Springer

Chinese Society of Pavement Engineering

International Journal of
Pavement Research and Technology

Journal homepage: www.springer.com/42947

Automatically detect and classify asphalt pavement raveling severity using 3D technology and machine learning

Yi-Chang (James) Tsai, Yipu Zhao, Bruno Pop-Stefanov, Anirban Chatterjee*

Georgia Institute of Technology, North Ave, Atlanta GA 30363, USA

Received 9 May 2020; received in revised form 19 September 2020; accepted 3 October 2020; available online 5 November 2020

Abstract

Raveling is one of the most common asphalt pavement distresses that occur on US highway pavements. Raveling results in safety concerns such as loose stones and hydroplaning; poor ride quality and road/tire noise; and shortened pavement longevity. Traditional raveling survey methods involve manual visual inspection, which is time consuming, subjective, and hazardous to highway workers. With the research project competitively selected and sponsored by the National Cooperative Highway Research Program (NCHRP) Innovation Deserving Exploratory Analysis (IDEA) program, the objective of this study is to develop an accurate raveling detection and classification algorithm using 3D pavement data that has become mainstream technologies for state Department of Transportations (DOTs) in the US for pavement condition evaluation, and to comprehensively validate these methods using large-scale, real-world data based on actual transportation agencies' distress protocol (Severity levels 1, 2, and 3). A total of 65 miles of 3D pavement data was collected on I-85 and I-285 in Georgia for training and testing. Three supervised machine learning techniques—AdaBoost with decision trees, support vector machine (SVM) and random forests—were developed for the detection and classification of raveling in the collected data. The random forest classifier had the best performance, with precision values ranging from 75.6% for level 3 raveling to 97.6% for level 0 (no) raveling and recall values ranging from 86.9% for level 1 raveling to 96.1% for level 0 raveling on real world large-scale data. The developed raveling detection and severity level classification method has been successfully implemented to entire Georgia's interstate highway system with1452.5 survey miles of asphalt pavements after the large-scale validation and refinement. The proposed method for raveling detection can be deployed to other transportation agencies for safer and more efficient assessment of roadway raveling conditions.

Keywords: Raveling; Road condition assessment; Machine learning; Decision trees; SVM; Random forest

1. Introduction

Raveling is defined as the "wearing away of the pavement surface caused by the dislodging of aggregate/stone particles and loss of asphalt binder" in the distress identification manual for the long-term pavement performance program [1], and it is one of the most common asphalt pavement distresses that occur on US highways. Fig. 1 shows some examples of raveling. Raveling will reduce pavement integrity, evenness, water tightness and skid resistance. This results in poor ride quality and road/tire noise. Raveling presents safety concerns, such as loose stones that may break windshield glass and can cause hydroplaning. Raveling also shortens pavement longevity. Thus, a raveling detection and classification is an important component of road infrastructure condition surveys for highway agencies so the appropriate preservation or rehabilitation treatment can be applied.

In most US highway agencies' practices, raveling is classified based on its severity, which is typically a qualitative definition. For example, Fig. 2 illustrates the raveling classification followed by the Georgia Department of Transportation (GDOT), which consists of three raveling severity levels [2] based on surface appearance.

It is critical for highway agencies to identify raveling in its early stage so that preventive maintenance treatments (e.g. fog seal) can

Fig. 1. Raveling on interstate highway and non-interstate road.

* Corresponding author
E-mail addresses: james.tsai@ce.gatech.edu (Y. Tsai); yipu.zhao@gatech.edu (Y. Zhao); brunopop@gatech.edu (B. Pop-Stefanov); anirban.chatterjee012@gmail.com (A. Chatterjee).
Peer review under responsibility of Chinese Society of Pavement Engineering.

ISSN: 1997-1400 DOI: https://doi.org/10.1007/s42947-020-0138-5

Fig. 2. Raveling classification in GDOT [2] (a) Severity Level 1, (b) Severity Level 2, and (c) Severity Level 3.

be applied before the pavement deteriorates further and requires more expensive treatments. However, current methods of raveling detection and classification suffer from the following shortcomings:

1. The current practices for state transportation agencies are a manual survey. The manual survey process is very time-consuming, labor-intensive, subjective, and error prone.
2. For high-traffic volume interstate highways, a manual raveling survey is often omitted by highway agencies because it is dangerous to highway workers.
3. Another approach is semi-automated surveys, where the pavement condition is recorded via vehicle-mounted cameras at highway speeds, then analyzed manually later. An video-logging based survey is unreliable because raveling is the change of pavement surface texture, and its appearance is susceptible to ambient lighting conditions. For example, under direct sunshine, it is hard to recognize lightly-raveled surfaces.
4. Automated surveys involve highway speed data collection followed by automated raveling detection and classification. Again, image processing based surveys are sensitive to lighting conditions. Algorithms for automated raveling detection and classification exist but these existing algorithms have their own shortcomings, which are discussed later in this section.

With the advancement of sensor technology, 3D laser technology with line-laser imaging and triangulation range computation, has become a mainstream technology to collect high-resolution, 3D pavement surface data. A survey in 2017 shows eighteen US state transportation agencies have used a 3D automated data collection system, and seventeen state transportation agencies said they plan to use it within two years [3]. These agencies expect to use 3D technology and 3D pavement surface data to automatically and semi-automatically extract different pavement distresses, including cracking, rutting, faulting, raveling, etc. The 3D pavement surface data has been used for detecting and measuring cracking [4,5] and its deterioration [6], rutting [7,8], concrete joint faulting [9], project-level micro-milling pavement surface texture construction quality control [10], automated raveling detection and classification [11], automatic pothole detection [12], and a new area-based faulting measurement with an enhanced accuracy [13].

