Robust Terrain Classification of High Spatial Resolution Remote Sensing Data Employing Probabilistic Feature Fusion and Pixelwise Voting

R. Derek West[a], Brian J. Redman[a], David A. Yocky[a], John D. van der Laan[a], and Dylan Z. Anderson[a]

[a]Sandia National Laboratories, Albuquerque, New Mexico, USA

ABSTRACT

There are several factors that should be considered for robust terrain classification. We address the issue of high pixel-wise variability within terrain classes from remote sensing modalities, when the spatial resolution is less than one meter. Our proposed method segments an image into superpixels, makes terrain classification decisions on the pixels within each superpixel using the probabilistic feature fusion (PFF) classifier, then makes a superpixel-level terrain classification decision by the majority vote of the pixels within the superpixel. We show that this method leads to improved terrain classification decisions. We demonstrate our method on optical, hyperspectral, and polarimetric synthetic aperture radar data.

Keywords: hyperspectral imagery, optical imagery, polarimetric synthetic aperture radar, terrain classification, probabilistic feature fusion, ensemble voting

1. INTRODUCTION

Terrain classification, from remotely sensed data, has been a topic of research for many years. Researchers have explored supervised and unsupervised[1] classification techniques and explored the utility of a variety of remote sensing technologies, such as ladar,[2,3] polarimetric synthetic aperture radar (PolSAR),[4,5] and hyperspectral imaging (HSI).[6,7] Furthermore many different classifiers have been proposed with varying levels of successful classification.

There are several factors that should be considered for robust terrain classification. Some of the factors include, handling seasonal variations, selecting the appropriate spatial resolution of the imagery, and generalizing well to unseen test data. For high spatial resolution systems, there exists the unique challenge that much more spatial variability is captured within a class, which leads to significant pixel-wise variability and an increased chance of mis-classification.

In this paper, we address the issue of high pixel-wise variability within terrain classes from high spatial resolution optical, PolSAR, and HSI[8] remote sensing modalities. As imagery from all three imaging modalities achieve spatial ground spacing of less than a meter, obtaining robust classification of regions of interest within remote sensing imagery can be a challenging problem. Some of the reasons that make robust classification challenging include: classes are subject to intra-class variability, pixels may have multiple classes, sensor noise, variability of sensor parameters, and determining robust features to represent the data for each class. Furthermore, in the case of vegetation and terrain classification, robust classification could also include reducing the spatial variability, as vegetation and terrain typically do not vary rapidly. In addition, because of the high spatial variability, employing terrain classification techniques on a pixel-wise level can lead to many false alarms. We propose a robust framework for dealing with the pixelwise variability that is inherent in high spatial resolution remote sensing data and show that it can help to make better terrain classification decisions.

In this paper we discuss the different data sets and how we processed them. Then, we introduce our proposed framework where we utilize the simple linear iterative clustering (SLIC) segmentation algorithm to decompose the

Further author information: (Send correspondence to RDW)
RDW: E-mail: rdwest@sandia.gov, Telephone: 1 (505) 284-5568
SAND20??-????C

images into superpixels that delineate localized contrasting regions (the pixels within a superpixel typically belong to a single terrain class). Next, we briefly describe the training process for probabilistic feature fusion (PFF) one-class classifiers.[9] Then we describe the performance metrics for a limited number of trained terrain classes and give the terrain classification performance results from making decisions at the pixelwise and superpixel levels and compare the results to making decisions by pixelwise voting within a superpixel. We show that in general voting within a superpixel leads to better terrain classification.

2. SENSOR IMAGERY

Sandia National Laboratories (SNL) collected optical, HSI, and PolSAR imagery during March and June 2016. The sensors, airborne vehicles, data sets, and the signal and image pre-processing are described in this section.

2.1 Optical and Hyperspectral

Two optical sensors were deployed by an unmanned aerial system (UAS) to collect imagery. For HSI products, a linear pushbroom visible and near infrared (VNIR) Headwall Photonics Nano-Hyperspec hyperspectral imager with an onboard global positioning system (GPS) and inertial measurement unit (IMU) was used. This system offers over 270 spectral bands over the wavelength range, $\lambda \approx 400$ - 1000 nm. The VNIR HSI data were collected on June 16, 2016. Figure 2 (a) illustrates a color-composite HSI image of a scene.

The high spatial resolution optical imagery were collected with a commercial digital single-lens reflex (DSLR) camera with a standard Bayer RGB filter (Sony Alpha II), integrated with the onboard navigation system. The optical imagery were collected on June 14, 16, and 18 2016. Figure 2 (b) illustrates an optical image of a scene.

Sixteen ground control points (GCP) were established in field areas primarily at the peripheral extremities of the collection area. GCP panels were established in a checkerboard pattern around monument points. The panels were designed to provide obvious contrast between target and background desert terrain in both optical and VNIR HSI. Three calibration targets were also laid out in proximity to one another and consisted of two gray scale 3×3 meter reflectance panels and one 8-foot by 10-foot Tyvek panel.

For both the optical images and the HSI products, geodetic control processing was performed. Observed control was processed according to the NGS Online Positioning User Service (OPUS) procedures.[10] Control coordinate values were incorporated into photogrammetric and geometric correction. In addition to control point measurements, a variety of additional geodetic point measurements were recorded in the field and subsequently used.[8]

2.1.1 Optical Processing

The RGB pre-processing consisted of the following steps. Automatic white balance adjustments were applied to the RGB images and were subsequently converted into JPEG files. Automatic mosaicking of the generated orthophotos revealed color balance issues due to the combined acquisition dates and variable illumination conditions over the three-day collection. To address these variations, mosaic seamlines were manually edited to provide consistent image composition over these areas. The orthorectified mosaic was output to GeoTIFF file format with a pixel size of 0.025 meters.

It is natural to want to define the feature vector for the RGB optical data with the color channels as,

$$\mathbf{v}_{RGB} = [R \ G \ B]^{T};$$

(1)

however, the color channels tend to be correlated. The rotation matrix generated in principle component analysis (PCA) can be used to de-correlate the color channels. This, however, cannot be done in general and must be done on a class-by-class basis. For example, if one of the terrain classes is vegetation, training data from vegetation can be used to generate a rotation matrix. Then, after applying the rotation matrix to the training data, the transformed data can be used to build a vegetation classifier. This technique requires retaining the class-specific rotation matrix for testing feature vectors through the trained classifier. Let Γ^{Class} represent the class-specific rotation matrix. The, the class-specific feature vector is given by,

$$\mathbf{v}_{RGB}^{Class} = \Gamma^{Class} \mathbf{v}_{RGB}.$$

(2)

2.1.2 Hyperspectral Processing

A total of 431 hyperspectral images were collected. The VNIR imagery were processed from raw to radiance products. Each radiance cube was processed using the sensor spectral radiometric calibration parameters to yield a radiance cube in units of $\mathrm{mW/(cm^2 \times sr \times \mu m)}$. Derived scale and offset values were applied to convert the radiance images to reflectance images. The transformed reflectance images were then orthorectified to a 0.10 m pixel orthoimage. A hyperspectral mosaic was then assembled using seam lining techniques based upon the calculation of mean reflectance values in overlapping image areas to minimize tonal variations in the mosaic.[11] The HSI data were collected for 274 wavelengths, ranging from $\lambda_1 = 397.893$ nm to $\lambda_{274} = 1004.378$ nm.

