

An Adaptive Noise Filtering Algorithm for AVIRIS Data with Implications for Classification Accuracy

Rhonda D. Phillips *Student Member, IEEE*, Christine E. Blinn *Member, IEEE*,
Layne T. Watson, *Fellow, IEEE*, and Randolph H. Wynne, *Member, IEEE*

Abstract—This paper describes a new algorithm used to adaptively filter a remote sensing dataset based on signal-to-noise ratios (SNRs) once the maximum noise fraction (MNF) has been applied. This algorithm uses Hermite splines to calculate the approximate area underneath the SNR curve as a function of band number, and that area is used to place bands into “bins” with other bands having similar SNRs. A median filter with a variable sized kernel is then applied to each band, with the same size kernel used for each band in a particular bin. The proposed adaptive filters are applied to a hyperspectral image generated by the AVIRIS sensor, and results are given for the identification of three different pine species located within the study area. The adaptive filtering scheme improves image quality as shown by estimated SNRs, and classification accuracies improved by more than 10% on the sample study area, indicating that the proposed methods improve the image quality, thereby aiding in species discrimination.

I. BACKGROUND

Hyperspectral images provide a powerful tool as the wave spectrum is finely discretized using hundreds of channels on a scanner. The large dimensionality of a hyperspectral dataset often requires a data transformation such as principal components analysis (PCA) or the singular value decomposition (SVD) to reduce the number of variables, or bands, within an image prior to further processing. Furthermore, these images tend to be noisy as a result of the fine discretization and other factors such as the method of acquisition (small aircraft) [1]. Green et al. first proposed the maximum noise transform (alternately called the minimum noise transform, minimum noise fraction, or MNF) to align a dataset in order of decreasing signal-to-noise ratio (SNR) using an eigenvalue decomposition similar to PCA [2]. Lee et al. equivalently defined the MNF (or noise adjusted PCA) as two PCA transformations, and used the MNF to reduce the noise level in an image [3]. The MNF can be used to reduce noise and the number of dimensions in an image. Reduction in noise in imagery is essential to many remote sensing applications, as Landgrebe has documented the relationship between noise in imagery and classification errors [4]. Furthermore, certain applications require a minimum SNR, for example, estimating foliar biochemical concentrations [5].

Noise can be reduced using a variety of filters defined on the frequency or spatial domains [6]. While certain frequency domain filters (using the Fourier transform) have been shown to be more effective than spatial filters with respect to specific types of noise, a spatial filter such as a median filter can produce similar results and requires significantly less computation [7]. An adaptive filter can alter the size of the filter kernel (spatial domain) or change the frequencies filtered (frequency domain) depending on image characteristics and noise levels. Lennon et al. used an adaptive median filter on data transformed to MNF coordinates [8], and Pok et al. vary the kernel size between three and five depending on the detected noise in a particular window in a three-band image [9]. King et al. present an adaptive frequency domain filter used on medical imagery [10].

The properties of the MNF are well suited to an adaptive filter, yet adaptive spatial filtering is not commonly used on MNF transformed data, although the idea was proposed by [8]. Typical data processing using the MNF truncates the data, resulting in loss of signal; uses all bands in the MNF coordinate system without noise removal, or applies a spatial convolution with a uniform kernel size across all bands despite all bands having drastically different SNRs. This paper introduces an algorithm for an adaptive median filter applied to data transformed using the MNF in which the filter support size varies with noise. To demonstrate the effectiveness of this technique, a real dataset is filtered using

R.D. Phillips and L.T. Watson are with the Departments of Computer Science and Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

C.E. Blinn and R.H. Wynne are with the Department of Forestry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

the algorithm presented in this paper, and the dataset is subsequently used to identify three different species of pines within a forest, an application for which the high spectral resolution imagery is well-suited.

II. MAXIMUM NOISE FRACTION

Data transformations such as PCA and SVD transform an image to a new coordinate system without taking factors such as noise into consideration [2]. PCA uses the eigenvectors (V) resulting from an eigen decomposition:

$$\Sigma = V\Lambda V^{-1},$$

where Σ is the covariance matrix of the image, and Λ is a diagonal matrix containing the eigenvalues corresponding to V , as a new coordinate basis for the image. This transformed image has the property that each successive band is aligned along an axis of decreasing overall variance in the original image; that is, as the component number increases, the variance within the component decreases. When PCA is used for data reduction, ideally these higher order bands with decreasing variances are not necessary to represent the majority of the original image, and these components can be removed, resulting in a data reduction. Unfortunately with datasets that are particularly noisy (the case with hyperspectral data), the first few components are not sufficient to represent the image as they capture much of the noise as well as the signal. The MNF is similar in spirit to PCA with the additional quality that it considers image noise when selecting a new coordinate system. While the PCA aligns the axes along directions of the maximum variance in the original image, the MNF aligns the axes along directions of the maximum SNR.

Theoretically, the MNF orders the data along the axes of maximum SNR using the eigen decomposition:

$$\Sigma_S \Sigma_N^{-1} = V\Lambda V^{-1}$$

where Σ_S is the covariance matrix of the signal, Σ_N is the covariance matrix of the noise, V is an (orthogonal) matrix containing the eigenvectors of $\Sigma_S \Sigma_N^{-1}$, and Λ is a diagonal matrix containing eigenvalues that correspond to V . V provides the basis for the transformed dataset. In practice, Σ_S and Σ_N are unknown and must be estimated from the data [2]. Σ_S is generally taken to be the covariance matrix of the image, and Σ_N can be estimated using various procedures [2][3]. The eigenvalues contained in Λ are the estimated variance of the signal (σ_S) divided by the estimated variance of the noise (σ_N), and therefore the diagonal element λ_b in Λ is an approximation for the SNR of band b in the transformed image.

III. ADAPTIVE FILTER

The MNF is commonly used in remote sensing for data reduction and noise removal. The MNF of an image can be truncated while still preserving most of the information within the image, which is especially useful in the hyperspectral image processing domain as images contain hundreds of highly correlated bands and noise. The higher order bands that are truncated commonly contain very low SNRs, and truncating the MNF can have the added effect of eliminating much of the noise without losing much signal. Determining the precise location to truncate the MNF is problematic, and a judgement call is often made by looking at a plot of the eigenvalues relative to the band number and determining where this eigenvalue curve begins to approach an asymptote ($\lambda = 1$). In practice, this truncation is performed as a means of reducing the overall noise within the image, but this method does not fully take advantage of the properties of the MNF. If the truncation includes too many bands, too much noise is left in the image, and if the truncation includes too few bands, useful signal may be excluded from the resulting image. A likely scenario would be that truncation includes noise in the bands that are kept while discarding good signal with the higher order bands that are discarded.

Green et al. suggest that with low SNR bands, all values can be replaced with the mean of the band, and the MNF image can be retransformed to the original subspace, resulting in a less noisy image [2]. This is an example of a rather extreme mean filter. Another approach to reducing the noise in an image is to apply a small (typically a three by three window) spatial filter such as a mean or median filter. However, applying a filter uniformly to all bands within the MNF will not take advantage of the specific ordering of the bands. Bands with lower SNRs might benefit from a filter with a larger window, while bands with high SNRs require little or no filtration. Bands with low SNRs have comparatively low signal relative to noise, yet may have enough signal to warrant smoothing of the noise. A large filter will degrade that signal, but will hopefully affect the noise more, resulting in a greater signal relative to noise.

Spatial median filters work by decreasing the variance within a small window (kernel) by assigning a pixel the median value of the surrounding pixels. For example, using a 3×3 window, a median filter would assign a pixel the median value of itself and its eight immediate neighbors (top, left, right, bottom, and four diagonal locations). As geographic data is highly correlated, the variance of the signal within such a window should be small, and noise should be random and not correlated within a neighborhood, making a large variance probable. With the assumption that the

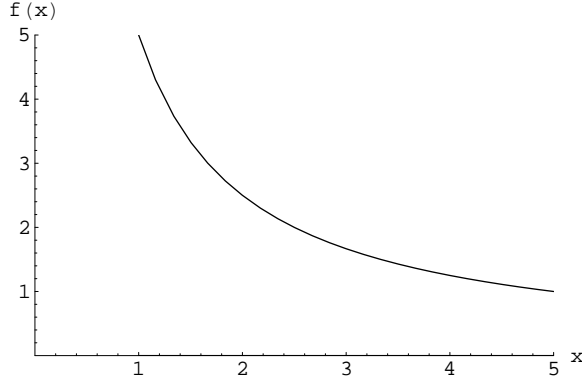


Fig. 1. Typical MNF eigenvalue curve shape.

