Feature Extraction with Rule-Based Classification Tutorial

In this tutorial, you will use Feature Extraction to extract asphalt surfaces from a multispectral GeoEye-1 scene of a residential area in Hobart, Tasmania. Feature Extraction provides a quick, automated method for identifying roads and parking lots, saving an urban planner or GIS technician from digitizing them by hand. You will learn how various attributes can help you build meaningful rules in classification, and you will export your classification results to a shapefile. If you need more information about a particular step, click the blue Help button to access ENVI® Help.

Files Used in This Tutorial

Tutorial files are available from our website or on the ENVIResource DVD in the `feature_extraction` directory. You will use the file named `Hobart_GeoEye_pansharp.dat` for this tutorial.

This is a pan-sharpened GeoEye-1 image (0.5-meter spatial resolution) of the Dowsing Point suburb of Hobart, Tasmania, acquired on 05 February 2009. GeoEye-1 images are courtesy of DigitalGlobe.

Background

Feature Extraction uses an object-based approach to classify imagery, where an object (also called a *segment*) is a group of pixels with similar spectral, spatial, and/or texture attributes. Traditional classification methods are pixel-based, meaning that spectral information in each pixel is used to classify imagery. With high-resolution panchromatic or multispectral imagery, an object-based method offers more flexibility in feature extraction types.

The workflow involves dividing an image into many segments, computing various attributes for the segments, then building rules to classify features of interest. Each rule contains one or more attributes such as area, length, or texture, which you constrain to a specific range of values. For example, we know that roads are elongated, some buildings approximate a rectangular shape, and trees are highly textured compared to grass.

Start the Workflow

1. From the menu bar, select **File > Open**.

2. Navigate to `feature_extraction`, and select the file `Hobart_GeoEye_pansharp.dat`. Click **Open**. The image is displayed at full resolution.

3. From the **Optimized Linear** drop-down list in the toolbar, select **Linear 5%**. This type of stretch brightens the image, making it easier to see individual features.

4. From the Toolbox, select **Feature Extraction > Rule Based Feature Extraction Workflow**. The Data Selection panel appears.

5. The filename is already listed in the **Raster File** field. Click **Next**. The Object Creation panel appears.

Segment the Image

Segmentation is the process of dividing an image into segments that have similar spectral, spatial, and/or texture characteristics. The segments in the image ideally correspond to real-world features. Effective segmentation ensures accurate classification.

1. Enable the **Preview** option. A Preview Window appears, showing the initial segments from the image, colored in green. The following image (zoomed to 200%) shows an example of a Preview Window centered over some asphalt roads:
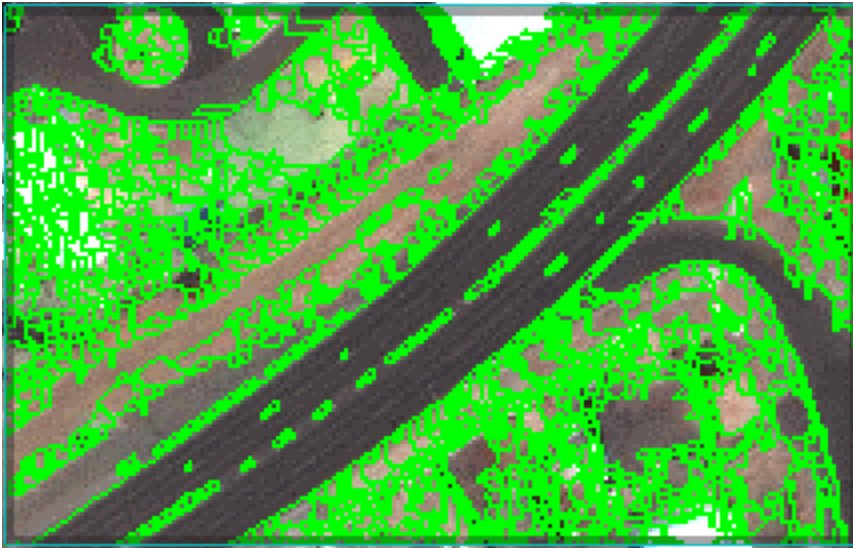
The initial segmentation shown in the Preview Window sharply delineates the edes of the roads, but there are too many small segments comprising each road. Use a combination of segmentation and merging to delineate the road boundaries with as few segments as possible. A general rule is to decrease the **Scale Level** (thus creating more segments) while increasing the **Merge Level** to merge adjacent segments. It is better to have more segments to delineate your features of interest, but realize that too many segments will result in slower processing.

