Remote Sensing of Roads and Highways in Colorado

Large-Area Road-Surface Quality Using Very-High Spatial Resolution Aerial and Satellite Data

Contract number: RITARS-12-H-CUB

Final Report, November 2014

Abstract

This project started with a very optimistic objective of being able to classify paved road surface conditions from high-resolution satellite imagery. Initially, the plan was to derive parameters from the satellite imagery that could be found to correlate with the parameters measured *in situ* by the Colorado Department of Transportation (CDOT) and various Municipal Planning Organizations (MPOs). Most of the first year was spent trying to find indices computed from the satellite data that correlated highly with *in situ* measurements of roughness, rutting, and cracking. Unfortunately, this analysis met with little success. During this effort, we learned that CDOT and the MPOs do not make their paving decisions at the granularity of the *in situ* data. Rather, they classify their roads holistically as good, fair, and poor (CDOT uses two classes of poor roads depending on how soon a road needs to be paved). With this realization, the focus of the project shifted to being able to estimate the condition of road surfaces in these broad categories.

Using a paving schedule provided by CDOT, we soon found that we were able to clearly discriminate between good and poor roads. This was both in terms of the digital brightness values of the imagery and in the texture metrics computed from the satellite imagery. We then extended this analysis to show that we could classify roads as good, fair, and the two classes of poor. Most of this analysis was carried out by human analysis of the satellite imagery. In an effort to automate the process, we began the development of an automated classification routine to be able to remove the non-asphalt surface components (paint lines, vehicles, shadows, road signs, concrete, etc.) in the satellite imagery. Thus, using the OpenStreetMap database, we can locate the roads in a satellite image. We can then use our classification software to select out those asphalt portions of the road surfaces for subsequent analyses.

We believe that this method has a lot of potential to be able to conduct large-scale censuses of road surface conditions, but this is where the project terminated. We have not had the time needed to do a large-scale test of the methods. We tried to link with companies that carry out the *in situ* surveys such as Pathway Services Inc. While they were helpful in discussing their present methods, none of them were interested in trying to use satellite data to improve their data collection. They felt that government requirements would always require someone on the ground measuring surface conditions.

Introduction

The initial approach to this project was based on the assumption that recent improvements in high spatial resolution satellite imagery provided by DigitalGlobe, Inc. should be able to detect changes in road surface conditions, thus indicating the health of these roads. The first step in determining if this was possible was to learn how the responsible agencies presently gather this information. We made contact with the Colorado Department of Transportation (CDOT) and were fortunate enough to get the active participation of Mr. Stephen Henry. He kindly provided past measurements of the International Roughness Index (IRI), rutting, and cracking for Colorado highways from 2007 to 2012. In addition, we developed a strong working relationship with Mr. Craig Casper of the Pikes Peak Area Council of Governments (PPACG), which is responsible for road management in the Colorado Springs area. Mr. Casper arranged for a new *in situ* sampling in his area and we arranged for a WorldView-2 (WV-2) image to be collected by DigitalGlobe.

Most of the first year of this two-year project was spent trying to find relationships between these *in situ* measurements and parameters derived from the satellite imagery. This was not a very rewarding experience as no correlations could be found. We developed very sophisticated modeling approaches to demonstrate these correlations but nothing conclusive was ever discovered. This was quite disturbing since we could easily see changes in the surface conditions of the roads and could not understand why correlations were not beginning to emerge.

We then started to learn more about the *in situ* data, how they are collected and their basic character. We came to realize that most of this data is human visual analysis of video images of the road surface. This makes it easy for human variability to introduce a lot of statistical noise into these measurements. We hoped this would not be the case with the special data collection carried out in Colorado Springs coincident with a WV-2 image. Detailed analysis of these data coincident with this image demonstrated that *in situ* values collected within 10 cm of each other and separated by less than 30 minutes were very different. When we contacted the contractor who performed this survey they refused to discuss it with us once they realized what we wanted to talk about. They would also no longer discuss things with Mr. Casper who paid ~$40,000 for the *in situ* survey.

Faced with this situation, we turned to see just what we could learn from the satellite images themselves. We initially started with even higher resolution aerial optical imagery and found that we could see both roughness and cracking in asphalt road surfaces. We then moved to examining the satellite imagery and indeed we could still see differences between poor and good surface conditions. All of this analysis was done by “hand” meaning that an operator would filter out the non-asphalt components of the road surface for subsequent analysis. Even with this limitation, we found that we could definitely see the differences.

To convince ourselves of this analysis, we went and took some digital photographs of various asphalt surfaces. We looked at parking lots with good (recently paved) surface

conditions versus parking lots with broken and disturbed asphalt elements. We found that we could use the same image analysis metrics with these digital photographs that we could with the satellite imagery. This helped to convince us that what we were seeing in the same type of information in the satellite imagery that we could see in the parking lot photographs.

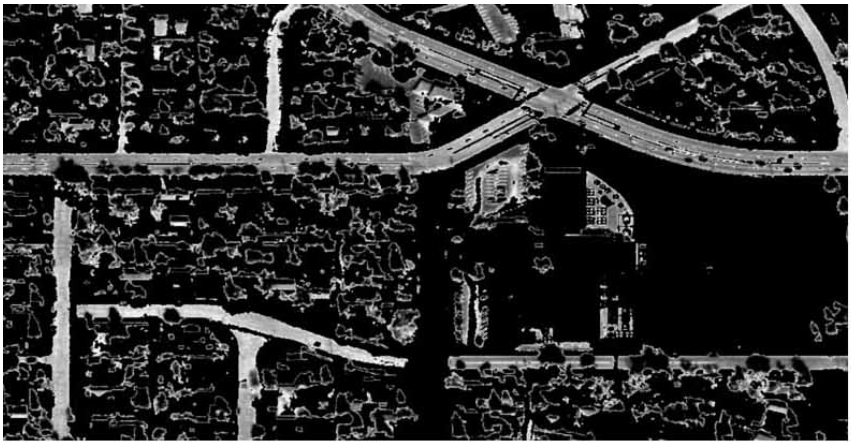
We then managed to acquire the paving schedules for CDOT. This made it possible to now analyze the satellite images in terms of the broad categories that CDOT actually used to make their paving decisions. Thus, our focus now switched to being able to estimate good, fair, and poor surface conditions to match the CDOT rating rather than trying to fit roughness, rutting, or cracking. Now we began to have considerable success. Using both just the brightness values of the imagery called digital number (DN) and a number of texture metrics, we were now able to find roads in good condition, fair condition, and both levels of poor conditions.

We reported on this analysis at the Transportation Research Board (TRB) Annual Meeting in January of this year. We were surprised that in the remote sensing session, we were the only people who talked about assessing road surface conditions. We got a number of questions from people in the audience who wanted to apply our methods to their area or apply our analytical techniques to other data types such as Synthetic Aperture Radar (SAR) data. We have not had time nor funding to follow up on any of these leads. We did spend much of our time at the TRB talking to some of the companies who collect the *in situ* data to learn more about how they perform this contract service.

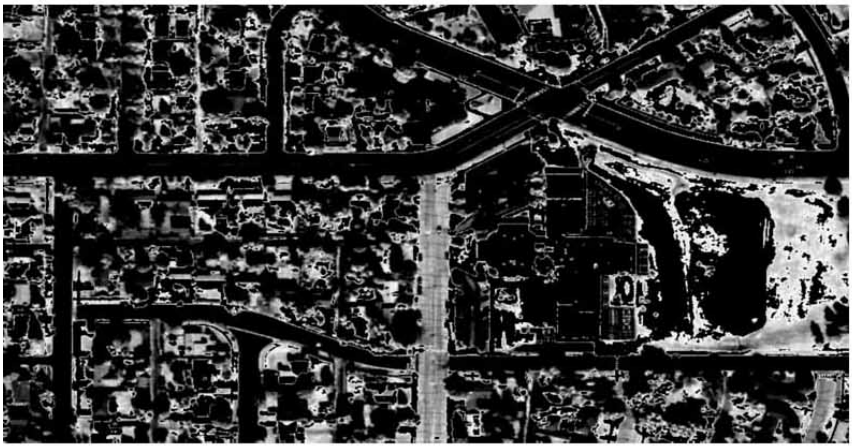
In the final phase of this project, we wanted to create software that would automatically remove the non-asphalt elements of the roads in each image. This turned out to be a real challenge, but we were able to define a software routine based on a “random forest” classification that did a pretty good job of identifying vehicles, paint lines, shadows, road signs, concrete, and other non-asphalt elements of the roads in the satellite imagery. We were in the development phase for this software when the project terminated. We believe that this approach has a lot of potential for developing a truly automated package for the selection of asphalt road surfaces and classifying them as good, fair, and two levels of poor. However, just how this type of an approach could be marketed was beyond the scope of this project.

In situ **Measurement Correlation**

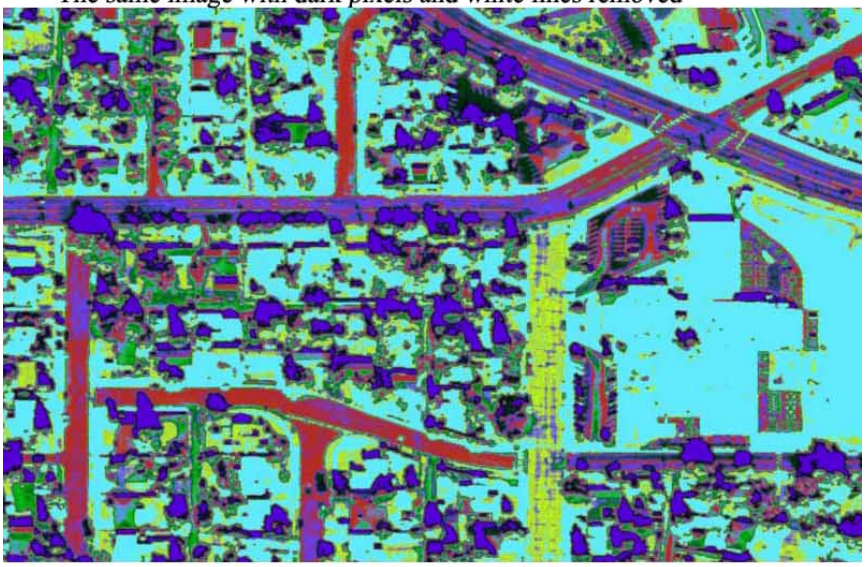
As mentioned in the introduction, the first phase of this project was to try and match indices extracted from the satellite imagery with *in situ* measurements of road surface conditions. The problems at this stage can be seen in the images in Figure 1. The top image is the unprocessed image while the middle panel is the same image with the paint lines and vehicles removed. Finally, the bottom image shows changes in road surface digital number.



Original satellite image of roads in Colorado Springs



The same image with dark pixels and white lines removed



Color image showing changes in road surface conditions

Figure 1: WorldView-2 image of roads in Colorado Springs

Unfortunately, when we extracted values from these images and compared them with the *in situ* measurements of road surface conditions the results were not very positive as seen here in Figure 2.

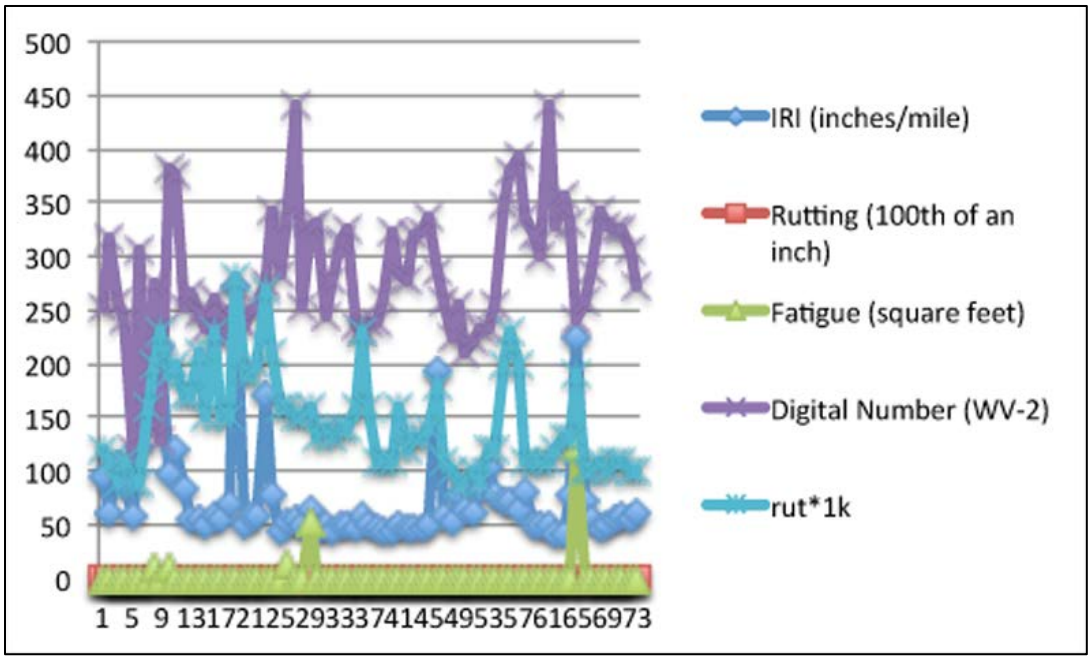


Figure 2: Comparison between satellite digital number (DN) and *in situ* measurements of roughness (IRI), rutting, and cracking (fatigue)

It is interesting here that there is much larger variation in the DN than in any of the *in situ* parameters. Still, there is very little correlation between the DN and any of the *in situ* parameters. Computed correlation values confirmed this lack of agreements between these parameters.

We continued to pursue this line of investigation thinking that this lack of correlation was due to our selection of features. We then looked at a smaller part of Colorado Springs as shown here in Figure 3. We felt that maybe if we were careful in analyzing a smaller region we would be able to make the satellite derived indices fit the *in situ* measurements. This smaller area in Colorado Springs does not contain the complexity of features that we had in the earlier images. We hoped that this would simplify the correlations. Unfortunately, this was also not successful as can be clearly seen in Figure 4, which are these parameters for the road in the center of the image in Figure 3.

Again, the variability is much greater in DN than in any of the *in situ* measured parameters. Correlations remained small and if anything were worse than in the larger area used previously. At this point, the entire image filtering for vehicles, paint lines, shadows and road signs was being done by hand leaving a potential for errors to creep into the analysis.

Figure 3: WV-2 image of a smaller region of Colorado Springs

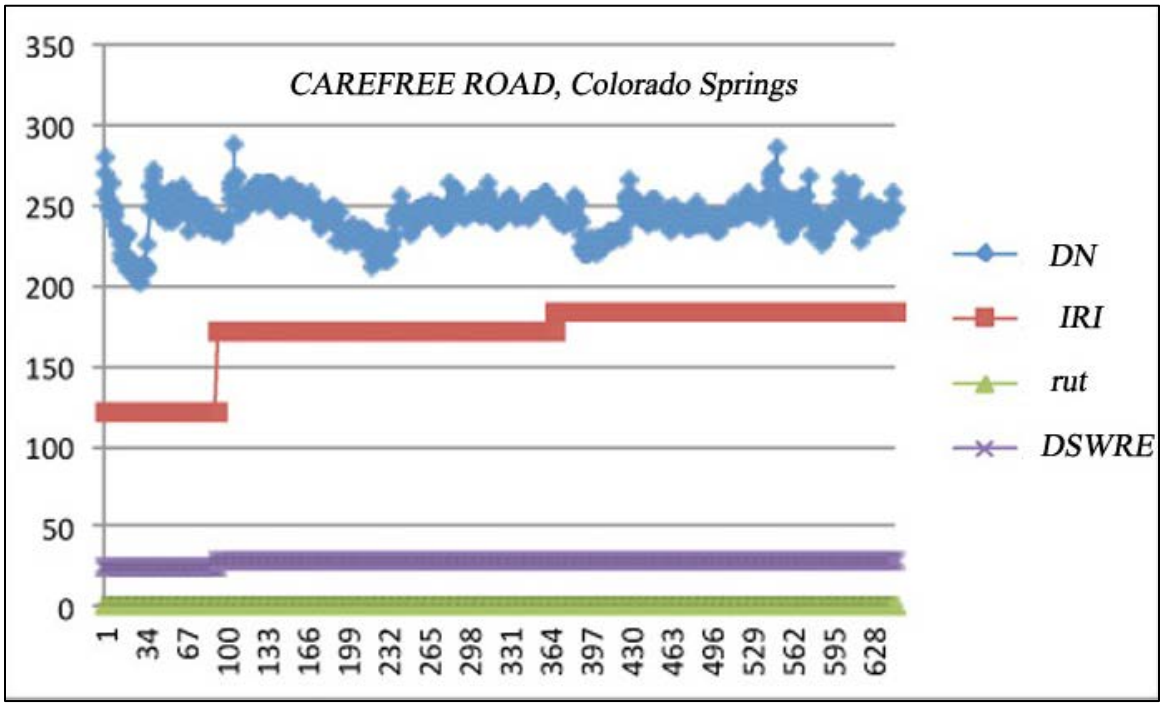


Figure 4: DN and *in situ* road surface parameters

Aerial Imagery Analysis

At this point of the project, it became pretty clear that fitting satellite values to the *in situ* measurements was not going to work. The basic question was then, could we see changes in surface conditions from satellite imagery? We therefore decided to examine even higher spatial resolution aerial imagery to learn what we could see in terms of road conditions. To guide this study, we used the fatigue (cracking) information from CDOT for US 34.

First, we looked at a road piece with a fairly low value of fatigue shown here in Figure 5. This piece had a fatigue value of 171 ft^2 in the region between the red dots.

