Technology Overview on Validating 3D Transverse Profile Data and Measurement of Pavement Surface Distresses

Yichang James Tsai, Ph.D., P.E. and Zhaohua Wang, Ph.D., P.E.

Georgia Institute of Technology

With the advancement of 3D line laser imaging technology (“3D technology”, hereafter for the sake of brevity), high-resolution 3D transverse profile data (“3D data”, hereafter for the sake of brevity), has gained widespread interest from researchers, industry and highway agencies and has been used for collecting pavement surface distresses. However, the study of data characteristics, e.g. data noise, and methods for validating the accuracy of data itself and the derived pavement distresses is limited in literature. A comprehensive understanding of 3D data characteristics, especially the noise impacting data accuracy and the validation of the collected pavement distress data using 3D data, is crucial for the successful implementation and utilization of 3D technology for automatic pavement distress data collection. The objective of this literature review is to summarize the existing research and to study 3D data characteristics, especially the data noise and the methods used to validate the accuracy and performance of pavement rutting and cracking measurement using 3D data. The review results will benefit researchers and transportation agencies by enabling them to identify future research needs for better implementation and utilization of 3D technology and data.

1. Introduction to 3D Transverse Profile Data

3D transverse profile data is collected using 3D line laser imaging technology. The triangulation principle, which is also commonly known as the laser triangulation, is used for data acquisition. Figure 1 illustrates the basic concept of the system diagram. In the 3D line laser system, the laser light projected is a fine laser line. A high-intensity, area-scanning camera with an image sensor such as a complementary metal oxide semiconductor (CMOS) sensor or charge-coupled device (CCD) sensor is placed at a known distance and at an oblique angle (θ) with respect to the laser line projector. The camera takes images of the laser light. Then, the deformations of the laser line on the object are analyzed to evaluate the elevation for each point with a known horizontal
position on the object. Because the laser sensors are coupled with a distance measurement instrument (DMI), the system can obtain point locations along the longitudinal direction. Consequently, a complete three-dimensional set of points of the object’s surface can be acquired. These 3D points can be used to identify pavement distresses and to measure the geometry of the distresses.

The triangulation principle has been developed for years for 3D data acquisition. Because of its simplicity, there are several similar systems that have been developed for pavement distress detection and measurement (Li, et al., 2010). However, constrained by the line width from the laser illumination (i.e. the fineness of the laser strip), the accuracy of the 3D range varies.

There are several commercially available 3D laser systems based on the laser triangulation used in the United States:

2) Pathway 3D Imaging (http://www.pathwayservices.com/3D_imaging.shtml)
3) Waylink (www.waylink.com)
4) Texas Department of Transportation (TxDOT) 3D Transverse Profiling System (Huang, et al., 2013a)
2. Triangulation and Calibration

Li et al. (2010) explained well the calculation of the triangulation process. The notations in Figure 2 followed the ones in Li’s paper (Li, et al., 2010). As shown in Figure 2, $Z_w$-O$_w$-$Y_w$ is the world coordinate system, and $Z_c$-O$_c$-$Y_c$ is the camera coordinate system. O$_c$ is the camera center, and P is the image plane. O$_p$ is the intersection of the camera optic axis Z$_c$ on the image plane. $f$ is the focal length. Y$_a$ is the point in the image coordinate system corresponding to the point P$_a$ in the world coordinate system. The angle between axis O$_c$O$_w$ and O$_c$O$_t$ is $\theta$; the angle between axis O$_c$O$_w$ and line Y$_a$P$_a$ is $\alpha$. According to the triangulation principle, the elevation, $\Delta h$, at point P$_a$ can be calculated by

$$
\alpha = \arctan\left(\frac{y_a}{f}\right)
$$

$$
h = l \cdot \tan(\theta)
$$

$$
m = l \cdot \tan(\theta - \alpha)
$$

$$
\Delta h = h - m = l \cdot (\tan(\theta) - \tan(\theta - \alpha))
$$

(1)

Where $l$ and $\theta$ are the parameters that can be determined by a calibration procedure.
The calibration procedure of 3D line laser system generally includes two parts, calibration of camera optical distortion and calibration of range measurement. For the calibration of camera optical distortion, a precisely fabricated chess board is normally used. To calibration range measurement, an object with known range difference is often used. It should be noted that for commercial products, this type of calibration is often performed by the device manufacturer and the actual calibration procedures are normally not released to the public. An end user can only verify the measurement repeatability and accuracy.

In the following sections, the literature review will focus on the evaluation of 3D data characteristics and the validation of repeatability and accuracy of 3D data using rutting measurement and crack detection.

3. Characteristics of 3D Range Data

Although there is little literature studying the fundamental characteristics of 3D data, it is crucial for utilizing 3D data to automatically extract pavement distresses. A study sponsored by the USDOT Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) and documented in Feng Li’s Ph.D. dissertation (2012) evaluated and identified major data uncertainty and outliers that were encountered during data collection and data analysis. Since data noise stems from the laser triangulation and the integrated mobile data collection system that is commonly adopted in the currently used 3D data acquisition system, corresponding validation methods need to be developed to verify the capability of a specific data acquisition system. Please note that some of the following content in this section and the following sections is to be published.

3.1 Data Uncertainty

The depth (i.e., range) measurement uncertainty (i.e. noise) is defined as the dispersion of the 3D range data from its theoretical value, which is given by the point on the corresponding fitted primitive. The standard deviation is used as a measure of the range measurement noise.

It was found that 3D data uncertainty exists when the 3D data capturing system is either stationary or moving. The following summarize the major sources of data noise in both stationary and moving modes, and the experimental tests conducted.
3.1.1 Uncertainty in Static Mode

Researchers have identified various sources of range measurement uncertainty when both the 3D line laser system and the survey object are stationary (Linares, et al., 2000). One main contributor to the uncertainty in the 3D range data is the speckle noise (Forest Collado, 2004; Dorsch, et al., 1994). The phenomenon of speckle, caused by the interference of waves of the same frequency, is inherent to the use of laser light (Wikimedia, 2012). Another potential source of the range measurement uncertainty is the brightness or color variation of the survey object’s surface (Whaite & Ferrie, 1990). Laser triangulation sensor uses Center of Gravity algorithm for sub-pixel laser line detection. Changes of surface brightness or color will create high laser intensity on bright side and push the detected laser line position to the bright side. As a result, the bright side of the surface may show data higher than its real elevation. Tests show some systems currently on the market could produce more than 6 inch spikes/noise on edges of a white lane stripe. This uncertainty has not been emphasized in the past literature, since most 3D line laser systems were applied to those surfaces with uniform intensity. However, pavement surfaces, especially the asphalt pavement surfaces, have significant brightness or color variation. Therefore, this uncertainty may become a major source of the range measurement uncertainty.

3.1.2 Uncertainty in Moving Mode

When the 3D line laser system is moving, the overall measurement uncertainty might be smaller than that in the static mode because the speckle noise caused by the interference of stationary laser light waves is reduced. Also, the survey vehicle’s vibration does not introduce additional uncertainty into the 3D range data because the profiling frequency of the 3D line laser system, which is up to 5,600 Hz (INO 2012) for the system used in Tsai’s work, is much higher than the frequency of vehicle body vibration, typically ranging from several Hz to several hundred Hz (Aoki, et al., 1998). The exposure time of the camera, which captures the laser line, is approximately 36.5 µs (INO 2012). This time is short compared to the time in which the position of the survey vehicle changes. Consequently, it can be reasonably assumed that the vehicle was not moving during the image capturing and the vehicle vibration has no impact on the quality of the 3D range data.
3.1.3 Experimental Tests

To quantify the range measurement uncertainty, experimental tests were conducted using a surface plate with known ground truth. As shown in Figure 3 (a), it is a Grade B surface plate made of granite. The tolerance of flatness for a Grade B surface plate is 0.0002 inches; that is, the difference between the lowest point and the highest point on the plate will be no more than 0.0002 inches. Because the surface plate is made of granite, its surface intensity is non-uniform, as shown in Figure 3 (b). To test the impact of material on the measurement uncertainty, both the bare surface plate and the plate covered with blue tapes, as shown in Figure 3 (c), were tested. In addition, to eliminate the impact of ambient lighting, all the tests were completed in the laboratory and in a short time period. The tests and results are presented below.

![Figure 3 Grade B Surface Plate (Li, 2012)](image)

(a) Grade B Surface Plate  
(b) Zom-in View  
(c) Birdseye View

Static Test

In the static test, the test object and the 3D line laser system were set in stationary positions. Four cross-sections on the test object were selected (locations A and B on the bare surface and locations C and D on the taped surface). Each cross-section was tested twice. In each test, the 3D line laser system scanned the object surface 1,000 times without interruption. Table 1 summarizes the static test.