2. Experiment with the different **Algorithm** options under **Segment Settings** and **Merge Settings**, and change the **Level** for each. The Preview Window automatically updates to show how your changes affect the segmentation.

You can achieve reasonable results with the following settings:

- Segment Algorithm: **Edge**
- Scale Level: **40**
- Merge Algorithm: **Full Lambda Schedule**
- Merge Level: **80**
- Click the **Select Segment Bands** button ⬚, and select only Band 4. Performing segmentation only on the near-infrared band produces sharper segments on asphalt roads.

The following image shows the resulting segmentation over a road intersection using these settings:
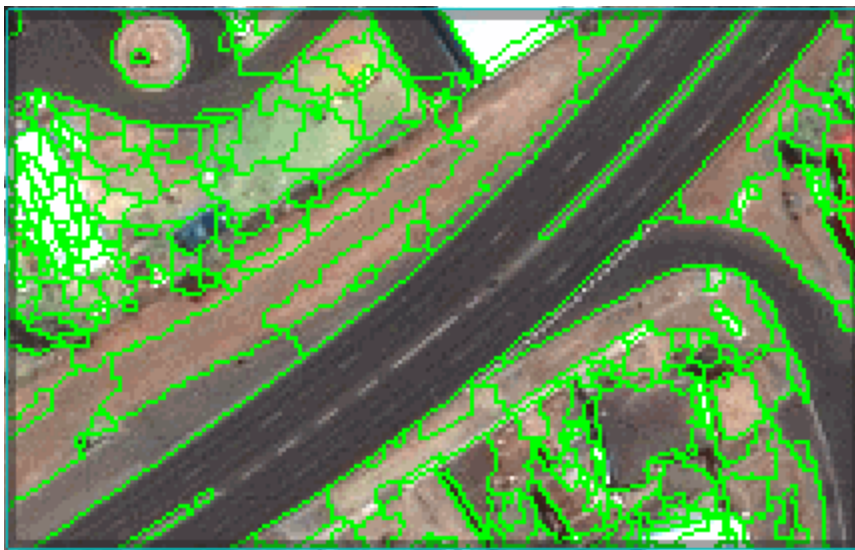
3. When you are satisfied with the segmentation, click **Next**. ENVI creates and displays a segmentation image (also called the Region Means image) in the Layer Manager. Each segment is assigned the mean spectral values of all the pixels that belong to that segment.

Build Rules

In this step, you will build several rules to define an Asphalt class. For each segment, ENVI computes various spatial, spectral, and texture attributes. See "List of Attributes" at the end of this tutorial for definitions of all available attributes.

To decide which attributes to use for identifying asphalt surfaces, consider some characteristics of asphalt in a remote sensing image. Here are some ideas:

- Asphalt roads appear very dark in a near-infrared image, so low spectral mean values in Band 4 should highlight the roads.

- Roads are elongated.

- Roads and parking lots have a uniform texture.

- Smaller features such as rooftops and shadows can also have low spectral values, so you can use an Area attribute to filter out these small segments.

The typical workflow for building rules is to begin with one attribute, test its confidence in extracting your feature of interest, then use more conditions and attributes to filter out all other features from the scene so that you are left only with your feature of interest.

When you combine multiple attributes for a given rule, *all* of the conditions must apply for the rule to be satisfied.

The Band 4 spectral mean is a good attribute to start with, since most of the roads are very dark in this band.

Spectral Mean

1. Disable the **Preview** option.

2. In the Rule-based Classification panel, click the **Add Class** button ✚.

3. In the Class Properties table, change the **Class Name** to **Asphalt** and press the **Enter** key.

4. A default rule set was created when you added the new class, with one attribute: `Spectral Mean(Band 1) > 0`. You need to change this from Band 1 to Band 4. Select this attribute name, then select **Band 4** from the **Band** drop-down list.