Figure 5: Aerial imagery mile 114.6-114.7, fatigue 171 ft^2

While it isn't immediately obvious, this road appears rather uniform indicating good surface conditions consistent with low a low fatigue value.

Next, we picked an intermediate fatigue value of 414 ft^2 for the road piece in Figure 6. Here, the road surface appears a bit brighter than in Figure 5, consistent with an older road surface.

Figure 6: Aerial imagery mile 114.1-114.2, fatigue 414 ft^2

This image is representative of an intermediate "fair" road surface condition consistent with the medium fatigue value.

Finally, in Figure 7 and Figure 8, we looked at areas with fatigue values > 5,000 ft^2 as representative of poor road surface conditions.

Figure 7: Aerial imagery mile 115.0-115.1, fatigue 5760 ft^2

In Figure 7, we can clearly see the crosswise cracks in the road surface pavement that resulted in the very high fatigue value of 5760 ft^2. These cracks are clearly visible in this aerial imagery, which has a spatial resolution of 30 cm.

While the fatigue value of Figure 8 is somewhat lower at 5014 ft^2, the road surface appears even more worn than that in Figure 7. The pavement in the right half of the image is very light but still heavily cracked. On the left, the pavement is darker but also severely cracked and broken.

Figure 8: Aerial imagery Mile 115.1-115.2, fatigue 5014 ft^2

There is a good possibility that the pavement on the right is concrete rather than asphalt. Still, the surface condition of both left and right areas is quite degraded and represents a poor quality.

To demonstrate even worse poor conditions, we look at other segments of this road with even higher fatigue values. In Figure 9, we have a road segment from the 114.8-114.9 mile markers with a fatigue value of 6567 ft^2.

Figure 9: Aerial imagery mile 114.8-114.9, fatigue 6567 ft^2

Here, the cracks are both larger and more numerous than in Figure 7 and Figure 8.

To assess whether or not the WV-2 satellite data would be able to detect these road surface cracks, we resampled an aerial image to have a spatial resolution of 50 cm, the

same resolution of the panchromatic WV-2 data. To our pleasant surprise, the resampled synthetic WV-2 image also clearly showed the same crosswise cracks in the road surface (Figure 10 right) that we observed in the original aerial imagery (Figure 10 left). The definition of the cracks is slightly distorted in the synthetic WV-2 image, but their presence is still clearly noted.

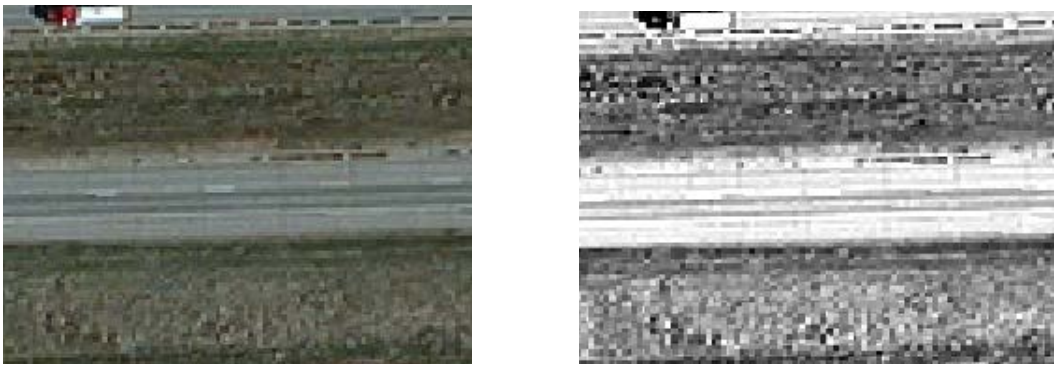


Figure 10: On the left is the original aerial image while on the right is the aerial image resampled to the 50 cm resolution of the WV-2 panchromatic imagery

As a result of these analyses, we felt that the satellite data could be used to classify road surface conditions if they were lumped into broader categories. What was not possible is to find relationships between the satellite data and the *in situ* ground information. Once we had realized this situation, the progress of the project increased greatly. We now had some firm evidence that we could estimate overall road surface conditions but not the specific *in situ* measurements of such conditions.

About this time in the project, the PPACG office sponsored a data collection in the Colorado Springs area. The location is marked here in Figure 11.

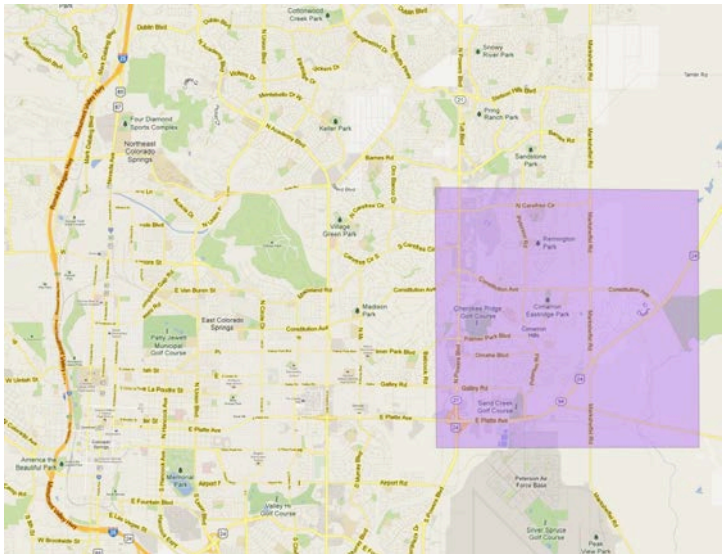


Figure 11: Ground survey area in Colorado Springs

Unfortunately, the *in situ* data collected during this survey appeared to be significantly flawed and all efforts to compare the newly collected WV-2 image with these data were inconclusive. As a consequence, this ideal opportunity was a failure in being able to establish connections between the satellite data and the *in situ* measurements.

Parking Lot Study

In an effort to evaluate the use of statistical image analysis methods, we carried out a study of known pavement conditions. This was done by taking digital camera images of two different parking lot surfaces, one of which was "good" in surface condition (uniform in appearance) and another, which was in very poor condition (varied in appearance and showing cracks, filled and broken segments). These images were collected using a digital camera and subsequently manipulated to simulate satellite imagery. Both the original digital images and the simulated satellite images were statistically analyzed to determine our ability to detect and map surface conditions. The results indicate that the homogeneity of the surface suggests the quality of the surface condition.

Pavement Photography

Photographs of two different parking lots were collected for this study. The first parking lot was located in the University of Colorado – Boulder campus south of Colorado Avenue between the Duane Physics & Astrophysics and Benson Earth Sciences buildings. The pavement of this lot had a dark and smooth appearance with no signs of wear and tear. If it were surveyed by CDOT, it would likely receive a condition rating of "good". The second parking lot was located in the East Aurora neighborhood of Boulder on the northwest corner of 29th Street and Bixby Lane directly north of the Spanish Towers building. The pavement of this lot had a speckled and rough appearance. It was also marked with numerous distresses such as cracking, potholes, and worn out sealant. If it were surveyed by CDOT, it would likely receive a condition rating of "poor". The stark contrast between these two parking lots would make for an interesting analysis.

The device that used to photograph the pavement was the digital camera built into a Samsung Galaxy S III mobile phone. The phone was held at a height of about 5 feet and the camera was pointed towards nadir, that is, facing the pavement. 25 photographs of each parking lot were collected and stored. Their original pixel resolution was 3264×2448. The digital pictures were initially stored as RGB JPG files, so they can be interpreted by image processing software to have 24 bits per pixel (8 for red, 8 for green, and 8 for blue). In order to be properly compared to panchromatic satellite imagery, they must be converted into 8-bit greyscale images. The luminosity algorithm, expressed in Equation 1, was applied to them using MATLAB.

$$L = 0.21R + 0.72G + 0.07B$$

Equation 1: Luminosity algorithm

The term L represents the 8-bit greyscale quantity that can be obtained from the 8-bit red (R), green (G), and blue (B) quantities. The RGB terms are weighted differently because

the sensitivity of the human eye varies according to color. Once the RGB pixels were converted to greyscale pixels, the images were saved as BMP files, which could be analyzed with image processing software.

It was previously believed that direct sunlight on cloudless days were necessary to illuminate the pavement properly. Photographs of these parking lots were thus initially collected on sunny days. However, preliminary analysis revealed that direct sunlight brought about uneven illumination of the pavement. The photographs were reshot on cloudy days and illumination of the pavement became more uniform. Therefore, pictures from cloudy days were used in this study instead.

Satellite Image Synthesis

The greyscale images of the parking lot pavement were manipulated to imitate imagery from DigitalGlobe's WorldView-2 spacecraft. The first step was to reduce their spatial resolution to match that of WV-2. In the panchromatic band, WV-2 has a spatial resolution of 46 cm per pixel. The scaling factor was determined by using measuring tape as a reference. A measuring tape was pulled out to 46 cm, placed on the ground, and photographed by the mobile phone camera. Recall that the photographs collected by this camera had a size of 3264×2448 pixels. In this picture, the tape had a length of approximately 500 pixels, which implies a scaling factor of 1/500. Using MATLAB, the greyscale images of the parking lot pavement were downsampled by this factor, resulting in new greyscale images with a size of 7×5 pixels.

The next step was to add artificial noise to these 7×5 images to simulate the noise characteristics of the WV-2 sensor. The WV-2 images provided by DigitalGlobe are represented by pixels carrying 11-bit digital number values. Brighter landscapes have higher DN values and darker landscapes have lower DN values. Experiments conducted by scientists at DigitalGlobe have determined that the noise variance associated with DN is linearly proportional to DN itself. This relationship is expressed in Equation 2 below.

$$\sigma^2 = a + b \cdot (\mathrm{DN})$$

Equation 2: Noise variance

The termσ^2 is the noise variance, with σ being the noise standard deviation. The terms a and b are the offset and slope coefficients respectively, which were determined experimentally through regression analysis of sensor data.

To measure the sensor's noise, the WV-2 radiometer was pointed at a cloudless scene of free space between the Moon and the Earth. The DN values in the two-dimensional image of this scene are constant in the horizontal direction but continuously varying in the vertical direction. Sample chips were collected along the horizontal direction at different vertical positions. The DN variances and DN count means of each chip were obtained and plotted against each other. The results of one such trial are shown in Figure 12 below. Through regression analysis of this data, the values for the offset a and slope b have been measured to be 2.0000 and 0.0143 respectively.

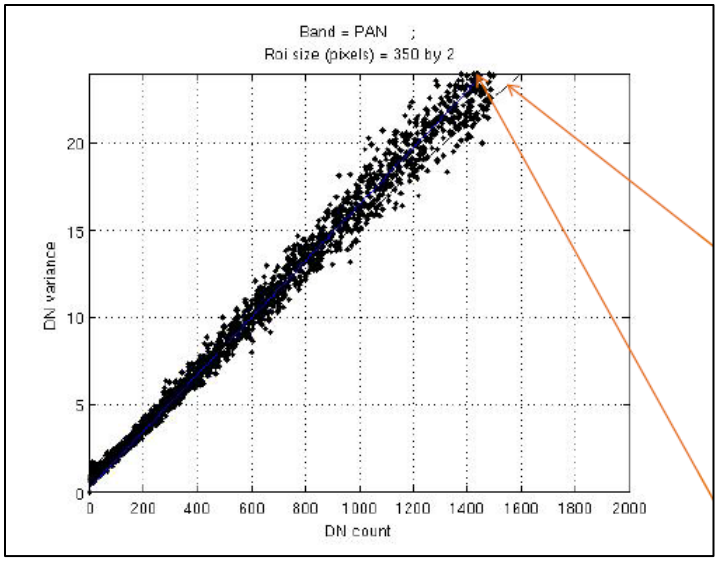


Figure 12: WV-2 DN variance vs. DN count

The 7×5 greyscale images of the parking lots were corrupted by artificial noise equivalent to that of the WV-2 sensor with MATLAB. The "clean" 8-bit greyscale values in the pixels of a given image were multiplied by 2^3 to simulate mock 11-bit DNs. The noise variances σ^2 were calculated from Equation 2 for each pixel's DN. A Gaussian random number generator was used to produce noise values at every pixel with zero as the mean and σ as the standard deviation of the noise distribution. The resulting noise values were added to the mock DNs to create "noisy" DNs. The "noisy" DNs were divided by 2^3 to simulate corrupted 8-bit greyscale values, which will henceforth also be referred to as DN values. The 8-bit greyscale images constructed from these artificially corrupted pixels finally represent the desired mock WV-2 imagery. They were saved as BMP files which could then be analyzed through image processing software.

Occurrence-Based Texture Filtering

A very powerful technique often used in the science of image analysis is the application of occurrence-based texture filters. When such a filter is applied to a given image, a small rectangular subsection of the image is initially selected. This rectangle, known as a sliding window, has dimensions of an odd number of pixels greater than or equal to three for both its length and width. The reason for odd-numbered window dimensions is to have a single pixel positioned exactly in the center of the window. All of the pixels within this window are considered to be part of a set. Statistical calculations are performed on the pixel brightness values in this set and the results are assigned to the central pixel. The window then shifts over to the next row or column of pixels and the statistical calculations are applied to the new set of pixels. This process is repeated until the entire image has been covered and a group of new images has been created.

The main occurrence-based texture filters are expressed below.

$$\text{Data Range} = \max I(i,j) - \min I(i,j)$$

$$\text{Mean} = \mu = \sum_{i,j} I(i,j) \cdot p(i,j)$$

$$\text{Variance} = \sigma^2 = E[(I(i,j) - \mu)^2]$$

$$\text{Entropy} = -\sum_{i,j} p(i,j) \ln p(i,j)$$

$$\text{Skewness} = E[((I(i,j) - \mu)/\sigma)^3]$$

Equation 3: Occurrence-based texture filters

The terms i and j represent the horizontal and vertical indices of the pixels contained within a sliding window. The term $I(i,j)$ is the brightness value of the pixel at (i,j). The term $p(i,j)$ is the probability mass function of the corresponding $I(i,j)$ value. The operators max and min respectively return the maximum and minimum values of a set. The operator E is the expectation value.

These texture calculations are applied to the set of pixel brightness values contained within the sliding window. The data range represents how far apart the lowest and highest pixel brightness values in the window are. The mean is the average of the brightness values. The variance describes how far the values in the window lie from the mean. The entropy is a statistical measurement of randomness within the set. Finally, the skewness describes the asymmetry of the spread. The results are assigned to the pixel located at the center of the window. The final product is a new set of images with its pixels containing the texture calculations. The image processing software ENVI features a tool, which quickly applies these occurrence-based texture filters to a given image. In this particular investigation, the data range, mean, variance, and entropy filters were applied to the greyscale photographs of the parking lots and their corresponding synthetic WV-2 images. Unfortunately, ENVI's algorithm to calculate skewness has been verified to be incorrect, so that texture filter shall be disregarded.

<u>Mobile Phone Camera Photograph Analysis</u>

While performing preliminary analysis of the original greyscale photographs of the parking lots taken by the mobile phone camera, the high pixel resolution of 3264×2448 posed a problem. Whenever the texture filters were applied to them, the ENVI software crashed before the computations could be completed. The dimensions of these pictures were thus downsampled in MATLAB by a factor of 1/5 each, resulting in images with a pixel resolution of 653×490. These images could be processed more easily by ENVI while at the same time still able to retain a respectable amount of the original information

regarding pavement quality.

These downsampled greyscale mobile phone images of the parking lot were examined first. Examples of these images are shown in Figure 13 below. Of course, they have been scaled down to fit within the page margins of this document. Notice the relatively smooth appearance of the good pavement (Figure 13 left) compared to the speckled and distressed appearance of the poor pavement (Figure 13 right).

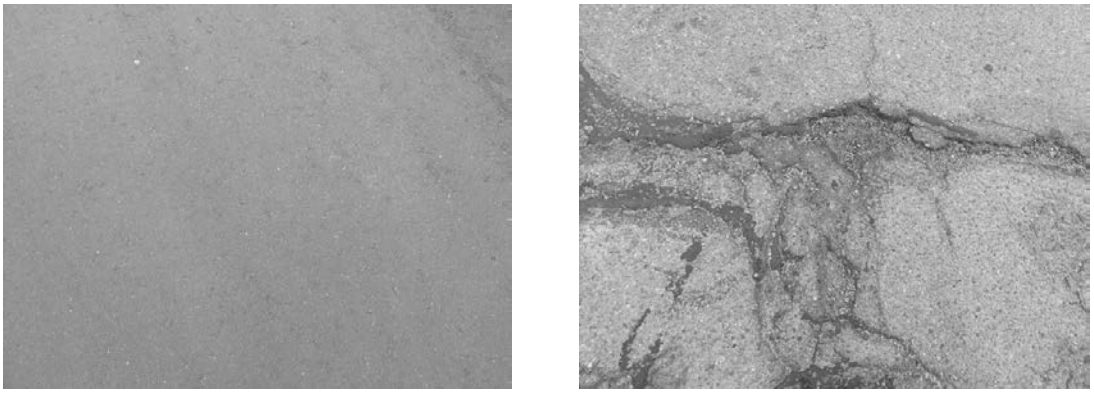


Figure 13: Digital photographs of good pavement (left) and poor pavement (right)

The 8-bit greyscale brightness values of these photographs' pixels can be interpreted to be 8-bit DN values. These DN values were extracted and plotted into a histogram. The histogram that shows these distributions is shown in Figure 14.