A compressed representation of the HSI measured data can be generated by fitting piecewise polynomials to selected bands of the measured spectrum. Figure 1 illustrates reducing the 274 measurements into eleven polynomial coefficients fit to various parts of the spectrum. The polynomial coefficients can be used to define a reduced-dimension feature vector,

$$\mathbf{v}_{HSI} = [b_1 \ c_1 \ a_2 \ b_2 \ c_2 \ b_3 \ c_3 \ b_4 \ c_4 \ b_5 \ c_5]^T, \tag{3}$$

where the a_n, b_n, and c_n polynomial coefficients are as defined in figure 1.
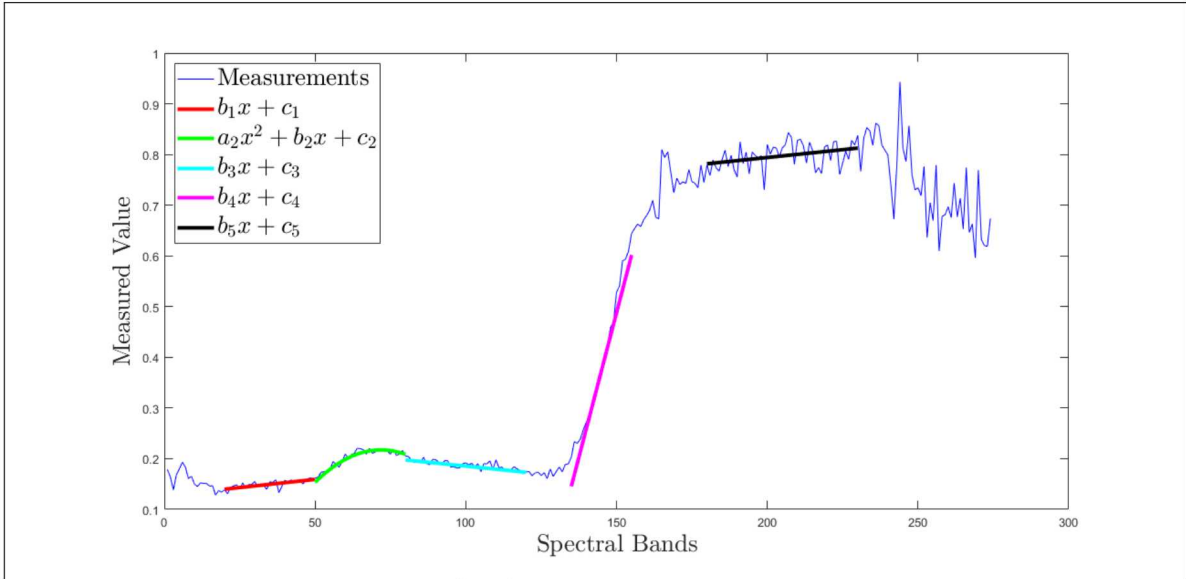

Figure 1. Illustration of an example spectrum (blue) and polynomial fits across various segments of the spectrum. The polynomial coefficients can be used to define a reduced-dimensional representation of the spectrum.

2.2 PolSAR

Synthetic aperture radar (SAR) is an active remote-sensing modality, in the sense that a SAR sensor provides its own illumination. A SAR sensor measures the reflectivity of a scene at a given nominal radio frequency (RF) center frequency and electromagnetic (EM) field polarization. The calibrated measured quantity is the radar cross-section (RCS).

A SAR sensor can be designed to measure the scene with multiple EM polarizations. A linearly fully-polarized SAR system will transmit horizontally-oriented EM energy and receive both horizontally- and vertically-oriented energy and similarly transmit vertically-oriented EM energy and receive both horizontally- and vertically-oriented energy. The full-polarized measurement set is often denoted as HH, VH, VV, HV, where the first and second letters indicates the received and transmitted polarizations, respectively.

The PolSAR data were collected using the Sandia National Laboratories-developed Facility for Advanced RF and Algorithm Development (FARAD) X-band radar system (9.6 GHz center frequency and a linear chirp bandwidth of 1.8 GHz), flown on a DHC-6 airplane. The PolSAR data were collected on March 17, 2016. The

radar was operated in videoSAR mode, along a circular flight path with a nominal range of 3460-m away from scene center and nominal grazing angle of 38°.

2.2.1 PolSAR Image Formation

The PolSAR videoSAR phase histories were formed into images at 5° azimuthal angles around the circular flight path using the polar format algorithm[12] and autofocused using the phase gradient autofocus algorithm.[13] The same autofocus correction was applied to all polarization channels. The PolSAR images were formed using an aperture extent and RF bandwidth so that, after applying a -35 decibel (dB), nbar = 5 Taylor window to the phase histories, the slant-plane range and azimuth resolutions of 0.2032 m and 0.1563 m pixel spacing were achieved. This resolution was chosen so that depth of focus blurring arising from the curved collection path was negligible. The image sets were also calibrated by compensating for the effective antenna patterns (magnitude and phase) of the different polarization channels, accounting for the differential phase that occurs from multiplexing the pulses, and radiometrically calibrated to balance the polarimetric channels and convert the images to units of RCS. The calibration technique used for this study does not attempt to compensate for co-/cross-polarization isolation issues. The measured two-way co-/cross-polarization isolation varies with azimuth in the images, which varies from better than 40 dB at the center of the beam at the scene center to approximately 10 dB at the scene extent.

The calibrated images were multi-looked to reduce speckle. The multi-looking approach employed consisted of selecting a reference image and co-registering the images formed ±5° and ±10° away from the reference. Co-registering the images gives pixelwise alignment of the images with mostly non-overlapped apertures; hence, the speckle statistics have reduced correlation and allow for speckle reduction. Multi-looking in this fashion preserves spatial resolution. The warp maps generated from the co-registration process were applied to the feature maps for the elements of the feature vector. Figure 2 (c) illustrates a multi-looked color-composite PolSAR image.
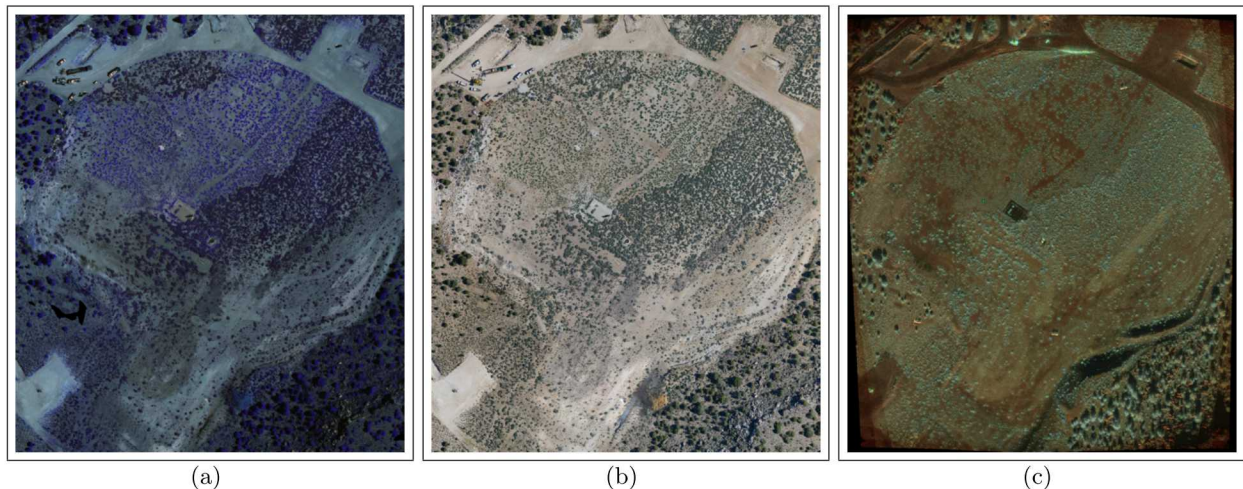

(a) (b) (c)

Figure 2. Illustration of a scene from (a) a color-composite HSI image, (b) an RGB optical image, and (c) a color-composite PolSAR image.

