Pothole detection in asphalt pavement images

Article *in* Advanced Engineering Informatics · August 2011

DOI: 10.1016/j.aei.2011.01.002 · Source: DBLP

2 authors:

Christian Koch
Bauhaus-Universität Weimar
149 PUBLICATIONS **4,592** CITATIONS
SEE PROFILE

Ioannis Brilakis
University of Cambridge
344 PUBLICATIONS **10,750** CITATIONS
SEE PROFILE

Pothole Detection in Asphalt Pavement Images

Christian Koch[a,*] and Ioannis Brilakis[b]

[a] Postdoctoral Fellow, Institute for Computing in Engineering, Ruhr-Universität Bochum, Universitätstraße 150, 44801 Bochum, Germany;
Phone: +49-234-32-26174; E-mail: koch@inf.bi.rub.de

[b] Assistant Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, 328 Sustainable Education Building, 788 Atlantic Drive NW, USA;
Phone:1-404-894-9981; Fax: 1-404-894-1641; E-mail: brilakis@gatech.edu

Abstract:

Pavement condition assessment is essential when developing road network maintenance programs. In practice, the data collection process is to a large extent automated. However, pavement distress detection (cracks, potholes, etc.) is mostly performed manually, which is labor-intensive and time-consuming. Existing methods either rely on complete 3D surface reconstruction, which comes along with high equipment and computation costs, or make use of acceleration data, which can only provide preliminary and rough condition surveys. In this paper we present a method for automated pothole detection in asphalt pavement images. In the proposed method an image is first segmented into defect and non-defect regions using histogram shape-based thresholding. Based on the geometric properties of a defect region the potential pothole shape is approximated utilizing morphological thinning and elliptic regression. Subsequently, the texture inside a potential defect shape is extracted and compared with the texture of the surrounding non-defect pavement in order to determine if the region of interest represents an actual pothole. This methodology has been implemented in a MATLAB prototype,

* Corresponding author

trained and tested on 120 pavement images. The results show that this method can detect potholes in asphalt pavement images with reasonable accuracy.

Keywords:

Pavement assessment, Pothole detection, Visual sensing, Image processing

1. INTRODUCTION

The highway system of a developed country contains thousands centerline kilometers of pavement. Such systems consist of asphalt, concrete or composite pavements ranging in condition, age and performance. In recent years, several road network maintenance programs have been established in order to monitor the on-going performance of the road network, to predict future pavement conditions and assess long term needs, to support investment planning and decision making, and to identify rehabilitation and maintenance treatment. Among others, the Long-Term Pavement Performance program (LTPP) established by the United States Department of Transportation is an ongoing and active program, which is concerned with data collection, data storage, data analysis and product development of the road network in the United States and Canada [1]. Within this program, pavement surface condition assessment is identified as a key component that requires reliable and good-quality measurements on pavement distresses like cracks, potholes, rutting, etc. [2].

In current practice, pavement image and video data collected by digital inspection vehicles are reviewed by technicians on computer monitors in order to manually detect and assess defects. Besides this being a time-consuming and costly task, the final results are influenced by the subjectivity and the experience of the raters [3]. To overcome these limitations, several research efforts towards automating defect detection have been undertaken. This paper focuses on pothole defects. Existing research work in pothole detection can be divided into 3D pavement reconstruction methods [4,5] employing 3D laser scanning [6] and stereo-vision [7,8], and

vibration-based approaches using acceleration sensors [9-12]. While the former basically suffer from high equipment costs and high computational effort, the latter are only intended for preliminary and rough condition surveys and therefore lack accuracy and reliability [9,11].

In this paper we present a novel approach for automated pothole detection based on asphalt pavement surface images. These images are either cropped from available survey videos, or, according to our future vision, cropped from pavement video data collected by ubiquitous passenger vehicles equipped with a high-speed camera. Our method starts with image segmentation to divide the pavement image into defect and non-defect regions. Based on the geometric properties of a defect region the potential pothole shape is approximated. Subsequently, the texture inside a potential defect region is extracted and compared with the texture of the surrounding non-defect pavement. If the texture inside the defect region is coarser and grainier than the one of the intact surface, it is assumed that the region of interest represents a pothole.

The methodology proposed in this paper has been implemented in a MATLAB prototype, manually trained and tested on 120 pavement images, which provide a high variety of potholes (shape and size), other defects (cracking, patching, discoloration, etc.) and non-defect pavements as well as diverse lighting conditions (shadows). Within the training procedure we conducted experiments to determine the most significant thresholds for shape extraction and texture comparison using precision-recall graphs. Based on the test results we found that the proposed method can successfully detect potholes with an overall accuracy of 86% involving 82% precision and 86% recall.

2. BACKGROUND

The process of pavement condition assessment is divided into three parts: data collection, distress identification and defect assessment. In current practice, the former is to a large extent automated, while the latter two are mostly performed manually. Sophisticated digital inspection vehicles

collect pavement data using several sensors, in particular video cameras for surface imaging, optical sensors for distance measurements, laser scanners for longitudinal and transverse profiling, additional ultrasonic sensors for rutting detection, and accelerometers for roughness measurements [13,14].

Pavement distresses are those defects visible on the pavement surface. They are symptoms indicating problems of deterioration like cracks, potholes and rutting. Traditionally, the captured video data is reviewed by technicians who visually detect and assess defects on computer screens based on their own experiences and a distress manual, which defines categories of severity and extent for each type of distress [15-17]. In cases where there are no digital inspection vehicles available, pavement condition surveys are still manually conducted from a van at a speed of 8 to 15 mph using data collection assistance software [16]. Lately, crack detection is performed semi-automated, off-line and overnight by available commercial software packages that implement digital image processing algorithms [13]. However, existing commercial software does not support the automated detection of some types of distresses like potholes, patching and discoloration. In order to assess and rate pavement condition, indices for both distresses and roughness have to be determined. Based on the severity and the extent of identified distresses, a surface distress index is calculated and combined with a roughness index to determine an overall pavement condition rating [18]. Unlike the distress identification, the pavement roughness is automatically determined in real-time based on collected accelerometer and laser scanner data [13].

