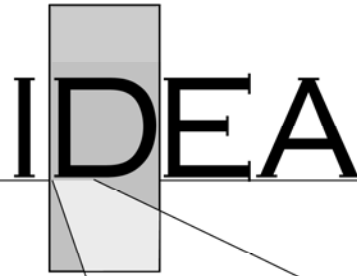


**Innovations Deserving  
Exploratory Analysis Programs**



---

***Highway IDEA Program***

---

*Automated Real-Time Pavement Crack Detection and  
Classification*

Final Report for Highway IDEA Project 111

Prepared by:

Kelvin C. P. Wang and Weiguo Gong, University of Arkansas, Fayetteville, AR

***May 2007***

---

**TRANSPORTATION RESEARCH BOARD**  
*OF THE NATIONAL ACADEMIES*

## **INNOVATIONS DESERVING EXPLORATION ANALYSIS (IDEA) PROGRAMS MANAGED BY THE TRANSPORTATION RESEARCH BOARD (TRB)**

This NCHRP-IDEA investigation by the University of Arkansas was completed as part of the National Cooperative Highway Research Program (NCHRP). The NCHRP-IDEA program is one of the four IDEA programs managed by the Transportation Research Board (TRB) to foster innovations in highway intermodal surface transportation systems. The other three IDEA program areas are Transit-IDEA, which focuses on products and results for transit practice, in support of the Transit Cooperative Research Program (TCRP), Safety-IDEA, which focuses on motor carrier safety practice, in support of the Federal Motor Carrier Safety Administration and Federal Railroad Administration, and High Speed Rail-IDEA (HSR), which focuses on products and results for high speed rail practice, in support of the Federal Railroad Administration. The four IDEA program areas are integrated to promote the development and testing of nontraditional and innovative concepts, methods, and technologies for surface transportation systems.

For information on the IDEA Program contact IDEA Program, Transportation Research Board, 500 5<sup>th</sup> Street, N.W., Washington, D.C. 20001 (phone: 202/334-3310, fax: 202/334-3471, <http://www.nationalacademies.org/trb/idea>).

The project that is the subject of this contractor-authored report was a part of the Innovations Deserving Exploratory Analysis (IDEA) Programs, which are managed by the Transportation Research Board (TRB) with the approval of the Governing Board of the National Research Council. The members of the oversight committee that monitored the project and reviewed the report were chosen for their special competencies and with regard for appropriate balance. The views expressed in this report are those of the contractor who conducted the investigation documented in this report and do not necessarily reflect those of the Transportation Research Board, the National Research Council, or the sponsors of the IDEA Programs. This document has not been edited by TRB.

The Transportation Research Board of the National Academies, the National Research Council, and the organizations that sponsor the IDEA Programs do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the investigation.

**AUTOMATED REAL-TIME PAVEMENT CRACK DETECTION  
AND CLASSIFICATION**

**IDEA PROJECT FINAL REPORT**

CONTRACT NCHRP-111

Principal Investigator: Kelvin C.P. Wang

Research Staff: Weiguo Gong

**IDEA Program**

**Transportation Research Board**

**May 31, 2007**

**University of Arkansas**

## TABLE OF CONTENT

Contract NCHRP-111 .....	3
Table of Content .....	4
Executive Summary .....	1
Introduction.....	3
Stage I Research .....	4
System Design .....	4
Mathematical Model.....	5
Calibration .....	6
Distortion Adjustment .....	16
Matching .....	17
3-D Reconstruct .....	21
Result .....	22
Stage II Research .....	22
System Design .....	23
Laser Road Imaging System (LRIS ).....	23
Automated Distress Analyzer (ADA).....	25
Algorithm.....	25
Speed.....	25
Quality Control .....	25
Reporting .....	26
Field Evaluation.....	27
Conclusions and Recommendations .....	27
Reference .....	28

## Executive Summary

This final report describes research results of the IDEA project, “Automated Real-Time Pavement Crack Detection and Classification,” contract No. NCHRP-111. The objective of this project was to apply the stereovision technique to detect pavement distresses based on a previously completed project NCHRP-88.

The research work was divided in two stages. Stage I focuses on the application of the stereovision technique in the pavement condition survey system. Stage II uses improved hardware and focuses on the development of automated distress survey algorithms.

In the first stage, key hardware of the system includes the lighting system and the area-scan cameras. A total of four cameras were used in two pairs to collect pavement surface images across a 4-meter wide pavement, each pair of images covering two meters of the road. The cameras are capable of capturing digital images at the resolution of  $1300 \times 1024$  with a speed of more than 25 frames per second. The lighting system includes a generator and a series of lighting bulbs.

Algorithm for 3-D pavement surface reconstruction was developed in stage I. It is based on the geometrical relation between two 2-D image planes and the object space coordinates. It includes four steps: calibration, distortion adjustment, matching and 3-D reconstruction. Different calibration models have been tested to obtain the highest possible accuracy. Computer code was written in Visual C++ to implement this algorithm.

The 3-D pavement surface reconstruction can be integrated with the pavement image acquisition system through two methods. One is to generate the 3-D geometrical structure for the entire road surface. This provides detailed and comprehensive information of the pavement. However, it does not meet the speed requirement of a real-time system since the coordinates of a large number of pixels need to be calculated. Another way is to choose a narrow strip of data around a specified location and use the 3-D structure for the smaller region. This can be applied in real-time processing to save computing time.

However, result from the first stage indicated that about 2-mm accuracy was obtained in the test which limited the applicability of the system. Accuracy level when the vehicle is in motion is worse. Several factors contribute to this unsatisfactory accuracy. First, the calibration may not be precise enough to provide an accurate model for 3-D reconstruction. Calibration should be conducted with high accuracy equipment in a professional optical lab which was not affordable within the project budget. Secondly, an illumination device providing uniform lighting across pavement surface for the area scan cameras was necessary for this research. Despite the fact that the Xenon based lighting was one of the most efficient lighting devices, there were still limitations of illumination uniformity and shadow. In addition, errors introduced due to motion and vibration of the moving vehicle was difficult to correct. A high-precision inertial gyro device should be used for correcting the errors, but was not used due to budget constraints.

Meanwhile, a new pavement image system, Laser Road Imaging System (LRIS), came to the market in late 2005. In the second stage, the laser lighting, line scan-based Laser Road Imaging System (LRIS) was applied to capture the pavement images. The advantages of applying LRIS are that the performance of laser imaging is invariant to the change of weather or shadow. Laser power consumption is about 1% of that for Xenon bulbs for covering the same area of pavement. Laser illumination from LRIS is uniform and substantially more powerful than sun light in the dedicated spectrum. However, with the line-scan camera, the stereovision technology is not directly applicable unless multiple LRIS devices are used. Due to time and budget constraint, this new setup with two or more LRIS devices was not possible during the research period.

The captured images with LRIS were then analyzed by Automated Distress Analyzer (ADA) in real-time. ADA is a real-time distress analyzer which is able to detect four common cracks (longitudinal, transverse, alligator and block). Road statistics according to predefined protocols are calculated to give the indices of the target segment pavement condition. The entire system has been demonstrated to be a successful real-time pavement survey system capable of traveling and collecting data at highway speeds. Three vehicles with LRIS and ADA made by WayLink Systems Co., minority-owned by the University of Arkansas, were delivered to users in US, Asia, and Europe.

In summary, through this project substantial insight and knowledge have been established in using 3-D surface reconstruction for building pavement surface models for comprehensive pavement distress survey. Detailed technical steps were thoroughly researched and software code was written for their implementations. It is anticipated that with additional funding, 1-mm level resolution of 3-D pavement surface models can be established with multiple laser imaging devices and precision Inertial Measurement Unit (IMU).

## INTRODUCTION

Billions of dollars are invested in pavements in US each year. Accurate data collection and interpretation of pavement data are critical for the decision-making process for pavement preservation and rehabilitation. Pavement condition surveys are one of the key components in a pavement management system.

Elements of consideration in pavement condition survey include cracking, rutting and other surface distortion and damages. Two-dimensional (2-D) imaging may not provide enough information to represent the entire pavement condition. For example, pavement rutting and roughness cannot be detected from 2-D pavement images. Thus, stereovision technique is applied in the project to enable a three-dimensional (3-D) reconstruction of the road surface which can provide enough information to conduct complete pavement condition survey for various types of distresses.

Stereovision is a computer vision technique which has been widely used for industrial inspection (Aguilar, 2005), robotic vision (Ayache, 1991), 3-D microscopy (Vignon, 2001) and medical imaging (Marchessoux, 2002). Its basic principle is to form a 3-D structure of the target object from images of the object taken from at least two different positions. The 3-D physical coordinate of the point on the object is obtained by matching the corresponding points in the images. Currently, most applications of stereovision technique are either in a static environment or with a small scale of measurement. Implementing the stereovision technique into a moving and vibrating system while maintaining an accuracy level of 1mm is a challenge. This project is to test the feasibility of this implementation.