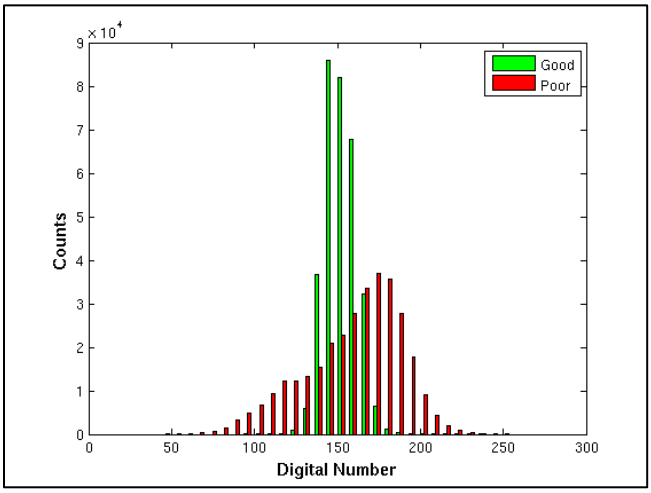


Figure 14: Pavement image DN histogram

The distribution of DNs in the good pavement image has a mean value of 152.3 with a standard deviation of 9.8. The distribution of DNs in the poor pavement has a mean value of 159.7 with a standard deviation of 29.0. Preliminary analysis concluded that nearly identical results were obtained for almost all of the photographs that were collected during the surveys for both the good and poor pavement conditions. So in the interest of conciseness, only the results for the images in Figure 13 will be presented and discussed in this report.

The occurrence-based texture filters were applied to these images using ENVI. A sliding window with dimensions of 3×3 pixels was used. The calculations were extracted from the resulting filtered images and plotted into histograms. These histograms are shown in Figure 15 to Figure 18.

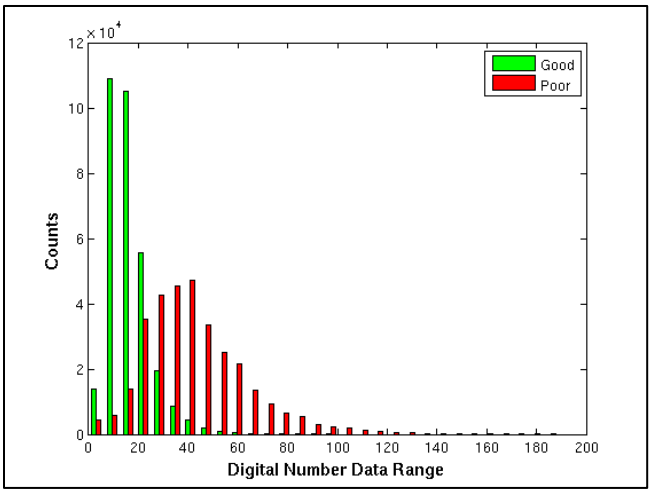


Figure 15: Pavement image data range histogram

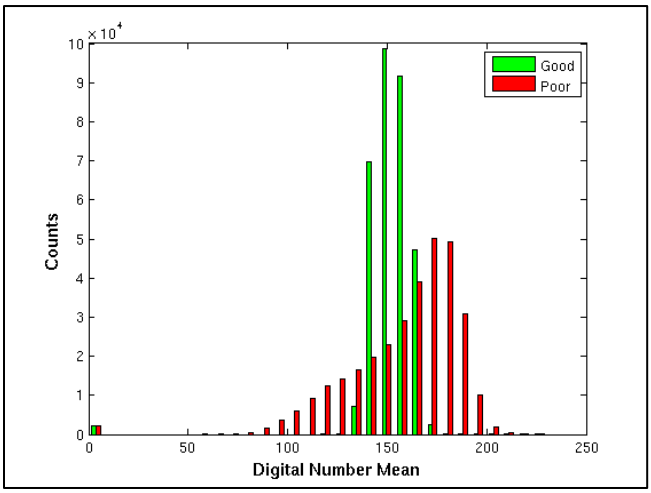


Figure 16: Pavement image mean histogram

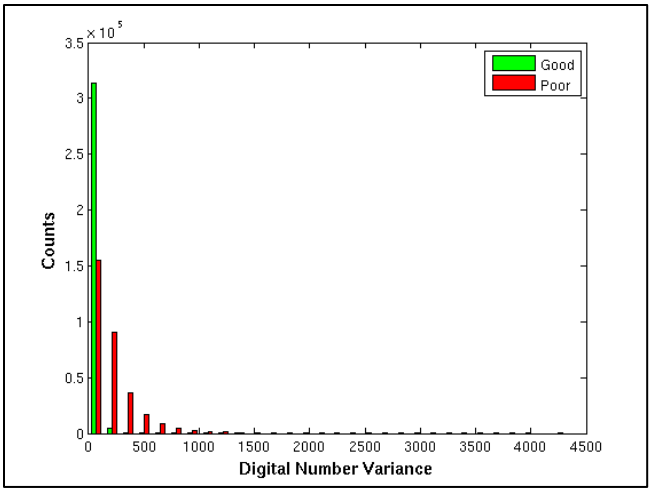


Figure 17: Pavement image variance histogram

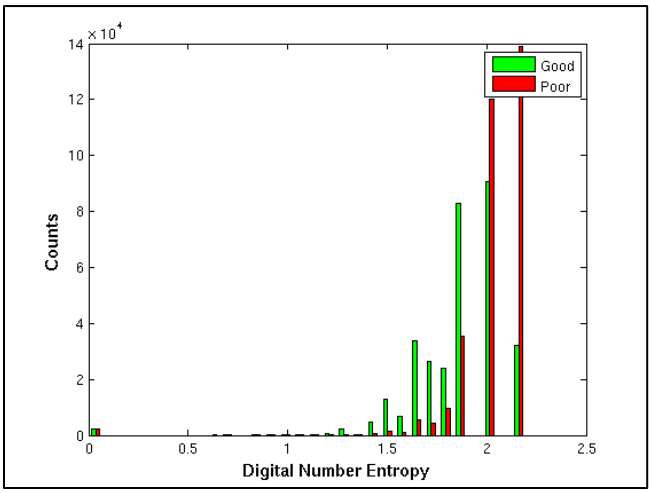


Figure 18: Pavement image entropy histogram

The means and standard deviations of the distributions in these histograms are shown in Table 1.

Texture Filter	Good Mean	Good STD	Poor Mean	Poor STD
Data Range	16.3	8.5	42.8	20.8
Mean	151.3	15.1	158.6	28.2
Variance	30.6	42.3	222.3	241.0
Entropy	1.9	0.3	2.1	0.2

Table 1: Filtered image histogram properties

Synthetic Satellite Image Analysis

A type of analysis similar to the one described above was performed on the set of synthetic WorldView-2 images created from the mobile phone photographs. The

simulated satellite images that correspond to the good and poor pavement photographs can be seen in Figure 19 below.

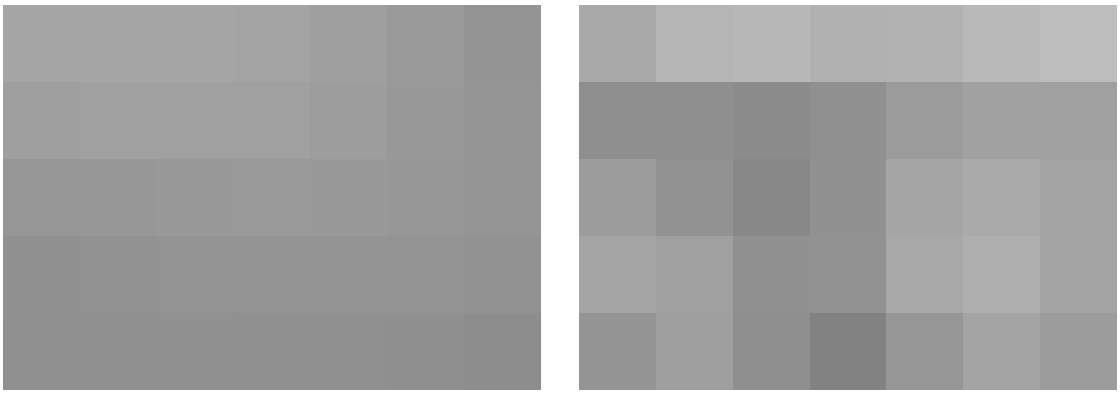


Figure 19: Pavement images (synthetic WV-2) for good (left) and poor (right) conditions

Recall that the pixel resolution of these images is only 7×5. They have been scaled up in this document so that they can be superficially compared to the images in Figure 13. Although much of the obvious information has been lost, there is still a slight difference between the good pavement (Figure 19 left) and the poor pavement (Figure 19 right) that can be recognized visually.

Each of the pictures in the set of synthetic WV-2 images has only 35 pixels. When occurrence-based texture filters with a sliding window of 3×3 pixels are applied to them, the resulting filtered images only have 15 pixels (the edge and corner pixels have no values assigned to them, so they have been trimmed out and only 5×3 pixel images are left). For a single image, these sample sizes are too small to build histograms analogous to the ones above. Fortunately, there are 24 more images in the set for each pavement condition. The DN and texture data from each of these images have thus been compiled to form two larger datasets, one for the good pavement and one for the poor pavement. Comparable histograms can be built from these datasets. The DN histogram is shown in Figure 20 below.

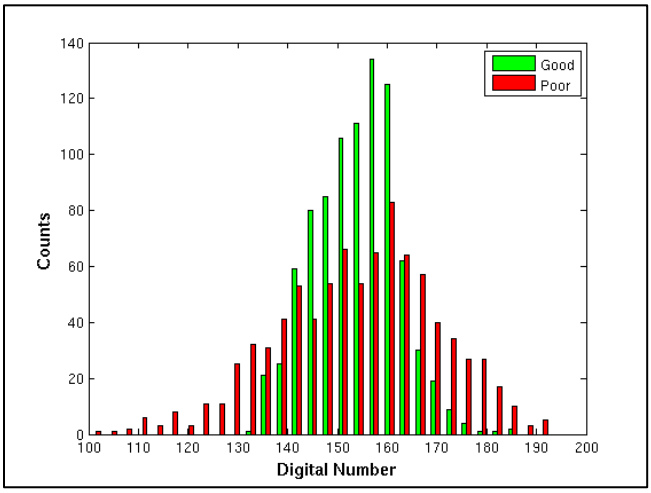


Figure 20: Pavement image DN histogram (synthetic WV-2)

The distribution of DNs in the good pavement images has a mean value of 153.9 with a standard deviation of 8.5. The distribution of DNs in the poor pavement images has a mean value of 154.2 with a standard deviation of 16.1.

The occurrence-based texture filter histograms for the simulated satellite images of the parking lot pavement are shown in Figure 21 to Figure 24 below.

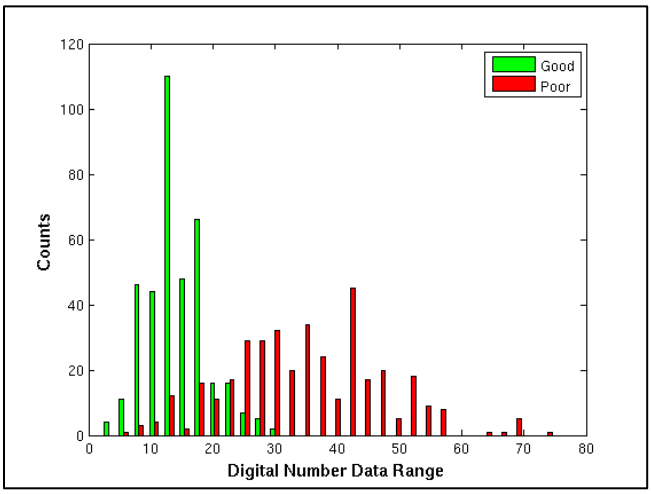


Figure 21: Pavement image data range histogram (synthetic WV-2)

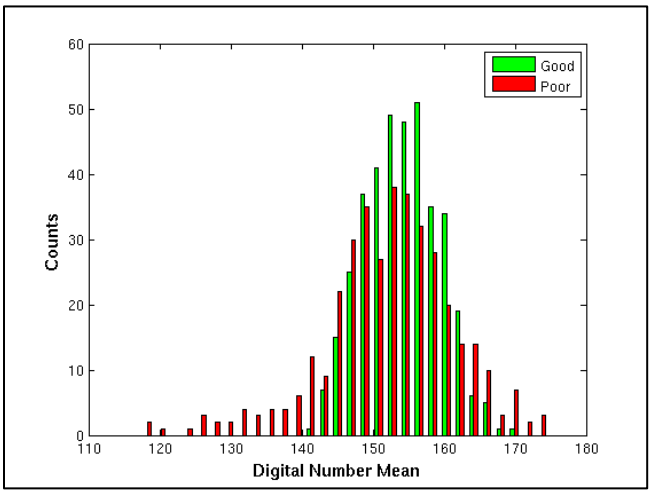


Figure 22: Pavement image mean histogram (synthetic WV-2)

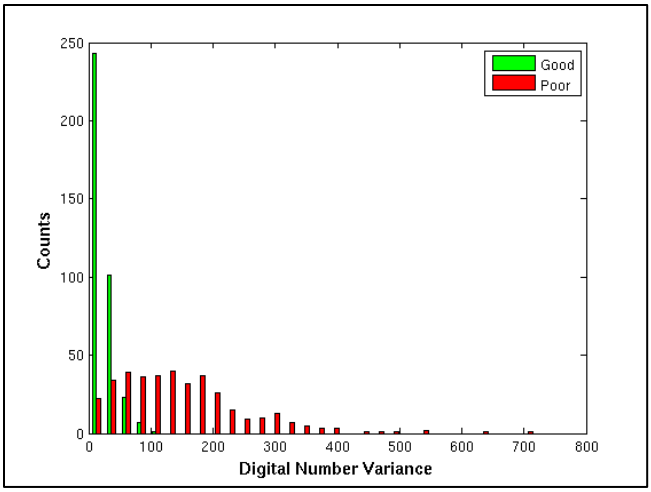


Figure 23: Pavement image variance histogram (synthetic WV-2)

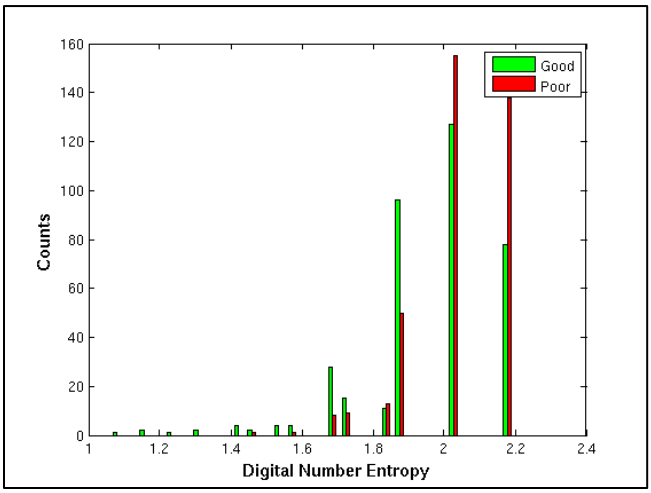


Figure 24: Pavement image entropy histogram (synthetic WV-2)

The means and standard deviations of the distributions in these histograms are shown in Table 2.

Texture Filter	Good Mean	Good STD	Poor Mean	Poor STD
Data Range	14.2	4.9	34.9	12.3
Mean	154.1	5.3	152.2	9.3
Variance	23.6	16.0	148.8	104.4
Entropy	2.0	0.2	2.1	0.1

Table 2: Filtered image histogram properties (synthetic WV-2)

Discussion and Conclusion for the Parking Lot Study

From the histograms shown above, there are notable differences in the statistical characteristics of the good and the poor pavement images. First, the statistics regarding the mobile phone photographs of the pavement shall be discussed. The distribution of the DNs is much narrower for the good pavement than for the poor pavement. This is

expected because just from cursory inspection of Figure 13, one can see that the pixels in the good pavement image seem to have more or less the same brightness values whereas the pixels in the poor pavement image have more varied brightness values. It is also worth noting that despite the presence of large dark cracks in the poor pavement image, the mean DN is actually higher than that of the good pavement image and the distribution is skewed towards brighter values. This could potentially imply that asphalt also becomes lighter as it degrades.

The statistics of the occurrence-based texture filtered images also have remarkable characteristics. The distribution of the DN data range has a higher mean and wider distribution for the poor pavement than for the good pavement. Because the poor pavement appears speckled, it is not surprising that there would be some differences in brightness between adjacent pixels. The DN mean distributions appear similar to the DN distributions. The sliding window size of only 3×3 pixels is rather small, so only the brightness values of pixels very close together would get averaged. Both of the DN variance distributions are skewed towards the left. However, the good pavement has nearly all of its values in the lowest bin whereas the poor pavement has significant presence in the higher bins. This again implies that the brightness of adjacent pixels is widely distributed for the poor pavement. Finally, both of the DN entropy distributions are skewed towards the right, but the good pavement has more presence in the lower bins than the poor pavement. Entropy is lesser when more of the pixels contained within the sliding window during the texture calculations have the same brightness values. From this result, there are more instances of adjacent pixels having the same brightness in the good pavement image than in the poor pavement image.

Next, the statistics of the simulated satellite images will be discussed. The DN distribution for the poor pavement is still wider than that of the good pavement. But strangely, it appears that the distributions' means have come closer together and that the skewness of the poor pavement DN distribution has been straightened out. The DN data range distribution of the poor pavement is still wider than that of the good pavement and it still has a higher mean. However, the means and standard deviations for both pavement types seem to have decreased. The DN mean distributions also have means that have come closer together. In fact, the mean of the DN mean distribution is actually now greater for the good pavement than the poor. But still, the DN mean distribution for the poor pavement is wider than the good, which is consistent from before. The DN variance distribution is wider and has a higher mean for the poor pavement, but the means and standard deviations themselves have decreased for both pavement types. Finally, the DN entropy distribution for the poor pavement still skewed more towards the right than the good pavement. The only notable difference is that the standard deviations became slightly smaller for both pavement types.

