4-2004

Application of Freeway Video Surveillance for Accident Investigation

Aemal Khattak
University of Nebraska-Lincoln, khattak@unl.edu

Geza Pesti
University of Nebraska - Lincoln

Yu Liu

Follow this and additional works at: http://digitalcommons.unl.edu/ndor

Part of the Transportation Engineering Commons

Khattak, Aemal; Pesti, Geza; and Liu, Yu, 'Application of Freeway Video Surveillance for Accident Investigation' (2004). Nebraska Department of Transportation Research Reports. 9.
http://digitalcommons.unl.edu/ndor/9

This Article is brought to you for free and open access by the Nebraska LTAP at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Nebraska Department of Transportation Research Reports by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
APPLICATION OF FREEWAY VIDEO SURVEILLANCE FOR ACCIDENT INVESTIGATION

Prepared for

Nebraska Department of Roads
1500 Nebraska Highway 2
Lincoln, NE 68509-4759

By

Aemal J. Khattak, Yu Liu, and Geza Pesti

Mid-America Transportation Center
University of Nebraska-Lincoln
W333.2 Nebraska Hall
Lincoln, Nebraska 68588-0530
Telephone: (402) 472-1974
Fax: (402) 472-0859
Website: http://www.matc.unl.edu

April 2004
DISCLAIMER

The contents of this report reflect the views of the authors and do not necessarily reflect the official views or policies of the Nebraska Department of Roads, the Federal Highway Administration, or the University of Nebraska-Lincoln. This report does not constitute a standard, specification, or regulation. Trade or manufacturers’ names are cited in this report because they are essential to the completeness of this report and are not endorsements by the research team.
EXECUTIVE SUMMARY

This report presents information on Research Project Number Z820 (UNL Ref. No. 26-1118-0057-001) titled “Application of Freeway Video Surveillance for Accident Investigation.” Traffic accident investigation requires detailed documentation of the accident scene, vehicles, and persons involved in the collision. To adequately document an accident scene, investigating officers accurately record various measurements (e.g., skid marks, dimensions of different highway geometric elements, etc.). Typically, a measuring tape and/or a total station (theodolite) are used to obtain the needed measurements. Advances in digital imaging and computer technology make it possible to develop cost-effective traffic accident investigation system alternatives that provide sufficient accuracy, significant time savings, and reduced costs. Utilizing stereo machine vision technology an accident investigation system, called the UNL Traffic Accident Investigation Tool (TAIT) consisting of two digital cameras and measurement-extraction software tool were developed in this project. A mock traffic accident was investigated using TAIT. Measurements provided by TAIT were compared to true values obtained by using a measuring tape. The accuracy of TAIT satisfied requirements of traffic accident investigation agencies. In addition, equipment costs as well as investigation time were reduced compared to the typical accident data collection methods. The UNL TAIT is a prototype and requires further development before it can be integrated into the day-today operations of agencies investigating traffic accidents.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER 1 - INTRODUCTION</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Objective</td>
<td>2</td>
</tr>
<tr>
<td>1.3. Report Contents</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2 - LITERATURE REVIEW</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Applications of Photogrammetry in The Field of Accident Investigation</td>
<td>3</td>
</tr>
<tr>
<td>2.2. Summary</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 3 – DIGITAL CAMERA ACCIDENT INVESTIGATION SYSTEM</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Image Projection and Stereo-Triangulation</td>
<td>5</td>
</tr>
<tr>
<td>3.2. Camera Parameters</td>
<td>11</td>
</tr>
<tr>
<td>3.3. Camera Calibration</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 4 – TAIT DESCRIPTION</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2. TAIT Image Analysis Software</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 5 – FIELD STUDY</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1. Field Study Procedure</td>
<td>24</td>
</tr>
<tr>
<td>5.2. TAIT Performance</td>
<td>31</td>
</tr>
<tr>
<td>5.2. TAIT Performance</td>
<td>31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 6 – CONCLUSIONS AND RECOMMENDATIONS</th>
<th>page</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>page</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>REFERENCES</th>
<th>page</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>APPENDIX A – THE CALCULATION OF STEREO TRIANGULATION</th>
<th>page</th>
</tr>
</thead>
</table>

LIST OF FIGURES

3.1. The Process of Projection and Back Projection.................................................................6
3.2. Ideal Stereo Triangulation...............................................................................................8
3.3. Real Stereo Triangulation ...............................................................................................9
3.4. Epipolar Geometry ..........................................................................................................10
3.5. Rotations between World Reference Frame and Camera Reference Frame ..................11
3.6. Translation between World Reference Frame and Camera Reference Frame ..............12
4.1. TAIT Hardware consisting of Digital Camera and Supporting Structure .......................15
4.2. TAIT Image Analysis Program Flow Chart .................................................................17
4.3. Select Images for Stereo Triangulation .......................................................................18
4.4. The View of Triangulation Software When Images Are Loaded ...............................19
4.5. Select the Starting Point on Left and Right Images ......................................................21
4.6. Select the Ending Point on Left and Right Images .......................................................22
4.7. The Distance between Starting and Ending Points is Given .......................................23
5.1. Calibration with Chess Board ......................................................................................25
5.2. Calibration with Pavement Markings ............................................................................27
5.3. Points Selected for Experiment ....................................................................................28

LIST OF TABLES

5.1. Camera Calibration Results with Chess Board and Any Known Points .......................29
5.2. Results of Mock Accident Investigation........................................................................30
CHAPTER 1 - INTRODUCTION

1.1. Background

Modern transportation technologies developed in recent years provide great mobility to the traveling public. Although transportation engineers put significant effort in making transportation systems safe, efficient, and environmentally friendly, traffic accidents continue to be a major national issue. Traffic accident analyses invariably utilize data that are collected at the scene of the accident. Since our understanding of traffic accidents significantly depends on the data collected at accident locations therefore, accurate measurements of elements (e.g., skid marks) present at the accident scene is of paramount importance.

