Automated Detection of Multiple Pavement Defects

Stefania C. Radopoulou[1] and Ioannis Brilakis[2]

[1] PhD student, Division of Civil Engineering, Department of Engineering, University of Cambridge, ISG-62, Trumpington Street, Cambridge, CB2 1PZ, UK (corresponding author). E-mail: scr58@cam.ac.uk

[2] Laing O'Rourke Lecturer of Construction Engineering, Department of Engineering, University of Cambridge, BC2-07, Trumpington Street, Cambridge, CB2 1PZ, UK. Email: ib340@cam.ac.uk

Abstract:

The World Bank reports that pavement networks carry more than 80% of a country's total passenger-km and over 50% of freight ton-km, justifying the importance of efficiently maintaining pavements. Knowing the pavement condition is essential for efficiently deciding on maintenance programs. Current practice is predominantly manual with only 0.4% of inspections happening automatically. All methods in the literature aiming at automating condition assessment focus on two defects at most, or are too expensive for practical application. In this paper, we propose a low-cost method that automatically detects pavement defects simultaneously using parking camera video data. The types of defects addressed in this paper are two types of cracks, longitudinal and transverse, patches and potholes. The method uses the Semantic Texton Forests (STFs) algorithm as a supervised classifier on a calibrated region of interest (myROI), which is the area of the video frame depicting only the usable part of the pavement lane. It is validated using data collected from the local streets of Cambridge, UK. Based on the results of multiple experiments, the overall accuracy of the method is above 82%, with a precision of over

91% for longitudinal cracks, over 81% for transverse cracks, over 88% for patches and over 76% for potholes. The duration for training and classifying spans from 25 minutes to 150 minutes, depending on the number of video frames used for each experiment. The contribution of this paper is dual: 1) an automated method for detecting several pavement defects at the same time, and 2) a method for calculating the region of interest within a video frame considering pavement manual guidelines.

Keywords: pavement assessment; pavement defect; automated detection;

INTRODUCTION

The US Society of Civil Engineers and the UK Institution of Civil Engineers have each graded their country's respective pavement infrastructure with a D, emphasizing the poor condition of existing pavements (ASCE 2013; ICE 2014). A survey held in the UK regarding the country's infrastructure showed that 52% of UK businesses reported a deterioration of highways, and 77% expect the same trend for the near future (CBI and URS 2014). More than 85% of respondents believe that the bad quality of pavements is a consequence of the current maintenance procedures. A similar survey held two years earlier revealed that this is a concern for the citizens as well, as 43% identify the urgency of revising the currently-followed maintenance process (Audit Commission 2011).

Pavement condition assessment is a prerequisite for efficiently designing, planning and deciding on maintenance programs. The initial requirements for an asset management system is to be aware of the existing assets, their status and the level of service they provide (NAMS Group 2006). The Department of Transport and Highways Agency in the UK report that current pavement condition data is insufficient and gaps exist in the collected information (National Audit Office 2014). Figure 1 shows a depiction of the current practice. The colored background boxes include the name of each step of the process. The white colored background boxes include the way that each method is performed, either automatically or manually. The steps of defect identification and assessment are mainly manual, however some road authorities own software for automatically detecting and assessing cracks.

The aim of the process is to capture the longitudinal and transverse profiles of the pavement, the condition at its edges, and the texture of the surface. At first, inspectors are collecting raw data either automatically or manually. Automated data collection uses specialized vehicles that are mounted with laser scanners, pavement profilers, accelerometers, image and video cameras, and positioning systems (DfT 2011). Several US states own such vehicles for automatically detecting pavement data (Attoh-Okine and Adarkwa 2013; Liosatos 2013; Rami and Kim 2015; Richardson et al. 2015; Rick Miller 2015; Zhou et al. 2013). The number and type of sensors on those vehicles determine their purchase cost, which usually starts at approximately £500,000 (Werro 2013). The choice of sensors also drives the operational costs, which is between £20 and £40 per kilometer. Due to these high costs, the use of automated data collection is restricted to the primary pavement network and only once per year (MnDOT 2009).

In the case of the UK, the primary road network constitutes almost 20% (major and 'B' roads correspond to 50,200 miles out of 245,800 (DfT 2015)) of the total pavement length. Inspectors are driving the primary network, for inspection purposes every week of the year. Hence, 52 times a year the primary network is inspected manually. In addition, automated inspection is applied on that part of the network only once a year. The above translates into 98% (52/53) of manual inspection and 2% (1/53) of automated inspection. As for the rest of the network, it is only inspected manually. So: a) Volume of manual inspection = 80% + 20%*98% = 99.6%, and b) Volume of automated inspection = 100 - 99.6 = 0.4%

Accredited surveyors who either walk or drive (Dye Management Group, Inc 2015; PublicWorksTraining 2014; Rami and Kim 2015; UKPMS 2005) along the road perform the other 99.6% of inspections. Inspectors insert all gathered data into the road authority's central database at the end of each inspection session. Such data includes images and descriptions of road defects encountered. "Before and after" images are required for repairs conducted on the spot along with a description of actions taken. The inspector is responsible for assigning a priority rating for repair based on the level of the defect's severity, in case he/she cannot address the defect on the spot. Hence, the second and third steps of the assessment process happen at the same time when collecting data manually. Manual visual surveys are time consuming, laborious and inefficient considering the amount of network that inspectors need to cover, in conjunction with the multiple tasks that he/she has to perform.

Technicians perform the second and third steps of the process for data collected using automated methods. Multiple screens are used to project video, images, and other sensor data in order for technicians to identify the defective areas and assess their level of severity (FHWA 2003; McTavish 2012; MnDOT 2009; Zhou et al. 2013). Image and camera data is mainly used as visual aid material to assist in the defect identification and assessment. The subjectivity of the technician inevitably affects the assessment results based on the level of his/her experience, even if well-written and reliable manuals are utilized during the assessment (Bianchini et al. 2010). It is also nearly impossible to analyze the vast amounts of collected data, so only 10% is typically post-processed (MnDOT 2003).

We conclude that the current pavement condition monitoring process is laborious, time consuming and subjective based on the limitations identified above. Hence, the aim of this paper is to present a method that is free of such limitations. The contributions of this paper are: 1) an automated method for simultaneously detecting longitudinal and transverse cracks, potholes and patches, and 2) a method for calculating the region of interest within a video frame taking into consideration the sizes of defects that inspectors are looking for according to pavement inspection manuals. The following section presents the current state of research for automated defect detection. The same section also discusses methods that are useful to this paper's research objective. Section 3 details the proposed method for automatically detecting pavement defects simultaneously. Section 4 discusses the implementation process and the results from the validation of the proposed method. Finally, section 5 includes the conclusions derived from this piece of research along with a discussion regarding future work.

BACKGROUND

Research on pavement defect detection

Research has focused on automating the detection of pavement defects, in order to overcome the limitations of the current practice. Figure 2 depicts the relevant current research in a three-dimensional graph and table 1 provides a list of all relevant references. The papers found in the literature are categorized using three criteria: 1) type of defect (x-axis), 2) type of data used for analysis (y-axis), and 3) level of detail reached (z-axis). The subcategories of the z-axis are presence, detection, and measurement. Presence is the sub-category that includes methods, which answer the simple question of whether a defect exists in the given data or not. Detection is the sub-category of methods that identify the exact position of the defect within the data. Finally, measurement includes methods that are capable of providing the spatial measurements of the detected defect, such as the width and depth of a pothole.

Many methods in the literature utilize 2D images as their input. A few have focused on differentiating images that depict pavement defects from those that do not. Several methods that focus on cracks have been proposed in the literature. Some have focused on offline or real-time crack detection. Efforts have been made for classifying the different crack types, such as alligator, longitudinal or transverse. Methods were also developed for estimating the depth of a pavement crack, and for automatically sealing them. A comparison study concluded that none are comprehensive and robust. 2D image-based methods that focus on other defects, such as patches and potholes also exist in the literature.

Other methods based on 2D images use stereo vision to reconstruct the captured scene. Researchers initially tested this idea in the area of pavement reconstruction, and used it later to detect highway assets (Balali and Golparvar-Fard 2015; Uslu et al. 2011) such as guardrails and pavement markings. This method, although accurate, does not concentrate on pavement defects. Some researchers have used 3D reconstruction for understanding the pavement surface's texture and for measuring the depth of potholes to calculate the necessary filling material. Others have applied it for the purpose of detecting and classifying cracks or for calculating the crack depth (Yu et al. 2007).

