Pothole Properties Measurement through Visual 2D Recognition and 3D Reconstruction

Conference Paper · June 2012

DOI: 10.1061/9780784412343.0070

Christian Koch
Bauhaus-Universität Weimar

Mani Golparvar-Fard
University of Illinois, Urbana-Champaign

Ioannis Brilakis
University of Cambridge

Pothole Properties Measurement through Visual 2D Recognition and 3D Reconstruction

G. M. Jog[1], C. Koch[2], M. Golparvar-Fard[3], I. Brilakis[4]

[1]PhD Student, Construction IT Lab, School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA; Email: gmjog@gatech.edu
[2]Akademischer Rat, Computing in Engineering, Ruhr-Universität Bochum, Germany; PH +49(0)176-64270897; Email: koch@inf.bi.rub.de
[3]Assistant Professor, Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA; PH (540) 231-7255; Email: golparvar@vt.edu
[4]Assistant Professor, School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA; PH (404) 894-9881; FAX (404) 894-1641; Email: brilakis@gatech.edu

ABSTRACT

Current pavement condition assessment methods are predominantly manual and time consuming. Existing pothole recognition and assessment methods rely on 3D surface reconstruction that requires high equipment and computational costs or relies on acceleration data which provides preliminary results. This paper presents an inexpensive solution that automatically detects and assesses the severity of potholes using vision-based data for both 2D recognition and for 3D reconstruction. The combination of these two techniques is used to improve recognition results by using visual and spatial characteristics of potholes and measure properties (width, number, and depth) that are used to assess severity of potholes. The number of potholes is deduced with 2D recognition whereas the width and depth of the potholes is obtained with 3D reconstruction. The proposed method is validated on several actual potholes. The results show that the proposed inexpensive and visual method holds promise to improve automated pothole detection and severity assessment.

INTRODUCTION

The roadway network of a developed or a developing country contains several thousand centerline kilometers of pavement, which consists of bituminous, concrete, or composite pavements. These pavements vary in age, condition, and performance. Since the majority of the pavement costs in the U.S. are related to road maintenance, programs such as the Long-Term Pavement Performance program (LTPP) are established. LTPP is concerned with data collection, storage, and analysis and product development of the road network (FHWA 2009). Such programs recognize surface condition assessment as a component that requires reliable, good-quality measurements for distresses such as cracks, potholes, and rutting (FHWA 2009).

Currently, roadways are usually maintained dedicated vehicles that collect pavement surface video and profile data. These data are then evaluated manually by technicians on computers. Manual analysis of data is a time-consuming task. Moreover, though well-defined criteria for assessment exist, subjectivity and experience of raters influences final assessment (Bianchini et al. 2010). To address these limitations, several research studies are being conducted for automated detection and analysis of various types of distresses. One particular type of distress that has received more attention in the recent years is pothole. Potholes are elliptical, bowl shaped depressions in the pavement surface. Current methods for pothole detection include 3D surface reconstruction methods using 3D laser-based scanning (Li et al. 2010, Wang et al. 2009, Chang et al. 2005), and stereovision (e.g Wang 2004) systems, vibration based systems that use accelerometers (Yu and Yu 2006), and image based 2D appearance detection approach (Koch and Brilakis 2011a and b). Despite the benefits, none of these techniques can simultaneously use appearance and depth information from images to detect and measure distress. To address the limitation of current state-of-the-art techniques, this paper presents a new approach for detection and measurement of potholes. Our approach is based on integration of image-based 2D recognition with 3D reconstruction algorithms to first verify the existence of potholes and then find geometric properties used for severity levels.

BACKGROUND

Current State of the Pavement Assessment Practice: Process of pavement assessment can be divided into three parts: 1) data collection, 2) distress identification and classification, and 3) distress assessment. Nowadays, inspection vehicles are quickly replacing traditional methods of data collection. These inspection vehicles equipped with several sensors such as cameras for surface imaging, optical sensors for distance measurement, laser scanners for profiling, ultrasonic sensors for rutting detection, and accelerometers for roughness measurements can collect data at speeds up to 60*m/h* (96 *Km/h*) (Fugro Roadware 2011, MNDOT 2009). Despite the automation of data collection process, the second two steps of distress classification and assessment are still predominantly manual.

