INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700  800/521-0600
Estimation of motion parameters for stereo-image sequences using data association of linear features

Habib, Ayman Fawzy, Ph.D.
The Ohio State University, 1994
Estimation of Motion Parameters for Stereo-Image Sequences
Using Data Association of Linear Features

Dissertation

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Ayman Fawzy Habib

The Ohio State University
1994

Dissertation Committee:
Anton Schenk
John D. Bossler
Kurt Novak

Approved by:
Adviser
Department of Geodetic Science and Surveying
To my Parents
Acknowledgments

My deep appreciation goes to my parents to whom I will always be in debt. Their continuous encouragement and support gave me the determination to excel in my work throughout my whole life.

I also thank Dr. Kurt Novak for his advice and support during the four years I spent at The Ohio State University. My work with him as a (GTA), (GRA) and advisee was of great benefit to me. Dr. Anton Schenk was the first one who introduced me to digital photogrammetry. His advice and comments were extremely valuable for the completion of this work. I met Dr. John Bossler only a few times before my defense but I am very grateful for his remarks and comments regarding this work.

Special thanks to Dr. Burkhard Schaffrin; without his advice through course work and personal consultation I would not have been able to complete this work.

Also, I will always be in debt to my adviser at Cairo University, Dr. Y. I. Abdulaziz who taught me how to think in a scientific way. Without the basics I received from Dr. Abdulaziz, I would not have reached this stage.

To my Sisters, Gyhan, Gakleen and their husbands I will always be grateful for their moral support during my studies here in the USA. Great Thanks to my friends Jarir Saleh and Dr. Abdulaziz Alobaida whom I consider as brothers. Their friendship is an honor to me and their encouragement to finish this work was a great motivation.
VITA

October 28, 1964 ....................................................... Born Sakha, Egypt.
1986 ............................................................................. B.S., Cairo University.
1989 ......................................................................... M.S., Cairo University.
1993 ......................................................................... M.S., The Ohio State University.

FIELDS OF STUDY

Major Field of Study  Geodetic Science and Surveying.
Photogrammetry   Professor Kurt Novak, Professor Anton Schenk,
                   Professor Burkard Schaffrin.
TABLE OF CONTENTS

| ACKNOWLEDGMENTS | iii |
| VITA | iv |
| LIST OF TABLES | vii |
| LIST OF FIGURES | viii |
| LIST OF SYMBOLS | xi |

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 General</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Research Objectives</td>
<td>3</td>
</tr>
<tr>
<td>1.3 System Description</td>
<td>4</td>
</tr>
<tr>
<td>1.4 Suggested Methodology</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Organization</td>
<td>7</td>
</tr>
<tr>
<td>II. Autonomous Navigation Systems and their Primitives</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Background</td>
<td>8</td>
</tr>
<tr>
<td>2.2 The GPSVan</td>
<td>10</td>
</tr>
<tr>
<td>2.3 Strategy</td>
<td>12</td>
</tr>
<tr>
<td>2.4 Extraction of Geometric Primitives</td>
<td>13</td>
</tr>
<tr>
<td>2.4.1 Edge Detection</td>
<td>14</td>
</tr>
<tr>
<td>2.4.2 Hough Transform</td>
<td>17</td>
</tr>
<tr>
<td>III. Representation and Perspective Transformation of the Geometric Primitives</td>
<td>21</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>21</td>
</tr>
<tr>
<td>3.2 Optimal Representation of 3-D lines</td>
<td>22</td>
</tr>
<tr>
<td>3.3 Perspective transformation of 3-D lines</td>
<td>27</td>
</tr>
<tr>
<td>3.3.1 General</td>
<td>27</td>
</tr>
<tr>
<td>3.3.2 The Intersection Problem</td>
<td>28</td>
</tr>
<tr>
<td>3.3.3 Perspective Transformation from Object to Image Space</td>
<td>30</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gate Threshold and Values of Probability mass in the gate</td>
<td>36</td>
</tr>
<tr>
<td>2. The cost matrix associated with the left image of data set 1</td>
<td>79</td>
</tr>
<tr>
<td>3. The cost matrix associated with the right image of data set 1</td>
<td>80</td>
</tr>
<tr>
<td>4. The cost matrix associated with the left image of data set 2</td>
<td>109</td>
</tr>
<tr>
<td>5. The cost matrix associated with the right image of data set 2</td>
<td>109</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The GPSVan of The Center For Mapping of The Ohio State University</td>
</tr>
<tr>
<td>2.</td>
<td>Differential edge detection</td>
</tr>
<tr>
<td>3.</td>
<td>Polar representation of the image line</td>
</tr>
<tr>
<td>4.</td>
<td>The parameter space image of The Hough Transform</td>
</tr>
<tr>
<td>5.</td>
<td>Minimal representation of 3-D lines using a point N, and an angle $\alpha$</td>
</tr>
<tr>
<td>6.</td>
<td>Pluckerian representation of 3-D lines</td>
</tr>
<tr>
<td>7.</td>
<td>Representation of 3-D line as the intersection of two planes parallel to x and y axes</td>
</tr>
<tr>
<td>8.</td>
<td>Representation of 3-D line as the intersection of two planes parallel to z and y axes</td>
</tr>
<tr>
<td>9.</td>
<td>Representation of 3-D line as the intersection of two planes parallel to x and z axes</td>
</tr>
<tr>
<td>10.</td>
<td>Polar parameters as derived from image plane</td>
</tr>
<tr>
<td>11.</td>
<td>Measurements in the validation region of a single target</td>
</tr>
<tr>
<td>12.</td>
<td>Observed and predicted edges within the search window of a single target</td>
</tr>
<tr>
<td>13.</td>
<td>Measurements in the validation regions of two targets</td>
</tr>
<tr>
<td>14.</td>
<td>Observed and predicted edges within the search window of two targets</td>
</tr>
<tr>
<td>15.</td>
<td>Relationship between model spaces $M_1$, $M_2$ and the object coordinate system $O$</td>
</tr>
<tr>
<td>16.</td>
<td>Relationship between $(X_1, Y_1, Z_1)$, $(X_2, Y_2, Z_2)$ and $(L_2, M_2, N_2)$</td>
</tr>
<tr>
<td>17.</td>
<td>Different configuration of linear features and expected singularities</td>
</tr>
<tr>
<td>18.</td>
<td>Left image of the first stereo-pair of data set 1</td>
</tr>
<tr>
<td>19.</td>
<td>Right image of the first stereo-pair of data set 1</td>
</tr>
<tr>
<td>20.</td>
<td>The left patch of the first stereo-pair of data set 1</td>
</tr>
<tr>
<td>21.</td>
<td>The right patch of the first stereo-pair of data set 1</td>
</tr>
<tr>
<td>22.</td>
<td>Edge image corresponding to the left patch of the first stereo-pair of data set 1</td>
</tr>
<tr>
<td>23.</td>
<td>Edge image corresponding to the right patch of the first stereo-pair of data set 1</td>
</tr>
<tr>
<td>24.</td>
<td>Hough space image and detected clusters associated with the left patch of the first stereo-pair of data set 1</td>
</tr>
</tbody>
</table>
25. Hough space image and detected clusters associated with the right patch of the first stereo-pair of data set 1 .......................................................... 66
26. Detected edges in the left patch of the first stereo-pair of data set 1 .......................................................... 67
27. Detected edges in the right patch of the first stereo-pair of data set 1 .......................................................... 68
28. Predicted road edges in the left image of the second stereo-pair of data set 1 .......................................................... 69
29. Predicted road edges in the right image of the second stereo-pair of data set 1 .......................................................... 70
30. Automatically extracted left patch in the second stereo-pair of data set 1 .......................................................... 71
31. Automatically extracted right patch in the second stereo-pair of data set 1 .......................................................... 72
32. Edge image corresponding to the left image of the second stereo-pair of data set 1 .......................................................... 73
33. Edge image corresponding to the right image of the second stereo-pair of data set 1 .......................................................... 74
34. Hough space image and detected clusters associated with the left patch of the second stereo-pair of data set 1 .......................................................... 75
35. Hough space image and detected clusters associated with the right patch of the second stereo-pair of data set 1 .......................................................... 76
36. Detected edges in the left patch of the second stereo-pair of data set 1 .......................................................... 77
37. Detected edges in the right patch of the second stereo-pair of data set 1 .......................................................... 78
38. Matched edges in the left image of the second stereo-pair of data set 1, NNSA .......................................................... 81
39. Matched edges in the right image of the second stereo-pair of data set 1, NNSA .......................................................... 82
40. Matched edges in the left image of the second stereo-pair of data set 1, MTIC .......................................................... 86
41. Matched edges in the right image of the second stereo-pair of data set 1, MTIC .......................................................... 87
42. Left image of the first stereo-pair of data set 2 .......................................................... 89
43. Right image of the first stereo-pair of data set 2 .......................................................... 90
44. The left patch of the first stereo-pair of data set 2 .......................................................... 91
45. The right patch of the first stereo-pair of data set 2 .......................................................... 92
46. Edge image corresponding to the left patch of the first stereo-pair of data set 2 .......................................................... 93
47. Edge image corresponding to the right patch of the first stereo-pair of data set 2 .......................................................... 94
48. Hough space image and detected clusters associated with the left patch of the first stereo-pair of data set 2 .......................................................... 95
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>49.</td>
<td>Hough space image and detected clusters associated with the right patch of the first stereo-pair of data set 2</td>
<td>96</td>
</tr>
<tr>
<td>50.</td>
<td>Detected edges in the left patch of the first stereo-pair of data set 2</td>
<td>97</td>
</tr>
<tr>
<td>51.</td>
<td>Detected edges in the right patch of the first stereo-pair of data set 2</td>
<td>98</td>
</tr>
<tr>
<td>52.</td>
<td>Predicted road edges in the left image of the second stereo-pair of data set 2</td>
<td>99</td>
</tr>
<tr>
<td>53.</td>
<td>Predicted road edges in the right image of the second stereo-pair of data set 2</td>
<td>100</td>
</tr>
<tr>
<td>54.</td>
<td>Automatically extracted left patch in the second stereo-pair of data set 2</td>
<td>101</td>
</tr>
<tr>
<td>55.</td>
<td>Automatically extracted right patch in the second stereo-pair of data set 2</td>
<td>102</td>
</tr>
<tr>
<td>56.</td>
<td>Edge image corresponding to the left image of the second stereo-pair of data set 2</td>
<td>103</td>
</tr>
<tr>
<td>57.</td>
<td>Edge image corresponding to the right image of the second stereo-pair of data set 2</td>
<td>104</td>
</tr>
<tr>
<td>58.</td>
<td>Hough space image and detected clusters associated with the left patch of the second stereo-pair of data set 2</td>
<td>105</td>
</tr>
<tr>
<td>59.</td>
<td>Hough space image and detected clusters associated with the right patch of the second stereo-pair of data set 2</td>
<td>106</td>
</tr>
<tr>
<td>60.</td>
<td>Detected edges in the left patch of the second stereo-pair of data set 2</td>
<td>107</td>
</tr>
<tr>
<td>61.</td>
<td>Detected edges in the right patch of the second stereo-pair of data set 2</td>
<td>108</td>
</tr>
<tr>
<td>62.</td>
<td>Matched edges in the left image of the second stereo-pair of data set 2, NNSA</td>
<td>110</td>
</tr>
<tr>
<td>63.</td>
<td>Matched edges in the right image of the second stereo-pair of data set 2, NNSA</td>
<td>111</td>
</tr>
<tr>
<td>64.</td>
<td>Matched edges in the left image of the second stereo-pair of data set 2, MTIC</td>
<td>113</td>
</tr>
<tr>
<td>65.</td>
<td>Matched edges in the right image of the second stereo-pair of data set 2, MTIC</td>
<td>114</td>
</tr>
<tr>
<td>66.</td>
<td>The intersection of the plane (x = az + p) with the x-z plane</td>
<td>122</td>
</tr>
<tr>
<td>67.</td>
<td>The intersection of the plane (y = bz + q) with the y-z plane</td>
<td>123</td>
</tr>
<tr>
<td>68.</td>
<td>Relationship between polar parameters of 2-D lines and the points of intersection with the x and y axes</td>
<td>125</td>
</tr>
</tbody>
</table>
LIST OF SYMBOLS

EOP .................................. Exterior Orientation Parameters.
NNSA ................................ Nearest Neighbor Standard Algorithm.
MTIC ................................ Multiple Targets In Clutter.

\((\rho, \theta)\) .................. Parameters describing the image line in its polar form.
\((a, b, p, q)\) .................. Four dimensional vector representing 3-D lines.
\((U, W, V, D)\) ............... Parameters describing the object plane that contains the perspective center and the object line.

\(\hat{z}(k+1|k)\) .................. Predicted measurement at epoch \((k+1)\) based on all observations up to epoch \(k\).

\(s_z(k+1)\) .................. Variance-covariance matrix of \(\hat{z}(k+1|k)\).
\(Z^k\) .......................... Cumulative set of observations up to and including epoch \(k\).
\(Z(k+1)\) .................. Set of observations within the search window at epoch \((k+1)\).

\(z_i(k+1)\) .................. \(i^{th}\) measurement in the set \(Z(k+1)\).

\(S_z(k+1)\) .................. Variance-covariance matrix of \(z_i(k+1)\).

\(u_i(k+1)\) .................. Difference between \(z_i(k+1)\) and \(\hat{z}(k+1|k)\).

\(S_{u_i}(k+1)\) .................. Variance-covariance matrix of \(u_i(k+1)\).

\(\tilde{V}_{k+1}(\gamma)\) .......... Validation region at epoch \((k+1)\).

\(n_x\) .................. The size of the observation vector.

\(Z_v(k+1)\) .................. The set of validated measurements at epoch \((k+1)\).

