Spectral Curve Shape Matching Using Derivatives in Hyperspectral Images

Delian Liu and Liang Han

***Abstract*— Owning to the absorption of radiation by ground materials on specific wavelengths, spectral curve shapes stay relatively stable in hyperspectral images under various conditions and can be used for material identification. Given a reference spectrum, ground materials can be identified by matching the spectral curve shape of the reference spectrum to that of observed spectra. In the past decade, the spectral curve shape matching approach of the Tetracorder system (SCSMT) has been widely used for material identification. However, the SCSMT approach is performed directly on the continuum-removed spectra. It only catches the shape features on a large scale, missing the fine structures on spectral curves. To consider this issue, a new spectral shape matching approach is proposed, which employs the first-order and second-order spectral derivatives to capture the fine structures on spectra. A new metric is designed to balance the contributions of spectral curves, their first-order and second-order derivatives. To evaluate the performance of the proposed approach, it is applied to synthesized spectra as well as to identify oil spills from two practical hyperspectral images. The experimental results show that the proposed approach can improve the detection performance of the SCSMT approach.**

***Index Terms*— Absorption band, derivatives, hyperspectral, spectral curve shape matching.**

I. Introduction

SPECTRAL feature analysis is widely used to detect, identify, and understand materials in a variety of applications, such as mineralogy, agriculture, and earth science. Many materials have their typical compositions, which absorb radiation on specific wavelengths due to the chemical bonds of the compositions. Thus, the reflection spectrum of an object exhibits absorption features on specific wavelengths and the spectral absorption features can be used for material identification.

At present, most hyperspectral sensors have spectral channels with full width at half maximum smaller than 10 *nm*, which exceeds the resolution of diagnostic spectra of many materials. Therefore, hyperspectral remote sensing can be used for ground materials identification and classification [1]–[3]. In general, the spectral curve shape in an absorption band contains information pertaining to absorption depth and area and persists relatively stable under a variety of conditions. Thus, the shape of the spectral curve is often used for material identification. Various spectral feature extraction approaches are present to identify specific materials. Van der Meer [4] presents a spectral curve shape matching method, which uses cross correlation to measure the similarity between a reference spectrum and an observed spectrum. In [5], the depth of absorption band is adopted to relate hydrocarbons of ground covers [5]. Employing the shape of the absorption band, Brown [6] develops a Gaussian fitting model to represent spectral features in order to identify ground materials. Ma *et al.* [7] present a novel similarity assessment index for spectral curve matching, which performs well in hyperspectral image classification. Clark *et al.* [8] and Kokaly *et al.* [9], [10] from the United States Geological Survey (USGS) have studied the characteristics of spectra of many materials in the past two decades and constructed a large spectrum database for material identification. To compare the similarity of two spectra, they utilize spectral curve shape matching approach of the Tetracorder system (SCSMT). Recently, the approach has been applied to oil spill detection in [11] and [12]. The SCSMT approach employs continuum removal to normalize spectral curves and eliminate the interference of background. After normalizing the spectral curves, a fit factor is calculated by scaling a reference spectrum to compare it to an observed spectrum. In fact, the SCSMT approach is similar to the correlation coefficient of two spectra. This means that the large-scale shape features are the primary factor for determining the similarity of two spectra. The fine structures of the two spectra have slight effect on the fit factor. For remote sensing, the spectra of a hyperspectral image may be affected by the ground material compositions, atmosphere effects, or imaging system effects. They have very complex fine structures in a given absorption band. The fine structures may relate to the complex compositions of ground objects. Lang *et al.* [13] examine the absorption features in an absorption band of plant leaves and found a new fine structure at 530 nm. Therefore, the SCSMT approach that only considers the large-scale shape features is inappropriate for spectral diagnostics in hyperspectral images.

