

# Wavelets for Computationally Efficient Hyperspectral Derivative Analysis

Lori Mann Bruce, *Member, IEEE*, and Jiang Li, *Student Member, IEEE*

**Abstract**—Smoothing followed by a derivative operation is often used in the analysis of hyperspectral signatures. The width of the smoothing and/or derivative operator can greatly affect the utility of the method. If one is unsure of the appropriate width or would like to conduct analysis for several widths, scale-space images can be used. This paper shows how the wavelet transform modulus-maxima method can be used to formalize and generalize the smoothing followed by derivative analysis and how the wavelet transform can be used to greatly decrease computational costs of the analysis. The Mallat/Zhong wavelet algorithm is compared to the traditional method, convolution with Gaussian derivative filters, for computing scale-space images. Both methods are compared on two points: 1) computational expense and 2) resulting scalar decompositions. The results show that the wavelet algorithm can greatly reduce the computational expense while practically no differences exist in the subsequent scalar decompositions. The analysis is conducted on a database of hyperspectral signatures, namely, hyperspectral digital image collection experiment (HYDICE) signatures. The reduction in computational expense is by a factor of about 30, and the average Euclidean distance between resulting scale-space images is on the order of 0.02.

**Index Terms**—Computation time, computer-aided analysis, remote sensing, signal processing, wavelet transforms.

## I. INTRODUCTION

DERIVATIVES play an important role in the analysis of spectral signatures. For many years, derivative analysis has been successfully applied to spectroscopy in analytical chemistry. More recently, derivatives have been applied to remotely sensed hyperspectral signatures. However, differences arise due to the fact that the remotely sensed data is not collected under controlled laboratory conditions. Viewing geometry, illumination, atmospheric effects, and spatial resolution are a few of the conditions that cause degradation of the remotely sensed data. One result of the degradation can be an additive high-frequency noise in the signature. This particularly causes havoc with derivative analysis since derivatives act as high-pass filters in the frequency domain. In order to overcome the high-frequency noise problem, many researchers have used smoothing prefilters, or other lowpass filters, before applying the derivative-based methods to the spectral signatures. When utilizing these types of methods, the widths of the derivative window and the smoothing window are critical. Typically, these widths have been selected experimentally and have been

application dependent. If one is unsure of the appropriate choice of window width or would like to conduct analysis for several window widths, a tool known as scale-space imaging can be used. Scale-space imaging is intimately related to wavelet transforms. This paper shows how the wavelet transform modulus-maxima method can be used to formalize and generalize the smoothing followed by derivative analysis and how the wavelet transform can be used to greatly decrease computational costs of said analysis.

Section II describes the use of derivative-based analysis in hyperspectral data. Section III describes the scale-space image and spectral fingerprint and how they are traditionally implemented. In Section IV, the authors describe the corresponding wavelet decomposition and the modulus-maxima representation and how the Mallat/Zhong algorithm can be used as an efficient implementation. Section V compares the traditional implementation of the scale-space image and the Mallat/Zhong wavelet algorithm in terms of their orders of complexity and experimental benchmarking. In Section VI, the authors compare the scale-space images produced by the two methods to determine how much error is introduced by using the less computationally expensive wavelet algorithm. Sections VII and VIII, respectively, provide a discussion of the results and conclusions drawn from this study.

## II. DERIVATIVE-BASED ANALYSIS

In remotely sensed multispectral/hyperspectral signatures, first, second, and higher order derivatives have been used for analysis and recognition of vegetation, soil, minerals, etc. Campbell and Esaias used second-order derivatives of remotely sensed spectral signatures for estimating oceanic chlorophyll concentrations [1]. Demetriades-Shah *et al.* used first derivatives to detect the “red edge” in spectral signatures and to decorrelate soil background from vegetation spectra [2]. They also used second derivatives of canopy reflectance spectral signatures as measurements for estimating the concentration of leaf chlorophyll. Blackburn also utilized derivatives of remotely sensed spectra for quantification of plant chlorophyll [3]. He showed that the first and second derivatives of a pseudo-absorbance spectral signature ( $\log 1/R$ , where  $R$  is the reflectance spectrum) are well correlated to the concentrations of chlorophyll a, chlorophyll b, and carotenoids at canopy and leaf scales. Gong *et al.* used first derivatives of remotely sensed hyperspectral signatures as input features for a neural network system [4]. The system was used to classify six conifer tree species.

Derivatives of spectral signatures have also been used in methods that are less application dependent. For example,

Manuscript received June 10, 2000; revised March 15, 2001. This work was supported by the Department of Energy Remote Sensing Laboratory.

The authors are with the Department of Electrical and Computer Engineering, Mississippi State University, Mississippi State, MS 39762 USA.

Publisher Item Identifier S 0196-2892(01)05465-1.