The cyan-colored histogram shows the frequency of Spectral Mean (Band 4) values for each segment in the segmentation image.
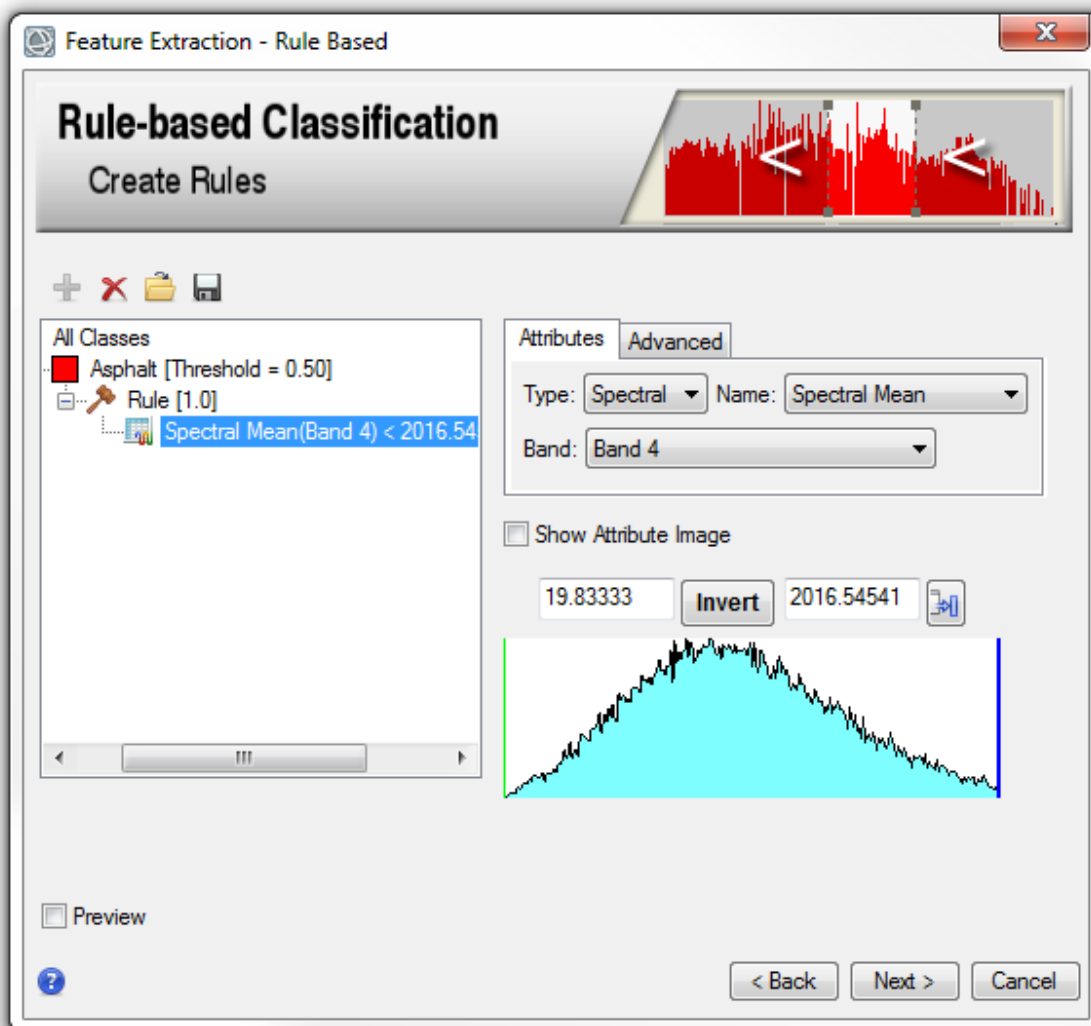
5. Click the **Dock/Undock Histogram Window** button to display the histogram in a separate dialog. This option is only available for Windows users.

6. Click and drag a corner of the Attribute Histogram window to enlarge it, so that you can better view the shape of the histogram.
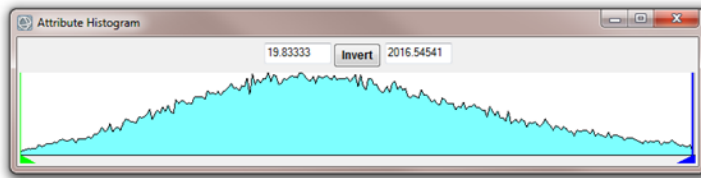
7. Enable the **Preview** option in the Rule-based Classification panel. A Preview Window appears, with a solid-red color.

8. Zoom out to approximately 50% using the **Fixed Zoom Out** button in the toolbar, so that you can view more of the image in the display.