Potholes, as one type of pavement distresses, are bowl-shaped depressions of various sizes in the pavement surface (as per definition in [16, 17]). Considering their visual impact, they can also be defined as almost elliptical pavement regions, which are fully or partially surrounded by a dark shadow (due to depression) and which have a granular and course textural appearance (due to fragmentation). Based on these visual characteristics they are identified and assessed manually within visual inspections of pavement image and video data. This is a time-consuming

and tedious process. Furthermore, a certain amount of subjectivity and the experience of the raters have an undoubted influence on inspection data repeatability and thus the final assessment [2,3]. These limitations demonstrate the need for automation. Our research aims to automate the process of pothole detection in visual pavement data. Consequently, we compare the ability to decide whether a pavement image contains a pothole or not versus to manually search for potholes in a collection of pavement images, as our goal is to automate the existing process, not improving it nor measuring geometry.

2.1 State of research in automated pothole detection

Current research efforts in automating the detection of potholes can be divided into 3D reconstruction-based, vibration-based and vision-based methods. Detection methods which are based on a 3D reconstruction of the pavement surface rely on 3D point clouds provided by laser scanners or by stereo-vision algorithms using a pair of video cameras. 3D laser scanning systems can be further distinguished into time-of-flight scanners that employ reflected laser pulses to directly create 3D point coordinates [4], and hybrid systems that use digital cameras to capture consecutive images of lines projected by infrared lasers [6]. Li et al. [6] have presented a real-time 3D laser scanning system for pavement rutting, shoving and pothole detection. Using the combined approach their system can collect, measure and visualize 3D pavement surface data, which is used to identify possible distortions. A stereo-vision based surface model for comprehensive pavement conditioning has been proposed by Wang [7] and Hou et al. [8]. With the availability of a 3D point cloud, Chang et al. [4] have presented a clustering approach that can quantify the severity and coverage of potholes and Jiaqiu et al. [5] have created a method for identifying, locating, classifying and measuring sag deformations like potholes and depression. The drawbacks of stereo-vision based approaches are that they require a complete 3D reconstruction of the pavement surface and that the procedure of matching points between the two views is quite challenging due to the very irregular texture and color of the pavement surface.

This results in a high computational effort precluding this approach from being used in a real-time environment. Also, although laser scanning systems collect, measure and visualize data in real-time and their costs have come down, they still suffer from high equipment costs (either time-of-flight laser system or combined laser projector and camera system) compared to a single camera systems. Moreover, current literature on employing 3D laser scanning to detect pavement distress is usually focused on the local accuracy of 3D measurements, validation in terms of distress classification results (e.g. precision and recall) for global survey test are not presented.

Vibration-based approaches use accelerometers in order to detect potholes. Yu and Yu [9] have proposed a preliminary pavement condition survey. In analogy to cameras "looking" at pavement surface, an accelerometer "feels" the ground conditions based on the vehicle's mechanical responses. The advantages of a vibration-based system have been identified as requiring small storage, being cost-effective and amenable for real-time processing [9]. The projects BusNet [10] and Pothole Patrol [11] share the same basic idea of combining vibration sensors with GPS and using mobile nodes to sense road conditions in terms of pothole detection and positioning. Rode et al. [12] have proposed an integrated pothole detection and warning system that distributes collected data (spatially tagged potholes) to the participating vehicles assisting drivers in avoiding potholes. However, Erikson et al. [11] found that vibration-based approaches could provide wrong results, e.g. in case of bridge expansion joints that have been detected as potholes by mistake. Moreover, potholes in the center of a lane are not hit by any of the vehicle's wheels and thus cannot be recognized using accelerometers. Nevertheless, these potholes must be captured as they are pavement defects and, in addition, are particularly dangerous obstacles for motorcycles.

A different approach is based on 2D vision (image or video data) since potholes usually are significantly different from the pavement background surface. Karuppuswamy et al. [19] have presented a solution to the detection and avoidance of simulated potholes integrating a vision and motion system. Although this approach can detect potholes in the center of a lane, it relies on

simulated potholes that are larger than 2 feet in diameter and white in color. However, these are very simplified assumptions that do not reflect realistic pavement conditions. The authors are not aware of further vision-based approaches.

3. METHODOLOGY

The overall objective of this research is to test whether a ubiquitous and comprehensive framework can detect, recognize, spatially locate and evaluate the magnitude of defects on pavements from a network of passenger vehicles for the purpose of simultaneously assessing the condition of road networks and its impact on the driving public. A rear-view (parking) camera on a vehicle is replaced with a high-speed fish-eye camera that can tilt downwards during driving forward. The video feed is processed by an in-car computer in 2 stages; at first with high speed (real time) algorithms for detecting frames that might contain evidence of defects, then with defect recognition algorithms only on those frames selected in the first stage. The result is frame regions tagged with the type of defect identified (cracks, potholes, patches etc.). Defect properties measurement algorithms are then applied to assess the severity of the recognized defects. The motivation behind the proposed framework lies in the deficiencies of current practice and the potential of gradually and inexpensively converting passenger vehicles into ubiquitous sensors and reporters of the roads' condition. The work presented in this paper is the first step to achieve our objective. We use a sensing robot that simulates an equipped passenger vehicle that collects visual pavement data. This data is then used to validate our method for automated pothole detection in pavement images.

The pothole detection approach proposed is based on the idea of creating a pattern recognition model that will automate the recognition of infrastructure-related elements [20]. Zhu and Brilakis [21,22] have successfully created vision-based recognition models to detect concrete columns in images and videos, and to assess the quality of concrete surfaces, respectively. Following this strategy, in this paper a recognition model for automating the detection of potholes

in images is presented. Considering the visual appearance of potholes three main distinctive visual characteristics have been identified by the authors:

a) A pothole includes one or more shadows that are darker than the surrounding area.
b) The shape of a pothole is approximately elliptical, due to a perspective view.
c) The surface texture inside a pothole is much coarser and grainier than the surface texture of the surrounding, intact pavement.

Based on these observations, the proposed model mainly includes three components: 1) image segmentation, 2) shape extraction, and 3) texture extraction and comparison (Fig. 1).

3.1 Image segmentation

In pavement surface images, color information, in particular RGB values are not essential when performing the segmentation process with regard to defect detection. Therefore, the first step is to transform original color images into gray-scale images. Subsequently, a 5 by 5 median filter is used to reduce noise that is always produced by hardware during image or video capturing processes.