The overall trends seem to be that the means and standard deviations for all distributions is lesser in the synthetic satellite images than in the original mobile phone images for both pavement types. The only exception is that the mean of the DN mean distribution increased slightly in the good pavement. The likely reason for these differences is that the statistics for the mobile phone images were only computed from one of the 25

photographs collected from each parking lot. On the other hand, the statistics for the synthetic satellite images were computed from all 25 photographs. Perhaps those two images in Figure 13 weren't as representative of the entire sets as the preliminary analysis may have suggested. A more apt procedure would have been to combine the data from all 25 mobile phone images and make histograms of the combined dataset. Then the statistical properties of the distributions of the two image types could have been more similar.

But without making direct comparisons between the two image types and just considering them on their own, the results appear to be quite promising. In both the mobile phone photographs and the synthetic satellite pictures, there are clear differences between the good and poor pavement in the way that the statistical characteristics of their images are distributed. All of the good pavement statistics imply that that the pixels are more homogeneous. Conversely, all of the poor pavement statistics imply that the pixels are more heterogeneous. This follows conventional intuition because that is how the pavement surfaces appear from visual examination.

The ultimate goal of this investigation was to demonstrate that pavement surface quality could be assessed from images of the pavement itself, even at the spatial resolution available through satellite remote sensing techniques. Although it may seem obvious when approached qualitatively, the results of this study serve as evidence that the concept can be implemented quantitatively. There are clear differences in the way that the statistical information of the good and poor pavement photographs is distributed. These statistics verify that good pavement appears uniform when compared to poor pavement, which appears more varied. These differences are not only present at the comparatively high-resolution mobile phone photographs, but also at the lower resolution and noisier synthetic satellite images. If this technique were to be applied to actual satellite imagery, then it is quite possible that similar results would ensue.

If this investigation were to hypothetically continue, there is some room for improvement. One suggestion would be to survey a wider variety of pavement conditions. For example, instead of just examining extremely good and extremely poor pavement, more intermediate quality pavement could be examined. If the statistics of this fair quality lie in between that of the good and poor, then that would strengthen the results of this study even further. Also, some better photograph capturing techniques could be used. Although the mobile phone camera has been quite reliable, it was difficult to keep steadily pointed towards nadir at a constant height. And since it was only held a few feet above the ground, its field of view was not very wide. In order to match the resolution of the WorldView-2 satellite sensor, the pictures needed to be scaled down by a large factor, which caused a drastic reduction in pixel samples. Perhaps an apparatus could be built to hold the camera steady at a higher height in future surveys to mitigate these issues. Nonetheless, this investigation has been fruitful overall and the results will undoubtedly aid the CCAR Imaging Laboratory as they continue to assess the feasibility of using remote sensing techniques to study pavement surface quality.

Satellite Imagery Analysis

All of these investigations taught us a number of things:

- Any effort to match satellite image derived indices with *in situ* measurements of roughness, rutting, or cracking are going to be unsuccessful. This is due to many things including the noise and/or error in the *in situ* data, the spatial mismatch between the *in situ* samples and the satellite imagery, and the difficulty of removing erroneous pixels from the satellite images (paint lines, vehicles, shadows, road signs, etc.).

- Optical imagery is capable of detecting and quantifying road surface conditions on a measurement scale consistent with that used by CDOT and PPACG.

- Statistical methods can be defined to reveal the condition of road surfaces and, using historical data, can be related to the need for repaving of a road segment.

To demonstrate these lessons learned, we developed the material presented at the TRB Annual Meeting in January 2014. This will be reviewed here.

Satellite Data

The following studies will use the panchromatic (black and white) bands of the QuickBird, WorldView-1, and WorldView-2 satellites. We use these data to both depict conditions and to compute statistical co-occurrence texture parameters that reveal the details contained in these satellite images.

Case 1: Roads in Larimer County

The first case we examined consists of two north-south highways in Larimer County to the north of Boulder, CO. One is highway US 287 that travels from Boulder, CO north through Longmont, CO and on up to Fort Collins, CO. The segment of the highway analyzed was last paved in 1973 and was last maintained in 1998. The condition of the road is considered “poor-0”, which is the worst rating that a highway can get from CDOT. We used this road because it was contained in the same WorldView-2 image from August 8, 2012, which also contained the segment of Interstate 25 (I-25) that we analyzed. This interstate highway segment was last paved in 2011 and the condition is considered “good” by CDOT. We present images of both highway segments here in **Figure** 25 with US 287 on the left and I-25 on the right.

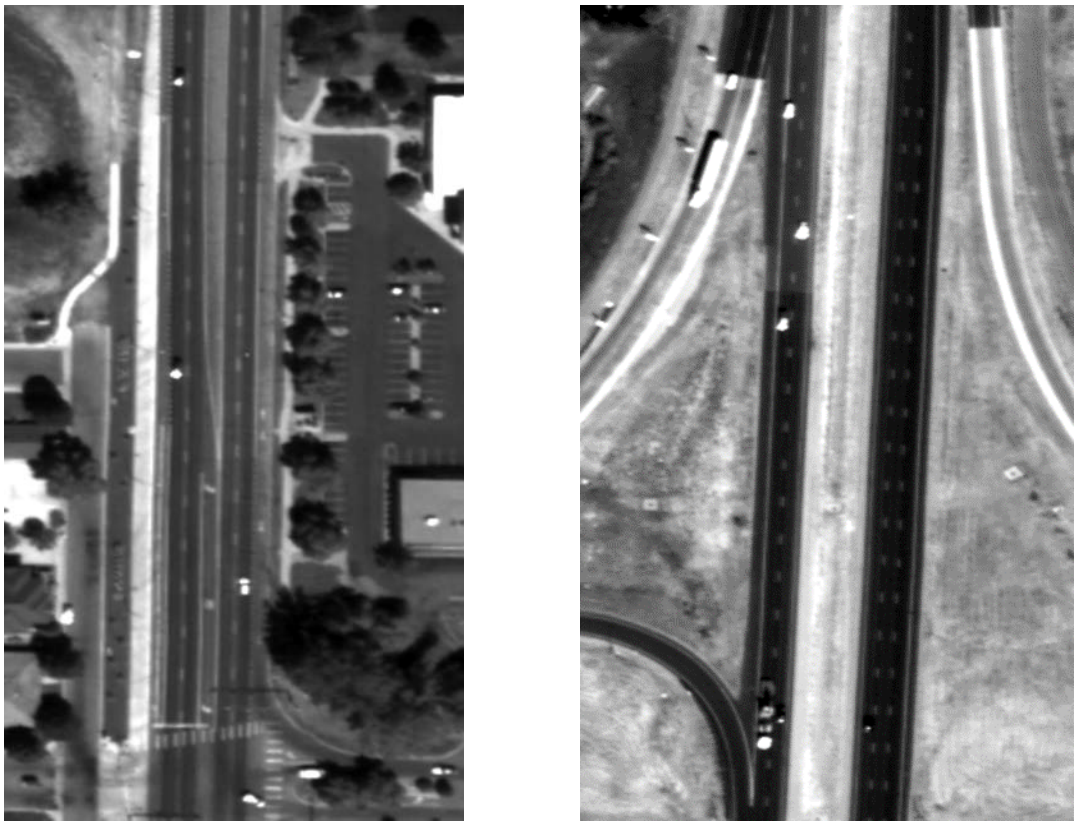


Figure 25: US 287 (left) and I-25 (right) from a WorldView-2 image collect on August 8, 2012

In these coincident images, we can clearly see what we know to be better surface conditions of I-25 as indicated by the I-25 (paved in 2011) corresponding to the darker color of the road surface. Looking at the right northbound lanes, we can really sense the degraded condition of US 287 in this area by the apparent lighter colors of the road. Still, the highway condition appears to be better than the condition of the frontage road that parallels US 287 at this location. To quantitatively assess the condition of these surfaces, we selected regions of interest (ROIs) indicated by the red and green colors on the roads in question in **Figure** 26.

ß

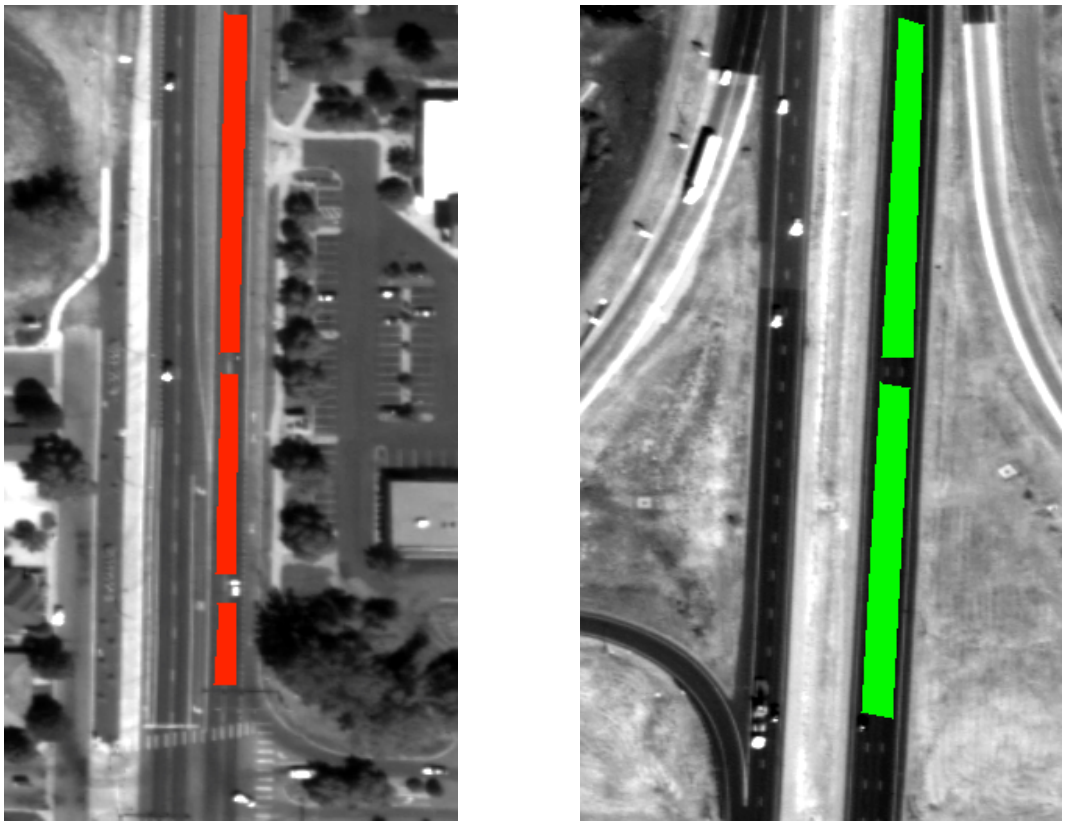


Figure 26: Regions of interest (ROIs) for the statistical calculations for US 287 (left) and I-25 (right)

An effort was made to keep the number of values used in the statistical computations somewhat similar. The first statistic examined was a simple mean computed over a 2-D pixel box passed over each image (**Figure** 27).

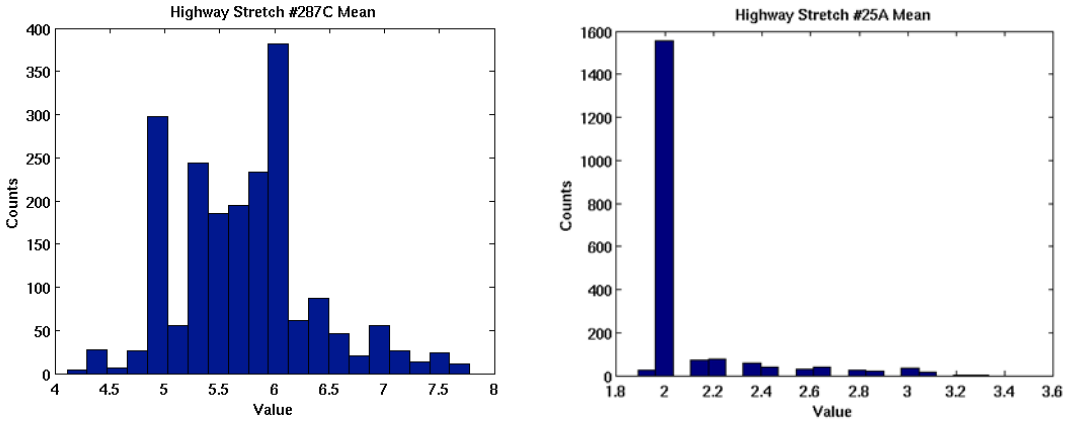


Figure 27: Histograms of the mean in a 2-D pixel moving region for US 287 (left) and I-25 (right)

In these histograms, the recently repaved I-25 has a single peak at a relatively low value, which indicates a uniform dark reflectance of the highway surface. US 287, on the other hand, exhibits a wide range of values with the maximum number of points at a value of 6, which is below the maximum at about 7.8. This indicates a serious degradation of the

highway surface leading to a wide range of increased reflectances and a lighter gray color of the highway.

Next, we looked at the variance computed in the same way as the mean. It also clearly showed the difference in these two road surfaces (Figure 28).

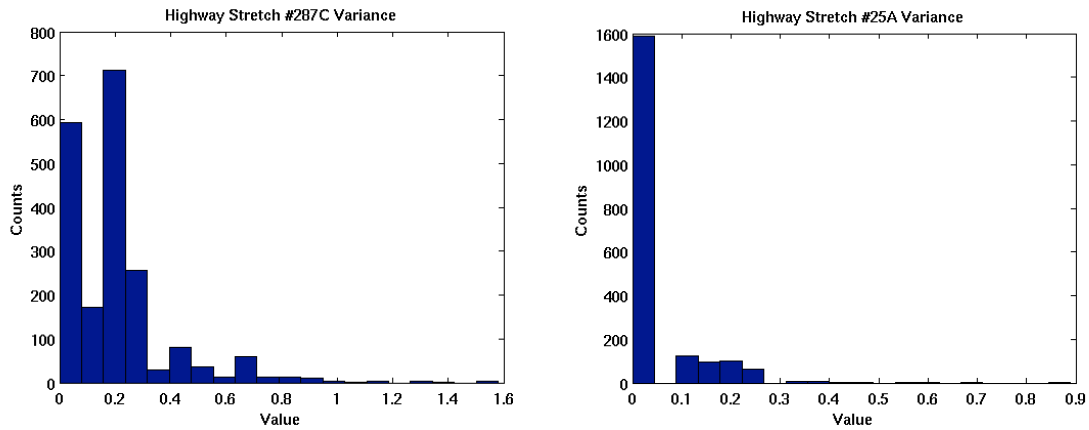


Figure 28: Histograms of the variance in a 2-D pixel moving region for US 287 (left) and I-25 (right)

Again, there is a marked difference in the variance of these two roads with I-25 having a very low and uniform variance as expected while the US 287 variance had larger values with a peak at about 0.25. An analysis of any highway would clearly show these differences indicating the quality of the road surface.

This effect is even more marked in a comparison of the contrast texture metric computed for this same moving area in **Figure** 29.

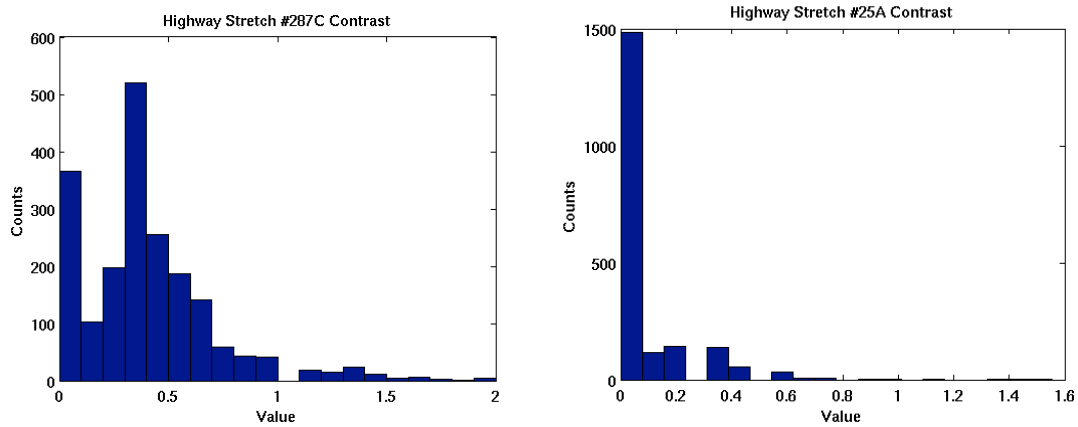


Figure 29: Histograms of the contrast in a 2-D pixel moving region for US 287 (left) and I-25 (right)

Here (**Figure** 29), the spread for the US 287 contrast is even greater than it was in the variance, which is very different than the nearly uniform near zero values for I-25. There are a few values away from zero, but these are likely due to the presence of road markers, which were not specifically removed for this analysis. Still, the very wide spread for US 287 can't be all explained by the presence of lane markers.

Moving on to homogeneity, we find (Figure 30) that I-25 is extremely homogeneous (values mostly at 1) while US 287 has its peak at about 0.85.