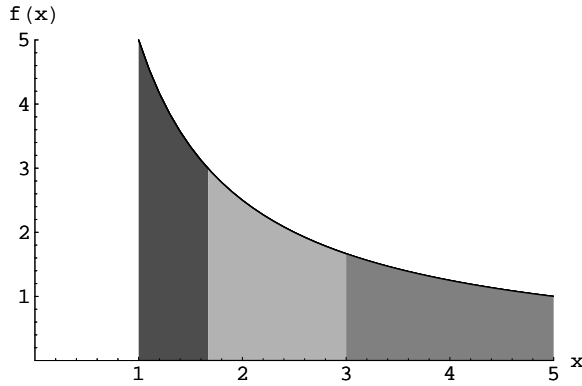


Fig. 2. Dividing area under curve into bins.

variance of the noise is larger than that of the signal in these small windows, a spatial filter such as a median filter will preserve most of the signal while eliminating much of the noise. A filter with a larger window has a more dramatic smoothing effect over a filter with a small window, resulting in a larger SNR at the expense of the signal. A median filter has the property of preserving original values unlike a mean filter.

The MNF is ordered such that for any two bands numbered m and n (assume $n > m$), $\text{SNR}_m \geq \text{SNR}_n$. Recall that the eigenvalues associated with the MNF are estimates of SNR, meaning $\lambda_m \geq \lambda_n$. Based on the reasoning offered above, the size K of the filter kernel for band m should not be greater than that for band n , $K_m \leq K_n$. Similarly, the same size filter should be applied to bands with the most similar eigenvalues.

Consider the shape of the typical MNF eigenvalue curve, shown in Fig. 1. The first few bands with the largest decrease in slope should be grouped together in smaller groups than the last bands with very similar small negative slopes. In order to divide the bands into bins in this manner, the area underneath the eigenvalue curve can be divided evenly into a number of bins corresponding to the number of different sized filters to be applied, as shown in Fig. 2. The three colors represent three different bins and three different kernel sizes. A formal algorithm for the adaptive filter (AF) is given below.

Algorithm AF(M, Λ , nb)

input/output:

M (image transformed to MNF coordinates,
filtered upon exit)

input:

Λ (B eigenvalues where B is number of bands in M)

nb (number of bins to use)

1 **begin**

2 approximate $f(b) = \lambda$ with $C(b) = \lambda$

```

3    $area := \int_1^B C(b)db$ 
4    $area\_per\_bin := \frac{area}{nb}$ 
5    $sarea := 0$ 
6   for  $i := 1$  step 1 until  $B - 1$  do
7     begin
8        $sarea := sarea + \int_i^{i+1} C(b)db$ 
9        $bin := \text{ceiling}\left(\frac{sarea}{area\_per\_bin}\right)$ 
10       $kernel_i := 2 \cdot (bin - 1) + 1$ 
11     end
12      $kernel_B := kernel_{B-1}$ 
13     for  $i := 1$  step 1 until  $B$  do
14       apply  $kernel_i \times kernel_i$  median filter
        to  $band_i$  in image  $M$ 
15  end

```

Calculating the area under the curve will require a function to approximate the eigenvalues as a function of band number, as indicated in line 2 of the above algorithm. A description of Hermite splines, which are recommended given their suitability to this particular application, is included in Section 3.1. The assignment of bins occurs in line 9, and warrants further explanation. Starting with the first band, the area under the curve is calculated as

$$\int_1^2 C(b)db.$$

The total area under the curve through band i is therefore

$$\int_1^{i+1} C(b)db.$$

The results of previously calculated integrals are stored in *sarea* to prevent redundant calculations. Taking the ceiling of the cumulative area under the curve divided by the area per bin results in bands one through $B - 1$ being placed in bins one through nb , and band B is placed in the same bin as $B - 1$. Line 10 continues with the conversion of a bin number to the size of a spatial filter kernel that corresponds to bin number. The bands in bin one should have no filter (equivalent to a kernel of size one) applied, and the bands in bin two should have a 3×3 filter applied.

This approach is valid for convex eigenvalue curves that are similarly shaped to Fig. 1, which is usually the case. The properties of the MNF dictate that the eigenvalue function is strictly decreasing, but in the event that the eigenvalue function is not convex, dividing up the area under the curve of the derivative of the function will group the most similar eigenvalues and their corresponding bands together. Consider finding the area under the curve of the derivative:

$$\int_a^b f'(x)dx = f(b) - f(a)$$

according to the fundamental theorem of calculus. This calculation requires no approximation of the function as the actual functions' values can be used. Because the eigenvalue function is monotonically decreasing, the area underneath the curve will be negative, and therefore the area will be negated to produce a positive result necessary for bin determination. The above algorithm is modified to produce the following adaptive filter with derivative (AFD) algorithm using the area underneath the curve of the derivative to determine the location of bins.

Algorithm AFD(M, Λ, nb)

input/output:

M (image transformed to MNF coordinates,
filtered upon exit)

input:

Λ (B eigenvalues where B is number of bands in M)

nb (number of bins to use)

```

1  begin
2     $area := \Lambda(1) - \Lambda(B)$ 
3     $area\_per\_bin := \frac{area}{nb}$ 
4     $sarea := 0$ 
5    for  $i := 1$  step 1 until  $B - 1$  do
6      begin
7         $sarea := sarea + \Lambda(i) - \Lambda(i + 1)$ 
8         $bin := \text{ceiling} \left( \frac{sarea}{area\_per\_bin} \right)$ 
9         $kernel_i := 2 \cdot (bin - 1) + 1$ 
10     end
11     $kernel_B := kernel_{B-1}$ 
12    for  $i := 1$  step 1 until  $B$  do
13      apply  $kernel_i \times kernel_i$  median filter
        to  $band_i$  in image  $M$ 
14  end

```

Finally, either of the above variations on the adaptive filtering algorithm may be used on a particular range of bands. For example, if prior knowledge or analysis of the MNF transformed dataset indicates that there is no usable signal beyond a specific band, the MNF image can still be truncated and filtered adaptively. The value of B would be changed from the total number of bands in the image to the number of bands desired after truncation. This is different from simply truncating the MNF because the bands that are kept would be filtered to decrease the noise, and the number of bands kept could be larger to ensure that very little signal is lost in the truncation.

A. Hermite Splines

The filtering algorithm requires a function that approximates the eigenvalue curve generated by the MNF. Cubic splines are piecewise cubic polynomials that produce a visually appealing curve and interpolate a given set of points. In particular, Hermite cubic splines have only one continuous derivative (standard cubic splines have two) and produce a monotone cubic spline curve interpolating a monotonic function, rendering this type of spline ideal for interpolating the monotonic SNR curve. The Hermite cubic spline $C(x)$ is composed of $2n$ basis functions, $c_i(x)$, $\hat{c}_i(x)$, $i = 1, \dots, n$, where n is the number of interpolation points. The function

$$C(x) = \sum_{i=1}^n y_i c_i(x) + d_i \hat{c}_i(x)$$

interpolates the points (x_i, y_i) , $i = 1, \dots, n$ if

$$\begin{aligned} c_i(x_i) &= 1, c_i(x_j) = 0, & j \neq i, \\ \hat{c}_i(x_j) &= 0, & \text{for all } i, j. \end{aligned}$$

Furthermore,

$$\begin{aligned} c'_j(x_i) &= 0, & \text{for all } i, j, \\ \hat{c}'_i(x_i) &= 1, \hat{c}'_i(x_j) = 0, & j \neq i, \end{aligned}$$

making

$$C'(x_i) = d_i.$$

Only x_i , y_i , and d_i , $i = 1, \dots, n$ are required to define a Hermite cubic spline, and the d_i are chosen to make $C(x)$ monotone (theoretically always possible for monotone data y_i). Refer to [11] for a more detailed description of Hermite cubic splines including definitions of the basis functions $c_i(x_i)$, $\hat{c}_i(x_i)$.

