

IDEA

**Innovations Deserving
Exploratory Analysis Programs**

NCHRP IDEA Program

**An Inexpensive Vision-Based Approach for the Autonomous
Detection, Localization, and Quantification of Pavement
Defects**

Final Report for
NCHRP IDEA Project 169

Prepared by:
Burcin Becerik-Gerber, Sami F. Masri
University at Southern California

Mohammad R. Jahanshahi
Purdue University

August 2015

Innovations Deserving Exploratory Analysis (IDEA) Programs Managed by the Transportation Research Board

This IDEA project was funded by the NCHRP IDEA Program.

The TRB currently manages the following three IDEA programs:

- The NCHRP IDEA Program, which focuses on advances in the design, construction, and maintenance of highway systems, is funded by American Association of State Highway and Transportation Officials (AASHTO) as part of the National Cooperative Highway Research Program (NCHRP).
- The Safety IDEA Program currently focuses on innovative approaches for improving railroad safety or performance. The program is currently funded by the Federal Railroad Administration (FRA). The program was previously jointly funded by the Federal Motor Carrier Safety Administration (FMCSA) and the FRA.
- The Transit IDEA Program, which supports development and testing of innovative concepts and methods for advancing transit practice, is funded by the Federal Transit Administration (FTA) as part of the Transit Cooperative Research Program (TCRP).

Management of the three IDEA programs is coordinated to promote the development and testing of innovative concepts, methods, and technologies.

For information on the IDEA programs, check the IDEA website (www.trb.org/idea). For questions, contact the IDEA programs office by telephone at (202) 334-3310.

IDEA Programs
Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

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 National Academies of Sciences, Engineering, and Medicine. 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 Academies of Sciences, Engineering, and Medicine; or the sponsors of the IDEA Programs.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; 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.

An Inexpensive Vision-Based Approach for the Autonomous Detection, Localization, and Quantification of Pavement Defects

**IDEA Program Final Report
NCHRP IDEA Project 169**

Prepared for the IDEA Program
Transportation Research Board
The National Academies

Submitted by

Burcin Becerik-Gerber, DDes (PI)
Sonny Astani Department of Civil and Environmental Engineering
University of Southern California
3620 S. Vermont Ave. KAP 224C
Los Angeles, CA 90089
becerik@usc.edu
(213) 740-4383

Sami F. Masri, PhD (Co-PI)
Sonny Astani Department of Civil and Environmental Engineering
University of Southern California
3620 S. Vermont Ave. KAP 206A
Los Angeles, CA 90089
masri@usc.edu
(213) 740-0602

Mohammad R. Jahanshahi, PhD (Co-PI)
Lyles School of Civil Engineering
Purdue University
550 Stadium Mall Drive HAMP 4116
West Lafayette, IN 47907-2051
jahansha@purdue.edu
(765) 494-2217

August 2015

NCHRP IDEA PROGRAM COMMITTEE

CHAIR

DUANE BRAUTIGAM
Consultant

MEMBERS

CAMILLE CRICHTON-SUMNERS
New Jersey DOT

AGELIKI ELEFTERIADOU
University of Florida

ANNE ELLIS
Arizona DOT

ALLISON HARDT
Maryland State Highway Administration

JOE HORTON
California DOT

MAGDY MIKHAIL
Texas DOT

TOMMY NANTUNG
Indiana DOT

MARTIN PIETRUCHA
Pennsylvania State University

VALERIE SHUMAN
Shuman Consulting Group LLC

L.DAVID SUITS
North American Geosynthetics Society

JOYCE TAYLOR
Maine DOT

FHWA LIAISON

DAVID KUEHN
Federal Highway Administration

TRB LIAISON

RICHARD CUNARD
Transportation Research Board

COOPERATIVE RESEARCH PROGRAM STAFF

STEPHEN PARKER
Senior Program Officer

IDEA PROGRAMS STAFF

STEPHEN R. GODWIN
Director for Studies and Special Programs

JON M. WILLIAMS
Program Director, IDEA and Synthesis Studies

INAM JAWED
Senior Program Officer

DEMISHA WILLIAMS
Senior Program Assistant

EXPERT REVIEW PANEL

TOM PYLE, *California DOT*

ARAMIS LOPEZ, *FHWA*

JEROME DALEIDEN, *Fugro Consultants, Inc.*

TIMOTHY MILLER, *Fugro Roadware, Inc.*

MICHEL ROBSON, *Fugro Roadware, Inc.*

Table of Contents

1. Executive Summary	1
2. IDEA Product.....	1
3. Concept and Innovation	2
4. Investigation.....	3
4.1 Design and Development.....	3
4.1.1 Hardware Development	4
4.1.2 Software Development.....	9
4.2 Experimental Results and Field Tests.....	17
4.2.1 Data Acquisition and Synchronization of Heterogeneous Sensors	17
4.2.2 Pavement Profile Measurement	17
4.2.3 Pavement Data Collection on California State Freeway and on City of Pasadena Streets	20
4.2.4 Pavement Images at Various Speeds.....	22
4.2.5 Localization of the Pavement Defects.....	29
4.2.6 Rectification of Rolling Shutter Distortion	32
4.2.7 Stroboscopic Solution for Motion Blur Problem	34
4.2.8 A Hybrid Pavement Crack Detection Algorithm	35
5. Plans for Implementation.....	37
6. Conclusions.....	37
6.1 Recommendations for Future Work	38
Acknowledgment	39
References.....	39

1. Executive Summary

This project involves developing a cost-effective, vision-based sensor system for autonomous data acquisition of road surface conditions and for autonomous detection, localization, and quantification of pavement defects. The proposed sensor system provides a large array of information such as vision, depth, inertial, and positioning information, which can be efficiently used for detecting and localizing defects on pavements, on an offline basis, at a lower cost. The sensor system can be used periodically to gather data that can be maintained in a main server. The proposed system provides useful pavement defect information at a fraction of the cost of an advanced sensor system.

The cost, ease of use, and computational aspects of this project offer unique benefits for the data acquisition and monitoring of pavement surfaces. The sensor system can be mounted on a variety of public and private vehicles. This provides continuous and long-term information that can be used for data analytics. Furthermore, increased data gathering can provide an improved understanding of the formation and evolution of defects in time. This can also lead to earlier measures to curtail the degradation of the pavement surface. The proposed technology is ideal for small cities and counties.

In this report, the results of the project are described in three stages. Work in stage I focused on the results of the design and development of the hardware and software modules. The hardware module included the selection of suitable depth cameras, accelerometers, and an Inertial Navigation System (INS). The software module consisted of the data and hardware synchronization schema, and color and depth camera calibration. In stage II, field tests and data analysis were performed. In the third part of this report, the results of the pavement crack segmentation algorithm are discussed.

2. IDEA Product

A relatively inexpensive sensor system for autonomous data acquisition of road surface conditions was designed and assembled, and a related crack segmentation algorithm was developed. There are a number of issues that limit the accuracy, reliability, and affordability of current pavement defect-detection solutions, such as environmental conditions, high equipment costs, and laborious manual or semi-manual quantification of defects. There is a need for detecting, locating, and quantifying pavement defects in an efficient, reliable and economic way at normal highway or city traffic speed. The proposed system is capable of recording the data at normal highway speed of 96.56 km/h (60 mph) using off-the-shelf heterogeneous sensors. The objective of this IDEA project is to develop a relatively inexpensive pavement data-collection tool and condition-assessment solution that can not only detect all critical types of defects (e.g., potholes and cracks), but also autonomously and reliably quantify and classify them

based on their location and severity. Implementation of the proposed approach will enable federal, state, and local agencies to:

1. Select and prioritize the operations necessary to repair and maintain the reliability of a system within acceptable limits over its desired lifespan;
2. Understand the contributing factors to pavement deterioration; and
3. Make better informed management strategies through increased data coverage and accessibility.

The solution will autonomously detect, localize, quantify, and classify all types of pavement defects in an inexpensive and reliable fashion. The outcomes of this research will benefit the industry in the following ways:

1. **Increased Coverage and Accessibility:** The vehicle-mounted hardware system will enable a relatively inexpensive data collection method;
2. **Prioritization and Optimization of Maintenance Work:** The assessment method will detect, quantify, and classify defects autonomously, providing reliable information for maintenance of critical defects;
3. **Better Informed Management Strategies:** The available data will inform optimization algorithms and pavement performance prediction models; and
4. **Understanding the Evolution of Pavement Defects:** The evolution of pavement defects will be monitored to understand the contributing factors (e.g., materials, traffic speed, traffic volume, weather, and environmental conditions).

3. Concept and Innovation

In the conventional pavement condition assessment process, trained personnel survey the condition of the road surface, detect the pavement defects, measure the severity, and finally classify defects. This manual assessment process is very time consuming and laborious; in addition, it poses practical safety threats to the personnel involved. Furthermore, since trained raters carry out the assessment, due to the subjectivity of the detection and classification processes, different raters could interpret the defects differently, and the results may not be repeatable. In automated pavement condition assessment, the process of data acquisition is carried out using vehicles equipped with various types of sensors such as cameras, laser sensors, infrared lighting, ultrasonic sensors and so on. Some of the widely used examples of these vehicles are Automated Road Analyzer (ARAN) and Digital Highway Data Vehicle (DHDV), which are prohibitively expensive (approximately \$1 million). These vehicles use image processing solutions for automated defect-detection. WiseCrax and Automated Distress Analyzer (ADA) are two examples of

solutions used in the industry for analysis of pavement data, captured by ARAN and DHDV, respectively. There are a number of issues that limit the accuracy, reliability, and affordability of these solutions. First, environmental conditions, such as light/shadow condition, different background textures, and non-crack patterns, can compromise the assessment outcomes. Second, the image processing solutions are limited in the detection of various types of defects, especially 3D defects such as potholes, and raveling, which are disruptive to the quality of the pavement surfaces. Third, currently there is no solution that can autonomously quantify the severity of various pavement defects. Although automated approaches, especially vision-based methods, have been the subject of various research studies, and have improved substantially in the last decade, manual survey methods are still the predominant approaches employed by the industry as the equipment cost is a barrier in adopting automated methods.

In an era when agencies are continually asked to do “more with less” in terms of both people and funding resources, it is critical to have new data collection tools that are effective but inexpensive, and methods that are reliable and robust to assist agencies in making the most efficient decisions. The project aims to achieve the following goals: (1) data collection with an inexpensive vehicle-mounted hardware system and pavement defect-detection at normal highway traffic speed; and (2) pavement data processing for quantification and localization of defects.

The concept of the relatively inexpensive sensor system for autonomous data acquisition of road surface conditions is similar to the high end data acquisition systems, but it is comprised of inexpensive sensors of an order less than \$15,000. Furthermore, these lower end sensors are effective enough to acquire the data at the desired rate. Primarily, the setup’s main focus is on the color and depth information capturing cameras. These cameras are connected and synchronized with an array of heterogeneous sensors such as accelerometers, Inertial Measurement Unit (IMU), global positioning system (GPS), etc.

4. Investigation

The work completed under this project involved the design and development of the hardware system, followed by the development of the software architecture for synchronization and online data recording. Lastly, field tests were conducted, at various speeds, and more than 300 gigabytes of data were recorded and analyzed.

4.1 Design and Development

In the design and development phase, a major effort was focused into selecting commercially available off-the-shelf cameras for acquiring color and depth information. In addition to this, a considerable amount of effort was put into the software design of the synchronization module. The detailed explanation of the

hardware and software designs is provided in the following subsections.

4.1.1 Hardware Development

A schematic diagram illustrating an overview of the vision-based approach for pavement inspection is shown in Figure 1. Based on the pavement field of view, there are eight RGBD (color + depth) cameras firmly attached to aluminum holders. Four high-frequency accelerometers are connected to a robust eight analog input channels (maximum 48,000 samples/second). Additionally, a GPS system is included so that the accurate location of pavement defects can be visualized on a map. Finally, all these hardware components are connected to one or more computers synchronized with a master clock to minimize the latency during data acquisition.

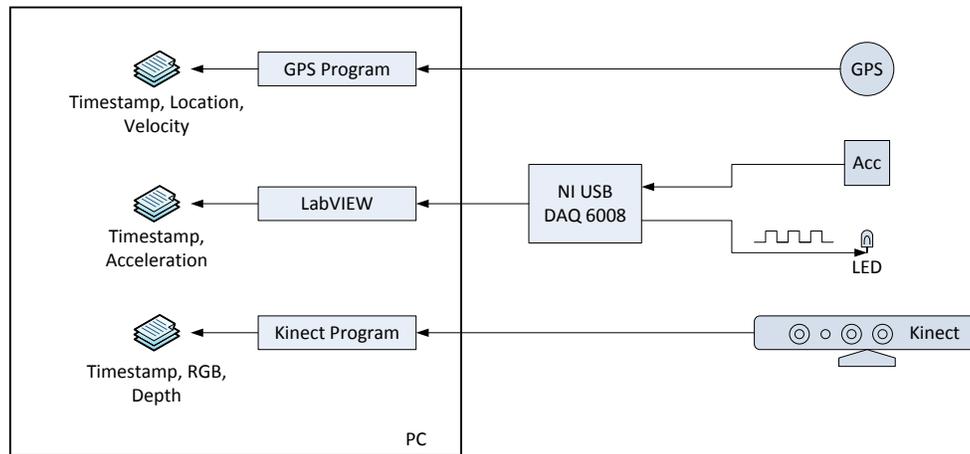


Figure 1 Overview of the 3D scanning system for pavement inspection.

4.1.1.1 Depth Sensors Selection

To select the color and depth camera, several depth sensors were evaluated for their pros and cons in capturing reliable data for detection and quantification of pavement defects, such as potholes, patches, raveling, etc.



Figure 2 Evaluated depth sensors.

**TABLE 1
COMPARISON OF THE SPECIFICATIONS OF INEXPENSIVE DEPTH SENSORS**

Specifications	Microsoft Kinect	SoftKinetic DS311	SoftKinectic DS325	SwissRanger SR4000
Range (short)	N/A	N/A	15 cm–1.5 m	N/A
Range (long)	0.8–4 m	1.5–4.5 m	N/A	0.8–8 m
Resolution (depth)	QVGA (640 x 480)	QVGA (320 x 240)	QVGA (160 x 120)	176 x 144
Field of View (H x V x D)	57° x 43° x N/A	57.3° x 42° x 73.8°	74° x 58° x 87°	43° x 34° x N/A
Technology (depth sensor)	Light coding	Depth sense	CAPD ToF	Time of Flight (ToF)
Frame Rate (depth sensor)	30	25–60	25–60	50
Resolution (RGB)	1280 x 960/640 x 480	640 x 480	1280 x 720 (HD)	N/A
Field of View (RGB)	57.3° x 42° x N/A	50° x 40° x 60°	63.2° x 49.3° x 75.2°	N/A
Frame Rate (RGB)	30	<25	<25	N/A
Power/Data Connection	USB 2.0 (1)	USB 2.0 (1)	USB 2.0 (1)	Lumberg M8 Male 3-pin
Size (W x H x D)	27.94 x 7.62 x 7.62 cm	24 x 5.8 x 4 cm	10.5 x 3.1 x 2.7 cm	6.5 x 6.5 x 6.8 cm
Price	\$99	\$299	\$249	\$4,295

To this end, commercially available structured light and time-of-flight depth sensors, including the Microsoft Kinect, SoftKinetic DS311, DS325, and SwissRanger SR4000 were evaluated for various outdoor conditions (Figure 2). Table 1 compares the specifications of the various commercially available off-the-shelf depth sensors used in this project. Kinect was used for the range sensing application because it belongs to a structured light sensing category which is roughly immutable for sunlight. In structured light, a known pattern of light is projected on to the scene. Thereby, the distance is measured when the light deforms by striking surfaces.

