BUILDING A HYPERSPECTRAL LIBRARY AND ITS INCORPORATION INTO SPARSE UNMIXING FOR MINERAL IDENTIFICATION

Thanh Bui[1,4], Beate Orberger[2,3], Simon B. Blancher[1], Ali Mohammad-Djafari[4], Henry Pilliere[5], Anne Salaun[1], Xavier Bourrat[6], Nicolas Maubec[6], Thomas Lefevre[5], Celine Rodriguez[1], Antanas Vaitkus[7], Saulius Grazulis[7], Cedric Duée[6], Dominique Harang[5], Thomas Wallmach[1], Yassine El Mendili[8], Daniel Chateigner[8], Mike Buxton[9], Monique Le Guen[10]

1) Eramet Research, Eramet Group, Trappes, France; 2) GEOPS-Université Paris Sud-Paris Saclay, Orsay, France; 3) Catura Geoprojects, Paris, France; 4) L2S, CNRS, Centrale Supélec, Université Paris-Saclay, France; 5) ThermoFisher Scientific (TFS), Artenay, France; 6) BRGM, Orléans, France; 7) Vilnius University Institute of Biotechnology, Vilnius, Lithuania; 8) CRISMAT-CNRS, Normandie Université, Caen, France; 9) Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands; 10) Eramet Nickel Division, Eramet Group, Trappes, France

ABSTRACT

The objective of the SOLSA project (EU-H2020) is to develop an analytical expert system for on-line-on-mine-real-time mineralogical and geochemical analyses on sonic drill cores. As one aspect of the system, this paper presents the building of the hyperspectral library and its incorporation into sparse unmixing techniques for mineral identification. Twenty seven spectra representing 14 minerals have been collected for the library. Three sparse unmixing techniques have been investigated and evaluated using simulated data generated from our hyperspectral library, and real hyperspectral data acquired from a serpentinized harzburgite sample. Among the three techniques, the collaborative sparse unmixing by variable splitting and augmented Lagrangian (CLSUnSAL) method provided the best accurate results on the simulated data. In addition, the results of the CLSUnSAL method show high correlation with that of the QEMSCAN® analysis on the harzburgite hyperspectral data.

Index Terms— Hyperspectral library, sparse unmixing, shortwave infrared (SWIR).

1. INTRODUCTION

Combined mineralogical and chemical analyses on drill cores are highly demanded by mining and metallurgical companies to speed up exploration, mining, and to define geometallurgical parameters for beneficiation and metal extraction. Currently, analyses are done by exploiting only a single technique, such as infrared, X-ray fluorescence (XRF), etc. The coupling of different analytical instruments is still a technological challenge. As a result, only a small section that is considered to be representative for drill cores is analyzed, leading to a lack of systematic mineralogical analyses on the whole drill cores. The EU-H2020 SOLSA project (www.solsa-mining.eu), targets to construct an expert system coupling sonic drilling with an on-line-real-time analytical system combining systematic mineralogical

and chemical analyses on drill cores. The analytical system of the project is composed of SOLSA IDA and SOLSA IDB. SOLSA IDA that is designed to perform a fast scan consists of a profilometer, a high resolution RGB camera, VNIR (Visible Near Infrared)/SWIR (Shortwave Infrared) hyperspectral cameras (SPECIM Ltd., Finland), and an XRF spectrometer. Its main goal is to define quickly the regions of interest (ROI) on the SOLSA drill cores in qualitative and semi-quantitative manners by using the information of hardness, color, textures, chemical composition and mineralogy. Those ROIs will then be further analyzed with XRD, XRF and Raman spectroscopy in the SOLSA IDB. The system will be validated for nickel laterites, which represent 70% of the Ni resources worldwide. Nickel laterites are soils that are very heterogeneous in grain size and textures. They host mainly hydrous minerals, which can be effectively detected by hyperspectral imaging techniques.

