Electronic Theses and Dissertations

12-2004

Using Linear Features for Aerial Image Sequence Mosaiking

Caixia Wang

Follow this and additional works at: http://digitalcommons.library.umaine.edu/etd

Part of the Databases and Information Systems Commons, and the Graphics and Human Computer Interfaces Commons

Recommended Citation

Wang, Caixia, "Using Linear Features for Aerial Image Sequence Mosaiking" (2004). *Electronic Theses and Dissertations.* 573.
http://digitalcommons.library.umaine.edu/etd/573

USING LINEAR FEATURES FOR AERIAL
IMAGE SEQUENCE MOSAICKING

By

Caixia Wang

B.S. Wuhan University, 1996

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Spatial Information Science and Engineering)

The Graduate School

The University of Maine

December, 2004

Advisory Committee:

Anthony Stefanidis, Assistant Professor of Spatial Information Science and Engineering, Advisor

Peggy Agouris, Associate Professor of Spatial Information Science and Engineering

Mary Kate Beard-Tisdale, Professor of Spatial Information Science and Engineering

USING LINEAR FEATURES FOR AERIAL
IMAGE SEQUENCE MOSAICKING

By Caixia Wang

Thesis Advisor: Dr. Anthony Stefanidis

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Spatial Information Science and Engineering)
December, 2004

With recent advances in sensor technology and digital image processing techniques, automatic image mosaicking has received increased attention in a variety of geospatial applications, ranging from panorama generation and video surveillance to image based rendering. The geometric transformation used to link images in a mosaic is the subject of image orientation, a fundamental photogrammetric task that represents a major research area in digital image analysis. It involves the determination of the parameters that express the location and pose of a camera at the time it captured an image. In aerial applications the typical parameters comprise two translations (along the x and y coordinates) and one rotation (rotation about the z axis). Orientation typically proceeds by extracting from an image control points, i.e. points with known coordinates. Salient points such as road intersections, and building corners are commonly used to perform this task. However, such points may contain minimal information other than their radiometric uniqueness, and, more importantly, in some areas they may be

impossible to obtain (e.g. in rural and arid areas). To overcome this problem we introduce an alternative approach that uses linear features such as roads and rivers for image mosaicking. Such features are identified and matched to their counterparts in overlapping imagery. Our matching approach uses critical points (e.g. breakpoints) of linear features and the information conveyed by them (e.g. local curvature values and distance metrics) to match two such features and orient the images in which they are depicted. In this manner we orient overlapping images by comparing breakpoint representations of complete or partial linear features depicted in them. By considering broader feature metrics (instead of single points) in our matching scheme we aim to eliminate the effect of erroneous point matches in image mosaicking. Our approach does not require prior approximate parameters, which are typically an essential requirement for successful convergence of point matching schemes. Furthermore, we show that large rotation variations about the z-axis may be recovered. With the acquired orientation parameters, image sequences are mosaicked. Experiments with synthetic aerial image sequences are included in this thesis to demonstrate the performance of our approach.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Automatic Mosaicking of Aerial Image Sequences

Image mosaics are generated by pasting together individual frames to generate a synthetic image with large field of view. In geospatial applications, image mosaics are commonly generated using aerial and satellite photographs to produce large scale, map-like coverage for the depicted areas. For example, a strip mosaic (assembled from a single strip of photography) is commonly complementing existing maps when any type of route study is underway, such as for a highway, railroad, transmission line, pipe line, canal, or set of flood-control levees (Moffitt and Mikhail, 1980). With the advancement of digital techniques, automatic image mosaicking has become an active area of research and a number of techniques have been developed for this task (Takeuchi et al., 1999; Shum and Szeliski, 2000; Su et al., 2004). We can identify two major steps involved in this process. First, a *geometric transformation* is required to link together multiple overlapping images, a task commonly accomplished through the identification and matching of conjugate (corresponding) points in the overlapping area of the processed imagery. Second, a *radiometric transformation* is required, in order to blend image intensities in successive images, and thus generate visually seamless mosaics. This thesis addresses the first issue, namely the geometric transformation of overlapping image sequences to generate image mosaics.

1

At the age of digital imaging, mosaicking is becoming a rather popular process, as many commercial off-the shelf (COTS) software packages typically include tools to blend individual images in mosaics. Considering the currently available capabilities to annotate image files using text and/or graphics, one can easily see the great potential of mosaics to serve as map substitutes in geospatial applications: they can convey updated information in a user-friendly and content-rich manner. However, even though the geospatial community realized this potential early on, the use of mosaics as map substitutes has remained rather low, hampered by the lack of efficient automated techniques to solve the correspondence problem in overlapping imagery. Identifying and matching the same feature in two images remains a challenging task, despite the efforts made in this direction within the scope of automated image orientation.

Image orientation is a fundamental task in photogrammetry and computer vision. It involves the determination of the parameters that express the position and pose in space of a camera at the instance it captured an image. Considering the various spaces (e.g. inside and outside the camera) and different types of coordinate systems (e.g. absolute or relative) involved, we can identify *interior*, *relative* and *absolute orientation* (please refer to Appendix A for a more in-depth presentation). Developing automated orientation processes remains a major scientific challenge, with notable work on relative orientation (Schenk et al., 1991), point transfer in photogrammetric block triangulation (Agouris and Schenk, 1996) and exterior orientation (Drewniok and Rohr, 1996). Despite these advances, automated image orientation remains a challenge, typically affected by the lack of efficient and robust object extraction techniques to support feature selection, and the lack of efficient techniques to match such features. Notable literature in these

topics tends to focus on rather limited tasks, e.g., road extraction (Baumgartner et al., 1997; Katartzis et al., 2001; Tupin et al., 2002; Alhichri and Kamel, 2003) or edge detection (Basu, 2002) in very limited environments (such as high resolution imagery, rural scenes or intensity variation).

The underlying problems that limit the success of automated orientation relate to the considerable variety of object types and scales in natural scenes, and the complexity of modeling and representing spatial relations between objects. In order to overcome these problems automated techniques focus on using conjugate points (instead of more complex features like linear elements) in massive amounts: often hundreds of points are matched in a single stereopair. This "brute force" approach to orientation works well for traditional photogrammetric applications as orientation parameters are estimated with high accuracy, but require substantial post-matching manual editing in order to identify and remove false matches. Thus, they remain in essence partially automated, and are therefore unsuitable for modern applications (e.g. processing motion imagery captured by digital cameras on-board unmanned aerial vehicles – UAVs) that involve very large numbers of images. Furthermore, UAV-type applications often involve rural terrain types where besides few linear features (e.g. roads) it is practically impossible to identify an adequate number of well-distributed distinguished points that may be used for matching. All these still unsolved issues and emerging capabilities are making the need for efficient automated image mosaicking processes a priority for geospatial applications.

1.2 Statement of Objective

The objective of this thesis is to automatically co-register sequences of aerial imagery using elongated features present in these photographs as control information. We are particularly interested in recovering and correcting rotation variations and translations in order to generate automatically mosaics of aerial imagery. We also consider mainly sequences of quasi-vertical imagery, with relatively small variations of flying height between successive frames.

We are motivated by emerging data collection schemes, and especially ones involving UAVs and motion imagery, and aim to contribute a model to automatically recover rotations and translations in areas traversed by linear features such as rivers or curvilinear road segments. By recovering such orientation differences we are able to rectify and mosaic image sequences.

The hypothesis of this thesis is that linear features provide an efficient alternative to single points for the recovery of orientation variations and the mosaicking of image sequences. We argue that the geometric information conveyed by linear features provides additional robust content for this process, minimizing the potential for blunders often associated with single point-based approaches.

1.3 Intended Audience

This thesis is presented from the perspectives of photogrammetry, remote sensing, and computer vision. A basic knowledge of photogrammetry is assumed, but to assist readers from the geospatial community at large we have included an overview of fundamental photogrammetric principles in the Appendix. The fields of photogrammetry,

remote sensing, and computer vision will be able to expand the concepts and model for further complicated applications.

1.4 Organization of Thesis

This thesis is comprised of six chapters. Chapter 1 gives a brief introduction of the problem and its significance. In chapter 2, we provide a review of literature related to automatic techniques in the image mosaicking and relative orientation. The focus in this section will address automatic relative orientation with various types of control elements. The characteristics of our proposed approach will be introduced at the end of this section. Our detailed approach for automatic image mosaicking is described in Chapters 3 and 4. We start in Chapter 3 to describe the principles of road extraction with the *active contour model* (i.e., *snake*) as the basis of our approach. Based on the extracted linear features, we address our matching method in Chapter 4 using linear features. Chapter 5 focuses on the experiments with the synthetic aerial image sequence using the proposed approach. The precision estimation and analysis will be discussed in this section. In Chapter 6, we conclude with a summary of the research and suggestions for the future work. This thesis includes an appendix with an overview of fundamental photogrammetric principles, and a second appendix with an overview of orientation techniques in photogrammetry, to assist readers.

Chapter 2

LITERATURE REVIEW

The objective of this section is to review previous work done within the scope of image mosaicking and image orientation. We focus on reviewing dominant methods to compute correspondences in the above two fields. The discussion will concentrate on automated techniques involving extraction and matching of control elements such as points, linear features or areas. We conclude this section with brief introduction of our approach and address the advantages of our approach compared with most closely related state-of-the-art works.

2.1 Computing Correspondences in Automatic Image Mosaicking

Image mosaicking creates a composite view or panoramic mosaic from a sequence or a collection of overlapping images with smaller fields of view. Composing has traditionally been done manually. But digital photography enabled new automatic implementations for mosaicking (Milgram, 1975; 1977; Peleg, 1981; Burt and Adelson, 1983; Irani et al., 1995; Shum and Szeliski, 2000) which were first applied to aerial and satellite images and later used for scene and object representation. Computing correspondences between image sequences is a key step in making panoramic image mosaics. Techniques can be mainly categorized into featureless methods and feature-based methods (Gong et al., 1999). Featureless methods transform images with parameters acquired by minimizing a sum of squared difference function. Existing

featureless techniques include cylindrical/spherical panoramas (Krishnan and Ahuja, 1996; Shum and Szeliski, 2000), affine or a planar -projective transform-based panoramas (Hansen et al., 1994; Irani et al., 1995; Sawhney et al., 1995). However, methods in this category typically restrict their applications to small change (translation, rotation, etc) from one image to another, and good initial values for the parameters of the transform. Feature-based methods utilize the feature correspondence between image pairs to find transforms that register the image pairs. The challenge of these methods exists in the acquisition and tracking of image features due to the available features and noise or occlusion on the imagery. Point features such as corners, as the most common features, have been widely used (Zoghiami et al., 1997; Capel and Zisserman, 1998; Kanazawa and Kanataniy, 2002; Mallick, 2002; Gracias et al., 2003). Recently, the extending applications of image mosaics, such as virtual travel (Chen, 1995), visual maps for autonomous navigation of mobile robots (Garcia et al., 2001; Gracias, 2003), highlight the need of alternative features other than points for correspondence problems, typically in environments where distinctive points are impossible to identify, e.g., low-contrast, suburban areas with little man-made architecture.

2.2 Conjugate Feature Detection in Digital Aerial Photogrammetry

Selecting appropriate features to compute correspondences have been studied in aerial photogrammetry for a long time. In aerial photogrammetry, relative orientation reconstructs the relative position of the two consecutive images (please refer to Appendix A for a detailed description). With this knowledge, the two images can be mosaicked through a variety of transformations, ranging from simple affine to complex projective ones in terms of the distortion between imagery. Automating the process of relative

orientation has been developed with the advent of digital photogrammetry and brought great challenge forward.

During the last decade or so, photogrammetry has moved to an increasingly digital spectrum, fostered by a wide range of technological advances (e.g. large format digital cameras and scanners, stereoscopic glasses for video games). Digital imagery has effectively replaced its analog counterparts for photogrammetric applications. High-performance computers and sophisticated software have replaced the cumbersome optical/mechanical stereoplotters or analytical stereoplotters. Even though this evolution has materialized over the past 15 years, its origins go as back far as 1959 with the pioneering work of Gilbert Louis Hobrough on image correlation (Hobrough, 1959). Back then, the correlation process was an analog one, with hardware used to compare the gray levels of imagery. From this experimental and impractical early step we have now reached the point where digital photogrammetric workstations (DPWS) have become the standard equipment of most photogrammetric firms (Heipke, 1995).

The automation of relative orientation is a fundamental process for digital stereo processing and thus has been the subject of substantial research activities in photogrammetry and computer vision (CV). Even though this work has resulted in substantial progress, often materialized in the form of competent software packages, we are still at a stage where existing solutions leave room for improvement.

Automatic relative orientation requires the identification of control primitives (through *feature extraction*) from digital images, and establishing correspondences among them (through *feature matching*) in pairs (or multiples) of overlapping images. These processes can be performed with various degrees of ease by human operators, and

8

their automation has been addressed in the development of numerous algorithms in photogrammetry and CV. The variety of information contained in aerial imagery provides a wealth for selecting control primitives in one hand, as well as a challenge in another hand for matching. Therein lies the fundamental difficulty for automatic relative orientation, e.g., point identification and matching (Schenk and Toth, 1993; Tang and Heipke, 1996), relational descriptions between points, lines or areas (Cho, 1995; 1996; Wang, 1996). We can identify two major types of strategies for solving the matching problem in the photogrammetric and CV literature, known as area-based and feature-based matching. It should be noted that when talking about area-based or feature-based matching in this chapter we not only refer to the whole matching process, but also the selection of the primitives.

2.2.1 Area Based Matching

In area based matching (ABM), the candidates for matching in overlapping images are windows of predefined size or even entire images. Matching is based on comparison of raw gray values (also referred to as pixel intensity) in a preset window (template) in one left image and of a search area in the other image. ABM can be implemented either through cross-correlation or least-square approaches. In cross-correlation, similarity is measured by assigning correlation values to each location by comparing the template window content to the corresponding matching candidate pixels (Ackermann, 1984; Hannah, 1989; Rottensteiner, 1993). Alternatively, the normalized spatial root mean square deviation, or absolute difference of normalized differences may be used to express correlation between these two windows. The local maxima (or minima) express best matching. Least-square approaches establish the correspondence

9

between the windows that minimize the squared sum of the differences of their gray values (Forstner, 1982; Bergen et al., 1992; Danuser, 1996; Berger, 1998). Although the precision of area based matching can be potentially high in well-textured image regions, it is affected by geometric distortions (e.g., due to the relief displacement). In addition, it is not invariant to rotation variations of two overlapping images. In a deviation from traditional orthogonal windows, some authors have recently proposed to use circular windows (Zitova and Flusser, 2003).

2.2.2 Feature-based Matching

Image features, from a photogrammetric and CV perspective, may range from local features such as points, edgelets (edge elements), and lines, to global features such as polygons and complex descriptions of the image content called structures. Even though the difference between the terms local and global is rather artificial (Heipke, 1996), it is customary to consider *local* features as those contained within highly localized image windows (e.g., in the order of 50*50 pixels). On the other hand, the term *global* is commonly reserved for features that span larger areas, or even a complete image. The most popular feature-based approaches may be classified in two broad categories: ones that focus on points as features, and ones that use more complex entities like lines and objects.

2.2.2.1 Point Features as Primitives

Point features are widely used in terms of their invariance to imaging geometry (Heipke, 1997) and their good perceptibility by a human observer. Distinct points such as road intersections, oil and gas pads, high variance points, sharp corners, are commonly used as control primitives in the photogrammetric community. Their automatic matching

usually starts by searching for interesting points that have distinct differences in contrast to their vicinities. They are detected by so-called interest operators in the extensive body of research, e.g., Moravec operators (Moravec, 1977), Förstner operator (Forstner, 1986), statistical operator (Hannah, 1980). The underlying algorithms of these detectors differ in the manner in which they define the concept of a 'point'. For example, the high variance points detectors such as Moravec operators, Förstner operator work by finding those points which have high variant gray values with respect to their background (Muller and Hahn, 1992; Haala et al., 1993; Hahn and Kiefner, 1994; Liang and Heipke, 1996). Corners are more complex. Intuitively, corners are recognized as points with high curvature along region boundaries. However, defining corners mathematically is not a trivial issue. A great deal of effort has been spent on this problem, in particular on developing precise, robust and fast methods for corner detection. Second-order partial derivatives of the image function are exploited for corner detection (Kitchen and Rosenfeld, 1982). Gray scale level based corner detectors search the corners as the local extrema of the Gaussian curvature (Dreschler and Nagel, 1981). However, corner detectors based on the second-order derivatives of the image function are sensitive to noise. Gray scale level based corner detectors do not detect the exact position of a corner. Deriche and Giraudon (1993) proposed to localize the exact position of a corner with two properties: a) the Laplacian image is zero at the exact position of the corners (zero-crossing property), and b) the property associated with the measure they proposed to use. The most up-to-date and exhaustive survey of corner detectors can be found in Rohr (2001).

In aerial photogrammetry, point feature matching is typically based on the powerful epipolar constraint (shown in Appendix B) in which the candidates of conjugate points should lie on epipolar lines. This assumption reduces the search space to a single dimension - the epipolar lines - and thus increases the speed and reliability of the matching algorithm. Furthermore, a radiometric check is incorporated in the algorithm such as the correlation coefficient (Liang and Heipke, 1996). The shortcoming is that the approximate orientation parameters should be known as initial values for matching. After feature matching, the use of a multitude of points provides higher redundancy and makes a least squares adjustment available to determine the relative parameters using the available model coordinates of the conjugate points as observations, which results in more reliable results with respect to the analytical relative orientation with few points.