\(d\{z_i(k+1)\}\) .......... Mahalanobis distance associated with \(z_i(k+1)\).

MHA .......................... Multiple Hypothesis Approach.

\(\Theta(i)\) .................. \(i^{th}\) association event.

\(\Omega(i)\) .................. Association matrix corresponding to \(\Theta(i)\).

\(\omega_{ji}(i)\) .................. The element in the \(j^{th}\) row and the \(i^{th}\) column of the association matrix \(\Omega(i)\).

\(C(i)\) .................. The cost of the association matrix \(\Omega(i)\).
\( \Lambda(\Theta_i) \) ........................................... The likelihood function of the association event \( \Theta(i) \).

\( \lambda(i) \) ............................................ The modified likelihood function of the association event \( \Theta(i) \).

\( \alpha \) ............................................. Probability of type I error, level of significance.
Chapter I

INTRODUCTION

1.1 General

During the last decade, the need for land-related information, especially in digital format, grew tremendously (Novak, 1992). The motivation for this need is the development of computer data-bases and especially Geographic Information Systems (GIS). Huge amounts of digital data are necessary for a GIS. The acquisition process is a major research area in Geodetic Science and especially Photogrammetry. A specific problem in this area is the acquisition and processing of data associated with road networks. Road network data can be obtained either by digitizing existing maps or by site surveying, e.g., the use of Mobile Mapping Systems (MMS). MMS proved to be faster, more accurate, and current than manual digitization, which introduces a number of secondary errors such as those due to redigitization of existing generalized maps.

The goal of this research is to develop an algorithm capable of utilizing information gathered by a multi-sensor navigation system to help the system make decisions about its movement in the surrounding environment. The proposed algorithm utilizes the observations collected by the system's sensors in a sequential manner to update the knowledge about its movements and the environment. The methodology proposed consists of three main steps: prediction, observation and updating.

Autonomous navigation systems are always equipped with sensors, such as radar and/or imaging systems, that observe the surrounding environment to guide their motion. First, the designer of the system must decide about the kind of primitives that
will be used to guide the system in the navigation process. The system will use these primitives to infer its location within the surrounding environment. The inference of this knowledge relies on the mathematical model that relates the observations and the corresponding object primitives. Therefore, the first issue to be considered in any navigation system is the choice of the proper primitives and the mathematical model that describes the transformation between observations and object space.

As it collects observations, the autonomous navigation system has to decide about the source of these observations. At different epochs the system must group its observations into sets that are believed to originate from the same target. Solving this problem is essential as updating the state estimate is based on discerning the change in the observed primitives as a result of the dynamic behavior of the observer and/or the target. This means that the system must group its observations into sets, each one of them representing the history of the observations that correspond to the same target. These sets are termed as tracks according to the terminology implemented in the computer vision literature. The system must make decisions about the origin of its observations, which is known as the data association problem (Bar-Shalom and Fortman, 1988). This research will investigate different data association techniques that vary in complexity and time/memory requirements to decide which suits this application in the best way.

After solving the data association problem, a mathematical model is developed that uses the changes in the observed primitives at different epochs to derive the motion parameters of the system as well as the object primitives between these epochs. The primitives observed at different viewpoints whose relationships are unknown or approximately known from other sensors installed in the system are used
to estimate the three dimensional displacements of the system and the world primitives.

All techniques mentioned above are implemented and tested with observations gathered by the GPSVan developed by the Center For Mapping of The Ohio State University. The GPSVan is equipped with GPS/INS system for vehicle positioning and a digital stereo-imaging system for capturing road features. It provides the main components of an autonomous vehicle navigation system. For this particular research the geometric primitives to be considered are line segments representing road markings. The optimal representation as well as the perspective transformation of these primitives between image and object space are developed. Using data association techniques the system groups the road edges detected in consecutive stereopairs into tracks. These tracks will be used to update the van's position as well as the parameters describing the road edges. Once a map of the road edges is available, the vehicle can use this information to make some decisions about how to move within the environment.

1.2 Research Objectives

The objective of this research is to investigate the feasibility of generating a map containing the road edges in real time with a vehicle mounted stereo-vision system and other navigation sensors. This goal incorporates the following tasks:

(1) Developing a system that integrates observations from different sensors onboard the navigation system to dynamically update the system's position and the parameters describing the surrounding environment.

(2) Implementing geometric primitives of a higher level of abstraction (i.e., using 3-D line segments instead of points to represent the system's
environment). Also, introducing (perspective) transformations between image and object space that are suitable for these geometric primitives.

3. Introducing data association techniques that group the temporal observations gathered by the sensors into sets (tracks) according to the origin of these observations.

4. Using the matched primitives at different epochs to update the motion parameters of the system.

1.3 System Description

At the Center for Mapping of The Ohio State University, a mobile mapping workstation (GPSVan) has been developed over the past four years, (The Center For Mapping, 1991). The GPSVan is equipped with a GPS receiver, an inertial navigation system, and a stereo-vision system for capturing highway data. The GPS observations are used to provide position fixes in global coordinates. The inertial system provides information about the attitude of the van as well as positional information in case of loss satellites' lock. The positional and attitude information is utilized by the stereo-vision system to extract spatial features visible in the field of view of both cameras.

Pseudo ranges and carrier phases are the fundamental measurements of the GPS system (Leick, 1990). Both are thought of as biased distance measurements between the GPS satellites and receiver antennas. Pseudo ranges are obtained by multiplying the travel time of the signal by the speed of light. Carrier phases are the measurements of the phase difference between the satellite and receiver oscillators. The synergistic combination of pseudo ranges and phases improves the positional accuracy and resolves the cycle slips that affect distance computation from phase measurements.
The inertial system, which is actually a dead-reckoning device, consists of two components: the gyro system and the wheel counter. The gyro system provides pitch, roll and direction data in a self-contained gyro package, while the wheel counter measures the distance traveled by the van.

The stereo-vision system consists of two digital CCD cameras that are connected to a tape storage unit. While the GPSVan is driving, the system captures digital image-pairs that are compressed and stored on digital tapes. The two digital cameras are mounted on the roof of the van facing forward. After data collection, 3-D coordinates are computed by intersection of image points. The image points are then related to a global coordinate system using the known GPS position and attitude of the van.

Data from the GPSVan can be transformed to a format acceptable for entry into any commercial Geographic Information System (GIS). This information can be used to monitor road and transportation features, and to set management priorities.

The main components of the GPSVan satisfy the basic hardware requirements of an Autonomous Land Vehicle (ALV), according to Barnard et. al. (1986). An ALV can use information from an extensive base of knowledge of its environment, including the locations of objects along its path. For the GPSVan, this could be an existing GIS. It can also rely on a kind of "dead reckoning" with its navigation system (INS), which allows for system position estimation by integrating sensor observations over time. To be truly effective, it must exploit information from its visual sensors to monitor and correct its motion through the world. Beyond these basic components, the GPSVan utilizes GPS receivers, which yield global positions of the vehicle. In other words the system's output will be referenced to a world wide not local reference system.
1.4 Suggested Methodology

The navigation process can be defined as follows (Faugeras and Hebert, 1986):

Suppose that the analysis of the first image-pair captured with the stereo-vision system yields a partial 3-D description of the environment in terms of points, lines and planes, each with some uncertainty attached to it. The uncertainty associated with the geometric primitives depends on the pixel uncertainty of the digital images and the uncertainty of the exterior orientation parameters (EOP) of the cameras. The perspective relationship between image and object space is responsible for error propagation.

Now, suppose that the vehicle on which the cameras are mounted moves by an imperfectly known displacement vector due to the uncertainties associated with GPS/INS observations. After this motion, a second stereopair is acquired. The second stereopair is analyzed to obtain another 3-D description of another part of the environment. If the displacement vector is small, it is likely that some of the geometric primitives identified in the first stereopair will also appear in the second one. By matching these primitives, it should be possible to recover a better estimate of the displacement and to construct a better model of the geometric primitives describing the environment.

According to the above-mentioned functionality of the proposed system, this research will address the following points:

- **Best Representation of the Environment**: This means that one has to choose the right geometric primitives to represent the van's surroundings. This also deals with modeling the perspective transformation of these primitives between image and object space.
• **Data Association**: This aims at matching geometric primitives in successive stereopairs according to their origin.

• **Updating State Estimates**: Using the matched primitives together with the transformation models, one can update the parameters describing the van's motion and its environment.

### 1.5 Organization

This study is organized in seven chapters. Chapter II deals with the issues of choosing and extracting the geometric primitives to be used throughout the research. Chapter III discusses the problem of optimal representation of the chosen geometric primitives together with their perspective transformation back and forth between image and object space.

Chapter IV is devoted to the explanation of the data association problem and how it will be implemented in this work for the purpose of identifying the edges in the second stereopair arising from lane boundaries. In this chapter, two main algorithms will be discussed. The first one is the Nearest Neighbor Standard Algorithm (NNSA). In this algorithm, the competition of several tracks—lane boundaries—for the same observation is ignored. On the other hand, this fact is considered for the second algorithm, Multiple Targets In Clutter (MTIC). Chapter V describes the suggested algorithm for motion parameters updating using the matched linear features in the stereopairs under consideration.

Chapter VI presents the results of all suggested algorithms. Results for two data sets which are representatives of the most common lane boundaries are shown. Finally, in the last chapter, VII, the conclusions and recommendations drawn from this work are presented.
2.1 Background

An autonomous vehicle needs to plan its action, perceive its surroundings, execute its plan, and adapt itself to the environment for survival (Kuan et al., 1986). Based on this plan, the autonomous vehicle starts to execute the plan in the real world. It collects information with the on-board sensors to perceive its environment, to follow a road, and to understand scenes. If some unexpected events happen that interfere with its current plan, the autonomous vehicle needs to adjust its plan to the current situation.

Visual navigation of autonomous vehicles on road networks is an important problem. A number of different systems exist. Among them is the one developed by FMC Corporation, Santa Clara, California (Kuan et al., 1986), which uses an M113 armored personnel carrier. This system acquires a color image from a single camera (mono-vision) and the current vehicle location from an Inertial Navigation System. An image segmentation module uses a pixel classification algorithm to segment the image into road and non-road regions. Then, a road boundary tracking module finds the most likely road region and traces the contour of the region. The contour is then represented as a sequence of line segments using a line fitting algorithm. These line segments are transformed from the image coordinate system to the local vehicle coordinate system.
and sent to the geometric reasoning module. The geometric reasoning module aggregates local geometric supports to constrain and increase the likelihood of the hypothesized model (e.g., road sides consistency constraint, smoothness constraint, continuity constraint), and assigns a consistent interpretation to these line segments. The resulting road interpretation is fused with other sensor interpretation results (e.g., an obstacle map from a range sensor) and sent to the pilot system to generate a local path and navigate the vehicle.

Another system was developed by the Artificial Intelligence Center, Menlo Park, California (Barnard et al., 1986), that is equipped with an Inertial Navigation System, an ERIM range sensor and an extensive base of knowledge of its environment, including the location of objects along its path. The autonomous vehicle knows its initial position and orientation in the world. It begins to move with some acceleration, measured by the INS. Information from the INS is used to estimate the new position and orientation of the vehicle. This estimate is passed along to a recognition module that attempts to locate known objects in the field of view of the range sensor. Knowledge of the relative positions and orientations of the vehicle and the objects aids the recognition module in its task. If an object is recognized, the parameters of its position are computed and used to correct the estimate of the current vehicle position.

The VaMoRs developed in Neubiberg, Germany (Dickmanns, 1986, 1992-a, 1992-b), succeeded in the demonstration of autonomous road following at a speed of 96 km/h in 1987. This is a 5-ton van equipped with an electro-mechanical pan-tilt platform carrying two CCD cameras mounted in the center behind the front windshield, hanging from the roof. These two cameras have different focal lengths, a wide angle camera for global feature analysis (such as road boundaries) and another one for detail inspection (e.g., for focusing on objects or obstacles further away). The
camera pointing capability allows active search and tracking, e.g., for initial self orientation, motion blur reduction and continuous road tracking while driving. The system does not have any positioning capabilities.

The GPSVan developed by the Center For Mapping of The Ohio State University, (Columbus, Ohio) is equipped with a GPS receiver and two CCD cameras to record highway features (The Center For Mapping, 1991). The imaging system is synchronized to GPS and registers the precise time and position of each image pair. To compensate for blocked satellite signals when traveling in urban areas, through tunnels, or under bridges, an inertial system is implemented that determines the distance and direction traveled. This system consists of horizontal and vertical gyroscopes for direction, pitch and roll measurements, and a wheel counter for distance measurements. The hardware component of the GPSVan can be used for automatically navigating the system by implementing the appropriate software.

2.2 The GPSVan

This research will investigate the navigation issue as it pertains to the GPSVan (Figure 1). The basic goal for the development of this system according to the proposal to the Federal Highway Administration, Application of the Global Positioning System for Transportation Planning:

"It is the goal of this project to build a prototype system that can economically, rapidly, and accurately collect, analyze, and process transportation data into a Geographic Information System. The system will include a satellite receiving station, mounted in a van, that can determine the vehicle's geographic position at highway speeds. The data will be post processed for loading into a GIS. Global three dimensional position accuracy will be 10 meters or better. It also the goal of the project to add remote sensors to the vehicle, such as a stereo-video system, to locate and record the condition of transportation features."
Contrary to the previously mentioned systems, the GPSVan is equipped with a stereo-vision system. Similar to the first and the third systems, the proposed system will use the road edges as extracted by edge detection operators as navigation primitives. In addition, this system will provide a map displaying the lane boundaries as a by-product. This map can be directly fed to any existing GIS as one of its layers. This is feasible since the GPSVan is equipped with a GPS receiver that yields a global positions.