A transverse profile example is the zigzag profile shown in Figure 4 (a). Since the test object is essentially flat, a straight line was obtained through robust linear regression analysis and used to
fit the transverse profile. The straight line was used as the true surface of the plate. Deviations of the range measurements from the straight fitting line were defined as the range measurement uncertainties. Figure 4 (b) shows the deviations of the range data from the straight line for the given example, which is denoted as an uncertainty profile.

Table 1 Summary of Static Tests (Li, 2012)

<table>
<thead>
<tr>
<th>Material</th>
<th>Location</th>
<th>Test ID</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stone</td>
<td>A</td>
<td>1</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td>Tape</td>
<td>C</td>
<td>1</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>1</td>
<td>Repeated for 1,000 times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Repeated for 1,000 times</td>
</tr>
</tbody>
</table>

Figure 5, Figure 6, and Figure 7 show the uncertainty profiles for four transverse profiles repeatedly collected in Test A-1, Test B-1, and Test C-1, respectively. All profiles were randomly selected. It can be found that the uncertainty profiles collected in the same test repeated each other very well. This high repeatability implies that the majority of the range measurement uncertainty is time-invariant. In addition, the uncertainty in Figure 7 ranges from -1 mm to 1 mm, which is much smaller than the range measurement uncertainty in Figure 5 and Figure 6. This is because granite is a partially translucent material into whose surface the laser may penetrate part way, which affects the accuracy of range measurements.
Figure 5: Repeatability of Range Measurement Uncertainty (Test A-1) (Li, 2012)

Figure 6: Repeatability of Range Measurement Uncertainty (Test B-1) (Li, 2012)

Figure 7: Repeatability of Range Measurement Uncertainty (Test C-1) (Li, 2012)

Figure 8 and Figure 9 compare the uncertainty profiles for data collected from the same location during different tests. The only change from one test (e.g., A-1) to another (e.g., A-2) is that the
3D line laser system was turned off after the first test and then turned on. As shown in Figure 8, the two upper profiles, i.e., uncertainty profiles from Test A-1 and Test A-2, are analogous. The bottom profile is the difference between two uncertainty profiles. Most of the difference is within +/-1 mm. Figure 9 shows the comparison of uncertainty profiles from Test C-1 and Test C-2. Similarly, two profiles follow a similar trend. The majority of the absolute differences are no more than 0.5 mm. Both figures show that the repeatability between two tests, although still high, is inferior to that between two repetitions within one test. This might be because the environment (e.g. temperature) between two tests changed slightly.

Further, the uncertainty profile collected from one location was compared with the one from another location. The comparison in Figure 10 and Figure 11 shows that the uncertainty profiles do not repeat each other. This evidence supports the idea that the range measurement uncertainty is location dependent. Even Locations C and D seem to have similar surface color and reflectivity; their uncertainty profiles are dissimilar. This might be caused by the unobservable features (e.g., texture) on the object’s surface.
Figure 10 Range Measurement Uncertainties at Locations A and B (Li, 2012)

Figure 11 Range Measurement Uncertainties at Locations C and D (Li, 2012)

Figure 12 shows the histogram of the range measurement uncertainty for Test A-1 and Test C-1. The histogram was plot based on 1,000 profiles collected in each test. Table 2 tabulates the standard deviations for all tests. The average standard deviation for the bare surface and taped surface is 0.73 mm and 0.43 mm, respectively. Therefore, based on limited tests, for surfaces made of opaque materials and with solid color, the range measurement uncertainty is 0.43 mm. This is consistent with the manufacturer’s specification for the testing device that the depth accuracy is 0.5 mm. The range measurement uncertainty is greater for surfaces made of partly-translucent materials, e.g., granite. Additional tests are needed to quantify (1) how much additional uncertainty will be introduced for surfaces made of partly-translucent materials and (2) how much additional uncertainty will be introduced for surfaces with varying colors or intensities.
Moving Test

In the moving test, range data of the test object’s surface were collected by moving the survey vehicle at a low speed. The uncertainty profiles are aligned and presented in Figure 13. It is clear that the range data collected from the bare surface contains much more uncertainty than the one from the taped surface.

Figure 14 shows the histogram of range measurement uncertainty for data from the bare surface and the taped surface. The standard deviations of data from the bare surface and the taped surface are 0.66 mm and 0.43 mm, respectively. The scale of this uncertainty is slightly smaller than the scale of the uncertainty captured in the static test. This shows that the vehicle's movement removes some of the speckle noises and, thus, reduces the range measurement uncertainty.
3.2 Missing Range Data

In some cases, a 3D line laser system could not obtain a valid range measurement, which was set at a value of -10,000 in Tsai’s tested system. Those 3D laser points whose range value equals -10,000 are, therefore, denoted as missing points. There are three possible causes for missing points:

- *The object surface gets out of the measurement range.*

The measurement range of the 3D line laser system tested in Tsai’s work is +/-125 mm from the calibrated ground surface. Any object, e.g., curb on the roadside, that is higher or deeper than
125 mm, will fall out of the region that can be captured by the CCD camera. Thus, it becomes “invisible” to the 3D line laser system, and the 3D line laser system cannot provide a valid range measurement for the object.

- **Occlusion.**

Because of the sensor configuration (the angle between the laser and the camera), some points are invisible to the camera. This phenomenon is named occlusion (Mavrinac, et al., 2010), and it is an intrinsic limitation of the 3D line laser system using laser lighting and a single camera. As shown in Figure 15, the edge part was occluded by the object above, i.e., a curved wood board. Those regions were invisible to the 3D line laser system. Thus, the range value for those regions was set as -10,000.

![Figure 15 Missing Data Points due to Occlusion (Li, 2012)](image)

- **Dim laser point.**

If the laser line looks dim on the CCD camera, it is difficult to differentiate it from the background, and, therefore, a reliable range value cannot be obtained. Ideally, the laser line will be the brightest object on the CCD image. However, because of the sunlight or the laser light reflection from the pavement surface, the background of the CCD image is not fully dark (Li, et al., 2010). When extracting the laser line, a threshold is set in the 3D line laser system to remove those background pixels. For any portion of the laser line that looks dim, i.e., the maximum intensity is lower than the predefined threshold, -10,000 will be assigned as the range value. For example, when the object surface is very dark, most energy of the input lighting may be absorbed by the surface, sufficient light cannot be captured by the camera, and, thus, missing points may be present in the 3D range data.

### 3.3 Unseemly Range Data

Besides missing data, unseemly points are also observed in the 3D range data. Unseemly points are defined as spikes abnormally higher or lower than the surrounding range data. Figure 16
gives an example. The single-point spike is around 35 mm higher than the surrounding data points, which is unlikely, part of the actual pavement surface.

![Figure 16 Example of Unseemly Point (Li, 2012)](image)

To verify the cause of the unseemly points, tests have been conducted. The following summarizes the tests and test results.

- **Effect of the Pavement Type**

Most unseemly points are found on asphalt pavement surfaces, and only a few are found on concrete pavement surfaces. Therefore, it is likely that some features that are common on the asphalt pavement surfaces are causing the unseemly points.

- **Effect of Loose Rocks**

A test at the Georgia Tech Savannah campus was conducted to verify whether unseemly points might be caused by loose rocks on the pavement surfaces. Approximately 40 rocks with different colors (e.g. black or white) and varying sizes were placed on the pavement surface, as shown in Figure 17. The size of the rocks ranged from 2 mm to more than 10 in. Then, the rocks were scanned using the 3D line laser system. The 3D range data was examined. The examination results show that none of the rocks, regardless of the size or color, caused unseemly points.
Figure 17 Effect of Loose Rocks (Li, 2012)

- **Effect of Selected Road Surface Characteristics**

Several characteristics, including small holes, white rocks, smooth rocks, and oil stains that could potentially result in unseemly points were identified through visual inspection on asphalt pavements. Figure 18 shows the selected characteristics. These characteristics were marked on the road, and data was collected using the 3D line laser system. The 3D range data was visually examined. None of the suspected characteristics was found to actually cause any of the unseemly points.

(a) Small holes    (b) White rocks    (c) Smooth stones    (d) Oil stains

Figure 18 Surface Characteristics Suspected to Cause Unseemly Points (Li, 2012)
• *Effect of Vegetation*

By examining the 3D range data, it was determined that some unseemly points, as shown in Figure 19, were caused by the vegetation that grew in the cracks between the asphalt surface and the concrete curb and gutter.