Hierarchical approaches are often employed to reduce the burden of computation since the number of detected points can be very high. In hierarchical approaches, images are often analyzed in multiple resolutions, from coarse to fine, using for example image pyramids (see Figure 2.1) or other multiresolutional analysis schemes. In such hierarchical approaches, feature extraction and matching are performed in each level separately, starting with coarse levels, and using the latest results to guide feature extraction and matching in finer resolutions.
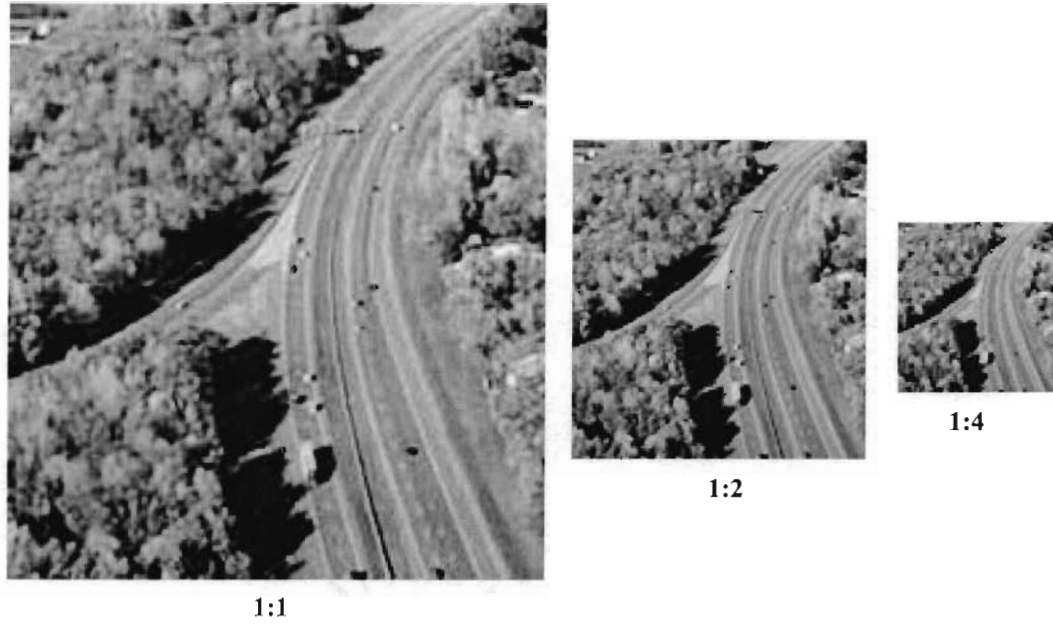
Figure 2.1: Example of an image pyramid

2.2.2.2 High Level Features as Primitives

High Level features, defined in the context of this thesis as all features other than points (e.g., roads, buildings, regions) are not only easier to find in the natural environments but also more meaningful than the point features. Nevertheless, they are more challenging from an algorithmic point of view, due to their substantial variations in nature, form, scale, and overall appearance characteristics. Furthermore, standard projection models used to support point matching in relative orientation are less suitable to support the deformations in the representation of longer and larger entities. Thus the use of higher-level features for relative orientation applications is rather limited.

Even so, much effort has been devoted over the past decades to devise efficient and robust algorithms for it. Relational matching (Shapiro and Haralick, 1987; Vosselman, 1992) is one direction. It relies on the similarity of topological relations of features which are stored in feature adjacency graphs rather than on the similarity of gray

levels or the similarity of point distributions. Topology is invariant under perspective transformation and thus relational matching is a rotation invariant solution for relative orientation and can be used to determine the image overlap. However, since it leads to rather complex search trees (Vosselman, 1995), the computational complexity is very high, thus limiting its use in practice.

Other high level features such as curved lines (Schenk et al., 1991), and straight lines confront great limitations for automatic relative orientation. For example, relative orientation can only use straight lines that are parallel to the epipolar lines for solving the orientation parameters (Schenk, 1999). The collinearity equations used to solve the orientation parameters are more appropriate for points instead of high-level features since they are constructed based on the corresponding points. High level feature based automatic relative orientation is still a subject of intensive research to overcome these limitations (Heipke, 1997; Jones and Oakley, 2000; David et al., 2003).

2.3 Most Related Work Review

Our proposed approach is presented in detail in chapter 3 and 4, but in order to better position it within the current state-of-the-art we will present a brief overview of road extraction and shape matching approaches in photogrammetry and CV in the following subsections.

2.3.1 Road Extraction

Road extraction is commonly regarded as a difficult problem. Firstly, images have a wide range of scales, which leads to a variety of road representations such as two parallel lines in large-scale images or only a single line otherwise. Secondly, roads do

not possess specific global shape. No signature shape can be specified as in template matching methods.

A great deal of effort has been devoted to devise efficient and robust algorithms for it (Gruen et al., 1995b; Gruen et al., 1997; Baltsavias et al., 2001). With respect to the level of automation, the methods are conventionally categorized into *semi-automation* (*SA*) and *full-automation* (*FA*). Examples of substantial works are shown in Table 2.1 (Doucette, 2002).

Table 2.1: Summary of approaches for automated road extraction

Extraction Models	Automation Level	Exemplary References
Road tracking	SA	Litton, 1993;
Template Matching	SA	Gruen et al., 1995a; Vosselman and Knecht, 1995; Park and King, 2001;
Snake-based models	SA	Trinder and Li, 1995; Fischler and Heller, 1998; Peteri et al., 2003a;
GIS database update	FA	Zhang and Baltsavias, 2000; Agouris et al., 2001b
Rigorous models	FA	Baumgartner et al., 1999; Oddo et al., 2000; Agouris et al., 2001a;

The implementations of methods in SA all require the real-time interaction of a human operator, for example, inputs of the road direction, width, initial seeds. It should be noted that the snake model introduced from Kass et al. (1987) is distinct from conventional methods (e.g., road tracking) and more recently developed. Similarly, as a deformable object deforms, the snake attaching itself to an edge location provides the final edge delineation until its energy reaches the minimum.

To contrast, full automation (FA) algorithms strive for little or no human operator interaction with automating the initialization. Existing cartography information facilities in GIS database algorithm for FA road extraction. Rigorous FA road extraction research even contributes to fulfilling self-sufficiency in seed finding and /or delineation.

2.3.2 Shape Matching

Shape matching is an important issue in visual information systems, computer vision, pattern recognition, and robotics. As such, there are various ways to approach the problem in computational geometry. Computational geometry is the subarea of algorithm design that deals with the design and analysis of algorithms for geometric problems involving operations on objects like points, lines, polygons, and polyhedra. Table 2.2 offers a good tabulation of some representative examples Veltkamp and Hagedoorn (1999).

Table 2.2: Examples of shape matching approaches

Approaches	Matching strategy	Exemplary References
Tree pruning	As a tree search procedure, the matching algorithm generates all maximum matchings satisfying a condition called delta - boundedness	Umeyama, 1993
Generalized Hough transform	Find the correspondence with a given criterion	Zaharan, 1997
Statistics	Mathematical theory of shapes	Small, 1996
Deformable templates	Determine a body-centered coordinate frame for each objects and then attempt to match up the feature points	Sclaroff and Pentland, 1995
Fourier descriptors	Use the Fourier transform to characterize the shape	Loncaric, 1998
Wavelet transform	Haar basis functions	Jacobs et al., 1995
Shape context	The shape context at a reference point captures the distribution of the remaining points relate to it. Corresponding points will have similar shape context	Belongie et al., 2002

However, the orientation parameters are required to be available prior to solving the correspondence problems. In addition, the majority of the existing literature only deals with the images having the same objects. It should be recalled that we are interested in overlapping images with unknown rotation difference. Thus, we are facing a challenge when matching the extracted edges based on the shape information, i.e., the proposed solution should be invariant to rotation and simultaneously detect the overlapping area.

2.4 Our Proposed Approach

As we discussed previously, high-level features carry more information (e.g., shapes, intensity) than point features for subsequent processes such as object recognition, while points are more appropriate for the solution for orientation parameters. This thesis proposes a robust method that takes advantage of both point and high level features (i.e., roads) to support the mosaicking of image sequences. We concentrate on roads primarily because roads in remotely sensed scenes are widely regarded as the most recognizable objects for humans, purely based on their shape and its variations. In our proposed method, we utilize the snake model to extract roads. Although the human interaction may place a time constraint on algorithm execution, semi-automation is reliable with the initialization as the constraints. In addition, the snake model delineates the edges with sets of points, which are appropriate for resolving relative orientation parameters. We proceed by making use of points of interest along road edges to perform feature matching and therefore bear no restrictions of shape matching (i.e., invariance to rotation variations). By combining feature- and point-based matching we aim at the development of a robust yet fast matching approach. Finally, with the detected orientation parameters, all images other than the first one (referred to as the anchor) in the sequence are transformed into the anchor space and generate the mosaic of the set of sequence. We determine rapidly the overlapping area of the image pairs by using linear features and reduce drastically the search space for our matching process. Furthermore, we avoid matching on a point-to-point basis, minimizing the potential for costly blunders that may require post-process manual editing, and would therefore minimize the potential use of our approach.

18

Chapter 3

ROAD EXTRACTION WITH SNAKES

In Chapter 3 and 4, we detail our proposed approach for mosaicking image sequence. The approach proceeds by determining the orientation parameters of image pairs based on the extracted roads (i.e., control features), followed by transforming images other than the first image (referred as anchor) into the anchor space with detected orientation parameters, described in Figure 3.1.
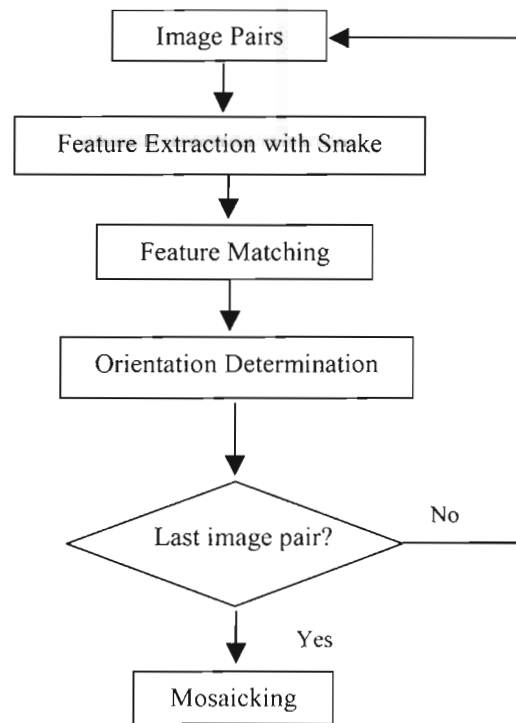

Figure 3.1: Mosaicking image sequence

In this section, we focus on extracting control features for matching. Different selected features require different mathematical models for automatic extraction and

matching (Zitova and Flusser, 2003). Hence, it is a critical issue to choose the most suitable features for performing such a task. Two major factors are considered in terms of feature selection: the availability of features in the related taken scenes and the distortion of the two involved images. In aerial photogrammetry, point features such as road intersections, corners, even the positions with high gray variance, typically are the preferred control features in terms of their stable geometry and traditional use as control points (Heipke, 1997). However, they convey little information other than the positions. In addition, within this thesis, we concentrate on aerial imagery covering rural areas with minimal numbers of salient point features (and sporadic distribution of these points) but with linear features (i.e., roads) that cross the images. Furthermore, roads are static during the image-capturing process. We assume two images are taken from different viewpoints given the movement of the UAVs between successive exposures. There is only a spatial difference between the two images, i.e., the roads are the same in the common coverage of the two images with translation and rotation misregistration. With respect to these aspects, we select roads as the control features to determine the orientation difference.

3.1 General Formulation of Snakes

An extensive body of research has dealt with extracting roads from digital imagery (Doucette, 2002). The snake model, as a distinguished method from conventions, has caught attention and been developed recently, for example, (Agouris et al., 2001b) advanced the differential model based on it to automatically detect the changes in GIS database with aerial imagery. Peteri et al. (2003b) makes use of snakes combined with a multiresolution analysis (MRA) for minimizing the problem of geometric noise to extract road networks.

Snakes, sometimes referred to as *active* or *deformable* contour models, are common tools to extract object boundaries from imagery in computer vision. They were introduced by the work of Kass et al. (1987) and have since been the subject of substantial research. The snake model combines radiometric (gray level) and geometric (shape description) information to extract features. In the snake model, a linear feature is extracted as a sequence of nodes and links among them. Within this thesis, snakes are applied as a means to extract road edges from aerial images for image mosaicking. Two essential advantages exist. As we know, aerial images include extensive information. It is extraordinarily difficult to extract a specific feature automatically from such rich information. In snakes, information from other mechanisms such as interaction with a user, interaction with some higher-level image understanding process can be used to limit the search area so as to decrease the uncertainty. The other advantage is that points, as the extraction results, are appropriate to use for solving the orientation parameters. Simultaneously, challenges exist in the further step to map the extraction results in order to solve relatively accurate orientation parameters, which will be detailed in the followed section.

The snake model is defined by an energy function, combining weighted internal and external forces as

$$E_{snake} = \alpha \cdot E_{cont} + \beta \cdot E_{curv} + \gamma \cdot E_{edge} \tag{1}$$

Where α, β and γ are relative weights describing the importance of each energy term. Commonly their selection is performed empirically.

- *Internal forces*

21

Internal forces consist of the first two energy terms, i.e., E_{cont} and E_{curv}. They emanate from the geometry of the contour. Suppose $v_i = (x_i, y_i)$ is a point on the contour. E_{cont} expresses the first order continuity constraint defined as

$$E_{cont} = \bar{d} - |v_i - v_{i-1}|$$ (2)

Where $\bar{d}$ is the average distance defined as Equation 3, which forces the snake nodes to be evenly spaced, avoiding grouping at certain areas, while at the same time minimizing the distance between them.

$$\bar{d} = \sum_{i=1}^{points-1} |v_{i+1} - v_i| / (\text{points-1})$$ (3)

Another internal force E_{curv} in Equation 1 denotes the second-order continuity constraint that represents the curvature of the snake contour, and allows us to manipulate its flexibility and appearance, i.e.,

$$E_{curv} = |v_{i-1} - 2v_i + v_{i+1}|^2$$ (4)

- *External force*

The third energy term E_{edge} in Equation 1 is referred to as the *external force*. It describes the relation of the contour to the radiometric content of the image. In general, it forces points to move towards image edges, defined as:

$$E_{edge} = -\nabla I(v_i)$$ (5)

With few seed points selected in an image manually or from higher-level image understanding process, the points along the contour move to new locations to progressively optimize the energy functions, i.e., minimizing its energy, guided by the discussed external forces and influenced by the internal forces. This is an iterative process. Various techniques have been proposed to compute the minimum of the snake

22

energy function. In (Kass et al., 1987) the Euler-lagrange equation is used to analyze and solve numerically the snake equations. (Amini et al., 1990) proposed an algorithm for the optimization using dynamic programming. We use an alternative faster approach i.e., greedy algorithm, suggested by (Williams and Shah, 1992) in this thesis.

3.2 Mapping Issues of the Extracted Point Sets

Using the above described snake model and optimization process, roads are extracted in successive images, and are represented by polygonal lines, defined by points (snake nodes) and segments connecting these points. In automating image orientation, matching plays a basic role. Also referred to as the correspondence problem, matching can be defined as the establishment of correspondences among two or more data sets (identifying the same feature in them). In our applications (sequences of aerial imagery), the two data sets that we analyze (i.e., extracted road segments and corresponding point sets) have the following particular characteristics with respect to the matching problem:

- Differences in coverage and coordinate system

Under typical aerial survey/monitoring missions (refer to Appendix A), two adjacent images are acquired from successive exposure positions as the sensor flies over an area including one or more linear features. Thus, the images have certain sidelap and endlap without covering the same zone of the ground. Correspondingly, the extracted points along the road edge cover different parts of a road edge. As in Figure 3.2, I_1 and I_2 represents the two acquired images on the successive exposure position A_1 and A_2.