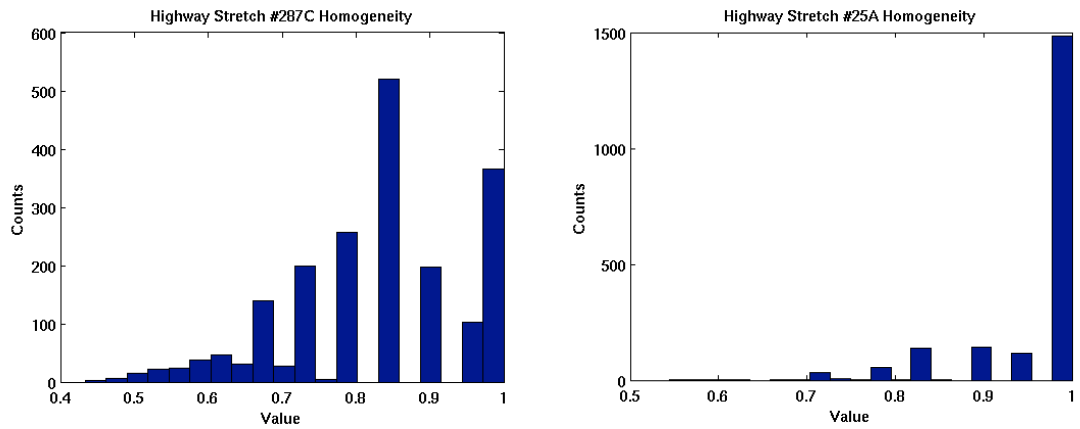


Figure 30: Histograms of the homogeneity in a 2-D pixel moving region for US 287 (left) and I-25 (right)

There is an interesting almost Gaussian series of histogram values surrounding this peak value. Again, the values away from 1 for I-25 are likely due to the presence of lane markers.

Finally, we examine the entropy calculated from these two road segments in the satellite images (**Figure** 31).

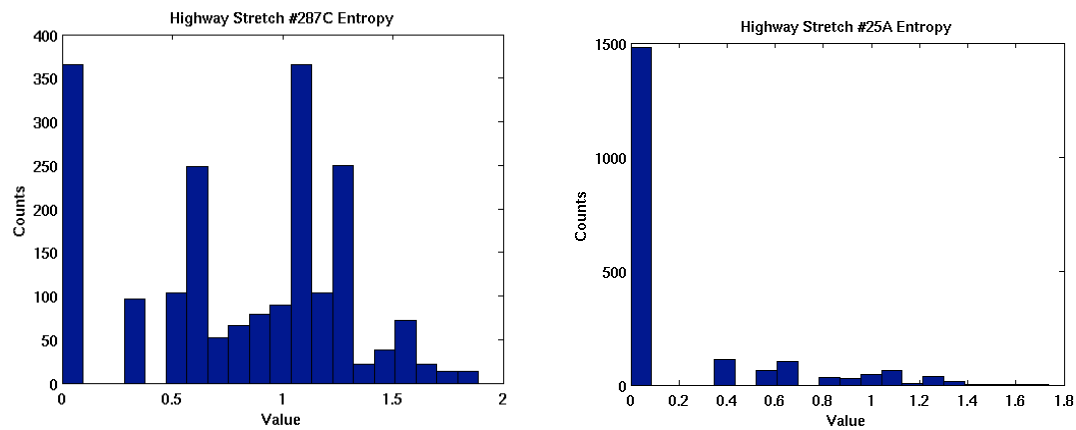


Figure 31: Histograms of the entropy in a 2-D pixel moving region for US 287 (left) and I-25 (right)

As a measure of disorder, this parameter is particularly telling of the condition of the road surface. The left side of **Figure** 31 shows a large spread again with a maximum of about 1.2 in spite of values at zero. This variety of entropy values suggests that the surface is not at all uniform and that a great variety of conditions are present. However, the I-25 values are mostly at zero with again some off-zero values due to the presence of lane markers in the analysis.

Discussion of Case 1

No matter what co-occurrence texture parameter we examine, there is a clear difference between these two roads. It is very likely with continued analysis of this type that we can develop a connection between the pavement quality and the texture parameters of the corresponding panchromatic band image. Whether we look at coherency measures (mean, homogeneity) or variability measures (variance, contrast, entropy) the results are very consistent and the indication is that this type of analysis of the satellite data would clearly provide the information needed to make a future paving decision.

Case 2: I-25 South of Colorado Springs

In an effort to verify these initial results, we looked for other areas where we found similar results. We found an area of I-25 just to the south of Colorado Springs where the highway had been paved and then stopped very abruptly. We found that simply from appearance, the comparison of these road surfaces should be able to indicate a dramatic difference (**Figure** 32). Note the large change in gray shade that takes place in the middle of this image on I-25. It should be noted that this image (collected on October 26, 2007) is from the older QuickBird satellite rather than from either of the WorldView satellites. The major change associated with this older satellite is a 60 cm resolution of the panchromatic band rather than the 50 cm panchromatic resolution of the WorldView satellites.

Figure 32: QuickBird 60 cm panchromatic image of I-25 south of Colorado Springs

The dramatic change in appearance of these road surfaces is due to a six-year difference in paving. The darker homogeneous surface was repaved in 2007 while the lighter degraded surface was repaved in 2001. This is a relatively short period to have a major degradation in surface condition. We selected portions of this roadway for our statistical analysis as shown here in **Figure** 33.

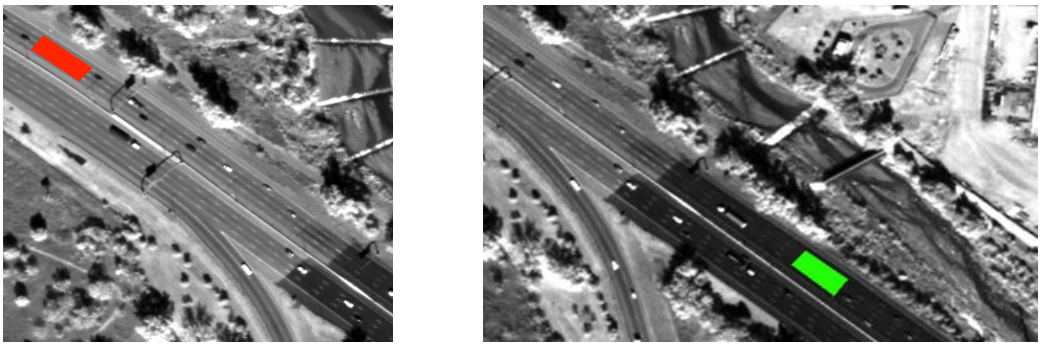


Figure 33: ROIs selected for statistical comparisons of the textures of the road surfaces in I-25

Looking at the same statistics that we looked at above, we present the mean values of the surface conditions of these two road segments in **Figure** 34.

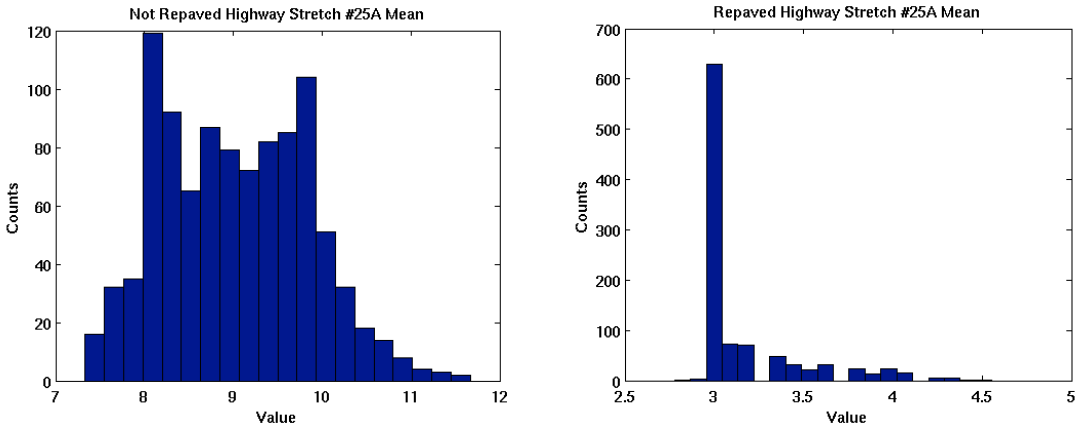


Figure 34: Histograms of reflectance mean for the 2001 paved segment of I-25 (left) and the 2007 paved segment of I-25 (right)

Here the 2001 paved surface is on the left and the 2007 paved surface is on the right. The degraded surface histogram on the left has a wide range of values with peaks at 8.1 and 9.9 while the good surface has only one peak at 3. The really surprising thing here is that the histogram of the good surface of I-25 on the right is similar to the histogram on the right of **Figure** 27, which is also from I-25 but much farther north. The main difference here is that the peak in **Figure** 34 right is at 3 while the previous peak in **Figure** 27 right is at 2. This might indicate that the better I-25 surface here has degraded slightly for the 2007 paved surface when compared with the 2011 farther north. At the same time, the histograms of the 2001 surface and the 1998 paved surface of US 287 (**Figure** 27) are very different which suggests that roads degrade differently. However, the left panels of both **Figure** 27 and **Figure** 34 exhibit large ranges in values that demonstrate the degraded highways.

Looking at the variance for the road surfaces, we again have the 2001 surface on the left and the 2007 surface on the right (Figure 35).

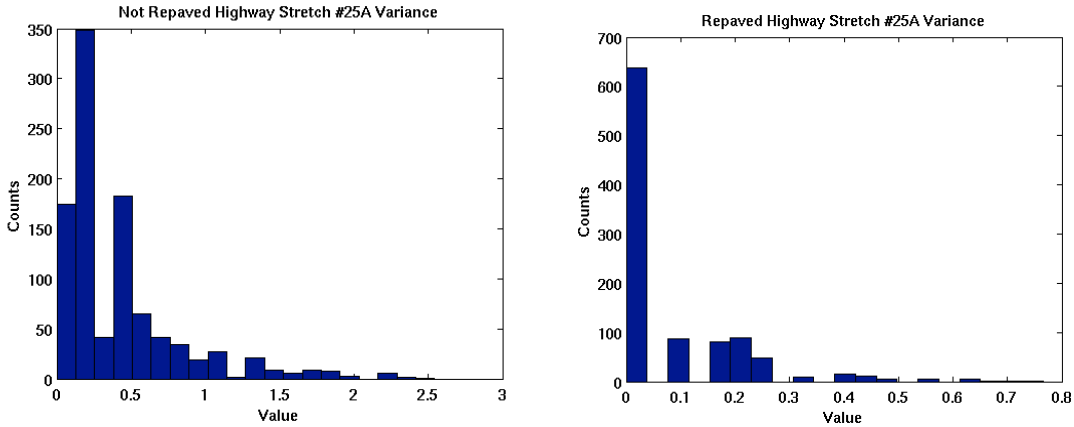


Figure 35: Histograms of reflectance variance for the 2001 paved segment of I-25 (left) and the 2007 paved segment of I-25 (right)

The better-conditioned surface of I-25 on the right of Figure 35 has a peak near zero as does the histogram of I-25 farther north on the right of **Figure** 28. The primary difference in these two histograms is that the higher values peaks in Figure 35 right are slightly larger than those for **Figure** 28 right. This is likely because the I-25 represented by Figure 35 on the right was paved in 2007, while that on the right of **Figure** 28 was paved in 2011. So this degradation appears as slightly higher contributions to the variance. The 2001 paved surface by comparison has two large peaks at 0.2 and 0.5 and a wider range of values than in the right panel of Figure 35.

The histograms of the contrast for the road segment in **Figure** 32 are presented here in Figure 36.

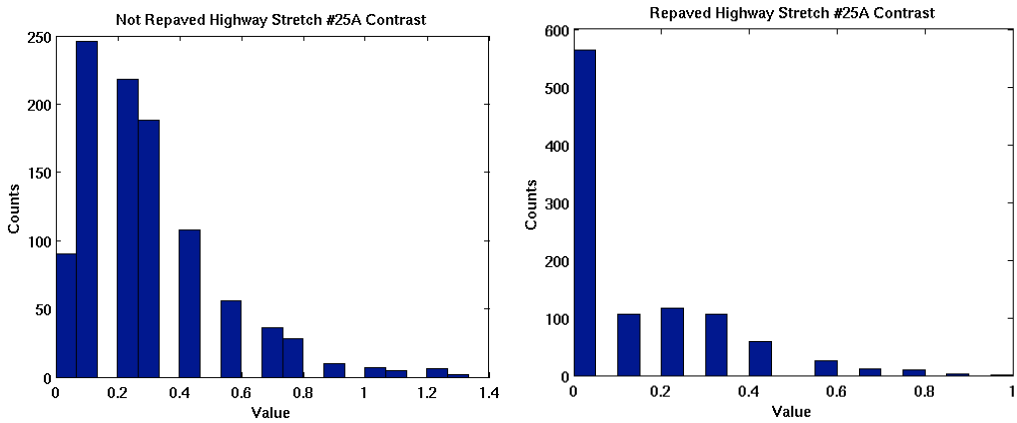


Figure 36: Histograms of reflectance contrast for the 2001 paved segment of I-25 (left) and the 2007 paved segment of I-25 (right)

Here again, we see a very large peak near zero in the more recently paved portion of I-25 (Figure 36 right) and a series of large peaks for the 2001 paved portion of the same

highway (Figure 36 left). In fact, the peak close to zero on the left is much smaller than that of four other peaks, which range up to 0.43. In addition, the histogram on the left of Figure 36 ranges all the way up to 1.35. A comparison of the right panel of Figure 36 with the right panel of **Figure** 29 shows some very interesting differences. While the peak is again near zero, there are a series of larger peaks strung across the x-axis up to about 0.6, which cannot be explained by the presence of lane markers. These histogram peaks may indicate the degradation that must have occurred in the 2011 paved surface of I-25 in **Figure** 29 versus the 2007 paved surface of the same highway in Figure 36.

There is again a very marked difference between the left and right panels with a number of higher histogram peaks in the left panel indicating the increased road surface deterioration in the segment of the road paved in 2001. More comparisons of this kind can likely develop in to an algorithm that could be used to indicate conditions leading up to a paving decision.

Turning to homogeneity, the histograms in Figure 37 again show the difference between these two surfaces. As suspected, the right panel of Figure 37 shows a very high homogeneity for the smooth appearing surface of I-25. However, this 2007 paved I-25 surface has many lower peaks when compared with the 2011 paved I-25 surface in the right panel of Figure 30. This may indicate its condition.

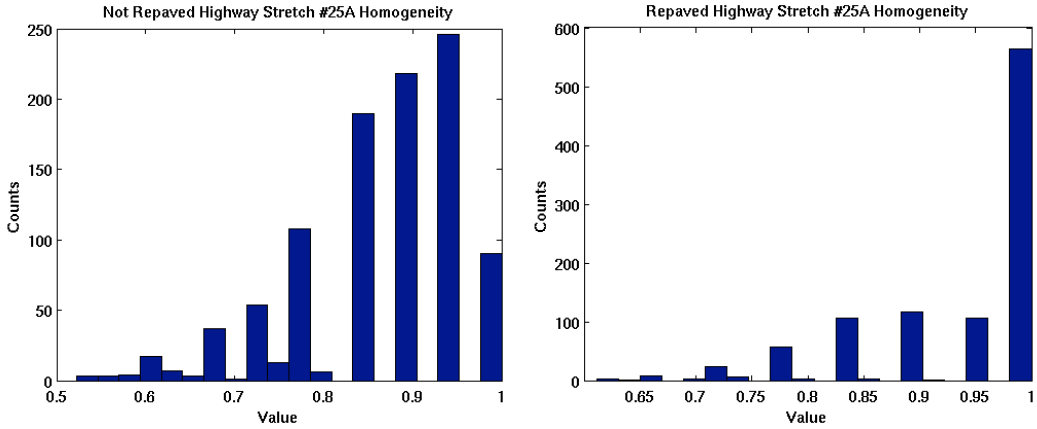


Figure 37: Histograms of reflectance homogeneity for the 2001 paved segment of I-25 (left) and the 2007 paved segment of I-25 (right)

There are significant differences again between the right and left panels of Figure 37 that are consistent with the degradation of the 2001 paved surface. A comparison of left panels in Figure 37 and Figure 30 do clearly suggest much worse surface conditions of the US 287 road consistent with its last maintenance in 1998 compared with the 2001 paving of the surface on the left of Fig. 37.

Finally we look at the statistical entropy for the road surface junction in **Figure** 32 and present them in Figure 38.

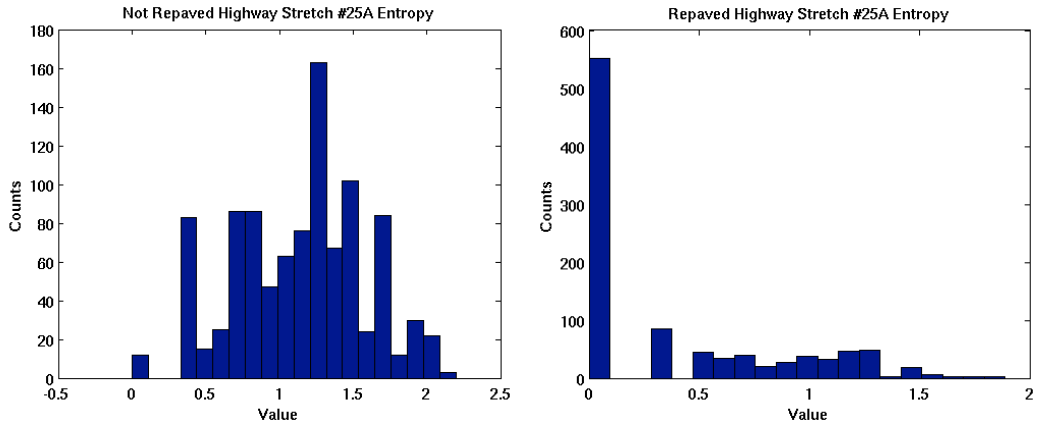


Figure 38: Histograms of reflectance entropy for the 2001 paved segment of I-25 (left) and the 2007 paved segment of I-25 (right)

Here again, the entropy shows a dramatic difference between the more recently paved portion and that paved back in 2001. Since entropy is a measure of disorder, you want the value to be near zero, which is the case for the 2007 paved surface on the right of Figure 38. A comparison with the right panel of **Figure** 31 suggests the difference between the 2007 paved surface and the 2011 paved surface. Here the differences aren't as great as they were in other statistical parameters, but there are still some higher peaks at larger values in Figure 38 when compared to **Figure** 31. It is interesting that both of them show a range of very small peaks.

