A MACHINE LEARNING TECHNIQUE FOR DRILL CORE HYPERSPECTRAL DATA ANALYSIS

Cecilia Contreras, Mahdi Khodadadzadeh, Laura Tusa, Pedram Ghamisi and Richard Gloaguen

Helmholtz-Zentrum Dresden-Rossendorf (HZDR),
Helmholtz Institute Freiberg for Resource Technology, Germany

ABSTRACT

Hyperspectral data are increasingly being used to map minerals in drill core samples allowing a non-invasive and non-destructive characterization of the mineral assemblages, and therefore, the mineralogical composition of a system, its variability, and structural features. The analysis of drill core hyperspectral data is traditionally carried out by a visual interpretation of the spectra and a comparison with reference libraries using spectral similarity measures. Although this approach produces good results it is time-consuming and subjective. In this work, we introduce, for the first time, an innovative automatic mineral mapping technique for drill core hyperspectral data by using a machine learning approach. More specifically, we propose to exploit detailed information coming from the Scanning Electron Microscopy (SEM)-based Mineral Liberation Analysis (MLA) to train a supervised classifier. For the extraction of input features, a traditional technique is explored, i.e., Principal Component Analysis (PCA). For the classification step, we suggest to use Random Forest (RF) because of its significant performance when there are few training samples available. Experimental results conducted on a VNIR-SWIR drill core hyperspectral dataset, show accurate classification results.

Index Terms— Drill core hyperspectral data, Mineral Liberation Analysis, Random Forest, mineral mapping.

1. INTRODUCTION

Hyperspectral imaging (HSI) is an emerging technique that is able to provide valuable information through a wide range of wavelengths. Nowadays, hyperspectral data are recorded in several tens of spectral bands and used in various fields. For the exploration of mineral deposits, especially in the mining industry, drill core hyperspectral data are usually acquired to provide a general overview of the mineralogical composition of a system, its variability, and structural features. The procedure consists in drilling a hole from the surface to a specific depth in the range of meters, extracting cores of rock samples and scanning them with hyperspectral drill core scanners. The hyperspectral sensors used in the drill core scanners work in the visible-near infrared (VNIR), short-wave infrared (SWIR) and thermal infrared (TIR) [1, 2].

Drill core hyperspectral data offer a non-invasive, non-destructive, and economical mineralogical analysis. A traditional approach to characterize and map minerals in drill core hyperspectral data is to visually select some representative samples, the so-called endmembers, and perform unmixing or classification by using a spectral similarity measure, e.g. Spectral Angle Mapper (SAM) [3, 4].

To have fast and objective analyses in geological applications, recently, the use of machine learning techniques is proposed [5, 6]. As an example, in [6] the authors compared the performance of four regression algorithms: artificial neural networks (ANNs), regressions trees (RTs), random forest (RF), and support vector machines (SVMs), for predictive modelling of epithermal gold potential. However, for drill core hyperspectral data very few machine learning techniques have been proposed in the literature. As one of the few attempts, in [7], Schneider et al., tested the performance of an SVM, Gaussian Process (GP), and Spectral Angle Mapper (SAM) algorithms for the analysis of drill cores using the spectral features of hyperspectral data.

Developing supervised machine learning techniques for the analysis of drill core hyperspectral data is a challenging task. In such techniques, it is fundamental to define meaningful classes and select some representative samples for training. In this work, for the first time, we explore the use of MLA for training a supervised classifier and mapping minerals in drill core hyperspectral data. MLA is a high-resolution imaging method that identifies minerals in polished thin sections. It includes automated mineralogy to collect quantitative mineralogical data and analyze the microstructure attributes of the thin sections by using automated Scanning Electron Microscopy (SEM). Backscattered electron (BSE) images are used to define mineral grains and are combined with X-ray analysis to identify the minerals [8, 9, 10]. In the literature, the use of MLA has been limited to validate hyperspectral image analysis results [11, 12], whereas we propose to exploit the MLA information to define training samples for a supervised classification system. Another important challenge for developing a supervised classification technique is feature extraction. In this work, we use the traditional PCA approach for feature extraction. For the classification, we suggest to use Random Forest (RF) since it has been shown to produce reliable results in other geological applications [6].