Another reason to initialize this project is automation. The automation level of the current available survey technologies is still limited. Most state highway agencies still use manual surveys to obtain measurements of many types of pavement distresses, including cracking surveys. Manual surveys are inconsistent, costly, tedious, labor intensive, subjective, and unsafe. Due to rapid technological advances of sensors and computing, it is time to study the technical implementation of using stereovision techniques for comprehensive survey of pavement surface distresses.

The long term goals of this project are:

- 1) Automated survey for most of the distress types identified in the LTPP distress manual (SHRP, 1993), and pavement profiles
- 2) Limiting the production cost of the final product to no more than \$500,000
- 3) Using off-the-shelf hardware sensors and systems
- 4) Fully automated data collection and interpretation at highway speeds
- 5) Providing pavement surface model for the entire pavement surface at 1-mm resolution in X, Y, and Z directions
- 6) Providing geometries of surface defects for inclusion into proper distress indices

## STAGE I RESEARCH

In the first stage of this project, focus was on the implementation of the stereovision technique. The following research work has been conducted:

- 1) Setting up Quad-Camera System in the Existing Vehicle
- 2) Developing Mathematical Models for Stereovision Based on Pairs of Images
- 3) Correction of Lens Distortion
- 4) Extracting 3-D Surface from Pavement
- 5) 3-D Surface Reconstruction

### System Design

The components of the system in Stage I include 2 industrial computers, GPS device, Digital Measurement Instrument (DMI), 4 area-scan cameras, lighting devices and a generator. Software includes capturing program (captures pavement images, GPS and DMI), calibration program, matching program, 3-D reconstruction program, rutting index calculation program and a light controller program. All programs were written in VC++.

Figure 1 shows the appearance of the vehicle. Figure 2 shows work principle of the stereovision in the system. The captured images are first compressed by JPEG technique and stored in the hard disks of computer 1. At the same time, the images are transferred to the second computer to do real-time processing. The processing includes 2D crack detection on each of the 2-meter images and 3-D surface reconstruction from the pairs of the images. The comprehensive condition survey was conducted and the results are written in a pavement surface condition database.



Figure 1. The DHDV in stage I



## Mathematical Model

The geometrical relationships are studied among cameras and the object in order to establish the mathematical model. The geometry of two cameras is illustrated in Figure 3. Here  $C_1$  and  $C_2$  are the optical centers of two cameras.  $M_1$  and  $M_2$  is the two normalized image plane.  $P$  is a 3-D point, and  $P_1$  and  $P_2$  are the image point of  $P$  on the two-image plane respectively. We call  $C_1C_2$  baseline, the plane  $C_1C_2P$  epipolar plane. The intersections of the epipolar plane with the two image planes are two lines, i.e.,  $l_1$  and  $l_2$ , which are called epipolar lines. Obviously,  $l_1$  is uniquely determined by the baseline and  $P_2$ . We call  $l_1$  the epipolar line associated with  $P_2$ , and vice versa. These terms will be used in the later explanation of the models.

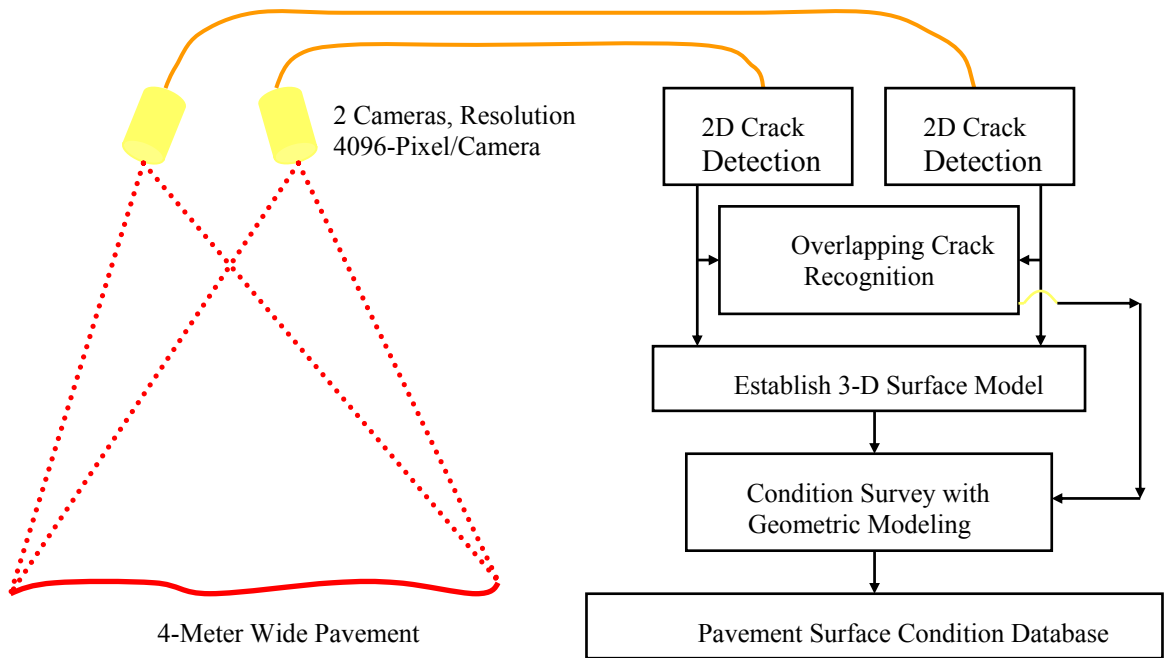


Figure 2. The basis of stereovision for condition survey

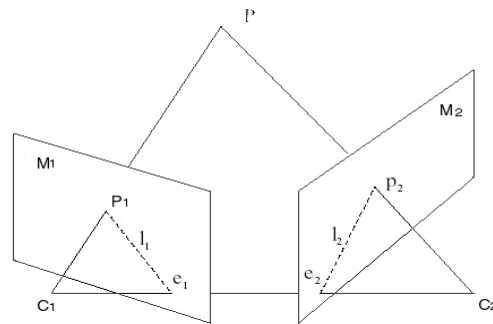


Figure 3. Epipolar geometry

It is important that the epipolar line  $l_2$  is the projection of line  $C_1P$  and  $l_1$  is the projection of  $C_2P$ , as shown in Figure 4. For any given image point  $P_1$  in the left image, the ray  $C_1P_1$  is fixed, since every point on the line  $C_1P_1$  will have the same image projection. As a result, the epipolar line  $l_2$  is fixed. The intersection of the baseline  $C_1C_2$  with the epipolar lines, i.e.,  $e_1$  and  $e_2$ , are epipoles.

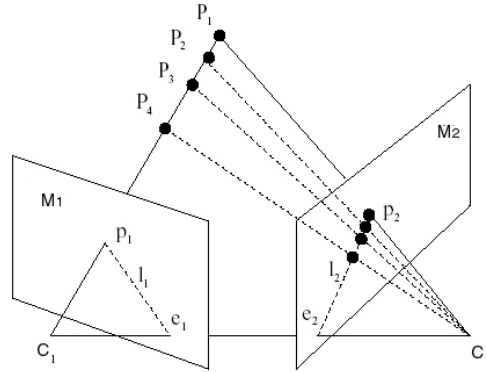


Figure 4. The relationship between point and its epipolar line

Another two key words in stereovision are correspondence and reconstruction. Given a point in one image and to find the corresponding point is the correspondence or matching problem. Reconstruction is to compute the 3-D coordinates of the corresponding point in space with a given correspondence.

At the very beginning, 3-D camera calibration is a necessary step in 3-D reconstruction. The purpose of the 3-D camera calibration is to decide the relationship between 2-D image planes and object space coordinates.

In this project, four algorithms have been developed and coded for the 3-D reconstruction of the pavement surface: calibration, distortion adjustment, matching and 3-D reconstruction.

### **Calibration**

Two calibration models have been tested to improve the accuracy: 3-D Modified DLT (Direct Linear Transformation) method (Hatzel, 1988) and Tsai Camera Calibration method (Tsai, 1986 & 1987).

### ***Standard DLT Equation***

The standard 3-D DLT equation is expressed as below

$$\begin{aligned} u &= \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \\ v &= \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1} \end{aligned} \quad (1)$$

Where  $u$  and  $v$  are the image plane coordinates.  $x$ ,  $y$  and  $z$  are the object-space coordinates. Equation (1) is derived by connecting the coordinates between the object space and image space.

### **Modified DLT Equation**

Standard DLT Equation only applies to the ideal situation. In order to improve the accuracy, the optical errors are considered. Equation (1) is changed as below:

$$\begin{aligned} u - \Delta u &= \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \\ v - \Delta v &= \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1} \end{aligned} \quad (2)$$

Where  $\Delta u, \Delta v$  are the optical errors, expressed as:

$$\begin{aligned} \Delta u &= \xi(L_{12}r^2 + L_{13}r^4 + L_{14}r^6) + L_{15}(r^2 + 2\xi^2) + L_{16}\xi\eta \\ \Delta v &= \eta(L_{12}r^2 + L_{13}r^4 + L_{14}r^6) + L_{15}\eta\xi + L_{16}(r^2 + 2\eta^2) \end{aligned} \quad (3)$$

Where,

$$\begin{aligned} [\xi, \eta] &= [u - u_0, v - v_0] \\ r^2 &= \xi^2 + \eta^2 \end{aligned} \quad (4)$$

$u_0, v_0$  are coordinates of the principal points in the image plane. Parameters L12 - L14 are related to the optical distortion while L15 & L16 are related to the de-centering distortion. Table 1 shows the implications of the parameters.