Philpot developed the derivative ratio algorithm (DRA) [5]. DRA is a method that uses the ratio of the  $n$ th derivative at a particular wavelength  $\lambda_1$  to the  $n$ th derivative at a second particular wavelength  $\lambda_2$  of a measured spectral signature. He showed that under certain conditions, the above ratio is equivalent to the ratio of the  $n$ th derivatives at the same wavelengths of the true reflected radiance of the target. Thus, under certain conditions, the DRA can be used to analyze spectral signatures without the need for careful atmospheric correction algorithms. Another important utilization of spectral derivatives is for the deconvolution of constituent absorption/reflectance bands in a spectral signature. In particular, Huguenin and Jones showed that if the second derivative were less than zero, if the fourth derivative were greater than zero, and if the fifth derivative were equal to zero for a particular wavelength  $\lambda_0$ , then the spectral signature contained a constituent reflectance band centered at  $\lambda_0$  [6]. They used a smoothing prefilter to reduce high frequency noise.

Consider the first derivative of a hyperspectral signature  $f(\lambda)$

$$\frac{df}{d\lambda} = \frac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1} = \frac{f(\lambda_2) - f(\lambda_1)}{\Delta\lambda}. \quad (1)$$

The term  $\Delta\lambda$  is referred to as the derivative window. In the applications of derivative analysis described above, the width of the derivative window must be determined and can greatly affect the results. By increasing the derivative window, the effects of high frequency components in the signal are decreased. A similar effect is achieved by applying a lowpass prefilter, typically an averaging or smoothing prefilter, followed by the derivative analysis. However, the challenge is to appropriately select the width of the derivative window and/or the width of the smoothing prefilter.

In Huguenin and Jones's work in hyperspectral signature unmixing, they used a smoothing prefilter to reduce the effects of high frequency noise [6]. They determined the width of their smoothing prefilter by optimizing the tradeoff between noise and signal attenuation. However, their approach required an estimate or prediction of the full-width-half-maximums of the constituent reflectance bands contained in the hyperspectral signature. In Demetriades-Shah *et al.*'s use of derivatives to detect the red edge and to decorrelate soil background from vegetation spectra, they used the Savitsky–Golay method for determining the derivative spectra [2]. They experimentally determined the width of the smoothing operation by maximizing the SNR of the smoothed signature. However, their approach required an estimate for the signal and the noise components of the signature. They estimated the signal component as the mean of 40 scans. They estimated the noise component as the difference between the signature being analyzed and the estimate of the signal component.

Piech and Piech have conducted extensive investigations into the use of scale-space images for analyzing hyperspectral signatures [7]–[9]. Scale-space images are produced by convolving the hyperspectral signature with Gaussian derivative filters, which have been scaled, or stretched, to have increasing variances. The Gaussian function smooths the signature and the derivative operator computes the derivative of the signature.

Thus, it is a systematic means for analyzing the hyperspectral signature with varying widths of the presmoothing filter. Piech and Piech have utilized the scale-space image for automated analysis of hyperspectral signatures to determine soil moisture content [7]. They have also extracted features from scale-space images to assess atmospheric absorption features in hyperspectral signatures [8]. Piech and Piech have also used scale-space images to discriminate between terrains in topographic/elevation images [9]. They used the method for automated discrimination between shale and granite, as well as between humid and arid shale.

Tsai and Philpot explored the use of several algorithms for smoothing hyperspectral signatures and calculating their derivatives: 1) Savitzky–Golay smoothing and derivative method; 2) Kawata–Minami's linear least mean square smoothing with finite approximation of derivative; and 3) mean filter smoothing with finite approximation of derivative [10]. They examined the effects of changing the width of the smoothing operation and the tradeoff between noise removal and ability to resolve fine spectral details. They discussed three effects of increasing the smoothing width: the magnitude of the spectral details in the derivative spectra decrease, the details in the derivative spectra are displaced from the location of the corresponding feature in the signature, and the range over which the derivatives can be computed is decreased due to the proximity of the endpoints of the finite length signature. Their conclusion was that the smoothing width should be as large as possible but remain less than the width of the spectral feature of interest. They utilized the second, fourth, and fifth derivatives to locate constituent band centers in the smoothed spectra. They stated that a key issue of the method is to allow users to adapt the smoothing width to match the scale of the spectral features of interest. Furthermore, they provided suggestions for future work: 1) develop procedures for selecting appropriate width smoothing prefilters; 2) formalize and extend the spectral decomposition procedure for more general applications; and 3) improve the algorithms to speed up computation for handling large hyperspectral images. Scale-space images can accomplish suggestions 1) and 2) but not suggestion 3). In this paper, however, the authors show how wavelet transform algorithms can be used to compute scale-space images in a cost-efficient manner, thus demonstrating how wavelet algorithms can accomplish all three suggestions.

