3D Pavement Surface Reconstruction Using An RGB-D Sensor †

Ahmadreza Mahmoudzadeh [1],*, Sayna Firoozi Yeganeh [2], Sara Arezoumand [3] and Amir Golroo [3]

[1] Zachry Department of Civil and Environmental Engineering, Texas A&M University,
   College Station, TX 77843-3136, USA

[2] School of Civil Engineering, College of Engineering, University of Tehran, Tehran 14155-6619, Iran;
   Sayna.Firoozi@ut.ac.ir

[3] Department of Civil and Environment Engineering, Amirkabir University of Technology,
   Tehran 159163-4311, Iran; Arezoumand.s@gmail.com (S.A.); Agolroo@aut.ac.ir (A.G.)

* Correspondence: A.Mahmoudzadeh@tamu.edu

† Presented at the 6th International Electronic Conference on Sensors and Applications,
   15–30 November 2019; Available online: https://ecsa-6.sciforum.net.

Abstract: Data collection plays an important role in pavement health monitoring, which is usually performed using costly devices, including point-based lasers and laser scanners. The main aim of this study measures pavement characteristics using an RGB-D sensor. By recording the depth and color images simultaneously, the sensor benefits the data fusion. By mounting the sensor on a moving cart, and fixing the vertical distance from the ground, data were collected along 100 m of the asphalt pavement using MATLAB. At each stop point, multiple frames were collected, the central region of interests was stored, and a low pass filter was subsequently applied to the data. To create a 3D surface of the pavement, sensor calibration was performed to map the RGB and depth infrared images. The SURF (speeded-up robust features) and MSAC (M-estimator sample consensus) algorithms were used to match the stitched images along the longitudinal profile. A case study of measuring roughness and rutting is applied to test the validity of the method. The result confirms that the proposed system is capable of measuring such indices with acceptable accuracy.

Keywords: pavement health monitoring; RGB-D sensor; Kinect sensor; pavement data collection; pavement rutting; pavement roughness

1. Introduction

Due to the cost burden of evaluating the pavement condition in a network, there has been a lot of attention in using cost-effective health-monitoring methods. Roughness and rutting, as two of the main characteristics of pavement, play an important role in the monitoring of pavement and they are measured using the longitudinal and transverse profiles.

Scholars developed new pavement automated data collection systems that use different types of technologies including ultrasonic sensors [1], point-based lasers [2], laser scanners [3–5], infrared sensors [6,7], digital cameras [8,9], and cellphones [10]. The most important factor besides the accuracy of using these technologies is their costs. Due to the extensive production of digital cameras as a tool in video games, the costs of RGB-D cameras have been significantly reduced. Microsoft Kinect is one of the most used of RGB-D cameras, which can project the rays and collect RGB and infrared images. In 2015 the second version of this camera was released which has higher accuracy (1920 × 1080 pixels for RGB image and 512 × 424 pixels for depth image). Figure 1 demonstrates both sensors' components, including the infrared sensor, the RGB camera, the microphone, and the tilt motor [11,12].

Figure 1. Microsoft Kinect components: (**a**) Kinect v1; (**b**) Kinect v2.

There are two types of technologies that Kinect sensors apply to measure the depth: Time of flight (ToF) and 3D scene reconstruction. The former employs radiation and reflection of a continual laser pulse that has two signals with different phases. ToF technology can be classified into two categories: Direct ToF (pulse) and Indirect ToF (continuous). The former measures the elapsed time between radiation and detection, which is not used in Kinect due to its cost. The latter technology benefits from irradiation of surface with sinus or square waves in a frequency between 10 to 100 MHz. Afterward, four reflection waves with 90 degrees phase differences are absorbed by the camera, and they are converted to distances. Microsoft Kinect v2 uses this technology. It is also worth noting that the phase offsets of radiation in Kinect v2, which uses this technology, are 0, $\frac{2}{3}\pi$ and $\frac{4}{3}\pi$. The 3D scene reconstruction technology uses two different cameras to reconstruct the 3D surface. One of them illuminates the object with a laser beam and the other absorbs the beam reflection. In this method, the vertical distance of a point from the camera could be measured regarding the known parameters such as the distance between two cameras, the angles between the horizontal axes of the sensors, and radiation and reflection waves.

Other scholars in other fields have used the sensors for an automated data collection, while a few studies have implemented this sensor in transportation engineering. The aim of this study is to use this sensor and develop a new image processing technique to compute roughness and rutting. In the next section, the methodology is described in details and the results of the case study of computing roughness and rutting will be provided with the newer version of Kinect called Microsoft Kinect v2.