The derivative and the definite integral of Hermite cubic splines can be easily obtained as the cubic polynomials (and basis functions) are easily differentiated or integrated analytically. Included in [11] is a set of subroutines designed to define, evaluate, and integrate Hermite cubic splines, PCHEZ, PCHEV, and PCHQA, respectively. PCHEZ defines continuous derivatives, d_i , that result in a visually appealing function, PCHEV evaluates the function and the derivative at a set of points, and PCHQA returns the definite integral of the function between two points, a and b .

IV. STUDY AREA

The study area (as described by [12]) is located in the Appomattox Buckingham State Forest in Virginia, USA. Three flight-lines of Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) 224-band imagery were acquired in the winter of 1999. The spatial resolution of this dataset is 3.4 m, and the spectral resolution is 10 nm, with the AVIRIS sensor range being 400–2500 nm contiguously. Training data consisted of 142 field collected locations surrounded by homogeneous areas of single pine species with differentially corrected global positioning system (GPS) coordinates. Three pine species, loblolly (*Pinus taeda*), shortleaf (*Pinus echinata*), and Virginia pine (*Pinus virginiana*), with 64, 30, and 48 locations, respectively, were collected in August of 1999. The image (shown in Figs. 10–13 and hereafter referred to as ABSF) contains various tree stands that include the three species of pines listed above, hardwoods, and mixed (evergreens and hardwoods).

V. SNR ESTIMATION

To support the conclusion that the adaptive filter is improving the image, evidence is provided to show that the filter is reducing the noise without drastically reducing the signal, effectively improving the SNR. Estimating the SNR in an image is nontrivial, as most noise estimates are variance-based and the naturally occurring variance within an image will lead to overestimated noise. In order to minimize the impact of signal variance on noise estimates, homogeneous portions within an image are identified for noise estimation. Curran and Dungan proposed SNR estimation of AVIRIS images using a homogeneous line within an image [13]. Gao proposed an alternative method that does not require identifying homogeneous regions within an image, but instead divides the entire image into small blocks and uses local standard deviations to estimate the standard deviation of the noise for the entire image [14]. Smith and Curran found both methods effectively estimated SNRs in AVIRIS images, although both methods overestimated SNR [5]. As the study area used for this paper lacks clearly defined homogeneous regions, Gao’s whole image method of SNR estimation will be used, as it does not require identification of homogeneous regions and it can be easily automated [14].

Gao’s method assumes that within a small block (4×4 or 8×8) the local standard deviation is either low because of noise and a small amount of natural variance or high because the block contains edges, etc. An image band is broken into small blocks of equal size, and the local standard deviation is calculated for each block. A histogram is created from the local standard deviations, and the most frequently occurring standard deviation provides a reasonable estimate of the standard deviation of the noise for the entire band. Gao generated data with a known amount of noise and showed that this method accurately estimated the noise [14].

In this SNR estimation method, the signal in AVIRIS images is generally estimated using the mean values of each band, however the mean values of MNF transformed bands do not have the same magnitude as the mean values in the original coordinate system. The original digital numbers within the image are all positive, and MNF digital numbers can be positive or negative, resulting in mean values that are much lower in the resulting MNF coordinate system. In order to study the effects of filtration on SNRs in MNF coordinates, a estimate suggested by Schowengerdt is used:

$$SNR = \frac{\sigma_S^2}{\sigma_N^2},$$

where σ_S^2 is the variance of the signal and σ_N^2 is the variance of the noise [15]. The variance of the signal is estimated by the variance of the entire image, and the variance of the noise is determined using the method described previously. Note that this method is not being suggested to estimate SNRs in order to evaluate whether an image can be used for a certain application, or to compare an image to other images whose SNRs are estimated using other procedures. This estimation of the signal, noise, and SNR is needed only to compare unfiltered and filtered data to provide evidence that the SNRs are improving as a result of filtration.

The behavior of the SNR was evaluated using the methods described above, and the test image was transformed to MNF coordinate space and filtered using a 3×3 , 5×5 , 7×7 , and 9×9 median filter. Notice in Fig. 3 that as the filter size increases, the SNR also increases compared to the unfiltered image (1×1). Also, notice in Figs. 4 and 5 that since

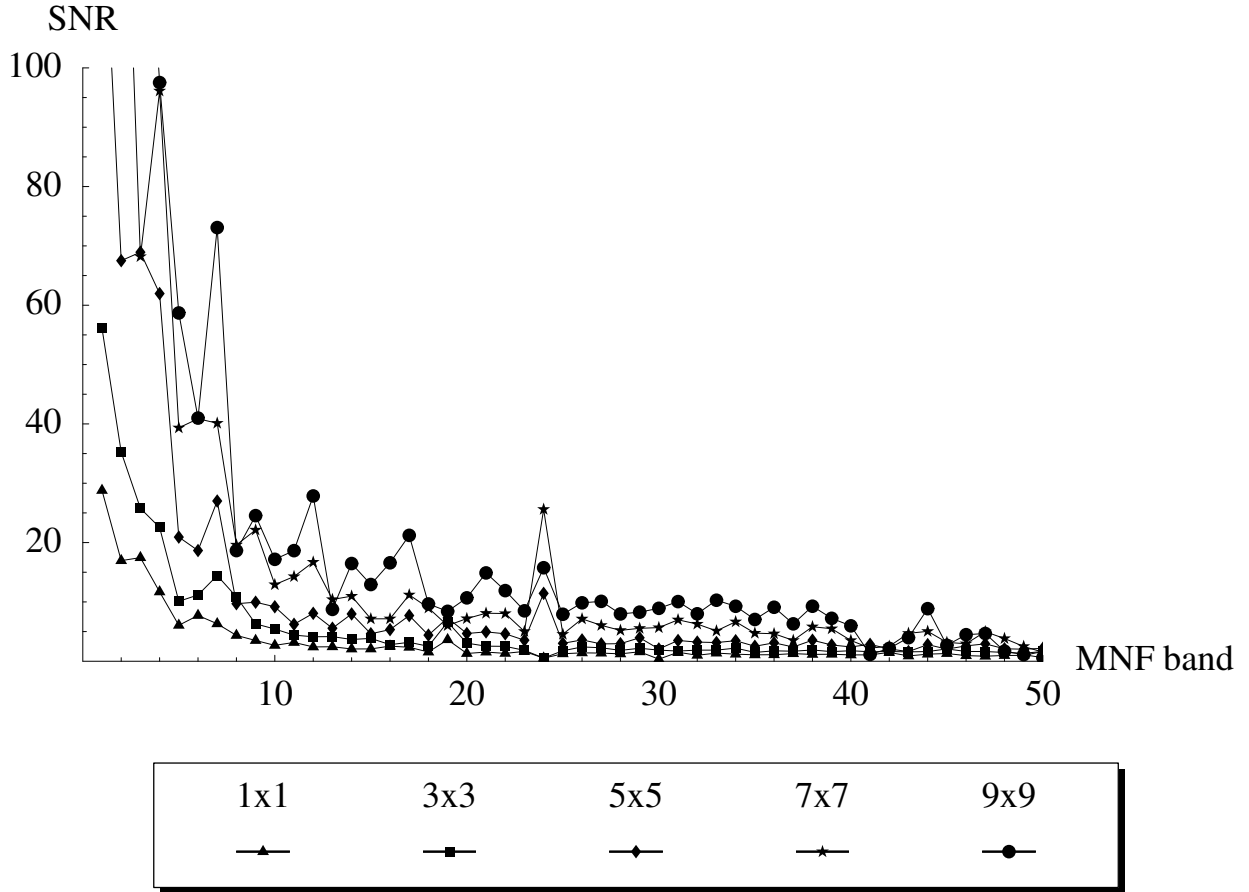


Fig. 3. SNR estimates of MNF bands 1–50 unfiltered (1×1) and after 3×3 through 9×9 median filters applied.

the magnitudes of the signal are very different in bands one and 50, degrading the signal in the first band has more impact on the total image signal degradation than removing some signal in the 50th band.