The entropy histogram of the 2001 paved surface in Figure 38 has a peak at about 1.25 with other significant peaks at values stretching from 0.4 to 2. A comparison with the poor road surface conditions of the 1998 maintained US 287 in **Figure** 31 show some surprising results. The older surface of US 287 still exhibits a large peak at 0 suggesting that in some ways this older surface is in better condition than that of the 2001 paved I-25 surface, which does not have very many values near zero. Again, there is a suggestion here that these texture statistics could be used to infer conditions of the road surface relevant to whether or not they need repaving.

Discussion of Case 2

Some of the most interesting results of Case 2 were comparisons with the I-25 portion of Case 1. Here we have I-25 paved in 2001, 2007, and 2011. While the 2007 and 2011 surfaces were similar, there were differences that showed up on the texture measures. The question is whether or not these texture measures can be sorted into categories that assist CDOT in making decisions about their paving schedules for these highways.

All of the different statistics are supportive of the general conditions, but there are subtle differences between statistics. This suggests that it may be possible to use a combination of these texture statistics to infer more specifics about the road surface than what is apparent from looking at only a single statistic.

Case 3: I-70 West of Vail Pass

As a final test of these co-occurrence texture metrics comparisons, we got the 2014 paving schedule from CDOT and picked some roads that they were getting ready to pave. One ideal case is the right or slow lane of the two eastbound lanes of I-70 just to the west of Vail Pass. A WorldView-1 image was available for this area that was collected August 17, 2011. Fortunately, WorldView-1 has the same 50 cm resolution panchromatic band as WorldView-2, which we used earlier for the Larimer County image.

As can be seen in the image of **Figure** 39, the poor condition of this right eastbound lane visually differs from the left eastbound lane and both of the west bound lanes.

Figure 39: WorldView-1 panchromatic image of I-70 just to the west of Vail Pass

In fact the right eastbound lane is only slight darker than the shoulder of the highway. We selected the ROIs for good and poor pavement as shown here in **Figure** 40.

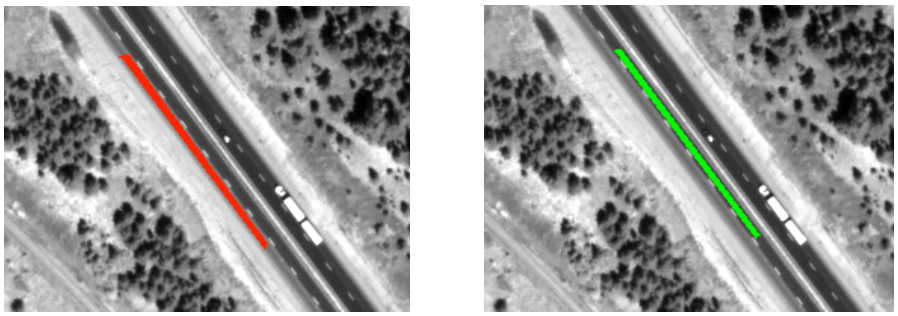


Figure 40: ROIs for statistical computations of the poor road surface (red) on the left and the good road surface (green) on the right

We have used the two eastbound lanes rather than mix between east and westbound lanes. We unfortunately do not know when either the left or right eastbound lanes were last paved.

As before, the first statistic that we will examine is the mean computed over a 2-D moving area. The histograms of these means are presented here in Figure 41 again with the poorer surface conditions on the left.

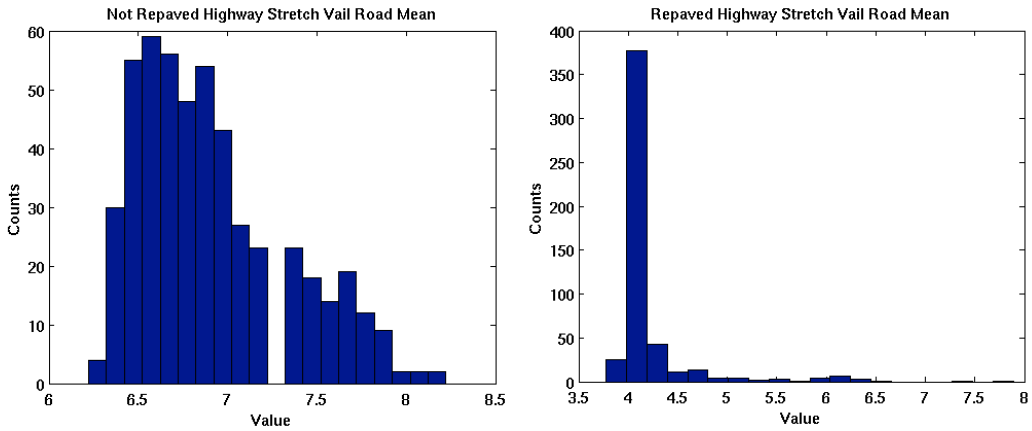


Figure 41: Histograms of reflectance mean from the right eastbound lane of I-70 (left) and the left eastbound lane of I-70 (right)

Again, there is a dramatic difference in these statistics. The "poor" surface values are all greater than the peak value of the "good" surface. In addition, the "poor" surface histogram has a very large range of values with peaks ranging from 3 to 8. It is this range of higher histogram values that seems to be very characteristic of poor road surfaces. The good surface has a single large histogram peak at about 4.1. Comparison between this good histogram with previous mean good histograms in **Figure** 27 and **Figure** 34 all show similar behavior with a single large peak but it is very interesting that the peak in **Figure** 27 is at 2 for a road paved in 2011, in **Figure** 34 is at 3 for road paved in 2007 and here in Figure 41 is at 4 for a road that may have been paved prior to 2007. The mean histogram peak may suggest the progressive degradation of the road according to its age.

A similar comparison of the poor road surfaces is not as revealing. All of the histograms have a very large range but there is not a consistent shift of the histograms. If we compare with the I-25 condition of the 2001 paved surface on the left of **Figure** 34, we find that its peaks range over larger values than does Figure 41 but the peaks in Figure 41 are larger. It may be once a highway has reached a poor condition, it is difficult to discern the exact condition from these metrics.

Turning now to variance, we present those in **Figure** 42.

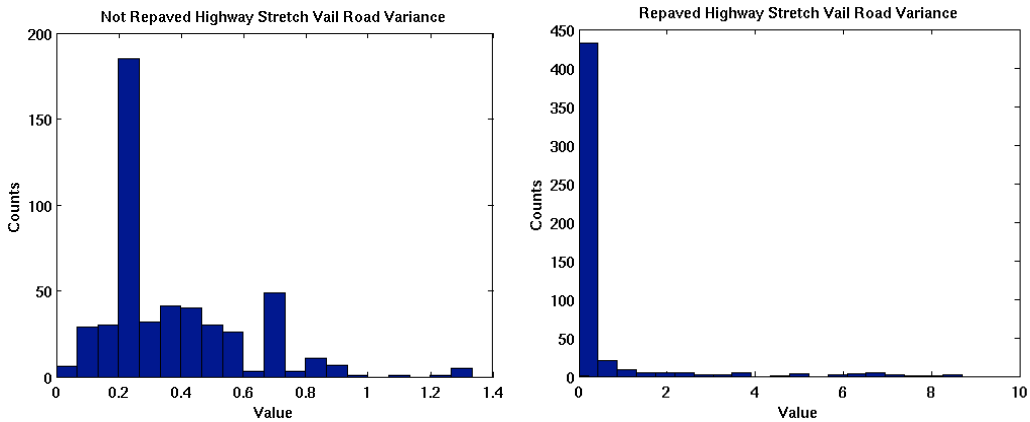


Figure 42: Histograms of reflectance variance from the right eastbound lane of I-70 (left) and the left eastbound lane of I-70 (right)

As expected, the good surface condition on the right has a single large peak near zero. The poor lane has its largest peak at 0.2 but again the peaks are all spread over the range of values up to about 1. The right panel of **Figure** 42 is very similar to the corresponding panels in **Figure** 28 and Figure 35 but the left panel is very different than the corresponding panels in **Figure** 28 and Figure 35.

We will not present the contrast and homogeneity statistics that were presented for the other cases and we will close with the entropies for this case in Figure 43.

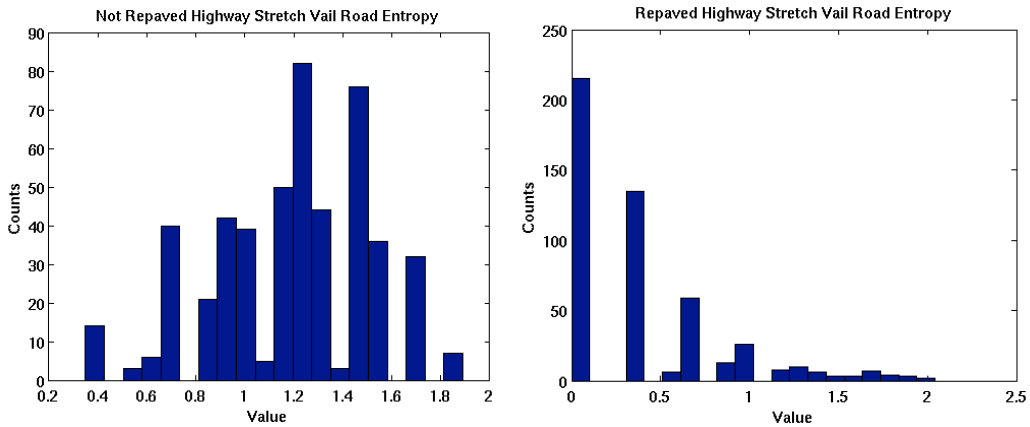


Figure 43: Histograms of reflectance entropy from the right eastbound lane of I-70 (left) and the left eastbound lane of I-70 (right)

As before, the main entropy peak is near zero for the good surface (right) and is widely spread for the poor surface on the left. There are, however, a lot more peaks on the right at higher values than what appeared in any of the earlier entropies for good road condition. Since we do not know the date of repaving of the good lane, this might reflect that this date is not particularly recent. This was suggested by the earlier comparisons of the mean values that appear to be confirmed here by the entropy.

Discussion of Case 3

Unfortunately, we do not know the dates of previous paving of either the good or bad surfaces of the eastbound lanes of I-70. We do know, however, that the right eastbound lane is on the schedule to be paved in 2014 so we can easily infer that it was last repaved much earlier than the left lane. This is evidenced by the appearance of each lane in the satellite image and by the co-occurrence texture statistics computed from the satellite data. Looking at the mean value of the texture, there is a suggestion that the good surface might have been paved sometime before 2007 when compared with the I-25 road just south of Colorado Springs.

General Discussion of All Cases

The analysis of the images of these highways has clearly indicated that good and poor road surface conditions can be distinguished in high-resolution satellite images. More importantly, texture statistics computed from these images provide insight into the conditions of the roads that can hopefully be used in the future to decide when certain roads may be in need of repaving. These results are a lot more promising than our earlier studies of the satellite images relative to the *in situ* data collected by CDOT and the various MPOs that we are working with. Perhaps if we restrict our analysis to good, fair, and poor road surface conditions, we will be much more successful in being able to estimate these conditions from the satellite data alone.

One might think that all we have achieved is a satellite measure of the age of the highway. But a comparison of the statistics for US 287, paved in 1998, with those from the 2001 paved segment of I-25 south of Colorado Springs suggests that is not the case. For example, the histogram of mean values for I-25 exhibits a wider range of values and a much broader histogram (**Figure** 34 left) than that for US 287 (**Figure** 27 left). Thus, a road paved in 1998 has a lower set of mean values than a road paved in 2001. Still, this seems consistent in that I-25 sees a greater volume of traffic traveling at much higher speeds than does US 287. This is supported by the entropy values which in Figure 38 left for I-25 (paved in 2001) that spread from 0.2 to 2.2 with a large mean peak at 1.2 and no values near zero while US 287 (paved in 1998) in **Figure** 31 left has a peak near zero, a central peak at about 1 and maximum value of 1.8. Other statistics confirm that the older road actually exhibits texture statistics that suggest a better road surface condition for US 287 than for I-25 (paved in 2001).

Special Study of Colorado Springs

Since the beginning of this project, we have worked closely with PPACG, the Colorado Springs MPO. They provided us with their last sample data set that was collected in 2007 and also arranged for a new collection of *in situ* data in 2013. For 2007, we have a QuickBird image and we arranged for WorldView-2 images to be collected during the 2013 survey. We were limited to an analysis of the 2007 data since it was reported as polylines that follow the tracks of the vans that collected the *in situ* data. Unfortunately, the 2013 data set was reported as polypoints and we can only extract the satellite data along the lines followed by the vans collecting the data.

We did, however, carry out an analysis of the 2007 dataset. Once again, we were frustrated in being able to regress extracted satellite measurements again the *in situ* data set. Realizing that much of this problem was due to scatter in the *in situ* data set we decided to again look only at the satellite data but now only extracted along the polylines that describe the tracks of the vans collecting data. Linear statistics along these lines were used to classify the conditions of the road surfaces. We basically had three classes. Good condition (bright green), poor (red), and intermediate (olive drab). These classes were used to classify the roads in central Colorado Springs in **Figure** 44. You can see that poor patches are juxtaposed to good and intermediate road conditions. This is typical of city streets, which are basically maintained according to citizen complaints about road conditions rather than by any predetermined maintenance schedule. Intersections are particularly susceptible to surface degradation as evidenced by the numerous red line segments at or around intersections.

It seems strange that the I-25 corridor has good, fair, and poor conditions. But in light of Case 2 above, this is an expected condition depending on the road's paving history. It is very likely that these surface conditions can be directly linked to the past paving schedule. In **Figure** 45, we present the same analysis for different parts of the satellite image that showed similar features. Intersections are again the sites of poor surface conditions even when the road surface itself is in good condition. Here, the I-25 surface exhibits largely poor (red) surface condition but in the north of the image the surface now appears to be in good (bright green) condition. Most of the rectangular crossing streets are in good condition with the caveat for the numerous intersections.

We are trying to find out how well this analysis compares with the PPACG's earlier analysis and paving decisions. Even though the *in situ* parameters (roughness, rutting, and cracking) are reported with high spatial resolution, we have learned that most paving decisions are made using a much broader designations of "good", "fair", and "poor". Thus, we have elected to analyze the satellite images with a similar set of classifications. This appears to overcome the problems we had in comparing with the *in situ* data.

Figure 44: Road surface conditions computed from a QuickBird image of Colorado Springs in 2007

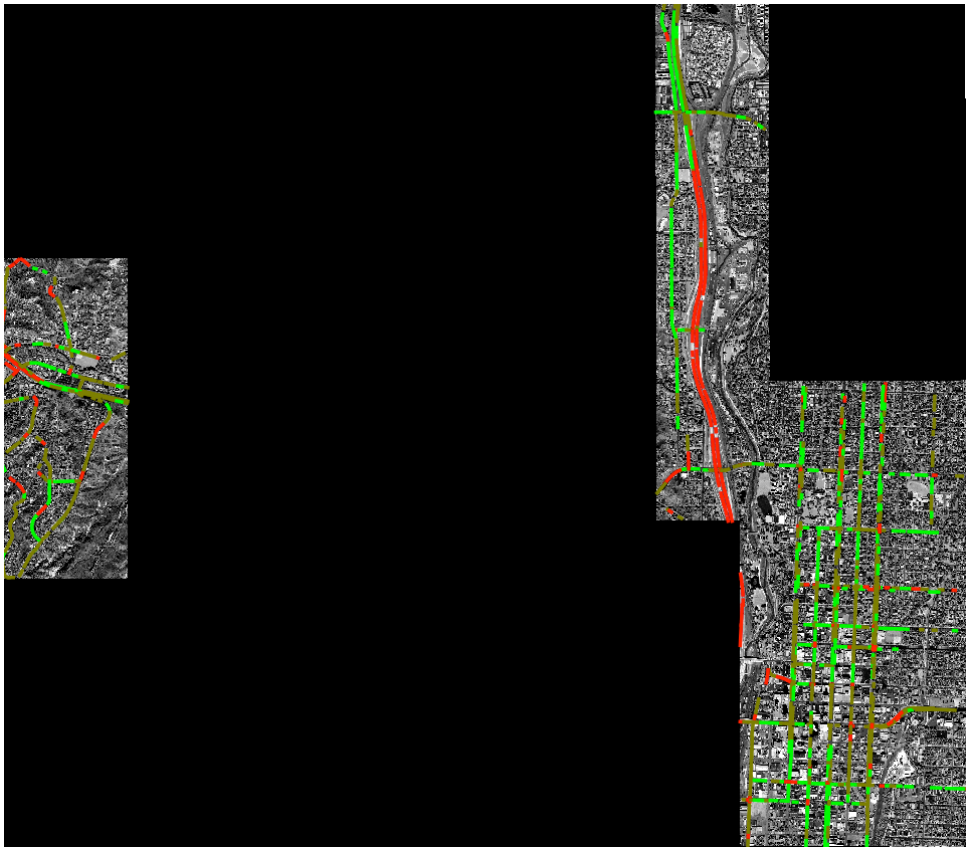


Figure 45: Same as Figure 44 but for a different area

Discussion of Colorado Springs Case Study

Once again, the real success appears to have come when we no longer attempted to regress the extracted satellite information against the *in situ* measurements. We want to now work with CDOT and our MPOs to determine if our satellite measurements could be used to assess future paving needs. If so, the satellite data represents a much simpler and more cost-effective method to determine when a road surface is in need of repaving or at least some form of maintenance. This can be done in a city or out on the open road wherever satellite data are available.