Figure 1: The GPSVan of the Center for Mapping of The Ohio State University.
2.3 Strategy

As will be described in the next section, the navigation primitives that will be used in this research are the road edges as represented by the lane boundaries. The proposed system works in the following manner:

(1) Initialization Process: During this process, the road edges of the first stereopair are initialized. This is done semi-automatically; in other words, some decisions by the user are required, for example, the user has to define the area of the image where the road edges are expected. Within the established area, an edge detection operator extracts the road edges.

(2) Prediction Process: Having some knowledge about the movement of the navigation system between the first and the second stereopairs, one can predict the location of the initialized road edges in the second stereopair. Information about the vehicle's movement can be obtained from the on-board positioning sensors, e.g., an GPS/INS system. The predicted edges can be obtained by projecting the road edges of the first stereopair into object space and then projecting them back into the image space of the next stereopair.

(3) Observation Process: At this stage, a search window is created surrounding the predicted edges obtained in step 2. The size of the search window is proportional to the confidence level of the vehicle's motion parameters between the stereopairs under consideration, as well as the reliability of the initialized edges. Inside the search windows, edge detection operators are applied to find all possible edges.

(4) Data Association Process: Some of the observed edges obtained in step (3) correspond to the road edges while the rest are due to shadows, noise, and texture. The function of the data association algorithm is to come up with a set of edges that most likely correspond to the road edges.
(5) **Motion Parameter Updating Process:** After matching the road edges in successive stereopairs, the vehicle's motion parameters are updated. For the next stereopair, steps 2 to 5 are repeated.

Therefore, the proposed system consists of four modules. The first one is the *representation module* of the geometric primitives. For this research, straight lines that correspond to lane boundaries are the geometric primitives of choice. Optimal representations are studied to choose the most suitable representation. The *perspective transformation* of these primitives from object to image space is another module. The transformation models are developed based on the chosen representation module. The *data association module*, the third one, solves the correspondence problem between the detected edges and their origins such as edges arising from road boundaries, noise, and shadows. Different data association techniques are studied and compared based on their complexity, computational requirements, and performance. Finally, the *motion parameter updating module* uses the matched road edges at successive stereopairs to come up with an estimate of the displacement between these stereopairs.

### 2-4 Extraction of Geometric Primitives

An autonomous vehicle always relies on some kind of primitives that help it in making decisions about its location within the environment. The first system mentioned in section 2.1 uses the road boundaries as its navigation primitives. The system implements a segmentation algorithm to come up with these primitives from a single color image. On the other hand, the second system built by the Artificial Intelligence Center, Menlo Park, California, uses an onboard range sensor together with an extensive base knowledge of its environment, including the location of objects along its path. The navigation primitives in this case are identified known objects in the field of view of the range sensor. The third system, VaMoRs, also implements road
boundaries that are detected via edge based algorithms from intensity images, region-based algorithms from intensity images, or region based algorithms from color images (Kuan et al., 1986).

In this study, road edges that are extracted through oriented edge detection operators will be used as the navigation primitives. The advantage of using the road edges as the navigation primitives lies in the fact that they can be mapped into a Geographic Information System (GIS) as the actual road way location. Also, the system can be taught to navigate itself between these lines using additional obstacle avoiding sensors such as radar.

Edge detection operators will be discussed in the remaining sections of this chapter. Classical edge detection operators are purely signal driven and lack scene descriptive criteria. They treat all edges, whether they are real road edges or caused by shadows, in the same manner. To overcome this drawback of low level processing, a higher level processing algorithm must be implemented. This algorithm will distinguish the edges corresponding to road boundaries from those arising from shadows, noise and texture. In this research, this task is performed through data association techniques that will be discussed in Chapter 4.

2.4.1 Edge Detection

The human visual and recognition system heavily relies on physical object boundaries (Attneaves, 1954). "Edge" is the word used to describe an object boundary. In this study, edges are assumed to arise from lane boundaries together with some other features such as shadows, texture, and/or noise. Boundaries of objects on an image usually manifest themselves as abrupt changes in intensity. Therefore, the identification of boundaries is based on the extraction of intensity changes. A variety of edge detection techniques have been introduced based on combination of studies
from many disciplines, e.g., psychological studies, image processing and artificial intelligence research. Another issue associated with edge identification is that after locations of sudden intensity changes have been located, they must be grouped together to form a meaningful description of the scene.

In this research, a differential edge detection operator is implemented to determine which pixels correspond to intensity changes on the images captured by the stereo-vision system of the GPSVan. The road edges always appear to be almost vertical in image space. Therefore, a directional edge detection operator that finds the gradient in a direction perpendicular to the edge orientation can be used. For this purpose, the vertical component of a Sobel operator is applied to the image. The 3x3 mask representing the operator, (2.1):

\[
\begin{pmatrix}
-1 & 0 & 1 \\
-2 & 0 & 2 \\
-1 & 0 & 1
\end{pmatrix}
\]  

(2.1)

is applied to 3x3 windows of the original image:

\[
\begin{pmatrix}
x_1 & x_2 & x_3 \\
x_4 & x_5 & x_6 \\
x_7 & x_8 & x_9
\end{pmatrix}
\]

where \(x_i\) is the gray value at the \(i^{th}\) location.

The gradient in the horizontal direction \(G_x\) for the central pixel of the mask is given by:

\[
G_x = (x_1 + 2x_6 + x_9) - (x_1 + 2x_4 + x_7)
\]

(2.2)

If \(G_x\) is greater than a prespecified threshold for the magnitude of the gradient as defined by the user, the corresponding central pixel of the mask is recognized as edge
point, otherwise, it is mapped as background. By assigning a gray value of 255 to the edge points and zero to the background, a binary edge image is generated. An original image, as well as the corresponding edge image, are displayed in Figure 2.

(a) Original image.

(b) Marked edge pixels.

Figure 2: Differential edge detection.
2.4.2 Hough Transform

The output of the Sobel operator is an edge image that highlights the pixel locations that correspond to intensity changes. However, these scattered pixels cannot give any meaningful description. Therefore, another operator is required that groups these pixels into lines. This process is called line segmentation.

For the purpose of this work, the Hough Transform will serve as the line segmentation operator (Hough, 1962). The attraction of using the Hough Transform as a line segmentation operator is due to its ability to overcome gaps, partial occlusions, and noise (Adamos and Fiag, 1992). The Hough Transform is considered an intermediate level vision task, since it converts an edge image that does not give any abstract information into an image with a meaningful symbolic description of the scene.

Duda and Hart (1972) introduced and modified the Hough Transform for the detection of lines. In this case, the line is described by the following equation in the image space:

\[ x \cos \theta + y \sin \theta = \rho \]  \hspace{1cm} (2.2)

where:

- \( \theta \) is the angle relative to the positive horizontal direction,
- \( \rho \) is the normal distance measured from the origin, (see Figure 3).
This form of line representation avoids the singularity arising when the lines are perpendicular to the horizontal axis. The distance $\rho$ ranges from -0.5 D to 0.5 D, where D is the length of the diagonal in the image space. The angle $\theta$ can be a value ranging from 0 to 180 degrees.

A transformation from image space $(x, y)$ to parameter space $(\theta, \rho)$ is performed. This means that image points that have been marked as edge pixels in the edge image are mapped into sinusoidal curves in parameter space. If a number of these edge points lie on a straight line in image space, the corresponding sinusoidal curves in parameter space will intersect in point $(\theta_k, \rho_k)$, representing the parameters of this particular line.

To segment the pixels detected in the edge image into meaningful lines, a parameter space image is generated (see Figure 4). The number of pixels in this image $n_\theta, n_\rho$ depends on how many increments one wants to quantize the 180 degrees and the distance D associated with the parameters $\theta, \rho$, respectively. Initially, all cells of
parameter space image are set to zero. Then, for each highlighted point \((x, y)\) in the edge image the quantized values for \(\theta\) are incremented by \(\delta \theta = \frac{180^\circ}{n_\theta}\) and for each value of \(\theta\), the corresponding \(\rho\) is calculated according to equation 2.2. The computed value of \(\rho\) is quantized according to the increments \(\delta \rho = \frac{D}{n_\rho}\) and the corresponding cell in the parameter space is incremented by 1. The cells in parameter space with the greater accumulated number of votes correspond to lines in the image space (Leavers, 1992).

![Parameter space image of Hough Transform](image)

If the original image is noise free, all points that lie on a straight line will result in sinusoidal curves that perfectly intersect in one point in parameter space. Since real road images are not free of noise, the intersection point will become a cluster with the local maxima corresponding to the line of interest. So, a cluster detection algorithm is
implemented to detect these agglomerations in parameter space. The clusters are extracted with the help of a threshold value that is provided by the user. This means that cells with accumulated values less than the prespecified threshold will be excluded from the cluster.

For each cluster, the mean values $m_\theta, m_\rho$ and the corresponding standard deviations $\sigma_\theta, \sigma_\rho$ are computed as follows:

$$m_\theta = \frac{1}{S} \sum \sum i f(i,j) \quad (2.3)$$

$$m_\rho = \frac{1}{S} \sum \sum j f(i,j) \quad (2.4)$$

$$\sigma^2_\theta = \frac{1}{S} \sum \sum (i-m_\theta)^2 f(i,j) \quad (2.5)$$

$$\sigma^2_\rho = \frac{1}{S} \sum \sum (j-m_\rho)^2 f(i,j) \quad (2.6)$$

where:

$$S = \sum \sum f(i,j)$$

with the summation goes over the whole cells belonging to that cluster,

$f(i,j)$ is the accumulated votes in cell $(i, j)$ of that cluster.

The mean values $m_\theta, m_\rho$ are the parameters of the extracted lines, while $\sigma_\theta, \sigma_\rho$ are the corresponding standard deviations.
CHAPTER III

Representation and Perspective Transformation of

the Geometric Primitives

3.1 Introduction

Most photogrammetric applications use distinct points to describe the world model. Unfortunately, it is often impossible to define a unique measuring point on an object (Kubik, 1991). Another disadvantage of using specific points is the deficiency in deriving higher level abstract representations of the world from these points. Therefore, it is mandatory to seek different geometric primitives that are useful for higher level processing goals (e.g., autonomous navigation, object recognition).

For the following reasons, a straight line representation is most suitable (Tommaselli, 1992):

- Images of man-made environments contain many straight lines;
- Straight lines are easier to detect than distinct points, and the correspondence problem between image and object space becomes easier;
- Straight line parameters can be obtained with sub-pixel accuracy.

For this particular research the features of interest are road edges. Therefore, the linear representation is the optimum choice. In this chapter, the optimal representations of 3-D lines which require the minimum number of parameters will be investigated. Also, the perspective transformation of these lines back and forth between image and object space will be addressed.
3-2 Optimal Representation of 3-D Lines

As mentioned above, 3-D lines are the geometric primitives of choice in this research. In this section, the question of how to represent 3-D lines will be answered.

The first and most obvious representation is one using two points \((x_1, y_1, z_1), (x_2, y_2, z_2)\) along the line. This representation is six dimensional. It is not minimal since one can impose constraints that limit the degrees of freedom of these six parameters. First, one can choose the first point \((x_1, y_1, z_1)\) to be the intersection of the perpendicular through the origin and the line under consideration. In that case, the 6 parameters must satisfy the following condition:

\[
x_1(x_2 - x_1) + y_1(y_2 - y_1) + z_1(z_2 - z_1) = 0 \quad (3.1)
\]

On the other hand, the second point \((x_2, y_2, z_2)\) can be chosen such that the length of the vector connecting the two points is unity, i.e., the following constraint must be satisfied:

\[
(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = 1.0 \quad (3.2)
\]

Another way of representing the 3-D line is through a point \((x_1, y_1, z_1)\) that belongs to the line and a unit vector along its directions. This representation is five dimensional. In the same way, one can choose the point \((x_1, y_1, z_1)\) to be the intersection of the perpendicular through the origin with the line under consideration.

From the above discussion, one comes to the conclusion that the minimal representation of a 3-D line requires only four parameters (Roberts, 1988). One way of minimal representation of 3-D lines is by considering the normal to a line from the origin. The point of intersection, \(N\), between the normal and the line can be represented by three parameters \((x_n, y_n, z_n)\). Using this point, one can define a plane through the point of intersection that is perpendicular to the normal:

\[
x_n(x - x_n) + y_n(y - y_n) + z_n(z - z_n) = 0 \quad (3-3)
\]
The line under consideration is contained in that plane and can be defined by one more angle, $\alpha$. This angle can be measured from an arbitrary direction (see Figure 5) where the arbitrary direction is the line connecting $N$ and the point of intersection of the plane with the $z$-axis.

