Therefore, vertical line segments are not useful for $b_1$ estimation regardless of their location in the images to be registered.

*Partial derivative with respect to* ($a_2$):

$$\frac{\partial f}{\partial a_2} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial a_2} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial a_2}$$

$$\frac{\partial f}{\partial a_2} = y \cos\theta \tag{4.7}$$

A closer look at Equation 4.7 reveals the following facts:

- Line segments in the reference image with relatively small *y*-coordinates ($y \approx 0$) would lead to no change in the normal distance ($\partial f / \partial a_2 \approx 0$). Therefore, line segments that are close to the x-axis are not useful for $a_2$ estimation, regardless of their orientation.

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to no change in the normal distance ($\partial f / \partial a_2 \approx 0$). Therefore, horizontal line segments are not useful for $a_2$ estimation, regardless of their location in the images to be registered.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to a change in the normal distance that is proportional to their *y*-coordinate in the reference image ($\partial f / \partial a_2 \approx y$). Therefore, vertical line segments that are far from the x-axis are useful for $a_2$ estimation.

*Partial derivative with respect to* ($b_2$):

$$\frac{\partial f}{\partial b_2} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial b_2} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial b_2}$$

$$\frac{\partial f}{\partial b_2} = y \sin\theta \tag{4.8}$$

A closer look at Equation 4.8 reveals the following facts:

- Line segments in the reference image with relatively small *y*-coordinates ($y \approx 0$) would lead to no change in the normal distance ($\partial f / \partial b_2 \approx 0$). Therefore, line segments that are close to the x-axis are not useful for $b_2$ estimation, regardless of their orientation.

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to a change in the normal distance that is proportional to their *y*-coordinate in the reference image ($\partial f / \partial b_2 \approx y$). Therefore, horizontal line segments that are far from the x-axis are useful for $b_2$ estimation.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to no change in the normal distance ($\partial f / \partial b_2 \approx 0$). Therefore, vertical line segments are not useful for $b_2$ estimation, regardless of their location in the images to be registered.

**Table 4.1:** The Influence of Different Image Regions on the Affine Transformation Parameters

| Region | | ① | ② | ③ | ④ | ⑤ |
|---|---|---|---|---|---|---|
| | | $x \approx 0$ $y \neq 0$ | $x \neq 0$ $y \approx 0$ | $x \approx 0$ $y \neq 0$ | $x \neq 0$ $y \approx 0$ | $x \approx 0$ $y \approx 0$ |
| $\frac{\partial f}{\partial a_0} = \cos\theta$ | $\theta = 0^0$ | 1 | 1 | 1 | 1 | 1 |
| | $\theta = 90^0$ | 0 | 0 | 0 | 0 | 0 |
| $\frac{\partial f}{\partial b_0} = \sin\theta$ | $\theta = 0^0$ | 0 | 0 | 0 | 0 | 0 |
| | $\theta = 90^0$ | 1 | 1 | 1 | 1 | 1 |
| $\frac{\partial f}{\partial a_1} = x\cos\theta$ | $\theta = 0^0$ | 0 | $x$ | 0 | $x$ | 0 |
| | $\theta = 90^0$ | 0 | 0 | 0 | 0 | 0 |
| $\frac{\partial f}{\partial b_1} = x\sin\theta$ | $\theta = 0^0$ | 0 | 0 | 0 | 0 | 0 |
| | $\theta = 90^0$ | 0 | $x$ | 0 | $x$ | 0 |
| $\frac{\partial f}{\partial a_2} = y\cos\theta$ | $\theta = 0^0$ | $y$ | 0 | $y$ | 0 | 0 |
| | $\theta = 90^0$ | 0 | 0 | 0 | 0 | 0 |
| $\frac{\partial f}{\partial b_2} = y\sin\theta$ | $\theta = 0^0$ | 0 | 0 | 0 | 0 | 0 |
| | $\theta = 90^0$ | $y$ | 0 | $y$ | 0 | 0 |

Table 4.1 summarizes the above discussion. It shows analysis of the partial derivatives of the similarity measure function with respect to the unknown parameters of the affine transformation function. As a result, one can derive the optimum sequence for parameter estimation in the following manner (Figure 4.4):

- Use vertical lines in Region 5 to estimate $a_0$.
- Use horizontal lines in Region 5 to estimate $b_0$.

- Use vertical lines in Regions 2 and 4 to estimate $a_1$.
- Use horizontal lines in Regions 2 and 4 to estimate $b_1$.
- Use vertical lines in Regions 1 and 3 to estimate $a_2$.
- Use horizontal lines in Regions 1 and 3 to estimate $b_2$.

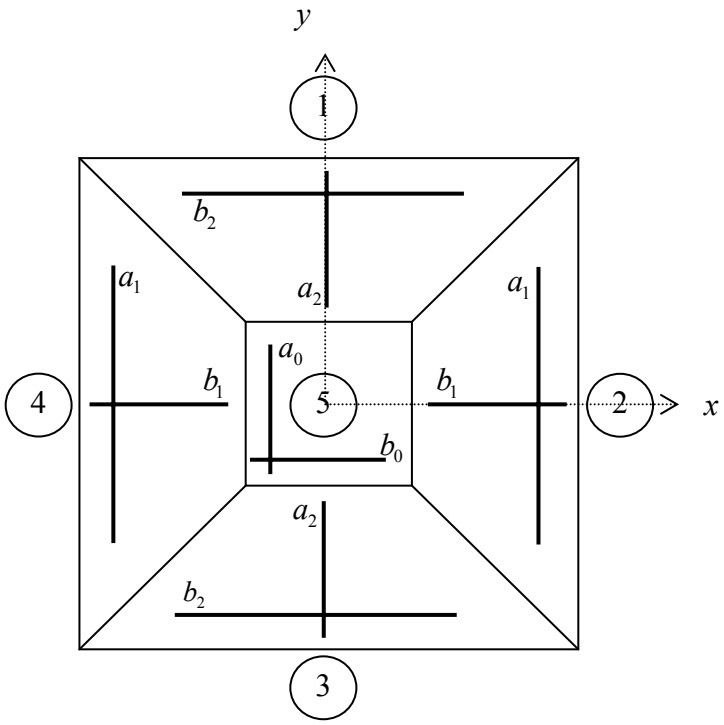



**Figure 4.4:** Optimal Sequence for Affine Transformation Parameters

As illustrated in Table 4.1, vertical line segments in various areas of the involved imagery can be used for estimating the shift component $a_0$ (as indicated by non-zero partial derivatives). However, only vertical lines in Region 5 are used. This restriction is imposed because variations in the normal distances between conjugate line segments, as expressed by the partial derivatives, from the optimum value (zero) in Region 5 are only

attributed to $a_0$. On the other hand, variations in Regions 1 and 3 can be ascribed to $a_0$ and $a_2$. Similarly, variations in Regions 2 and 4 are attributed to $a_0$ and $a_1$. Therefore, vertical line segments in Region 5 are the only alternative for ensuring the separation among $a_0$, $a_1$, and $a_2$. Thus, a certain region would be useful for the estimation of a certain parameter in either of the following cases: 1) the parameter has a large influence at that region, while other parameters have minor or no influence at the same region; 2) other parameters affecting that region have been already estimated. Following the same argument, one can see that only horizontal lines in area 5 would allow for the separation among $b_0$, $b_1$, and $b_2$.

The evaluated partial derivatives in Table 4.1 assume that the $x$-coordinates in Regions 1 and 3 are small ( $x_1 \approx x_3 \approx 0$ ) and the $y$-coordinates in Regions 2 and 4 are small as well ( $y_2 \approx y_4 \approx 0$ ). Even if these assumptions might not be true, the sequential procedure would still work. However, more iteration cycles would be required until convergence. This is mainly because the partitioning is only needed to identify regions in the image that are affected significantly more than others by incremental changes in the parameters of the transformation function. It should be noted that the requirement for vertical and horizontal line segments is not stringent. The suggested procedure can be implemented through the use of predominantly horizontal or vertical segments or both. The deviation from being truly vertical or horizontal would only lead to a slower convergence. The lack of features in any of the five regions will only slow the convergence process, because all parameters affect all regions, but with different magnitudes.

### 4.3 2-D Similarity Transformation

Let us assume that the mathematical model between the coordinate systems shown in Figure 4.1 can be represented by 2-D similarity transformation.

In Figure 4.5, 2-D similarity transformation function can be expressed as:

$$\begin{aligned} x' &= x_T + s\cos\kappa\, x + s\sin\kappa\, y \\ y' &= y_T - s\sin\kappa\, x + s\cos\kappa\, y \end{aligned} \tag{4.9a}$$

or

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x_T \\ y_T \end{bmatrix} + s \begin{bmatrix} \cos\kappa & \sin\kappa \\ -\sin\kappa & \cos\kappa \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{4.9b}$$

where

$(x, y)$ : Coordinates of a point in the reference image

$(x', y')$ : Coordinates of the conjugate point in the input image

$x_T$ : Shift in x-axis

$y_T$ : Shift in y-axis

$\kappa$ : Rotation angle

$s$ : Scale factor.

The previous equation can be rewritten as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \tag{4.9c}$$

where

$a_1 = s\cos\kappa$ $\quad$ $b_1 = s\sin\kappa$

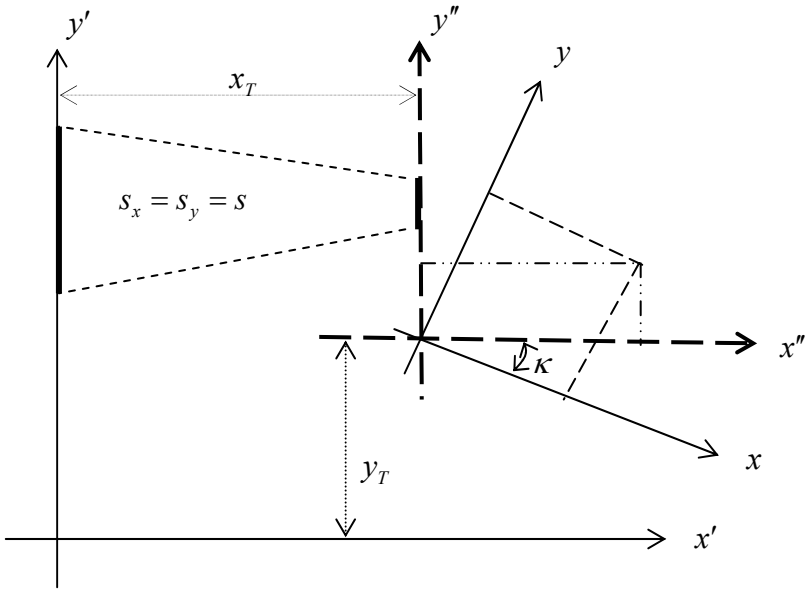



**Figure 4.5:** 2-D Similarity Transformation Parameters

The similarity measure in Figure 3.8 ensuring that a line segment in the reference image is conjugate to another line segment in the input image is expressed by two constraints of the form in Equation 4.10:

$$f = x' \cos\theta + y' \sin\theta - \rho = 0 \tag{4.10}$$

where

$$x' = x_T + s \cos\kappa\, x + s \sin\kappa\, y$$

$$y' = y_T - s \sin\kappa\, x + s \cos\kappa\, y$$

As with affine transformation (Section 4.1), the optimal sequence can be evaluated on the basis of the effect of an incremental change in $(x_T, y_T, s, \kappa)$, as expressed by the

respective partial derivative on the normal distance described by the similarity measure as follows:

*Shift in x-axis* ($x_T$):

$$\frac{\partial f}{\partial x_T} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial x_T} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial x_T}$$

$$\frac{\partial f}{\partial x_T} = \cos\theta \tag{4.11}$$

A closer look at Equation 4.11 reveals the following facts:

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to no change in the normal distance ($\partial f / \partial x_T \approx 0$). Therefore, horizontal line segments are not useful for $x_T$ estimation.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to a change in the normal distance ($\partial f / \partial x_T \approx 1$). Therefore, vertical line segments are useful for $x_T$ estimation.

*Shift in y-axis* ($y_T$):

$$\frac{\partial f}{\partial y_T} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial y_T} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial y_T}$$

$$\frac{\partial f}{\partial y_T} = \sin\theta \tag{4.12}$$

A closer look at Equation 4.12 reveals the following facts:

- Horizontal line segments in the input image ($\theta = 90^0$) would lead to a change in the normal distance ($\partial f / \partial y_T \approx 1$). Therefore, horizontal line segments are useful for $y_T$ estimation.

- Vertical line segments in the input image ($\theta = 0^0$) would lead to no change in the normal distance ($\partial f / \partial y_T \approx 0$). Therefore, vertical line segments are not useful for $y_T$ estimation.

*Rotation angle* ($\kappa$):

$$\frac{\partial f}{\partial \kappa} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial \kappa} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial \kappa}$$

$$\frac{\partial f}{\partial \kappa} = \cos\theta \left[ -s \sin\kappa\, x + s \cos\kappa\, y \right] + \sin\theta \left[ -s \cos\kappa\, x - s \sin\kappa\, y \right] \tag{4.13}$$

Assuming small rotation angle, $\kappa \approx 0$, Equation 4.13 can be re-written as:

$$\frac{\partial f}{\partial \kappa} = s\, y \cos\theta - s\, x \sin\theta \tag{4.14}$$

A closer look at Equation 4.14 reveals the following facts:

- Line segments in the reference image with relatively small *x*-, *y*-coordinates ($x \approx y \approx 0$) would lead to no change in the normal distance ($\partial f / \partial \kappa \approx 0$).

Therefore, line segments that are close to the center are not useful for $\kappa$ estimation regardless of their orientation.

- Line segments in the reference image with relatively small *y*-coordinates ($y \approx 0$) would have the following effects:

a. Horizontal line segments in the input image ($\theta = 90^0$) would lead to a change in the normal distance that is proportional to their *x*-coordinate in the reference image ($\partial f / \partial \kappa \approx -s\, x$). Therefore, horizontal line segments with relatively small *y*-coordinates and far from the y-axis are useful for $\kappa$ estimation.

b. Vertical line segments in the input image ($\theta = 0^0$) would lead to no change in the normal distance ($\partial f / \partial \kappa \approx 0$). Therefore, vertical line segments with relatively small *y*-coordinates are not useful for $\kappa$ estimation.

- Line segments in the reference image with relatively small *x*-coordinates ($x \approx 0$) would have the following effect:

a. Horizontal line segments in the input image ($\theta = 90^0$) would lead to no change in the normal distance ($\partial f / \kappa \approx 0$). Therefore, horizontal line segments with relatively small *x*-coordinates are not useful for $\kappa$ estimation.

b. Vertical line segments in the input image ($\theta = 0^0$) would lead to a change in the normal distance that is proportional to their y-coordinate in the reference image ($\partial f / \kappa \approx s\, y$). Therefore, vertical line segments with relatively small *x*-coordinates and far from the x-axis are useful for $\kappa$ estimation.

