Road Defect Detection in Street View Images using Texture Descriptors and Contour Maps

David Abou Chacra   University of Waterloo, ON, Canada
Henry Leopold   University of Waterloo, ON, Canada
Jeremy Pinto   University of Waterloo, ON, Canada
Nolan Lunscher   University of Waterloo, ON, Canada
Georges Younes   University of Waterloo, ON, Canada
John Zelek   University of Waterloo, ON, Canada

Abstract

Road quality assessment is a crucial part in municipalities' work to maintain their infrastructure, plan upgrades, and manage their budgets. Properly maintaining this infrastructure relies heavily on consistently monitoring its condition and deterioration over time. This can be a challenge, especially in larger towns and cities where there is a lot of city property to keep an eye on. We review road quality assessment methods currently employed, and then describe our novel algorithm aimed at identifying distressed road regions from street view images and pinpointing cracks within them. We predict distressed regions by computing Fisher vectors on local SIFT descriptors and classifying them with an SVM trained to distinguish between road qualities. We follow this step with a comparison to a weighed contour map within these distressed regions to identify exact crack and defect locations, and use the contour weights to predict the crack severity. Promising results are obtained on our manually annotated dataset, which indicate the viability of using this cost-effective system to perform road quality assessment at a municipal level.

1 Introduction

Road maintenance is a significant concern at municipal, provincial, and national levels. In the United States alone, the American Society of Civil Engineers (ASCE) calculated that \$91 billion a year is invested in road infrastructure [9]. However, at this rate of investment, roads are predicted to decline in quality, as maintaining or improving their current condition would require an additional \$10 - \$79 billion in investment [9]. This decline is also predicted to happen in Canada, as the average reinvestment rate in roads and bridges is around 50% of what is required to maintain them at high quality [5]. Roads must be kept in good condition to prevent damage to cars, car accidents, and potholes among others. The ASCE predicted that poor roads cost US motorists a total of \$67 billion in additional operating and repair costs every year, which amounts to \$324 a year per motorist.

Municipalities often have asset management plans in place that allow them to maintain their roads and, given that their budget is already stretched, allocate their funds properly. For Canadian municipalities, 71% perform data collection about their roads at least every five years[5]. Road quality assessment is currently usually done via surveyors who drive along the roads and check their conditions but semi-automated and fully automated methods have begun to emerge. Road quality assessment is done both to prevent injuries and further damage from occurring as a result of infrastructure failure, and since it is can be more cost effective to maintain property rather than have to replace it. Road surveying can be an expensive and time-consuming task, moreover, manual surveying can be inaccurate, while a large portion of automatic surveying techniques rely on expensive equipment. We aim to use computer vision to automate road quality assessment.

We can split road quality assessment into two main tasks: gathering road data and analyzing that data. Fortunately, there is a comprehensive and useful collection of data that is already publicly available, namely street view images. Road image databases, like Google Street View, contain large amounts of data and are updated quite frequently (approximately every two years for large cities). These databases are accessible to the public, and we can automate mining of street view data to assess municipal assets and infrastructure.

The long-term vision is that other assets such as bridges, tunnels, and sidewalks (which are all a part of the transportation infrastructure) will be similarly mined. Google Street View data also provides a history of images that can be used to track the degradation of municipal assets and determine the optimal time and method of addressing the deterioration of these assets. The most common type of data in street view databases is natural images, i.e., ordinary color images. For our work, we limit the scope to road assessment and provide a demonstration of the applicability of the proposed method on this section of infrastructure with the intention to expand to other assets in the future.

On the data analysis side, we explore various methods to best generate an appropriate quality assessment for roads. We build upon previous work utilizing SIFT descriptors and Fisher vectors to segment road from background in a street view image [3]. After segmenting the street view image we focus on its road portions and try to find cracks, potholes, or any other road degradation artefacts. Our main focus is on road texture, since good quality road and damaged road can be differentiated using texture descriptors. As such, we built upon the describable texture dataset (DTD) methods which achieve the state-of-the-art in texture detection [4]. Our road quality assessment algorithm proves to be accurate and properly detects cracks and damaged road. We perform our tests on a manually annotated subset of Google Street View, as we could not find any publicly available road quality assessment databases to test on.

2 Background

2.1 Road Quality Assessment

In this section, we will discuss the way road quality is currently assessed. The methods used can be split between manual and automated methods.

