Automatic Detection of Pavement Surface Defects
Using Consumer Depth Camera

Chenxi YUAN¹ and Hubo CAI²

¹Graduate student, Division of Construction Engineering & Management, School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907; email: yuan89@purdue.edu
²Assistant Professor, Division of Construction Engineering & Management, School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907, PH (765) 494-5028; email: hubocai@purdue.edu

ABSTRACT
The detection of pavement surface defects is a critical task in assessing and monitoring pavement condition. A number of image-based approaches have emerged in recent years to automate the process of surface defect detection. A prominent technology is the consumer depth camera, a camera that captures not only images, but also the varying distance between scene objects and the camera itself in a format of three-dimensional (3D) point clouds, similar to 3D point clouds obtained via Light Detection and Ranging (LiDAR). The main challenge in applying a consumer depth camera to pavement defects detection lies in the low resolution and high random noise of the raw data. This paper presents a method that effectively addresses this challenge. It starts by filtering the signal noise using disciplined convex optimization and then enhances the resolution via moving least square up-sampling. This newly created method was tested in the field to detect different types of pavement defects including cracks, ruts, and potholes. It was found to be effective in detecting and quantifying pavement defects to provide quantitative information to support informed decision making in pavement management.

KEYWORDS: Pavement defects detection, consumer depth camera, point clouds, disciplined convex optimization, moving least square up-sampling.

INTRODUCTION
The detection of pavement surface defects is a critical task in assessing and monitoring pavement condition. Pavement defects affect the traffic safety and efficiency as well as riding quality. Reliable and efficient data collection for pavement condition will make it possible to repair the pavement defects in the early age and hereby to reduce the maintenance cost and prolong the life span of pavement.

Compared to the traditional on-site visual inspection and manual classification approach, automated data collection and processing of pavement defects is much safer and more efficient. Over the past three decades, a number of sensing technologies have emerged to automate this data collection and processing process. In the 1990s, the most popular technology of automated defects detection were based on very high speed (VHS) analog photo-logging or videotaping (Gramling and Hunt 1993; Wang and Elliott 1999). However, it was difficult to digitize analog signals for computer storage and automated processing (Wang 2003). Owing to the rapid
development of digital sensing technologies in recent years, digital cameras/video cameras and laser scanners have been well incorporated into pavement defects detection systems. Bursanescu et al. (2001) introduced a system developed by GIE technologies Inc., which integrated both cameras and infrared laser scanners. Automated road analyzing and inventory vehicles with corresponding data processing software packages and tools have been well received by state transportation agencies (STAs) and are widely used (Wang 2012). Meanwhile, a lot of studies aimed to improve the existing automated detection sensor systems by modifying or optimizing data processing algorithms. For instance, Jahanshahi et al. (2013) put forward an innovative stereo image processing method to detect crack patterns using morphological operation. Sun et al. (2012) proposed a new method based on the sparse representation to detect cracks via analyzing the profile signal using laser scanners.

Despite the proven benefits of automated pavement defects detection and quantification via sensing technologies, high expense remains to be the main barrier to practice in (McGhee 2004; Salari and Bao 2011). Consumer depth camera has emerged as an inexpensive range sensing technology that combines the merits of both 2D cameras and 3D laser scanners. It captures not only images, but also the varying distance between scene objects and the camera itself in a format of three-dimensional (3D) point clouds at the same time. No extra effort or additional processing is needed to align resulting 2D images and 3D point clouds. Another merit is that a depth camera can capture relatively high frame rate video. For instance, SwissRanger SR4000 can reach up to 54 frames per second, which is ideal to be installed on vans for mobile data collection. Based on these merits, Jahanshahi et al. (2012) put forward a new approach using Microsoft Kinect, one of the consumer depth cameras, to detect and quantify the defects in pavements.