### III. SCALE-SPACE IMAGES

The computation of a scale-space image requires multiple convolutions with Gaussian derivative filters [11]. This can be represented mathematically by letting  $f(\lambda)$  denote a hyperspectral signature and  $G'_\sigma(\lambda)$  denote a first-derivative zero-mean Gaussian filter with standard deviation  $\sigma > 0$

$$G'_\sigma(\lambda) = \frac{-\lambda}{\sqrt{2\pi}\sigma^3} \exp(-\lambda^2/2\sigma^2) \quad (2)$$

and the convolution

$$F(\lambda, \sigma) = (f * G'_\sigma)(\lambda) = \int_{-\infty}^{\infty} f(\lambda - \tau) G'_\sigma(\tau) d\tau \quad (3)$$

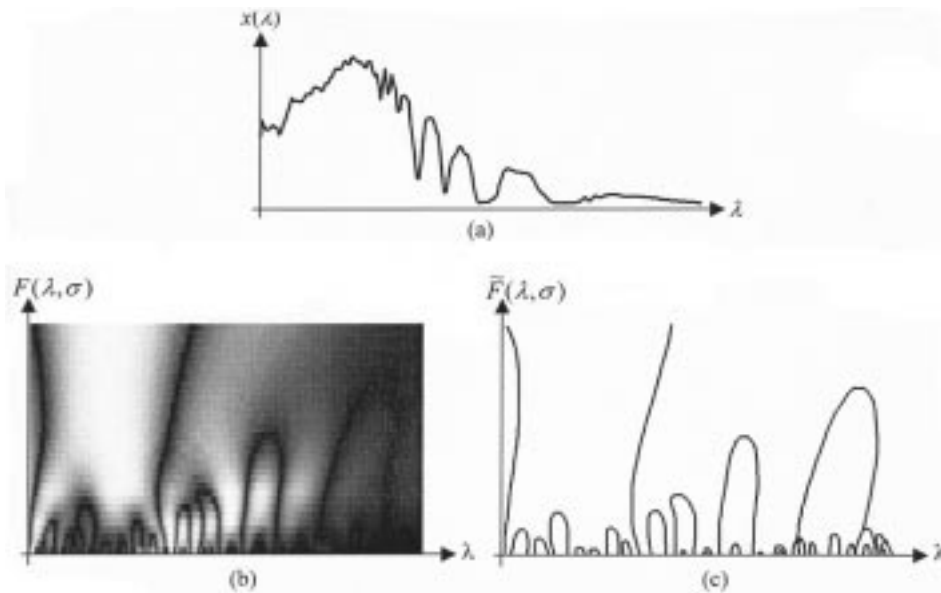


Fig. 1. (a) Example hyperspectral signature, (b) scale-space image using first-derivative Gaussian filter, and (c) spectral fingerprint.

produces the smoothing of  $f(\lambda)$ . Smoothing progresses as  $\sigma$  increases. The symbol  $\sigma$  is called the scale, and  $F(\lambda, \sigma)$  is the scale-space image. Typically, when using scale-space images, a spectral feature refers to a hill or valley contained between two adjacent inflection points. A scale-space contour refers to an image of the inflection points of the filtered curve for varying scales. As the scale increases, the number of inflection points decreases. As a result, contours usually appear as an assemblage of closed arches, each corresponding to a spectral feature of the curve [7], [12]. When this technique is applied to a hyperspectral curve, the characteristic shape of the contours leads to the nickname of “spectral fingerprint.” The exact locations of the arch closures and the rate at which they close can provide important information about the spectral features.

In order to locate the inflection points, it is typical to use first-derivative or second-derivative of Gaussian filters. This stems from the fact that when one convolves a first-derivative Gaussian filter with a hyperspectral signature, inflection points will produce extrema in the output curve (or spectral fingerprint). The extrema are tracked across the varying scales, and a plot of the extrema forms the spectral fingerprint  $\bar{F}(\lambda, \sigma)$ . If a second-derivative Gaussian filter is used, the inflection points will produce zero-crossings in the output curve. Either way, the resulting spectral fingerprints are theoretically identical. Fig. 1 provides a typical hyperspectral signature, its corresponding scale-space image, when using first derivative Gaussian filters, and the corresponding spectral fingerprint.

#### IV. WAVELET-BASED ALGORITHMS

The scale-space image can be easily computed using wavelet transforms [12]–[14]. The continuous wavelet transform can be defined as the following convolution:

$$W_s f(\lambda) = f(\lambda) * \psi_s(\lambda) = f(\lambda) * \frac{1}{s} \psi\left(\frac{\lambda}{s}\right) \quad (4)$$

where  $f(\lambda)$  is the hyperspectral signature,  $s$  is the scale, and  $\psi_s(\lambda)$  is the scaled or dilated wavelet function. It is easy to see that if the wavelet function is the first-derivative Gaussian function, this equation is identical to the traditional scale-space image implementation described above. When the mother wavelet (scale  $s = 1$ ) has a variance of  $\sigma^2$ , the scale factor  $s$  is equivalent to the standard deviation of the Gaussian function. Note that certain mathematical restrictions apply to the choice of the wavelet function, and the first-derivative Gaussian function satisfies those restrictions [14].