Using automated surveys on 3D pavement data overcomes most of the shortcomings discussed above. Ooijen et al. [14] started to use laser data (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 in raveling detection and classification. McRobbie et al. [15-17] used laser data with 3.6 m FOV and 25 points per scan. However, 25 data points in one transverse profile (4 meter wide pavement) cannot provide sufficient data resolution to detect raveling. Laurent et al. [18,19] worked on range data with 4 m FOV and 4,096 points per scan. Laurent et al. [18,19] 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). The use of a laser scanner with longitudinal profiles collected 1mm apart enabled a much better representation of the pavement surface. However, systematic validation of the RI algorithm using a large-scale dataset is needed. In addition, raveling severity classification is still not available.

Ooijen et al. [14] developed the "Stoneway" algorithm to detect raveling on porous asphalt pavement, where raveling was detected by analyzing each longitudinal laser profile for gaps that were both above a length and depth threshold, indicating a possible loss of aggregate. It was found that this approach generally underestimated raveling severity, scheduling maintenance operations later than recommended by visual condition surveys. Additionally, this method analyzed longitudinal profiles 500 mm apart which may be too sparse to obtain the representative condition of the road. Another method based on a similar concept was developed by Scott [17]. These approaches only worked when the assumption that the pavement surface was flat (no incline) held.

McRobbie used a feature called Root Mean Square Texture (RMST) derived from 3D pavement data [15,16]. The comparison of the distribution of RMST values at two different scales provided an estimate of the raveling condition. A group of sites totaling approximately 90 km was selected, representing a combination of different surface types (thin surface course, porous asphalt, hot rolled asphalt, etc.) and surface conditions. However, the pair of scales needed to be calibrated for each site which made this method not practical for implementation. Further research is needed to obtain the pair of scales which work for all pavement types.

Mathavan et al. [20] presented a method to detect raveling from 3D pavement image (intensity and range). First, a texture descriptor method called Laws' texture energy measure is used in conjunction with the Gabor filter and other morphological operation to distinguish road areas from others. Then raveled road areas are detected by estimating the standard deviation (STD) on the corresponding range data. By heuristically setting the thresholds for STD values, the raveling condition (within a limited grid) can be characterized into good, average, or bad. However, there is a lack of comprehensive validation in this paper. Detailed information on the validation dataset, such as the location of data collection, the distribution of raveling conditions in these data, is not mentioned in this paper. Moreover, the outcome of raveling quantification is not compared with the ground truth (e.g. visual survey results) specified based on transportation agencies' distress protocol (e.g. severity levels 1, 2, 3 or low, medium and high).

An enhanced raveling detection and severity level classification method is urgently needed to replace the current visual inspection practices. Although there are some raveling detection and severity level classification algorithms developed, it still remains a technical challenge to reliably and accurately detect and classify raveling when they are validated with real-world pavement surface data based on actual transportation agencies distress protocols (e.g. raveling severity levels 1, 2, 3 or low, medium and high). Thus, it is still difficult for transportation agencies to implement any of these algorithms. Therefore, there is an urgent need to develop robust algorithms for automatic pavement raveling detection, classification, and measurement.

To address the problems in existing raveling detection and its severity level classification methods, the objectives of this study are to develop successful and effective raveling detection, classification, and measurement algorithms using 3D pavement data and macro-texture analysis, and to comprehensively validate these methods using large-scale, real-world data.

Our research group has manually classified raveling on 12 miles of diverse 3D pavement data. This allowed our research group to explore supervised machine learning based approaches for raveling classification, which generally require a large number of labeled data points for training and testing. In this study, features based on the 3D pavement data were provided to classifiers with the task to predict the GDOT raveling severity level (0, 1, 2 or 3). Multinomial logistic regression, Naive Bayes and support vector machines (SVMs) are commonly used linear classifiers. Naive Bayes classifiers assume that the features are mutually independent, which was not the case in this study. SVMs were explored in this study as they are more robust against outliers in the data than multinomial logistic regression. Among non-linear classifiers, nearest neighbors and decision trees are often used. In manual testing, nearest neighbors and k-nearest neighbors classifiers were found to be infeasible for this problem due to the presence of outliers. Decision trees are often combined with the concept of AdaBoost and random forests (RF) to improve their performance. Therefore, the scope of this study was limited to three popular classifiers: SVMs, AdaBoost with decision trees and random forests.

Without loss of generality, the developed algorithms were tested and validated using the pavement condition survey protocol in the Georgia Department of Transportation [2]. The developed raveling detection and severity level classification method was successfully implemented on Georgia's entire interstate highway system with 1452.5 survey miles of asphalt pavements after a large-scale validation and refinement. The algorithms can be extended to other highway agencies' pavement condition survey protocols by re-training the classification components using corresponding ground truth data. This paper is organized as follows. This section presents the research need and objectives. The following section presents the proposed methodology. The third section presents validation and test results. Finally, conclusions and recommendations for future research are made.