4.1.1.2 Accelerometers and Data Acquisition Modules

Accelerometers play an important role in aligning the sensor platform, calculating the vertical motion of a moving body (e.g., car), computing the three-dimensional translation and twist of the vehicle chassis, which in turn provides three-dimensional kinematics of the cameras. Figure 3(a) shows a typical piezoelectric accelerometer (gas damping) used in this study. Input base voltage of < 2.5 v and excitation current of < 5 mA are required to trigger the measuring process. Each accelerometer has a maximum frequency response range of 0–1500 Hz.

To obtain digital data through these accelerometers, a National Instruments DAQ module was connected to a host computer, and this module was controlled by a software interface (LabVIEW™). Figure 3(b)

shows the National Instruments' hardware module for the acquisition of the analog input. It should be noted that there were eight differential analog input channels, and in total 32 channels were embedded in the DAQ module including digital input and outputs.



Figure 3 (a) High-frequency accelerometer, (b) National Instruments USB DAQ module.

4.1.1.3 Global Positioning System (GPS)

A universal serial bus (USB) supported GPS device was added to the data acquisition system. The GPS program was triggered by the LabVIEW™ system to collect data. The time-stamping of the GPS data was consistent with the global host computer. Position, velocity, and time data were recorded at every second. The resulting data were in compliance with the National Marine Electronics Association (NMEA) standards.

4.1.1.4 Multiple Kinect Array

Using a single Kinect limits the maximum traveling speed of a data collection vehicle to 35 mph. In addition, a single Kinect does not cover the whole width of a traffic lane. To address these issues, an array of four Kinects that almost doubles the field of view (FOV) and the data collection speed of a single Kinect were built. Figure 4 shows the Kinect array separated by 700 mm x 500 mm. Each Kinect is placed at the corner of the 700 mm x 500 mm rectangular aluminum channel.

4.1.1.5 Development of Inexpensive Vision-Based Data Collection System

The relatively inexpensive vision-based data collection system for pavement inspection includes four Microsoft Kinects, which capture color and depth images. In theory, the arrangement of four Kinects could obtain each image covering about a 1.52 x 1.21 m area, and a video without gaps between two consecutive frames with vehicle speed up to approximately 96.56 km/h (60 mph).

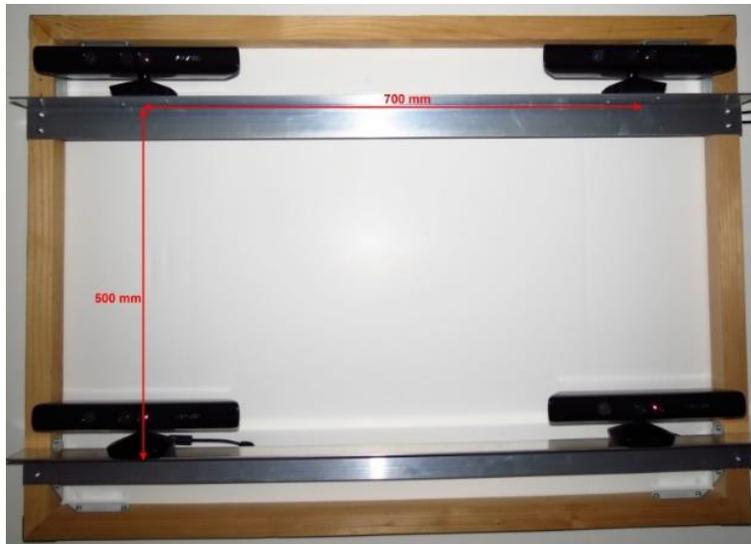


Figure 4 A typical multiple Kinect array.

The Kinect sensors are connected to a desktop computer, which provides four dedicated USB controllers for each Kinect to maintain a suitable data acquisition bandwidth. A high-performance computer consisted of an AMD FX-8350 4.0Hz eight-core processor and 16GB memory. A 250GB solid state hard drive (SSD) is used to manage and archive the large amount of captured data. A touch screen monitor is used to control the data acquisition modules.

The data acquisition system is equipped with a GPS to record spatial information, and six accelerometers for Kinect array attitude estimation and pavement roughness evaluation. These sensors are synchronized with Kinects using an external LED light, which is mounted in front of one of the Kinect's RGB cameras. A USB data acquisition board sends a pulse signal to flash the LED when acquiring data from the sensors.

The above data acquisition system requires mobile power supplies to perform recording tasks when the test vehicle is in motion. The computer needs about 200 watts of power to run the system and about 300 watts to start up; therefore, it is connected to a 12-volt automotive battery through an 800-watt DC-AC converter. Other devices such as Kinects and accelerometers consume power from a 12-volt vehicle's outlet through a 400-watt DC-AC converter.

Based on some initial tests, this power scheme could provide electricity for the data acquisition system to operate for one and half hours. Figure 5 shows different components of the proposed data acquisition system.

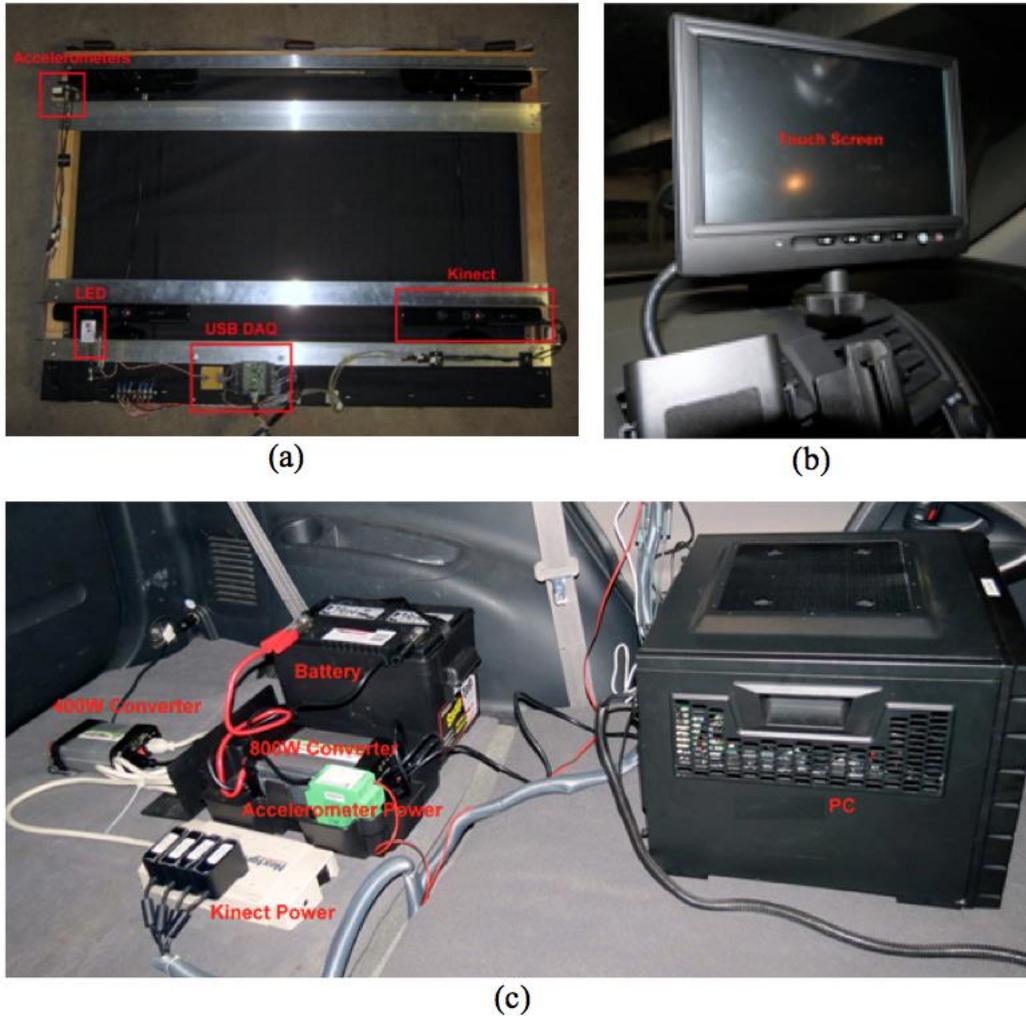


Figure 5 Components of the data acquisition system.

4.1.1.6 Design of Sun Shades

Since an off-the-shelf Kinect range sensor is developed primarily for indoor gaming purposes, it is easily interfered with by direct sunlight when operated outdoors. Based on the outdoor tests, a Kinect can obtain good depth images under indirect sunlight (i.e., in shadow). A top-cover sun shade, which minimizes pavement exposure to sunlight, was designed and tested [see Figure 6(a)]. However, depending on the location of the sun in the sky, the shadow area shifts from the area under scan by the Kinect sensors. Figure 6(a) shows the sensor systems when mounted on a top-cover sun shade frame.

The full-cover sun shade shown in Figure 6(b) was made to create an indoor environment for Kinect sensors while operating in outdoor environments. It is a bottom-open box-like structure to carry four

Kinects inside the box. When the box-like sun shade is mounted on a vehicle, there is a space between the structure and ground to prevent the structure from directly bumping into the road. The space is covered by flexible rubber sheets to block sunlight that saturates the scanning area. LED lights were installed inside the box to illuminate the area under scan, so that the color cameras could take good pictures. However, it was observed during the field tests that there were significant motion blur effects on the color images when the vehicle speed exceeded 24.14 km/h (15 mph). The default settings of Kinect enable automatic exposure control for the RGB camera. It extends exposure time automatically when luminance is not enough and this phenomena causes motion blur. Automatic exposure adjustment also disturbs the color image recording when luminance has rapid change (e.g., moving from a shaded area to a bright area). Thus, the automatic exposure feature of Kinect sensors should be disabled for the type of application under discussion.



**Figure 6 Different types of shades used in this project. (a) Full-cover sun shade
(b) Top-cover sun shade.**

4.1.2 Software Development

Software development for the proposed system was comprised of the data acquisition modules for color and depth camera (Microsoft Kinect), accelerometers, Garmin GPS, and VectorNav VN-200 INS (see Figure 7 top portion). In addition, the synchronization module included the triggering of the events at the specified time, logging the time stamps and transferring the output analog signal to control the light emitting diodes (LEDs) blinks. Furthermore, calibration of the color and depth cameras was necessary to remove the spatial distortions in the images, and to obtain a good quality image registration. Additionally,

the color and depth camera combined with accelerometers and linear variable differential transformer (LVDT) was used to calibrate the displacement at a known point on a checkerboard pattern in a controlled laboratory setup. Next, aligning of the four color and depth images were carried out by using the open source software called CloudCompare. Lastly, the pavement crack detection using a hybrid algorithm based on anisotropic diffusion filtering and Eigen analysis of Hessian matrix was performed on few original images, using color camera information and other publically available images.

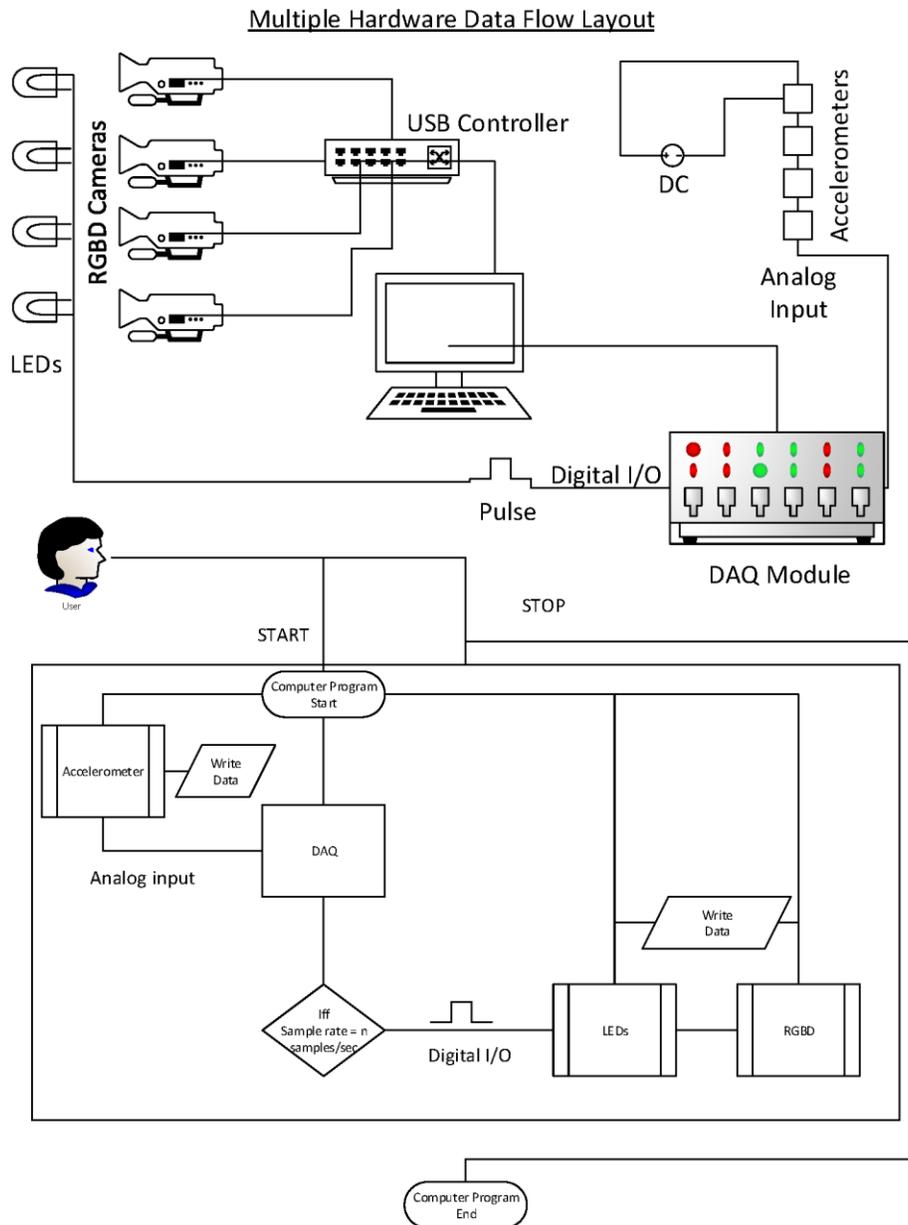


Figure 7 Data read/write flow chart or layout for the DAQ, RGBD cameras, accelerometers output, and pulse train for LEDs.

4.1.2.1 Kinect Data Acquisition

A Kinect data acquisition program was developed using C/C++ and OpenNI library to record an RGB image stream and a depth data stream simultaneously, with a resolution of 640 x 480 pixels with 30 frames per second. When the software is running, it writes an RGB image stream and a depth stream produced by a Kinect sensor to an ONI video file. The ONI file can be read later to extract a series of RGB image frames and depth data frames for post processing. As there is no timing information recorded in the ONI file, the acquisition program also saves frame numbering and timestamp information to a text file.