VI. RESULTS

The AF, AFD, truncated AF, truncated AFD, and uniform 9×9 median filters were applied to the MNF of ABSF. Results for the 9×9 filter are reported as the 9×9 filter produced the best classification accuracies for all uniform spatial filters. Further evidence of improved image quality as a result of adaptive filtering is provided in a classification of ABSF to identify loblolly, shortleaf, and Virginia Pines. An advantage of the technique used (discriminant analysis) is that individual bands are selected by the method, showing that high order noisy MNF bands contain signal that impacts applications such as classification.

A. Bin Creation

The adaptive filtering algorithms (using the approximation of the curve and using the derivative of the curve to establish bin locations) described above were applied to ABSF. Fig. 6 shows the eigenvalues (SNRs) generated by the MNF. Figs. 7 and 8 show the resulting eigenvalue function and derivative curves (for the entire image and for the image truncated to 80 MNF bands) approximated by Hermite splines, with the area underneath the respective curves divided into five equal partitions, representing median filter kernels of size one through nine.

B. Classification

A discriminant analysis was used to generate classification functions for each filtered and unfiltered dataset using SAS^(R) 9.1.3 [16]. First the STEPDISC procedure was used to identify bands that contribute to the discrimination

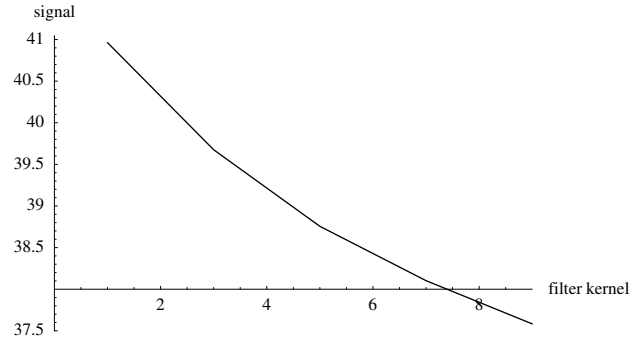


Fig. 4. Signal estimate of MNF band 1 with various sized median filters applied.

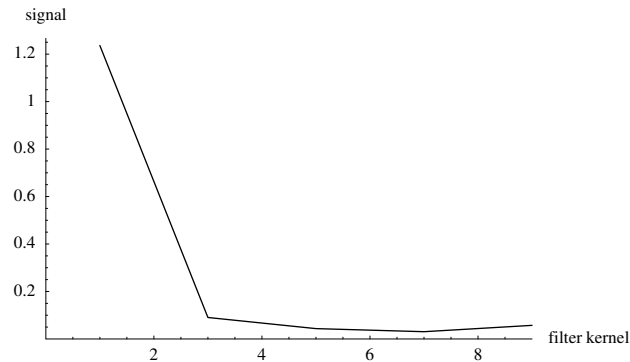


Fig. 5. Signal estimate of MNF band 50 with various sized median filters applied.

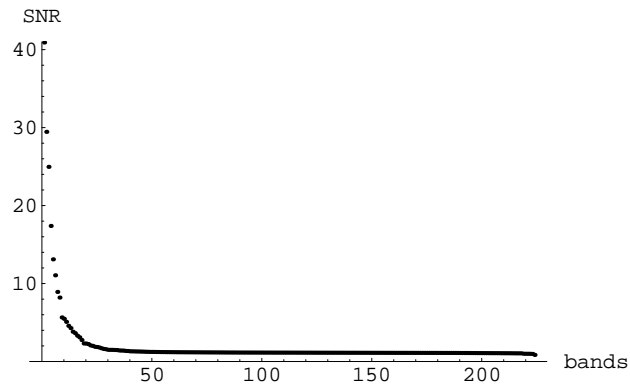


Fig. 6. Study area MNF SNRs (eigenvalues).

between the three pine species. The STEPDISC procedure (using stepwise selection) starts with no variables in the model. The variable that contributes most to species discrimination enters the model if the significance level of an F-test is above the specified threshold. In each subsequent step, the variable within the model that contributes least to discrimination is first considered for removal (if the significance level of an F-test is beyond the specified threshold to leave the model). If no variable is removed, of the remaining variables not included in the model, the variable that best contributes to the discrimination of the model is added, if it meets the criteria to enter. At each step, either one variable leaves the model or one variable enters the model until the procedure terminates. The procedure terminates when no more variables are eligible to enter or leave the model, or a maximum number of steps is reached.

DISCRIM generates a discrimination function (classification) when given a set of quantitative variables and corresponding classifications for each observation. Classification accuracies are determined by using the generated discriminant function to classify each observation, which is a biased test. A better measure of accuracy is the

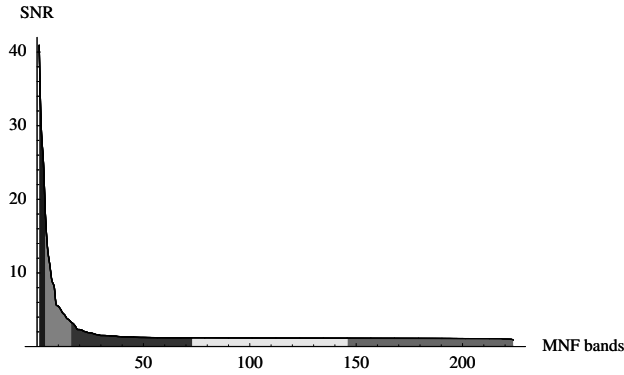


Fig. 7a. Bins generated by AF on study area.

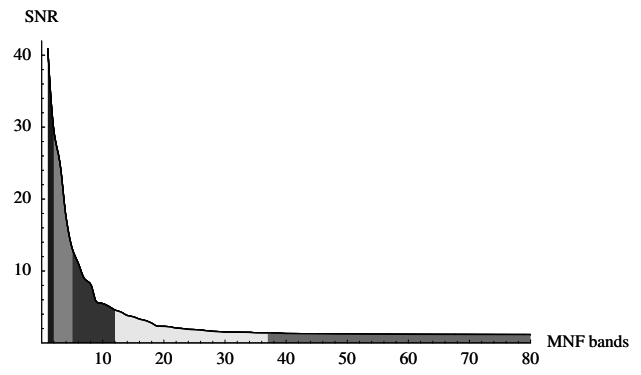


Fig. 7b. Bins generated by AF (truncated to 80 MNF bands) on study area.

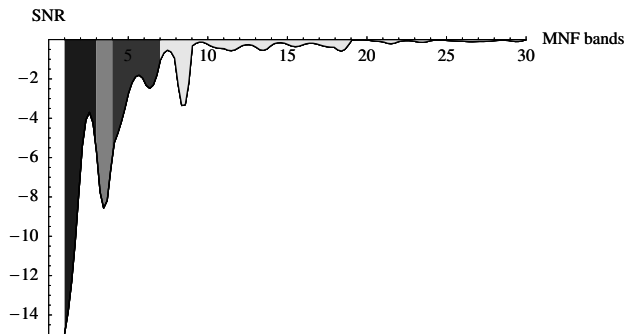


Fig. 8a. Bins generated by AFD on study area, bands 1–30 shown.

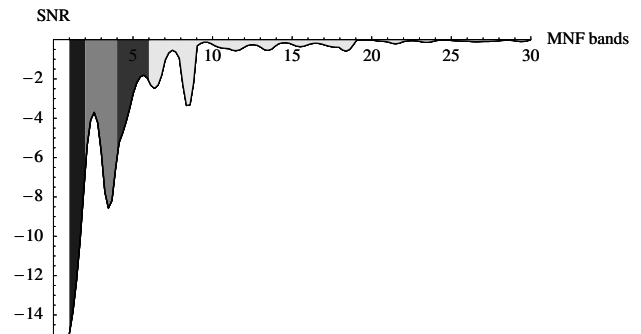


Fig. 8b. Bins generated by AFD (truncated to 80 MNF bands) on study area, bands 1–30 shown.

cross validation accuracy, where each observation is classified using the classification model derived from the other observations (not including itself).

The classification results in this paper were generated using STEPDISC with stepwise selection, and the variables (bands) identified were used to generate the discriminant function using DISCRIM. The default value of .15 was used as the threshold for entering and leaving the model in STEPDISC, and the default assumption of normality was used to generate the discriminant function in DISCRIM. In order to select different numbers of hyperspectral bands, the α -level for the F-test was varied between .1 and .00001 for different classifications reported in Tables 1 and 2. The data and discriminant analysis procedure are identical to those used to obtain maximum cross validation classification accuracies of roughly 85% using a maximum of 10 bands in [12].

