IDEA

Innovations Deserving
Exploratory Analysis Programs

Highway IDEA Program

Automated Pavement Distress Survey through Stereovision

Final Report for Highway IDEA Project 88

Prepared by:
Kelvin C.P. Wang, University of Arkansas, Fayetteville, AR

August 2004

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

INNOVATIONS DESERVING EXPLORATORY ANALYSIS (IDEA) PROGRAMS
MANAGED BY THE TRANSPORTATION RESEARCH BOARD (TRB)

This NCHRP-IDEA investigation 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 and 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 5th Street, N.W., Washington, D.C. 20001 (phone: 202/334-1461, fax: 202/334-3471, http://www.nationalacademies.org/trb/idea)

ACKNOWLEDGEMENTS

The research team would like to thank Dr. Inam Jawed and his staff of Transportation Research Board for their assistance in coordinating this project. Review comments from the IDEA panel helped researchers revise the final report. The University of Arkansas provided matching funds for this study. Engineers from the Arkansas State Highway and Transportation Department (AHTD) provided various supports, including Chief Engineer Mr. Bob Walters, and Division Head Mr. Alan Meadors. Their insight and encouragement made this project possible. Last but not the least, individuals from the University of Arkansas gave moral and material support in carrying out the tasks of the project, including Dr. Robert Elliott, Dr. Collis Geren, Admiral Jack Buffington, Dr. Melissa Tooley, and others.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Pavement condition survey includes detection of surface distresses, such as cracking, rutting, and other surface defects, and can also include survey of pavement roughness in certain cases. The research conducted within this IDEA project targets at the feasibility study of comprehensive survey of pavement condition in its broad definition through the use of stereovision technology. It is a proof-of-concept study.

In the past two decades, various agencies tried to develop automated methods to identify, classify, and quantify pavement surface distresses. Collection of many types of distresses, particularly cracking, still cannot be fully automated in production mode. Most federal and state highway agencies still use manual labor to collect distress data by either patrolling pavements, or rating pavements via looking at pavement surface images collected through automated method. A small number of highway agencies are using semi-automated survey systems that are post-processing based, and require substantial human interaction.

Based on the Digital Highway Data Vehicle (DHDV) and associated automated distress survey technologies developed at the University of Arkansas, this IDEA project demonstrates that it is feasible to apply stereovision techniques to establish pavement surface in three-dimensional (3D) space with off-the-shelf hardware. This new development can lead to the use of an alternative, more cost-effective and more powerful method to collect and analyze pavement condition data at sufficient resolution.

The basic techniques in this proof-of-concept study are based on traditional stereovision methods, some of which are successfully used in aerial photography. It is known now that such techniques with certain modifications are applicable in an environment which demands several magnitudes better accuracy, even though further study is needed to demonstrate the implementation of this new technology in the field.

In view of the fact that all necessary hardware platforms are high-performance and high-resolution, and off-the-shelf, algorithm development and their implementations would be the key to fulfilling the long term goal of establishing 3D pavement surface at one millimeter resolution in all three directions, and produce comprehensive condition results at real-time in a cost-competitive environment.

Similar to the application of inertial profiling techniques to the measurement of pavement roughness, the basic techniques for such measurement were known many years. It took nearly several decades for widespread and successful application of laser based profiling techniques. For this stereovision technology to be used in automated pavement distress survey, it will take substantial resources and time. This IDEA project gave us the opportunity to start this process.

IDEA PRODUCT

Pavement condition survey normally includes surface distresses, such as cracking, rutting, and other surface defects. Broadly, pavement roughness is also included as a condition survey item. The research conducted within this IDEA project targets at the feasibility study of comprehensive survey of pavement condition in its broad definition through the use of stereovision technology.

Based on the Digital Highway Data Vehicle (DHDV) and associated technologies developed at the University of Arkansas, this IDEA project demonstrates that it is feasible to apply stereovision techniques to establish pavement surface in three-dimensional (3D) space with off-the-shelf hardware. This new development can lead to the use of an alternative, more cost-effective and more powerful method to collect and analyze pavement condition data at sufficient resolution.

CONCEPT AND INNOVATION

BACKGROUND

Results from pavement condition survey are critical for pavement engineers to determine the needs of maintenance and rehabilitation for both network and project level studies. Pavement condition survey includes identification and classification of various types of surface cracks, identification of patching and potholes, and quantification of rutting and shoving. Determination of other types of distresses, such as bleeding, polished aggregate, and raveling, are also made during certain pavement condition surveys. For pavements with jointed Portland cement concrete surfaces and continuously reinforced concrete surfaces, in addition to cracks, spalling, faulting, corner breaks, and other types of distresses may be also determined during condition survey (1). In this proposal, pavement roughness is also broadly defined as a part of condition survey; even though, it is highly correlated with serviceability.

In the past two decades, various agencies tried to develop automated methods to identify, classify, and quantify pavement surface distresses (2). It is generally agreed that rutting and roughness measurements can be automatically collected at highway speed. Collection of many other types of distresses, particularly cracking, still cannot be fully automated in production mode. Most federal and state highway agencies are using manual labor to collect distress data by either patrolling pavements, or rating pavements via looking at pavement surface images. A limited number of highway agencies are using semi-automated survey systems that are post-processing based, require substantial human interaction, and slow as their processing speed is about a fraction of normal driving speed.

In the past few years, researchers at the University of Arkansas made substantial progress in automatically identifying and classifying pavement surface cracks at highway speed using a data collection system with one high-resolution digital camera and parallel processing. Currently, the developed system can collect two-dimensional pavement surface images, identify and classify four types of cracks at over 60 MPH. The four types of cracks are longitudinal, transverse, alligator and block. The spatial resolution of the system is 4,096 pixels per lane. The size of cracks that can be identified and classified is about 1-millimeters. In this IDEA Phase-I project, the researchers with technical background of automated cracking survey studied the feasibility of using stereovision techniques in establishing 3D surface model of pavement, with the long-term goal of conducting comprehensive pavement condition survey at high-resolution and at high-speed.

OTHER DEVELOPMENTS

It should be noted that there have been numerous publications on image processing algorithms for pavement crack survey. The literature search concentrates on system design and implementation, and most recent development. In the past two decades, there have been a number of efforts at automating pavement condition survey. One direction of the research is on establishing 3D surface models of pavement by using laser technologies and shadow Moire optical interference method. The other direction is using imaging or laser approaches to detecting pavement surface cracks based on 2D characteristics of pavement surface.

3D Approaches

Phoenix Scientific Inc. in California (http://phnx-sci.com) developed a phase-measurement Laser Radar (LADAR) to measure the 3D properties of pavements. The Laser Radar uses scanning laser and reflector to measure the reflecting times across pavement surface, therefore establishing a 3D pavement surface after the Laser Radar moves longitudinally along the traveling direction. Its system, as claimed, is able to produce roughness and rutting data at this time.

Another company, GIE Technologies Inc. in Canada (http://www.gietech.com/), has the LaserVISION system, which also models the 3D surface of pavements. The lasers are stationary and four of them are used to cover full lane-width. The service it provides is primarily for roughness and rutting survey. At this time, there is no independent evidence that laser based technologies are able to provide usable data for pavement cracking survey and other condition survey.

In addition, researchers at Illinois Institute of Technology developed a 3D technology for pavement distress survey through a NCHRP-IDEA grant (Project 13) in the mid-1990s (3). That technology uses the shadow Moire optical interference method and can only identify off-the-plane crack. That is, it can only find cracks, if both sides of which are not on the same plane. The limitation of this technology is apparent, as relative to the size of cracks, sides of most cracks are on the same plane.