Road Identification

At this stage, we realized that we had a good method to evaluate road surface conditions. We then wanted to make it an automated system that could be used in a variety of situations to evaluate road health. To do this, we needed to develop a method to automatically locate the roads in the satellite image first. We tried a variety of methods first using some software known as RoadTracker to find the roads in a satellite image. Unfortunately, experience showed us that this software was only able to detect and map about 30% of the roads. We then turned to the United States Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) files of roads to located road surfaces in Colorado. Although most of the roads were present in these files, they did fit the road centers in satellite images to the desired level of accuracy. Finally, we turned to an open source database known as OpenStreetMap to locate the roads in a satellite image. This proved to be the most useful approach and since it didn't cost us anything, it was an advantage in developing our general system.

There was only a few months left in the project, but we continued to develop the image classification software that would allow us to remove pixels containing paint lines, vehicles, concrete, shadows, and road signs. We managed to develop a useful package that appeared to satisfy all of the requirements. We could only do superficial testing on this software package, but it appeared to do what we wanted. We could use at least a year of additional funding to further develop this software and evaluate an automated system for making paving decisions from high-resolution satellite imagery.

Time Series Analysis

Over the last two years, we have been investigating the concept of using high-resolution satellite imagery to determine the health of roads paved with asphalt. The most salient discovery from our studies is that asphalt roads in poorer condition generally have a brighter and an overall more heterogeneous appearance than roads in better condition. These characteristics are discernable in satellite imagery. Therefore, by analyzing the brightness and heterogeneity of road surfaces present in a satellite image, the relative quality of these roads can be assessed.

These conclusions were arrived at only after a considerable amount of time was spent trying to make some calculated parameters from the satellite imagery match *in situ* data

collected by the Colorado Department of Transportation and the Colorado Springs MPO. None of these studies gave us any usable information and we later learned that these organizations do not use the *in situ* data at the full resolution. Instead, they reduce the wide-ranging information down to broad designation of good, fair, and poor (CDOT actually has poor and poor-0). Once we realized that these were the road surface designations that we needed to target, it totally changed our approach. Now we were able to use the high-resolution satellite imagery effectively to find good, fair, and poor road surface conditions. In this regard, the most important information provided us was the paving schedule of CDOT.

In our previous experiments, we observed many highways throughout the state of Colorado that are maintained by CDOT. The results of these experiments have shown that there is a strong relationship between the health of a road and its visible appearance. Most of these experimental trials have only involved three types of highways: one freshly paved highway with brand new asphalt (good condition), one dilapidated highway with asphalt older than ten years (poor condition), and one more highway of intermediate health with asphalt a few years old (fair condition). When the image pixel statistics of these highways were compared, it was clear that the brightness and the heterogeneity (measured by the texture of the image) of the roads' appearances increase as the roads' conditions worsen.

The contrast between a new road and an old road is indeed quite stark. However, the contrast between roads with only small differences in asphalt age is not as obvious. In our latest experiment, we observed the pixel statistics of roads with asphalt that is closer in age than what we have studied in the past. The intention was to see whether these subtler differences are detectable in satellite images.

We acquired a series of satellite images over Longmont, CO from DigitalGlobe (DG). These images were collected by the WorldView-2 spacecraft from 2010-2013, with one image per year. They are panchromatic images with a ground spatial resolution of ~50 cm. The specific properties of these images are listed in Table 3 below.

Image Date	Maximum Off-Nadir Angle	Minimum Sun Elevation Angle
June 24, 2010	8.01°	69.90°
August 20, 2011	20.78°	61.36°
October 17, 2012	10.90°	40.08°
July 16, 2013	24.87°	66.73°

Table 3: WV-2 Longmont Image Properties

The reason why Longmont was chosen for this investigation is because DG is currently based there. Their satellite data are frequently collected over the area for sensor calibration and experimental image processing purposes. An additional benefit is that DG processes these images of Longmont to include surface reflectance giving us an even more precise measurement of surface conditions rather simple satellite radiances represented by the raw digital number. Reflectance is likely to be much more representative of surface conditions than DN.

Unlike the image sets from previous experiments in which the pixel data was represented in the form of raw DN values, these images contained pixel data represented by surface reflectance. That is, each pixel is represented by a number equivalent to the ratio of the amount of light reflected by the target to the amount of light incident to the target. Objects with higher reflectance appear brighter than those with lower reflectance. The raw DN values that have been used in prior investigations are only proportional to the amount of light striking the satellites' detector. They have not been compensated for effects such as the scattering and absorption of the light as it propagates through the Earth's atmosphere. However, the surface reflectance values in this investigation have been compensated for such effects. Therefore, they provide more information regarding the inherent physical properties of the observed targets, in this case road asphalt.

We also acquired GIS shape files from the City of Longmont Public Works & Natural Resources from the years 2009-2013 for ground-truth information. Each shape file contained the roads that have been repaved during its respective year. For example, the 2009 shape file contains the roads that have been repaved in 2009 and so on. Figure 46 displays these shape files laid out over the WV-2 image of Longmont captured in 2013.

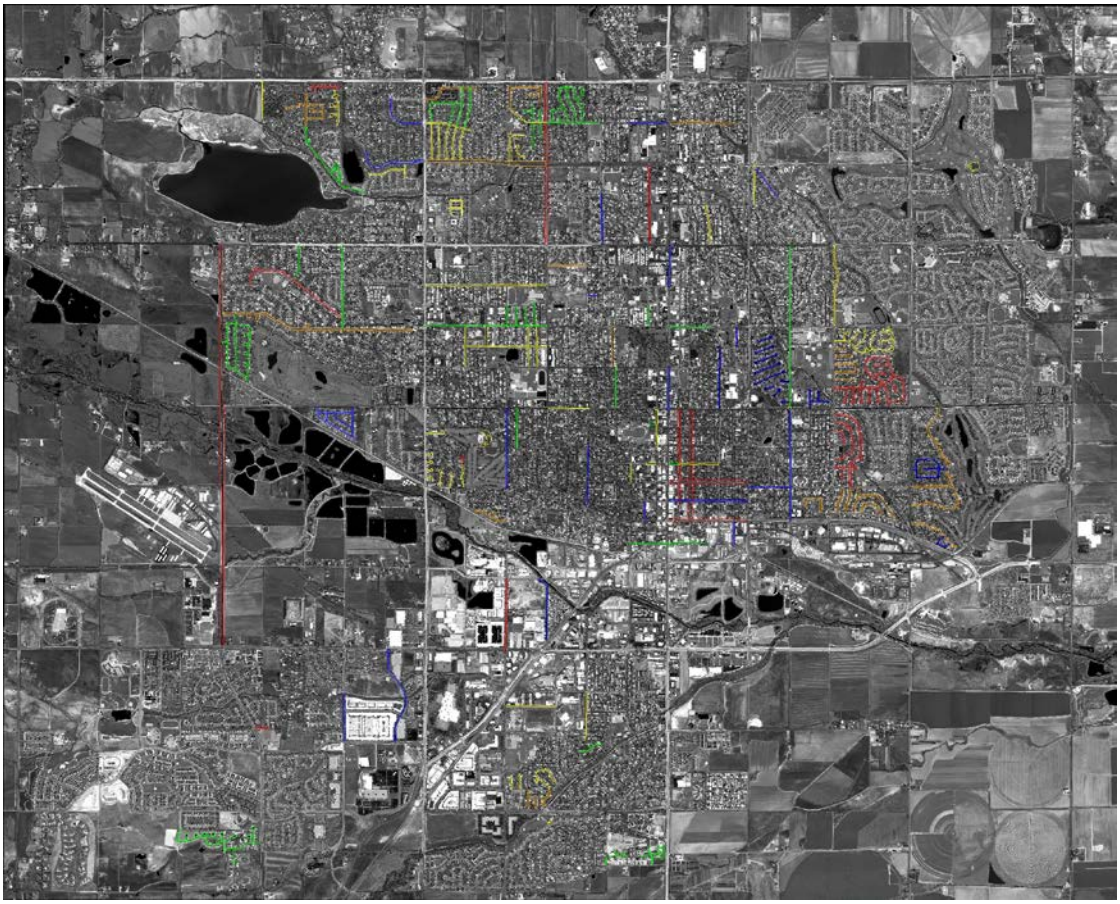


Figure **46: Longmont r**epaved roads

Although it might be difficult to see from Figure 46, the repaved roads are color-coded according to their year of repavement. The red roads were repaved in 2009, the orange ones were repaved in 2010, the yellow ones in 2011, the green in 2012, and finally the

blue in 2013. This same color-code will be used for the histograms shown later in this report. By explicitly showing which roads have been repaved and when, these shape files made it easier to select regions of interest, which we could use to perform our analysis.

Surface Reflectance Analysis

As the roads degraded over the years, the statistics of their surface reflectances were examined using the information from the satellite imagery. Regions of interest were first selected over the roads. Each ROI covered roads that were repaved in a given year. Both neighborhood streets and city streets from the entire scene were sampled. In this particular experiment, the ROIs were selected manually because we wanted to carefully and meticulously avoid non-asphalt objects commonly found in roads. This is one of the biggest challenges of the satellite image based method and that is to remove anything from the image is not illuminated asphalt. Thus we had to select out paint lines, cars, shadows, street signs and all other elements that blocked direct access to the road surface.

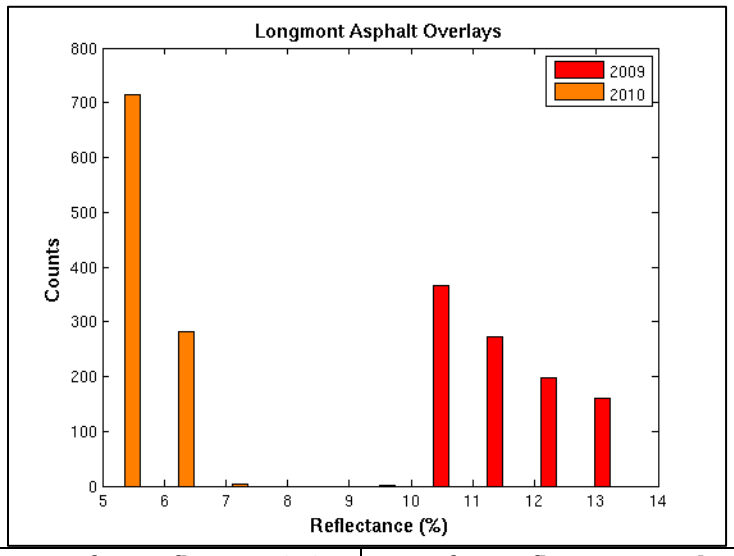
As mentioned previously, we are currently developing an automated classification technique to separate the asphalt pixels from the non-asphalt pixels. In this technique, we first use OpenStreetMap road shapefiles to mask out the non-road parts of the scene. Then, we manually select just a few asphalt and non-asphalt pixels. The information within these pixels serves as a training set from which a random forest classifier can be built. This classifier can be applied to the rest of the road pixels to determine which ones represent asphalt and which ones do not. The asphalt pixels are kept as part of the ROI, while the non-asphalt ones are discarded. We hope that we can develop this technique to the point where we can quickly create road ROIs in future experiments.

It is important to study diverse samples of roads for two reasons. The first reason is that we can account for any variations in asphalt degradation caused by differing traffic levels. Traffic tends to be heavier in city streets than in neighborhood streets. As a result, city streets tend to wear down faster. By considering all these different kinds of streets, we can get a broader idea of how asphalt road deterioration appears in satellite imagery. In addition, the non-asphalt elements contaminating our road images are very different in various geographic areas. In dense urban areas, shadows of high-rise buildings are a real source of contamination. Trees also provide shadows, but since many are deciduous it is possible to view the target surface in the respective winter season where the trees have lost their leaves and provide far less shadow to the road surface.

The second reason is so that we can account for any variations in the asphalt's observed surface reflectance as it relates to the satellite's viewing angle. The reflection pattern of light as it strikes a surface is called the bidirectional reflectance distribution function (BRDF). The BRDF is dependent upon both the angle of the light incident to the surface and the angle at which the observer (satellite) views the surface. Since the predominant light source is the Sun, it is estimated that the incident light is at the same angle across the whole scene in a given satellite image, so that factor can be assumed constant. However, there might be slight variations in the satellite's viewing angle for each pixel in the scene. These slight variations may skew the observed surface reflectance of the

asphalt. Therefore, in order to understand how asphalt reflects incident light, it is necessary to consider roads from multiple viewing angles all over the scene.

In the first trial of this experiment, roads in the WV-2 image of Longmont captured in 2010 were examined. The ROIs extracted from the roads repaved in 2009 and 2010 were laid over the image. The surface reflectance values contained within the pixels underneath both ROIs were extracted and plotted. The histogram and table below (Figure 47) show and describe the distributions of the surface reflectance values for these ROIs. Recall that this compares a recently paved surface reflectance from 2010 with a slightly older surface from 2009.



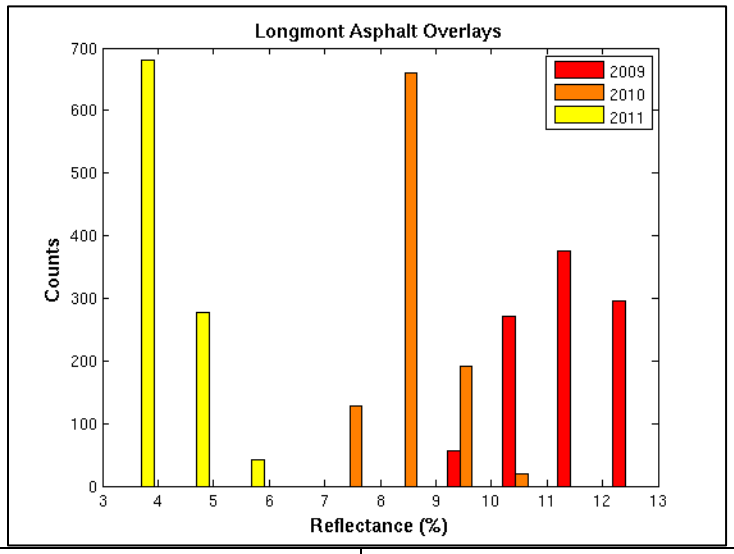
Year	Mean Surface Reflectance (%)	Surface Reflectance Standard Deviation (%)
2009	11.6	0.9
2010	5.6	0.3

Figure 47: Road Surface Reflectance in 2010 WV-2 Image

The reflectance distribution of the asphalt from 2009 has a higher mean and standard deviation than the asphalt from 2010, the year when the image was captured. This phenomenon implies that the asphalt from 2009 deteriorated slightly in the one year span from when it was laid down. For practical purposes, both classes of asphalt would be considered “good” by most organizations that deal with road quality management. However, from these results, it is clear that the quality of asphalt can noticeably change in just one year. Thus, the satellite image analysis is capable of identifying fairly subtle changes even between roads that would generally be classed in the same broad condition. Here we can use the satellite imagery to detect what has happened to an asphalt surface in just a single year.

The second part of this experiment was similar to the first. But this time, the WV-2 image captured in 2011 was used and the asphalt laid down during that year was also examined.

The histogram and table below (Figure 48) show and describe the distributions of the surface reflectance values for the roads repaved from 2009-2011.

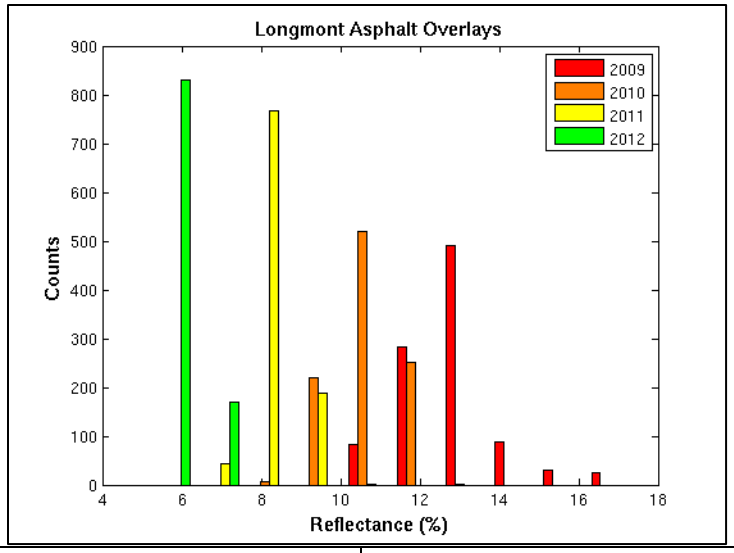


Year	Mean Surface Reflectance (%)	Surface Reflectance Standard Deviation (%)
2009	11.4	0.6
2010	8.7	0.5
2011	3.9	0.4

Figure 48: Road Surface Reflectance in 2011 WV-2 Image

The reflectance distribution of the 2011 asphalt has the lowest mean and standard deviation, followed by the 2010 asphalt and then the 2009 asphalt. This is consistent with 2009 being the oldest of the road surfaces. Although it is not as clear, there is a wider spread of surface reflectance values in 2009 than in either 2010 or 2011. This spread is related to the spatial heterogeneity that we will discuss later.