Derivative analysis is widely used to extract fine structure features in hyperspectral images [14], [15]. In [16], the first-order and fifth-order derivatives are used to find the absorption wavelength of a spectrum. In [17], the second-order derivative is used to estimate carbonate sediments. Bao *et al.* [14] deeply examine the traits of spectral derivatives and did outstanding works for hyperspectral classification by incorporating spectral derivatives. Since spectral derivatives have attractive performance in extracting fine structures on a spectrum, the spectral derivatives are introduced here to extract

Manuscript received July 22, 2016; revised November 10, 2016; accepted January 6, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61301290 and in part by the Fundamental Research Funds for the Central Universities under Grant NSIY151410 and Grant JB160502. *(Corresponding author: Delian Liu.)*

The authors are with the School of Physics and Optoelectronic Engineering, Xidian University, Xi'an 710071, China (e-mail: dlliu@xidian.edu.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2017.2651060

fine structures on spectra. In this letter, a new spectral curve shape matching approach is proposed, which employs the first-order and second-order derivatives of a spectrum to extract its fine structures.

This letter is organized as follows. In Section II, this letter introduces the continuum removal method for spectrum normalization and in Section III, the spectral curve shape matching approach of the SCSMT approach is introduced. The new spectral curve shape matching approach is then given in Section IV. In Section V, the proposed approach is evaluated and the conclusion is discussed in Section VI.

II. Continuum Removal

In hyperspectral images, each spectrum is a mixture of several pure material spectra. To identify a specific spectrum, interfering background spectra of other compositions must be removed. To eliminate the background interference, the continuum removal method is proposed. The continuum is defined as the envelope of a given absorption band of a spectrum.

A spectrum of a hyperspectral image is denoted as $x(i)$. i is the index of spectral channels. For a given absorption band the continuum of the band is the line between the two endpoints of the band

$$x_l(i) = k(i - I_l) + x(I_l) \tag{1}$$

where k is the slope of the continuum

$$k = \frac{x(I_r) - x(I_l)}{I_r - I_l} \tag{2}$$

where I_l and I_r are the indices of the endpoints. The continuum-removed spectrum is thus

$$C_r[x(i)] = \frac{x(i)}{x_l(i)}. \tag{3}$$

The continuum-removed spectrum is shown in Fig. 1.

As shown in Fig. 1, the spectrum in Fig. 1(a) appears like an absorption feature superimposed on a background spectrum. After continuum removal, the irregular background spectrum is eliminated. The spectral feature is normalized and more symmetrical.

III. Spectral Curve Shape Matching

To measure the similarity of two normalized spectra, a spectral shape matching algorithm is applied. Clark *et al.* [8] propose a spectral curve shape matching algorithm by scaling a reference spectrum to compare it to an observed spectrum. An observed spectrum is denoted by $y(i)$, the continuum-removed spectrum of $y(i)$ in a given absorption band is denoted by $y_c(i)$ calculated using (3). The continuum-removed spectrum of the observed spectrum is

$$y_c(i) = C_r[y(i)]. \tag{4}$$

The reference spectrum is denoted by $r(i)$, the continuum-removed spectrum of $r(i)$ in the given absorption band is denoted by $r_c(i)$ calculated using (3). The continuum-removed spectrum of the reference spectrum is

$$r_c(i) = C_r[r(i)]. \tag{5}$$

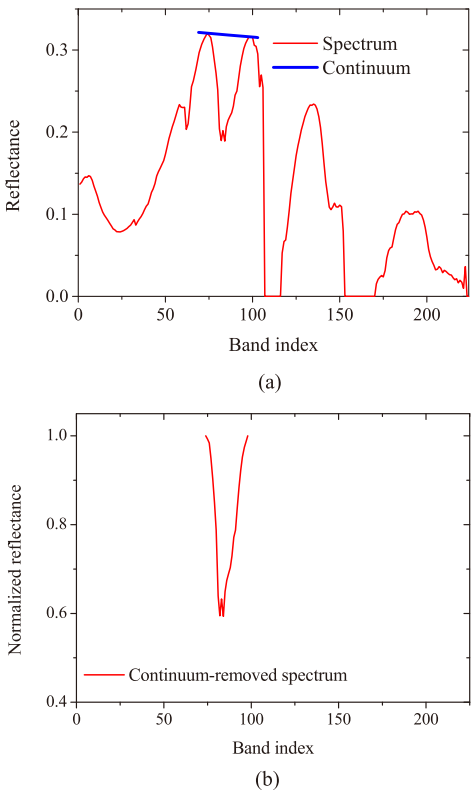

Fig. 1. Continuum removal. (a) Continuum. (b) Continuum-removed spectrum.