Spatial data methods utilize range sensors to detect elevation defects such as rutting and shoving. These defects are not detectable in standalone images. The advantages of those methods are: 1) they are not disruptive, since the vehicle that carries the necessary equipment and performs the data collection can travel up to 100km/hr, and 2) they are insensitive to lighting conditions, which allows their application at any time of the day. These sensors are quite expensive though, which restricts their extensive/regular use in practice.

Methods that use vehicle dynamic sensor data aim at either understanding the roughness of the pavement surface or estimating the pavement profile. An accelerometer is such a sensor and its advantage is the small storage it requires for saving the collected data, which allows easy real-time processing. However, it is necessary to calibrate the vehicle with the sensors so the results are possible to compare.

In summary, no method addresses all, or even most, pavement defects simultaneously, as shown in the research cube by the empty "all defects" column. Such methods are necessary in order to address the limitations of current practice. Methods that focus on one or a few defects are appealing, but still require the manual detection of the rest. In other words, unless a method that automatically detects all types of defects at the same time is used, inspectors would need to assess the network manually. Having inspectors perform their job for some defects, while other are detected automatically invalidates the practical use of the method for cost reasons. Hence, current practice limitations remain.

Machine learning for object detection

Machine learning multi-classifier algorithms enable the simultaneous segmentation and recognition of several objects in images (Shotton et al. 2009; Uijlings et al. 2010; Zhang 2000). There are three different categories of such algorithms, and those are supervised, semi-supervised and unsupervised. Supervised are the algorithms that use multiple manually annotated data/images to train themselves how to detect certain patterns. Training images typically depict several poses of the object(s) in interest, to cover all possible appearances. Such algorithms create a codebook of visual words during training, and each word corresponds to a region of the image. This is achieved with the extraction of feature descriptors using algorithms such as SIFT

(Scale Invariant Feature Transform) (Lowe 2004) and SURF (Speed-Up Robust Features) (Bay et al. 2008).

During road condition assessment the aim is to identify road defects and distinguish them from each other. Thus, both the input and the output are known in advance. Road data is easy to find and collect, so there is no need to engage unsupervised training, which is usually meant for cases where data is insufficient or difficult to obtain. Another parameter of categorizing learning algorithms is by considering the way they are operating. This is with respect to whether they make a generalization based on the training data and build a rule for classifying new data, or whether they use all of the training data for every classification decision. The former is the so-called eager learning, whereas the latter is named lazy learning. Lazy learning techniques require a large storage space and are quite slow while classifying data, and thus are not selected for the purpose of this paper.

Artificial Neural Networks (ANNs) are a widely used family of classification algorithms (Zhang 2000) and are based on the notion of perceptrons, consisting of a large number of units (neurons) connected in different patterns. Researchers have used ANN methods for road condition related problems such as crack detection (Wu et al. 2016; Xu et al. 2008), defects and road roughness reconstruction (Ngwangwa et al. 2010) and road profile estimation (Solhmirzaei et al. 2012). The main disadvantages of the ANN methods are: 1) they are quite slow and require much time for training, 2) designing the hidden layer and its nodes is difficult because an underestimate in the number of neurons can lead to poor results (Kotsiantis et al. 2007), and 3) they underperform in noisy data.

Support Vector Machines (SVM) is another supervised classification method. The main idea of SVMs is to construct a set of hyperplanes for classifying data based on their distance from them (Wu et al. 2008). Usually, a range of potential settings are tested and cross validated to identify the best option in each problem. For that reason, SVMs have low speed in the training phase (Kotsiantis et al. 2007). On the other hand, the complexity of the model is unaffected from the number of features selected for the training phase and this constitutes a benefit of the method. They are very popular for binary classifications. However, they do not seem suitable for the classification of multiple defects.

Superpixel algorithms are quite popular recently within the computer vision community for image segmentation applications. Such algorithms segment images into groups of pixels that are meaningful atomic regions. Many approaches exist in the literature (Felzenszwalb and Huttenlocher 2004; Levinshtein et al. 2009; Veksler et al. 2010), each one with its own advantages and limitations, and the characteristics of each application define which one is the best to be applied. However, some considerations/limitations that affect the quality of a superpixel algorithm are the following: 1) many parameters need to be tuned, which can result in lost time and poor performance, 2) providing the option to specify the amount of superpixels, which isn’t a characteristic of all such algorithms, and 3) providing the ability to control the compactness (compactness refers to a regular shape and size of the superpixels along with smooth boundaries (Schick et al. 2012)) of superpixels, which is desirable but not always possible (Achanta et al. 2012).

Semantic Texton Forests (STFs) is a supervised learning algorithm (Johnson and Shotton 2010) which uses kernel features instead of feature points during classifier training. STFs consist of randomized decision forests, which are classifiers formed by several decision trees (Geurts et al. 2006). Decision trees are trained using the bag of semantic textons that is created during training. At that phase, features are extracted using a squared patch of pixels with predefined dimensions. Additionally, randomly selected subsets of features are utilized to assign a class distribution and a binary function at each tree node. The class distribution represents the probability of the tree node. The binary function is formed using the raw pixel values. The advantage of this tactic is that it ensures greater speed and avoids over-fitting (Johnson and Shotton 2010).

In general, there is no best learning technique (Kotsiantis et al. 2007; Wu et al. 2008). The No Free Lunch Theorems of Optimization (Wolpert and Macready 1997) show that a unique optimal method is impossible and the best technique always depends on the nature of the problem. Accuracy is a characteristic that is highly desirable for the aim of this paper.

From image to world coordinates

One of the types of data that inspectors collect when inspecting the pavement network is video of the lane and its surroundings. For those cases, it is useful to know the world coordinates of the objects depicted. This is achievable by projecting the objects in the video frame from the camera's optical plane to the pavement plane. This process is known as Inverse Perspective Mapping (IPM) and it has seen application in pavement lane extraction (Aly 2008; Tapia-Espinoza and Torres-Torriti 2013). IPM uses the pinhole camera model and the following assumptions in order to be constructed:

a) The world coordinate system is fixed to the vehicle; $\{x^w, y^w, z^w\}$, and

b) The camera is positioned at the rear of the vehicle (in the middle) at a specific height h from the ground and is tilted towards the pavement plane forming an angle θ_0 with an axis parallel to x^w going through the focal point.

Figure 3 depicts the IPM model and equations (1) and (2) (Tapia-Espinoza and Torres-Torriti 2013) show how to calculate the x and y coordinates of a point P in the world using its position within the image. The image plane is assumed to be of size $m \; x \; n$ pixels. The point p can be represented with the coordinate pair (u, v) when considering the reference system of the camera, where u and v are the horizontal and vertical axes of the image sensor. It can also be represented with the pair (r, c) of the standard image row-column.

In conclusion, based on the state of research, although methods that automate the detection of defects do exist, those are restricted to just one or a couple of defects at a time. Hence, the necessity of applying laborious and time-consuming manual detection methods remain. Another limitation of current methods is that some require expensive sensors for data collection, which makes them unattractive for regular usage. On the other hand, methods that use cheap sensors, such as accelerometers, are restricted to the lowest level of detail (presence) which is not enough for practitioners. Given the limitations of the current practice and state of research, we consider the following question: How can we efficiently detect most pavement defects simultaneously? Our objective for this paper is to propose such an approach.

PROPOSED SOLUTION

There are three main parts of the research question that the authors are concentrating their focus. One is the key word “efficiently”, next is “most pavement defects”, and last is “simultaneously”. In order to meet the objective of proposing an efficient solution, the authors aim to propose an approach that is both low-cost and automated. Such a method could not only be appealing to practitioners, but also easily and widely adopted. For that reason, the proposed method (figure 4 depicts a diagram of the overall vision of this research) utilizes parking cameras.

The idea of using such a sensor originates from the motivation of transforming everyday road users into ubiquitous pavement condition reporters. Parking cameras already exist in many cars, and they are gradually becoming a standardized feature, so there is no additional equipment cost required. It is also worth mentioning that all cars in the USA are mandated to have such a sensor installed by 2018 (NHTSA 2014).