The primary problem with laser based 3D systems is that their resolution is too coarse to be usable for pavement analysis, particularly in identifying cracks and other surface distortions at necessary resolution. We believe the spatial resolution of laser based systems is larger than five millimeters. The limitations of laser based 3D systems prompted us to examine a different approach to establishing 3D surfaces of pavements.

2D Approaches

The following are abbreviated discussions on some significant developments based on 2D approaches in chronological order.

In the late 1980s, the Japanese consortium Komatsu built an automated-pavement-distress-survey system (4), comprising a survey vehicle and data-processing system on board to simultaneously measure cracking, rutting, and longitudinal profile. Maximum resolution of 2048 x 2048 is obtained at the speed of 10 km/h. The Komatsu system worked only at night to control lighting conditions. The system represents an implementation of the most sophisticated hardware technologies at that time. However, it does not output the types of cracking and only works during the night. There are no known operating units at this time.

From late 1980s to early 1990s, Earth Technology Corporation (2) created a research unit called Pavement Condition Evaluation Services (PCES). The automated system created by PCES was the first to use line-scan cameras at 512-pixel resolution to collect pavement data. One important factor for discontinuing the effort is that the necessary technologies associated with the image capturing and processing were not mature enough. PCES designed, produced their own hardware and made their own system level software, which were not only costly, but also limited the research team from obtaining higher performance equipment from third parties at a later time.

In the early 1990s, NCHRP also funded a research called *Video Image Processing for Evaluating Pavement Surface Distress,* project 1-27 (5). The research in that project used traditional imaging algorithms with 2D images of pavements. The research was experimental and the final report was not formally published.

The Swedish PAVUE (2) acquisition equipment includes four video cameras, a proprietary lighting system, and four S-VHS videocassette recorders. The image collection subsystem is integrated into a Laser RST van. The off-line workstation is based on a set of custom designed processor boards in a cabinet to analyze continuous pavement data from the recorded video images. Surface images are stored on S-VHS tapes in analog format. Due to some limitations of the PAVUE system, it is not used in US any more.

Dr. Max Monti of the Swiss Federal Institute of Technology (EPFL) completed his Ph.D. on the design and implementation of a new pavement imaging system (6). With the developed system, Crack Recognition Holographic System (CREHOS), the pavement surface is scanned with a focused laser beam along a straight line in the lateral direction, while the longitudinal scan is conducted with the movement of the vehicle. The hardware system was dismantled due to performance limitations.

From 1995, the US Federal Highway Administration awarded contracts to LORAL Defense Systems in Arizona, now a unit of Lockheed-Martin, to provide an Automated Distress Analysis for Pavement, or ADAPT for short. The researchers used techniques developed for military purposes. The data source is digitized images from PASCO's 35-mm film. The delivered system after project completion could not be used and Arizona DOT currently is funding Lockheed-Martin to continue the research.

Since late 1996, RoadWare Corporation has been actively using a new product, WiseCrax, for automated survey of pavement surface. The data collection uses two cameras synchronized with a strobe illumination system, with each camera covering about half-width of a pavement lane. The image processing is done in the off-line office environment relying on the host CPUs to conduct image processing with substantial operator assistance.

In the late 1990s, NCHRP IDEA funded a study (Project 81) titled Automated Real-Time Pavement Crack Detection and Classification System (7). This project was primarily focused on algorithm development in processing images for cracks. Based on what the PI knows about Project 81, the project did not include data acquisition of images and relied on data provided by third parties for analysis. It also appears that the technology developed with the IDEA project was not field demonstrated.

In all, systems based on the Swedish PAVUE technology were used in US briefly in the mid 1990s. At this time, WiseCrax of RoadWare is used system in US with automated features of analyzing cracks. Technologies developed by the PI and staff at the University of Arkansas on automated cracking survey are marketed by WayLink Systems Corporation and its partners.

OUR INNOVATIONS

Research Goals

The primary goal of this project is to study the feasibility to apply stereovision technology to develop a three-dimensional (3D) surface model of pavement surface. The first step is to refine the current system to have an acquisition system to have the capability of 1-millimeter resolution, and of multi-camera capability. The second step is to develop algorithms to establish 3D surface of pavements. When the resolution of the 3D surface model is sufficiently high, such as 1-millimeter, the vast majority of surface distresses for both flexible and rigid pavements can then be identified and quantified through geometric modeling. Roughness can be also determined by using the 3D surface model, as longitudinal profiles of pavement surface can be also established through the use of the 3D surface model. The existing hardware platform at the university, including the Digital Highway Data Vehicle (DHDV) and associated hardware and software, are fully utilized for the research.

Stereovision with a pair of simultaneous image sources is not new. Highway agencies have been using the principle in photogrammetry in highway design and planning. In the recent decade, the failure of newer technology to provide solutions to comprehensive pavement condition survey prompted us to examine the possibility of using stereovision for that purpose. With rapid advances in digital camera and computing power of microcomputers, we believe that the time is right to integrate such a system with commonly available devices and develop efficient and accurate algorithms in software sub-systems to achieve the following long term goals:

1. High-resolution. The 3D surface model will provide 1-millimeter resolution in all three coordinate axis directions, X, Y, and Z.
2. Comprehensive survey. The information obtained from the 3D surface model includes cracking, roughness, rutting, potholes, and other surface defects associated with both flexible and rigid pavements.
3. Real-time at highway speed for both image acquisition and processing. Based on the new real-time technology developed at the University of Arkansas for pavement cracking with 2D images, the realization of this capability will provide condition results while images are being acquired at real-time.

Current pavement condition data collection uses a mixture of automated devices and manual methods. For pavement roughness and rutting, the automated method is using static laser sensors to determine longitudinal and transverse profiles of pavements. This technology is mature, but, costs over $100,000 in materials (not including the vehicle and R&D cost) to make such a system with five lasers for the transverse profiling.

For surveying other pavement surface defects, the predominant method is to use manual labor to walk pavements or examine visual information collected in the field to determine the conditional parameters. The speed for such data analysis is slow, normally 1 to 2 miles per hour per person if complete pavement surface is covered with the visual examination of collected images. It usually costs over $20 per mile for manual surface condition survey, not including the cost to collect the images.

When the long-term goals are achieved, the cost to assemble such as a system, including all components and labor involved in assembling, is about $100,000 without considering the cost of the vehicle, precision gyro, R&D, and other monetary issues. The system is to be able to collect all necessary data and produce comprehensive survey results, including cracking, roughness, rutting, and other surface deficiencies, at the spatial resolution of 1-millimeter.

Current Development with DHDV

As a part of a research effort to develop a digital highway data vehicle started in the mid 1990s, the researchers at the University of Arkansas focused on the development of a real-time automated system for distress survey. After examining the feasibilities of using neural net and fuzzy sets, traditional imaging algorithms and customized approaches have been developed to achieve the goal of real-time processing at highway speed. Figure 1 shows the digital highway vehicle at the university. Its basic features are:

(1) The vehicle is based on a full-digital design.

(2) The vehicle includes two sub-systems: sub-system of pavement surface image collection, and sub-system of automated distress analyzer to identify and classify pavement cracks at real-time.

(3) The vehicle collects right-of-way video or still frame images and save them in digital format in real-time.

(4) The vehicle acquires the location through the use of a GPS device and a distance-measuring instrument (DMI).

(5) Real-time relational database engine, inter-computer communication techniques, and multi-computer and multi-CPU based parallel computing.