9. Since the roads have a low spectral mean, you want to evaluate only the lowest values in the histogram. Click the blue triangle in the lower-right corner of the Attribute Histogram dialog, and drag the blue vertical bar to the left. The Preview Window will update to show the segments that remain. The following image shows an example:
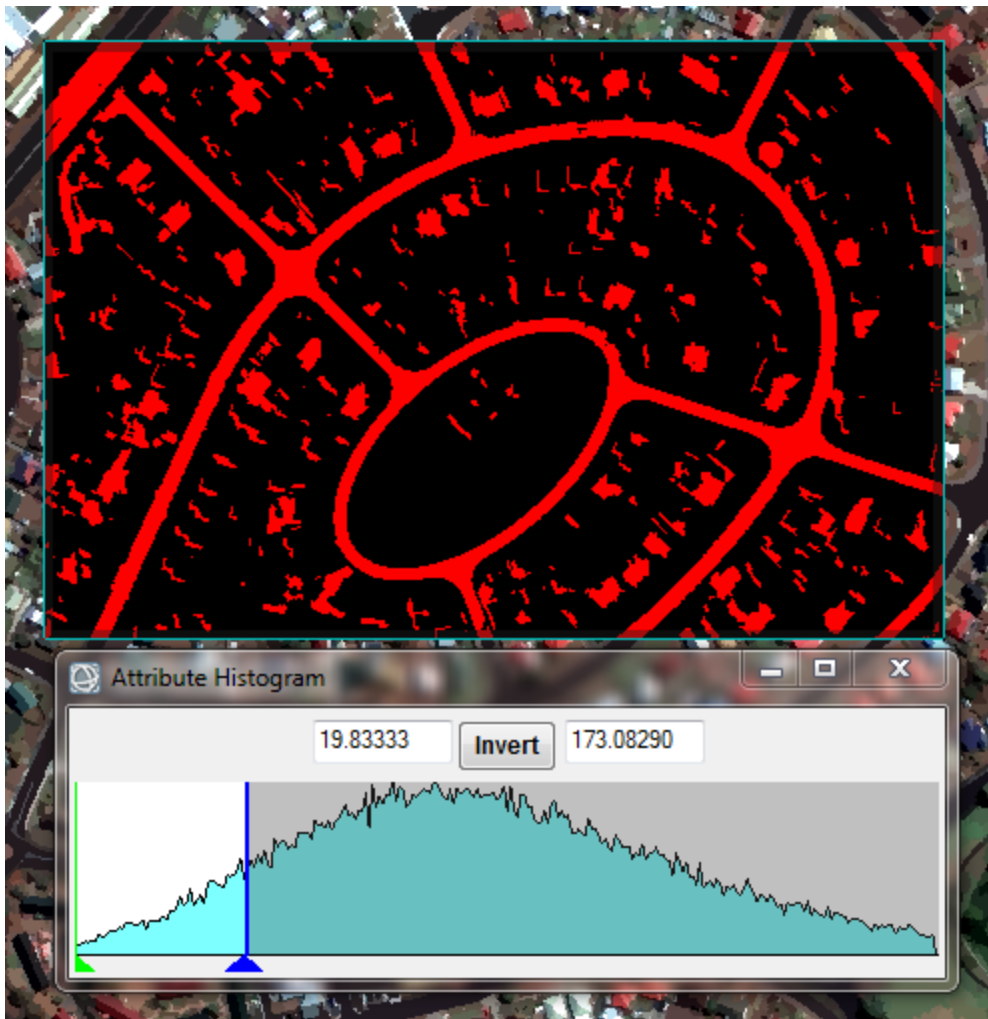
Constraining the Band 4 spectral mean values to less than 174 seems to effectively highlight the asphalt roads. (The slider only allows fractional values; type **174** into the right-most field and press **Enter**.)