Thanks to its non-invasive measurement, rapid data acquisition and simple instrumentation, the reflectance spectroscopy using the wavelength regions in the SWIR (1000 – 2500 nm) ranges is considered to be a useful mineral analysis tool, particularly for mineral exploration and geometallurgy. When a sample is illuminated by a light source, e.g., halogen light, certain wavelengths of light in the SWIR ranges are absorbed by the minerals in the samples due to sub-molecular vibrations involving bending and stretching of molecular bonds in the minerals. The bonds that engender absorption features in the SWIR wavelength include OH, H_2O, AlOH, FeOH, MgOH and CO_3. These molecules are found as major components in phyllosilicates, hydroxylated silicate, tectosilicates, sulphates, carbonates [1].

In practice, due to the characteristics of the samples and the resolution of the imaging system, a SWIR spectrum acquired from a sample surface may contain a mixture of several minerals. Therefore, spectral unmixing techniques have been developed to handle the problem associated with

mineral mixtures in an acquired spectrum [2]. Two kinds of spectral unmixing have been commonly used in the literature: linear and nonlinear [2], [3]. The linear spectral unmixing which exhibits practical advantages assumes that the spectra collected by the hyperspectral camera can be expressed in a linear combination of endmembers (pure spectral signatures) weighted by their corresponding fractional abundances (proportions).

Most spectral unmixing methods can be categorized as statistical or geometrical frameworks [2]. The statistical Bayesian framework relies on the posterior probability density (specifically, the likelihood formalizes the assumed data generation model, while the prior imposes natural constraints on the endmembers as well as model spectral variability) of the unknowns, given the observations. Whereas, the geometrical approaches exploit the fact that, under the linear mixing model, the observed hyperspectral vectors belong to a simplex set whose vertices correspond to the endmembers.

Sparse unmixing [4], [5] which has connections with both the statistical and geometrical frameworks is another direction that have been recently explored for spectral unmixing. Sparse unmixing techniques aim at finding the optimal subset of signatures in a spectral library that can best model each mixed pixel in the scene. These methods take into account the fact that a spectrum always contains a mixture of a small number of endmembers (up to 4 as documented in [6]).

Spectral unmixing techniques rely on the spectral library. There exist some hyperspectral libraries such as the USGS and the commercial library built by CSIRO [6]. However, several factors influence the spectra, such as the light source, illumination, particular instruments, spectra purity, minerals of interest, etc. Therefore, it is important to build our own library for the SOLSA system that is under construction at TFS, Artenay, France.

This paper presents one aspect of our system by exploiting the use of the SWIR hyperspectral camera (Specim Ltd., Finland) for mineral identification. Specifically, the main contributions of this work include: (1) build a new hyperspectral library and plan to open it to public; (2) integrate the hyperspectral library into sparse unmixing techniques for mineral identification; (3) evaluate unmixing results by using simulated data and by comparing them with QEMSCAN® results.

2. HYPERSPECTRAL LIBRARY AND SPARSE UNMIXING TECHNIQUES

2.1. Building the hyperspectral library

The spectral library is a collection of spectra of pure minerals that are of interest to the SOLSA project and research communities. To build the library, collecting the samples is an essential first step. Rock and mineral samples

have been provided by BRGM, ERAMET and the National Museum of Natural History (MNHN), France. Rock samples are composed of silicates and oxides, while mineral samples mainly consist of one phase. All samples are analyzed by conventional laboratory methods (XRD, Raman spectroscopy, SEM and EPMA).

The samples were scanned to acquire the hyperspectral data as illustrated in Fig. 1. Subsequently, a preprocessing step is necessary to convert the data into relative reflections (reflectance) by using white and dark references. The white

Fig. 1. Acquiring the hyperspectral data from selected samples using our system

reference is a measurement of material with almost 100% reflection, whereas the dark reference is acquired by closing the shutter to capture the equipment noise. The reflectance is then computed by dividing the sample spectra by the white reference spectra after subtracting the dark reference spectra from both spectra.
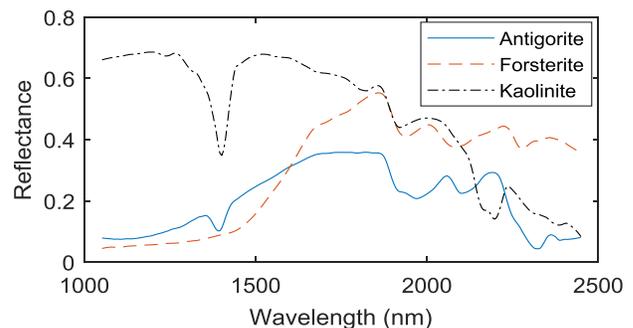

Fig. 2. SWIR spectra acquired from three minerals.

