Continuous Road Surface Distress Detection

PI: Christoph Mertz, The Robotics Institute, Carnegie Mellon University
Research team: Martial Hebert, Sobhagya Jose, Karan Sharma, Srivatsan Varadharajan, Lars Wander, and Jinhang Wang

Introduction

State and local maintenance departments are tasked with keeping roads in good repair. One part of this task is to monitor the degradation of road surfaces, which manifests itself with the presence of cracks, potholes, and other distress.

Currently, this is done by (Figure 1): (i) inspectors who visually judge the road condition, (ii) specialized vehicles which measure the distress with cameras[1,3] or laser devices[2,4], or (iii) citizens who call in their observations. The first method is tedious and often inconsistent if several inspectors do the inspections. The second method does not have these problems, but it is generally expensive. The cost can range between $40 and $230 per mile[3]. Because of budget constraints one stretch of road is usually only traversed by the specialized vehicle once every two years. This method is effective for interstates and highways, but it is not practical for the inner roads in a city. As for the third method, reports by citizens are generally only about severe problems, like large potholes on main roads. Each of the three methods has its strength, but none of them is able to monitor the road on a continuous basis at low cost with minimal human intervention.

Figure 1 Methods of road inspection: Camera based[3] (left) and laser based[4] (middle) specialized vehicles, hotline (right top) and road inspector (right bottom).

The approach of the road inspection system developed in this project is to use images or videos collected by commodity devices such as smartphones. The devices are mounted in vehicle that travels the roads for other purposes and therefore no dedicated vehicles or drivers are needed. The captured

---

[1] http://www.pavemetrics.com/
[2] http://www.greenwood.dk/profiler.php, http://www.roadware.com/products/ARAN-Subsystems/
[3] http://www.grbj.com/articles/76949-pavement-inspection-keeps-improving
[4] LCMS - Laser Crack Measurement System

images are analyzed automatically with machine vision algorithms and the resulting distress scores are passed to the asset management system.

Methodology

Data collection

The data collection system is shown in Figure 2. A smartphone is mounted on the windshield and is powered through the cigarette lighter. A device is plugged into the OBDII and sends out readings via Bluetooth. While the vehicle is driving the smartphone collects images or videos of the outside and tags them with time, GPS, accelerometer, gyro, speed from the OBDII and other selected information. The data is manually downloaded or transmitted to a central computer via Wi-Fi where it is analyzed.

Figure 2 Smartphone mounted inside a vehicle. It is powered by the cigarette lighter. An OBDII-to-Bluetooth devise is installed under the dashboard.

For the data collection we developed an app for Android phones. We implemented it on the Galaxy Camera. A screen snapshot of the app is shown in Figure 1. After starting the app the user only needs to tap on one button to start or stop the data collection. The user can also go to the settings menu and change the update rate, the resolution, or switch between collecting images or videos.
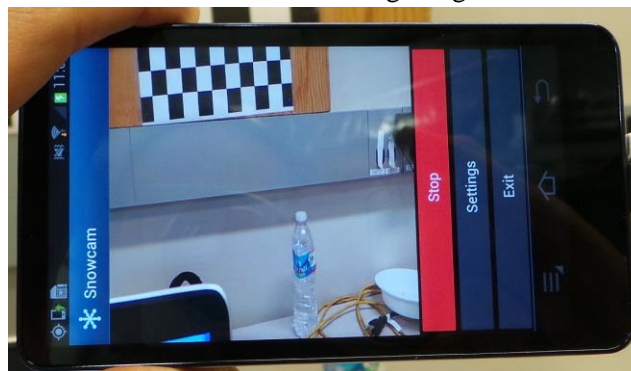

Figure 3 Screen snapshot of the collection app.

Furthermore, the user can download the images through Wifi and Dropbox by tabbing one button in the settings menu. We made the app open source. A simplified version without Dropbox and OBDII can be found at https://bitbucket.org/lwander/snowcam. The full version is at https://bitbucket.org/lwander/vehicle-state.

One of the key ideas behind our data collection system is that it can be easily mounted on any vehicle. Of particular interest are vehicles that drive on the roads on a regular basis (Figure 4), e.g. garbage trucks drive through every neighborhood once a week. It is therefore possible to collect data frequently without the need for a dedicated vehicle or a dedicated driver.

Figure 4 Examples of vehicles that drive on the road on a regular basis.