2.1.1 Manual Methods

Manual methods are somewhat primitive, quite time consuming, and expensive. There is no single universal method to do road quality assessment, however, the process is smilar across different methods with the most variation coming about with the way the road defects are graded. Surveyors drive along every road in a municipality and note any defects they see, the severity is also noted. Operators are usually aided by a device made specifically for this purpose for them to quickly be able to mark the defects they see [10]. Additionally, some systems include an on-board IMU (inertial measurement unit) or accelerometer that records surface roughness [1]. Some systems also use a camera to record the road which allows viewing the data and evaluating it offline.

The surface distress types are ranked based on their severity (low, medium, high), or on a point scale. Depending on the convention followed, different weights are assigned to different surface distress types and their severities, and these are used to determine the surface condition rating (SCR). In the case where an IMU is present, a measure of roughness called the roughness condition index (RCI) is also calculated. The SCR and RCI are combined to generate the pavement condition index (PCI). The United States Department of Transportation follows the convention in equation 1 [1].

$$PCI = 0.6 * SCR + 0.4 * RCI \qquad (1)$$

After the city is surveyed, operators compile collected data and give every city block a PCI based on the number and severity of defects in that block. The PCI is calculated differently depending on the firm conducting the assessment. This results in a map where stretches of road between intersections share a common condition index which roughly describes their condition. This alerts the municipality to roads that require maintenance. Road assessment is done in different time intervals depending on the agency conducting the evaluation and the municipality. There are many flaws to manual road surveying. A condition index shared along a stretch of road is not truly indicative of its quality, a pothole in an otherwise
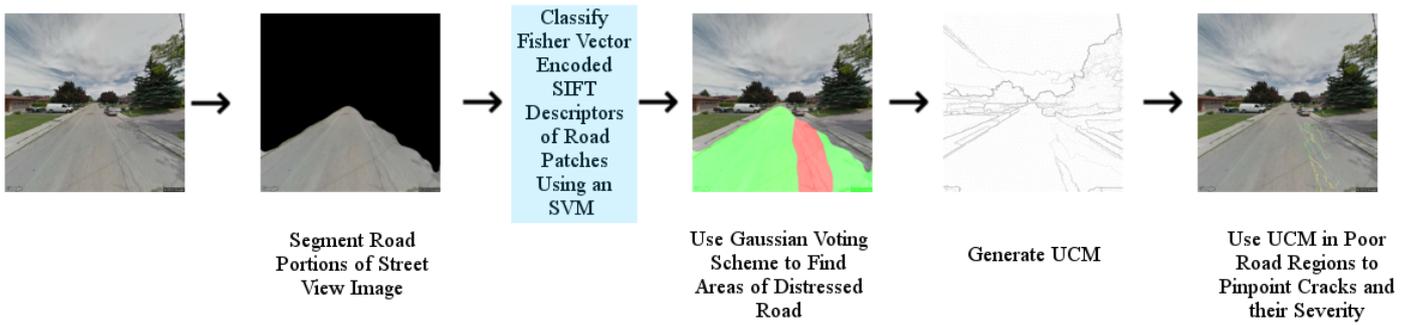
Fig. 1: A visualization of our crack detection algorithm. We first segment the image into 'road and 'background' using the method of Abou Chacra et al. [3]. We then single out the Fisher vectors [11] of the road pixels only. The Fisher vectors are passed through an SVM to classify the vectors as belonging to 'good road' or 'distressed road'. A voting scheme using patches of Gaussians is employed to generate a segmentation. An ultrametric contour map [6] is computed for the image, and cracks are detected by finding locations of high UCM response within 'distressed road' regions. nThese cracks are further distiguished by color-coding them according to severity which is also deduced from the UCM response.

good road, will have a similar PCI to a very poor quality road that is arguably more urgent to fix. Also, there are usually several survey-ors that go around parts of the city, and their condition evaluations can vary significantly as surveyors can assess defects differently or even miss defects completely leading to variability between the cited PCI and the actual PCI [7]. More inefficiency is introduced when traversing larger roads and highways, as surveyors have to traverse these multiple times in the same direction, then in the other direction.

This method is still very popular and used in most road assess-ments. It is relatively archaic, but was very useful in the recent past when the technology was not yet available to automate this pro-cess. It gave municipalities useful insight and helped them allocate their budget properly. The output map is a potent representation of road quality, and our method aims to improve it by adding finer de-tails that are more accurate and helpful in pinpointing problematic sections of road. Our automated method strives to be a more ob-jective one, minimizing the subjectivity introduced by human error and observation. Municipalities can use this data to more efficiently plan road repairs.