One camera is not enough for capturing all pavement defects, and those related to the z-axis of the road (e.g. depressions and rutting) are particularly susceptible to this limitation. The proposed solution utilizes an additional sensor to account for this limitation, allowing detection of most defect types. Specifically, a vehicle dynamic sensor is used, which is capable of capturing defects such as pavement elevations and depressions. Additionally, a GPS device assists in the geo-tagging of all collected data in order to provide the location information of detected defects. The suggested sensors are low-cost, providing a significantly cheaper automated way of collecting data in comparison to current practice. Finally, after the detection of defects, the solution includes the automatic assessment of their severity. Both defect detection and assessment are proposed to be fully automated in contradiction to the mainly-manual current practice. The proposed system does not require any lightning support since it is designed for use under daytime fair weather conditions, which is consistent with the current practice.

This paper's scope is limited to the detection and classification of surface defects, defining how parking camera feeds are used in support of the overall solution. The black-dotted rectangle in Figure 4 provides a visual indication of how this paper's scope fits within the framework of the larger solution. For that step of the overall vision, we hypothesize that applying a supervised learning algorithm can detect several defects occurring in video frames in a more efficient way than standalone algorithms. In particular, we propose the use of Semantic Texton Forests (Johnson and Shotton 2010). The scope is restricted to the following pavement defects: longitudinal and transverse cracks, patches and potholes. However, this method can address additional defects (if trained accordingly) to cover them all when combined with vehicle dynamic sensor data. The method proposed in this paper automates the first and second steps of the pavement condition assessment which can be seen in figure 1.

RESEARCH METHODOLOGY

Pavement defects' multi-classifier

The flowchart of figure 5 depicts the research activities followed for testing the hypothesis of this paper. We initially collect pavement video data, and then process each frame separately to prepare the ground truth. This step is performed manually and it is necessary for the following step of the methodology. Ground truth video frame data include the following metadata: 1) whether they are defective, 2) the type(s) of defects they include, and 3) the location of each defect within the frame (coordinates of a polygon surrounding the defect). Once a defective frame is prepared, we save two copies for training and testing purposes. One copy is the plain image of the video frame and the other is a blank copy of the frame showing the designated defective areas. The part of the frame that corresponds to areas other than defects is marked as void. The first and second columns of figure 9 are examples of such copies. A specific color represents each defect (see table 2).

The parameters that affect the performance of the method are set before the training step. During training, the algorithm "learns" how to detect each defect. Video frames are randomly selected from the previously prepared ground truth data. Only a portion of the ground truth data is used in this step and the rest is used in the following one. At this stage, the plain image copy facilitates the identification of the characteristic features of each defect, and the copy marked with the designated defective areas directs the algorithm to search in the right part of the image. STFs perform segmentation based on bag of semantic textons that groups decision trees and act directly on the video frame pixels. Textons and priors are used as features for labeling pixels.

After the training stage, we apply the trained STFs to the rest of the video frames (the ones that have not been used in the previous stage) in order to test their performance. Both training and testing are fully automated and don't need any human intervention. The outcome of the process is segmented versions of the testing video frames produced by the algorithm. Last, we calculate the statistics by comparing the results of the STFs with the ground truth to measure the applicability of the algorithm and compare the combinations of parameters that affect its performance.

Finding the Region of Interest

Parking cameras have wide angles of view, usually greater than 90 degrees, both horizontally and vertically. For this reason, each video frame depicts more than just the travelled pavement lane. Surroundings such as the sky, following vehicles, trees, etc. are also depicted (see example in figure 6).

Since this study focuses on detecting specific types of pavement defects, the useful part of the video frame is that which depicts the pavement lane only. We are naming this area myROI (my Region of Interest), an example of which can be seen in figure 6. In order to calculate this region, the following are used: 1) Equations of IPM, 2) Camera's position and specifications (image analysis and lens' angles of view), 3) Pavement lane width, which is the other component for calculating the side boundaries of myROI, and 4) Inspection guidelines, which uses the sizes of defects that inspectors are looking for to define the upper bound of myROI.

First, the image coordinates are mapped to world coordinates using the equations of IPM. The characteristics that are used at this step are the camera's position and specifications. Then the real world distance that is represented by consecutive video frame rows is calculated. This information is then used, along with the size of defects that need to be reported based on pavement defect manuals and the width of the road that is being inspected, in order to calculate the vertices of myROI.

IMPLEMENTATION & RESULTS

Experimental setup

We collected data using two cameras: an HP Elite Webcam, chosen to simulate a low-resolution parking camera, and a Point Grey Blackfly 05S2M-CS that meets the standards of parking cameras available in the market. Research on commercially available parking cameras and car manufacturers' websites highlighted the specifications required to simulate existing parking camera models. Parking cameras typically have low resolution (maximum 0.4MP) and wide angles of view. Compared to the HP Elite, the Blackfly has higher resolution and a wider horizontal angle of view. Table 3 includes both cameras' specifications. We mounted the cameras on the test vehicle in a position consistent with car manufacturer specifications; that is on the rear of the vehicle above or below the sign plate (see figure 7). Some vehicles have the parking camera close to the trunk handle. However, we chose to position it below the sign plate. The collected videos were saved locally to the laptop used in the field. The ground truth was prepared afterwards in the office.

We used four metrics to measure the performance of the algorithm. Two metrics, overall and average accuracies, correspond to the overall performance of STFs, and the other two, average precision and area under curve, correspond to the performance of STFs in respect to each defect. The total proportion of correctly detected pixels corresponds to the overall accuracy (OA). Average accuracy (AA) refers to the average proportion of correctly detected pixels per defect. Average precision (AP) is the fraction of correctly detected pixels (True Positive, TP) over the sum of correctly and incorrectly detected pixels (False Positive, FP). The area formed when we plot TP versus FP represents the area under the curve (AuC). Good performance corresponds to high AuC.

Many parameters affect the performance of STFs, so several parameter combinations were tested. Specifically, the parameters changed at each test were the patch pixel size and the maximum depth that a tree can reach during the training of the algorithm. Tables 4 - 7 summarize the parameter combinations of each test, along with the produced results.

We performed the first round of tests (table 4) using the data collected with the HP camera. The ground truth was marked using four categories (one for each defect). In the second round of tests (table 5), which was performed using the same dataset, an additional category called “healthy pavement” was added in the ground truth data. The third round of tests (table 6) was performed using the data collected with the PG camera and the ground truth was prepared using 5 categories (4 defects and healthy pavement). Finally, we performed the last round of tests (table 7) using the data collected with the PG camera, and considering the calculated myROI. myROI was calculated using MATLAB (see figure 8). The parameters were: 1) Camera resolution - 800 x 500 pixels. As shown in figure 7, we did not position the camera in the middle of the car, but slightly left from its center (~5cm). 2) Lane width - 2.4m, and 3) Detection of transverse cracks greater than 3.175mm. All copies of video frames (both plain image and image with designated defective areas) produced during the ground truth preparation of the previous round of tests were cropped using the above calculated myROI and used for this round of tests.

In summary, the control variables tested through our experiments were: 1) Image color: color or monochrome, 2) Number of categories in ground truth data: 4 or 5, 3) STFs parameters: Patch pixel size and maximum tree depth, and 4) Use of myROI.

We collected data twice from the local streets of Cambridge, UK. Data collection was performed during daytime and the weather was sunny, cloudy or slightly rainy. The vehicle's speed was 10-15km/hr. Unexpected vibrations of the vehicle were minimal due to the low speed and did not affect the quality of the data. We saved the video data locally and post-processed it using a desktop computer (Intel Core i7 @ 3.4 GHz, 8GB Ram). The method was implemented using C# in the Visual Studio .NET framework. Right-click options and keyboard selection functions were created in order to facilitate the step of preparing the ground truth and improve the efficiency of the process. A pop-up menu was created for inserting the values of the parameters that were tested.