10. Move the Preview Window to the lower-right corner of the image over the water:
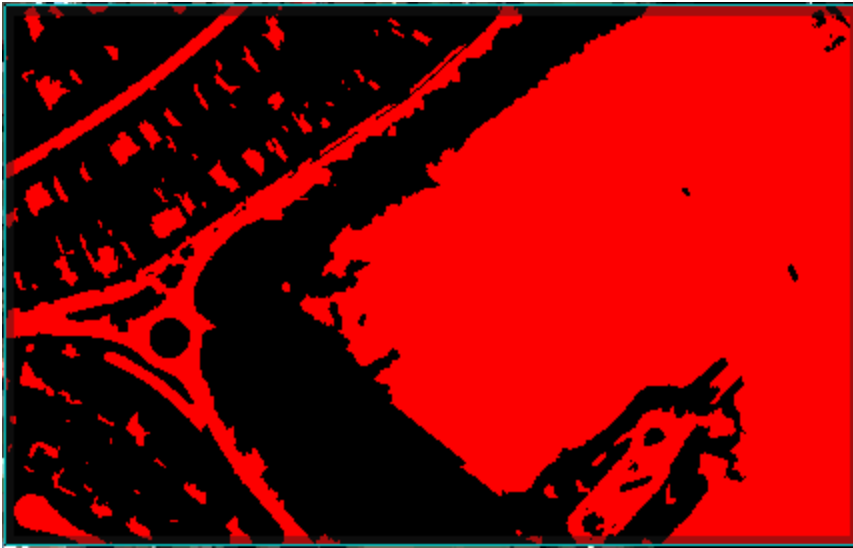
11. The large red area in the right side of the Preview Window is water, which has an even lower spectral mean in Band 4 than the asphalt. So, water is also included in this histogram range. You want to exclude these extremely low values. A minimum value of **85** seems to reasonably exclude the water from consideration, so type **85** into the left-most field of the Attribute Histogram dialog, and press **Enter**.

12. Click the red X button in the Attribute Histogram dialog to close it.

Major Length

The roads and most of the parking lots in this image are lengthier than other features such as houses, so you should test one of the length attributes to see if it improves your rule set.

1. In the Rule-based Classification panel, select the rule name **Rule [1.0]**.

2. Click the **Add Attribute To Rule** button.

3. Select the default **Spectral Mean(1)** attribute that was added.

4. From the **Type** drop-down list, select **Spatial**.

5. From the **Name** drop-down list, notice that there are three different attributes related to length: Length, Elongation, and Major Length. If you are unsure which attribute to test, an attribute image can help you:

6. Select **Length**, and enable the **Show Attribute Image** option. The Preview Window displays a greyscale image of Length attribute values, with the highest values shown in

white. If your feature of interest shows a high contrast (either bright or dark) relative to other features, then that attribute may be a good choice.

7. Move the Preview Window around the image. Notice how some of the roads have high Length attribute values (indicated by white), but other roads—and most parking lots— have lower values (indicated by grey). So this particular attribute may not effectively identify all of the asphalt.
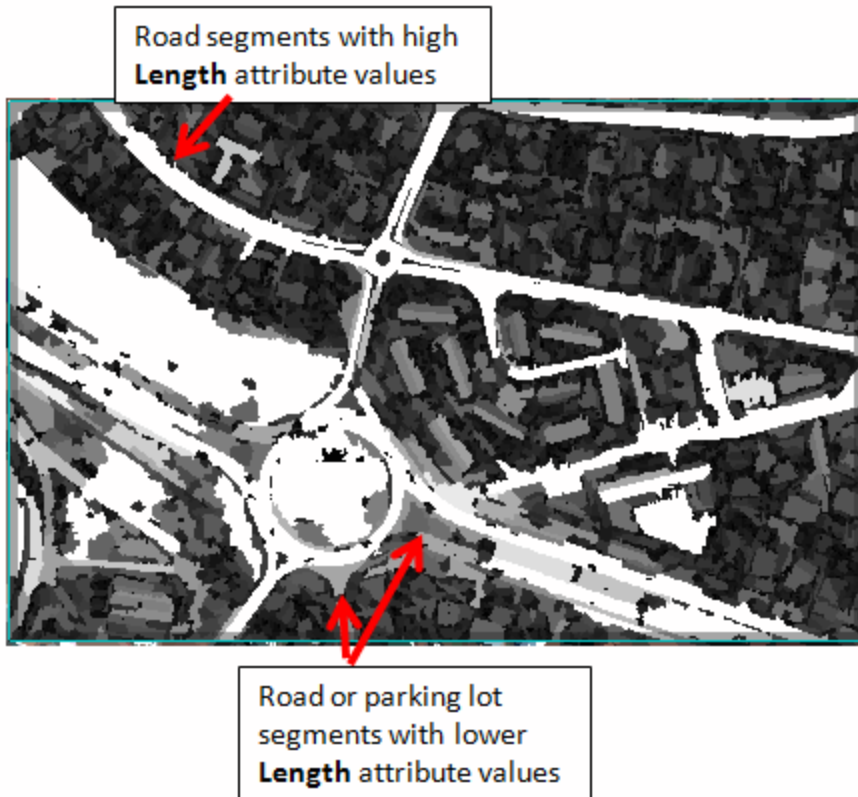
Tip: You can check the actual attribute values for these segments by clicking the **Cursor Value** button 📍 in the main toolbar. In the Cursor Value dialog, look for the "Data" line underneath `Preview Window: FX_CLASSIFY_RULE_STEP`. This is the Length attribute value for the segment underneath your cursor. Use these values to guide how much you want to limit the histogram.

