Sparse approximation, coherence and use of derivatives in hyperspectral unmixing.

Jakub Bieniarz, Rupert Müller, Xiaoxiang Zhu, Peter Reinartz
German Aerospace Center (DLR),
Earth Observation Center (EOC)
Remote Sensing Technology Institute (IMF)
Wessling, 82234, Germany

***Abstract*—Recently, it has been shown that the spectral unmixing can be regarded as a sparse approximation problem. In our studies we employ predefined dictionaries containing the measured spectra of different materials in a hyperspectral image, where for each pixel the abundance vector can be estimated solving the ℓ_1 optimization problem. This results in an automation of the unmixing procedure and enables using complex overcomplete dictionaries. However, the reflectance spectra of most materials are highly coherent and this could result in confusion in the mixture estimation. In this work we present a novel approach for spectral dictionary coherence reduction and discuss the feasibility of the methodologies in terms of mutual coherence and approximation error values using overcomplete dictionaries. We compare standard sparse unmixing procedures with our novel derivative method. The presented method was tested on both simulated hyperspectral image as well as on a AVIRIS data.**

***Index Terms*—Hyperspectral image, unmixing, sparse approximation, basis pursuit, spectral derivative, coherence.**

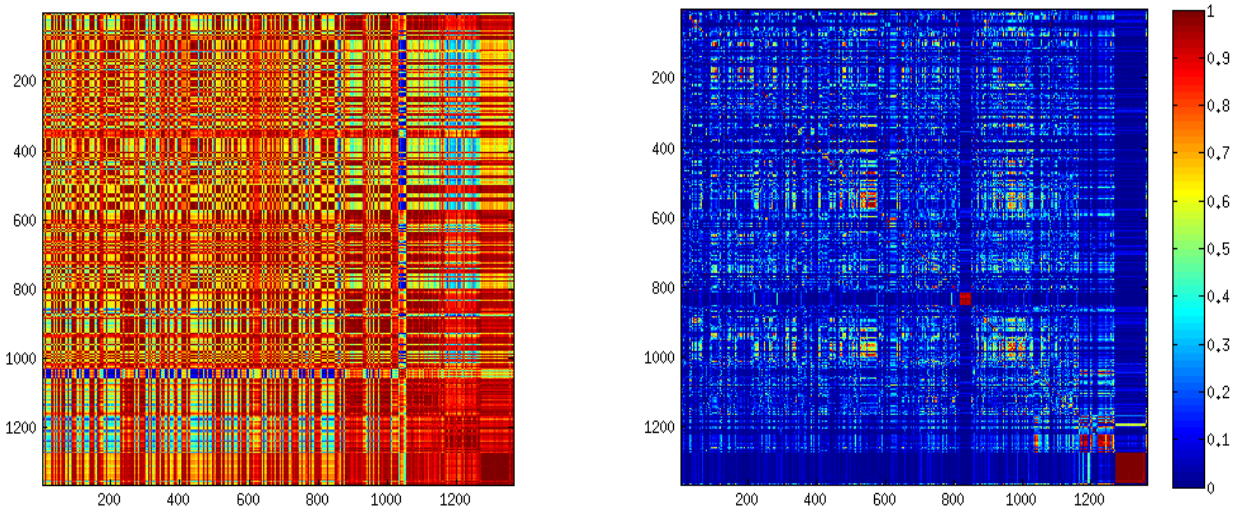
I. Introduction

Hyperspectral imaging and algorithms for the hyperspectral data processing are recently one of the most studied techniques in the remote sensing. Each hyperspectral pixel consists of hundreds of contiguous and narrow spectral bands covering typically the visible to short infrared wavelength range. So, hyperspectral image is typically L - dimensional data cube where L is number of channels. This huge amount of information provided by the hyperspectral sensors presents new challenges to the standard algorithms for image processing and analysis. Wide spectral range and resolution takes advantage of classical spectroscopy enabling detailed spectral investigation for each pixel including material detection and quantitative analysis of material content for each pixel. However, the compactness of channels results in the limitation of the pixel size. Modern hyperspectral space borne cameras have normally a ground resolution of about 30 m. Due to this relatively low spatial resolution, mixing of several sources (materials) may occur in one hyperspectral image pixel which makes detailed pixel based analysis difficult. On the other hand, the high spectral resolution enables the use of source separation methods in order to retrieve basic sources of which the pixel is composed. After separation of the basis sources, known as endmembers in imaging spectroscopy, a mixture of endmembers for each pixel can be approximated by the fractional abundances. The method for recovery of endmembers and estimation of their amount for each pixel is called spectral unmixing.