2. Methodology

2.1. Data collection

The Georgia Tech Survey Vehicle (GTSV) [21], sponsored by the US DOT, was used to collect 3D pavement surface data (Fig. 3(a)). This vehicle is equipped with a 2D imaging system, a 3D laser system (for collecting 3D pavement surface data), a mobile LiDAR system, an Inertial Measurement Unit (IMU), and a Differential GPS. IMU and GPS systems establish very high-accuracy location references. Laser profilers use laser beams with fixed directions in the frame of reference of the vehicle. As a result, laser profilers collect only a few longitudinal profiles of the pavement surface as the vehicle moves forward. In contrast, the 3D line laser imaging system uses a rapidly spinning laser beam to capture transverse profiles of the pavement surface. As the vehicle moves forward, the collected transverse profiles are stacked to provide 3D full-lane-width pavement surface data. There are 4,096 points at each transverse line, and the interval between two adjacent points is 1 mm. In the longitudinal direction, the interval between two adjacent transverse lines is 5 mm if the vehicle is operated at 100 km/hr. A sample surface texture image is visualized in Fig. 3(b). The full-lane-width-coverage 3D pavement surface data has already been used to automatically detect and measure cracking [22,23] and its deterioration [24], rutting [25,26],

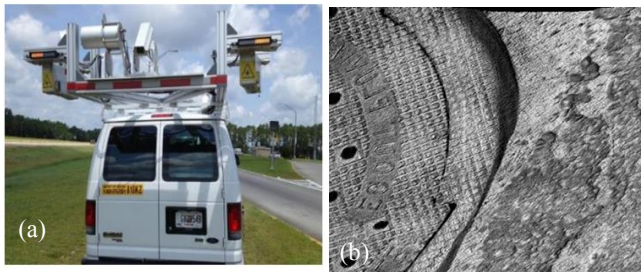

Fig. 3. The Georgia Tech Sensing Vehicle (GTSV) and 3D pavement data (a) Georgia Tech Survey Vehicle and (b) 3D Image from laser scanner.

concrete joint faulting [27], project-level micro-milling pavement surface texture construction quality control [28], and automatic pothole detection [29].

2.2. Data processing

3D pavement data is stored in individual files; each image covers a 5 m pavement section. To consider the non-uniformity of a 3D pavement image, it is divided into six equal-size sub-sections: three in each wheel path. The number of subsections in each wheel path is determined as a balance between two factors: the non-uniformity of raveling, and the manual rating effort. Each image is processed independently and outputs subsection-level raveling severity levels. The flowchart in Fig. 4 summarizes the data processing steps. The following paragraphs describe each step in detail.

2.2.1. Data preprocessing

Before the detection algorithms can be applied, the raw 3D laser data needs to be preprocessed. First, the invalid data points, which are indicated by invalid depth values in the data file are removed. Second, the 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. Therefore, they can be easily detected by using intensity data which is collected alongside the depth data. The area of the image in the pavement edge drop-off area are also removed because they might trigger false-positives.

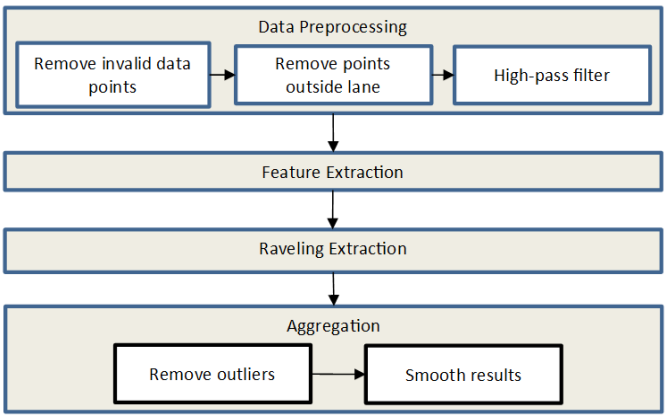

Fig. 4. The Georgia Tech Sensing Vehicle (GTSV) and 3D pavement data.

Finally, the preprocessing algorithm rectifies the range data in order to eliminate the impact of cross slope of the pavement. Asphalt pavement surfaces are curved along the transverse direction for drainage. They can also have depressions under the wheel path due to rutting. The curvature of the transverse pavement surface profile can induce false positives and negatives. A high-pass filter based on a normalized box filter is used to remove the curvature of the pavement surface. This operation removes the local mean from the data and makes edges and raveling easier to identify. Fig. 5 illustrates this preprocessing step. The blue line represents the depth of points in a transverse profile of the range image. The red line represents the depth filtered with a moving average filter to demonstrate the effect of the high-pass filter.

2.2.2. *Feature extraction*

As previously discussed, each 5-m pavement section (4 meter wide and 5 meter long), which is stored in a data file, is divided into six cells in a 2×3 grid, as show in Fig. 6. In each cell, two types of statistical factors (i.e. features) are calculated based on the range data that indicate the pavement surface texture:

1. Pavement surfaces with light raveling (e.g. severity level 1) have the isolated aggregate loss: the distributions of range data collected on these surfaces will be less uniform than pavement surface without raveling. As the raveling conditions deteriorate to severity level 2, more aggregate loss occurs and gets channelized. Therefore, the distributions of range data become non-uniform. When the pavement surfaces have severe raveling (e.g. severity level 3), the distribution of range data on these surfaces will be uniform again (since the entire surface layer is lost). Thus, the selected statistical features need to capture the characteristics of surface texture changes under different severity levels of raveling. Based on the characteristics of different raveling severity levels, 7 statistical features are selected and extracted from each cell in our study, as listed in Table 1.
2. To better capture the statistical characteristics of a raveled surface, the distribution of the indicators on small patches of a cell are calculated and applied as features. For example, the distribution of standard deviation values along all 0.1m × 0.1m patches within a cell can be used to distinguish raveling severity levels 0 and 1. Each cell had 333 patches with each patch consisting of 2,000 data points. This provided sample data points to calculate the indicators for each patch and provide a rich distribution for each indicator in each cell. The distributions of STD on raveled cell are likely to expand wider than those of non-raveled sub-sections, which might be because the raveled surface is more non-uniform. The distributions are estimated for each one of the 7 features mentioned above. The indicators are all scalar values (one-dimensional). Each distribution is discretely represented by a histogram where the distribution is discretized into 100 equally sized intervals. So each distributions consist of 100 scalar values. Therefore, in total, 700 features are computed to capture the distributions of the 7 defined features.

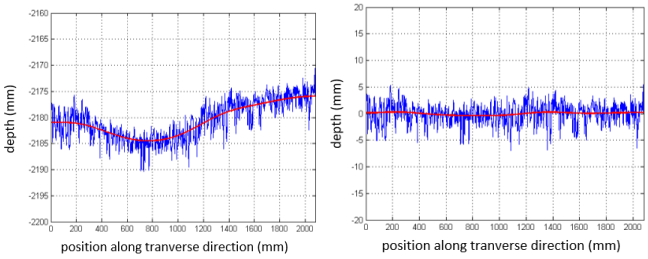

Fig. 5. Before and after rectification of range data (from left to the right figures) using high-pass filtering.

2.2.3. *Raveling classification*

According to GDOT's pavement condition survey protocol [2], raveling is classified as three types of severity levels (Level 1, Level 2, and Level 3). For convenience, we used Level 0 to indicate the conditions of no raveling. A ground truth dataset with 23,467 feature vectors (15,118 vectors for raveling severity level 0; 5,091 for raveling severity level 1; 3,053 for raveling severity level 2; and 205 for raveling severity level 3) was used for the testing and training of the classifiers. The development of the ground truth dataset is described in more detail in the next section. Three popular classification techniques using supervised

Table 1
Features estimated using range data of each cell.

Feature	Physical meaning	Dimension of statistical value	Dimension of distribution
Standard deviation	Standard deviation of range values: $\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$	1	100
Interquartile range	Distance between 75th percentile and 25th percentile range values.	1	100
Arithmetic average of absolute values	Average of absolute range values: $\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert$	1	100
Root mean square	Root of mean of square of range values: $\sqrt{\frac{1}{n}\sum_{i=1}^{n}{x_i}^2}$	1	100
Skewness	A measure of symmetry, or more precisely, the lack of symmetry of the range values: $\frac{1}{n\sigma^3}\sum_{i=1}^{n}(x_i - \bar{x})^3$	1	100
Kurtosis	A measure of whether the range values are peaked or flat relative to a normal distribution: $\frac{1}{n\sigma^4}\sum_{i=1}^{n}(x_i - \bar{x})^4$	1	100
Aggregate loss volume	Directly estimate the volume of aggregate loss by differentiating range image with reference surface (assumed to have no raveling): $\frac{1}{n}\sum_{i=1}^{n} max(\bar{x} - x_i, 0)$	1	100

Note: n = number of samples; x_i: ith sample; $\bar{x}$ = mean of samples.

learning—AdaBoost with decision trees, support vector machines (SVM) and random forests—were trained and tested for their ability to accurately provide the raveling severity level (0, 1, 2 or 3) of a cell given the features of the cell.

A decision tree classifier repeatedly splits the feature space based on the value of one of the features. This generates a flowchart of decisions, resembling a tree, where the ends (leaf nodes) belong to one of the classes. Thus, for any given input data point, the flowchart of decisions can be followed to reach a leaf node that will give the predicted class of that data point. A random forest classifier is a classifier which reports the consensus from a group of decision trees. AdaBoost with decision trees is similarly a strategy for combining the predictions from multiple decision trees based on the performance of each tree on the target classes. An SVM assumes that there exists a hyperplane in the feature space that separates two classes and classifies input data points based on which side of the hyperplane their features lie within. When there are more than two classes, by training a binary SVM classifier for each target class, the strongest positive classification (or least negative one) can be taken as the predicted class.

2.2.4. Aggregation

Each cell represents an approximately 6 ft × 6 ft pavement surface area as shown in Fig. 6. According to GDOT's raveling survey protocol, one predominant raveling severity level and the raveling extent has to be assigned to every 1 mile segment. Incorrectly classified cells can be identified and removed by checking isolated cells based on the assumption that raveling pavements are continuous to some certain extent. In addition, small spot of raveling (e.g. raveling in an isolated cell) is normally neglected in a practical survey, which would not affect the decision making on network-level pavement maintenance. Therefore, in the proposed aggregation method, an isolated cell with raveling is not counted at the segment level. Based on the extensive discussion with GDOT's engineers, the aggregation algorithms were developed to aggregate the cell-level raveling into the one at the segment level (say 1-mile). The algorithms are divided into two phases. The first phase removes outliers, such as the isolated cells with raveling and the second phase smoothens the raveling distribution. Finally, the outcomes are aggregated to 1-mile segments to support GDOT's pavement management system. The steps for outlier removal are described below:

1. For a given sub-section, compare its assigned severity level to the severity levels of its direct neighbors.
2. Each cell has 5 neighbors, as shown in Fig. 6. A neighbor can be in the next or previous image. For cells at the boundary (first and last image), there are only 3 neighbors instead of 5.
3. If the severity level of the cell is isolated among its neighbors (e.g., Level 1 surrounded by five at Level 0), then it is considered an outlier. Its severity level is changed to the majority severity level in the neighbors.
4. Repeat the above steps for all cells. Fig. 7 shows examples of outlier removal.