*Scale factor* ( $s$ ):

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x'} \cdot \frac{\partial x'}{\partial s} + \frac{\partial f}{\partial y'} \cdot \frac{\partial y'}{\partial s}$$

$$\frac{\partial f}{\partial s} = \cos\theta \left[ x \cos\kappa + y \sin\kappa \right] + \sin\theta \left[ - x \sin\kappa + y \cos\kappa \right] \tag{4.15}$$

Assume dealing with a small rotation angle ( $\kappa \approx 0$ ), Equation 4.15 can be reduced to

$$\frac{\partial f}{\partial s} = x \cos\theta + y \sin\theta \tag{4.16}$$

A closer look at Equation 4.6 reveals the following facts:

- Line segments in the reference image with relatively small x-, y-coordinates ( $x \approx y \approx 0$ ) would lead to no change in the normal distance ( $\partial f / \partial s \approx 0$ ). Therefore, line segments that are close to the center are not useful for scale estimation, regardless of their orientation.

- Line segments in the reference image with relatively small *y*-coordinates ( $y \approx 0$ ) would have the following effect:

  a. Horizontal line segments in the input image ( $\theta = 90^0$ ) would lead to no change in the normal distance ( $\partial f / \partial s \approx 0$ ). Therefore, horizontal line segments with relatively small *y*-coordinates are not useful for scale estimation.

b. Vertical line segments in the input image ($\theta = 0^0$) would lead to a change in the normal distance that is proportional to their *x*-coordinate in the reference image ($\partial f / \partial s \approx x$). Therefore, vertical line segments with relatively small *y*-coordinates and far from the y-axis are useful for scale estimation.

- Line segments in the reference image with relatively small *x*-coordinates ($x \approx 0$) would have the following effects:

  a. Horizontal line segments in the input image ($\theta = 90^0$) would lead to a change in the normal distance that is proportional to their *y*-coordinate in the reference image ($\partial f / \partial s \approx y$). Therefore, horizontal line segments with relatively small *x*-coordinates and far from the x-axis are useful for scale estimation.

  b. Vertical line segments in the input image ($\theta = 0^0$) would lead to no change in the normal distance ($\partial f / \partial s \approx 0$). Therefore, vertical line segments with relatively small *x*-coordinates are not useful for scale estimation.

**Table 4.2:** The Influence of Different Image Regions on the 2-D Similarity Transformation Parameters

| Region | | ① | ② | ③ | ④ | ⑤ |
|---|---|---|---|---|---|---|
| | | $x \approx 0$ $y \neq 0$ | $x \neq 0$ $y \approx 0$ | $x \approx 0$ $y \neq 0$ | $x \neq 0$ $y \approx 0$ | $x \approx 0$ $y \approx 0$ |
| $\frac{\partial f}{\partial x_T} = \cos\theta$ | $\theta = 0^0$ | 1 | 1 | 1 | 1 | 1 |
| | $\theta = 90^0$ | 0 | 0 | 0 | 0 | 0 |
| $\frac{\partial f}{\partial y_T} = \sin\theta$ | $\theta = 0^0$ | 0 | 0 | 0 | 0 | 0 |
| | $\theta = 90^0$ | 1 | 1 | 1 | 1 | 1 |
| $\frac{\partial f}{\partial \kappa} = s\ y\cos\theta - s\ x\sin\theta$ | $\theta = 0^0$ | $s\,y$ | 0 | $s\,y$ | 0 | 0 |
| | $\theta = 90^0$ | 0 | $s\,x$ | 0 | $s\,x$ | 0 |
| $\frac{\partial f}{\partial s} = x\cos\theta + y\sin\theta$ | $\theta = 0^0$ | 0 | $x$ | 0 | $x$ | 0 |
| | $\theta = 90^0$ | $y$ | 0 | $y$ | 0 | 0 |

Table 4.2 summarizes the above discussion for analyzing the partial derivatives of the similarity measure function with respect to the unknown parameters of the transformation function. As a result, one can derive the optimum sequence for parameter estimation as follows, Figure 4.6:

- Use vertical lines in Region 5 to estimate $x_T$.
- Use horizontal lines in Region 5 to estimate $y_T$.
- Use horizontal lines in Regions 2 and 4 and vertical lines in Regions 1 and 3 to estimate $\kappa$.
- Use horizontal lines in Regions 1 and 3 and vertical lines in Regions 2 and 4 to estimate $s$.

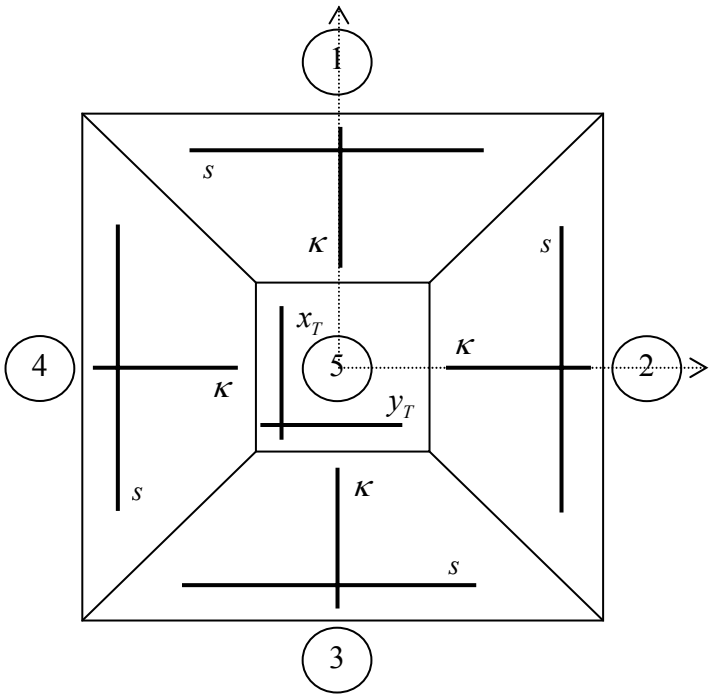



**Figure 4.6:** Optimal Sequence for 2-D Similarity Transformation Parameters

Once again, vertical line segments in Region 5 are used for $x_T$ estimation while horizontal lines in the same area are used for $y_T$ estimation, since they will allow for the separation among $x_T$, $y_T$, and the remaining parameters ($\kappa$ and s). Also, the partial derivatives in Table 4.2 assume a small rotation angle ($\kappa$). Even if this assumption might not be true, the sequential procedure would still work. However, more iteration cycles would be required until convergence. This is mainly because the main objective is to identify the optimal sequence that is based on the relative contribution of line segments with different orientations at different image regions towards the estimation of the parameters of the registration transformation function.

# CHAPTER 5

# CHANGE DETECTION METHODOLOGY

## 5.1 Introduction

Applications utilizing multi-temporal remotely sensed images are dependent on accurate registration of the involved images. Change detection, as one of the most important applications of image registration, can be defined as the process of identifying differences in land cover over time. As human and natural forces continue to alter the landscape, it is important to develop monitoring methods to assess and quantify these changes. Such changes have to be accurately and reliably inventoried to understand fully the physical and human processes at work (Estes, 1992). Recent advances in satellite imagery, in terms of improved spatial and temporal resolutions, are allowing for efficient identification of change patterns and the prediction of areas of growth. Change detection analysis might involve multi-spectral, multi-source, and multi-resolution images that have been captured at different times. Accurate co-registration of these datasets is a prerequisite step and a key factor in the developmental of a reliable change detection procedure.

In general, the uncertainty of change detection outcome depends on two factors: geometric and radiometric differences in the involved images. Sections 5.2 and 5.3, will discuss the impact of potential geometric and radiometric differences on the change

detection results. Section 5.4 demonstrates the proposed methodology of change detection.

## 5.2 Geometric Differences

High resolution overlapping scenes captured by space-borne platforms (e.g., LANDSAT-7, IKONOS, QUICKBIRD, ORBVIEW, EROS-A1, and SPOT-5) and aerial images are becoming more available at a reasonable cost. These images represent the main source of recent and historical information necessary for change detection application. Because of different imaging systems, spatial resolutions, viewing points and perspective geometry between these temporal images, geometric differences should be expected. Reliable change detection is contingent on accurate compensation of these differences among the involved images.

To ensure accuracy in the performance of change detection, one must apply a co-registration process so that pixels in the same position in the two images (input image and resampled reference image) belong to the same object on the ground. If accurate registration between images is not achieved, then change detection techniques that are based on image differencing will cause spurious changes; different locations are compared instead of locations belonging to the same object space. In other words, Pixel-by-pixel differencing methodologies for change detection are sensitive to registration errors. On other hand, change detection techniques, which are based on classification methods, will be able to tolerate geometric differences. However, multi-spectral data is needed for such a methodology.

In this work, the proposed technique for change detection is designed to cope with possible geometric differences among the scenes in question. The proposed registration methodology described in Chapters 3 and 4 will accurately align the images in question, regardless of the possible geometric differences. Experimental results set out in Chapter 6 will show that the range of sub-pixel to a few pixels accuracy has been achieved. Moreover, the suggested change detection methodology (Section 5.4) should tolerate possible remaining geometric differences (in the order of few pixels) after performance of the registration procedure.

### 5.3 Radiometric Differences

The basic premise in using remote sensing data for change detection application is that changes in land cover will result in changes in radiance values. Moreover, changes in radiance due to land cover change must be larger than radiance changes caused by other factors (Ingram et al., 1981). These other factors might include differences in atmospheric conditions, sun angle, and soil moisture (Jenson, 1983). One should expect that these factors will affect the reliability of change detection algorithms, especially when one is considering images with varying radiometric and spatial resolutions, which are captured by different sensors.

To alleviate the effect of these factors, intensity normalization is traditionally used as a pre-processing technique, compensating for possible illumination variations between the involved images. In this type of pre-processing, the intensity values in the second image are normalized to have the same mean and variance values as those in the first image.

Assuming that the involved images are co-registered relative to a common reference frame, one can proceed by applying image-differencing methods to create a new image that represents the change. The comparison results are based on the assumption that the differences between the radiometric properties of corresponding pixels are due to actual changes in the object space. However, these differences could be the result of other factors, such as different atmospheric conditions, noise, different imaging sensors, and errors in registration. Moreover, the difference image is usually binarized by thresholding, where thresholds are empirically selected. In these cases, traditional approaches to change detection that are based on the differencing of intensity images fail.

This problem is overcome through the use of derived edges from the registered images as the basis of the proposed change detection methodology. Edges are used because they are invariant with respect to possible radiometric differences between the images in question.

In summary, uncertainty in the change detection outcome relies on two factors. First, it is affected by possible radiometric differences due to atmospheric changes and different sensor types. Second, the detected changes might be biased by inaccurate registration procedure. The effect of radiometric differences between the images in question can be mitigated by using image derivatives that are robust to such differences (e.g., edge images). However, accurate co-registration remains a necessary pre-processing step for all change detection algorithms. In fact, accurate registration of multi-source imagery can be considered one of the most important components of an accurate and reliable change detection procedure.

## 5.4 Change Detection Methodology

The proposed method for change detection is as follows:

- Resampling (see appendix B) the input image into the reference frame associated with the reference image. The parameters of the registration transformation function (Chapters 2 and 3) are used in the resampling process. After resampling, corresponding pixels are assumed to belong to the same object space feature.
- Applying intensity normalization techniques to the images in question (e.g., to ensure that they have the same intensity mean and variance values) in order to reduce the radiometric differences between the involved images. However, this procedure would not be enough to eliminate radiometric differences in the involved images.
- Extracting edge cells from the resampled images using the Canny edge detector (Canny, 1986). As mentioned before, utilizing the edge images has two advantages. First, derived edges are robust to possible radiometric differences between the registered images (e.g., due to noise and different spectral properties). Also, the edges would correspond to interesting features (e.g., building boundaries, roads, trails, etc.). Therefore, comparing edge images will be useful in outlining the amount of urbanization activities, which is one of the most important objectives of change detection exercises. The final output of the edge extraction process will be binary images in which white pixels refer to edges while black pixels refer to non-edges.
- Applying the majority filter (also known as mode filter) to the edge images. In the proposed methodology for change detection, the majority filter is implemented for the following reasons:

- To allow for binary transformation of edge images, thus compensating for any registration errors (in the order of few pixels).
- To balance the effect of varying edge densities in the registered images, especially when dealing with multi-source images.
- To fill small gaps within an area with numerous edges, and smooth object boundaries without expanding or shrinking objects, Figure 5.1-a.
- To eliminate isolated edges, Figure 5.1-b.

As a result, filtered images will highlight areas with interesting features since they would lead to a dense distribution of edge cells.

In general, the majority filter is applied to binary images where a window is centered at each pixel and the value of this pixel is changed or maintained according to the majority of the pixels within this window (Lillesand and Kiefer 2000). The advantages of a majority filter can be summarised as follows:

- It does not generally shrink or expand objects.
- It smoothes object boundaries.
- It removes small peninsulas, bays, small objects, and small holes.
- It is less biased than close-open or open-close filtering techniques.
- Unlike a median filter, it selects the pixel value with highest frequency and assigns it to the output pixel. (Chidumayo et al., 1999)

- Subtracting filtered images (pixel-by-pixel) in order to highlight areas of change.

- Applying a majority filter to the difference image to eliminate small areas (since changes/no-changes are expected to be locally spread – i.e., they are not isolated).

The workflow of the proposed method of change detection is illustrated in Figure 5.2.