TABLE 1 the implication of the parameters

Parameters	Remarks
L1 - L11	Standard DLT parameters
L12 - L14	3rd-, 5th-, and 7th-order optical distortion terms
L15 & L16	de-centering distortion terms

### Calibration

Based on the equations (2) and (3), the following equation is derived:

$$\begin{bmatrix} \frac{x_1}{R_1} & \frac{y_1}{R_1} & \frac{z_1}{R_1} & \frac{1}{R_1} & 0 & 0 & 0 & 0 & \frac{-u_1x_1}{R_1} & \frac{-u_1y_1}{R_1} & \frac{-u_1z_1}{R_1} \\ 0 & 0 & 0 & 0 & \frac{x_1}{R_1} & \frac{y_1}{R_1} & \frac{z_1}{R_1} & \frac{1}{R_1} & \frac{-v_1x_1}{R_1} & \frac{-v_1y_1}{R_1} & \frac{-v_1z_1}{R_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{x_n}{R_n} & \frac{y_n}{R_n} & \frac{z_n}{R_n} & \frac{1}{R_n} & 0 & 0 & 0 & 0 & \frac{-u_nx_n}{R_n} & \frac{-u_ny_n}{R_n} & \frac{-u_nz_n}{R_n} \\ 0 & 0 & 0 & 0 & \frac{x_n}{R_n} & \frac{y_n}{R_n} & \frac{z_n}{R_n} & \frac{1}{R_n} & \frac{-v_nx_n}{R_n} & \frac{-v_ny_n}{R_n} & \frac{-v_nz_n}{R_n} \end{bmatrix} \begin{bmatrix} \frac{\xi_1^2 R_1}{R_1} & \frac{\xi_1^4 R_1}{R_1} & \frac{\xi_1^6 R_1}{R_1} & \frac{(\xi_1^2 + 2\xi_1^2) R_1}{R_1} & \frac{\xi_1 \eta_1 R_1}{R_1} \\ \frac{\eta_1^2 R_1}{R_1} & \frac{\eta_1^4 R_1}{R_1} & \frac{\eta_1^6 R_1}{R_1} & \frac{\eta_1 \xi_1 R_1}{R_1} & \frac{(\eta_1^2 + 2\eta_1^2) R_1}{R_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\xi_n^2 R_n}{R_n} & \frac{\xi_n^4 R_n}{R_n} & \frac{\xi_n^6 R_n}{R_n} & \frac{(\xi_n^2 + 2\xi_n^2) R_n}{R_n} & \frac{\xi_n \eta_n R_n}{R_n} \\ \frac{\eta_n^2 R_n}{R_n} & \frac{\eta_n^4 R_n}{R_n} & \frac{\eta_n^6 R_n}{R_n} & \frac{\eta_n \xi_n R_n}{R_n} & \frac{(\eta_n^2 + 2\eta_n^2) R_n}{R_n} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_{15} \\ L_{16} \end{bmatrix} = \begin{bmatrix} \frac{u_1}{R_1} \\ \frac{v_1}{R_1} \\ \vdots \\ \frac{u_n}{R_n} \\ \frac{v_n}{R_n} \end{bmatrix} \quad (5)$$

Where

$$R = L_9x + L_{10}y + L_{11}z + 1 \quad (6)$$

Numbers 1 to n indicate the calibration control points. In the above equation,  $[x_i, y_i, z_i]$  are the object-space coordinates for the control points which are all known. The control points must not be co-planar. In other words, the control points must form a volume. The control points are typically fixed to a calibration frame. The DLT parameters can be obtained using the least square method and iteration. Table 2 shows the minimum required number of control points for specific number of parameters.

**TABLE 2 Requirement of the Number of the Control Points**

No. of Parameters	Minimum No. of Control Points
11	6
12	6
14	7
16	8

## Experimental Results

To test the modified DLT calibration model, two cameras are used to take photos for the same calibration board simultaneously. Figure 5 shows the images from the left and right cameras. There are 22 control points in the calibration board. The coordinates of the 22 points in the object space and image space can be measured and substituted into equation (5). The parameters L1-L16 are then obtained from equation (5) by iteration and least square method. Figure 6 shows the interface of the calibration program.

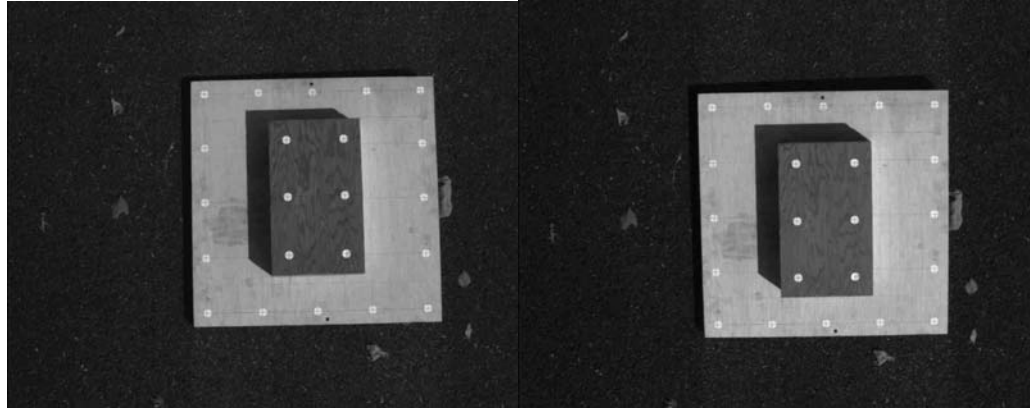


Figure 5. Left & Right images for calibration model

The main procedure of the algorithm is as follows:

- 1 In the first iteration, it is impossible to compute R since L9, L10, and L11 are not available. The coordinates of the principal point are not available as well. Therefore, the standard DLT with 11 parameters must be used.
- 2 From the second iteration on, compute R using the L9, L10, and L11 obtained from the first iteration. Compute  $[u_0, v_0]$  based on the L0 -L10 from the first iteration.
- 3 Repeat this procedure until all the L converges sufficiently. A reasonable tolerance level must be set for the convergence check in each iteration.

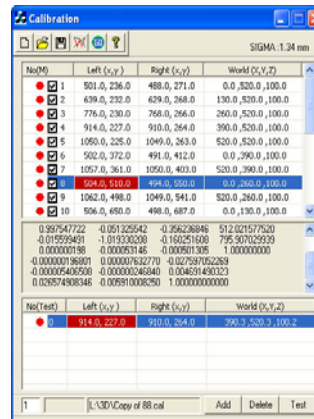


Figure 6. The interface of the calibration program

### ***Experimental Result by Using the DLT Method***

By using DLT method, the results including projection matrix (3 \*4) are listed below:

**TABLE 3. the DLT Results**

<b>Projection matrix for left Camera</b>			
0.918929618	-0.013814117	-0.147250562	140.695207360
0.007227535	-0.268132554	-0.874826848	650.785051092
0.000070426	0.000032920	-0.000459940	1.000000000
<b>Projection matrix for right Camera</b>			
0.941100636	0.011050581	-0.297224940	85.168048653
-0.003490740	-0.926320269	-0.248243646	610.847698963
0.000011345	0.000023721	-0.000480914	1.000000000
<b>Fundamental matrix</b>			
-.000000027995	0.000004191081	-0.000759195091	
-.000004922170	0.000000214244	0.013894849155	
0.000585871946	-0.014438358354	1.000000000000	
<b>Mean Error</b>	2.44 mm		

The average mean space error is 2.44 mm.

### ***Tsai Camera Calibration***

To improve the accuracy, another calibration method, Tsai calibration method is tested. It is a popular method in stereovision application. It is suitable for a wide area of application since it can deal with coplanar and non-coplanar points. It also offers the possibility to separately calibrate internal and external parameters. This option is particularly useful since it gives the possibility to fix the internal parameters of the camera, when known, and carry out only pose estimation.