Pavement images are stored in JPEG format in RAID disks when the images are being acquired. The resolution of the pavement surface images is 4,096-pixel per lane. The imaging system is able to cover the complete surface of several thousand miles of pavement without changing storage device or downloading. The data collection and processing are conducted at highway speed.

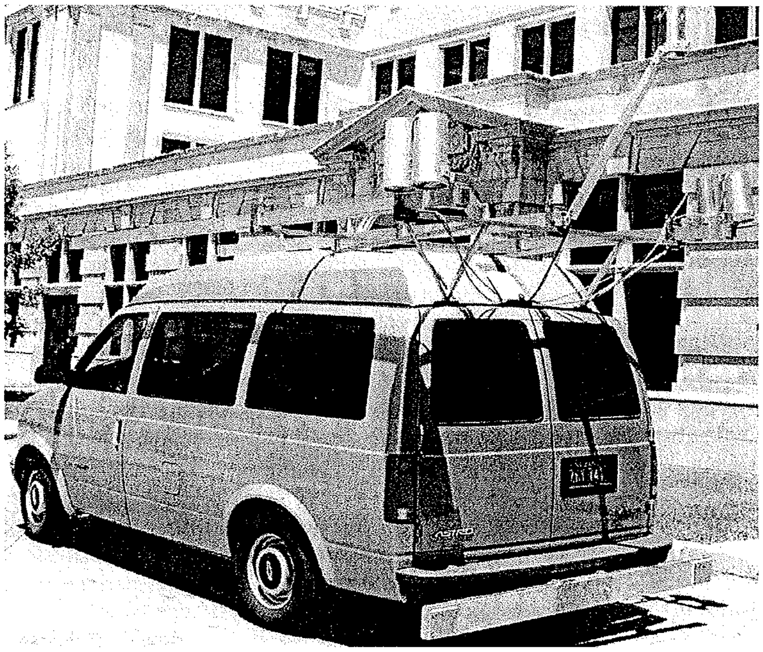


Figure 1 Exterior of the Data Vehicle

Two key issues in automated survey include improving processing speed, and developing sufficiently accurate algorithms. Highway pavement surface contains numerous foreign objects, such as oil residue, dirt, lane markings, vehicle's tire mark, tree limbs, and other non-distress related items. The task of the research was to develop algorithms to correctly distinguish cracks from these non-distress noises. The implementation of the developed algorithms in a parallel computing environment produced a real-time system that can automatically identify and classify pavement surface cracks while high-resolution digital images are being acquired and archived into a multimedia database at highway speed. This image processing system for cracks is called Distress Analyzer.

The first step in the image processing process is to distinguish any cracks from other non-distress noises. The primary method in this step relies on analytical descriptions of distresses' characteristics. The second step is to connect and vectorize the detected cracks, and establish a distress database related to location, orientation, and other geometry information of each crack. Based on the geometric information obtained in the second step, cracks can be classified using any pre-defined distress categorization protocols. Several distress protocols are incorporated into the system, including AASHTO Interim Distress Protocol, Texas distress manual, and Universal Cracking Indicator (CI) from World Bank. These indices are immediately computed and available for analysis when the vehicle completes a data collection field trip.

Figure 2 illustrates the framework of the data acquisition and processing in a parallel environment. A dual-CPU computer is used for data acquisition of GPS data, DMI data, and images. These data sets are moved to a multi-CPU computer at real-time for the distress analyzer. The distress analyzer has a project manager for parallel processing, which coordinates the processing of images among the *n* processors. The current implementation is using two CPUs for the distress analyzer. Figure 3 shows a screen shot of the distress analyzer working at real-time with two processors. At the bottom of each process are the analysis results for crack geometry and basic classification. Each process shows the original image on the left and processed binary image on the right. Each binary image shows each identified crack in a bounding box with a unique integer number. The right-most window illustrates the processing status of the analyzers.

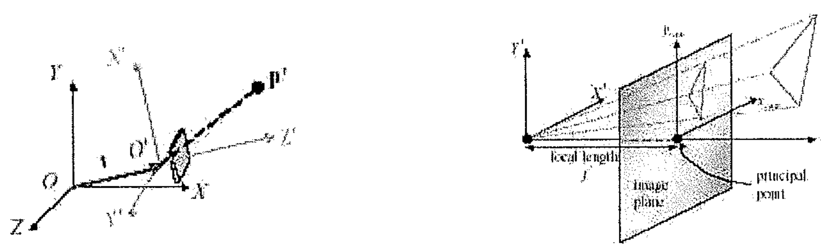
STEREO GEOMETRY

Viewing a scene simultaneously from two different positions can make inferences about three-dimensional (3D) structure of the scene, provided that corresponding points can be matched up in the images. The visual systems of humans make use of this technique. The following list shows the sequence of steps to be performed in order to recover the 3D properties from given pairs of 2D image on pavement surface:

- Camera Calibration
- Distortion Correct
- Corner Detection
- Matching Corner
- Interpolation
- 3D Reconstruct
- Profile Report

The Single Camera Model

The coordinate systems for setting up a model for a simple pinhole camera are shown in Figure 8 (9):



(a) (b)

Figure 8 Camera model (a) and (b)

***World Coordinate System (Oxyz)*:** This is the coordinate system in which the coordinates of points in the 3D world are measured. It is the common frame to communicate position and orientation information. Coordinates with respect to this frame are indicated as (X, Y, Z).

***Camera Coordinate System (O'x'y'z')*:** The retina (sensor surface) is assumed to be perpendicular to the optical axis and intersecting it at distance f (focal length) from the origin.

***Retina (Image) Coordinate System*:** The retina or image coordinate system is a frame that is used to index the image sample points (i.e. the pixels). This frame is on the retina surface and assumed to be aligned with the camera coordinate system.

A point with real world coordinates (X, Y, Z) is depicted on the retina in the point with retina coordinates (x, y). The relationship is shown in Equation 3 (9):

$$\begin{vmatrix} x \\ y \\ 1 \end{vmatrix} = \begin{vmatrix} a_u & 0 & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} R^T \vdots -T \end{vmatrix} \begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} = M_{\text{int}} M_{ext} \begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} = M \begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} \quad \text{.........................} (3)$$

Where:

a_u, a_v: scale factor to account for any uncertainty due to horizontal and vertical scanline resampling

u_0, v_0 :coordinates of center of radial lens distortion -and- the piercing point of the camera coordinate frame's Z axis with the camera's sensor plane

R : rotation angles for the transform between the world and camera coordinate systems

T : transnational components for the transform between the world and camera coordinate frames

M_{int} : the internal camera parameters which describe the optical and mechanical construction of the camera;

M_{ext} : the external camera parameters, which describe the position and orientation of the camera frame with respect to the world frame;

M : the camera projection matrix.

Geometry of Two Views

The geometry of two views from two cameras is illustrated in Figures 9 and 10. Here ***C1*** and ***C2*** are the optical centers of two cameras. ***M1*** and ***M2*** is the two normalized image planes. **P** is a 3D point, and **P1** and **P2** are the image point of **P** on the two-image plane respectively. **C1C2** are called baseline, the plane **C1C2P** epipolar plane. The intersections of the epipolar plane with the two image planes are two lines, i.e., ***l1*** and ***l2***, which are called epipolar lines. Obviously, ***l1*** is uniquely determined by the baseline and ***P2***. We call ***l1*** the epipolar line associated with ***P2***, and vice versa.