The collected data is currently analyzed manually by technicians who identify the existence of distresses and assess their severity from the computer screen. Such a manual procedure is labor-intensive and due to the large size of the collected data can make the process non-systematic, ultimately affecting the quality of the assessment. For example, the Georgia Department of Transportation employees 60 full time engineers to assess the condition of its 18,000-mile (29,000*km*) centerline highways (GDOT 2011). Although there are well-defined guidelines for asphalt distress assessment, manual identification and assessment, the experience of the technicians has an impact on the final assessment (Bianchini et al. 2010).

An automated solution for pavement distress classification and assessment is based on 3D surface profiles from time-of-flight laser scanners (Fugro Roadware 2011) or using hybrid imaging devices that integrate digital cameras with infrared lasers to capture consecutive images of lines projected by infrared lasers (Li et al. 2010). Nonetheless, these commercial software applications do not identify and compute the total number of detected potholes. Pavemetrics (2011) claims to use 3D data from laser scanners to automatically detect and analyze potholes but its performance is not documented or validated. Moreover, the cost of laser scanning equipment is significant at vehicle-level and suffers from mixed pixel phenomena that can affect the quality of the results.

Current State of the Research in Pothole Detection: Research related to pothole detection can be classified into 1) 3D reconstruction based, 2) vibration based, and 3) 2D vision-based approaches. 3D reconstruction based methods can further be divided into 3D point clouds obtained from laser scanners or stereovision methods. 3D point clouds obtained from laser scanning are based either on time-of-flight based 3D point coordinates (Chang et al. 2005) or from hybrid systems that use digital cameras to capture lines projected by infrared lasers (Li et al. 2010). However, the cost of laser scanning equipment is still significant at vehicle-level, and suffers from mixed pixel phenomena that can affect the quality of the results. In addition, these works only focus on the accuracy of the 3D measurements, and no particular method for classification of distress is presented.

Another line of research for comprehensive pavement assessment proposes stereovision based surface modeling (e.g. Wang 2004). Using a 3D point cloud, Chang et al. (2005) used a clustering approach is used to compute severity and extent of potholes on the pavement and Wang et al. (2009) proposed a method to identify, locate, classify, and measure depressions such as potholes. Stereovision based methods require both cameras to be very accurately aligned. Given the vibration of the vehicle motion, the cameras may misalign and affect the quality of the outcome.

Vibration based methods use acceleration sensors to "*feel*" the ground conditions as opposed to "*seeing*" them with cameras. Yu and Yu (2006) present an initial evaluation of pavement condition based on a vibration based method. These systems require small storage, are cost-effective, and can be used for real-time processing. However, response of the sensors is modulated by the vehicle response and results are not comparable unless the vehicle service condition is calibrated. In addition, the ratio of false positives in current algorithms is still significantly high (Eriksson et al. 2008). For example, in several cases bridge joints have been detected as potholes. Furthermore, potholes that appear in the middle of a lane (which usually fall between the wheels of a vehicle) are also not identified by this method.

2D vision-based approaches include detection of computer-generated (simulated) potholes that are larger than 2*ft* (60*cm*) and appear white in color (Karuppuswamy et al. 2000) but these assumptions are simplistic and do not reflect real pavement conditions. Recently, a method for automated 2D detection of potholes in images has been presented which includes image segmentation, shape and texture extraction, and finally comparison of the textures from a potential pothole area and healthy pavement area (Koch and Brilakis 2011a). Texture obtained for the healthy pavement area is a small, scattered area due to the presence of a pothole and related fatigue cracking. Hence, healthy pavement texture is not appropriately represented. Moreover, though this method identifies potholes, scattered images do not enable systematic counting of the number of potholes that is a metric used to assess severity.