The Tsai model is based on a pinhole perspective projection model (Figure 7) and the following eleven parameters are to be estimated:

- F:** Focal length of camera
- K:** Radial lens distortion coefficient
- C<sub>x</sub>, C<sub>y</sub>:** Coordinates of centre of radial lens distortion
- S<sub>x</sub>:** Scale factor to account for any uncertainty due to imperfections in hardware timing for scanning and digitization
- R<sub>x</sub>, R<sub>y</sub>, R<sub>z</sub>:** Rotation angles for the transformation between the world and camera coordinates
- T<sub>x</sub>, T<sub>y</sub>, T<sub>z</sub>:** Translation components for the transformation between the world and camera coordinates

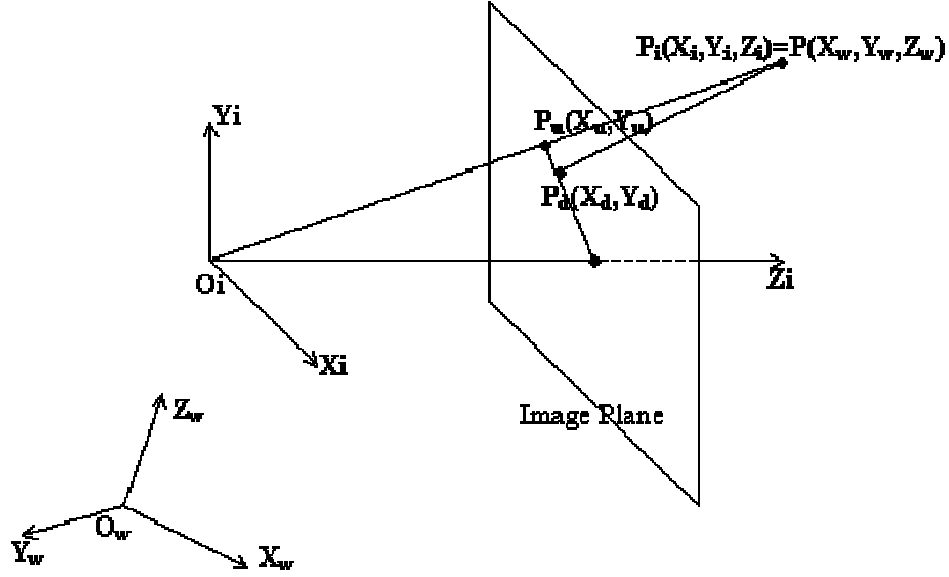


Figure 7. Pinhole Mode

The transformation from world  $(X_w, Y_w, Z_w)$  to image  $(X_i, Y_i, Z_i)$  coordinates considers the extrinsic parameters of the camera (Translation  $T$  and Rotation  $R$ ) within the equation:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T \quad \dots\dots\dots (7)$$

Where  $R$  and  $T$  characterize the 3-D transformation from the world to the camera coordinate system and are defined as follows:

$$T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}$$

$$\begin{aligned}
r_1 &= \cos(R_y) \cos(R_x) \\
r_2 &= \cos(R_x) \sin(R_y) \sin(R_z) - \cos(R_x) \sin(R_z) \\
r_3 &= \sin(R_y) \sin(R_z) + \cos(R_x) \cos(R_z) \sin(R_y) \\
r_4 &= \cos(R_y) \sin(R_x) \\
r_5 &= \sin(R_y) \sin(R_z) \sin(R_x) + \cos(R_x) \cos(R_z) \\
r_6 &= \cos(R_x) \sin(R_y) \sin(R_z) - \cos(R_x) \sin(R_z) \\
r_7 &= -\sin(R_y) \\
r_8 &= \cos(R_y) \sin(R_x) \\
r_9 &= \cos(R_x) \cos(R_y)
\end{aligned} \tag{8}$$

$(R_x, R_y, R_z)$  are the Euler angles of the rotation around the three axes.  $(T_x, T_y, T_z)$  are the 3-D translation parameters from world to image coordinates. The transformation from 3-D position (in the image coordinate frame) to the image plane is then computed through the following steps:

Transformation from 3-D world coordinates  $(X_i, Y_i)$  to undistorted image plane  $(X_u, Y_u)$  coordinates:

$$X_u = f \frac{X_i}{Z_i} \tag{9}$$

$$Y_u = f \frac{Y_i}{Z_i} \tag{10}$$

Transformation from undistorted  $(X_u, Y_u)$  to distorted  $(X_d, Y_d)$  image coordinates

$$X_d = X_u(1 + kr^2) \tag{11}$$

$$Y_d = Y_u(1 + kr^2) \tag{12}$$

Where:

$$r = \sqrt{X_d^2 + Y_d^2}, \text{ and } k \text{ is the lens distortion coefficient.}$$

Transformation from distorted coordinates in image plane  $(X_d, Y_d)$  to the final image coordinates  $(X_f, Y_f)$  are:

$$X_f = \frac{s_x X_d}{d_x} + C_x \tag{13}$$



$$Y_f = \frac{Y_d}{d_y} + C_y \dots\dots\dots (14)$$

$(d_x, d_y)$ : distance between adjacent sensor elements in the X and Y direction.  $d_x$  and  $d_y$  are fixed parameters of the camera. They depend only on the CCD size and the image resolution.  $(X_f, Y_f)$  are the final pixel position in the image.

### ***Experiment Result***

Figure 8 shows the calibration module used in the calibration process. In total, 59 labels are made on this calibration box. In the calibration process, let each camera in the stereo camera pair to see these dots and apply Tsai's approach to compute camera parameters. The labels of world coordinates in 3-D have been correctly determined.

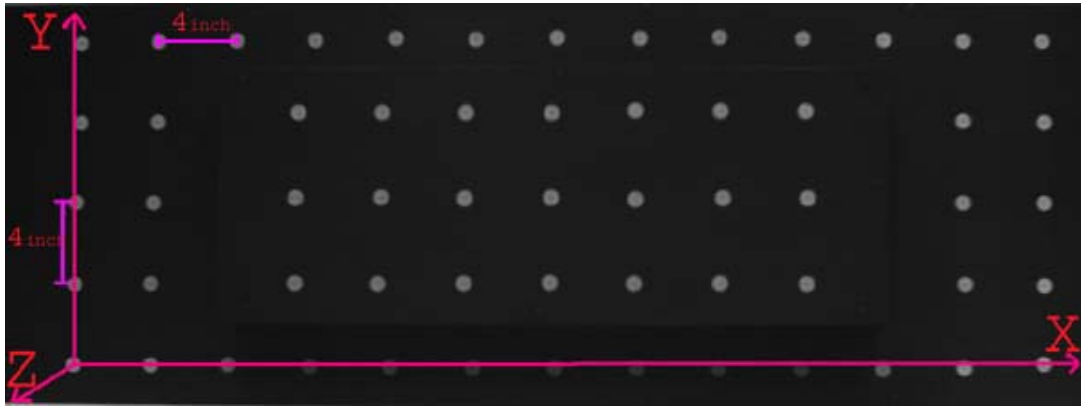


Figure 8. Calibration Module with 59 circle labels

For a pair of images, let each camera see the calibration module in Figure 8 and take a series of images to get image coordinate information. Then the Tsai method is applied to conduct 3-D calibration. Results include each camera parameters (intrinsic parameters and extrinsic parameters), rotation matrix and translation matrix for transforming 3-D coordinate system into 3-D image coordinate.

### ***Calibration Results by Using the Tsai Method***

To achieve the best results, the Tsai method with full optimization is implemented to analyze the groups of data.

Results concerning the Left Camera:

Data file: LeftCD.dat (59 points)

The intrinsic parameters

$$f = 15.230637 \text{ [mm]}$$

$$k = 1.144457\text{e-}03 \text{ [1/mm}^2\text{]}$$

$$\mathbf{Sx} = 0.991164$$

$$\mathbf{Cx} = 629.322071,$$

$$\mathbf{Cy} = 452.333599 \text{ [pixels]}$$

The extrinsic parameters

$$\mathbf{Tx} = -1497.952655,$$

$$\mathbf{Ty} = 1247.943535,$$

$$\mathbf{Tz} = 2268.498650 \text{ [mm]}$$

$$\mathbf{Rx} = 176.109399,$$

$$\mathbf{Ry} = -9.546206,$$

$$\mathbf{Rz} = -1.608673 \text{ [deg]}$$

$$\mathbf{Tz} / \mathbf{f} = 148.943120$$

$\mathbf{R}$

$$0.985764 \ -0.039257 \ 0.163491$$

$$-0.027684 \ -0.996986 \ -0.072470$$

$$0.165843 \ 0.066912 \ -0.983880$$

Normalized calibration error: 2.392541

Results concerning the Right Camera

Data file: rightCd.dat (59 points)