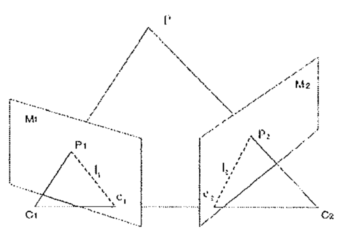


Figure 9 Epipolar geometry

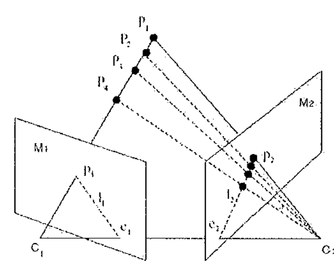


Figure 10 Point and its Epipolar Line

An important thing is that the epipolar line ***l2*** is the projection of line ***C1P*** and ***l1*** is the projection of ***C2P***, as shown in Figure 10. For any given image point ***P1*** in the left image, the ray ***C1P1*** is fixed, since every point on the line ***C1P1*** will have the same image projection. As a result, the epipolar line ***l2*** is fixed. The intersections of the baseline ***C1C2*** with the epipolar lines, i.e., **e1** and **e2**, are epipoles. Based on the principles in Figures 9 and 10, the search for matching points from pair of images may then be focused on identifying a point on a image (such as left image) and finding the epipolar line on the right image for this particular point on the left image. Therefore, computation process may be simplified in matching pixels in pairs of images.

THE ESSENTIAL MATRIX AND FUNDAMENTAL MATRIX

There are two sets of parameters that characterize optical properties and positioning of a camera: the intrinsic and extrinsic parameters (10). The intrinsic parameters relates to the camera's optical, geometrical and digital characteristics, while the extrinsic ones characterize the position of the camera in a known world reference frame. The accurate determination of these parameters is critical for the vision system for the 3D surfaces. The process of estimating these parameters is also called camera calibration.

The intrinsic parameters include the focal length f which is the distance from the focus point of the lens to the lens center, the location of image center in pixel coordinates (**Ox, Oy**), the effective pixel size in the horizontal and vertical direction (**Sx, Sy**) and the radial distortion coefficient introduced by the optics. The location and orientation of the camera in a known reference coordinate system can be modeled as a rotation

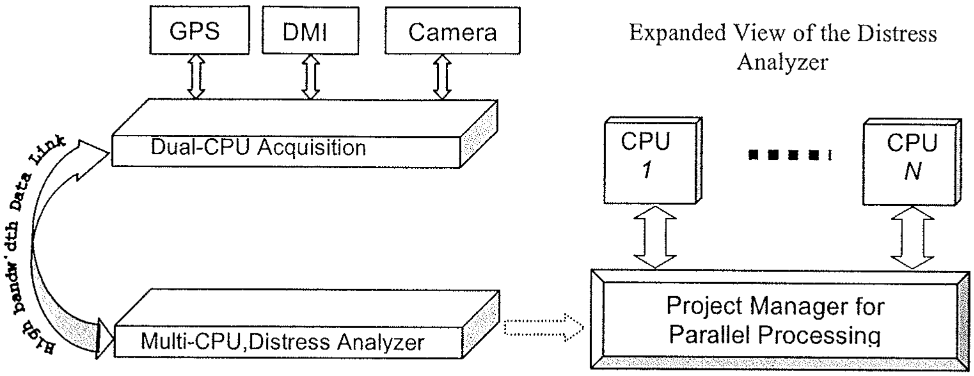


Figure 2 The Framework of the Data Acquisition and Processing with One Image Source

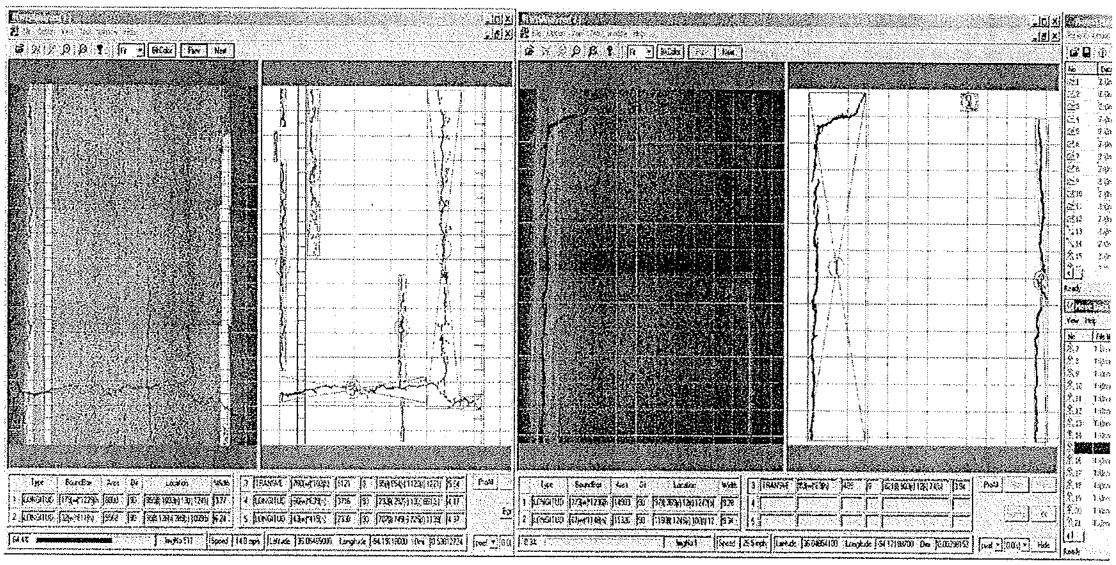


Figure 3 Dual-Processing for Cracks in a Parallel Environment with the Crack Analyzer

INVESTIGATION

ACQUISITION SYSTEM WITH DUAL-CAMERA

A lesson learned from past endeavors is that developing proprietary hardware not only requires high capital, but also limits the research from using better and cheaper hardware in the future. The fundamental approach taken by us in the past few years and used in this research is to develop algorithms and their implementations in software with general-purpose computing devices, and to use digital systems only. Data acquisition hardware components, such as digital cameras, acquisition boards, sensors, and others, have been supplied by third parties who engage in constant performance improvements.

The general technical approach of the proposed research is illustrated in Figure 4. In order to achieve 1-millimeter resolution, two digital cameras are used to cover half of a lane-width, approximately two meters. The first step is to analyze 2D images from each of the two cameras to detect and classify any cracks. The results from analyzing two image sources of the same pavement are then combined so that cracks missed by one analysis are still counted, therefore potentially achieving higher accuracy.

The pair of images on the same pavement surface is also used to establish 3D surface model, which is then used to detect other deficiencies of the pavement through straightforward geometric modeling. The geometric modeling including establishing the longitudinal and transverse profiles of the pavement, and detect abnormalities on pavement surface based on vertical variations on pavement surface. Rutting information is obtained by scanning transverse profiles with a virtual straight line. With the established longitudinal profile values, roughness is obtained through computing International Roughness Index (IRI) and Ride Number (RN) based on common accepted procedures developed at the University of Michigan. Computer code has been already realized by the authors of the research on calculating IRI and RN based on known profile values.

Potholes are detected based on the characteristics of shape and depth of defects. Other deficiencies, particularly those associated with rigid pavements, can also be found based on the 3D characteristics of the deficiencies. Furthermore, cracking information obtained based on 2D imaging can be refined based on the 3D information by either eliminating noises that were classified as cracks, or adding new cracks that exhibit depths and were missed in 2D image analysis.

To cover a complete lane surface plus a portion of shoulder area, a four-camera setup is necessary with two systems, each of which is similar to Figure 4 in design.