Traffic accident investigation and reconstruction requires detailed documentation of the accident scene and the vehicles involved in the accident. To adequately document a traffic accident, a series of accurate on-site measurements have to be made. Investigating officers record coordinates and determine measurements for selected objects (e.g., skid marks, extent of vehicular damage, etc.) in some reference system. The measurements are typically made by using either the so-called coordinate method, or a total station (theodolite). The coordinate method measures perpendicular-off-set distances from a base line to all points of interest. The baseline is tied to some fixed object at the accident scene. The coordinate method is time consuming, and often requires that traffic lanes be closed for relatively long time periods (an hour or more) to allow the investigating officers to make the needed measurements. Lane closures may increase traffic delay and the risk of secondary collisions.

The second method, which has gained popularity in recent years, uses a total station to make measurements at traffic accident scenes. The total station surveying system requires less than half the time needed by the coordinate method to make the same number of measurements. Therefore, it reduces the duration of accident-related lane closures, which in turn reduces traffic delay, and the potential of secondary accidents. A disadvantage of total station surveying systems is their relatively high equipment price.

In this research, photogrammetry is applied to traffic accident investigation. Photogrammetry refers to the technique of measuring objects from photographs. Previous efforts of using photogrammetry to accident investigation required expensive professional equipment, such as metric cameras, to ensure the required accuracy. However, recent developments in digital
imaging and computer vision technology have made it possible to make measurements with sufficient accuracy using commercially available equipment.

1.2. Objective

This project is oriented to develop accident investigation functions that are based on machine vision technology, and to provide knowledge for the integration of NDOR freeway video surveillance with the traffic accident investigation functions of the Nebraska State Patrol (NSP). Specific objectives are: (1) to develop a system using digital cameras for traffic accident investigations, and (2) to evaluate the performance of the digital camera accident investigation system.

1.3. Report Contents

This report documents the results of the NDOR Research Project Number Z820 (UNL Ref. No. 26-1118-0057-001) titled “Application of Freeway Video Surveillance for Accident Investigation” The research was conducted by researchers at the Mid-America Transportation Center at the University of Nebraska-Lincoln. Following this introduction, Chapter 2 provides a review of pertinent literature on the application of digital cameras to traffic accident related analysis (i.e., vehicle deformation). Chapter 3 introduces the principles of the UNL TAIT system. Its hardware and software portions are described in Chapter 4. The next chapter presents information on TAIT testing during a simulated accident. Chapter 6 provides the research conclusions based on TAIT’s performance and recommendations for further development.
2.1. Applications of Photogrammetry in The Field of Accident Investigation

In recent years, the application of photogrammetry in various aspects of transportation has become the focus of a number of research efforts (1-5). Researchers from different countries applied photogrammetry in determining vehicle deformation. For example, Duignan et. al. (2) evaluated a deformation measurement system developed in Sweden. They studied 3,000 vehicle crashes, and found that the accuracy of the system was within a range of +/-15~30 mm. It was necessary to take 12 pairs of photos from 24 fixed positions to record a vehicle involved in an accident. The average time required to setup the system and take all photographs was between 20 to 30 minutes. The system requirements included expensive equipment, highly trained operators, specialized vehicles and a dedicated building space. A study by Rentschler and Uffenkamp (3) documented another application of digital photogrammetry to vehicle deformation analysis. Their method did not require fixed camera positions. The accuracy of their photogrammetric measuring system was at the millimeter-level. The time to complete more than 50 measurements was significantly less than using traditional coordinate techniques. However, all measurements had to be taken in a special room.

Several studies (1, 4 - 6) focused on the application of digital cameras in traffic accident scene investigation. McCoy et. al. (1) defined the functional requirements of a digital camera traffic accident investigation system, and developed a mono-camera (i.e., single camera) system. They applied the system to the investigation of a mock accident scene, and compared its performance to the coordinate method and the method using total stations. The system was found to have an accuracy of +/- 1 foot, but its on-site camera setup and calibration procedure was time-consuming and required at least two persons.

Roper (4) reported that each minute saved in clearing the accident will result in four to five minutes of savings in travel time delays. Cooner and Balke (5) reviewed previous applications of photogrammetric systems for traffic accident investigation. State highway patrols in Arizona, California, Oregon, Utah, and Washington were interviewed regarding their experience using photogrammetry for crash investigations. Software package PhotoModeler was used by these states and it was reported to require less training time than the total station method.
They acquired accuracy higher than required by accident investigation, but one of these states spent approximately $40,000 on three officers’ training and one set of the equipment.

Dechant Consulting Services (6) conducted experiments comparing the performances of photogrammetry and total station on straight and curved portions of interstate highway in Utah State. PhotoModeler Pro and Nikon AIMS Total Station measuring system were used representing photogrammetry and total station methods. On both straight and curved segments, the photogrammetry method required about 25 minutes to record the whole accident scene. Measurements made by total station required one and half hours on the same straight scene and 45 minutes on the curved stage. Both methods required traffic cones and evidence markers. The setup of these additional devices accounted to about 10 minutes in the total time.