After outliers are removed, the next step retains a continuous, predominant raveling portion by mimicking the actual field survey practice. Based on the field survey practice, it is assumed that the length of pavement with a uniform raveling condition is approximately 200 ft. This interval can be adjusted for other state DOTs. Therefore, a window of approximately 202 ft., with 37 continuous cells, is used for smoothing. The major steps are described below:

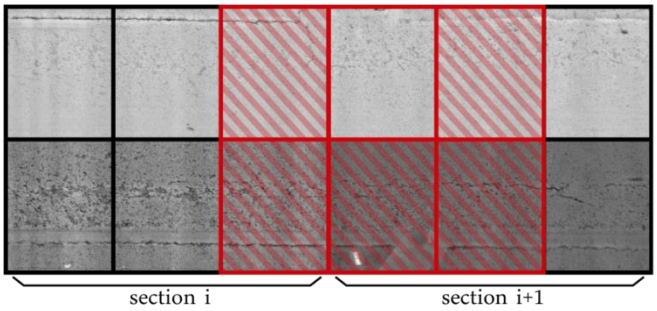

Fig. 6. Cells and its neighbors (each 5-meter image is sub-divided into 6 cells).

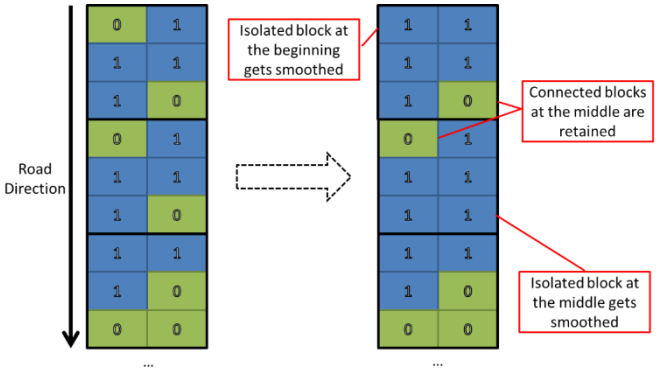

Fig. 7. Outlier removal example.

1. Treat the left wheel path and right wheel path separately. For each cell in a wheel path, as shown in Fig. 8, compute a weighted average on a (2×18)+1 window centered on the cell (18 cells backwards, the cell itself, and 18 cells forward; therefore, the total length of the window is 202 ft.).
2. The weights are defined as a Gaussian distribution along the window with 37 cells, so that the cells that are further away from the center cell have less influence in the weighted average. The generated weighted average will be a real number in [0, 3] that is further discretized into 0, 1, 2, or 3. The key parameter here is the variance of Gaussian, which determine how much the nearby section influences the raveling severity level of the center cell. The variance was determined via manual trial-and-error with the objective to minimize the difference between the resultant severity levels of the blocks and the ground truth as described in the next section.
3. For cells at boundary (i.e., cells that are less than 18 positions away from the beginning or the end of the mile), the number of cells within the side of the windows will be less than the required number of 18. In order to generate consistent weighted average on boundary cells, a technique that is commonly used in signal processing is applied here; some cells are padded over the boundary using mirroring; then the weighted average can be consistently applied over the entire wheel path. Mirroring consists of extending the length of the array by reversing the data.
4. Repeat the above steps for each cell. Fig. 9 shows an example of cell smoothness.

Finally, within each one-mile section, the outcomes are aggregated and summed–up; the total percentage of each raveling severity level is generated.

3. Validation and test results

The developed algorithms have been tested and validated using the 3D pavement data, collected on I-85 and I-285 near Atlanta, Georgia. All the asphalt pavements are open-graded friction course (OGFC), which is the most common type of pavement used in Interstate Highways in 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 the one obtained from GDOT's pavement condition database. On I-285, raveling detection was conducted on the entire highway in the clockwise and counterclockwise directions. A GDOT engineer also performed an in-field validation. To ensure the richness of the ground truth, data from 65 miles (4 miles on I-85 and 61 miles on I-285) of Asphalt Concrete pavements to be rated manually, were selected. The ground-truth data with cell level severity level

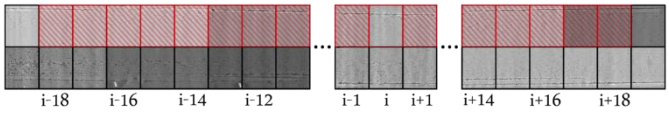

Fig. 8. Cell smoothness.

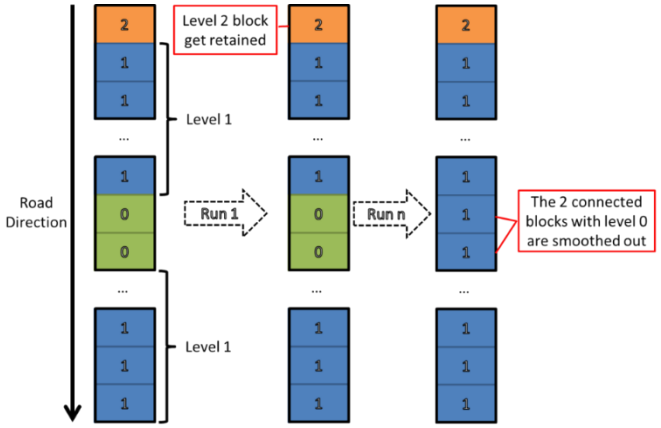

Fig. 9. Smoothness example.

labeling were then randomly divided into training and testing sets. 90% of the cells were used for training the models. The trained models were then tested on the remaining 10% of the cells which the models had never encountered beforehand.