The main challenge in the use of inexpensive consumer depth camera in automatic detection and quantification of pavement defects is data quality – low resolution, high level of noise of the raw data. Efficient and effective signal processing and data analysis algorithms must be developed to enhance the data quality and analyze the data to produce high quality information in terms of where defects exist and what are their measures. This paper presents an effective approach to enhance the data quality in two aspects: 1) noise filtering – removing noise using disciplined convex optimization; and 2) resolution enhancing – enhancing the resolution via moving least square up-sampling. The remainder of the paper is organized as follows. First, key contributions are highlighted. Following the contributions section, technical details regarding the preprocessing of the point clouds, the depth-based defects detection and the result’s quantification are discussed in the Methodology section. Finally, the authors discuss research findings, draw conclusions, and point out future research directions.

CONTRIBUTIONS

This study created a new workflow tailored for the use of consumer depth cameras in pavement surface defects detection and quantification. It consists of three main contributions: 1) the use of the disciplined convex optimization method that leads to the effective removal of high levels of random noise that typically exists in
consumer depth camera data; 2) the customization of the moving least square
approach based on the characteristics of local/neighboring points that enables high
quality densification and resolution enhancement; and 3) the creation of effective
algorithms that determine and measure the defect extent.

METHODOLOGY

The main objective of this study is to create a new workflow for processing
low resolution, high noise level consumer depth camera data to accurately detect and
measure pavement defects. Figure 1 illustrates the overall workflow for data
processing and defects detection and quantification. This workflow is detailed below.

1) The system collects low resolution image data and 3D point clouds for
pavement surface. The image data and the point cloud data are registered
together on a common spatial referencing system.

2) The raw image and point clouds data are preprocessed separately. For image
data (i.e. intensity and color), it is up-sampled by matching them with higher
resolution images captured by high definition (HD) cameras, and the
alignment between low resolution image and high resolution image can follow
the method outlined by Mac Aodha et al. (2012). For point clouds data, noise
is reduced by using disciplined convex optimization method; up-sampling and
smoothing are achieved through moving least squares in 3D visualization.

3) The up-sampled image data is used to extract the edges of cracks and potholes.

4) From the previous steps, the raw image links both the up-sampled image and
the de-noised point cloud, therefore identified cracks and potholes from step 3)
are combined with step 2) de-noised point clouds data to measure the depth of
potholes and rutting.

5) Quantitatively measured and detected defects visualized. Results can also be
reported in a compatible format with most of the practical standards as

Figure 1. Proposed workflow for automated pavement defects detection
The consumer depth camera used in this study is SwissRanger SR4000, a phase-shift Time of Flight consumer depth camera. The distance is calculated by measuring the phase-shift between a reference signal and the reflected signal (Fig. 2). The sensor has a resolution of 176×144 pixels. Its standard field of view is 43° (h) × 34° (v). Its detection range is 0.1~5.0m at 30MHz with an absolute accuracy of ±1cm. The methodology discussed in Fig. 1 is also applicable to other types of consumer depth cameras.

Figure 2. Working principle of Phase-shift ToF cameras

Preprocessing of Organized Point Clouds

This section describes two pre-processing steps for the point cloud – de-noising and up-sampling along with their associated algorithms.

De-noising – Disciplined convex optimization

This method was first put forward by Schuon et al. (2009) for depth image super-resolution, which employed the cvx toolbox developed by M. C. Grant (2008). Here, it is used for the organized point cloud de-noising. The target function (Equation 1) aims at finding an optimized point cloud set S, which can achieve the best agreement with several point cloud sets S_i with slightly displaced viewpoints.

Minimize \[ E_{\text{dist}}(S) + \lambda E_{\text{grad}}(S) \]  

(1)

Where \( \lambda \) is the trade-off parameter between \( E_{\text{dist}} \) and \( E_{\text{grad}} \) (\( \lambda = 0.04 \) in this study).