The fit between the reference spectrum and the observed spectrum is defined as

$$f = A[y_c(i), r_c(i)] = \sqrt{B[y_c(i), r_c(i)]B_s[y_c(i), r_c(i)]} \tag{6}$$

where f is less than or equal to 1. The larger the f value, the greater probability that the observed spectrum fits the reference spectrum

$$B[u(i), v(i)] = \frac{\sum u(i)v(i) - [\sum u(i) \sum v(i)]/N_b}{\sum v^2(i) - [\sum v(i)]^2/N_b} \tag{7}$$

$$B_s[u(i), v(i)] = \frac{\sum u(i)v(i) - [\sum u(i) \sum v(i)]/N_b}{\sum u^2(i) - [\sum u(i)]^2/N_b} \tag{8}$$

where N_b is the number of channels in the given absorption band.

IV. Shape Matching Using Derivatives

The aforementioned spectral curve shape matching approach is based on the least squares method and, therefore, catches large-scale spectral curve shape information while missing the fine structures on spectral curves. As discussed in Section I, spectral derivatives show excellent performance in extracting fine structures of spectra. To better capture the fine structures of spectra, we present a new spectral curve shape matching approach by employing the first-order and second-order spectral derivatives. For the continuum-removed spectra, the first-order derivatives are

$$r_c'(i) = r_c(i + 1) - r_c(i) \tag{9}$$

$$y_c'(i) = y_c(i + 1) - y_c(i). \tag{10}$$

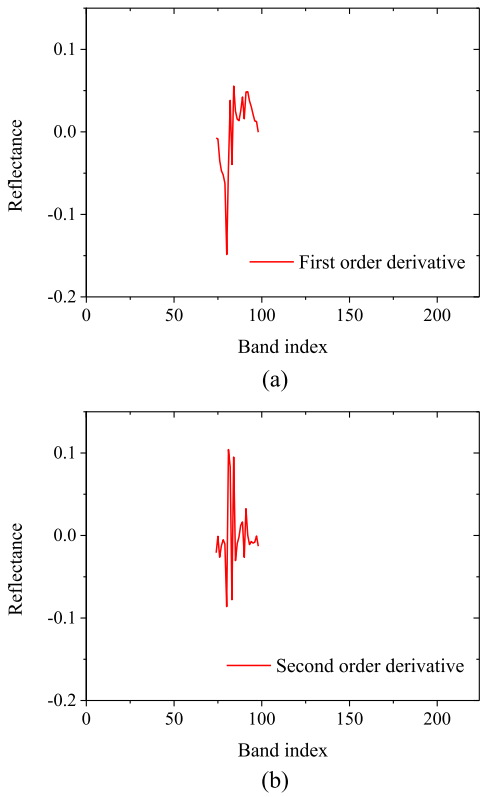

Fig. 2. Derivative spectra. (a) First-order derivative. (b) Second-order derivative.

Their second derivatives are

$$r_c''(i) = r_c(i+1) + r_c(i-1) - 2r_c(i) \tag{11}$$

$$y_c''(i) = y_c(i+1) + y_c(i-1) - 2y_c(i). \tag{12}$$

The first-order and second-order derivatives of the continuum-removed spectrum are shown in Fig. 2.

As shown in Fig. 2, the first-order and second-order derivatives fluctuate more severely than the continuum-removed spectrum in Fig. 1. The fluctuations indicate the small-scale fine structures of the spectrum, which can be extracted by using the first-order and second-order derivatives.

