GEORGIA DOT RESEARCH PROJECT 13-18

FINAL REPORT

DETECTING ASPHALT PAVEMENT RAVELING USING EMERGING 3D LASER TECHNOLOGY AND MACROTEXTURE ANALYSIS

Georgia Department of Transportation

OFFICE OF RESEARCH
15 KENNDY DRIVE
FOREST PARK, GA 30297-2534

Contract Research

GDOT Research Project No. 13-18

Final Report

DETECTING ASPHALT PAVEMENT RAVELING USING EMERGING 3D LASER TECHNOLOGY AND MACROTEXTURE ANALYSIS

By

Yichang (James) Tsai, Ph.D., P.E.

Zhaohua Wang, Ph.D., P.E.

Georgia Institute of Technology

Contract with

Georgia Department of Transportation

In cooperation with

U.S. Department of Transportation

Federal Highway Administration

August, 2015

<table>
<tr><td>1.Report No.:
GA-14-1308</td><td>2. Government Accession No.:</td><td colspan="2">3. Recipient's Catalog No.:</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle:
Detecting Asphalt Pavement Raveling Using Emerging 3D Laser Technology and Macrotexture Analysis</td><td colspan="2">5. Report Date: August 2015</td></tr>
<tr><td colspan="2">6. Performing Organization Code:</td></tr>
<tr><td colspan="2">7. Author(s): Yichang (James) Tsai; Zhaohua Wang</td><td colspan="2">8. Performing Organ. Report No.:</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address:
Georgia Institute of Technology
790 Atlantic Drive
Atlanta, GA 30332-0355</td><td colspan="2">10. Work Unit No.:</td></tr>
<tr><td colspan="2">11. Contract or Grant No.:
RP 13-18</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address:
Georgia Department of Transportation
Office of Research
15 Kennedy Drive
Forest Park, GA 30297-2534</td><td colspan="2">13. Type of Report and Period Covered:
Final; July 2013 – July 2015</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code:</td></tr>
<tr><td colspan="4">15. Supplementary Notes:
Prepared in cooperation with the U.S. Department of Transportation, Federal Highway Administration.</td></tr>
<tr><td colspan="4">16. Abstract:
This research project comprehensively tested and validated the automatic raveling detection, classification, and measurement algorithms using 3D laser technology that were developed through a project sponsored by the National Cooperative Highway Research Program (NCHRP) Innovations Deserving Exploratory Analysis (IDEA) program. The raveling condition survey protocol used in the Georgia Department of Transportation (GDOT) was adopted in the testing and validation, though it can be easily extended to other highway agencies' protocols. Four miles of test sections on I-85 and the entire asphalt pavements (61 miles) on I-285 were selected for validating the developed algorithms. The ground truth data was established by in-field investigation and in-office review of videolog images and 3D laser data with the help of GDOT's pavement engineers. The results have demonstrated the promising capabilities of automatically detecting and measuring asphalt pavement raveling using the developed algorithms and 3D laser technology to assist in transportation agencies' raveling data collection. Potentially, it will save tremendous manual effort for field surveys, improve data accuracy, and help highway agencies make more informed decisions on pavement maintenance and rehabilitation.</td></tr>
<tr><td colspan="3">17. Key Words:
3D laser technology; raveling; detection; classification; measurement</td><td colspan="2">18. Distribution Statement: No Restriction</td></tr>
<tr><td>19. Security Classification (of this report):
Unclassified</td><td>20. Security classification (of this page): Unclassified</td><td>21. Number of Pages: 64</td><td>22. Price:</td></tr>
</table>

Acknowledgements

The work described in this final report was supported by the Georgia Department of Transportation (GDOT) research project 13-18. The algorithms used in this research were developed through a research project sponsored by the National Academy of Sciences, National Cooperative Highway Research Program (NCHRP) Innovations Deserving Exploratory Analysis (IDEA) program. We would like to thank Mr. Timothy Dale Brantley, Ms. Ernay Robinson, Mr. Eric Pitts (retired), and Mr. Thomas Mims from GDOT for their strong support and heavy involvement in selecting test sections and collecting ground truth data. We would like to thank the research team, Mr. Hadrien Glaude, Mr. Bruno Pop-Stefanov, Mr. Jinqi Fang, Ms. Josephine M. Simon, and Mr. Yipu Zhao for their diligent work. We would like to thank Mr. Yusuf H. Ahmed for his assistance in managing this project.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

In Georgia, raveling has become the most predominant distress of concern on interstate highways with open graded friction course (OGFC). Raveling shortens pavement life and also results in various safety concerns, such as flying stones that damage vehicle windshields and bodies, and it causes rough and uneven pavement surfaces that increase road/tire noise and degrade riding safety.

On interstate highways, due to the large volume of traffic and large scope of distribution, an in-vehicle windshield survey is normally used. This method has shortcomings: results are subjective and have large variations. Also, the survey procedure is very time-consuming and labor-intensive. Thus, there is an urgent need for an automatic survey method. Compared to the method using digital images captured under ambient lighting conditions, laser technology has the advantage of being able to directly acquire pavement surface textures. However, due to the poor resolution of point laser profilers, the related study is very limited. With the advancement of sensing technology, 3D line laser imaging technology can be employed to acquire full-lane-width, high-resolution pavement surface laser data upon which raveling detection, classification, and measurement algorithms can be developed. Although some raveling detection algorithms using 3D laser data have been developed, they were not validated and the classification of raveling severity was not developed either. Thus, it has become difficult for transportation agencies to implement such algorithms because an automatic raveling data collection includes raveling detection, classification, and measurement. The raveling detection, classification, and measurement algorithms presented in this final report were the first ones that have been comprehensively validated using real-world, large-scale pavement data.

To address the above urgent need, the Georgia Tech research team developed new raveling detection, classification, and measurement algorithms using 3D laser technology in a project sponsored by the National Academy of Sciences, National Cooperative Highway Research Program (NCHRP) Innovations Deserving Exploratory Analysis (IDEA) program. This research project further tested and validated the developed algorithms using GDOT's pavement condition survey protocol. The algorithms can easily be extended to other highway agencies' pavement condition survey protocols.

The algorithms were comprehensively tested and validated near Atlanta, Georgia, on I-85 and I-285 because they are surfaced with OGFC. Four test sections, each of which is one mile long, were selected on I-85; the entire outer lane of asphalt pavements (61 miles) on I-285 were also selected for the testing and validation.

The testing results on I-85 showed that the developed algorithms are very accurate for GDOT's use, and the lump sum of all types of raveling is very accurate. The automatic classification results on each of the test sections were compared with the ground truth (the ones measured by the GDOT pavement expert). The difference of total raveled percentage on test section #1 is about 1.34%, which is less than the 0.2% for the other 3 test sections. The predominant severity levels for Test Sections #1 and #2 are also correctly classified. For Test Sections #3 and #4, there is, essentially, no raveling, and the classification errors are 0.15% and 0.06%.

The testing on I-285 showed promising results for automatic raveling detection, classification, and measurement. All the pavements (with or without raveling) were 100% correctly detected and classified at the segment level (each segment is one mile long). However, due to the difficulty of correctly labeling all the raveling areas using videolog images and 3D laser data and due to the impact of cracking and flat-tire scratches, the raveling extent (percentage) showed a certain level of variation in comparison with the manually labeled ground truth. The difference between the surveyed results that conducted by the experienced GDOT pavement engineer and the automatically detected and measured results is less than 15% and most of them are less than 10%.

In summary, the proposed algorithms and validation results have demonstrated promising capabilities in being able to automatically detect and measuring asphalt pavement raveling. Using the proposed algorithms will, potentially, save tremendous amounts of manual effort in field surveys, improve data accuracy, and help highway agencies make more informed decisions about pavement maintenance and rehabilitation.

Chapter 1 Introduction

1. Research Background and Research Need

Raveling is one of the most common asphalt pavement distresses that occur on U.S. highways. It is defined as the "wearing away of the pavement surface caused by the dislodging of aggregate particles and loss of asphalt binder" in the distress identification manual for the long-term pavement performance program (LTPP) (FHWA, 2003). In Georgia, raveling has become the most predominant distress of concern on interstate highways with open graded friction course (OGFC). Figure 1.1 shows two examples of raveling that occurs on interstate highway and non-interstate roads.

Figure 1.1 Raveling on Interstate Highways and Non-interstate Roads

Raveling shortens pavement life and results in various safety concerns, such as flying stones that damage vehicle windshields and bodies; it creates rough and uneven pavement surface that increase road/tire noise and degrade riding safety. Georgia has experienced significant raveling issues on all the major interstate highways around the greater Atlanta area. . Similarly, practitioners from the Netherlands and the United Kingdom (UK) found that raveling is the predominant distress for porous pavements in the Dutch trunk network (70% were paved with porous asphalt) and for the hot rolled asphalt (HRA) wearing course used on UK motorways. If the raveled asphalt pavement is not sealed in a timely manner, the raveling problem can develop very fast. Consequently, pavement surface layers should be replaced quickly after the first observation of raveling (Miradi, 2004). Nowadays, due to its stringent pavement budget, GDOT aggressively applies cost-effective preventive maintenance treatments, such as fog seal, on interstate highways to deter raveling development. It is very crucial to identify raveling and treat it in its very early stages using low-cost surface coating methods. Otherwise, much more expensive corrective treatments will be needed; this situation would seriously deplete highway agencies' already stringent budgets.

The commonly used manual survey method has hindered the early discovery of pavement raveling due to the following reasons: 1) the manual survey process is very time-consuming, and it is difficult for highway agencies to conduct a full-coverage survey, 2) the survey protocol is subjective, and survey results vary from rater to rater and from time to time due to the impact of ambient lighting conditions, 3) for high-traffic volume interstate highways, a raveling survey is often omitted by highway agencies due to the high demand of traffic control, and 4) digital-image-based surveys are very unreliable because raveling is the change of pavement surface texture, and the ability to see it, recognize it, and accurately assess it heavily depends on ambient lighting conditions. Therefore, there is an urgent need to develop an objective, reliable method for automatic asphalt pavement raveling detection, classification, and measurement.

Being able to conduct an automatic pavement condition survey at highway speed has attracted more and more interest from highway agencies due to its objectiveness and cost-effectiveness. Digital videolog images and laser profilers have been widely used for pavement cracking and roughness surveys. Some of them are even commercially available. However, to the best of our knowledge, no successful method using digital video log images and a laser profiler for pavement raveling surveys has been reported. One reason is that the use of digital videolog images or point-based laser profilers cannot effectively capture the characteristics of pavement ment of Transportation, which is the essential representation of pavement raveling. A digital image has only two-dimensional information of a pavement surface, while a point-base laser profiler only captures the pavement texture on a single longitudinal line. Thus, they are not suitable for detecting the area-based pavement raveling at the macrotexture level.

However, emerging 3D line laser imaging technology can capture 3D pavement surface macrotexture with full coverage. This technology has brought new opportunities for developing a more accurate and reliable automatic pavement condition survey. Currently, the commercially available 3D line laser imaging device can capture continuous pavement transverse profiles at highway speed with 0.04 in. (1 mm) transverse resolution and 0.2 in. (5 mm) longitudinal resolution (driving direction). Figure 1.2 shows the Georgia Tech Sensing Vehicle (GTSV) implemented by Dr. James Tsai and his research team at Georgia Tech; the van is equipped with a high-definition line laser imaging device and will be used for this research.

Figure 1.2 Georgia Tech Sensing Vehicle

2. Research Objectives

Sponsored by the National Academy of Sciences, National Cooperative Highway Research Program (NCHRP) Innovations Deserving Exploratory Analysis (IDEA) program, algorithms for automatic raveling detection, classification, and measurement using 3D laser technology and macrotexture analysis were developed by the Georgia Tech research team. The proposed algorithms provide an effective means to automatically extract raveling by taking advantage of the high-resolution, full pavement lane-width coverage, and 3D pavement surface range data that have already been collected for rutting and crack detection.

The objective of this research project is to comprehensively validate the capability and effectiveness of the developed algorithms on Georgia's highways and further refine the algorithms based on the testing results. The research outcome of this study will have a significant impact on enabling GDOT to perform an accurate asphalt pavement raveling survey, especially on high-traffic-volume interstate highways. The following work tasks are proposed to achieve this goal:

- Selecting test roadways and collecting video log images and 3D laser data;
- Conducting field and in-office surveys to establish ground truth;
- Performing automatic raveling detection, classification, and measurement, and validating test results;
- Refining raveling detection, classification, and measurement algorithms;
- Summarizing research findings.

3. Report Organization

This report is organized into five chapters. Chapter 1 introduces the research background, need, research objective, and major tasks. Chapter 2 summarizes the literature review of highway agencies' practices on raveling surveys and the automatic raveling detection and classification methods. Chapter 3 presents the proposed methodology for an automatic raveling detection, classification, and measurement using 3D laser data and macrotexture analysis. Chapter 4 presents the validation procedures and validation results using data collected on I-85 and I-285. Chapter 5 summarizes the project, presents conclusions, and offers recommendations for future research.

References

FHWA (2003). "Distress identification manual for the Long-Term Pavement Performance Project." Federal Highway Administration, publication number FHWA-RD-03-031.

Miradi, M. (2004). "Neural network models for analysis and prediction of raveling." *Proc., 2004 IEEE Conference on Cybernetics and Intelligent Systems (CIS).*

Chapter 2 Literature Review

This chapter summarizes GDOT's pavement condition survey protocol on asphalt pavement raveling. Other highway agencies' practices for conducting raveling surveys are also reviewed. Finally, the state-of-the-art of algorithms for automatic raveling detection and classification are presented.

1. Raveling Survey in GDOT

In the asphalt pavement condition survey manual used by GDOT, raveling is defined as the progressive disintegration of the surface layer (GDOT, 2007). It is characterized by the loss of stones constituting the surface layer and happens over time when the surface binder is eroded by the friction caused by vehicle tires or when the pavement is damaged by an accident. It ranges from the loss of a few stones to the loss of an entire portion of the surface layer. Once the surface loses a few stones, adjacent stones break loose because they have no more support. Stones are then lost exponentially until the surface layer has disappeared, and the lower pavement layers are exposed and become bare. Raveling usually occurs in the wheel path and should not be confused with other types of distresses, such as cracks or potholes.

In GDOT, raveling is classified into Severity Levels 1, 2 and 3 based on different raveling conditions. GDOT's severity levels are as follows:

- **Level 1**: loss of substantial number of stones (see Figure 2.1 (a))
- **Level 2**: loss of most surface (see Figure 2.1 (b))
- **Level 3**: loss of substantial portion of surface layer (>1/2 depth) (see Figure 2.1 (c))

In field surveys, raveling is closely observed, and an estimate (to the nearest 5%) is made of the extent and the predominant severity of the distress within the rated segment. The percent of the length of the rated segment (mile or partial mile) that contains raveling is recorded along with the predominant severity level. On two-lane and multi-lane undivided highways, the rater should determine which lane is in the worst general shape and base his/her estimate of the extent and severity of the pavement distress on what is observed in the lane selected. Likewise, on divided highways, only the lane in the worst condition in a given direction is to be rated; each direction is rated separately for divided highways.

(a) Severity (b) Severity (c) Severity

Figure 2.1: Raveling Classification in GDOT (GDOT, 2007)

2. Other Highway Agencies' Practice

As one of the most common asphalt pavement distresses, raveling data is collected by highway agencies for evaluating pavement conditions and determining proper treatment. Though other highway agencies define raveling in almost the same way as GDOT does (FHWA, 2003; NYSDOT, 2000; ODOT, 2010), rating methods for severity levels and extents change from agency to agency. This section reviews the current practices in different agencies and summarizes the challenges and needs for improvement.