2. Methods

The methodology of this paper is divided into three main steps as they are described in Figure 2.
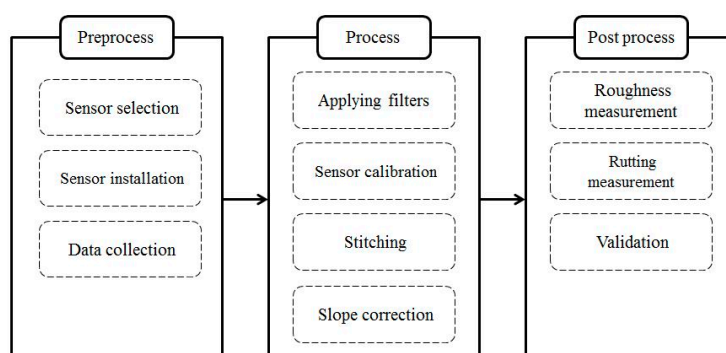

Figure 2. Methodology flowchart.

2.1. Preprocess

A cart was designed and fabricated (schematic view in Figure 3), which was capable of mounting multiple sensors on it and providing different vertical distances between the sensor and the surface.

This study is limited to focusing on one sensor, as it covers almost the width of a lane in the vertical distance of above 1 m from the ground [13].
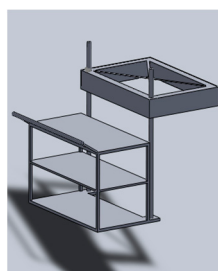

Figure 3. The schematic view of the fabricated cart.

A 100-m section of the pavement was used as the main field of data collection, which has different types of ruttings and different levels of roughness. Five hundred frames of RGB-D images were captured at multiple stations along this road. The stations were defined considering the required amount of coverage between each of the two successive images.

2.2. Process

Image processing techniques consist of four steps, including applying filters, sensor calibration, stitching images, and slope correction. Each of these steps will be discussed in detail in the following sections.

2.2.1. Applying Filters

To smooth the collected images and reduce the noise level, a low pass filter (mean filter), as it is visualized in Figure 4, was applied followed by a Gaussian to the data. The 60 peripheral pixels from each dimension were removed from each matrix and then each pixel of these 500 matrices was averaged to have a high-resolution final matrix.
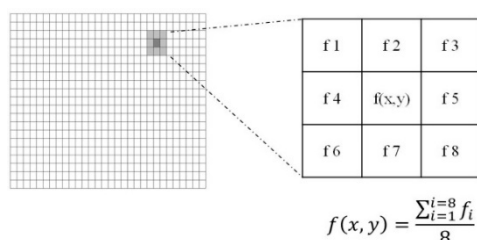

$$f(x, y) = \frac{\sum_{i=1}^{i=8} f_i}{8}$$

Figure 4. Mean filter.

The Gaussian filter was used with the window size of 5 × 5 and the deviation of 2 as two dimensions of the filter. The authors decided to choose this filter instead of the mean filter to make the pixels become more homogeneous within each window. In other words, the filter was used to blur everything that is less than the window size of the filter. The standard deviation of this filter determines the shape of the Gaussian function and the size of the mass was about three sizes of the standard deviation. The Gaussian and Mean filters were used to compute rutting and roughness, respectively.

2.2.2. Sensor Calibration

The MATLAB Toolbox was used in order to determine the intrinsic (focal length, principal points, and the skew coefficient) and extrinsic parameters of the camera, which can compensate for the distortion effect by capturing multiple images from a checkerboard. Fifty pairs of points from the checkerboard in a constant distance were used to develop the matrices. This transformation matrix converts the color image coordinate (x, y) into the infrared image coordinate (X, Y).

By measuring the intrinsic parameters, the point cloud was developed and normalized for each of the cameras. Then, the extrinsic parameters were measured to compute the rotation and transformation matrices in order to have a connection between the right and left cameras (infrared and the RGB cameras). Equation (1) is the transformation matrix and Equation (2) is the rotation matrix in which R is a 3 × 3 rotation matrix, The X_R and X_L are the right and left cameras coordination, respectively, and T is the transformation matrix. Figure 5 shows a 3D view of the camera location and the images from the checkerboard [14].

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{1}$$

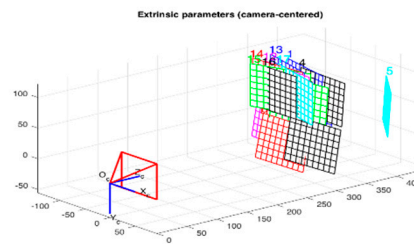$$X_R = R \times X_L + T, \tag{2}$$

Figure 5. 3D view of the camera location and the images from the checkboard.