Recently, spectral unmixing is intensively studied by many researchers resulting in many methods where the endmembers can be retrieved from the image vertex component analysis - VCA, pixel purity index - PPI, N-FINDR, iterative error analysis - IEA, independent component analysis ICA-EA [1], [2]) or manual selection from existing databases of endmembers. Alternatively endmembers can be preselected from the spectral library. However, this method requires expert knowledge. The quantitative approximation of the abundances in the HSI data is often regarded as least squares problem. While many existing algorithms are excellent tools for the community and experienced users these methods can however work only with a limited number of endmembers restricted by the dimensionality of the HSI data. If the number of endmembers exceeds the dimensionality of the spectral feature space, fewer equations than unknowns lead to an infinite number of solutions for this problem. Constrained versions of least squares, proposed in [3] and studied in [4], can be used to solve these linear problems, however with some limitations which will be further discussed in this paper.

To deal with these shortcomings, spectral unmixing can be also addressed as a sparse approximation problem. If we assume that the mixing of endmembers is a linear process, then each measured pixel can be described as a product of a known dictionary and a sparse coefficients vector using ℓ_1 norm as minimizer. In this case the dictionary is overcomplete, which means that it consists of more endmembers than the dimensionality of the signal. Dictionaries can be created from existing databases containing hundreds of spectra of different materials, measured using field or laboratory spectrometers. Iordache at al. [5] tested several sparse approximation methods for spectral unmixing with different preselected spectral libraries.

In this work we discuss the use of sparse approximation methods for hyperspectral unmixing and the influence of the coherence between the elements of the overcomplete spectral dictionaries. We introduce a novel usage of spectral derivatives to increase the endmember detection rate. Our approach was evaluated using both, simulated hyperspectral image and real AVIRIS image scene and compared with one of the state of the art method.



(a) Original dictionary.

(b) Transformed dictionary.

Fig. 1. Coherence matrices for the dictionaries: a) containing original spectra and b) containing derivatives of original spectra.

II. Methods

In the following section we describe the methods used in our studies. E.g. sparse approximation methods for hyperspectral unmixing. This method enables semi-automatic endmember selection as well as estimation of the fractional abundances for each pixel. We also discuss the usage of spectral derivative as a dictionary coherence reduction method and its application to sparse spectral unmixing.

A. Spectral unmixing

Usually most HSI pixels are mixed i.e. they consist of signals from more than one source (endmember). The mixing process can be modeled as a transformation T over the matrix $M^{[L\times N]}$ containing N endmembers to describe the pixel y

$$y = T\{M\} + \varepsilon, \tag{1}$$

where ε is an additive noise and L is the dimensionality of the pixel y. If we assume that the transformation T is linear i.e. only single scattering occur for each detected photon, then the mixing model can be formulated as

$$y = Ma + \varepsilon, \tag{2}$$

where $a \geqslant 0$ is a coefficient vector which defines the fractional abundance of each material [1]. Linear spectral unmixing aimsrecovery of the endmembers and their abundances contained in a pixel.