2.1 Raveling Classification and Survey Practice

In highway agencies' practices, raveling is classified based on its severity. It can further be used for rating the overall pavement conditions and facilitate the determination of maintenance or rehabilitation treatments. The definitions of different raveling classifications and the survey practices of different state DOTs are presented below.

1) Florida Department of Transportation (FDOT) (FDOT, 2009)

In FDOT, raveling is classified into three categories: low, moderate, and severe. Only significant areas of raveling are considered. An isolated area is not counted in a long section if it is not representative of the rated section. The predominant severity level and percent affected area of raveling are recorded. FDOT's definitions of different severity levels are as follows:

- **Low**: The aggregate and/or binder has begun to wear away but has not progressed significantly; there is some loss of aggregate.

- **Moderate**: The aggregate and/or binder have worn away, and the surface texture is becoming rough and pitted; loose particles generally exist; loss of aggregate has progressed.
- **Severe**: The aggregate and/or binder have worn away, and the surface texture is very rough and pitted; loss of aggregate is very noticeable.

2) Minnesota Department of Transportation (MDOT) (MDOT, 2000)

The raveling classification in MDOT is similar to FDOT's; it is categorized as low, moderate, and high.

- **Low:** The aggregate or binder has begun to wear away but not progressed significantly. Some loss of fine aggregate is visible.
- **Moderate:** Aggregate and/or binder is worn away, and the surface texture is becoming rough and pitted. Loose particles exist; there is loss of fine aggregate and some loss of coarse aggregate.
- **High:** Aggregate and binder have worn away, and the surface texture is very rough and pitted due to the loss of coarse aggregate.

3) Nebraska Department of Roads (NDOR) (NDOR, 2002)

The terms for raveling classification in NDOR are similar to MDOT's. However, the definitions differ as follows:

- **Low**: Minimal loss of aggregate or binder (see Figure 2.2 (a)).
- **Moderate**: Some aggregate loss; small areas may be stripped away (see Figure 2.2 (b)).
- **High**: Sections greater than one square foot may be pitted, stripped, or eroded away (see Figure 2.2 (c)).

(a) Low (b) Moderate (c) High

Figure 2.2 Raveling Classification in NDOR (NDOR, 2002)

4) Oregon Department of Transportation (ODOT) (ODOT, 2010)

ODOT's classifications are the same as MDOT's. But, the definitions are more quantitative based on the percentage of aggregate loss.

- **Low**: The aggregate has worn away, resulting in 25% to 50% aggregate loss in a 1ft-wide longitudinal strip of pavement surface (see Figure 2.3 (a)).
- **Moderate:** The surface texture is noticeably rough and/or pitted with 50% to 75% aggregate loss in a 1ft-wide longitudinal strip of pavement surface (see Figure 2.3 (b)).
- **High:** The surface texture is very rough and/or pitted with 75% or more aggregate loss in a 1ft-wide longitudinal strip of pavement surface (see Figure 2.3 (c)).

In field surveys, raveling can be identified by a roughened or pitted texture on the pavement surface. Mechanical abrasion from tire chains, studs, snowplows, or dragging equipment that can significantly roughen the texture should be rated as raveling. Raveling tends to be most often found in the wheel paths but can be elsewhere on the pavement surface. To measure raveling, the number of linear feet for each severity level in each path (inside, outside, and between wheel paths) must be recorded.

(a) Low (b) Moderate (c) High

Figure 2.3 Raveling Classification in ODOT (ODOT, 2010)

5) Texas Department of Transportation (TxDOT) (TxDOT, 2009)

In TxDOT's pavement condition survey manual (TxDOT, 2009), raveling is classified into low, medium, and high based the percentage of raveled area as follows:

- **Low**: The percent of raveled pavement area is from 1% to 10%.
- **Medium:** The percent of raveled pavement area is from 11% to 50%.

- **High:** The percent of raveled pavement area is greater than 50%.

6) Washington State Department of Transportation (WSDOT) (WSDOT, 1999)

The terms for raveling classification in WSDOT are same as those in TxDOT, but, the definition is not based on quantitative measures.

- **Low:** The aggregate and/or binder has started to wear away but has not progressed significantly. The pavement only appears slightly aged and slightly rough (see Figure 2.4 (a)).
- **Medium:** The aggregate and/or binder have worn away and the surface texture is moderately rough and pitted. Loose particles may be present, and fine aggregate is partially missing from the surface (see Figure 2.4 (b)).
- **High:** The aggregate and/or binder have worn away significantly, and the surface texture is deeply pitted and very rough. Fine aggregate is essentially missing from the surface, and pitting extends to a depth approaching one half the coarse aggregate sizes (see Figure 2.4 (c)).

In field surveys, raveling is measured or observed differently depending on whether the road surface is bituminous surface treatment (BST) or asphalt concrete pavement (ACP). Care should be exercised when rating chip-sealed pavements, as they tend to look raveled because of the inherent nature of the chip-seal surface. However, raveling in chip-sealed pavements (loss of aggregate) actually results in a condition of excess asphalt and should be rated as flushing. In practice, the raveling severity and extent are both estimated and recorded. The extent of raveling is estimated and expressed relative to the surface area of the surveyed lane. Recommended ranges for estimated extent include the following:

- Localized —Patchy areas, usually in the wheel paths.
- Wheel Path —The majority of wheel tracks are affected, but there is little or none elsewhere in the lane.
- Entire Lane —Most of the lane is affected.

(a) Low (b) Medium (c) High

Figure 2.4 Raveling Classification in WSDOT (WSDOT, 1999)

2.2 Summary

Based on the above review on the current practice of highway agencies for raveling classification and survey procedures, the following summarize the major challenges and the needs for research on automatic raveling detection and classification.

Raveling is by nature the change of asphalt pavement surface texture due to the disintegration of coarse aggregates. It develops exponentially after it starts. Hence, it is critical for highway agencies to know the early stages of raveling at low severity levels and some preventive maintenance treatments (e.g. fog seal) that can be applied before a pavement deteriorates to higher severity levels and much more expensive corrective treatments will be needed. For a low severity level of raveling, the appearance of surface texture dramatically changes under different lighting conditions and depends on how one observes it. Under direct sunshine, it is hard to recognize lightly-raveled surfaces. Also, if one conducts a windshield survey in a moving vehicle at highway speed, it is hard to recognize low-severity raveling. Thus, visual inspection under natural lighting conditions is not a good means for a raveling survey, especially for low-severity level raveling. To overcome this shortcoming, 3D laser data is a better means for capturing pavement surface texture because it is independent of ambient lighting conditions and can be collected at highway speed. Thus, raveling data extracted from 3D laser data is more reliable than the one under ambient lighting conditions.

In all major highway agencies, raveling levels are defined for in-field visual inspection, which is qualitative and subjective. This will cause large variations in the survey data. Some state DOTs, such as ODOT and TxDOT, use a quantitative method for raveling classification, but, there is no direct relationship between raveling severity levels and surface textures. In most state highway agencies, only a limited number of sample images are used for training inspectors, so there is no real consistency among different inspectors and, therefore, the data the inspectors collect varies greatly. To overcome this shortcoming, an objective measure is needed to detect and classify raveling. Since raveling appears as the change of asphalt pavement surface texture, the texture-related parameters need to be considered for use in a computerized, automatic raveling detection and classification system.

There are many problems in evaluating raveling. Raveling is the change of a pavement's surface texture and can be continuously distributed. Additionally, the survey procedure for raveling is tedious, time-consuming, and error-prone. For example, in GDOT, raveling is surveyed by means of windshield survey. Also, raveling appears different when it is observed when standing still on the ground than when it is observed from a moving vehicle. Raveling extent is hard to accurately measure from a moving car. Thus, though raveling is one of the most common and critical asphalt pavement distresses, current visual inspection methods are insufficient. Consequently, there is an urgent need to develop an automatic method for detecting and classifying raveling. Fortunately, with the advancement of sensing technology and machine learning theory, it has become possible nowadays.

3. Review of Algorithms for Automatic Raveling Detection

Before the proposed algorithms for automatic raveling detection are presented, the state-of-the-art of the existing methods will be reviewed.

3.1 Sensing Data for Raveling Detection

Due to their merit of being independent from ambient lighting conditions, laser sensors were used for collecting pavement surface texture data. In 2004, Ooijen et al. (Ooijen, 2004) started to use laser data (10.5 ft. (3.2 m) Field of View (FOV), 25 points per scan) in detecting and classifying raveling. Since then, laser sensors with increasing FOV and resolution have been applied to raveling detection and classification. McRobbie et al. (McRobbie, 2008; McRobbie, 2012; Scott, 2008) used laser data with 11.8 ft. (3.6 m) FOV and 25 points per scan. Laurent et al. (Laurent, 2012a; Laurent, 2012b) worked on range data with 13.1 ft. (4 m) FOV and 4,096 points per scan.

3.2 Algorithms

Ooijen et al. (Ooijen, 2004) developed the “Stoneway” algorithm to detect raveling on porous asphalt pavement. The Stoneway model calculates the percentage of lost stones per meter. The idea is that if somewhere on the surface there is a “Stoneway,” it will show up on the texture profile in one way or another. Two main parameters in the model are the height and the length of a gap (i.e. the imprint of a lost stone) which are referred to as the “highgap” and “greatgap,” respectively. Furthermore, the severity of raveling is classified by the percentage of aggregate missing from the surface.

In this method, the raveling regions are defined as gaps in the longitudinal sampling data that are large and deep enough (shown in Figure 2.5). Due to the nature of 25-line laser sensing, the method runs in one dimension per line. The parameter “greatgap” is used as the threshold between large and small. The other one, “highgap,” is used to help judge if the gap is deep enough.

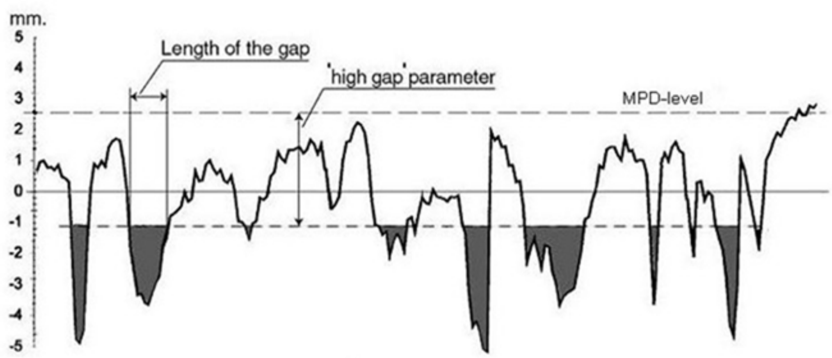


Figure 2.5 “Stoneway” Algorithms (Ooijen, 2004)

Validation was performed in a test in which two datasets consisting of the visual condition survey (VCS) data and the Stoneway data were correlated. The VCS method does not assess the actual amount of raveling; rather, it estimates the intervention year directly during the survey. Therefore, the comparison was made on the intervention year predicted by 2 approaches.

About 500 sections, each 328 ft. (100 m) long, of different age classes and of the same age class but with differing severities of raveling were chosen to validate the Stoneway derived intervention years, derived by applying the SHRP-NL propagation model against the direct VCS estimated year of intervention. Note that the testing data were from a porous asphalt surface.

To arrive at the validation result, analyses of standard deviations were run on both outputs. The result, $S_{Stoneway} / S_{VCS} = 0.87$, suggests that the standard deviation of the Stoneway model is significantly smaller than the VCS standard deviation.

The second comparison was run on the means of the estimated intervention years. From the results shown in Figure 2.6, it can be seen that the Stoneway model tends to schedule the intervention later than the VCS estimation, except for the first planned years.

Standard deviations (years) degrees of freedom		Intervention year according to last monitoring cycle (2001)		
Increasing age	Construction year	2005-2006	2004	2001-2003
	1989-1992	0.88 140	0.52 138	0.52 140
	1993-1996	0.64 140	0.55 136	0.72 146
	1997-2000	0.59 76	0.73 98	0.29 8
		increasing distress		

Table 1 *Standard deviations of intervention years according to VCS*

Standard deviations (years) degrees of freedom		Intervention year according to last monitoring cycle (2001)		
Increasing age	Construction year	2005-2006	2004	2001-2003
	1989-1992	0.44 140	0.57 138	0.81 140
	1993-1996	0.56 140	0.50 138	0.58 146
	1997-2000	0.61 76	0.26 98	0.37 10
		increasing distress		

Table 2 *Standard deviations of intervention years according to Stoneway-model*

Figure 2.6 Results Comparison between "Stoneway" Method and VCS Method (Ooijen, 2004)

Two challenges are observed for the Stongway method. First, the road surface is assumed to be flat at the horizontal direction. Therefore, it may not work on inclined surfaces. Second, the sampling rate of the road profiles is quite low. The transverse sampling rate is 19.7 in. (500 mm) per point. At such low rate, the collected profile may not be sufficient to represent the whole surface. Therefore the overall raveling detection and classification results can be easily influenced.

Laurent et al. (Laurent, 2012a; Laurent, 2012b) developed a raveling index (RI) to quantify raveling. The RI is calculated by measuring the volume of aggregate loss (holes due to missing aggregates) per unit of surface area (square meter). 3D line laser imaging technology was used for surface range data collection. This high-resolution 3D laser data allows for the detection of missing aggregates. The formula for RI estimation is given below:

$$RI = V_{ravelling} / A_{total}$$

Limited tests have been done on raveling detection. The general method used here is to run raveling detection on the same road section repeatedly. If the results tend to be similar, then the robustness of the detection approach is proved (on a limited level, though). The results of a repeatability test (3 passes) on road sections in the Netherlands are shown below (see Figure 2.7).