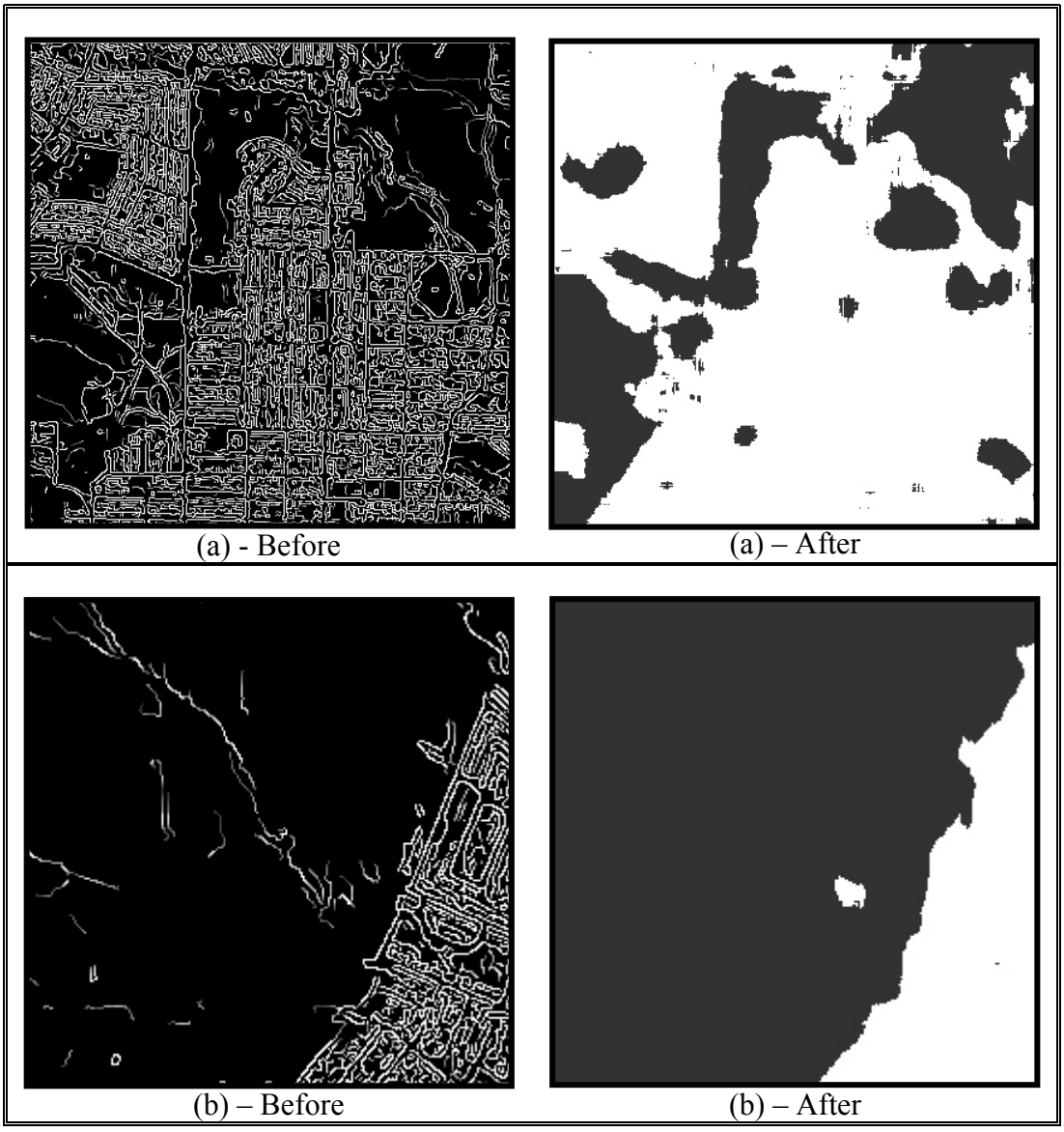



**Figure 5.1:** Majority Filter: (a) Filling Gaps among Dense Edges (b) Removing Isolated Edges

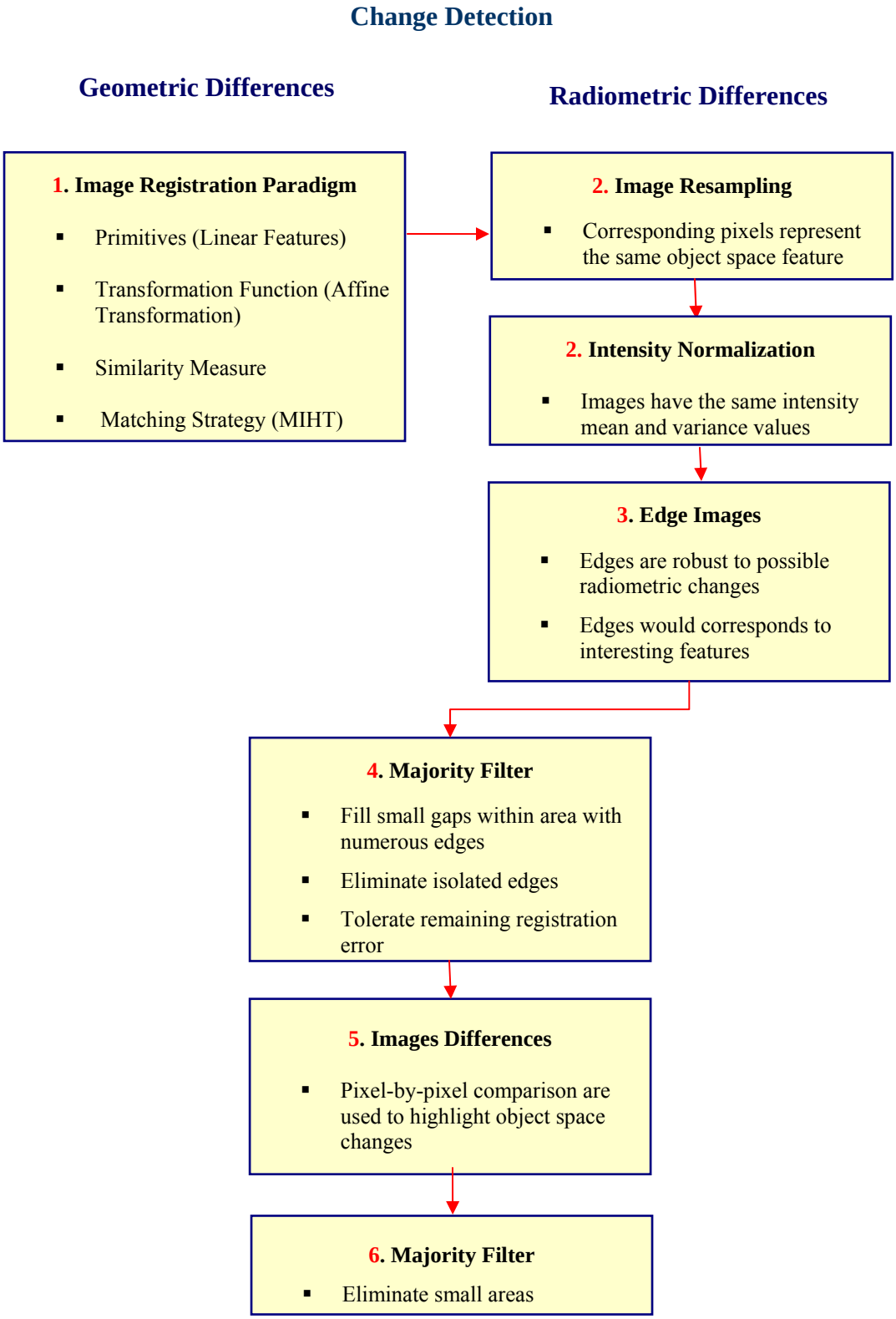



**Figure 5.2:** The Workflow of the Proposed Method of Change Detection

# CHAPTER 6

# EXPERIMENTS AND RESULTS

## 6.1 Introduction

This chapter deals with real data used to conduct a series of experiments. The purpose is to demonstrate the robustness and feasibility of the proposed algorithm as it is applied for image registration and change detection purposes. Section 6.2 presents the experiments conducted for image registration purposes. It starts with data description, then provides details of the conducted experiments and ends with the analysis of results. The resulting resampled images are then used as an input for change detection experimentation (Section 6.3).

Experiments were conducted using real data and primarily focused on achieving the following objectives:

- To compare the performance of points versus straight lines as the primitives of choice for the image registration process. The criteria for comparison are based on the ability to identify the primitive in multi-resolution satellite imagery as well as the accuracy of the image registration results as obtained from a least squares adjustment procedure.
- To analyze the validity of various transformation functions (2-D similarity, affine and projective) for representing the mathematical relationship between scenes to be registered.

- To evaluate the performance of MIHT for establishing the correspondence between the digitized primitives and simultaneously solving for the parameters involved in the registration transformation function.
- To verify the role of accurate image registration as an essential prerequisite for reliable and accurate change detection.
- To explore the use of edges as the bases of a change detection methodology.

## 6.2 Image Registration Experiments

To illustrate the feasibility and robustness of the suggested registration process, experiments were conducted using two real datasets. The first dataset covers the city of Daegon, Korea while the second covers the city of Calgary, Canada.

The Daegon dataset is composed of different satellite scenes, namely, 1500 rows × 1500 columns LANDSAT scene (15m); 1500 rows × 1500 columns SPOT scene (10m); 1500 rows × 1500 columns KOMPSAT scene (6m); and 6000 rows × 6000 columns IKONOS stereo-pair (1m). Figure 2.2 shows sample image patches. These scenes were captured at different times (multi-temporal) and exhibit significantly varying geometric and radiometric properties.

First, the parameters of 2-D similarity and affine registration transformation functions were estimated with the use of thirty-six well distributed tie points, manually identified in the scenes. The selection of common points in the various scenes proved to be a difficult and time-consuming task. The variance component ($\hat{\sigma}_o^2$ ) derived from the least squares adjustment procedure summarizes the quality of fit between the involved primitives in the

registration process. A smaller variance component indicates a better fit between the registration primitives. In analyzing the results in Table 6.1, one can see that the estimated variance component improved with the use of an affine transformation when compared to that derived for the 2-D similarity transformation. A projective transformation with eight parameters was tested as well. It was noticed that the estimated variance component did not improve over the six-parameter affine transformation; this implies the sufficiency of the affine transformation to represent the mathematical relationship between the involved scenes.

Considering the estimated variance component resulting from the registration of the two IKONOS scenes using a 2-D similarity transformation ($105.6437^{\wedge 2}$ $pixel^{\wedge 2}$), it can be concluded that such a transformation function is not a valid one. This can be attributed to the large scale associated with IKONOS scenes. Moreover, 2-D similarity transformation assumes that the image is parallel to the object space that is far from being true for IKONOS scenes. On the other hand, using an affine transformation resulted in a much more reasonable variance component ($9.8179^{\wedge 2}$ $pixel^{\wedge 2}$), and this result signified the validity of the affine transformation.

Note that the translation in x and y directions are represented by $a_o$ and $b_o$ parameters, respectively, for both 2-D similarity and affine transformations. These values should be the same regardless of the implemented transformation function. As expected, comparing the values of $a_o$ and $b_o$ between 2-D similarity and affine transformations in Table 6.1 shows that such values are very close for all images except for IKONOS / IKONOS scenes. This difference for IKONOS attributed to the invalidity of 2-D as the registration transformation function.

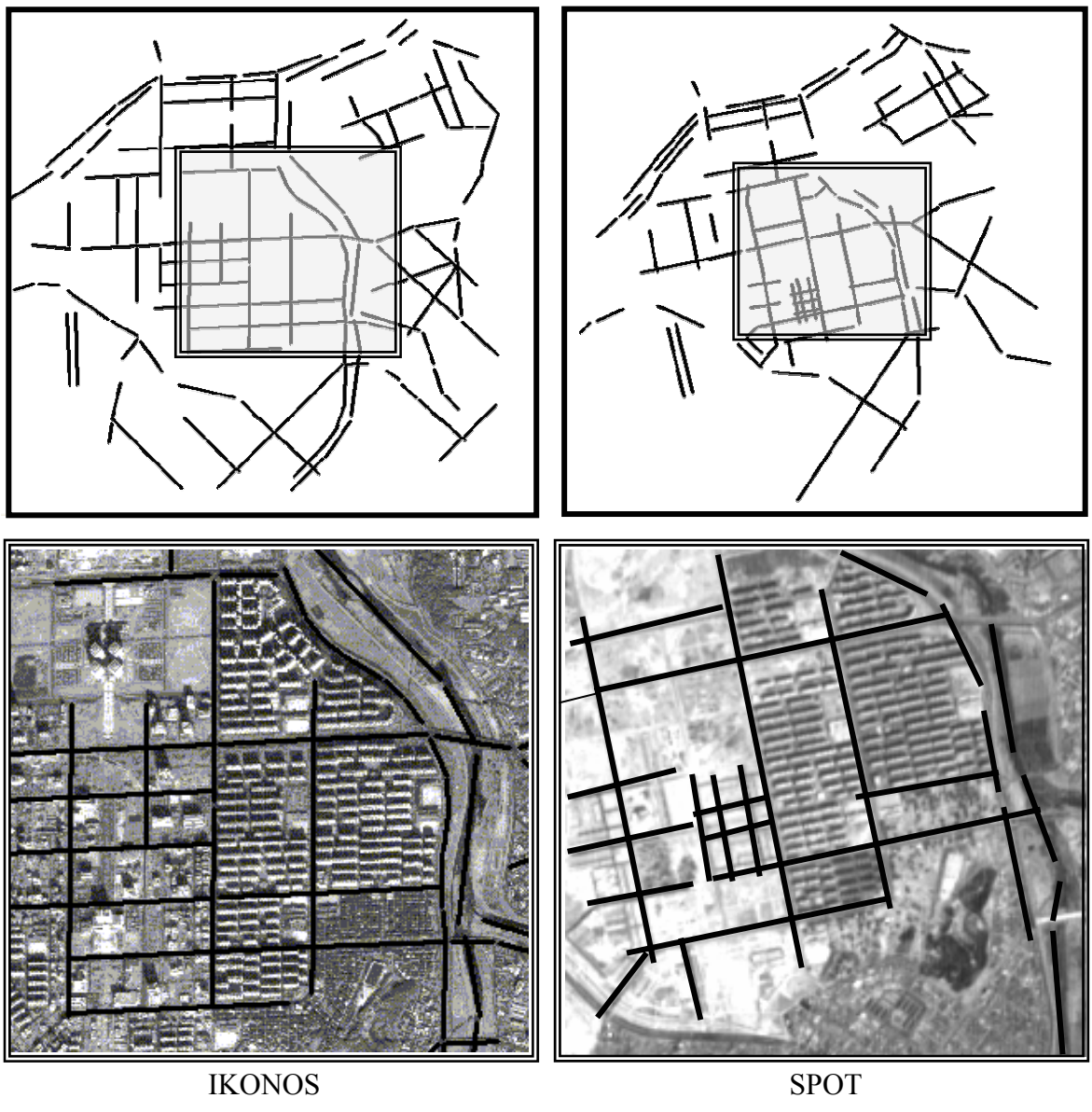
**Table 6.1:** Transformation Parameters Based on Manual Point Measurements - Daegon

| 2-D Similarity | IKONOS/IKONOS | IKONOS/KOMPSAT | IKONOS/SPOT | IKONOS/LAND |
|---|---|---|---|---|
| $\hat{\sigma}_o^2$ (Pixel ^2) | 105.6437^2 | 4.6154^2 | 7.6691^2 | 7.4872^2 |
| $a_0$ (Pixel) | 56.81489 | -99.65358 | 19.30487 | 0.13623 |
| $b_0$ (Pixel) | -18.69259 | -26.27758 | -6.58108 | -9.58130 |
| $a_1$ | 1.029469 | 0.013085 | 0.08846 | 0.03292 |
| $b_1$ | 0.071629 | 0.03187 | -0.01589 | -0.00472 |
| Affine | IKONOS/IKONOS | IKONOS/KOMPSAT | IKONOS/SPOT | IKONOS/LAND |
| $\hat{\sigma}_o^2$ (Pixel ^2) | 9.8179^2 | 2.2249^2 | 6.6021^2 | 6.5063^2 |
| $a_0$ (pixel) | 72.48928 | -97.42270 | -19.59451 | 0.04353 |
| $a_1$ | 1.051263 | 0.12707 | 0.08756 | 0.03051 |
| $a_2$ | -0.001246 | -0.03174 | 0.018210 | 0.00319 |
| $b_0$ (pixel) | -2.419632 | -25.58517 | -6.49936 | -9.85226 |
| $b_1$ | 0.140353 | 0.03153 | -0.01341 | -0.00545 |
| $b_2$ | 1.005484 | 0.13352 | 0.09020 | 0.03521 |

Afterwards, straight-line segments can be manually digitized or automatically extracted in the available scenes. Manual digitization was adopted in this research since the main objective is focused on image-to-image registration through the use of straight lines, not the extraction method. Automatic extraction of straight lines was beyond the objective of this study and will be investigated in future work.