In order to meet the first identified characteristic, darker regions that indicate defects have to be separated from the background within each pavement image. For this purpose, we use a histogram shape-based thresholding algorithm, which is based on the triangle algorithm presented in [23]. In order to remove histogram peaks that might interfere with the subsequent method, a 1D median filter with order 5 is applied on the histogram. Then the threshold T is determined as the intensity value of a histogram point $P_T = [T, h(T)]$, which has the maximum distance to a line $l = [P_0, P_{max}]$ that intersects the histogram's origin P_0 and the point P_{max} indicating the maximum intensity (Fig. 2). Based on the threshold T the enhanced image G_{enh} is converted into a binary image B using Equation 1. According to Figure 1, B contains the defect areas in the image considered. Figure 3 depicts the image segmentation process for two selected images each containing a pothole

$$B(i,j) := \begin{cases} 1 & if \;\; G_{\text{enh}}(i,j) \leq T \\ 0 & otherwise \end{cases} \quad (1)$$

3.2 Shape extraction

After filling holes within the remaining regions of the binary image *B*, small regions (e.g. short cracks, artifacts), regions, which have a linear shape (e.g. long cracks, joints, curb shades) and regions, which are connected to the image boundary are assumed not to be potholes and thus are removed. The absolute size s of a region of interest R (measured in number of pixels) is dependent on the image resolution, whereas the linearity of a region's shape can be defined by its eccentricity ε, which is a relative parameter. In this study, a region is determined to be linear if its eccentricity ε is larger than 0.99. Within the next step, the shapes of the remaining defect candidate regions are extracted. For this purpose the following additional region properties are determined:

- Length of the major axis (l_{max})
- Position of the centroid (P_{cent})
- Orientation angle (α)

Based on these properties it is distinguished between regions that (a) might represent a pothole shade (e.g. Fig. 3a) and regions, which (b) could represent an entire pothole (e.g. Fig. 3c). A shade region is identified if the region's centroid does not reside inside the region, or if both its eccentricity exceeds a threshold ε_{max} and the ratio of the major axis length (l_{max}) and the region size (s) exceeds a threshold r_{max}. In section 4.2.2. we explain how these thresholds have been experimentally determined. Equation 2 describes how a region of interest R is distinguished.

$$type(R) := \begin{cases} \text{pothole shade candidate} & if \;\; P_{\text{cent}} \notin R \;\vee\; (\varepsilon > \varepsilon_{\text{max}} \wedge (l_{\text{max}}/s) > r_{\text{max}}) \\ \text{entire pothole candidate} & otherwise \end{cases} \quad (2)$$

In order to approximate the elliptical shape of a pothole based on its detected shade, the following procedure is proposed. First, morphological thinning is used to shrink the shade region to its

minimally connected stroke or skeleton (Fig. 4a-b). Subsequently, the skeleton's branching points are identified (Fig. 4c) and connected in order to determine the major path of the shade region (Fig. 4d). In cases where the number of skeleton end points is smaller than 5, meaning the boundary of the shape is almost smooth, the entire skeleton defines the major path. Finally, based on the major path elliptic regression [24] is used to approximate an ellipse (Fig. 4e-f).

3.3 Texture extraction and comparison

In order to meet the third identified visual characteristic of potholes, the surface texture inside a pothole candidate has to be described and compared with the texture of the surrounding region. This step is essential in order to distinguish between false candidates (small repair patches, spot shadows and discoloration) and true potholes. Figure 5 illustrates typical pavement textures inside (a) and outside (b) a pothole. In the subsequent description the inside region is denoted as R_i, whereas R_o refers to the outside region. The major approaches used in image processing to describe the texture of a region are statistical, structural, and spectral. In order to determine the texture's smoothness, coarseness, granularity, etc. statistical approaches are best suited [25,26]. Therefore, we propose to use the standard deviation of gray-level intensity values as a statistical measure to describe the texture of both the inside and the outside region. The standard deviation value for the inside region is denoted as $std(R_i)$, for the outside region $std(R_o)$, respectively.

In order to emphasize structural texture characteristics, respective filters are applied to the original gray-level image. In this case, filter responses are used as the basis for statistical measures. According to the authors' observation that the region inside a pothole is mostly granular, so-called spot filters are utilized to create high filter responses. Leung and Malik [27] and Schmid [28] used collections of filters (filter banks) to classify materials based on their textural appearance and to construct texture-based feature descriptors for content-based image retrieval, respectively. These filter banks were successfully tested for images with varying lighting and viewing conditions, which also applies to pavement surface images. For our purpose,

three spot filters of the Leung-Malik (LM) filter bank [27] and one spot filter of the Schmid (S) filter bank [28] are selected (Fig. 6) and applied on the entire image. In order to cover a certain range, we use LM filters of three different sizes. The absolute size of each filter is set to 31 by 31 pixels, which has yield the best results within our training experiments (see section 4.2.2). Since all pavement images are assumed to be shot from nearly the same distance and angle, the absolute filter size doesn't have to be adjusted in terms of scale.

In section 3.2 we have described how the shape of a pothole (inside region) can be extracted. If the shape of a defect candidate region is determined to be a pothole shade, then the approximated ellipse defines the inside region R_i (Fig. 7d, top). In case of an assumed entire pothole region, the entire defect region is used to define R_i (Fig. 7d, bottom). As mentioned above, the outside region R_o is supposed to represent the intact and healthy part of the pavement image. In order to determine the outside region R_o we propose the following procedure. First, morphological dilation is applied on the segmented binary image B (Fig. 7b), which contains the assumed defect areas. The result is the binary image B_D (Fig. 7c). This step is necessary in order to remove unintentional filter responses in defect regions at the boundary of the outside region. For example, if an 11 by 11 pixel spot filter is applied on an image region that contains a 2 pixel wide crack, the response region will be larger than the original crack region and will thereby adulterate the response in the close-by non-defect region. For this reason, we dilate the defect region by half of the spot filter size. In this way we make sure that there are no wrong filter responses in the non-defect outside region resulting from nearby defects. Next, we combine the binary image B_D with the inside region R_i resulting the total defect region. Finally, the outside region R_o is determined as the complement of the total defect region (Fig. 7e). Figure 7 exemplifies this procedure for two pothole images.

