Image Segmentation of Acme, WA Using the Orfeo Toolbox and Classification with R and ENVI

An Analysis by Skyler Elmstrom for ESCI 442

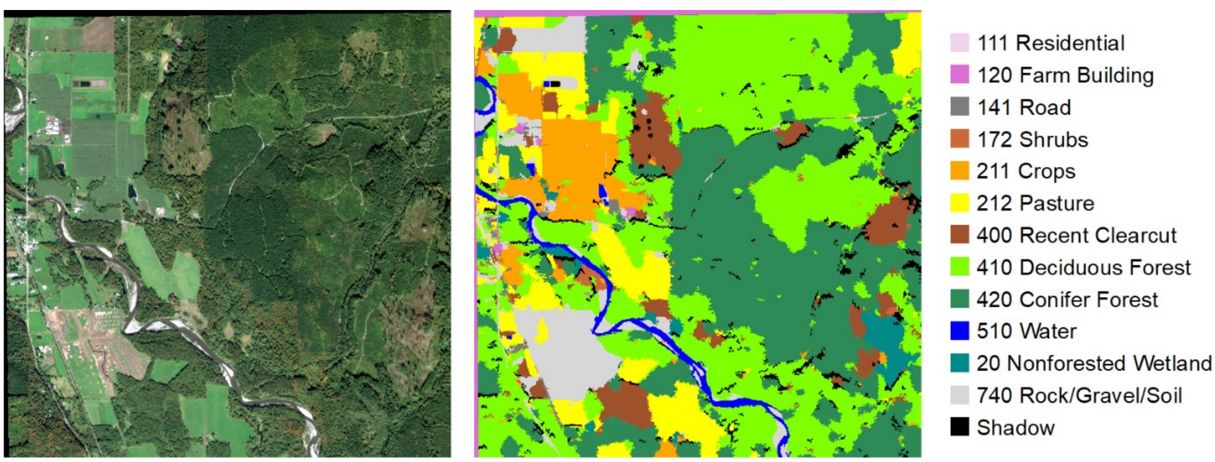

Figure 1. True-color composite of Acme, WA (left) and an image classification created using image segmentation in Orfeo Toolbox, classified using K-means clustering in R, and identified in ENVI (right). The classified image consists of 12 information classes plus shadow. The parameters of the segmentation follow the guidance of Wallin (2018): spatial radius of 23, range radius of 285, and minimum region size of 40.

Abstract

High spatial resolution imagery requires some preprocessing for extracting specific information of value; pixel-based classification techniques used in low to medium spatial resolution are often inadequate for classifying land use and land cover (LULC) in high resolution imagery. In this example, image segmentation – object-based grouping of pixels by comparing the spectral and spatial properties of surrounding pixels— is applied to high resolution imagery prior to classification. I achieved an overall accuracy of 51.9% for 12 information classes. This technique and the parameters used performed well for identifying features with regular geometry such as crops and pastures but struggled with natural vegetation types that blend between each other. Additional refinements to the size of segments during segmentation and improving my qualitative differentiation of LULC types should allow me to make moderate improvements to my accuracy for roads, buildings, forests, and wetlands.

Methods

My methods follow the procedure established by Wallin (2018). I ran a Large-Scale Mean-Shift (LSMS) segmentation in Orfeo Toolbox (OTB) on a WorldView-2 image of Acme, WA using the updated segmentation parameters added on February 21, 2018. The vectorized output values of the segmentation were converted to z-scores and classified in RStudio using K-means clustering. Finally, the classified segments were assigned to information classes in ENVI and an accuracy assessment was performed using ground-truth data provided by Wallin.

Results

I completed the four-step LSMS segmentation[1] provided in OTB as described by Wallin (2018) – the results are described in figure 2. Following segmentation and classification into twelve Land Use/Land Cover (LULC) codes, I created a color-coded, classified raster image to visualize my results (figure 1). Overall, my classification corroborates trends observed in the original image: this scene of Acme, WA is primarily forested area and pasture/crop lands –about 301 acres or 78.2% of the image area— with sparse urban development (table 1). When comparing my classification to ground-truth data provided by Wallin, I achieved an overall accuracy percentage of 51.9%. Crops, pasture, and rock/gravel/soil LULC types were the most accurately identified with production accuracies greater than 85% while farm buildings, residential areas, roads, and wetlands had production accuracies below 38% (table 2). There also appears to have been some confusion identifying several similar LULC types such as shrubs, clear-cuts, and conifer forest with deciduous forest, wetlands with clear-cuts, and roads with rock/gravel/soil (table 2).