2.2.2 PolSAR Processing

The complex-valued PolSAR measurements can be represented by the scattering matrix,

$$S = \left[\begin{array}{cc} s_{HH} & s_{HV} \\ s_{VH} & s_{VV} \end{array} \right],$$ (4)

where s_{XY} is the complex-valued measurement from polarization channel XY. The total-power, or Span, can be computed at each pixel location as,

$$Span = |s_{HH}|^2 + |s_{VH}|^2 + |s_{HV}|^2 + |s_{VV}|^2.$$ (5)

We further compute the square-root as,

$$\zeta = \sqrt{Span}. \tag{6}$$

There is also utility in processing coherent combinations of the channels, which leads to so-called polarimetric decompositions. The Pauli feature vector is one such coherent combination,

$$\mathbf{k} = \frac{1}{\sqrt{2}} \begin{bmatrix} s_{HH} + s_{VV} \\ s_{HH} - s_{VV} \\ 2s_{HV} \end{bmatrix}. \tag{7}$$

The Pauli feature vector can then be processed into a coherency matrix by computing a spatial average of the Pauli feature vectors,

$$T = \left\langle \mathbf{k}\mathbf{k}^H \right\rangle, \tag{8}$$

where the notation $\langle \cdot \rangle$ denotes a spatial average and the superscript ^{H} denotes a complex-conjugate transpose. Several polarimetric decompositions are based on the coherency matrix.

The H/A/α decomposition is an information theoretic decomposition.[14] The H/A/α decomposition computes the eigenvectors and eigenvalues of the coherency matrix. The entropy parameter, H, is computed from the eigenvalues; the anisotropy parameter, A, is the relative strength of the second and third eigenvalues; the roll-invariant scattering mechanism parameter, α, is computed from the eigenvectors and normalized eigenvalues.

The general four-component scattering power decomposition with unitary transformation of coherency matrix (G4U) decomposition computes scattering powers from the coherency matrix.[15] The decomposition algorithm computes the surface, dihedral, volumetric, and helical scattering powers, which are denoted as P_s, P_d, P_v, and P_c respectively.

One other SAR image product that we use is the coherenct change detection (CCD) image. A CCD image is computed using the absolute value of the complex-valued cross-correlation between two temporally-separated co-registered SAR images of a scene, collected with similar imaging geometries.[12] CCD images convey the loss of coherence between two images, which is typically an indication of changes within a scene, such as trees blowing in the wind or a vehicle driving down a dirt road. The coherence parameter is denoted as γ.

Collecting the various PolSAR data descriptors on a pixel-by-pixel basis creates a pixel-wise feature vector. From the parameters discussed above, the following feature vector is formed at each pixel location:

$$\mathbf{v}_{SAR} = [H \ \alpha \ P_s \ P_d \ P_v \ P_c \ \zeta \ \gamma]^T. \tag{9}$$

The anisotropy parameter, A is not included in the feature vector; this is because it tends to capture too much of the antenna pattern and for the data collected, and as explained above, the scene size is on the order of the antenna pattern. Note, the H and γ quantities are natively in the interval [0 1], and the other parameters are normalized to it; this is accomplished by converting the power and ζ quantities to dB and then limiting the dynamic range and rescaling, and dividing the α parameter, which is naturally in the range [0° 90°], by 90°.

3. FRAMEWORK

In this section we introduce the terrain classes and our framework for extracting features for terrain classification, walk through the probabilistic feature fusion (PFF) modeling of the selected features, and discuss the pixelwise voting method.

3.1 Terrain Classes

The imagery from the three imaging modalities contain three distinct types of vegetation. Figure 3 illustrates sample images of the three different vegetation classes, denoted as VEG1, VEG2, and VEG3, from the three imaging modalities. As can be seen from the figure, the representations of the three different classes are very different between each imaging modality. Some readily observable differences include: the lay-over effect of the trees in the PolSAR images, the direction of the shadows between the optical modalities and the PolSAR images, and the resolution (the optical has the highest resolution).
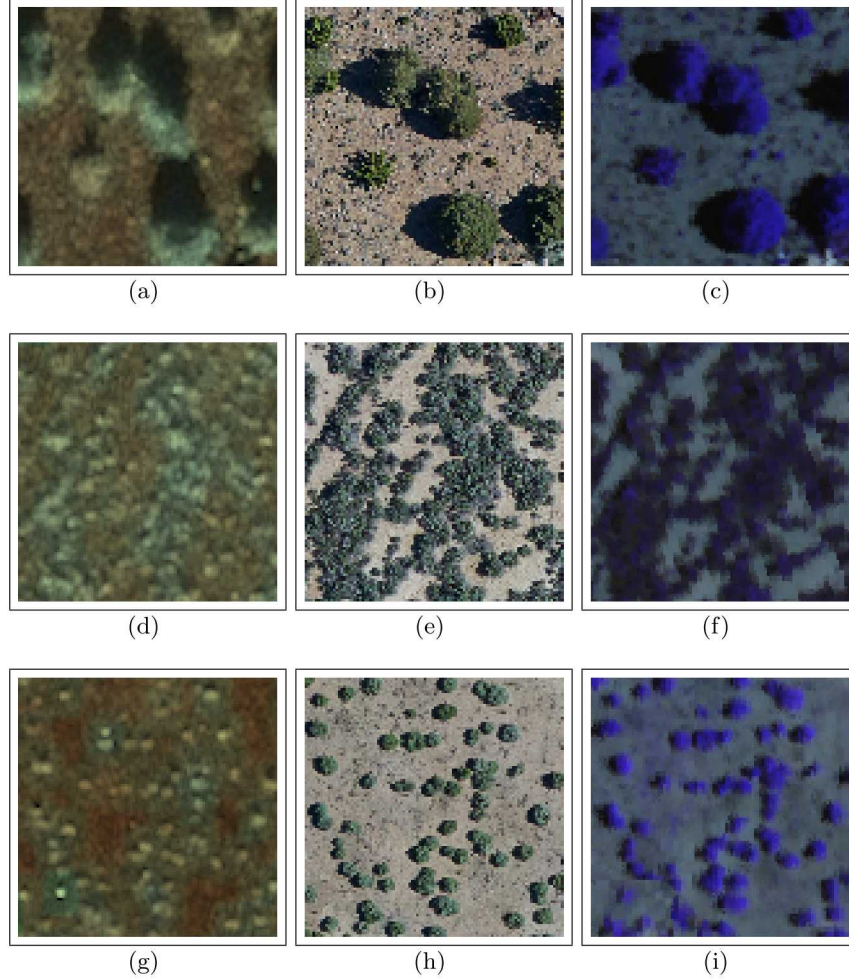
Figure 3. Illustration of three vegetation classes observed from the three different imaging modalities. The rows from top to bottom are trees (VEG1), shrubs (VEG2), and low-shrubs (VEG3). The columns from left to right are PolSAR, optical, and HSI.