The Kinect cameras have limited field of views (see Table 1). At a range of 800 mm, the RGB camera only covers an area of about 961 mm x 722 mm, and the IR camera only covers an area of about 869 mm x 672 mm. To extend the viewing region for pavement condition assessment, adding more Kinect sensors is a possible approach. A program was developed to implement multiple Kinect data acquisition. It was found that a single Kinect required at least 50% of USB bandwidth; therefore, it was decided to dedicate an individual USB controller to each Kinect. The other issue was interference caused by multiple Kinects' infrared emissions on the same area.

4.1.2.2 Data Acquisition from Other Sensors

Sensors such as accelerometers, GPS, and INS used a third-party software for the data acquisition. The accelerometer(s) is coupled to the National Instruments Data Acquisition (DAQ) module such as NI 6008 USB. The data were sampled at the rate of 1 kHz using National Instruments LabVIEW™ software environment, whereas the GPS used built-in application programming interface (API) to record the data at 1 Hz. Lastly, the INS used VectorNav's custom-designed Graphical User Interface (GUI) and libraries to acquire the data at 200–800 Hz.

4.1.2.3 Data and Hardware Synchronization

Synchronization of independent processes or sensors is necessary to achieve accurate readings from external hardware system. To synchronize heterogeneous sensor systems, two methods, naive and heuristic synchronizations, were implemented. In the naive synchronization approach, separate processes were triggered by an external application and synchronization was solely dependent on the host computer clock cycles. A clock system of a host computer was used as a reference for each device. For example, assume that there are three external hardware devices A, B, and C, and all of them are triggered at time t .

If device A starts at time $t + x$, and devices B and C start at time $t + y$, their relative time difference from the referential time (host time) would provide an accurate time measurement, even if there is latency or clock drift within each device. An external virtual interface such as LabVIEW™ is used to trigger independent processes and the latency between the processes was observed to be around 250 milliseconds (ms).

In the heuristic synchronization approach, the synchronization process is based on the cue marks of the external hardware such as laser diodes or LEDs or hyper flashlights (xenon), controlled by a master program in a host computer. A heuristic method for synchronization of the hardware was successfully implemented in a laboratory environment. The heuristic synchronization consisted of a data acquisition schema and syncing of the external devices, such as the LED–5 mm, accelerometers, RGBD cameras, and also LVDT for measuring the true displacement. Figure 7 (bottom portion) shows a typical data flow in the Data Acquisition (DAQ) and Synchronization (SYNC) modules. Data acquisition starts at an arbitrary time. At every n th sample, a digital pulse is triggered. To achieve a satisfactory resolution of millisecond, a sampling rate of 1 kHz was adopted. Boolean states of color and IR LEDs, analog inputs/outputs, and their corresponding samples were recorded. It is worthwhile to note that the color and depth sensors were triggered at the same time, while any differences in their respective frames were corrected and aligned for sub-frame accuracy.

During the data processing, the first saturated color and depth frames were detected to define the starting point of the synchronization. Subsequent frames were corrected for their timestamps. In other words, the frames were aligned in time with respect to the first frame and the first sample of the DAQ module. The same concept was applied for multiple cameras, and the relative difference between two or more cameras were corrected. Another strategy is to use the master and slave method for the cameras (RGBD). Consider a camera that is master and all other cameras are slaves. Each timestamps of the frames of the slave camera are corrected with respect to the nearest frame of the master camera. This process is repeated until the last frame of the complete set.

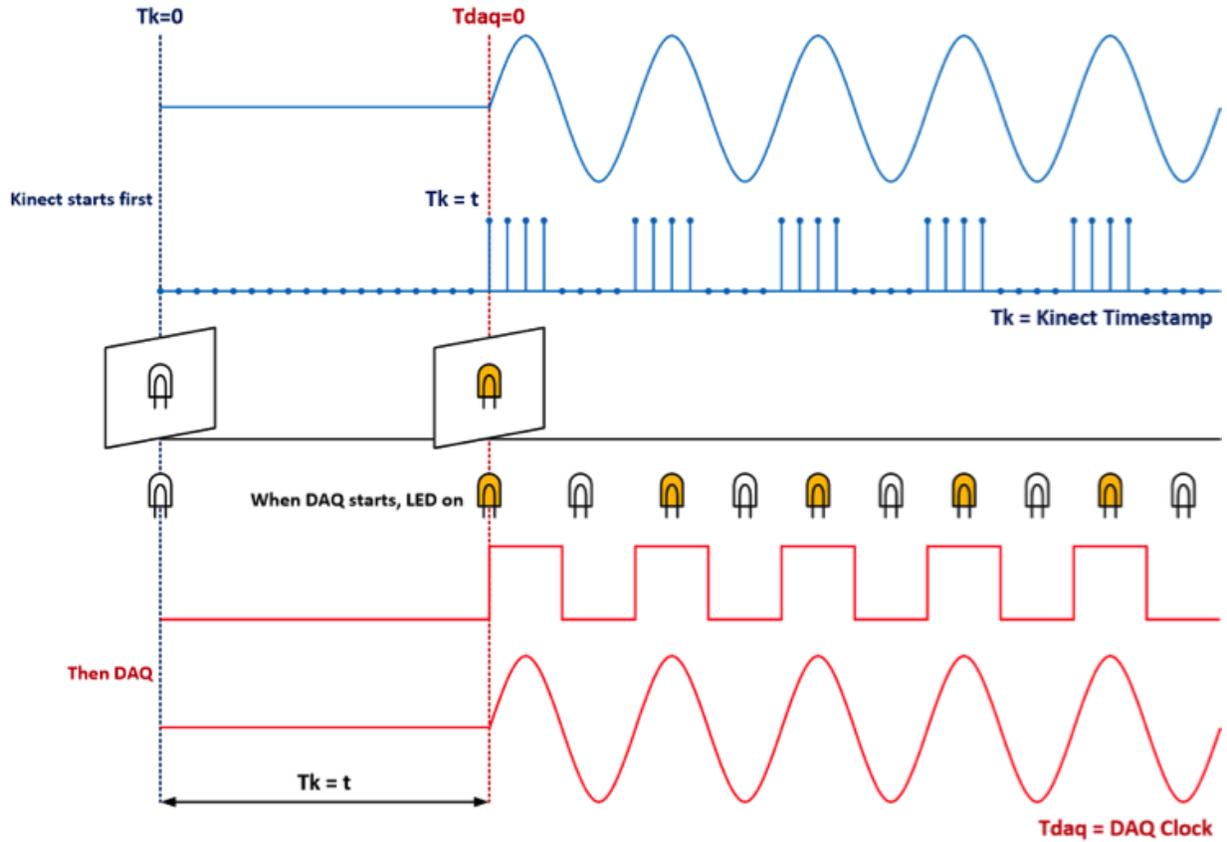


Figure 8 Sync mechanism of hardware with LEDs.

Figure 8 shows a simple and inexpensive synchronization procedure for one or more RGBD cameras. To minimize the ambiguity that arises due to software-based timestamps, the Kinect unit was triggered first and, after a delay of n milliseconds, the DAQ module with accurately synchronized LED digital pulses was triggered. After n milliseconds, the i th color and depth frames were saturated with an LED and IR LED cues. This process continued until the end of the execution of the program.

4.1.2.4 Kinect Calibration

Based on the experimental results, the Microsoft Kinect, a structured light depth sensor, performed robustly compared to the others in capturing the data needed for pavement condition assessment. In addition, the Kinect's RGB and depth cameras needed to be calibrated to superpose color and depth data to generate colored point clouds, which was then used to perform point clouds registration.

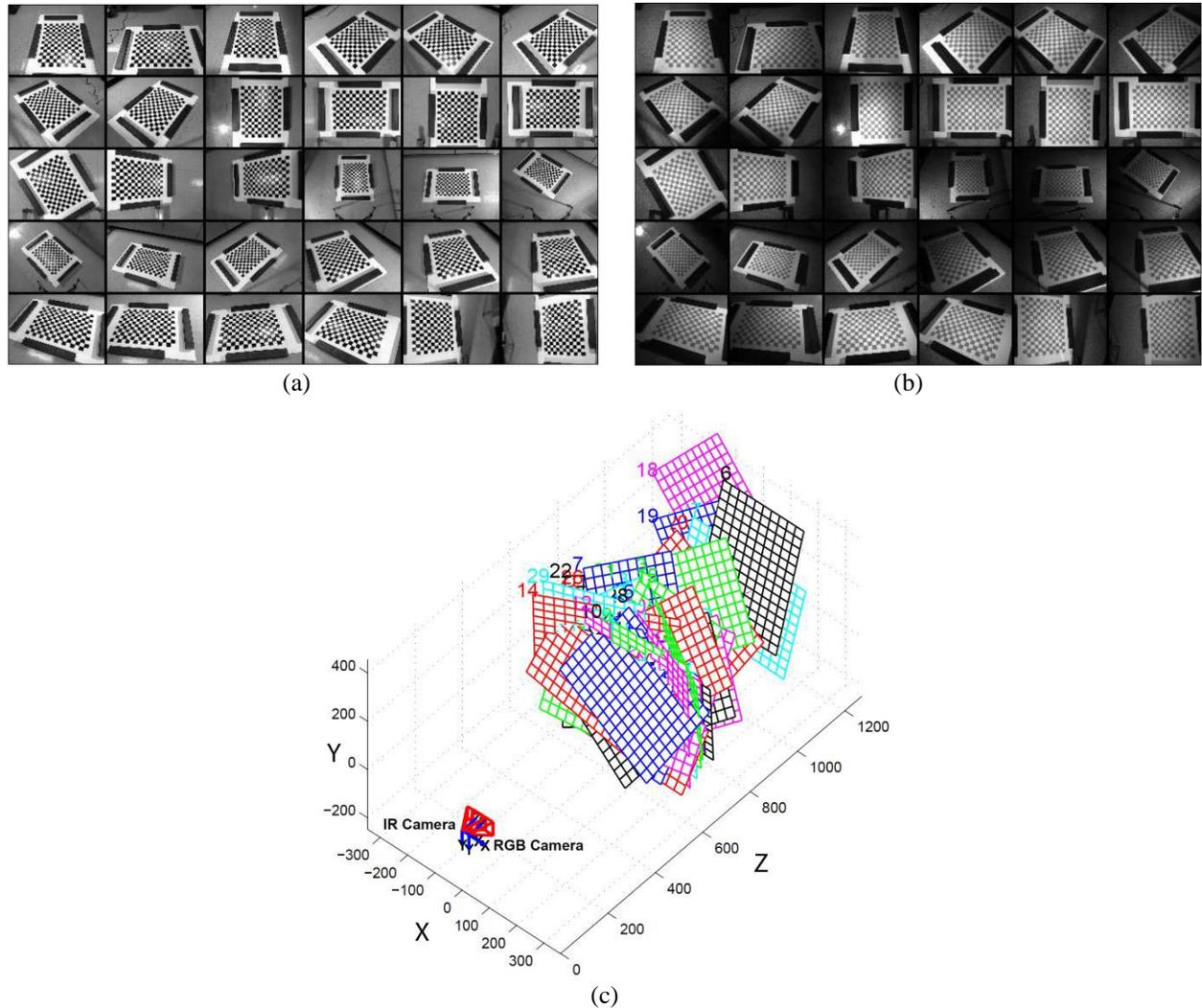


Figure 9 (a) Calibration images for the RGB camera, (b) calibration images for the IR camera, and (c) the 3D plot of the spatial configuration for stereo calibration.

The camera calibration processes were performed to compute intrinsic and extrinsic parameters for the IR camera and the RGB camera. The intrinsic parameters included focal length, principal point, and distortion coefficients for each camera. The extrinsic parameters contained a rotation matrix and a translation matrix for a stereo system. To estimate these parameters, images of a chessboard were captured by the RGB and IR cameras simultaneously, from different viewing positions [Figure 9(a) and Figure 9(b)]. The spatial relationship between the two cameras and the various positions of the chessboard is shown in Figure 9(c). After obtaining the camera calibration parameters, the color data were mapped onto the point cloud captured by the depth sensor. This is a critical step in stitching multiple point clouds to generate road surface maps.

4.1.2.5 Kinect Calibration for Displacement

The performance of Kinect for accurate depth/displacement measurements was scrutinized by using a shaking table and a LVDT. Furthermore, to determine the kinematics of the sensors attached to a vehicle, high-frequency accelerometers were incorporated into the data acquisition system. The shaking table was used to evaluate the performance of these accelerometers.

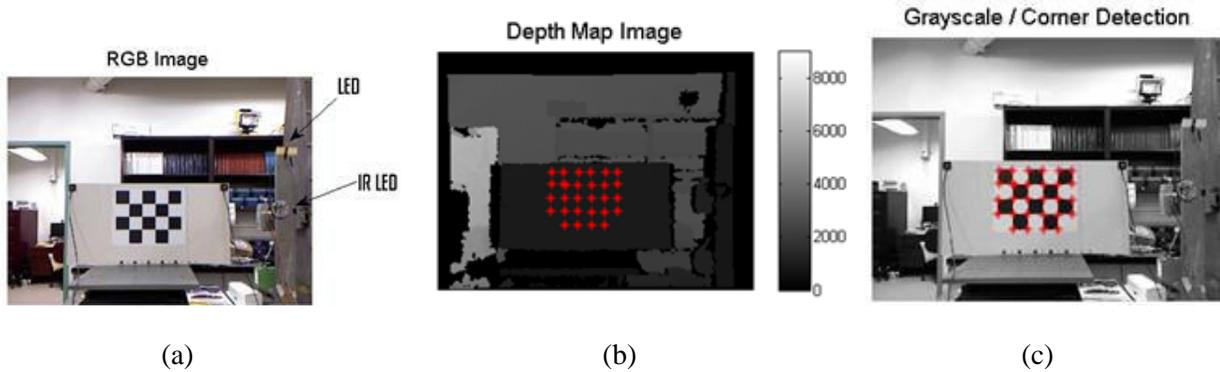


Figure 10 Experimental setup: (a) a color image of the experimental scene, (b) a depth map of the world coordinate system (black pixels are saturated or void), and (c) the checkerboard corners tracked as the rigid panel moved back and forth.

A typical laboratory experimental setup of the data acquisition and synchronization module is as shown in Figure 10. Figure 10(a) shows a color image of the experimental scene. Figure 10(b) displays a depth map of the world coordinate system (black pixels are saturated or void). Figure 10(c) illustrates the checkerboard corners tracked as the rigid panel moved back and forth. To simulate exact submillimeter-level displacement accuracy, a low- to medium-frequency range shaking table was used to produce predefined oscillations in a rigid motion. Given the known sinusoidal waveform or Gaussian white noise digital output, which mimics the vibro-interaction phenomenon between a vehicle and pavement distress, the data were converted to analog input and transformed mechanically to a shaking table. For a precise location of the rigid points, a checkerboard pattern was used as a referential measurement. Then, the corresponding depth values of red crosshairs were averaged and the waveform was filtered. It is worth noting that the camera was placed at least 0.8 m away from the moving plate. Once the shaker table was triggered with or without zero initial condition, the DAQ module was also triggered and their respective data were written to a uniformat data file.

Next, the color and depth sensors were executed simultaneously and a large dataset was recorded. Accelerometers are highly sensitive to the low/high-frequency noise; therefore it was necessary to filter the DC content, linear trend, and unwanted frequency content of the accelerometer signal before and after the integration to obtain the displacements. This was preceded by the low-pass filtering to smooth the

signal content.

Lastly, displacement signals obtained from the various sensors were compared and evaluation tests were performed. Note that the GPS is of little use within a laboratory environment; hence, it was not included in the experiments at this stage.