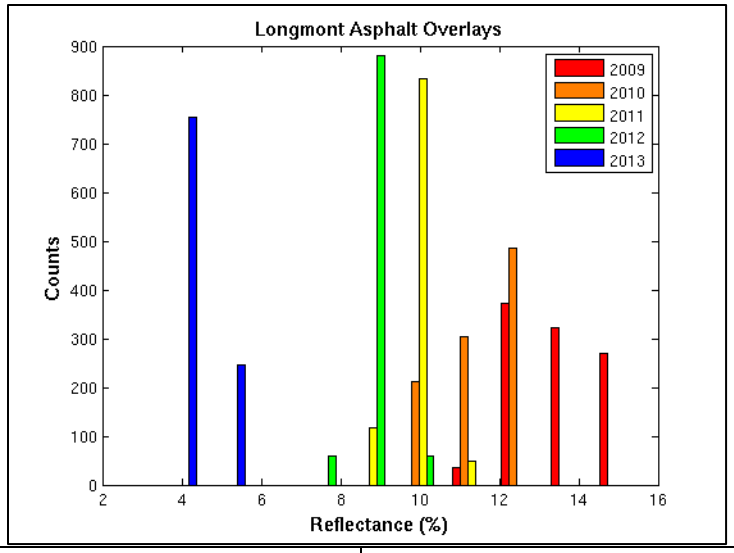
The next two trials involved the satellite images ranging up to 2012 and 2013 respectively. The trial with the 2012 image involved the roads that were repaved from 2009-2012. The surface reflectance histogram and table are below in Figure 49. It clearly demonstrates again the ability of the image analysis to separate out the age classes of the road surfaces descending from the newest in green down to the oldest in red. The spread again dramatically increases from the green (2012) to the red (2009). In Figure 50, we add 2013 to this analysis. Once again, there is strong agreement with all of our earlier conclusions. The most recently paved surface has the lowest surface reflectance and the narrowest spread while the oldest pavement exhibits higher values spread over a larger number of lower histogram peaks.



Year	Mean Surface Reflectance (%)	Surface Reflectance Standard Deviation (%)
2009	12.8	1.2
2010	10.7	0.8
2011	8.4	0.5
2012	6.0	0.4

Figure 49: Road Surface Reflectance in 2012 WV-2 Image

Finally, the trial with the 2013 image considered roads repaved from 2009-2013. The results are displayed below.



Year	Mean Surface Reflectance (%)	Surface Reflectance Standard Deviation (%)
2009	13.5	1.0
2010	11.7	0.8
2011	10.0	0.5
2012	8.8	0.4
2013	4.4	0.3

Figure 50: Road Surface Reflectance in 2013 WV-2 Image

In each of the trials, the mean and standard deviation increase with respect to asphalt age. Asphalt can only wear down as it ages. Thus, there appears to be a negative correlation between brightness and heterogeneity of the asphalt's appearance and the quality of the roads.

Surface Reflectance Discussion

Road asphalt is typically composed of a mixture of bitumen and mineral aggregates. Bitumen is an organic substance in the form of a black, sticky, highly-viscous liquid or semisolid. It binds the aggregates together. The aggregates themselves are primarily tiny pieces of rock, sand, and gravel. They provide bulk and robustness to a road that allows it to withstand stress from passing vehicles.

Because pure bitumen is black in color, a freshly paved road has a dark and smooth appearance. But over time, the bitumen wears away leaving the mineral aggregates exposed. By contrast, the aggregates are generally lighter in shade. Therefore, as an asphalt road deteriorates, it tends to get brighter as we have seen in our results.

There are however some other interesting results from this experiment that have not yet been discussed. They can be seen most clearly in the histogram for the 2013 image. Notice that there is a large difference between the mean surface reflectance values for the roads paved in 2012 and the roads paved in 2013. This value is larger than the differences

between the mean reflectances for the roads paved in other pairs of consecutive years. Not only that, but in the histogram, it is not as easy to distinguish the roads repaved in the earlier years because their reflectance distributions are so close to each other. The reasonable conclusion is that the rate at which an asphalt road becomes brighter appears to slow down with time.

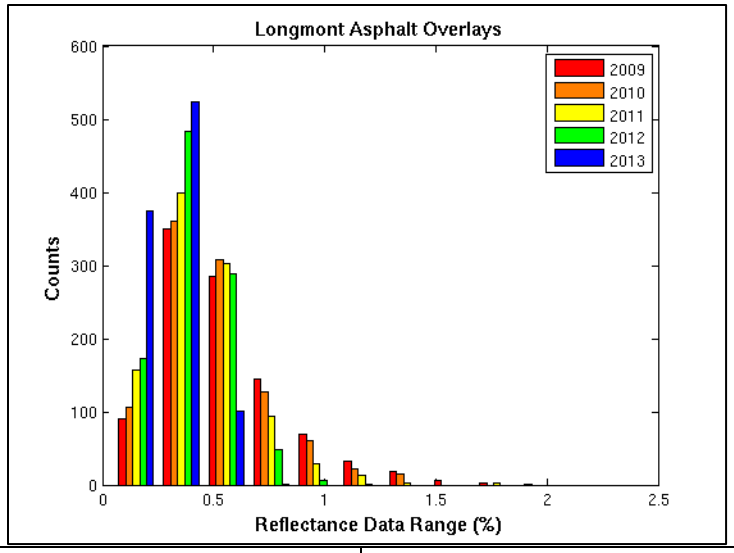
This phenomenon seems to imply that a disproportionately large percentage of the bitumen used in the original asphalt mixture gets worn out in just the first year. Then in the remaining years, less and less bitumen gets depleted. But eventually, the amount of bitumen in a road is not enough to hold the mineral aggregates together. As a consequence, the road becomes weaker and more susceptible to damage. The amount of bitumen in a road (and thus the road's structural integrity) can possibly be measured through this satellite remote sensing technique.

Texture Analysis

Although the analysis of the road surface reflectance by itself seems to give a good indication of road quality, texture analysis of the road imagery supports the idea that the heterogeneity of an asphalt road's appearance and the road's health are interrelated. The surface reflectance satellite images of Longmont were processed through the three main occurrence-based texture filters: data range, variance, and entropy.

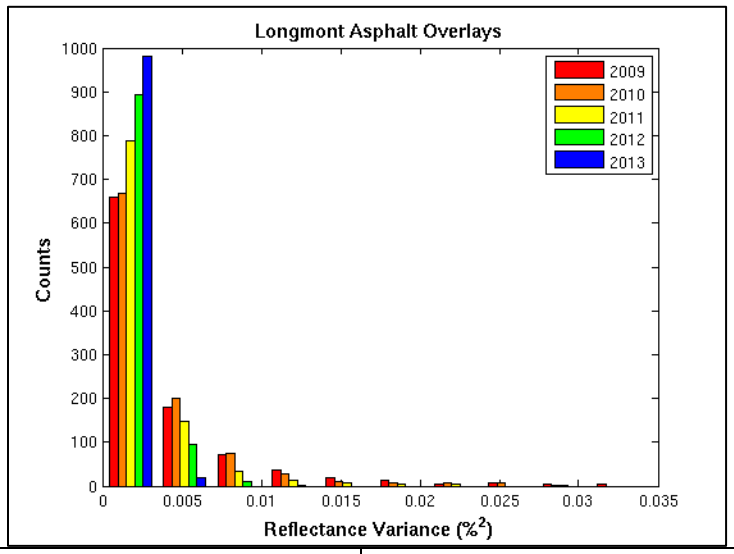
The data range, variance, and entropy texture parameters all give an indication of the relative uniformity of an object's appearance in an image. For all three parameters, higher values represent less uniformity. In the previous section with the surface reflectance, the roads with older asphalt had a wider distribution of reflectance values. This was demonstrated by the larger standard deviations, which implies that older asphalt is less uniform. If a related observation were to be seen in texture filtered imagery, then it can also likely be used to assess road quality.

The same procedure of extracting the pixels underneath the ROIs and plotting the values as histograms was repeated for the textured filtered imagery. Similar results were obtained from the images for all the years. But to avoid cumbersome repetition, only the results from the 2013 image will be presented here. The histograms and tables that show and describe the pixel statistics are below (Figure 51, Figure 52, and Figure 53).



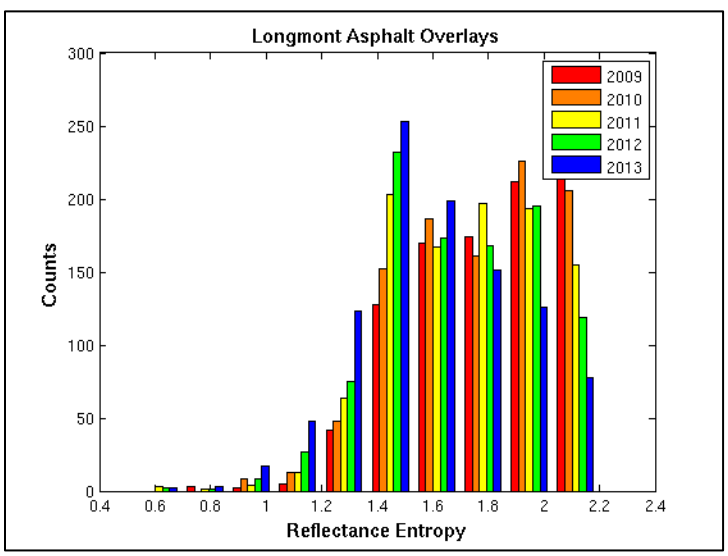
Year	Mean Surface Reflectance Data Range (%)	Surface Reflectance Data Range Standard Deviation (%)
2009	0.54	0.28
2010	0.52	0.25
2011	0.44	0.21
2012	0.40	0.15
2013	0.30	0.11

Figure 51: Road Surface Reflectance Data Range in 2013 WV-2 Image



Year	Mean Surface Reflectance Variance ($\%^2$)	Surface Reflectance Variance Standard Deviation ($\%^2$)
2009	0.0398	0.0484
2010	0.0362	0.0408
2011	0.0252	0.0313
2012	0.0184	0.0145
2013	0.0102	0.0072

Figure 52: Road Surface Reflectance Variance in 2013 WV-2 Image



Year	Mean Surface Reflectance Entropy	Surface Reflectance Entropy Standard Deviation
2009	1.78	0.25
2010	1.75	0.25
2011	1.70	0.25
2012	1.68	0.26
2013	1.60	0.26

Figure 53: Road Surface Reflectance Entropy in 2013 WV-2 Image

For all three parameters, the results once again demonstrate that asphalt heterogeneity and road quality are negatively correlated. The entropy (Figure 53) is highest for the oldest road, while the reflectance variance (Figure 52) and the reflectance data range (Figure 51) are both low for the oldest road surfaces. Since road quality can only worsen with time, it is assumed that the roads with older asphalt are in poorer health. The roads paved in 2013 have the lowest mean data range, variance, and entropy values. These mean values progressively increase for the roads paved in earlier years. Even when much shorter timescales are considered, this analysis technique reinforces the already established link between asphalt uniformity and road health.

Texture Discussion

The bitumen component of asphalt is the primary contributor to a freshly paved road's dark and smooth appearance. But as it is weathered away over time, the mineral aggregates become more exposed. As a consequence, not only does the road's brightness increase, but its overall appearance becomes rougher.

In addition, the road becomes more vulnerable to damage such as rutting and cracking. These kinds of distress are visible on a road's surface when viewed up close. Individual instances of road distress may be too small to detect from space even with WV-2's high-resolution capabilities. However, they can accumulate on a road's surface over time. As they do, they can make a road look more varied in the satellite imagery.

Older roads that have accumulated more damage will appear to be less uniform than newer roads. This phenomenon was verified by our results above. For all three occurrence-based texture parameters, higher values represent a more heterogeneous landscape. As the age of the roads increased, the mean values of the texture parameter distributions also increased. Therefore, this technique can also be used to determine road health from satellite imagery. It can supplement the results from the surface reflectance analysis.

Asphalt Composition Caveat

In the analysis described above, results involving only asphalt laid down during the years 2009-2013 were presented. In actuality though, asphalt laid down during 2008 was also examined in the original analysis. It was hypothesized that its deterioration would be

marked by increased brightness and heterogeneity. For the most part, such a trend was indeed observable in the data. However, the rate at which this asphalt became brighter did not seem to match that of the asphalt from the later years. Rather, it was much slower. The histogram below (Figure 54) shows the surface reflectance distributions of the asphalt from the years 2008-2013 as seen in the 2013 WV-2 image of Longmont with the 2008 data in grey.

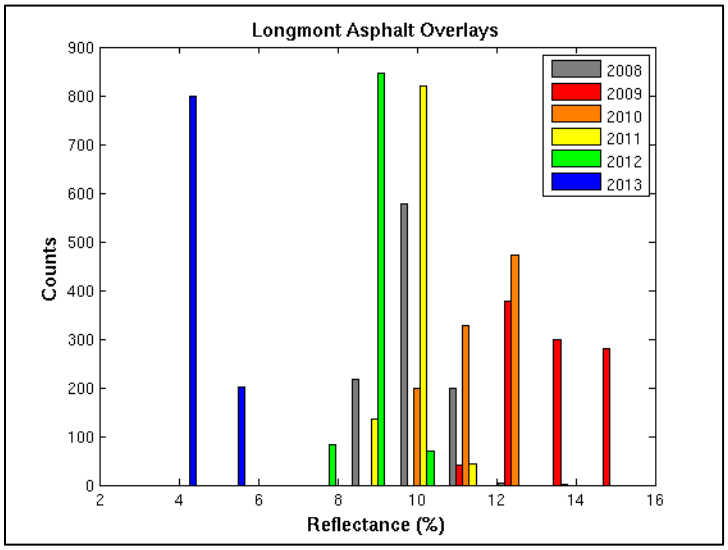


Figure 54: Road Surface Reflectance Data Range in 2013 WV-2 Image (w/ 2008 Roads)

Notice that the mean reflectance values increase as the asphalt age increases. However the mean surface reflectance for the 2008 asphalt in the image is 10.0%. It is not higher than that of 2009 like we might have expected. Similar results were present in the imagery from the earlier years. This was a very unusual anomaly. We contacted the City of Longmont Public Works & Natural Resources to ask for any insight into this matter. They replied that a different contractor was employed to repave the roads in 2008 than in 2009-2013. Therefore, it is highly likely that the composition of the asphalt from the two contractors was different. This could possibly explain why the 2008 asphalt did not become brighter at the same rate. At face value, this result would suggest that the 2008 asphalt is more durable than the asphalt from later years. As of now, it is unclear whether it truly is. But if satellite remote sensing techniques were to be used to assess road quality, it is vital that this kind of information be available *a priori*.

Findings and Conclusions

An important realization in the first year of this project is that no parameter extracted from the high-resolution satellite imagery could be made to correlate with the *in situ* measurements of roughness, rutting, and cracking. There were a number of reasons for this lack of correlation, but the main reason was the noisy and error-prone nature of the *in situ* measurements. Most of these measurements are based on human interpretation of video images collected by a van in the curb lane of most roads and highways. Human analysis is very prone to reading error and this curb lane is definitely not representative of

the entire road surface. Still, all paving decisions are presently made based on some form of these *in situ* measurements.

Another reason for this lack of correlation is the complexity of removing non-asphalt pixels from the satellite images. This opens the door to having the asphalt pixels influenced by the surrounding non-asphalt pixels. We later developed good, automated techniques for removing these non-asphalt surfaces in the satellite imagery.

Once we realized that trying to fit the satellite data to the *in situ* measurements was not going to work, we managed to get paving schedules and past paving records. This offered a much more useful set of conditions and recommendations with which we could compare our analysis of the satellite data. In addition, we carried out an analysis of photographs of good and poor asphalt surfaces to better determine how our spatial texture analysis was functioning.

We also developed an automated classification procedure (using random forest simulations) to remove the non-asphalt surfaces from our images. Armed with this tool, our knowledge of the inverse relationship between visible brightness and asphalt age along with our spatial texture analysis, we examined a number of roads and highways and we were clearly able to say what the asphalt surface condition was at a specific location. This analysis corresponded well with the paving schedules from Larimer County and for the Colorado Department of Transportation (CDOT) for recent years.

Thus, we believe that we now have a viable method to estimate asphalt road surface condition from high-resolution satellite images. The problem is whether any government body responsible for road maintenance would use this method either to compliment their present *in situ* measurement program or to replace it. Conversations we have had with personnel from CDOT and Pathway Services Inc. have suggested that DOT will still require all government agencies to make these traditional *in situ* measurements and would not trust parameters derived from high-resolution satellite imagery.

Future Plans

Although technically this project is at its conclusion, there is still much more work that could be done. We now have sound prototypical methods to identify road asphalt from satellite imagery and determine its relative quality. The next steps involve developing these techniques further and streamlining them into a workflow that can be implemented quickly and efficiently. This would result in an automated software system that could be applied to any high-resolution satellite image of a road area of interest. More applications to diverse road conditions are needed to test this system and evolve it into a robust software package that could be used routinely by a wide variety of people and organizations.

Our interactions with companies such as Pathway Services Inc. have convinced us that these groups are not motivated to incorporate satellite data analysis into their measurement programs as that could potentially put people out of work and lead to a

change in their business model. The only way our new satellite-based methods could find acceptance in the wider road transportation community would be if we could find a state DOT that was willing to try our methods in parallel with their funded contractor *in situ* measurements. This DOT would then use our results independently of the *in situ* data to develop their paving schedule for the next year, which would make it possible to evaluate the value of the satellite derived information relatively to the *in situ* data collected. If successful, we would then work with the state DOT to implement a satellite observation based system into their paving schedule.

Any of these efforts would require additional funding to be able to finish the development of our methods and to find and work with one or more state DOTs on the testing of our satellite derived indices for the practical application to their paving decisions. Under its new management, CDOT is no longer a good candidate for this type of effort. But with additional funding, we can finish off the software and look for DOT partners to evaluate our methods and whether they can be used operationally to determine paving schedules.