B. Spectral unmixing using overcomplete dictionaries

In order to automate the unmixing process we could use a large overcomplete mixing matrix containing all available spectra (e. g. measured in field or in laboratory). If each column of the mixing matrix is normalized to ℓ_2 unit length it is called dictionary D with the columns as atoms φ. Then a hyperspectral pixel $y^{[L\times 1]}$ can be interpreted as the measurement of a sparse signal. Hence, the redundant dictionary $D^{[L\times N]}$ is in this case a measurement matrix,

$$y = Dx, \tag{3}$$

and $x^{[N\times 1]}$ is the coefficient vector. Through D with $N > L$, the system of equations is underdetermined and has normally an infinite number of solutions. Therefore, instead of having $y = Dx$ we can find such a solution x that minimizes $\|Dx - y\|_2$ where $\|\cdot\|_2$ is the euclidean ℓ_2 norm with $\|x\|_p = \sqrt[p]{\sum |x_i|^p}$.

Since we expect that each pixel is a mixture of only few endmembers we can look for a sparse approximation of x. If we denote by $\|x\|_0 = k$ the pseudo-norm ℓ_0 indicating the number of non-zero elements of the vector x, then the x vector is k-sparse.

$$min\ \|x\|_0\ s.t.\ \|Dx - y\|_2 \leqslant \xi, \tag{4}$$

where $\xi > 0$ is the tolerance value. The solution of the above problem is called the sparse solution to equation. 3.

1) Basis Pursuit: Since equation 4 is a non-convex, combinatorial optimization problem, it is difficult to solve. Instead we can use the ℓ_1 norm as a minimizer which also promotes the sparsest solution. This minimization problem can be written as

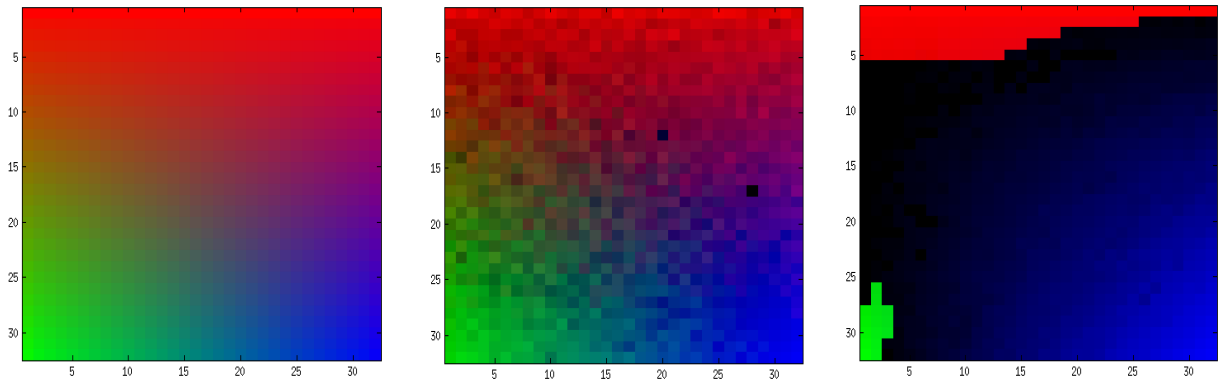
$$min\ \|x\|_1\ s.t.\ \|Dx - y\|_2 \leqslant \xi, \tag{5}$$

known as basis pursuit (BP) [6] or LASSO (least absolute shrinkage and selection operator) [7].

2) Non negative least squares: Non negative least squares (nnls) is the state of the art method for spectral unmixing and abundance estimation [1], [2], [3], [4].

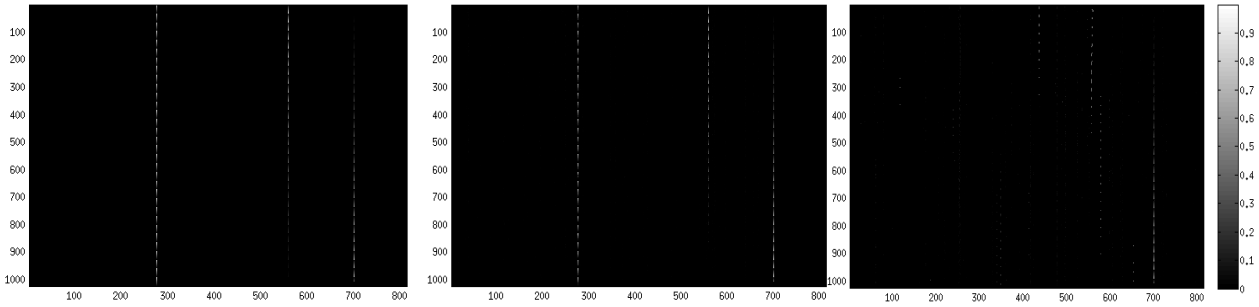
$$min\ \ \|Dx - y\|_2\ s.t.\ x \geq 0, \tag{6}$$