As an example, Figure 6.1 shows the digitized segments in IKONOS and SPOT scenes. In this figure, one can see that there is no complete (i.e., one-to-one) correspondence between the digitized primitives in the input and reference images. The digitized segments are then incorporated in the MIHT strategy to automatically determine the correspondence between conjugate line segments as well as the parameters involved in

the registration transformation function. The estimated registration transformation parameters as well as the corresponding variance component for all the datasets are listed in Table 6.2.



IKONOS SPOT

**Figure 6.1:** Digitized Linear Features in IKONOS and SPOT Scenes

**Table 6.2:** Transformation Parameters Based on Automatically Matched Linear Features Using MIHT - Daegon

| 2-D Similarity | IKONOS/IKONOS | IKONOS/KOMPSAT | IKONOS/SPOT | IKONOS/LAND |
|---|---|---|---|---|
| $\hat{\sigma}_o^2$ (Pixel^2) | No Conversion | 4.2431^2 | 4.2587^2 | 0.8947^2 |
| $a_0$ (pixel) | | -103.94052 | -19.69236 | 2.81575 |
| $b_0$ (pixel) | | -28.15586 | -8.77077 | -16.96265 |
| $a_1$ | | 0.13150 | 0.08704 | 0.02985 |
| $b_1$ | | 0.03197 | -0.01583 | -0.00435 |
| Affine | IKONOS/IKONOS | IKONOS/KOMPSAT | IKONOS/SPOT | IKONOS/LAND |
| $\hat{\sigma}_o^2$ (Pixel^2) | 9.7022^2 | 1.3567^2 | 1.1634^2 | 0.7193^2 |
| $a_0$ (pixel) | 70.17578 | -97.95137 | -18.87100 | 2.20314 |
| $a_1$ | 1.05151 | 0.12695 | 0.08738 | 0.02924 |
| $a_2$ | -0.00037 | -0.03193 | 0.01905 | 0.00510 |
| $b_0$ (pixel) | -22.33391 | -27.23188 | -8.24337 | -16.94389 |
| $b_1$ | 0.14591 | 0.03196 | -0.01358 | -0.003795 |
| $b_2$ | 1.00904 | 0.13332 | 0.08881 | 0.029800 |

On the bases of results shown in Table 6.2, the following observations and can be made:

- As with the results from the point datasets, the affine transformation produced better results than the 2-D similarity transformation. This shows the validity of the affine transformation as the registration transformation function relating the scenes under consideration. As mentioned before, the 2-D similarity transformation does not constitute a proper registration transformation function between the IKONOS scenes. Therefore as expected, the MIHT procedure did not converge for this dataset.
- Comparing the results in Tables 6.1 and 6.2, one can see that utilizing linear features led to a better fit between the scenes than the fit derived through the use of point features. This should be expected, since identifying linear features in multi-resolution

imagery is much more reliable and accurate than identifying distinct points. As mentioned earlier, the affine transformation is valid when relatively flat terrain is assumed. In this context, linear features are advantageous since they restrict the selected primitives along relatively flat terrain as represented by the road network. This might not be the case for point primitives that might have significant relief distortions (e.g., simultaneous considerations of points along the terrain as well as high-rise buildings).

- Observing the estimated shift components among the registered scenes ($a_0$, $b_0$), one can see that the proposed strategy successfully converged without the need for approximate registration of these scenes.

Figure 6.2 depicts established correspondences between the digitized primitives in the IKONOS and SPOT. The estimated transformation parameters are then used to resample the reference image to the coordinate system associated with the input image. Figure 6.3 shows a mosaic image derived by combining IKONOS and SPOT scenes (where every other square patch in the reference image has been replaced by the corresponding resampled patch in the input image). Features (e.g. roads, rivers, and buildings) in the derived mosaic accurately fit each other. (A smooth transition can be observed along the features within the resampled patches-solid circles). This proves the validity of the estimated parameters of the transformation function relating these scenes. However, one can also note that there are some discontinuities along the boundaries between some of the resampled patches (highlighted by dotted circles).These discontinuities are attributable to physical changes in the object space between the epochs of capture of the involved scenes. (The SPOT scene was captured a few years earlier than the IKONOS.)

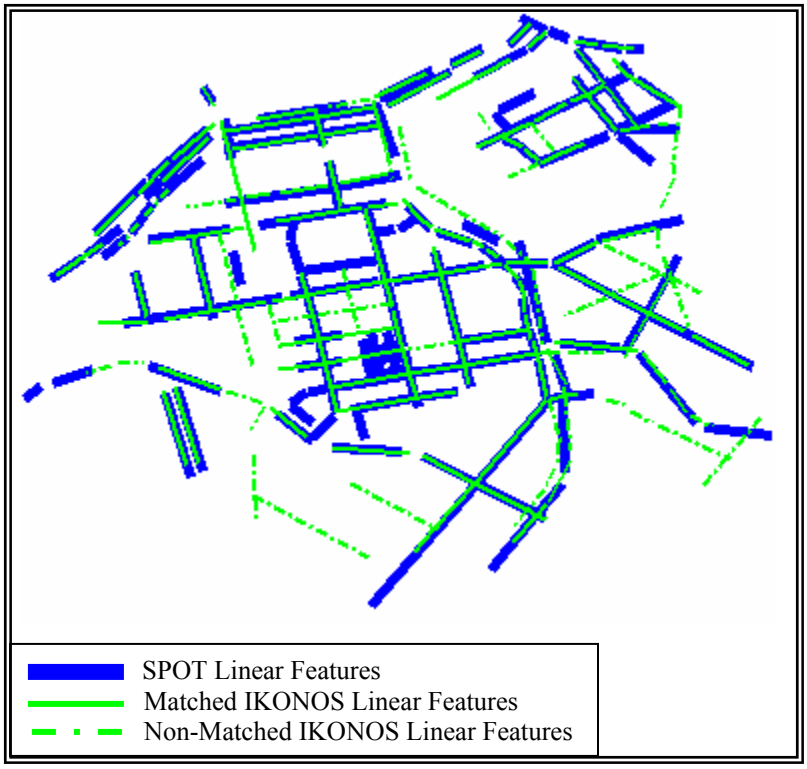



**Figure 6.2:** Established Correspondences between IKONOS and SPOT Primitives

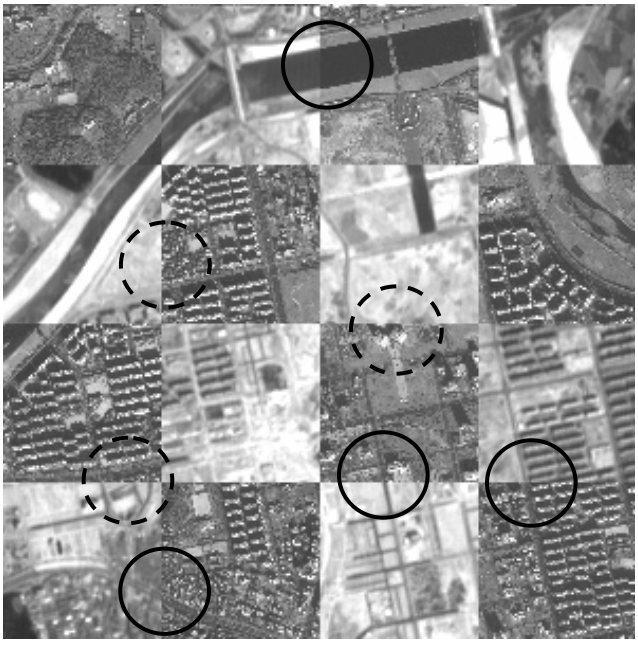


**Figure 6.3:** IKONOS-SPOT Mosaic with Highlighted Continuities (Solid Circles) and Highlighted Discontinuities (Dotted Circles) Resulting from Physical Changes in the Object Space**.**

The previous discussion and analysis are also valid for all the involved images in this dataset. For example, Figure 6.4 shows the established correspondences between IKONOS & KOMPSAT primitives, while Figure 6.5 shows the mosaic image derived by combining IKONOS and KOMPSAT scenes.

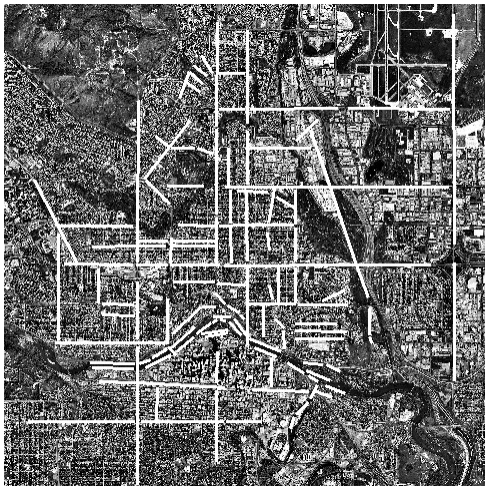

**Figure 6.4:** Established Correspondences between IKONOS and KOMPSAT Primitives.

**Figure 6.5:** IKONOS-KOMPSAT Mosaic

The second set of experiments has been conducted using multi-source, multi-resolution, and multi-temporal imagery over the city of Calgary, Canada where aerial and satellite images are involved. The experiments incorporated 1374 rows × 1274 columns aerial photo (5.0m) captured in 1956, 1374 rows × 1274 columns aerial photo (3.5m) captured in 1972, 2000 rows × 2000 columns ortho-image (5.0m) created from an aerial image captured in 1999, 500 rows × 500 columns LANDSAT image captured in 2000, and 300 rows × 300 columns LANDSAT image (30m) captured in 2001, Figure 6.6.

Aerial 1956 (5m)

Aerial 1972 (3.5m)



Ortho-photo 1999 (5m)

LANDSAT 2000 (15m)

LANDSAT 2001 (30m)

**Figure 6.6**: Digitized Linear Features in the Calgary Dataset

The sequence of experiments applied to Daegon's dataset was repeated for Calgary's. Results showed the superiority of straight lines features over points and the suitability of affine transformation over 2-D similarity. These results were compatible with those of the previous experiments, and verified the robustness and flexibility of the suggested approach to handle multi-source, multi-resolution, and multi-temporal imagery. The estimated registration transformation parameters as well as the corresponding variance components using points and straight lines primitives are listed in Tables 6.3 and 6.4, respectively.

**Table 6.3:** Transformation Parameters Based on Manual Point Measurements - Calgary

| 2-D Similarity | ORTHO 1999 AERIAL 1956 | ORTHO 1999 AERIAL 19 72 | ORTHO 1999 LANDSAT 2000 | ORTHO 1999 LANDSAT 2001 |
|---|---|---|---|---|
| $\hat{\sigma}_o^2$ (Pixel$^{\wedge 2}$) | 4.3580$^{\wedge 2}$ | 3.1635$^{\wedge 2}$ | 2.2987$^{\wedge 2}$ | 1.847$^{\wedge 2}$ |
| $a_0$ (pixel) | 95.0619 | 65.9943 | 91.4917 | 53.1679 |
| $b_0$ (pixel) | -105.2252 | 270.2409 | 75.4917 | 29.5500 |
| $a_1$ | 0.9164 | 1.3008 | 0.3369 | 0.1586 |
| $b_1$ | -0.0185 | 0.0546 | 0.0123 | -0.0532 |
| Affine | ORTHO 1999 AERIAL 1956 | ORTHO 1999 AERIAL 19 72 | ORTHO 1999 LANDSAT 2000 | ORTHO 1999 LANDSAT 2001 |
| $\hat{\sigma}_o^2$ (Pixel$^{\wedge 2}$) | 4.1231$^{\wedge 2}$ | 2.6313$^{\wedge 2}$ | 2.4148$^{\wedge 2}$ | 1.8650$^{\wedge 2}$ |
| $a_0$ (pixel) | 93.8898 | 64.9483 | 91.8455 | 53.3285 |
| $a_1$ | 0.9120 | 1.2945 | 0.3361 | 0.1577 |
| $a_2$ | 0.0162 | -0.0573 | -0.0107 | 0.05505 |
| $b_0$ (pixel) | -105.5540 | 270.9716 | 77.7104 | 30.8239 |
| $b_1$ | -0.0216 | 0.05179 | 0.0138 | -0.0519 |
| $b_2$ | 0.9196 | 1.3056 | 0.3402 | 0.1624 |

**Table 6.4:** Transformation Parameters Based on Automatically Matched Linear Features Using MIHT - Calgary.

| 2-D Similarity | ORTHO 1999 AERIAL 1956 | ORTHO 1999 AERIAL 19 72 | ORTHO 1999 LANDSAT 2000 | ORTHO 1999 LANDSAT 2001 |
|---|---|---|---|---|
| $\hat{\sigma}_o^2$ (Pixel^2) | 2.2298^2 | 2.7774^2 | 1.7599^2 | 0.8977^2 |
| $a_0$ (pixel) | 94.0756 | 65.4424 | 87.9770 | 53.1336 |
| $b_0$ (pixel) | -106.6365 | 269.8632 | 75.8580 | 30.9736 |
| $a_1$ | 0.9195 | 1.3041 | 0.3341 | 0.1595 |
| $b_1$ | -0.0210 | 0.0562 | 0.0132 | -0.0507 |
| Affine | ORTHO 1999 AERIAL 1956 | ORTHO 1999 AERIAL 19 72 | ORTHO 1999 LANDSAT 2000 | ORTHO 1999 LANDSAT 2001 |
| $\hat{\sigma}_o^2$ (Pixel^2) | 2.1785^2 | 2.0657^2 | 1.6761^2 | 0.8522^2 |
| $a_0$ (pixel) | 94.0991 | 64.6135 | 89.5263 | 52.7716 |
| $a_1$ | 0.9181 | 1.3018 | 0.3355 | 0.1589 |
| $a_2$ | 0.0181 | -0.0592 | -0.0105 | 0.0500 |
| $b_0$ (pixel) | -106.6896 | 270.2862 | 75.7333 | 31.3885 |
| $b_1$ | -0.0229 | 0.0542 | 0.0142 | -0.0506 |
| $b_2$ | 0.9204 | 1.3053 | 0.3334 | 0.1612 |

Figure 6.7 shows the derived correspondences between the digitized primitives in the Ortho-photo 1999 and aerial 1956 images while Figure 6.8 shows the established correspondences between the digitized primitives in the Ortho-photo 1999 and LANDSAT 2000 scenes. Even in the presence of small overlap between the Ortho-photo 1999 and the LANDSAT 2000, the results of image registration are accurate enough Figure 6.9 (b).