Results

In the first round of experiments, the OA ranged between 0.69 and 0.79, and AA ranged from 0.55 to 0.73. In the second round of experiments, where the additional category of healthy pavement was used in the preparation of the ground truth data, the OA increased to between 0.86 and 0.89. AA still remained quite low, ranging from 0.56 to 0.67. The computational cost for both rounds of experiments varied from 23 to 35 minutes. The algorithm performed better in the third round of experiments, where we used the data collected from the PG camera. OA was above 0.74 in all tests and the AA never fell below 0.7. In the final round of experiments, we considered myROI and the results were further improved. OA ranged between 0.80 and 0.88 and AA ranged between 0.71 and 0.8. The third column of figure 9 shows some examples of the derived results. The first row corresponds to an example from the first round of experiments, the second row to the second round of experiments etc. The computational cost for the third and fourth rounds of experiments varied from 120 to 150 minutes. The third and fourth round experiments were performed 5 times each in order to ensure repeatability due to the fact that video frames are randomly selected both in training and in testing. The results shown in tables 6-7 constitute the average values and variance of the results produced from all the runs of the experiments.

Tables 4-7 also show the performance of each defect individually on each test run. The best results are highlighted in each table. Several successful combinations can accurately detect longitudinal cracks. However, the best combinations are: 1) monochrome videos - 5 categories - patch pixel size of 11, and max tree depths 12 & 14, and 2) monochrome videos - 5 categories - use of myROI - patch pixel size of 9, and max tree depths 10 & 15. For transverse cracks the best combination is: monochrome videos - 5 categories - patch pixel size of 13, and max tree depths of 10 & 14. For patches, the best combination is: monochrome videos - 5 categories - use of myROI - patch pixel size of 11, and max tree depths of 10 & 14. Finally, the best combinations for detecting potholes are: 1) colored videos - 5 categories - patch pixel size of 15, and max tree depths of 10 & 14, and 2) colored videos - 5 categories - patch pixel size of 13, and max tree depths of 12 & 16. However, the following combination is worth mentioning due to its high performance: monochrome videos - 5 categories - use of myROI - patch pixel size of 15, and max tree depths of 10 & 14.

Tables 8-11 show the confusion matrix for segmentation of each defect. The confusion matrices correspond to the best performing combination of parameters based on the OA and AA. For the first round of experiments the average accuracy for region segmentation is 59%. In the second round of experiment, the average accuracy increases to 60%. In the third round of experiment the average region segmentation accuracy is 72%, and in the final round of experiments it is 74%.

CONCLUSIONS & FUTURE WORK

The current practice in pavement condition monitoring suffers from limitations such as subjectivity and time consumption. Multiple research efforts have focused on automating this task. However, all proposed methods focus on only one or a couple of defects. Even if automated methods exist for detecting some defects, the remaining defect types need to be detected manually, and the limitations and issues of the current practice remain.

In this paper, we tested the application of Semantic Texton Forests, a supervised learning algorithm, to detect several pavement defects in video frames. STFs was selected due to the multiple features it uses for segmentation, which are texture, layout and context. Superpixel algorithms were rejected because of the existing concerns regarding controlling the amount of superpixels and their compactness. Each pavement defect has its own size, which varies significantly, so it would have been very challenging or even impossible to decide on a "universal" superpixel shape and/or size to ensure compactness.

The main challenge was the preparation of the ground truth which was manual. However, the several options built in the platform for this step made it easy and quick. The idea is to test the usage of parking cameras for potentially crowdsourcing the task of pavement monitoring to everyday pavement users. We used a camera that follows vehicle manufacturer standards for parking cameras in the experiments. Several combinations of parameters, such as the patch pixel size and the max tree depth, were tested. Those parameters affect the performance of the algorithm. The built-in pop up menu for inserting the parameters affected the applicability of the method positively, since it provides a friendly user interface. Additionally, we applied the theory of Inverse Perspective Mapping for isolating the pavement lane in the video frame and restricting the application of the algorithm in that area only, while considering the size of each defect that inspectors are looking for.

The initial results of the experiments with the HP camera were quite low. This is probably due to the low resolution of the camera and the restricted information that such a camera can capture. Additionally, in that round of experiments the detection of the transverse crack was very low. This is explained by the smaller sample that was available in the data in comparison to the other defects. This shows that more samples are necessary for the algorithm to "learn" the object.

The additional information of healthy pavement in the ground truth data resulted in an improvement of the performance. This shows that the use of more categories is beneficial to the improvement of the algorithm's performance. The performance of the algorithm was even better on the data collected with the PG camera, which follows parking camera standards. This is due to the higher camera resolution. Those data also allowed the creation of a larger database. The database with the HP camera consists of 230 video frames, whereas the second one includes 546 video frames. Finally, we derived even better results when we considered myROI. This is because the algorithm was restricted to the area were defects are expected to be found. In regards to each defect detection individually, different combinations of the control variables are achieving the best performing results.

The method was slower in the experiments using the PG camera data. The difference can be explained due to the following reasons: 1) the database created with the PG camera was almost double the size of that created with the HP camera, and 2) the image resolution of the second database is higher than the first, which means that the total number of pixels is much bigger. The performance gain can be viewed in the results that the method produced. In the initial experiments the overall accuracy varies from 69% to 79%, whereas in the final experiments it improved up to 85%. The same holds for the segmentation of each region in the video frame, which has an accuracy of 59% in the first round of experiments and increases to 74% in the last one. The initial dataset proved the practicability of low-resolution cameras for the automation pavement defects. The second dataset and the produced results show the applicability of the method.

The results show that the method performs well under fair weather. STFs uses texture as one of the features for segmentation and this assists in the differentiation amongst the different defects. For example potholes are coarser than patches and the can be detected even in direct sunlight. Intensity values are also incorporated in the segmentation. Even if asphalt is already dark, the defects' intensities are usually darker and the difference assists the detection as well. Also, the results show that the method performs well when data is collected in low speeds. Hence, the concept of using parking cameras for detecting pavement defects is proved. In order though for this framework to be applied commercially, it should be tested in higher speeds and that consists part of our future work.

To conclude, STFs perform well for the detection of surface pavement defects. However, other defects such as rutting, depressions and elevations also need to be incorporated for a fully automated pavement condition monitoring method. These defects are related to the z-axis of the road profile and could be detected in vehicle dynamic sensor data as suggested in the proposed solution. The type and number of the sensors needed to capture this type of information needs to be investigated. The same holds for the positioning of those sensors on or within the vehicle. The measurement of the detected defects is also necessary for their evaluation. Although the scope of this paper is restricted to the level of detection, it could be extended to the next level of detail. However, it would be necessary to eliminate the distortion that wide angles are causing. The method is still practical, since it can direct inspectors to the spots where defects should be further investigated and save the time of searching for them. Another interesting research problem is the transfer of the collected data from the 'inspection' vehicle(s) to the pavement maintenance authority. Hence, our future work will be directed towards these additional problems.

Acknowledgements

This material is based in part upon work supported by the National Science Foundation under Grant Number 1031329. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), 2274–2282.

Adu-Gyamfi, Y. O., Okine, N. A., Garateguy, G., Carrillo, R., and Arce, G. R. (2011). "Multiresolution information mining for pavement crack image analysis." *Journal of Computing in Civil Engineering*, 26(6), 741–749.

Aly, M. (2008). "Real time detection of lane markers in urban streets." *Intelligent Vehicles Symposium, 2008 IEEE*, IEEE, 7–12.

Amarasiri, S., Gunaratne, M., and Sarkar, S. (2009). "Modeling of Crack Depths in Digital Images of Concrete Pavements Using Optical Reflection Properties." *Journal of Transportation Engineering*, 136(6), 489–499.

ASCE. (2013). "2013 report card for America's infrastructure." <http://www.infrastructurereportcard.org/> (Jul. 20, 2013).

Attoh-Okine, N., and Adarkwa, O. (2013). *Pavement Condition Surveys - Overview of Current Practices*. Project Report, Delaware Center for Transportation, Newark, DE.

Audit Commission. (2011). *Going the distance - Achieving better value for money in road maintenance*. Local government report, London, UK.

Austroads. (2011). *Pavement Rutting Measurement with a Multi-Laser Profilometer*. *Austroads Test Method AG*.

Balali, V., and Golparvar-Fard, M. (2015). "Segmentation and recognition of roadway assets from car-mounted camera video streams using a scalable non-parametric image parsing method." *Automation in Construction*, 49, Part A, 27–39.