Since the least squares approximation is not capable of unmixing signals from the overcomplete dictionaries, it is proved that



(a) Original mixture. (b) Presented method. (c) Non negative least squares.

Fig. 2. RGB representation of the mixtures. Red channel represents abundance of the endmember no. 275, green no. 559 and blue no. 700.



(a) Original mixture. (b) Presented method. (c) Non negative least squares.

Fig. 3. RGB representation of the mixtures. Red channel represents abundance of the endmember no. 275, green no. 559 and blue no. 700.

its constrained non-negative version, under certain condition can recover sparse solution for the underdetermined system of equations [8], [9].

3) Coherence of the dictionary: BP does not recover the solution for problem 4 directly, but under certain condition it is able to recover the sparsest solution for the underdetermined system. One of the most important condition in sparse approximation methods is the mutual coherence. Mutual coherence is a measure of maximum coherence between pairs of atoms φ of the dictionary.

$$\mu = \underset{i \neq j}{max} \; |\varphi_i^{\mathsf{T}} \cdot \varphi_j|. \tag{7}$$

However, mutual coherence is a worst-case measure, in principle the smaller the coherence between the atoms, the sparser abundance vector which can be approximated using the ℓ_1 minimization [10].

C. Derivative of spectrum

The differentiation of the function estimates the slope over the changing independent variable. In our case the independent variable is the band number. We calculate the derivative of a spectrum in the following way