\( E_{\text{dist}}(X) \) measures Root Mean Square (RMS) value of the vertex distance and is calculated in Equation 2:

\[
E_{\text{dist}}(X) = \sum_{i=1}^{N} \| W_i \ast T_i \ast (S_i - S) \|_2
\]

(2)

Where \( \ast \) denotes element-wise multiplication, \( W_i \) and \( T_i \) are confidence matrixes related to amplitude and confidence map to filter the points with low confidence, \( N = 10 \).

\( E_{\text{grad}}(X) \) measures the RMS value of gradient within a certain range. Both \( E_{\text{dist}}(X) \) and \( E_{\text{grad}}(X) \) employ \( \ell^2 \) norm in Euclidean space.
\[ E_{\text{grad}}(X) = \sum_{r,c} \| \nabla S_{r,c} \| = \sum_{r,c} \left\| \left[ G_{r,c}(0,1), G_{r,c}(1,0), \ldots, G_{r,c}(l,w) \right]^T \right\| 
\]

Where \( G_{r,c}(l,w) = \frac{S(r,c) - S(r + l,c + w)}{\sqrt{l^2 + w^2}} \), \( r \) and \( c \) are the index for each point in the organized 176x144 grid, \( l=w=3 \).

One of the profile signals before and after de-noising in Fig. 3 shows that after using disciplined convex optimization, the profile signal is smoother with the random noise level significantly reduced.

Figure 3. Profile signal before and after de-noising

**Up-sampling – Moving Least Square (MLS)**

Moving Least Square is a method of local approximation using polynomials from a set of points (Alexa et al. 2003; Rusu et al. 2008). It can be used in reconstructing surfaces from the unorganized or noisy points captured by 3D scanning devices. As the pavement defects present different geometric features according to varying defects or different pavement surface roughness, MLS is a promising approach to render a smoother 3D surface. Figure 4 gives a simple illustration in 2D about how MLS in Point Cloud Library works. Given the point cloud set \( \{ p_i \} \) (the rectangle points), a smooth surface \( S_p \) (the solid curve) is derived based on the input points \( \{ p_i \} \) using MLS (Alexa et al. 2003), and a set of representative points \( \{ \bar{p}_i \} \) (the dot points) will replace \( \{ p_i \} \). Next a smoother surface \( S_R \) (the dash curve) will display instead of the original surface \( S_p \) by using MLS for a second time. Finally the smooth surface \( S_R \) can be reconstructed by a set of dense points \( \{ r_i \} \) (the triangle points).

The merits of using MLS twice recommended by Point Cloud Library are: (1) remove some outliers and noisy points; (2). reduce the computing time by using a lighter set of control points \( \{ \bar{p}_i \} \).
Relevant parameters (Point Cloud Library) used in this study are summarized in the table below:

Table 1. Parameter List

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Parameter Value</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>polynomial order</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>square Gauss parameter</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>up-sampling method</td>
<td>SAMPLE_LOCAL_PLANE</td>
<td></td>
</tr>
<tr>
<td>up-sampling radius</td>
<td>0.01m</td>
<td></td>
</tr>
<tr>
<td>up-sampling step size</td>
<td>0.01m</td>
<td></td>
</tr>
<tr>
<td>search method</td>
<td>k-dimensional tree</td>
<td>k=3</td>
</tr>
<tr>
<td>search radius</td>
<td>0.03m</td>
<td>The closest distance of two points is 0.004m when camera height is 0.88m.</td>
</tr>
</tbody>
</table>

Figure 5 illustrates that the point cloud turns to be denser and smoother after up-sampling, and some coarse point holes are filled via its local surface estimation (see the red circles in detailed view below).
This section describes the approach for measuring the depth of potholes and rutting via profile analysis of the organized point cloud data.

Profile analysis – Depth of Potholes

First, the result of image-based defects detection is referred so that the location of potholes can shrink into several candidate regions (see the red closed curves in Fig.6c).