2.2. Summary

Based on the reviewed literature, photogrammetry appears to have significant potential for traffic accident investigation. Its use results in measuring accident scenes much faster than the coordinate and total station methods. However, existing photogrammetry systems for traffic accident investigation require professional equipment, complicate calculation software, and multiple highly trained operators. Therefore, a photogrammetry system that utilizes off-the-shelf technology and utilizes a single investigator for traffic accident investigation is highly desirable.
A system using digital images for accident investigation is expected to determine location coordinates and distances for objects selected by the investigating officers at accident scenes. Since photographs record visual information in two-dimensions only, the third dimension (i.e., depth) of objects cannot be recovered from a single image. However, stereo images (image pairs) taken simultaneously by two cameras slightly apart can be used to determine depth information. To make a stereo vision system work properly, a relationship between physical objects at a scene and their views on the stereo images has to be known. This relationship can be defined by coordinate transformations as described later in this chapter. The coordinate transformations include several system parameters, which need to be determined through a calibration process using images of objects with known dimensions.

With a calibrated stereo vision system, various scenes can be studied, and the three-dimensional positions of any objects on the scene can be calculated based on their digital stereo images. The process of recovering three dimensional (3D) information from a pair of two dimensional (2D) images is called stereo triangulation. This chapter introduces the principles of image projection and stereo triangulation that were used in the development of the TAIT. The parameters used by the projection and stereo triangulation models are also discussed.

### 3.1. Image Projection and Stereo-Triangulation

Three-dimensional coordinates define the exact position of a point in a given reference system. When a point is photographed, it is going through two main processes: (1) transformation, and (2) projection. In the transformation phase, the three-dimensional coordinates of a point are defined in the camera’s reference system. The transformation is a linear conformal transformation process. Every point in the three-dimensional space has its unique correspondence in the camera reference frame, and vice versa.

However, projection, unlike transformation, is not a conformal process. A point on an image plane may correspond to an infinite number of points located along the path of a projecting light ray in the three-dimensional space. Therefore, depth information is lost after projection. The processes of projection and back-projection are illustrated in Figure 3.1.
Figure 3.1. The Process of Projection and Back Projection

Legend:
1. Downward direction is imaging process;
2. Upward direction is recovering 3D information;
3. ↑ means conformal transformation;
4. ↑ Means only part of the information can be recovered (i.e., a point on image plane corresponds to a line in 3D space.)
As shown in Figure 3.1, the downward arrows indicate the transformation and projection process, which determines the position of a point in space on the image plane. The upward arrows denote back projection and inverse transformation, through which recovering the space position of an image point is attempted. Solid arrows are processes whose outputs can be uniquely defined by its inputs, while the dashed arrow represents an undetermined output. It is apparent that the depth information of 3D space is lost where the pair of arrows consists of both solid and broken lines.

Given the position of any point in the camera’s field of view, its image position on the camera’s image plane can be expressed using the perspective projection model.

\[
sp_p = A[R \ t]p_w
\]  

(1)

where,

- \(p_w\) is the homogeneous vector \([p_{wx}, p_{wy}, p_{wz}, 1]^T\) of a selected point in the world reference frame, its Euclidean form is \([x, y, z]^T\);
- \(p_p\) is the homogeneous vector \([p_{px}, p_{py}, 1]^T\) of the image of \(p_w\) on the image plane, its Euclidean form is \([fx/z, fy/z]^T\);
- \(R\) and \(t\) are rotation matrix and translation vector for the transformation from the world coordinate system to the camera coordinate system;
- \(s\) is the scale factor; and
- \(A\) is the matrix of the internal parameters of the camera.

Using the projection model, all points along the path of the projecting light ray are projected to one image point.

Although the depth information is lost when a 3D space point is projected onto a 2D image plane, it is possible to recover it if two images of the same point from different view points and orientations are provided, and stereo triangulation applied. The concept of stereo triangulation is illustrated in Figure 3.2.
Points $p_l$ and $p_r$ are the images of an arbitrarily selected point $p$ in the left and right image planes. Points $O_l$ and $O_r$ are the projection centers of left and right cameras. The lines connecting $O_l$ and $p_l$, $O_r$ and $p_r$ are defined by vectors $\vec{u}$ and $\vec{v}$. Under ideal conditions these two lines intersect and their intersection specifies the exact location of point $p$. If this were the case, it would make the recovery of 3D coordinates from a stereo-image pair fairly straightforward and relatively easy. The main task is to accurately locate point $p_l$ and $p_r$ on the image planes (i.e., computer monitors). However locating these two image points depends on several factors, such as screen resolution, image quality, and human error.

Because of the inaccuracies in determining image coordinates and possible errors in the calibration of camera parameters, the two vectors $\vec{u}$ and $\vec{v}$ typically do not lie in a common plane, and therefore do not intersect, as it is illustrated in Figure 3.3. The distance between any points on vectors $\vec{u}$ and $\vec{v}$ is specified by the length of a random vector $\vec{P}$ that can be expressed as:

$$\vec{P} = -a\hat{u} + \vec{T} + b\hat{v}$$

(2)

where,

$$\hat{u} = \frac{\vec{u}}{|\vec{u}|} \text{ and } \hat{v} = \frac{\vec{v}}{|\vec{v}|}$$

are unit vectors and $a$ and $b$ are constant scalars.
Assuming that the errors in determining the image coordinates and camera parameters are comparable for both the right and left cameras, the object will be located in the mid-point of the shortest vector \( \vec{p} \). The values of \( a \) and \( b \) that minimize vector \( \vec{p} \) can be determined as:

\[
\begin{align*}
\frac{\partial |\vec{p}|}{\partial a} = \frac{\partial |\vec{p}|}{\partial b} = 0 & \quad \rightarrow \quad a = \frac{- (\hat{u} \cdot \hat{v})(\hat{v} \cdot \vec{t}) + (\hat{v} \cdot \vec{t})}{1 - (\hat{u} \cdot \hat{v})^2} \\
& \quad \quad \quad b = \frac{(\hat{u} \cdot \hat{v})(\hat{u} \cdot \vec{t}) - (\hat{v} \cdot \vec{t})}{1 - (\hat{u} \cdot \hat{v})^2} \quad (4)
\end{align*}
\]