![Figure 5: Minimal representation of 3-D lines using one point, $N$, and angle, $\alpha$.](image)

The Pluckerian line representation (Navab and Zhang, 1992) is another minimal representation of 3-D lines (see Figure 6). Two parameters define the unit vector, $l$, that describes the direction of the line under consideration. The other two parameters define the vector, $h$ perpendicular to the plane containing the 3-D line and the origin. The latter two parameters can be substituted by the parameters of the vector perpendicular to the 3-D line through the origin, $n$. 
A third minimal representation of 3-D lines is given by defining the line as the intersection of two planes, one parallel to the y-axis and the other parallel to the x-axis (Ayache and Faugeras, 1989). Each of these two planes contains two parameters as follows:

\[
\begin{align*}
x &= a z + p \quad \text{(plane parallel to the y-axis)} \quad (3-4) \\
y &= b z + q \quad \text{(plane parallel to the x-axis)} \quad (3-5)
\end{align*}
\]

Thus, the line is defined by the four dimensional vector \((a, b, p, q)\). The parameters have the following geometric interpretation (see Appendix A):

- The direction of the line is given by the vector \((a, b, 1)\).
- The point of intersection of the line with the x-y plane has the coordinates \((p, q, 0)\) (see Figure 7).
Since the direction vector of this line is \((a, b, 1)\), this form cannot be used to represent 3-D lines that are parallel to the \(xy\)-plane. To avoid this singularity, one can represent the 3-D lines by the intersection of the following planes:

\[
\begin{align*}
  y &= ax + p \quad \text{(plane parallel to the z-axis)} \quad (3-6) \\
  z &= bx + q \quad \text{(plane parallel to the y-axis)} \quad (3-7)
\end{align*}
\]

In that case, the representation vector \((a, b, p, q)\) has the following geometric interpretation:

- The direction of the line is given by the vector \((1, a, b)\).
- The point of intersection of the line with the \(yz\) plane has coordinates \((0, p, q)\) (Figure 8).
This representation can not represent lines that are perpendicular to the x-axis, for which one can use:

\[ z = b \, y + q \quad \text{(plane parallel to the x-axis)} \quad (3-8) \]
\[ x = a \, y + p \quad \text{(plane parallel to the z-axis)} \quad (3-9) \]

Therefore, the vector \((a, b, p, q)\) signifies that:

- The direction of the line is defined by the vector \((a, 1, b)\).
- The point of intersection of the line with the x-z plane has the coordinates \((p, 0, q)\)

(see Figure 9):
For this study, the last representation is used. This is due to its simplicity when transforming the 3-D lines back and forth between image and object spaces with the perspective equations.

3.3 Perspective Transformation of 3-D Lines

3.3.1 General

In most photogrammetric applications, the relationship between the image coordinates \((x, y, -f)\) and the corresponding object coordinates \((X, Y, Z)\) of a single point is described by the collinearity equations. They constrain the perspective center, the image point, and the corresponding object point to a straight line. Unfortunately, it is often difficult to establish this correspondence. This is due to the fact that some points are occluded in some of the images. Another disadvantage of this model appears
when one attempts to locate the conjugate point automatically by a matching algorithm. The matching process using features is more reliable than matching distinct points. It has to be mentioned that the ultimate aim of digital photogrammetry is to understand scenes captured by imaging systems for the purpose of decision making. Single points contribute very little towards the achievement of this task. On the other hand, features in general and linear ones in particular contribute a lot towards higher level tasks since they can be associated with objects. Therefore, it is desirable to establish a mathematical model that is based on feature-to-feature rather than point-to-point correspondence (Tankovich, 1991).

Substantial work has been done in this area by Mulawa and Mikhail (1988), Tommaselli and Luganani (1988), Tommaselli (1992) and Ayache and Faugeras (1989). In this research, the mathematical model that is based on the minimal representation of 3-D lines will be investigated; specifically, the model suitable for the representation of 3-D lines as the intersection of two planes (Ayache and Faugeras, 1989) will be applied.

3.3.2 The Intersection Problem

In the intersection problem, one tries to locate an object feature knowing its corresponding image features in two or more overlapping images, together with the Exterior Orientation Parameters (EOP) of the photos. EOP can be obtained from the onboard positioning sensor, e.g., GPS receiver. In the following section, the intersection problem using linear features instead of distinct points is explained. The outlined model is based on work by Ayache and Faugeras (1989).

One starts with the parameters that represent the same line in two or more images. These parameters are the polar representations of the image lines under
consideration (Eq. 3-10). They can be obtained directly from the Hough segmentation algorithm explained in section 2-4-2.

\[ x \cos \theta_i + y \sin \theta_i = \rho_i \]  

(3-10)

where \( i \) is the image index.

The EOP \( (X_o, Y_o, Z_o, \omega_o, \phi_i, \kappa_i) \) of the involved images are usually available. The objective is to estimate the four-dimensional vector \((a, b, p, q)\) of the corresponding 3-D object line. Throughout the rest of this chapter, the plane containing the perspective center and the image line is referred to as the image plane and the plane containing the perspective center and the object line is called the object plane.

First, the parameters representing the image plane are derived (Eq. 3-11). These parameters are calculated with respect to the object coordinate system.

\[ U_i X + V_i Y + W_i Z + D_i = 0 \]  

(3-11)

The four-dimensional vector \((U_i, V_i, W_i, D_i)\) can be obtained from the parameters of the image line \((\rho_i, \theta)\) in the \(i^{th}\) image as well as the EOP of the same image. The procedure of deriving this vector is explained in Appendix B.

Assume that the corresponding 3-D object line is represented as the intersection of the two planes of equations (3-4) and (3-5). As explained before, this 3-D line passes through the point \((p, q, 0)\) and has the direction vector \((a, b, 1)\). Therefore, the point \((p, q, 0)\) must satisfy the equation of the object plane. This condition yields the following:

\[ U_i p + V_i q + W_i q + D_i = 0 \]  

(3-12)

Vector \((U_i, V_i, W_i)\) represents the direction of the normal to the object plane. This means that this vector is perpendicular to the direction vector \((a, b, 1)\) of the object line. Thus, the dot product of these two vectors must equal zero,
Having two overlapping images, one can write four equations, two equations of the form (3-12) and (3-13) for each image. Solving these equations, the four dimensional vector \((a, b, p, q)\) is estimated. To derive the dispersion matrix of these parameters, a variance-covariance propagation can be carried out using the variance-covariance matrices of the image lines\' parameters and the EOP of the involved images.

### 3.3.3 Perspective Transformation from Object to Image Space

According to the methodology outlined in section (2-3), the location of the corresponding image lines in the second stereopair should be predicted using the 3-D object lines projected from the first stereopair and the EOP of the second pair.

Having the 3-D line \((a, b, p, q)\) and the EOP \((X_o, Y_o, Z_o, \omega_i, \phi_i, \kappa_i)\), one can derive the parameters \((U_i, V_i, W_i, D_i)\) of the object plane with respect to the object coordinate system, (see appendix C). Once again the parameters \((U_i, V_i, W_i)\) represent the direction vector of the normal to the object plane. The corresponding vector \((u_i, v_i, w_i)\) given in the image coordinate system can be obtained through the rotation matrix of the involved image according to Equation (3-14):

\[
\begin{pmatrix}
    u_i \\
    v_i \\
    w_i
\end{pmatrix} = R_i^T 
\begin{pmatrix}
    U_i \\
    V_i \\
    W_i
\end{pmatrix}
\]  

(3-14)

where:

\[R_i^T = f(\omega_i, \phi_i, \kappa_i)\]

Now, the equation of the image plane with respect to the image coordinate system can be written as:

\[u_i x + v_i y + w_i z + d_i = 0\]  

(3-15)

Since this plane passes through the origin of the image coordinate system, the perspective center, \(d_i\) must equal to zero, yielding:
\[ u_i x + v_i y + w_i z = 0 \] (3-16)

The intersection of the image plane with the focal plane \((z = -f)\) produces the image line whose equation will be:

\[ u_i x + v_i y = w_i f \] (3-17)

The parameters \((\rho_i, \theta)\) of that line are found by (3-18) and (3-19) (see Figure 10):

\[ \theta_i = \tan^{-1}(v_i / u_i) \] (3-18)

\[ \rho_i = (w_i f / u_i) \cos \theta_i \] (3-19)

Figure 10: Polar parameters as derived from image plane parameters.

Error propagation can be used to derive the variance-covariance matrix of these parameters using the dispersion matrices of both the object line and the EOP of the involved image.
Chapter IV
The Data Association Problem

4.1 Introduction

According to the strategy explained in Chapter 2, the following tasks have been performed so far:

(1) The road edges in the first stereopair were extracted by applying differential edge detection operators and Hough transform for line segmentation.

(2) Using the EOP of the first stereopair, the edges can be projected into object space. Then, the object road edges can be projected back into the image space of the second stereopair using the corresponding EOP and the model explained in Chapter 3. The projected road edges serve as the predicted road edges in the second stereopair.

(3) Search windows are established around the predicted road edges. Within these windows, edge detection operators are applied to extract all possible edges.

The problem at hand stems from the fact that the detected edges within the search windows either correspond to road edges or other linear features caused by shadows and noise. The task addressed in this chapter is finding the edges that most likely correspond to the road boundaries. This task is called the data association problem.

4.2 Problem Definition

The objective of data association techniques is to label the sensor observations, direct or derived, into sets or tracks according to their origin. Established tracks are used for state estimates. In this study, the edges derived from the stereo-imaging
system are classified into the ones originating from lane boundaries and those due to clutter, such as noise or shadows. Before these techniques, one has to distinguish between the following kinds of uncertainties associated with the observations (Cox and Leonard, 1993):

1. **Noise Uncertainties** are errors associated with the measurements. This kind of uncertainty is common to photogrammetrists. The effect of these errors can be reduced by implementing any minimization technique such as a least squares adjustment, i.e.:

   \[ e^T P e = \text{minimum} \]  

   where: \( e \) is the noise uncertainty.

2. **Data Association Uncertainties** are the uncertainties in the origin of the measurements. Solving these kinds of uncertainties can be viewed as classification, clustering, discrimination, detection, recognition, or registration (Broida, 1992).

   It is not an easy task to deciding where measurements originate from. Some of the problems that make data association problematic are (Cox and Leonard, 1993):

   (1) **The input data are noisy**: As applied to this research, the EOP of the involved images as well as the image lines' parameters are contaminated with noise.

   (2) **False or missing measurements are frequent**: This is the case when the edge detection algorithm extracts edges that do not correspond to road boundaries or when it fails to detect some of these edges.

   A "track" is defined as a sequence of measurements that are assumed to originate from the same geometric feature. In the literature, the term "beacon" is sometimes used instead of track (Durant and Leonard, 1989). In this study, the track is a static representation of a stationary geometric feature, which is the observations arising from the road boundaries. Once the track has been established, it can be used to estimate the parameters of interest, the vehicle's motion parameters between the
involved stereopairs. In summary, data association techniques are the procedures that select the measurement to be incorporated into the state estimate from among several candidates (Bar-Shalom and Fortman, 1988).

Consider a target whose track has been initialized. From this point on, the predicted measurement vector at epoch \((k+1)\) based on all observations up to epoch \(k\) will be denoted as \(\hat{z}(k+1|k)\) with the associated variance \(s_p(k+1)\). For this study, the predicted measurement vector \(\hat{z}(k+1|k)\) contains the predicted image-line parameters \((\rho, \theta)\) in the \((k+1)st\) stereopair based on all observations up to the \(k^{th}\) stereopair. \(Z^k\) will be used to denote the cumulative set of observations up to and including the \(k^{th}\) epoch. \(Z(k+1)\) is used to denote the observed quantities at the \((k+1)st\) epoch. In this research, \(Z(k+1)\) represents the set of the image line parameters corresponding to those lines detected within the search window in the \((k+1)st\) stereopair established around the predicted lines \(\hat{z}(k+1|k)\). \(z_i(k+1)\) denotes the \(i^{th}\) line in this particular set with a variance-covariance matrix of \(s_{\rho_i}(k+1)\).

Assume that all previous measurements \(Z^k\) are target originated, then the predicted measurement \(\hat{z}(k+1|k)\) will be normally distributed as follows:

\[
p[\hat{z}(k+1|k) | Z^k] = N(\hat{z}(k+1|k); z(k+1), s_p(k+1))
\]  
(4-2)

where:

\[
z(k+1) = E[\hat{z}(k+1|k)] \text{ the true measurement at the } (k+1)^{st} \text{ epoch}
\]

\[
s_p(k+1) = D(\hat{z}(k+1|k))
\]

For each observed line within the search window the difference between the predicted and the observed parameters, which will be denoted the "innovation", can be established as :

\[
v_i(k+1) = z_i(k+1) - \hat{z}(k+1|k)
\]  
(4-3)
Assuming that $z_i(k+1)$ is the target originated measurement, i.e., the one that caused by road edges, one can say that it is normally distributed,

$$p[z_i(k+1)|i^{th} \text{ measurement is target originated}] = N[z_i(k+1); z(k+1), s_z(k+1)]$$ (4-4)

From equations (4-2) and (4-4), it appears that the innovation $v_i(k+1)$ defined by equation (4-3), conditioned on the fact that the $i^{th}$ measurement is target originated, is also normally distributed:

$$p[v_i(k+1)|i^{th} \text{ measurement is target originated}] = N[v_i(k+1); 0, s_v(k+1)]$$ (4-5)

where:

$$s_v(k+1) = s_z(k+1) + s_r(k+1)$$ (4-6)

It has to be noticed that $z_i(k+1)$ and $\hat{z}(k+1|k)$ are statistically independent. This is why the variance of the innovation, $s_v(k+1)$, is written as the sum of the corresponding variances $s_z(k+1)$ and $s_r(k+1)$, respectively.

Now, one can define a region in the measurement space, $\tilde{V}_{k+1}(\gamma)$, where the target originated measurement will be found with high probability (Bar-Shalom and Fortman, 1988):

$$\tilde{V}_{k+1}(\gamma) = \{z_i(k+1); v_i^T(k+1)s_v^{-1}(k+1)v_i(k+1) \leq \gamma\}$$ (4-7)

The region $\tilde{V}_{k+1}(\gamma)$ will be denoted as the "validation region or gate". It is the ellipse of probability concentration, the region of minimum value that contains a given probability mass under Gaussian assumptions. The function of the validation region is as follows: measurements that lie outside this region are discarded and measurements that are inside are considered as candidates for being target originated. The value of the parameter $\gamma$ can be obtained from Chi-square distribution tables. This is due to the fact that the weighted norm of innovation, $v_i^T(k+1)s_v^{-1}(k+1)v_i(k+1)$- which is called
statistical or Mahalanobis distance—follows a Chi-square distribution with the number of degrees of freedom equal to $n_z$, the size of the observation vector (Koch, 1988). For this research, $n_z$ equals 2 corresponding to $(\rho, \theta)$ of the line under consideration.