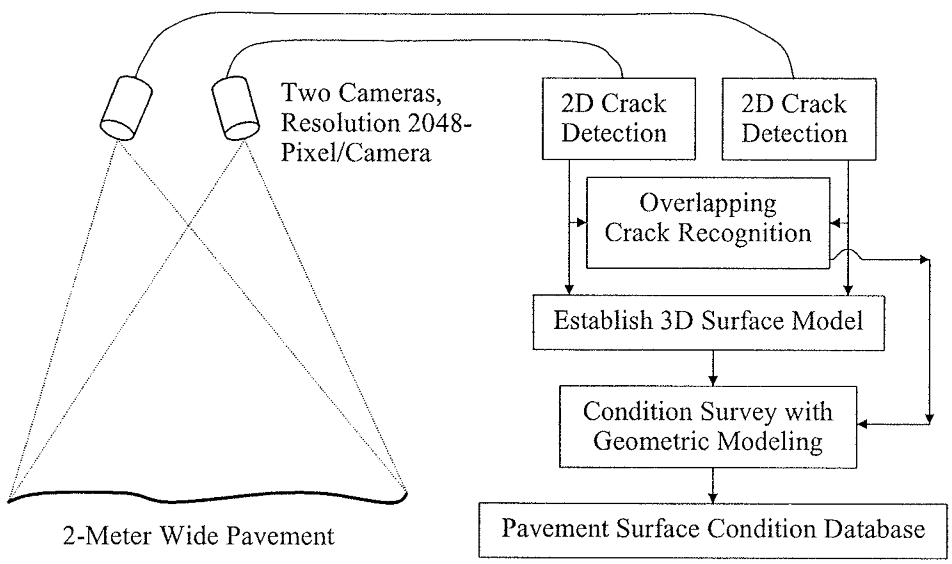


Figure 4 General Procedures for Automated Condition Survey with Stereovision

Application of stereovision has long been used in remote sensing, particularly with airborne and deep space based imaging devices. There were several obstacles for using it for pavements. First, image resolution had to be sufficiently high for pavement surveys, for instance, 1-millimeter. Digital cameras capable of this resolution at reasonable cost became only available in the past couple of years. Second, the required input bandwidth of the computing device is more than sustained 100 megabytes per second. The limitation of I/O sub-systems and microprocessor power prohibited a desktop computer from being used for acquiring and processing images from two camera sources. The proposed research is based on the fact that today's digital cameras are capable of providing needed resolution at reasonable cost, and I/O of most recent microcomputers is able to provide over 100 megabytes per second per channel. Furthermore, based on our work on real-time 2D imaging processing it is evident that technology advances already provide computing hardware platforms for a real-time system at highway speed.

Figure 5 shows the prototype of a dual-camera setup in the back of the DHDV that was used for the testing and calibration of the feasibility study.

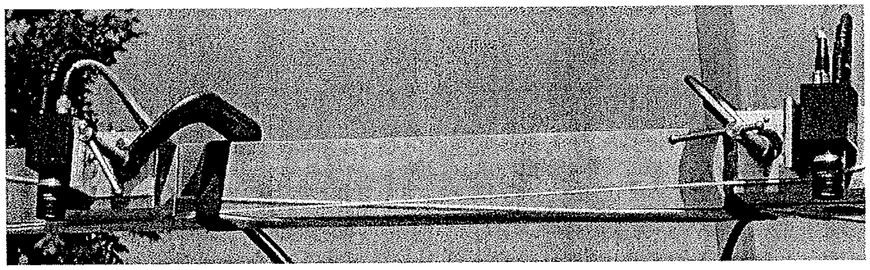


Figure 5 Detailed view of the dual camera sub-system

CONCEPTS IN STEREOVISION

The first step in finding height (H) of each pixel to establish 3D surface is to correctly identify the same point on the pavement surface in each of the pair images. The positioning difference, *d*, for the same point on the pavement surface in the two images is therefore used as input to determine H in Figure 6. The method to recognize similar image characteristics in small regions in a pair of images is called digital image matching (8). The technique of area-based matching is used in our test.

Area-based methods perform the image match by a numerical comparison of digital numbers, commonly called gray-scale, in small arrays from each image. The mathematical technique chosen for area-based matching is Normalized Cross-Correlation. In this approach, a statistical comparison is computed from digital numbers taken from same-size sub-arrays in both images. A correlation coefficient is computed by the following equation, using digital numbers from sub-arrays A and B (8):

$$c = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left[(A_{ij}-\overline{A})(B_{ij}-\overline{B})\right]}{\sqrt{\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij}-\overline{A})^2\right]\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(B_{ij}-\overline{B})^2\right]}} \quad \text{.. (1)}$$

Where:

c: the correlation coefficient,

m and n: the numbers of rows and columns in the sub-arrays,

A_{ij} and B_{ij}: digital number (gray value) from sub-arrays A and B at row i, column j,

$\overline{A}$ and $\overline{B}$: the average of all digital numbers in sub-arrays A and B,

The correlation coefficient, *c*, can range from −1 to +1, with 1 indicating a perfect match, −1 indicating a negative correlation, and zero indicating a non-match. Negative-correlation means identical images from a photographic negative and positive are compared. The image matching process is conducted by using a candidate sub-array from one image, and a search is performed for its corresponding sub-array in the second image. A search array is selected from the second image with dimensions much larger than those of the candidate sub-array. A moving window approach is then used to compare the candidate sub-array from the first image with all possible window locations within the search array from the second image. At each window location in the search array, c is computed; resulting in a correlation matrix C after the search computation is completed. The largest correlation value in C is tested against a threshold. If the test is positive, the corresponding location in the search array is considered a match.

(a) and (b) in Figure 7 show the concept of a Digital Elevation Model (DEM) model and the illustration of the positions of the sub-arrays for the pair images when the digital image matching is completed. Note the irregular positioning of sub-arrays in the second image (right image) after the processing. The differences in the positioning of sub-images from the pair images are directly used to produce height, H, in the second step.

From Figure 6, the height of any given point is determined as follows:

$$H = d \times h_2 \times \frac{Sin(\alpha_2)}{Sin(\beta) \times d_2} \qquad (2)$$

Where:

H: height of a given point on the pavement surface,
h2: the height of the second camera from a reference point on the pavement surface,
d2: the projected distance from camera 2 to the point of interest based on image 2,
α_2 and β : known angles determined by p_1, p_2, p_3, and p_4.