In most cases, the mid-point of the shortest vector \( \vec{P} \) is an acceptable estimation of the location of the true point in space. However, this approach gives good estimates only when the length of vector \( \vec{P} \) is short. The length of \( \vec{P} \) primarily depends on the error in locating the image points on the screen. To assist the user in accurately locating the matching images of a point, an epipolar line function was developed (Figure 3.4).
The epipolar geometry between left and right views, illustrated in Figure 3.4, essentially consists of the intersections of the image planes with an epipolar plane having the baseline connecting the two camera centers $o_l$ and $o_r$. Point $P$ has an image $p_l$ in the left camera and an image $p_r$ in the right camera. The epipolar plane intersects the image plane of right camera in the epipolar line $p_re_r$, where $e_r$ is the projection of $o_l$ through $o_r$. When the image of point $P$ is located in the image plane of the left camera, point $P$ has to be on the epipolar line $p_re_r$. The coordinates of point $o_l$ and $p_l$ in the right camera coordinate system can be determined by applying a rotation and translation to their coordinates in the left camera frame. With the coordinates of these two points in the right camera frame, their images on the right camera’s image plane can be obtained as $e_r$ and $p_r$. The line connecting $e_r$ and $p_r$ is the epipolar line on the right image. This line is also the projection of $o_l p_l$ and its extension on the image plane of the right camera. Thus, locating $p_r$, the matching image of $p_l$ on the right image plane, can be restricted to a narrow band around the epipolar line on the right image plane. With the assistance of epipolar line, the stereo triangulation procedure becomes more accurate. However the approach described here requires proper calibration of all camera parameters, which is described next.
3.2. Camera Parameters

A primary objective of traffic accident investigation is to obtain measurements in a three-dimensional reference coordinate system, which can provide accurate data for accident analysis and reconstruction. In most cases, the reference coordinate system is established with its origin fixed on a permanent object such as a light pole, a utility pole or a fire hydrant close to the accident scene. One axis of the coordinate system is often aligned with a lane line, a curb, or other longitudinal object. However, a digital image of an accident scene provides special information in the camera’s coordinate system, whose origin is normally located in the projection center of the camera.

3.2.1 External Camera Parameters

External parameters uniquely define the transformation from the world coordinate system (world reference frame) to the camera coordinate system (camera reference frame). Typically, these parameters can be determined by: (i) finding the rotation matrix that brings the corresponding axes of both coordinate systems into coincidence, and (ii) finding the translation vector between the relative positions of the origins of the two coordinate systems. This process is illustrated by Figures 3.5 and 3.6.

Figure 3.5. Rotations between World Reference Frame and Camera Reference Frame
The coordinates of point P in a reference coordinate system, shown in Figure 3.5, are defined by vector $\vec{p}$. Rotating the reference coordinate system about axes x, y and z, the new coordinates of point P will be defined by vector $\vec{p}_{\text{new}}$:

$$\vec{p}_{\text{new}} = R \vec{p} = R_\kappa R_\phi R_\omega \vec{p}$$

(5)

where,

$T = \text{rotation angle about x axis}$ and

$$R_\kappa = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{bmatrix}$$

$n = \text{rotation angle about y axis}$ and

$$R_\phi = \begin{bmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix}$$
\[ \theta = \text{rotation angle about } z \text{ axis and } R_\kappa = \begin{bmatrix}
\cos \kappa & \sin \kappa & 0 \\
- \sin \kappa & \cos \kappa & 0 \\
0 & 0 & 1
\end{bmatrix} \]

* Note: all rotations are clockwise.

In addition to the rotation matrix, a translation vector \( \vec{t} \), defining the coordinates of the origin of world reference frame in the camera reference frame, has to be determined:

\[ \vec{t} = [t_x \ t_y \ t_z]^T \]  \hspace{1cm} (6)

Then, as illustrated in Figure 3.6, \( p_w \) in the world coordinate system can be expressed by vector \( \vec{p}_c \) in the camera coordinate system as:

\[ \vec{p}_c = R \vec{p}_w + \vec{t} \]  \hspace{1cm} (7)

Knowing the external parameters of a camera (i.e., rotation matrix \( R \) and translation vector \( \vec{t} \)), the coordinate transformation can be performed.

### 3.2.2 Internal Camera Parameters

Internal parameters characterize the optical, geometric, and digital characteristics of a camera. More specifically, these parameters include horizontal and vertical magnifications, the skew between horizontal and vertical axes, the position of the image center and other parameters accounting for lens distortions. Arranging the internal parameters in the following matrix form:

\[ A = \begin{bmatrix}
f_1 & \alpha f_1 & c_1 \\
0 & f_2 & c_2 \\
0 & 0 & 1
\end{bmatrix} \]  \hspace{1cm} (8)

enables the projection model formulated as equation (1).

Here,

- \( f_1 \) and \( f_2 \) are the focal length expressed in units of horizontal and vertical pixels, they are equal to a unique value in mm;
- \( c_1 \) and \( c_2 \) are horizontal and vertical coordinates corresponding to the true image center;
- " is the skew coefficient defining the angle between x and y pixel axes, which should be perfectly horizontal and vertical.
Image plane coordinates do not account for lens distortion. Image point displacement caused by lens distortion is corrected before converting camera coordinates to pixel coordinates. The parameters required for the correction include radial distortion factors \(k_{r1}, k_{r2}, k_{r3}\) and tangential distortion factors \(k_{t1}, k_{t2}\).