![Figure 19 Unseemly Points Caused by Vegetation (Li, 2012)](image)

• *Reflective Surface Test*

Specular components on the pavement surface may cause unseemly points. Because of the specular components, noisy reflections of the laser line may appear in the images observed by the camera. These reflections can be easily confused with the primary signal, in which case false 3D range values will result (Trucco, et al., 1994). The floor, painted with reflective materials, of the laboratory located in the ELAB building at the Georgia Tech Savannah campus was tested to confirm this.

![Figure 20 Reflective Surface (Li, 2012)](image)
Three runs of data were collected and inspected. The inspection results indicate that few unseemly points were present in the 3D range data, although a number of specular components were observed on the physical surface. Also, in some runs of data collection, unseemly points repeatedly occurred in some locations, but no unseemly points were observed in the same location for the third run. Figure 21 shows the location that had an unseemly point in one run but not in the other runs. This may be because the 3D line laser system may sample the pavement surface slightly differently in different data collection runs. The specular components, i.e., small glass beads in the painting material, were too small to be captured by the 3D line laser system in every run of data collection. Thus, unseemly points may be caused by the specular components on the pavement surface, and their occurrence may appear to be random.

![Figure 21 Results from Reflective Surface Test (Li, 2012)](image)

Because the unseemly points adversely impact the 3D range data quality, it is desirable to estimate how frequently the points appear and how much they impact the data quality. The number of unseemly points in the 3D range data was estimated by randomly inspecting 3 data files that were collected by the 3D line laser system. The number of unseemly points was determined by counting spikes that were 10 mm or higher. A conservative threshold of 10 mm was selected to count changes in the 3D range data that are, unlikely, caused by pavement surface texture. The results from this analysis are shown in Table 3. The number of unseemly points varies from file to file. However, compared to 4.16 million (= 2000 profiles * 2080 points) 3D range data points in one file, the average 90 unseemly points form only a trivial portion. Further review on hundreds of 3D range data files has confirmed this finding. Thus, the
unseemly points, although present, are only a trivial portion of the whole 3D range data, and they have limited impact on the overall data quality.

### Table 3 Amount of Unseemly Points (Li, 2012)

<table>
<thead>
<tr>
<th>File #</th>
<th>Sensor</th>
<th># of Unseemly Points</th>
<th>Sensor</th>
<th># of Unseemly Points</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Left</td>
<td>47</td>
<td>Right</td>
<td>27</td>
<td>74</td>
</tr>
<tr>
<td>2</td>
<td>Left</td>
<td>92</td>
<td>Right</td>
<td>65</td>
<td>157</td>
</tr>
<tr>
<td>3</td>
<td>Left</td>
<td>19</td>
<td>Right</td>
<td>21</td>
<td>40</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>--</td>
<td></td>
<td></td>
<td>90</td>
</tr>
</tbody>
</table>

In summary, test results show that unseemly points are possibly caused by reflective components, vegetation, and, perhaps, other features on pavement surfaces. The occurrence of unseemly points may appear somehow random since the 3D line laser system samples the pavement surface differently in different data collection runs. The number of unseemly points is limited, and their impact on the data quality and the repeatability of derived rutting information is expected to be minimal. Additional tests should be conducted to obtain more conclusive results.

### 3.4 Location Uncertainty

An important aspect of the quality of the 3D range data is the location uncertainty of the 3D range data points in the 2D plane. Figure 23 illustrates how vehicle movement, including yaw, pitch, and roll, changes the location of the laser line from its expected location on the pavement surface. It may shift the laser line from the expected location to a partially or totally new location, and it may pick up a paritially new transverse profile, as shown in Figure 23 (a), or a totally new transverse profile, as shown in Figure 23 (b), (c), and (d). Consequently, the 3D line laser system does not take equal-spacing samples of the roadway surface, as described in Figure 22(a), because of the vehicle's movement. This is the uncertainty in the location (X and Y coordinates) of the 3D range data. An extreme case of the location uncertainty due to vehicle movement is shown in Figure 22 (b). The location uncertainty in the driving direction is ignorable compared to the scale of smallest pavement inspection unit, e.g., 0.1 mile. However, the location uncertainty in the transverse direction may become critical when the 3D line laser system is used to measure the rut depth.
3.5 Other Cases

3.5.1 Bump Test

Bumps, such as "rumble strips" or "speed breakers" are usually used as a safety counter measurement on city streets to slow moving traffic. The presence of bumps, however, will adversely impact the quality of 3D range data collected by the sensing vehicle. To assess this impact, bump tests have been conducted on the Osprey Point Circle near Savannah, Georgia.
Figure 24 shows the typical 3D range data collected when there is a bump. As observed in the data, when the front wheel crosses a bump, it has ignorable impact on the 3D range data quality; however, when the rear wheel crosses the bump, it leaves a dark band before the bump in the 3D range data, as shown in Figure 24. It looks like that the road surface before the bump was depressed, which, in reality, it is not. Missing points are also present in the 3D range data, since the portion of the bump was out of the measurement range.

![Figure 24 Data Collected When Passing a Bump (Li, 2012)](image)

Figure 24 Data Collected When Passing a Bump (Li, 2012)

Figure 25 explains the cause of the dark band. When the rear wheels cross the bump, they are pushed upwards by the bump, as is the 3D line laser system. Thus, the distance between the laser emit point and the road surface is elongated and, thereby, the measured range data increases. In other words, the increase of the range data is caused by the position change of the laser profiler, not a depression in the road surface. However, the current 3D line laser system is not coupled with a sensor unit, such as an inertial measurement unit (IMU), to detect its own position change. Any change in the range data is simply assumed to be caused by features on the road surface. Consequently, the increase in the 3D range data because of lifted profilers is “wrongly” recognized as a depression in the road surface. In the future, this can be improved by integrating an IMU and the 3D line laser system.
3.5.1 Data Collection on New Pavements

The 3D range data collected from newly-paved open graded friction course (OGFC) contain a significant number of missing points and unseemly points, since its surface is dark and specular. Figure 26 shows an example. The black points observed in Figure 26 are missing points and unseemly points that are abnormally lower than surrounding range data, and the white points are unseemly points that are abnormally higher than surrounding range data. A 3D transverse profile highlighted with a blue straight line in Figure 26 (a) is shown in Figure 26 (b). As shown in Figure 26 (b), the 3D transverse profile, which is the blue one, contains a number of spikes that are unlikely part of the actual pavement surface. The red line is the smoothed transverse profile obtained after applying a low-pass filter to the blue profile. In summary, the quality of the 3D range data collected from newly-paved OGFC using a 3D line laser system is inferior to the quality of 3D range data collected from aged pavements.
4. Data Quality and Data Utilization

Data quality is critical for data utilization. The above section discussed all the major data characteristics, which should be considered in data processing in order to enhance data quality. More importantly, data quality should be controlled before each field data collection is started. After all, post processing can only deal with a certain type of data uncertainty. The following two sub-sections will introduce a calibration/validation board developed by Tsai et al. and the major types of distresses that can be collected using 3D laser data.

4.1 Calibration/Validation Board

Calibration/validation is a critical part of 3D transverse profile data collection and should be performed before each field data collection to ensure the desirable data accuracy. Through the research project sponsored by USDOT/OST-R, Tsai et al. developed a calibration/validation board as shown in Figure 27. The board is used to quantitatively evaluate the accuracy of 3D transverse profile data, especially in depth direction, which consists of three known steps of 0.5 mm, 1.0 mm, and 2.0 mm and 27 grooves with known dimensions shown in Figure 27.

3D transverse profile data will be collected using the developed calibration/validation board before data collection to ensure 3D data accuracy and quality. The depth values of the steps and grooves will be measured using the 3D data and compared with the ground truth. If the difference is greater than a specified value (that was not standardized yet), a calibration/validation process, including sensor configuration checking and system setup checking, such as checking the tire pressure and the alignment of all sensors.

A separate study using this validation board was conducted to evaluate the accuracy of the micro-milled surface texture measurement by means of ridge-to-valley (RVD) value (Tsai, et al.,
2014) with 3D transverse profile data. In addition, another study, using the validation board, was conducted to evaluate the accuracy of concrete joint faulting measurement accuracy (Tsai, et al., 2012b) with 3D transverse profile data.