Battiato, S., Cafiso, S., Di Graziano, A., Rizzo, L., and Stanco, F. (2006). "Pavement Surface Distress by Using Non-linear Image Analysis Techniques."

Battiato, S., Stanco, F., Cafiso, S., and Di Graziano, A. (2007). "Adaptive Imaging Techniques for Pavement Surface Distress Analysis." *Communications to SIMAI Congress*.

Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). "Speeded-Up Robust Features (SURF)." *Computer Vision and Image Understanding*, 110(3), 346–359.

Bianchini, A., Bandini, P., and Smith, D. W. (2010). “Interrater reliability of manual pavement distress evaluations.” *Journal of Transportation Engineering*, 136(2), 165–172.

Cafiso, S., Di Graziano, A., and Battiato, S. (2006). “Evaluation of pavement surface distress using digital image collection and analysis.” *Seventh International Congress on Advances in Civil Engineering*.

CBI, and URS. (2014). *Taking the long view: a new approach to infrastructure,*. Infrastructure survey, UK.

Chang, K. T., Chang, J. R., and Liu, J. K. (2005). “Detection of pavement distresses using 3D laser scanning technology.” *Proc. of the 2005 ASCE Int. Conf. on Computing in Civil Engineering*, 105.

Cheng, H. D., Shi, X. J., and Glazier, C. (2003). “Real-time image thresholding based on sample space reduction and interpolation approach.” *Journal of computing in civil engineering*, 17(4), 264–272.

Cord, A., and Chambon, S. (2012). “Automatic Road Defect Detection by Textural Pattern Recognition Based on AdaBoost.” *Computer-Aided Civil and Infrastructure Engineering*, 27(4), 244–259.

DfT, D. for T. U. (2011). “SCANNER User Guide and Specification.” <http://www.pcis.org.uk/index.php?p=6/8/0/list,0,58> (Jul. 3, 2013).

DfT, D. for T. U. (2015). *Statistical Release - Road Lengths in Great Britain 2014*.

Doumiati, M., Victorino, A., Charara, A., and Lechner, D. (2011). “Estimation of road profile for vehicle dynamics motion: experimental validation.” *American Control Conference (ACC), 2011*, 5237–5242.

Dye Management Group, Inc. (2015). “Level of Service Condition Assessments - Data Collection Manual.” Alabama Department of Transportation.

Felzenszwalb, P. F., and Huttenlocher, D. P. (2004). “Efficient Graph-Based Image Segmentation.” *International Journal of Computer Vision*, 59(2), 167–181.

FHWA. (2003). *Distress Identification Manual for the Long-Term Pavement Performance Program*. Federal Highway Administration.

Gavilán, M., Balcones, D., Marcos, O., Llorca, D. F., Sotelo, M. A., Parra, I., Ocaña, M., Aliseda, P., Yarza, P., and Amírola, A. (2011). “Adaptive road crack detection system by pavement classification.” *Sensors*, 11(10), 9628–9657.

Georgieva, K., Koch, C., and König, M. (2015). “Wavelet Transform on Multi-GPU for Real-Time Pavement Distress Detection.” *Computing in Civil Engineering 2015*, ASCE, 99–106.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). “Extremely randomized trees.” *Machine learning*, 63(1), 3–42.

Ghanta, S., Birken, R., and Dy, J. (2012). “Automatic road surface defect detection from grayscale images.” *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, 83471E–83471E.

González, A., O’brien, E. J., Li, Y.-Y., and Cashell, K. (2008). “The use of vehicle acceleration measurements to estimate road roughness.” *Vehicle System Dynamics*, 46(6), 483–499.

Haas, C. (1996). “Evolution of an automated crack sealer: a study in construction technology development.” *Automation in construction*, 4(4), 293–305.

Harris, N. K., Gonzalez, A., OBrien, E. J., and McGetrick, P. (2010). “Characterisation of pavement profile heights using accelerometer readings and a combinatorial optimisation technique.” *Journal of Sound and Vibration*, 329(5), 497–508.

Huang, Y., and Xu, B. (2006). “Automatic inspection of pavement cracking distress.” *Journal of Electronic Imaging*, 15(1), 013017–013017.

ICE. (2014). *The state of the nation - Infrastructure 2014*. Institution of Civil Engineers.

Imine, H., and Delanne, Y. (2005). “Triangular observers for road profiles inputs estimation and vehicle dynamics analysis.” *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 4751–4756.

Imine, H., Delanne, Y., and M’sirdi, N. K. (2006). “Road profile input estimation in vehicle dynamics simulation.” *Vehicle System Dynamics*, 44(4), 285–303.

Imine, H., M Sirdi, N. K., and Delanne, Y. (2003). “Adaptive observers and estimation of the road profile.” *SAE SP*, 175–180.

Islam, S., Buttlar, W., Aldunate, R., and Vavrik, W. (2014). “Measurement of Pavement Roughness Using Android-Based Smartphone Application.” *Transportation Research Record: Journal of the Transportation Research Board*, 2457, 30–38.

Jahanshahi, Mohammad R., Jazizadeh, Farrokh, Masri, Sami F., and Becerik-Gerber Burcin. (2013). “Unsupervised Approach for Autonomous Pavement-Defect Detection and Quantification Using an Inexpensive Depth Sensor.” *Journal of Computing in Civil Engineering*, 27(6), 743–754.

Jiang, C., and Tsai, Y. J. (2015). “Enhanced Crack Segmentation Algorithm Using 3D Pavement Data.” *Journal of Computing in Civil Engineering*, 04015050.

Jing, L., and Aiqin, Z. (2010). “Pavement crack distress detection based on image analysis.” *Machine Vision and Human-Machine Interface (MVHI), 2010 International Conference on*, 576–579.

Jog, G. M., Koch, C., Golparvar-Fard, M., and Brilakis, I. (2012). “Pothole Properties Measurement through Visual 2D Recognition and 3D Reconstruction.” *Computing in Civil Engineering (2012)*, 553–560.

Johnson, M., and Shotton, J. (2010). “Semantic texton forests.” *Computer Vision*, Springer, 173–203.

Johnsson, R., and Odelius, J. (2012). “Methods for road texture estimation using vehicle measurements.”

Kamaliardakani, M., Sun, L., and Ardakani, M. K. (2014). “Sealed-crack detection algorithm using heuristic thresholding approach.” *Journal of Computing in Civil Engineering*, 30(1), 04014110.

Kaul, V., Tsai, Y., and Mersereau, R. (2010). “Quantitative Performance Evaluation Algorithms for Pavement Distress Segmentation.” *Transportation Research Record: Journal of the Transportation Research Board*, 2153, 106–113.

Kim, Y. S., Yoo, H. S., Lee, J. H., and Han, S. W. (2009). “Chronological development history of X–Y table based pavement crack sealers and research findings for practical use in the field.” *Automation in Construction*, 18(5), 513–524.

Koch, C., and Brilakis, I. (2011). “Pothole detection in asphalt pavement images.” *Advanced Engineering Informatics*, 25(3), 507–515.

Koch, C., Jog, G. M., and Brilakis, I. (2012). “Automated Pothole Distress Assessment Using Asphalt Pavement Video Data.” *Journal of Computing in Civil Engineering*, 27(4), 370–378.

Kotsiantis, S. B., Zaharakis, I. D., and Pintelas, P. E. (2007). “Supervised machine learning: A review of classification techniques.”

Lakusić, S., Brcić, D., and Tkalcević Lakusić, V. (2011). “Analysis of Vehicle Vibrations–New Approach to Rating Pavement Condition of Urban Roads.” *PROMET-Traffic&Transportation*, 23(6), 485–494.

Laurent, J., Hebert, J. F., Lefebvre, D., and Savard, Y. (2012). “Using 3D Laser Profiling Sensors for the Automated Measurement of Road Surface Conditions.” *Mechanisms, Modeling, Testing, Detection and Prevention Case Histories*, Springer, 157–159.

Levinshtein, A., Stere, A., Kutulakos, K. N., Fleet, D. J., Dickinson, S. J., and Siddiqi, K. (2009). “Turbopixels: Fast superpixels using geometric flows.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(12), 2290–2297.

Lin, J., and Liu, Y. (2010). “Potholes detection based on SVM in the pavement distress image.” *Distributed Computing and Applications to Business Engineering and Science (DCABES), 2010 Ninth International Symposium on*, 544–547.