### 3.3. Camera Calibration

Camera calibration is a essential procedure in machine vision. Through calibration, each camera’s internal and external parameters are obtained and the relationship between different cameras is also determined using controlled field measurements. The camera calibration procedure is discussed in Chapter 5.
CHAPTER 4 – TAIT DESCRIPTION

The UNL TAIT has both hardware and software components that are described in this chapter. The hardware portion (Figure 4.1) includes two digital cameras and a supporting tripod with a dual-camera stand extension structure. The software part involves a prototype stereo triangulation software that was specifically developed for this research.

![Figure 4.1. TAIT Hardware consisting of Digital Cameras and a Supporting Structure](image)

4.1. TAIT Hardware Configuration

The distance between cameras and their orientations were adjustable in order to achieve the best common field of view for stereo images. The two cameras were mounted about 40 cms apart on the dual-camera stand extension centered on the tripod. The left camera is a Canon
PowerShot G2 and the right one is a Kodak DC120. The dual-camera extension allowed the two cameras to be mounted on a single tripod with a maximum distance of 60 cms between them. All equipment used in the system are off-the-shelf products. Unlike other methods using total station or previous photogrammetric applications, no expensive professional instruments are required for this system. The financial investment in hardware was significantly reduced and the total cost of the hardware was less than $1,000 in this case. With the rapid development of digital imaging technology, cameras that are capable of capturing higher resolution images are available in the market. Although finer images improve the accuracy of the measurements, the software module developed for this research has no special requirements for the choice of digital cameras.

The image storage media is determined by the digital cameras used. With the above hardware configuration, Compact Flash (CF) cards were used. These storage cards can be operated as removable disks, making it possible to easily transfer digital images of accident scenes to notebook or desktop computers for data extraction and analysis.

4.2. TAIT Image Analysis Software

To determine actual distances between points selected from digital images, a computer program, which is capable of implementing the principles of stereo triangulation and the calibration process described above, was developed. The program was expected to import multiple images, record image coordinates, solve an optimization algorithm for camera calibration, and calculate three-dimensional coordinates for any object of interest. It was developed under the Matlab Graphical User Interface Development Environment (GUIDE). The main program steps are summarized in the flow chart shown in Figure 4.2. The analysis of accident scene images begins with the loading of camera parameters and image pairs. This step is illustrated by Figures 4.3 and 4.4. The camera and system parameters are determined by a separate program. The software package Camera Calibration Toolbox for Matlab (7) developed by a research team at the California Institute of Technology was adopted and modified to accommodate the particular requirements of calibration process for the application of accident investigation. After loading the system parameters and image pairs, relevant points are identified by the user on both the right and left images. It is done by selecting the appropriate pixel on the computer monitor using a pointing device (e.g., computer mouse).
Read in system parameters

Load left and right images

Locate the first point in the left image

Locate the same first point in the right image

Locate the second point in the left image

Locate the same second point in the right image

Determine the 3D coordinates of both points and calculate the distance between them

Continue?

Yes

No

Exit

New Images

Same Images

Figure 4.2. TAIT Image Analysis Program Flow Chart
Welcome to use accident investigation tool,
Please select the left and right images.

Figure 4.3. Select Images for Stereo Triangulation
Accident Investigation Tool

Figure 4.4. The View of Triangulation Software When Images Are Loaded

Instruction: To start measuring dimensions, click select point button.

Status:

SELECT POINT  ZOOM  Another  Step Over
It was observed that accurately locating the same objects from the image pair was
difficult and it introduced significant errors in the subsequent estimation of object coordinates.
Therefore, an epipolar line function was developed to assist users in accurately identifying the
same objects on both right and left images. This step is illustrated in Figures 4.5 and 4.6. The
three-dimensional coordinates of the selected image points, and the distances between them are
calculated in a program module using stereo triangulation. The stereo triangulation principle was
discussed in Chapter 3. A more detailed description of the method used in the stereo
triangulation program module is given in Appendix A. The output of the program includes three-
dimensional coordinates for any single point, and the distance between any two points picked by
the user in the image pairs. This last step indicating distance output is illustrated in Figure 4.7.

The program can be terminated at any time, by clicking the “Cancel and Close” button
and confirming that the user really wants to exit. Instructions assisting users through the program
operation are displayed in message boxes.
Instruction: Click the starting point of measurement in the RIGHT image. It should be some point on the epipolar line appeared on the right.

Status: Starting point of measurement in the LEFT image has been selected.

Figure 4.5. Select the Starting Point on Left and Right Images
Figure 4.6. Select the Ending Point on Left and Right Images

Instruction: Click the ending point of measurement in the RIGHT image.

Status: Ending point of measurement in the LEFT image has been selected.
Figure 4.7. The Distance between Starting and Ending Points is Given (unit: mm)
CHAPTER 5 – FIELD STUDY

Field experiments were conducted to assess the feasibility of using TAIT for traffic accident investigation. A hypothetical accident situation, involving a single passenger car, was staged in an empty parking lot. The objective of the experiments was to determine distances between various points, arbitrarily selected on the vehicle and in its surrounding area, using digital photographs, and compare them with the actual distances determined by physical measurements. The parking lot was selected for the location of the experiments because of the availability of highly visible grid-like pavement markings delineating the parking stalls. These pavement markings made camera calibration much easier.

5.1. Field Study Procedure

Each experiment began with the assembly of the system by mounting the digital cameras to the supporting structure. Between experiments, the system components were completely disassembled because of limited storage space. Therefore, the external and internal camera parameters had to be re-calibrated for each experiment. Re-calibration of the external parameters was necessary because the relative positions and orientations of the two cameras changed between experiments. Internal parameters had to be re-calibrated because the focal lengths of digital cameras with zooming capabilities might have changed every time the cameras were turned off and on between experiments. If solid-state cameras (i.e., with fixed focal length) permanently mounted to the supporting structures were used, both internal and external parameters would remain constant, and therefore re-calibration would not be required. Once such a system is fully calibrated, it can be used as ‘Point and Shoot’ cameras.