By pre-specifying the probability $P_G$, that the true measurement $z(k+1)$ will fall within the validation region, (Eq. 4-8); one can obtain the value of $\gamma$ by:

$$P_G = P\{z(k+1) \in \tilde{V}_{k+1}(\gamma)\} \quad (4-8)$$

Table 1 gives some options for $\gamma$ for different probability mass quantities, $P_G$, within the validation region or gate.

Table 1: Gate threshold and values of probability mass in the gate.

<table>
<thead>
<tr>
<th>$P_G$</th>
<th>0.9</th>
<th>0.95</th>
<th>0.99</th>
<th>0.995</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>4.61</td>
<td>5.99</td>
<td>9.21</td>
<td>10.60</td>
</tr>
</tbody>
</table>

The above mentioned validation procedure limits the region within the measurement space where the data association technique looks for the measurement that is most likely target-originated. Usually, more than one measurement is found in the validation region due to clutter or background noise. The set of measurements that belong to the validation region at the $(k+1)^{th}$ epoch is denoted:

$$Z_v(k+1) \triangleq \{z_v(k+1): z_v(k+1) \in \tilde{V}_{k+1}(\gamma)\} \quad (4-9)$$

Now, the problem is to associate each validated measurement with the appropriate target or to discard it as arising from noise, such as shadows. The solution to this problem is offered by data association techniques.
4.3 Single Target in a Clutter

4.3.1 Introduction

Up to now, the validated measurements that lie within the validation region have been identified according to the procedure outlined in the previous section. The validated measurements include:

(1) The correct measurement, i.e., the one that is originated from the target being tracked- the road boundaries.

(2) The incorrect measurement arising from noise or shadows.

Figure 11 shows a two-dimensional validation region which is defined by an ellipse centered around the predicted measurement, \( \hat{z}(k+1|k) \). Once again, the validation region is two-dimensional, since the measurement vector has only two components \((\rho, \theta)\). The size and the shape of the ellipse is determined by the eigen values and the eigen vectors of the variance-covariance matrix of the innovation, \( s_c(k+1) \). Let us assume that the variance-covariance matrix of the observed measurement, \( s_c(k+1) \) is identical for all the measurements. This assumption is only made for simplifying the graphical interpretation, since it leads to the same validation ellipse for all the measurements, see Equation (4-6).
Figure 11 can be thought of as a graphic of the situation depicted in Figure 12, where three edges are observed within the search window established around the predicted edge corresponding to the road boundary.
The crux of the problem is to determine which one of the observed edges will most likely correspond to the road boundary.

### 4.3.2 Nearest Neighbor Standard Algorithm

In the Nearest Neighbor Standard Algorithm (NNSA), the validated measurement nearest to the predicted one is assumed to be target originated. The measure to be used as indication of the distance is the weighted norm of innovation:

\[
2 \{z_i(k+1)\} = d^2_i(k+1) = \{z_i(k+1) - \tilde{x}(k+1|k)\}^T s_i^{-1}(k+1) \{z_i(k+1) - \tilde{x}(k+1|k)\} = v_i(k+1) s_i^{-1}(k+1) v_i(k+1) \quad (4-10)
\]

The validated measurements with the minimum distance, \(d_i(k+1)\), are used for updating the state estimate, namely the motion parameters of the vehicle between epoches k, k+1.

This procedure is performed twice for each target: first in the search window of the left image, and then in the right image of the stereopair under consideration. The problem of this algorithm is that it deals with each track separately without considering the fact that some of the tracks may compete for the same observations, i.e., the same observation falls in the validation region of two different targets (see Figure 13). Fortunately, this problem is considered in multiple targets in clutter data association algorithms.

### 4.4 Multiple Targets in a Clutter

#### 4.4.1 Introduction

The major disadvantage of NNSA is that it ignores the possibility of having nearby targets compete for the same measurement. In this case, the term clutter has to
be extended to include false observations arising from nearby targets. Figure 13 shows two targets $\hat{z}_1(k+1|k)$ and $\hat{z}_2(k+1|k)$ with overlapping validation regions with one common validated measurement. Figure 13 can correspond to the situation depicted in Figure 14. The following can be concluded from Figure 13:

1. Measurements $z_1(k+1)$ and $z_2(k+1)$ could have originated from either the target represented by the predicted measurement $\hat{z}_1(k+1|k)$ or from noise.
2. Measurement $z_4(k+1)$ could have originated from either the target represented by $\hat{z}_2(k+1|k)$ or noise.
3. Measurement $z_3(k+1)$ could have originated from either targets represented by $\hat{z}_1(k+1|k)$, $\hat{z}_2(k+1|k)$, or noise.

Figure 13: Measurements in the validation regions of two targets.

Figure 13 could result from the observed edges depicted in Figure 14. In the single target in clutter algorithms, the following situation may take place: measurement $z_3(k+1)$ is statistically the closest validated measurement to both $\hat{z}_1(k+1|k)$ and
In this situation, the algorithm assigns the same observation to the same target, which is impossible. In such a case, multiple targets in a clutter solve the problem by creating association hypotheses that associate the measurements to all targets at the same time. In each association hypothesis the same measurement cannot be assigned to more than one target.

![Diagram of observed and predicted edges within the search window of two targets.](image)

**Figure 14:** Observed and predicted edges within the search window of two targets.

### 4.4.2 Multiple Hypothesis Approach

After the discussion in the previous section, it is evident that the association of measurements has to be done by considering all targets simultaneously when several targets are being tracked. This is due to the fact that there might be cases where one measurement falls in the validation regions of two or more other targets indicating that this measurement could originate from either one of them.

The key to the Multiple Hypothesis Approach (MHA) is forming several association hypotheses that represent all possible assignments of the observed edges to the targets being tracked (Bar-Shalom and Fortman, 1988). Then one has to find the hypothesis which most likely is correct by assigning some kind of cost for each
hypothesis. The cost indicates the amount of uncertainty of a particular hypothesis. The association hypothesis with the minimum cost is considered for state estimate updating.

In MHA, one starts by building the association events pertaining to the current epoch $k+1$:

$$\Theta(i) = \bigcap_{j=1}^{n_{k+1}} \theta_{j,i}$$

(4-11)

where,

$$\theta_{j,i} \doteq \{\text{measurement } 'j' \text{ originated from target } 't' \}$$

$j = 1, \ldots, n_{k+1}$ is the index for the measurement under consideration,

$k_{k+1}$ is the number of validated measurements,

$t = 1, \ldots, T$ is index for the targets,

$T$ is the number of targets involved.

Here, one can consider that each measurement is validated for each target, i.e., the validation gate of each target coincides with the whole surveillance region, the whole image in our case. The disadvantage of this approach is that the number of association events, $N$:

$$N = \frac{m_{k+1}!}{(m_{k+1} - T)!}$$

(4-12)

can become very large when too many measurements are available. This also means that association events with minute probabilities, extremely high cost, will be still considered. For this research, the number of observed edges within the search window is not too big and does not lead to an unmanageable number of association events. If this is the case, the validation gates can be used to only consider feasible events. For
example, referring to Figure 4-3, an association hypothesis that assumes the measurement \( z_{k+1} \) originated from the target \( \hat{z}_1(k+1|k) \) will be discarded.

Now, let us define the association matrix:

\[
\Omega_j = [\omega_{jt}]
\]

with \( m_{k+1} \) rows corresponding to the observed measurements and \( T \) columns corresponding to the involved targets. The binary element \( \omega_{jt} \) indicates whether measurement number \( j \) is assumed to originate from target \( t \), \( \omega_{jt} = 1 \), or not, \( \omega_{jt} = 0 \), for the association event under consideration, \( \Theta(t) \). Each association matrix represents an association event. Some possible association matrices that pertain to Figure 13 are:

\[
\Omega_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}
\]

\[
\Omega_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}
\]

A feasible association matrix \( \Omega_j \) is one satisfying the following conditions:

(1) A measurement can have only one source. In other words, it can be assigned to at most one target:

\[
\sum_{t=1}^{r} \omega_{jt}(i) \leq 1 \quad \text{for } j = 1, \ldots, m_{k+1}
\]

(4-15)

(2) Each target can give rise to only one measurement:
\begin{align}
\sum_{j=1}^{m_{k+1}} \omega_{jt}(i) = 1 \quad \text{for } t = 1, \ldots, T \quad (4-16)
\end{align}

This means that the association matrix is only valid, if no more than one unit can be found in each row or column.

The next step is to build the cost matrix \( C \):

\[ C \doteq [c_{jt}] \quad (4-17) \]

where:

\begin{align*}
j &= 1, \ldots, m_{k+1} \\
t &= 1, \ldots, T \\
c_{jt} &\text{ is the weighted norm of innovation resulting from associating the } j^{th} \text{ measurement to the } t^{th} \text{ target (Eq. 4-18).}
\end{align*}

\begin{align}
c_{jt} &= \{ z_{j} (k + 1) - \hat{z}_{t} (k + 1|k) \}^{T} \{ s_{j0} (k + 1) + s_{pt} (k + 1) \}^{-1} \\
&\quad \{ z_{j} (k + 1) - \hat{z}_{t} (k + 1|k) \} \\
&\quad (4-18)
\end{align}

In other words, \( c_{jt} \) indicates the cost arising from associating target number 'j' with the \( t^{th} \) target.

For each of the association matrices, the association cost \( C(i) \) can be computed as follows:

\[ C(i) = \sum_{j=1}^{m_{k+1}} \sum_{t=1}^{T} \omega_{jt}(i) \cdot c_{jt} \quad (4-19) \]

The association event \( \Theta(i) \) with the minimum cost \( C(i) \) will be assumed to be the correct one, and it will be used for state estimate updating.
Next, the statistical meaning of the association cost, \( C(i) \), is being explored. The likelihood function of the association event \( \Theta(i) \) being the correct one can be expressed by the joint probability density function:

\[
\Lambda(\Theta_i) = \prod_{j=1}^{n_{\text{all}}} P\{z_j(k+1)|\Theta_i\} \tag{4-20}
\]

Under the Gaussian assumption, the measurement \( z_j(k+1) \) is normally distributed with a mean of \( z_j(k+1) \) and a variance-covariance matrix of \( s_{j_i}(k+1) \) defined as:

\[
s_{j_i}(k+1) = s_{j_i}(k+1) + r_{j_i}(k+1) \tag{4-21}
\]

if in the event \( \Theta(i) \), the \( j^\text{th} \) measurement has been associated with the \( t^\text{th} \) target:

\[
P\{z_j(k+1)|\omega_{ji} = 1\} = N[z_j(k+1); z_{ji}(k+1), s_{ji}(k+1)] \tag{4-22}
\]

or using the corresponding innovation:

\[
P\{v_{ji}(k+1)|\omega_{ji} = 1\} = N[v_{ji}(k+1); 0, s_{ji}(k+1)] \tag{4-23}
\]

From equations (4-20) and (4-23), it follows that:

\[
\Lambda(\Theta_i) = c_k \exp[-1/2 \sum_{j=1}^{n_{\text{all}}} \sum_{t=1}^{T} v_{ji}(k+1) s_{ji}^{-1}(k+1) v_{ji}(k+1) \omega_{ji}(i)] \tag{4-24}
\]

where \( c_k \) is constant.

The modified likelihood function corresponding to (4-24) is shown in (4-25):

\[
\lambda(i) = -2 \log \left\{ \frac{\Lambda(\Theta_i)}{c_k} \right\} = \sum_{j=1}^{n_{\text{all}}} \sum_{t=1}^{T} v_{ji}^T(k+1) s_{ji}^{-1}(k+1) v_{ji}(k+1) \omega_{ji}(i) \tag{4-25}
\]

From equation (4-19), one gets:

\[
\lambda(i) = C(i) \tag{4-26}
\]

It has to be noted that the association cost \( c(i) \) is a Chi-squared distribution with \( (T \times n_z) \) degrees of freedom where:

\( T = \) the total number of targets being tracked,

\( n_z = \) the size of the measurement vector.
This means that the modified log-likelihood function of the right association event \( \Theta(i) \) can be expressed by \( C(i) \). This is actually nothing but a measure of the "goodness of fit" (Bar-Shalom and Fortman, 1988). The statistical testing for accepting an association event is that the log-likelihood function satisfies:

\[
\lambda(i) \leq a \tag{4-27}
\]

where, the threshold \( a \) follows from the chi-square tables for \( (T \times n) \) degrees of freedom:

\[
P\{\chi^2_{n} \geq a\} = \alpha \tag{4-28}
\]

with \( \alpha \) being the probability of making an error of the first type (Rice, 1988), i.e., a true track will be rejected.
CHAPTER V

Updating of Motion Parameters

5.1 Introduction

In the previous chapter, corresponding linear features at successive stereopairs were found by data association algorithms. The matched features are used to update the motion parameters of the mapping vehicle. Substantial work has been done on the estimation of sensor orientation in image sequences using distinct points Bandopadhay and Ballard, 1986, Kersten and Baltsavias, 1992, Gruen and Kersten, 1992, and Edmundson and Novak, 1992.

In this study, a new algorithm was developed for updating the motion parameters using linear features. Linear features that correspond to the road boundaries can be extracted from road images. Faugeras et al. (1986) introduced a model that can be used to find relative displacements by matching straight lines. In this model, each straight line contributes two equations. Since 3-D lines have four degrees of freedom in space, more information should be available from lines. Therefore, constraining two straight lines to be the same in consecutive image pairs should result in four independent constraints that depend on the parameters of the lines, as well as the displacement under consideration.

Ayache and Faugeras (1989) described the changes in the four-dimensional representation due to rotation and translation. The suggested algorithm is based on their work. The methodology of the suggested algorithm works as follows.