The first set of entries in Table 1 correspond to classifications applied to transformed data in the MNF coordinate system. Accuracies for MNF data with no filter applied are recorded for comparison purposes. The second set of Table 1 entries are classification accuracies and parameters using data that was filtered in the MNF coordinate system and then inverse transformed to the original coordinate system. The classifications in Table 2 are applied to data that are transformed to MNF space, truncated or truncated and filtered to remove noise, and inverse transformed. The unfiltered MNF data are truncated at 40 or 60 MNF bands because these thresholds appear to be located near where the eigenvalue curve approaches one (see Fig. 6). As more noisy bands are included, the accuracies of the classification of truncated MNF data decrease, and therefore results are not reported for the inverse of MNF data truncated at 80 bands, corresponding to the adaptive filters applied to MNF truncated 80 band images.

VII. DISCUSSION

Classification results in Table 1 indicate that using any of the mentioned variations of adaptive filtering ultimately improves the classification accuracy over not using any filtering method, indirectly indicating that the adaptive filter has improved the image quality, enabling better classification results. Furthermore, considering the first six table entries

Table 1. Results of stepwise discriminant analysis and discriminant analysis on study area using various filtering schemes in both original and MNF coordinates (Legend: CV Acc. = Cross Validation Accuracy, Class Acc. = Classification Accuracy).

Filter	CV Acc. (%)	Class Acc. (%)	α -level	# Bands	Bands
MNF coordinates					
none	85.92	88.03	.01	9	1,2,3,5,7,13,16,18,173
AF	91.55	93.66	.001	13	1,2,6,7,9,10,14,16,19,36,44,52,148
AFD	98.59	100.00	.00001	11	5,10,12,16,18,19,34,36,45,48,53
Trunc. AF	98.59	98.59	.0001	10	2,7,14,16,19,29,36,44,45,53
Trunc. AFD	98.59	100.00	.00001	11	5,10,12,16,18,19,34,36,45,48,53
9x9	97.89	100.00	.00001	11	5,10,12,16,18,19,34,36,45,48,53
original coordinates					
none	76.76	79.58	.0001	4	19,24,116,213
9x9	91.55	92.25	.00001	4	3,19,27,164
9x9	90.85	92.96	.0001	5	3,19,26,27,164
AF	83.10	83.80	.001	5	19,26,27,126,158
AFD	91.55	92.25	.00001	4	3,19,27,164
AFD	87.32	92.25	.0001	6	4,9,19,23,27,44
9x9	95.07	97.89	.001	11	3,19,23,37,105,120,140,144,147,183,200
AF	92.25	97.89	.01	14	11,16,19,26,27,30,43,120,158,161,183,194,197,222
AFD	92.96	96.48	.01	10	4,9,19,23,27,40,41,42,164,218
AFD	93.66	95.07	.001	8	4,9,19,23,27,41,44,164

Table 2. Results of stepwise discriminant analysis and discriminant analysis on study area using various truncated filtering schemes in original coordinates (Legend: CV Acc. = Cross Validation Accuracy, Class Acc. = Classification Accuracy).

Filter	# MNF Bands Kept	CV Acc. (%)	Class Acc. (%)	α -level	# Bands
none	40	88.03	92.25	.1	12
none	60	84.51	88.03	.1	14
AF	80	95.78	97.89	.01	12
AFD	80	93.66	94.37	.01	10
none	60	73.24	77.47	.001	4
AF	80	86.62	89.45	.0001	4
none	40	83.10	85.92	.01	8
AFD	80	94.37	95.07	.001	7

corresponding to classification in the MNF axes, the more accurate classifications make use of high order MNF bands. The least accurate classification (applied to unfiltered data) primarily included bands from the first 20 MNF bands, whereas the more accurate classifications consistently used higher order bands in the 30–55 range. Many of the bands identified to be important for pine species discrimination are consistent across filtering schemes, but are not present in the classification of the unfiltered dataset. These results support the hypothesis that there is signal in higher order MNF bands that can be important for applications such as classification, and that applying a filter to reduce the levels of noise can make the weak signal in these bands usable.

Recall that the MNF maximizes SNR in each band where the signal is estimated by the variance of the image, making the transformation similar to PCA, which aligns the data along axes of maximum variance. In applications such as this where the goal is discrimination between pine species, higher order MNF bands (and PCA bands) might be expected to play a vital role as the variance between pine species is likely small. Unfortunately these high order bands in the MNF transformation that have little signal (variance) have relatively large amounts of noise. Simply truncating the MNF to remove this noise will result in also removing signal.

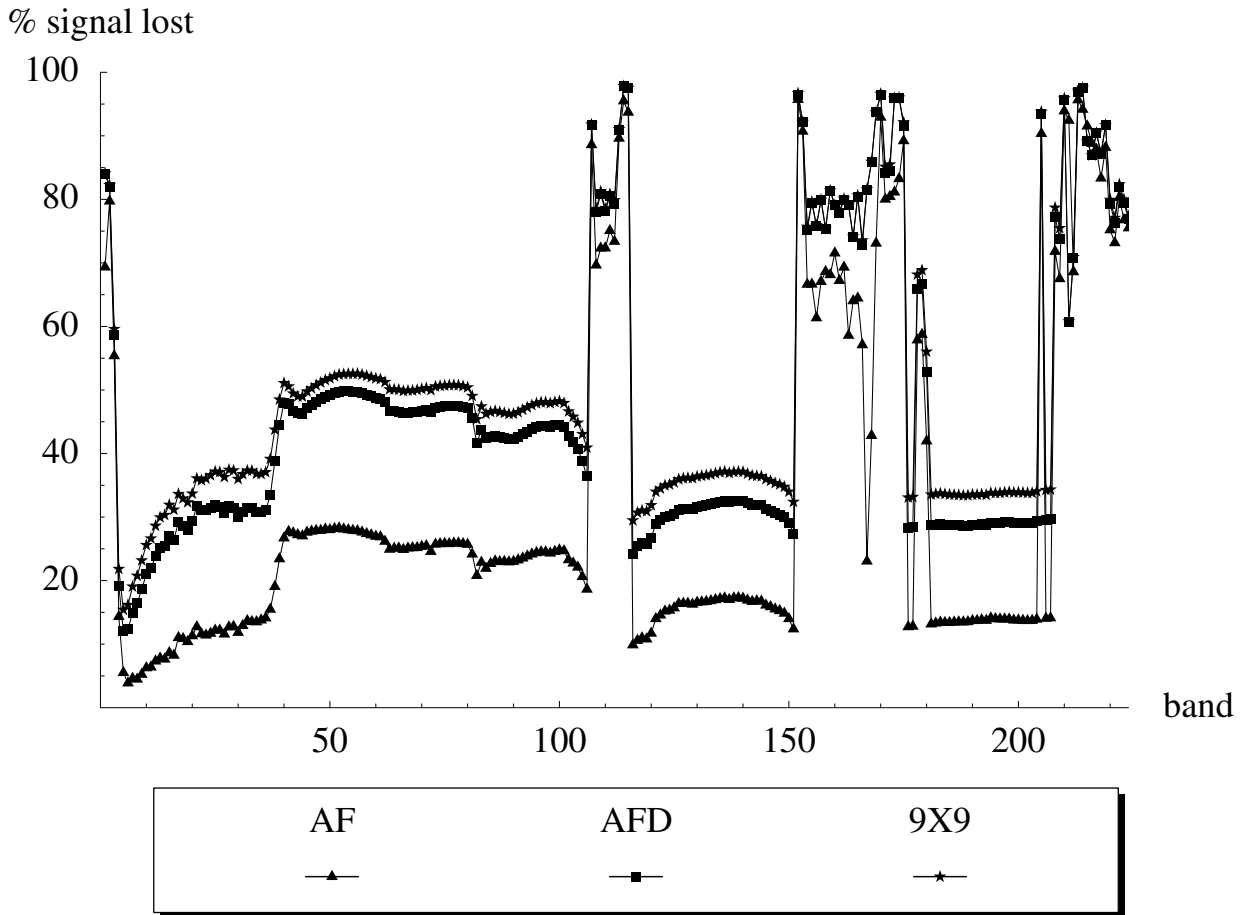


Fig. 9. Percentage of signal lost by AF, AFD, and 9×9 median filters.