$$\frac{\partial y}{\partial b_i} = \frac{y(b_i) - y(b_j)}{\Delta b}, \tag{8}$$

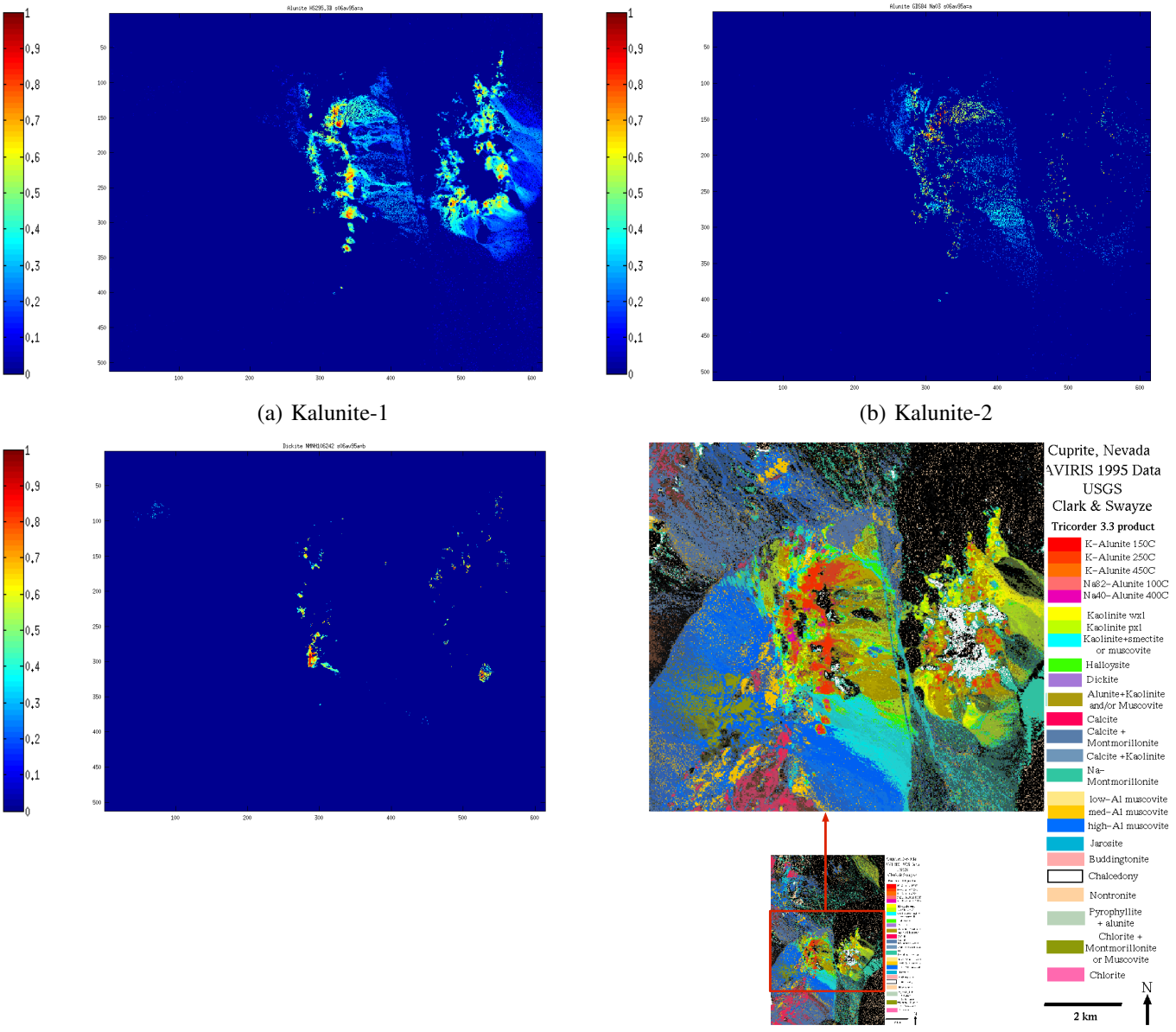
where b is a hyperspectral band, $\Delta b = b_i - b_j$ and $b_i > b_j$. The differentiation of spectra does not result in more information that is contained in the in original bands but it is possible to decrease background reflectance and can therefore considerable improve detection of convoluted weaker absorption features [11]. We calculate the derivatives for each atom in the library. The comparison of the coherence matrices for original and transformed one using derivative method can be seen on fig. 1.

III. Experimental results

We have tested our method with simulated as well as real AVIRIS data [12]. The algorithm has been assessed for the detection rate, the ℓ_2 approximation error. We have compared results of our method with one of the state of the art method nnls.

A. Spectral dictionary

For the experiment we use the USGS spectral library [13]. The library consists of 1365 spectra, but due to missing readings of many spectra, and incorrect values we have reduced the size of the library atoms to 813. Each atom was resampled to 224 bands according to the AVIRIS airborne sensor as described in [14]. From the dictionary we have excluded atmospheric absorption and noisy bands. After selection 154 spectral bands were used for further experiments. The final dictionary size is $D^{[154 \times 813]}$.



(a) Kalunite-1 (b) Kalunite-2

(c) Dickite (d) Classification result from [15]

Fig. 4. Abundance maps created using the proposed method.

B. Simulated image

The simulated image was created using three spectra selected from the dictionary described above. The image consist of 32×32 pixels and each pixel is sampled with 153 channels. The endmembers are distributed and mixed like in the figure 2. (a). A Gaussian noise is added to the simulated image with a SNR = 400, which is a realistic value for the hyperspectral sensors. Abundances values are calculated for this image, for both original and derivatives of spectra. Results for this test are shown in fig. 2, 3. The selected spectra are Microcline (no. 275), Kaoline (no. 559) and Plastic PTE (no. 700). The coherence between the endmembers no. 275 and no. 559 is 0.99, between no. 275 and no. 700 is 0.90 and coherence for the pair no. 559, no. 700 equals 0.9. The coherence between the spectra derivatives pairs is 0.43, 0.04 and 0.02, respectively. The mean approximation error for the whole image is 0.054 using our method, while using non negative least squares for the original spectra the error is 0.005. The reconstruction error using state of the art method is significantly lower and the abundances are not recovered correctly (fig. 2 (c)). The fig. 3 (b) and (c) shows all recovered endmembers using both methods. The derivative method provides sparser results and detects endmembers correctly which is not the case for the nnls. To assess the detection accuracy we calculate the ℓ_2 error between the original abundance vectors and the estimated ones. The detection error for our method is 0.20, while for nnls