**Figure 6.7:** Established Correspondences between Ortho-photo 1999 and Aerial 1956 Primitives.

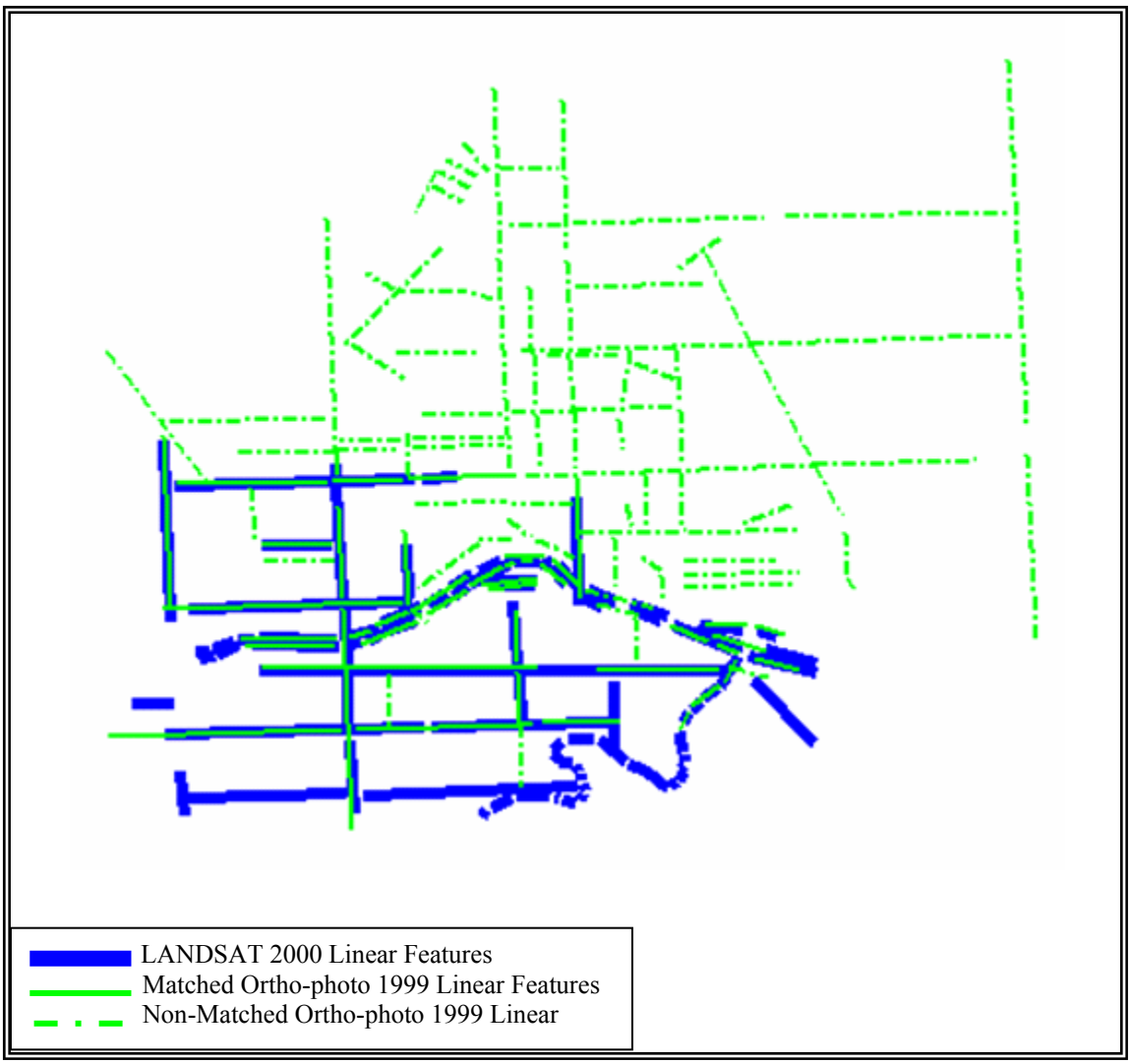



**Figure 6.8:** Established Correspondences between Ortho-photo 1999 and LANDSAT 2000 Primitives.

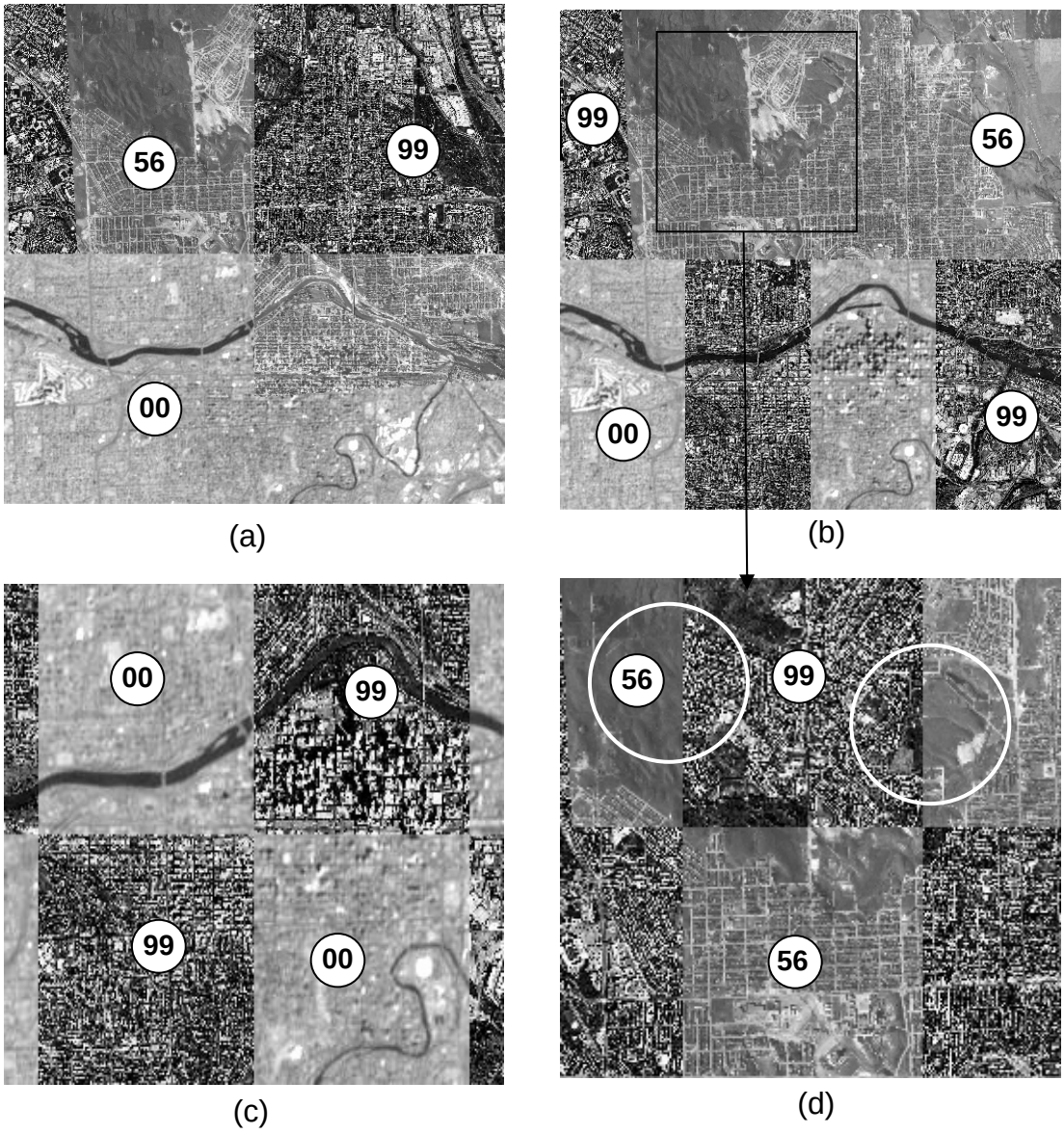



**Figure 6.9:** Multi-Image Mosaic for Calgary Dataset

To verify these results, the estimated transformation parameters are used to resample the reference image into the coordinate system associated with the input image. Figure 6.9 shows mosaic images derived by combining LANDSAT 2000, Ortho-photo 1999, and aerial 1956. A closer look at this figure reveals the following facts:

- Due to the limited area covered by the LANDSAT 2000 scene, Figure 6.6, image completion concept has been applied to obtain full coverage for the city of Calgary. Aerial 1956 and Ortho-photo 1999 were used to achieve such a task Figure 6.9(a, b). Multi-image integration has been accomplished. This important process is needed to cope with the large diversity of contemporary available images.

- In Figure 6.9(c), every other square patch in the reference image (Ortho-photo, 1999) has been replaced by the corresponding resampled patch in the input image (LANDSAT, 2000). Features in the derived mosaic accurately fit each other. This proves the validity of the estimated parameters of the transformation function relating these scenes.

- Discontinuities appear along the boundaries between some of the resampled patches in Figure 6.9(d) (highlighted by hollow circles). These discontinuities are attributed to real changes in the object space between the epochs of capture of the involved scenes. (The aerial image was captured forty-three years earlier than the Ortho-photo scene). This is significant for change detection applications, since accurate image registration is a prerequisite for accurate and reliable change detection output.

## 6.3 Change Detection Experiments

Once the transformation function has been established between the images, the input image can be resampled into the reference frame associated with the reference image. As explained in Chapter 5, the resampling is followed by the application of the Canny edge detection technique and majority filter to both images. Then, the resulting images are

subtracted to produce a change image, which is enhanced by re-applying the majority filter.

An experiment was conducted using aerial image 1956 and Ortho-photo 1999 for the city of Calgary dataset. Figures 6.10 (a) and (b) show the input image and resampled reference image, respectively, with the same number of rows and columns. Figures 6.10 (c) and (d) show the derived edges after applying the Canny edge detector for aerial 1956 image and resampled Ortho-photo image respectively. A closer look to the edge images shows that linear features (rivers, roads, and buildings), which represent the main source of changes in urban areas, have been successfully detected in each image.

Afterward, a majority filter was applied on the resulting edge images in order to fill small gaps within an area with numerous edges as well as to eliminate isolated edges. As a result of applying the majority filter, edge cells were densified and areas with interesting features were highlighted (see Figure 6.11). Then, the filtered images were subtracted to highlight areas of change, as in Figure 6.12 (a). White areas indicate changes while black areas indicate parts with no change. Finally, a majority filter was applied to the difference image to eliminate small areas which do not reflect a real object space change, Figure 6.12 (b). Figure 6.13 shows the areas of change in the city of Calgary between 1956 and 1999.

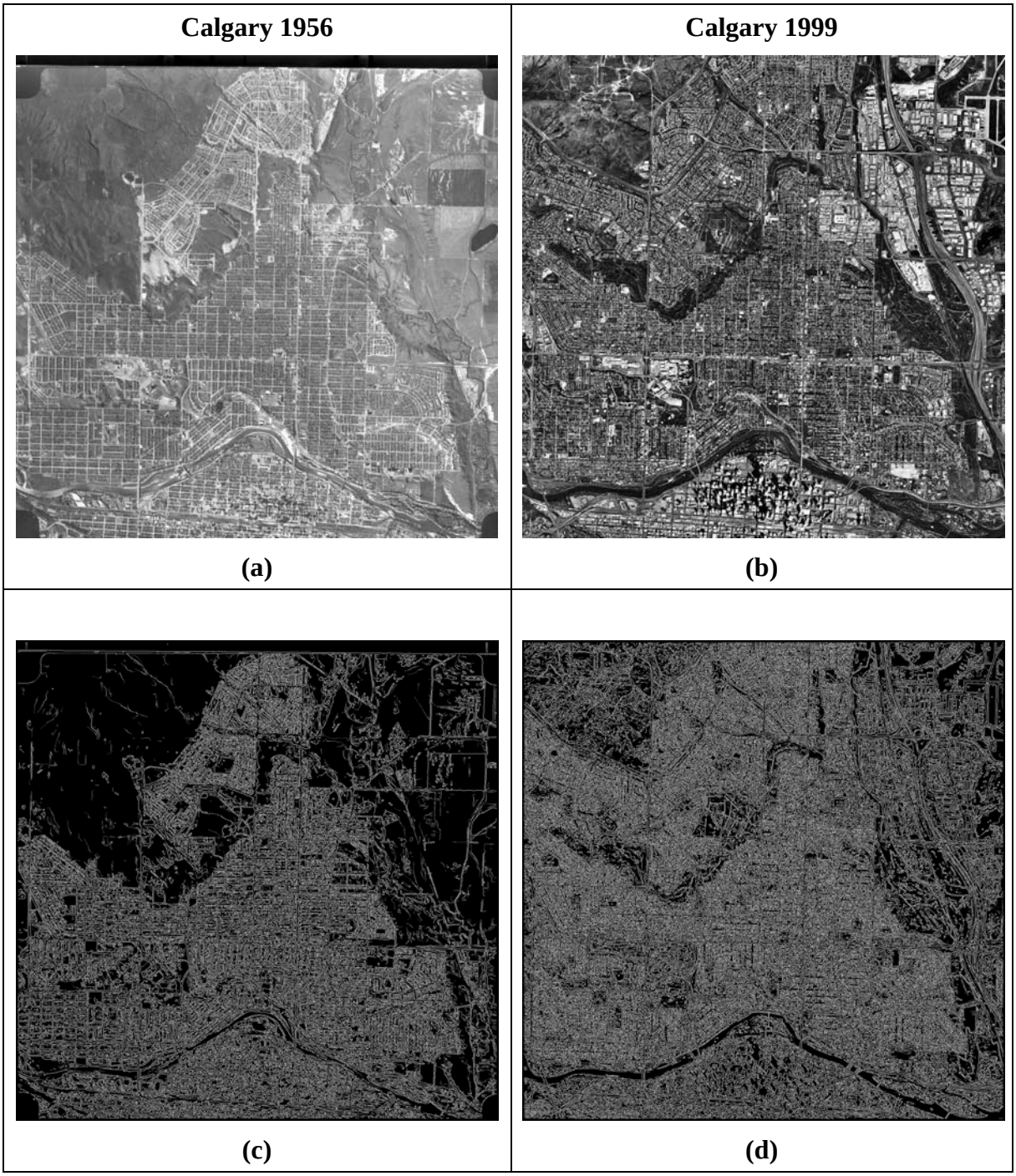



**Figure 6.10:** Resampled and Edge Images for the City of Calgary Dataset: (a) Aerial, 1956 (b) Resampled Ortho-photo, 1999 (c) Edge Image for Aerial, 1956, and (d) Edge Image for Resampled Ortho-photo, 1999.