Data display

The collected data is already useful by itself to the road maintenance department without any further analysis. The GPS tagged images can be display in the asset management system of the department or with free software. An example is shown in Figure 5 where the data is displayed on Google Earth. One can see the satellite view of a neighborhood and small yellow and white arrows. The arrows point in the direction the vehicle was driving. If one clicks on one of the arrows, the corresponding image appears. This will allow the user to visually inspect the road from a computer instead of having to go out. With the slider on the top left the user can also choose the time window of interest if he or she likes to see how the road surfaces were in the past.
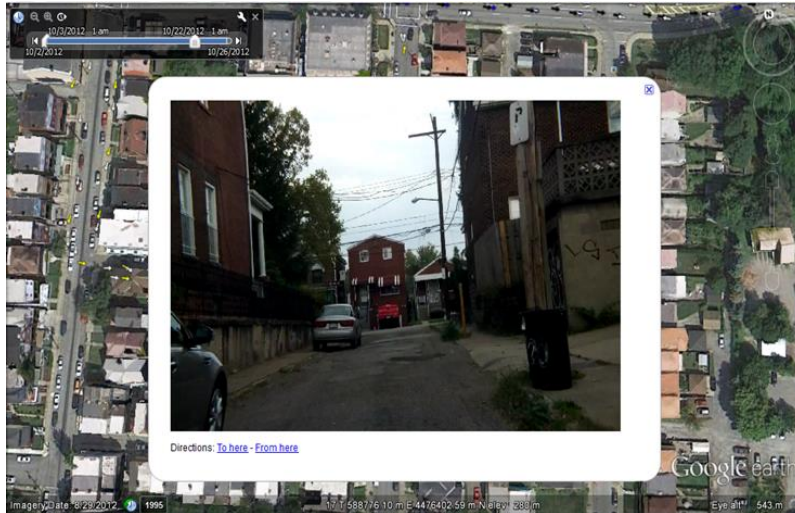
Figure 5 Example of road image displayed on Google Earth. The small yellow and white arrows on the street are markers. When clicking on the marker the corresponding image appears.

Data analysis

The analysis of the images to determine the amount of damage of the road is an interesting and challenging computer vision problem. A paper[5] describing the detailed methods was accepted for publication. Here we will present an outline.

Automatic selection of good images

The images are taken during different weather and lighting conditions. Some images also have blur caused by the movement and vibrations of the vehicle and exaggerated by the rolling shutter of the camera. Many of the roads are traversed many times by the test vehicle and therefore multiple images of the same road are in the data base. We automatically select the most favorable images. Since the images have time and GPS stamps, one can look up the weather conditions when the image was taken. It is best when it is overcast or cloudy. Precipitation creates unwanted artefacts in the image and sunshine results in shadows that can look like cracks or patches. The blur is measured[6] and images with too much blur are discarded.

Computer vision pipeline

A typical image of the road that was captured from inside the car is shown in Figure 6. A significant part of the image is sky, buildings, and other objects that are not of interest to find the road distress. It is necessary to determine the part of the image that is road.

---

[5] S. Varadharajan, S. Jose, K. Sharma, L. Wander, and C. Mertz, "Vision for Road Inspection" IEEE Winter Conference on Applications of Computer Vision, March 2014.

[6] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas. The blur effect: perception and estimation with a new no-reference perceptual blur metric. Human vision and electronic imaging XII, 6492:64920I, 2007.

Figure 6 A typical road image.

An algorithm that finds geometric context[7] is used to determine the ground plane in the image. The resulting segmentation into ground plane, objects, and sky can be seen in Figure 7.

Figure 7 Image segmented into ground plane (green), objects (red) and sky (blue).

It is advantageous to further segment the ground plane into so-called superpixels. These are patches with similar texture. A modified SLIC[8] algorithm creates the superpixels shown in Figure 8. Notice how the regions with cracks are separated from the regions without cracks.

Figure 8 The ground plane segmented into superpixels.

---

[7] Hoiem, Derek, Alexei A. Efros, and Martial Hebert. "Geometric context from a single image." Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. Vol. 1. IEEE, 2005.

[8] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 34(11):2274–2282, 2012.