3.2 Pixels vs. Superpixels

Depending on the imaging modality, there can be significant pixel-to-pixel variability in resulting image products, as can be seen in figure 3. In general, the pixel-to-pixel variability depends on many factors such as, the local signal-to-noise ratio throughout the image, the texture of the objects within the scene being imaged, etc. Furthermore, in the case of SAR images, spatially distributed speckle is intrinsic to the imaging modality and is apparent throughout images, unless a speckle mitigation strategy is employed.

Due to the pixel-to-pixel variability, the distribution of measured values within relatively homogenous regions may have more spread than desired for terrain classification. There are many methods to deal with this issue; one simplistic method is to process the image with a moving average filter. This, however, has the undesirable effect of blurring edges within the image and reducing the resolution. (For terrain classification, blurred edges are likely to be more detrimental than loss of resolution.) We have found that the SLIC superpixel segmentation algorithm does an excellent job delimiting contrasting terrain boundaries; thus, in this paper we perform a superpixel segmentation of the images with a nominal setting of 40 pixels per superpixel. We retain the pixel membership for each superpixel, for each image, which enables extracting feature vector information from derived image products and computing statistics on the pixel data within each superpixel.

3.3 Feature Selection

In the previous section we introduced our feature vector representation for each imaging modality. Depending on the terrain class, it may be the case that only a few feature vector elements are required to sufficiently discriminate it from other classes. It is worth taking the time to find the discriminating feature vector elements. To this end, generating pairwise combinations of the training data feature vector elements, for each terrain class, and displaying as scatterplots, helps to visualize the separability of the classes. As an example, consider the pairwise feature vector scatterplot from the PolSAR training data in figure 4 (a); the different colors represent the three different terrain classes. It can be seen in the figure that there is good separation of the classes with the (α, P_c) feature vector pair.

We selected the features for each class and each imaging modality by viewing the pairwise combinations of feature vector element scatterplots. At worst case, this required viewing 55 scatterplots from the HSI feature vector data.

3.4 Probabilistic Feature Fusion Modeling

The PFF classifier is a one-class classifier.[9] With the PFF classifier, it is up to the practitioner to define and model the features to be combined (fused) for the decision process. The features must be uniquely descriptive of the class in order to properly pass in-class data and reject out-of-class data. One of the most important attributes of PFF is that the classification decision is completely traceable back to the features selected to discriminate. This attribute allows the practitioner to fully understand the decision process.

The training data from the selected feature vector elements, for each class, and for a given imaging modality, are used to generate a PFF model. For the HSI and PolSAR discriminative pairings of feature vector elements, the pairwise data are used to define a Mahalanobis distance between the centroid of the pairwise data to any other point in the pairwise space. An illustration of a selected pairwise set of in-class feature vector elements (blue dots), and a contour of the Mahalanobis distance (black ellipse), is given in figure 4 (b). It is important to note that the elliptical contour is not an in-class / out-of-class boundary, but is just to illustrate the shape of the distance surface; values along a contour all have the same Mahalanobis distance. There are only three feature vector elements for the optical data; thus, selecting a single pairwise feature vector combination would exclude one feature or include one feature twice if two pairwise combinations were used. Instead of a two-dimensional Mahalanobis distance, the absolute distances from the mean value of the selected in-class features were used.
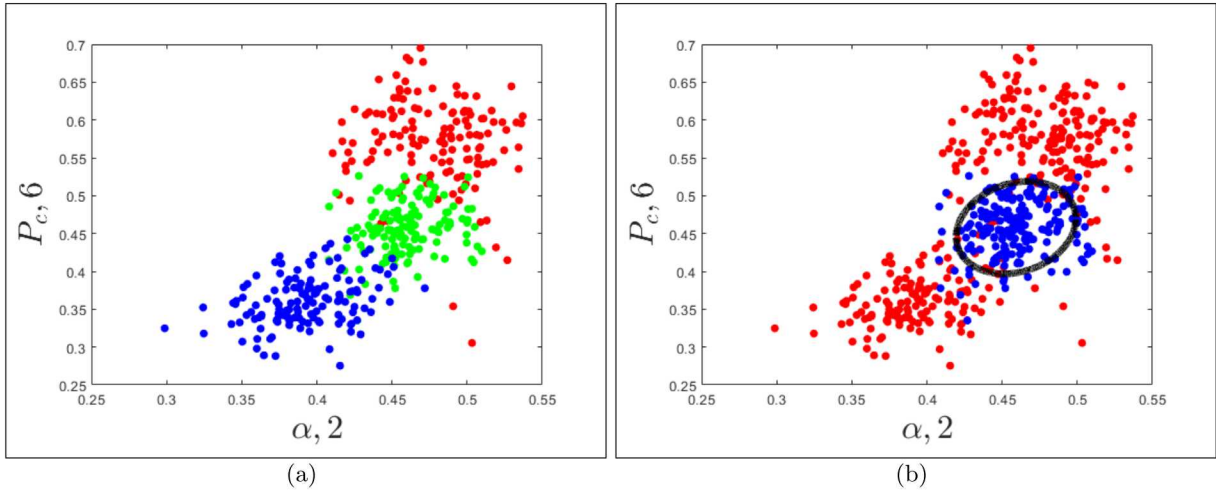

Figure 4. (a) Illustration of the scatterplot of the (α, P_c) PolSAR feature vector elements from the training data. The colors represent the three classes. It can be seen that there is good class separation with these two feature vector elements. (b) Illustration of a selected pairwise feature vector combination. The black ellipse illustrates a contour of the Mahalanobis distance metric that is computed about the centroid of the selected feature pair. Values along the ellipsoidal contours have the same Mahalanobis distance from the centroid.

The next step in the PFF classifier modeling is to fit probability distributions to the distances for the in-class data. Gamma distributions were fit to the metric data computed, as illustrated in figure 5 (a). The upper-left plot in the figure illustrates the gamma distribution fit to the in-class metric histogram data. The plot in the upper-right illustrates the in-class p-values, which should be uniformly distributed if the gamma distribution fits the data well. The plot in the lower left compares the histogram of the out-of-class data to the in-class gamma distribution; as can be seen there is a some overlap between in-class and out-of-class. The plot in the lower-right illustrates the histogram of the p-values computed from the out-of-class data through the feature model; as can be seen, the p-values computed from the out-of-class data are low, which is desirable. The fusion of the modeled features does not require any training or tuning. The fused features for one of the PolSAR classes is illustrated in figure 5 (b), where it can be seen that the fused out-of-class p-values are lower, as compared to the single modeled feature vector pair. The final step in defining a PFF classifier is to set a p-value threshold for determining whether a sample is in-class, or not.