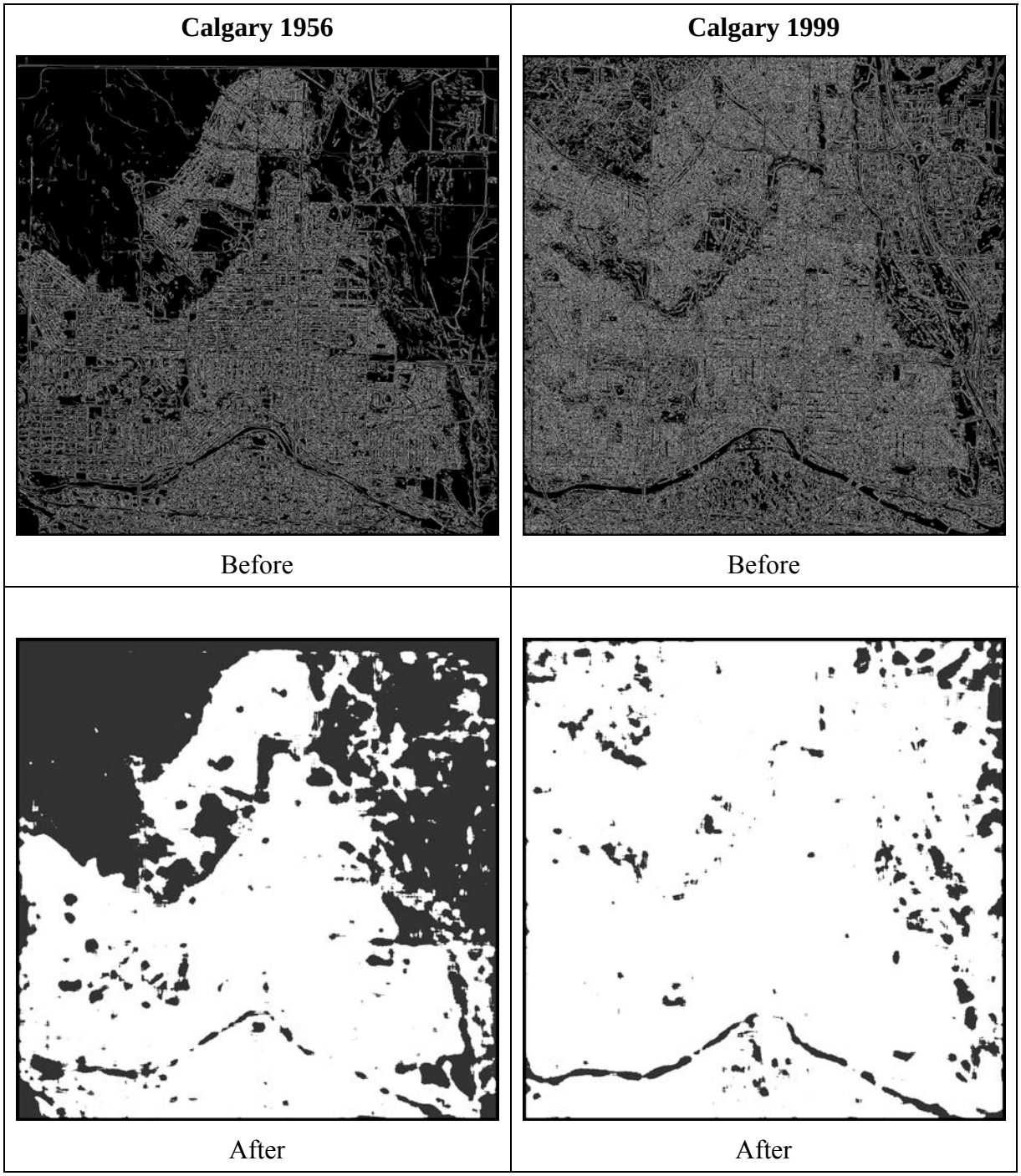



**Figure 6.11:** Edge Images Before and After Application of the Majority Filter

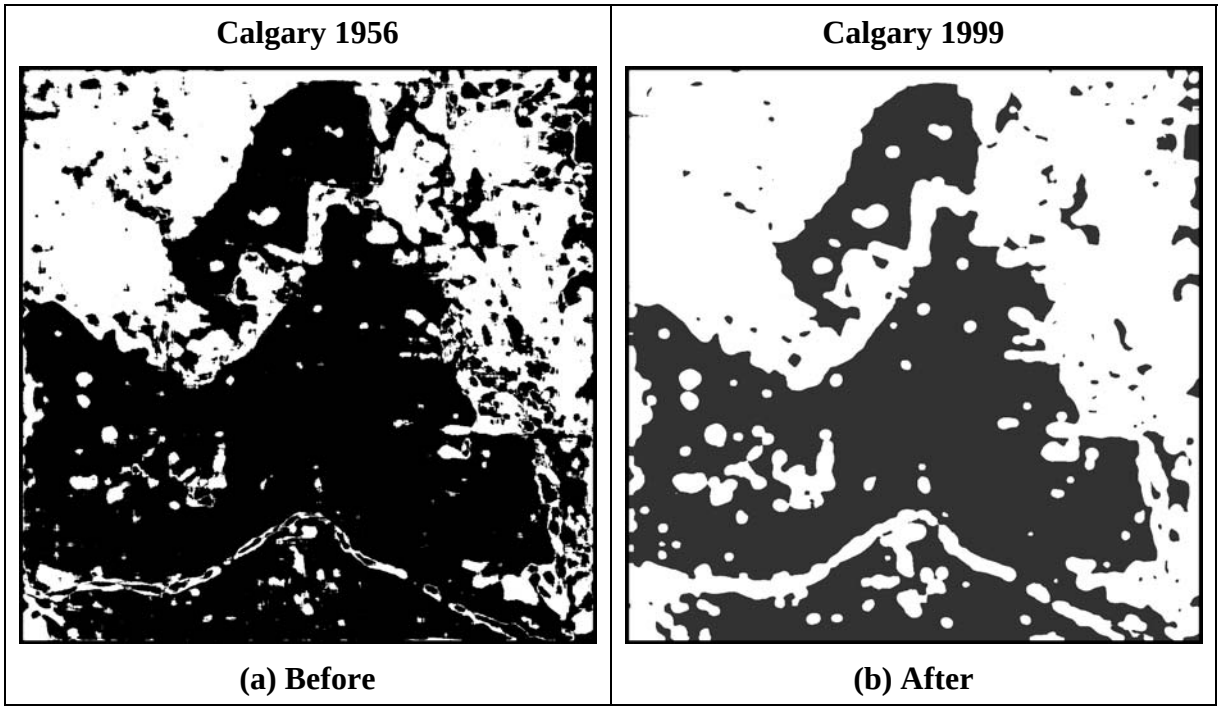



**Figure 6.12:** Difference Image Before and After Application of the Majority Filter

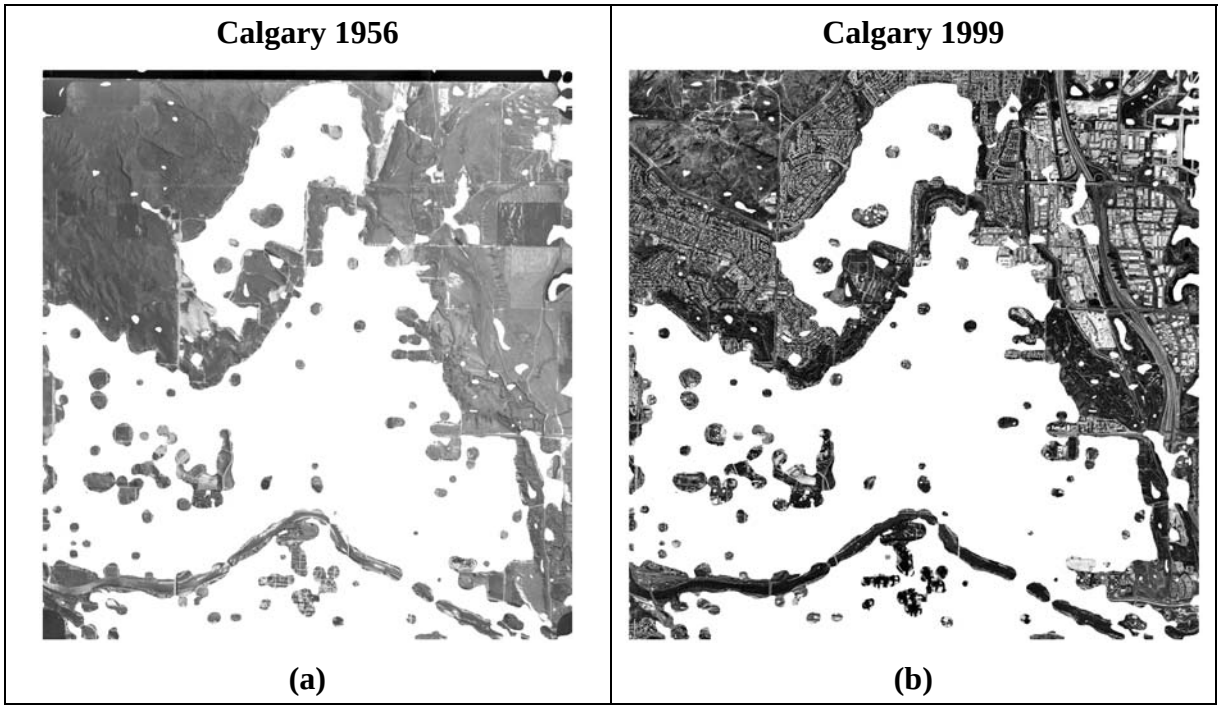



**Figure 6.13:** Areas of Change for the City of Calgary between 1956 and 1999

Simple statistics for the change image in Figure 6.14 show that there is an overall 50.6% change between the 1956 and 1999 imagery. Dividing the area into four quadrants shows that the percentages of change that occurred in the northwest, northeast, southeast and southwest parts of the image are 74.8%, 66.4%, 34.4%, and 26.8%, respectively. The sub images (b, c, d, and e) in Figure 6.14 show different types of changes that took place. Sub image 6.14 (b) shows changes as a result of an urbanization activity. (A new residential community was built). Sub image 6.14 (c) shows changes caused by trails in newly developed parks. Changes resulting from the construction of a new highway along the east side of the city are shown in sub image 6.14 (d). Finally, sub image 6.14(e) shows the changes due to shadowing effects caused by newly erected high-rise buildings in the downtown area.

In order to examine the efficiency of the suggested change detection approach compared to traditional methods, image differencing based on supervised classification approach was conducted. For this purpose, texture images were derived from the original images using PCI Geomatics software. A filter size of 25 x 25 pixels was used to derive the texture images based on homogeneity and variance of the original image. Then, the aerial 1965 image was classified, based on the original image and its derived texture image, into two main classes, urban and non-urban areas. The same process was carried out for Ortho-photo 1999 image. Finally, image differencing was applied to highlight the areas of change. The results indicated that small isolated regions which do not belong to real change still appear in the change image. However, the overall results can still be considered relatively compatible with the results of the approach suggested in this research. The disadvantages of image differencing based on supervised classification

mentioned in the literature were clearly observed. These drawbacks (including the great computational and labelling efforts required for classification purposes, The need for at least two bands for each image, and the thresholding problem) were avoided by using edges as a base for change detection. Figure 6.15 shows the classification for both images as well as the change detection results based on the classified images.

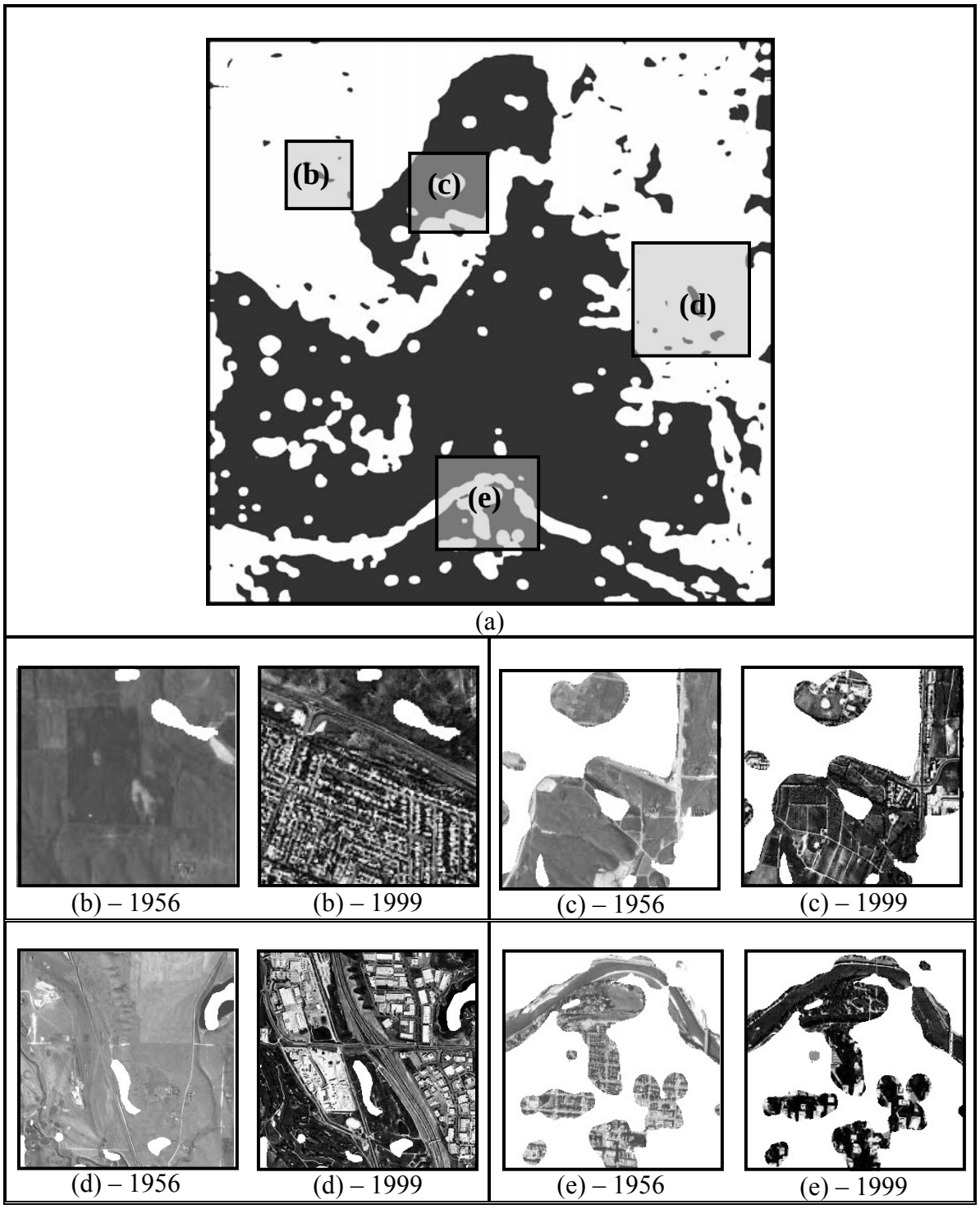



**Figure 6.14:** Change Detection Image (a), White Pixels Represent Changes. Sub-Figures b, c, d, and e have been Cropped and Closely Examined.