8. Try a different length-related attribute. From the **Name** drop-down list, select **Major Length** and enable the **Show Attribute Image** option. More of the roads and parking lots should have a slightly higher contrast than surrounding features, compared to the Length attribute.

9. Uncheck the **Show Attribute Image** option.

10. Since roads are lengthier than other features, you want to constrain the histogram to the highest Major Length values. Click the green diamond in the histogram, and drag the green vertical slider toward the right. Many different features are still included (shown in red) in the Preview Window. To filter out these features, you must constrain the histogram to include only extremely high values.

11. Leave the maximum value as-is, but change the minimum value to **22** in the Attribute Histogram dialog, and press **Enter**.

12. To make the following steps quicker, disable the **Preview** option.

Texture Mean

Asphalt roads and parking lots generally have a uniform texture, so you can experiment with the Texture attributes to see if any of them work well in extracting these features. Texure Range, Texture Mean, Texture Variance, and Texture Entropy are the four texture attributes that are calculated for every band in the image.

Prior tests show that the Texture Mean attribute for Band 4 (near-infrared) works well in identifying asphalt because it shows a high contrast in road and parking lot segments, relative to surrounding features.

1. In the Rule-based Classification panel, select the rule name **Rule [1.0]**.

2. Click the **Add Attribute To Rule** button.

3. Select the default **Spectral Mean(1)** attribute that was added.

4. From the **Type** drop-down list, select **Texture**.

5. From the **Band** drop-down list, select **4**.

6. From the **Name** drop-down list, select **Texture Mean**.

7. You want to identify segments with the lowest Texture Mean Band 4 values. A maximum value of 155 effectively identifies these segments, so enter **155** in the right-most field of the histogram and press **Enter** on your keyboard.

Area

So far, your rule consists of the following attributes:

- **85** < Band 4 Spectral Mean < **174**, AND
- Major Length > **22**, AND
- Texture Mean < **155**

Next, you will use the Area attribute to filter out smaller features that remain such as shadows and rooftops.
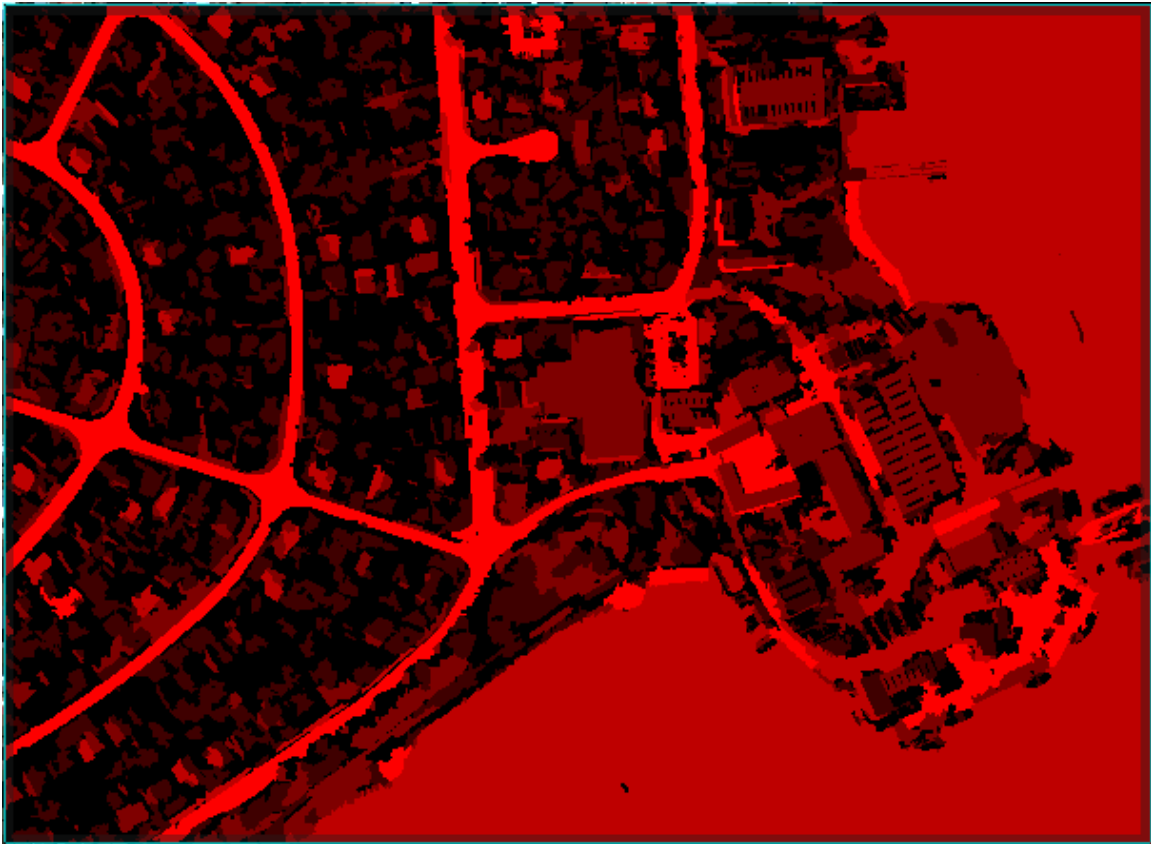
1. In the Rule-based Classification panel, select the rule name **Rule [1.0]**.