Each spectrum is examined and extracted from the data using ENVI 5.4 and the spectral interpretation field manual G-MEX [1] by taking into account the wavelength positions and the relative intensities of the absorption features in the spectrum. Currently, 27 spectra representing 14 minerals (i.e., asbolane, chromite, diaspore, olivine (forsterite), clay minerals (kaolinite, saponite, pimelite), magnesite, pyroxene (enstatite), serpentine (lizardite, nepouite, antigorite), talc, calcite) have been collected for our library. Three spectra are shown in Fig. 2 as spectra examples. The hyperspectral library is going to be made available as a spectral open database at https://solsa.crystallography.net/sod/.

2.2. Sparse unmixing

For each observed pixel $\boldsymbol{y}$ of the hyperspectral data, sparse unmixing finds a linear combination of spectral signatures in a $L \times M$ spectral library $\boldsymbol{A}$ using $\boldsymbol{y} = \boldsymbol{Ax} + \boldsymbol{n}$, where $\boldsymbol{y}$ is a $L \times 1$ vector with L spectral bands, M is the number of spectral signatures in $\boldsymbol{A}$, $\boldsymbol{x}$ is a $M \times 1$ vector containing the estimated abundances, and $\boldsymbol{n}$ is a $L \times 1$ vector collecting the errors affecting the measurements at each spectral band. Assuming that the data set contains P pixels organized in a matrix $\boldsymbol{Y} = [\boldsymbol{y_1}, \cdots, \boldsymbol{y_P}]$, we can write: $\boldsymbol{Y} = \boldsymbol{AX} + \boldsymbol{N}$, where $\boldsymbol{X} = [\boldsymbol{x_1}, \cdots, \boldsymbol{x_P}]$ is the abundance matrix and $\boldsymbol{N} = [\boldsymbol{n_1}, \cdots, \boldsymbol{n_P}]$ is the noise matrix.

Three sparse unmixing methods including Full Constrained Least Squares (FCLS), Sparse Unmixing by Variable Splitting and Augmented Lagrangian (SUnSAL), and Collaborative Sparse Unmixing by Variable Splitting and Augmented Lagrangian (CLSUnSAL) [4], [5] have been investigated and implemented. To find the abundance matrix $\boldsymbol{X}$, the optimization problems for the three methods are as follows:

$$\min_{\mathbf{X}} \; \|\mathbf{AX} - \mathbf{Y}\|_{\mathbf{F}}^2 \text{ subject to: } \mathbf{X} \geq \mathbf{0}, \mathbf{1}^T\mathbf{X} = \mathbf{1}$$

$$\min_{\mathbf{X}} \; \|\mathbf{AX} - \mathbf{Y}\|_{\mathbf{F}}^2 + \lambda\|\mathbf{X}\|_{1,1} \text{ subject to: } \mathbf{X} \geq \mathbf{0}, \mathbf{1}^T\mathbf{X} = \mathbf{1}$$

$$\min_{\mathbf{X}} \; \|\mathbf{AX} - \mathbf{Y}\|_{\mathbf{F}}^2 + \lambda\|\mathbf{X}\|_{2,1} \text{ subject to: } \mathbf{X} \geq \mathbf{0}, \mathbf{1}^T\mathbf{X} = \mathbf{1}$$

The FCLS imposes the abundance non-negativity constraint and the abundance sum-to-one constraint into the model. The differences of the two remaining methods correspond to the way to impose the sparsity in the solution of the abundance matrix $\boldsymbol{X}$. Specifically, the SUnSAL employs pixelwise independent regressions, while the CLSUnSAL enforces joint sparsity among all the pixels. The CLSUnSAL globally assumes that all the pixels in the hyperspectral image share the same active set of endmembers.

The solutions $\boldsymbol{X}$ of the above optimization problems were derived by employing the alternating direction method of multipliers [7] that basically decomposes a difficult problem into a sequence of simple ones.