Liosatos, J. (2013). *Road Maintenance in the PAG Region: Challenges and Opportunities*. Tuscon, Arizona.

Li, Q., and Liu, X. (2008). “Novel approach to pavement image segmentation based on neighboring difference histogram method.” *Image and Signal Processing, 2008. CISP’08. Congress on*, 792–796.

Li, Q., Yao, M., Yao, X., and Xu, B. (2010). “A real-time 3D scanning system for pavement distortion inspection.” *Measurement Science and Technology,* 21(1), 015702.

Liu, F., Xu, G., Yang, Y., Niu, X., and Pan, Y. (2008). “Novel approach to pavement cracking automatic detection based on segment extending.” *Knowledge Acquisition and Modeling, 2008. KAM’08. International Symposium on*, 610–614.

Lokeshwor, H., Das, L. K., and Goel, S. (2013). “Robust method for automated segmentation of frames with/without distress from road surface video clips.” *Journal of Transportation Engineering*, 140(1), 31–41.

Lowe, D. G. (2004). “Distinctive image features from scale-invariant keypoints.” *International journal of computer vision*, 60(2), 91–110.

Ma, C., Wang, W., Zhao, C., Di, F., and Zhu, Z. (2009). “Pavement cracks detection based on FDWT.” *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*, 1–4.

Maode, Y., Shaobo, B., Kun, X., and Yuyao, H. (2007). “Pavement crack detection and analysis for high-grade highway.” *Electronic Measurement and Instruments, 2007. ICEMI’07. 8th International Conference on*, 4–548.

Mathavan, S., Rahman, M., Stonecliffe-Jones, M., and Kamal, K. (2014). “Pavement Raveling Detection and Measurement from Synchronized Intensity and Range Images.” *Transportation Research Record: Journal of the Transportation Research Board*, 2457, 3–11.

McTavish, T. H. (2012). *Performance Audit of the Measurement of State Highway Pavement Conditions*. Audit Report, Michigan Department of Transportation, Lansing, Michigan.

MnDOT. (2003). “Mn/DOT Distress Identification Manual.” Minnesota Department of Transportation.

MnDOT. (2009). *Pavement Condition Executive Summary*. MnDOT/OMRR-PM--2009-01, Minnesota Department of Transportation.

NAMS Group. (2006). *International infrastructure management manual*. National Asset Management Steering Group.

National Audit Office. (2014). *Maintaining stategic infrastructure: roads*. Summary, National Audit Office, UK.

Nejad, F. M., and Zakeri, H. (2011). “An optimum feature extraction method based on Wavelet–Radon Transform and Dynamic Neural Network for pavement distress classification.” *Expert Systems with Applications*, 38(8), 9442–9460.

Nguyen, T. S., Avila, M., and Begot, S. (2009). “Automatic detection and classification of defect on road pavement using anisotropy measure.” *Proceeding of EUSIPCO*, 617–621.

Ngwangwa, H. M., Heyns, P. S., Labuschagne, F. J. J., and Kululanga, G. K. (2010). “Reconstruction of road defects and road roughness classification using vehicle responses with artificial neural networks simulation.” *Journal of Terramechanics*, 47(2), 97–111.

NHTSA, 2014. (2014). “Federal Motor Vehicle Safety Standards; Rear Visibility.” *FEDERAL REGISTER-The Daily Journal of the United States Government*, <https://www.federalregister.gov/articles/2014/04/07/2014-07469/federal-motor-vehicle-safety-standards-rear-visibility> (May 6, 2014).

Nishiyama, S., Minakata, N., Kikuchi, T., and Yano, T. (2015). “Improved digital photogrammetry technique for crack monitoring.” *Advanced Engineering Informatics*, Collective Intelligence Modeling, Analysis, and Synthesis for Innovative Engineering Decision MakingSpecial Issue of the 1st International Conference on Civil and Building Engineering Informatics, 29(4), 851–858.

Nitsche, P., Stütz, R., Kammer, M., and Maurer, P. (2012). “Comparison of Machine Learning Methods for Evaluating Pavement Roughness Based on Vehicle Response.” *Journal of Computing in Civil Engineering*, 28(4), 04014015.

Peng, B., Jiang, Y., and Pu, Y. (2015). “Review on Automatic Pavement Crack Image Recognition Algorithms.” *Journal of Highway and Transportation Research and Development (English Edition)*, 9(2), 13–20.

PublicWorksTraining. (2014). *PASER Data Collection Best Practices Manual - Indiana LTAP PASER Training 2014*. Houghton, MI.

Radopoulou, S. C., and Brilakis, I. (2015). “Patch detection for pavement assessment.” *Automation in Construction*, 53, 95–104.

Rami, K. Z., and Kim, Y.-R. (2015). *Nebraska Data Collection*. Lincoln, NE.

Richardson, D. N., Lusher, S. M., and Luna, R. (2015). *MoDOT Pavement Preservation Research Program. Volume II, Data Collection for Pavement Management: Historical Data Mining and Production of Data*. Missouri University of Science and Technology for Missouri Department of Transportation.

Rick Miller. (2015). *Condition Survey Report*. Kansas Department of Transport.

Salari, E., and Bao, G. (2011). “Pavement distress detection and severity analysis.” *IS&T/SPIE Electronic Imaging*, 78770C–78770C.

Shotton, J., Winn, J., Rother, C., and Criminisi, A. (2009). “Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context.” *International Journal of Computer Vision*, 81(1), 2–23.

Solhmirzaei, A., Azadi, S., and Kazemi, R. (2012). “Road profile estimation using wavelet neural network and 7-DOF vehicle dynamic systems.” *Journal of mechanical science and technology*, 26(10), 3029–3036.

Sorncharean, S., and Phiphobmongkol, S. (2008). “Crack detection on asphalt surface image using enhanced grid cell analysis.” *Electronic Design, Test and Applications, 2008. DELTA 2008. 4th IEEE International Symposium on*, 49–54.

Subirats, P., Dumoulin, J., Legeay, V., and Barba, D. (2006). “Automation of pavement surface crack detection using the continuous wavelet transform.” *Image Processing, 2006 IEEE International Conference on*, 3037–3040.

Sun, Y., Salari, E., and Chou, E. (2009). “Automated pavement distress detection using advanced image processing techniques.” *Electro/Information Technology, 2009. eit’09. IEEE International Conference on*, 373–377.

Sy, N. T., Avila, M., Begot, S., and Bardet, J.-C. (2008). “Detection of defects in road surface by a vision system.” *Electrotechnical Conference, 2008. MELECON 2008. The 14th IEEE Mediterranean*, 847–851.

Tapia-Espinoza, R., and Torres-Torriti, M. (2013). “Robust Lane Sensing and Departure Warning under Shadows and Occlusions.” *Sensors*, 13(3), 3270–3298.

Teomete, E., Amin, V. R., Ceylan, H., and Smadi, O. (2005). “Digital image processing for pavement distress analyses.” *Proceedings of the 2005 Mid-Continent Transportation Research Symposium*, 1–13.

Tsai, Y.-C. J., and Li, F. (2012). “Critical assessment of detecting asphalt pavement cracks under different lighting and low intensity contrast conditions using emerging 3D laser technology.” *Journal of Transportation Engineering*, 138(5), 649–656.

Tsai, Y.-C., Kaul, V., and Mersereau, R. M. (2009). “Critical assessment of pavement distress segmentation methods.” *Journal of Transportation Engineering*, 136(1), 11–19.

Tsai, Y. J., Li, F., and Wu, Yiching. (2013). “A New Rutting Measurement Method Using Emerging 3D Line-Lase-Imaging System.” 6(5), 667–672.

Uijlings, J. R., Smeulders, A. W., and Scha, R. J. (2010). “Real-time visual concept classification.” *Multimedia, IEEE Transactions on*, 12(7), 665–681.

UKPMS. (2005). “The UKPMS user manual.” United Kingdom Pavement Management System.

Uslu, B., Golparvar-Fard, M., and de la Garza, J. M. (2011). “Image-based 3D reconstruction and recognition for enhanced highway condition assessment.” *Proceedings of the 2011 ASCE Intl. Workshop on Computing in Civil Engineering, Miami, FL*, 67–76.