Fast algorithms can be used for the computation of the wavelet transform in order to avoid the computational expense of direct convolution. One such algorithm is described by Mallat and Zhong [13]. It can be used for computing a discrete wavelet transform which utilizes dyadic scales, i.e.,  $s = 2^j, \forall j \in \mathbf{I}$ . For this algorithm, the wavelet function is characterized by two discrete filters: a lowpass filter  $H$  and a highpass filter  $G$ . The notation  $H_j$  and  $G_j$  denote the upsampling of filters  $H$  and  $G$  by the factor of  $2^j$ . The following algorithm computes the discrete wavelet transform,  $W_{2^{j+1}} f(\lambda)$ , at scale  $s = 2^{j+1}$  of the input signature  $f(\lambda)$ :

$$\begin{aligned} & j = 0 \\ & S_{2^j} f(\lambda) = f(\lambda) \\ & \text{while } (j < J) \\ & \quad W_{2^{j+1}} f(\lambda) = \frac{1}{\alpha_j} \cdot S_{2^j} f(\lambda) * G_j \\ & \quad S_{2^{j+1}} f(\lambda) = S_{2^j} f(\lambda) * H_j \\ & \quad j = j + 1 \\ & \text{end of while.} \end{aligned}$$

The computation of the wavelet transform is an iterative process, where the input signature is decomposed into approximation and detail signals. At each iteration, the approximation is then further decomposed into coarser approximation and detail signals. For example, at each scale  $2^j$ , the current

TABLE I  
FINITE IMPULSE RESPONSE OF THE FILTERS  $H$  AND  $G$  FOR USE IN THE  
MALLAT/ZHONG WAVELET ALGORITHM

n	H(n)	G(n)
-1	0.125	
0	0.375	-2.0
1	0.375	2.0
2	0.125	

TABLE II  
NORMALIZATION COEFFICIENTS TO COMPENSATE FOR DISCRETIZATION EFFECT  
IN THE MALLAT/ZHONG WAVELET ALGORITHM

$j$	$\alpha(j)$
1	1.5
2	1.12
3	1.03
4	1.01
5	1.00

approximation signal  $S_{2^j}f(\lambda)$  is decomposed into the next higher scale approximation signal  $S_{2^{j+1}}f(\lambda)$  and a detail signal  $W_{2^{j+1}}f(\lambda)$ . Due to discretization, the wavelet extrema corresponding to an inflection point in the original spectral curve do not have the same amplitude at all scales as they should in the continuous model. The coefficient  $\alpha_j$  compensates for this discrete effect and corresponds to the choice of filters  $H$  and  $G$ . The filters  $H$  and  $G$  are lowpass and highpass filters. They can be appropriately selected such that the detail signals correspond to the output of equation (3). For scale  $2^j$ ,  $W_{2^j}f(\lambda) \approx F(\lambda, \sigma = s)$ . That is, each iteration of the algorithm produces a dyadic level of the scale-space image. Table I provides the finite impulse response of filters  $H$  and  $G$  such that the wavelet decomposition is equivalent to a scale-space image, and Table II provides the normalization coefficients  $\alpha_j$  that are used to compensate for the discretization effect [13].

Just as the scale-space image can be used to produce a spectral fingerprint, so can a wavelet decomposition. In fact, in wavelet terminology the spectral fingerprint could be referred to as a special case of the wavelet modulus-maximum (mod-max) representation. Let the wavelet decomposition of a hyperspectral signature be denoted as the following sequence of signals:

$$\{S_{2^j}f(\lambda), (W_{2^j}f(\lambda))_{1 \leq j \leq J}\} \quad (5)$$

including a final approximation signal at scale  $2^J$  and detail signals from scale 1 to  $2^J$ . The wavelet mod-max representation is then obtained by detecting the local extrema of the wavelet decomposition at each scale. The values of the wavelet decomposition are then set to zero if the locations do not coincide with local extrema. Thus, the mod-max representation is defined as

$$\{\bar{S}_{2^J}f, (\bar{W}_{2^j}f)_{1 \leq j \leq J}\} \quad (6)$$

where

$$\bar{S}_{2^j}f(\lambda) = \begin{cases} S_{2^j}f(\lambda), & \text{where } |S_{2^j}f(\lambda)| \\ & \text{is a local maxima} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

and

$$\bar{W}_{2^j}f(\lambda) = \begin{cases} W_{2^j}f(\lambda), & \text{where } |W_{2^j}f(\lambda)| \\ & \text{is a local maxima} \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

Note that the wavelet mod-max representation is a more generalized signal decomposition than the spectral fingerprint. However, by choosing the wavelet decomposition's lowpass and highpass filters ( $H$  and  $G$ ) appropriately, the two representations are virtually the same. The only difference, theoretically, is that the wavelet mod-max representation takes into account only those scales that are a power of 2, where the spectral fingerprint is typically computed over a continuum of scales. This point is discussed in more detail later in the paper.