First, the linear features matched in successive stereopairs are projected into model space with its origin at the perspective center of the left camera of the imaging
The origin of the model system could be chosen differently; for example, the phase center of the GPS antenna could be chosen as the origin of the model space. It has to be noted that the projection procedure is performed separately for the stereopairs under consideration. For each 3-D line in the model space of the second stereopair, a rotation 'R' and displacement 'T' are defined, which represent the relative displacement between the first and second stereopairs. The resulting 3-D line should be equivalent to the corresponding 3-D line projected into the model space of the first stereopair. The equivalency is obtained by imposing four constraints on the four degrees of freedom of the two lines of interest.

5.2 The Algorithm

As mentioned before, one starts by projecting the matched features into the model space associated with the stereopair, e.g., defined at the perspective center of the left camera. This is done separately for each of the successive stereopairs. As explained in Chapter 3, the 3-D line is represented by a point, (p, q, 0), which is defined by the intersection of the line with the x-y plane and a direction vector (a, b, 1). The point and the direction vector can be chosen differently to overcome some singularities, which exist if the 3-D line is parallel to the x-y plane or perpendicular to the z-axis. For the sake of generality let (l, m, n) and (x, y, z) denote the chosen point and the direction vector, respectively. Assume that \((l_1, m_1, n_1)\) and \((x_1, y_1, z_1)\) represent a specific 3-D line with respect to the model space of the first stereopair. Also, suppose that the corresponding 3-D line in the model space of the second stereopair is represented by \((l_2, m_2, n_2)\) and \((x_2, y_2, z_2)\).
Figure 15: Relationship between the model spaces $M_1$, and $M_2$ and the object coordinate system $O$.

Figure 15 shows the relationship between the model space of the first stereopair $M_1$, the model space of the second stereopair $M_2$, and the object coordinate system $O$. This relationship can be established as follows:

(1) The model associated with $M_1$ can be transferred into the object system through a rotation $R_i$ and a displacement $T_i$. This means that vector $(l_i, m_i, n_i)$ can be transformed into the object coordinate system as follows:

$$
\begin{pmatrix}
L_i \\
M_i \\
N_i
\end{pmatrix}
= R_i
\begin{pmatrix}
l_i \\
m_i \\
n_i
\end{pmatrix} \quad (5-1)
$$

Point $(x_1, y_1, z_1)$ will become:
(2) In a similar fashion, the model associated with \( M_2 \) can be transformed into the object system through the rigid displacement defined by \( R_2 \) and \( T_2 \). Vector \( (L_2, M_2, N_2) \) and point \( (X_2, Y_2, Z_2) \) that correspond to \( (l_2, m_2, n_2) \) and \( (x_2, y_2, z_2) \), respectively, can be computed as follows:

\[
\begin{align*}
\begin{pmatrix}
L_2 \\
M_2 \\
N_2
\end{pmatrix} &= R_2 \begin{pmatrix}
l_2 \\
m_2 \\
n_2
\end{pmatrix} + T_2 \quad (5-3)
\end{align*}
\]

The two direction vectors described by equations (5-1) and (5-3) represent the same 3-D line with respect to the object coordinate system. The relationship between their components can be described analytically by introducing a scale factor \( \lambda \):

\[
\begin{align*}
\begin{pmatrix}
L_2 \\
M_2 \\
N_2
\end{pmatrix} &= \lambda \begin{pmatrix}
L_1 \\
M_1 \\
N_1
\end{pmatrix} \\
\begin{pmatrix}
l_2 \\
m_2 \\
n_2
\end{pmatrix} &= \lambda \begin{pmatrix}
l_1 \\
m_1 \\
n_1
\end{pmatrix} \quad (5-5)
\end{align*}
\]

Substituting equations (5-1) and (5-3) in equation (5-5), the following expression is obtained:

\[
\begin{align*}
R_2 \begin{pmatrix}
l_2 \\
m_2 \\
n_2
\end{pmatrix} &= \lambda R_1 \begin{pmatrix}
l_1 \\
m_1 \\
n_1
\end{pmatrix} \\
(5-6)
\end{align*}
\]
Since the rotation matrices are orthogonal, equation (5-6) can be rewritten as:

\[
\begin{pmatrix}
  l_1 \\
  m_1 \\
  n_1
\end{pmatrix}
= R_1^T R_2
\begin{pmatrix}
  l_2 \\
  m_2 \\
  n_2
\end{pmatrix}
\]  
(5-7)

By examining Equation (5-7), one can conclude that the product \( R_1^T R_2 \) represents the rotation matrix that relates the model space of the second stereopair to that of the first one. This product will be denoted as \( R \).

To eliminate scale factor \( \lambda \) in equation (5-7), the first and the third equations are divided by the second one:

\[
\frac{l_1}{m_1} = \frac{r_{11} l_2 + r_{12} m_2 + r_{13} n_2}{r_{21} l_2 + r_{22} m_2 + r_{23} n_2}
\]
(5-8)

\[
\frac{m_1}{m_1} = \frac{r_{11} l_2 + r_{12} m_2 + r_{13} n_2}{r_{21} l_2 + r_{22} m_2 + r_{23} n_2}
\]
(5-9)

where: \( r_{ij} \) are the elements of rotation matrix \( R \).

Equations (5-8) and (5-9) represent the first two equations that represent the equivalency between the two direction vectors \((l_1, m_1, n_1)\) and \((l_2, m_2, n_2)\). These constraints involve only the rotation matrices \( R_1 \) and \( R_2 \).

Now another constraint that incorporates the displacement vectors \( T_1, T_2 \) is introduced. Point \((X_1, Y_1, Z_1)\) must belong to the 3-D line, which is given by point \((X_2, Y_2, Z_2)\) and the direction vector \((L_2, M_2, N_2)\) (see Figure 16). This can be expressed as:

\[
\begin{pmatrix}
  X_1 \\
  Y_1 \\
  Z_1
\end{pmatrix}
= \begin{pmatrix}
  X_2 \\
  Y_2 \\
  Z_2
\end{pmatrix}
+ s \begin{pmatrix}
  L_2 \\
  M_2 \\
  N_2
\end{pmatrix}
\]
(5-10)

where \( s \) is a scale factor.
Substituting equations (5-2), (5-3) and (5-4) in equation (5-10), one gets:

\[
R_1 \begin{bmatrix} x_1 \\ y_1 \\ x_1 \end{bmatrix} + T_1 = R_2 \begin{bmatrix} x_2 \\ y_2 \\ x_2 \end{bmatrix} + T_2 + s R_2 \begin{bmatrix} l_2 \\ m_2 \\ n_2 \end{bmatrix}
\]  

(5-11)

Once again, due to the orthogonality of the rotation matrices, equation (5-11) can be written as:

\[
\begin{bmatrix} x_1 \\ y_1 \\ x_1 \end{bmatrix} - R_1^T (T_2 - T_1) - R_1^T R_2 \begin{bmatrix} x_2 \\ y_2 \\ x_2 \end{bmatrix} = s R_1^T R_2 \begin{bmatrix} l_2 \\ m_2 \\ n_2 \end{bmatrix}
\]  

(5-12)

The term \( R_1^T (T_2 - T_1) \) represents the translation vector of the second model space with respect to the first one. This translation vector is denoted as \( T \). Scale factor \( s \) can be
eliminated by dividing the first and third equations of (5-12) by the second equation, resulting in:

\[
\begin{align*}
\frac{x_1 - (r_{11} x_1 + r_{12} y_1 + r_{13} z_2)}{t_x} &= \frac{r_{11} l_2 + r_{12} m_2 + r_{13} n_2}{r_{21} l_2 + r_{22} m_2 + r_{23} n_2} \\
y_1 - (r_{21} x_1 + r_{22} y_1 + r_{23} z_2) - t_y &= \frac{r_{21} l_2 + r_{22} m_2 + r_{23} n_2}{r_{21} l_2 + r_{22} m_2 + r_{23} n_2} \\
z_1 - (r_{31} x_1 + r_{32} y_1 + r_{33} z_2) - t_z &= \frac{r_{31} l_2 + r_{32} m_2 + r_{33} n_2}{r_{31} l_2 + r_{32} m_2 + r_{33} n_2} \\
y_1 - (r_{21} x_1 + r_{22} y_1 + r_{23} z_2) - t_y &= \frac{r_{21} l_2 + r_{22} m_2 + r_{23} n_2}{r_{21} l_2 + r_{22} m_2 + r_{23} n_2}
\end{align*}
\]

(5-13) (5-14)

where \((t_x, t_y, t_z)\) are the components of translation vector \(T\).

Equations (5-8), (5-9), (5-13) and (5-14) serve as the constraints that ensure the correspondence of the matched feature in successive stereopairs. In those equations \((l_1, m_1, n_1), (x_1, y_1, z_1), (l_2, m_2, n_2), (x_2, y_2, z_2)\), and the rotation angles forming \(R_1\) and \(T_1\) are observed quantities while \(T_2\) and the three rotation angles of the rotation matrix \(R_2\) are unknowns. The suitable adjustment model is "condition equations with parameters", (Koch, 1988):

\[
f(y - e, \Xi) = 0 \quad \text{with} \quad e \sim N(0, \sigma_e^2 P^{-1})
\]

(5-15)

where:

- \(y\) is the observation vector,
- \(e\) is the noise contaminating the observation vector,
- \(\Xi\) is the vector of unknowns, and
- \(\sigma_e^2 P^{-1}\) is the variance-covariance matrix of the noise.

This model can be linearized through Taylor's expansion. Ignoring second and higher order terms yields:

\[
\frac{\partial f}{\partial y} |_{y, \Xi_o} (-e) + \frac{\partial f}{\partial \Xi} |_{y, \Xi_o} (\Xi - \Xi_o) + f(y, \Xi_o) = 0
\]

(5-16)

where, \(\Xi_o\) contains the approximate values for the unknown parameters.
Denoting $\frac{\partial f}{\partial y_{i,\Xi_o}}$ with $B$, $f(y,\Xi_o)$ with $-w$, $(\Xi-\Xi_o)$ with $\xi$ and $\frac{\partial f}{\partial \Xi_{i,\Xi_o}}$ with $A$, equation (5-16) can be written as:

$$w = A \xi + Be$$

with $w \sim N(\xi, \sigma_o^2 BP^{-1}B^T)$ (5-17)

Each feature contributes four equations (5-8), (5-9), (5-13) and (5-14). If prior information is available for the unknown parameters, it can be introduced in the form of pseudo observations. The solution vector and the corresponding variance-covariance matrix are computed as:

$$\hat{\xi} = (A^T BP^{-1}B^T)^{-1} A^{-1} (A^T BP^{-1}B^T)^{-1} w$$

(5-18)

$$\sigma_o^2 = \frac{\left(w-A \hat{\xi}\right)^T \{BP^{-1}B^T\}^{-1} \{w-A \hat{\xi}\}}{r}$$

(5-19)

An estimate of the variance component is obtained by:

where, $r$ is the redundancy of the system.

### 5.3 Expected Singularity

Each 3-D line tracked in successive stereopairs contributes four equations towards the computation of the displacement parameters between consecutive stereopairs. This means that by matching one 3-D line in two consecutive stereopairs, two of the six displacement parameters cannot be solved for. One can expect that these two parameters would be the displacement in the direction of the line, as well as the rotation angle around this direction, roll angle. This expectation was confirmed by an eigen-value analysis of a normal equation matrix that considered only one line parallel to the direction of the vehicle's movement. The displacement in the movement direction together with the roll angle could not be recovered. On the other hand, considering two or more parallel lines, the singularity drops to 1. This is due to the fact
that the system can solve for all parameters except the displacement in the direction parallel to those lines. By adding cross lines, one can resolve all displacement parameters. The expected singularities arising from different configurations are explained in Figure 17 (a-c). Unfortunately, the available configuration for this research is the one depicted in Figure 17 (b). This means that the displacement in the direction of the linear feature could not be resolved. This displacement can be obtained from another sensor onboard the mapping vehicle, such as the wheel counter in the case of the GPSVan.
(a) Two singularities corresponding to the roll angle and displacement in the direction of movement.

(b) One singularity corresponding to the displacement in the direction of movement.

(c) No singularity.

Figure 17: Different configurations of linear features and expected singularities.
CHAPTER VI
FINAL RESULTS

6-1 Introduction

In this chapter, the algorithms outlined in Chapters 3, 4 and 5 are implemented. The detection, perspective transformation, data association, and motion parameter updating modules were incorporated in a computer program. This program is used for processing image data captured by the GPSVan of the Center For Mapping of The Ohio State University. In the following sections, the results obtained from two data sets are presented. In the first data set, the edges bounding the central lane boundary are tracked. In the second one, the edges of the left and right lane boundaries are being tracked.

6-2 Results from First Data Set

In this data set the edges corresponding to the central lane boundary serve as targets for tracking. The left and right images of the first stereopair are shown in Figures (18) and (19), respectively. In this stereopair the user is prompted to select patches in the left and right images where the targets of interest are expected. The selection process is based on supplying the row and the column numbers of the upper left corner as well as the number of rows and columns of the patch to be selected. The patches selected in the left and right images are displayed in Figures (20) and (21), respectively.

An edge detection operator is applied to both patches to mark edge pixels. The resulting edge images are shown in Figures (22) and (23) for the left and right patches,
respectively. In the following step, the Hough transform is applied to segment the edge pixels and find the ones arising from straight lines. As mentioned in Chapter 2, the edge pixels along straight lines will appear as clusters in Hough space. The centroids of these clusters represent the lines of interest. The Hough space images as well as the detected

Figure 18: Left image of the first stereopair of data set 1.
Figure 19: Right image of the first stereopair of data set 1.
Figure 20: The left patch of the first stereopair of data set 1.
Figure 21: The right patch of the first stereopair of data set 1.
Figure 22: Edge image corresponding to the left patch of the first stereopair of data set 1.
Figure 23: Edge image corresponding to the right patch of the first stereopair of data set 1.
clusters can be seen in Figures (24) and (25). The centroid of each cluster corresponds to the point of intersection of the black cross. Also, the dispersion of each centroid is chosen as the variances of the parameters of the line represented by this cluster. The lines corresponding to the detected centroids are shown in Figures (26) and (27).