3. RESULTS

3.1. Results on simulated data

Using our hyperspectral library $\boldsymbol{A}$ with $L = 251$ bands and $M = 27$ members, simulated data were generated to evaluate the performance of the three unmixing methods. We used the same data simulation mechanism as described in [4], [5] with the number of endmembers $K = 2, 3, 4$. The fractional abundances of the endmembers follow a Dirichlet distribution; the data were contaminated with Gaussian noise that was quantified by signal-to-noise (SNR) ratio; 300 pixels were generated with the SNR of 40dB. The quality of the reconstruction of mixed spectra was measured using the signal-to-reconstruction error (SRE) because this measure, as opposed to the classical root-mean-squared error, gives more information regarding the power of the signal in relation with the power of the error [5].

As presented in Tab. 1, the CLSUnSAL method provides the most accurate unmixing results for 3 different endmember numbers by providing the highest SRE values, followed by the SUnSAL. In all cases, the regularization parameter λ was tuned to achieve the best results. In terms of computational time, the FCLS provides the fastest solution. This is an important factor, considering the real-time aspects of our project. For visual evaluation, the CLSUnSAL provides a sparse solution corresponding well to the original abundance matrix as shown in Fig. 3.

K	FCLS		SUnSAL		CLSUnSAL	
	SRE	time	SRE	time	SRE	time
2	14.24	0.022	14.94	0.254	16.74	0.228
3	6.41	0.019	7.45	0.259	11.95	0.230
4	5.25	0.022	7.07	0.499	7.16	0.453

Tab. 1. SRE values (dB) and time (second) of the three methods with different endmembers K obtained from the simulated data (SNR = 40dB).
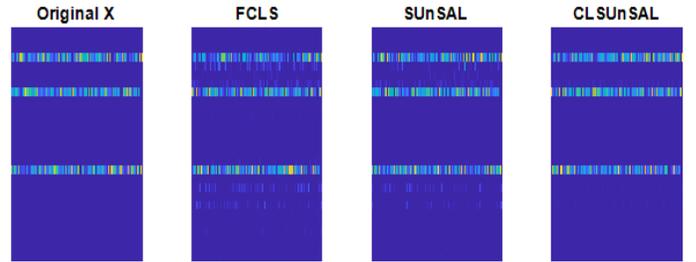

Fig. 3. Generated abundance matrix ($K = 3$) and estimated abundance matrices with three unmixing methods.

3.2. Results on a serpentinized harzburgite sample

Because the CLSUnSAL provided the most accurate unmixing results on simulated data, we applied it to the unmixing of the hyperspectral data acquired from a serpentinized harzburgite sample (SOLSA label of ER-MB00-0012). The unmixing result was compared with the mineralogical mapping provided by QEMSCAN® technique carried out at Eramet Research. QEMSCAN® [8] uses a scanning electron microscopy platform and is an automatic mineral and phase characterization technique. The technique takes into account physical, chemical and texture properties of minerals and phases in order to generate a high-resolution (up to $2\mu m$) mineral maps and porosity structures.

Fig. 4 presents the RGB image of the serpentinized harzburgite sample, the unmixing result by the CLSUnSAL and the mineralogical mapping provided by QEMSCAN® analysis. As indicated by QEMSCAN® analysis, the serpentinized sample contains four phases including pyroxene (in blue), chromite (in red), serpentine (green) and

olivine (grey). The unmixing technique successfully identifies the same mineral phases as the QEMSCAN® analysis; note that the estimated abundance values of olivine are rather low.

Fig. 4 also shows that the mineral phases identified by the unmixing technique are well correlated with those provided by the QEMSCAN® analysis. To facilitate comparison with the QEMSCAN® results, the abundance matrix of the unmixing results was applied thresholding to find a major component in each pixel.

4. CONCLUSIONS AND PERSPECTIVES

This paper presents a preliminary result on building the hyperspectral library and incorporates it into sparse unmixing techniques for mineral identification. Among investigated unmixing methods, the CLSUnSAL provided the most accurate unmixing results on simulated data and also provide results that are well correlated with the QEMSCAN® results.