## V. COMPUTATIONAL EXPENSE ANALYSIS

The "traditional" means of computing the scale-space image, via convolution with Gaussian derivative filters as in (3) is computationally expensive when compared to the Mallat-Zhong algorithm for computing the wavelet mod-max representation. The computational costs of these two methods were analyzed and compared in two manners: 1) analytical assessment and 2) experimental benchmarking.

### A. Analytical Assessment

The traditional method of computing a scale-space image requires multiple convolutions with dilated, or scaled, Gaussian filters. That is, at each scale  $s$ , a direct convolution must be performed. Let  $N$  be the number of bands in the hyperspectral signature and  $L$  be the length of the original first-derivative Gaussian filter. At the first scale, the operations required are  $L$  multiplies per band of the spectral curve and  $L-1$  additions per band of the spectral curve. Therefore, for the first scale, a constant of  $C = 2L - 1$  operations must be performed per band. At each scale, the filters are dilated by the scaling factor  $s$  or  $\sigma$ . Thus, the length of the Gaussian filter becomes  $Ls$ . At scale  $s$ , the number of operations per band is  $2Ls - 1$ , which becomes  $2Ls$  for very large scales. Thus, the total number of operations per scale is  $2LsN$ . For a total of  $P$  scales, the number of operations must be summed, and the number of operations becomes

$$2L \sum_{s=1}^P Ns. \quad (9)$$

The term  $2L$  is a constant that is dependent on the type of Gaussian filter selected and not dependent on the length of the hyperspectral signature being analyzed. Thus, the order of complexity for direct convolution can be simplified to the following:

$$O_d = O\left(\sum_{s=1}^P Ns\right) \quad (10)$$

where  $N$  is the number of bands in the original spectral curve, and  $P$  is the total number of scales utilized.

For a standard continuous wavelet transform (CWT) without the use of fast algorithms, the order of complexity is the same

as for the traditional method of direct convolution. That is, the order of complexity for the CWT is the following:

$$O_{\text{CWT}} = O_d. \quad (11)$$

This stems from the discussion in Section IV about the similarities between the direct convolution and the continuous wavelet transform.

Multiple convolutions are also required for the Mallat/Zhong algorithm for computing the wavelet mod-max representation. At each scale, a direct convolution must be performed to compute the detail signal  $W_{2^{j+1}}f(\lambda)$  and a direct convolution must be performed to compute the approximation signal  $S_{2^{j+1}}f(\lambda)$ . Let the length of the original spectral curve be  $N$ , the length the original highpass filter  $G$  be  $L$ , and the length of the original lowpass filter  $H$  be  $M$ . At the first scale, the operations required are  $L + M$  multiplies per band of the spectral curve and  $(L - 1) + (M - 1)$  additions per band of the spectral curve. Therefore, for the first scale, a constant of  $C = 2L + 2M - 2$  operations must be performed per band. At each scale  $s$ , the filters are upsampled by a factor of  $2^s$ . This introduces  $2^s - 1$  zeros between each sample of the filters. Theoretically, this would increase the lengths of the filters  $L$  and  $M$  to  $Ls$  and  $Ms$ . However, the algorithm can be implemented in an intelligent way as to avoid the multiplies and additions with the zero-samples of the filters. For example, when implementing this algorithm on a DSP chip, one would use an indirect addressing mode indexed by the amount of upsampling at each scale. In this way, the order of complexity is independent of the scale. Thus for each scale, the order of complexity is  $(CN)$ , or  $O(N)$ . That is, the order of complexity is constant across scales. Once this is summed for all scales being utilized, the order of complexity becomes the following:

$$O_{MZ} = O\left(\sum_{s=1}^P N\right) = O(NP). \quad (12)$$

Since this method uses dyadic scales, the total number of utilized scales is  $P = \log_2 N$ . As a result, the total order of complexity is

$$O_{MZ} = O(N \log_2 N) \quad (13)$$

where  $N$  is the total number of bands in the hyperspectral signature.

### B. Experimental Benchmarking

The experimental assessment of the computational expense of the two methods was completed by computing the number of floating-point operations during the computation of the scalar decompositions, i.e., scale-space images or wavelet mod-max representations. The assessment was conducted using hyperspectral digital image collection experiment (HYDICE) signatures. The authors constructed a database of 100 hyperspectral signatures. The signatures were extracted from a set of four HYDICE data cubes. Each data cube was of the size (320 pixels)  $\times$  (320 pixels)  $\times$  (210 spectral channels). The four HYDICE data cubes included a mixture of vegetative and urban scenes.