It should be mentioned that the size and field of view (FoV) of the RGB and the infrared cameras are different, which are 1920 × 1080 and 512 × 424 pixels, respectively. The calibrations' output has reduced to 370 × 460 pixels, which have the matched images.

2.2.3. Stitching

In order to create a 3D reconstruction of pavement surface, the successive images were stitched together. Thus, the colored images were mapped first, followed by mapping the depths matrices. Then, the RGB images and the depth matrices were matched by using the transformation matrix. The speeded-up robust features (SURF) algorithm was used to match the corresponding features between two RGB images. The captured images had overlaps, which enable the algorithm to identify the corresponding features. Figure 6a shows the application of the SURF algorithm. In order to reduce the noise of corresponding features between two successive points, the M-estimator sample consensus (MSAC) algorithm was used. This algorithm is a robust estimation procedure for finding the transformation matrix between sets of images. As it is visible in Figure 6a, detection of two corresponding points was associated with some level of noise. The white lines are associated with the first image and the red correspond with the second one, which is aligning some places. The transformation matrix connects the maximum pairs of points, associated with the images, between two successive images. The random sample consensus (RANSAC) algorithm was used to compute this transformation matrix and to reduce the effect of outlier detection. This algorithm assumes that at least 50% of the points can be mapped together after using the algorithm. Figure 6b shows the images after applying MSAC algorithm.
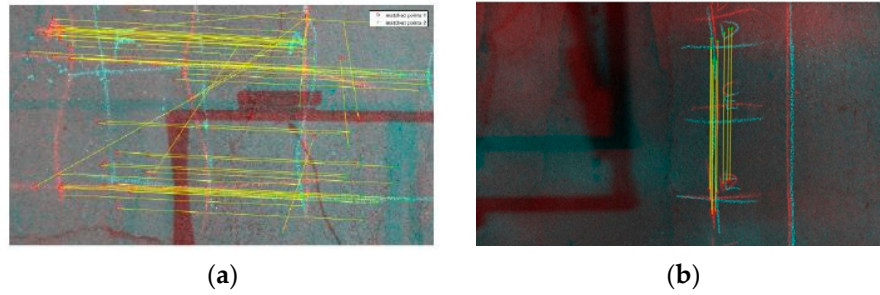
(**a**) (**b**)

Figure 6. Applying the stitching algorithm: (**a**) Speeded-up robust features (SURF) algorithm; (**b**) M-estimator sample consensus (MSAC) algorithm.

2.2.4. Slope Correction

The collected depth data shows a trend, which can cause erroneous data. Some algorithms, including random sample consensus (RANSAC) and singular-value decomposition (SVD), can be applied to the dataset to overcome this issue. The RANSAC algorithm selects some random points from the dataset that corresponds to a surface and fits a plane through the selected points. Then, the amounts of residuals between the fitted plane and the original data points are calculated. By iterating this procedure, the plane that has the least sum of squares of residuals is selected. However, the output of implementing this algorithm is not unique since the points are selected randomly, at each iteration.

The second approach to overcome this issue could be implementing a linear regression method to create a plane that corresponds with the dataset and rotating the plane to create a horizontal plane, which has a zero slope on the z dimension. While the results of these approaches are not unique, this clockwise or counterclockwise rotation changes the 3D fitted plane into a 2D plane. In this study, in order to compensate for the erroneous slope generated by the sensor, the SVD algorithm was applied to the dataset. This algorithm detects the prominent direction of the data spread and rotates the fitted plane corresponds with the data in the two dimensions.

2.3. Post Processing

The International roughness index (IRI) was computed using the equation explained in ASTM 1364-95 (American society for testing and materials). Roughness is defined as the irregularities in the pavement surface, which declines ride quality and increases vehicle depreciation. Moreover, the surface depression, which usually occurs in the path wheel, is called rutting. In this study, the rutting depth was measured by fitting a polynomial equation in the transverse profile dataset.