4.1.2.6 Pavement Crack Detection

As part of the data processing tasks, a new crack detection algorithm was developed and evaluated to segment the pavement defects. A hybrid filtering approach, consisting of several advanced and low-level filters, was developed (Figure 11). The change of curvature is a unique signature of a class of pavement cracks including surface (alligator) cracks.

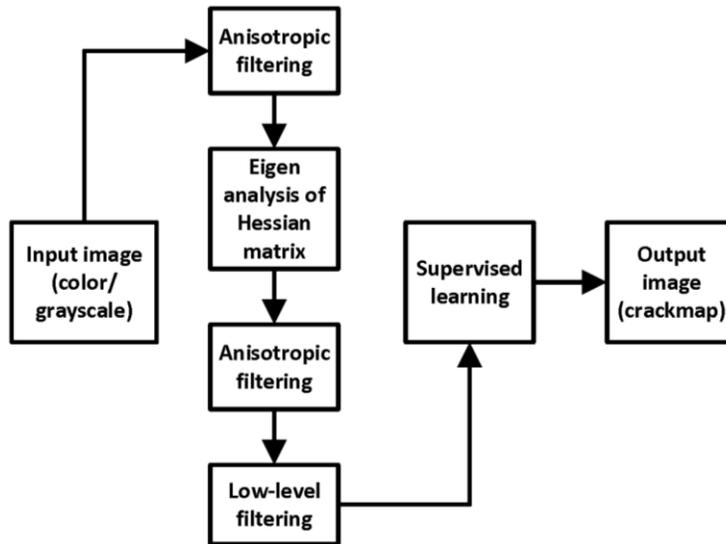


Figure 11 A hybrid crack detection algorithm flowchart.

The proposed hybrid crack detection algorithm is based on this key feature. In this approach, an image was filtered through an anisotropic diffusion procedure to ensure that most of the small curvature-dominant non-crack patterns were eliminated. Next, the multi-scale Eigen decomposition of the Hessian matrix was performed to guarantee that the curvature-dominant cracks were highlighted. Then, image binarization and post-processing steps were implemented to remove isolated pixels and undesirable lines and blobs, and to group the broken segments where a strong connectivity existed. Next, a decision-making classifier such as Support Vector Machine (SVM), neural network (NN), and k-Nearest Neighbor (k-NN) was used to remove the remaining non-crack pattern. Lastly, the output image was saved in a binary image format where white pixels represent crack and non-crack entities and black pixels

correspond to the image background.

4.2 Experimental Results and Field Tests

In the following subsections various results of the experiments conducted during the design and development of the IDEA product are discussed. Some of the experimental and field tests include:

- Data acquisition and synchronization of the multiple sensors in laboratory environment.
- Indoor test facility to measure pavement profile measurement.
- Data collection on California state freeway and close-by city of Pasadena.
- Comparison of the image at various speed for overlapping.
- Image stitching using the open source software called CloudCompare.
- Comparison of the moving depth camera outcome to a stationary LIDAR RIEGL VZ-400.
- Comparison of the GPS and VectorNav INS in field and at a stationary point.
- Rolling shutter distortion and correction for it.
- Stroboscopic setup to remove the blurring effect.
- Results of the hybrid crack detection algorithm.

4.2.1 Data Acquisition and Synchronization of Heterogeneous Sensors

Several cases of sinusoidal and random vibration environment were evaluated. It was concluded that the result of primary displacement-based sensors, including LVDT (ground truth), accelerometers, and depth sensors (Microsoft Kinect) agreed strongly in amplitude. In case of phase, there existed a delay of less or equal to 1/33rd of a second due to the maximum resolution of the frame rate which was 33 ms per frame (i.e., the sample rate of 30 fps). In fact, this error can be further minimized to sub-frame accuracy of less than 5–10 ms by using an external millisecond timer, while controlled and triggered by a DAQ module. Figure 12 shows the results of a sample experiment of random vibration measurements using synchronized LVDT, accelerometer, and Kinect sensors.

4.2.2 Pavement Profile Measurement

Pavement profile or roughness is an important factor for ride quality. A test was performed to evaluate displacement measurements obtained by a Kinect and an accelerometer for future pavement profile rating applications.

The sensor platform test facility was mounted on a large shaker shown in Figure 13. A large board was placed in front of the sensor platform for range measurement. The ranges were obtained according to the

center of the depth maps. The sensor platform mounted on the shaker was moved in sinusoidal motion. Figure 14 shows the displacement acquired by a Kinect and an accelerometer, where the two measurements are highly correlated.

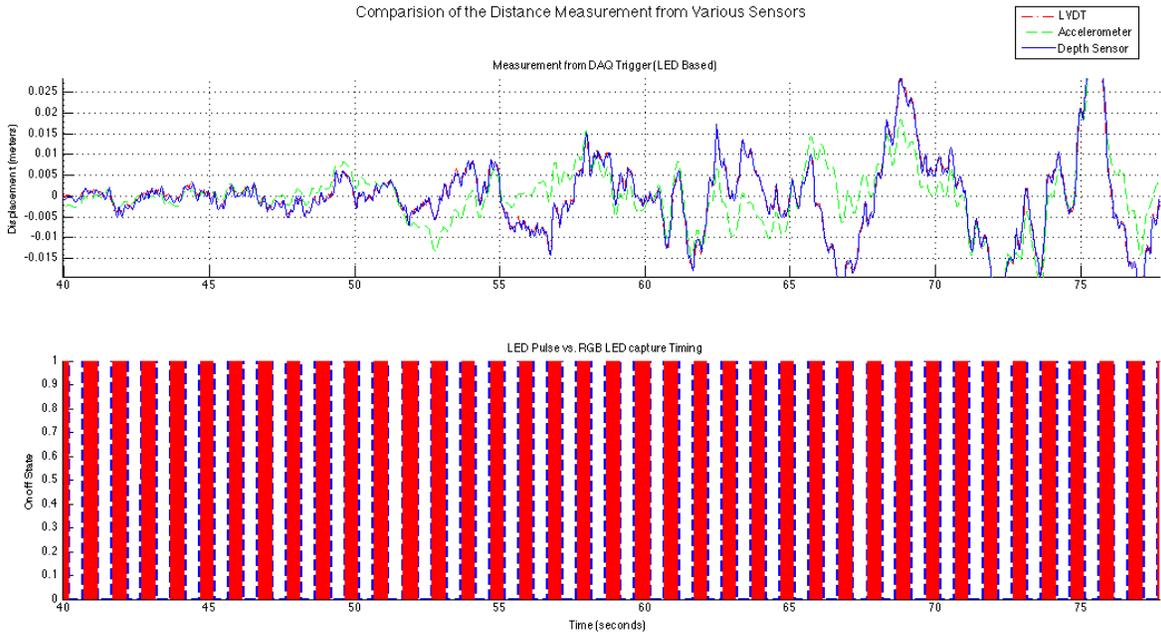


Figure 12 Comparison of the measured distance from the various sensors (LVDT, accelerometer, and Microsoft Kinect) and pulse train displaying the LED blinks in blue and recorded frames in red.

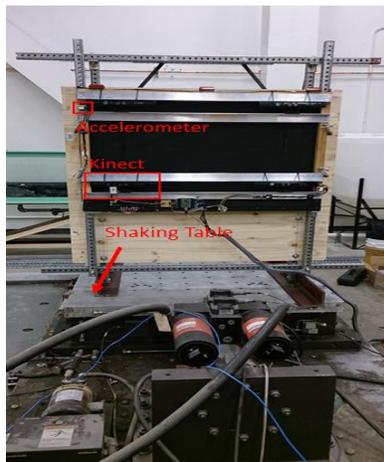


Figure 13 Indoor test facility for pavement profile measurement.

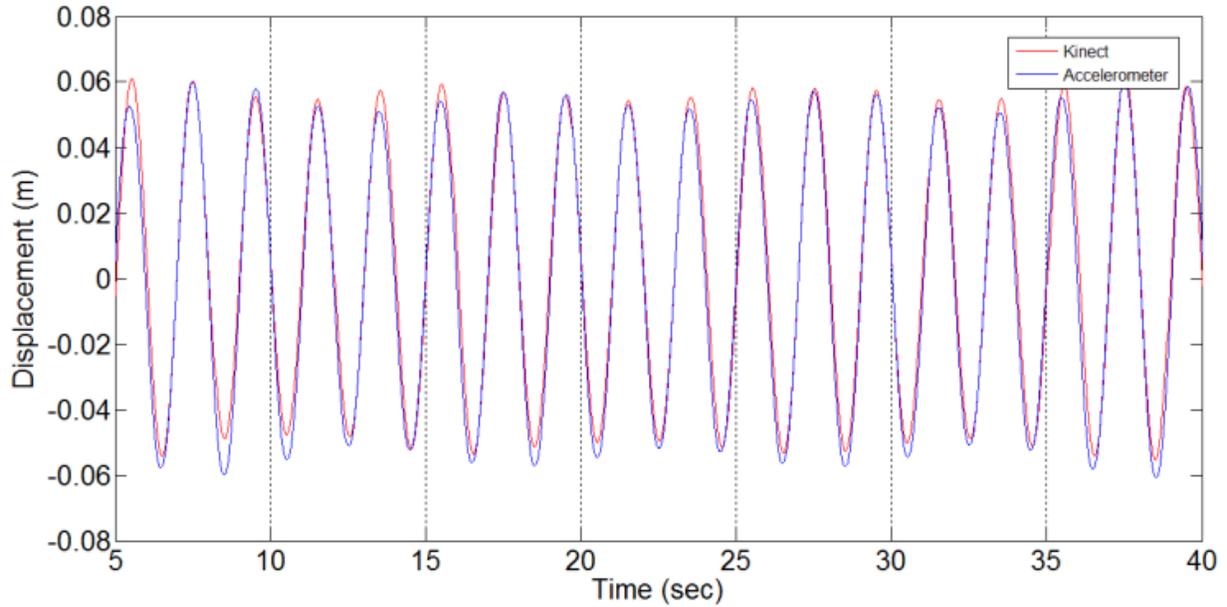


Figure 14 Comparison of the displacement measurements obtained from a Kinect and an accelerometer.

In addition, the depth data, captured from California State Highway 110 with traffic flow, were compared to the displacement computed from the accelerometers. The sensor platform was placed 0.67 m above the ground. This measurement provides the initial condition for the computed displacements. Figure 15 shows the road profile with respect to the sensor platform obtained from the accelerometers and Kinect readings. The means of the depth values over a designated area were used as the depth values obtained from the Kinect array (i.e., four Kinects). As seen in Figure 15, the Kinect depth data are closely correlated with the displacements obtained from the accelerometers (particularly in the early stage of the experiment; see Figure 15b).

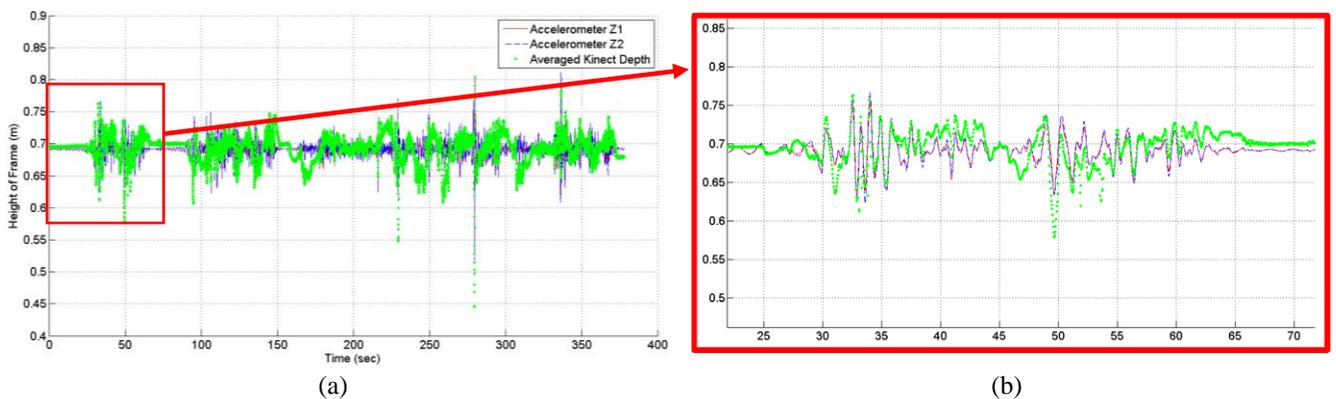


Figure 15 (a) Displacement obtained from Kinect data and accelerometers for the field experiment on California State Highway 110 while the baseline (i.e., the initial distance between the sensor platform and the ground) was 0.67 m; and (b) an enlarged portion of (a).

The cumulative error in computing the displacement from the accelerometer data grows as time passes, which leads to less correlation between the displacements obtained from the accelerometers and Kinects. In addition, higher vehicle speeds could have affected the quality of the depth data captured by Kinects, as there exists distortions in depth information (refer to rolling shutter distortion).

4.2.3 Pavement Data Collection on California State Freeway and on City of Pasadena Streets

To elaborate on the data collection and processing, several field tests were performed to evaluate the capabilities of the aforementioned data acquisition system. The majority of the tests were close to the vicinity of the University of Southern California (USC) campus, at a low speed, and the rest were close to the city of Pasadena, California, at a normal highway speed.

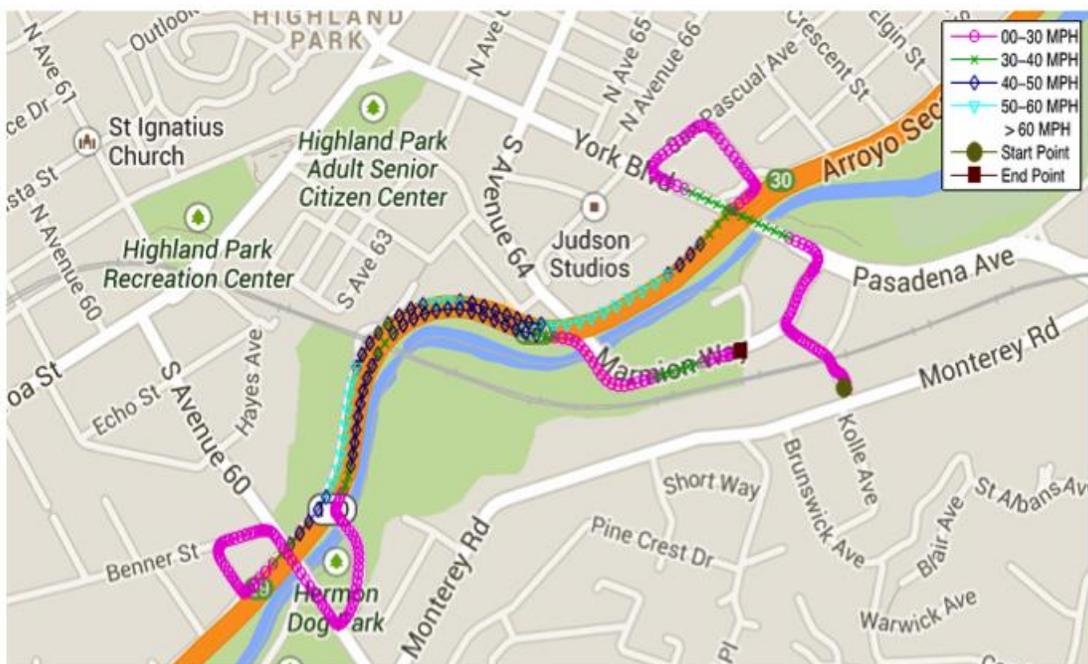


Figure 16 An example of field tests: data collection route with different vehicle speeds on California State Highway 110.