By substituting y_c' and $r_c'(i)$ into (6), the fit of the first-order derivatives is

$$f' = A[y_c'(i), r_c'(i)]. \tag{13}$$

Similarly, the fit of the second-order derivatives is

$$f'' = A[y_c''(i), r_c''(i)]. \tag{14}$$

The new metric of the proposed approach is

$$\eta = f[\alpha f' + (1-\alpha) f''] \tag{15}$$

where α is ratio of the first-order derivative and second-order derivative

$$\alpha = \frac{p_1}{p_1 + p_2} \tag{16}$$

$$p_1 = \sum r_c'^2(i) \tag{17}$$

$$p_2 = \sum r_c''^2(i). \tag{18}$$

Here, the metric can be extended to other spectral matching approaches

$$\eta = M[y(i), r(i)]\{\alpha M[y'(i), r'(i)] + (1-\alpha) M[y''(i), r''(i)]\}. \tag{19}$$

$M(\cdot)$ can also be the spectral angle measure (SAM), spectral correlation measure (SCM), spectral information divergence (SID), or other approaches which can measure the similarity of two spectral vectors [18].

The new proposed shape matching method considers the spectral curve and its first and second derivatives, including both large-scale and fine structure features, leading to a better performance than the SCSMT approach.

V. Experimental Results and Discussion

In this letter, the spectral derivatives are introduced to extract the fine structures of spectra for spectral curve matching. However, derivative analysis may be affected by spectral noise. To evaluate the sensitivity of the proposed approach in the presence of noise, we synthesize two new spectra by adding different levels of Gaussian noise. The two spectra are generated by using the spectra of lane grass and jarosite downloaded from the USGS, having jarosite spectrum of 80% and 50% and land grass spectrum of 20% and 50%, respectively. The spectra are shown in Fig. 3(a). The band in gray is the selected band. We use the jarosite spectrum as a reference spectrum and use the synthesized spectra as observed spectra.

To illustrate the discrimination performance of the SCSMT approach and the proposed approach, the contrast of the fit values is calculated

$$C_t = \frac{f_t - f_b}{f_b} \tag{20}$$

where f_t denotes the fit values of the synthesized spectra and jarosite spectrum. f_b denotes the fit value of the jarosite spectrum and lane grass spectrum. Fig. 3(b) and (c) is the contrast of the fit values with different Gaussian noise levels. As shown in Fig. 3(b) and (c), our proposed approach offers a higher contrast than the SCSMT approach as the SNRs are higher than 100. When SNRs decrease below 100, the proposed approach does not demonstrate discernable improvement over the SCSMT approach. However, given that most hyperspectral sensors have SNRs exceeding 200 at present, the proposed approach has better discrimination performance than the compared approach for most hyperspectral data.

In order to quantitatively assess the proposed approach, three widely used hyperspectral data sets are selected. The first hypespectral data set is the Indian Pines data set gathered by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) sensor over the Indian Pines test site in Northwestern Indiana and consists of 145 × 145 pixels and 224 spectral channels in the wavelength range 0.4–2.5 μm. The scene is a subset of a larger one. We select four classes of the scene for test. The selected classes are the grass-pasture, grass-pasture-mowed, soybean-mintill, and woods and stone-steel-towers. Only the 9th to 35th channels are selected, because atmosphere has high transmission in the band, reducing the