TABLE III  
EXPERIMENTAL BENCHMARKING RESULTS

		"Traditional" method		Wavelet-based method	
		386,325 adds	390,144 mults	10,901 adds	18,766 mults
TI TMS320C30	60ns instructions	46.59 ms		1.78 ms	
333MHz Pentium II	3ns instructions	2.33 ms		0.09 ms	

For each of the four data cubes, 25 signatures were randomly selected, and their coordinates were determined by using the following two-dimensional (2-D) uniform distribution:

$$u(x, y) = \begin{cases} 1/320^2, & \text{for } 0 < x, y < 320 \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

These signatures have a length of  $N = 210$ . The following dyadic scales were used:  $s = 1, 2, 4, \dots, 2^8$ . First derivative Gaussian filters were used, and the original length of the filters were  $L = 2$  and  $M = 4$ .

For the traditional method of direct convolution with Gaussian derivative filters, the total number of addition operations per signature was approximately 386 325, and the total number of multiply operations per signature was 390 144. Similarly, for the Mallat/Zhong algorithm, there were approximately a total of 10 901 addition operations and 18 776 multiply operations per signature. For a practical view of the difference of the computational expense between the traditional method and the Mallat/Zhong wavelet method, typical microprocessor chips were chosen to evaluate the computational efficiency based on the criteria of time costs. For a Texas Instruments TMS320C30 chip, it takes 60 ns to implement a floating-point addition or a floating-point multiply. For a 333 MHz Pentium II chip, it takes 3 ns to finish a floating-point addition or a floating-point multiply. Table III shows experimental results of the computational expense and illustrates the contrast of the computational efficiency. Note that for the TMS320C30 processor, many features related to the DSP chip were not used, such as direct memory access and parallel multiply-add instructions. These types of operations could greatly further reduce the time requirements.

## VI. OUTPUT COMPARISON ANALYSIS

The outputs of the two methods, traditional direct convolution with Gaussian derivative filters and the Mallat/Zhong algorithm were compared to determine if any differences existed in the scalar decompositions. The measurements used to compare the two methods' outputs included normalized cross correlation and normalized Euclidean distance measures. The comparison was performed at each level or scale of the decompositions. The comparison analysis was not conducted on the actual spectral fingerprint or mod-max representation. In order to create a spectral fingerprint from the scalar decompositions, the extrema (or zero-crossings) of the image must be tracked across the varying scales. The technique used to go from the scalar decomposition to the spectral fingerprint is the same regardless of the method used to obtain the scalar decomposition. Furthermore, in the

spectral fingerprints and wavelet mod-max representations, only the extrema (or zero-crossings) are retained, i.e., the fingerprint is simply a plot of the extrema (or zero-crossing) locations. If the comparison were made between the spectral fingerprint and the wavelet mod-max representation, errors in locations that did not correspond to extrema (or zero-crossings) would not be accounted for. Thus, comparing the outputs of the two methods before converting them to spectral fingerprints results in a more conservative measurement of the errors.

The normalized cross correlation and normalized Euclidean distance at each scale is defined as

$$\varphi(\lambda) = \frac{1}{\sqrt{\sum_{\lambda} |T(\lambda)|^2 \sum_{\lambda} |W(\lambda)|^2}} \cdot \sum_{\tau=-\infty}^{\infty} T(\tau)W(\tau + \lambda) \quad (15)$$

and

$$Eucl = \frac{\sqrt{\sum_{\lambda} |T(\lambda) - W(\lambda)|^2}}{\sqrt{\sum_{\lambda} |T(\lambda)|^2 \sum_{\lambda} |W(\lambda)|^2}} \quad (16)$$

respectively, where  $T(\lambda)$  is the traditional Gaussian filtered signal and  $W(\lambda)$  is the result of the Mallat/Zhong wavelet algorithm. Note that if the two signals are identical except for a linear shift by  $\tau$  samples  $T(\lambda) = W(\lambda + \tau)$ , then  $\varphi(\tau) = 1$ . At each scale, the entire cross correlation is computed, and the maximum value and its shift  $\tau$  is recorded. The Euclidean distance is also measured for each scale. Furthermore, it is computed for various linear shifts in  $W$ . Thus, it is similar to the normalized cross correlation, except that the Euclidean distance is minimized to zero when the two methods produce identical results.

The database of 100 HYDICE signatures, as described in Section V-B, was used for the output comparison analysis. The normalized cross-correlation and the normalized Euclidean distance were applied to the scalar decompositions resulting from all 100 hyperspectral signatures. For each error measure, the mean  $\mu$  and variance  $\sigma^2$  of the 100 errors were computed. The resulting means and variances are listed in Table IV, where  $\mu_{Corr}$  and  $\sigma_{Corr}^2$  refer to the normalized cross correlation mean and variance, respectively. Likewise,  $\mu_{Euc}$  and  $\sigma_{Euc}^2$  refer to the normalized Euclidean distance mean and variance, respectively.