Each experiment included four major steps: 1) hardware setup, 2) camera calibration, 3) taking pictures, and 4) taking control measurements. Two sets of experiments were conducted; they only differed in their calibration methods.

In the first set of experiments, camera parameters were calibrated using a method developed by Bouguet (7). Calibration was accomplished by taking several pictures of a calibration grid, which is basically a chess board-like image consisting of 3x3-cm black and white squares arranged in seven rows and nine columns, as shown in Figure 5.1.
The pictures were taken from different angles and camera orientations. An important criterion of taking the pictures for this calibration method is that the calibration grid has to fill a large portion of both cameras’ fields of view for both cameras. Therefore, the distance between the digital cameras was reduced to 10 centimeters, and the calibration grid was placed within the one-meter range in front of the double-camera system.

Thirteen pairs of images of the calibration board were taken, and then processed. Using the thirteen image pairs, the internal camera parameters such as focal lengths, corrected image center, and lens distortions were determined for both digital cameras separately. Then the output of internal calibration results of both cameras was used as input to their external calibration process. In the external calibration process, relative position and orientation of the left and right cameras were determined and the results of the internal calibrations were optimized.

After the camera parameters were calibrated, images containing measurements of various objects of interest were taken. During image acquisition the internal and external parameters of the two digital cameras remained the same. The zooming functions of both cameras were disabled to keep the internal parameters constant. To ensure the same external parameters, no adjustment to the connection between cameras and the double stand extension was made. However, the whole system was moved around the mock accident scene, and pictures were taken from different angles to achieve the best views for the accident scene images.

The digital image-pairs taken by the dual-camera system were analyzed on site using a laptop computer. Three-dimensional coordinates of each selected point were determined using the TAIT software (discussed in Chapter 4). Distances between any two points were calculated from their known three-dimensional coordinates. A comparison of the calculated and measured distances revealed that there were significant differences between some of the calculated results and their corresponding true values. The level of error was unacceptable for the accuracy requirement of +/- 1 foot. After a review of all steps of the first set of experiments, and a thorough analysis of the possible causes of the error, it was concluded that the calibration process used above needed revision. In the first set of experiments, the relatively small chess board-like calibration grid was placed within 1-meter range in front of the cameras, while the measured objects were at least 3 meters from the cameras. Thus, the calibration board filled only a small part of the measuring space, which led to intolerable errors. For improved accuracy, the calibration process was updated and a second set of experiments was conducted.
Figure 5.2: Calibration with Pavement Markings
A larger scale calibration grid was required to improve measurement accuracy. However, a calibration board of the size, which would fill most of the measuring space in the cameras’ field of view at an accident scene would be impractical. Objects of such dimensions are not feasible for on-site calibration. As an alternative, the pavement markings in a remote area of the parking lot were used to provide true dimensions for camera calibration. Five pairs of images, similar to those shown in Figure 5.2, were taken and the end points of the markings were used as control points. Their actual dimensions were determined with a measuring tape. Camera calibration was performed in a similar manner as was accomplished by using the chess board-like calibration grid. The parameter values obtained using the two calibration methods are summarized in Table 5.1. With improved camera calibration, distances in the second set of experiments were determined in the same manner as in the first set using the TAIT stereo triangulation software (Figure 5.2). Results of the second set of experiments on a mock single-vehicle accident are presented in Table 5.2. Distances measured with tape ruler were considered as true values, since tape measurement is a common tool used in accident investigation. Three trials were carried out for each measurement.

![Figure 5.3. Points Selected for Experiment](image-url)
### TABLE 5.1 Camera Calibration Results with Chess Board and Any Known Points

<table>
<thead>
<tr>
<th>Camera Parameters</th>
<th>1st: With chess calibration board</th>
<th>2nd: Using any known points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left Cam</td>
<td>Focal Length (pixels)</td>
<td>1021</td>
</tr>
<tr>
<td></td>
<td>Principle Point (pixels)</td>
<td>(522.2, 390.3)</td>
</tr>
<tr>
<td></td>
<td>Distortion ((k_{r1}, k_{r2}, k_{r3}, k_{t1}, k_{t2}))</td>
<td>-0.2072 0.2355 0.0000 -0.0003 -0.0003</td>
</tr>
<tr>
<td>Right Cam</td>
<td>Focal Length (pixels)</td>
<td>1121</td>
</tr>
<tr>
<td></td>
<td>Principle Point (pixels)</td>
<td>(499.1, 404.1)</td>
</tr>
<tr>
<td></td>
<td>Distortion ((k_{r1}, k_{r2}, k_{r3}, k_{t1}, k_{t2}))</td>
<td>-0.2505 0.5757 0.0000 -0.0009 0.0001</td>
</tr>
<tr>
<td><strong>External Parameters</strong></td>
<td>Rotation Vector</td>
<td>0.0199 0.0946 -0.0148</td>
</tr>
<tr>
<td></td>
<td>Translation Vector (mm)</td>
<td>-137.6227 -8.3679 -188.3513</td>
</tr>
</tbody>
</table>
Table 5.2. Results of Mock Accident Investigation