Veksler, O., Boykov, Y., and Mehrani, P. (2010). “Superpixels and supervoxels in an energy optimization framework.” *Computer Vision–ECCV 2010*, Springer, 211–224.

Vilacca, J. L., Fonseca, J. C., Pinho, A. C. M., and Freitas, E. (2010). “3D surface profile equipment for the characterization of the pavement texture–TexScan.” *Mechatronics*, 20(6), 674–685.

Wang, K. C., and Gong, W. (2005). “Real-time automated survey system of pavement cracking in parallel environment.” *Journal of infrastructure systems*, 11(3), 154–164.

Wang, K. C., Hou, Z., and Williams, S. (2010). “Precision test of cracking surveys with the automated distress analyzer.” *Journal of Transportation Engineering*, 137(8), 571–579.

Wang, Q., McDaniel, J. G., Sun, N. X., and Wang, M. L. (2013). “Road profile estimation of city roads using DTPS.” *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, 86923C–86923C.

Wei, L., Fwa, T. F., and Zhe, Z. (2005). “Wavelet analysis and interpretation of road roughness.” *Journal of transportation engineering*, 131(2), 120–130.

Werro, P. (2013). “SCANNER surveys.”

Wolpert, D. H., and Macready, W. G. (1997). “No free lunch theorems for optimization.” *Evolutionary Computation, IEEE Transactions on*, 1(1), 67–82.

Wu, L., Mokhtari, S., Nazef, A., Nam, B., and Yun, H.-B. (2016). “Improvement of Crack-Detection Accuracy Using a Novel Crack Defragmentation Technique in Image-Based Road Assessment.” *Journal of Computing in Civil Engineering*, 30(1), 04014118.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., and Philip, S. Y. (2008). “Top 10 algorithms in data mining.” *Knowledge and Information Systems*, 14(1), 1–37.

Xu, G., Ma, J., Liu, F., and Niu, X. (2008). “Automatic recognition of pavement surface crack based on Bp neural network.” *Computer and Electrical Engineering, 2008. ICCEE 2008. International Conference on*, 19–22.

Yao, X., Yao, M., and Xu, B. (2008). “Automated Detection and Identification of Area-based Distress in Concrete Pavements.” *Seventh International Conference on Managing Pavement Assets*.

Ying, L., and Salari, E. (2010). “Beamlet Transform-Based Technique for Pavement Crack Detection and Classification.” *Computer-Aided Civil and Infrastructure Engineering*, 25(8), 572–580.

Yousefzadeh, M., Azadi, S., and Soltani, A. (2010). “Road profile estimation using neural network algorithm.” *Journal of mechanical science and technology*, 24(3), 743–754.

Yu, B. X., and Yu, X. (2006). “Vibration-based system for pavement condition evaluation.” *Applications of Advanced Technology in Transportation. The Ninth International Conference*.

Yun, H.-B., Mokhtari, S., and Wu, L. (2015). “Crack Recognition and Segmentation Using Morphological Image-Processing Techniques for Flexible Pavements.” *Transportation Research Record: Journal of the Transportation Research Board*, (2523), 115–124.

Yu, S.-J., Sukumar, S. R., Koschan, A. F., Page, D. L., and Abidi, M. A. (2007). “3D reconstruction of road surfaces using an integrated multi-sensory approach.” *Optics and lasers in engineering*, 45(7), 808–818.

Zhang, G. P. (2000). “Neural networks for classification: a survey.” *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 30(4), 451–462.

Zhou, H., Jalayer, M., Gong, J., Hu, S., and Grinter, M. (2013). *Investigation of Methods and Approaches for Collecting and Recording Highway Inventory Data*.

Zhou, J., Huang, P. S., and Chiang, F.-P. (2003). “Wavelet-aided pavement distress image processing.” *Optical Science and Technology, SPIE’s 48th Annual Meeting*, 728–739.

Table 1 Reference list on methods for automating pavement defect detection & measurement

Presence

1. Cheng et al. 2003
2. Li and Liu 2008
3. Battiato et al. 2006, 2007
4. Zhou et al. 2003
5. Doumiati et al. 2011
6. González et al. 2008
7. Harris et al. 2010
8. Imine and Delanne 2005
9. Imine et al. 2003, 2006
10. Johnsson&Odelius 2012
11. Ngwangwa et al. 2010
12. Solhmirzaei et al. 2012
13. Wang et al. 2013
14. Wei et al. 2005
15. Yousefzadeh et al. 2010
16. Yu and Yu 2006
17. Lakusić et al. 2011
18. Georgieva et al. 2015
19. Lokeshwor et al. 2013

Detection

20. Jiang and Tsai 2015
21. Kaul et al. 2010
22. Yun et al. 2015
23. Adu-Gyamfi et al. 2011
24. Cord & Chambon 2012
25. Gavilán et al. 2011
26. Ghanta et al. 2012
27. Huang and Xu 2006
28. Jing and Aiqin 2010
29. Kamaliardakani et al. 2014
30. Liu et al. 2008
31. Ma et al. 2009
32. Nejad and Zakeri 2011
33. Peng et al. 2015 (r)
34. Sornchearean and Phiphobmongkol 2008
35. Subirats et al. 2006
36. Sy et al. 2008
37. Tsai et al. 2009 (r)
38. Tsai and Li 2012
39. Wu et al. 2016
40. Xu et al. 2008
41. Rado and Brilakis 2015
42. Chang et al. 2005
43. Jog et al. 2012
44. Koch and Brilakis 2011
45. Koch et al. 2012
46. Mathavan et al. 2014
47. Jahanshahi et al. 2013
48. Cafiso et al. 2006
49. Lin and Liu 2010
50. Yao et al. 2008
51. Uslu et al. 2011

Measurement

52. Wang et al. 2010
53. Amarasiri et al. 2009
54. Haas 1996
55. Kim et al. 2009
56. Laurent et al. 2012
57. Maode et al. 2007
58. Nejad and Zakeri 2011
59. Nguyen et al. 2009
60. Nishiyama et al. 2015
61. Salari and Bao 2011
62. Sun et al. 2009
63. Teomete et al. 2005
64. Wang and Gong 2005
65. Ying and Salari 2010
66. Liu et al. 2008
67. Li et al. 2010
68. Austroads 2011
69. Islam et al. 2014
70. Nitsche et al. 2012
71. Tsai et al. 2013
72. Vilacca et al. 2010

Table 2 Pavement defect and its representative color in the ground truth data

Type of defect	*Color*
Longitudinal crack	Red
Transverse crack	Blue
Patch	Yellow
Pothole	Pink
Healthy pavement	Grey
Void	Black

Table 3 Specifications of cameras used for collecting data

	HP Elite Autofocus Webcam	*Point Grey Blackfly 05S2M-CS*
Image resolution	640 x 480	800 x 500
Horizontal angle of view	~50^{o}	133^{o}
Frame rate per second	30	50
Color	RGB	Monochrome

Table 4 Tested parameters and results of STFs (data captured by HP camera using 4 categories)

		Test 1	*Test 2*	*Test 3*	*Test 4*	*Test 5*	*Test 6*	*Test 7*	*Test 8*
	Box size	15	11	13	9	17	15	15	15
	Max tree depth	10 & 14	10 & 14	10 & 14	10 & 14	10 & 14	11 & 13	12 & 16	15 & 16
	Ov.Acc.	0.78	0.69	0.73	0.78	0.76	0.78	0.78	0.79
	Av.Acc	0.64	0.55	0.73	0.62	0.60	0.60	0.64	0.65
Longitudinal crack	*Av.Pr.*	0.95	0.95	0.95	0.97	0.96	0.96	0.90	0.96
	AuC	0.86	0.80	0.80	0.90	0.90	0.90	0.90	0.89
Transverse crack	*Av.Pr.*	0.20	0.04	0.01	0.28	0.02	0.27	0.35	0.01
	AuC	0.85	0.76	0.26	0.73	0.68	0.93	0.75	0.29
Patch	*Av.Pr.*	0.75	0.86	0.68	0.81	0.80	0.69	0.62	0.84
	AuC	0.88	0.91	0.81	0.92	0.88	0.86	0.80	0.92
Pothole	*Av.Pr.*	0.89	0.99	0.84	0.90	0.90	0.82	0.81	0.92
	AuC	0.96	0.99	0.95	0.96	0.96	0.90	0.96	0.96