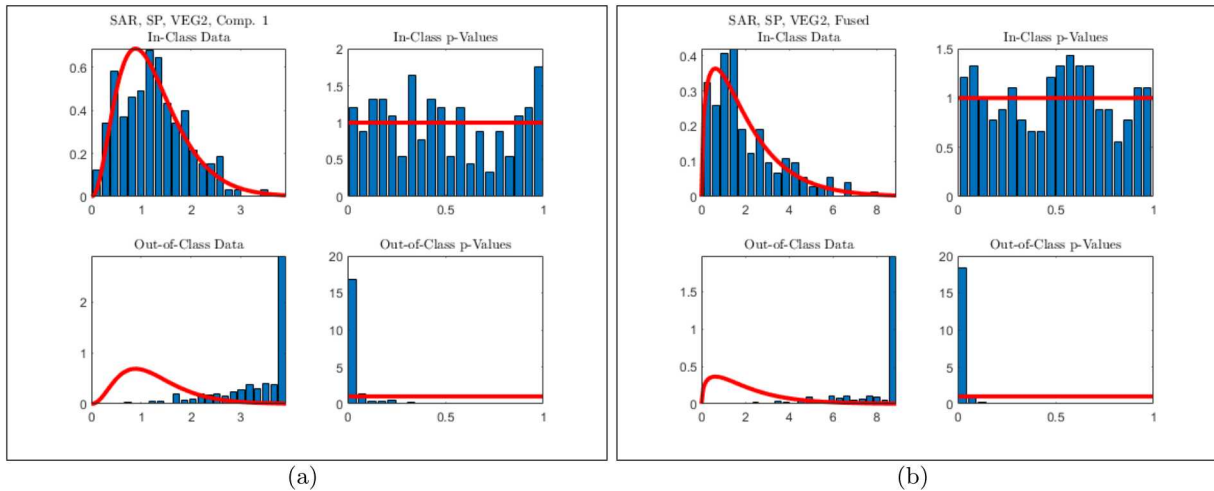(a)   (b)

Figure 5. (a) Illustration of the the PFF model of a selected feature. The upper-left plot shows a gamma distribution fit to the computed in-class Mahalanobis distances. The plot in the upper-right illustrates the in-class p-values, which should be uniformly distributed. The lower-left plot shows the overlay of the in-class gamma distribution fit to the out-of-class histogram. Finally, the lower-right plot illustrates that the out-of-class data generates low p-values, as desired. (b) Illustration of the same for the fused set of feature vector pairings.

3.5 Pixelwise Voting

Depending on the terrain class, there can be a fair amount of measurement variability on a pixel-by-pixel basis. Figure 6 (a) illustrates the pixel-by-pixel measurement variability in the VEG2 class from the optical data. This variability will cause histogram spread and runs the risk of increasing the overlap between closely spaced classes thereby increasing the number of mis-classifications. There are multiple ways to deal with this issue; in this paper we explore two approaches. The first is to average the feature values within a superpixel, to reduce histogram spread, then make a class determination from the mean values. The second approach is to make pixelwise classification decisions, then to take the majority vote for an overall classification determination for the superpixel.

To demonstrate the difference of the approaches, consider figure 6 (b) that illustrates the pixelwise class declarations from the PFF classifier; it can be seen that there is significant variability in the class decisions. Figure 6 (c) illustrates the decisions made by averaging the pixels within a superpixel; the results are more homogenous than the pixelwise method, but still contain some mis-classifications. Finally, figure 6 (d) illustrates the voting-based class determinations by assigning the superpixel the label from the mode of the pixel decisions within the superpixel; as can be seen, the voting method has the fewest mis-classifications (the 'Unknown' label is due to a tie vote).
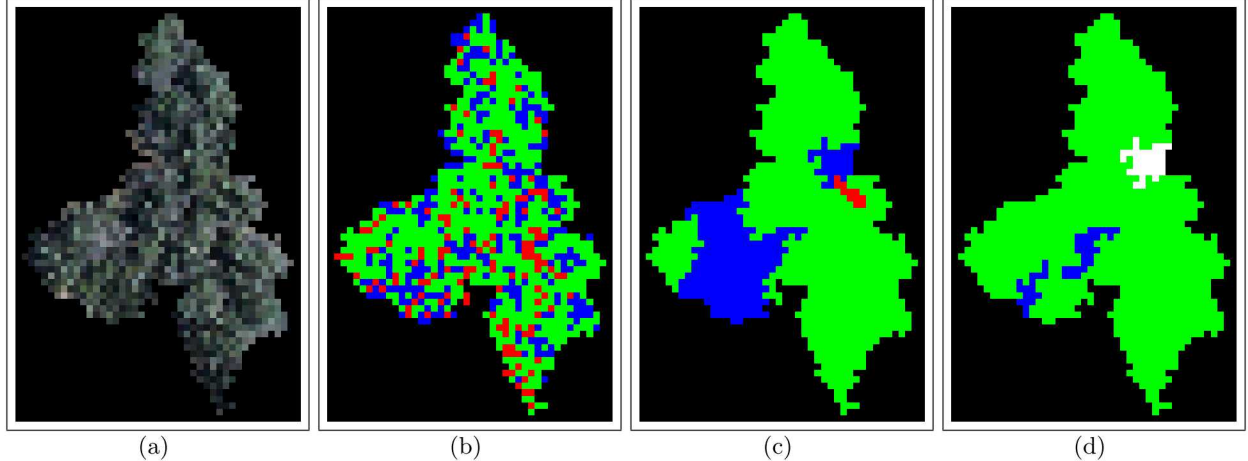
Figure 6. (a) Illustration of 28 superpixels composed primarily of the VEG2 class. (b) and (c) The terrain classification as determined by the pixelwise and superpixel PFF classifiers, respectively. (d) The terrain classification decisions determined by pixelwise voting within a superpixel. In (b)-(d), the mapping is as follows: red → VEG1, green → VEG2, blue → VEG3, and unique to the voting method white → 'Unknown' (*i.e.* a tie vote). Given that the VEG2 class should be declared, it can be seen that the voting method has the best performance.

4. PERFORMANCE METRICS

In this section we demonstrate the performance of the voting method from the pixelwise decisions of the PFF classifier. There are a variety of methods that help to understand the performance of a classifier, such as receiver operating characteristic (ROC) curves and confusion matrices, which we show the ROC curves for the pixelwise and superpixel PFF models. To demonstrate the performance of the voting method, we generate confusion matrices and give a visual comparison of the terrain classification of the scenes illustrated in figure 2 for the different imaging modalities.

It is important to note that the performance metrics are computed on the training data; we did not select test data to compute the performance metrics. Our expectation is that the performance metrics will look better with training data and would fully expect a drop in performance on blind test data; however, computing the performance metrics on the training data should be sufficient to show the *relative* performance between the terrain classification decision approaches.

4.1 ROC Curves

ROC curves are computed by sweeping the classifier metric threshold across its domain of values and determining either: the normalized count (typically given as a percentage) of both in-class and out-of-class values that are above the threshold as it sweeps; or, if the in-class and out-of-class probability distributions are modeled, computing the area of the distributions that are above the threshold as it is varied. The PFF classifiers generate p-values; thus, the threshold domain is bounded between zero and one. Furthermore, with the PFF classifier, only the in-class probability distributions are modeled; thus the first method is employed to generate the ROC curves.