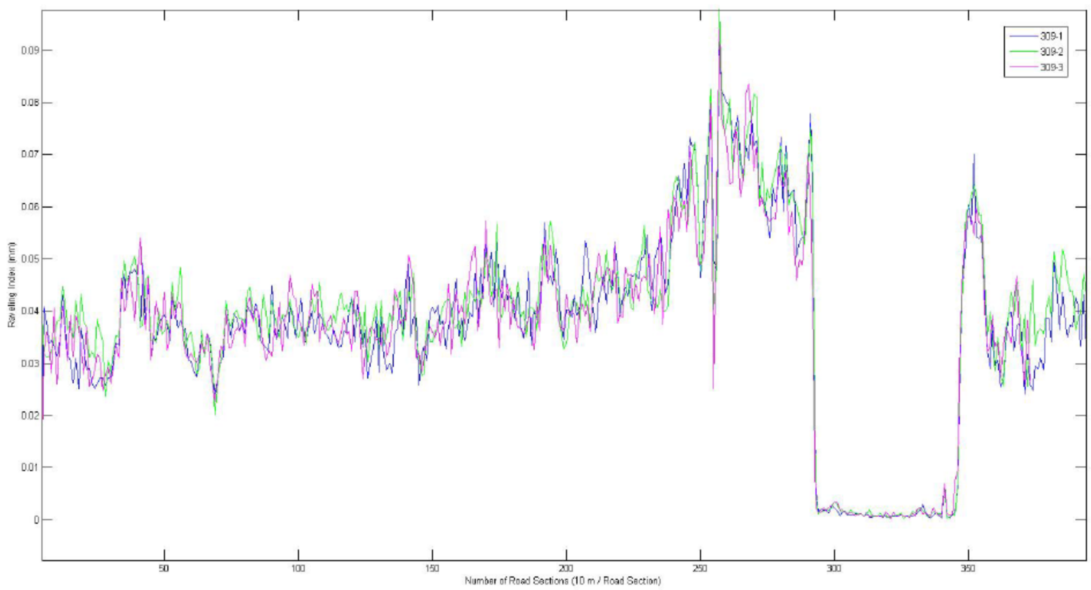
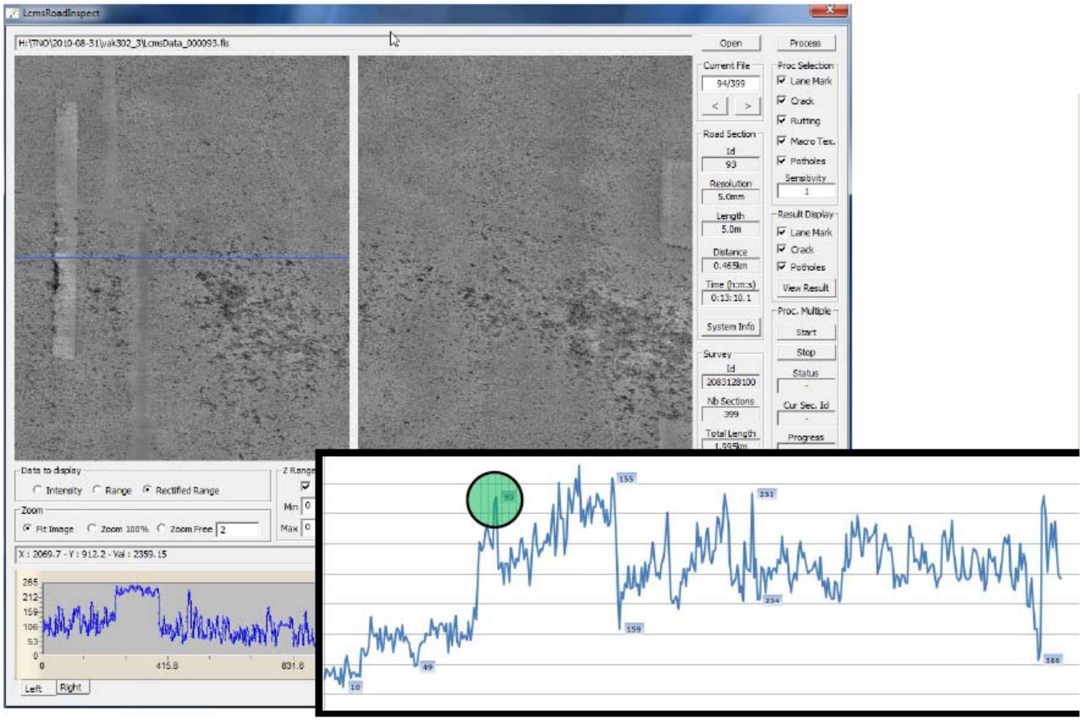
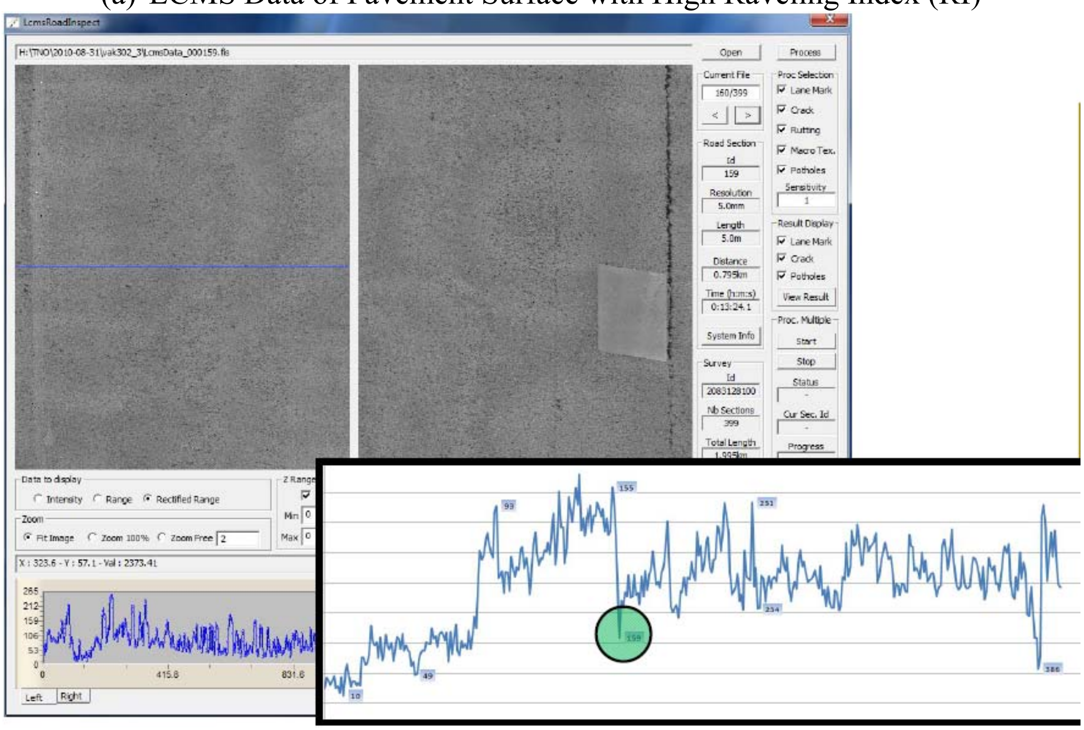


Figure 2.7 Repeatability Test Results (3 Passes in 3 Different Colors) (Laurent, 2012b)

In addition, some visual comparison was carried out. By watching the RI on surfaces with different raveling severities, the RI seems to be relevant to raveling severity. A comparison figure is given below in Figure 2.8. Though LCMS (laser crack measurement system developed by Pavemetrics) has developed raveling detection algorithm, it has not been fully validated using large-scale data set. In addition, there is no raveling severity level that can be classified using LCMS software because it is developed so far due to its complexity.



(a) LCMS Data of Pavement Surface with High Raveling Index (RI)



(b) LCMS Data of Pavement Surface with Low Raveling Index (RI)

Figure 2.8 Results Comparison with Different Raveling Index (RI) (Laurent, 2012b)

Different from Ooijen (2004), the density of 3D laser data used here was high enough to cover the whole lane transversely. Therefore, the performance of raveling detection and classification was expected to be better. However, the tests for raveling detection are very limited and without systematic validation of using a large-scale dataset. Also, the RI is not related to raveling survey protocol. Thus, it is difficult for highway agencies to directly use.

McRobbie et al. developed two raveling detection and classification methods (McRobbie, 2008; McRobbie, 2012; Scott, 2008). The first method is based on mean profile depth (MPD) (Scott, 2008). Locations that differ from the characteristic level by a sufficient depth and over a significant length, are deemed to be raveled. The proportion of the road affected by raveling is reported. Two parameters used here are *D* (the required difference, which must be observed between the baseline and the filtered profile before fretting can be reported) and *L* (the length of profile over which *D* must be exceeded before fretting can be reported).

In the second approach, the root mean square texture (RMST) was calculated and reported (McRobbie, 2008; McRobbie, 2012). By assigning a color to each of the RMST values in the data it was possible to produce a visual representation of the surface texture in which features such as road markings, metalwork, surface changes, potholes, and raveling could be seen (see Figure 2.10). Based on RMST, a raveling detection algorithm is introduced. The basic underlying concept for the algorithm is the comparison of the distribution of RMST values in a small ("Local") area against those from a much larger surrounding ("Global") area (see Figure 2.10).

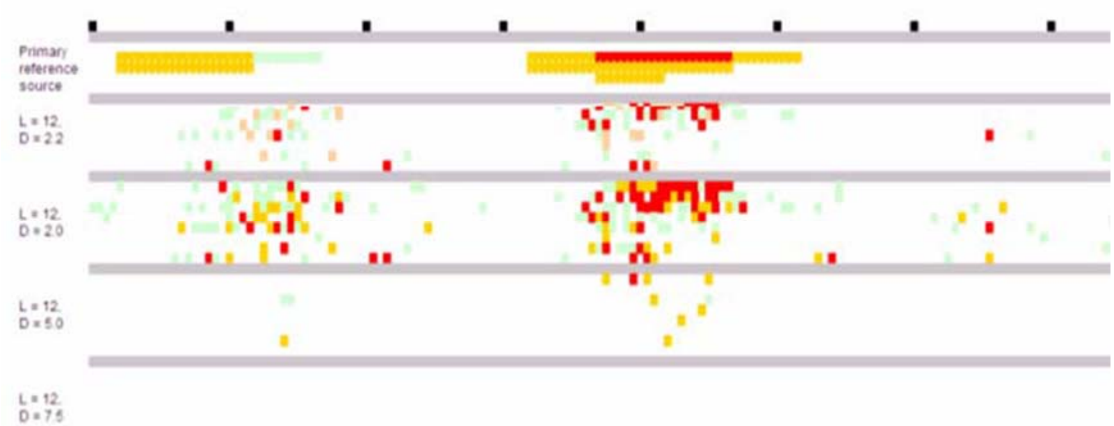


Figure 2.9 Visual Representation of Surface Using Root Mean Square Texture (RMST) (McRobbie, 2008)

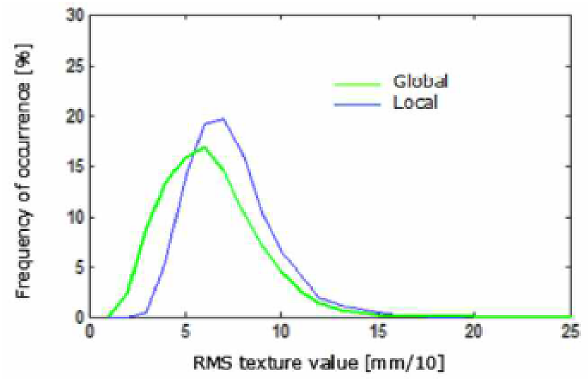


Figure 2.10 An Example of the Distribution Difference between "Local" Area and "Global" Area (McRobbie, 2008)

To provide a good range of reference (i.e. ground truth) data, a selection of sites totaling a length of approximately 90 km were selected (see Figure 2.11), representing a combination of different surface types (thin surface course, porous asphalt, hot rolled asphalt, etc.) and surface conditions. Then the coarse visual inspection (CVI) method was proposed as a suitable means of collecting larger volumes of reference data.

Site	Description	Site lane length [m]
1	A28. Godmersham to Chilham	7102
2	A28. St. Michaels, Tenterden	6490
3	A228. Colts Hill to Hale Street Bypass	9227
4	A249. Detling to Stockbury	8985
5	A256. Betteshanger to Guston	9583
6	A259. Rhee Wall	5186
7	A2070. Rhee Wall to Kingsnorth, Ashford	14305
8	A299. Thanet Way	23621
9	A28/A268 Sandhurst to Newenden	5281
	Total length	89780

Figure 2.11 Sites Selected for Testing the RMST Method (McRobbie, 2008)

The validation was conduct by the comparison between the output of RMST method and the reference data (see Figure 2.12). This generally showed a good agreement with the same areas usually being picked out by higher values, and a few areas where the local trends and shapes of the lines follow each other well.

The first method, MPD, is an extension of the Stoneway method proposed by Ooijen et al. (Ooijen, 2004). Similar to Stoneway, MPD suffers from its assumption that the surface should be horizontally flat. Though the baseline used in MPD is calculated on relatively short lengths (e.g. 200 m), it still fails on surfaces with large inclines (obvious in short lengths). As for the second method, RMST, some challenges exist due to its basic assumption. The first one is how to decide local and global. The criteria should not be the same for different surfaces and under different road conditions. The other challenge is how to accurately estimate surface conditions based on the inaccurate representation, i.e. on an RMST histogram. Many different surface conditions may appear as similar in the RMST histogram.

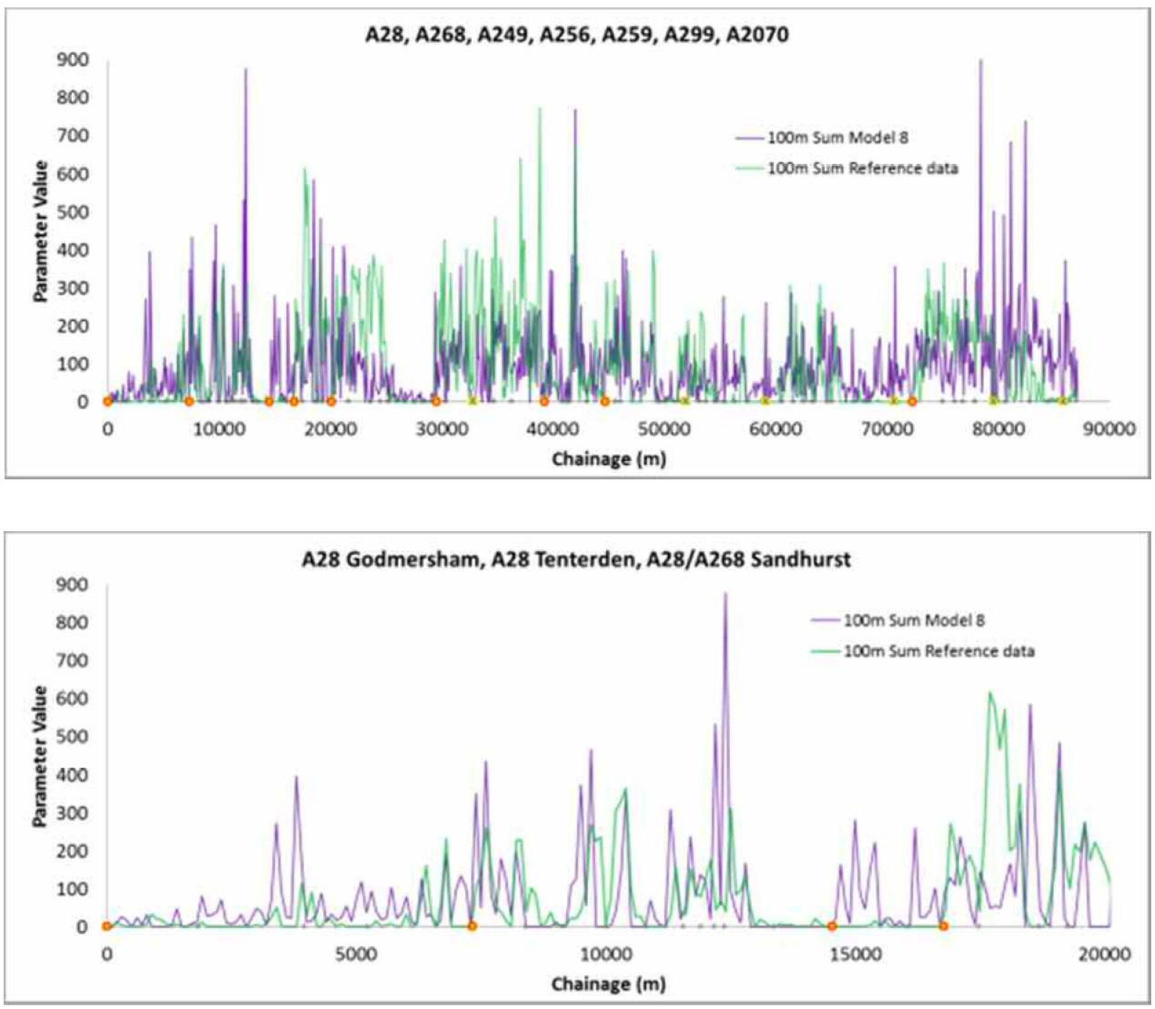


Figure 2.12 Results Comparison between Proposed Method Output and Reference Data (McRobbie, 2008)

3.3 Summary

First, most of the raveling detection and classification research is still at the research stage and many questions remain. Compared to the extensive studies for other pavement distresses, e.g. cracking, rutting, etc., there are only a very limited number of studies on raveling detection and classification. All the existing automatic methods use only pavement intensity images to distinguish the raveled pavement from the non-raveled pavement and then further classify their severity.