On average, each on-road test lasted for three to five minutes in duration. Figure 16 shows the data collection route, and the corresponding GPS-measured speed at 0 to 96.56 km/h (0 to 60 mph), for one of the tests, where the vehicle was driven along California State Highway 110. The complete road trip was 3.86 km (2.4 miles), as obtained from the Garmin GPS system. In total, 300 gigabytes of data were

recorded for more than two and half hours for the road trip.

4.2.3.1 Acceleration Data

The collected data from high-frequency accelerometers were synchronized with GPS data, and depth and color frames. This data were used to enhance the image alignment process while stitching the multiple images. Since the local axes of the RGBD cameras are crucial in image registration processes, each single axis accelerometer was exactly aligned along the axes of the RGBD cameras, as shown in Figure 17.

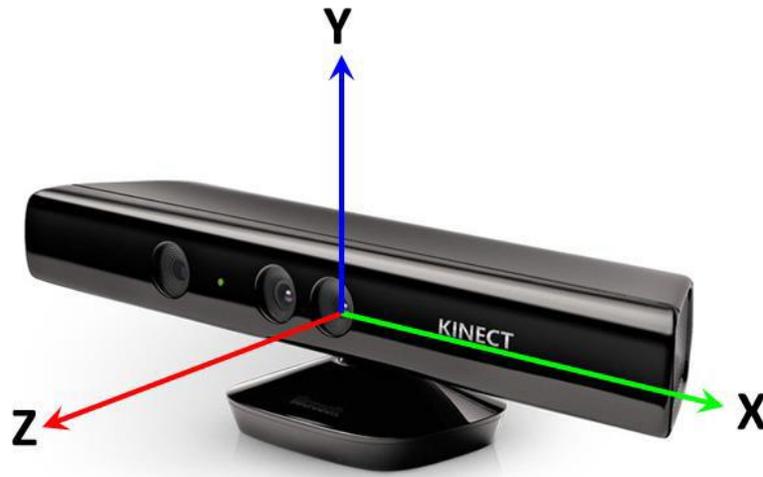
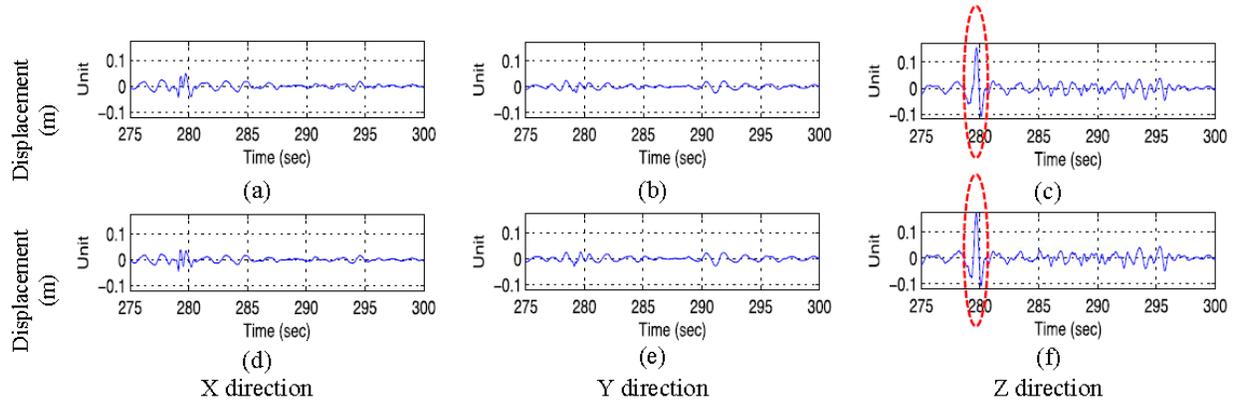


Figure 17 Camera/accelerometer convention axes.

By knowing the rotation and translation caused in the camera/accelerometers plane, image alignment and multiple stitching can be enhanced. Furthermore, accelerometers were used to measure the sudden impacts caused by the potholes or road bumps while the vehicle is in motion. Figure 18 shows the displacements from 275 to 300 seconds of a field test. The data were obtained by double-integrating the acceleration data captured by the six accelerometers. Note that three accelerometers were clustered to form a triple axis accelerometer, and they were mounted on opposite corners of the frame [see Figure 5(a)]. Time histories in Figure 18 in Z direction (i.e., the gravity direction) show a subtle vertical displacement of 10 cm at 280 seconds. By using the information obtained by the cameras to quantify, classify, and localize the defect, accelerometers are useful in filtering the false positive and negative alarms. Table 2 summarizes the synchronized inertial quantities and GPS data corresponding to each depth and color frame.



**Figure 18 Time histories of displacements for 275 to 300 seconds.
A pothole or road bump is identified at 280 seconds.**

**TABLE 2
SAMPLE OF SYNCHRONIZED MEASURED QUANTITIES FOR EACH RGBD FRAME**

Frame No.	Time (sec)	Inertial Quantities									GPS Data				LED State
		Acceleration (m/s ²)			Velocity (m/s)			Displacement (m)			Latitude (n ^o)	Longitude (n ^o)	Speed		on-1/ off-0
		1	2	..	1	2	..	1	2	...			mph	km/h	
1	0	0.20	-0.27	..	-0.00	-0.07	..	-0.00	-0.00	...	34.110487	-118.175180	0.20	0.33	1
...
213	7.07	-0.03	0.70	..	-0.00	-0.00	..	0.00	0.00	...	34.110511	-118.175232	0.10	0.17	1
...
5555	189.77	0.07	-0.06	..	0.02	0.08	..	-0.00	0.00	...	34.111264	-118.181375	47.82	76.96	0
...
8623	296.11	-0.64	-0.64	..	-0.07	0.01	..	-0.00	-0.01	...	34.106884	-118.184912	18.75	30.18	1
...
11031	378.04	-0.07	-0.09	..	0.00	0.00	..	0.00	0.00	...	34.110831	-118.177154	0.10	0.17	1

4.2.4 Pavement Images at Various Speeds

The color and depth frames were recorded under various speeds to study if there is enough overlap (i.e., at least 10%) between sequences of captured frames to stitch them together. Figures 19 (a) and (b) show overlaps on two images for planar motion stitching, taken under speeds of 10 mph and 50 mph, respectively, where the data collection rate was 30 fps. An image frame acquired from a higher speed has less overlapping area.

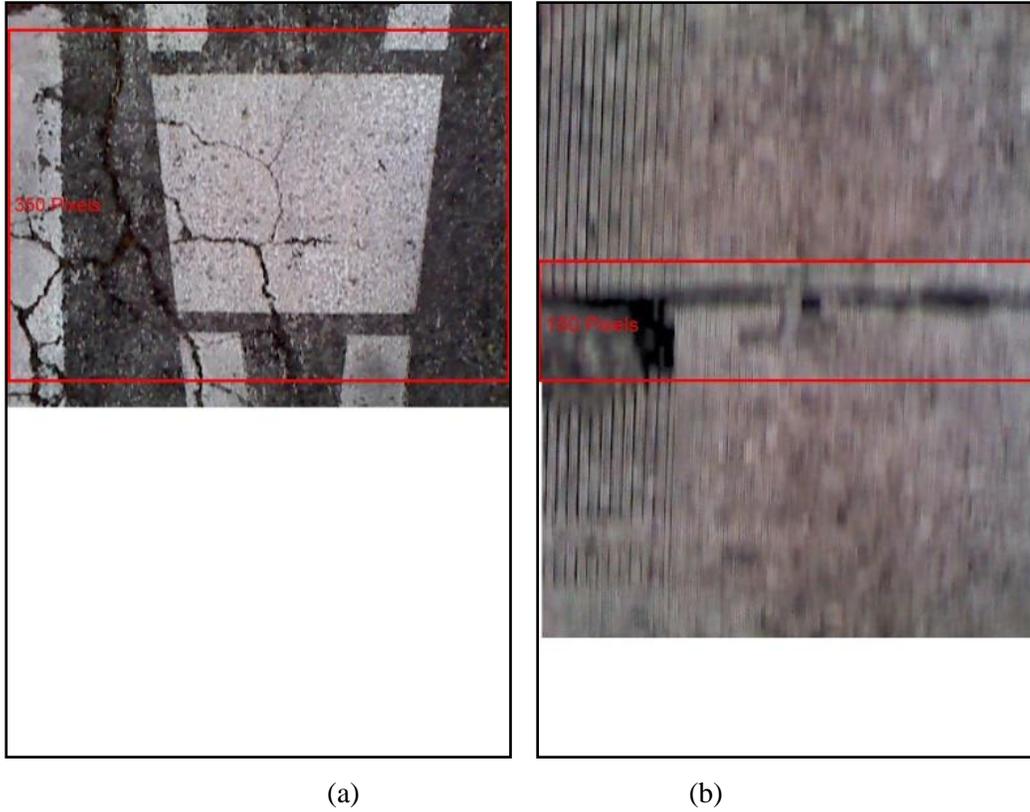


Figure 19 Planar image stitching under (a) 15 mph and (b) 50 mph.

Figure 20 shows the relationship between consecutive frame overlaps and various speeds. Based on this figure, a sequence of color frames still have about 30% overlaps when the motion speed reaches 50 mph, which makes the planar image stitching possible.

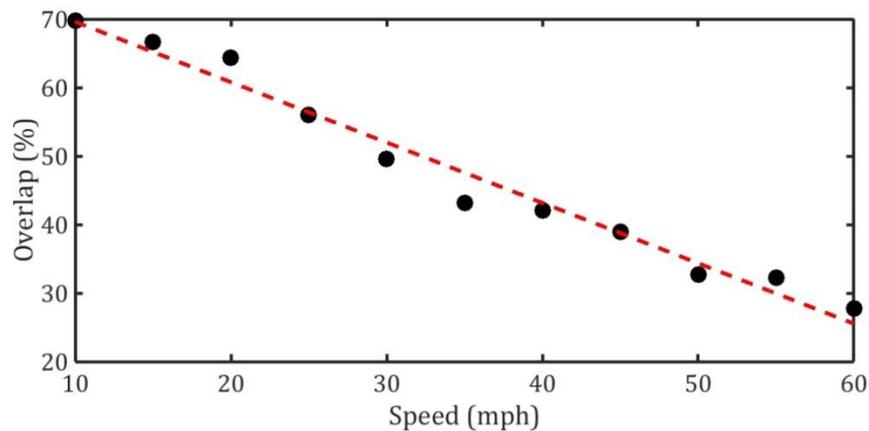


Figure 20 Overlap percentages of the sequential color frames under various vehicle speeds.

Since these images were acquired at about 5 p.m. in November, the low luminance triggered Kinect to extend its exposure time and, consequently, it introduced motion blur in the captured frames. Figure 21

shows the difference between image frames captured of the same road segment under high and low luminance, on California State Highway 110.

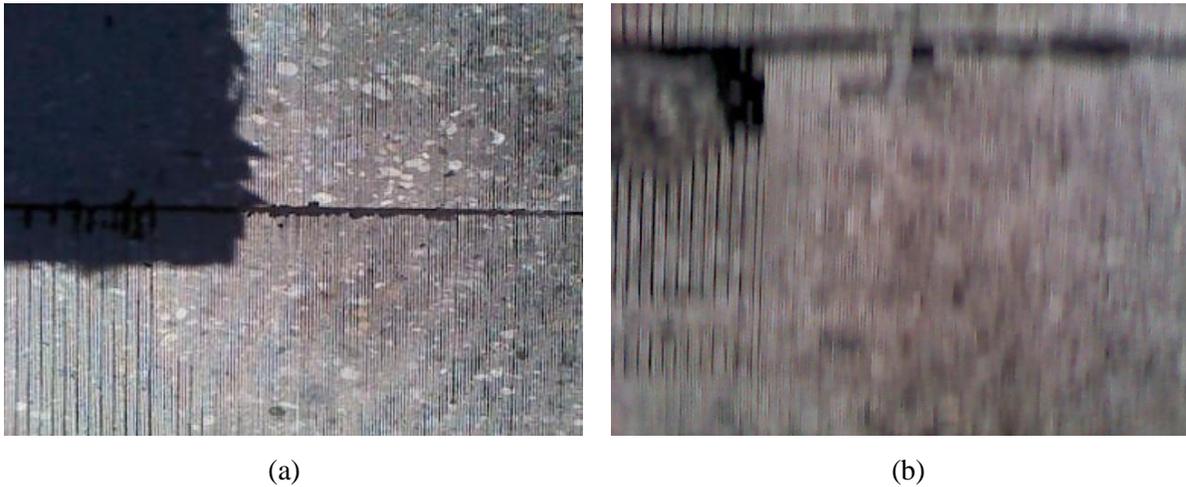


Figure 21 Comparison of two images captured from the same road segment under (a) high luminance and (b) low luminance.

This figure shows that the bright illumination has led to a sharper image. The motion blur problem could be solved by disabling automatic exposure control of Kinect.

To generate color points, the depth frames and the corresponding color frames were superposed by the prior calibration of the Kinect sensing modules. Figure 22(a) shows the color map representation of a captured depth frame, where colors correspond to the sensor-pavement for each pixel. Figure 22(b) shows the point cloud, where the true color data is superimposed on the depth data.

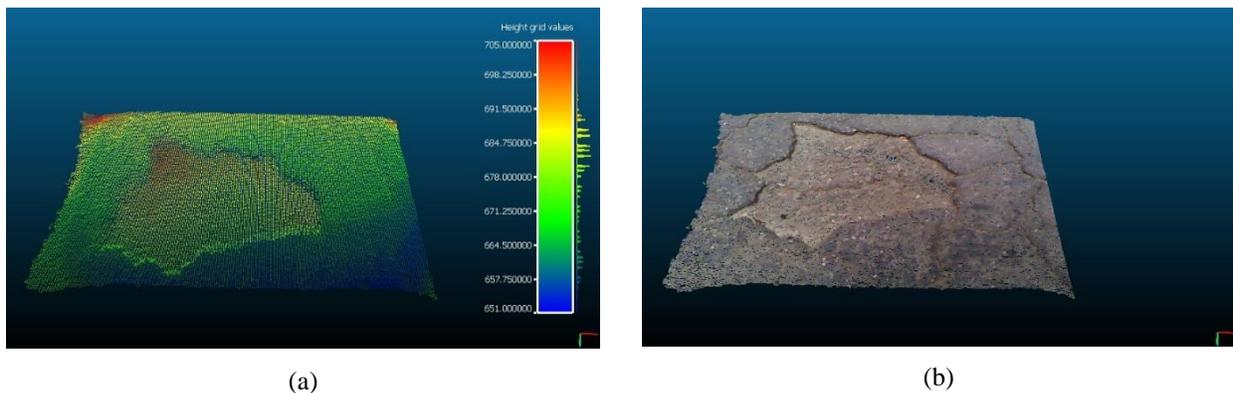


Figure 22 (a) Point cloud with height colormap and (b) point cloud with true color values.

After collecting the four individual depth and color frames at a given time, these frames have to be stitched together to generate a larger image and point cloud of the road, for further damage assessment

processing. To this end, tests were performed to stitch color and depth frames. Figure 23 shows the stitching of four colored images, which had at least a 10% overlap by utilizing a feature-based image composition approach. As an initial effort to generate large-scale color point clouds, the four point clouds corresponding to color images in Figure 23 were composed manually with a simple planar translation to form a full-view point cloud displayed in Figure 24.