TxDOT developed its own 3D scanning system, called VTexture, to measure pavement textures. 12 test sections with different pavement type were selected for validation (Huang, et al. 2013b). In comparison with the mean profile depths (MPD) measured by sand patch method and circular texture meter (CTM), they found the correlation is very high. In this study, Huang, et al. (2013b) also validated the MPD measurement accuracy with the change of 3D scanning system exposure time under different travel speeds. The ground truth textures were acquired from several precisely machined test pads as shown in Figure 28. These test pads were placed in a line in field. And, the 3D laser data was captured by the vehicle-mounted VTexture at different speeds

---

**Figure 27 Calibration/Validation Board Developed by Tsai, et al.**

TxDOT developed its own 3D scanning system, called VTexture, to measure pavement textures. 12 test sections with different pavement type were selected for validation (Huang, et al. 2013b). In comparison with the mean profile depths (MPD) measured by sand patch method and circular texture meter (CTM), they found the correlation is very high. In this study, Huang, et al. (2013b) also validated the MPD measurement accuracy with the change of 3D scanning system exposure time under different travel speeds. The ground truth textures were acquired from several precisely machined test pads as shown in Figure 28. These test pads were placed in a line in field. And, the 3D laser data was captured by the vehicle-mounted VTexture at different speeds.
(0, 10, 20, 30, 40, 50, 60, and 70 mph) with different camera exposure times (20, 30, 40, 50, 100, 150, 200, and 250 µs). Figure 29 shows the MPD measurement error with different exposure times at different travel speeds. It can be seen that long sensor exposure time is the main reason causing texture data error with vehicle speed, especially on small surface features. The maximum exposure time of 20 µs is recommended to maintain MPD errors under 10% over the full speed range.

![Figure 28 Calibration/Validation Board Developed by Huang, et al. (2013b)](image)

<table>
<thead>
<tr>
<th>Travel Speed</th>
<th>2.5 mm block</th>
<th>5.5 mm block</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exposure time in µs</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>3.0</td>
<td>2.4</td>
</tr>
<tr>
<td>20</td>
<td>4.3</td>
<td>4.0</td>
</tr>
<tr>
<td>30</td>
<td>6.6</td>
<td>4.2</td>
</tr>
<tr>
<td>40</td>
<td>9.6</td>
<td>7.1</td>
</tr>
<tr>
<td>50</td>
<td>9.6</td>
<td>11.4</td>
</tr>
<tr>
<td>60</td>
<td>9.8</td>
<td>10.2</td>
</tr>
<tr>
<td>70</td>
<td>9.2</td>
<td>12.6</td>
</tr>
</tbody>
</table>

![Figure 29 MPD Percentage Errors with Different Exposure Time at Different Travel Speed (Huang, et al. 2013b)](image)

### 4.2 Using 3D Transverse Profile Data

Continuous 3D transverse profile data captures pavement surface geometry in three dimensions, which can be used to collect all the major types of pavement surface distresses. Pavements,
including flexible pavements and rigid pavements, have three major types of surface distresses, such as deformation (caused by pavement material or foundation), cracking, and loss of aggregates that are suitable for automatic data collection using 3D laser data. Among these three types of pavement distresses, rutting, cracking, and raveling are common for flexible pavements; cracking, faulting, and spalling are common for rigid pavements. Nowadays, more and more researchers and vendors in industry are using 3D data to collect these types of pavement distresses. Thus, the validation of pavement distress data collection accuracy is very useful for highway agencies to justify their adoption of 3D laser technology.

In the following sections, two types of pavement distresses, rutting and cracking, are selected for this literature review due to their importance to highway agencies and their popularity in literature. However, it doesn’t mean 3D laser data cannot be used for collecting other types of pavement distresses.

5. Validation of Rut Depth Measurement

Rut depth measurement is one of the most suitable applications using 3D transverse profile data. Validating its measurement accuracy can directly check its applicability.

5.1 In-Lab Test

Tsai et al. (2013) performed a comprehensive laboratory test to validate the rut depth measurement accuracy using 3D transverse profile data. The laboratory test was performed under a controlled environment to exclude the effect of external factors, such as vehicle vibration. The rut depth is calculated using a simulated 1.8 m straight edge method, which is suggested in the ASTM 1703 Standard (2010) and commonly adopted by researchers (Laurent, et al., 1997; Li, et al., 2009, 2010). Two rut depth calculation algorithms based on the principles of the simulated straightedge method have been applied. One comes with the 3D line laser, and the other one has been developed by the Georgia Tech research team. Both algorithms are presented and compared in this section. The following are the procedures:

1) Identify the highest points at both ends of a 3D transverse profile that are the standing points of the straightedge;
2) Connect the two standing points;
3) The maximum distance between the simulated straightedge and the profile is the rut depth, as shown in Figure 30.

![Figure 30 1.8 m Straightedge Method (Tsai and Wang, 2013)](image)

In the laboratory tests performed by Tsai et al. (2013), 11 profiles were fabricated using a curved wood board and a curved metal bar to simulate ruts at different severity levels that are defined by the Oregon Department of Transportation (ODOT) (Figure 31). In ODOT (2010), rut depth for the low severity rutting is 1/4 in. (6.35mm) to 1/2 in. (12.7mm); the medium severity is from 1/2 in. to 3/4 in. (19.1mm); and the high severity is greater than 3/4 in. As shown in Figure 31 (a) and (b), 10 profiles were marked with blue tape on the wood board. The rut depths of those profiles varied from several millimeters to several centimeters. The curved metal bar was used to simulate a rut of the high severity level.

![Figure 31 Simulated Rutting: (a) Wood Board; (b) Metal Bar (Tsai and Wang, 2013)](image)

The ground truth was established by using the straightedge method, and the data collection procedure followed the standard specified in ASTM E1703 (2010). As shown in Figure 32 (a), a steel angle bar was used as the straightedge. The rut depth was measured using a vernier caliper with a precision of 0.02 mm. During the measurement, the vernier caliper is set to be perpendicular to the steel bar. To identify the maximal distance between the steel bar and the wood board surface, sufficient measurements were made along the steel bar. To reduce the
measurement error, the measurement for each profile was repeated 3 times. The average rut depth of these three runs was considered as the ground truth.

Figure 32 (b) shows the setup of laser sensor. One sensor was used to cover about 2 meters. The infrared camera was used to observe the invisible laser line. The measurement procedure for each profile was repeated twice. During each measurement procedure, the wood board or the metal bar was placed under the laser profiling unit, and its position was fine-tuned until the laser line was right on the marked profile. After that, 2,000 repetitive data profiles were collected using the 3D line laser. For testing the 11 ruts, 44,000 (=11×2×2,000) profiles were obtained.

![Rutting Measurement](image1)

![3D Line Laser System](image2)

**Figure 32 Laboratory Test (Tsai and Wang, 2013)**

Table 4 shows the rut depth measurement results on the 11 simulated rutting profiles. The average manual measurements vary from 7.9 mm to 43.4 mm and cover low to high severity levels. They are considered as the ground truth. Two runs of 3D line laser measurements were performed. The difference between these two runs ranges from 0.1 mm to 1.3 mm, which is comparable to the manual measurement error. The difference between the 3D-line-laser-measured results and the ground truth varies from -0.4 mm to 0.7 mm, which is less than 1 mm.

### 5.2 Field Test

Tsai and Wang (2013) performed a field test to validate the rut depth measurement accuracy using 3D transverse profiles. Two roadway sections were selected in Pooler, Georgia. As shown in Figure 33 (a), a 725 m roadway section was chosen on Benton Boulevard. A 45 m roadway section was selected on Towne Center Ct., shown in Figure 33 (b). There were six test transverse profiles marked with paint, which can be seen from the laser intensity data, on the Benton Blvd.
test section. On the Towne Center Ct. test section, 4 test profiles were marked. These test profiles are non-uniformly distributed over the test section. To establish the rut depth ground truth of those 10 test profiles, a 1.8m straightedge method was performed as shown in Figure 33 (c). The same measurement procedure was followed as that in the laboratory test, which was repeated 3 times for each transverse profile. The average rut depth for each transverse profile was considered as the ground truth.