<table>
<thead>
<tr>
<th>DISTANCE</th>
<th>TRUE VALUE</th>
<th>Measured</th>
<th>Deviation</th>
<th>Squared Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>unit</td>
<td>ft</td>
<td>mm</td>
<td>mm</td>
<td>mm^2</td>
</tr>
<tr>
<td>Front/Rear Wheels</td>
<td>8.2</td>
<td>2499</td>
<td>2536</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2533</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2534</td>
<td>35</td>
</tr>
<tr>
<td>Point A, B</td>
<td>8.6</td>
<td>2621</td>
<td>2665</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2713</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2680</td>
<td>59</td>
</tr>
<tr>
<td>Front wheel, A</td>
<td>5.91</td>
<td>1801</td>
<td>1876</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1896</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1890</td>
<td>89</td>
</tr>
<tr>
<td>Rear wheel, B</td>
<td>12.19</td>
<td>3716</td>
<td>3743</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3702</td>
<td>-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3778</td>
<td>62</td>
</tr>
<tr>
<td>Radius of rear wheel</td>
<td>0.9</td>
<td>274</td>
<td>291</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>318</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>309</td>
<td>35</td>
</tr>
<tr>
<td>Front wheel center,</td>
<td>4.68</td>
<td>1426</td>
<td>1490</td>
<td>64</td>
</tr>
<tr>
<td>Front keyhole</td>
<td></td>
<td></td>
<td>1494</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1504</td>
<td>78</td>
</tr>
</tbody>
</table>

RMSE = 59.27 mm
5.2. TAIT Performance

Six measurements were chosen to test TAIT performance. Each distance was determined using TAIT software three times in a row to test the precision of the system. Comparing the three computed distances to the true values measured by tape ruler, it was found that TAIT overestimated real distances. In only one trial, the calculated distance was 14 mm shorter than manually measured result.

The maximum difference between the distances calculated from the stereo images and true distances was 95 mm. This error was far less than the required accuracy of +/- 1 foot (+/- 304.8 mm) specified by police departments. Deviations of calculated results from true values were analyzed to study the performance of the system. The mean difference between calculated values and true dimensions was 52.3 mm, and the standard deviation of differences was 28.7 mm. It can be concluded that on average, the output from TAIT overestimated the real distances by about 5.92 cms (the root mean square error of measurements). Considering the relatively small standard deviation, TAIT provides a fairly precise tool for accident investigations.

Compared to measuring with tape rulers and total station theodolites, the time required to investigate the accident scene was significantly reduced. It only took 15 minutes for one person to record an accident scene using the digital camera accident investigation system.
The studied digital camera accident investigation system is capable of measuring three-dimensional coordinates from stereo images with an accuracy typically required in accident investigation. The three-dimensional coordinates can be used to determine the distances between any points of interest at the accident scene.

It takes about 15 minutes to set up the system, which also includes the time required for calibration. Depending on the number of stereo images that needed to be taken, additional 2 to 5 minutes may be required. Compared to the coordinate and total station methods, which take more than one hour, and half an hour, respectively, TAIT will significantly reduce lane closure duration during accident investigation. Shorter lane closure times result in reduced delays and road user costs. In addition to time savings, TAIT is also labor-efficient since only one person is required to complete the accident investigation process. Savings from the reduction in road user expense, secondary accident costs, and labor costs are expected to make TAIT cost-effective. An additional benefit of TAIT is that it results in a permanent record of the accident scene that can be retrieved at any time in the office and measurements made on elements that might not have been made on the accident scene.

The accuracy of the digital camera measuring system is approximately +/- 6 cms, which satisfies the requirement of most law enforcement agencies. However, it was observed that the system had a tendency to overestimate distances. The differences between calculated and true distances were fairly constant. Refining the system, particularly the calibration module, can further improve its accuracy. All components of the system are off-the-shelf products. The cost of the whole system is low (i.e., less than $1,000 for field equipment) and within the affordable range to equip most agencies dealing with traffic accidents.

Recommendations

For future applications, it is recommended that solid-state cameras should be used, and a system with permanent camera mounting function be developed. More accurate camera calibration can be secured by either designing a standard process to be performed at police departments or having camera parameters associated with each permanently mounted solid-state cameras. To reduce vibration-related errors, shutters of the cameras should be triggered remotely.
and simultaneously rather than directly pushing camera buttons one after the other. Such system would always be ready for use and hopefully more accurate.
ACKNOWLEDGMENTS

The research team is thankful to various persons who contributed to this research. In particular, Milo Cress (FHWA), John Perry (FHWA), Paul Cammack (NDOR), the late Dr. Patrick McCoy (UNL), and the NDOR Research Division staff significantly helped with various aspects of this research project.
REFERENCES


APPENDIX A – THE CALCULATION OF STEREO TRIANGULATION

In the reconstruction of the 3D coordinates of a point, two vectors, obtained from user inputs, are used to determine the space position of the point. The user locates the point on both left and right digital images. The origins of the two vectors are the projection centers of the two cameras, and their destinations are the corresponding two image points located by the user. Ideally, the intersection of extensions of these two vectors specifies the true location of the selected point in the 3D space. However, in practice these two vectors do not intersect. To find the best approximation of the true location of the point, the middle point of the shortest distance between the two vector has to be found. Since only the minimal distance between the two vectors or their extensions needs to be found, the orientations rather than lengths of the two vectors are necessary inputs to the stereo triangulation program.