The three classifiers were trained on this field data and used to predict the raveling severity level. The hyperparameters of each classifier were optimized using brute force search i.e. each classifier was trained several times using different hyperparameters and only the best performing trained model from each classifier was saved. The confusion matrix of all three classification techniques are given in Table 2. Each quantity in Table 2 denotes the number of cells with actual ground truth raveling severity level given by the row that were classified as the raveling severity level given by the column. The last two columns provide the precision and recall for each class. The precision and recall are defined as follows:

$$Precision\ of\ class\ x = \frac{Number\ of\ cells\ correctly\ classified\ as\ class\ x}{Total\ number\ of\ cells\ predicted\ as\ class\ x}$$

$$Recall\ of\ class\ x = \frac{Number\ of\ cells\ correctly\ classified\ as\ class\ x}{Total\ number\ of\ cells\ with\ ground\ truth\ class\ x}$$

From the precision and recall rates listed above, we can see that the performance of SVM and RF are both much better than the AdaBoost classification. Based on the fact that most of the asphalt pavements have no raveling and the rest of them have mostly with level 1 raveling, the number of misclassification cases between Class 0 (no raveling) and Class 1 (level 1 raveling) provide us a hint for the comparison between SVM and RF. RF slightly outperforms SVM in terms of the misclassification cases between Class 0 and Class 1 (577+375=952 vs. 653+359=1012). Therefore, RF is selected as the classification technique in our raveling detection and severity level classification algorithm.

The RF classifier was then used to classify the raveling severity level of pavement data collected on I-285, followed by the aggregation step to obtain a raveling extent for each 1-mile segment. In the GDOT protocol, the raveling extent is defined as the total percentage length of roadway affected by raveling of any severity level. This can be easily calculated from the proposed method by calculating the percentage of subsections classified as raveling severity level 1, 2 or 3. The raveling extent estimated on I-285 is shown in Figs. 10 and 11 (I-285 is a divided Interstate Highway which loops around Atlanta, hence the directions are

Table 2
Evaluation results of AdaBoost, SVM and random forest classification.

Classified ground truth	Class 0	Class 1	Class 2	Class 3	Precision	Recall
Adaboost						
Class 0	14,531	576	11	0	0.961	0.933
Class 1	1,014	3,388	689	0	0.665	0.833
Class 2	5	97	2,951	0	0.966	0.772
Class 3	25	8	172	0	0	0
SVM						
Class 0	14,757	359	0	2	0.976	0.956
Class 1	653	4,129	304	5	0.811	0.878
Class 2	3	208	2,817	25	0.923	0.897
Class 3	19	4	19	163	0.795	0.836
Random forest						
Class 0	14,756	375	1	4	0.976	0.961
Class 1	577	4,169	327	2	0.819	0.869
Class 2	3	241	2,791	16	0.914	0.888
Class 3	15	11	24	155	0.756	0.876

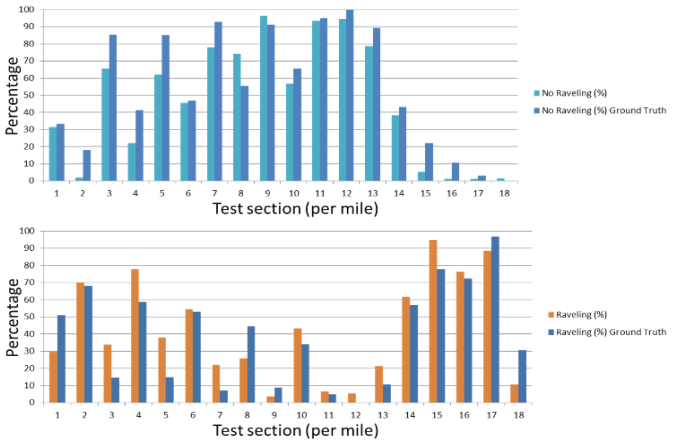

Fig. 10. Segment-level comparison for I-285 clockwise test sites (one mile per segment).

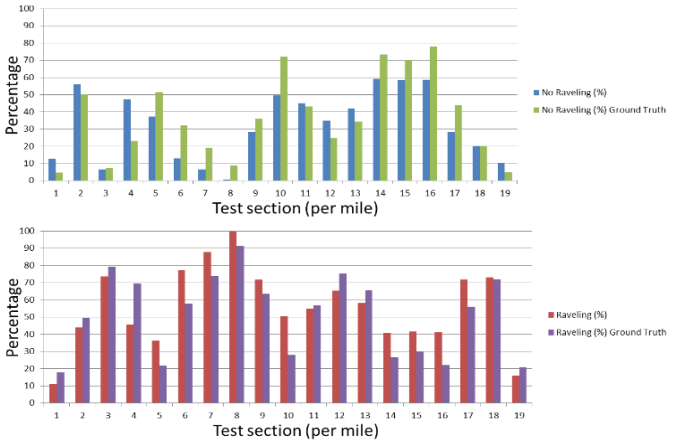

Fig. 11. Segment-level comparison for I-285 counter clockwise test sites.

clockwise and counterclockwise). It can be subjectively observed that the predicted raveling extent closely approximates the actual raveling extent. This highlights the usefulness of the proposed method from an engineering perspective. The proposed method can be used to accurately estimate the raveling severity level from a highway speed survey, independent of lighting conditions.