23
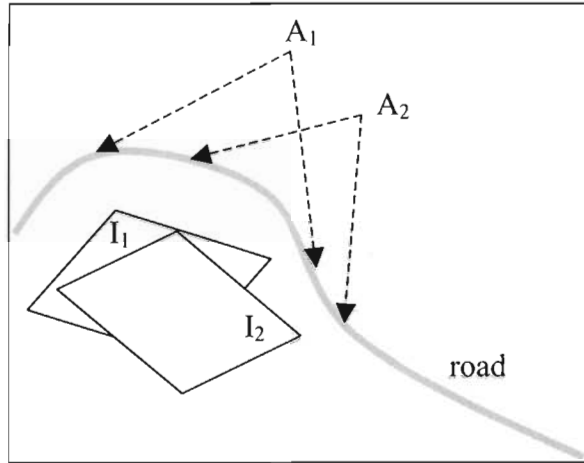
Figure 3.2: Coverage of successive Images

Simultaneously, as the UAVs fly along the linear features, the image coordinate systems of image I_1 and I_2 are not the same, with translation and rotation differences. Figure 3.3 shows the translation D and rotation κ between system $o_1x_1y_1$ and $o_2x_2y_2$, where y_1' is the translated axis y_1 by D.
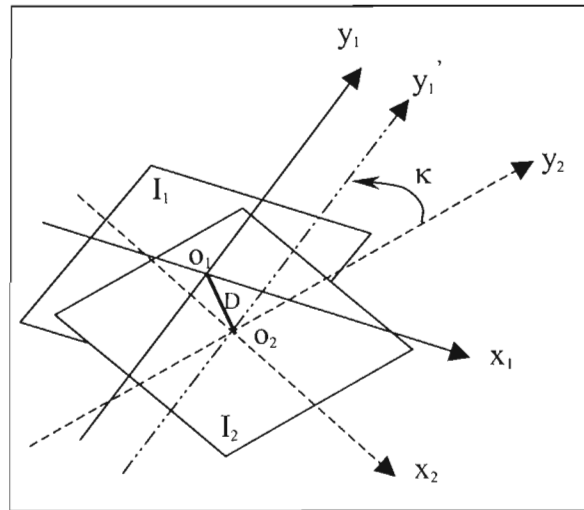
Figure 3.3: Translation and rotation difference of image I_1 and I_2

- Noise

In aerial photogrammetry, noise within the context of road extraction primarily refers to the adverse effects of shadows and visual obstructions (e.g. canopy, large

buildings) adjacent to road segments. As these obstructions tend to be three-dimensional in nature (e.g. a tall tree or bridge), their effects on each image may be slightly misplaced due to differences in the viewing angle as the UAV moves along its path. The snake model may be affected by noise (if noise has substantial spatial extent), following for example locally the outline of a canopy shadow instead of the road in question (e.g. Figure 3.4). Such noise effects manifest themselves as spurious concave or convex extremities at the noise locations, deviating slightly by the desired noise-free result (Figure 3.5). Furthermore, noise may cause snake nodes to be distributed differently in two conjugate scenes. This is demonstrated in Figure 3.5, where the reader can easily see how points extracted with or without noise effects do not have a one-to-one correspondence.
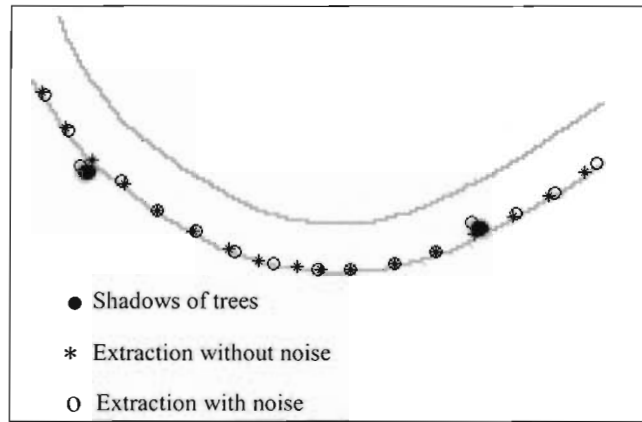
Figure 3.4: Tree shadows effects to the snake models
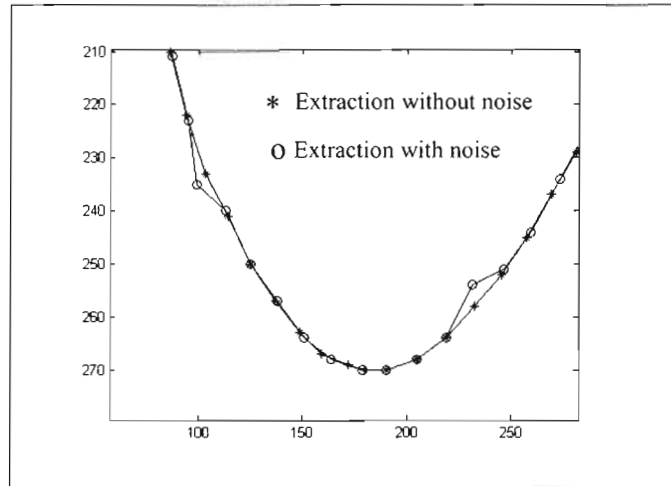
Figure 3.5: Concave and Convex caused by noise

Consequently, two sets of road nodes extracted with the snake model are not one-to-one mapped. And there is no guarantee that a solution exists, is unique, and /or is stable with respect to small variations (e.g., noise) in the input data, leading to an ill-posed problem. Additional knowledge is commonly required to find a solution for an ill-posed problem (Heipke, 1997), such as initial values of the unknown parameters. In this thesis, we develop a new approach for mapping based on both the extracted points and road shapes to overcome this problem, without the stringent prerequisites of additional information.

Chapter 4

ORIENTATION DETERMINATION AND MOSAICKING

As we mentioned in Chapter 3, the snake model is a powerful tool to extract linear features from digital imagery. Using snakes, roads are extracted from aerial imagery and are represented by polygonal lines, defined by points (snake nodes) and segments connecting them. By matching these polygonal lines we can establish correspondences among overlapping imagery, and thus recover orientation and mosaic image sequences, which will be described specifically in this chapter (please refer to Figure 3.1).

Matching snake-derived polygonal representations of linear features from overlapping imagery is a rather unique matching problem. As mentioned previously in this thesis, there exists substantial work in the photogrammetric and CV communities on point matching (Alt et al., 1988; Heffernan and Schirra, 1994; Belongie et al., 2002), region matching (Flusser and Suk, 1994; Bhatacharya and Sinha, 1997), and even linear matching in applications where matched images differ only in simple translation (Fonseca and Manjunath, 1996; Althof et al., 1997). Our approach contributes to this literature by addressing the matching of polygonal descriptions (extracted road outlines) where due to noise and algorithmic effects we do not have a direct one-to-one correspondence among snake nodes in overlapping imagery. Furthermore, we address translation and/or large rotation differences among overlapping image pairs. In order to

meet these particularities we introduce a method that departs from direct point matching which is regarded as a time-consuming and error-prone process. Through this matching process we can determine the orientation parameters relating multiple overlapping images, information that can be used to transform them into a common reference frame, and generate their mosaic.

4.1 Feature Matching in Image Pairs

Let L and R respectively denote the left and right images in a stereo pair. The sets of points that describe an extracted edge in each image are represented by $L_i(x_{L_i}, y_{L_i})(i = 1,2,...,n_L)$ for the L image and $R_i(x_{R_i}, y_{R_i})(i = 1,2,...,n_R)$ for the R image, where n_L and n_R are the overall numbers of extracted points for each image. It is easily understood that the number of points may differ between the left and right image, as their content varies (besides their common overlapping area). Furthermore, the object extraction process itself may result in differences in the number and placement of points (nodes) used to describe the same object under variations of the viewing conditions. Thus, in general, n_L is not equal to n_R, and we do not have a direct, one-to-one point correspondence.

Our approach for feature matching is briefly outlined in Figure 4.1. It is implemented in image pairs: image i is compared to image $i+1$, and image $i+1$ is in turn compared to image $i+2$, and this process is repeated until all image pairs are compared. Snakes are used to extract roads on each image (i.e., left and right images in an image pair). With the on-the-fly input curvature threshold by users, we detect critical points (CRPs, i.e., snake nodes having larger curvature than the threshold) in the right image. Then a polygonal template is constructed, customized by users with a subset of detected

28

CRPs. We proceed in our approach by identifying the similarity between the customized template and extracted road polyline from the left image. The outcome is a point-to-point matching between the two data sets (i.e., L_i and R_i).
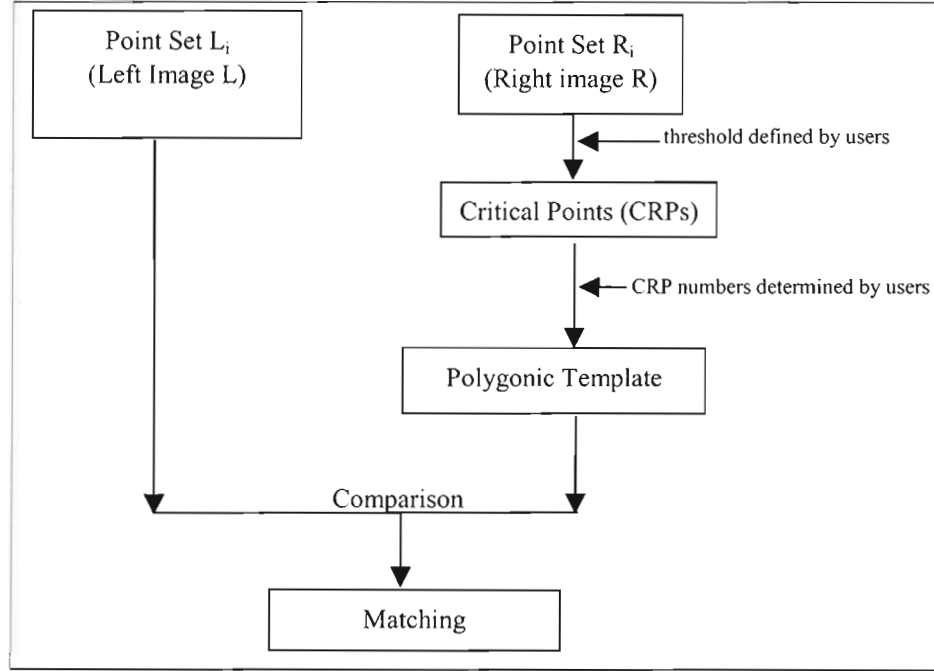
Figure 4.1: Shape matching

4.1.1 Critical Point Detection

When humans recognize an object as one that they have seen before, objects are decomposed to primitives called *geons* (David, 1982) combined with their spatial relationships in an object-centered frame. Similarity comparison is based on these geons and their relationships (David, 1982; Biederman, 1987; Hummel and Biederman, 1992). Similarly, when comparing two scenes, humans mentally deconstruct each scene to its characteristic elements (features within it) and then the elements are decomposed to more subtle properties (i.e., convex parts) to identify correspondences among them. In our case, snakes yield object outlines from the distribution of radiometric content in each

29

image separately. In order to support the subsequent comparison of these extracted linear objects, we first need an algorithm to analyze their geometric properties.

Our objective is to identify *notable* geometric characteristics of these lines (e.g. breakpoints, transitions to/from concave or convex components), and thus reduce the representation of a line to its characteristic information. These characteristics then provide the information that will be compared to match these two lines. Thus, determining characteristic or critical points is an essential step for our matching algorithm.

From linear algebra, we know that the angle φ between two vectors as shown in Figure 4.2 can be calculated with the following formula:

$$\varphi = \arccos(\frac{\overrightarrow{PO} \cdot \overrightarrow{PQ}}{|PO||PQ|}) \tag{6}$$

Where $\overrightarrow{PO}$ and $\overrightarrow{PQ}$ are the vectors starting from point P and ending at point O and Q respectively. $|PO|$ and $|PQ|$ are the norms of those vectors. The range of φ is between $0°$ and $180°$.
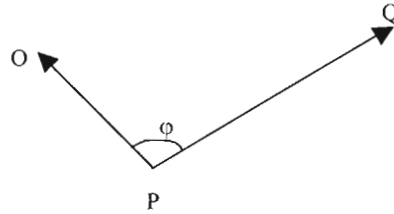

Figure 4.2: Angle between two vectors

After extraction, road edges are represented with polygonal lines defined by extracted points and segments connecting these points. In Figure 4.3, we used stars to show the extracted points, and dotted lines to show the connecting segments. Let φ_i

30

denote the angle between vectors $\overrightarrow{P_{i-1}P_i}$ and $\overrightarrow{P_{i+1}P_i}$, where P_{i-1} and P_{i+1} are the adjacent

prior point and next point to the current point P_i respectively. The *local curvature* θ_i for

each point P_i is defined in our approach as the complementary angle of φ_i:

$$\theta_i = 180° - \varphi_i \tag{7}$$
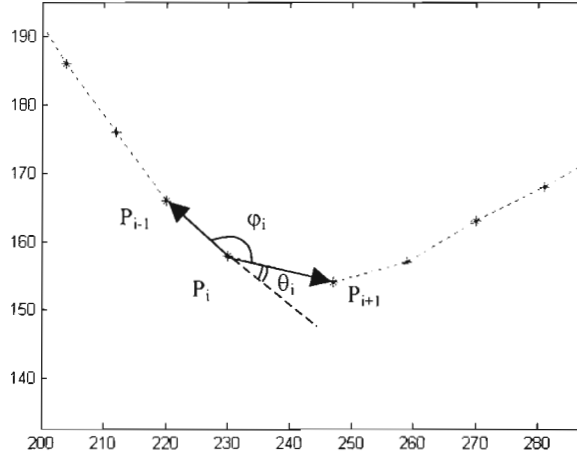
Figure 4.3: Local curvature θ_i for point P_i

Accordingly, the smaller the vector angle, the larger θ_i. The local curvature θ_i

of point P_i represents the abruptness at this point the variation in road direction. In Figure

4.4 we see that point P_i is a point of high curvature, as its corresponding angle θ_i is

rather large. Accordingly, we mark point P_i as a point of *local maximum curvature* and

thus consider it to be a *critical point*. When making such assignments we can use a

predefined threshold to evaluate sharpness (i.e. if an angle is higher than this threshold

the corresponding node is considered to be a critical point). By increasing this threshold

value, we force critical point selection to identify only a few nodes where road outlines

display very abrupt changes in orientation, thus selecting a limited number of highly

distinguishable points. By setting a lower threshold we can have more points detected as critical points, but among them there will exist some points of relatively low orientation variation. To a certain extent this process may be viewed as the equivalent of a multiresolutional decomposition of the outline, with fewer (more) points used to describe the outline when high (respectively low) threshold values are selected. Threshold selection may be performed empirically according to the degree of curvatures of roads. Automating threshold selection is a different research topic that is beyond the scope of this thesis.

The above presented process has a shortcoming as it fails to mark locations where a road rather slowly changes its curvature from a concave to a convex curve (or vice versa). This is demonstrated in Figure 4.4, where point set S_i ($i = 1,2...12$) shows the detected critical points. The variation of local curvature after point S_2 is smaller than a specific threshold set for local maxima. Only point S_2 (pointed by the solid arrow) is detected as a critical point using local curvature as the sole criterion. The information of curvature change from convex to concave, which is an important geometric characteristic to support shape matching, is lost. Even though decreasing the threshold value can detect more points (e.g., S_3, S_4, S_5, S_9, S_{10}), there are more chances to select critical points that locate at the almost unbent section (e.g., S_3, S_4, S_5) and thus increase the computation burden in later matching.
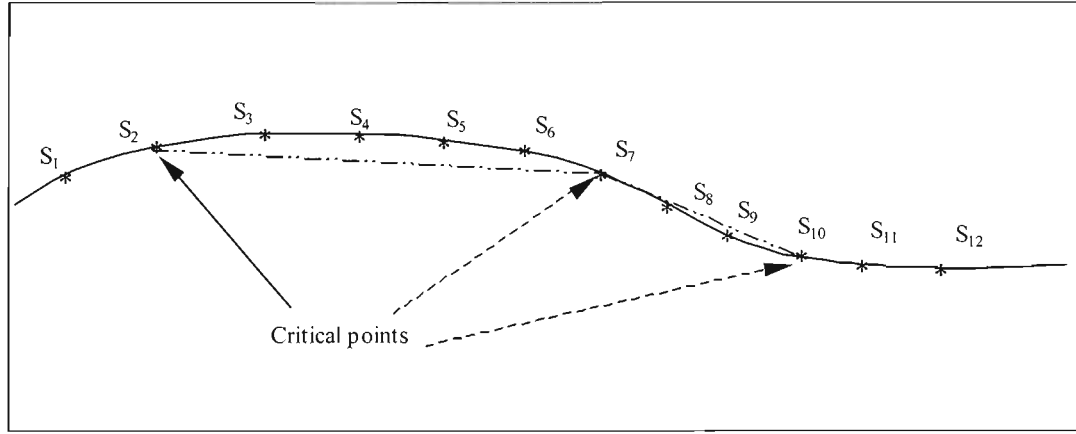
Figure 4.4: Poorly approximating with local maximum curvature

In order to overcome this problem we enhanced our curvature analysis by introducing a *global maximum curvature* criterion. It is defined as the complementary angle, as θ_i shown in Figure 4.5, of vectors generated by current point P_i, its adjacent prior critical point (CP_j) and the next snake node (P_{i+1}). However, if the points before current point P_i have smaller curvature than the threshold, the concave characteristic shown on point P_i is more likely to be lost. Thus, in our approach, the first snake node is always considered as the first CP.
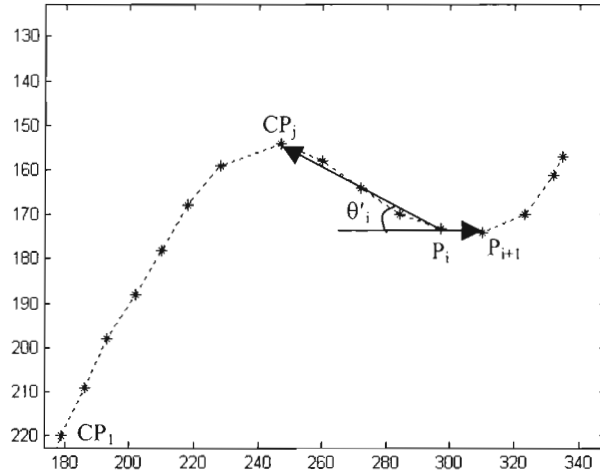
Figure 4.5: Global curvature for point P_i

Using the above described processes, critical points are determined using *local and global maxima curvature* criteria. The process starts by using the local criterion. As if a critical point is detected, the global maxima curvature criterion is executed and checks the rest of the points. The detected critical points form a polyline that represents "essential" curvature of the linear features. For the example shown in Figure 4.4, the points denoted by dotted arrows are detected with global maxima curvature. Thus, the polyline established by point S_2, S_7 and S_{10} retains the critical information lost otherwise.