**Table 4 Laboratory Testing Results (Tsai and Wang, 2013)**

<table>
<thead>
<tr>
<th>Profile #</th>
<th>Severity Level</th>
<th>Ground Truth</th>
<th>3D Line Laser Measured</th>
<th>Difference to Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>8.0</td>
<td>8.3, 7.1</td>
<td>1.2, 7.7, 0.3</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>7.9</td>
<td>8.2, 8.0</td>
<td>0.2, 8.1, -0.2</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>7.9</td>
<td>6.8, 7.6</td>
<td>0.8, 7.2, 0.7</td>
</tr>
<tr>
<td>4</td>
<td>Medium</td>
<td>13.2</td>
<td>13.2, 13.1</td>
<td>0.1, 13.2, 0.0</td>
</tr>
<tr>
<td>5</td>
<td>Low</td>
<td>12.3</td>
<td>12.3, 11.5</td>
<td>0.8, 11.9, 0.4</td>
</tr>
<tr>
<td>6</td>
<td>Medium</td>
<td>14.2</td>
<td>13.8, 14.0</td>
<td>0.2, 13.9, 0.3</td>
</tr>
<tr>
<td>7</td>
<td>Medium</td>
<td>15.5</td>
<td>15.0, 14.8</td>
<td>0.2, 14.9, 0.6</td>
</tr>
<tr>
<td>8</td>
<td>Medium</td>
<td>16.2</td>
<td>15.4, 16.7</td>
<td>1.3, 16.1, 0.1</td>
</tr>
<tr>
<td>9</td>
<td>Medium</td>
<td>17.5</td>
<td>17.6, 17.1</td>
<td>0.5, 17.4, 0.1</td>
</tr>
<tr>
<td>10</td>
<td>Medium</td>
<td>10.0</td>
<td>11.0, 9.7</td>
<td>1.3, 10.4, -0.4</td>
</tr>
<tr>
<td>11</td>
<td>High</td>
<td>43.4</td>
<td>43.2</td>
<td>--, 43.2, 0.2</td>
</tr>
</tbody>
</table>

--- Invalid testing data

The 3D line lasers were mounted on a van as shown in Figure 33 (d). The vehicle made 3 runs on both test sections and collected 3D continuous transverse profile data. The mean vehicle speeds for the first test section were 37.8 mph, 34.4 mph, and 32.8 mph; they were 13.5 mph, 17.8 mph, and 15.3 mph for the second test section. Because the road sections were chosen on local roads to facilitate the manual measurement, no highway speed higher than 60 mph was tested in this study.

The field test results on the local roads are summarized in Table 5. Ten manually marked profiles on the test roadway sections were examined. The manually measured rut depths, which are considered as the ground truth, vary from 6.4 mm to 21.1 mm. They correspond to the low-to-high severity levels of rutting. Compared to the average of 3D-line-laser-measured results,
the difference varies from -1.0 mm to 2.3 mm, which is apparently higher than the one in the well-controlled laboratory test. Several factors could contribute to this. First, for a profile-based comparison, it is very critical to make sure that the location of each extracted profile from 3D transverse profile data is exactly the same as the manually marked and measured one. In the harsh field testing environment, it is very difficult to make this happen. Second, unlike in the well-controlled laboratory test, vehicle wandering is inevitable in a field test, which will impact the rut depth measurement. Besides, from the comparison result, a trend shows that the ground truth is greater than the 3D-line-laser-measured result except for Profile #3. There is an approximate 1.5 mm measurement bias for the 3D line laser as compared to the ground truth. In real applications, for the network level rutting survey, profile-based rutting data is normally aggregated for every fixed interval, say 10 ft., which will help reduce the random measurement error. In this study performed by Tsai et al., only profile-based testing was conducted.

A separate study on evaluating the rut depth measurement of laser bars with point laser measurements (e.g. 3-point, 5 points, 13 points, etc.) was conducted using 3D transverse profile data with more than 4000 points in one transverse profile (Tsai, et. al., 2015). Results show that
the 3-point laser measurements currently used by state DOTs could, potentially, result in more than 64% rut depth measurement error, which could be a problem in pavement condition evaluation.

Table 5 Field Testing Results (Tsai and Wang, 2013)

<table>
<thead>
<tr>
<th>Profile #</th>
<th>Severity Level</th>
<th>Ground Truth</th>
<th>3D Line Laser Measured</th>
<th>Difference to Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Medium</td>
<td>14.5</td>
<td>12.1</td>
<td>14.0</td>
</tr>
<tr>
<td>2</td>
<td>Medium</td>
<td>15.8</td>
<td>13.4</td>
<td>14.6</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>9.6</td>
<td>10.7</td>
<td>10.8</td>
</tr>
<tr>
<td>4</td>
<td>Medium</td>
<td>14.2</td>
<td>12.9</td>
<td>12.1</td>
</tr>
<tr>
<td>5</td>
<td>Low</td>
<td>8.5</td>
<td>6.0</td>
<td>6.7</td>
</tr>
<tr>
<td>6</td>
<td>Low</td>
<td>9.5</td>
<td>7.3</td>
<td>7.2</td>
</tr>
<tr>
<td>7</td>
<td>Low</td>
<td>7.8</td>
<td>5.9</td>
<td>6.0</td>
</tr>
<tr>
<td>8</td>
<td>Low</td>
<td>9.4</td>
<td>7.2</td>
<td>7.1</td>
</tr>
<tr>
<td>9</td>
<td>High</td>
<td>21.1</td>
<td>19.8</td>
<td>20.8</td>
</tr>
<tr>
<td>10</td>
<td>Low</td>
<td>6.4</td>
<td>5.7</td>
<td>4.7</td>
</tr>
</tbody>
</table>

TxDOT also performed a large-scale field test to validate the rut depth measurement accuracy using its rut measurement system called VRUT (Huang, et al., 2013a). There were 6,000 manual rut data points from different pavement test sections at 3 m (10-ft) intervals. At each manual rut data point, a straight bar and a rut depth measurement wedge/tool with a 1.6 mm increment at each step (for 16 steps), as shown in Figure 34, were used to measure the rut depth.

Figure 34 Manual Rut Measurement: (a) Rut Block Measurement Wedge/Tool with 1.6-mm (0.0625-in.) Steps; (b) Rut Depth Reading with Straightedge (Huang, et. al., 2013)
To conduct a repeatability test, three runs of 3D transverse profile data were collected on road FM165 in Texas. Figure 35 shows strong linear relationships between these runs and indicates good repeatability.

![Figure 35 Repeatability Tests on FM165 (Huang, et. al., 2013)](image)

Other than the repeatability test, the manual results were also compared with the VRUT measurements to validate the rut depth measurement accuracy. TxDOT has done tests on both smooth asphalt pavement roads and rough major roads. It was believed that smooth pavements introduced less aggregate noise. Figure 36 shows the comparison on a smooth asphalt concrete pavement road (Manda Road, a county road). Because there is less aggregate noise, the VRUT rut results show very high correlation to manual rut data. On the other hand, the correlation shown in Figure 37 that was tested on US183 and US290 in Texas is not as good as the one on the smoother county road.

![Figure 36 Manual and VRUT Comparison on Smooth Pavement (Huang, et. al., 2013)](image)
Erskine validated the transverse profiles captured by the Laser Rut Measurement System (LRMS) against the ones measured by the Transverse Profile Beam (TPB), and the derived rut depth using LRMS against the ones derived by both TPB and the Greenwood Laser Profilometer (GLP) (Erskine, 2012). Six sites with different surface types, geometry and age were selected. The rut depth ranges from less than 5 mm to greater than 20 mm; and the texture characteristics ranges from less than 0.5 mm to greater than 2 mm. Figure 38 shows an example transverse profile comparison between LRMS and TPB, in which the $r^2$ of the two profiles is about 99%. Erskine compared the derived rut depth using the transverse profiles captured by LRMS and TPB. In this comparison, the profiles were captured by LRMS statically. The results showed good match with the absolute difference less than 2 mm. Another comparison was also made by Erskine among LRMS, down-sampled LRMS, and GLP. The down-sampled LRMS simulated GLP with 13 laser sensors using a Multi-point Transverse Profile (MPTP) filtering algorithm. Figure 39 shows the comparison of calculated rut depth at 6 test sites. It can be seen that the rut depth decreased when fewer points (13 points) were used for calculation (see the green line and the blue line). The rut depth calculated by using the down-sampled LRMS profiles (i.e., after MPTP filtering) are consistent with the GLP results with the absolute difference less than 1.2 mm (see the blue line and the brown line).
Figure 38 Comparison of Transverse Profiles Captured by LRMS and TPB (Erskine, 2012)

Figure 39 Rut Depth Comparison of LRMS and GLP (Left Wheel Path) (Erskine, 2012)

5.3 Rut Depth Measurement Issues
Several issues were identified during the tests conducted by Tsai and Wang (2013) as follows:

- **Issues of vehicle wandering**

An example is given in Figure 40. Figure 40 (a) is the 3D laser data collected in a normal driving. Figure 40 (b) is the data collected when the vehicle wandered. Because of the vehicle
wandering, the left lane marking is missing, and the right lane marking is almost in the middle of the right sensor data, as shown in Figure 40 (b). For the road section in Figure 40 (b), both rut depths for the right and left wheelpaths can be underestimated when they are calculated using the half-lane profiles.