Second, there are no global indicators that can be universally accepted for reliable raveling classification. In addition, many indicators are based on certain assumptions about the surface that might not be applicable to other cases. For example, MPD, the commonly used indicator, employs two parameters to describe the volume of losing aggregates to classify raveling. However, it only works on horizontally flat surfaces due to the nature of its definition. Another indicator, RMST, relies on the concept that raveled areas have a different texture pattern than non-raveled ones. When applying RMST on a long stretch of consistently fretted pavement, this indicator will fail to identify the raveled areas.

Third, even with only a limited number of indicators that can be potentially used for raveling detection and classification, the existing methods frequently require parameter tuning and adjustment based on empirical experiment. These empirical trial and error approaches might constrain the existing algorithm from a wide application for different surfaces, different raveling condition, or even different data sources. The principle of using intensity images for identifying raveling is in lieu of associating the depth variance with the intensity variance. However, some of other roadway conditions, such as pavement edge drop, sudden intensity changes caused by pavement marking, etc., may introduce challenges to the existing algorithms.

Fourth, most of the validation methods used in the literature only contain a limited quantity of data. More importantly, among the limited number of data, the diversity of the data might not be adequate to objectively reveal the true performance of the automatic method (e.g. only on 1 or 2 types of pavement) or its limitations.

To summarize, automatic raveling detection and classification is still in its early stage of development. There have been a few attempts at processing intensity images with the above-mentioned algorithms. However, their performance still needs to be improved. With the advancement of the 3D laser technology, true 3D data become available for pavement distress identification. The physical features of raveling, i.e. aggregate lost, can be more realistically and precisely captured by 3D data. Therefore, there is an opportunity to develop a new raveling detection and classification algorithm using 3D data. Though LCMS has developed raveling detection algorithm using 3D laser data, it has not been fully validated using large-scale data set. In addition, there is no raveling severity level that can be classified using LCMS software because it is developed so far due to its complexity.

References

FDOT (2009). “Flexible Pavement Condition Survey Handbook.” Florida Department of Transportation

FHWA (2003). “Distress identification manual for the Long-Term Pavement Performance Project.” Federal Highway Administration, publication number FHWA-RD-03-031

GDOT (2007). “Pavement Condition Survey Manual.” Georgia Department of Transportation

Laurent, J., Hebert, J. F., Lefebvre, D., Savard, Y. (2012a). “Using 3D Laser Profiling Sensors for the Automated Measurement of Road Surface Conditions.” 7th RILEM International Conference on Cracking in Pavements

Laurent, J., Hebert, J. F., Lefebvre, D., Savard, Y. (2012b). “High-Speed Network Level Road Texture Evaluation Using 1mm Resolution Transverse 3D Profiling Sensors Using A Digital Sand Patch Model.” 7th International Conference on Maintenance and Rehabilitation of Pavements and Technological Control

McRobbie, S., and Furness, G. (2008). “Automated Detection of Fretting on HRA Surfaces.” TRL Published Project Report

McRobbie, S., Iaquinta, J., Wright, A., Trumper, P., Kennedy, J. (2012). “Development and Validation of Algorithms for the Automatic Detection of Fretting Based On Multiple Line Texture Data.” Research into Pavement Surface Disintegration. Phase 2 Interim Report

MDOT (2000). “Best Practices Handbook on Asphalt Pavement Maintenance.” Minnesota T2/LTAP Program

NDOR (2002). “Pavement Maintenance Manual.” Nebraska Department of Roads

NYSDOT (2000). “Comprehensive Pavement Design Manual.” New York State Department of Transportation

ODOT (2010). “Pavement Distress Survey Manual.” Oregon Department of Transportation

Ooijen, V. W., Bol, V. D. (2004). “High-Speed Measurement of Raveling on Porous Asphalt.” Symposium on Pavement Surface Characteristics of Roads and Airports. Toronto

Scott, P., Radband, K., Zohrabi, M., Sanders, P., McRobbie, S., and Wright A. (2008). “Measuring Surface Disintegration (Raveling or Fretting) Using Traffic Speed Condition Surveys.” 7th International Conference on Managing Pavement Assets

TxDOT (2009). “Pavement Management Information System: Rater's Manual.” Texas Department of Transportation

WSDOT (1999). “Pavement Surface Condition Field Rating Manual for Asphalt Pavements.” Northwest Pavement Management Association

Chapter 3 Automatic Raveling Detection and Classification Method

Based on the literature review and the identified research need, Georgia Tech's research team has developed a method to automatically detect raveling using 3D laser technology and macrotexture analysis. This research has been sponsored by the NCHRP IDEA program through a research project entitled "Detecting Asphalt Pavement Raveling Using Emerging 3D Laser Technology and Macrotexture Analysis."

This chapter presents the general framework of the developed automatic raveling detection method and the classification of raveling severities and section-based aggregation method to support transportation agencies' pavement condition assessment.

The overall procedure for raveling detection and classification is as follows. First, each 3D laser data image/file (16 ft. by 12 ft. in length and width) is divided into six blocks, each of which is 5 ft. by 6 ft. Then, raveling detection and classification algorithms are applied on each block to detect and classify raveling. Based on GDOT's pavement condition survey protocol, raveling is classified as Severity Levels 0, 1, 2, and 3. Severity Level 0 means there is no raveling. After raveling is detected and classified in each block of a 3D laser data image, an image-level aggregation method will be applied to determine a single severity level for an image. After that, the raveling survey results can be aggregated for each one-mile segment based on GDOT's pavement condition survey protocol. In GDOT, only the predominant severity level of raveling in each mile is recorded. It should be noted that the intermediate results shown in the above procedures can also be used to fit in with other highway agencies' survey protocols.

1. General Framework of Proposed Raveling Detection and Classification Method

Figure 3.1 illustrates the general framework of the proposed raveling detection and classification method. 3D laser data is stored in individual files; each image covers a 16 ft. pavement section. To consider the non-uniformity of a 3D laser data image, it is divided into equal-size blocks, each of which is about 5 ft. by 6 ft. Each image is processed independently and outputs image-level raveling severity levels. The process is divided into the following four steps:

- First, a "detection" algorithm quantifies raveling by outputting a set of features, i.e. statistical characteristics of the pavement surface texture are calculated. These features are calculated after splitting the section into six blocks, three for each side of the lane.
- Second, a properly trained classifier labels each block with a severity level (0, 1, 2 or 3), given the features calculated by the detection algorithm for that block. Severity Level 0 indicates no raveling. Severity levels 1 to 3 comply with GDOT's definitions of raveling severity levels.
- Third, for each 3D laser data image, a single severity level will be assigned based on a predefined aggregate rule.
- The last step of the process is to aggregate the classified image-level labels into one-mile segment-level ratings, which are compatible with GDOT's raveling standard. For each segment, the percentage of each raveling level is generated, along with the predominant raveling level of the section.

In the developed algorithms, the most important step is to classify raveling based on pre-defined pavement surface texture features. This is implemented by using a classification (or prediction) model. Figure 3.2 shows the procedures to find a prediction model. Finding a prediction model that accurately assigns a label to new data required Georgia Tech's team to manually label a significant number of examples (i.e. ground truth data).

2. Introduction to Raveling Detection and Classification Method

The following subsections will briefly introduce the major components of the developed raveling detection and classification algorithms, including data pre-processing, raveling detection, raveling classification, and aggregation rules.

2.1 3D Laser Data Pre-Processing

Before the detection algorithms can be applied, the raw 3D laser data needs to be pre-processed. First, the invalid data points, which are indicated by invalid depth values in the data file as shown in Figure 3.3, should be removed. Second, pavement marking needs to be detected because only the portion between two pavement markings is used for raveling detection and classification. Because of their high reflectivity, pavement markings produce higher laser reflectance values, so they can be detected by using intensity data (a simple grayscale picture aligned with the 3D range data). The pixels in the pavement edge drop-off area are, also, removed because they might trigger false-positives. Figure 3.4 highlights the detected pavement marking and edge drop-off in green.

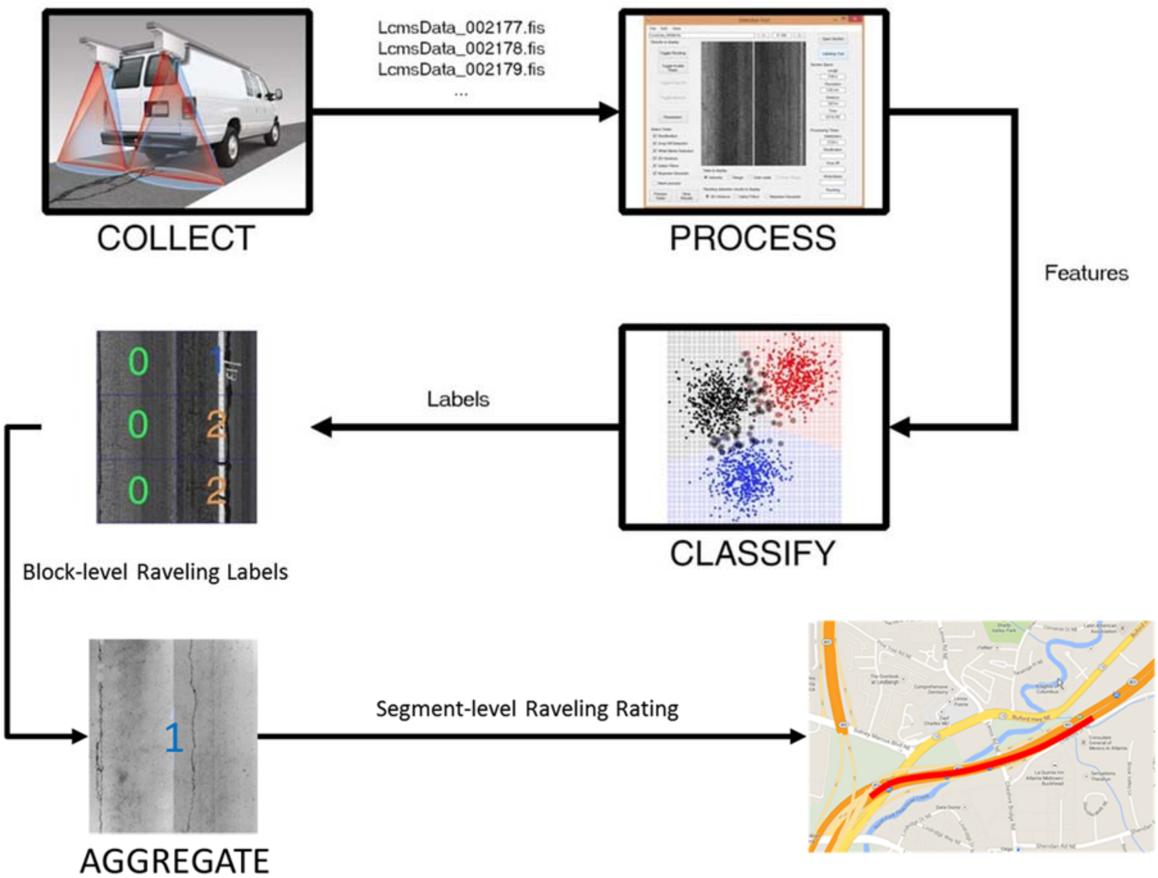


Figure 3.1 General Framework of Raveling Detection and Classification

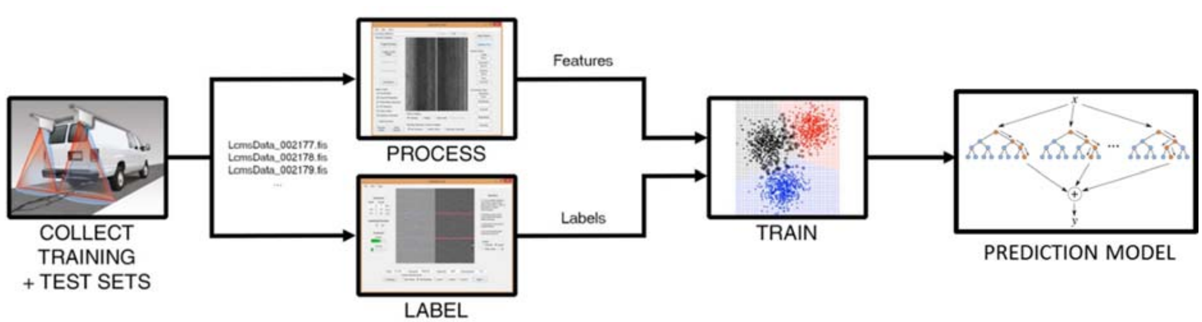


Figure 3.2 Methodology Used to Obtain Classifiers

The third required pre-processing algorithm rectifies the range data in order to eliminate the cross slope of the pavement. The curvature of the pavement surface can induce false positives and negatives. The rectification algorithm blurs the range image with a normalized box filter and subtracts the blurred image from the original. This operation removes the local mean from the data and makes edges and raveling easier to identify. Figure 3.5 illustrates this pre-processing step.