The intrinsic parameters

$$\mathbf{f} = 16.738583 \text{ [mm]}$$

$$\mathbf{k} = 1.200057\text{e-}03 \text{ [1/mm}^2\text{]}$$

$$\mathbf{Sx} = 0.994377$$

$$\mathbf{Cx} = 679.980691,$$

$$\mathbf{Cy} = 469.169413 \text{ [pixels]}$$

$$\mathbf{Tz} / \mathbf{f} = 151.126583$$

The extrinsic parameters

$$\mathbf{Tx} = -1625.970581,$$

$$\mathbf{Ty} = 1159.708667,$$

$$\mathbf{Tz} = 2529.644807 \text{ [mm]}$$

$$\mathbf{Rx} = 178.273696,$$

$$\mathbf{Ry} = -1.999355,$$

$$\mathbf{Rz} = -0.498954 \text{ [deg]}$$

**R**

0.999353 -0.009755 0.034609

-0.008703 -0.999499 -0.030428

0.034888 0.030107 -0.998938

Normalized calibration error: 2.594057

**TABLE 4 3-D Coordinates of Dots on the Calibration Module**

#	Xw	Yw	Zw	#	Xw	Yw	Zw
1	1000.000	1406.400	100.000	31	1711.200	1203.200	199.000
2	1101.600	1406.400	100.000	32	1812.800	1203.200	199.000
3	1203.200	1406.400	100.000	33	1914.400	1203.200	199.000
4	1304.800	1406.400	100.000	34	2117.600	1203.200	100.000
5	1406.400	1406.400	100.000	35	2219.200	1203.200	100.000
6	1508.000	1406.400	100.000	36	1000.000	1101.600	100.000
7	1609.600	1406.400	100.000	37	1101.600	1101.600	100.000
8	1711.200	1406.400	100.000	38	1304.800	1101.600	199.000
9	1812.800	1406.400	100.000	39	1406.400	1101.600	199.000
10	1914.400	1406.400	100.000	40	1508.000	1101.600	199.000
11	2016.000	1406.400	100.000	41	1609.600	1101.600	199.000
12	2117.600	1406.400	100.000	42	1711.200	1101.600	199.000
13	2219.200	1406.400	100.000	43	1812.800	1101.600	199.000
14	1000.000	1304.800	100.000	44	1914.400	1101.600	199.000
15	1101.600	1304.800	100.000	45	2117.600	1101.600	100.000
16	1304.800	1304.800	199.000	46	2219.200	1101.600	100.000
17	1406.400	1304.800	199.000	47	1000.000	1000.000	100.000
18	1508.000	1304.800	199.000	48	1101.600	1000.000	100.000
19	1609.600	1304.800	199.000	49	1203.200	1000.000	100.000
20	1711.200	1304.800	199.000	50	1304.800	1000.000	100.000
21	1812.800	1304.800	199.000	51	1406.400	1000.000	100.000
22	1914.400	1304.800	199.000	52	1508.000	1000.000	100.000
23	2117.600	1304.800	100.000	53	1609.600	1000.000	100.000
24	2219.200	1304.800	100.000	54	1711.200	1000.000	100.000
25	1000.000	1203.200	100.000	55	1812.800	1000.000	100.000
26	1101.600	1203.200	100.000	56	1914.400	1000.000	100.000
27	1304.800	1203.200	199.000	57	2016.000	1000.000	100.000
28	1406.400	1203.200	199.000	58	2117.600	1000.000	100.000
29	1508.000	1203.200	199.000	59	2219.200	1000.000	100.000
30	1609.600	1203.200	199.000	60			

**TABLE 5 Distorted Image Plane Error (Left Camera)**

Mean	Stddev	Max [pix]	Sse[pix^2]
0.962660	0.447636	2.093695	= 6.298005

**TABLE 6 Undistorted Image Plane Error (Left Camera)**

Mean	stddev	Max[pix]	Sse [pix^2]
0.976639	0.453129	2.096669	68.184559

**TABLE 7 Object Space Error: (Left Camera)**

Mean	stddev	Max [mm]	Sse [mm^2]
1.064684	0.489675	2.264307	80.786841

**TABLE 8 Distorted Image Plane Error (Right Camera)**

Mean	stddev	max[pix]	sse[pix^2]
1.038042	0.615425,	2.630565	85.541721

**TABLE 9 Undistorted Image Plane Error (Right Camera)**

Mean	stddev	max[pix]	sse[pix^2]
1.059238	0.633672	2.791428	89.486532

**TABLE 10 Object Space Error (Right Camera)**

Mean	stddev	max[mm]	sse[mm^2]
1.055019	0.634765	2.722671	89.040602

From the above results, the mean space error is 1.064 mm for the left camera and 1.055 mm for the right camera. The average mean space error is 1.059 mm which is almost half of the error from DLT method.

### **Distortion Adjustment**

Because lenses are not flat surfaces, the world points are therefore not projected on a plane, but rather in a surface, which can be considered to be spherical. This has the effect that straight lines are mapped to parabolas in the image. For example, the grid of Figure 9b is actually seen as shown at the Figure 9a. We therefore should consider lens distortions of a camera

According to the literature, the distortion function is dominated by the radial components, and especially by the first term. If  $(u,v)$  are the ideal image coordinates of a point,  $(u',v')$  are the distorted image coordinates, and  $(x,y)$  are the ideal normalized image coordinates (coordinates of the point  $m$ ), then:

$$u' = u + k(u-u_0)(x^2+y^2)$$

$$v' = v + k(v-v_0)(x^2+y^2)$$

The center of distortion is the same as the principal point  $(u_0, v_0)$ . Therefore, we also have to estimate the value of  $k$ , in the process of camera calibration.

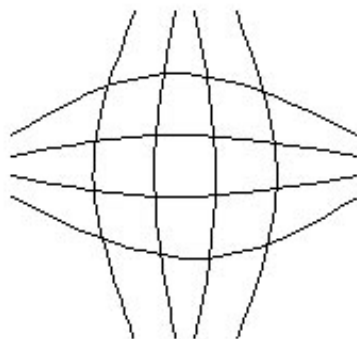


Figure 9a. Actual image of grid.

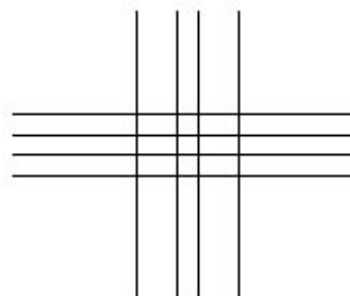


Figure 9b. Ideal image of grid.

Pictures of a grid with the cameras (Figure 10) were taken. By knowing that the actual distance of the horizontal and the vertical lines of the grid remain constant, we could estimate the radial distortion term. Because of the radial distortion, the image would be similar to Figure 9a, and the distance between the lines would not be the same. The principal point was estimated by the grid image. In particular, the shape of the "barrel" (the grid pixels look like barrels) was used to see where the lines started curving towards different directions. By this method estimates for the radial distortion term were made.



Figure 10. Captured Distorted Image

### Matching

Matching one single point in two images can be difficult due to noisy surfaces of pavements, since there could be many potential matches for a single point. To reduce the ambiguity, the spatial relationship among image features should be exploited. For example, a set of points along an epipolar line all-together can be matched instead of just matching one single point in images.

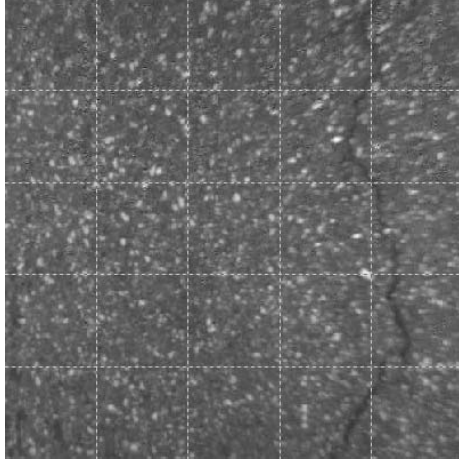


Figure 11a. Left image (500\*500)

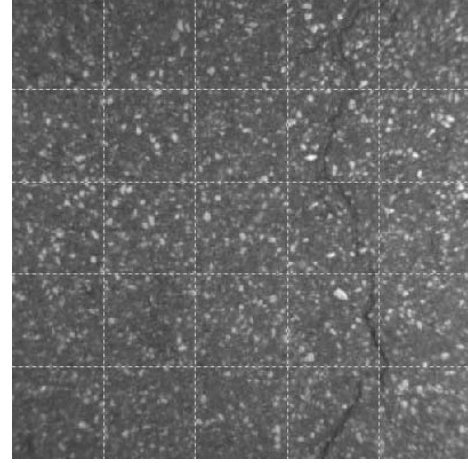


Figure 11b. Right image (500\*500)

### ***1) Edge Detection***

The  $3 \times 3$  Sobel filter matrix is applied on every  $3 \times 3$  area. Two values were obtained when H and V filters are applied on  $3 \times 3$  pixels. After getting magnitude of each point, one value was compared against the given threshold to determine the point whose magnitude is greater than the threshold. This process continues until all the  $3 \times 3$  areas in an image are scanned. The output image contains the edge feature (Figure 12a and Figure 12b).