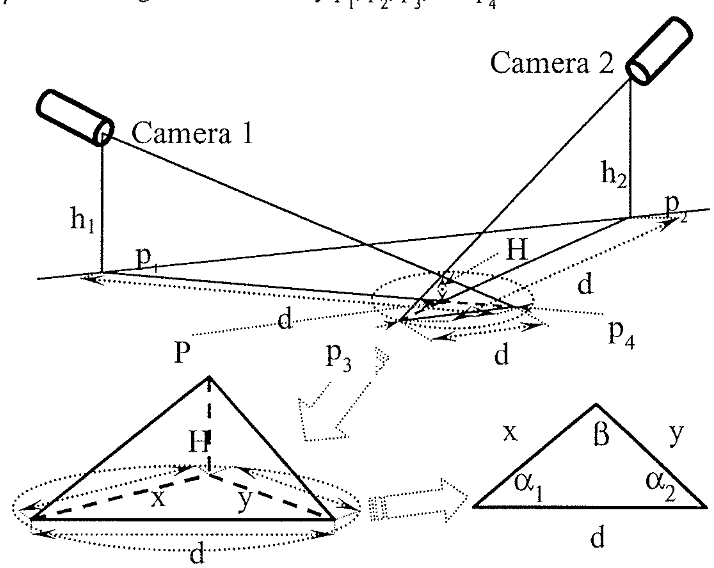


Figure 6 Using Two Cameras to Establish Height, H

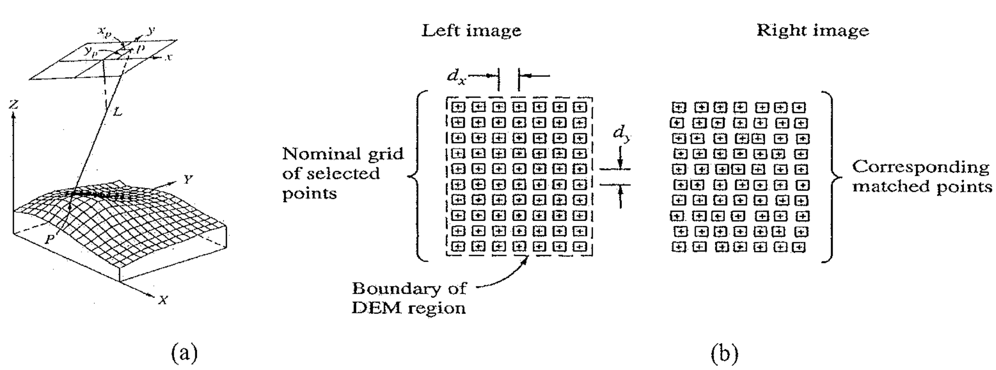


Figure 7 DEM Digital Image Matching Principles (based on (8))

followed by a translation. Therefore, the extrinsic parameters include a 3x3 rotation matrix and a 3D translation vector.

The Essential Matrix and the Fundamental Matrix are used to determine relationships between matched points in the coordinate system. The Essential Matrix is the link between the epipolar constraint and the extrinsic parameters of the system. The Fundamental Matrix is similar to the Essential Matrix. The most important difference between them is that the Essential Matrix is in terms of camera coordinates, and the Fundamental Matrix is in terms of pixel coordinates.

$$E = RS \quad (4)$$

$$F = (M_r^{-1})^T E M_l^{-1} \quad (5)$$

Where:

M_r and M_l: two matrices of intrinsic parameters. $a_u = -fK_u$

$a_v = -fK_v$: the focal lengths in horizontal and vertical pixels, respectively (f is the focal length in millimeters, K_u and K_v are the effective number of pixels per millimeter along the *u* and *v* axes)

(u_0, v_0): the coordinates of the *principal point*, given by the intersection of the optical axis with the retinal plane.

The camera position and orientation (extrinsic parameters) are encoded by the 3 ×3-rotation matrixes *R* and the translation vector *T*. The rotation and translation represent the rigid transformation that brings the camera reference coordinate system onto the world reference coordinate system.

$$T = \begin{vmatrix} Tx \\ Ty \\ Tz \end{vmatrix} \quad (6)$$

$$R = \begin{vmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{vmatrix} \quad (7)$$

$$S = \begin{vmatrix} 0 & -Tz & Ty \\ Tz & 0 & -Tx \\ -Ty & Tx & 0 \end{vmatrix} \quad (8)$$

$$M = \begin{vmatrix} a_u & 0 & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{vmatrix} \quad (9)$$

CALIBRATION

In our research, 22 points in a physical 3D model are used for calibration. The 3D model is shown in Figure 11. The input data for calibration contains 22 pairs point in the 3D model within the world coordinate system and corresponding 22 pairs in the 2D planes. The orientation of the world coordinate system was chosen so that the Z-axis is vertical and oriented upwards. A calibration model can be found in (11).

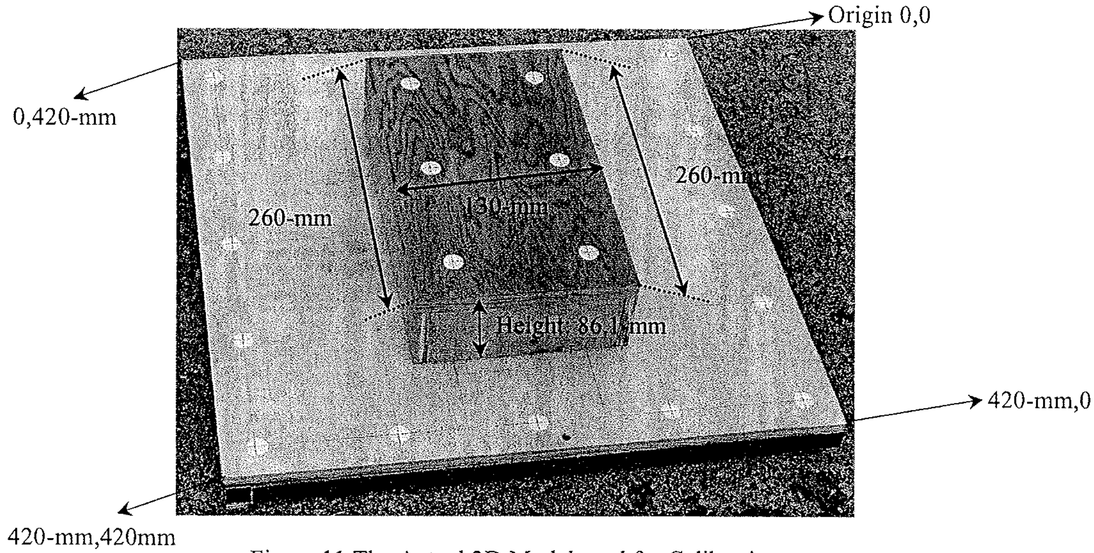


Figure 11 The Actual 3D Model used for Calibration

Figure 12 shows the left and right images from the pair cameras of the 3D model. The input data for calibration is acquired by computing the world 3D coordinates of the chosen points and manually getting the correspondent 2D points in the two images in Figure 12 through mouse clicks. The input data for calibration is shown in Table 1. The calibration routines provide the internal and external parameters as follows:

$$Z_{ci}\begin{vmatrix} u_i \\ v_i \\ 1 \end{vmatrix} = \begin{vmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{vmatrix} \begin{vmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{vmatrix} \quad \text{.........} (10)$$

$$\begin{cases} Z_{Ci}u_i = m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14} \\ Z_{Ci}v_i = m_{21}X_i + m_{22}Y_i + m_{23}Z_i + m_{24} \\ Z_{Ci} = m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34} \end{cases} \quad \text{.........} (11)$$

$$\begin{cases} X_i m_{11} + Y_i m_{12} + Z_i m_{13} + m_{14} - u_i X_i m_{31} - u_i Y_i m_{32} = u_i m_{34} \\ X_i m_{21} + Y_i m_{22} + Z_i m_{23} + m_{24} - v_i X_i m_{31} - v_i Y_i m_{32} - v_i Z_i m_{33} = v_i m_{34} \end{cases} \quad \text{.........} (12)$$

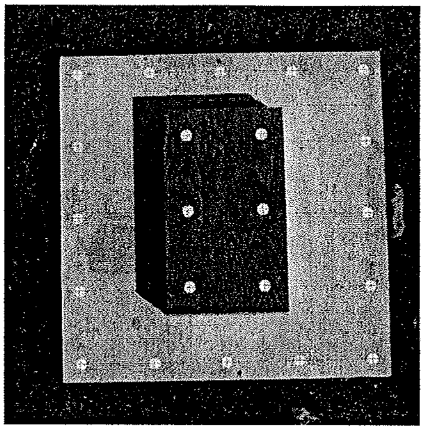
Where:

(u_i, v_i) : A coordinate in image frame.