To further study the performance of our critical point detection method, we compare it with Douglas-Peucker (DP) algorithm (Douglas and Peucker, 1973), the most popular method and even the standard by which all others are judged in cartography to reduce the number of vertices in a digital polyline (Ebisch, 2002). McMaster ranks DP algorithm as "mathematically superior" (McMaster, 1986). With a study of three simplification algorithms, based on Marino's work (Marino, 1979) on critical points as a psychological measure of curve similarity, White reported that the Douglas-Peucker

method was best at choosing critical points and representing perceptually the original lines. Other than angle tolerances used in our addressed approach, the DP algorithm is based on a distance tolerance. It defines a general direction of the line by the link between its start and end points. The intermediate points are tested to find the one with the greatest vertical distance between it and the link. If this distance is less than the tolerance, a straight line suffices to represent the whole line. Otherwise, split the link at this point and recursively approximate the two pieces.
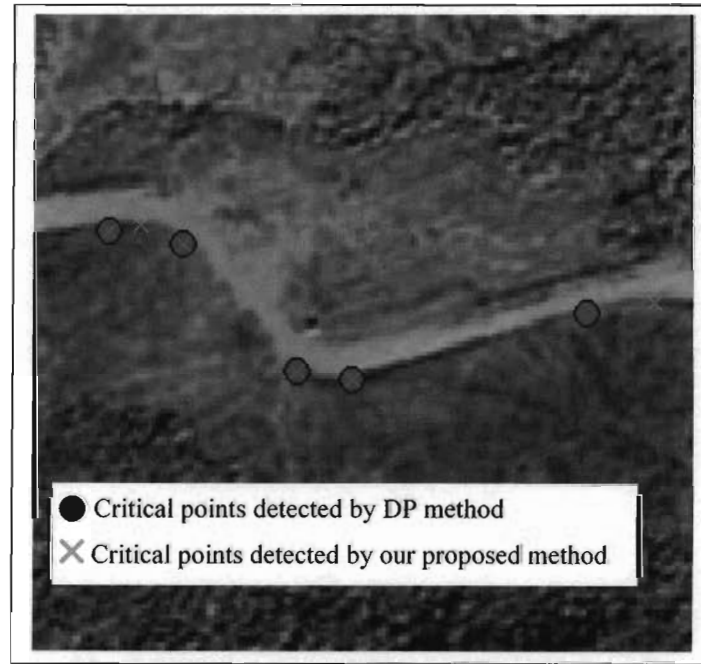
Figure 4.6: Detected critical points

Figure 4.6 shows the detection results respectively by DP algorithm and our method for the same point set. Due to different mechanism and tolerances, the identified critical points are not exactly coincident (e.g. the second point). However, our method detects the same crucial critical points as DP algorithm. As such, the polyline connecting critical points detected by our method describes the similar characteristics from DP algorithm, shown in Figure 4.7.
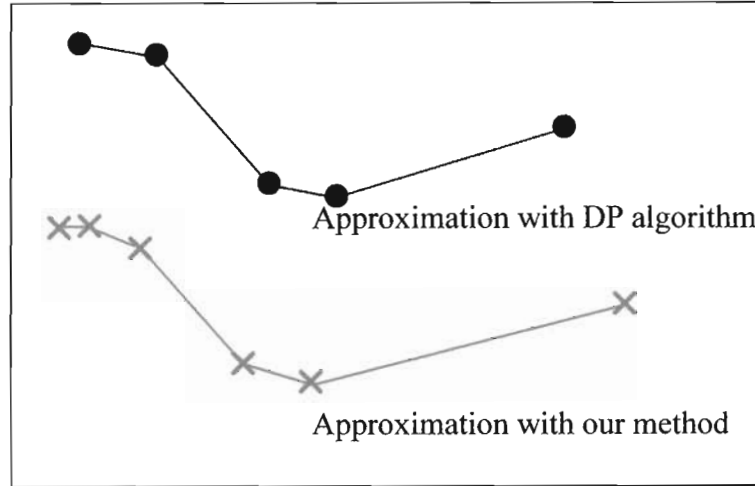
Figure 4.7: Approximating the line with detected critical points

4.1.2 Polygonic Template Construction

Through the detection of critical points a road is represented by a reduced set of nodes (critical points). One can intuitively consider comparing the two polylines connecting critical point sets from the involved images to find the correspondence between two road segments. However, as discussed in Chapter 3, there is no one-to-one correspondence in the extracted points due to image noise, viewing variations, and the nature of the snake algorithm. Some points may be detected as critical points in the right image (e.g., noise), while in the left image the points located on the corresponding places may not be the critical points, or vice versa. Under such circumstances, pursuing point-to-point matching of polylines would fail. In order to overcome this problem we make use of a polygonic template by connecting all extracted points within a pre-defined range (e.g. determined by certain critical points) from the right image. Thus a polygonic line connecting all the snake nodes within the range represents the dominant shape properties. By comparing the template from the right image with all the extracted points in the left image, the similarity between the corresponding road segments is assessed.

36

Users specify how many critical points are used to determine the range for snake nodes to construct the template. The intervention by users at this step is significant. It helps to define a range with little or even no noise and thus minimizes the noise effects. Nevertheless, using more critical points results in higher accuracies at the cost of higher computational requirements (and corresponding execution time).

4.1.3 Similarity Matching

Given the list of extracted points on the left image and the template generated from the right image, the objective of the feature matching is to find correspondences between these two collections of points, and thus match the corresponding road segments. This is performed by using a mathematical model to transform features from one image to the other and a similarity measure to form and evaluate the matches. Specifically, the mathematical model transforms the constructed template from the right image to the left one. The similarity measure detects the most similar part of the road on the left image with the transformed template. As we discussed previously, the involved images have rigid motion (i.e., translation and rotation) caused by differences in their acquisition conditions, e.g., exposure points, and direction variations of the UAVs at different instances. In order to preserve shapes and angles in the transformation for mosaicking, a linear conformal transformation (Equation 8) is used to transform the template to the left image. Thus, straight lines remain straight, and parallel lines are still parallel (Mathworks, 2000). Figure 4.8 shows the transformation between the two images given in Equation (8):

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\omega & -\sin\omega & X_T \\ \sin\omega & \cos\omega & Y_T \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (8)$$
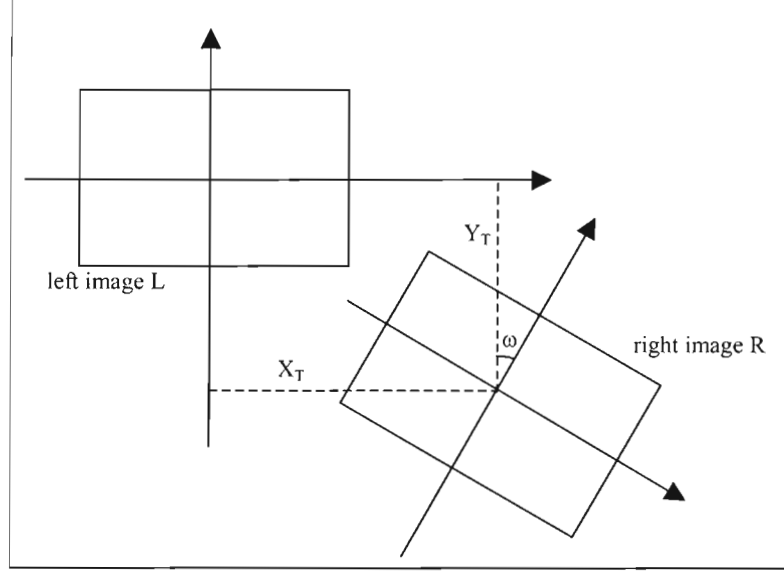
Figure 4.8: Geometric relationship of the image pair

Where X_T, Y_T are the translation values along X and Y axes respectively; ω is the relative rotation of the image pair.

Let subset M ($x_{m_i}, y_{m_i}, p \le i \le q$) denote the extracted points that constitute the template, where p and q are the sequence order numbers of M in the point set R_i. We convolve the template along the corresponding road edge on the left image. All road sections which have the same number of points on image L are considered to be the potential conjugates with the template. For each pair thus established, the distance D between them is calculated as the similarity measure. The smallest D indicates the match to the template. There are many combinations of road sections with the same point number. To reduce the burden of computation, the overlapping of photograph requirement is considered for the start of the convolution. Since the two images must

overlap on their border according to the aerial photogrammetry requirements, we start to convolve M from the border of the left image and end where the number of available extracted points is less than the dimension of subset M.

As shown in Figure 4.9, we move the template (M) as a unit to the left image with the relative translation between the last points M_q (e.g. the last point of the template) and L_{n_L} so that M_q overlap with L_{n_L}. Then M is rotated as a whole so that the direction of the vector $\overline{M_p M_q}$ is coincident to $\overrightarrow{L_{n_L} L_k}$ where $k = n_r - (q\text{-}p+1)$. The similarity measure D_{kn_L} between the template and the current matching candidate is calculated with the Euclidean distances of all involved points according to their sequence:

$$D_{kn_L} = \sum_{\substack{i=n_L \\ j=q}}^{\substack{i=k \\ j=p}} \sqrt{(x_{L_i} - x_{M_j})^2 + (y_{L_i} - y_{M_j})^2} \tag{9}$$

Then M moves to the next matching candidate starting at L_{n_L-1} and computes another similarity measure $D_{(k-1)(n_L-1)}$. As such the template is compared with all the matching candidates and has their correspondent similarity measures D_i ($q - p + 1 \le i \le kn_L$). The smallest D_i indicates the best matching.
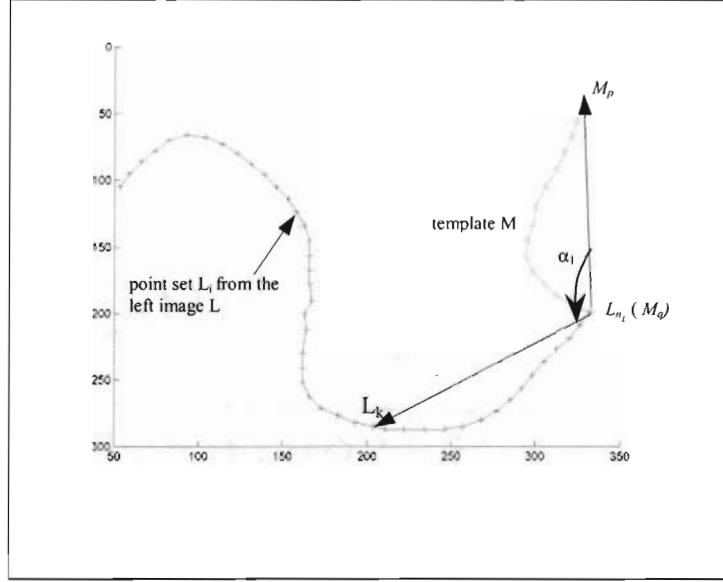
Figure 4.9: Detect the conjugate of the template on left image

After the conjugate of the template is detected, we proceed to match the road in order to acquire maximum point matching pairs for orientation determination in terms of the point order numbers in the sets. The points L_m (x_{L_m}, y_{L_m}) in image L are regarded as the conjugate of points R_n (x_{R_n}, y_{R_n}) in image R if the order number of L_m relative to the conjugate of M in point set L_i is the same as the order number of R_n relative to the template M in point set R_i. Thus, we detect the correspondence of the two point sets, i.e., the point matching, from the matching of the linear feature shape.

4.2 Orientation Parameter Determination and Mosaicking

Through this process, we determine a large number of conjugate points in each image pair. There are three unknown orientation parameters, i.e., translation coordinates X_T, Y_T and rotation angle ω shown in Figure 4.8. At least two conjugate points are needed to solve three unknowns. With more than two matched points in our approach, the optimal parameters are solved by a least squares adjustment in view of the redundancy

(Anderson et al., 1999). Linear conformal transformation is used to map the conjugate points in the adjustment process.

The acquired orientation parameters represent the relative status of the each image pair. The left image is regarded as the reference image. The right image is transformed with orientation parameters to the coordinate system of the left image by linear conformal transformation. It is recalled that we do not consider the radiometric transformation for mosaicking images in this thesis. Yet there exist radiometric differences in the mosaicked image. Thus it is not suitable to mosaic another image to the generated mosaic based on our presented method. To mosaic a sequence of images with our method, we design a strategy to overcome such a radiometric problem. Suppose there are five images in the image sequence. We construct the image pairs from the beginning to the end shown in Figure 4.10:

```
                        ┌─────────────────┐
                        │ Image sequence  │
                        └─────────────────┘
    ┌──────────┬──────────┬──────────┬──────────┐
┌────────┐ ┌────────┐ ┌────────┐ ┌────────┐ ┌────────┐
│1st     │ │2nd     │ │3rd     │ │4th     │ │5th     │
│image   │ │image   │ │image   │ │image   │ │image   │
└────────┘ └────────┘ └────────┘ └────────┘ └────────┘
```

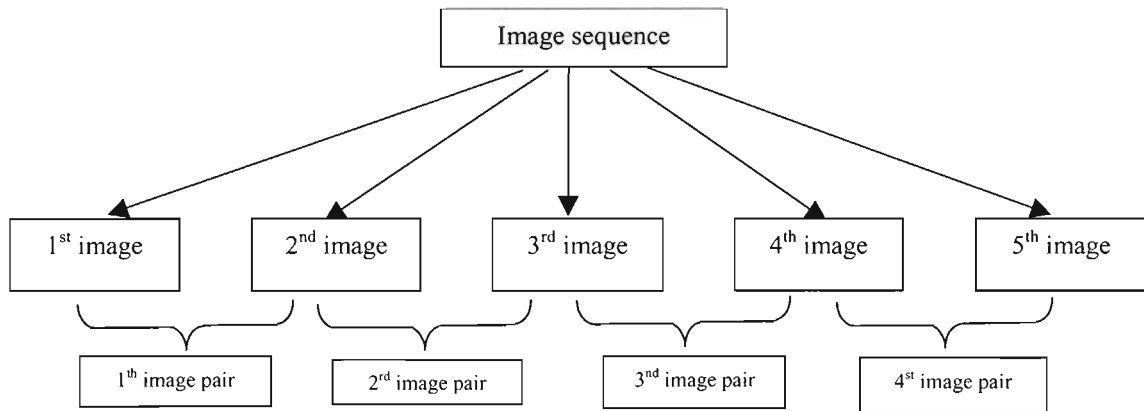| 1th image pair | 2rd image pair | 3nd image pair | 4st image pair |

Figure 4.10: Image pairs in the image sequence

In every image pair, the orientation parameters are calculated with the left image as the reference image. We refer to the parameters for each pair as IOPs (Image Orientation Parameters). The first image is regarded as anchor for the whole image sequence. For images that are not the reference images of involved pairs, we transform

41

them into their own reference image space first, then into their previous image pair. For reference images, they are transformed into their previous image pair directly. As such, we transform images other than the first image into the anchor space from their current image pair with IOPs. For example, the 4th image, the reference image in 4th pair, is transformed to the 4th image pair directly with the 4th IOPs, then to the 3rd image pair, 2nd image pair, until the 1st image pair (i.e., the anchor space)

Obviously, there are residuals generated by the matching accuracy, the least square calculation for the orientation parameters, etc. In our strategy for image sequence mosaicking, we use orientation parameters of each image pair to mosaic, thus propagating residuals from each image pair. As the number of image pairs increases, the residuals would become larger and may exceed the requirements of accuracy in certain applications. Thus when we mosaic large number of images in the sequence, we have to blend their radiometry to generate visually seamless image. Then the mosaic is used as one image to mosaic with the rest of the images by our proposed method. Since the radiometric transformation is out of the scope of this thesis, we focus on mosaicking sequence having few numbers of images. The next section will discuss this for practical use with experiments.

Chapter 5

EXPERIMENTAL STUDY

In this section, we demonstrate the performance of our algorithm introduced in chapters 3 and 4 to mosaic an aerial image sequence. The experiments have been implemented under the MATLAB (6.5 version) software environment. We simulated the acquisition of image sequences in digital aerial surveillance by selecting three partially overlapping sub-images from a larger aerial image (termed *original* image here) with a ground pixel resolution of 1m. The rotations and translations in this imagery are similar to the variations one would expect from a UAV flying over an area and collecting images. Thus our image datasets resemble UAV-type imagery, as the focus of this thesis is on mosaicking UAV image sequences. In Figure 5.1, we show our test images, as they partially overlap to represent the variation of exposure positions, while at the same time rotate about the vertical axis to demonstrate their acquisition at different camera poses. Image pairs are formed by considering two adjacent images at a time, resulting in two image pairs for our experiments. *Pair1* comprises images1 and 2; while *Pair2* comprises images2 and 3. Figure 5.2 illustrates the relative pose of image pairs in space prior to registration.