As the relative orientation parameters of the stereo-vision system are known from previous calibration, one can project the detected edges into the model space associated with the first stereopair. The perspective transformation of the image lines into model space was explained in Section 3-3. The EOP of the second stereopair can be approximated by assuming no relative rotation between the first and the second stereopair and a displacement only in the direction of the road. The displacement along road direction can be obtained by another sensor onboard the mapping system, such as the wheel counter of the GPSVan or merely by multiplying the speed of the vehicle by the time period between the moments of exposure. Now, one can transform the local world model of the road associated with the first stereopair into one associated with the second stereopair. This model can be projected into the image space of the second stereopair using the algorithm outlined in Section 3-4. This projection yields the predicted road edges in the second stereopair. These edges, in the left and right images of the second stereopair, are displayed in Figures (28) and (29), respectively.

The predicted road edges are used to automatically extract the search windows in the second stereopair, where the tracked targets are expected. The extracted search windows for the left and right images of the second stereopair are shown in Figures (30) and (31). As before, edge detection and segmentation algorithms are applied to these patches. The results of these processes are portrayed in Figures (32) to (37). As can be seen in Figures (36) and (37) nine and eight edges were observed inside the search windows associated with the left and right images of the second stereopair.
Figure 24: Hough space image and detected clusters associated with the left patch of the first stereopair of data set 1.
Figure 25: Hough space image and detected clusters associated with the right patch of the first stereopair of data set 1.
Figure 26: Detected edges in the left patch of the first stereopair of data set 1.
Figure 27: Detected edges in the right patch of the first stereopair of data set 1.
Figure 28: Predicted road edges in the left image of the second stereopair of data set 1.
Figure 29: Predicted road edges in the right image of the second stereopair of data set 1.
Figure 30: Automatically extracted left patch in the second stereopair of data set 1.
Figure 31: Automatically extracted right patch in the second stereopair of data set 1.
Figure 32: Edge image corresponding to the left image of the second stereopair of data set 1.
Figure 33: Edge image corresponding to the right image of the second stereopair of data set 1.
Figure 34: Hough space image and detected clusters associated with the left patch of the second stereopair of data set 1.
Figure 35: Hough space image and detected clusters associated with the right patch of the second stereopair of data set 1.
Figure 36: Detected edges in the left patch of the second stereopair of data set 1.
Figure 37: Detected edges in the right patch of the second stereopair of data set 1.
Now, the problem at hand is to find out which of the observed edges in the second stereopair correspond to the two targets previously initialized in the first stereopair (the boundaries of the dashed road centerline). As mentioned in Chapter 4, the solution to this problem is offered through data association techniques. For this purpose a cost matrix is created for both images, with the number of rows equal to the number of observed edges and the number of columns equal to the number of targets being tracked. The element in the $i^{th}$ row and the $j^{th}$ column is the weighted norm of innovation, which was explained in section 4-3-2, between the $i^{th}$ observation and the prediction associated with the $j^{th}$ target. The cost matrices for the left and right images are shown in Tables (2) and (3), respectively.

Table 2: The cost matrix associated with the left image of data set 1.

<table>
<thead>
<tr>
<th>Observation #</th>
<th>Target # 1</th>
<th>Target # 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26.18</td>
<td>28.15</td>
</tr>
<tr>
<td>2</td>
<td>9.65</td>
<td>10.95</td>
</tr>
<tr>
<td>3</td>
<td>11.21</td>
<td>12.49</td>
</tr>
<tr>
<td>4</td>
<td>10.82</td>
<td>11.55</td>
</tr>
<tr>
<td>5</td>
<td>12.92</td>
<td>13.44</td>
</tr>
<tr>
<td>6</td>
<td>4.74</td>
<td>3.11</td>
</tr>
<tr>
<td>7</td>
<td>1.50</td>
<td>1.04</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>1.12</td>
</tr>
<tr>
<td>9</td>
<td>0.94</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Table 3: Cost matrix associated with right image of data set 1.

<table>
<thead>
<tr>
<th>Observation #</th>
<th>Target # 1</th>
<th>Target # 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>108.94</td>
<td>105.67</td>
</tr>
<tr>
<td>2</td>
<td>195.23</td>
<td>182.88</td>
</tr>
<tr>
<td>3</td>
<td>138.57</td>
<td>133.39</td>
</tr>
<tr>
<td>4</td>
<td>154.96</td>
<td>145.58</td>
</tr>
<tr>
<td>5</td>
<td>19.33</td>
<td>18.57</td>
</tr>
<tr>
<td>6</td>
<td>0.55</td>
<td>0.43</td>
</tr>
<tr>
<td>7</td>
<td>0.63</td>
<td>0.47</td>
</tr>
<tr>
<td>8</td>
<td>41.63</td>
<td>39.96</td>
</tr>
</tbody>
</table>

The first data association technique mentioned in Chapter 4 is the Nearest Neighbor Standard Algorithm (NNSA). This algorithm chooses the observation that is statistically closest to each predicted target. By inspecting Table 2, one can see that the NNSA would choose the ninth observation to be the one arising from both targets being tracked, which is evidently wrong. Similarly, the NNSA algorithm will select the sixth observation to be the one arising from the left and right boundaries of the central dividing line, see Table 3. The results of these matches are shown in Figures 38 and 39. The wrong results are due to the fact that the NNSA selects the observation that is believed to be target originated for each target individually without considering the competition of several targets for the same observation. Ignoring this fact will lead to the possibility of assigning the same observation to several targets which is exactly what is happening here.
Figure 38: Matched edges in the left image of the second stereopair of data set 1, NNSA.
Figure 39: Matched edges in the right image of the second stereopair of data set 1, NNSA.
This severe drawback of the NNSA is overcome by implementing the Multiple Targets In Clutter (MTIC) association technique, see Section 4-4. In the MTIC association algorithm, several binary association matrices are generated considering the following two conditions:

(1) Each target can give rise to only one observation,

(2) Each observation can have only one source, see Section 4-4-2.

These conditions lead to the creation of $N$ association matrices, where $N$ is defined as:

$$N = \frac{m!}{(m-T)!}$$  \hspace{1cm} (6-1)

where: $T$ is the number of targets being tracked, and $m$ is the number of observed edges.

This means that 72 association matrices are generated for the left image of the second stereopair, while 56 association matrices are built for the right image. For each association matrix, a corresponding cost is evaluated according to Equation 4-20. Then the association matrix with the minimum cost is chosen as the association event that is believed to be true.

Among the 72 association events generated for the left image, the association matrix:
with a cost of 1.96 is found to be the optimum one. This association matrix suggests that the eighth observation originated from the first target (the left boundary of the central dividing line) and the ninth observation originated from the second target (the right boundary).

Similarly, the association matrix:

\[
\begin{bmatrix}
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 0 \\
\end{bmatrix}
\]  

(6-3)
with a cost of 1.02 is chosen as the optimum association matrix for the right image. Again, this matrix suggests that the sixth observation is believed to originate from the first target while the second target gave rise to the seventh observation.

The matched lines for the left and right images of the second stereopair can be seen in Figures 40 and 41 respectively. As mentioned in Section 4-4-2, the cost of an association matrix is the modified likelihood function of the corresponding association event being the right one, i.e., its components are target originated. Also, it was mentioned that the cost follows a Chi-square distribution. Its degree of freedom equals the number of targets multiplied by the number of observations corresponding to each target. For this data set, the cost follows a Chi-square distribution with four degrees of freedom. Both association events for the left and right images are accepted with a level of significance of 0.1, \( \chi^2_4(0.9) = 7.78 \).

Accepting the association events corresponding to the above mentioned association matrices, one can use the matched lines in the successive stereopairs to estimate the motion parameters. This is done by the algorithm suggested in Chapter five. Using an eigen value analysis of the normal equation matrix, it was found that for this data set the displacement in the direction of the road and the roll angle cannot be solved for. This deficiency conforms with the expected singularities outlined in Section 5-3, since the targets being tracked are very close to each other and can be considered a single line (see Figure 17, a).

By constraining the displacement in the road direction and the roll angle to the values obtained from the wheel counter and the gyros on board the GPSVan, one can solve for the remaining four parameters of the relative displacement between the stereopairs under consideration. They were found to be:
Figure 40: Matched edges in the left image of the second stereopair of data set 1, MTIC.
Figure 41: Matched edges in the right image of the second stereopair of data set 1, MTIC.
tx  tz  Azimuth  Pitch
0.26m  0.46m  -2.07°  0.64°

with a variance-correlation matrix of:

\[
\begin{pmatrix}
0.000637m^2 & 0.84 & -0.025 & 0.011 \\
0.00045m^2 & -0.144 & 0.130 \\
0.000208rad^2 & -0.94 \\
0.000159rad^2 &
\end{pmatrix}
\]

where, the y axis is pointing in the road direction, and the z axis is pointing upwards.

6-3- Results Obtained from Second Data Set

In the second data set, the road edges of the lane boundaries to the right and left of the mapping system are tracked. The results of the detection and perspective transformation modules are displayed in Figures 42 to 61. Here, it has to be mentioned that the edges detected in Figure 50 corresponding to the far left lane boundary were not tracked to the second stereopair. This means that only five targets are being tracked from the first to the second stereopair. Within the search windows established in the left image of the second stereopair, eight edges were observed, see Figure 60. Seven edges were detected in the search window of the right image of the second stereopair (Figure 61). The corresponding cost matrices for the left and right images can be seen in Tables 4 and 5, respectively.
Figure 42: Left image of the first stereopair of data set 2.
Figure 43: Right image of the first stereopair of data set 2.
Figure 44: The left patch of the first stereopair of data set 2.
Figure 45: The right patch of the first stereopair of data set 2.
Figure 46: Edge image corresponding to the left patch of the first stereopair of data set 2.
Figure 47: Edge image corresponding to the right patch of the first stereopair of data set 2.
Figure 48: Hough space image and detected clusters associated with the left patch of the first stereopair of data set 2.
Figure 49: Hough space image and detected clusters associated with the right patch of the first stereopair of data set 2.
Figure 50: Detected edges in the left patch of the first stereopair of data set 2.
Figure 51: Detected edges in the right patch of the first stereopair of data set 2.
Figure 52: Predicted road edges in the left image of the second stereopair of data set 2.
Figure 53: Predicted road edges in the right image of the second stereopair of data set 2.
Figure 54: Automatically extracted left patch in the second stereopair of data set 2.
Figure 55: Automatically extracted right patch in the second stereopair of data set 2.
Figure 56: Edge image corresponding to the left image of the second stereopair of data set 2.
Figure 57: Edge image corresponding to the right image of the second stereopair of data set 2.
Figure 58: Hough space image and detected clusters associated with the left patch of the second stereopair of data set 2.
Figure 59: Hough space image and detected clusters associated with the right patch of the second stereopair of data set 2.
Figure 60: Detected edges in the left patch of the second stereopair of data set 2.
Figure 61: Detected edges in the right patch of the second stereopair of data set 2.
Table 4: Cost matrix associated with the left image sequence of the second data set.

<table>
<thead>
<tr>
<th>Observation #</th>
<th>Target #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.36</td>
<td>0.53</td>
<td>117.7</td>
<td>128.5</td>
<td>1.92</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>0.30</td>
<td>138.3</td>
<td>152.0</td>
<td>2.39</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>28.85</td>
<td>28.33</td>
<td>31.42</td>
<td>35.85</td>
<td>38.42</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>31.35</td>
<td>30.53</td>
<td>41.61</td>
<td>47.36</td>
<td>40.68</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>32.22</td>
<td>31.26</td>
<td>50.48</td>
<td>57.26</td>
<td>41.25</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>151.4</td>
<td>148.0</td>
<td>2.83</td>
<td>4.29</td>
<td>14.64</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>166.0</td>
<td>162.5</td>
<td>1.49</td>
<td>2.56</td>
<td>11.30</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>202.0</td>
<td>198.1</td>
<td>0.02</td>
<td>0.24</td>
<td>5.07</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Cost matrix associated with the right image sequence of the second data set.

<table>
<thead>
<tr>
<th>Observation #</th>
<th>Target #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.18</td>
<td>0.18</td>
<td>175.2</td>
<td>174.5</td>
<td>172.5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>0.08</td>
<td>165.0</td>
<td>164.3</td>
<td>163.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>45.80</td>
<td>41.05</td>
<td>17.72</td>
<td>17.94</td>
<td>23.96</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>68.64</td>
<td>62.49</td>
<td>2.53</td>
<td>2.59</td>
<td>5.56</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>73.77</td>
<td>67.12</td>
<td>2.20</td>
<td>2.25</td>
<td>5.19</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>84.89</td>
<td>77.3</td>
<td>1.00</td>
<td>1.02</td>
<td>3.35</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>197.4</td>
<td>182.9</td>
<td>32.02</td>
<td>31.60</td>
<td>19.74</td>
<td></td>
</tr>
</tbody>
</table>

The matched lines arising from NNSA for both left and right images are displayed in Figures 62 and 63, respectively. As in the first data set, the problem of associating the
same observation with several targets appeared in this data set. Implementing the MTIC techniques, the optimum association matrix for the left image was found:

Figure 62: Matched edges in the left image of the second stereopair of data set 2, NNSA.
Figure 63: Matched edges in the right image of the second stereopair of data set 2, NNSA.
with a cost of 11.12.