Figure 6. Field test of image-based cracks and pothole edges detection
(a: Original 2D Image; b: After edge recognition and enhancement; c: Classification of cracks in blue and pothole edges in red using morphological operation(SDC Information Systems))

According to the distribution of possible potholes from the image processing, the region within row 35~93 and column 65~115 (Fig.6c) is selected for profile analysis. Figure 7 shows the profile of row 51 with the deepest value of -0.9727m. According to the intersections between column boundary and profile itself, the highest value can be found that is -0.9201m. Therefore the depth is (-0.9201)-(-0.9727) = 0.053m. As the pothole shrinks into a smaller region in the image, the computing time is no more than \((93-35)\times(115-65)/(144\times176) = 12\%\) of the original computing time which requires searching all the 144×176 points.

Figure 7. Pothole depth measurement via profile analysis

Profile analysis – Depth of Rutting

Compared to the local defect of a pothole, a rut is a global longitudinal surface depression. Therefore, in order to measure the depth of rutting, the original
(as-built) profile of the road section should be known. As the field of view of one depth camera can’t cover the whole longitudinal surface profile, a local profile of row 110 in Fig.8 without potholes is used as the assuming ground truth to demonstrate the basic principle of rutting depth measurement.

The yellow dash line (Fig.8) represents the horizontal ruler for on-site measurement (Fig.9), and the red dot line is used to refill the depression of the pothole. The vertical green dot line represents the vertical ruler (Fig.9), which slides from left to right to find the maximum vertical difference between refilled row 51 and linearized row 110, and the result is 0.0143m. The whole process is based on simple sliding operation without extra searching algorithms; therefore the computing time is proportional to the number of the columns, which is very efficient.

![Figure 8. Rutting depth measurement via multiple profiles analysis](image)

Figure 8. Rutting depth measurement via multiple profiles analysis

According to the results of field test, the depth of pothole and rutting are 0.0490m and 0.0165m respectively. Compared with the profile analysis result, the relative error of pothole depth is about $\frac{|0.053-0.049|}{0.049} = 8\%$, and for rutting depth it is $\frac{|0.0165-0.0143|}{0.0165} = 13\%$. As we use local profile to represent the global profile of rutting, the 13% error includes the errors from this substitution.

Therefore, the accuracy can be considered at the level of 0.01m ±10% and the profile analysis results can be easily reported in a compatible format specified by AASHTO (2010), ASTM (2011) and FHWA (2003).
3D Visualization--- TIN Model

Triangulated Irregular Network (TIN) Model represents a surface as a set of contiguous, non-overlapping triangles, which are directly made from a set of point cloud. The reconstruction and rendering process are both efficient. Therefore, it’s a good choice using TIN model for the whole road section’s visualization. Figure 10 shows the TIN models before and after MLS up-sampling. The key defects features are remained and meanwhile the surface is smoothed after MLS.

![Figure 10. TIN model (left: Before MLS; right: After MLS)](image)

SUMMARY AND CONCLUSIONS

This paper presents a new workflow tailored for consumer depth cameras in detecting pavement surface defects. The main challenge of the low resolution and high noise level of the raw data is discussed and solved in this study, and a field test is created and investigated for each step of the workflow.

The experimental results reveal that the newly created method is effective in detecting pavement surface defects including cracks, potholes and rutting. First, the newly created method can measure depth of potholes with an accuracy level of 0.01m ±10% according to the results comparison between profile analysis and the field test. Second, the candidate areas for pothole depth measurement shrink to small regions.

Further work will include multiple depth cameras to cover the profile information of the entire lanes in order to measure the global longitudinal surface depression of the rut accurately. Future studies will also aim at integrating defects quantification results to TIN model to make the information retrieval for pavement defects more intuitive and interactive.

ACKNOWLEDGEMENTS

The authors thank Shuai Li, Peng Mao and Jason Park for their valuable suggestions for both field test and paper composition. Any opinions, findings and conclusions in this paper are those of the authors and do not necessarily reflect the views of Purdue University.

REFERENCES