4. Conclusions and recommendations

An enhanced raveling detection and severity level classification method is urgently needed to replace the current visual inspection practices. Although there are some raveling detection and severity level classification algorithms developed, it still remains a technical challenge to reliably and accurately detect and classify raveling when they are validated with real-world pavement surface data with actual transportation agencies distress protocols (e.g. raveling severity levels 1, 2, 3 or low, medium and high). Thus, it is still difficult for transportation agencies to implement any of these algorithms. Therefore, there is an urgent need to develop robust algorithms for automatic pavement raveling detection, classification, and measurement. Contributions of this paper include 1) development of raveling detection, severity level classification and measurement algorithm based on actual transportation agencies' distress protocols with real-world pavement surface data, and 2) validation of raveling detection and severity level classification with real-world data with performance comparison of different raveling detection and severity level classification methods.

The objectives of this study were to develop an accurate and reliable raveling detection and classification algorithm using the 3D pavement data that has already collected by majority of state DOTs in the US for other pavement condition evaluation (e.g. cracking, rutting, faulting, etc.), and to comprehensively validate these methods using large-scale, real-world data. A total of 65 miles of 3D pavement data was collected on I-85 and I-285 in Georgia for training and testing. Three supervised machine learning techniques—AdaBoost with decision trees, SVM and random forests—were developed for the detection and classification of raveling in the collected data. The random forest classifier had the best performance, with precision values ranging from 75.6% for level 3 raveling to 97.6% for level 0 (no) raveling and recall values ranging from 86.9% for level 1 raveling to 96.1% for level 0 raveling on real world large-scale data.

The proposed algorithms have demonstrated promising capabilities to automatically detect and measure asphalt pavement raveling. The proposed algorithms can be used to first quickly assess the pavement condition. Then, manual effort can be drastically reduced by ignoring segments with little to no raveling and focusing on segments with significant raveling detected. Using the proposed algorithms will save tremendous amounts of manual effort in field surveys, improve data accuracy, and help highway agencies make more informed decisions on pavement maintenance and rehabilitation using the 3D data already available for transportation agencies in detecting cracking, rutting, faulting, etc.

The developed raveling detection and severity level classification method has been successfully implemented to entire Georgia's interstate highway system with1452.5 survey miles of asphalt pavements after the large-scale validation and refinement. The proposed method for raveling detection can be deployed to other transportation agencies for safer and more efficient assessment of roadway raveling conditions.

The recommendations for future research are as follows:

1. 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.
2. Although the developed method is promising and already been successfully implemented to detect and classify raveling on Georgia's interstate highway by the Georgia Department of Transportation, more advanced deep learning methods can be explored in the future.
3. Beyond state DOT's qualitative pavement condition survey protocol, it is recommended to develop a quantitative raveling indicator. For example, percentage of aggregate loss, is recommended. The current raveling classification method (Severity Levels 1, 2, and 3) is somewhat 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.

Acknowledgments

This research project is sponsored by the National Academy of Science National Cooperative Highway Research (NCHRP)

Innovation Deserving Exploratory Analysis (IDEA) program (NCHRP IDEA-163). The authors would like to thank the research project (Research Project 15-11: Implementation of Automatic Sign Inventory and Pavement Condition Evaluation on Georgia's Interstate Highways), sponsored by the Georgia Department of Transportation (GDOT) for us to successfully implement the developed method on the entire Georgia's interstate highway system with 1452.5 survey miles of asphalt pavements. The authors would like to thank GDOT, the Florida Department of Transportation (FDOT), the North Carolina Department of Transportation (NCDOT), the South Carolina Department of Transportation (SCDOT), the Kansas Department of Transportation (KDOT), the Department of Public Works, Nashville, Tennessee, and Fugro Roadware Inc. for their valuable inputs.

We would like to thank Mr. Timothy Dale Brantley (retired), Ms. Ernay Robinson, Mr. Eric Pitts (retired), and Mr. Thomas Mims (retired) from GDOT for their heavy involvement in selecting test sections and collecting ground truth data. We would like to thank the members of the Georgia Tech research team, Dr. Zhoahua Wang, Dr. Hadrien Glaude, Mr. Jinqi Fang, and Dr. Chengbo Ai, for their diligent work.

References

[1] Federal Highway Administration, Distress identification manual for the Long-Term Pavement Performance Project. Publication number FHWA-RD-03-031. FHWA, VA, USA, 2014.

[2] Georgia Department of Transportation, Pavement Condition Survey Manual, GDOT, GA, USA, 2007.

[3] K. Zimmerman, Pavement Management Systems: Putting Data to Work. NCHRP Synthesis 501. Transportation Research Board of the National Academies, Washington DC, USA, 2017.

[4] Y. Tsai, F. Li, Critical Assessment of Detecting Asphalt Pavement Cracks under Different Lighting and Low Intensity Contrast Conditions Using Emerging 3D Laser Technology, J. Transp. Eng. 138 (5) (2012) 649–656.