Figure 7 illustrates the ROC curves for the three vegetation terrain classes for the three different imaging modalities. For reasonable operating points on the ROC curves, it is clear that the superpixel PFF classifiers significantly outperform the pixelwise PFF classifiers for both the optical and HSI data and moderately outperforms the pixelwise PFF classifier for the PolSAR data. The implication is that the variability of the terrain classes, and hence the class overlap, has been dramatically reduced by averaging the feature values over a superpixel. The results of the ROC curves build a strong case for sacrificing resolution for improved classification, which in the case of high resolution terrain, may be an appropriate tradeoff.
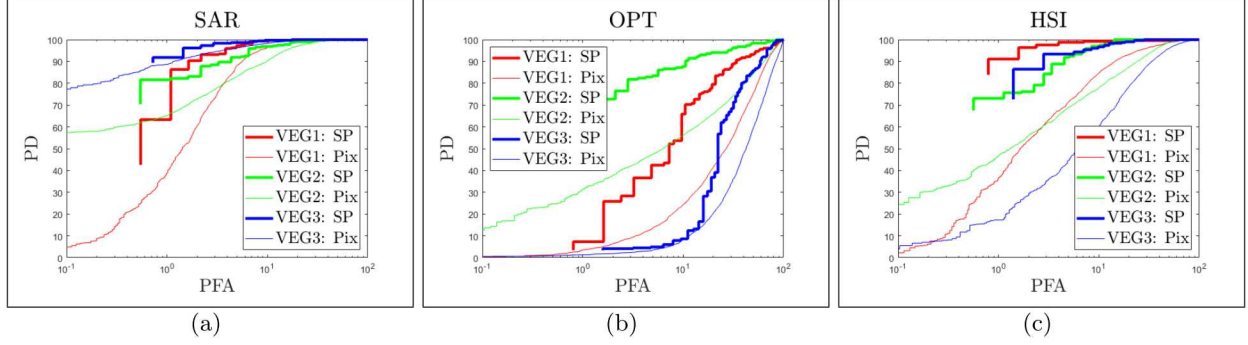
Figure 7. Illustration of the ROC curves generated for the terrain classes for the different imaging modalities. The ROC curves for the PolSAR, optical, and HSI modalities are illustrated in (a), (b), and (c), respectively. In the plots the colored lines indicate the performance of the different terrain classes. The bold lines are from the superpixel PFF models and the thinner lines are from the pixelwise PFF models. As can be seen, for reasonable operating points on the ROC curves, the performance of the superpixel PFF models varies from moderately to dramatically better performance over the pixelwise PFF models.

4.2 Confusion Matrices

The first step in generating confusion matrices is to process data samples from known classes through each of the different classifiers. For a given data sample, the classifier metric scores are compared and the class decision is made based on which classifier has the highest metric score. An 'Unknown' class can be introduced and a data sample is binned as such if the highest metric score is below a determined threshold.

Figures 8, 9, and 10 illustrate the confusion matrices for the PolSAR, optical, and HSI data, respectively, for the pixelwise, superpixel, and voting methods. We defined an 'Unknown' class and populated it if the maximum p-value across all classifier outputs was less than 0.05. In the figures, the values were rounded to a tenth of a percent and a '-' was placed for entries less than 0.1%. As can be seen from the correct declarations in the figures, the PolSAR data set benefited from the voting method across all classes, the optical data benefited from the voting method for two classes and tied with the superpixel method for the third, and the HSI data only benefited from the voting method for one class.

4.3 Applied Terrain Classification

An important qualitative measure of performance is to consider the performance of the trained terrain classifiers on an image. A visual inspection of the results helps to gain a practical sense of usability. Furthermore, an inspection of the by-feature p-value maps from the PFF classifier helps to determine how useful a given feature is in the classifier and give a sense of what other terrain types are confusers for a given class.

As an example of the utility of visual inspection, consider the illustrations in figure 11. The image in (a) illustrates the p-values generated from the metric computed from the (P_d, γ) feature pair from the PolSAR data for the VEG1 terrain class. The image in (b) illustrates the same for the (P_s, P_c) feature pair, also for the VEG1 terrain class. By inspecting (a) and (b), it is clear that they have agreement in some regions. The image in (c) illustrates the fused p-values from the two feature pairs; it is evident in (c) that the fused p-values have rejected the non-VEG1 regions. Finally, (d) illustrates the VEG1 decisions by thresholding the fused p-values at 0.10. There is a horizontal streak in (c) and (d) in the center of the image towards the top of the image that clearly does not belong to the VEG1 class and it is evident that it appears in both of the paired features. With PFF, the maps of other features can easily be viewed to see if inclusion of another feature will remove mis-classifications.

The output of the pixelwise, superpixel, and voting methods for the SAR data are illustrated in figure 12(a-c), for the optical data in (d-f), and for the HSI data in (g-h). In all cases, the decision threshold was set to a p-value of 0.10.
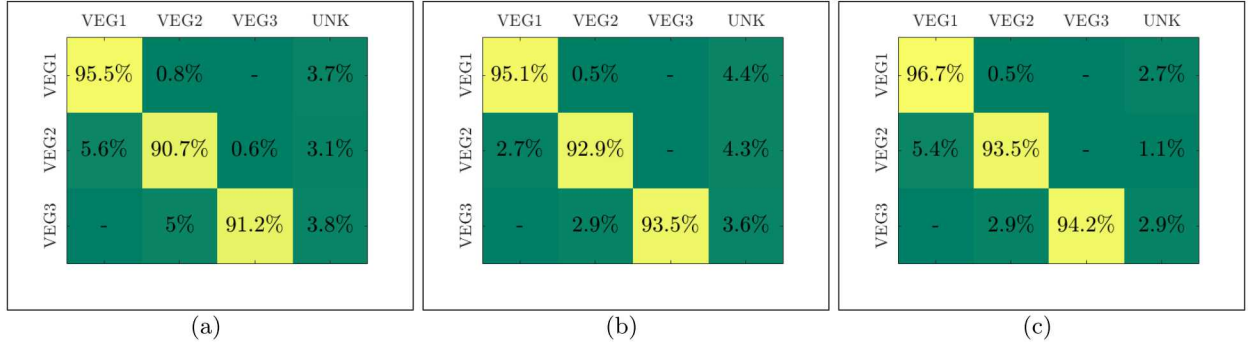
Figure 8. Illustration of the confusion matrices computed from the PolSAR data. The pixelwise, superpixel, and voting methods are given in (a), (b), and (c), respectively. There is one case where the pixelwise method slightly outperformed the superpixel method in terms of in-class declarations; however, the voting method showed improved performance for in-class classification compared to the pixelwise and superpixel methods across all classes.
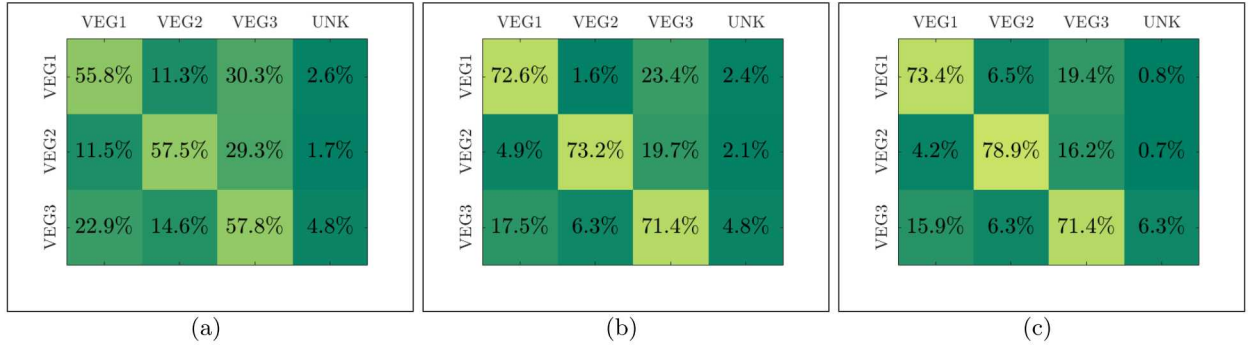