Based on the filter responses inside the regions R_i and R_o (e.g. $LM40(R_i)$, $LM40(R_o)$), corresponding standard deviation values are determined and used to set up two feature vectors f_i and f_o, describing the texture of the regions R_i and R_o, respectively (Eqn. 3).

$$f_i := \left(std(R_i),\ std\left(LM40(R_i)\right),\ std\left(LM42(R_i)\right),\ std\left(LM44(R_i)\right),\ std\left(S4(R_i)\right)\right)$$
$$f_o := \left(std(R_o),\ std\left(LM40(R_o)\right),\ std\left(LM42(R_o)\right),\ std\left(LM44(R_o)\right),\ std\left(S4(R_o)\right)\right) \quad (3)$$

Following this, the lengths of the feature vectors f_i and f_o are calculated and compared. If $|f_o|$ is smaller than $|f_i|$, meaning the texture inside R_i is coarser and grainier than the one inside R_o, it is assumed that the region R_i is an inside region of a pothole (Eqn. 4).

$$type(R_i) := \begin{cases} \text{pothole} & if \quad |f_o| < |f_i| \\ \text{no pothole} & otherwise \end{cases} \quad (4)$$

4. IMPLEMENTATION AND EXPERIMENTS

4.1 Implementation

In order to test its performance, the pothole detection method presented in this paper has been implemented in MATLAB version 7.5.0 (R2007b) utilizing the embedded Image Processing Toolbox. Figure 8 shows two screenshots of the prototype presenting the final images with a detected pothole marked red (Fig. 8a), and a correctly classified non-pothole region (green). The image processing was performed on a desktop PC (Intel Core2 Quad CPU, 2.40 GHz, 2GB RAM).

4.2 Experiments

4.2.1 Data collection

Within the first step, a database of 120 images was collected in order to train (50 images) and test (70 images) the effectiveness of the pothole recognition method. Assuming our sample set properly represents a common population of pavement images captured in a pavement survey, the minimum sample size (N) required for testing can be determined using Equation 5 [29]. In this Equation, z_{crit} is the standard normal deviate that is fixed to be 1.645 when the confidence interval (CI) is 0.90 (90%); D is the total width of the expected CI, which is set to be 0.02 (±10%); and p

297 is a pre-determined estimate of accuracy of the test, which is assumed to be 0.50 (50%, most
298 pessimistic case for sample size calculation). Based on this calculation (including a 90%
299 confidence in our method) we conclude that we need at least 68 test images, which is less than the
300 number of images we actually used for testing.

301 $$N = \frac{4(z_{\text{crit}})^2 p(1-p)}{D^2} \quad (5)$$

302 Most of the images that contain a pothole were found online using the Google search engine,
303 others were cropped from video test files. These video files were captured around the Georgia
304 Tech campus using a remote-controlled robot vehicle prototype being equipped with a HP Elite
305 Autofocus Webcam and a Viliv S5 Ultra Mobile PC (Intel Atom CPU, 1.33 GHz, 1GB RAM).
306 As illustrated in Figure 9, the camera is installed at an altitude of about 2 feet (H≈0.6m) above the
307 pavement surface and its view is directed down and backwards (α≈45°). With regard to our
308 overall research objective described above, our intention is to simulate a high-speed, fish-eyed
309 and tilt-down parking camera mounted on to a common passenger vehicle's rear bumper. The
310 images collected are characterized by a high variety of potholes (shapes and sizes), other defect
311 (cracking, patching, discoloration, etc.) and non-defect asphalt pavements, and lighting conditions
312 (shadows). The average image resolution was 640 by 480 pixels (0.3 megapixels).

313 ***4.2.2 Training***

314 In order to determine proper thresholds for the proposed shape extraction procedure and the
315 appropriate filter size for the described texture comparison procedure, 50 pavement images were
316 randomly selected to manually train our method. Based on varying values for the thresholds and
317 the filter size, precision, recall and accuracy are calculated to measure the performance of each
318 procedure separately. For this purpose, we tested different thresholds/ filter sizes and manually
319 counted the number of true positives (TP, correctly detected shapes textures), false positives (FP,
320 wrongly detected shapes/textures), true negatives (TN, correctly detected as false shapes/textures)

and false negatives (FN, wrongly detected as false shapes/textures). While precision describes the detection exactness or fidelity (TP/(TP+FP)), recall is a measure for detection completeness (TP/(TP+FN)). Accuracy is used to describe the average correctness of a classification process ((TP+TN)/(TP+FP+TN+FN)).

Within the shape extraction procedure proposed above, defect regions identified within the upstream image segmentation process are to be classified as pothole shade regions or entire pothole regions based on the values ε_{max} and r_{max} (Eqn. 2). In order to determine appropriate thresholds for both values we performed a simple optimization strategy depicted in Figure 10. First, we tested different values for ε_{max}, without taking r_{max} into account, since we consider ε_{max} to have the major impact on the results. With regard to Figure 10a, the most suitable value for ε_{max} is 0.85. This is concluded because the recall is very high, the accuracy is at a maximum and the precision is at a satisfying level. Moving to 0.90 would results in a significant recall and accuracy loss without gaining much precision. Within the next step, ε_{max} is fixed to be 0.85 and the value for r_{max} is modified to fine-tune our method. The respective graph presented in Figure 10b identifies 0.02 to be the best choice for the threshold r_{max}. Trying to increase the precision would again result in a significant loss of recall and accuracy. This sequential strategy might not arrive at an optimal solution, because ε_{max} and r_{max} are interrelated in terms of l_{max}. However, our current choices have reached a good solution and prove the method proposed. Moreover, this paper rather focuses on the general method than on parameter optimization to improve the method, which is the scope of future work.