Similarly, the optimum association matrix for the right image is:

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]  

(6-3)

with a cost of 8.40.

The cost in this data set follows a Chi-square distribution with ten degrees of freedom; five targets are being tracked. Both association events are accepted under a level of significance of 0.1, \( \chi^2_{10}(0.9)=15.99 \). The matched lines for the left and right images of the second stereopair can be seen in Figures 64 and 65, respectively.
Figure 64: Matched edges in the left image of the second stereopair of data set 2, MTIC.
Figure 65: Matched edges in the right image of the second stereopair of data set 2, MTIC.
Using these association events, the motion parameters between the stereopairs under consideration were determined. Once again, the eigen value analysis of the normal equation matrix proved that all motion parameters, except for the displacement in the road direction can be recovered. This complies with the expected singularity outlined in Section 5-3 (see Figure 17-b). The motion parameters are:

<table>
<thead>
<tr>
<th>tx</th>
<th>tz</th>
<th>Azimuth</th>
<th>Pitch</th>
<th>Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.24m</td>
<td>-0.02m</td>
<td>-0.37°</td>
<td>-1.0°</td>
<td>-1.72°</td>
</tr>
</tbody>
</table>

with a variance-correlation matrix of:

\[
\begin{pmatrix}
0.00836m^2 & 0.72 & 0.44 & -0.12 & 0.97 \\
0.00258m^2 & 0.17 & 0.34 & 0.67 & \\
2.15*10^{-6} rad^2 & -0.5 & 0.6 & 4.68*10^{-6} rad^2 & -0.21 & 1.23*10^{-3} rad^2
\end{pmatrix}
\]

6-4 Considering Additional Criteria

Up to now, the recognition of observations that are believed to be target originated is only based on the distance between the observed edges and the predicted ones. In MTIC data association techniques, the association event with the minimum cost is chosen to be the event whose components are target originated. There might be situations where the targets are close to each other and the minimum cost is not significantly different from the second one. For example, the optimum association event corresponding to the right image of the first data set was found to be:
with a cost of 1.02. The second best association matrix is:

\[
\begin{pmatrix}
0 & 0 \\
0 & 0 \\
0 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 0 \\
\end{pmatrix}
\]

(6-5)

with a cost of 1.06.

The costs of the above mentioned association events are extremely close to each other. Due to round off errors the order of the cost could be easily reversed. Therefore, the choice of the association event with the minimum cost is not robust enough. This means that an additional criterion has to be added for making a correct
and reliable decision. It has to be noted that this criterion should be problem dependent.

For this particular work, the targets being tracked in the first stereopair are ordered according to their angular parameters, \( \theta \), e.g., the angular parameter of the first target is less than the one of the second target, which in turn is smaller than that of the third one. Due to the fact that the motion between the stereopairs is relatively small, this property will be maintained in consecutive stereopairs. Thus, one can use this criterion in conjunction with the association cost to choose the optimum association event more reliably.

Let us consider the above mentioned association events. The first one suggests that the sixth observation originated from the first target and the seventh observation originated from the second target. On the other hand, the second association event suggests the opposite, i.e., the second target gave rise to the sixth observation and the seventh observation originated from the first target. Since both events have almost the same cost, one has to consider the angular parameter criterion. The angular parameter of the sixth observation is 121.80°, while that of the seventh observation is 124.24°. Therefore, the association event that associates the sixth observation with the first target and the seventh observation to the second one is chosen as the optimum one. Fortunately, this association event happens to be the one with the minimum cost.

In summary, one chooses the association event that has the minimum cost and also satisfies the criterion of increasing angular parameters for successive targets. This method was the basis for choosing the optimum association matrices for the previously mentioned data sets.
CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

7-1 Summary

The goal of this dissertation was to develop a system that tracks road edges between successive stereopairs captured by a mobile mapping platform. The system uses the tracked road edges for the purpose of estimating its motion parameters between stereopairs.

The first step to achieve this goal was to develop an algorithm for the extraction of geometric primitives, such as road edges. Directly associated with this is the choice of the optimum representation module for these primitives. In this research, a minimal representation of 3-D lines that only requires four parameters was investigated, which is most suitable for the problem at hand. The representation of the 3-D lines as the intersection of two planes parallel to any two of the three coordinate axes describing object space was implemented. The development of the mathematical model that describes the perspective projection of the geometric primitives between image and object spaces is completely dependent on the representation issue. Previously developed models were incorporated and extended to perform this task.

The choice of linear features as the geometric primitives is of great importance, since the main mathematical model used in most photogrammetric applications are the collinearity equations, which are based on the identification of distinct points in the images. The identification of conjugate points in overlapping photos (point matching) is extremely difficult and unreliable, as compared to that of linear features. Another advantage offered by linear features is that they raise the level of abstractness, since
they are usually associated with physical objects in the system's environment, e.g., in this research linear features correspond to the lane boundaries of the road. Therefore, they are of great help for high level decision making in areas like the autonomous navigation of mapping systems.

The matching of road edges in consecutive stereopairs is achieved via data association techniques. They are implemented to label the observed edges in the second stereopair according their source of origin, i.e., making a decision about which edges correspond to the lane boundaries previously initialized in the first stereopair and others arising from clutter, e.g., noise and shadows.

In this work, two data association techniques were implemented. The first one, the Nearest Neighbor Standard Algorithm (NNSA), merely chooses the observed edge that is statistically closest (has the minimum Mahalanobis distance to the predicted road edge) as the one that is target-originated. Experimenting with this algorithm revealed a severe drawback: namely, it does not consider the competition of different targets for the same observation. Ignoring this fact usually leads to the association of the same observation with several targets.

This disadvantage of the NNSA was handled by incorporating Multiple Targets In Clutter (MTIC) association algorithms. A multiple hypotheses approach that simultaneously considers all observed edges and the tracked targets is used. In this algorithm, several association hypotheses are generated. Each one of these hypotheses must satisfy the following conditions:

1. Each target can give rise to only one observation.
2. Each observation can have only one source.

A cost is estimated for each one of these association events. The association hypothesis with the minimum cost is chosen as the one whose components are target-
originated. It was statistically proven that this cost follows a Chi-square distribution and represents the modified likelihood function of this association event being the right one. This means that the decision making process is supported by a statistical measure showing the probability of the correctness of this event.

Whenever close targets are involved, which is the case here, it is recommended to use an additional criterion to improve the decision making process. It was found that in this particular case, the optimum association event has a cost that is not significantly smaller than the second one. This additional criterion is problem dependent.

For all data sets tested during this research, the algorithm was successful. Since data association techniques can be viewed as classification, clustering, discrimination, detection, recognition or registration techniques, they can be applied for different photogrammetric activities, especially those related to high level processing.

Finally, the matched road edges at successive stereopairs are used to estimate motion parameters between the image pairs involved. A new algorithm was developed and tested for this task. The proposed algorithm was feasible for recovering all motion parameters except the displacement in the road direction. This is due to the fact that the features being tracked are parallel to the motion of the vehicle and do not provide any information about the advancement of the system. If transverse features were implemented the system would be able to estimate all motion parameters. All other displacement components were estimated with centimeter level accuracies, while the rotation angles were accurate up to 1 or 2 arc-minutes.

The developed system offers a great advantage when other positioning systems fail, e.g. when satellite signals are interrupted when using GPS. It is also helpful for estimating drift parameters of Inertial Navigation Systems. In summary, the suggested algorithm was successful with the tested data sets. The developed system offered an
environment capable of integrating observations from different sensors installed onboard a mobile mapping platform.

7-2 Recommendations for Further Research

Throughout this research, the motion parameter estimation was based on tracking road edges that were assumed to be straight lines. This algorithm could be extended to track linear features of higher order. In this work, the matching process was based on the geometric properties of the edges. An extension can be done by considering other edge attributes such as the gradient of the edge. Another addition to the implemented data association techniques would allow for the generation of association hypotheses that accommodate tracking an unknown number of targets with the possibility of terminating or initializing new tracks.

The proposed system could serve as the base for building an autonomous navigation vehicle. After recognizing the relative position of the system with respect to its surrounding environment, represented here by lane boundaries, the vehicle could navigate within its surroundings with the help of an obstacle avoiding sensor, such as radar. This is of great importance for the production of robots that can self-navigate in industrial environments unsafe for human operators. Thus, this research may serve as a bridge linking different fields, such as photogrammetry, computer vision and artificial intelligence, and will be of great value for the advancement of digital photogrammetry.
Appendix A

Geometric Interpretation of the Parameters Describing 3-D lines

In this appendix, the geometric interpretation of the four dimensional vector \((a, b, p, q)\) representing the 3-D line will be explained. The four parameters define the two planes:

\[
\begin{align*}
x &= az + p & \text{(plane parallel to the y-axis)} \\
y &= bz + q & \text{(plane parallel to the x axis)}
\end{align*}
\]

(A-1) (A-2)

whose intersection yields the line under consideration.

Since, the 3-D line belongs to the plane defined by (A-1), the x-coordinate of the intersection point with the x-y plane \((z = 0)\) must be equal to \(p\). Using the same argument with the plane defined by (A-2), one can say that the y-coordinate of the point of intersection with the x-y plane must be \(q\). Thus, the point of intersection of the line under consideration with the x-y plane has coordinates \((p, q, 0)\).

Figure 66: The intersection of plane \((x = a z + p)\) with x-z plane.
Inspecting Figure 66, one can conclude that the line under consideration has the direction vector \((p, y, p/a)\) for some arbitrary 'y'. This vector can be rewritten as \((a, y, 1)\). With the help of Figure 67, this line must have the direction vector \((x, b, 1)\) for some arbitrary x. Since the two vectors \((a, y, 1)\) and \((x, b, 1)\) are essentially the same, the line of interest will have the direction vector \((a, b, 1)\). It has to be noted that since the z-component of the direction vector is 1, this representation cannot be used to describe 3-D lines that are parallel to the x-y plane or perpendicular to the z-axis.

Figure 67: The intersection of the plane \(y = b z + q\) with y-z plane.
Appendix B

Derivation of the Parameters Describing the Object Plane

Using the Image Line and the Exterior Orientation Parameters

In this appendix, the vector \((U_i, V_i, W_i, D_i)\) representing the plane containing the perspective center \(O_i\), whose EOP are \((X_o, Y_o, Z_o, \omega, \phi, \kappa)\) and the image line \((\rho_i, \theta)\) are derived. The parameters of the plane will be referred to in the object coordinate system with the vector \((U_i, V_i, W_i)\) representing the normal \(N_i\) to this plane.

The image line, defined by:

\[
x \cos \theta_i + y \sin \theta_i = \rho_i
\]  

(B-1)

passes through the points \((0, \rho_i / \sin \theta_i, -f)\) and \((\rho_i / \cos \theta_i, 0, -f)\) (see Figure 68). These points are given with respect to the image coordinate system. The image plane through the perspective center and the image line \((\rho_i, \theta)\) contain vectors \((0, \rho_i / \sin \theta_i, -f)\) and \((\rho_i / \cos \theta_i, 0, -f)\) which connect the perspective center \((0, 0, 0)\) and the above mentioned two points. The normal to the image plane \(n_i\) can be obtained through the cross product of these vectors:

\[
\begin{pmatrix}
0 \\
\rho_i / \sin \theta_i \\
-f
\end{pmatrix}
\times
\begin{pmatrix}
\rho_i / \cos \theta_i \\
0 \\
-f
\end{pmatrix}
= 
\begin{pmatrix}
\rho_i f / \sin \theta_i \\
\rho_i f / \cos \theta_i \\
\rho_i^2 / (\sin \theta_i \cos \theta_i)
\end{pmatrix}
\]  

(B-1)

Vector \(N_i\) is obtained by multiplying vector \(n_i\) with rotation matrix \(R_i\), as follows:
\[
N_i = \begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix} = R_i \cdot n_i
\]  \hspace{1cm} (B-2)

where: \( R_i = f(\omega_i, \phi_i, \kappa_i) \)

Since the object plane passes through the perspective center \( O_i \), the coordinates \((X_o, Y_o, Z_o)\) must satisfy the equation of the object plane:

\[
D_i = -(U_i X_o + V_i Y_o + W_i Z_o)
\]  \hspace{1cm} (B-3)

Figure 68: Relationship between polar parameters of 2-D lines and the points of intersection with x and y axes.
Appendix C

Derivation of the Object Plane Parameters

Using the Object Line and the Exterior Orientation Parameters

In this appendix, the vector $\{U^, V^, W^, D^\}$ defining an object plane through the perspective center $(X^, Y^, Z^)$ and the object line $(a, b, p, q)$ will be derived. From the geometric interpretation of the parameters describing the 3-D object line, it has been established that this line passes through point $(p, q, 0)$ and has the direction vector $(a, b, 1)$.

The object plane also contains the vector $(p-X^, q-Y^, -Z^)$ connecting the perspective center and the intersection point of the line under consideration with the x-y plane. Thus, the normal vector to the object plane $N_i$ can be obtained as the cross product:

$$
N_i = \begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix} = \begin{pmatrix} p-X^ \\ q-Y^ \\ -Z^ \end{pmatrix} \times \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} = \begin{pmatrix} q-Y^ + b Z^ \\ -p+X^ - a Z^ \\ (p-X^) b - (q-Y^) a \end{pmatrix}
$$

Since the object plane passes through the perspective center $O_i$, point $(X^, Y^, Z^)$ must satisfy its equation yielding:

$$
D_i = -(U_i X^ + V_i Y^ + W_i Z^)
$$

(C-2)
References


Bar-Shalom, Y., Fortmann, T., 1988, Tracking and Data Association, Academic Press, INC.


Tankovich, J., MS thesis, 1991, Multiple Photograph Resection and Intersection Using Linear Features, Department of Geodetic Science and Surveying, The Ohio State University.