The second set of entries in Table 1 corresponding to classification results in the original coordinate system show that the improved MNF image quality is consistent when the image is inverse transformed. This set of classification accuracies are not as high as those corresponding to the MNF coordinate system, even when the same number of bands is used to build each of these models. Although the numbers of bands used in these classifications are similar to those used on the MNF coordinate data, the MNF coordinate data still results in better classifications. However, some classifications require that data not be transformed, and the results in Table 1 demonstrate that the MNF-based adaptive filtering methods are relevant to images that are inversely transformed from MNF coordinate space.

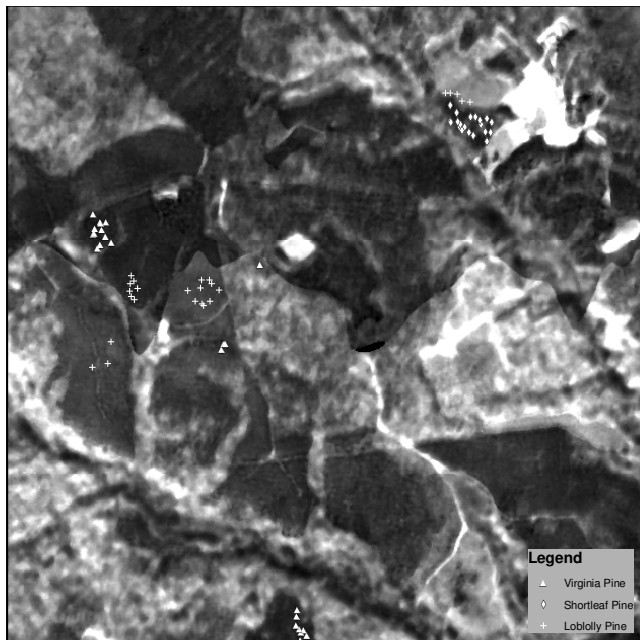
For this particular dataset, the AFD usually produced better classifications (indicated by classification accuracy and cross validation classification accuracy) than the AF, however, for other applications, the AF may be the better choice. The following discussion shows that preserving image quality (signal) and achieving high classification accuracy are *not* equivalent — AF has better image quality (more preserved signal), but worse classification/cross validation accuracy, than AFD. Referring back to Figs. 7 and 8, recall that using the derivative curve to determine bin placement resulted in larger spatial filters being applied to more bands within the image. In fact, the classification results for the 9×9 filter are consistently similar to results obtained using the AFD filter because in this case, the filters are very similar. As alluded to earlier, the low order bands containing most of the signal (variance) of the entire image likely do not greatly aid discrimination between spectrally similar pine species. These filters are substantially degrading the signal in the image, as shown in Fig. 9. Because the first MNF bands contain the majority of the signal, applying a filter with such a large window drastically degrades the signal of the overall image. Fig. 9 shows that while the AF filter decreases the variance of the original image by around 20% for most bands, the AFD and the 9×9 filters decrease the variance by around 40% for many bands. There is a noticeable difference in the signal (variance) degradation between the AF,



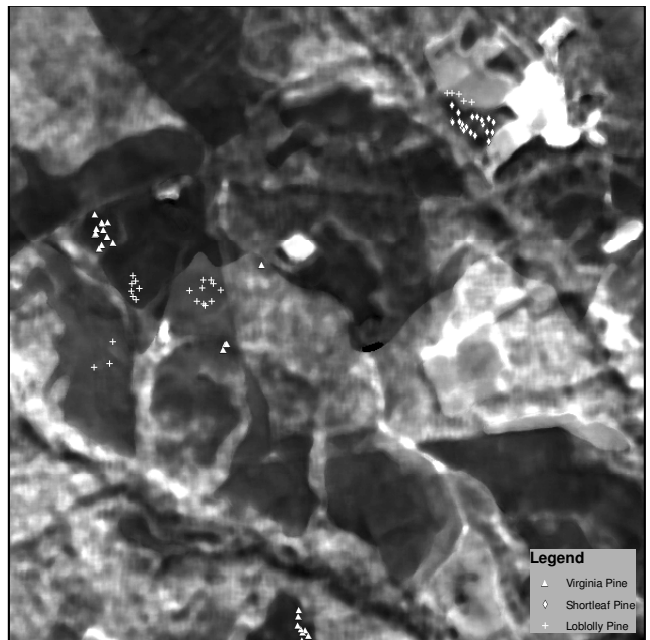
a. Unfiltered.



b. AF filtered.



c. AFD filtered.

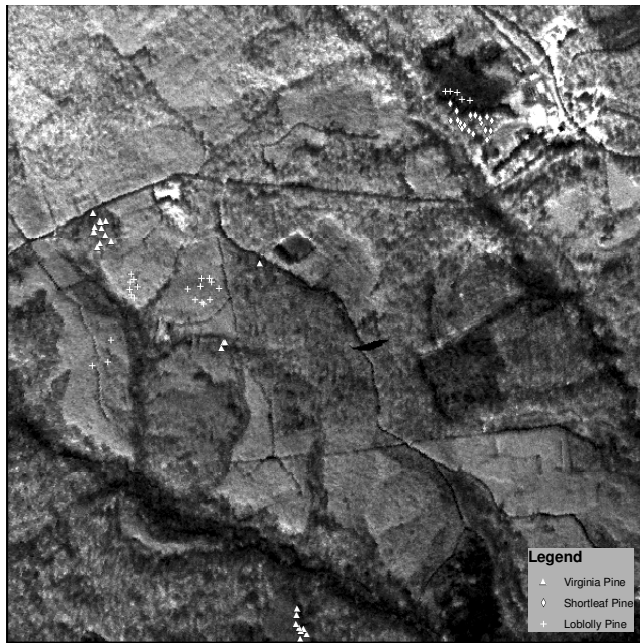


d. 9x9 filtered.

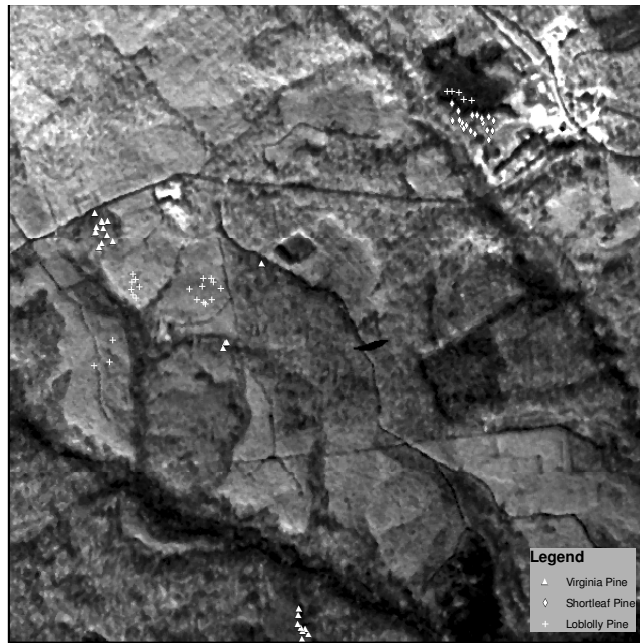
Fig. 10. Band 27 of ABSF in the original coordinate system (zoomed).

AFD, and 9×9 filters, indicating that using the AFD and, especially, a large uniform filter substantially reduces the image signal.