In this paper, we propose a fully automatic technique to map minerals where no previous knowledge is required. It considers not only spectral information of the SWIR active minerals but also the possibility to map SWIR non-active minerals based on mineral liberation analysis.

The rest of the paper is structured as follows: section 2 contains the methodology proposed in this work. Section 3 presents data description, experimental results and discussions. Finally, the conclusions are drawn in Section 4.

2. METHODOLOGY

Fig. 1 shows the proposed machine learning-based framework to map minerals in drill core hyperspectral data. It includes the acquisition of the VNIR-SWIR hyperspectral dataset as well as MLA. In the first step, the MLA image is carefully coregistered with the hyperspectral data and resampled with respect to the hyperspectral spatial resolution by considering the most dominant minerals in each HSI pixel. For the purpose of mapping both VNIR-SWIR active and non-active minerals, two MLA maps are generated from the MLA results, i.e., one with all MLA detected minerals and the other considering only the SWIR active minerals. As the final MLA reference map, the combination of the two previous maps is used for classification. This is to guarantee that the reference map is labeled based on the minerals detectable in the VNIR-SWIR hyperspectral data.

In the second step, PCA-based features are extracted from the drill core hyperspectral data. The last step of the proposed technique is to predict the class label of each pixels of the drill core HSI. To this aim, the extracted features together with the MLA reference map are given as input to a supervised classifier. The well-known machine learning classifier RF is used since it has been proven that it performs well when a few number of training samples are available. This is the case in drill core data when usually detailed a priori information, such as the MLA reference map, is only available for a small area of the drill core.

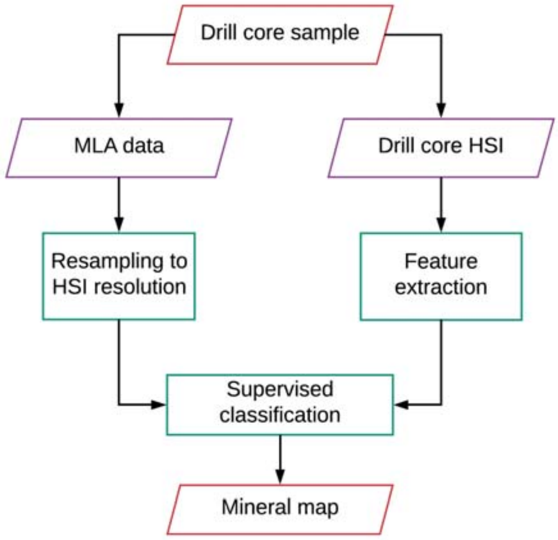


Fig. 1: Flowchart of the proposed machine learning technique to map minerals in drill core hyperspectral data. MLA stands for Mineral Liberation Analysis. HSI refers to hyperspectral image.

3. EXPERIMENTAL RESULTS

To show the performance of the suggested method, we used a SisuRock drill core scanner equipped with an AisaFenix VNIR-SWIR hyperspectral sensor to acquire the hyperspectral data from the unpolished half of the drill core sample (see Fig.2). An RGB image of the drill core sample can be seen in Fig.3a. The spectral range of the camera is from 380 to 2500 nm, with a spectral resolution of 3.5 nm in the VNIR and 12 nm in the SWIR. The total number of bands is 450 and the pixel size is 1.7 mm/pixel. Preprocessing of the data was carried out with the toolbox presented in [13], where radiometric corrections, corrections of the sensor shift, and geometric corrections for the effect of lenses were applied.

MLA was performed in a polished thin section, of about 30 μm, of the drill core sample (red rectangle in Fig.3a). It was carried out using an FEI Quanta 650 F field emission SEM with two Bruker Quantax X-Flash 5030 energy-dispersive X-ray (EDX) detectors equipped with the MLA software. Polishing removed about 300 μm of rock surface, therefore, the thin section may not have an identical mineralogy after polishing. However, this difference is negligible in our work, thus, the mineralogy is assumed to remain the same. For the development of the MLA high resolution mineral map the grain-based X-ray mapping (GXMAP) method was applied (see Fig.3b). It consists of the discrimination of minerals based on their grayscale level in the BSE images and the identification of the different phases by performing X-ray measurements on a closely-space grid. The used MLA operating conditions are: pixel size 3 μm, the resolution is 1000×1000 pixels with a step size of 6×6 pixels, the acquisition time is 5 ms, the BSE trigger 26-255, the minimum grain size is 3 pixels.

