Image Classification

Waiparous, AB

June 9, 2016

David Bouchard

8 Pages

1 Introduction

The Western Sky Land Trust (WSLT) is striving to conserve open and natural areas within the Calgary region, focusing on watershed lands associated to agricultural, natural, heritage, scenic, and recreational values (WSLT, 2015). They have approached our team to help them to produce a sampling strategy for a species inventory program on a land parcel near Waiparous, AB – including vegetation, wildlife, and aquatic species.

Producing an effective sampling program will require a good understanding of the land cover types found within the land parcel. Gaining this knowledge can be done with supervised land cover classification of multispectral satellite imagery. To produce this sort of classification, some field data was required: multiple polygons of each classification type were captured with GPS units, and were used to train the algorithms used to produce the supervised land cover classification.

The following sections detail the imagery used, the workflow method, the classification results, and the accuracy analysis.

2 Imagery Description

2.1 Provided Imagery

The imagery provided for the assignment was taken by the SPOT 6 Satellite on July 2, 2014 and September 6, 2014. SPOT 6 Satellite uses the NAOMI (New AstroSat Optical Modular Instrument) sensor which captures both panchromatic and multispectral imagery. The multispectral portion is four bands consisting of blue, green, red and near-infrared. The sensor has a spatial resolution of 1.5m for the panchromatic imagery and 6m for the multispectral imagery (Satellite Imaging Corporation, 2015).

2.2 Acquired Imagery

Imagery was found on an open data source provided by the Canadian Government. The imagery found was taken by SPOT 5 Satellite on August 21, 2009. SPOT 5 Satellite uses the HRG, HRS and Vegetation sensors which capture panchromatic and multispectral imagery. The HRG captures panchromatic and green, red, near-infrared and mid-infrared in the multispectral range. HRS captures panchromatic to be used for stereo imagery and the Vegetation sensor captures blue, red, near-infrared and mid-infrared. HRG has a panchromatic spatial resolution of 2.5m or 5m and a multispectral spatial resolution of 10m for green, red and NIR and 20m for MIR. HRS has a spatial resolution of 5 – 10m and vegetation has a spatial resolution of 1000m. An interesting aside from SPOT 5 imagery is that with the HRG sensor a true colour image cannot be obtained as it does not have a blue band, but one can be done using the Vegetation sensor but with a very poor spatial resolution (Government of Canada, 2009).

3 Workflow Method

After gathering field polygons for trees and shrubs, grass, and developed areas, these shapes were imported into shapefiles useable by ENVI software. These were then combined with provided training polygons, and then used in the classification functions: Maximum Likelihood and Minimum Distance. The first functions by assuming the statistics of each class given are normally distributed, and calculates the probability that a given pixel belongs to a specific class (Harris, 2016). Minimum distance functions by determining the Euclidean (straight line) distance from unknown points to classified points, classifying each pixel by its nearest class, unless a standard deviation threshold is provided; where some pixels might not be classified.

3.1 Provided Imagery Workflow

For the imagery provided, the classifications proceeded smoothly with mostly default values used for both the Minimum Distance and Maximum Likelihood Classification Methods.

3.2 Acquired Imagery Workflow

For the imagery we gathered (SPOT5) the classification did not proceed smoothly; the accuracy of the produced classifications was poorer than with the provided imagery. This is likely due to the cell size of the pixels in the acquired imagery: larger pixels lead to fewer pixels to train the classification algorithms. The default values for the classification functions were used as well.

3.3 Finishing Touches

After the classification raster files were created, they were then checked for accuracy with Confusion Matrices, using Test-points provided, providing the classification results in the next section.

Once the error matrices were produced, the four classification raster files produced were resized to the extent of the study area, and stored separately.

These raster files were then used to produce the area calculations that each imagery and classification method yielded – this was done by multiplying the resulting percent land use (Statistics View in ENVI), and multiplying it by the total area of the study area in hectares.

A map layout was then created in ArcMap 10.3 to provide WSLT with information on the land use in their study area, near Waiparous. Maximum likelihood classification with the provided imagery was used, since it provided the most accurate results, as seen in the next section.

4 Results Analysis

Table 1 - Confusion Matrix for Minimum Distance Classification of the Provided Imagery

Confusion Matrix: ~\ENVI_Outputs\WAIP_pro_Mind

Overall Accuracy = (64/69) 92.7536%
Kappa Coefficient = 0.9012

Class	Ground Truth (Pixels) Test_Dev	Test_Grass	Test_TreeShrub	TestPoint_Water	Total
Unclassified	0	0	0	0	0
Merged_Dev	15	0	0	2	17
Merged_Grass	2	15	0	0	17
Merged_TreeSh	0	0	24	0	24
Merged_Water	1	0	0	10	11
Total	18	15	24	12	69

Table 2 - Confusion Matrix for the Maximum Likelihood Classification of Provided Imagery