Recalling chapter 3, two factors $a$ and $b$ are used to define the lengths of the vectors. In the triangulation algorithm, vectors $\vec{u}$ and $\vec{v}$ are converted to unit vectors before using them in the equation. Vector $\vec{p}$ connecting any two points on the lines of vectors $\vec{u}$ and $\vec{v}$ can be expressed as:

$$\vec{p} = -a \frac{\vec{u}}{|\vec{u}|} + \vec{T} + b \frac{\vec{v}}{|\vec{v}|}$$

The values of $a$ and $b$, which minimize $\vec{p}$ can be obtained by setting the partial derivatives of the length of $\vec{p}$ equal to zero. The calculation process is shown as follows:
\[ \frac{\partial |\vec{p}|}{\partial a} = 0; \quad \frac{\partial |a \vec{u} + \vec{T} + b \vec{v}|}{\partial a} = 0 \]

\[ \frac{\partial}{\partial a} \left[ (-a \frac{\vec{u}_x}{|\vec{u}|} + \vec{T}_x + b \frac{\vec{v}_x}{|\vec{v}|})^2 + (-a \frac{\vec{u}_y}{|\vec{u}|} + \vec{T}_y + b \frac{\vec{v}_y}{|\vec{v}|})^2 + (-a \frac{\vec{u}_z}{|\vec{u}|} + \vec{T}_z + b \frac{\vec{v}_z}{|\vec{v}|})^2 \right] = 0 \]

\[ (-2 \frac{\vec{u}_x}{|\vec{u}|})(-a \frac{\vec{u}_x}{|\vec{u}|} + \vec{T}_x + b \frac{\vec{v}_x}{|\vec{v}|}) + (-2 \frac{\vec{u}_y}{|\vec{u}|})(-a \frac{\vec{u}_y}{|\vec{u}|} + \vec{T}_y + b \frac{\vec{v}_y}{|\vec{v}|}) + (-2 \frac{\vec{u}_z}{|\vec{u}|})(-a \frac{\vec{u}_z}{|\vec{u}|} + \vec{T}_z + b \frac{\vec{v}_z}{|\vec{v}|}) = 0 \]

\[ 2a \frac{\vec{u}_x^2 + \vec{u}_y^2 + \vec{u}_z^2}{|\vec{u}|^2} - 2 \frac{\vec{u}_x \vec{T}_x + \vec{u}_y \vec{T}_y + \vec{u}_z \vec{T}_z}{|\vec{u}|} - 2b \frac{\vec{u}_x \vec{v}_x + \vec{u}_y \vec{v}_y + \vec{u}_z \vec{v}_z}{|\vec{u}| |\vec{v}|} = 0 \]

\[ a - \frac{(\vec{u} \cdot \vec{T})}{|\vec{u}|} - b \frac{(\vec{u} \cdot \vec{v})}{|\vec{u}| |\vec{v}|} = 0 \]

\[ |\vec{u}| |\vec{v}| a - |\vec{v}|(\vec{u} \cdot \vec{T}) - b(\vec{u} \cdot \vec{v}) = 0 \]

Similarly,

\[ |\vec{u}| |\vec{v}|b + |\vec{u}|(\vec{v} \cdot \vec{T}) - a(\vec{u} \cdot \vec{v}) = 0 \]

\[ \begin{aligned}
   a &= \frac{|\vec{u}| |\vec{v}|^2 (\vec{u} \cdot \vec{T}) - |\vec{u}|(\vec{u} \cdot \vec{v})(\vec{v} \cdot \vec{T})}{|\vec{u}|^2 |\vec{v}|^2 - (\vec{u} \cdot \vec{v})^2} \\
   \therefore \quad b &= \frac{|\vec{u}|^2 |\vec{v}|(\vec{v} \cdot \vec{T}) - |\vec{v}|(\vec{u} \cdot \vec{v})(\vec{u} \cdot \vec{T})}{|\vec{u}|^2 |\vec{v}|^2 - (\vec{u} \cdot \vec{v})^2}
\end{aligned} \]

Note that vector \( \vec{u} \) and \( \vec{v} \) are defined in different coordinate systems, and therefore vector \( \vec{u} \) in the above equations should be replaced by:

\[ R \times \vec{u} \]

where, \( R \) is the rotation matrix bringing the left camera coordinate system into coincidence with the right camera coordinate system; and the equations to calculate \( a \) and \( b \) are:
\[
\begin{align*}
    a &= \frac{|\vec{u}|^2 (R \times \vec{u} \cdot \vec{T}) - |\vec{u}|(R \times \vec{u} \cdot \vec{v})(\vec{v} \cdot \vec{T})}{|\vec{u}|^2 |\vec{v}|^2 - (R \times \vec{u} \cdot \vec{v})^2} \\
    b &= \frac{|\vec{u}|^2 |\vec{v}|(\vec{v} \cdot \vec{T}) - |\vec{v}|(R \times \vec{u} \cdot \vec{v})(R \times \vec{u} \cdot \vec{T})}{|\vec{u}|^2 |\vec{v}|^2 - (R \times \vec{u} \cdot \vec{v})^2}
\end{align*}
\]

Then, the middle point of the shortest vector \( \vec{p} \) is:

\[
middle of \vec{p} = a\vec{u} + \frac{1}{2} \vec{p}
\]

\[
= \frac{1}{2}(a\vec{u} + \vec{T} + b\vec{v})
\]

**Distance Calculation**

When considering \( \vec{u} \) and \( \vec{v} \) in the left camera coordinate system, vector \( b\vec{v} \) in the right camera coordinate system is expressed as \((R^{-1} \cdot (b\vec{v} - \vec{T}) - \vec{T})\). The object’s position in the left camera coordinate system is \( \vec{p}_L \):

\[
middle of \vec{p} = \frac{1}{2}(a\vec{u} + R^{-1} \times (b\vec{v} - \vec{T}))
\]

and its coordinate in the right camera coordinate system is:

\[
\vec{p}_R = R \times \vec{p}_L + \vec{T}
\]

With the 3D positions of any two points, the distance between them can be easily determined as:

\[
D = |\vec{p}_{L1} - \vec{p}_{L2}| = |\vec{p}_{R1} - \vec{p}_{R2}|
\]