To support the analysis of our mosaicking, we manually identified six 'control points' on the original image (CPs, represented with crosses in Figure 5.1). CPs are evenly distributed. Each image in the sequence has its own unique control point (OCP),

i.e., a control point that is located outside the overlapping area of the image pair. The linear conformal transformation we used for mosaicking preserves object shapes (i.e., geometries including distances and angles) on imagery. In addition, we have ensured that other potential error sources (e.g. human errors while digitizing points on the screen, erroneous interpolations) have been kept at a minimum, to avoid potential contamination of our results. As such, we consider their effects on our mosaicking solution to be negligible. Thus, the changes of feature shapes and positions on images are due to errors from transformation parameters used to generate the mosaic.
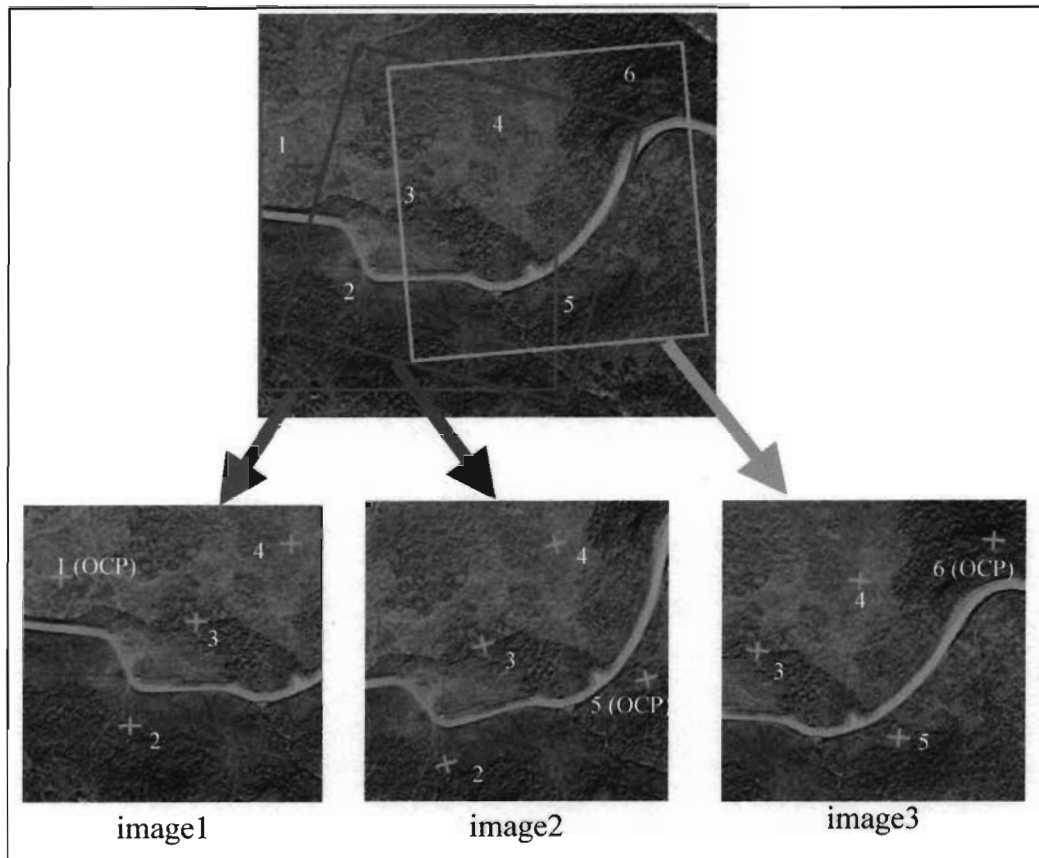
Figure 5.1: Generating image sequence with control points
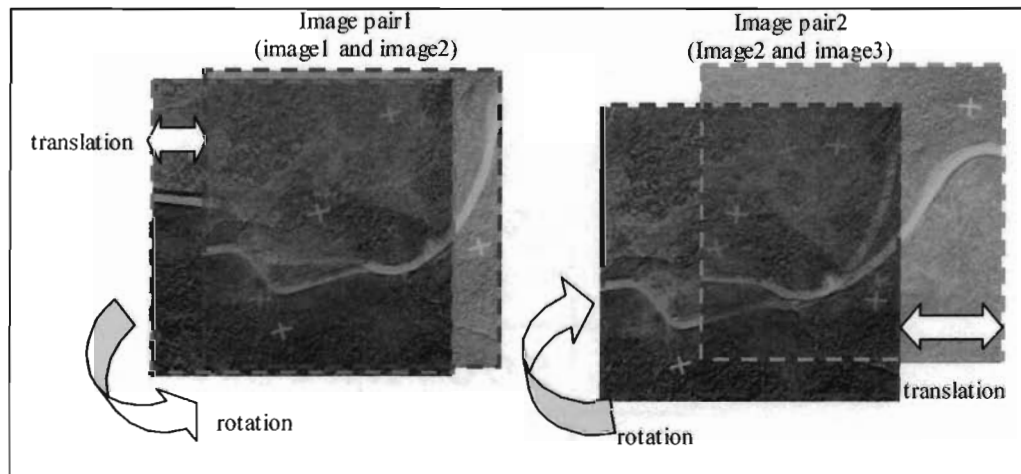
Figure 5.2: Geometric difference of images in the sequence

5.1 Experiment Implementation

For our experiment, the following processing steps were performed:

- The active deformable model (i.e., snake) has been applied on imagery to extract road segments on each image independently. Points and polylines connecting these points, are shown in Figure 5.3, and represent the extracted road segment.

Figure 5.3: Road extraction on image1

- On the right image of each pair, template construction was implemented, considering the geometric content of the road segment on that image. In view of the extent of road curvature, we selected a smaller threshold for the second pair. Figure 5.4 shows the parameters input by users and the templates constructed for each image pair using this input information.

Template for image pair1   Template for image pair2

Threshold for the curvature: 15 degree   Threshold for the curvature: 10 degree
Number of used Critical points: all   Number of used Critical points: all

Figure 5.4: Constructed templates

- Using the proposed similarity-matching algorithm, we detect the corresponding road segments for each image pair. Accordingly, extracted conjugate points are constructed. By observing these conjugate points, one can see that some of them do not correspond. However, such discrepancy is restricted under the shape of the road in the proposed approach so that its effect to mosaicking is relatively small. As the later result demonstrated, mosaicking imagery can be successfully accomplished under such discrepancies. The orientation parameters are solved using these correspondences in a least squares adjustment.

Image pair1
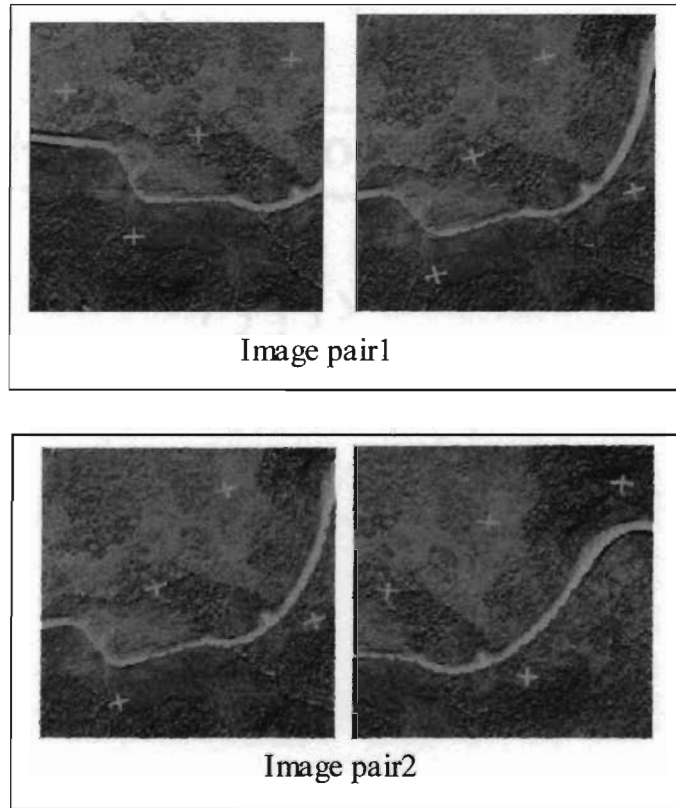
Image pair2

Figure 5.5: Conjugate road segment in image pairs

- Images were mosaicked in the following order: image3 was transformed to the space of image2 with the acquired orientation parameters of image pair2. Then, transformed image3 and original image2 were transformed into the space of image1 with the orientation parameters of image pair1. Image1 was then mosaicked with the transformed image2 and image3. Figure 5.6 shows the mosaicked result.
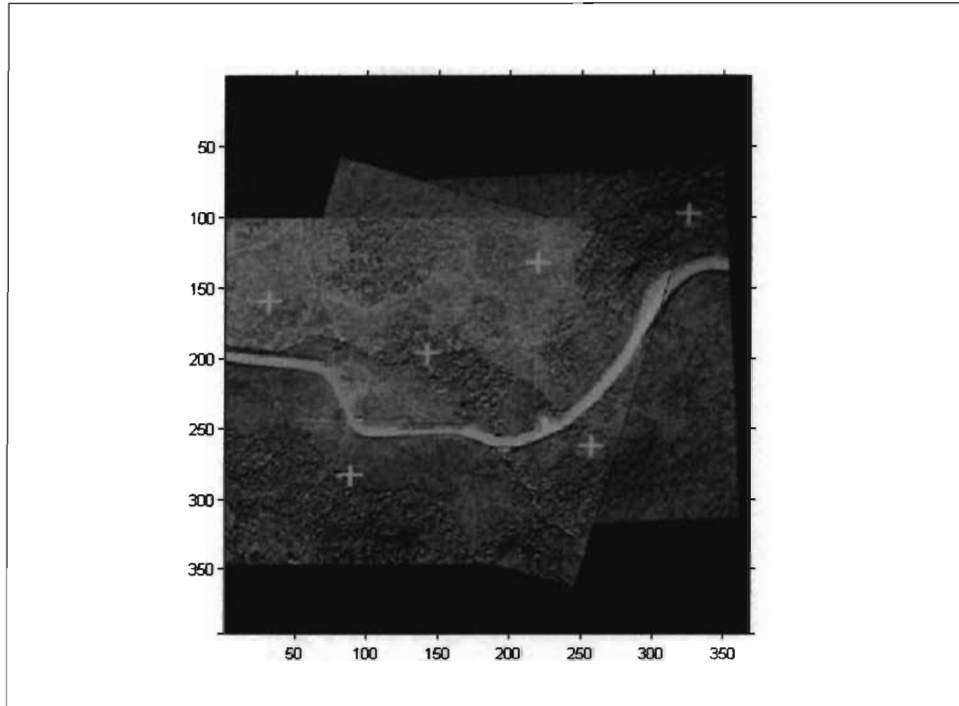
Figure 5.6: Mosaicked image sequence

5.2 Precision Study

By observing the mosaicked image, one can visually see the performance of our proposed algorithm. In order to further quantitatively study the result, we proceed on two accuracy aspects typically considered in geospatial applications: feature positions and feature shapes. We note here that the three sub-images in Figure 5.1 have rotation variations: image 2 reflects a clockwise rotation from image 1, while image3 reflects a counterclockwise rotation from image 2. In order to better investigate the effects of rotation variations we created another set of overlapping imagery (see Figure 5.7) with a different type of rotation between images 2 and 3. We examine accuracy aspects in those two situations, shown in Figure 5.1 and Figure 5.7.

49
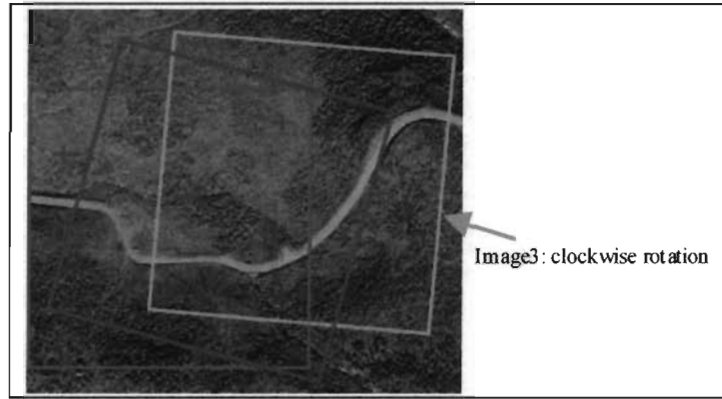
Figure 5.7: Second set of images

In aerial photogrammetry, points used to orient image pairs (named relative orientation points) are typically required to be located near the von Gruber locations in order to reach desirable accuracy. The objective of this practice is to select points that cover to the greatest possible extent the overlapping area (Figure 5.8). By matching points in these locations we optimize the potential accuracy of the orientation process, resulting in robust solutions. In our approach, we determine orientations by matching points on the edge of extracted linear features. It is easily understood that the location and topology of these points may deviate substantially from the von Gruber pattern, as one does not have control over the location and distribution of features in an image sequence. This is reflected in our test image selection: in pair 1 the common road segment is spanning only a small strip along the middle of the images, and is thus expected to produce less robust solutions. Conversely, in pair 2 the common part of the road segment spans a larger part of the overlapping area, and is expected to lead to more precise orientation.
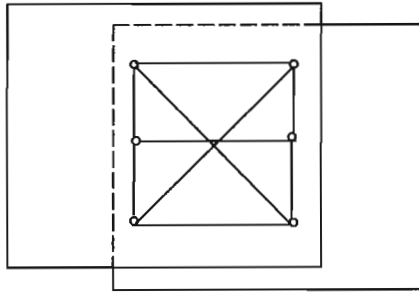
Figure 5.8: Von Gruber locations in an image pair

5.2.1 Position Accuracy Analysis

Positioning accuracy is one of the major considerations in geospatial applications. In the case of mosaicking, it is expressed by comparing the transformed location of a point to the position it should occupy. This is typically evaluated by using control points of known coordinates. In our case we used the CPs identified in Figure 5.1 to analyze the accuracy of our mosaic. Using the procedure presented in this thesis, images were mosaicked and we compared the transformed location of a CP in a transformed image of a stereopair to the location of the same point in the untransformed image of the same stereopair. In an ideal transformation these two locations should match. In Figure 5.9 we offer a visualization of errors resulting from our mosaicking solution.
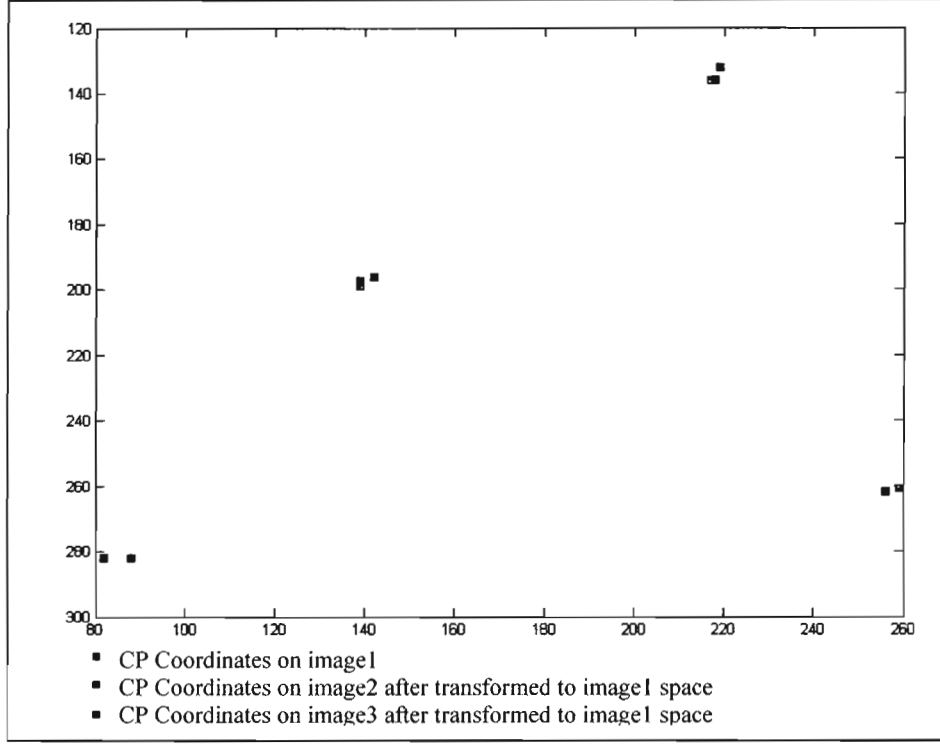
Figure 5.9: Deviations in CP mosaicking

We carried out position analysis in mosaicking with single image pairs (1-2 and 2-3) and two image pairs (1-2-3) and the results are summarized in Table 5.1. The position difference is described in terms of root mean square error (RMSE). As expected, RMSE increased with the number of image pairs, indicating errors accumulate when mosaicking longer image sequences. It is not surprising to see that RMSE in image pair2 is much smaller than that in image pair1 for single image pair mosaicking. As we argued previously, the distribution of linear features significantly affects the mosaicking accuracy. As shown in Figure 5.1, the distribution of linear features for image pair1 is relatively narrow and centers on the common area of the pair. However, the linear feature in image pair2 traverses almost from the lower left corner to the upper right corner in the common area of the pair. It covers much larger area than in image1 and image2 and

52

matches more closely the von Gruber locations, thus giving higher mosaicking accuracy for image pair2.

Table 5.1: Position difference of common CPs

		Common CP No.	$\lvert \Delta x \rvert$ (pixel)	$\lvert \Delta y \rvert$ (pixel)	RMSE $\sqrt{\dfrac{\Delta x^2 + \Delta y^2}{n}}$
Single image pair	Image1, Image2 (image1 space)	2	6	0	
		3	3	1	
		4	1	4	4.58
	Image2, image3 (image2 space)	3	2	1	
		4	0	1	
		5	1	2	1.91
Two image pairs	Image1, Image3 (image1space)	3	5	2	
		4	3	4	5.20

In order to consider the rotation direction effects, we implement the same experiment in the situation shown in Figure 5.7. Results are shown in Table 5.2. We can see that position accuracy is not significantly affected by rotation direction.
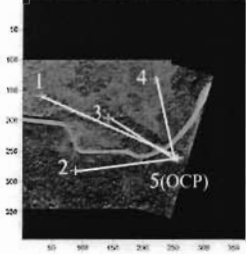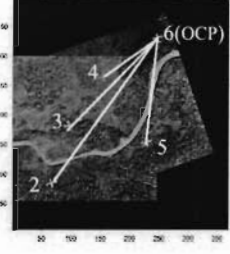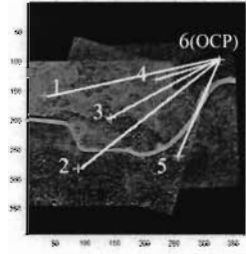
Table 5.2: Rotation direction study on position accuracy

		Common CPs No.	$\lvert \Delta x \rvert$ (pixel)	$\lvert \Delta y \rvert$ (pixel)	**Position RMS** $\sqrt{\dfrac{\Delta x^2 + \Delta y^2}{n}}$
One image pair	Image1, Image2 (image1 space)	2	6	0	
		3	3	1	
		4	1	4	4.58
	Image2, image3 (image2 space)	3	2	3	
		4	1	1	
		5	3	2	3.06
Two image pairs	Image1, Image3 (image1space)	3	3	3	
		4	2	4	4.36

5.2.2 Feature Shape Accuracy Analysis
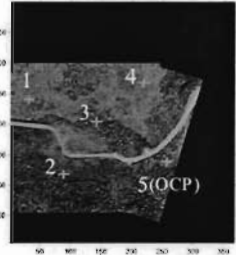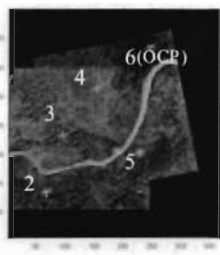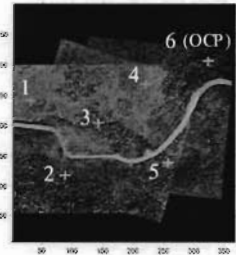
In geospatial applications, maintaining feature shape is another concern in evaluating algorithms. We perform the analysis with CP from reference image and OCPs in image pairs. As section 5.1.1, we consider single image pair and two image pair mosaicking. In each stage, OCP is transformed into the reference space. The errors of transformation parameters affect its new position, while the errors do not affect the CPs as they are digitized in the reference image. Demonstrated by figures in Table 5.3, by connecting CPs from the reference image and OCP from the to-be-mosaicked image we construct mosaicked lines and angles, which contain the mosaicking errors. Their true values can be regarded as measurements of corresponding lines and angles on the original image. RMSE are calculated from the difference to evaluate the precision of the result in each stage, as shown in Table 5.3.