[5] C. Jiang, Y. Tsai, Enhanced Crack Segmentation Algorithm Using 3D Pavement Data, ASCE J. Comput. Civ. Eng. 30 (3) (2015) 04015050.

[6] C. Jiang, Y. Tsai, Z. Wang, Crack Deterioration Analysis Using 3D Pavement Surface Data: A Pilot Study on Georgia State Route 26, Transp. Res. Rec. 2589 (2016) 154-161.

[7] Y. Tsai, F. Li, Y. Wu, A New Rutting Measurement Method Using Emerging 3D Line-Laser Imaging System, Inter. J. Pavement Res. Technol. 6 (5) (2013) 667-672.

[8] Y. Tsai, Z. Wang, F. Li, Assessment of Rut Depth Measurement Accuracy of Point-based Rut Bar Systems using Emerging 3D Line Laser Imaging Technology, J. Marine Sci. Technol. 23 (3) (2015) 322-330.

[9] Y. Tsai, Y. Wu, C. Ai, E. Pitts, Feasibility Study of Measuring Concrete Joint Faulting Using 3D Continuous Pavement Profile Data, ASCE J. Transp. Eng. 138 (11) (2012) 1291-1296.

[10] Y. C. Tsai, Z. Wang, A Remote Sensing and GIS-Enabled Asset Management System (RS-GAMS) Phase 2. Final Report for USDOT project: RITARS-11-H-GAT. Washington DC, USA, 2014.

[11] Y. Tsai, Z. Wang, Development of an Asphalt Pavement Raveling Detection Algorithm Using Emerging 3D Laser Technology and Macrotexture Analysis. NCHRP IDEA-163 Final Report , National Academy of Science, Washington DC, USA, 2015.

[12] Y. Tsai, A. Chatterjee, Pothole Detection and Classification Using 3D Technology and Watershed Method, ASCE J. Comput. Civ. Eng. 32 (2) (2017) 04017078.

[13] G. Geary, Y. Tsai, Y. Wu, An Area-Based Faulting Measurement Method Using Three-Dimensional Pavement Data, Transp. Res. Rec. 2672 (40) (2018) 41-49.

[14] V. W. Ooijen, V. D. Bol, High-Speed Measurement of Raveling on Porous Asphalt. Symposium on Pavement Surface Characteristics of Roads and Airports, Toronto, Ontario, Canada, 2004.

[15] S. McRobbie, G. Furness, Automated Detection of Fretting on HRA Surfaces. Report no. PPR299. TRL, Berkshire, UK, 2008.

[16] S. McRobbie, J. Iaquinta, A. Wright, P. Trumper, J. Kennedy, 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. Report no. PPR628. TRL, Berkshire, UK, 2012.

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

[18] J. Laurent, J. F. Hebert, D. Lefebvre, Y. Savard, Using 3D Laser Profiling Sensors for the Automated Measurement of Road Surface Conditions, 7th RILEM International Conference on Cracking in Pavements, Delft, the Netherlands, 2012.

[19] J. Laurent, J. F. Hebert, D. Lefebvre, Y. Savard, 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, Auckland, New Zealand, 2012.

[20] S. Mathavan, M. Rahman, M. Stonecliffe-Jones, K. Kamal, Pavement Raveling Detection and Measurement from Synchronized Intensity and Range Images, Transp. Res. Rec. 2457 (2014) 3-11.

[21] Y. C. Tsai, Z. Wang, A Remote Sensing and GIS-Enabled Asset Management System (RS-GAMS). Final Report for USDOT project: DTOS59-10-H-0003. Washington DC, USA, 2013.

[22] Y. Tsai, F. Li, Detecting Asphalt Pavement Cracks under Different Lighting and Low Intensity Contrast Conditions Using Emerging 3D Laser Technology. ASCE J. Transp. Eng. 138 (5) (2012) 649–656.

[23] C. Jiang, Y. J. Tsai, Enhanced crack segmentation algorithm using 3D pavement data, J. Comput. Civ. Eng. 30 (3) (2015) 04015050.

[24] C. Jiang, Y. Tsai, Z. Wang, Crack Deterioration Analysis Using 3D Pavement Surface Data: A Pilot Study on Georgia State Route 26, Transp. Res. Rec. 2589 (2016) 154-161.

[25] Y. Tsai, F. Li, Y. Wu, A New Rutting Measurement Method Using Emerging 3D Line-Laser Imaging System, Inter. J. Pavement Res. Technol. 6 (5) (2013) 667-672.

[26] Y. Tsai, Z. Wang, F. Li, Assessment of Rut Depth Measurement Accuracy of Point-based Rut Bar Systems

using Emerging 3D Line Laser Imaging Technology, J. Marine Sci. Technol. 23 (3) (2015) 322-330.
[27] Y. Tsai, Y. Wu, C. Ai, E. Pitts, Feasibility Study of Measuring Concrete Joint Faulting Using 3D Continuous Pavement Profile Data. ASCE J. Transp. Eng. 138 (11) (2012) 1291-1296.
[28] Y. Tsai, Y. Wu, Z. Lewis, Full-Lane Coverage Micromilling Pavement-Surface Quality Control Using Emerging 3D Line Laser Imaging Technology, J. Transp. Eng. 14 (2) (2014) 04013006.
[29] Y. Tsai, A. Chatterjee, Pothole Detection and Classification Using 3D Technology and Watershed Method, ASCE J. Comput. Civ. Eng. 32 (2) (2017) 04017078.