Within the proposed texture comparison procedure, the texture inside a pothole candidate region and the texture of the intact, healthy pavement have to be extracted and compared in order to classify regions as pothole regions. As described above, we propose to use the responses of spot filters to create the feature vectors f_i and f_o, which are the basis for classification (Eqn. 4). Figure 11 presents the precision-recall graph for different filter sizes varying from 23x23 to 35x35 pixels. Since the texture inside a pothole is not related to the pothole size, the size of the

filter only depends on the image resolution, which was adjusted within our image collection. Based on Figure 11 we determined the filter size of 31 pixels to be best suited, because it was the smallest size yielding highest precision and recall. The smaller the filter size, the faster the filter can be applied on an image, which has a positive impact on the performance of our method.

4.2.3 Results

The proposed method was tested on the remaining 70 images, of which 30 contain one or two potholes, 13 have cracks, 5 include repair patches, 7 show several discoloration spots, 4 have dark spot shadows, 1 contains a manhole cover, and 10 do not include any defect areas. Figure 12 presents some examples of detected potholes in our test images. Table 1 illustrates the final test results in terms of the numbers of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN), which have been determined manually. In addition, Table 1 shows that our method reaches an overall accuracy of 86% with 82% precision and 86% recall. All three measures validate that most potholes in asphalt pavement images can be correctly detected.

The results are promising, nevertheless, our approach is vision-based and therefore relies on normal lighting. Like in vision-based crack detection methods, ordinary daylight conditions (ambient light) are enough to create shadows that appear darker than the intact pavement surface. Based on our current test we found that only in a very special case (direct sunlight from the top) potholes might be completely illuminated, so that there are only weak shadows hard to be recognized. In most cases, sunlight falls from any inclined direction resulting in strong shadows, which supports our pothole candidate recognition and shape extraction procedure. Within the texture extraction and comparison step we intentionally use spot filters that are not sensitive to varying lighting conditions and viewing directions [27].

5. CONCLUSION AND FUTURE WORK

Pavement condition assessment is a key component when developing road network maintenance programs. In current practice, pavement image and video data collected by digital inspection

372 vehicles are reviewed by technicians on computer monitors in order to manually detect and assess
373 pothole defects. Since this is time-consuming, costly and highly influenced by the subjectivity
374 and the experience of raters, several research studies focused on automating the procedure of
375 pothole detection. Existing approaches are either based on complete 3D pavement surface
376 reconstruction, which comes along with high equipment and computation costs, or they utilize
377 acceleration sensors that can only facilitate preliminary and rough pavement condition surveys.

378 This paper presented an automated method for pothole detection in asphalt pavement
379 images, which, for example, can be cropped from common pavement survey videos. Under the
380 proposed method, the image is first segmented into defect and non-defect regions. According to
381 the geometric properties of a defect region the potential pothole shape is then approximated.
382 Subsequently, the texture inside a potential pothole region is extracted and compared with the
383 texture of the surrounding non-defect pavement. The region of interest is assumed to be pothole if
384 the texture inside the defect region is coarser and grainier than the one of the surrounding intact
385 surface.

386 The method presented in this paper was implemented in a MATLAB prototype, manually
387 trained and tested on 120 pavement images, which provided a high variety of potholes (shape and
388 size), other defects (cracking, patching, discoloration, etc.) and non-defect pavements as well as
389 diverse lighting conditions (shadows). Within the training procedure we conducted experiments
390 in order to determine the thresholds for shape extraction and texture comparison. Accuracy,
391 precision and recall were calculated to measure the detection performance. The resulting accuracy
392 of 86%, coming along with 82% precision and 86% recall, indicate that most potholes in asphalt
393 pavement images can be correctly detected.

394 In the future, we envisage to improve the pothole detection rate by (a) taking additional
395 visual characteristics into account when extracting the pothole shape (e.g. also use the region of
396 defect texture to approximate the pothole shape), (b) performing parameter optimization to
397 determine true optimal thresholds (ε_{max} and r_{max}) for shape classification, and (c) utilizing

machine learning techniques to more accurately classify pavement textures. For example, Zhu and Brilakis [30] used machine learning classifiers for concrete region detection. In contrast to manually train our method, we want to enable machine learning to automatically train and classify pothole and non-pothole pavement textures. In addition, a sensitivity analysis can help to make our method more robust to varying lighting and viewing conditions. With regard to our overall research objective we will focus on automated pothole detection in pavement videos, rather than in separate images. Instead of processing each video frame separately we want to employ vision tracking techniques to track a pothole, once it is detected, in subsequent video frames. Another important future task is dedicated to the assessment of detected potholes based on estimations and measurements regarding pothole depth and area.

Acknowledgements

The authors gratefully acknowledge the support of this research by the German Academic Exchange Service (DAAD). Furthermore, the authors would like to thank Erick Rivera, Rebekah Myrick and Parth Patel who contributed to this work in terms of constructing the robot vehicle and collecting pavement image and video data.

References

[1] Federal Highway Administration, LTPP – Beyond FY 2009: What needs to be done?, Technical Report, Publication No. FHWA-HRT-09-052, 2009, http://www.fhwa.dot.gov/pavement/ltpp/pubs/09052/09052.pdf

[2] Federal Highway Administration, Variability of Pavement Distress Data From Manual Surveys, Technical Report, Publication No. FHWA-RD-00-160, 2000, http://www.tfhrc.gov/pavement/ltpp/pdf/00160.pdf

[3] A. Bianchini, P. Bandini, D.W. Smith, Interrater Reliability of Manual Pavement Distress Evaluations, J. Transp. Engrg. 136:2 (2010) 165–172.

[4] K.T. Chang, J.R. Chang, J.K. Liu, Detection of Pavement Distresses Using 3D Laser Scanning Technology, in: Proc. of the 2005 ASCE Int. Conf. on Computing in Civil Engineering, 2005, 105pp.

[5] W. Jiaqiu, M. Songlin, J. Li, Research on Automatic Identification Method of Pavement Sag Deformation, in: Proc. of the 9th Int. Conf. of Chinese Transportation Professionals (ICCTP), 2009, pp. 2761–2766

[6] Q. Li, M. Yao, X. Yao, B. Xu, A real-time 3D scanning system for pavement distortion inspection, Meas. Sci. Technol. 21 (2010) 8pp.