2.1.2 Automated Methods

The frequency of performing road quality assessment led to a good deal of work aimed at automating it. The prevalent automation methods involve attaching an apparatus to a truck that also drives along the city [2, 10]. This apparatus can be attached to several locations on the vehicle, for example in tow, on the top or bottom of the vehicle. The main sensing portion contains a combination of one or more cameras pointed orthogonal to the road, LIDAR or LASER scanners, GPSs, IMU's, and accelerometers, among others. Notable among these is the Laser Crack Measurement System (LCMS) produced by Pavemetrics. The information from these sensing apparati is aggregated and different vision and sig-nal processing methods are used to determine the pavement qual-ity. These methods are quite powerful and yield highly precise re-sults.

Automated methods are being used more and more often, and, depending on the algorithm, they can be quite precise. Gabor fil-ters have been convolved with images from a Pavemetrics system, and cracks are found by thresholding the output of this convolu-tion, which can pinpoint crack pixels with an average precision of 81% and an average recall of 84% [12]. CrackIT [10] merges multi-ple preprocessing techniques and clustering algorithms (K-means, Gaussian Mixture Models, among others) to detect cracks and their types and also yields favourable results with a 93.5% F-measure for their best performing algorithm. The LCMS comes with pro-prietary software to detect defect type and severity, it utilizes all aspects of the LCMS, from the texture and LIDAR information (to detect and measure crack depth), to the IMU information that de-tects rutting. The main caveat is that the input sensory data can

be expensive and tedious to procure. The sensing apparatus has a limited scope of vision, specifically 4 meters wide for the Pave-metrics devices. It is usually "looking down" at the road (pointing orthogonal to the road), so the vehicle still needs to drive up and down every road lane in the area to be checked. This also means parts of the road that are not surveyed by the vehicle will not be sensed, such as portions with parked cars, or portions blocked off, or even road shoulders. Sensing requires high resolution cameras (almost 1 pixel per millimetre of road in the case of the LRIS), and expensive LIDAR scanners, so there are usually only a few sensing vehicles that can do the sensing, which would require more time in large areas. Another drawback is that the raw data is somewhat large in size; the Pavemetrics devices use up to 1 GB per Km. While this method definitely has the benefit of having extra detail going into road analysis, it comes at a cost that is both monetary, in the form of expensive sensory devices in the range of a hundred thousand dollars, and temporal, in the form of long data acquisition and processing times.

A few other methods utilize street view images for road quality assessment. An earlier work [8] utilizes a simplified (and cheaper) version of the LIDAR methods mentioned above attached to city vehicles, namely buses and police cars, along with a dashboard camera. While interesting, it falls prey to sensor noise, however it is an order of a hundred times cheaper than the above methods. In their work, Varadharajan et al. use a dashboard camera and drive around the city of Pittsburgh, Pennsylvania to collect their own data, essentially similar to street view data [14]. They detect the ground plane, then over-segment it into superpixels and generate various descriptors for these superpixels which give way to a binary classifier used to identify cracks inside the superpixels. A different approach filters the input image rigorously by background subtrac-tion followed by wavelet-based de-noising, then, locally thresholds patches from the image (by utilizing Otsu's method to quickly find these local thresholds) resulting in a binary image [2]. Horizon-tal and vertical histograms are used to classify the crack type; the system can distinguish between longitudinal, transverse and alli-gator cracks. These methods are effective, but slightly simplistic and can be improved with better road segmentation, and a more powerful classifier. A survey of other methods [13] also tests the performance of older methods, and highlights the fact that these crack detection methods operate locally and can be fooled by the low signal-to-noise ratio (SNR) in pavement images.

Unfortunately, we could not find a standardized road defect dataset to train or test our method, and we resorted to creating one ourselves by manually annotating defect regions in street view images. This prevented us from being able to compare our results to any of the methods presented in this review as the datasets they tested on are not publicly available. With this in mind, our work focuses on finding any and all road defects in an input image, with the knowledge that classifying these defects can be done per the methods mentioned here.
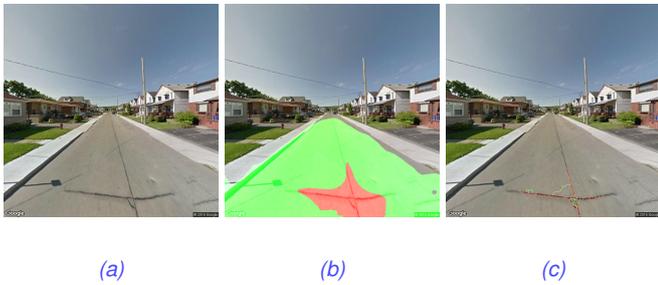
<div align="center">(a)   (b)   (c)</div>

Fig. 2: Sample crack detection results of our algorithm. (a) The street view image is shown. (b) The image is segmented using the initial region segmentation scheme [3] is annotated with a green overlay for detected 'good' road and red for detected 'distressed' road. (c) The crack detection scheme displays severities of detected defects in green for 'mild' defects, orange for 'medium' and red for 'severe' defects.