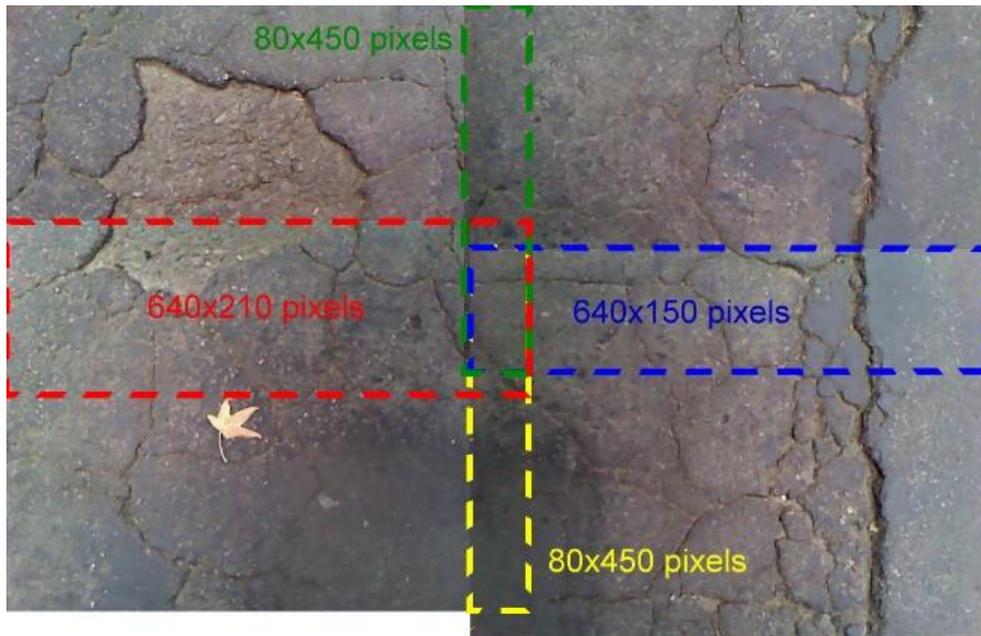


Figure 23 Color image stitching of four RGB frames.

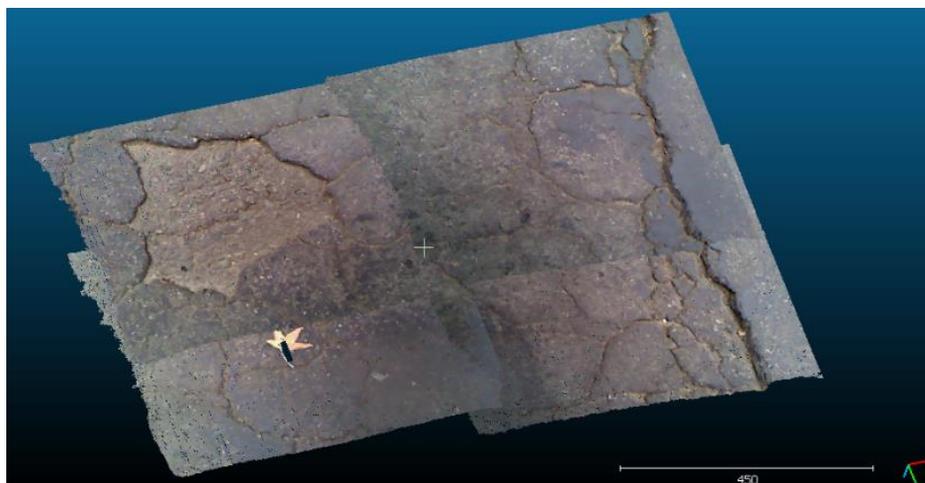


Figure 24 Composition of four color point clouds.

4.2.4.1 Defect Ground-truthing at Various Speeds

Two sets of tests were performed to evaluate the performance of the Kinect data acquisition system covered by the full-cover sun shade. To obtain the ground-truth data, point clouds of potholes were acquired by a high-accuracy RIEGL VZ-400 LIDAR (Figure 25).



Figure 25 RIEGL VZ-400 LIDAR used for ground truthing.

Figure 26(a) shows a pothole on a pavement surface, which was selected as a test target. The color and depth images were captured while the vehicle was stationary. Next, the color and depth images were captured while the vehicle was moving at 15 to 20 mph. Figure 26(b) shows the distance between the two points on the point cloud corresponding to the pothole in Figure 26(a) that was captured by a stationary Kinect. This estimated value is 448 mm, which is close to the manual reading in Figure 26(a). Figures 26(c), (d), and (e) illustrate depth images acquired by a stationary Kinect, a moving Kinect at a speed of about 15 mph, and the stationary RIEGL VZ-400 LIDAR, respectively.

The depth map obtained by a moving Kinect has distortion in shape, which is caused by the rolling shutter mechanism of the infrared Complementary Metal-Oxide Semiconductor (CMOS) inside the Kinect. A comparison was made to analyze error between two aligned point clouds captured by a stationary Kinect and a high-accuracy RIEGL VZ-400 LIDAR. Open source CloudCompare software was used to align the two point clouds. The procedure to register two point clouds includes two steps. First, the two point

clouds are roughly aligned by manually picking at least three corresponding points in each point cloud, and then performing fine tuning using the Iterative Closest Point (ICP) algorithm. Figure 27(a) shows a color map of point pair differences between the point clouds obtained by a stationary Kinect and VZ-400 LIDAR. Figure 27(b) displays a histogram of point pair distances in the Z direction. Point pair distances of 48,144 points (out of 69,711 total points) are less than 2.5 mm, which is 70% of the population.

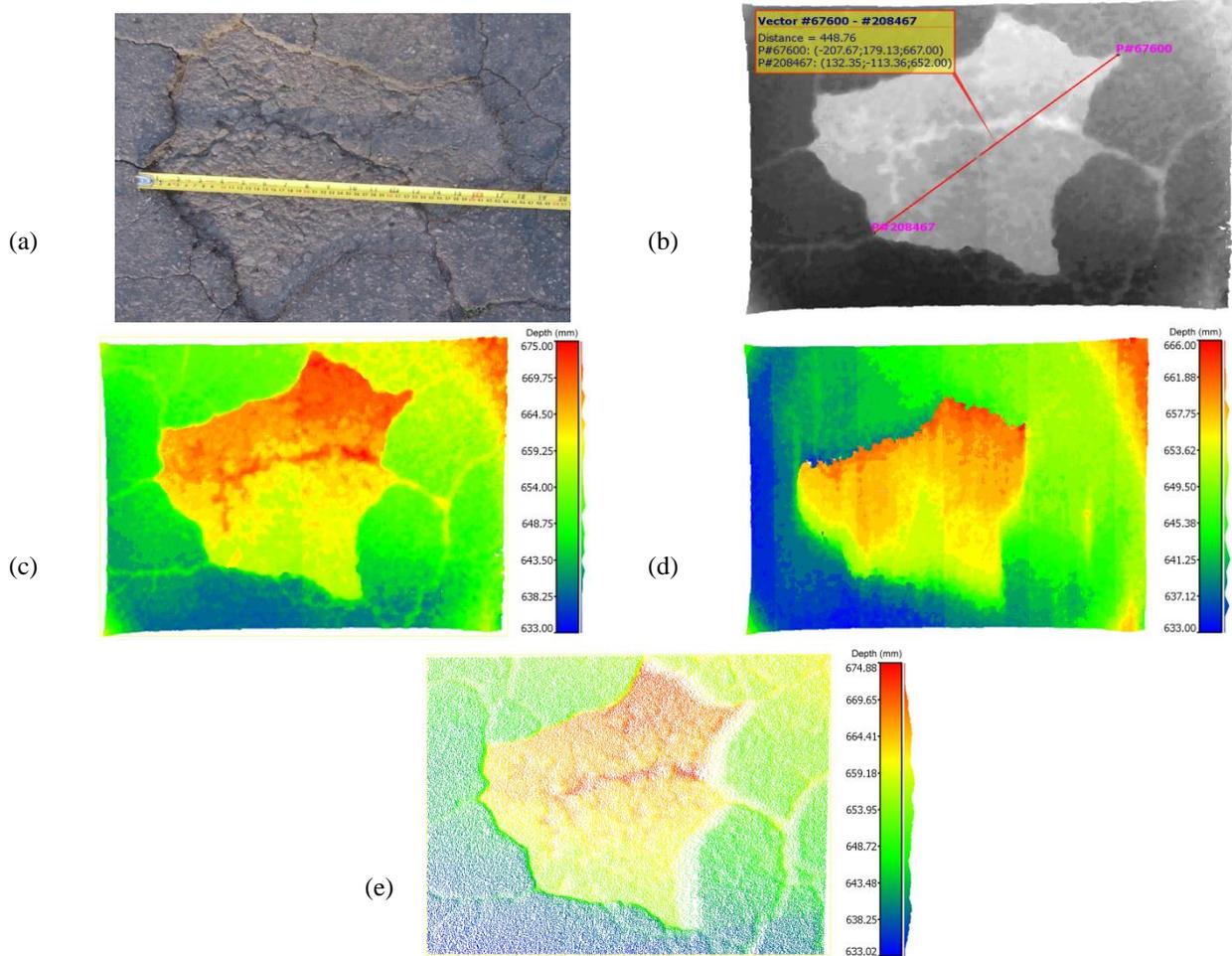


Figure 26 (a) Manual measurement of a pothole (tape measurement is 448 mm); (b) distance between two points on the point cloud acquired by a stationary Kinect; (c) depth values acquired by a stationary Kinect; (d) depth values acquired by a moving Kinect at 15 mph; and (e) depth values acquired by the RIEGL VZ-400 LIDAR.

The same registration and point pair comparison methods were also applied to the two point clouds acquired by a stationary Kinect and a moving Kinect [Figures 26(b) and (d)]. Figure 27(c) shows the color map of point pair distances. It also illustrates that an area colored in yellow, red, and orange has large point pair distance in Z direction distributing from 17 to 27 mm.

This finding reveals that the size of the pothole is shrinking in the longitudinal direction when a mobile Kinect is used. Figure 27(d) shows the histogram of the depth difference between two aligned clouds, obtained from stationary and mobile Kinects.

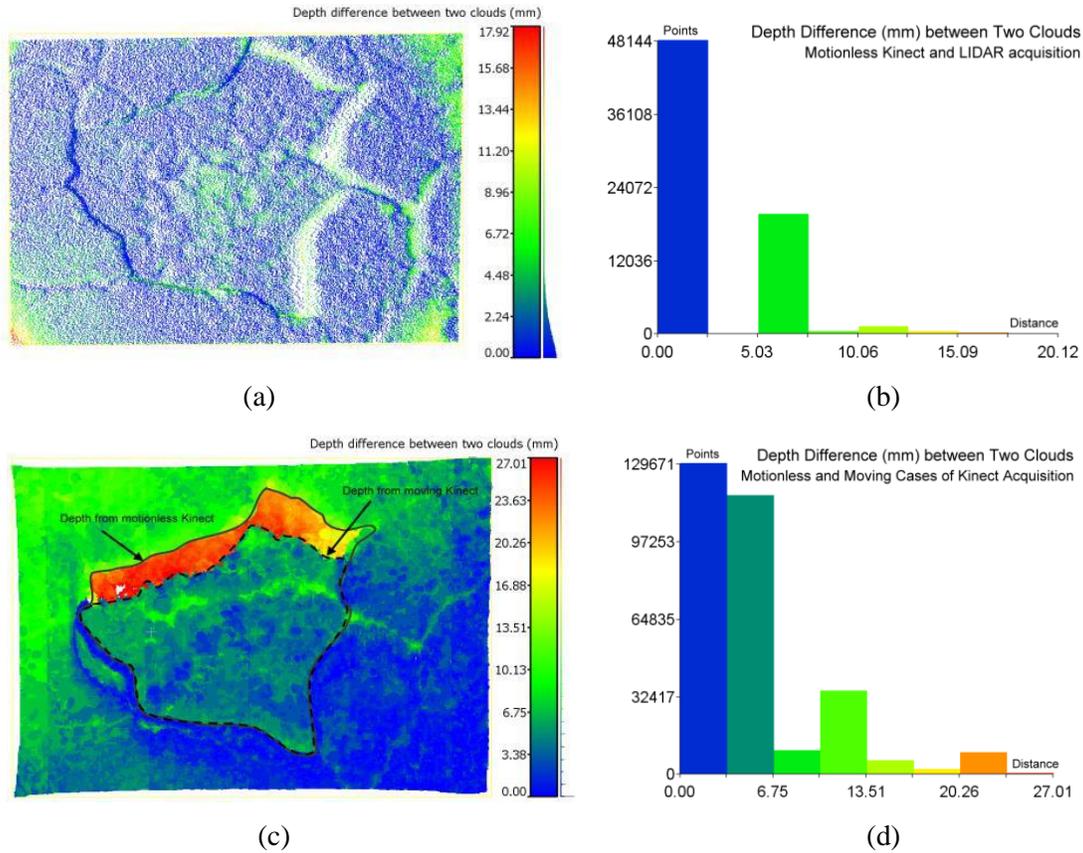


Figure 27 (a) depth difference between two aligned point clouds obtained by a stationary Kinect and the VZ-400 LIDAR; (b) the histogram of depth difference between two aligned clouds in (a); (c) depth difference between two aligned point clouds obtained by a stationary and a moving Kinect at 15 mph; and (d) the histogram of depth difference between two aligned clouds in (c).

4.2.4.2 Effect of Speed on Autonomous Defect Detection and Quantification

Figures 28(a) and (b) show a sample of the results from autonomous processing by an algorithm, developed by the authors, for defect detection and quantification using the depth maps captured by stationary and mobile Kinects (3). Table 3 summarizes the quantified parameters and shows the effect of speed on the quantification process.

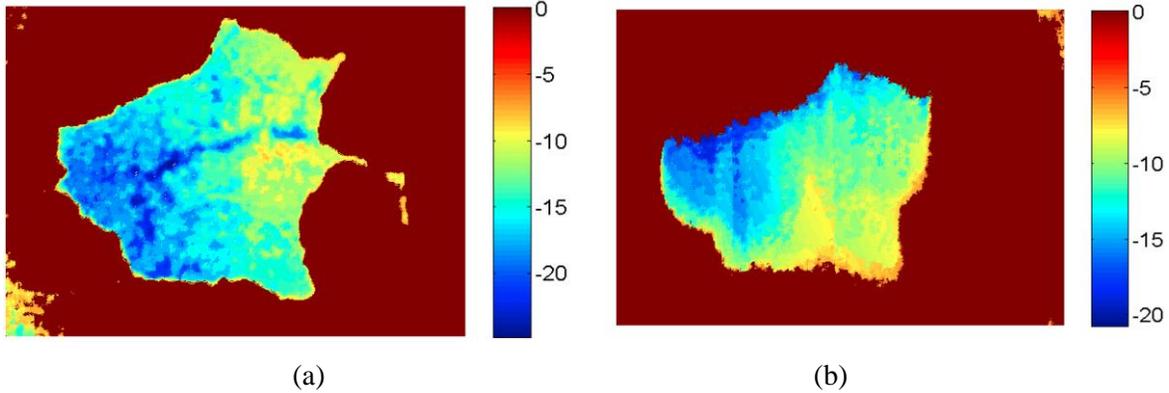


Figure 28 Effect of speed on autonomous defect detection and quantification: (a) data captured by a stationary Kinect, and (b) data collected by a mobile Kinect.