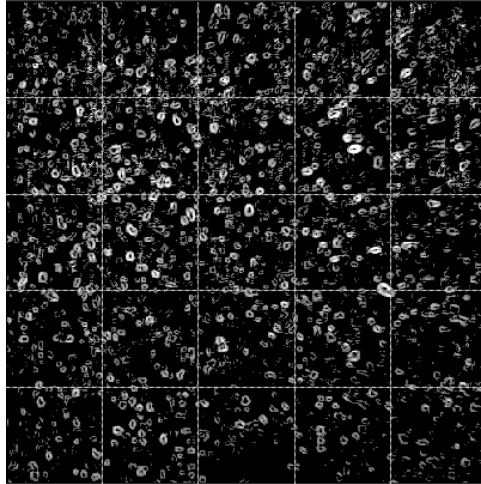


Figure 12a. Left Image

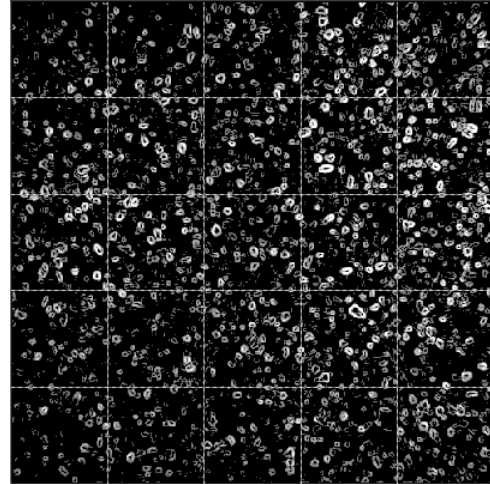


Figure 12b. Right Image

### ***2) Compute the Equation of Epipolar Line***

It is clear at this point that for any given image point in one image, its correspondence on the other image must lie on the epipolar line associated with this image point. This fact suggests that there is no need to search the whole 2D image for matching, but just search the 1D epipolar line, which greatly reduces the computation. The following issue is to determine the algebraic forms for the epipolar constraints.

The Fundamental Matrix will be used to establish a natural link between the pixel coordinate of a point in the left image and the projective epipolar line that correspond to that point. So an epipolar line is determined, given the pixel coordinate of a point of the left image. This is the key technique of finding the matching point. The epipolar line is always given by the equation:

$$P_r = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad P_l = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad P_r^T F P_l = 0$$

$$\Rightarrow \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \Rightarrow xU_0 + yU_1 + U_2 = 0 \quad \dots\dots\dots (15)$$

Where:

$$U_0 = F_{11}x' + F_{12}y' + F_{13}, \quad U_1 = F_{21}x' + F_{22}y' + F_{23}, \quad U_2 = F_{31}x' + F_{32}y' + F_{33}$$

But in practice, the straight line is always represented by using the equation  $Y=m \cdot X + c$  and transform the epipolar line into this form.

$$m = -\frac{U_0}{U_1} \quad C = -\frac{U_2}{U_1} \quad \dots\dots\dots (16)$$

### 3) Corresponding Matching

For each point in the left image, an epipolar line is determined on the right image. If the line is properly drawn, the matching point must lie on this line. For each point on the epipolar line, the distance is calculated between it and the point of the left image. Then three different weights are applied on the three differences that determine the similarity of the original point and possible match point. The equation is

$$D = w_1(l_L - l_R)^2 + w_2(\theta_L - \theta_R)^2 + w_3|m_L - m_R|^2 \quad \dots\dots\dots (17)$$

Where:

$w_1$  is the weight of the distance

$w_2$  is the weight of the difference in direction

$w_3$  is the weight of the difference in magnitude.

For each point on the epipolar line, a different value  $s$  is obtained. The smaller the  $s$  is, the more probably the point is the matching point in the right image. So the point with the smallest  $s$  value will be regarded as the matching point. The entire epipolar line is not searched, but only the point on epipolar line that are near the original point of the left image is searched. The search range can be enlarged or reduced by changing the range value.

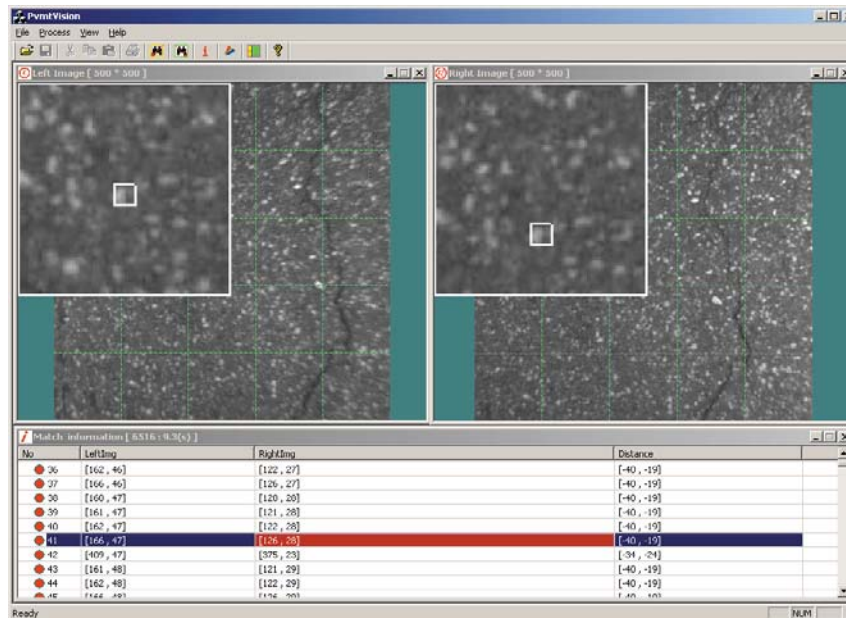


Figure13. Matching program

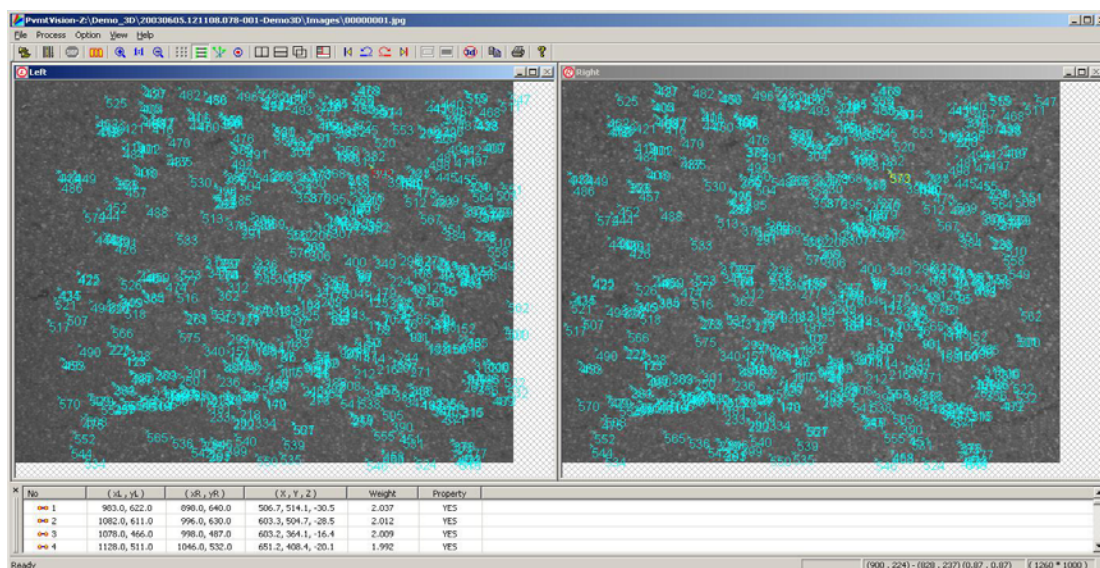


Figure 14. Matching Program



If only the pixels on the epipolar line are searched, the matching point may not be accurate if there is any error in finding epipolar lines. Therefore, the scheme of searching the matching points needs to be improved. For example, a “window” can be sliding along the epipolar line and all the pixels in that window should be searched to reduce the error in finding the epipolar line (Figure 13).

### 3-D Reconstruct

After finding the matching points, then the 3-D position of these points can be reconstructed. First the pixel coordinates are converted into image plane coordinates. This can be achieved by the following equation:

$$P_l = M_l^{-1} \bar{p}_l \quad P_r = M_r^{-1} \bar{p}_r \quad \dots\dots\dots (18)$$

The 3-D point will be achieved by using triangulation. Assume  $ap_l$  as the ray, which goes through the center of left camera like  $p_l$ ; And set  $T + bR^T P_r$  be the ray, which goes through the center of right camera. Due to that these two rays may not actually be intersecting, the point with minimum distance from each other is regarded as the intersecting point, which is the 3-D point being pursued after.

Let  $P_l \times R^T P_r$  be the vector which is perpendicular to these two rays, the equation bellow is obtained:

$$ap_l - bR^T P_r + c(p_l \times R^T P_r) = T \quad \dots\dots\dots (19)$$

In fact the 3-D point is the midpoint of the segment, which joins the two rays and parallel to the vector  $P_l \times R^T P_r$ . After solving the equation, then the 3-D point coordinates are then calculated:

$$P_w = aP_l + \frac{c}{2}(P_l \times R^T P_r) \quad \dots\dots\dots (20)$$

Figure 15 shows the interface of the 3-D construction program based on the above theory.