Table 5.3: Distance and angle difference between mosaicked and original images

	Mosaicked image1 with transformed image2 in **image1** space			Mosaicked image2 with transformed image3 in **image2** space			Mosaicked image1 with transformed image2 and image3 in **image1** space		
	Points	Δdistance (Pixel)	Δangle (Degree)	Points	Δdistance (Pixel)	Δangle (Degree)	Points	Δdistance (Pixel)	Δangle (Degree)
	1,5	2.5772		2,6	2.9307		1,6	1.5893	
	2,5	4.6431	0.0556	3,6	1.0088	0.0671	2,6	1.419	0.0021
	3,5	2.5548	1.0682	4,6	1.8639	0.54219	3,6	2.583	0.5592
	4,5	3.2815	0.17586	5,6	1.868	0.30281	4,6	0.8742	0.7703
RMSE		3.3735	0.6259		2.0353	0.3606		1.7303	0.5496

According to the mosaicking steps, image3 has been transformed to image2 space first and then transformed to image1 space. It is assumed that the transformation errors for image3 should be accumulated from image2 space to image1 space. Nevertheless, we notice that the Δdistance RMSE value (1.7303pixels) of image3 in image1 space are smaller than that (2.0353pixels) in image2 space. The transformation errors have not been accumulated for the distance measurements. What if we consider the situation shown in Figure 5.7? Table 5.4 shows the analysis results.

Table 5.4: Distance and angle difference for mosaicked second sets of images

	Mosaicked image1 with transformed image2 in **image1** space			Mosaicked image2 with transformed image3 in **image2** space			Mosaicked image1 with transformed image2 and image3 in **image1** space		
	Points	Δdistance (Pixel)	Δangle (Degree)	Points	Δdistance (Pixel)	Δangle (Degree)	Points	Δdistance (Pixel)	Δangle (Degree)
	1,5	2.5772		2,6	1.4042		1,6	1.3418	
	2,5	4.6431	0.0556	3,6	0.3313	0.3086	2,6	3.1485	0.3827
	3,5	2.5548	1.0682	4,6	0.0755	0.1837	3,6	3.3613	0.3133
	4,5	3.2815	0.1759	5,6	2.4863	0.2817	4,6	2.7527	1.1636
RMSE		3.3725	0.6258		1.4378	0.2635		2.7654	0.7300

In Table 5.4, the RMS values of image3 in image1 space (2.7654pixels, 0.73degree) are bigger than those in image2 space (1.4378pixels, 0.2635degree). The transformation errors have been accumulated in this situation. Consequently, we can conclude that we may reduce error accumulation by acquiring images in opposite direction of rotation variation (i.e., clockwise or anti-clockwise).

For both situations shown in Table 5.3 and Table 5.4, errors are smaller in image pair 2 than those in image pair1 as the position errors due to the distribution difference of linear features.

Chapter 6

CONCLUSIONS

This thesis presented a new approach to automatically mosaic image sequences with natural linear features. This comprises determining the orientation difference between images and transforming them with the acquired parameters into a mosaic.

The methodology to determine the orientation parameters is a novel automated approach in the perspective of integrating points and linear feature geometry. By representing the roads with extracted points and polylines connecting them, a customized template can be constructed with detected critical points to characterize the shape of the roads. The customization enables users to reduce the blunders in detected critical points and therefore improves the matching reliability. The matching of two overlapping image pairs is implemented by measuring with Euclidean distance the geometric similarity between the template and the road shape on the to-be-matched imagery. The geometric constraint to matching minimizes the blunders, which are usually associated with point matching approach and take substantial effort to prune. Our approach takes advantage of point and geometry of linear features to result in a robust mosaicking of the image pairs. Furthermore, initial rotation variations are not required as prior information in our approach, which is critical in point matching in order to obtain convergence.

In the mosaicking process, we present a framework to mosaic image sequences with the orientation parameters of each image pair. Using the first image as an anchor,

the other images in the sequence are transformed to the first image space. By controlling the direction of orientation variations on purpose for successive images, the error propagation may be reduced. This allows us to mosaic several images at one time, improving the efficiency under a certain precision requirements.

The future work includes extending our core ideas and major results to solve more complicated orientation difference between images. One potential need in geospatial applications is to mosaic aerial imagery and the existed orthophoto maps over the same zone for change detection or updating the data in GIS database. There are six orientation parameters to represent the spatial relationship between the two images (please refer to the Exterior Orientation in Appendix B). By introducing the effects of these orientation differences to the geometry of linear features on images, the similarity measurement can be extended to find the correspondence under the effects. Furthermore, research in the future can be directed toward the goal of higher level of automation by minimizing the interaction with users.

REFERENCES

Ackermann, F., 1984. High Precision Digital Image Correlation. In: *The 39th Photogrammetric Week*, Vol. 9, Schriftenreihe der University Stuttgart, Stuttgart.

Agouris, P., Doucette, P. and Stefanidis, A., 2001a. Spatialspectral Cluster Analysis of Elongated Image Regions. In: *International Conference on Image Processing*, pp. 789-792.

Agouris, P. and Schenk, T., 1996. Automated Aerotriangulation Using Multiple Image Multipoint Matching. *Photogrammetric Engineering and Remote Sensing*, **62**(6), pp. 703-710.

Agouris, P., Stefanidis, A. and Gyftakis, S., 2001b. Differential Snakes for Change Detection in Road Segments. *Photogrammetric Engineering and Remote Sensing*, **67**(12), pp. 1391-1399.

Alhichri, H. S. and Kamel, M., 2003. Virtual Circles: A New Set of Features for Fast Image Registration. *Pattern Recognition Letters*, **24**, pp. 1181-1190.

Alt, H., Mehlhorn, K. and Wagener, H., 1988. Congruence, Similarity and Symmetries of Geometric Objects. *Discrete and Computational Geometry*, **3**, pp. 237-256.

Althof, R. J., Wind, M. G. J. and Dobbins, J. T., 1997. A Rapid and Automatic Image Registration Algorithm with Subpixel Accuracy. *IEEE Transactions on Medical Imaging*, **16**, pp. 308-316.

Amini, A., Weymouth, T. and Jain, R., 1990. Using Dynamic Programming for Solving Variational Problems in Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**(9), pp. 855-867.

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Croz, J. D., Greenbaum, A., Hammarling, S., McKenney, A. and Sorensen, D., 1999. LAPACK User's Guide (Third Edition), SIAM, Philadelphia.

Baltsavias, E. P., Gruen, A. and van Gool, L., 2001. Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), A.A. Balkema.

Basu, M., 2002. Gaussian-based Edge-detection Methods - a Survey. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, **32**(3), pp. 252-260.

Baumgartner, A., Steger, C., Mayer, H. and Eckstein, W., 1997. Semantic Objects and Context for Finding Roads. In: *SPIE*, Vol. 3072, pp. 98-109, Orlando, Florida.

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W. and Ebner, H., 1999. Automatic Extraction Based on Multi-scale, Grouping, and Context. *Photogrammetric Engineering and Remote Sensing*, **65**(7), pp. 777-785.

Belongie, S., Malik, J. and Puzicha, J., 2002. Shape Matching and Object Recognition Using Shape Contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(24), pp. 509-522.

Bergen, J. R., Anandan, P., Hanna, K. J. and Hingorani, R., 1992. Hierarchical Model-Based Motion Estimation. In: *Computer Vision ECCV' 92*, pp. 237-252, Santa Margherita Ligure, Italy.

Berger, M., 1998. The Framework of Least Squares Template Matching. Technical Report 180, Image Science Lab, ETH Zfirich.

Bhatacharya, D. and Sinha, S., 1997. Invariance of Stereo Images via Theory of Complex Moments. *Pattern Recognition*, **30**, pp. 1373-1386.

Biederman, I., 1987. Recognition-by-components: a Theory of Human Image Understanding. *Psychological Review*, **94**(2), pp. 115-147.

Burt, P. J. and Adelson, E. H., 1983. A Multiresolution Spline with Application to Image Mosaics. *ACM Trans. Graphics*, **2**(4), pp. 217-236.

Capel, D. and Zisserman, A., 1998. Automated Mosaicing with Super-resolution Zoom. In: *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 885--891, Santa Barbara.

Chen, S. E., 1995. QuickTime VR - an Image-based Approach to Virtual Environment Navigation. *Computer Graphics (SIGGRAPH'93)*, **8**, pp. 29-38.

Cho, W., 1995. Relational Matching for Automatic Orientation. Ph.D. thesis, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio.

Cho, W., 1996. Relational Matching for Automatic Orientation. *International Archives of Photogrammetry and Remote Sensing*, **31**(B3), pp. 111-119.

Danuser, G., 1996. Stereo Light Microscope Calibration for 3D Submicron Vision. In: *Proceedings of the 18th ISPRS Congress*, 31/B5, pp. 101-108, Vienna, Austria.

David, M., 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, Freeman, San Francisco.

David, P., DeMenthon, D., Duraiswami, R. and Samet, H., 2003. Simultaneous Pose and Correspondence Determination Using Line Features. In: *Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, pp. 424-431.

Deriche, R. and Giraudon, G., 1993. A Computational Approach for Corner and Vertex Detection. *International Journal of Computer Vision*, **10**(2), pp. 101-124.

Doucette, P. J., 2002. Automated Road Extraction from Aerial Imagery by Self-Organization. Ph.D. thesis, Spatial Information Science and Engineering, University of Maine, Orono.

Douglas, D. H. and Peucker, T. K., 1973. Algorithms for the Reduction of the Number of Points Required to Represent a Line or Its Caricature. *The Canadian Cartographer*, **10**(2), pp. 112-122.

Doyle, F., 1964. The Historical Development of Analytical Photogrammetry. *Photogrammetric Engineering*, **XXX**(2), pp. 259-265.

Dreschler, L. and Nagel, H., 1981. Volumetric Model and 3-D Trajectory of a Moving Car Derived from Monocular TV-frame Sequence of a Street Scene. In: *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 692-697, Vancouver, Canada.

Drewniok, C. and Rohr, K., 1996. Automatic Exterior Orientation of Aerial Images in Urban Environment. *International Archives of Photogrammetry and Remote Sensing*, **30**(B3), pp. 146-152.

Ebisch, K., 2002. A Correction to the Douglas-Peucker Line Generalization Algorithm. *Computers & Geosciences*, **28**(8), pp. 995-997.

Falkner, E. and Morgan, D., 2002. *Aerial Mapping: Methods and Applications* (2nd edition), Lewis Publishers, Boca Raton, Florida.

Fischler, M. and Heller, A., 1998. Automated Techniques for Road Network Modeling. In: *DARPA Image Understanding Workshop*, pp. 501-516.

Flusser, J. and Suk, T., 1994. A Moment-based Approach to Registration of Images with Affine Geometric Distortion. *IEEE Transactions on Geoscience and Remote Sensing*, **32**, pp. 382-397.

Fonseca, L. M. G. and Manjunath, B. S., 1996. Registration Techniques for Multisensor remotely sensed Imagery. *Photogrammetric Engineering and Remote Sensing*, **62**, pp. 1049-1056.

Forstner, W., 1982. On the Geometric Precision of Digital Correlation. *International Archives of Photogrammetry and Remote Sensing*, **24**(3), pp. 176-189.

Forstner, W., 1986. A Feature Based Correspondence Algorithm for Image Matching. *International Archives of Photogrammetry and Remote Sensing*, **26**(3/3), pp. 150-166.

Garcia, R., Batlle, J., Cufi, X. and Amat, J., 2001. Positioning an Underwater Vehicle through Image Mosaicking. In: *IEEE International Conference on Robitics and Automation*, 3, pp. 2779 - 2784.

Gong, Y., Proietti, G. and LaRose, D., 1999. A Robust Image Mosaicing Technique Capable of Creating Integrated Panoramas. In: *IEEE International Conference on Information Visualization*, pp. 24-29, London.

Gracias, N., 2003. Mosaic-Based Visual Navigation for Autonomous Underwater Vehicles. Ph.D. thesis, Instituto Superior Tecnico, Lisbon, Portugal.

Gracias, N. R., van der Zwaan, S., Bernardino, A. and Santos-Victor, J., 2003. Mosaic-Based Navigation for Automous Underwater Vehicles. *IEEE Journal of Oceanic Engineering*, **28**(4), pp. 609-624.

Gruen, A., Agouris, P. and Li, H., 1995a. Linear Feature Extraction with Dynamic Programming and Globally Enforced Least Squares Matching. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhauser Verlag, Basel, pp. 83-94.

Gruen, A., Baltsavias, E. P. and Henricsson, O., 1997. Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Springer Verlag.

Gruen, A., Kuebler, O. and Agouris, P., 1995b. Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhauser Verlag, Basel.

Haala, B., Hahn, M. and Schmidt, D., 1993. Quality and Performance Analysis of Automatic Relative Orientation. In: *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision, SPIE Proceedings (1944)*, pp. 140-150.

Hahn, M. and Kiefner, M., 1994. Relative Orientierung durch Digitale Bildzuordnung. *Z. Photogramm. Fernerkundung*, **62**(6), pp. 223-228.

Hannah, M. J., 1980. Bootstrap Stereo. In: *Proceedings Image Understanding Workshop*, pp. 201-208.

Hannah, M. J., 1989. A System for Digital Stereo Image Matching. *Photogrammetric Engineering and Remote Sensing*, **55**(12), pp. 1765-1770.

Hansen, M., Anandan, P., Dana, K., van der Wal, G. and Burt, P. J., 1994. Real-Time Scene Stabilization and Mosaic Construction. In: *Proceedings of DARPA Image Understanding Workshop' 94*, pp. 457-465.

Heffernan, P. J. and Schirra, S., 1994. Approximate Decision Algorithms for Point Set Congruence. *Computational Geometry Theory and Applications*, **4**, pp. 137-156.

Heipke, C., 1995. State-of-the-Art of Digital Photogrammetric Workstations for Topographic Applications. *Photogrammetric Engineering and Remote Sensing*, **61**(1), pp. 49-56.

Heipke, C., 1996. Overview of Image Matching Techniques. In: *OEEPE Workshop on the Application of Digital Photogrammetric Workstation*, Lausanne.

Heipke, C., 1997. Automation of Interior, Relative, and Absolute Orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, **52**, pp. 1-19.

Hobrough, G. L., 1959. Automatic Stereo Plotting. *Photogrammetric Engineering*, **25**(5), pp. 763-769.

Horn, B. K. P., 1989. Relative Orientation. Memo 994-A, Artificial Intelligence Laboratory, MIT, Cambridge, Massachusetts.

Hummel, J. E. and Biederman, I., 1992. Dynamic Binding in a Neural Network for Shape Recognition. *Psychological Review*, **99**, pp. 480-517.

Irani, M., Anandan, P. and Hsu, S., 1995. Mosaic Based Representations of Video Sequences and Their Applications. In: *5th International Conference of Computer Vision*, pp. 605-611.

Jacobs, C., Finkelstein, A. and Salesin, D., 1995. Fast Multiresolution Image Querying. In: *Computer Graphics Proceedings SIGGRAPH*, pp. 277-286.

Jin, W., Yang, X., Zhao, H. and Cui, R., 2001. Photogrammetry, Wuhan University publisher, Wuhan, P.R.China.

Jones, M. and Oakley, J. P., 2000. Registration of Image Sets Using Silhouette Consistency. *IEEE Proceedings Vision, Image and Signal Processing*, **147**(1), pp. 1-8.

Kanazawa, Y. and Kanataniy, K., 2002. Image Mosaicing by Stratified Matching. In: *Proc. Statistical Methods in Video Processing Workshop*, pp. 31-36, Copenhagen, Denmark.

Kass, M., Witkin, A. and Terzopoulos, D., 1987. Snakes: Active Contour Models. In: *1st International Conference on Computer Vision*, pp. 259-268, London.

Katartzis, A., Sahli, H., Pizurica, V. and Cornelis, J., 2001. A Model-based Approach to the Automatic Extraction of Linear Features from Airborne Images. *IEEE Transactions on Geoscience and Remote Sensing*, **39**(9), pp. 2073-2079.

Kitchen, L. and Rosenfeld, A., 1982. Gray-level Corner Detection. *Pattern Recognition Letters 1*, pp. 95-102.

Konecny, G., 1985. The International Society for Photogrammetry and Remote Sensing - 75 Years Old, or 75 Years Young. *Photogrammetric Engineering and Remote Sensing*, **51**(7), pp. 919-933.

Krishnan, A. and Ahuja, N., 1996. Panoramic Image Acquisition. In: *Proceedings of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 379-384.

Liang, T. and Heipke, C., 1996. Automatic Relative Orientation of Aerial Images. *Photogrammetric Engineering and Remote Sensing*, **62**(1), pp. 47-55.

Litton, A., 1993. Automated Data Capture: the Scanning Solution. In: *Proceedings of ARC/INFO User Conference*.

Loncaric, S., 1998. A Survey of Shape Analysis Techniques. *Pattern Recognition*, **31**(8), pp. 983-1001.

Mahajan, S. K. and Singh, V., 1972. Comparison of Analytical Relative-orientation Methods. *Surveying and Mapping Division*, pp. 73-85.

Mallick, S. P., 2002. Feature Based Image Mosaicing. http://www.cs.ucsd.edu/classes/fa02/cse252c/smallick.pdf

Marino, J. S., 1979. Identification of Characteristic Points along Naturally Occurring Lines: An Empirical Study. *The Canadian Cartographer*, **16**, pp. 70-80.

Mathworks, T., 2000. Image Processing Toolbox User's Guide. Natick, Massachusetts.

McMaster, R. B., 1986. A Statistical Analysis of Mathematical Measures for Linear Simplification. *American Cartographer*, **13**, pp. 103-116.

Milgram, D. L., 1975. Computer Methods for Creating Photomosaics. *IEEE Trans. Computers*, **24**, pp. 1113-1119.