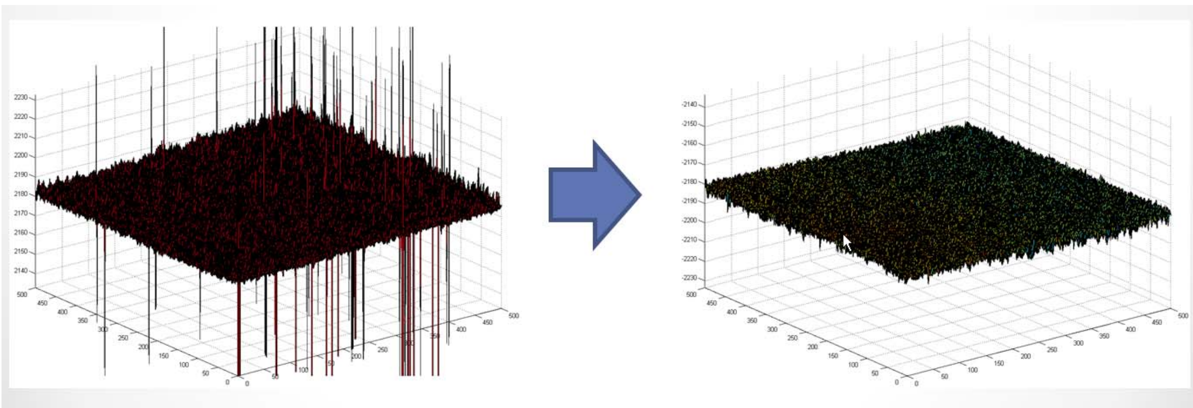


Figure 3.3 Removal of Invalid Data Point

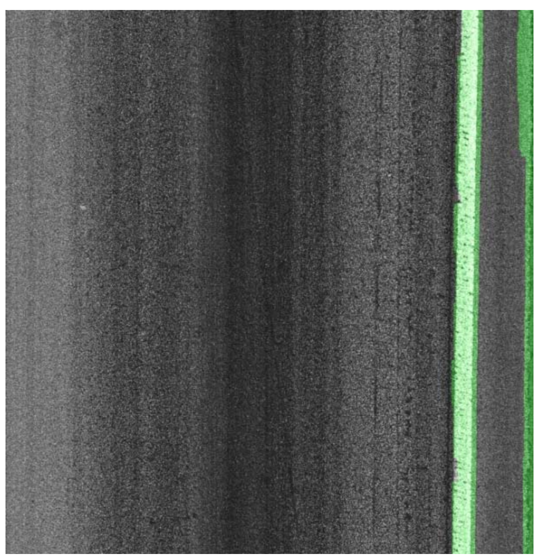


Figure 3.4 Pavement Marking and Edge Drop-off Detection

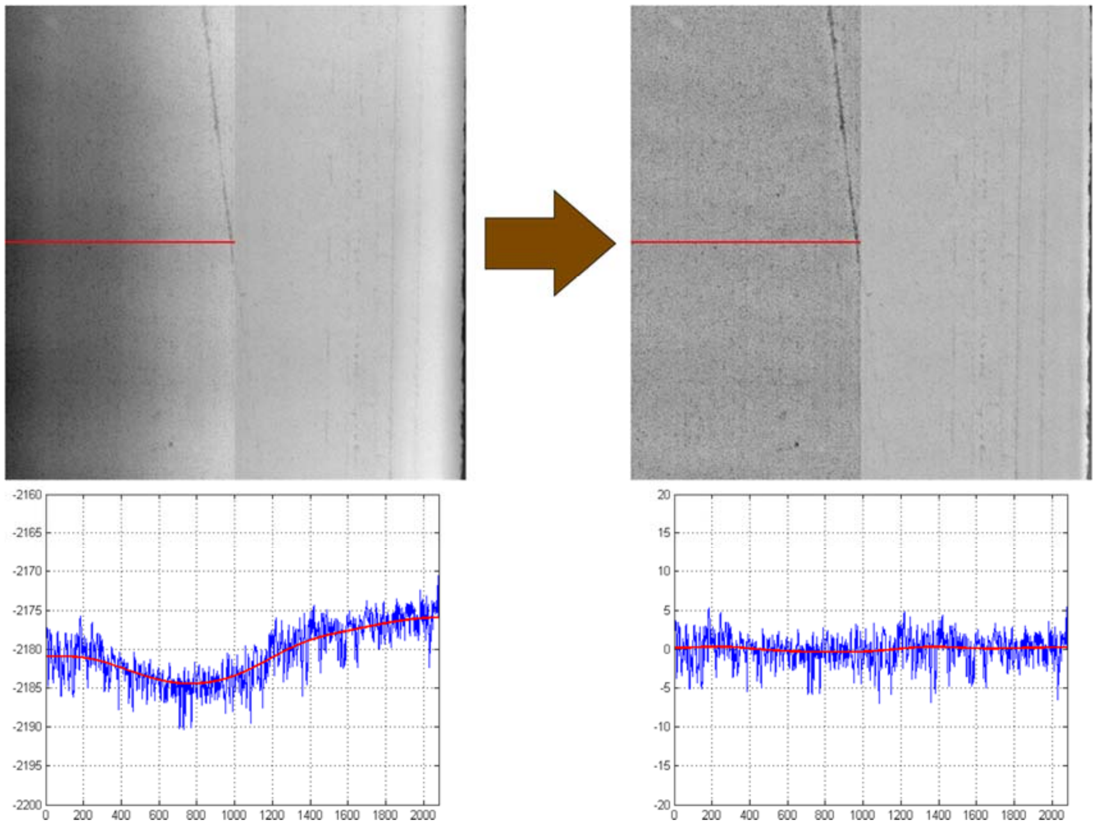


Figure 3.5 Before and After Rectification of Range Data

2.2 Raveling Detection Algorithms

As previously discussed, each 16-foot pavement section, which is stored in a data file, is divided into 6 blocks as shown in Figure 3.6. In each block, two types of statistical factors (i.e. features) are calculated based on the range data that indicate the pavement surface texture:

1) Intuitively, rougher surfaces will have more severe raveling. Thus, statistical features that can indicate the surface roughness are used; these include standard deviation of range, minimal value, maximal value, root mean square (RMS) value, and mean profile depth (MPD) etc. There are 11 statistical features extracted from each block; some of them are displayed in Figure 3.6.
2) To better capture the statistical characteristics of a raveling surface, the distribution of some indicators are approximated and applied as features. For example, the distribution of standard deviation along a block can be used to distinguish raveling levels 0 and 1

(Figure 3.7). We selected 8 indicators to get the distributions. Each distribution is represented by a histogram. Therefore, in total, 800 features are computed.

After summing up the two types of features, there are total 811 features for each block. With such a large feature vector extracted from the range data, the key is to establish a relationship between the feature vector and the true raveling severity level, which will be presented in the next sub-section.

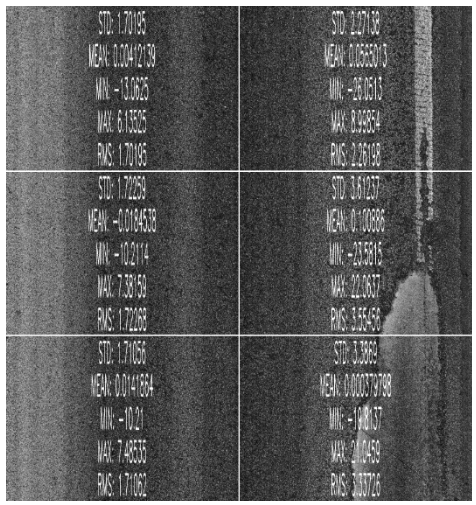


Figure 3.6 Statistical Feature Output for a Pavement Section

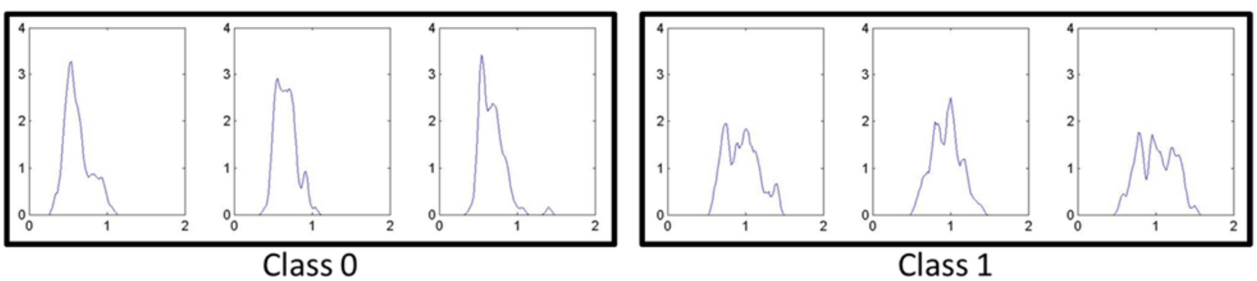


Figure 3.7 Distribution of Indicators for Pavement with Different Raveling Level

2.3 Raveling Classification Algorithms

In this sub-section, the raveling classification algorithms are presented. According to GDOT's pavement condition survey protocol, raveling is classified as Level 1, Level 2, and Level 3. For convenience, we use Level 0 to indicate the conditions of no raveling. A supervised learning technique, Random Forest (RF), was adopted in the raveling classification algorithms. Random Forest is one of the most commonly used supervised learning techniques (Breiman, 2001; Cutler, 2007). It is an ensemble learning method for classification and regression that builds many decision trees at training time and combines their output for the final prediction.

In supervised learning, a classifier is trained on a correctly (manually) labeled set (i.e. ground truth dataset). The numerical value corresponding to each labeled point is called a feature. In the proposed algorithms, 811 statistical values form a feature vector.

2.4 Raveling Aggregation Algorithms

The above algorithms detect and classify raveling for each block. The results can be used to summarize the segment-level raveling. In GDOT, each segment is about 1 mile long. This bottom-up process might not be consistent with GDOT's process, since an engineer evaluates raveling in field based on a bigger area rather than each block. Thus, an isolated block with raveling might not be counted at the segment level. Based on extensive discussion with GDOT's engineers, the following aggregation algorithms were developed to aggregate the block-level raveling into the one at the segment level. The algorithm is divided into three phases. The first phase removes outliers, such as the isolated block with raveling; the second phase smoothens the raveling distribution; the final phase performs aggregation from block-level raveling classification results to image-level and segment-level levels. The following steps describe the steps for outlier removal:

1) For a given block, compare its assigned severity level to the severity levels of its direct neighbors.
2) Each block has 5 neighbors, as shown in Figure 3.8. A neighbor can be on the next or previous image. For blocks at the boundary (first and last image), there are only 3 neighbors instead of 5.
3) If the severity level of the block is isolated among its neighbors (e.g., Level 1 surrounded by five at Level 0), then change the severity level to the majority severity level in the neighbors.
4) Repeat the above steps for all blocks. Figure 3.9 shows an example of outlier removal.

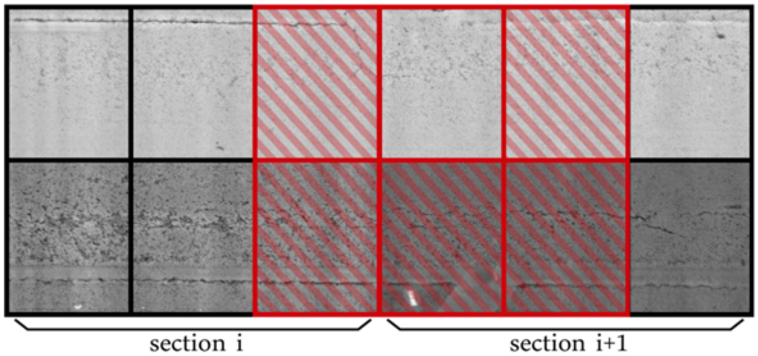


Figure 3.8 Sub-section and its Neighbors

After outliers are removed, the next step will conduct smoothness to remove isolated raveling spots. The following description describes the major steps:

1) Treat the left wheel path and right wheel path separately. For each block in a wheel path, as shown in Figure 3.10, compute a weighted average on a 2*18+1 window centered on the sub-section (18 blocks backwards, the block itself, and 18 blocks forward);
2) The weights can be those of a Gaussian or can be linear so that blocks that are further away from the current block have less influence in the weighted average. This will produce a real number in [0, 3]. Assign to the current subsection the nearest integer. 2*18+1 = 36 blocks represents 202 ft.;
3) For boundary blocks (i.e., blocks that are less than 18 positions away from the beginning or the end of the mile), use mirroring to make up for blocks that don't exist (because they are out of range). Mirroring consists of extending the length of the array by reversing the data. Example: Suppose the size of the window is 2*4+1=9; to compute the weighted average at the red position in the following array, the data would be extended such that |0|1|0|2|3|0|1|1|2|3|... becomes |1|0|0|1|0|2|3|0|1|1|2|3|... (0,1 mirrored into 1,0,0,1);

Repeat the above steps for each block. Figure 3.11 shows an example of block smoothness.

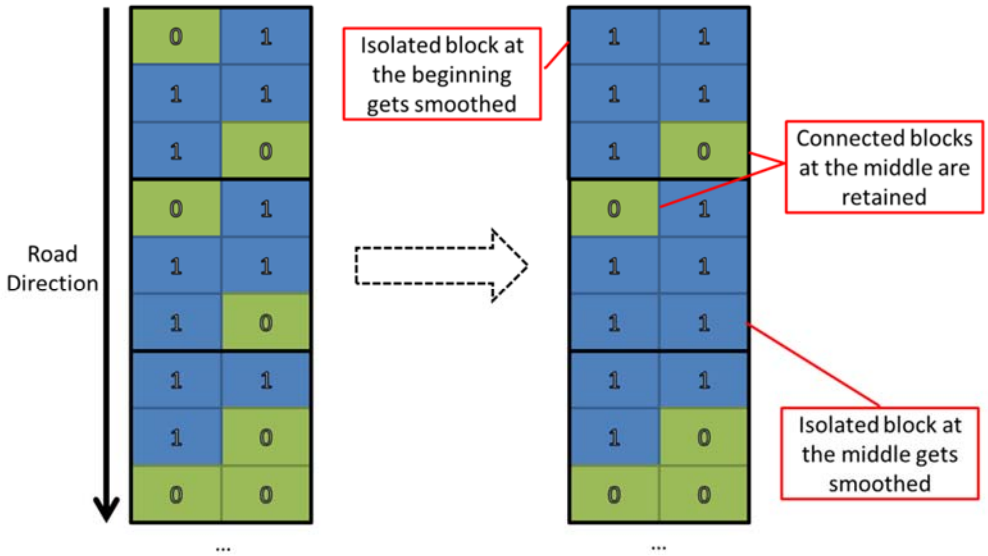


Figure 3.9 Outlier Removal Example

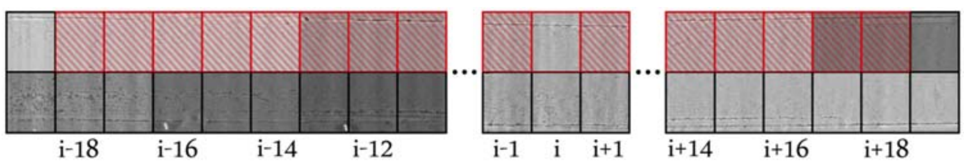


Figure 3.10 Block Smoothness

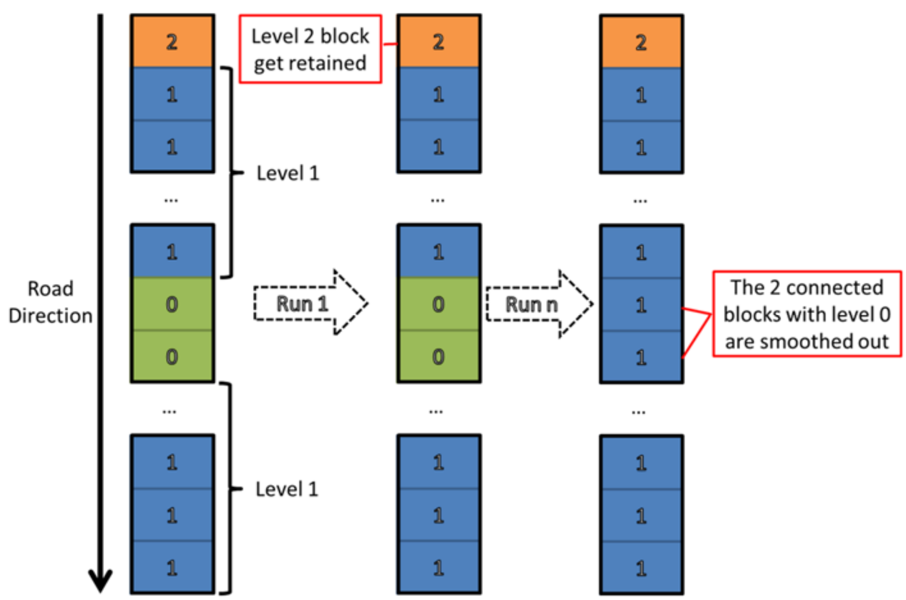


Figure 3.11 Smoothness Example

Finally, a single severity level will be assigned to each 3D laser data image through a set of aggregation rules. The following steps describe the aggregation process for each image:

1) Count the number of blocks with each level in an image,
2) If there are at least 4 (out of 6) blocks at Level 0, then the image is assigned as Level 0,
3) Otherwise, the level with the maximum number of appearances in the image is assigned as the image level.
4) When multiple levels appear the maximum number of times, the level with higher severity is chosen.
5) Sum up the aggregated image-level results in a one-mile section; get the total percentage of each raveling level. Figure 3.12 shows an example of image-based aggregation.