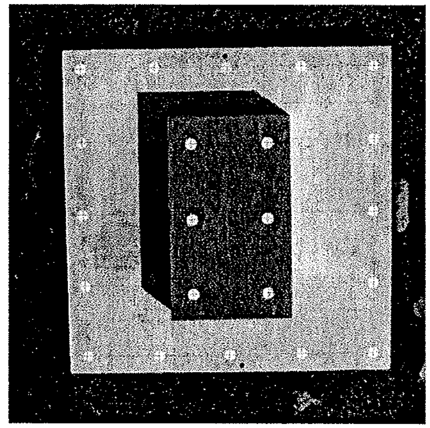
(X_i, Y_i, Z_i) : A coordinate in 3d world system.

m_{ij} : An element of the projection matrix.

Z_{ci} : An unknown constant parameter



Left image



Right image

Figure 12 Pair Images of the 3D Model for Calibration

Table 1 Input Data for Calibration, in mm

No	x_L	y_L	x_R	y_R	X	Y	Z
1	473.0	220.0	453.0	239.0	0.0	520.0	100.0
2	612.0	220.0	589.0	242.0	130.0	520.0	100.0
3	751.0	222.0	725.0	245.0	260.0	520.0	100.0
4	888.0	225.0	859.0	250.0	390.0	520.0	100.0
5	1023.0	227.0	991.0	254.0	520.0	520.0	100.0
6	472.0	360.0	451.0	375.0	0.0	390.0	100.0
7	1021.0	364.0	992.0	387.0	520.0	390.0	100.0
8	472.0	498.0	449.0	510.0	0.0	260.0	100.0
9	1018.0	499.0	993.0	520.0	520.0	260.0	100.0
10	473.0	636.0	447.0	649.0	0.0	130.0	100.0
11	1015.0	634.0	994.0	656.0	520.0	130.0	100.0
12	473.0	772.0	445.0	389.0	0.0	0.0	100.0
13	609.0	771.0	585.0	790.0	130.0	0.0	100.0
14	745.0	769.0	723.0	791.0	260.0	0.0	100.0
15	879.0	769.0	860.0	792.0	390.0	0.0	100.0
16	1011.0	767.0	994.0	792.0	520.0	0.0	100.0
17	615.0	436.0	592.0	433.0	130.0	325.0	186.0
18	760.0	436.0	736.0	435.0	260.0	325.0	186.0
19	903.0	437.0	876.0	438.0	390.0	325.0	186.0
20	616.0	580.0	592.0	576.0	130.0	195.0	186.0
21	761.0	579.0	737.0	578.0	260.0	195.0	186.0
22	900.0	579.0	877.0	580.0	390.0	195.0	186.0

In the model, there are 22 points with known coordinates in the 3D world system and known coordinates in both images. These 22 markers will determine 44 equations as below ($n=1,\ldots 22$):

- **Ordering.** Based on the work by Baker and Binford (13). If $m \leftrightarrow m'$ and $n \leftrightarrow n'$ and if *m* is *to the left* of *n* then *m*' should also be *to the left* of *n*' and vice versa. That is, the ordering of features is preserved across images.
- **Epipolar**. Given a feature point *m* in the left image, the corresponding feature point *m*' must lie on the corresponding epipolar line.
- **Relaxation**. A global matching constraint to eliminate false matches.

In this project, Harris Corner Detector (12) is used to detect corners and a modified version of the algorithm proposed by Zhang (10) is used to solve for correspondence. Figure 16 shows the matching results from a pair of images, where matched pixels in both images have the same integer values in blue.

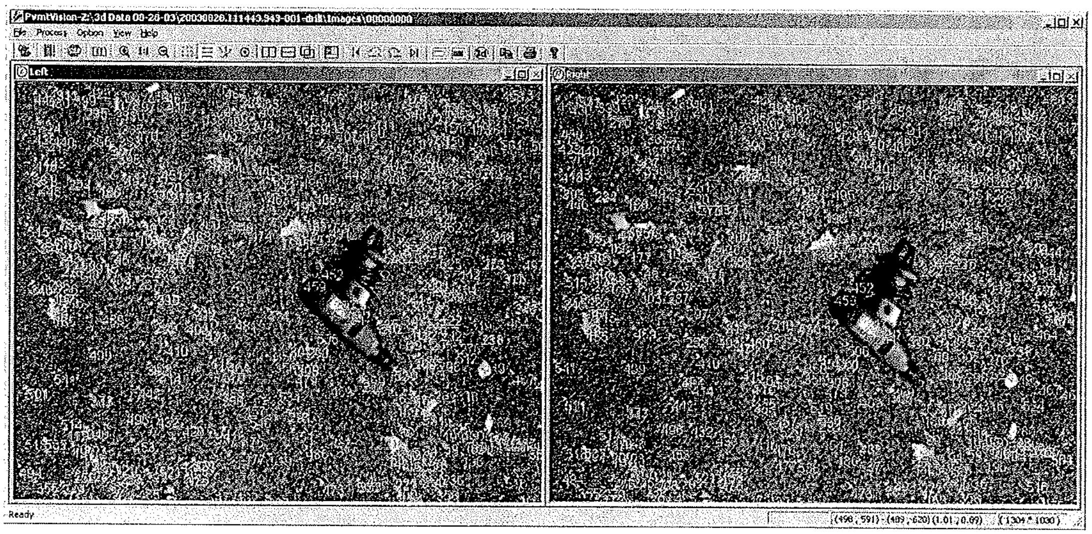


Figure 16 Matched Points in a Pair of Images

INTERPOLATION

From Figure 16, a large number of points on one image are identified with matching points on the other image. However, certain areas in the images do not contain matched points. Interpolation technique may be used to further identify pair of points in the two images.

A new technique is proposed as the interpolation method. A three-step process is used as follows.

Compute the Equation of Epipolar Line

This step determines the small area in the second image that may contain the matched point with a point on the first image. This step employs the equation of epipolar line on the second image for the point on the first image. The small area is contained in a general area where previously determined matched points exist.

Edge Detection

Edge detection is applied to the point in the first image, resulting in an edge detection feature value. Same edge detection algorithm is then applied to all points contained in the small area determined in the previous step on the second image, resulting in a number of feature values of edge detection.

Simple Corresponding Matching

The closest value of the feature values from the second image to the single value from the first image is considered a match being found. The following equation sets the criterion to determine the best match.

$$D = w_1(l_L - l_R)^2 + w_2(\theta_L - \theta_R)^2 + w_3|m_L - m_R|^2 \quad (21)$$

Where:

l : the distance
θ : the orientation value
m : the magnitude value
w_1 : the weight of the distance
w_2 : the weight of the difference in direction
w_3 : the weight of the difference in magnitude.

The smaller the D is, the more probable the point is the matching point.

3D RECONSTRUCT

After finding the matched points, 3D position of these points can be established. First the pixel coordinates are converted into image plane coordinates. This can be achieved by the following equation (11):

$$P_l = M_l^{-1} \bar{p}_l \quad (22)$$

$$P_r = M_r^{-1} \bar{P}_r \quad (23)$$

The 3D coordinates of a point are derived through triangulation. ap_l is the ray going through the center of left camera like p_l. Set $T + bR^T P_r$ is the ray going through the center of right camera. Due to the fact that these two rays may not actually intersect, the two points in the two rays with minimum distance between the two rays are considered containing the intersection point, which is the 3D point that has been pursued since the beginning.