Discussion

My classification accuracy of 51.9% is similar to my pixel-based classification accuracies of LANDSAT images; I typically achieved 53-57% accuracy with LANDSAT scenes before combining classes or conducting post-processing. Despite the similarity in accuracy, I expected my high-resolution image to have more accurate results given that I was classifying data into similar numbers of information classes. It is worth noting that all my segments were larger than the 30-meter x 30-meter pixel area of LANDSAT imagery. Image segmentation introduces several different sources of error compared to purely pixel-based classification. For example,

[1] Information on LSMS segmentation and the additional steps performed can be found at https://www.orfeo-toolbox.org/CookBook/Applications/app_LargeScaleMeanShift.html

segmentation often created fuzzy boundaries between observable areas of deciduous and conifer forest or between road, residential, and rock or soil. This often meant that features of differing information classes were lumped together. These segments likely had similar spectral and spatial properties: roads are made of similar material as rock and are flat like cleared bare ground areas; mixed forest can appear to be like both conifer and deciduous forest; farm buildings may appear spectrally similar to residential areas and spatially similar to residential areas when found in close proximity. Another source of confusion likely occurred during my assignment of information classes due to my lack of experience identifying LULC types, particularly wetlands. The Acme, WA scene is a summer image which adds to the difficulty of accurately identifying wetlands. My image segmentation excelled in the differentiation of crops versus pasture; nearly all the crops and pasture were correctly classified according to the ground-truth data (table 2). My previous attempts using pixel-based classification on LANDSAT imagery for identifying crops and pasture were poor due to the likelihood of mixed pixels between the two LULC types. An increase in spatial resolution from 30 meters to less than 1 meter separated the two types very well. This may suggest that high resolution imagery outperforms mid to low resolution imagery in classifying features with unnatural geometry such as manicured farmlands, clearings, and built environments. One feature that did not benefit from the increase in spatial resolution in my analysis was recent forest clear-cuts. However, this was not usually due to misidentification as I have seen with pixel-based classification of LANDSAT scenes, but rather from confusion between clear-cuts and other forest classes, suggesting that some of my scenes clear-cuts got lumped in with forest cover types. Decreasing the size of the segments created during image segmentation may help improve my clear-cut class accuracy. Mixed-forest was also problematic for my classification accuracies of forest types, particularly for conifer forest where half of the ground truth conifer forests areas were misclassified as deciduous. In future analyses of high-resolution imagery, I will consider adding a new forest information class for mixed-forest as I was able to see them clearly in the original image being analyzed. If this classification were to be repeated, special attention should be paid to residential areas, roads, and wetlands as they were more likely to be misidentified than to be confused with a similar information class (table 2).

Conclusion

My image segmentation produced moderately accurate results that were comparable to pixel-based classification techniques. However, there is significant room for improvement to reduce confusion between similar information classes and to correct misidentified classes. Reducing the size of my segments should allow me to reduce the confusion in forest types and improving my personal skill in identifying shrubs and wetlands should improve misidentification in those two classes but I expect to see only modest improvements in my classification accuracy without combining similar information classes. Generally, image segmentation performed well identifying regular geometric shapes and spectrally similar areas such as crops and pasture but less well with forest types and clear-cuts following the parameters used by Wallin (2018).

Sources

Wallin, D. 2018. *Lab V: Image Segmentation with Monteverdi, the Orfeo Toolbox and ENVI: WorldView-2, 8-band image of Acme, WA*. Retrieved from http://myweb.facstaff.wwu.edu/wallin/envr442/ENVI/442_segmentation_ENVI__Orfeo_acme4.htm. Accessed on 3/7/2018.

Figures and Tables

Table 1. LULC class and total area derived from image segmentation of Acme, WA.

LULC Type	# Pixels	Acres	Percent Area
Deciduous	1447510	132.9	34.51
Conifer	1224052	112.4	29.18
Pasture	377495	34.7	9.00
Rock/Gravel/Soil	278680	25.6	6.64
Crops	231152	21.2	5.51
Clearcut	230017	21.1	5.48
Farm Buildings	120520	11.1	2.87
Shadow	117653	10.8	2.81
Wetland	66920	6.1	1.60
Water	63475	5.8	1.51
Road	18950	1.7	0.45
Shrubs	17880	1.6	0.43
Residential	0	0.0	0
Total	4194304	**385.2**	100

Table 2. Confusion Matrix for Image Segmentation and Classification of Acme, WA. Columns show ground-truth classes and the number of ground truth points in each predicted class. The number of Ground-truth data matching the predicted class are indicated by the shaded diagonal region and are used to calculate the overall accuracy. Values near this line indicate confusion between spectrally/spatially similar classes whereas values farther away likely indicate misidentification.