![Figure 40 Issue of Vehicle Wandering (Tsai and Wang, 2013)](image)

- **Issue of half-lane rut depth calculation**

Currently, rut depths are calculated for half-lanes in Tsai’s study. An assumption is made that the transverse profile is symmetrical to the center line. However, this assumption may not hold, and the rut depth for some cases may be underestimated. Figure 41 shows such an example. Figure 41 (a) is the half-lane profile collected by the left sensor, and Figure 41 (b) is the one by the right sensor. If pinning these two half-lane profiles, a full-lane profile can be obtained. As shown in Figure 41, the hump in the middle was captured by the left sensor, not the right sensor. Therefore, the rut depth for the right wheelpath can be underestimated if it is calculated using the right half-lane profile only.

![Figure 41 Issue of Half-lane Rut Depth Calculation (Tsai and Wang, 2013)](image)

- **Issue of lane marking detection**

When using the 3D line laser, the lane marking detection is performed on fixed-length road sections. Each of such road section is 5 m long for this study, and only the portion of transverse
profiles between the lane markings is used for rut depth calculation. Thus, it is crucial to accurately detect lane markings for rut depth calculation. However, current 3D line laser software is not robust enough to detect the lane marking correctly for all road sections. An example is shown below.

Figure 42 shows the lane marking detection results for three adjacent road sections, S1, S2, and S3. The purple straight lines are the detected lane markings. Compared to S1 and S3, S2’s detected lane markings shift to the right significantly. This shift causes abrupt changes of rut depth and rut width. As shown in Figure 43, both abrupt changes occur at boundaries between adjacent sections.
• **Issue of rut depth calculation in the presence of other distresses (e.g. cracking) or objects**

Based on Tsai’s study (Tsai & Wang, 2013), rut depth calculation accuracy is hard to achieve when rutting is accompanied by transverse cracks (especially for cracks wider than 5mm), potholes and patches, raised pavement markings, rail tracks, and other objects (e.g. tree branches on the road).

### 6. Validation of Crack Detection

Continuous transverse profile data is often used to extract pavement cracks using automatic crack detection algorithms. The following section reviews the performance measure and various tests performed by researchers.

#### 6.1 Crack Detection Performance Measure

To validate the crack detection accuracy (or performance) by both sensing data and algorithms, various performance measures were developed by researchers. Nazef et al. (2006) conducted a qualitative evaluation of pavement distress detection algorithms under different lighting conditions, but the performance of segmentation algorithms was not analyzed quantitatively. Others, like Koutsopoulos (1993), Wang (2007a), and Wang (2007b), did a comparison of segmentation algorithms to show the superiority of their particular algorithms, but the evaluations are, again, qualitative. Huang (2006) and Zhou (2006) measured the performance of their algorithms by devising a scoring criterion based on statistical correlation. Mean square error is a metric that is extensively used in image comparison studies. Another metric, called the Mahalanobis distance (Gonzalez & Woods, 2002), has also been described in literature. All the above evaluation methods use the entire image data for image comparison and do not target the crack regions specifically. This can obscure the results because crack pixels are typically only a small percentage of the total number of image pixels, and these scoring measures are not specifically sensitive to crack information. In addition, information about crack locations is not used in these evaluation methods. This may lead to the erroneous conclusion that two segmented crack images having different crack locations are the same simply because they have the same number of crack pixels in an image.
Different quantification methods are, also, used in medical imaging and machine vision. These methods include Receiver Operator Characteristic (ROC) (Tagashira, et al., 2008; Song, et al., 2007; and Kerekes, 2008) and Hausdorff distance (Beauchemin, et al., 1998; Wang, 2002). Hausdorff distance enables the user to measure the distance between objects of different sizes. This can be helpful in the case of cracks, as the number of crack pixels in the ground truth image can be different from the crack pixels in the segmented images.

Lee and Kim (2006) have done a comprehensive study to analyze the error during the establishment of ground truth for evaluating the accuracy of automatic crack detection algorithms. Three individuals were selected to conduct both manual field survey and manual image-based data collection on a 1,500-ft test section. Three types of cracks, longitudinal crack, transverse crack, and block crack, were surveyed. By conducting repeatability analysis using the total crack length (for longitudinal crack and transverse crack) and total area (for block crack), Lee and Kim (2006) concluded that the field survey result should not be used as ground truth due to its poor repeatability. The image-based data collection is more suitable for establishing ground truth. Based on Lee and Kim’s results, Kaul and Tsai (Kaul, et al., 2010; Tsai, et al., 2010) used digital pavement images to establish the ground truth for their study on quantitatively evaluating crack detection performance, which will be discussed as follows.

The most recent progress was made by Kaul and Tsai by proposing a new quantification method based on a buffered Hausdorff distance metric (Kaul, et al., 2010; Tsai, et al., 2010). The buffered Hausdorff distance method is developed to provide researchers and transportation agencies a quantitative method to evaluate the performance of different crack detection algorithms and applications. The proposed buffered distance method incorporates the strengths of both mean square error and Hausdorff distance by modifying the Hausdorff distance metric. The Hausdorff distance is among the most popular distance measures; it measures the distance between two curves and is a metric. For any two sets of points \( A = a_1, a_2, \ldots, a_n \) and \( B = b_1, b_2, \ldots, b_m \),

\[
H(A, B) = \max(h(A, B), h(B, A))
\]  

Where
\[ h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \]  

(3)

\( h(A, B) \) is the greatest of all the small distances from points of A to B and \( h(B, A) \) is the greatest of all the small distances from points of B to A. Figure 44 effectively illustrates this distance measurement.

\[ (,)|h_{AB}|\]  

Figure 44 Illustration of Hausdorff Distance (Kaul, et al., 2010)

The value of Hausdorff distance is large, even if one crack pixel in the segmented image is far from the ground truth image crack pixels. Seeing this as a limitation of the Hausdorff distance metric, a new metric was developed that does not suffer from the shortcoming of the Hausdorff distance. The intuitive development of this measure is described next. A better distance measure than the Hausdorff distance is the modified Hausdorff distance given by \( MH(A, B) \):

\[ MH(A, B) = \max(h_1(A, B), h_1(B, A)) \]  

(4)

Where

\[ h_1(A, B) = \frac{1}{m} \sum_{a \in A} \min_{b \in B} \|a - b\| \]  

(5)

Once a crack pixel in the automatically segmented image falls substantially away from the closest pixel in the ground truth image, it no longer makes sense to heavily penalize this distance. Wrong detections beyond a certain distance should be penalized equally. This leads to a new distance measure: the buffered Hausdorff distance measure given by \( BH(A, B) \).

\[ BH(A, B) = \max(h_2(A, B), h_2(B, A)) \]  

(6)

Where
Here, \( sat \) indicates that when the distance of the crack pixel to the closest crack pixel in the other image exceeds a saturation value \( L \), we use a constant value of \( L \) for the distance. The proposed method was compared with four other possible quantification methods (mean square error, statistical correlation, Receiving Operating Characteristic, and Hausdorff distance) and demonstrated its superior capability to distinguish the performance of different segmentation algorithms. Again, the developed method is crucial for researchers and transportation agencies to evaluate and validate the performance of the various crack detection applications provided by various service providers.

6.2 Crack Detection

Different 3D laser devices have different resolutions that determine their capability for crack detection. Finer cracks can only be detected by using the 3D transverse profile data with better resolution. In addition, in comparison with naturally lit digital images, 3D transverse profile data are not affected by changing lighting conditions because the data consists of pavement surface elevation changes instead of different intensities.

Tsai et al. (2012a) performed controlled laboratory tests on simulated cracks with known crack widths and depths. The objective was to assess the capability of the 3D transverse profile data to detect cracks of different widths under different lighting conditions. Four crack widths (1 mm, 2 mm, 3 mm, and 5 mm) under two extreme lighting conditions (daytime and nighttime) were tested. The crack depth was about 19 mm. The crack was simulated by a controlled gap between two solid wood boards as shown in Figure 45. The width of the gap was measured before and after the test using a caliper.
The simulated crack was placed on a flat floor and the operator drove the sensing vehicle to collect the 3D transverse profile data from two wood boards as shown in Figure 46. The tests were conducted during daytime and nighttime. A dynamic optimization method was employed to detect the simulated cracks (Alekseychuk, 2006). Meanwhile, the ground truth cracks was manually digitized and extracted from the 3D transverse profile data. The aforementioned buffered Hausdorff distance scoring method was applied to quantitatively assess the performance of the crack segmentation results by comparing the segmented outcomes with the ground truth. The test results are summarized in Table 6 and Figure 47. Cracks with widths of 1mm, as shown in Figure 47, are detected partially. The scores are approximately 64. For cracks with widths equal to or greater than 2mm, the scores are much better, about 93. Daytime and nighttime tests result in similar scores.