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

$$ap_l - bR^T P_r + c(p_l \times R^T P_r) = T \quad (24)$$

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

$$P_w = aP_l + \frac{c}{2}(P_l \times R^T P_r) \quad (25)$$

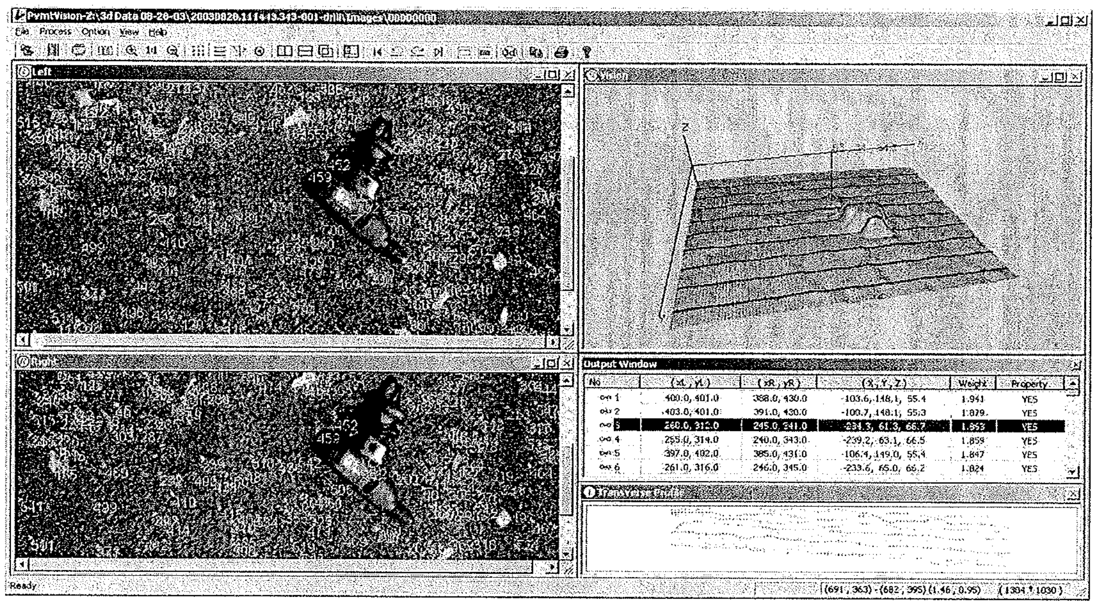


Figure 17 Pavement program (3D reconstruct)

PLANS FOR IMPLEMENTATION

Based on the feasibility study conducted in this project, stereovision techniques have the potential to be applied to pavement condition survey. From technology development point view, the implementation of this technique might not be possible several years back for two main reasons. The first reason is hardware platforms to support such implementation were not mature enough. This limitation included camera resolution and performance, computer's processing speed, computer storage, and I/O throughput. Today, these constraints of hardware are no longer there. The second reason is developing algorithms at these hardware platforms would take time and proper approaches. Our successful research on 2D imaging technology for real-time cracking survey provided much needed unique knowledge and experience for the implementation of stereovision technique.

The proposed implementation plan calls for IDEA Phase-II study. This study will implement the algorithms developed in this project and install necessary hardware to support an operational sub-system within the Digital Highway Data Vehicle (DHDV) owned by the University of Arkansas. The duration of the Phase-II study is 12-month. The proposed funding level from NCHRP for the Phase-II study is $80,000. It is anticipated that an equivalent matching fund will be provided from the University of Arkansas, which has made a commitment to NCHRP about this matching. In addition, the chief engineer of the Arkansas highway department has endorsed the implementation for Phase-II funding. The following tasks are proposed to be carried out in the Phase-II study:

1. Hardware acquisition, including cameras at 2,048-pixel resolution or higher, a precision gyro system; laser distance sensor for measuring vehicle body's displacement, and integration of the new hardware into DHDV
2. Implementation of the algorithms into DHDV described in this report
3. Development of the sub-system into a real-time device for pavement stereovision and surface reconstruction
4. Development of real-time feature extraction algorithms to establish cracking, rutting, and possibly longitudinal roughness data sets

CONCLUSION

Various efforts at Federal, state, and private levels in the past decades did not produce a fully automated system for pavement condition survey. Due to advances in off-the-shelf hardware and proper approaches in our recent research with DHDV, it is evident from this completed IDEA study that fully automated comprehensive survey of pavement surface has the potential for implementation. The general principles of measuring pavement longitudinal profile and roughness were known long before the actual application of such technology in the field. Similarly, stereovision techniques are well established long time ago. The researchers hope that application of these techniques for pavement survey at sufficient resolution and speed may well provide the industry a cost-effective solution to a long-lasting problem of relying on disparate techniques and methods to collect pavement condition data.

The long-term goal is to assemble a highway data vehicle that can collect and analyze comprehensive pavement information at highway speed, including surface condition (primarily cracking, rutting, and other surface defects), roughness, and others. The assembly cost can be as low as $250,000, including a base vehicle, conversion, all necessary vision and computing hardware, DMI and differential GPS receiver, and a high precision gyro and displacement sensor. The proposed IDEA Phase-II study will produce a fully functional system based on an existing DHDV vehicle. Technologies developed out of the Phase-II study may lead to the application of stereovision for pavement condition survey.

REFERENCE

1. Strategic Highway Research Program National Research Council (1993), Distress Identification Manual for the Long-Term Pavement Performance Project
2. Wang, K.C.P. (2000) "Design and Implementation of Automated Systems for Pavement Surface Distress Survey", ASCE Journal of Infrastructure Systems, Vol.6, No1, March, pp. 24-32
3. Guralnick S. A. and Suen, E.S. (1995). Advanced Testing of An Automated NDE System for Highway Pavement Surface Condition Assessment. Project Report of the Idea Program, Transportation Research Board, Washington, D.C.
4. Fukuhara, T., et al. (1990). "Automatic Pavement-Distress-Survey System," ASCE Journal of Transportation Engineering, Vol. 116, No. 3, May/June, pp.280-286.
5. Fundakowski, R. A., et al. (1991). *Video Image Processing for Evaluating Pavement Surface Distress*, Final Report, National Cooperative Highway Research Program 1-27, September, Washington, D.C.
6. Monti, Max. (1995), "Large-Area Laser Scanner with Holographic Detector Optics for Real-Time Recognition of Cracks in Road Surfaces" Optical Engineering, July, Vol. 34, No. 7, pp.2017-2023.
7. NCHRP, An Annual Progress Report of the NCHRP IDEA Program, 2003, N-81, pp.96, Transportation Research Board, National Research Council
8. Wolf and Dewitt (2000). Elements of Photogrammetry with Applications in GIS, 3rd Edition, McGraw Hill
9. Christy, Stéphane, 3D Localization and Modeling by successive approximations of the projective camera model, 1998
10. Zhang, Z., Deriche, R., Faugeras, O. and Luong, Q. A robust technique formatching two uncalibrated images through the recovery of the unknown epipolar geometry *Artificial Intelligence Journal, 78:87-119, October 1995*
11. Mohr, R. and Triggs B. Tutorial on projective geometry given at International Symposium of Photogrammetry and Remote Sensing, Vienna, 1996
12. Harris, C. and Stephens, M. "A combined corner and edge detector," in Proc. Alvey Vision Conference, pp. 147-151, 1988
13. Baker, H. and Binford, T. ``Depth from Edge and Intensity Based Stereo", Int. Joint Conf. on Artificial Intelligence, Vancouver, Canada, pp. 631-636,1981.