Figure 9. Illustration of the confusion matrices computed from the optical data. The pixelwise, superpixel, and voting methods are given in (a), (b), and (c), respectively. There is one case where the voting method in-class declaration tied with the superpixel method; otherwise, the voting method showed improved performance for in-class classification, compared to the pixelwise and superpixel methods.
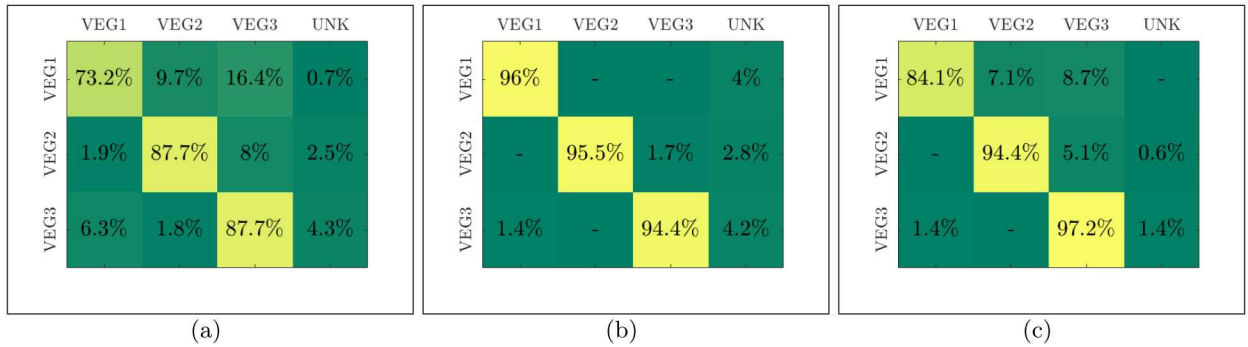

Figure 10. Illustration of the confusion matrices computed from the HSI data. The pixelwise, superpixel, and voting methods are given in (a), (b), and (c), respectively. There are two cases where the voting method in-class declarations dropped below the superpixel method; the voting method showed improved performance for in-class classification for the remaining class, compared to the pixelwise and superpixel methods.
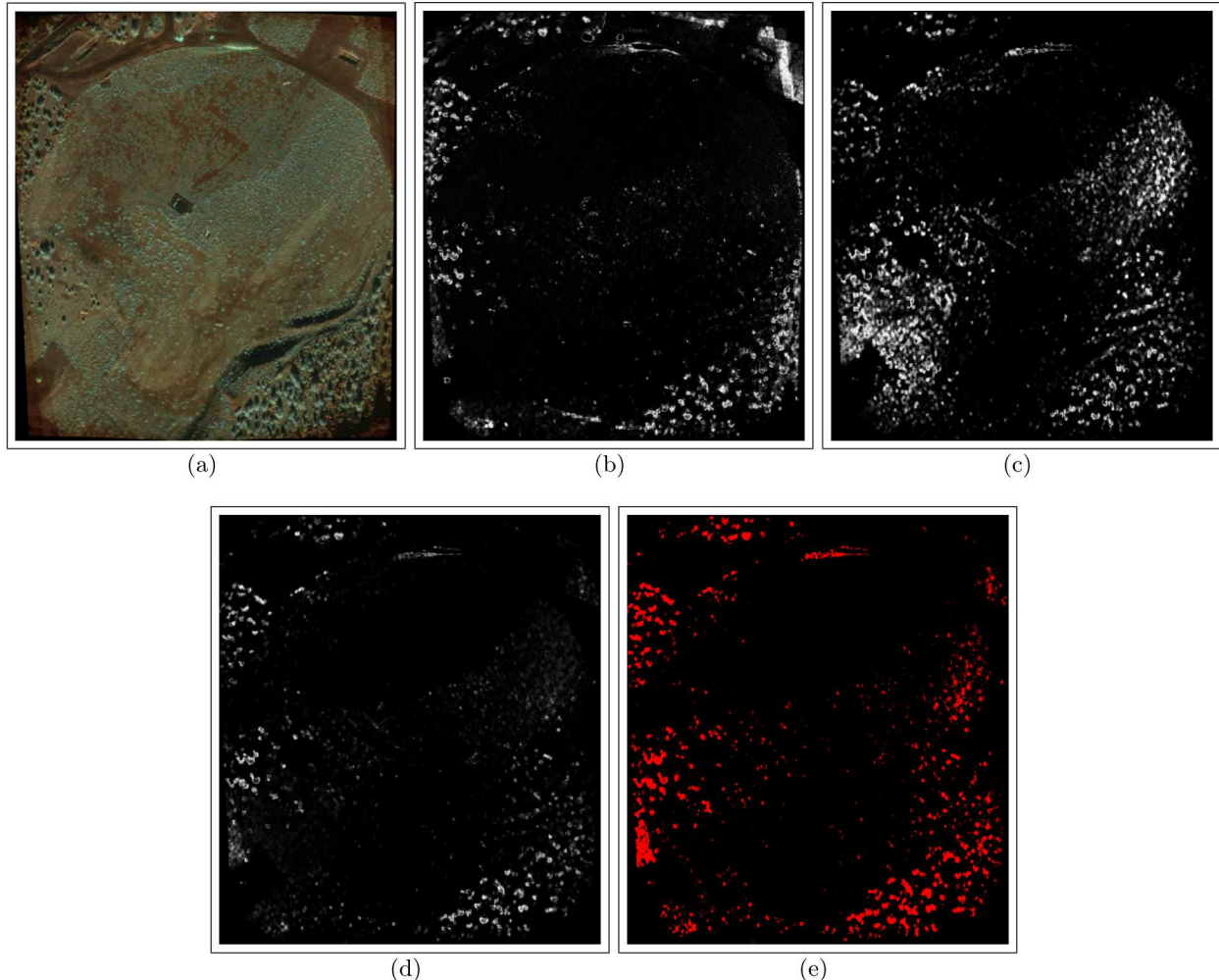
Figure 11. Illustration of (a) the color-composite PolSAR image (for reference), (b) the p-value map generated from the (P_d, γ) feature pair through, (c) the p-value map generated from the (P_s, P_c) feature pair, (d) the p-value map from the fused features, and (e) the regions in the PolSAR image that were declared as VEG1 with the p-value threshold set to 0.10. In (b)-(d), white indicates a p-value close to one and represents good agreement with the training data used to model the PFF feature.

5. CONCLUSION

In this paper we introduced a robust framework for high-resolution terrain classification that employs pixelwise voting and the probabilistic feature fusion (PFF) one-class classifier. The pixelwise voting is accomplished by first, segmenting the image into superpixels using the simple linear iterative clustering (SLIC) algorithm, second processing the pixel feature data through the PFF models, and finally assigning each superpixel the classification label from the majority vote of the pixelwise output of the PFF model. In this manner, the pixelwise variability in high-resolution remote sensing images can be reduced.

The importance of using the PFF classifiers is that they allow the practitioner to define meaningful discriminative features and offers complete traceability for the decisions that are made.

6. ACKNOWLEDGMENTS

REFERENCES

[1] Jong-Sen Lee, Grunes, M. R., Pottier, E., and Ferro-Famil, L., "Unsupervised terrain classification preserving polarimetric scattering characteristics," *IEEE Transactions on Geoscience and Remote Sensing* **42**, 722–731 (April 2004).