### Table 6 Scores for the Controlled Tests (Tsai, et. al., 2012a)

<table>
<thead>
<tr>
<th>Score</th>
<th>Crack width</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1mm</td>
<td>2mm</td>
<td>3mm</td>
<td>5mm</td>
</tr>
<tr>
<td>Day-time</td>
<td>63.9</td>
<td>93.6</td>
<td>93.1</td>
<td>93.3</td>
</tr>
<tr>
<td>Night-time</td>
<td>64.1</td>
<td>93.4</td>
<td>93.0</td>
<td>93.1</td>
</tr>
</tbody>
</table>
Figure 47 Crack Segmentation Results on Simulated Cracks: the Left is Raw 3D Data, and the Right is Crack-Segmented Results (Tsai, et. al., 2012a)

Other than the controlled test, two field tests on actual roadways were also conducted by Tsai et al. (2012a) The first field tests evaluated the capability of the 3D transverse profile data to detect cracks under low-intensity contrast conditions. The second field test evaluated the capability to detect cracks under different lighting conditions, including nighttime, daytime with shadow, and daytime without shadow.

Figure 48 (a) shows a roadway image with low-intensity contrast between a crack (approximately 1mm to 6mm wide) and its pavement background. The low-intensity contrast makes the crack difficult to detect, even with the human eye, on an intensity-based digital image. However, the same image collected by 3D line laser device shows a more distinct contrast between the crack and the pavement background. This is illustrated by Figure 48 (b) and (d),
collected during the day and at night, respectively, and Figure 48 (c) and (e), which represent the corresponding crack segmentation results. The high scores from this first test, 98.3 for daytime and 98.0 for nighttime, demonstrate the advantage of using 3D transverse profile data to detect cracks under low intensity contrast conditions.

![Figure 48 Test Results on Crack with Low-Intensity Contrast: (a) Roadway Image; (b) 3D Laser Data Collected during the Daytime; (c) Crack Segmentation Result (Daytime; Score = 98.3); (d) 3D Laser Data Collected during the Nighttime; (e) Crack Segmentation Result (Nighttime; Score = 98.0) (Tsai, et al., 2012a)](image)

The second field test was conducted on SR 80 in Savannah, Georgia, to evaluate the consistency of using the 3D line laser system to detect cracks under three different lighting conditions: nighttime, daytime with shadows, and daytime without shadows. Eleven test segments, including ten longitudinal cracks (cracks A to J) and a transverse crack (crack T), were labeled in the field. Examples of the three lighting conditions are shown in Figure 49. All eleven crack segments were analyzed using the dynamic optimization-based crack segmentation algorithm. Figure 50 shows an example of 3D raw data collected under three lighting conditions and the corresponding crack segmentation results for the crack J. Each sub-figure contains the 3D raw data on the left and the segmented crack on the right. Visual observation shows that the crack can be clearly captured by the 3D laser system and segmented well using the dynamic optimization-based method that was evaluated as one of the best crack detection algorithms (Tsai, et al., 2010). Then, the detected cracks were compared with the ground truth using the buffered Hausdorff
scoring method to obtain the score; the ground truth was established by observing the cracks in
the field and digitizing them manually onto the crack images. The scores for each 3D raw data
image captured under three lighting conditions are listed in Table 7. As observed, the three
scores for each crack are very close to each other. The score range was calculated for each crack
and is shown in Table 7. The average score difference for those eleven cracks is 1.9. This
difference is very small and may easily be caused by a tiny deviation, e.g. a 2-pixel deviation,
from the actual crack pixels when digitizing the ground truth. Therefore, the preliminary results
demonstrate that the proposed 3D laser system can perform consistently under different lighting
conditions in the field.

![Examples of Three Lighting Conditions](image)

**Figure 49 Examples of Three Lighting Conditions (Tsai, et. al., 2012a)**

![3D Laser Data and Corresponding Crack Segmentation Results](image)

**Figure 50 3D Laser Data and Corresponding Crack Segmentation Results on the Crack J for Three Lighting Conditions (Tsai, et. al., 2012a)**

Ouyang et al. (2013) used the laser scanning system developed at University of Texas at Austin
to validate the crack detection algorithms. Three series of field tests were conducted on a 240 m,
two-way road on campus. The first test was to evaluate the repeatability of multiple runs at the
same driving speed (30 km/hr; see Table 8); the second test was to test the repeatability at
different driving speeds (25, 30, 40, and 50 km/hr; see Table 9); and the third one was to
evaluate the repeatability under different weather conditions (cloudy and sunny; see Table 10).
Unlike the scoring method used by Tsai et al. (2010), Ouyang et al. used the total detected crack length in longitudinal and transverse directions to evaluate the repeatability.

**Table 7 Scores for the Second Field Tests (Tsai, et. al., 2012a)**

<table>
<thead>
<tr>
<th>Crack Name</th>
<th>Score Night-time</th>
<th>Score Day-time with Shadow</th>
<th>Score Day-time no Shadow</th>
<th>Score Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>95.8</td>
<td>97.4</td>
<td>97.2</td>
<td>1.6</td>
</tr>
<tr>
<td>B</td>
<td>95.5</td>
<td>96.1</td>
<td>95.4</td>
<td>0.7</td>
</tr>
<tr>
<td>C</td>
<td>93.6</td>
<td>96.8</td>
<td>97.2</td>
<td>3.6</td>
</tr>
<tr>
<td>D</td>
<td>95.0</td>
<td>97.2</td>
<td>96.9</td>
<td>2.2</td>
</tr>
<tr>
<td>E</td>
<td>96.5</td>
<td>97.8</td>
<td>97.3</td>
<td>1.3</td>
</tr>
<tr>
<td>F</td>
<td>96.5</td>
<td>98.0</td>
<td>97.5</td>
<td>1.5</td>
</tr>
<tr>
<td>G</td>
<td>95.1</td>
<td>97.7</td>
<td>97.5</td>
<td>2.6</td>
</tr>
<tr>
<td>H</td>
<td>95.4</td>
<td>96.6</td>
<td>97.6</td>
<td>2.2</td>
</tr>
<tr>
<td>I</td>
<td>96.3</td>
<td>96.3</td>
<td>97.4</td>
<td>1.1</td>
</tr>
<tr>
<td>J</td>
<td>95.6</td>
<td>97.6</td>
<td>97.7</td>
<td>2.1</td>
</tr>
<tr>
<td>T</td>
<td>95.9</td>
<td>96.9</td>
<td>97.6</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Average score difference 1.9

**Table 8 Repeatability of Three Runs at Same Driving Speed (Ouyang, et al., 2013)**

<table>
<thead>
<tr>
<th>Run</th>
<th>Crack length on the west way (m)</th>
<th>Crack length on the east way (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Longitudinal</td>
<td>Transverse</td>
</tr>
<tr>
<td>1</td>
<td>280.29</td>
<td>138.49</td>
</tr>
<tr>
<td>2</td>
<td>281.94</td>
<td>142.99</td>
</tr>
<tr>
<td>3</td>
<td>281.16</td>
<td>141.05</td>
</tr>
<tr>
<td>Difference</td>
<td>0.59%</td>
<td>3.19%</td>
</tr>
</tbody>
</table>

**Table 9 Repeatability at Different Driving Speeds (Ouyang, et al., 2013)**

<table>
<thead>
<tr>
<th>Driving speed (km h⁻¹)</th>
<th>Int_y (mm)</th>
<th>Crack length on the west way (m)</th>
<th>Crack length on the east way (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Longitudinal</td>
<td>Transverse</td>
</tr>
<tr>
<td>25</td>
<td>1.39</td>
<td>289.07</td>
<td>148.12</td>
</tr>
<tr>
<td>30</td>
<td>1.67</td>
<td>281.94</td>
<td>142.99</td>
</tr>
<tr>
<td>40</td>
<td>2.22</td>
<td>289.19</td>
<td>143.81</td>
</tr>
<tr>
<td>50</td>
<td>2.78</td>
<td>286.38</td>
<td>141.93</td>
</tr>
<tr>
<td>Difference</td>
<td>2.53%</td>
<td>4.29%</td>
<td>2.85%</td>
</tr>
</tbody>
</table>

**Table 10 Repeatability under Different Weather Conditions (Ouyang, et al., 2013)**
Wang et al. (2011) used the cracking index defined in the UK Surface Condition Assessment of the National Network of Roads, also known as UK SACNNER, to evaluate results from manual, semi-automated and fully automated procedures. Using this protocol, pavement surface is divided into 200 mm × 200 mm grids. The percentage of the cracked grids out of the total number of the grids is calculated for each 50-meter roadway section, which is the UK SCANNER index:

\[
\text{Index} = \frac{n_c}{N} \times 100\% 
\]  

