Remote Sensing of Roads and Highways in Colorado

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

Contract No. **RITARS-12-H-CUB**

Quarterly Progress Report #6

Quarter from 01/01/2014 to 03/31/2014

Principal Investigator
Dr. William Emery
Professor
Aerospace Engineering Science Department,
University of Colorado at Boulder

Program Manager
Mr. Caesar Singh
Research Innovative Technology Administration,
U.S. Department of Transportation

Contents

GLOSSARY

CDOT	Colorado Department of Transportation
CU	University of Colorado
DG	DigitalGlobe
DN	Digital Number
IRI	International Roughness Index
MPO	Municipal Planning Office
PPACG	Pikes Peak Area Council of Government
QB	QuickBird
WV-1	WorldView-1
WV-2	WorldView-2

EXECUTIVE SUMMARY

In this report, the process that will be implemented to identify the road pixels in a given image is described (asphalt surfaces only). During this process, information from selected regions of the image itself is used to build a random forest classifier. The classifier can then be applied to the rest of the image to determine whether its pixels belong to roads or not. By being able to automatically identify road pixels, assessment of road surface conditions can be carried out more quickly.

I — TECHNICAL STATUS

Automated Road Identification

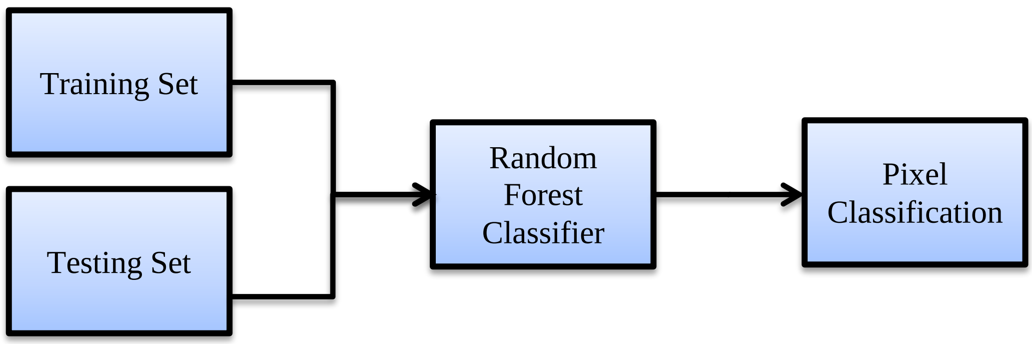
In the most recent project quarters, we focused the majority of our efforts on developing ways to assess the quality of road surface pavement from satellite imagery. We believe that we now have a method involving texture analysis that allows us to perform this assessment. However, in order to properly implement this method, the asphalt pavement must first be identified in the images. In this quarter, we worked on developing a technique, which we believe can be used to automatically separate the asphalt roads from the rest of the image features.

The image products provided by DigitalGlobe's WorldView-2 spacecraft come in two types: panchromatic and multispectral. The panchromatic images have a spatial resolution of ~0.5 m but only one spectral band. The multispectral images have a spatial resolution of only ~2 m but eight spectral bands. In order to identify the asphalt in the imagery, both the high spatial resolution of the panchromatic images and the multiple bands of the multispectral resolution are necessary. Both of these separate advantages were combined by pansharpening the multispectral images with the panchromatic images. The results were ~0.5 m resolution images with eight spectral bands. The figure below illustrates the concept of pansharpening. Here the 50 cm resolution of the panchromatic band on the left is merged with the 2 m multispectral image in the center. The result is a ~50 cm resolution multi-spectral image on the right.

In addition to the pansharpened multispectral images, texture filtered images of the panchromatic data were created for asphalt identification. The five main occurrence-based texture filters were used for this step: data range, mean, variance, entropy, and skewness. The filtered images also have a spatial resolution of ~0.5 m. The five texture filter bands along with the eight multispectral bands lead to a total of thirteen bands at a common spatial resolution that can be used for automated asphalt identification.

A scene from Colorado Springs was used to test out how well road pavement can be automatically identified through computerized means. Using ENVI's Region of Interest (ROI) tool, a testing set was manually selected. The testing set included two classes: "roads" and "not roads". The roads class consisted of highways and neighborhood streets varying in quality. The not roads class consisted of vegetation, soil, buildings, shadows, and other features. This testing set will be used to score how well the classification scheme will work. Three training sets were created with roads as one class and not roads as the other. The roads and not roads classes were identified manually for the first training set. The second training set used the US

Census Bureau's Topologically Integrated Geographic Encoding & Referencing (TIGER) shapefiles as the roads class and randomly selected pixels around the scene as the not roads class. The third training set used the output shapefile from the RoadTracker software, which is currently being developed by DigitalGlobe, as the roads class and randomly selected pixels around the scene as the not roads class. For all three training sets, the information contained in the pixels at the thirteen bands was used to build a random forest classifier. This random forest then classified the pixels in the testing set as either roads or not roads. The flowchart below illustrates this process.