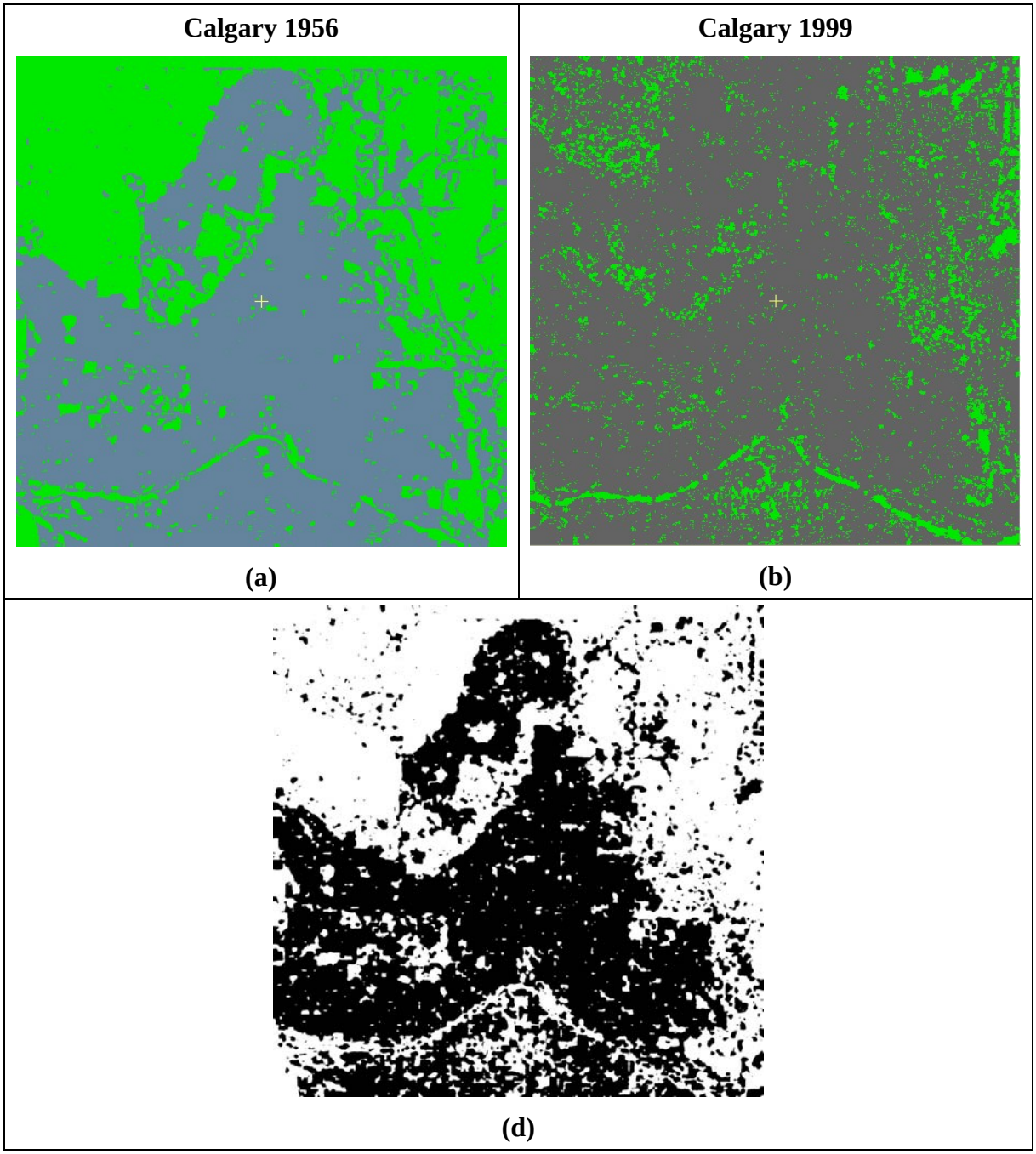



**Figure 6.15:** Change Detection Based on Supervised Classification: (a) Classification of Aerial 1956 (b) Classification of Ortho-photo 1999 and (c) Difference Image with White Pixels Representing Changes.

# CHAPTER 7

# CONCLUSION AND FUTURE WORK

## 7.1 Conclusion

With the flux of high resolution scenes captured by space-borne platforms (e.g., LANDSAT-7, IKONOS, QUICKBIRD, ORBVIEW, EROS-A1, KOMPSAT-I, and SPOT-5), there is an increasing need for a robust registration technique that can tolerate varying geometric resolutions among the available scenes. This research has comprehensively addressed the key issues of an efficient semi-automatic registration methodology that can handle such scenes. First, straight-line segments have been chosen as the registration primitives. The rationale for selection is that such primitives can be reliably identified in multi-resolution scenes. Then, the registration transformation function is analyzed to determine the mathematical relationship between the scenes to be registered. It has been established that affine transformation can be used as the registration transformation function for scenes captured by high altitude imaging satellite systems with narrow angular field of view. Moreover, 2-D similarity transformation can be used as another alternative for some applications with less demanding accuracy requirements. Afterwards, the geometric attributes of conjugate primitives are manipulated to derive a similarity measure describing the necessary constraints for the coincidence of these primitives after establishing the registration procedure. It is important to note that the similarity measure has been developed while considering the

fact that the end points of conjugate line segments are not identical. Finally, the primitives and the similarity measure are manipulated in a MIHT procedure to sequentially solve for the parameters involved in the registration transformation function while establishing the correspondence between conjugate primitives. The MIHT procedure proved to be helpful in verifying the validity of the registration transformation function, since it will only converge if the transformation function is valid. Within the MIHT, an optimum sequence with the use of a 2-D similarity and affine transformation functions has been derived through the analysis of deviations from the similarity measure constraints associated with line segments with different orientations at various regions within the imagery as a result of incremental changes in the transformation parameters.

Experimental results showed the feasibility and the robustness of the suggested approach that could tolerate possible discrepancies between the imagery due to varying sensor operational principles as well as changes in the object space without the need for approximate registration of the involved imagery. Moreover, the results proved the superiority of straight-line segments over distinct points. This should be expected since linear features can be identified more accurately than distinct points. In addition, the results verified the fact that affine transformation yields better registration when compared with 2-D similarity transformation. The proposed technique could be used to robustly and simultaneously estimate the parameters of the registration transformation function as well as the feature-to-feature correspondence between multi-temporal, multi-resolution, and multi-source satellite imagery.

The importance of accurate image registration as an essential prerequisite for reliable and accurate change detection has been established. To avoid the effect of possible radiometric differences between the registered images, due to different atmospheric conditions, noise, and different spectral properties, the change detection is based on derived edge images. The use of edge images is attractive; since it would lead to an effective detection of urbanization activities since they would appear as a dense distribution of edge cells. Also, a majority filter has been applied to compensate for small registration errors as well as eliminating small gaps and isolated edges. The images are then subtracted to produce a change image, which could be enhanced through the application of a majority filter to remove small regions. The change detection results are found to be consistent with these visually identified.

## 7.2 Recommendations for Future Work

Future research will focus on automatic extraction of registration primitives from input imagery as well as the utilization of free-form linear features, represented as a sequence of straight line segments (polylines). In addition, the impact of various generalization levels of these primitives in terms of the processing time and the quality of the registration outcome should be investigated. Furthermore, research should be conducted to evaluate the limits for the validity of the affine transformation as the registration transformation function. The proposed strategy can be used to establish the registration of satellite scenes with vector data in existing GIS databases for change detection and for updating applications where the nature of detected changes (e.g., new residential community, new roads, etc.) is investigated as well.

# REFERENCES

Abbasi-Dezfouli, M., and T. Freeman, 1994. Stereo-image registration based on uniform patches, *International Archives of Photogrammetry and Remote Sensing*, 30(3): 1-8.

Abdel-Aziz, Y., and H. Karara, 1971. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry, *American Society of Photogrammetry, Symposium on Close-Range Photogrammetry*, Urbana, Illinois 433, pp. 1-18.

Ackermann, F., 1995. Sensor and data integration-the new challenge, *ISPRS Journal of Photogrammetry and Remote Sensing, Work-shop "integrated Sensor Orientation",* Barcelona, pp. 2-10.

Agouris, P., G. Mountrakis, and A. Stefanidis, 2000, Automated spatiotemporal change detection in digital aerial imagery, *SPIE Proceedings*, 4054: 2-12.

Banner, A., and T. Lynham, 1981. Multitemporal analysis of LANDSAT data for forest cut over mapping- a trial of two procedures, *Proceeding of the 7$^{th}$ Canadian Symposium on Remote Sensing,* Winnipeg, pp. 233-240.

Boardman, D., I. Dowman, A. Chamberlain, D. Fritsch, and W. Newton, 1996. An automated image registration system for SPOT data, *International Archives of Photogrammetry and Remote Sensing*, 31(4): 128–133.

Brown, L., 1992. A survey of image registration techniques, *ACM Computing Surveys* 24(4): 325-376.

Bruzzone, L., and D. Prieto, 2000. Automatic analysis of the difference image for unsupervised change detection, *IEEE Transactions on Geoscience and Remote Sensing*, 38(3): 1171-1182.

Bruzzone, L., and S. Serpico, 1997. An iterative technique for the detection of land-cover transitions in multitemporal remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing*, 35(4): 858-867.

Canny, J., 1986. A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):679-698.

Cavallaro, A., and E. Touradj, 2001. Change detection based on color edges, *Proceeding of IEEE International Symposium on Circuits and Systems* (ISCAS-2001).

Chidumayo, E., O. Kalumiana, H. Ntalasha, and I.Masialeti, 1999. Individual partner annual report: Zambia, UNZA Zambia annual report.

Cho, S., 2000. Digital change detection by post-classification comparison of multtemporal remotely-sensed data, *Journal of the Korean Society of Remote Sensing*,16(4): 367-373.

Coppin, P., and M. Bauer, 1994. Processing of mutitemporal LANDSAT TM imagery to optimise exatraction of forest cover change features, *IEEE Transactions on Geoscience and Remote Sensing*, (32) 4: 918-927.

Dare, P., and I. Dowman, 2001. An improved model for automatic feature-based registration of SAR and SPOT images, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56 : 13-28.

Dowman, I., 1998. Automated procedures for integration of satellite images and map data for change detection: The archangel project, *GIS-between visions and applications*. IAPRS, 32(4): 162-169.

Dowman, I., D. Boardman and W. Newton, 1996. A system for automatic registration of two SPOT images, *Proceedings of 22nd Annual Conference of the Remote Sensing Society*, Durham. 704: 169-175.

Estes, J., 1992. Technology and policy issues impact global monitoring, *GIS world*, 5(10): 52-55

Flusser, J., 1992. An adaptive method for image registration, *Pattern Recognition*, 25(1): 45-54.

Flusser, J., and T. Suk, 1994. A moment based approach to registration of images with affine geometric distortion, *IEEE Transactions on Geoscience and Remote Sensing,* 32(2): 382-387.

Fonseca, L., and B. Manjunath, 1996. Registration techniques for multisensor remotely sensed imagery, *Photogrammetric Engineering and Remote Sensing*, 62(9):1049-1056.

Fonseca, L., and M. Costa, 1997. Automatic registration of satellite images, *Proceedings. Los Alamitos, IEEE Computer Society*, pp. 219-226.

Förstner W., 1986. A feature based correspondence algorithm for image matching, *International Archives of Photogrammetry and Remote Sensing*, 26 (3): 150-166.

Förstner, W., and E. Gulch, 1987. A fast operator for detection and precise location of distinct points, corners and centers of circular features, *Proceedings of the ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data,* 2-4 June 1987, Interlaken, Switzerland, pp. 281–305.

Fraser, C., H. Hanley, and T. Yamakawa, 2001. Sub-metric geopositioning with IKONOS geo imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, workshop"high resolution mapping from space 2001", 19-21 September, Hanover, Germany, pp. 61-68

Fung T., 1990. An assisment of TM imagery for l;and-cover change detection, *IEEE Transactions on Geoscience and Remote Sensing*, 28 (4): 681-684.

Gonzalez R., and R. Woods, 1992. *Digital Image Processing*, Addison-Wesley, Reading, MA, USA, 716 p.

Goshtasby, A., 1988. Registration of images with geometric distortions, *IEEE Transactions on Geoscience and Remote Sensing*, 26(1): 60-64.

Goshtasby, A., G. Stockman and C. Page, 1986. A region based approach to digital image registration with subpixel accuracy, *IEEE Transactions on Geoscience and Remote Sensing,* 24(3): 390-399.

Habib, A., A. Asmamaw, D. Kelley, and M. May, 2000. Linear features in photogrammetry. *Report No. 450*, D*epartment of Civil and Environmental Engineering and Geodetic Science*, the Ohio State University, Columbus, Ohio, USA, 80p.

Habib, A., and D. Kelley, 2001a. Single photo resection using the modified Hough transform, *Photogrammetric Engineering and Remote Sensing*, 67(8): 909-914.

Habib, A., and D. Kelley, 2001b. Automatic relative orientation of large scale imagery over urban areas using modified iterated Hough transform, *International Journal of Photogrammetry and Remote Sensing*, 56 (2001): 29-41.

Habib, A., and M. Morgan, 2002. Epipolar image resampling from push-broom imagery: investigation and preliminary implementation, *Korean Electronics and Telecommunications Research Institute (ETRI)*, Daejeon, Korea, 107 p.

Habib, A., B. Beshah, 1998. Multi sensor aerial triangulation, *ISPRS Journal of Photogrammetry and Remote Sensing,* Commission III Symposium in Columbus, Ohio, pp. 37-41.