2. Click the **Add Attribute To Rule** button .

3. Select the default **Spectral Mean(1)** attribute that was added.

4. From the **Type** drop-down list, select **Spatial**.

5. From the **Name** drop-down list, select **Area**.

6. You want to identify segments with higher Area values. A minimum value of 40 effectively identifies these segments, so enter **40** in the left-most field of the histogram and press **Enter** on your keyboard.

Preview the Classification

Now that you have built a rule set using four different attributes, you can preview how well the rule identifies asphalt surfaces.

1. In the Rule-based Classification panel, select the rule name **Rule [1.0]**.

2. Enable the **Preview** option. The Preview Window shows a *rule confidence image*. The brightest segments have a high likelihood of satisfying the conditions of your rule, based on the attributes you selected. Dark or black features have little or no likelihood of meeting the conditions. In the following example, the roads and parking lots are bright, which means the rule set you built will most likely classify these segments as Asphalt.

3. On the left side of the Rule-based Classification panel, select **All Classes**. The Preview Window shows a classification image with only one class (Asphalt). This allows you to preview the final classification in a small area before classifying the entire image.

4. Click once inside the Preview Window, then lower its transparency using the Transparency slider in the main toolbar. This lets you preview the Asphalt class over the Region Means image.

5. Click the **Save Rule Set** button 💾 in the Rule-based Classification panel.

6. Select an output location and filename for the rule set. You can restore this rule file later if you want to continue where you left off.

7. Click **Next** to proceed to the Output panel.

Set Output Options

The Output panel lets you export various components of the rule classification to raster images and/or shapefiles. For this tutorial, you will only export a classification image of the Asphalt class to ENVI raster format, then overlay the class on the original image.

1. In the Export Vector tab, disable the **Export Classification Vectors** option.

2. Click the Export Raster tab.

3. Enable the **Export Classification Image** option, and select an output directory to save the file (in ENVI raster format). Change the output filename to `asphalt_class`.

4. Click **Finish**. ENVI adds a new layer called `asphalt_class.dat` to the Layer Manager and displays it over the original GeoEye-1 image.

5. Use the Pan tool to navigate around the image and evaluate the asphalt class.

The rule set did a fairly good job of extracting asphalt, but like any automated method, it did not extract every segment that represents asphalt. Some roads or parking lots were mixed with dirt, or they displayed much brighter signatures than other similar features. Successful results depend heavily on good segmentation, followed by effective rule building. If you are not satisfied with your results, try experimenting with different segmentation and merging values, or try a different set of attributes.

List of Attributes

Spectral Attributes

Spectral attributes are computed on each band of the input image. The attribute value for a particular segment is computed from input data band where the segmentation label image has the same value (i.e., all pixels in the same segment contribute to the attribute calculation).

Attribute	Description
Spectral_Mean	Mean value of the pixels comprising the region in band x
Spectral_Max	Maximum value of the pixels comprising the region in band x
Spectral_Min	Minimum value of the pixels comprising the region in band x
Spectral_STD	Standard deviation value of the pixels comprising the region in band x

Texture Attributes

Texture attributes are computed on each band of the input image. Texture attribute computation is a two-step process where the first pass applies a square kernel of pre-defined size to the input image band. The attributes are calculated for all pixels in the kernel window and the result is referenced to the center kernel pixel. Next, the attribute results are averaged across each pixel in the segment to create the attribute value for that band's segmentation label.