## VII. DISCUSSION

From the experimental assessment of the two methods, it was clear that the computational expense of the Mallat/Zhong wavelet algorithm was much less than that of the traditional direct convolution with Gaussian derivative filters. Utilizing the Mallat/Zhong wavelet-based method, the computational expense was reduced by a factor of about 26. This coincided with the result of the analytical assessment, where the reduction

TABLE IV  
ERROR ANALYSIS RESULTS FOR 100 RANDOMLY SELECTED  
HYDICE SIGNATURES

Scale	$\mu_{Corr}$	$\sigma_{Corr}^2 (\times 10^{-3})$	$\mu_{Euc}$	$\sigma_{Euc}^2 (\times 10^{-3})$
1	1.0000	0.0000	0.0000	0.0000
2	0.9966	0.0061	0.0263	0.5183
3	0.9986	0.0002	0.0120	0.0546
4	0.9972	0.0023	0.0155	0.0636
5	0.9984	0.0013	0.0111	0.0259
6	0.9984	0.0021	0.0081	0.0206
7	0.9945	0.0080	0.0154	0.0198
8	0.9641	0.2088	0.0382	0.0438
Average	0.9935	0.0286	0.0157	0.0997

was by a factor of about 30. In this study, the number of sampling points of each hyperspectral signature was 210, since HYDICE data was used. When using dyadic scales for the wavelet-based method, the total number of utilized scales was eight. If the signatures were longer (more spectral bands), the number of utilized scales would increase. With an increase in signature length and number of scales, the computational efficiency gained from using the wavelet-based algorithms would also increase.

The outputs of the traditional and the Mallat/Zhong wavelet-based methods were compared to ensure that the “fast” algorithms did not cause any major errors in the spectral fingerprints. From this, we see that the average scalar normalized cross correlation was 0.9935, and the average normalized Euclidean distance was 0.0157. While relatively little error was induced by using the Mallat/Zhong algorithm, the computational expense was greatly reduced.

One shortcoming of the Mallat/Zhong wavelet algorithm, however, is that it does not utilize a continuum of scales. In fact, it utilizes dyadic scales, i.e.,  $s = 2^j$ . Traditional methods do have the advantage that any rational scale may be analyzed and included in the scale-space image or spectral fingerprint. The usefulness of nondyadic scales to remote sensing applications, such as automated signature anomaly detection or signature classification, is unclear. However, other wavelet-based algorithms do exist where nondyadic scales can be utilized. For example, Unser *et al.* have developed an algorithm for computing wavelet decompositions over integer scales [15]. Their algorithm has been shown to have a order of complexity of  $O(N)$  per scale, similar to the Mallat/Zhong algorithm [15], [16]. Their algorithm requires that the wavelet function,  $\psi(\lambda)$ , be represented as a polynomial spline function. However, spline functions could easily be used to approximate the Gaussian derivative filters. This is just one alternative to the Mallat/Zhong algorithm if the need arises to use nondyadic scales.

It should be noted that by using wavelet-based algorithms, one gains the mathematical tractability that is one of the strengths of wavelet methods. Moreover, the use of wavelets for analyzing and processing hyperspectral data provides access to

many existing methods for signal denoising, estimation, and compression [14], which are all highly applicable to hyperspectral data processing. The following is one example of the use of existing knowledge with regards to wavelet transforms for analyzing hyperspectral data. When using the smoothing followed by derivative analysis of hyperspectral signatures, one important question is "How much smoothing can be applied before pertinent spectral features disappear from the signature?" The answer to this question depends on the type of hyperspectral signature and the application at hand. However, Bruce showed that with the use of the wavelet mod-max representation, one can predict exactly when (at which scale) a spectral feature will disappear given the width of the spectral feature in the original signature is known [12]. This is just one example of how the existing knowledge base regarding wavelet transforms can be easily applied to hyperspectral imaging.

### VIII. CONCLUSIONS

The results of this study indicate that a wavelet-based method is feasible and practical for derivative analysis of hyperspectral signatures, specifically for computing scale-space images and spectral fingerprints. Furthermore, the results indicate that wavelet-based methods can be computationally efficient as compared to traditional methods of smoothing followed by derivative analysis, including Gaussian derivative filtering methods. The use of the Mallat/Zhong wavelet algorithm along with the computation of the mod-max representation offers a generalized method where hyperspectral signatures may be analyzed across many scales, i.e., smoothing prefilter widths, in a computationally efficient manner. Once a scale-space image or spectral fingerprint is produced, the appropriate scale (or combination of scales) may be more easily identifiable, and its selection may be automated. Thus, the wavelet-based method accomplishes three goals that have been set forth in hyperspectral derivative analysis: 1) develop procedures for selecting appropriate width smoothing prefilters; 2) formalize and extend the spectral decomposition procedure for more general applications; and 3) improve the algorithms to speed up computation for handling large hyperspectral images.