	Ground Truth (Reference Data)												**Row Total**	User Acc (%)
Predicted Class	Residential	Farm Bldg	Road	Shrubs	Crops	Pasture	Clearcut	Deciduous	Conifer	Water	Wetland	Rock/Gravel		
111 Residential	0	0	0	0	0	0	0	0	0	0	0	0	0	0
120 Farm Bldg	0	3	0	0	0	0	0	0	1	0	1	0	5	60
141 Road	1	3	0	0	0	0	0	0	0	0	0	0	4	0
172 Shrubs	0	0	0	0	0	0	0	0	0	0	0	0	0	0
211 Crops	2	0	1	0	15	0	0	0	0	0	0	0	18	83.33
212 Pasture	0	0	0	0	0	16	0	0	0	0	0	0	16	100
400 Clearcut	0	0	0	0	0	0	5	0	0	0	8	0	13	38.46
410 Deciduous	0	0	1	8	0	0	13	14	15	1	0	1	53	26.42
420 Conifer	0	0	0	0	0	0	2	5	15	0	0	0	22	68.18
510 Water	0	0	0	0	0	0	0	0	0	7	0	0	7	100
20 Wetland	0	0	0	0	0	0	0	0	0	0	0	0	0	0
740 Rock/Gravel	1	2	7	0	1	0	0	0	0	1	0	6	18	33.33
Column Total	4	8	9	8	16	16	20	19	31	9	9	7	156	**Total Samples**
													81	**Total # Correct**
Prod Acc (%)	0	37.5	0	0	93.8	100	25	73.68	48.39	77.78	0	85.71	51.9	**Overall Accuracy**

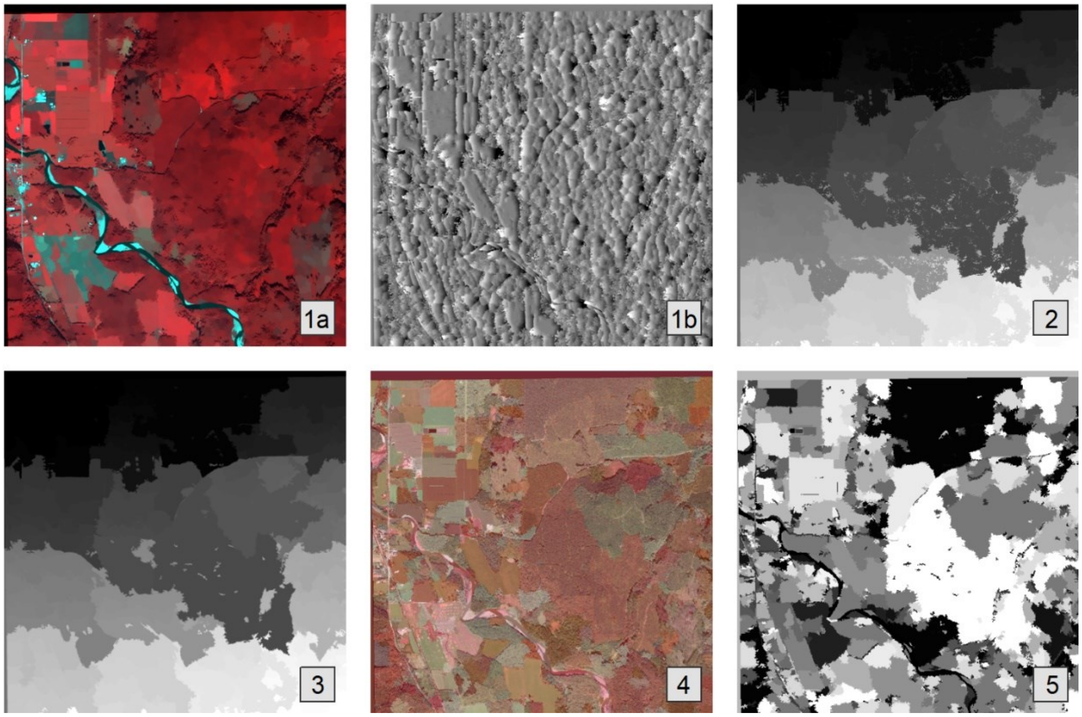

Figure 2. Orfeo Toolbox output during four-step LSMS segmentation and final segmented raster. From top-left: step 1 filtered output showing spectral groupings, step 1 spatial output showing spatially similar groupings, step 2 segmentation based on spectral and spatial groupings, step 3 regions merging smooths and combines segments, step 4 vectorization creates vector polygons based on segmentation and is shown overlaid with the true color image of the study area, and finally the resulting classification-ready raster after polygon to raster conversion in ArcMap.