Milgram, D. L., 1977. Adaptive Techniques for Photomosaicing. *IEEE Trans. Computers*, **26**, pp. 1175-1180.

Moffitt, F. H. and Mikhail, E. M., 1980. Photogrammetry (3rd edition), Harper & Row, Publishers, New York.

Moravec, H. P., 1977. Towards Automatic Visual Obstacle Avoidance. In: *Proc. 5th Int. Joint Conference on Artificial Intelligence*, pp. 584, MIT, Cambridge, MA, USA.

Muller, B. and Hahn, M., 1992. Parallel Processing - the Example of Automatic Relative Orientation. *International Archives of Photogrammetry and Remote Sensing*, **29**(B2), pp. 623-630.

Oddo, L., Doucette, P. and Agouris, P., 2000. Automated Road Extraction via the Hybridization of Self-organization and Model Based Techniques. In: *29th Applied Imagery Pattern Recognition Workshop, IEEE Computer Society*, Washington, D.C.

Park, S.-R. and King, T., 2001. Semi-automatic Road Extraction Algorithm from IKONOS Images Using Template Matching. In: *the 22nd Asian Conference on Remote Sensing*, Singapore.

Peleg, S., 1981. Elimination of Seams from Photomosaics. *Computer Graphics and Image Processing*, **16**, pp. 90-94.

Peteri, R., Celle, J. and Ranchin, T., 2003. Detection and Extraction of Road Networks from High Resolution Satellite Images. In: *International Conference on Image Processing*, 1, pp. 14-17.

Rohr, K., 2001. Landmark-based Image Analysis: Using Geometric and Intensity Models. Kluwer Academic Publishers.

Rottensteiner, F., 1993. Area-based Matching of Fiducial Marks in Scanned Images. In: *Image Analysis and Synthesis, Proceedings of the 17th OeAGM Workshop*, volume 68 of Schriftenreihe der OCG, pp. 163-172, R. Oldenbourg Verlag, Vienna.

Sailor, S., 1965. Demonstration Board for Stereoscopic Plotter Orientation. *Photogrammetric Engineering*, **31**(1), pp. 176-179.

Sawhney, H. S., Ayer, S. and Gorkani, M., 1995. Model-based 2D and 3D Dominant Motion Estimation for Mosaicing and Video Representation. In: *Proceedings of Fifth Int'l Conf. Computer Vision*, pp. 583-590.

Schenk, T., 1999. Digital Photogrammetry, TerraScience, Laurelville, Ohio.

Schenk, T., Li, J.-C. and Toth, C., 1991. Towards an Autonomous System for Orienting Digital Stereopairs. *Photogrammetric Engineering and Remote Sensing*, **57**(8), pp. 1057-1064.

Schenk, T. and Toth, C., 1993. Towards a Fully Automated Aerotriangulation System. In: *Proceedngs ACSM/ASPRS Annual Spring Convention*, 3, pp. 340-347, New Orleans.

Sclaroff, S. and Pentland, A. P., 1995. Modal Matching for Correspondence and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**(6), pp. 545-561.

Shapiro, L. G. and Haralick, R. M., 1987. Relational Matching. *Applied Optics*, **26**(10), pp. 1845-1851.

Shum, H. Y. and Szeliski, R., 2000. System and Experiment Paper: Construction of Panoramic Image Mosaics with Global and Local Alignment. *International Journal of Computer Vision*, **36**(2), pp. 101-130.

Small, C. G., 1996. The Statistical Theory of Shapes. Springer-Verlag, New York.

Su, M.-S., Hwang, W.-L. and Cheng, K.-Y., 2004. Analysis on Multiresolution Mosaic Images. *IEEE Transactions on Image Processing*, **13**(7), pp. 952-959.

Takeuchi, S., Terashima, N. and Tominaga, H., 1999. Image Mosaicing without Distortion Using Projected Mask for Image Digitization. In: *Proceedings International Conference on Image Processing*, 4, pp. 123-127, Tokyo, Japan

Tang, L. and Heipke, C., 1996. Automatic Relative Orientation of Aerial Images. *Photogrammetric Engineering and Remote Sensing*, **62**(1), pp. 47-55.

Trinder, J. and Li, H., 1995. Semi-automatic Feature Extraction by Snakes. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhauser Verlag, Basel, pp. 95-104.

Tupin, F., Houshmand, B. and Datcu, M., 2002. Road Detection in Dense Urban Areas Using SAR Imagery and the Usefulness of Multiple Views. *IEEE Transactions on Geoscience and Remote Sensing*, **40**(11), pp. 2405-2414.

Umeyama, S., 1993. Parameterized Point Pattern Matching and Its Application to Recognition of Object Families. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**(1), pp. 136-144.

Veltkamp, R. C. and Hagedoorn, M., 1999. State-of-the-Art in Shape Matching. Technical Report UU-CS-1999-27, Utrecht University, Netherlands, Utrecht.

Vosselman, G., 1992. Relational Matching. *Lecture Notes in Computer Science*, vol. 628, Springer, Berlin.

Vosselman, G., 1995. Applications of Tree Search Methods in Digital Photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, **50**(4), pp. 29-37.

Vosselman, G. and Knecht, J., 1995. Road Tracing by Profile Matching and Kalman Filtering. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhauser Verlag, Basel, pp. 265-274.

Wang, Y., 1996. Structural Matching and Its Applications for Photogrammetric Automation. *International Archives of Photogrammetry and Remote Sensing*, **31**(B3), pp. 918-923.

Williams, D. J. and Shah, M., 1992. A Fast Algorithm for Active Contours and Curvature Estimation. *CVGIP: Image Understanding*, **55**(1), pp. 14-26.

Zaharan, M., 1997. Shape-Based Multi-Image Matching Using the Principles of Generalized Hough Transform. Ph.D. dissertation, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio.

Zhang, C. and Baltsavias, E., 2000. Knowledge-based Image Analysis for 3D Edge Extraction and Road Reconstruction. *International Archives of Photogrammetry and Remote Sensing*, **33**(B3/2), pp. 1008-1015.

Zitova, B. and Flusser, J., 2003. Image Registration Methods: a Survey. *Image and Vision Computing*, **21**(11), pp. 977-1000.

Zoghiami, I., Faugeras, O. and Deriche, R., 1997. Using Geometric Corners to Build a 2D Mosaic from a Set of Images. In: *Proceedings of CVPR*, pp. 420-425, Puerto Rico.

APPENDICES

Appendix A.

A Review of Relevant Basic Photogrammetric Principles

As the definition by the American Society for Photogrammetry and Remote Sensing (ASPRS), photogrammetry is 'the art, science, and technology of obtaining reliable information about physical objects and the environment, through processes of recording, measuring and interpreting images and patterns of electromagnetic radiant energy and other phenomena' (Falkner and Morgan, 2002). Invented in 1851 by Laussedat, photogrammetry has advanced from analogue to analytical, and to current digital photogrammetry with the improvements of the computational power available on cheap machines and the availability of digital techniques. High-resolution digital cameras with CCD (Charge-Coupled Device) sensors are used to record an image as a matrix of pixels along with a computer data storage device to record a group of image data sets. Furthermore, a desktop computer, stereo viewing glasses and sophisticated softwares take the place of the large and expensive stereo-comparators and analytical machines in digital aerial photogrammetry. The use of digital imagery and numerical techniques allow us to automate more and more tasks, even if a large part of the automation is still at a research level.

Principally photogrammetry is divided into terrestrial photogrammetry and aerial photogrammetry. Each serves the needs of users from distinct categories. Terrestrial photogrammetry typically satisfies the needs of architects, civil engineers (to supervise buildings, document their current state, deformation or damages), archaeologists, surgeons (plastic surgery), etc. Aerial photogrammetry is often used in mapping community with an emphasis on obtaining quantitative information from aerial

71

photographs. The photographs of the interested area in aerial photogrammetry are acquired with a metric camera mounted in an aircraft flying over the area in an orderly sequence. A metric camera is one in which focal length and internal dimensions are exactly known or can be determined through calibration. The proposed method introduced later tries to solve the challenging issue within aerial photogrammetry area. Therefore, in this thesis, we focus our attention on aerial photogrammetry environment.

In the following sections, we present an overview of basic photogrammetric principles, relevant to the discussion in this thesis.

A.1. Image Formation

Photogrammetric operations typically proceed in two distinct steps: imagery is oriented in order to determine its location and pose in a georeferenced coordinate system, and then analyzed in order to measure the location of objects in it. Orientation information allows us to transfer these image measurements into a survey coordinate system, thus populating geospatial database (e.g. maps, GIS layers). Photogrammetric applications typically assume that imagery was captured through a perspective projection (Figure A.1). In a perspective projection all rays connecting image points to their object space counterparts pass through a common point in the camera lens, the perspective center. The CCD (or films) serves as the reference *focal* plane and the image is captured behind the perspective center as shown in Fig.1. In terms of notation, the area between the perspective center and the CCD (or film) is called the *image space*. The area in front of the lens is the *object space*. The Geometrical axis of the lens system is the *optical axis*, which is perpendicular to the focal plane of the camera. The *Principal point* (PP) is the intersection of optical axis and focal plane. Generally when discussing the image space, it

is convenient to use the positive position of the photograph in front of the perspective center instead of the photograph negative position (Moffitt and Mikhail, 1980).
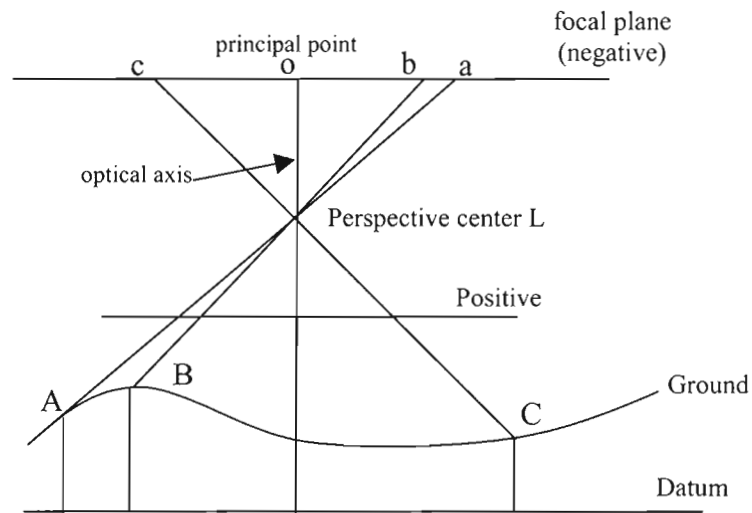

Figure A.1: Perspective projection

A.2. The Overlapping of acquired photographs

Aerial photographs are used to form stereoscopic view that enables us to make photographic maps, determine elevation of terrain points without ever setting foot in the field. Therefore, photos have to be taken as pairs, called *stereo-pairs*, with overlapping coverage of the scene photographed. Flight planning guarantees these requirements to be satisfied when taking the photographs. The basic elements in flight planning are the flying height above a datum, normally sea level, which determines the scale of photographs; the ground distance between successive exposures; and the ground spacing between the flight lines (Moffitt and Mikhail, 1980). Those elements are determined by the purpose of aerial photogrammetry, the survey block, etc. *Endlap* and *sidelap* are two significant features between photographs. *Endlap*, also called *forward overlap*, is the overlapping area between consecutive images along a flight strip. It creates the three-

dimensional effect necessary for mapping. Normally the average endlap covers 60% of the previous photographs. The ground spacing between adjacent flight lines will contribute to the amount of overlapping areas of images obtained in adjacent flight lines. It is called *sidelap*, sometimes side overlap and serves that there are no gaps in the three-dimensional coverage of a multiline project. Usually sidelap ranges between 20% and 40% of the width of a photo, with a nominal average of 30%. The shadows in Figure A.2 represent the endlap area on the adjacent photos taken along a flight line and the sidelap between flight strips.
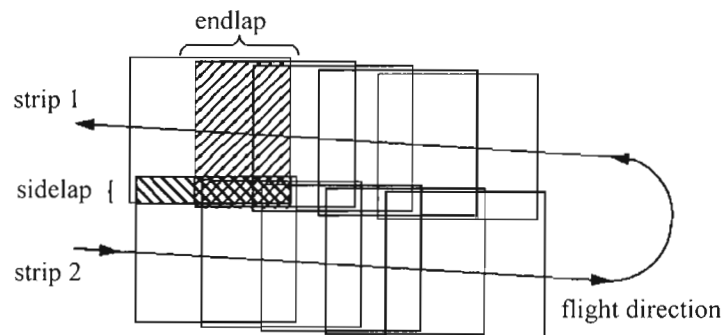

Figure A.2: A regular block of aerial photos

A.3.   Plotting Principles in Aerial Photogrammetry

In aerial photogrammetry, the widely applied techniques to taking 3D measurements off photographs are based on the geometry of perspective scenes and on the principles of stereovision. *Stereoscopy* is a term to describe the following phenomenon: When a viewer observes two photographs of the same scene taken from two different viewpoints, the viewer can visualize the depicted scene in three dimensions. This principle is demonstrated in Figure A.3, where S_1 and S_2 denote our left and right eyes. When both eyes gaze at point A, the lines joining the point A and the two eyes form

74

an angle φ_1 called parallactic angle. Similarly, the parallax angle for point B is φ_2. Points A and B forms a_1 and b_1 in the left eye, a_2 and b_2 in the right eye. The viewer's brain interprets the difference between a_1b_1 and a_2b_2 as the difference in the two angels (φ_1, φ_2) and forms a spatial construction of the scene. Point B is closer than point A.
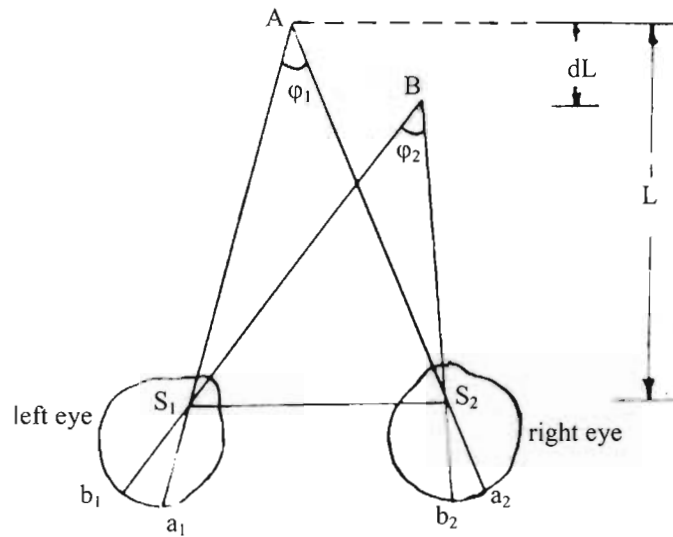

Figure A.3: Stereoscopic view by human eyes

If we locate a piece of glass in front of each eyes as P_1 and P_2 shows in Figure A.4. We record the images on the glass as a_1, b_1 and a_2, b_2. Then move away the objects (A, B). By observing the images on the glass, our eyes can still intersect the spatial positions where the real objects A and B locate.
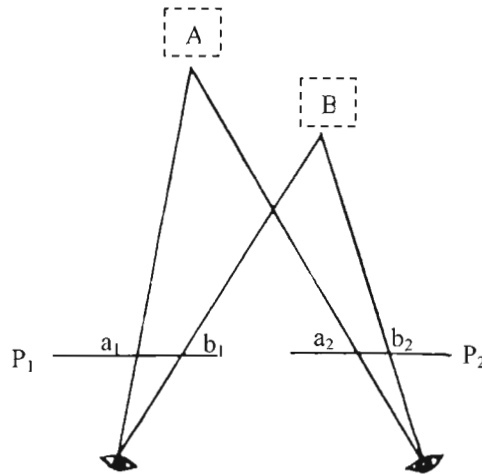
Figure A.4: Man-made stereoscopic view

In aerial photogrammetry, we apply the foregoing stereoscopic principles to measure the 3D information of ground objects off photographs. After acquiring photographs of the same objects from different viewpoints that act like our two eyes, the first essential work is to orientate images, which defines the status and location of images relative to the ground when they are taken. Afterwards, the stereo-space model is constructed as the real world by observing the overlapped images defined by the former step. Then measurements can be fulfilled for the ground objects through this model.