To classify the superpixels as cracked or non-cracked regions 138 different features are calculated:

Features	Indices
Mean RGB	1-3
HSV for mean RGB	4-6
Hue histogram	7-11
Saturation histogram	12-14
Centroid (x,y)	15-16
10th, 90th percentile x	17-18
10th, 90th percentile y	19-20
Texton histogram	21-70
Multi-scale LBP histogram	71-138

These are color and contrast (RGB, HSV, Hue, Saturation), spacial (x, y), and texture (texton and LBP) based features. The textons[9] and LBP[10] (linear binary pattern) features are illustrated in Figure 9.
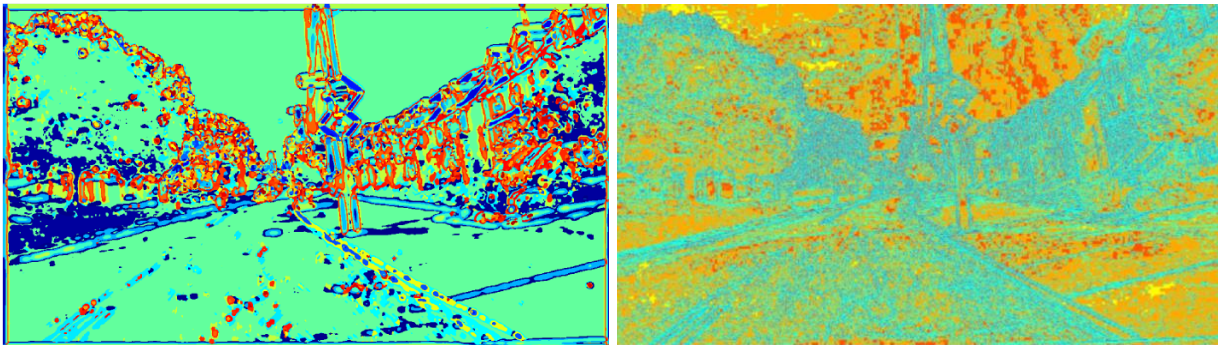

Figure 9 Illustration of textons (left) and linear binary patterns (right).

Finally, a Support Vector Machine and Multiple Instance Learning[11] are employed to classify the superpixels. The result can be seen in Figure 10. Patches with cracks are shown in red, the intensity of red indicates the severity of the cracks.

---

[9] Jitendra Malik, Serge Belongie, Thomas Leung and Jianbo Shi, "Contour and Texture Analysis for Image Segmentation". International Journal of Computer Vision, 43(1), 7-27, June 2001.

[10] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24.7 (2002): 971-987.

[11] Andrews, Stuart, Ioannis Tsochantaridis, and Thomas Hofmann. "Support vector machines for multiple-instance learning." Advances in neural information processing systems. 2002.

Figure 10 Classification result. The color indicates the severity of distress: blue = no cracks, red = cracks.

Computation of distress score

After classifying the superpixels in the image according to their crack content, it is now possible to calculate a distress score. The score is the ratio of the area with cracks and the total area in front of the vehicle (Figure 11). The area with cracks is also weighted with the severity.

Figure 11 The area in front of the vehicle (outlined with a white line in the left image) is projected into a bird's eye view (right). The distress score is the area with cracks (red) divided by the total area. In this case the distress score is 0.012.

Further use of the data

There are many more things that can be done with the data. To make better use of the many successive images they can be projected into bird's-eye-view and stitched together. An example is shown in Figure 12. We anticipate that the detection of cracks will be more efficient on such an image. It is also easier for an user to judge the amount of cracks.
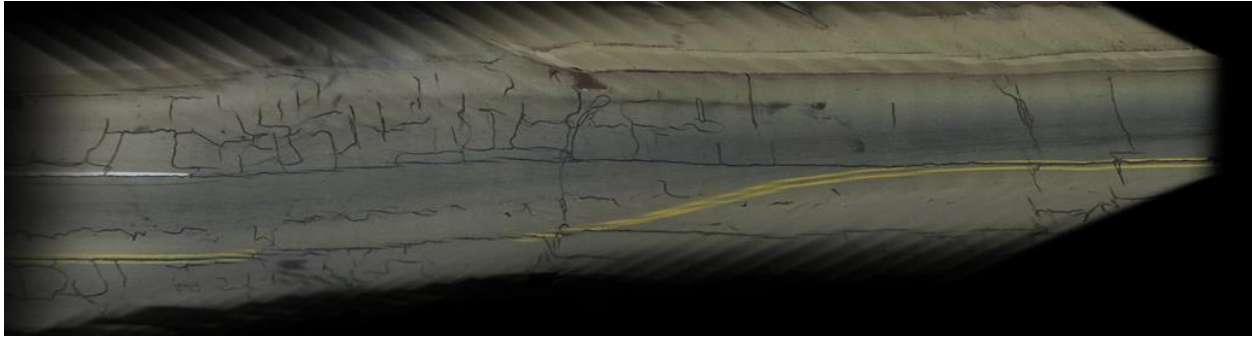
Figure 12 A bird's-eye-view of a road stitched together from about 60 images.