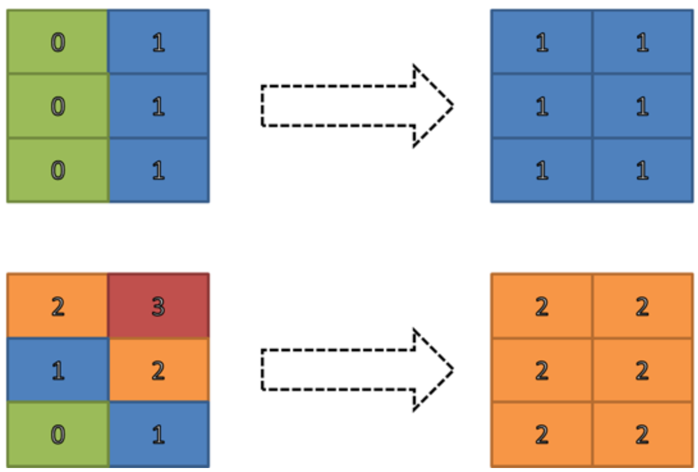


Figure 3.12 Image-level Aggregation Example

References

Breiman, L. (2001). “Random forests.” *Machine learning*, *45*(1), 5-32

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, *88*(11), 2783-2792.

Chapter 4 Testing and Validation of Developed Algorithms

The developed algorithms have been tested and validated using the 3D laser data collected on I-85 and I-285 near Atlanta, Georgia. On I-85, four 1-mile test sections were selected. In each test section, a 500-ft sample section was further marked and investigated with a GDOT pavement engineer's assistance. The aggregated test results were compared with GDOT's pavement condition survey database. On I-285, raveling detection has been conducted on the entire highway clockwise and counter clockwise. A GDOT engineer also performed an in-field validation. The following sections present the detailed test results.

1. Test Sections

1.1 I-85 Test Sections

Four 1-mile test sections were selected on I-85. In GDOT's pavement condition survey practice, the basic survey unit is about 1 mile. For example, a 10-mile survey project will be divided into ten 1-mile segments. The pavement condition survey will be conducted on each segment. For raveling, the windshield survey method was used, and the percentage of raveled sections and the predominant severity level were recorded. In the Georgia Tech test, four test sections were selected so that each severity level of raveling (including 0, i.e., no raveling) occurred. Figure 4.1 shows the four selected test sections on I-85. The following describes these four test sections:

- Test Section #1

  This test section is located on I-85 South from milepost 87 to 88. The majority of this test section has Severity Level 1 raveling, but some spots show Severity Levels 2 and 3 raveling. A 500-ft sampling section was selected for field investigation. A 50-foot-long surface layer loss can be seen clearly from the photos, as shown in Figure 4.2.
- Test Section #2

  This test section is located on I-85 South from milepost 99 to 100. The majority of this test section shows no raveling. A 500-ft sampling section was selected for field investigation. Figure 4.3 shows the typical pavement surface.

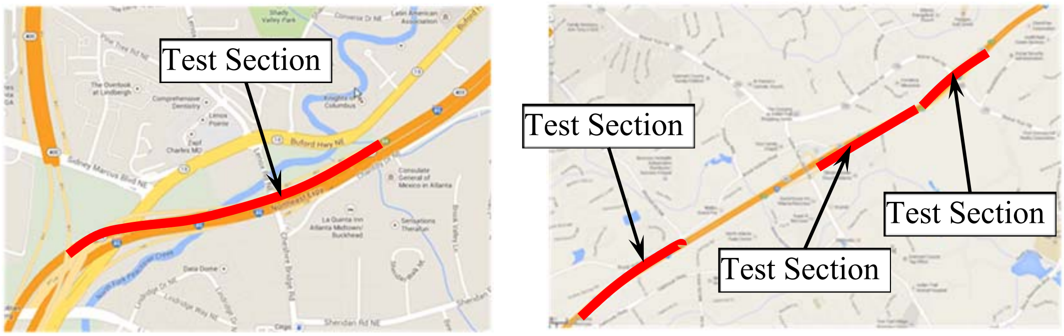


Figure 4.1 Test Sections on I-85

Figure 4.2 Severe Raveling on Test Section #1

Figure 4.3 Typical Pavement Surface on Test Section #2

- Test Section #3

This test section is located on I-85 South from milepost 101 to 102. The majority of this test section shows no raveling. A 500-ft sampling section was selected for field investigation. Figure 4.4 shows the typical pavement surface.

Figure 4.4 Typical Pavement Surface on Test Section #3

- Test Section #4

  This test section is located on I-85 South from milepost 102 to 103. The majority of this test section shows no raveling. A 550-ft sampling section was selected for field investigation. Figure 4.5 shows the typical pavement surface.

Figure 4.5 Typical Pavement Surface on Test Section #4

1.2 I-285 Test Section

I-285 is a major bypass around Atlanta for 18-wheel trucks to use rather than three major interstate highways through the heart of Atlanta, I-75, I-85, and I-20. As shown in Figure 4.6, I-285 is about 64 centerline miles. About 47.6% of pavements on I-285 are asphalt concrete (AC) pavements; the other 52.4% are Portland cement concrete (PCC) pavements.

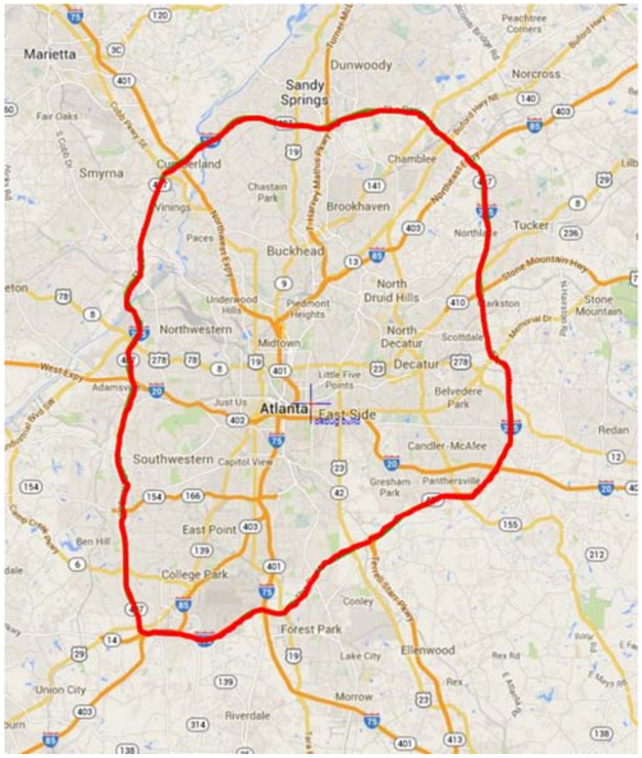


Figure 4.6 I-285 in Atlanta

To validate the developed raveling detection and classification algorithms, a large scale test has been done on the entire I-285 AC pavements in two directions. The total length is about 61 lane miles. The automatically detected results were validated by Mr. Mims using a field drive-through.

2. Collecting Ground Truth Data using a Labeling Tool

As already discussed, the developed raveling detection and classification algorithms employed a supervised machine learning method that needs to be trained using data with known raveling conditions. This type of data is well known as ground truth. The algorithm has to learn how to recognize raveling Severity Levels 0, 1, 2 and 3 from a large number of ground truth data. Therefore, the correctness and richness of ground truth are very important for the project.

Mr. Thomas Mims, a liaison engineer from GDOT, closely worked with Georgia Tech's team to establish the ground truth data. The following major steps were used in the process of establishing ground truth:

1) Data preparation. 3D laser data collected at some representative sections were used for establishing the ground truth. The sections chosen include 4 miles of asphalt pavement on I-85 and 61 miles of asphalt pavement on I-285. Sufficient raveling areas, from Severity Level 0 to level 3, were covered in the selected sections. For reference, we also collected videolog images of the pavement surface using the GTSV.
2) Field raveling survey. After picking the sections, the Georgia Tech team went to the field with Mr. Mims (Figure 4.7). By looking at raveling areas closely, the Georgia Tech team determined raveling condition of the selected sections.
3) Drive through evaluation. To validate the automatic raveling detection and classification results, the reference data was collected by the GDOT pavement expert (Mr. Mims). GDOT's pavement condition survey protocol was adopted. Table 4.1 shows the raveling percentage and the predominant severity level for each test section on I-85.

Table 4.1: Raveling Survey Conducted by GDOT

Test Section #	Percentage (%)	Predominant Severity Level
1	21	1
2	0	0
3	10	1
4	0	0

4) Manual labeling. With the knowledge provided by expert, the Georgia Tech team manually labeled the 3D laser data with different severity levels. The labeling process was repeated by several people. This allowed comparison of the manual labels among different people and identified "difficult" cases (uncertain levels) from "easy" ones (certain levels).
5) Cross checking. Again, the GDOT pavement expert helped to double-check "difficult" cases. By providing both video log (mimicking the input of the manual survey) and 3D laser data (input of automatic algorithm), the raveling level of most "difficult" cases could be decided.

Figure 4.7 Field Raveling Investigation on I-85 with GDOT Pavement Experts

To ensure the richness of the ground truth, data from 65 miles (4 miles on I-85 and 61 miles on I-285) of AC pavements to be labeled manually was selected. There were over 22,000 3D laser images to be reviewed. To perform the manual labeling process faster and more easily, two labeling applications were developed.

1) The 1st application was used for block labeling. It displayed the entire 16-ft image (left image in **Figure 4.8**) or an individual block (right image in **Figure 4.8**). It allowed users to choose a severity level and to switch between blocks using keyboard shortcuts.
2) The 2nd one was for image-level labeling. It displayed both the 3D range data (left image in Figure 4.9) and the video log image (right image in Figure 4.9). To simulate the drive-through survey for a manual raveling survey, the application can play the videolog images and 3D laser data at an adjustable speed. When a raveling section was observed, the reviewer pressed a button indicating the raveling level. All the images following are labeled as the select level until the raveling section ends or the raveling condition changes.

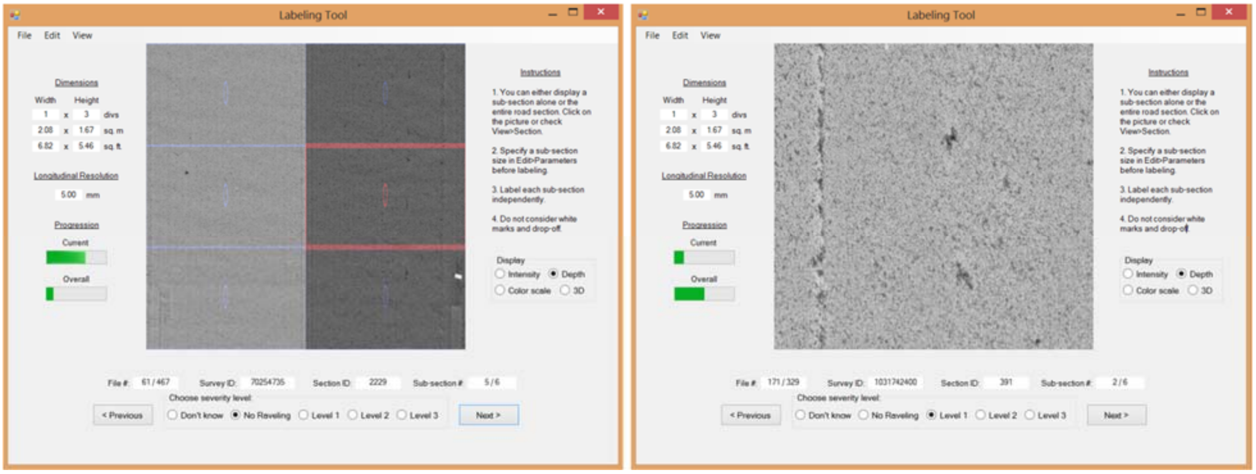


Figure 4.8 Application for Block-level Raveling Labeling

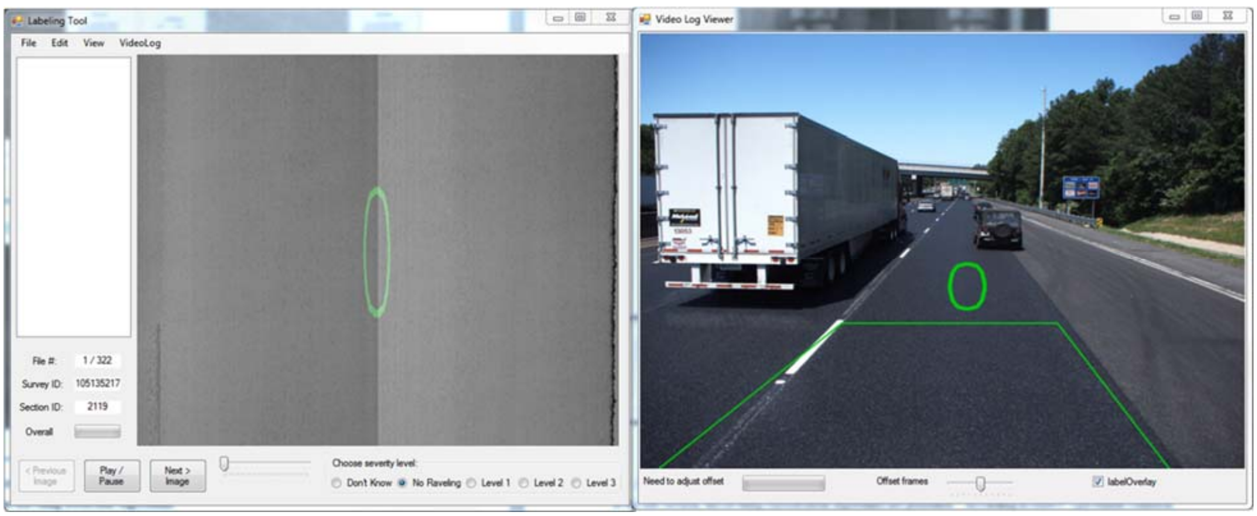


Figure 4.9 Application for Image-level Raveling Labeling

3. Testing Results on I-85

In highway agencies' practices, the details at block level are not needed for a raveling condition survey. Normally, raveling is recorded for a certain length of pavement section. In GDOT, the basic unit for raveling survey is a segment is normally 1 mile long. More importantly, in field visual investigation, an engineer often checks a large area for raveling, rather than counting all the small raveled areas.

To mimic the field visual inspection procedure and ensure the raveling condition data was consistent with the past engineering practices, an aggregation algorithm was developed to aggregate all the block-level raveling data and report the raveling conditions at segment levels (i.e., 1 mile long pavement section).

3.1 Comparison between In-office Labeling and In-field Investigation

For validation purpose, manually labeling was done for each block in the 4 test sections. After smoothing using the aggregation algorithm, segment-level ground truth was obtained. This ground truth data was acquired by in-office labeling rather than field investigation. Thus, a comparison was needed to assess its accuracy.

As shown in Table 4.2, though the predominant severity level for each test section was consistent with the field investigation result, the percentage of raveling identified in office was much less than that acquired in the field. To further validate the results, a forensic study was conducted with the GDOT pavement engineers. After careful review of every single pavement image, GDOT engineers agreed that the in-office result should be more accurate than the in-field investigation because field investigation was performed by a windshield survey. Because of the difficulties of perceiving pavement texture change accurately from a vehicle traveling at highway speed, the in-office ground truth was adopted.

Table 4.2: Ground Truth Comparison between In-office and In-filed Results

Test Section #	Test Method	Percentage (%)	Difference	Predominant Severity Level
1	In-Field	21	15.47	1
	In-Office	5.53		1
2	In-Field	0	0	0
	In-Office	0		0
3	In-Field	10	9.95	1
	In-Office	0.05		1
4	In-Field	0	0	0
	In-Office	0		0

3.2 Comparison between In-office Ground Truth and Automatic Classification Results

The following discussion compares the aggregated raveling data for each test section based on the automatic detection and classification results with the ground truth acquired from in-office labeling.