Confusion Matrix: ~\ENVI_Outputs\Waip_pro_Maxl

Overall Accuracy = (64/69) 92.7536%
Kappa Coefficient = 0.9015

Class	Ground Truth (Pixels) Test_Dev	Test_Grass	Test_TreeShrub	TestPoint_Wat	Total
Unclassified	0	0	0	0	0
Merged_Dev	15	0	1	1	17
Merged_Grass	2	15	0	0	17
Merged_TreeSh	0	0	23	0	23
Merged_Water	1	0	0	11	12
Total	18	15	24	12	69

Table 3 - Confusion Matrix for Maximum Likelihood of Acquired Imagery

Confusion Matrix: ~\ENVI_Outputs\Waip_acq_Maxl

Overall Accuracy = (42/47) 89.3617%
Kappa Coefficient = 0.8276

Class	Ground Truth (Pixels) Test_Dev	Test_Grass	Test_TreeShrub	TestPoint_Wat	Total
Unclassified	0	0	0	0	0
Merged_Dev	0	0	0	0	0
Merged_Grass	0	8	0	0	8
Merged_TreeSh	4	0	24	0	28
Merged_Water	1	0	0	10	11
Total	5	8	24	10	47

Table 4 - Confusion Matrix for Minimum Distance of Acquired Imagery

Confusion Matrix: ~\ENVI_Outputs\Waip_acq_Mind

Overall Accuracy = (47/69) 68.1159%
Kappa Coefficient = 0.5690

Class	Ground Truth (Pixels) Test_Dev	Test_Grass	Test_TreeShrub	TestPoint_Wat	Total
Unclassified	0	0	0	0	0
Merged_Dev	1	1	1	3	6
Merged_Grass	3	14	0	0	17
Merged_TreeSh	4	0	23	0	27
Merged_Water	10	0	0	9	19
Total	18	15	24	12	69

5 Classification Results

Table 5 - Area of Land Use Types, by Imagery and Classification Method (Maximum Likelihood & Minimum Distance)

Imagery & Classification Method	Developed Area (ha)	Grasslands Area (ha)	Tree/Shrub Area (ha)	Water Area (ha)	Total Land Use (ha)
Provided, Maximum Likelihood	17.0	207.0	453.4	0.3	**677.8**
Provided, Minimum Distance	6.6	260.3	410.9	0.0	**677.8**
Acquired, Maximum Likelihood	293.7	110.9	261.1	12.1	**677.8**
Acquired, Minimum Distance	163.1	262.1	243.6	9.0	**677.8**

6 Summary

The study area's land use was quite varied, and, as such, supervised image classification functions were very useful to save time. Yet, the accuracy varied depending on the imagery source, resolution, and the classification method used. Overall, maximum likelihood classification was more accurate than minimum distance; this is probably due to the wide variety of different land use within a small area, causing confusion (error) in the resulting 'minimum distance' raster files. To contrast, maximum likelihood fared much better: its statistical approach to determining classifications, based on training pixels, is better able to predict what unknown pixels might be – this resulted in significant improvements over minimum distance classifications.

The resolution of imagery made a big difference as well: since the captured training polygons were relatively small – having more pixels in a particular training shape improved the accuracy of results dramatically. This can be seen when contrasting the classification results between the provided imagery, and the acquired imagery – the provided imagery had much better resolution, leading to a significant increase in useable pixels to train the algorithms with.

Regarding the provided imagery's classification: the accuracy assessment was within acceptable standards. Yet, due to cloud cover in the imagery used, the results were lower accuracy than if imagery was used without clouds. In both of the classifications made using the provided imagery, the cloud cover shows up as the developed class. The cloud cover was larger and more prominent in the Maximum Likelihood Classification than the Minimum Distance Classification method.

For future imagery classifications, a few method improvements could improve the accuracy of results. Capturing larger polygons, for instance, would provide more information for the classification functions, likely improving the accuracy of the results. Also, being mindful of what the satellite is able to view from orbit, as opposed to a human with a GPS unit on foot; since capturing accurate training data is vital to producing high-quality results.

7 References

Government of Canada. (2009). Open Data Portal. Retrieved from: http://open.canada.ca/data/en/dataset?keywords=spot+image on Jun 8, 2016.

Harris Geospatial. (2016). Minimum Distance Classification. Retrieved from: https://www.harrisgeospatial.com/docs/minimumdistance.html on Jun 7, 2016.

Harris Geospatial. (2016). Maximum Likelihood. Retrieved from: https://www.harrisgeospatial.com/docs/maximumlikelihood.html on Jun 7, 2016.

SAIT Polytechnic. (2016). Written class notes from GEOS 451.

Satellite Imaging Corporation. (2015). SPOT-6 Satellite Sensor. Retrieved from: http://www.satimagingcorp.com/satellite-sensors/spot-6/ on Jun 7, 2016.

WSLT (Western Sky Land Trust). (2015). Conserving Our Sense of Place – Western Sky Land Trust. Retrieved from: http://westernskylandtrust.ca/ on Jun 6, 2016.