(8)

Where \( n_c \) is the total number of cracking grids in a 50-m roadway section, and \( N \) is the total number of grids in the same 50-m section. This index doesn’t consider distress severity and location, but it is objective and easy to implement by both computer and human raters. In Wang’s study, a precision test in accordance with ASTM E691-99 was first performed to analyze the repeatability and reproducibility of the manual results where 4 pavement sections were surveyed by 9 raters, and each rater repeated the survey 4 times for each test section. Then, the average of the manual results in each test section was used as the reference value to compare with the semiautomated and fully automated results; the repeatability and reproducibility acceptable range was used to determine whether the semiautomated and fully automated results were acceptable. Figure 51 shows the results.
Crack width is a common and important crack classification factor in most DOTs' pavement surface condition survey protocols, especially when differentiating severity levels. It is also crucial information for determining pavement maintenance operations, such as crack sealing/filling. However, crack width has rarely been used in the past crack classification studies. Considering the properties of 2D digital images, the accuracy of crack width measurement (measured pixel by pixel) is limited; even for high resolution images, crack width measurement is still influenced by other factors, such as lighting conditions and pavement noise (e.g. oil stains). The 3D line laser technology provides a better opportunity to measure crack width more accurately.

Tsai et al. (2013) conducted an experimental test to validate the crack width measurement using 3D transverse profile. In the experimental test, a total of 12 spots were selected from SR 275 Mile 1-2 for crack width measurement validation (as shown in Figure 52 (a) and (b)). The left image is the range image (based on elevation information of the pavement surface) with a detected crack map overlaid, and the automatically measured crack width information using 3D line laser data is labeled beside the corresponding crack elements. The right image is the intensity image with the crack map overlaid, and the selected 12 locations are marked for reference.
In order to validate the automatic crack width measurement accuracy for the 12 selected spots, the ground truth crack width in the field (as shown in Figure 53) was manually measured. The validation results are shown in Table 11. From the results, we can see that, based on the 3D transverse profile data, it is difficult to detect hairline cracks with a width of about 1 mm. This is reasonable, considering the properties of the laser technique used in this test. For the detected cracks, the automatically measured width information is relatively consistent with field-
measured ground truth. The maximum absolute difference is 1 mm, and the average absolute difference is 0.4 mm.

Figure 53 Crack Width Measurement for Ground Truth (Tsai and Wang, 2013)

Table 11 Crack Width Measurement Validation for 12 Locations (Tsai and Wang, 2013)

<table>
<thead>
<tr>
<th>Location No.</th>
<th>Computed Crack Width (mm)</th>
<th>Manually Measured Crack Width (mm)</th>
<th>Absolute Difference (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.5</td>
<td>3.5</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2.8</td>
<td>3.0</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3.5</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>Not detected</td>
<td>1.5</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>Not detected</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>3.8</td>
<td>3</td>
<td>0.8</td>
</tr>
<tr>
<td>7</td>
<td>Not detected</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>8</td>
<td>3.1</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>9</td>
<td>4.8</td>
<td>4</td>
<td>0.8</td>
</tr>
<tr>
<td>10</td>
<td>2.9</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>11</td>
<td>Not detected</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Avg. (detected)</td>
<td>3.6</td>
<td>3.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

The above validation results show good accuracy for measuring longitudinal crack width. To further assess the measurement accuracy for transverse crack, 12 spots on three transverse cracks on SR 275 were selected. The manually measured crack width, which is considered as the ground truth, varies from 1mm to 7mm. Figure 54 shows an example for transverse crack spot AT1. Roughly, the manually measured crack width for AT1 is 1 mm, as shown in Figure 54 (b); the 3D-line-laser-measured result is around 5mm, as shown in Figure 54 (c). Table 12 lists the comparison results for all 12 spots. It can be seen that all three transverse cracks can be detected
by using 3D line laser, but the crack width measurement is very inaccurate for fine cracks, e.g., AT1 and AT2. This type of inaccuracy is caused by the coarser resolution of 3D line laser data collected in this test in the driving direction, which is about 5 mm.

![Image](image-url)

**Figure 54 Measurement of Transverse Crack Width (Tsai and Wang, 2013)**

**Table 12 Crack Width Validation for Transverse Crack (Tsai and Wang, 2013)**

<table>
<thead>
<tr>
<th>Location No.</th>
<th>Computed Crack Width (mm)</th>
<th>Manually Measured Crack Width (mm)</th>
<th>Absolute Difference (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1</td>
<td>5.1</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>BT1</td>
<td>6.2</td>
<td>2</td>
<td>4.2</td>
</tr>
<tr>
<td>CT1</td>
<td>5.4</td>
<td>3</td>
<td>2.4</td>
</tr>
<tr>
<td>DT1</td>
<td>5.1</td>
<td>4</td>
<td>1.1</td>
</tr>
<tr>
<td>T1</td>
<td>4.8</td>
<td>3</td>
<td>1.8</td>
</tr>
<tr>
<td>T2</td>
<td>4.8</td>
<td>3</td>
<td>1.8</td>
</tr>
<tr>
<td>T3</td>
<td>5.5</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>T4</td>
<td>5.1</td>
<td>3</td>
<td>2.1</td>
</tr>
<tr>
<td>T5</td>
<td>5.5</td>
<td>3-4</td>
<td>1.5-2.5</td>
</tr>
<tr>
<td>T6</td>
<td>5.3</td>
<td>6</td>
<td>0.8</td>
</tr>
<tr>
<td>T7</td>
<td>4.9</td>
<td>7</td>
<td>2.1</td>
</tr>
<tr>
<td>T8</td>
<td>5.4</td>
<td>6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

### 6.4 Discussion

The above literature review showed the capability of using 3D laser data for crack detection and crack width measurement. 3D laser data is insensitive to different lighting conditions, low intensity contrast, and pavement oil marks, unlike the traditional line scan camera. However, the resolution of 3D laser data determines the thinnest cracks can be detected. The current technology should be capable of detecting cracks with widths greater than 1mm. However, a hairline crack with a width of approximately 1 mm is hard to detect due to the current resolution of the 3D line laser.
Crack width is very critical information for highway agencies to make pavement maintenance decisions. For example, crack filling can only be applied to cracks with widths greater than 1/8 inch in GDOT. The accuracy of measuring crack width using 3D laser data depends on the data resolution. In Tsai’s study (Tsai and Wang, 2013), longitudinal cracks with widths greater than 2 mm can be measured correctly. Cracks equal to and less than 1 mm cannot be detected and measured correctly because the resolution in the transverse direction is about 1 mm. Since the longitudinal resolution is about 5 mm in Tsai’s study (Tsai and Wang, 2013), the transverse crack width cannot be reliably measured.

7. **Summary and Road Map**

3D transverse profile data has gained more and more interest from highway agencies, industry, and researchers due to its advantages over 2D digital images in collecting pavement condition features, such as rutting, cracking, raveling, macrotexture, and smoothness. However, there is a lack of a standard calibration/validation method for end users (e.g. highway agencies, data collection service providers, and researchers) to conduct quality assurance and quality checking (QA/QC) on the collected data.

This literature review first briefly introduced the system diagram of laser triangulation that was used in all the current data acquisition systems and the triangulation principle that was used to calculate the surface elevation change. Then, major types of data noises and their characteristics identified by Tsai’s study were discussed. A calibration/validation board developed by Tsai was also introduced. The identification of data noise is very critical for the following development of calibration/validation methods.

Other than the data itself, validation through its application was normally applied in literature. Thus, this literature review also included the data validation by evaluating the accuracy of collecting pavement rutting and cracking. Basically, laboratory tests and field tests were performed to compare the derived/detected pavement distresses (rut depth, crack location, and crack width) to the known reference (i.e. ground truth). This type of validation cannot reveal the fundamental data characteristics and, to some extent, depend on the testing data. However, it can directly show the performance of using 3D transverse profile data to derive/detect/measure pavement distresses, which can be easily understood by end users.
To conclude the calibration/validation of 3D transverse profile data, a road map was developed by Tsai’s research group as shown in Figure 55 (to be published). First, the instrumental level of calibration/validation should be performed, including laser/illuminator, camera, DMI, and Inertial Measurement Unit (IMU). Then, raw data accuracy in three dimensions should be validated. In addition, algorithm accuracy needs to be validated for automatic pavement distress detection through the validation of the derived/detected/Measured pavement features/distresses.
Figure 55 Road Map of Calibrating/Validating 3D Transverse Profile Data (Developed by Tsai et al.)
References