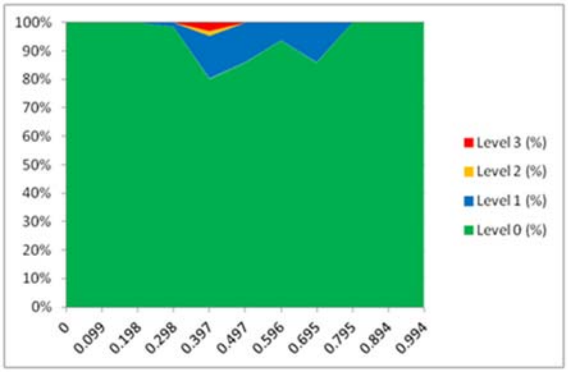
1) Test Section #1

Table 4.3 shows the validation results for Test Section #1. Based on GDOT's protocol, the predominant severity level is 1, and the total raveling percentage is 5.53% (5.07%+0.15%+0.31%). The automatic classification results show 6.87%, a difference of 1.34%. In considering the subjective factor in the manual labeling, this difference should not be significant.

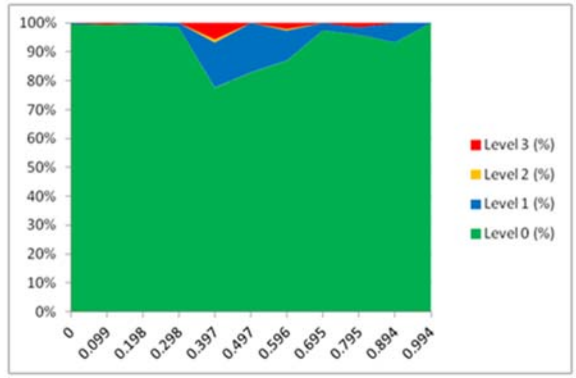
Table 4.3: Segment-level Comparison for Test Section #1

	Level 0 (%)	Level 1 (%)	Level 2 (%)	Level 3 (%)	Predominant
Ground Truth	94.46	5.07	0.15	0.31	1
Automatic Results	93.11	5.74	0.15	0.98	1
Absolute Error	1.35	0.67	0	0.67	-

Figure 4.10 compares the distribution of raveling between the ground truth and the automatic results, which were aggregated every 0.1 mile. The distribution of the automatically detected and classified raveling is close to that of the ground truth data. Thus, localized severe raveling can be identified, such as the location around the 0.397 mile.



(a) Ground Truth



(b) Automatic Results

Figure 4.10 Raveling Distribution in Test Section #1

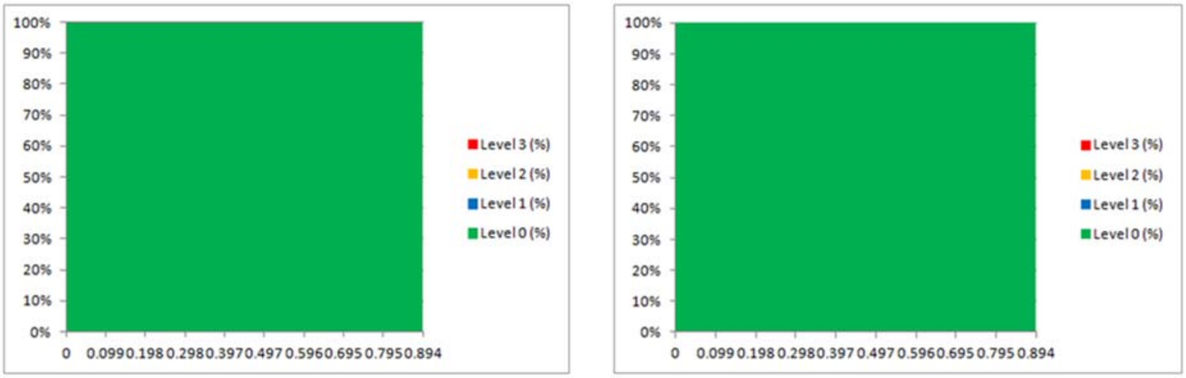
2) Test Section #2

Table 4.4 shows the validation results for Test Section #2. The automatic results exactly match the ground truth, in which no raveling appears.

Table 4.4: Segment-level Comparison for Test Section #2

	Level 0 (%)	Level 1 (%)	Level 2 (%)	Level 3 (%)	Predominant
Ground Truth	100	0	0	0	0
Automatic Results	100	0	0	0	0
Absolute Error	0	0	0	0	-

Figure 4.11 compares the distribution of raveling between the ground truth and the automatic results, which were aggregated every 0.1 mile. The distribution of the automatic detected and classified raveling is exactly same as the ground truth data.



(a) Ground Truth (b) Automatic Results

Figure 4.11 Raveling Distribution in Test Section #2

3) Test Section #3

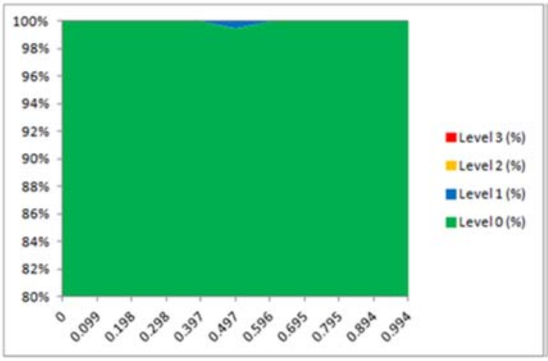
Table 4.5 shows the validation results for Test Section #3. Based on GDOT's protocol, the predominant severity level is 1; however, the raveling extent is very small, 0.05%, which is essentially 0. The automatic results show 0.15% at Severity Level 3 raveling, which is very small. In GDOT's pavement condition survey protocol, the required accuracy for raveling extent is 5%. Thus, the difference for Test Section #3 can be ignored. The automatic results can be considered as coinciding with the ground truth.

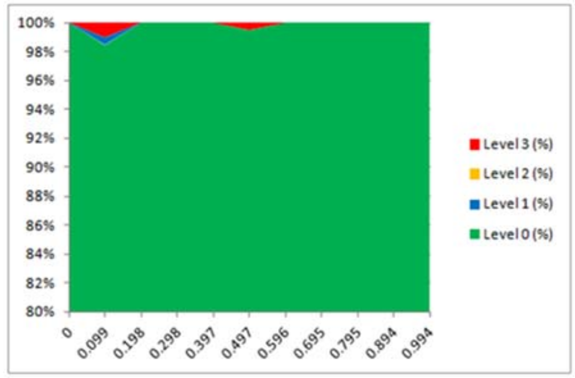
Table 4.5: Segment-level Comparison for Test Section #3

	Level 0 (%)	Level 1 (%)	Level 2 (%)	Level 3 (%)	Predominant
Ground Truth	99.94	0.05	0	0	1
Automatic Results	99.79	0.05	0	0.15	3
Absolute Error	0.15	0	0	0.15	-

Figure 4.12 compares the distribution of raveling between the ground truth and the automatic results, which were aggregated every 0.1 mile. It can be seen that the

automatically detected and classified results overestimated the raveling at mile 0.099 and 0.497 by a small value.





(a) Ground Truth (b) Automatic Results

Figure 4.12 Raveling Distribution in Test Section #3

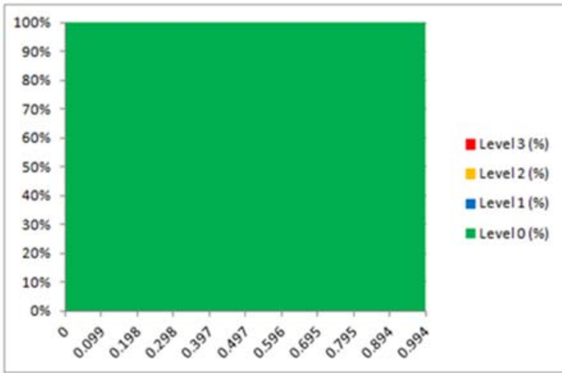
4) Test Section #4

Table 4.6 shows the validation results for Test Section #4. The ground truth shows no raveling in this section, but, the automatic results show 0.05% of Severity Level 2 raveling. Since the number is very small, the automatic results and the ground truth are, essentially, the same.

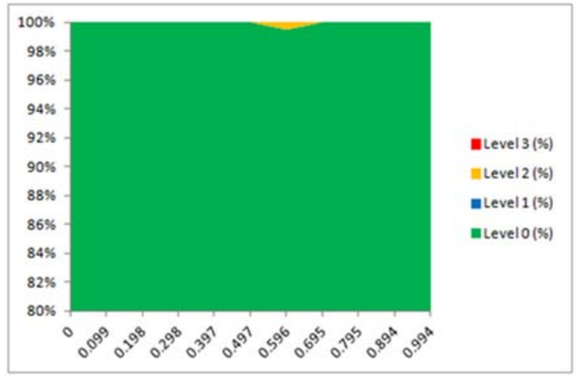
Table 4.6: Segment-level Comparison for Test Section #4

	Level 0 (%)	Level 1 (%)	Level 2 (%)	Level 3 (%)	Predominant
Ground Truth	100	0	0	0	0
Automatic Results	99.94	0	0.05	0	2
Absolute Error	0.06	0	0.05	0	

Figure 4.13 compares the distribution of raveling between the ground truth and the automatic results. The automatically detected and classified results are very close to the ground truth.



(a) Ground Truth



(b) Automatic Results

Figure 4.13 Raveling Distribution in Test Section #4

4. Testing Results on I-285

Similarly, automatic raveling detection and classification were performed on I-285. The following compares the aggregated raveling data for each test section with the ground truth acquired from in-office labeling. For better visualization, only the comparison results of test sections with raveling (concert sections and no-raveling sections are neglected) are shown. In general, all segments without raveling were 100% detected and classified. The severity level (Severity Level 1) of all raveled segments (shown in Figure 4.14 and Figure 4.15) was also 100% classified.

Figure 4.14 and Figure 4.15 compare the extents (i.e. percentage) of non-raveled and raveled portions in each segment that contains raveling with manually labeled results. In most sections, the difference in the raveling percentage between aggregation results and ground truth is around 10%. However, there are some cases in which the differences are larger than 15%, which may due to several reasons:

1) Ground truth labeling: As mentioned before, there are several "difficult" cases for manual labeling. Although they are assigned some label, the consistency between these cases is hard to ensure. There are two challenges: a) definition of severity level 1 or in the border; and b) the measurement of a mixed raveling condition, including no raveling.
2) Noise in 3D data: Other surface distresses, such as scratches and cracking, have strong effects on the raveling classification algorithm. Although some noise removal modules

have been developed and applied, it is still not possible to remove all the noise from the testing data.

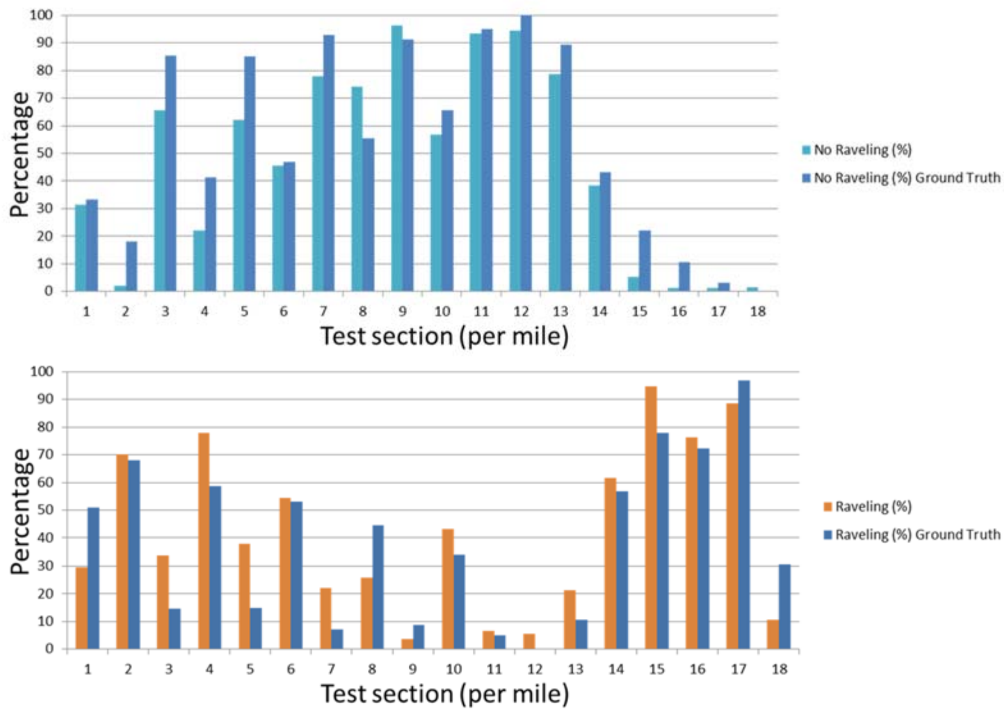


Figure 4.14 Segment-level Comparison for I-285 Clockwise Test Sites

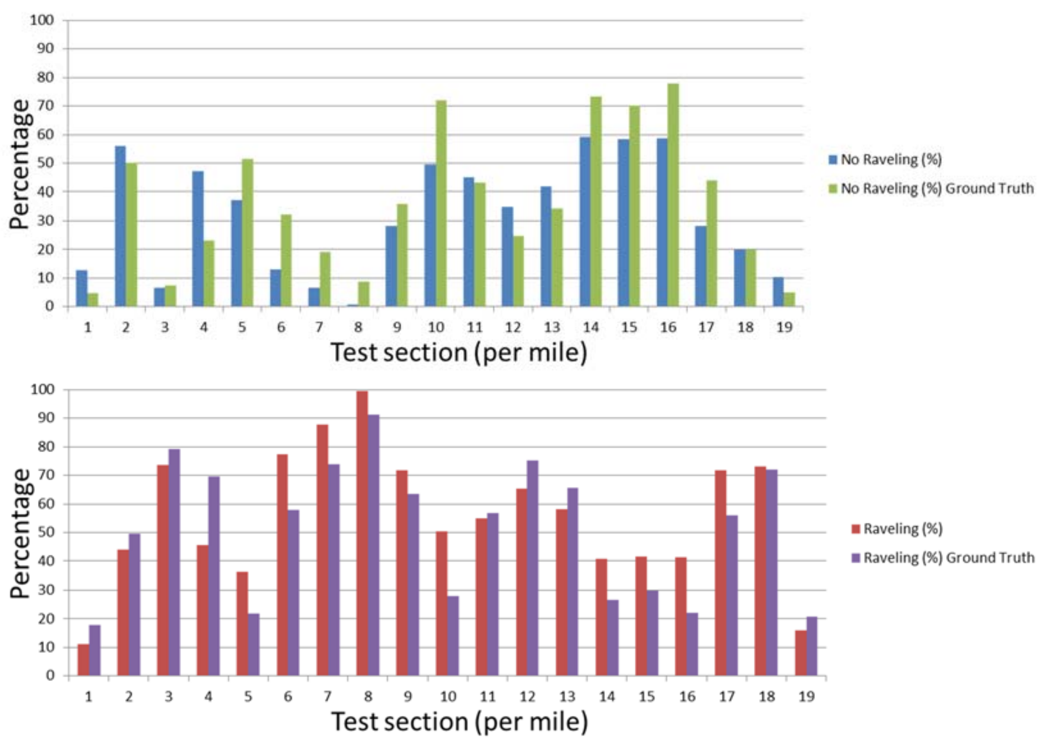


Figure 4.15 Segment-level Comparison for I-285 Counter Clockwise Test Sites

To get a better idea of the raveling aggregation results, the aggregation results of I-285 test sections on a map (Figure 4.16 and Figure 4.17) were overlaid. On the left part of the figure, a map is displayed with raveling aggregation results. Sections with red lines are classified as raveling sections (Level 1 in this case). Green sections are considered as no-raveling. On the right part of the figure, a histogram showing the raveling percentage along driving direction of I-285 is given. According to a previous field survey, the southbound lane of I-285 highway has severe Level 1 raveling. Similar observation can be seen in both the map-representation and the histogram-representation.