Habib, A., M. Morgan, and Y. Lee, 2001a. Integrating data from terrestrial mobile mapping systems and aerial imagery for change detection purposes, *Proceedings of the Third Mobile Mapping Symposium,* 3-5 January, Cairo, Egypt, unpaginated CD-ROM.

Habib, A., S. Shin, and M. Morgan, 2002a. Automatic pose estimation of imagery using free-form control linear features. ISPRS commission III symposium, photogrammetric computer vision, Graz, Austria, September 9 – 13, unpaginated CD-ROM.

Habib, A., Y. Lee, and M. Morgan, 2002b. Bundle adjustment with self-calibration using straight lines, *Photogrammetric Record*, 17(100): 635-650.

Habib, A., Y., Lee, and M. Morgan, 2001b. Surface matching and change detection using modified iterative Hough transform for robust parameter estimation, *Photogrammetric Record*, 17 (98): 303-315.

Hough, P., 1962. Methods and means for recognizing complex pattern, *US Patent* 3,069,654. 18th December.

Hsieh J., H. Liao, K. Fan, M. Ko, and Y. Hung, 1997. Image registration using a new edge-based approach, *Computer Vision and Image Understanding*, 67 (2): 112-130.

Ingram, k., E. Knapp and, J. Robinson, 1981. Change detection technique development for improved urbanized area delineation, *technical memorandum CSC/TM-81/6087, Computer Sciences Corporation*, Silver Springs, Maryland, USA.

Jenson, J., 1983. Urban/Suburban land use analysis. *Manual of Remote Sensing*, Vol 2, second addition, pp. 1571-1666.

Kraus, K., 1992. *Photogrammetry Fundamentals and standard processes*, Vol. 1. Dūummler, Bonn, Germany, 397 p.

Kubik, K., 1991. Relative and absolute orientation based on linear features, *ISPRS Journal of Photogrammetry and Remote Sensing*, 46(1): 199-204.

Li, D., S. Haigang, and X. Ping, 2002. Automatic change detection of geo-spatial data from imagery. *Mapping and Remote Sensing,* commission II, IC WG II/ IV:245-252.

Li, H., B. Manjunath, and S. Mitra, 1995. A contour-based approach to multi-sensor image registration, *IEEE Transaction on Image Processing*, 4(3): 320-334.

Li, J., and R., Narayanan, 2003. A shape-based approach to change detection of lakes using time serious remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing*, 41(11): 2466-2477.

Lillesand, T., and R. Kiefer. 2000. *Remote sensing and image interpretation*. 4$^{th}$ed. John Wiley and Sons, New York, NY, 724 p.

Mas, J., 1999. Monitoring land-cover changes: a comparison of change detection techniques, *International Journal of Remote Sensing,* (20) 1: 139-152.

Mikhail, E., and J. Bethel, 2001. *Introduction to modern photogrammetry*, J. Wiley, New York, 479 p.

Moravec, H., 1977. Toward automatic visual obstacles avoidance, *Proceedings of the 5$^{th}$ Joint Conference on Artificial Intelligence*, August, Cambridge, Massachusetts, USA. pp. 584-592.

Morgado, A., and I. Dowman, 1997. A procedure for automatic absolute orientation using aerial photographs and a map, *ISPRS Journal of Photogrammetry and Remote Sensing*, 52: 169-182.

Nelson, R., 1982. Detecting forest canopy change using LANDSAT, *NASA Technical memorandum 83918,* Goddard space flight center, Greenbelt, Maryland, USA.

Nelson, R., 1983. Detecting forest canopy change due to insect activity using LANDSAT MSS, *Photogrammetric Engineering and Remote Sensing*, 49:1303-1314

Palandro, D., S. Andrefouët, P. Dustan, and F., Karger, 2003. Change detection in coral communities using IKONOS satellite sensors imagery and historical aerial photographs. *International Journal of Remote Sensing,* (24) 1: 873-878.

Pratt, W., 1991. *Digital Image Processing*, 2nd edition, Wiley, New York, USA, 698 p.

Richards, A., 1993. *Remote Sensing Digital Image Analysis,* 2nd edition, Springer-Verlag,Berlin Heidlberg New York, 340 p.

Schenk, T., 1999. *Digital Photogrammetry*. Vol. I, Terra Science, Laurelville, 428 p.

Seedahmed, G., and L. Martucci, 2002. Automated image registration using geometrically invariant parameter space clustering (GIPSC), *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, 34(3A):318-323.

Singh, A., (1989). Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing*, 10(6):989-1003.

Tao, V., and Y. Hu, 2001. A comprehensive study for rational function model for photogrammetric processing, *Journal of Photogrammetric Engineering and Remote Sensing,* 67(12): 1347-1357.

Townshend, J., C. Justice, C. Gurny, and J. McManus, 1992. The impact of misregistration on change detection, *IEEE Transactions on Geoscience and Remote Sensing*, 30(5): 1054-1060.

# APPENDIX A

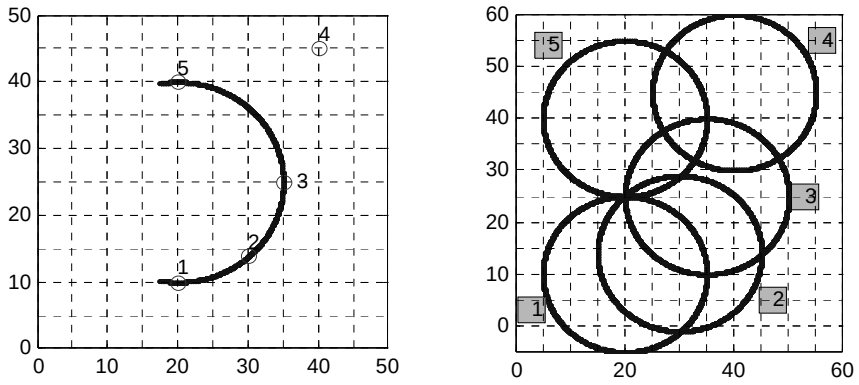
## Hough Transform

Hough introduced a method of determining parameters by way of a voting scheme [Hough 1962]. The Hough transform algorithm is a detection and segmentation technique. This technique is used if the location of a curve is not known but its shape is known as a parametric curve. Usually after automatic point extraction in digital photogrammetry, one has a list of points in image space, which are assumed to represent a certain analytical function. The Hough transform searches for the extracted points which satisfy this given function. The parameters of this function are the results of the Hough transform algorithm. The basic principle of Hough approach is to switch the roles of parameters and spatial variables. To illustrate this approach, consider the following example (Habib and Kelley, 2001a.). Suppose that we want to detect points that lie on a circle of known radius, r. A circle can be defined by Equation A1:

$$(x-u)^2+(y-v)^2-r^2=0 \tag{A1}$$

With x, y being the spatial variables and u, v the parameters (center) of the circle in the spatial domain. Now, let us introduce the parameter space, represented by the coordinate system u, v. A point with coordinate $x_i$, $y_i$ in the spatial domain corresponds to a circle in the parameter space centered at $x_i$, $y_i$. For every point in the spatial domain, there exists a circle in the parameter space, and vice versa. The intersection of circles in the parameter

space identifies centers of circles in the spatial domain. The number of intersecting circles in the parameter space is directly related to the number of points that lie on this circle, (see Figure A1). A point in the spatial domain (a) corresponds with a circle in the parameter space (b) and vice versa. The intersection of circles in the parameter space determines the center of the sought circles in the spatial domain. The intersection of four circles at u = 20, v = 25 identifies points 1,2,3 and 5 as belonging to a circle whose center c in the spatial domain is (20, 25).




(a) (b)

**Figure A1:** Illustration of Finding Circles through Data Points in Hough Transform

The Hough method is usually implemented by an accumulator array, which is an n-dimensional, discrete space, where n is the number of parameters. In this example with circles of known radii, the parameter space is two-dimensional. Each circle is discretely represented in the parameter space. To keep track of all the circles, we simply increment all of the cells that are turned on by every circle. After having processed all points in this fashion, we analyze the accumulator array and determine the number of hits per cell.

Every hit casts one vote for a point lying on that particular circle. The cell with the maximum number of hits, m, yields the center of the circle in the spatial domain that passes through m points. Similarly, other peaks in the accumulator array identify additional circle centers. Tracking the points contributing to the peak in the accumulator array identifies the points lying on the circle of known radius.

# APPENDIX B

## Image Resampling Techniques

After determining the geometric relation between the input and reference images, the relation can be used to resample one image to the space of the other. The resampling process involves the extraction and interpolation of gray levels from pixel locations in the original distorted image (input image) and their relocation to the approximate matrix coordinate location in the rectified (reference) image. After transformation, the gray value at an integer location in the input image will not be projected to an integer location in the reference image. Therefore, a decision has to be made about the gray value in the new non-integer location. There are several methods that can be used for this purpose, which include nearest neighbour, bilinear interpolation and cubic convolution resampling techniques.

The *nearest neighbour* approach uses the value of the closest input pixel for the output pixel value. As shown in Figure B1, among the four pixel gray values ($g_{11}$, $g_{12}$, $g_{21}$, $g_{22}$) the one closest to (x,y) is determined and its gray value is used as the gray value at (x,y).

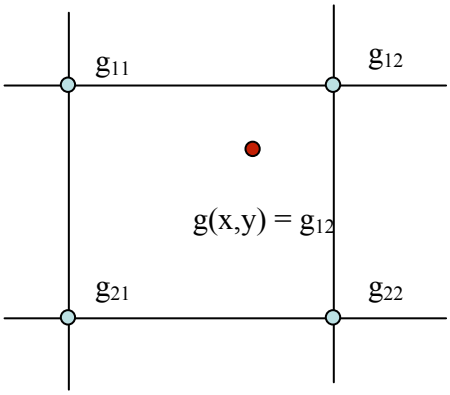
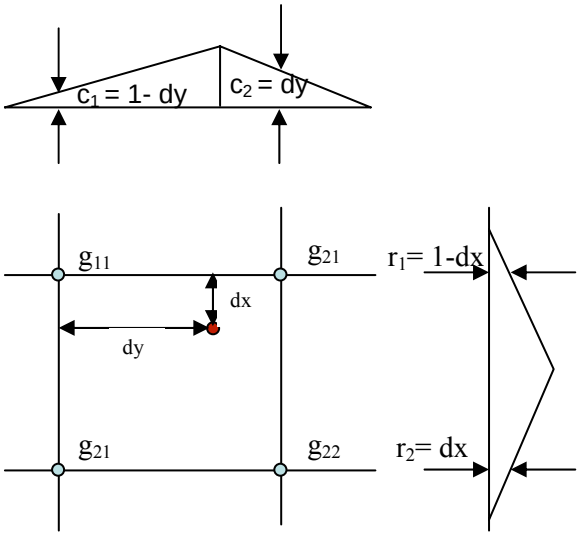



**Figure B1:** Nearest Neighbour Resampling

The *nearest neighbour* method is considered the most efficient in terms of computation time. Because it does not alter the gray level value, nearest neighbour interpolation is preferred if subtle variations in the gray levels need to be retained, or if classification will follow the registration. Nearest neighbour interpolation introduces a small geometric error into the newly registered image. The image may be offset spatially by up to 1/2 a pixel, causing a jagged, blocky or stair-step appearance.

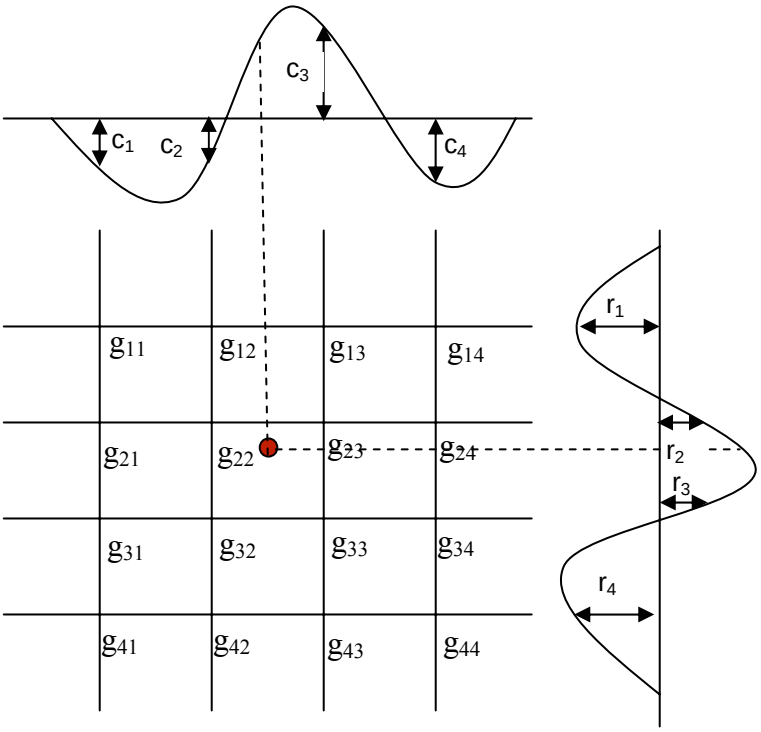
*Bilinear interpolation* determines the gray level from the weighted average of the four closest pixels to the specified input coordinates, and assigns that value to the output coordinates, see Figure B2. This method generates an image of smoother appearance than that of nearest neighbour, but the gray level values are altered in the process, resulting in blurring or loss of image resolution. Because of these changes in the gray level values, any image classification processes should be performed before the interpolation. Bilinear interpolation requires 3 to 4 times the computation time of the nearest neighbour method.




$$g = g_{11}\, r_1\, c_1 + g_{12}\, r_1\, c_2 + g_{21}\, r_2\, c_1 + g_{22}\, r_2\, c_2$$

**Figure B2:** Bilinear Resampling

*Cubic convolution* determines the gray level from the weighted average of the 16 closest pixels to the specified input coordinates, and assigns that value to the output coordinates. This method is closer to the perfect sin(x)/x resampling method than nearest neighbour or bilinear interpolation. The image is slightly sharper than that produced by bilinear interpolation, and it does not have the disjointed appearance produced by nearest neighbour interpolation. Because the gray level values are altered by this method, any image classification processes should be performed before the interpolation. Cubic convolution requires about 10 times the computation time required by the nearest neighbour method. The process is illustrated in Figure B3 (Richards, 1993).




$$g = g_{11} r_1 c_1 + g_{12} r_1 c_2 + g_{13} r_1 c_3 + g_{14} r_1 c_4$$
$$g_{21} r_2 c_1 + g_{22} r_2 c_2 + g_{23} r_2 c_3 + g_{24} r_2 c_4$$
$$g_{31} r_3 c_1 + g_{32} r_3 c_2 + g_{33} r_3 c_3 + g_{34} r_3 c_4$$
$$g_{41} r_4 c_1 + g_{42} r_4 c_2 + g_{43} r_4 c_3 + g_{44} r_4 c_4$$

**Figure B3:** Cubic Convolution Resampling