Table 5 Tested parameters and results of STFs (data captured by HP camera using 5 categories)

		Test 1	*Test 2*	*Test 3*	*Test 4*	*Test 5*	*Test 6*	*Test 7*	*Test 8*
	Box size	15	11	13	9	17	13	13	13
	Max tree depth	10 & 14	10 & 14	10 & 14	10 & 14	10 & 14	11 & 13	12 & 16	15 & 16
	Ov.Acc.	0.84	0.87	0.89	0.87	0.86	0.89	0.86	0.89
	Av.Acc	0.64	0.65	0.56	0.60	0.60	0.60	0.58	0.57
Longitudinal crack	*Av.Pr.*	0.95	0.95	0.92	0.97	0.96	0.96	0.96	0.96
	AuC	1.00	0.76	0.85	0.56	0.87	0.85	0.75	0.62
Transverse crack	*Av.Pr.*	1.00	0.77	0.77	0.53	0.02	0.78	0.04	0.53
	AuC	1.00	0.97	0.97	0.96	0.19	0.98	0.69	0.87
Patch	*Av.Pr.*	0.75	0.86	0.80	0.75	0.69	0.79	0.72	0.63
	AuC	0.85	0.93	0.89	0.82	0.85	0.89	0.83	0.81
Pothole	*Av.Pr.*	1.00	0.76	0.98	0.96	0.92	0.99	1.00	0.93
	AuC	1.00	0.94	0.99	0.99	0.96	0.99	1.00	0.96
Healthy pavement	*Av.Pr.*	0.97	0.99	0.98	0.98	0.98	0.98	0.96	0.97
	AuC	0.07	0.60	0.21	0.19	0.38	0.43	0.10	0.32

Table 6 Tested parameters and results of STFs (data captured by PG camera using 5 categories)

			Test 1	*Test 2*	*Test 3*	*Test 4*	*Test 5*	*Test 6*	*Test 7*	*Test 8*
		Box size	15	11	9	13	17	11	11	11
		Max tree depth	10 & 14	10 & 14	10 & 14	10 & 14	10 & 14	11 & 13	12 & 16	15 & 16
Average values		*Ov.Acc.*	0.83	0.84	0.82	0.82	0.82	0.82	0.84	0.86
		Av.Acc	0.74	0.76	0.74	0.72	0.76	0.74	0.74	0.72
	Longitudinal crack	*Av.Pr.*	0.93	0.94	0.94	0.92	0.93	0.95	0.94	0.92
		AuC	0.96	0.96	0.96	0.95	0.96	0.96	0.96	0.96
	Transverse crack	*Av.Pr.*	0.84	0.73	0.83	0.86	0.81	0.83	0.85	0.85
		AuC	0.95	0.94	0.97	0.93	0.94	0.93	0.94	0.94
	Patch	*Av.Pr.*	0.96	0.94	0.94	0.95	0.96	0.95	0.96	0.96
		AuC	0.94	0.92	0.93	0.93	0.92	0.91	0.95	0.95
	Pothole	*Av.Pr.*	0.84	0.82	0.82	0.77	0.79	0.81	0.81	0.76
		AuC	0.93	0.92	0.93	0.92	0.90	0.91	0.95	0.89
	Healthy pavement	*Av.Pr.*	0.97	0.97	0.97	0.98	0.98	0.92	0.98	0.96
		AuC	0.65	0.62	0.60	0.69	0.65	0.66	0.70	0.55
Variance		*Ov.Acc.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Av.Acc.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Longitudinal crack	*Av.Pr.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		AuC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Transverse crack	*Av.Pr.*	0.00	0.01	0.01	0.01	0.02	0.02	0.01	0.01
		AuC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Patch	*Av.Pr.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		AuC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Pothole	*Av.Pr.*	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.01
		AuC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Healthy pavement	*Av.Pr.*	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
		AuC	0.02	0.01	0.03	0.01	0.01	0.03	0.02	0.02

Table 7 Tested parameters and results of STFs (data captured by PG camera using 5 categories and myROI)

			Test 1	*Test 2*	*Test 3*	*Test 4*	*Test 5*	*Test 6*	*Test 7*	*Test 8*
		Box size	15	13	11	17	9	9	9	9
		Max tree depth	10 & 14	10 & 14	10 & 14	10 & 14	10 & 14	11 & 13	10 & 15	12 & 14
Average values		*Ov.Acc.*	0.83	0.83	0.83	0.83	0.83	0.83	0.85	0.84
		Av.Acc	0.75	0.73	0.74	0.74	0.74	0.75	0.74	0.73
	Longitudinal crack	*Av.Pr.*	0.92	0.91	0.92	0.90	0.92	0.91	0.92	0.93
		AuC	0.95	0.94	0.96	0.94	0.96	0.96	0.96	0.96
	Transverse crack	*Av.Pr.*	0.89	0.92	0.83	0.83	0.81	0.82	0.83	0.87
		AuC	0.95	0.98	0.94	0.95	0.93	0.89	0.95	0.95
	Patch	*Av.Pr.*	0.88	0.90	0.88	0.88	0.91	0.91	0.89	0.88
		AuC	0.89	0.89	0.88	0.87	0.90	0.90	0.88	0.87
	Pothole	*Av.Pr.*	0.71	0.71	0.66	0.66	0.62	0.62	0.68	0.56
		AuC	0.90	0.92	0.83	0.93	0.90	0.89	0.93	0.87
	Healthy pavement	*Av.Pr.*	0.96	0.87	0.96	0.98	0.96	0.96	0.97	0.97
		AuC	0.67	0.62	0.59	0.66	0.54	0.48	0.62	0.66
Variance		*Ov.Acc.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Av.Acc.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Longitudinal crack	*Av.Pr.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		AuC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Transverse crack	*Av.Pr.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		AuC	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	Patch	*Av.Pr.*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		AuC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Pothole	*Av.Pr.*	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.01
		AuC	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.00
	Healthy pavement	*Av.Pr.*	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
		AuC	0.01	0.03	0.04	0.01	0.03	0.03	0.02	0.01

Table 8 Confusion matrix for 2D segmentation of defects (data captured with HP camera using 4 categories - results from test 1)

	Longitudinal crack	*Transverse crack*	*Patch*	*Pothole*
Longitudinal crack	0.80	0.00	0.17	0.00
Transverse crack	0.67	0.02	0.12	0.00
Patch	0.21	0.00	0.78	0.00
Pothole	0.06	0.00	0.20	0.74

Table 9 Confusion matrix for 2D segmentation of defects (data captured with HP camera using 5 categories - results from test 6)

	Longitudinal crack	*Transverse crack*	*Patch*	*Pothole*	*Healthy pavement*
Longitudinal crack	0.28	0.01	0.05	0.00	0.66
Transverse crack	0.00	0.71	0.00	0.05	0.24
Patch	0.14	0.00	0.44	0.03	0.39
Pothole	0.06	0.00	0.06	0.66	0.23
Healthy pavement	0.04	0.01	0.03	0.00	0.92

Table 10 Confusion matrix for 2D segmentation of defects (data captured with PG camera using 5 categories - results from test 8)

	Longitudinal crack	*Transverse crack*	*Patch*	*Pothole*	*Healthy pavement*
Longitudinal crack	0.69	0.01	0.01	0.00	0.29
Transverse crack	0.02	0.63	0.01	0.00	0.34
Patch	0.02	0.01	0.61	0.00	0.36
Pothole	0.06	0.00	0.03	0.78	0.13
Healthy pavement	0.03	0.02	0.05	0.00	0.91

Table 11 Confusion matrix for 2D segmentation of defects (data captured with PG camera using 5 categories and myROI - results from test 7)

	Longitudinal crack	*Transverse crack*	*Patch*	*Pothole*	*Healthy pavement*
Longitudinal crack	0.75	0.01	0.02	0.00	0.22
Transverse crack	0.02	0.63	0.01	0.00	0.34
Patch	0.02	0.01	0.63	0.00	0.34
Pothole	0.05	0.00	0.04	0.80	0.11
Healthy pavement	0.02	0.02	0.07	0.00	0.89