Fig.3c shows the resampled MLA map (Fig.3b), where the minerals in each pixel represent the most dominant minerals in pixels of 1.7mm (HSI pixel size). Fig.3d is the MLA map showing only the most dominant SWIR active minerals (without considering SWIR non-active minerals). The final MLA reference map can be seen in Fig.3e, which is the combination of the maps in Fig.3c and Fig.3d. Therefore, 10 main classes were recognized: *Muscovite+Muscovite* (51pixels), *Gypsum+Gypsum* (43 pixels), *Plagioclase+Muscovite* (115 pixels), *Plagioclase+Biotite* (3 pixels), *Plagioclase+Chlorite* (66 pixels), *Plagioclase+Hornblende* (10 pixels), *Plagioclase+Gypsum* (13 pixels), *Quartz+Muscovite* (62 pixels), *Quartz+Chlorite* (2 pixels), and *Quartz+Gypsum* (25 pixels).

For the feature extraction, we set the PCA algorithm so as to retain 99.9% of spectral information which resulted in 21 principal component bands. Moreover, in the RF algorithm, we consider an ensemble of 200 decision tree classifiers.

Fig.4c shows the mineral map obtained by the proposed system, using the final MLA reference map (Fig.4b) as training samples and PCA-based features. As can be observed, this map correctly illustrates the mineral distributions detectable in the drill core sample by using the proposed technique. The veins and alterations halos present in the drill core sample are well characterised. Fig.5 shows the center spectrum of three main classes in the training set (dashed line) as well as the test set producing the resultant mineral map (continues line). In Fig.5a and Fig.5b, for the *Quartz+Muscovite* and *Plagioclase+Muscovite* classes, respectively, it can be seen that the main absorption features (at around 1400 nm, 2192 nm, 2342 nm and 2435 nm) corresponds to muscovite and are present not only in the training set spectra, as it is expected, but also in the test set spectra. Moreover, spectra for the *Plagioclase+Muscovite* class show a higher mixture with chlorite (at around 2241 nm and 2334 nm) than the *Quartz+Muscovite* class. This is due to the fact that the *Quartz+Muscovite* class corresponds to the alteration vein whereas the *Plagioclase+Muscovite* class is related to the alteration halo and in direct contact with the matrix characterized by the *Plagioclase+Chlorite* class. Center spectra for the *Plagioclase+Chlorite* class can be seen in Fig.5c. An important remark is that each pixel of the MLA reference map is assigned to the most dominant mineral, while it might contain different minerals. Moreover, we need to consider the inherent characteristic of hyperspectral data in which pixels might contain mixtures of different mineral responses.

From the previous visual analysis and comparison, it can be seen that the information from the MLA reference map has been accurately spread and generalized through the entire sample. In general, the use of MLA mineral maps as reference data to train an RF classifier produces good results in terms of the quality of the classification map. However, results can be affected by (1) the error produced from the visual corregistration between the MLA and HIS data and (2) the availability of a limited number of samples in the reference map for certain classes.

4. CONCLUSIONS AND REMARKS

In this paper, we explore, for the first time, the use of MLA within a supervised machine learning technique to map minerals available in drill core hyperspectral data. The combination of the information coming from MLA and hyperspectral data is complementary and can be exploited to accurately map minerals. The use of the MLA mineral map as reference data allows mapping mineral species that may be difficult to identify using only hyperspectral data due to the chemical variability that can occur within the mineral. The present approach considers a PCA-based feature extraction technique as input for a random forest classification approach.

As part of our future developments, we will test the proposed machine learning approach on different datasets and explore other feature extraction and classification techniques.

Fig. 2: SisuRock drill core scanner equipped with an AisaFenix VNIR-SWIR hyperspectral sensor.