3. The Summary of Results

Figure 7a,b are showing the stitched images along the road and the transverse profile, which visualizes three different types of rutting.
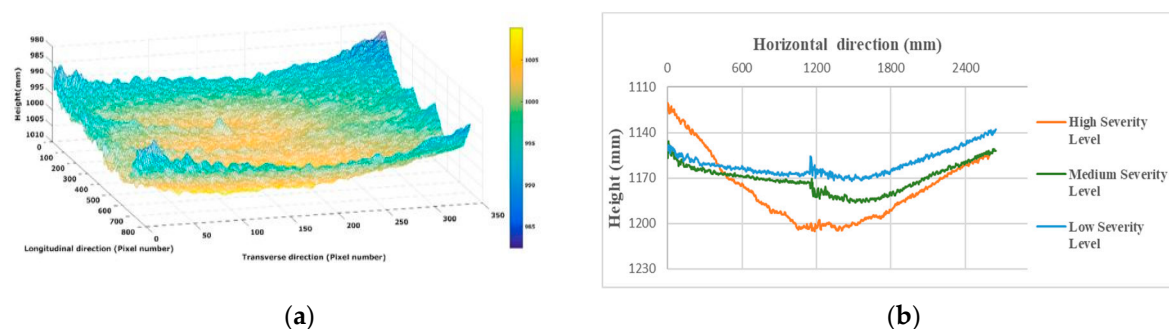

(**a**) (**b**)

Figure 7. Stitched images: (**a**) A 3D point cloud, (**b**) transverse profiles [7].

References

1. Pan, W.H.; Sun, X.D.; Wu, L.M.; Yang, K.K.; Tang, N. Damage Detection of Asphalt Concrete Using Piezo-Ultrasonic Wave Technology. *Materials* **2019**, *12*, 443.
2. Hui, B.; Tsai, Y.J.; Guo, M.; Liu, X. Critical assessment of the impact of vehicle wandering on rut depth measurement accuracy using 13-point based lasers. *Measurement* **2018**, *123*, 246–253.
3. Ye, C.; Li, J.; Jiang, H.; Zhao, H.; Ma, L.; Chapman, M. Semi-Automated Generation of Road Transition Lines Using Mobile Laser Scanning Data. *IEEE Trans. Intell. Transp. Syst.* **2019**, 1–14. doi:10.1109/TITS.2019.2904735.
4. Sollazzo, G.; Wang, K.C.P.; Bosurgi, G.; Li, J.Q. Hybrid Procedure for Automated Detection of Cracking with 3D Pavement Data. *J. Comput. Civ. Eng.* **2016**, *30*, 4016032.
5. Cabo, C.; Kukko, A.; García-Cortés, S.; Kaartinen, H.; Hyyppä, J.; Ordoñez, C. An Algorithm for Automatic Road Asphalt Edge Delineation from Mobile Laser Scanner Data Using the Line Clouds Concept. *Remote Sens.* **2016**, *8*, 740.
6. Chen, Y.L.; Jahanshahi, M.R.; Manjunatha, P.; Gan, W.; Abdelbarr, M.; Masri, S.F. Inexpensive multimodal sensor fusion system for autonomous data acquisition of road surface conditions. *IEEE Sens. J.* **2016**, *16*, 7731–7743.
7. Firoozi Yeganeh, S.; Golroo, A.; Jahanshahi, M.R. Automated rutting measurement using an inexpensive RGB-D sensor fusion approach. *J. Transp. Eng. Part B Pavements* **2019**, *145*, 04018061.
8. Fu, T.; Stipancic, J.; Zangenehpour, S.; Miranda-Moreno, L.; Saunier, N. Automatic Traffic Data Collection under Varying Lighting and Temperature Conditions in Multimodal Environments: Thermal versus Visible Spectrum Video-Based Systems. *J. Adv. Transp.* **2017**, *2017*, 1–15.
9. Radopoulou, S.C.; Brilakis, I. Automated Detection of Multiple Pavement Defects. *J. Comput. Civ. Eng.* **2017**, *31*, 4016057.
10. Yeganeh, S.F.; Mahmoudzadeh, A.; Azizpour, M.A.; Golroo, A. Validation of smartphone based pavement roughness measures. *arXiv* **2019**, arXiv:1902.10699.
11. Corti, A.; Giancola, S.; Mainetti, G.; Sala, R. A metrological characterization of the Kinect V2 time-of-flight camera. *Rob. Auton. Syst.* **2016**, *75*, 584–594.
12. Mahmoudzadeh, A.; Yeganeh, S.F.; Golroo, A. Kinect, a novel cutting edge tool in pavement data collection. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives, Kish Island, Iran, 23–25 November 2015; Volume 40, No. 1W5.
13. Mahmoudzadeh, A.; Golroo, A.; Jahanshahi, M.R.; Firoozi Yeganeh, S. Estimating pavement roughness by fusing color and depth data obtained from an inexpensive RGB-D sensor. *Sensors* **2019**, *19*, 1655.
14. Kim, C.; Yun, S.; Jung, S.W.; Won, C.S. Color and depth image correspondence for Kinect v2. In *Advanced Multimedia and Ubiquitous Engineering*; Springer, Berlin/Heidelberg, Germany, 2015; pp. 111–116.