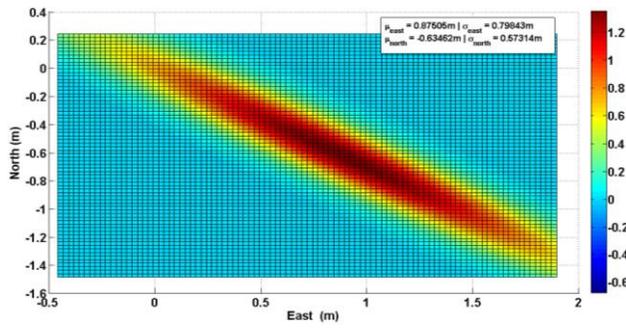
TABLE 3
QUANTIFIED CHARACTERISTICS OF DEFECTIVE REGIONS IN FIGURE 28

Sensing	Area (mm ²)	Volume (mm ³)	Length (mm)	Width (mm)	Max Depth (mm)	Mean Depth (mm)
Static	1164.9	1619.1	396.7	387.7	24.9	13.9
Mobile	887.4	1000.9	412.7	233.0	20.8	11.3

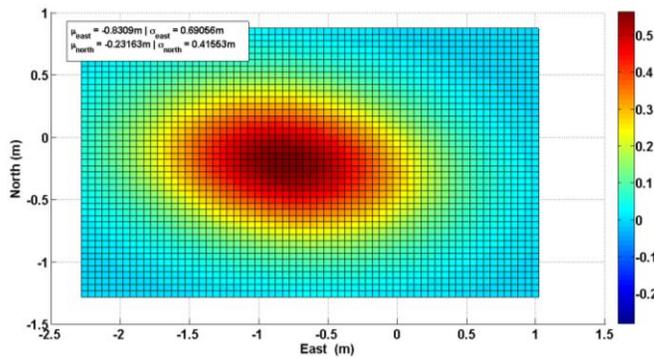
4.2.5 Localization of the Pavement Defects

Despite segmenting the defects from pavement background, it is also very important to know the location of the defects. Knowledge of the location can be used for efficient operation and maintenance. For example, knowing the location can help in finding and fixing a pothole or other defects. For the purpose of localization, two hardware systems are used. First, a Garmin 18x USB GPS; second, a VectorNav's VN-200 INS, combined with Extended Kalman Filter (EKF).

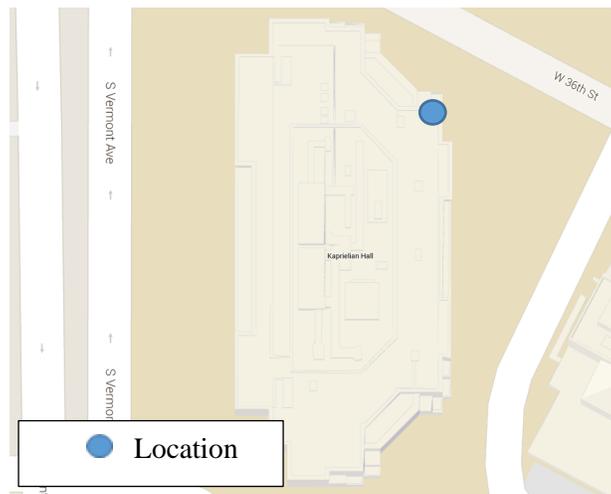
Two of the above sensors underwent two different types of the field test: a stationary test and a mobile test. In the stationary test, the two sensors are placed on a plane surface, each separated at a distance of 152.4 mm. Almost 2,100 and 420,000 data samples were collected from the Garmin GPS and VectorNav's INS, respectively, for 35 minutes. The data collected are longitude and latitude values. The location of the stationary test was at USC, on the roof of Kaprielian Hall [see Figure 29(c)]. Figure 29(a) and (b) shows the results of the stationary test.



(a)



(b)



(c)

Figure 29 Stationary test of Garmin GPS and VectorNav INS: (a) Probability density function of Garmin GPS; (b) Probability density function of VectorNav INS; and (c) Location of the test at USC, KAP Hall.

From Figures 29(a) and (b) it is evident that the standard deviation of the VectorNav unit is less than that of the Garmin GPS. The VectorNav unit is embedded with EKF and an incorporated INS unit, and has higher number of samples to be averaged. Thus, the standard deviation is better than the Garmin GPS.

Usually, a conventional GPS could provide long-term position accuracy but low bandwidth measurements of 1 Hz. The inertial sensors could provide high bandwidth acceleration and angular data but suffer from inaccurate long-term position and orientation measurements because of drift and integration errors. A study was conducted and tested with a miniature high-performance GPS-aided inertial navigation system (VectorNav VN-200) which was composed of MEMS (microelectromechanical systems) inertial sensors, a GPS receiver, and Kalman filtering algorithms. This navigation system generates optimal estimations of position, velocity, and orientation, at a rate of up to 200 Hz. The results were compared to a conventional GPS (Garmin 18x) and a high-performance GPS/INS (VectorNav VN-200) through a 50-minute, 9.65 km (6 miles) road test shown in Figure 30(a). Most waypoints generated by the Garmin GPS and VectorNav VN-200 GPS/INS were matched, and could be mapped on the lanes. However, Figure 30(b) shows that the VectorNav VN-200 GPS/INS suffered from serious drifts due to the presence of tall buildings. This anomaly is known as multipath because the minimum requirements of GPS satellites are unavailable to the GPS receiver; therefore erroneous trilateration causes a drift from the true position. However, an aided INS device is usually hundreds of times the price of a conventional GPS.

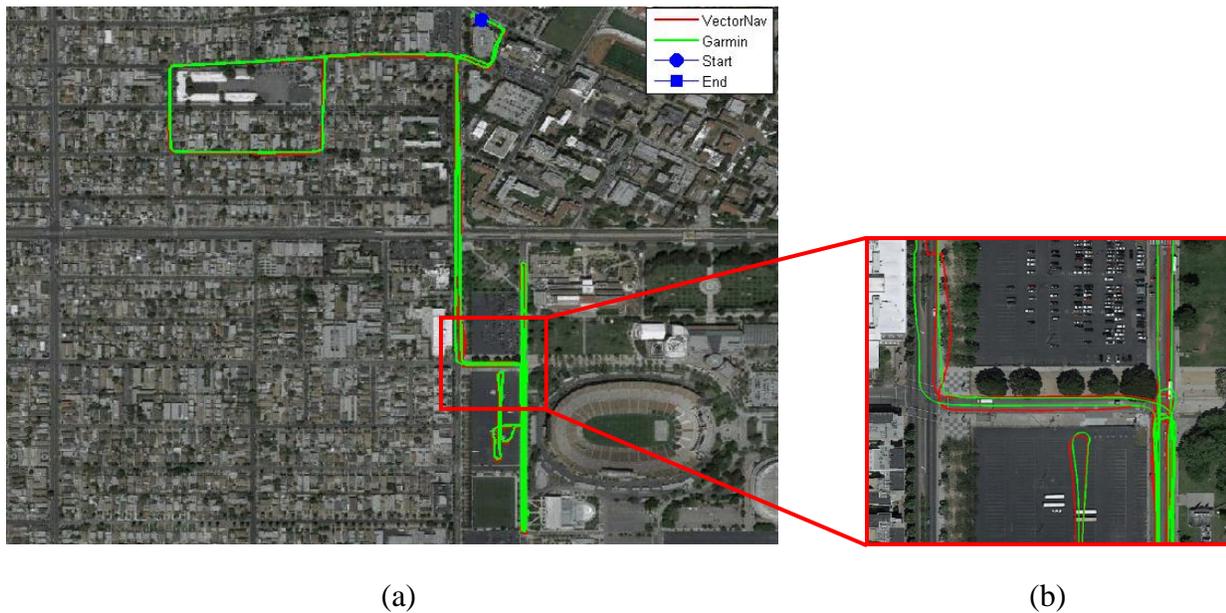


Figure 30 (a) A comparison between a Garmin GPS and a VectorNav GPS/INS tracking performance near USC and the Coliseum in the city of Los Angeles. (b) A close-up of a tracking route close to a tall building and open space.

4.2.6 Rectification of Rolling Shutter Distortion

Both the color camera and depth camera of a Kinect sensor have rolling shutter CMOS image sensors. When a rolling shutter CMOS sensor records each frame of a video stream, images are not captured by the entire pixel array simultaneously, but rather by scanning row-by-row or column-by-column, in the pixels. Rolling shutter CMOS sensors provide low-noise, low-power, fast data processing, and an inexpensive solution for most commercial cameras; however, this image acquisition method creates distortions when shooting moving objects.

The phenomenon is illustrated in Figure 31(a). To study and solve this problem, an experiment was conducted where a square brick was scanned with a Kinect sensor at different vehicular speeds, and then rectified the distortion using a rolling shutter camera model (1, 2). The moving direction of the Kinect is perpendicular to the scanning direction of the rolling shutter; therefore, the color and depth image of the square brick deforms into the shape of a parallelogram [see Figure 31(b) and 31(c)]. If the speed of motion increases, the acute angle of the acquired parallelogram decreases. The plot of the relation between speed and acute angle is shown in Figure 32. The measurements for acute angles of distorted depth images of different speeds are listed in Table 4. Currently, the assumption made for the motion of Kinect was a one-direction planar translation, and there is no consideration of rotations. Therefore, the rectification algorithm requires input of the speed only, which was estimated using a car's speedometer in this sequence of road tests. After applying the rectification algorithm on the deformed depth images that were acquired by a Kinect sensor, the results displayed in Table 4 demonstrate that the algorithm is able to restore the distortion.

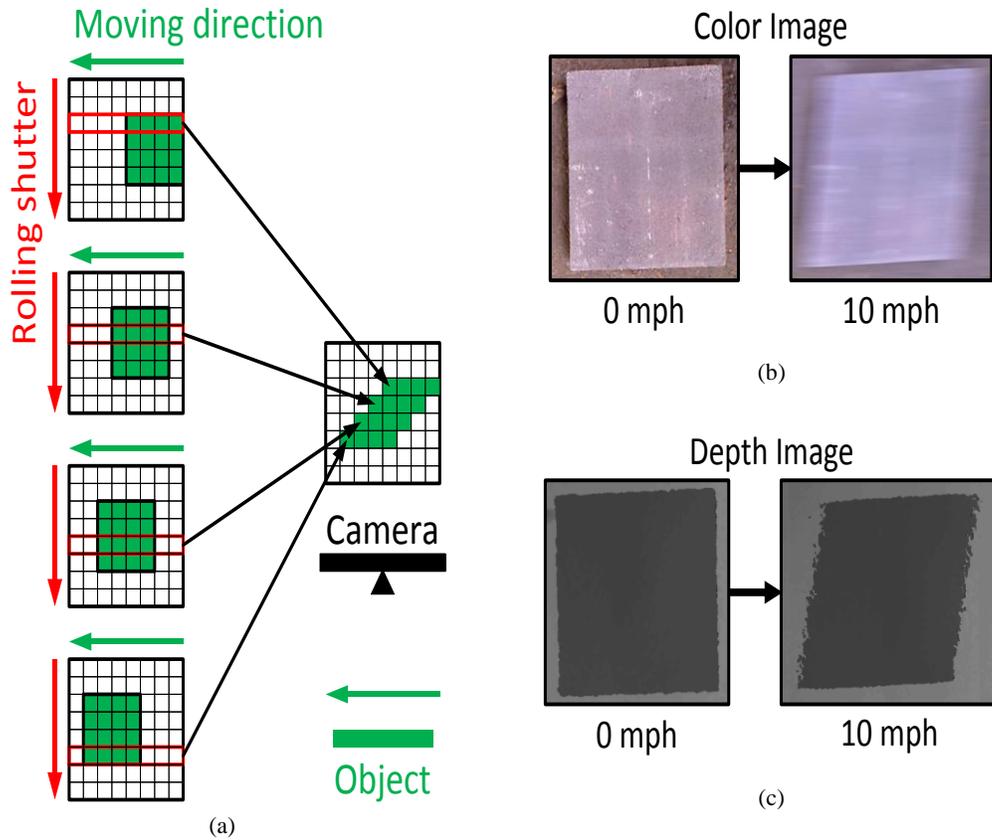


Figure 31 (a) Rolling shutter image acquisition and distortion; (b) the color image of a square brick captured by a Kinect's color camera at 10 mph became a parallelogram; (c) the depth image of a square brick captured by a Kinect's depth camera at 10 mph became a parallelogram.

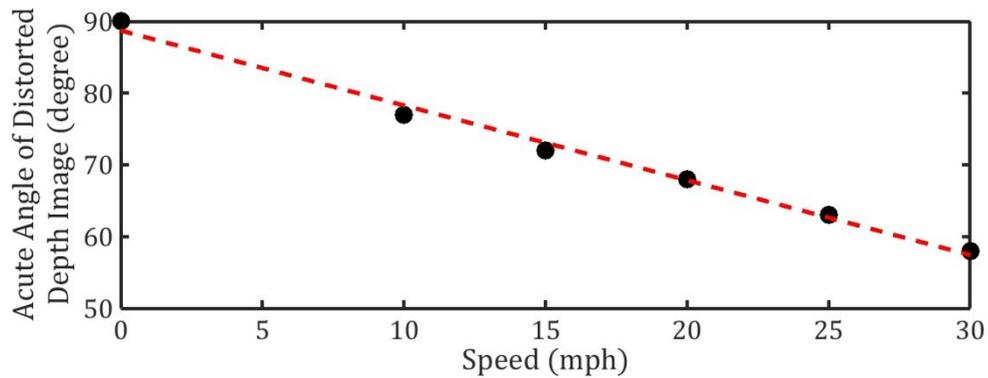
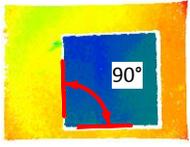
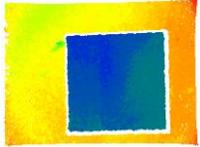
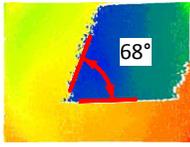
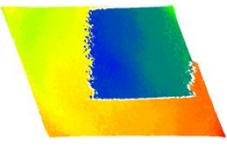
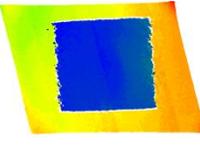
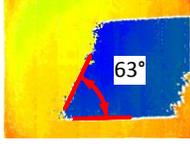
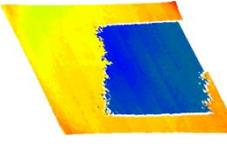
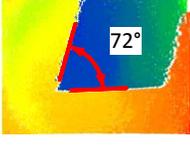
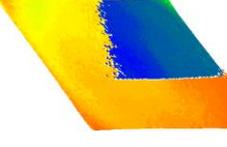


Figure 32 Relationship between rolling shutter distortions for depth images and motion speeds according to Table 4.

TABLE 4
DISTORTION AND RECTIFICATION OF DEPTH IMAGES FOR DIFFERENT SPEEDS

Speed	Distortion	Rectification	Speed	Distortion	Rectification
0 MPH			20 MPH		
10 MPH			25 MPH		
15 MPH			30 MPH		

4.2.7 Stroboscopic Solution for Motion Blur Problem

A stroboscopic technique is used to capture slow-motion pictures. This approach was used to solve the motion-blur problem for Kinect color image acquisition. Three LED work lights were mounted inside the full-cover sun-shade platform, which is shown in Figures 33(a) and (b). The low-power LED is an ambient light source which will not interfere with the depth sensor of a Kinect; therefore, it is a desirable illumination option for this application. Moreover, LEDs are easily controlled by digital signals. The pulse signal sent from an Arduino board triggered a transistor to switch the electrical current flowing into the LEDs to enable strobe lighting. The setup of strobe controller is shown in Figure 33(c).