To address the limitations of the above method, an automated video-based pothole detection method was presented by Koch and Brilakis (2011b). In this method, a pothole is recognized over a sequence of frames that facilitate calculating the total number of potholes. Additionally, representative healthy pavement texture is progressively generated over the sequence of video frames. The experimental results from this method indicate 75% precision and 84% recall in recognition, wherein the precision is the ratio of correctly detected potholes to the number of detected potholes and recall is the ratio of correctly detected potholes to the actual number of potholes. To further improve the automated video-based pothole detection performance and provide the required measurements (width, depth, and number) for severity assessment, the results of the 2D vision-based pothole detection are combined with 3D reconstruction based on the videos from a single camera.

METHODOLOGY

Figure 1 shows the process and data of the proposed method. Videos recorded by a High Definition camera are used initially to detect potholes in videos. Simultaneously, the same video is used for an initial sparse 3D reconstruction. Based on the results of 2D detection and 3D sparse reconstruction the existence of potholes is verified to reduce the number of areas wrongly identified as potholes. Next, the outcome of the sparse 3D reconstruction is improved using a dense reconstruction algorithm. The dense 3D point cloud model and the results generated from the 2D appearance-based recognition are fused to measure the geometrical properties of the potholes and assess their severity.

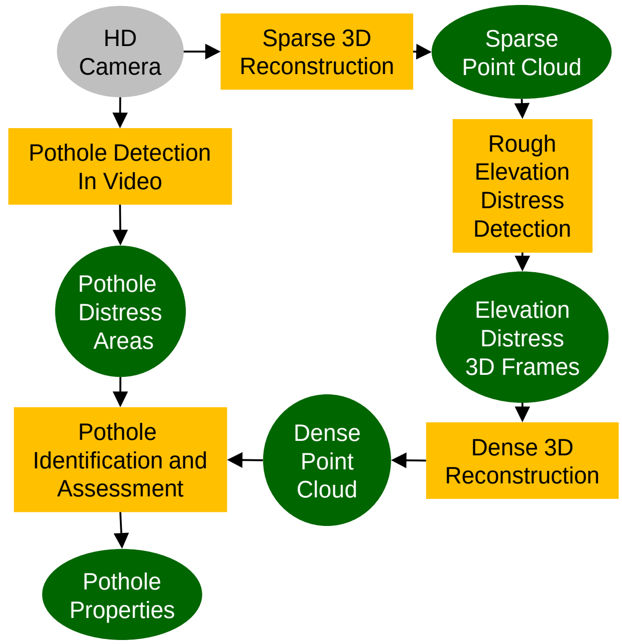

Figure 1. Proposed Methodology

2D Recognition of Potholes from the Video Streams: This step in our methodology builds upon the previous work of Koch and Brilakis (2011a and b). First, the potholes are recognized as bowl-shaped depressions that have a darker surrounding border (low intensity). The inner texture of a pothole is coarser and grainier as compared to the surrounding healthy pavement. Using these properties of a pothole and intensity histogram, shape-based thresholding, based on triangle algorithm (Zack et al. 1977), the image is segmented into distress and non-distress areas. Using the low intensity (darker shadow) regions of a pothole, the skeleton of the shadow is calculated. Then an ellipse shape is approximated using geometrical properties of the skeleton. To differentiate between stains/shadows on the road and potholes, the texture inside the pothole is compared with our training dataset that includes healthy pavement texture obtained from several frames in the video. Once a pothole is identified, it is tracked using a kernel based tracking method proposed by Ross et al. (2008) until it leaves the view. Pothole detection is suspended while a pothole is tracked. Once a pothole leaves the view, pothole detection is reinitiated.