it is 0.63.

C. AVIRIS data

For this test we have used freely available AVIRIS hyperspectral image data [12]. From this image we have selected a region of 612×512 pixels with 154 out of 224 spectral bands. The abundances for each differentiated pixel have been approximated using BP technique. The elements of the dictionary $D^{[154\times 813]}$ have been differentiated, too. Mean $\|x\|_0$ for all pixels in the image is about 3.06. In this study case we concentrated on material detection and approximation of abundance values. Here we imply that all material on ground are available in the dictionary. Selected abundance maps are shown in fig. 4.

IV. Conclusions

In this work we propose to use sparse approximation methods with large overcomplete dictionaries for spectral unmixing and material detection. To deal with detection confusion caused by high coherence of the dictionary, we introduce derivatives of spectra. This transformation of the signal significantly decreases overall coherence of the dictionary in comparison to the original spectra.

Experiments on simulated data show that the reduced coherence of differentiated dictionary enables sparser approximation and thus smaller detection confusion in comparison to nnls. However, use of the derivative method increases the approximation error. Hence, the improvement of the trade off between error and detection of actual endmembers is an important issue for future work.

We also show the applicability of the proposed method using real hyperspectral image data producing reasonable sparse results.

References

[1] J.M. Bioucas-Dias and A. Plaza, "An overview on hyperspectral unmixing: Geometrical, statistical, and sparse regression based approaches," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, july 2011, pp. 1135 –1138.

[2] Nirmal Keshava, "A survey of spectral unmixing algorithms a survey of spectral unmixing," *Lincoln Laboratory Jurnal*, vol. 14, pp. 55–78, 2003.

[3] D.C. Heinz and Chein-I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, no. 3, pp. 529 –545, mar. 2001.

[4] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," *ArXiv e-prints*, Feb. 2010.

[5] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 6, pp. 2014–2039, 2011.

[6] Scott Shaobing Chen, David L. Donoho, Michael, and A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998.

[7] Robert Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

[8] A. M. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of non-negative sparse & redundant representations," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing ICASSP 2008*, 2008, pp. 5145–5148.

[9] Michael A. Lexa, Mike E. Davies, and John S. Thompson, "Compressive and noncompressive power spectral density estimation from periodic nonuniform samples," *CoRR*, vol. abs/1110.2722, 2011.

[10] Jean-Luc Starck, Fionn Murtagh, and Jalal M. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*, New York : Springer, 2010.

[11] Gerhard Talsky, *Derivative Spectrophotometry of First and Higher Orders*, VCH Verlagsgesellschaft mbH, Weinheim (Federal Republic of Germany) VCH Publishers, Inc., New York, NY (USA), 1994.

[12] JPL, "http://aviris.jpl.nasa.gov/html/aviris.freedata.html," 1996.

[13] R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. R. Kok, and S.J. Sutley, "Usgs digital spectral library splib06a," Digital Data Series 231, 2007.

[14] F.D. van der Meer, S.M. De Jong, and W. Bakker, *Imaging spectrometry: basic principles and prospective applications*, Kluwer Academic Publishers, 2001.

[15] Roger N. Clark, Gregg A. Swayze, K. Eric Livo, Raymond F. Kokaly, Steve J. Sutley, J. Brad Dalton, Robert R. McDougal, and Carol A. Gent., "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems," *Journal of Geophysical Research*, vol. 108, pp. 5131, 2003.