The work can be extended in several ways. First, we are collecting new samples and also requesting new samples from the MNHN museum so that the number of spectra (currently 27) will be extended. Second, we plan to perform quantitative evaluation between the hyperspectral unmixing results and the QEMSCAN® results on the serpentinized harzburgite sample; similar evaluation is also going to be done on a serpentinized dunite sample and a saprolite sample from the lower Ni-laterite profile. Third, after establishing the quantitative evaluation mechanism, we will evaluate the FCLS method as the real-time processing is an important factor of our project. Forth, the Bayesian solutions that offer the automatic estimation of the regularization parameter λ will be considered to apply. Finally, a comprehensive comparison between hyperspectral unmixing results, QEMSCAN® mineral mapping and Raman spectroscopy mineral mapping is going to be done.

5. REFERENCES

[1] AusSpec, *GMEX1-Spectral Interpretation Field Manual*. AusSpec International Ltd., 2008.

[2] J. M. Bioucas-dias *et al.*, "Hyperspectral Unmixing Overview : Geometrical , Statistical , and Sparse Regression-Based Approaches," *IEEE Geosci. Remote Sens. Mag.,* vol. 5, no. 2, pp. 354–379, 2012.

[3] S. Zhang, J. Li, K. Liu, C. Deng, L. Liu, and A. Plaza, "Hyperspectral Unmixing Based on Local Collaborative Sparse Regression," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 5, pp. 631–635, 2016.

[4] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," *2nd Work. Hyperspectral Image Signal Process. Evol. Remote Sensing, WHISPERS 2010 - Work. Progr.*, 2010.

[5] M. Iordache, J. M. Bioucas-dias, and A. Plaza, "Collaborative Sparse Regression for Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 1–14, 2014.

[6] M. Berman *et al.*, "A Comparison Between Three Sparse Unmixing Algorithms Using a Large Library of Shortwave Infrared Mineral Spectra," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3588--3610, 2017.

[7] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011.

[8] B. Ayling, P. Rose, S. Petty, E. Zemach, and P. Drakos, "QEMSCAN® Quantitative evaluation of minerals by scanning electron microscopy: capability and application to fracture characterization in geothermal systems.," *Thirty-Eighth Work. Geotherm. Reserv. Eng. Stanford Univ.*, vol. Proceeding, p. 11, 2012.
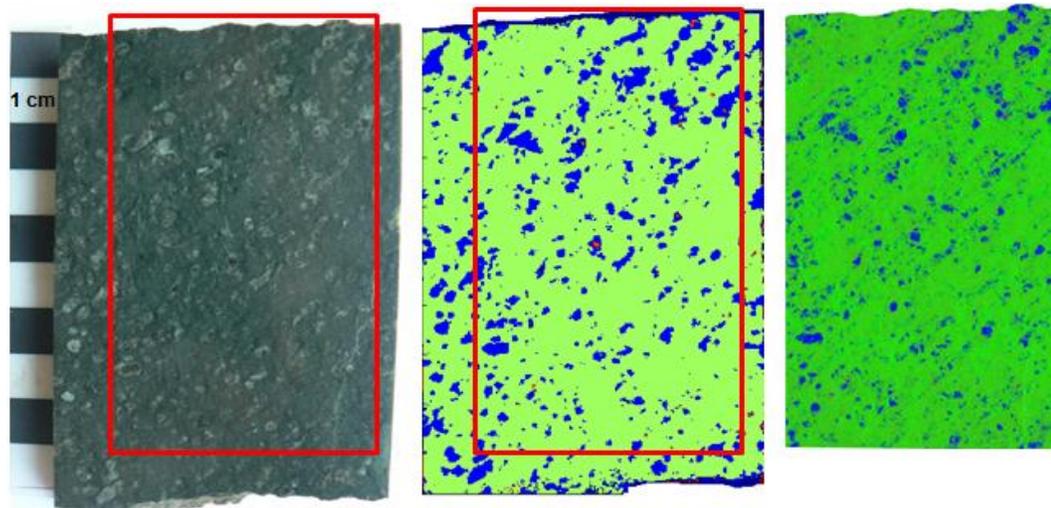
Fig. 4. Results on a serpentinized harzburgite sample. RGB image of the sample (left), unmixing result (middle), QEMSCAN® mineralogical mapping (right). Note that the QEMSCAN® analysis was done only in the region limited by the rectangle with red border. For the unmixing and QEMSCAN results, the pyroxene, serpentine, chromite and olivine phases are presented in blue, green, red and grey colors, respectively.