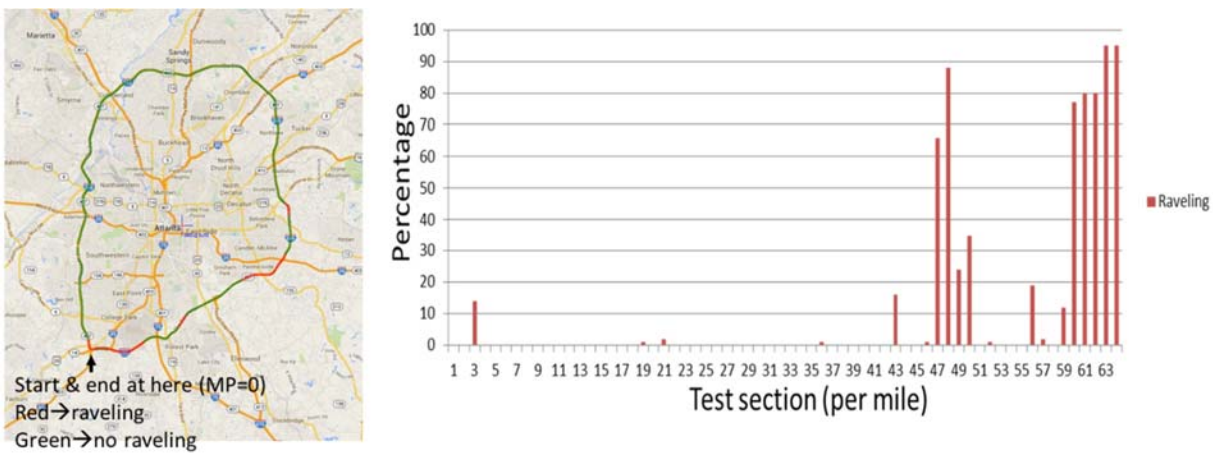


Figure 4.16 Percentage of Predominant Raveling (Level 1) along I-285 Clockwise Testing Sites

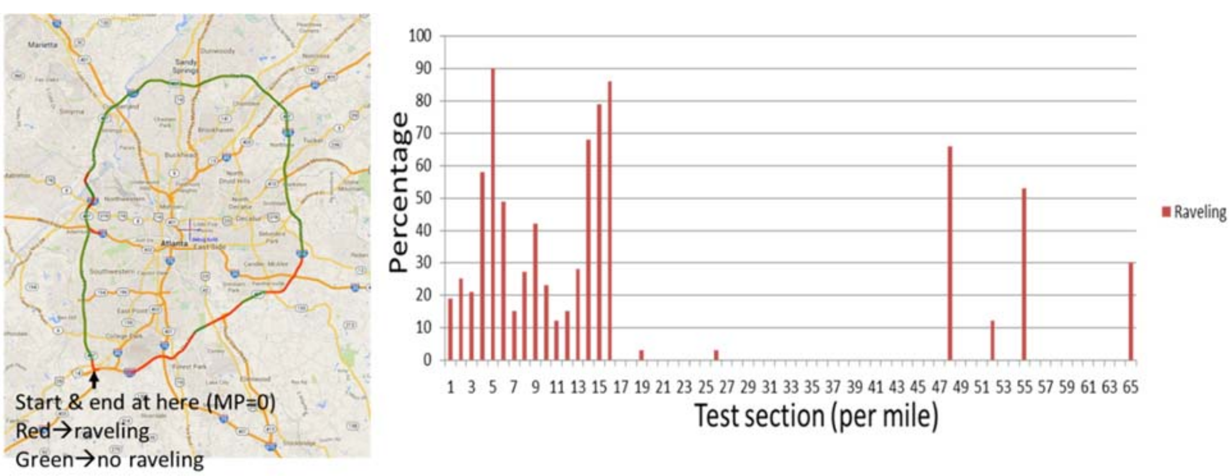


Figure 4.17 Percentage of Predominant Raveling (Level 1) along I-285 Counterclockwise Testing Site

To further check the correctness of the raveling classification results, a visualization tool can be developed to allow reviewers to see the data at different scales. The idea is shown in Figure 4.18 and Figure 4.19. On the left, an overview of the aggregation results is provided. In the middle, a zoom-in view of a 1-mile test section is displayed. At the right, a video-log image of a specific 16-ft section is given for in-detail validation. Such a tool can be very helpful for both results validation and visualization.

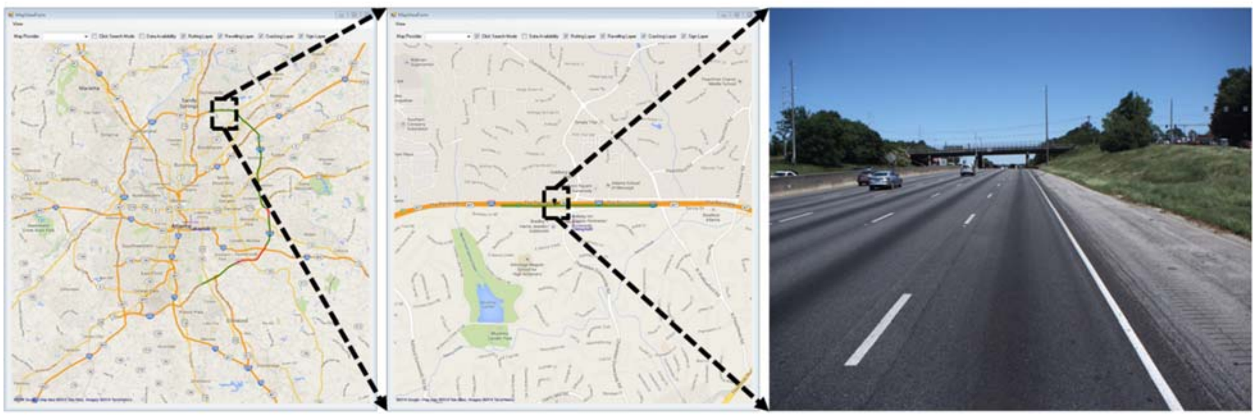


Figure 4.18 Visualization Example of No-raveling Spot

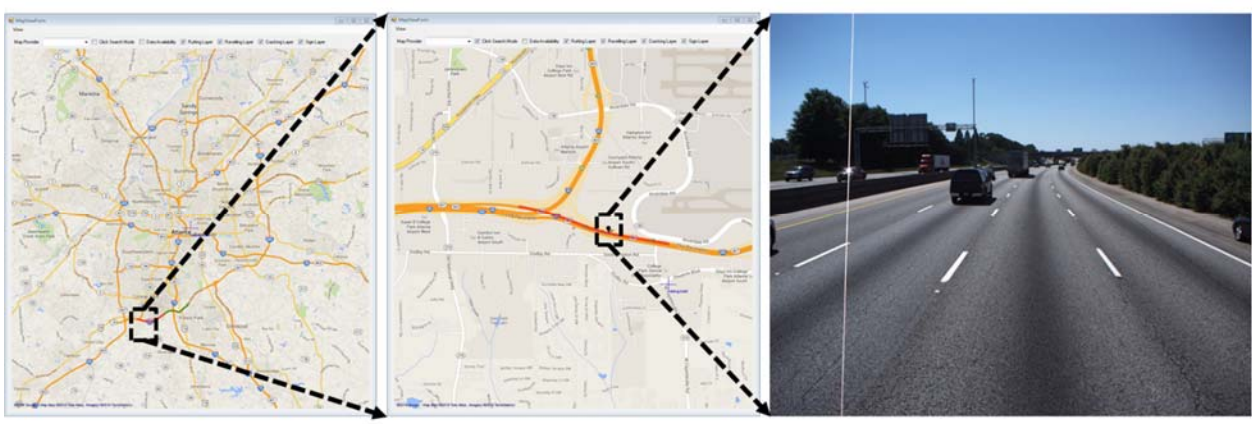


Figure 4.19 Visualization Example of Raveling Spot

5. Summary

This chapter presented the testing and validation results by using real-world data collected from the test sections on I-85 and I-285. The automatic classification results on each test sections on I-85 were compared with the ground truth. The difference of total raveled percentage on Test

Section #1 is about 1.34%; the difference is less than 0.2% for the other 3 test sections. The predominant severity levels for Test Sections #1 and #2 are also correctly classified. For Test Sections #3 and #4, since there is essentially no raveling and the classification errors are 0.15% and 0.06%, respectively, the automatic classification results are considered very close to the ground truth.

The testing on 61 lane-miles of I-285 AC pavements show promising results for automatic raveling detection and classification. All the pavements with or without raveling were 100% correctly detected and classified at the segment level; each segment is one mile long. However, due to the difficulty of correctly labeling all the raveling areas using videolog images and 3D laser data and the impact of cracking and flat-tire scratches, the raveling extent (percentage) shows a certain level of variation in comparison with the manually labeled ground truth. The difference between the surveyed results that conducted by the experienced GDOT pavement engineer and the automatically detected and measured results is less than 15% and most of them are less than 10%.

In summary, the proposed algorithms and validation results have demonstrated the promising capabilities of automatically detecting and classifying asphalt pavement raveling by taking advantage of the high-resolution, full pavement lane-width coverage, and 3D pavement surface range data that have already been collected for rutting and crack detection. It will potentially save tremendous amounts of manual effort in a field survey, improve the data accuracy, and help highway agencies make more informed decisions on pavement maintenance and rehabilitation.

Chapter 5 Conclusions and Recommendations

As one of the most common asphalt pavement distresses, raveling increases pavement roughness, which results in poor ride quality, road/tire noise, and safety issues. Besides safety concerns, including loose stones that may break windshield glass, and potential hydroplaning, raveling shortens pavement longevity. Thus, a raveling condition survey is critically needed so highway agencies know where and how severe their raveling is. Then, appropriate preservation or rehabilitation treatments can be applied.

In-field visual inspection is the most common method used for raveling condition surveys. On interstate highways, due to the large volume of traffic and large scope of distribution, an in-vehicle windshield survey is normally used. This method has the problems of the results being subjective and having large variations. Also, the survey procedure is very time-consuming and labor-intensive. Thus, there is an urgent need for an automatic survey method. In comparison with the method using digital images captured under ambient lighting conditions, laser technology has the advantage that pavement surface texture can be directly acquired. However, due to the poor resolution of point laser profilers, the related study is very limited. With the advancement of sensing technology, 3D line laser imaging technology can be employed to acquire full-lane-with high resolution pavement surface laser data, based on which raveling detection and classification algorithms can be developed. Although some raveling detection algorithms using 3D laser data have been developed, they were not validated and the classification of raveling severity was not developed either. Thus, it has become difficult for transportation agencies to implement such algorithms because an automatic raveling data collection includes raveling detection, classification, and measurement. The raveling detection and classification algorithms presented in this final report were the first ones that have been comprehensively validated using real-world, large-scale pavement data.

To address the above urgent need, the Georgia Tech research team developed new raveling detection and classification algorithms using 3D laser technology, which was sponsored by the NCHRP IDEA program. This research project further tested and validated the developed algorithms using GDOT's pavement condition survey protocol. Nevertheless, they can be easily extended to the protocols of other highway agencies. The algorithms were comprehensively tested and validated on I-85 and I-285 near Atlanta, Georgia, which is surfaced with OGFC. The following summarizes the research outcomes and major findings:

1) Using 3D laser data and the accompanied 2D intensity data, the developed raveling detection and classification algorithms consist of four major steps, a) data pre-processing, b) pavement texture feature calculation, c) block-level raveling classification, d) post-processing for raveling data aggregation.
2) To validate the developed algorithms, four test sections were selected on I-85, and the entire AC pavements were selected on I-285 in Atlanta, Georgia. A total of 65 miles (4 miles on I-85 and 61 miles on I-285) of pavement sections were selected to establish the ground truth. Working with GDOT pavement engineers, ground truth data was established through in-field survey and in-office videolog image and 3D laser data review.
3) The following are the comprehensive testing results of four test sections on I-85:
    a. In GDOT's raveling survey protocol, only the predominant severity level and the total raveled percentage is recorded. Given the fact that Severity Level 1 is the most predominant one in most cases on interstate highways, the currently trained algorithms are very accurate for GDOT's use, since the lump sum of all types of raveling is very accurate.
    b. A comparison was performed between the in-field survey and the in-office manual labeling. The results showed significant difference. Working with GDOT's experts, the Georgia Tech team carefully reviewed the entire captured 3D laser data and the videolog images. The in-office results are considered to be more accurate because the perception of pavement surface texture change by a person sitting in a vehicle at highway speed can be very different. Thus, the in-office manual labeling results were considered as the ground truth.
    c. After aggregating the classified block results, the automatic classification results on each test section were compared with the ground truth. The difference of total raveled percentage on Test Section #1 is about 1.34%, which is less than 0.2% for the other 3 test sections. The predominant severity levels for Test Sections #1 and #2 were also correctly classified. For Test Sections #3 and #4, since there is essential no raveling and the classification errors are 0.15% and 0.06%, respectively.

4) The testing on 61 lane-miles of I-285 AC pavements shows promising results of automatic raveling detection and classification. All the pavements with or without raveling were 100% correctly detected and classified at the segment level; each segment is one mile long. However, due to the difficulty of correctly labeling all the raveling area using videolog images and 3D laser data, and the impact of cracking and flat-tire scratches, the raveling extent (percentage) shows a certain level of variation in comparison with the manually labeled ground truth. The difference between the surveyed results that conducted by the experienced GDOT pavement engineer and the automatically detected and measured results is less than 15% and most of them are less than 10%.

In summary, the proposed algorithms and validation results have demonstrated the promising capabilities of automatically detecting, classifying, and measuring asphalt pavement raveling. This will potentially save tremendous amounts of manual effort for field surveys, improve the data accuracy, and help highway agencies to make more informed decisions on pavement maintenance and rehabilitation.

The following are recommendations for future research:

1) More testing and validation are suggested to evaluate the performance of the developed algorithms on pavements of different raveling conditions and different ages. The ground truth data need further study, especially for those cases that are difficult to manually recognize using videolog images and/or 3D laser data.
2) Further refinement is suggested to reduce the impact of other distresses, such as cracking and flat-tire scratches, on raveling detection and classification. It will require the detection of those unrelated distresses and performance of a removal process.
3) Beyond GDOT's pavement condition survey protocol for raveling, a finer indicator developed for raveling, e.g. percentage of aggregate loss, is recommended. The current raveling classification method (Severity Levels 1, 2 and 3) is pretty coarse for depicting the loss of aggregate on asphalt pavements, which might not be sufficient to indicate the best timing for a preventive maintenance method, e.g. fog seal. Thus, a finer indicator is desirable.

4) It is suggested an automatic raveling condition survey using the developed algorithms for all the interstate highway surfaced with OGFC be conducted. The results can, also, be imported into the current COPACES database.