[7] K.C.P. Wang, Challenges and Feasibility for Comprehensive Automated Survey of Pavement Conditions, in: Proc. of 8th Int. Conf. on Applications of Advanced Technologies in Transportation Engineering, 2004

[8] Z. Hou, K.C.P. Wang, W. Gong, Experimentation of 3D Pavement Imaging through Stereovision, in: Proc. of Int. Conf. on Transportation Engineering (ICTE 2007), 2007, pp. 376–381.

[9] B.X. Yu, X. Yu, Vibration-Based System for Pavement Condition Evaluation, in: Proc. of the 9th Int. Conf. on Applications of Advanced Technology in Transportation (AATT), 2006, pp. 183–189

[10] K. De Zoysa, C. Keppitiyagama, G.P. Seneviratne, W.W.A.T. Shihan, A public transport system based sensor network for road surface condition monitoring, in: Proc. of the 2007 workshop on Networked systems for developing regions (NSDR) ,2007, pp. 1–6.

[11] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, H. Balakrishnan, The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring, in: Proc. of the 6th Int. Conf. on Mobile systems, applications, and services (MobiSys), 2008, pp. 29–39.

[12] S.S. Rode, S. Vijay, P. Goyal, P. Kulkarni, K. Arya, Pothole Detection and Warning System: Infrastructure Support and System Design, in: Proc. of the 2009 Int. Conf. on Electronic Computer Technology, 2009, pp. 286–290.

[13] Fugro Roadware, Application: Pavement Condition Assessment, Data sheets available at http://www.roadware.com/_lib/pdf/pavement_assesment.zip

[14] Minnesota Department of Transportation, 2009 Pavement Condition Executive Summary, Report Number: MnDOT/OMRR-PM--2009-01, 2009, www.dot.state.mn.us/materials/pvmtmgmtdocs/execsumm_2009.pdf

[15] Minnesota Department of Transportation, Distress Identification Manual, 2003, http://www.dot.state.mn.us/materials/manuals/pvmtmgmt/distressmanual.pdf

[16] South Dakota Department of Transportation, Visual Distress Manual, 2009, http://www.sddot.com/pe/projdev/docs/distressmanual.pdf

[17] Federal Highway Administration, Distress Identification Manual for the Long-Term Pavement Performance Program, Publication Number: FHWA-RD-03-031, 2003, http://www.tfhrc.gov/pavement/ltpp/reports/03031/03031.pdf

[18] Minnesota Department of Transportation, An Overview of Mn/DOT's Pavement Condition Rating Procedures and Indices, 2006, http://www.dot.state.mn.us/materials/pvmtmgmtdocs/Rating_Overview_State.pdf

[19] J. Karuppuswamy, V. Selvaraj, M.M. Ganesh and E.L. Hall, Detection and Avoidance of Simulated Potholes in Autonomous Vehicle Navigation in an Unstructured Environment, in: D. P. Casasent (Ed.), Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision, Proc. of SPIE Vol. 4197, 2000, pp. 70–80.

[20] I. Brilakis, S. German, Pattern Recognition Models for Smarter Infrastructure Sensing, Proc. of the Workshop on Data Mining for Smarter Infrastructure, in: Proc. of SIAM Int. Conf. on Data Mining, 2010.

[21] Z. Zhu, I. Brilakis, Concrete Column Recognition in Images and Videos, Journal of Computing in Civil Engineering (2010), in print.

[22] Z. Zhu, I. Brilakis, Machine Vision based Concrete Surface Quality Assessment, Journal of Construction Engineering and Management 136:2 (2010) 210–218.

[23] G.W. Zack, W.E. Rogers, S.A. Latt, Automatic measurement of sister chromatid exchange frequency, J. Histochem. Cytochem. 25:7 (1977) 741–53.

[24] A.W. Fitzgibbon, M. Pilu, R.B. Fischer, Direct least-squares fitting of ellipses, IEEE Trans. on Pattern Analysis and Machine Intelligence 22:5 (1999) 476–480.

[25] R.C. Gonzalez, R.E. Woods, Digital Image Processing, third ed., Pearson Education International, Upper Saddle River, 2008

[26] I. Brilakis, L. Soibelman, Y. Shinagawa, Construction site image retrieval based on material cluster recognition, Advanced Engineering Informatics, 20:4 (2006) 443–452

[27] T. Leung, J. Malik, Representing and recognizing the visual appearance of materials using three-dimensional textons, International Journal of Computer Vision, 43 (2001) 29-44.

[28] C. Schmid, Constructing models for content-based image retrieval, in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, volume 2, 2001, pp. 39-45.

[29] Eng, J., Sample Size Estimation: How Many Individuals Should Be Studied?, Radiology 227:2 (2003) 309-313.

[30] Z. Zhu, I. Brilakis, Parameter optimization for automated concrete detection in image data, Automation in Construction, 2010, in press.

List of tables

Table 1

Table 1: Method performance for the test of 70 pavement images

Performance	
Total TP	31
Total FP	7
Total TN	42
Total FN	5
Accuracy	85.9%
Precision	81.6%
Recall	86.1%