3D Sparse Reconstruction of the Potholes: This step in our methodology builds upon the previous work of Golparvar-Fard et al. (2009) (Figure 2). First, a video stream is broken down into a number of consecutive frames. Next, these frames are analyzed independently to find a series of distinct and scale-invariant feature

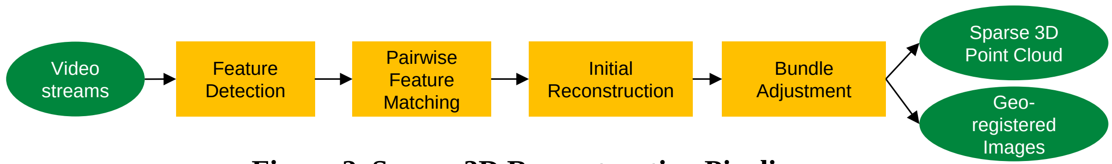

Figure 2. Sparse 3D Reconstruction Pipeline

points. In our work, we modified the existing SIFT algorithm in Golparvar-Fard et al. (2009) with SIFTGPU algorithm (Wu 2007) for rapid feature detection. Once the features are detected, there are matched in a pairwise fashion. Among all matches, a pair, which has a minimum of 200 matches and for which the ratio of percentage of inliers with respect to Homography to the percentage of inliers with respect to Fundamental Matrix is minimal, is selected and this pair is used for initial reconstruction. Finally more frames are added to the bundle adjustment stage and the position of the points, and the cameras configurations are optimized accordingly. The outcome of this 3D sparse reconstruction pipeline is a sparse 3D point cloud along with location and orientation of the images in the scene.

3D Dense Reconstruction and Mesh Modeling of the Potholes: The outputs of the sparse reconstruction are fed into the dense 3D point cloud reconstruction of Golparvar-Fard et al. (2012) to generate dense 3D point cloud models. Next, using the Poisson surface reconstruction approach (Kazhdan et al. 2006), colored mesh models are created. The resulting mesh models can enable automated detection and measurement of the depth within the reconstruction surface.

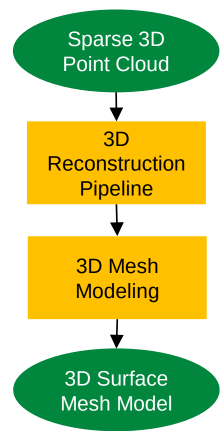

Figure 3. 3D Mesh Modeling Pipeline

EXPERIMENTAL RESULTS

In our work, the main idea is to use monocular cameras for the purpose of surface data collection. In the U.S., the Cameron Gulbransen Kids Transportation Safety Act of 2007 has required all new cars to have rearview cameras by 2014 (NHTSA 2010). We benefit from the availability of these cameras on all vehicles and will use them as a sensing device for pavement monitoring and assessment. In our experiments, a Canon Vixia HD camera was used for data collection at approximately 20*fps* at approximately 20*mph*. Figure 4 shows how the camera was mounted on a car at an approximate location and orientation of a backup camera. Using this camera, a few series of video streams were collected from a road, which had multiple potholes on its surface. Prior to the experiments, the camera was internally calibrated to initiate an approximate estimation for the focal length of the camera. This calibration only needs to be done once for each camera. An approximate value for the focal length helps with a better initialization of the 3D sparse reconstruction pipelines and minimizes the chances of the bundle adjustment optimization step to converge to local minima.

Pothole 2D Recognition and 3D Reconstruction – Several experiments were conducted using the video streams collected from the mounted camera were placed into our 2D recognition and reconstruction systems. Figure 5 shows detection results from video based pothole recognition.

Figure 4. Camera mounted on vehicle

Frame 830 Frame 840 Frame 845

Frame 1720 Frame 1731 Frame 1745

Figure 5. 2D pothole recognition. Left-most column shows detection and the other two columns show subsequent tracking

Figure 6a shows the outcome of the sparse 3D reconstruction for the pothole detected in frames 1720 to 1745 of figure 5. In this figure, 6b shows the outcome of the dense 3D point cloud reconstruction, while 6c and 6d are the Poisson mesh model and the textured Poisson mesh model. In this case, 77 1440×810px frames were used to generate these point clouds model. The sparse and the dense point cloud have a total of 166,134 and 726,538 points and the mesh model contains 86,791 vertices. The generated mesh model was used to measure the geometry of the pothole. As observed in Figure 6e, the depth of the pothole was 0.12*m*, and the width (Figure 6f) was 1.06*m*. Figure 7 shows another pothole detected using 77 1440×810px frames. Table 1 shows the results of the 2D pothole detection that are verified using the outcome of the 3D mesh modeling approach. This table also presents the geometrical properties of the potholes.