Further evidence of image quality can be observed qualitatively by viewing and comparing the filtered images to the original image and the unfiltered MNF image. Fig. 10 contains band 27 of the image in the original coordinate system, a band that was used for the pine species discrimination in many of the images. While Fig. 10b (the inverse of the MNF ASF image) does not appear as crisp as the original image, the texture is much greater than that of either Fig. 10c (ASFD) or 10d (9×9). Fig. 11 compares the four filtering schemes applied to band 5 of the MNF image, providing insight into the obvious signal degradation in band 27. While the AF image (Fig. 11b) is relatively crisp (3×3



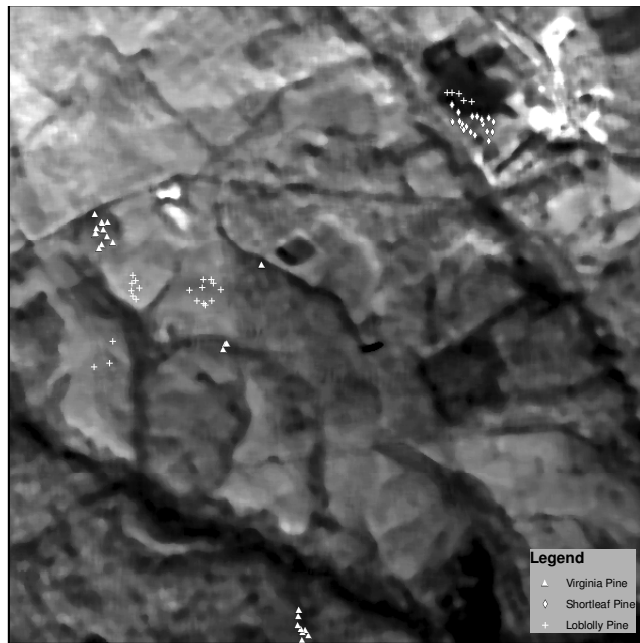
a. Unfiltered.



b. AF filtered (3x3).



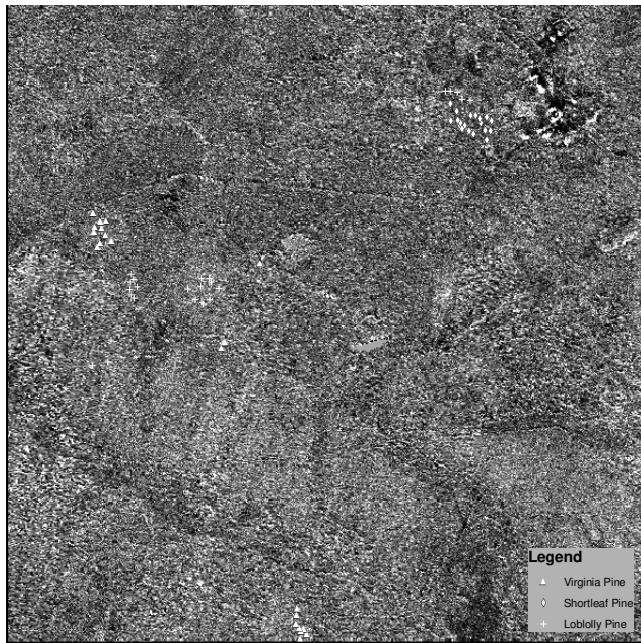
c. AFD filtered (7x7).



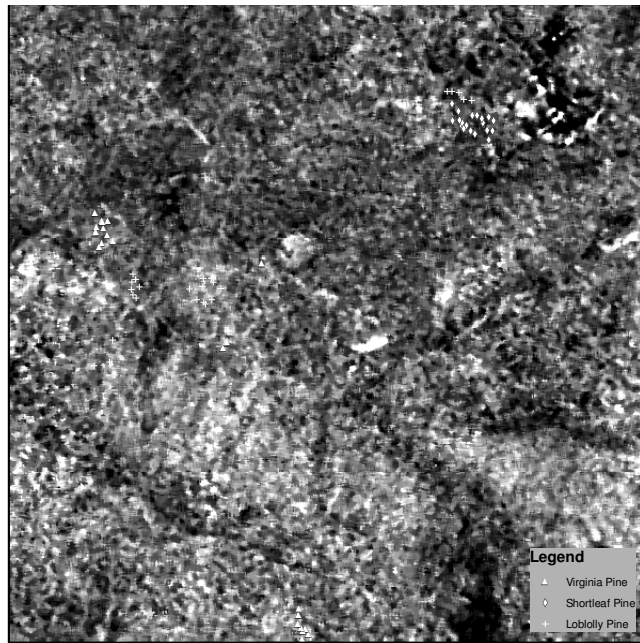
d. 9x9 filtered.

Fig. 11. Band 5 of the ABSF MNF (zoomed).

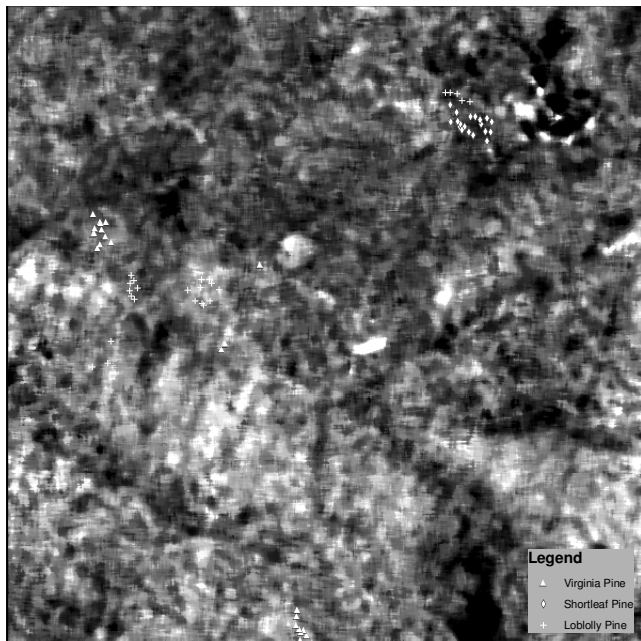
filter used), the AFD image (7x7) and the 9x9 image have substantially degraded signal, resulting in substantially degraded signal for the overall image (band 5 contains more of the image's signal than each subsequent band). A much higher order band, band 36, is shown in Fig. 12. This band was selected for each MNF filtered classification, but was not part of the less accurate classification of the unfiltered image. Although it is possible to distinguish a small signal in Fig. 12a, the noise is clearly dominating this band. The spatial filters are reducing the signal of the image, but importantly reduce the noise to a level that reveals variance between individual portions of the image. The areas marked by ground truth can be spectrally distinguished in this band (which is important to the classification) even if the signal does not appear to be strong. While this band was important for this particular image and classification, the



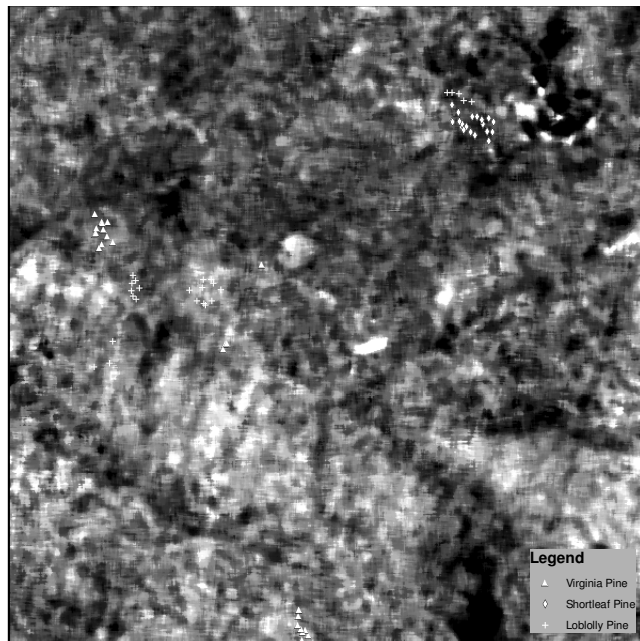
a. Unfiltered.



b. AF filtered (5x5).



c. AFD filtered (9x9).



d. 9x9 filtered.

Fig. 12. Band 36 of the ABSF MNF (zoomed).

obvious signal contained in this noisy band supports the claim that high order MNF bands contain signal that may be important to an application, but the signal is difficult to use without reducing the noise. The noisy band 36 was not selected to discriminate pine species in the unfiltered image (see Table 1), but was selected in *all* of the filtered MNF images. Fig. 13 compares a 9x9 filter for band 36 to the original band 36 for the full scene, showing that the variance between features in the image is noticeable once the level of noise is reduced.