Attribute	Description
Texture_Range	Average data range of the pixels comprising the region inside the kernel (whose size you specify with the Texture Kernel Size parameter in segmentation)
Texture_Mean	Average value of the pixels comprising the region inside the kernel
Texture_Variance	Average variance of the pixels comprising the region inside the kernel
Texture_Entropy	Average entropy value of the pixels comprising the region inside the kernel

Spatial Attributes

Spatial attributes are computed from the polygon defining the boundary of the segment, so band information is not required.

Attribute	Description
Area	Total area of the polygon, minus the area of the holes.
	If the input image is pixel-based, the area is the number of pixels in the segmented object. For a segmented object with 20 x 20 pixels the area is 400 pixels.
	If the input image is georeferenced, the area is in the map units of the input image. For a segmented object with 20 x 20 pixels, where the input image pixel resolution is 2 meters, the total area is 1600 square meters (400 pixels x 2 meters x 2 meters).
Length	The combined length of all boundaries of the polygon, including the boundaries of the holes. This is different than the Major_Length attribute.
	If the input image is pixel-based, the length is the number of pixels. For a segmented object with 20 x 20 pixels, the length is 80 pixels.
	If the input image is georeferenced, the length is in the map units of the input image. For a segmented object with 20 x 20 pixels, where the input image pixel resolution is 2 meters, the length is 160 meters (80 pixels x 2 meters).
Compactness	A shape measure that indicates the compactness of the polygon. A circle is the most compact shape with a value of 1 / pi. The compactness value of a square is 1 / 2(sqrt(pi)).
	Compactness = Sqrt (4 * Area / pi) / outer contour length
Convexity	Polygons are either convex or concave. This attribute measures the convexity of the polygon. The convexity value for a convex polygon with no holes is 1.0, while the value for a concave polygon is less than 1.0.
	Convexity = length of convex hull / Length
Solidity	A shape measure that compares the area of the polygon to the area of a convex hull surrounding the polygon. The solidity value for a convex polygon with no holes is 1.0, and the value for a concave polygon is less than 1.0.
	Solidity = Area / area of convex hull

Attribute	Description
Roundness	A shape measure that compares the area of the polygon to the square of the maximum diameter of the polygon. The "maximum diameter" is the length of the major axis of an oriented bounding box enclosing the polygon. The roundness value for a circle is 1, and the value for a square is 4 / pi. Roundness = 4 * (Area) / (pi * Major_Length^2)
Form_Factor	A shape measure that compares the area of the polygon to the square of the total perimeter. The form factor value of a circle is 1, and the value of a square is pi / 4. Form_Factor = 4 * pi * (Area) / (total perimeter)^2
Elongation	A shape measure that indicates the ratio of the major axis of the polygon to the minor axis of the polygon. The major and minor axes are derived from an oriented bounding box containing the polygon. The elongation value for a square is 1.0, and the value for a rectangle is greater than 1.0. Elongation = Major_Length / Minor_Length
Rectangular_Fit	A shape measure that indicates how well the shape is described by a rectangle. This attribute compares the area of the polygon to the area of the oriented bounding box enclosing the polygon. The rectangular fit value for a rectangle is 1.0, and the value for a non-rectangular shape is less than 1.0. Rectangular_Fit = Area / (Major_Length * Minor_Length)
Main_Direction	The angle subtended by the major axis of the polygon and the x-axis in degrees. The main direction value ranges from 0 to 180 degrees. 90 degrees is North/South, and 0 to 180 degrees is East/West.
Major_Length	The length of the major axis of an oriented bounding box enclosing the polygon. Values are map units of the pixel size. If the image is not georeferenced, then pixel units are reported.
Minor_Length	The length of the minor axis of an oriented bounding box enclosing the polygon. Values are map units of the pixel size. If the image is not georeferenced, then pixel units are reported.
Number_of_Holes	The number of holes in the polygon. Integer value.
Hole_Area/Solid_Area	The ratio of the total area of the polygon to the area of the outer contour of the polygon. The hole solid ratio value for a polygon with no holes is 1.0. Hole_Area/Solid_Area = Area / outer contour area