### ACKNOWLEDGMENT

The authors would like to thank C. Morgan, W. I. Ginsberg, and R. Pollina for numerous discussions and critical remarks on the manuscript.

### REFERENCES

- [1] J. W. Campbell and W. E. Esaias, "Basis for spectral curvature algorithms in remote sensing of chlorophyll," *Appl. Opt.*, vol. 22, no. 7, pp. 1084–1093, Apr. 1983.
- [2] T. H. Demetriades-Shah, M. D. Steven, and A. C. Clark, "High resolution derivative spectra in remote sensing," *Remote Sens. Environ.*, vol. 33, pp. 55–64, July 1990.
- [3] G. A. Blackburn, "Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches," *Remote Sens. Environ.*, vol. 66, pp. 273–285, Dec. 1998.

- [4] P. Gong, R. Pu, and B. Yu, "Conifer species recognition: An exploratory analysis of *in situ* hyperspectral data," *Remote Sens. Environ.*, vol. 62, pp. 189–200, Nov. 1997.
- [5] W. D. Philpot, "The derivative ratio algorithm: Avoiding atmospheric effects in remote sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 350–357, May 1991.
- [6] R. L. Huguenin and J. L. Jones, "Intelligent information extraction from reflectance spectra: Absorption band positions," *J. Geophys. Res.*, vol. 19, no. B9, pp. 9585–9598, Aug. 1986.
- [7] M. A. Piech and K. R. Piech, "Symbolic representation of hyperspectral data," *Appl. Opt.*, vol. 26, no. 18, pp. 4018–4026, Sept. 1987.
- [8] —, "Hyperspectral interactions: Invariance and scaling," *Appl. Opt.*, vol. 28, no. 3, pp. 481–489, Feb. 1989.
- [9] —, "Fingerprints and fractal terrain," *Math. Geol.*, vol. 22, pp. 457–485, May 1990.
- [10] F. Tsai and W. Philpot, "Derivative analysis of hyperspectral data," *Remote Sens. Environ.*, vol. 66, pp. 41–51, Oct. 1998.
- [11] A. Witkin, "Scale space filtering," in *Proc. Int. Joint Conf. Artificial Intelligence*, 1983, pp. 1019–1022.
- [12] L. M. Bruce, "Isolation criteria for the wavelet transform mod-max method," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 1084–1087, Aug. 1998.
- [13] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 14, pp. 710–732, July 1992.
- [14] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1998.
- [15] M. Unser, A. Aldroubi, and S. J. Schiff, "Fast implementation of the continuous wavelet transform with integer scales," *IEEE Trans. Signal Processing*, vol. 42, pp. 3519–3523, Dec. 1994.
- [16] L. M. Bruce and J. Li, "Fast wavelet-based algorithms for multiresolutional decomposition and feature extraction of hyperspectral signatures," *Proc. SPIE*, vol. 3717, pp. 72–81, Apr. 1999.



**Lori Mann Bruce** (S'91–M'96) was born in Fayetteville, TN, on June 26, 1968. She received the B.S.E. degree from the University of Alabama, Huntsville, in 1991, the M.S. degree from the Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree from the University of Alabama, Huntsville, in 1996, all in electrical engineering.

She has served as Member of Technical Staff, U.S. Army Strategic Defense Command, Huntsville, and from 1996 to 2000, she was an Assistant Professor in the Electrical and Computer Engineering Department, Howard R. Hughes College of Engineering, University of Nevada, Las Vegas. Currently, she is an Assistant Professor in the Department of Electrical and Computer Engineering, Mississippi State University, Mississippi State. Her research interests include wavelet transforms and automated pattern recognition in hyperspectral remote sensing and mammography.

Dr. Bruce is a member of Eta Kappa Nu, Phi Kappa Phi, and Tau Beta Pi.



**Jiang Li** (S'99) was born in Shannxi Province, China, in 1973. He received the B.S.E. degree from Beijing University of Aeronautics and Astronautics, Beijing, China in 1994, and the M.S. degree from the University of Nevada, Las Vegas (UNLV), in 1999, both in electrical engineering. He is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, Mississippi State University, Mississippi State.

From 1994 to 1997, he was an electrical engineer with the China Aero-Space Corporation, China. From 1997 to 2000, he was a Research Assistant with the Digital Image Processing Group, Electrical and Computer Engineering Department, UNLV. His current research interests include the wavelet transform, the adaptive wavelet algorithms, and their applications to hyperspectral remote sensing and mammography.