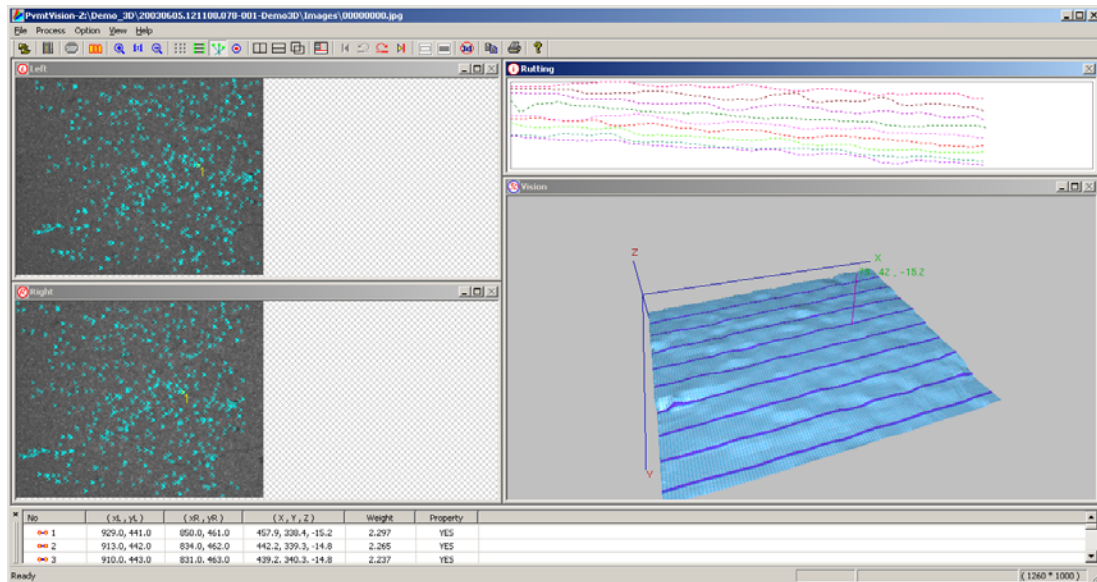


Figure 15. 3-D reconstruction program

## Result

The results obtained from Stage I are promising. The raw image resolution is within 1-mm accuracy. Pavement images can be captured with high quality with a speed above 100km/h (60mph). However, the resolution in 3-D reconstruction can only be made as 2-mm to 5-mm when the vehicle is not in motion. During the field capturing, the accuracy level of the 3-D reconstruction is greater than 5-mm.

Despite the fact that this accuracy level may not be sufficient for pavement survey of surface defects, a full set of algorithms and software have been established in this project for 3-D surface reconstruction of pavements. It is particularly worth noting that in 2005, the computing cost from the pixel matching program was expensive based on the observed processing speed. Throughout 2006, due to algorithm improvements, and rapid advance of microprocessor, such as dual-core and quad-core technologies, computing requirement was no longer excessive for the implementation of the developed algorithms.

The difficulties encountered during the research are largely due to the fact the image acquisition system including illumination was not advanced enough, and a high-end precision Inertial Measurement Unit (IMU), or gyro system, was not used due to high cost. In addition, a high precision calibration pattern needs to be made for the research. The current calibration pattern may have attributed to the low accuracy level.

## STAGE II RESEARCH

During the Stage II, the design of the system has been changed. The effort was put on the development of the software working with the newly adopted Laser Road Imaging System (LRIS). A series of software include the image capturing, automated distress analyzing and reporting programs have been developed. The hardware and the software

form a real-time automated pavement condition survey with a reliable and stable performance.

### **System Design**

The pavement imaging system in Stage II is based on the INO Laser Road Imaging System (LRIS). With two line scan cameras and lasers, this pavement imaging system is capable of constructing 4096x2048 images of full-lane width (12 ft, 3.657 m) pavement surface with complete and continuous coverage of pavement surface. The cameras and lasers reside inside a water-tight, aluminum container, which mounts on the external DHDV frame.

The cameras and lasers are powered and triggered through the LRIS Control Chassis. The LRIS Control Chassis connects to in-car computer and another custom Control Chassis box. The cameras mount on an aluminum alignment beam spaced equidistant from factory calibrated readings. The calibrated spacing of the cameras ensures full four meter coverage of pavement surface. The height has been previously calibrated. The frame has been designed to withstand all dynamic and static loads incurred during normal driving conditions. For 1,000-mile pavement Images, the captured data is about 600 GB after compression.



Figure 16. The Stage II DHDV

### **Laser Road Imaging System (LRIS )**

The LRIS system is composed of two high resolution line-scan cameras and lasers that are configured to image full 4m transverse road sections with 1mm resolution at speeds that can surpass 100 km/h (60mph).

This imaging system was designed to increase the contrast and visibility of both small longitudinal and lateral road cracks. Using high power laser line projectors and special collection optics, the LRIS system can operate in full daylight because it is immune to variations in outside lighting conditions and shadows cast by roadside objects, viaducts and the inspection vehicle itself. Other advantages of this system include its compact size and lower power consumption of less than 200 watts. Figure 17 shows the LRIS sensors mounted on the DHDV.



Figure 17. LRIS

LRIS use line-scan camera instead of area-scan camera. Line-scan imaging (Figure 18) uses a single line of sensor pixels (effectively one-dimensional) to build up a two-dimensional image. The second dimension results from the motion of the object to the camera, here it comes from the motion of the camera which is moving with the vehicle. Two-dimensional pavement images are acquired line by line by successive single-line scans while the image sensor in the vehicle moves (perpendicularly) past the line of pixels in the pavement.

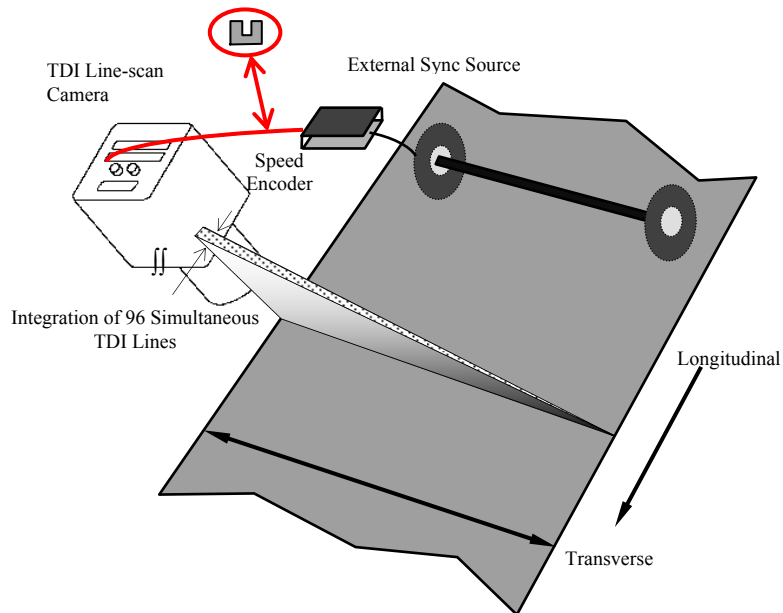


Figure 18. Line Scanning Method

### **Automated Distress Analyzer (ADA)**

A dedicated crack identification and categorization software (Figure 19) named Automated Distress Analyzer (ADA) was developed. As an addition to the pavement image capturing system, ADA was designed to detect longitudinal, transverse, block, and alligator cracks, calculate their lengths, widths, and other general information sets. ADA was designed for crack detection in asphalt pavement, and also works for detection of rigid pavement cracks. It works in real-time with a speed up to 60 mph. ADA also has a post-processing function. With several quality control features, ADA can give very good results for all types of pavements with limited human interference.

### **Algorithm**

The image processing in ADA includes the following steps:

- Image enhancement: in order to produce binary image with cracking information, images are all enhanced to eliminate noises. ADA uses histogram equalization and adaptive thresholding to generate binary images for further processing (Russ 1999). Particularly, image segmentation is used to divide images into smaller areas and different thresholding values are used for image enhancement.
- Feature analysis: in order to determine geometric characteristics of cracks, such as length, width, orientation (angle), algorithms such as closing, thinning, and Hough transformations are conducted to the preprocessed images.
- Classification and analysis of cracks: based on results from previous steps, methods such as projection are used to determine the main types of cracks, such as longitudinal and transverse cracks, and fatigue and block cracks.

### **Speed**

The real-time processing of the pavement images requires a very efficient algorithm and computing structure. For example, as DHDV drives at 60mph (26.8m/s), about 13 frames of 2048\*4096 pixels gray scale pavement images are captured in a second which means each image has to be processed in less than 100ms. This fast speed is achieved by improving the efficiency on both software and hardware through parallel computing. The first parallel feature of the real-time system is based on industry standard Symmetrical Processing (SMP). The dual-processor SMP system used for DHDV is programmed with multi-thread techniques, so that each CPU has its own resources and instructions to carry out the image analysis tasks. The second parallel feature is based on the within-CPU parallel capability called Single Instruction Multiple Data (SIMD). Through the multiple parallel computing mechanisms, ADA is able to maintain its performance at high speed.