CONCLUSIONS

This paper presents a new approach based on 2D recognition and 3D reconstruction for detection and measurement of potholes using a monocular camera. The preliminary experimental results show the promise of applicability of the proposed method. Future work includes real-time detection of the pothole and assessment of the total number of frames that need to be used for 3D reconstruction. This will significantly expedite the 3D reconstruction, as only a part of the surface will need to be meshed. More experiments also need to be conducted for validation. These are part of ongoing research and the results will be presented soon.

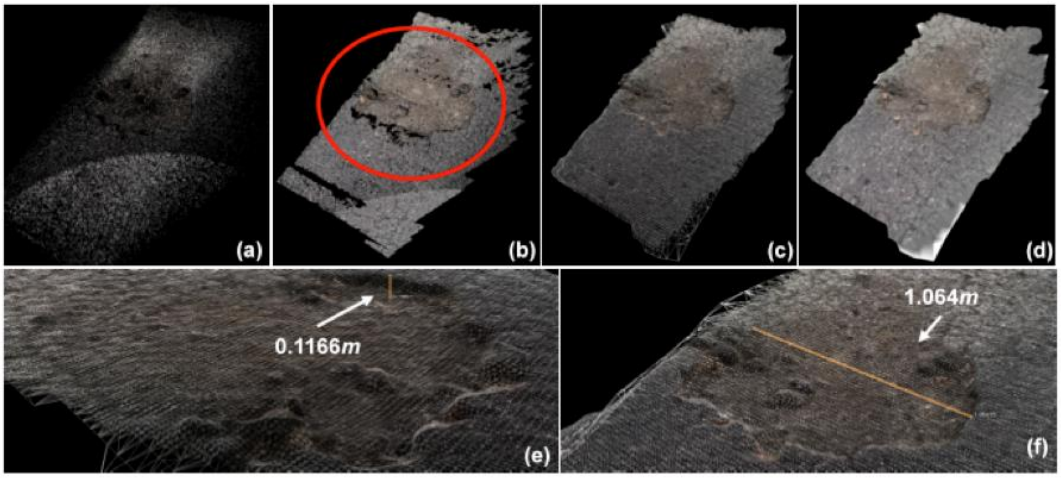

Figure 6. 3D pothole surface reconstruction

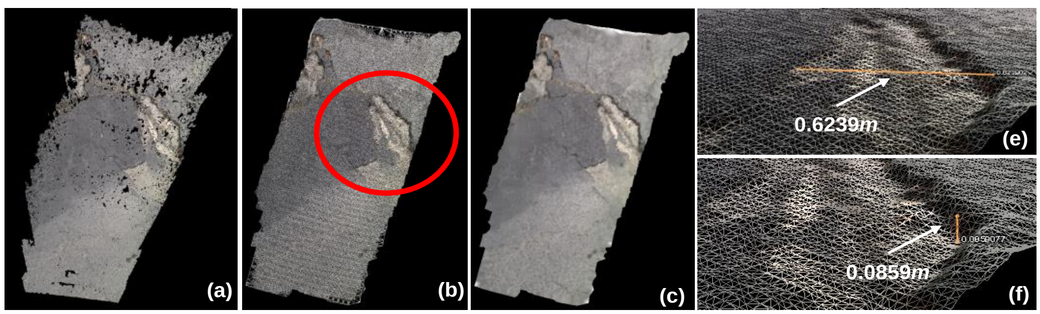

Figure 7. 3D pothole surface reconstruction

Table 1. 2D recognition verified/corrected by 3D reconstruction and pothole measurement (TP- True Positive, FP- False Positive, TN- True Negative)

Pothole #	Frame #	TP or FP	Verification/ Correction by 3D reconstruction	Width (*m*)	Depth (*m*)
1	840	TP	TP	0.630	0.086
2	1720	TP	TP	1.064	0.117
3	2070	FP	TN	--	--
4	2441	FP	TN	--	--

ACKNOWLEDGMENTS

Authors would like to thank Mr. Abbas Rashidi for help in data collection.