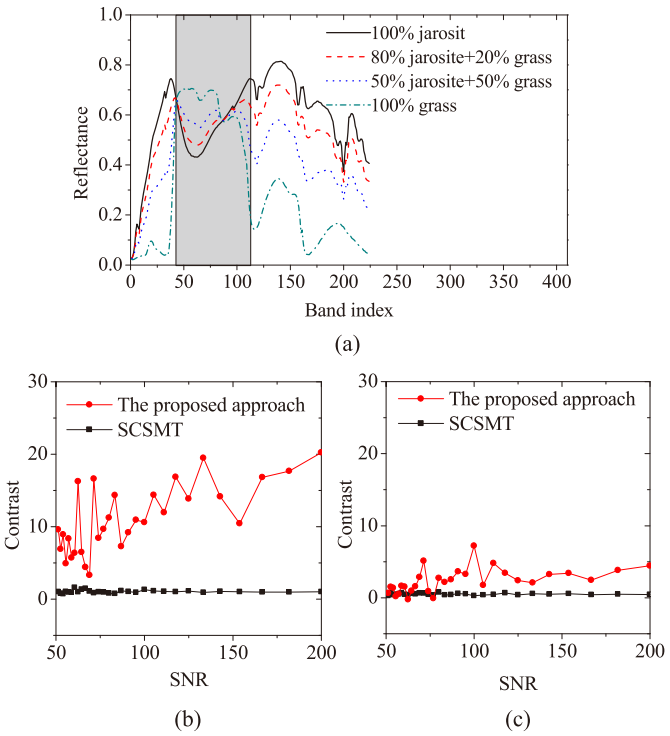

Fig. 3. Sensitivity of the proposed approach to noise. (a) Spectra of jarosite, lane grass, and synthesized. (b) Contrast of the jarosite spectrum and synthesized spectrum (80% jarosite) with different Gaussian noise levels. (c) Contrast of the jarosite spectrum and synthesized spectrum (20% jarosite) with different Gaussian noise levels.

effect of atmosphere. The SAM, SCM, SID, and the SCSMT are introduced as reference approaches for comparison. First, the reference spectrum of each class is generated by averaging spectra in a selected 3×3 window. Then, the SAM, SCM, SID, and SCSMT approaches are applied to calculate the distances of each spectrum from the reference spectra. Each spectrum is classified by finding its nearest reference spectrum. Next, the derivative features of each spectrum are extracted by the proposed approach. The SAM, SCM, SID, and SCSMT approaches are also applied to measure the new distance of two samples, and named SAMD, SCMD, SIDD, and SCSMTD. Finally, each spectrum can be classified to its nearest reference spectrum using the new distance. To reduce the effect of noise in spectral dimension, a Gaussian filter is applied before using the derivative-based approach. The overall accuracy (OA) and Kappa coefficient are then calculated to measure the discrimination performance of the evaluated approaches. The result is shown in Table I.

As shown in the leftmost two columns of Table I, the OA and the Kappa coefficients of the SAM, SCM, SID, and SCSMT approaches are smaller than that of the approaches that consider the derivative features of spectra. The results indicate that considering the first-order and second-order derivatives of spectra can improve spectral discrimination.

The second hyperspectal data set is the Salinas scene also collected by the AVIRIS sensor. It was collected over the Salinas Valley in California with a high spatial resolution of

TABLE I

OVERALL ACCURACY (OA) (%) AND KAPPA COEFFICIENT (κ) OF THE EVALUATED APPROACHES

	Indian Pines		Salinas		Pavia University	
	OA	κ	**OA**	κ	**OA**	κ
SAM	82.61	0.6097	94.66	0.9295	74.70	0.6721
SAMD	83.10	0.6150	98.67	0.9823	86.46	0.8182
SCM	67.31	0.4327	89.63	0.8656	69.33	0.5775
SCMD	79.47	0.5664	92.60	0.9033	72.03	0.6271
SID	84.24	0.6395	94.43	0.9265	73.98	0.6626
SIDD	85.13	0.6562	96.06	0.9480	80.83	0.7478
SCSMT	73.46	0.4931	96.91	0.9591	74.82	0.6644
SCSMTD	83.23	0.6168	98.77	0.9836	80.46	0.7372

3.7 m. The selected classes are the stubble, celery, soil vinyard develop, lettuce romaine 7wk, and vinyard untrained. The selected bands are the same as those used to analyze the Indian Pines set. The result is shown in the two middle columns of Table I. After introducing the spectrum derivative, the OA and the Kappa coefficients become larger, again showing that introducing derivative spectra can enhance classification performance.

The third hyperspectral data set is the Pavia University data captured by the Reflective Optics System Imaging Spectrometer (ROSIS-03) sensor. The selected classes are trees, painted metal sheets, bare soil, self-blocking bricks, and shadows. This data set is a reflectance data. We select the 50–80th channels, because it covers the red edge band of vegetation. The result is shown in the rightmost two columns of Table I. The OA and the Kappa coefficients of the evaluated approaches are also smaller than that of the approaches that consider derivative spectra. The results of the three test hyperspectral data sets reveal that our proposed approach can enhance the discrimination performance of the compared spectral curve matching approaches.