A.4. Coordinate systems in photogrammetry

Different coordinate systems are defined to facilitate describing the positions of objects on photographs and the real world. The next parts briefly introduce the definitions of common used coordinate systems in aerial photogrammetry: *image coordinate system, image space coordinate system* and *ground coordinate system.*

The image coordinate system (x, y) is defined by four fiducial marks. These marks are small permanent marks located in the middle of the sides of the focal plane opening or on its corners, or in both locations. They are exposed onto the negative when a

photo is taken. Their positions relative to the camera body are calibrated. Thus, they define the image coordinate system with respect to the camera. The x and y axes are defined by the opposite fiducial marks, with the fiducial center taken as the origin, shown in Figure A.5.
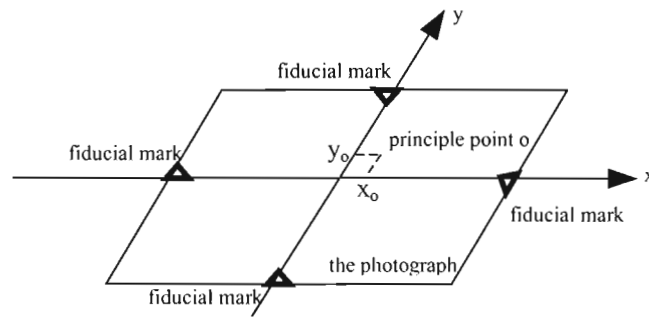

Figure A.5: Image coordinate system

Image Space Coordinate System (x, y, z) is similar to the image coordinate system except that it adds a third axis (z) as Figure A.6 shows. The origin of image coordinate system is located at the perspective center (L). The projection of the perspective point that is vertical to the photograph is called principle point. Normally the x-y axes are parallel to the axes in the image coordinate system, but sometimes they can be defined as needed.
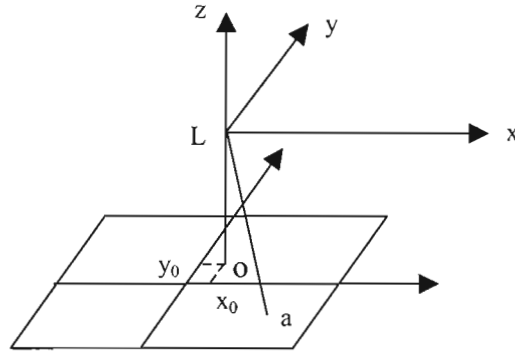
Figure A.6: Image space coordinate system

As Figure A.7 shows, generally the ground coordinate system is defined as a three-dimensional, right-handed Cartesian coordinate system (X, Y, Z), which utilizes a map projection. The Z value is the elevation above the mean sea for a given vertical datum.
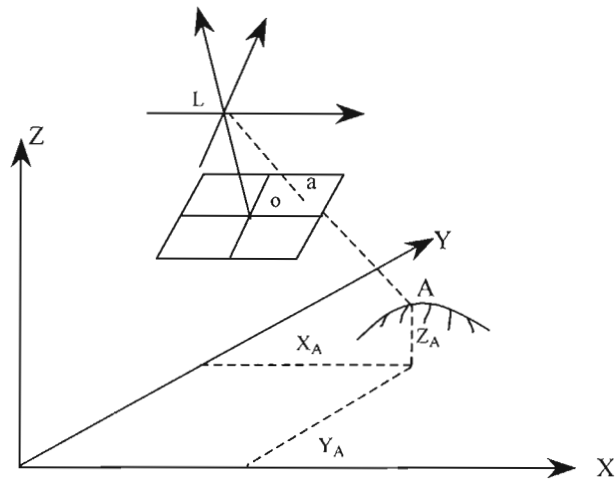

Figure A.7: Ground coordinate system

A.5. Image Orientation in Aerial Photogrammetry

In order to stereoscopically view the overlapped images for measuring 3D knowledge off photographs, the photographs have to be recovered to the status as they are taken. The procedure is regarded as *image orientation*. The parameters acquired in the

process are called *image orientation parameters* that describe the relative positions of photographs and their locations at the time of exposure with regard to the ground coordinate system. Automating the image orientation process still remains to be tackled for the photogrammetry community although considerable progress has been made (Heipke, 1997). There are two different types of orientation: *interior orientation* and *exterior orientation*. The former recovers the relative relationship between the camera and acquired photographs as the eyes and the glass demonstrated before, while the latter reconstructs the relative relationship between the photographs and the ground coordinate system as the two pieces of glass and the ground.

A.5.1. Interior Orientation

In detail, interior orientation recovers the geometry of the bundle of rays inherent in each photograph as that which existed at the instant of exposure. Its parameters include the calibrated focal length (f), the image coordinates of the principal point (x_0, y_0) and the lens distortion parameters. In fact, these parameters are from the results of the camera calibration procedure carried out prior to image acquisition.
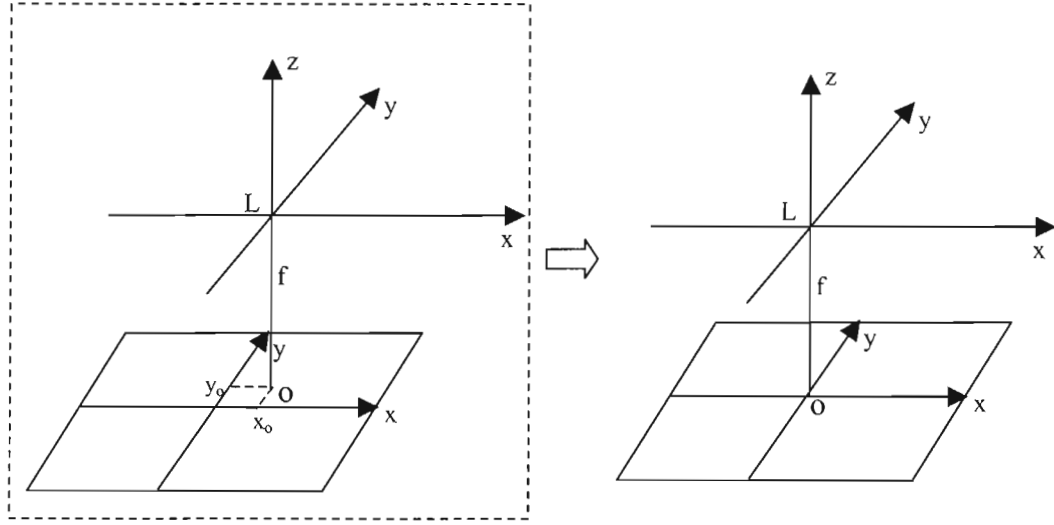
Figure A.8: Before (left) and after (right) interior orientation

With interior orientation corrected, the z-axis of image space coordinate system intersects the origin of image coordinate system shown as Figure A.8. The bundle of rays emerging from the lens will be identical to that which entered the camera at the instant of exposure.

A.5.2. Exterior Orientation

To acquire precise 3D information, stereoscopic measurements with two overlapped photographs are widely used as human beings observe 3D views with both eyes demonstrated before. Exterior orientation fixes the spatial location denoted by perspective center coordinates (X_L, Y_L, Z_L) in the ground coordinate system and the rotations (φ, ω, κ) of each photograph to every ground coordinate axis. Two procedures consist exterior orientation procedure for stereo-pairs: *relative orientation* and *absolute orientation*.

After the interior orientation has been accomplished, the rays from the corresponding image objects on the two overlapped photographs will not generally intersect one another when projected into the model space. The mismatch is called *parallax*. Relative orientation reconstructs the relative position of the two overlapped photographs at the time of photography and eliminates the parallax in the model space. It forms a three-dimension model precisely similar to the spot ground in an arbitrary space and at an arbitrary scale. Relative orientation can be accomplished either by moving two photographs or by holding one photograph fixed and only moving the second photograph. The second method is referred to *dependent relative orientation* and used in photogrammetric aerotriangulation that greatly reduces costly field surveys. With respect to this advantage, dependent relative orientation is applied in this thesis (we call it relative orientation in this thesis). Five orientation elements are solved as Figure A.9 shown. These are three rotational (φ, ω, κ) and two translational elements (b_y, b_z), which describe the relative orientation of the right image with respect to the left image and the spatial distance between the two perspective centers.
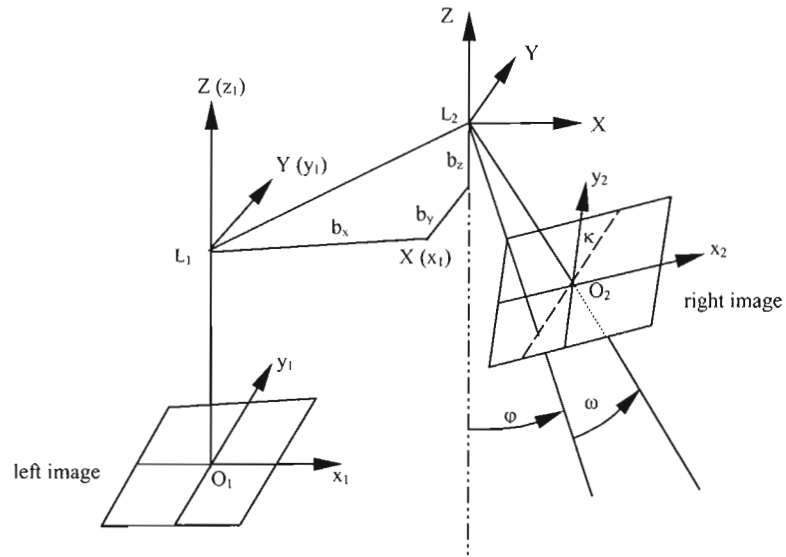
Figure A.9: Dependant relative orientation

Absolute orientation relates the pair of relative oriented photographs, and hence the three-dimension model, to the ground coordinate system. Seven absolute orientation elements are solved to represent the position status of the model to the ground. They are the scale of the model, three translations of the model and three rotations of the model. After the relative orientation accomplished, the model coordinates of ground control points that have known coordinates in the ground coordinate system are measured or computed. These are then used to determine the seven absolute orientation parameters necessary to transform the model coordinates into the ground coordinate system.

Finally, the other object points without ground coordinates can be solved by measuring their corresponding model points and then transforming the measured model coordinates into ground coordinates with the acquired orientation elements.

Appendix B.

The Development of Relative Orientation Techniques

B.1. Analog Relative Orientation

With interior orientation achieved, the prerequisite task is the relative orientation for obtaining correctly scaled three-dimensional information of the terrain by means of aerial photographs. Relative orientation restitutes the geometric relationship between an image pair at the instant of their exposure. Before a relative orientation is carried out, the rays originated from the conjugate image points normally do not intersect and exhibit disparities. Shown in Figure B.1, x-*parallax* (P_x) is defined as the difference in ray direction parallel to the *baseline* (the connection between the left and right projection centers). The difference in ray direction orthogonal to the baseline is called *y-parallax* (Q). X-parallaxes can be removed at any point by manipulation of the height adjustment of the projection, because height and x-parallax are closely allied, while y-parallaxes have to be eliminated through relative orientation.
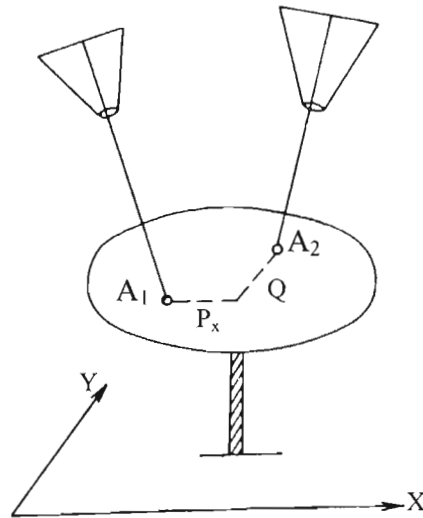
Figure B.1: Parallax

Relative orientation requires five degrees of freedom to be fixed in space by giving the position (b_x, b_y) and attitude (φ, ω, κ) of one photograph with respect to the other (the latter being fixed in space). Since the rays from the left and right projection centers must lie in a plane that contains the baseline (an epipolar plane) if they are to intersect, the removal of y-parallax is the criterion used for establishing correct relative orientation. As such, relative orientation will result in a *stereomodel*, a scaled down version of the real terrain, formed by intersecting conjugate rays. Once the stereomodel has been formed it has to be scaled and located in the ground coordinate system with the aid of known ground control points, which is part of the absolute orientation and is beyond the scope of this thesis.

From about 1900 to 1960 in analog photogrammetry (Konecny, 1985), Analog stereoplotters were commonly used, which are a optical devices that permit viewing of image pairs and superimposed synthetic feature called *floating marks* (Horn, 1989). Basically the two overlapped diapositives are put into the two projectors that have the same distance as that between the principle points. The lamps over the projectors are then

84

turned on. By means of physically operating the analog devices as in a mechanical gear system, the two overlapping photographs are relocated to the relative status at the time of exposure. Conjugate rays intersect and the images from the two projectors will form the model similar to the terrain with arbitrary scale and location.

Accurate interior orientation is first to be accomplished so that the ray bundles generated by the central projection of the photographs have the correct shape. To eliminate y-parallaxes, operations are manipulated on the setting devices of the projectors. Five conjugate points on the overlapping photographs, therefore five pairs of rays, are applied to determine the five degrees of freedom. Practically one more point are used for checking. These points are arranged in one or another specially designed pattern (Sailor, 1965; Moffitt and Mikhail, 1980). The reduction of vertical disparity at one point by means of an adjustment of a single parameter of the relative orientation disturbs the vertical disparity at the other points. Therefore, successive adjustments are applied to eliminate the y-parallaxes at each of five or six image points. Convergence is usually rapid if a good initial guess is available (Horn, 1989). However, such an analog procedure greatly depends on the skill of the operators, tedious repetitions of the projector manipulation and has limitations in function by the physical constraints of the analog mechanism. With the availability of modern digital computers with large storage capacity and the ability to compute at high speeds, analytical relative orientation gradually took the place of the analog approach.

B.2. Analytical Relative Orientation

Instead of obtaining relative orientation by analog devices that represent numerical quantities by means of physical variables, e.g., by translation; by rotation; as in

85

a mechanical gear system, analytical relative orientation uses a mathematical model to describe the criterion used for accomplishing relative orientation and acquire the orientation parameters by satisfying the criterion for all the conjugate rays. One of the photogrammetric pioneers is the German Sebastian Finsterwalder. He described the principles of modern double-image photogrammetry and the methodology of relative and absolute orientation. In addition, he introduced the necessity of redundant rays to recreate the proper geometry and used least squares theory to describe the relationship of the vectors between corresponding rays (Doyle, 1964). In 1924, Otto von Gruber derived the perspective equations and their differentials, which are fundamental of analytical photogrammetry. Uki Helava developed analytical stereoplotter in 1957, which used servocontrol instead of the optical or mechanical construction of previous instruments (Konecny, 1985).

Although various mathematical models are developed, *coplanarity condition* is the most precise and widely used (Mahajan and Singh, 1972). With reference to Figure B.2, the rays S_1a_1 from left image and S_2a_2 from the right image are conjugate rays. The coplanarity condition states simply that the base line B between the left and right perspective centers (S_1, S_2 respectively) and the two conjugate rays should all be coplanar.
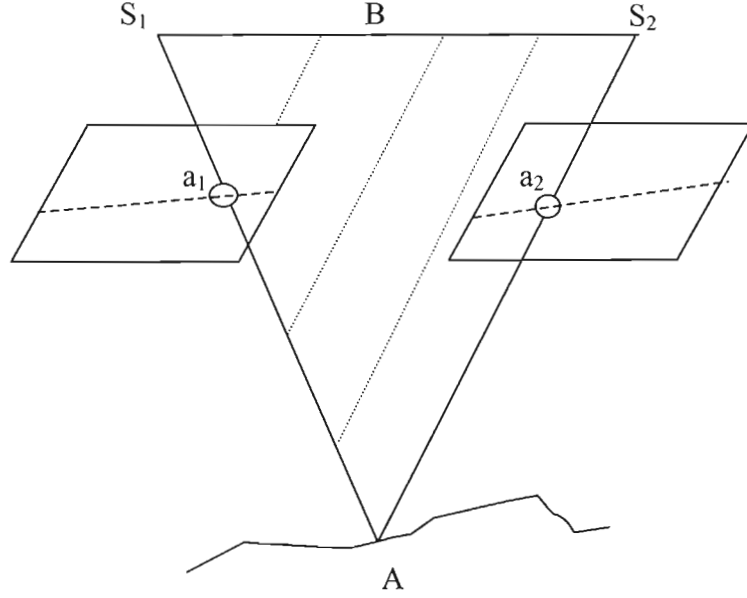
Figure B.2: Coplanarity Condition

This condition can be stated as follows:

$$\vec{B} \bullet (\overrightarrow{S_1 a_1} \times \overrightarrow{S_2 a_2}) = 0 \quad \text{(Jin et al., 2001)}$$

Let us denote the space image coordinate system for the left photograph with S_1-$X_1 Y_1 Z_1$, which is coincident with the arbitrary reference coordinate system. S_2-$X_2 Y_2 Z_2$ denotes the space image coordinate system of the right photograph and each axis is parallel to each corresponding axis in S_1-$X_1 Y_1 Z_1$. X_1, Y_1, Z_1 and x_1, y_1 respectively are the coordinates of a_1 in S_1-$X_1 Y_1 Z_1$ and the image coordinate system of the left photograph. X_2, Y_2, Z_2 and x_2, y_2 are correspondingly the coordinates of a_2 in S_2-$X_2 Y_2 Z_2$ and the image coordinate system of the right photograph. Moreover, bx, by and bz are the coordinates of S_2 in S_1-$X_1 Y_1 Z_1$. The coplanarity condition, therefore, can be expressed with coordinates:

$$\begin{bmatrix} bx & by & bz \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{bmatrix} = 0 \quad \text{(Jin et al., 2001)}$$

87

In which $\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ -f \end{bmatrix}$, $\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = R \begin{bmatrix} x_2 \\ y_2 \\ -f \end{bmatrix}$

R is the rotation transformation matrix of the image coordinate system with respect to the space image coordinate for the right photograph and is determined by φ, ω, κ. The scale of the model is determined by bx, which is not concerned in relative orientation. As such the unknowns are by, bz, φ, ω, κ. One set of conjugate points result in one equation. A minimum of five conjugate points is necessary to solve the five parameters by linearising with a Taylor expansion and differentiation with respect to the five orientation parameters. In practice, more than five points are more frequently applied to perform analytical relative orientation. The method of least squares is therefore used to adjust the redundant measurements.

BIOGRAPHY OF THE AUTHOR

Caixia Wang was born in the Xichong City in Sichuan province, People's Republic of China (P. R. China). After her graduation from Xichong high school in 1992, she enrolled at Wuhan University in Wuhan Province, P. R. China. There she earned a B.S. in Surveying Engineering in 1996 and graduated with outstanding graduate reward.

From 1996 to 1999, Caixia was employed as assistant engineer by Southwest Electric Design Institute in Chendu, Sichuan province, P. R. China. Then she joined Beijin SiWei Company in Beijing, P. R. China, to work in the Geospatial field. In the fall of 2003, Caixia entered the University of Maine as a research assistant investigating automatic image orientation with natural linear features. In the fall of 2004, besides continuing the research, Caixia worked as a teaching assistant for the course of ISE304/SIE534 Digital Image Processing. Upon completion of her M.S., Caixia plans to pursue a Ph.D position in the Digital Image Processing & Analysis Group, under the direction of Dr. Tony Stefanidis in the Department of Spatial Information Science and Engineering (SIE) at the University of Maine. She is a candidate for the Master of Science degree in Spatial Information Science and Engineering from The University of Maine in December, 2004.