Results

We have mounted the system on a personal vehicle and collected data for over a year. The routes included daily commutes and other typical day-to-day driving. Figure 13 shows the data overlaid on a satellite view of Pittsburgh.
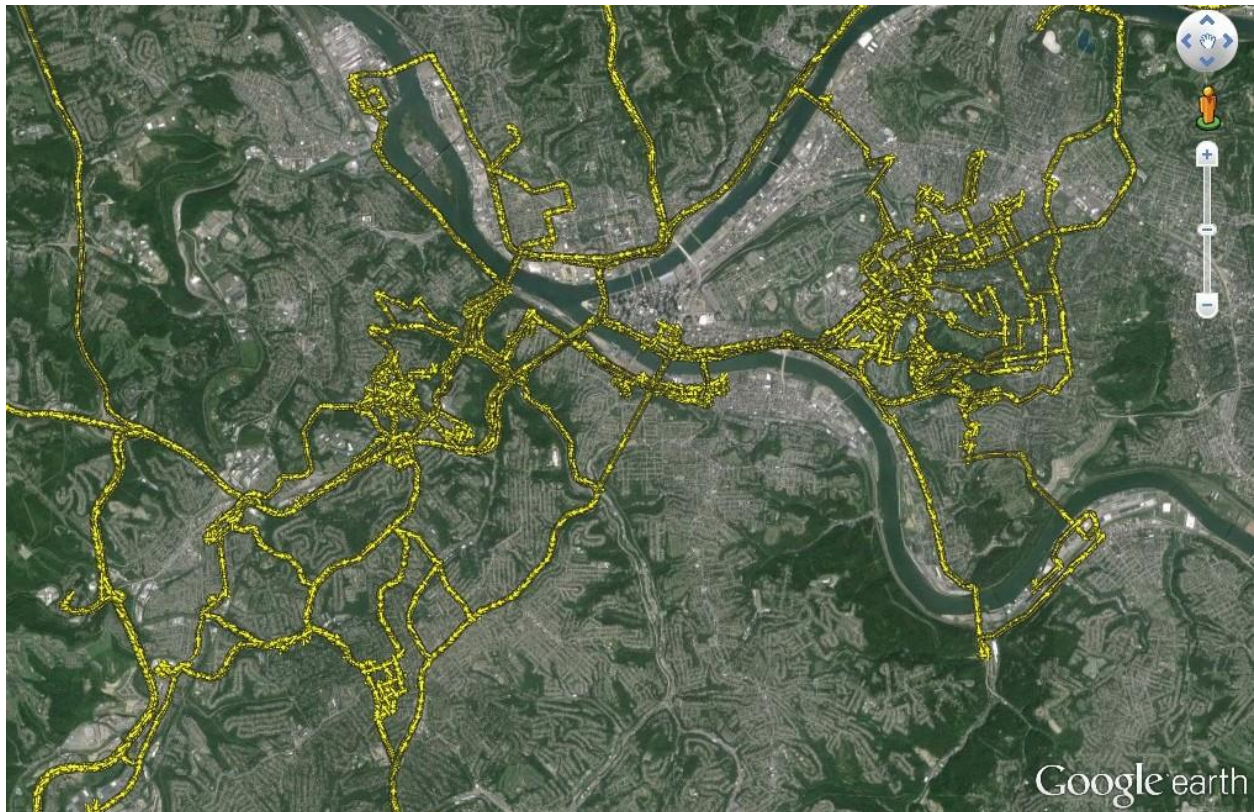

Figure 13 The data collected in over one year (yellow) overlaid on a satellite view of Pittsburgh, PA.