List of figures

Figure 1

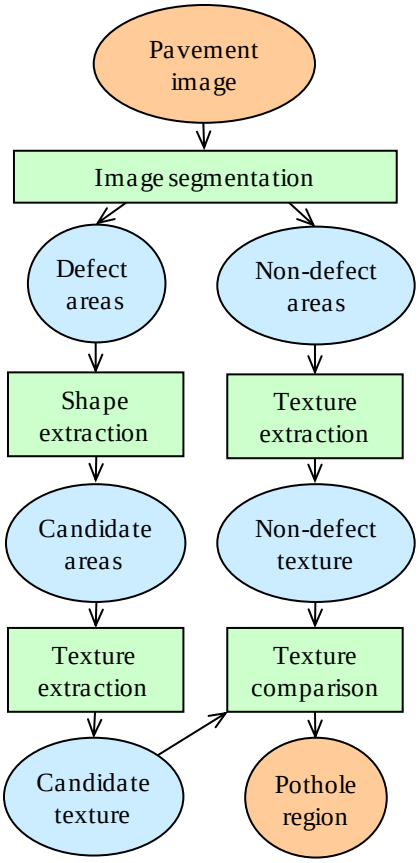


Fig. 1: Pothole detection model

Figure 2

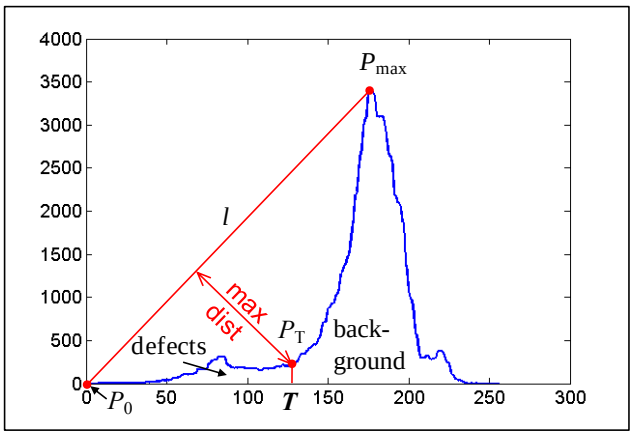


Fig. 2: Histogram shape-based thresholding to separate the background

Figure 3

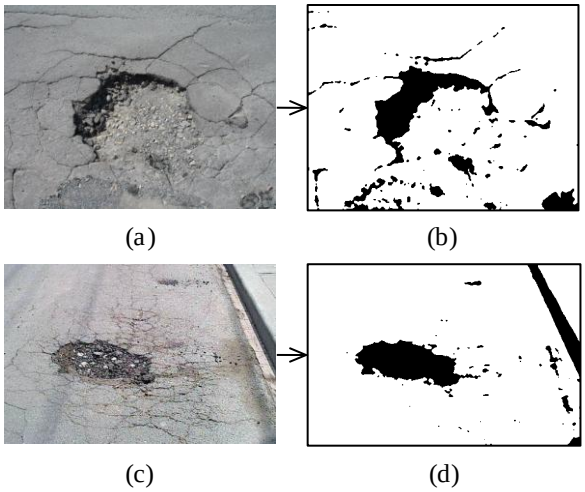


Fig. 3: Result of image segmentation for two potholes (Images courtesy of (a) Chiri Enterprice Inc. and (c) Pinto Potts LLP Solicitor)

Figure 4

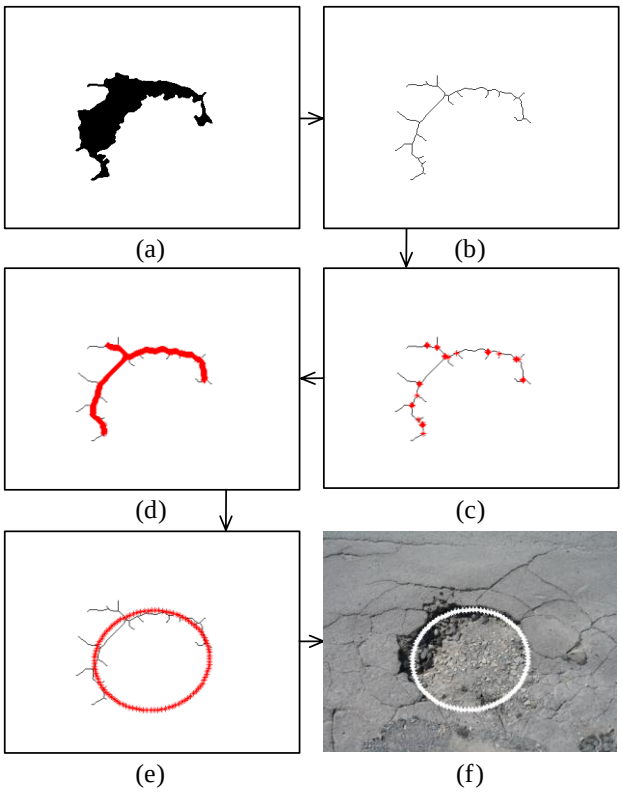
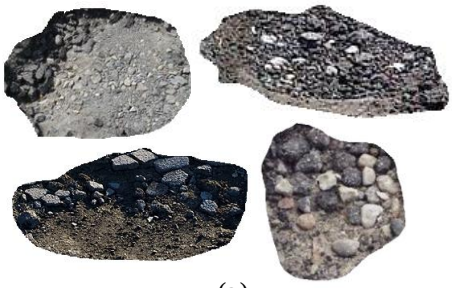


Fig. 4: Shape extraction procedure based on a pothole shade

Figure 5



(a)

(b)