To further evaluate the proposed approach, we apply the proposed approach for oil spill detection using AVIRIS hyperspectral images. In this letter, two hyperspectral image subsections are selected. They were captured on May 18, 2010 after the Gulf oil spill from an ER-2 aircraft flying at an altitude of 9000 m and downloaded from the American National Aeronautics and Space Administration. The two hyperspectral images are shown in false color generated using the 5th, 24th, 38th bands of the images. The AVIRIS radiance data were converted into surface reflectance data. The selected oil spill reference spectrum is shown in Fig. 1. The absorption band selected is the 1.2-μm band. The result is shown in Fig. 4.

As shown in Fig. 4, the false color images of the test images are in the leftmost column. Oil spills are in the two scenes. They show up as brown in the false color images. The first scene has oil spills on the top of the scene. The second scene has oil spills in the bottom of the scene. Fig. 4(d) is the detection result of SID. As this metric contains a logarithm operation, the values are compressed to a small range so the difference is not evident. For the SAM, SCM, and

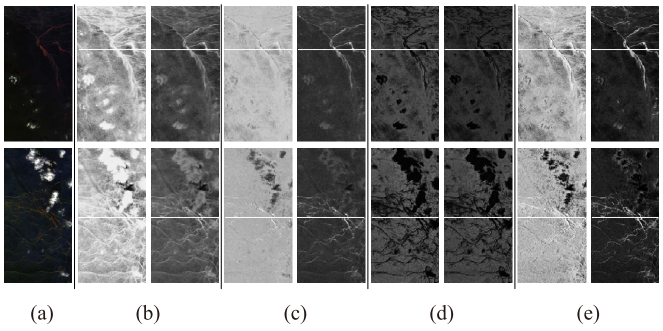

Fig. 4. Oil spill detection results. (a) False color images of scenes 1 and 2. (b) SAM approach. (c) SCM approach. (d) SID approach. (e) SCSMT approach. (Left) Results of the original approach. (Right) Results of considering spectral derivatives.

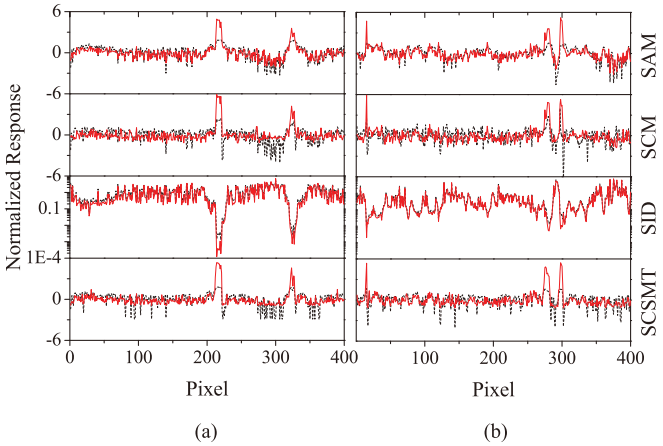

Fig. 5. Normalized fit values in a horizontal line of the evaluated approach. Solid line: fit values of the proposed approach. Dashed line: fit values of the compared approaches. (a) Test scene 1. (b) Test scene 2.

SCSMT approaches, the contrast of oil spills and background are improved after considering derivative spectral features. To show the details of the oil spills detection results, a line of the fit values is selected. The selected line is emphasized by a white line. The detection responses are normalized by $(x-u)/\sigma$ for the SAM, SCM, and SCSMT approaches. The result of SID approach is normalized by x/σ. The results are shown in Fig. 5. The proposed approach has higher values on the oil spill pixels for the SAM, SCM, and SCSMT approaches and has lower values for the SID approach. Since the fluctuation of background is normalized to a similar level, the results indicate that the proposed approach has a better performance in oil spill detection.