REFERENCES

Bianchini, A., Bandini, P., Smith, D. W. (2010). “Interrater Reliability of Manual Pavement Distress Evaluations.” *J. Transp. Engrg.*. 136(2), 165–172.

Chang, K. T., Chang, J. R., Liu, J.K. (2005). “Detection of Pavement Distresses Using 3D Laser Scanning Technology.” *Proc. ASCE Int. Conf. Computing in Civil Eng.*, 105-115.

Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., Balakrishnan, H. (2008). “The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring.” *Proc. MobiSys*, 29–39.

Fugro Roadware (2011). “Application: Pavement Condition Assessment.” Data sheets available at <http://www.fugroroadware.com/related/english-alldatasheets> (Last Accessed Nov. 2011)

Federal Highway Administration (2009). “LTPP–Beyond FY 2009: What needs to be done?” *Technical Report No. FHWA-HRT-09-052*.

Georgia Department of Transportation (2011). “Pavement Preservation in Georgia” (Last Accessed Jul 2011).

Golparvar-Fard, M., Peña-Mora, F., and Savarese, S. (2012). “Automated model-based progress monitoring using unordered daily construction photographs and IFC as-planned models.” *ASCE J. of Computing in Civil Eng.*. (In press).

Golparvar-Fard, M., Peña-Mora, F., and Savarese, S. (2009). “D^4AR model for automating construction progress data collection, processing and communication.” *ITcon*14, 129-153.

Karuppuswamy, J., Selvaraj, V., Ganesh, M. M., Hall, E. L. (2000). “Detection and Avoidance of Simulated Potholes in Autonomous Vehicle Navigation in an Unstructured Environment.” *Proc. Intelligent Robots and Computer Vision XIX*, 4197, 70–80.

Kazhdan, M., Bolitho, M., and Hoppe, H. (2006). “Poisson surface reconstruction.” *Proc., Symposium on Geometry Processing*, 61-70.

Koch, C., Brilakis, I. (2011a). “Pothole Detection in Asphalt Pavement Images.” *J. of Advanced Engineering Eng. Informatics*, 25(3), 507–515.

Koch, C., and Brilakis, I. (2011b). “Improving Pothole Detection through Vision Tracking for Automated Pavement Assessment.” *Proc. of 18th EG-ICE Workshop*,pp. 1-8.

Li, Q., Yao, M., Xu, B. (2010). “A real-time 3D scanning system for pavement distortion inspection.” *J. of Meas. Sci. and Technol.*, 21(015702), pp. 1- 8.

Minnesota Department of Transportation MNDOT (2009), “2009 Pavement Condition Executive Summary.” *Report No.:MnDOT/OMRR-PM-2009-01*.

NHSTA (2010), *Federal Motor Vehicle Safety Standard, Rearview Mirrors*;. Federal Motor Vehicle Safety Standard, Low-Speed Vehicles Phase-in Reporting Requirements, 11.

Ross, D., Lim, J., Lin, R.-S., Yang, M.H. (2008). “Incremental Learning for Robust Visual Tracking”, *Journal of Computer Vision*, 77, 125-141.

Wang, J., Ma, S. Jiang, L. (2009). “Research on Automatic Identification Method of Pavement Sag Deformation.” *Proc., ICCTP 2009*

Wang, K. C. P. (2004). “Challenges and Feasibility for Comprehensive Automated Survey of Pavement Conditions.” *Proc., Conf. on Applic of Adv Tech in Transpt Eng*, 531–536.

Yu, B. X., Yu, X. (2006). “Vibration-Based System for Pavement Condition Evaluation.” *Proc., Conf. on Applic of Adv Tech in Transpt Eng* 183–189.

Zack, G.W., Rogers, W. E., and Latt, S. A. (1977). “Automatic measurement of sister chromatid exchange frequency.”*J. of Histochemistry & Cytochemistry*. 25(7), 741 – 753.