Fig. 5: Typical surface textures inside (a) and outside (b) a pothole

Figure 6

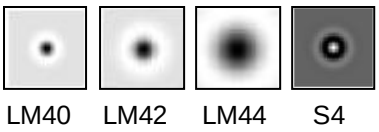


Fig. 6: Leung-Malik (LM) [27] and Schmid (S) [28] spot filters

Figure 7

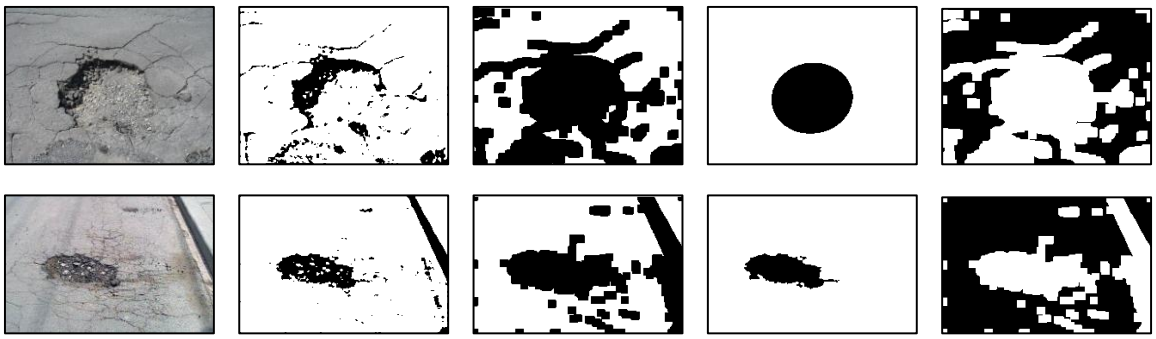
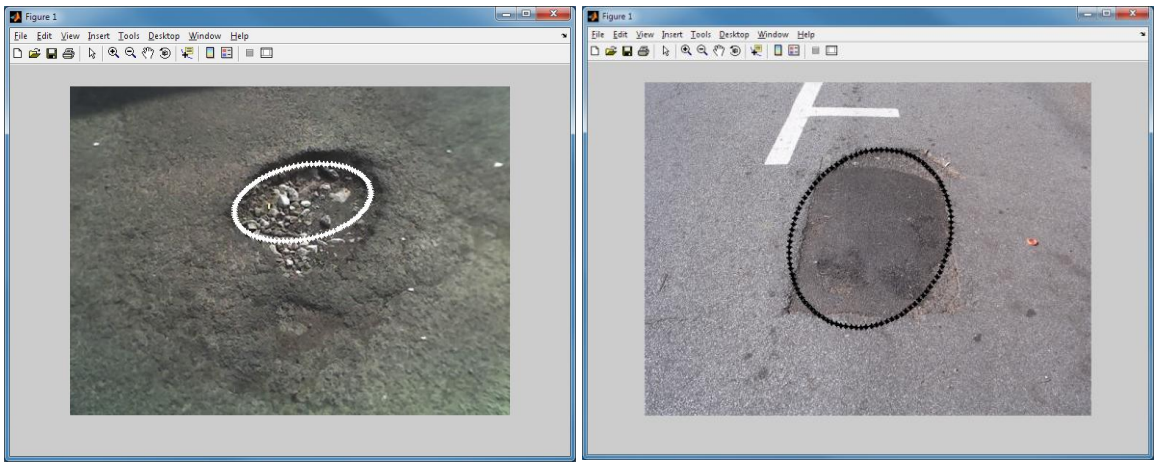


Fig. 7: Procedure to determine the inside region R_i and the outside region R_o : (a) original image; (b) binary image B after image segmentation (defects); (c) binary image B_D after morphological dilation; (d) inside region R_i after shape extraction; (e) outside region R_o (in the binary images the regions of interest are black)

Figure 8



(a) (b)

Fig. 8: Screenshot of the prototype: (a) detected pothole in an image (white ellipse); (b) detected pothole candidate after image segmentation and shape extraction that is correctly identified as a non-pothole region within the texture comparison procedure (black ellipse)

Figure 9

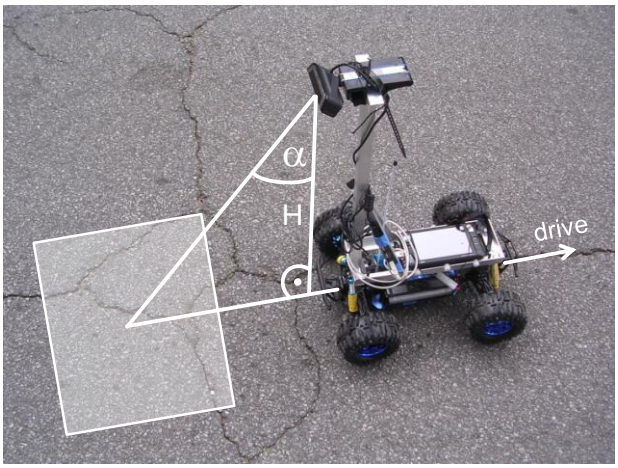
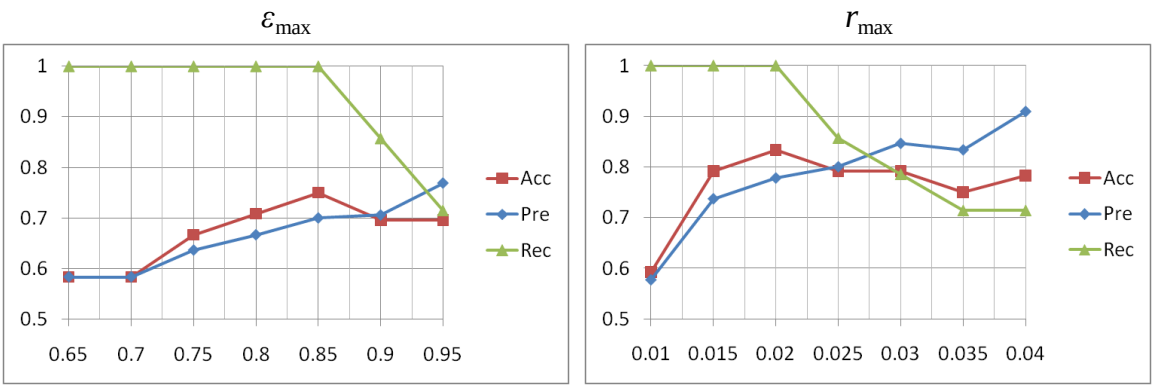


Fig. 9: Robot vehicle for visual pavement sensing. The robot simulates an ordinary passenger vehicle equipped with a high-speed camera mounted on the rear bumper (H ≈0.6m, α≈45°).

Figure 10



(a) (b)

Fig. 10: Precision-recall graphs for the determination of shape extraction thresholds used in Equation 2: (a) max. eccentricity ε_{max} and (b) max. ratio r_{max} of major axis length and area size (ε_{max}=0.85)

Figure 11

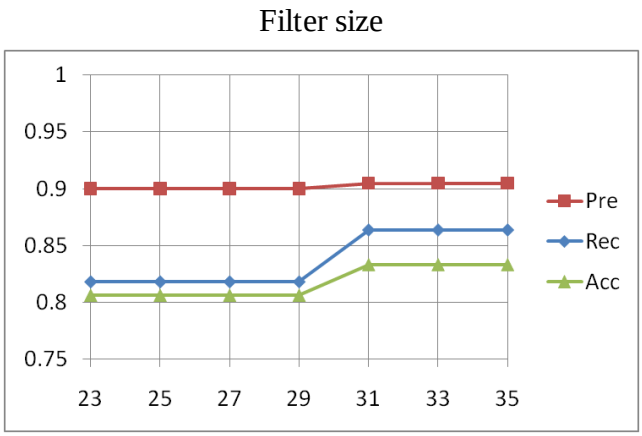


Fig. 11: Precision-recall graph for the determination of the best filter size for texture extraction

Figure 12

Fig. 12: Examples of detected potholes in asphalt pavement images