Table 1: Results on crack detection for images of a Google Street View subset.

Method	F-measure	Precision	Recall	FPR
Our Method	92.84%	86.64%	100%	0.02%

3 Algorithm and Results

Our crack detection algorithm is visualized and described in Figure 1. In brief, our algorithm densely samples windows from the detected road in the input image, then finds SIFT descriptors in those windows. These descriptors are then encoded using the Fisher Vector formulation [11]. The encoded features are then classified using an SVM classifier into "good quality" and "damaged" road. This is followed by a window-by-window voting scheme utilizing Gaussian windows to generate a segmentation that detects damaged road regions. We then utilize an ultrametric contour map (UCM) [6], to detect defects precisely inside damaged road regions.

Sample results are shown in Figure 2, and our numerical results are shown in Table 1.

4 Conclusion

Despite not having quantitative results to compare against, we saw that the method was able to single out cracks and poor road regions well, and perform quite favourably. The recall of 100% is an indication of the performance of the initial crack segmentation step. The high precision of the road segmentation algorithm [3] provided us with a reliable segmentation on which to detect cracks. We can also see that the portions of the road the segmentation algorithm missed on, in the distance and on the boundaries, were not detrimental to the crack detection algorithm. The precision could be improved by making the crack detection algorithm stricter, however, we did not deem that necessary after qualitatively assessing the algorithm output. Despite the low quality and coarseness of our manually labelled data, the classifier was able to truly learn what 'distressed' road looks like. Texture was the key cue in detecting poor roads, and it is the one we utilized, and it predictably generated good results. Whether the data is collected from an online database like Google Street View, or on a municipal level, the method is still applicable. As long as the images show these distress artefacts, the algorithm can be trained to detect them.

References

[1] Federal Highway Administration. *Pavement Distress Identification Manual*. United States Department of Transportation, 4 edition, 2009.

[2] Guanqun Bao. *Road Distress Analysis Using 2D and 3D Information*. PhD thesis, The University of Toledo, 2010.

[3] David Abou Chacra and John Zelek. Road segmentation in street view images using texture information. In *13th Conference on Computer and Robot Vision (CRV) 2016*, 2016.

[4] Mircea Cimpoi, Subhransu Maji, Iasonas Kokkinos, Sammy Mohamed, and Andrea Vedaldi. Describing textures in the wild. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3606–3613, 2014.

[5] CIRC. Canadian infrastructure report card. Technical report, Canadian Construction Association, Canadian Public Works Association, Canadian Society for Civil Engineering, Federation of Canadian Municipalities, 2016.

[6] Piotr Dollár and Lawrence C Zitnick. Structured forests for fast edge detection. In *IEEE International Conference on Computer Vision*, pages 1841–1848, 2013.

[7] Christian Koch, Kristina Georgieva, Varun Kasireddy, Burcu Akinci, and Paul Fieguth. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Advanced Engineering Informatics*, 29(2):196–210, 2015.

[8] Christoph Mertz. Continuous road damage detection using regular service vehicles. In *Proceedings of the ITS World Congress*, 2011.

[9] American Society of Civil Engineers. Report card for america's infrastructure. "http://www.infrastructurereportcard.org/a/documents/Roads.pdf", 2013.

[10] Henrique José Monteiro Oliveira. *Crack Detection and Characterization in Flexible Road Pavements using Digital Image Processing*. PhD thesis, Universidade de Lisboa - Instituto Superior Técnico, 2013.

[11] Florent Perronnin, Jorge Sánchez, and Thomas Mensink. Improving the fisher kernel for large-scale image classification. In *Computer Vision–ECCV 2010*, pages 143–156. Springer, 2010.

[12] M Salman, S Mathavan, K Kamal, and M Rahman. Pavement crack detection using the gabor filter. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 2039–2044. IEEE, 2013.

[13] Yi-Chang Tsai, Vivek Kaul, and Russell M Mersereau. Critical assessment of pavement distress segmentation methods. *Journal of transportation engineering*, 136(1):11–19, 2009.

[14] Srivatsan Varadharajan, Sobhagya Jose, Karan Sharma, Lars Wander, and Christoph Mertz. Vision for road inspection. In *IEEE Winter Conference on Applications of Computer Vision*, pages 115–122. IEEE, 2014.