Evidence supporting adaptive filters over simply truncating the MNF to reduce the noise in an image is shown in Table 2. The classification accuracies for the MNF truncated, unfiltered images are consistently less than the classification accuracies in Table 1 for both filtering schemes when comparing classifications built using similar numbers

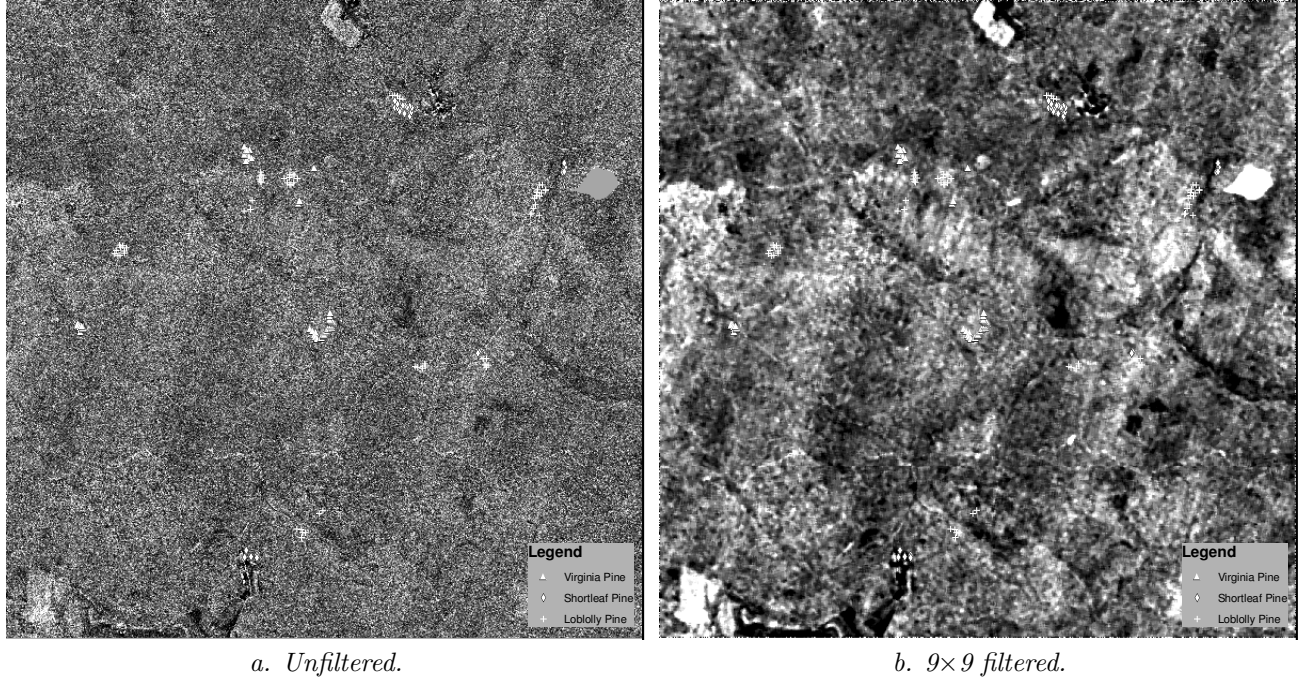


Fig. 13. Band 36 of the ABSF MNF (full scene).

of bands. Combining the truncation of MNF with an adaptive filtering method to entirely remove bands that have no discernible signal while filtering bands containing signal heavily degraded by noise may be another viable noise reduction technique. Since the truncated MNF will be filtered in this case, the number of bands kept can be higher to ensure that little or no signal is removed from the image. Notice in Table 2 that each adaptive filtering method using 80 MNF bands produces far better classification accuracies than keeping 40 or 60 unfiltered MNF bands. Also notice that accuracies are lower for the image truncated at 60 bands than the image truncated at 40 bands. Bands 40–60 have low SNRs, and including these bands without filtering them introduces far more noise than signal to the image, resulting in lower classification accuracies.

Figs. 10–13 show why adaptive filtering is necessary over one uniform filter size for data transformed by the MNF. While a large filter degrades strong signal in low order bands, it can be useful to remove noise from a high order band with a weak signal. Even though the weak signal is inevitably degraded, it is practically useless when the noise is comparatively strong. Using an adaptive filter allows the strong signal in the first few bands to be preserved while the dominating noise is removed by a much larger filter in high order bands. The results presented for this particular dataset are intended to demonstrate a general technique, and are not intended to indicate the ideal filter sizes and number of bins for other datasets and applications. This particular dataset and classification appeared to benefit from spatial filters with large windows, explaining why the AFD filter with a maximum filter window size of nine (very similar to a uniform 9×9 filter) produced such accurate classification results.

VIII. CONCLUSIONS AND FUTURE WORK

This work introduced an adaptive filter based on the MNF that exploits the ordering of the bands to apply median filters of different sizes. This filtering scheme greatly enhanced the MNF image for the purposes of identifying pine species, and accuracies were improved by more than 10% for certain variations of the filtering algorithm applied to AVIRIS data in the original and MNF coordinate systems. The results in this paper are substantially more accurate than previously reported results for the same application and analysis performed on the same data in which no spatial filters were applied [12].

The AFD version of the adaptive filter produced more accurate classification results and higher estimated SNRs than the AF version, however, there are indications that the AFD degraded the signal quality significantly (while also degrading the noise), perhaps making the AF more appropriate in certain applications. Both variations led to classifications that were substantially better than classifications performed on unfiltered data. This paper does not indicate how the number of bins should be selected. More work is needed to identify the number of bins and the

maximum size of filters that should be used for certain applications and on certain imagery. Furthermore, more study is necessary to determine the relationship between SNR estimates and ideal filter size, and the suitability of adaptive frequency domain filters in this context should be examined.

REFERENCES.

- [1] B.L. Dickson, "Recent advances in aerial gamma-ray surveying," *Journal of Environmental Radioactivity*, vol. 76, no. 1-2 pp. 225–236, 2004.
- [2] A.A. Green, M. Berman, P. Switzer, and M.D. Craig, "A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, January, 1988.
- [3] J.B. Lee, A.S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 295–304, May, 1990.
- [4] D.A. Landgrebe, and E. Malaret, "Noise in remote-sensing systems: the effect on classification error," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no. 2, pp. 294–300, March, 1986.
- [5] G.M. Smith and P.J. Curran, "The signal-to-noise (SNR) required for the estimation of foliar biochemical concentrations," *International Journal of Remote Sensing*, vol. 17, no. 5, pp. 1031–1058, March, 1996.
- [6] N.A. Quarmby, "Noise removal for SPOT HRV imagery," *International Journal of Remote Sensing*, vol. 8, no. 8, pp. 1229–1234, August, 1987.
- [7] M. Lennon, G. Mercier, and L. Hubert-Moy, "Nonlinear filtering of hyperspectral images and anisotropic diffusion," in *IGARSS 2002: IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing, Vols I-VI, Proceedings - Remote Sensing: Integrating Our View of the Planet*, pp. 2477–2479, 2002.
- [8] G. Pok, J. Liu, and A.S. Nair, "Selective removal of impulse noise based on homogeneity level information," *IEEE Transactions on Image Processing*, vol. 12, no. 1, pp. 85–92, January, 2003.
- [9] M.A. King, P.W. Doherty, R.B. Schwinger, D.A. Jacobs, R.E. Kidder, and T.R. Miller, "Fast count-dependent digital filtering of nuclear medicine images: concise communication," *Journal of Nuclear Medicine*, vol. 24, no. 11, pp. 1039–1045, 1983.
- [10] A.B. Carlson, P.B. Crilly, and J.C. Rutledge, *Communication Systems: An Introduction to Signals and Noise in Electrical Communication, fourth ed.*, New York: McGraw-Hill, 850 pp., 2002.
- [11] D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software*, New Jersey: Prentice Hall, pp. 100–113, 1989.
- [12] J.A.N Van Aardt and R.H. Wynne, "Examining pine spectral separability using hyperspectral data from an airborne sensor: An extension of field-based results," *International Journal of Remote Sensing*, vol. 28, no. 2, pp. 431–436, 2007.
- [13] P.J. Curran and J.L. Dungan, "Estimation of signal-to-noise: a new procedure applied to AVIRIS data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, no. 5, pp. 620–628, September, 1989.
- [14] B. Gao, "An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers," *Remote Sensing of Environment*, vol. 43, pp. 23–33, 1993.
- [15] R.A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing, second ed.*, San Diego: Academic Press, pp. 127–135, 1997.
- [16] SAS Institute Inc., *SAS OnlineDoc^(R)*, Version 8, Cary, NC: SAS Institute Inc., 1999.