[2] Vandapel, N., Huber, D. F., Kapuria, A., and Hebert, M., "Natural terrain classification using 3-d ladar data," in [*IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*], **5**, 5117–5122 (April 2004).

[3] Lalonde, J.-F., Vandapel, N., Huber, D. F., and Hebert, M., "Natural terrain classification using three-dimensional ladar data for ground robot mobility," *Journal of Field Robotics* **23**(10), 839–861 (2006).

[4] West, R. D., LaBruyere III, T. E., Skryzalin, J., Simonson, K. M., Hansen, R. L., and Van Benthem, M. H., "Polarimetric SAR Image Terrain Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **12**, 4467–4485 (Nov 2019).

[5] Chia-Tang Chen, Kun-Shan Chen, and Jong-Sen Lee, "The use of fully polarimetric information for the fuzzy neural classification of SAR images," *IEEE Transactions on Geoscience and Remote Sensing* **41**, 2089–2100 (Sep. 2003).

[6] Bau, T. C., Sarkar, S., and Healey, G., "Hyperspectral region classification using a three-dimensional gabor filterbank," *IEEE Transactions on Geoscience and Remote Sensing* **48**, 3457–3464 (Sep. 2010).

[7] Kuang, D. and Xu, J., "Combined multiple spectral-spatial features and multikernel support tensor machine for hyperspectral image classification," *Journal of Applied Remote Sensing* **14**(3), 1 − 16 (2019).

[8] Anderson, D., Craven, J. M., Dzur, R., Briggs, T., Lee, D. J., Miller, E., Schultz-Fellenz, E., and Vigil, S., "Using unmanned aerial systems to collect hyperspectral imagery and digital elevation models at a legacy underground nuclear test site," in [*SPIE Conference on Image Sensing Technologies: Material, Devices, Systems, and Applications V*], **10656**, UNSP 1065605 (Apr 2018).

[9] Simonson, K. M., West, R. D., Hansen, R. L., LaBruyere, T. E., and Van Benthem, M. H., "A statistical approach to combining multisource information in one-class classifiers," *Statistical Analysis and Data Mining: The ASA Data Science Journal* **10**(4), 199–210 (2017).

[10] NOAA, N., "OPUS: Online positioning user service," (February 2018).

[11] "OTB Version 5.6.1."

[12] Jakowatz, Jr., C. V., Wahl, D. E., Eichel, P. E., Ghiglia, D. C., and Thompson, P. A., [*Spotlight-mode Synthetic Aperture Radar: A Signal Processing Approach*], Springer (1996).

[13] Wahl, D. E., Eichel, P. H., Ghiglia, D. C., and Jakowatz, C. V., "Phase gradient autofocus-a robust tool for high resolution SAR phase correction," *IEEE Transactions on Aerospace and Electronic Systems* **30**, 827–835 (Jul 1994).

[14] Cloude, S. R. and Pottier, E., "A review of target decomposition theorems in radar polarimetry," *IEEE Transactions on Geoscience and Remote Sensing* **34**, 498–518 (Mar 1996).

[15] Singh, G., Yamaguchi, Y., and Park, S. E., "General four-component scattering power decomposition with unitary transformation of coherency matrix," *IEEE Transactions on Geoscience and Remote Sensing* **51**, 3014–3022 (May 2013).
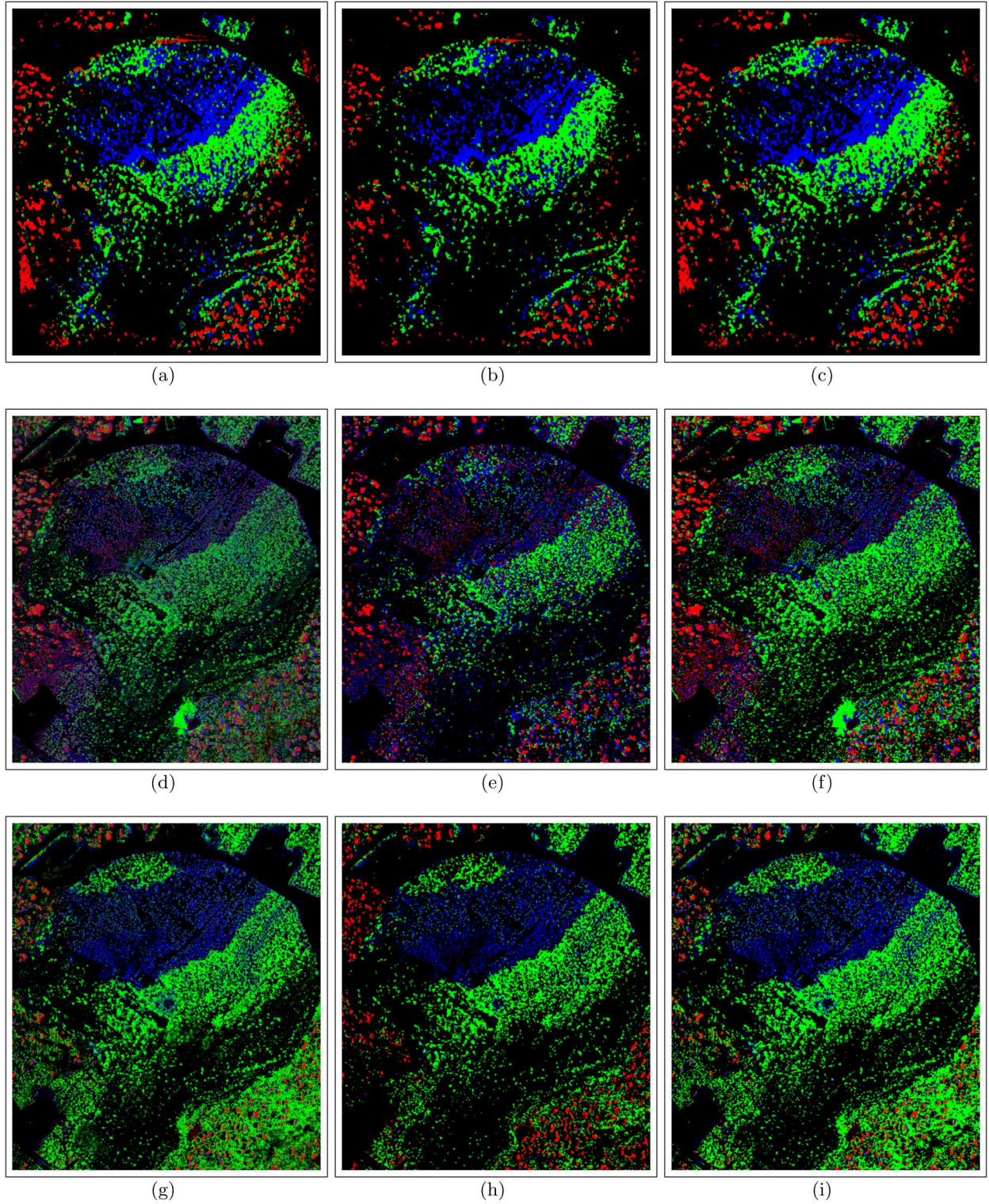
Figure 12. Illustration of the terrain classifications declared by the PFF models for each imaging modality and each of the decision methods. The rows from top to bottom are PolSAR, optical, and HSI. The columns from left to right are the pixelwise, superpixel, and voting methods. The decision threshold, for all cases, was set to a p-value of 0.10. Not in every case, but in general, the voting method tends to have less spatial variability than the pixelwise method and is not as conservative as the superpixel method.