The road test results corresponding to motion speeds less than 56.32 km/h (35 mph) using a Kinect with strobe light assistance are shown in Figure 34. The flash rate of the strobe light was 60 Hertz (Hz), and the pulse duration was 1 ms. The quality of the color images was improved when using the stroboscopic technique under discussion.

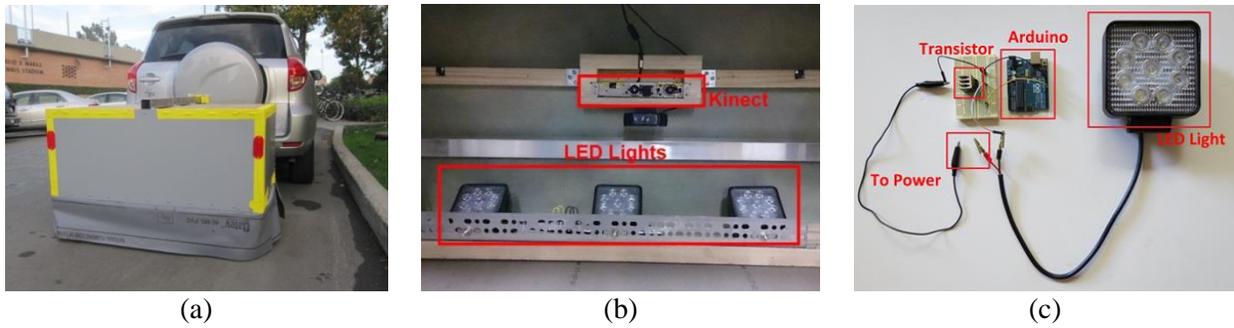


Figure 33 (a) The full-cover sunshade sensor platform; (b) representation of a Kinect, and LED strobe lighting mounted inside the full-cover sunshade; (c) the scheme of the strobe light controller.

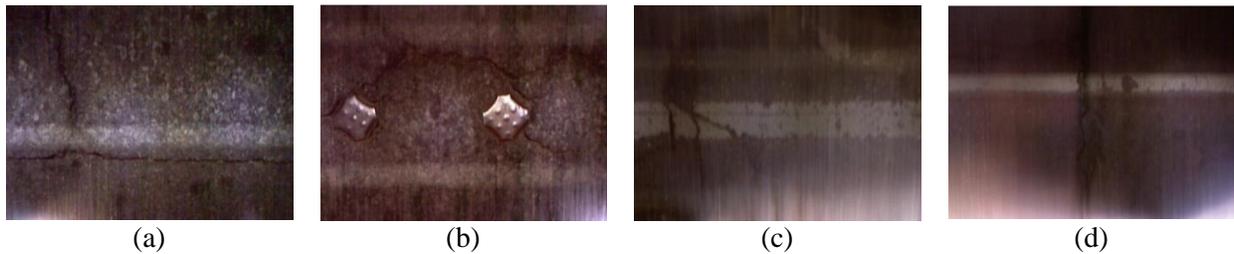


Figure 34 Road testing results for Kinect image acquisition using the strobe lighting. (a) and (b) were captured at 24.14 to 40.23 km/h (15 to 25 mph), (c) and (d) were taken at 24.14 to 56.32 km/h (25 to 35 mph).

4.2.8 A Hybrid Pavement Crack Detection Algorithm

Figures 35 and 36 show the results obtained from the proposed approach and an alternate morphological crack detection algorithm. As it is seen from these figures, the proposed hybrid algorithm performs better than the modified bottom-hat morphological method (4–6). In addition, the proposed algorithm is 2.5 times faster than the mathematical morphological method. Note that the proposed algorithm generates false positive alarms in the presence of shadow; however, since the full-shade data collection platform is used, no shadows will be present in the captured images and the issue is automatically resolved. Furthermore, to be fair for both the methods, no decision making system is included, as one needs to explicitly train and test both the methods. Instead, in this report both methods are compared based on their segmentation capabilities only.

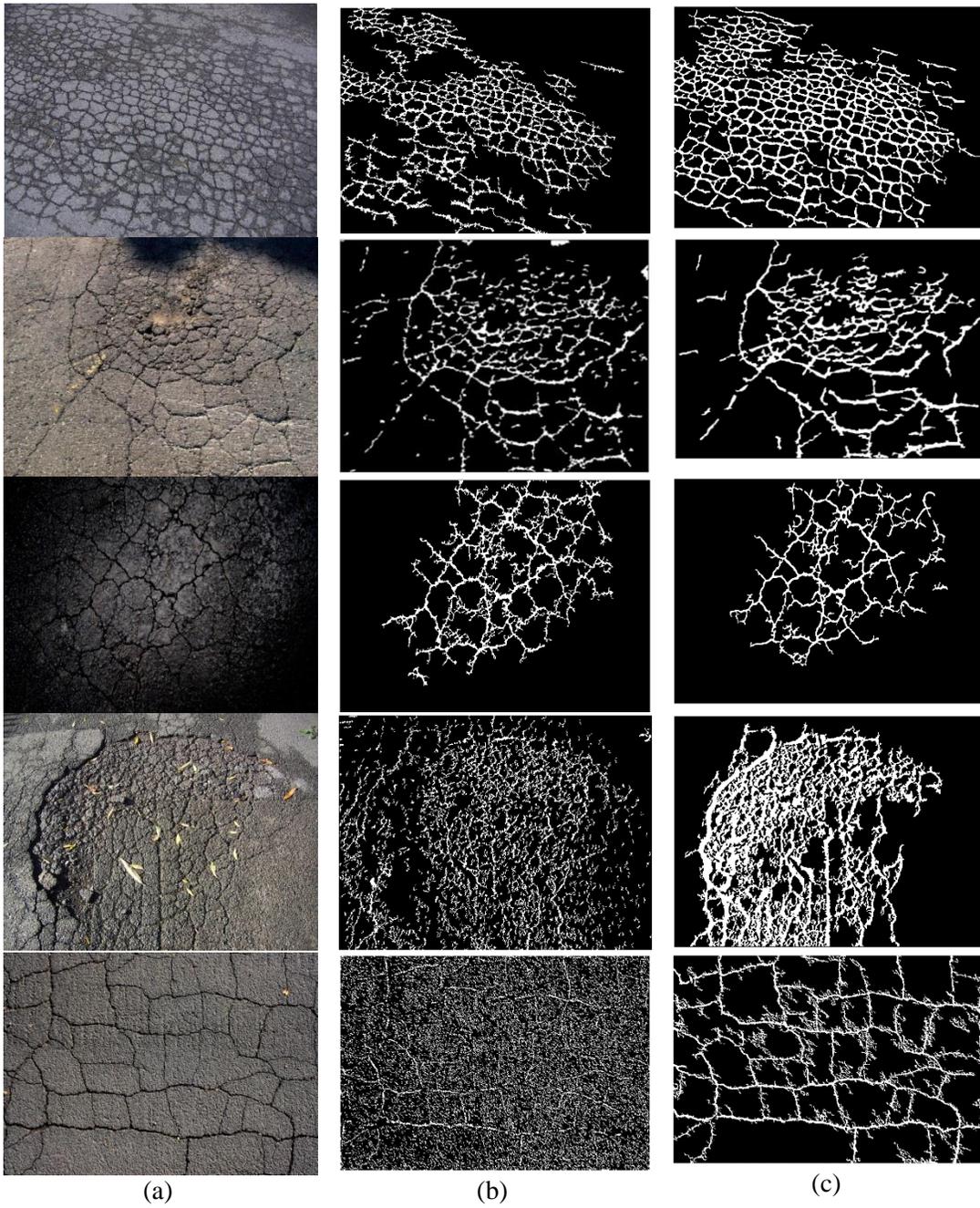


Figure 35 (a) The original high resolution images; (b) crack map obtained using a modified bottom-hat morphological method; (c) crack map obtained using a hybrid algorithm.

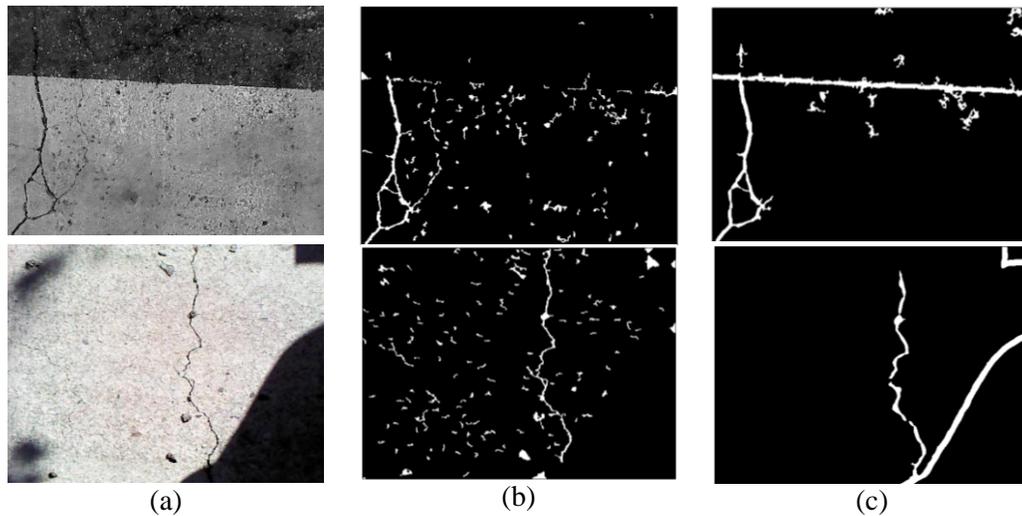


Figure 36 (a) The original low resolution (640 x 480 pixels) images, (b) crack map obtained using a modified bottom-hat morphological method, and (c) crack map obtained using the proposed hybrid algorithm.

5. Plans for Implementation

The research team is having discussions with the FHWA, Fugro Consultants, Inc. (Fugro), and the California Department of Transportation (Caltrans) to implement the results of this study in real-world applications.

6. Conclusions

This study resulted in the development of a relatively inexpensive vision-based approach to perform autonomous detection, localization, and quantification for pavement distress. The inexpensive vision-based approach uses off-the-shelf Microsoft Kinect, which costs under \$200, to collect color images and 3D point clouds of roadway surface. This study resulted in the building of a compact-size pavement data collection system which can be installed on a car easily and collects data at highway speed. The pavement data collection system was designed to compose multiple Microsoft Kinect sensors to cover a lane width and reach to ideal scanning speed. The data collection system also includes 3-axis accelerometers to record orientations of the system and GPS to obtain location and velocity.

This study performed several road tests on local streets and freeways using the developed pavement data collection system. According to those road tests, various challenges were discovered and corresponding solutions were also identified and implemented. The main challenges that this study encountered and redesigned are summarized as following:

- ***Sunlight interference:*** Based on outdoor tests, a Kinect is easily interfered with by sunlight. A top-cover sun shade and a full-cover sun shade were designed and tested. Using the full-cover sun shade can minimize sunlight interference and obtain good depth image; however, insufficient illumination inside the full-cover sun shade causes the RGB camera to generate blurred images.
- ***Motion blur problem:*** A stroboscopic technique was used to capture slow-motion pictures. This study used this approach to solve the motion-blur problem for Kinect's color image acquisition. The low-power LED is an ambient light source, which will not interfere with the depth sensor of a Kinect, is a desirable illumination option for this application. The LED is flashed by a pulse signal generated from a microcontroller. The quality of the color images was improved when using the stroboscopic technique
- ***Rolling shutter distortion:*** Both the color camera and depth camera of a Kinect sensor consist of rolling shutter CMOS image sensors. The rolling shutter CMOS sensors provide low-noise, low-power, fast data processing, and inexpensive solution for most commercial cameras; however, this image acquisition method creates distortions when shooting moving objects. This study investigated the relation between rolling shutter distortion and speed, and used a rectification algorithm to correct distorted images.

After improving the system according to the solutions listed above, the relatively inexpensive pavement data collection system could obtain good imaging results when moving at less than 48.28 km/h (30mph) (residential speed limit in most states).

Lastly, pavement crack detection using a hybrid algorithm based on anisotropic diffusion filtering and Eigen analysis of Hessian matrix showed a promising outcome in segmenting the cracks compared to a modified bottom-hat morphological method.

6.1 Recommendations for Future Work

Based on several road tests on local streets and freeways, the recommendations to improve the relatively inexpensive vision-based pavement data collection system in the future can be summarized as following:

1. Integrate Kinect sensors with Inertial Measurement Units (IMU) to rectify rolling shutter distortion and facilitate the accurate stitching of multiple images.
2. Use a GPS/INS sensor, which is composed of an IMU, a GPS receiver, and a Kalman filtering algorithm, to provide location data with higher bandwidth, even when the GPS signal is not available.

3. Improve the data acquisition software to fuse and synchronize different sensor data sets more efficiency.
4. Improve the hybrid pavement crack detection algorithm to localize the crack pixels accurately, as it is currently thicker than the real crack width. Also, a shadow removal algorithm should be used to assist hybrid pavement crack detection algorithm in segmenting high quality cracks.

Acknowledgment

The Principal Investigators would like to acknowledge the contribution of Mr. Yulu Chen, Mr. Preetham Aghalaya Manjunatha, and Mr. Mohamed Abdelbarr, three Ph.D. students at the University of Southern California, who helped conducting the tasks and preparing this report. Discussion with members of the California Department of Transportation (Caltrans) and Fugro, Inc. are appreciated.

References

1. Ait-Aider, O., N. Andreff, J.-M. Lavest, and P. Martinet, "Exploiting Rolling Shutter Distortions for Simultaneous Object Pose and Velocity Computation Using a Single View," *Computer Vision Systems, ICVS '06, IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 35, 35, Jan. 4–7, 2006.
2. Ait-Aider, O., A. Bartoli, and N. Andreff, "Kinematics from Lines in a Single Rolling Shutter Image," *CVPR '07, IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1, 6, June 17–22, 2007.
3. Jahanshahi, M.R., F. Jazizadeh, S.F. Masri, and B. Becerik-Gerber, "Unsupervised for Autonomous Pavement-Defect Detection and Quantification Using an Inexpensive Depth Sensor," *Journal of Computing in Civil Engineering*, No. 27, No. 6, 2012, pp. 743–754.
4. Jahanshahi, M.R., S.F. Masri, C.W. Padgett, and G.S. Sukhatme, "An Innovative Methodology for Detection and Quantification of Cracks Through Incorporation of Depth Perception," *Machine Vision and Applications*, Vol. 24, No. 2, 2013, pp. 227–241.
5. Salembier, P., "Comparison of Some Morphological Segmentation Algorithms Based on Contrast Enhancement. Application to Automatic Defect Detection," In *Proceedings of the EUSIPCO-90—Fifth European Signal Processing Conference*, Barcelona, Spain, Sept. 18–21, 1990, pp. 833–836.
6. Salembier, P. and J.C. Serra, "Morphological Multiscale Image Segmentation," In *Applications in Optical Science and Engineering*, International Society for Optics and Photonics, Bellingham, Wash., 1992, pp. 620–631.