### **Quality Control**

Other than the efficient algorithms implemented in the automated operation of ADA, manual intervention is also allowed to micro-adjust the parameters to further improve the accuracy of distress detection. The parameters in the quality control program include the pavement types, the pavement noise level, and the road border. The quality control

interface is user friendly and can be conducted during real-time capturing of pavement images.

## Reporting

The crack information from ADA can be read by the ReportWriter software (Figure 20). It is a type of utility software to generate crack statistics using predefined crack survey protocols such as CI, AASHTO, UK SCANNER, Texas DOT protocol etc.

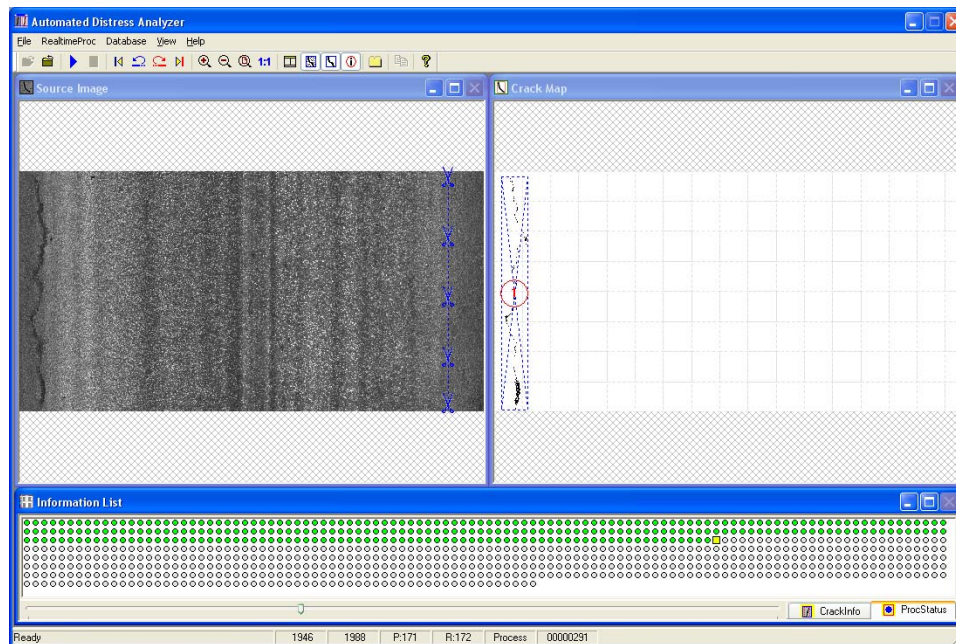


Figure 19. Automated Distress Analyzer (ADA) overall interface

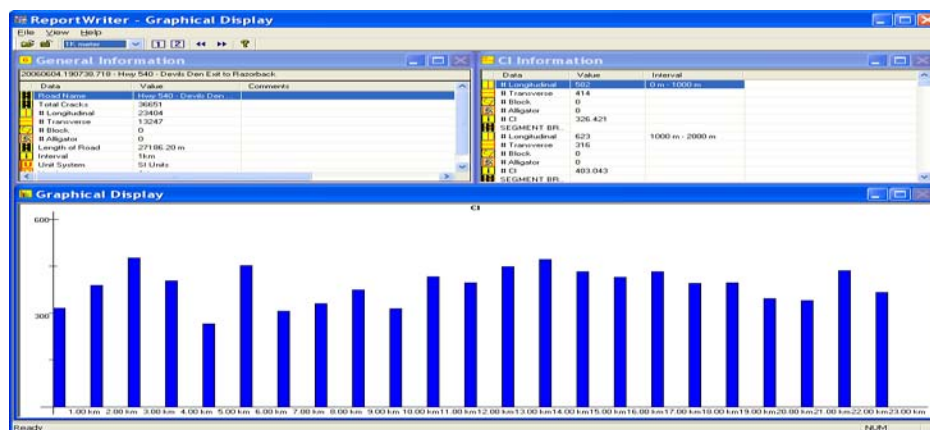


Figure 20. A road statistics reporter

## Field Evaluation

The stage II DHDV has been put in the market for nearly one year. Thousands of miles of pavement condition survey have been conducted. The performance of this system has been demonstrated to be stable, efficient and accurate.

## CONCLUSIONS AND RECOMMENDATIONS

It has been demonstrated in this report that the research team has successfully developed a set of algorithms and software for the implementation of establishing 3-D reconstruction of pavement surface for the purpose of comprehensive automated distress survey. Due to stated reasons, the accuracy level of the implemented system is not production worthy at this time.

In order to overcome the limitations, it is recommended that further research be conducted to foster a technological ‘quantum leap’ to provide a single, multi-function vehicular platform for real-time data acquisition and analysis for various data sets for pavement and associated asset management. The proposed new project could be characterized as high-risk; however, the payoff from successful completion of the project is certainly high-reward, providing vital data-based asset management tools for the roadway industry worldwide.

The major intended tasks of the recommended effort include:

1. Establish a three-dimensional (3-D) surface model of the entire pavement at one-millimeter resolution and in real-time using multiple laser-based imaging devices;
2. Provide an automated, real-time *condition survey* of pavements through imaging analysis of one-millimeter resolution 3-D surface models of pavements;
3. Provide an automated, real-time, multi-path *roughness survey* of pavements through an inertial measurement sub-system and one-millimeter resolution 3-D surface models of pavements;
4. Provide an automated, real-time *roadway asset survey* through the use of High-Definition cameras and stereovision techniques for signs, guardrails, pavement markings, and other roadside elements;
5. Develop and provide a comprehensive database and user-friendly interface for easy and efficient data retrieval, analysis, and reporting;

The five tasks will be based on the development of software algorithms, along with the following three efforts using off-the-shelf technologies:

1. Integrating a high-precision inertial measurement sub-system and a differential GPS positioning sub-system;
2. Leveraging existing and advanced sensor technologies in inertial measurements and laser-based imaging;
3. Exploiting 64-bit terra computing capabilities at the desktop level through multi-chip, multi-core CPU, and streaming platforms for data and image processing.

Development of a 3-D surface model of pavements requires multiple laser-based imaging sensors, resulting in combined and sustained data rate of over 300 megabyte per



second (MBPS) into the acquisition computer. In 2005, the Arkansas research team was able to acquire and analyze pavement 2D images for cracking analysis at the sustained data rate of 120 MBPS. There is currently no technical difficulty in acquiring sustained data at 300 MBPS with an off-the-shelf high end workstation and peripherals. The challenge, then, is the development of software algorithms to process and analyze the data at high speed. It is anticipated that before 2009, workstation class computers that possess 16 computing cores (dual-CPU-socket) and terra computing add-on board will not be uncommon. It is reasonable to envision that such computers will be available for the recommended research in 2008 to process 300 MBPS of data and images for several hours at a time.

The proposed deliverables from the recommended new research include a fully equipped and functional next-generation Digital Highway Data Vehicle (DHDV), which possesses the following real-time capabilities of providing:

1. Performance, condition, and location data sets based on pavement indices and summaries of roughness, rutting, surface cracking, pothole, other surface deformation and defects for both asphalt and concrete surfaces.
2. Roadway asset information on location, type, and categories of signs, guardrails, pavement markings, etc. Condition analysis on these elements shall be also conducted. The Manual on Uniform Traffic Control Devices (MUTCD) shall be the basis for sign recognition and identification. Other Federal standards shall be used for inventory purposes of other elements.

For many years, there was a lack of coordinated funding at the Federal level for developing technologies for automated pavement distress surveys. Based on the outcomes of the IDEA projects and rapid technological advances of sensors and computing, it is foreseeable that in the next few years, the proposed new system will be production worthy, only if research funding at adequate level can be sustained.

## REFERENCE

- Aguilar, J.J., Lope, M., Torres, F. and Blesa, A. (2005). "Development of a stereo vision system for non-contact railway concrete sleepers measurement based in holographic optical elements", *Journal of Measurement*, Volume 38, Issue 2, Pg. 154-165.
- Ayache, N and Lustman, F (1991). "Trinocular Stereo Vision for Robotics", *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 13, No. 1, January.
- Hatze, H. (1988), "High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach". *Journal of Biomech* 21, Pg. 533-538.
- Marchessoux, C., Richard, N., and Fernandez C. (2002). "Description of Simple Method in 3D Reconstruction in Medical Imaging", *IEEE Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission*.
- National Research Council(1993), "Distress Identification Manual for the Long-Term Pavement Performance Project", SHRP-P-338, Washington, DC, 1993.



- Russ, J. C. (1999). *The Image Processing Handbook*, CRC Press/IEEE Press, Boca Raton.
- Tsai, R. Y. (1986). "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, Pg. 364-374.
- Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, August.
- Vignon, F., Le Besnerais, G., Boivin, D., Pouchou, J.L., and Quan, L.(2001). "3D Reconstruction from Scanning Electron Microscopy Using Stereovision and Self-calibration," in *Physics in Signal and Image Processing*, Marseille, France, January.