This is about 1200 hours of driving on 500 miles of unique roads. The distress score for a subset of the data is shown in Figure 14 overlaid on the GIS street center lines used by the city of Pittsburgh.
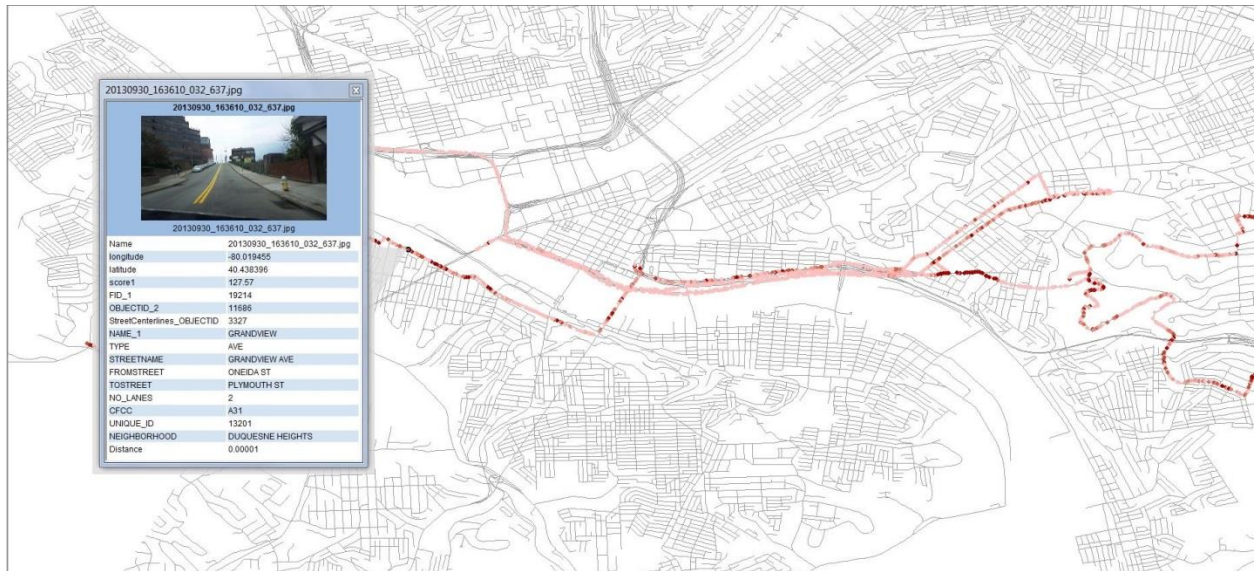
Figure 14 Distress score overlaid on a GSI street map of Pittsburgh. Light red is little or no cracks and dark red is lots of cracks. A click on a point brings up the corresponding image and other information about the road.

We compared the performance of our system with that of human inspection. We had several people[12] label the cracked regions of images and we compared their performance to each other and to our system. The overall performance is depicted in Figure 15.
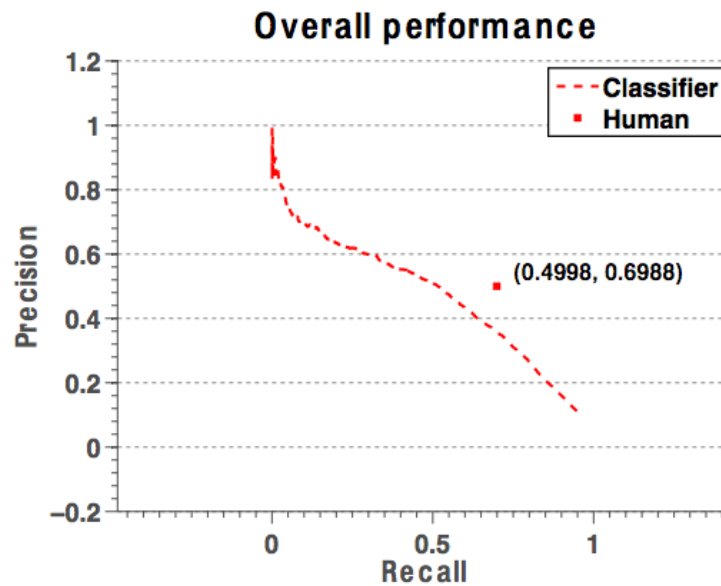

Figure 15 Crack detection performances of humans and our classifier. Higher recall or precision means better performance.

With such a comparison it is not possible for a classifier to be better than the human performance. The goal is to come close to the human performance. As can be seen in Figure 15, we achieved this goal.

---

[12] For this task we used Amazon Mechanical Turk.

Conclusion and Outlook

We have shown that it is possible to build a road distress monitoring system that can evaluate the surface quality of roads on a regular basis, close to human performance and at very low cost . The equipment can be purchased for less than $1000, the labor required is small, and the analysis can be done on a standard computer.

Having completed the development of a prototype system, we are in the process of testing it in a pilot project with the City of Pittsburgh. We will mount the data collection system on their vehicles, analyze the collected data and load the results into their asset management system. First indications are that the system will be very useful for them.

Further into the future we want to expand it into an all-purpose infrastructure inventory and monitoring system. We want it to be able to detect other things like traffic signs, lane markers, vegetation overgrowth or manholes to name just a few. It can also be converted into a system that monitors snow cover or weather conditions live.