VI. Conclusion

In this letter, we propose a new spectral curve shape matching approach for material identification in hyperspectral images. To incorporate fine structure features on spectral curves that are missed in the SCSMT approach, the proposed approach employs the first-order and second-order spectral derivatives. A new metric is designed to balance the contributions of the spectral curve, its first-order derivative, and second-order derivative. To evaluate its performance, the proposed approach is applied to synthesized spectra, and is also applied to identify oil spills from two practical hyperspectral images collected during the Deepwater Horizon oil spill. The experimental results show that the proposed approach can improve the detection performance of the SCSMT approach.

References

[1] S. Shanmugam and P. SrinivasaPerumal, "Spectral matching approaches in hyperspectral image processing," *Int. J. Remote Sens.*, vol. 35, no. 24, pp. 8217–8251, 2014.

[2] K. Wang, B. Yong, X. Gu, P. Xiao, and X. Zhang, "Spectral similarity measure using frequency spectrum for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 130–134, Jan. 2015.

[3] F. Mirzapour and H. Ghassemian, "Improving hyperspectral image classification by combining spectral, texture, and shape features," *Int. J. Remote Sens.*, vol. 36, no. 4, pp. 1070–1096, 2015.

[4] F. van der Meer, "Spectral curve shape matching with a continuum removed CCSM algorithm," *Int. J. Remote Sens.*, vol. 21, no. 16, pp. 3179–3185, 2000.

[5] F. Kühn, K. Oppermann, and B. Hörig, "Hydrocarbon index—An algorithm for hyperspectral detection of hydrocarbons," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2467–2473, 2004.

[6] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1601–1608, Jun. 2006.

[7] D. Ma *et al.*, "Spectral similarity assessment based on a spectrum reflectance-absorption index and simplified curve patterns for hyperspectral remote sensing," *Sensors*, vol. 16, no. 2, p. 152, 2016.

[8] R. N. Clark *et al.*, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res., Planets*, vol. 108, no. E12, p. 5131, 2003.

[9] R. F. Kokaly *et al.*, "Spectroscopic remote sensing of the distribution and persistence of oil from the Deepwater Horizon spill in Barataria Bay marshes," *Remote Sens. Environ.*, vol. 129, pp. 210–230, Feb. 2013.

[10] R. F. Kokaly and A. K. Skidmore, "Plant phenolics and absorption features in vegetation reflectance spectra near 1.66 μm," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 43, pp. 55–83, Dec. 2015.

[11] S. H. Peterson, D. A. Roberts, M. Beland, R. F. Kokaly, and S. L. Ustin, "Oil detection in the coastal marshes of Louisiana using MESMA applied to band subsets of AVIRIS data," *Remote Sens. Environ.*, vol. 159, pp. 222–231, Mar. 2015.

[12] D. Liu, J. Zhang, and X. Wang, "Reference spectral signature selection using density-based cluster for automatic oil spill detection in hyperspectral images," *Opt. Exp.*, vol. 24, no. 7, pp. 7411–7425, 2016.

[13] M. Lang, P. Siffel, Z. Braunová, and H. K. Lichtenthaler, "Investigations of the blue-green fluorescence emission of plant leaves," *Plant Biol.*, vol. 105, no. 6, pp. 435–440, 1992.

[14] J. Bao, M. Chi, and J. A. Benediktsson, "Spectral derivative features for classification of hyperspectral remote sensing images: Experimental evaluation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 594–601, Apr. 2013.

[15] F. Mirzapour and H. Ghassemian, "Multiscale Gaussian derivative functions for hyperspectral image feature extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 4, pp. 525–529, Apr. 2016.

[16] F. Tsai and W. Philpot, "Derivative analysis of hyperspectral data," *Remote Sens. Environ.*, vol. 66, no. 1, pp. 41–51, 1998.

[17] E. M. Louchard *et al.*, "Derivative analysis of absorption features in hyperspectral remote sensing data of carbonate sediments," *Opt. Exp.*, vol. 10, no. 26, pp. 1573–1584, 2002.

[18] F. van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery," *Int. J. Appl. Earth Observat. Geoinf.*, vol. 8, no. 1, pp. 3–17, 2006.