After the pixels in the testing set were classified as either asphalt roads or not roads by the random forest, they were scored for correctness. The statistical measurement used to score how well the classifier performed is Cohen's kappa coefficient. The formula to determine Cohen's kappa coefficient, κ, is shown below.

$$\kappa = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

$\Pr(a)$ is the relative observed agreement and $\Pr(e)$ is the hypothetical probability of chance agreement. The coefficient ranges from 0 to 1. The closer it is to 1, the better the classifier. When the first training set, which was selected manually, was used to build the random forest, the κ value was about 0.93. The κ values from the other two training sets were about 0.82. Although the classifiers built from the latter two training sets are not as good as the one from the first, they can be created automatically whereas the first one cannot. Because we want to keep the manual steps in this project to a minimum, automation is necessary. So for now, we will continue to work with the TIGER and RoadTracker shapefiles as an independent method of defining the road locations where we can apply our classifier to extract the asphalt surfaces. These surfaces can then be assessed for their pavement condition using the texture metrics.

Future Plans

We intend to use the output of the classifier to keep only the road pixels in the images while masking out the pixels that are not roads. Then, we will apply the occurrence-based texture filters to these masked images. We can then use the texture-filtered images to determine the conditions of the roads within them. The ultimate product will be a map of the roads with their conditions outlined. The figure below shows an example of what we are aiming for.

II — BUSINESS STATUS

Please see Appendix.

FEDERAL FINANCIAL REPORT

(Follow form instructions)

1. Federal Agency and Organizational Element to Which Report is Submitted	2. Federal Grant or Other Identifying Number Assigned by Federal Agency (To report multiple grants, use FFR Attachment)	Page 1	of pages
3. Recipient Organization (Name and complete address including Zip code)			

4a. DUNS Number	4b. EIN	5. Recipient Account Number or Identifying Number (To report multiple grants, use FFR Attachment)	6. Report Type ☐ Quarterly ☐ Semi-Annual ☐ Annual ☐ Final	7. Basis of Accounting ☐ Cash ☐ Accrual

8. Project/Grant Period From: (Month, Day, Year)	To: (Month, Day, Year)	9. Reporting Period End Date (Month, Day, Year)

10. **Transactions**	Cumulative
(Use lines a-c for single or multiple grant reporting)	
Federal Cash (To report multiple grants, also use FFR Attachment):	
a. Cash Receipts	
b. Cash Disbursements	
c. Cash on Hand (line a minus b)	
(Use lines d-o for single grant reporting)	
Federal Expenditures and Unobligated Balance:	
d. Total Federal funds authorized	
e. Federal share of expenditures	
f. Federal share of unliquidated obligations	
g. Total Federal share (sum of lines e and f)	
h. Unobligated balance of Federal funds (line d minus g)	
Recipient Share:	
i. Total recipient share required	
j. Recipient share of expenditures	
k. Remaining recipient share to be provided (line i minus j)	
Program Income:	
l. Total Federal program income earned	
m. Program income expended in accordance with the deduction alternative	
n. Program income expended in accordance with the addition alternative	
o. Unexpended program income (line l minus line m or line n)	

11. Indirect Expense	a. Type	b. Rate	c. Period From	Period To	d. Base	e. Amount Charged	f. Federal Share
				g. Totals:			

12. *Remarks: Attach any explanations deemed necessary or information required by Federal sponsoring agency in compliance with governing legislation:*

13. Certification: By signing this report, I certify that it is true, complete, and accurate to the best of my knowledge. I am aware that any false, fictitious, or fraudulent information may subject me to criminal, civil, or administrative penalties. (U.S. Code, Title 18, Section 1001)

a. Typed or Printed Name and Title of Authorized Certifying Official	c. Telephone (Area code, number and extension)
	d. Email address
b. Signature of Authorized Certifying Official	e. Date Report Submitted (Month, Day, Year)
	14. Agency use only:

Standard Form 425
OMB Approval Number: 0348-0061
Expiration Date: 10/31/2011

Paperwork Burden Statement

According to the Paperwork Reduction Act, as amended, no persons are required to respond to a collection of information unless it displays a valid OMB Control Number. The valid OMB control number for this information collection is 0348-0061. Public reporting burden for this collection of information is estimated to average 1.5 hours per response, including time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the Office of Management and Budget, Paperwork Reduction Project (0348-0060), Washington, DC 20503.