5. REFERENCES

[1] R. N Clark, "Spectroscopy of rocks and minerals, and principles of spectroscopy," in *Remote sensing for the earth sciences: Manual of remote sensing*, vol. 3, pp. 3–58. John Wiley & Sons, Inc, 1999.

[2] F. van der Meer, "Analysis of spectral absorption features in hyperspectral imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 5, no. 1, pp. 55–68, 2004.

[3] F. A. Kruse, R. L. Bedell, J. V. Taranik, W. A. Peppin, O. Weatherbee, and W. M. Calvin, "Mapping alteration minerals at prospect, outcrop and drill core scales using imaging spectrometry," *International Journal of Remote Sensing*, vol. 33, no. 6, pp. 1780–1798, 2012.

[4] M. Mathieu, R. Roy, P. Launeau, M. Cathelineau, and D. Quirt, "Alteration Mapping on drill cores using HySpex SWIR-320m hyperspectral camera: application to the exploration of an unconformity-related uranium deposit (Saskatchewan, Canada)," *Journal of Geochemical Exploration*, vol. 172, pp. 71–88, 2017.

[5] W. Sun, J. Ma, G. Yang, B. Du, and L. Zhang, "A Poisson nonnegative matrix factorization method with parameter subspace clustering constraint for endmember extraction in hyperspectral imagery," *ISPRS Journal of*

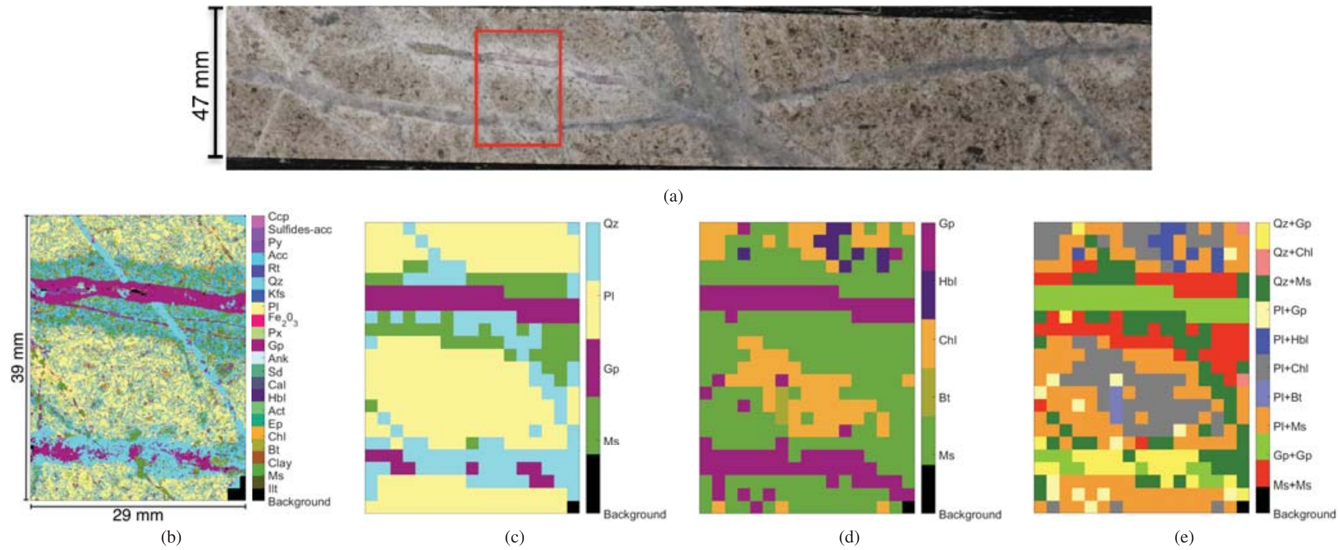


Fig. 3: (a) RGB image of the drill core sample. The red rectangle represents the area where the MLA was performed. (b-e) MLA mineral maps: (b) Original MLA mineral map. Pixel size is 3 μm. (c-e) MLA mineral maps resampled to the hyperspectral data resolution (pixel size is 1.7 mm): (c) Original MLA map (considering both SWIR active and non-active minerals), (d) MLA map with SWIR active minerals, (e) Final MLA reference map (combination of (c) and (d)). Mineral abbreviations after [14].

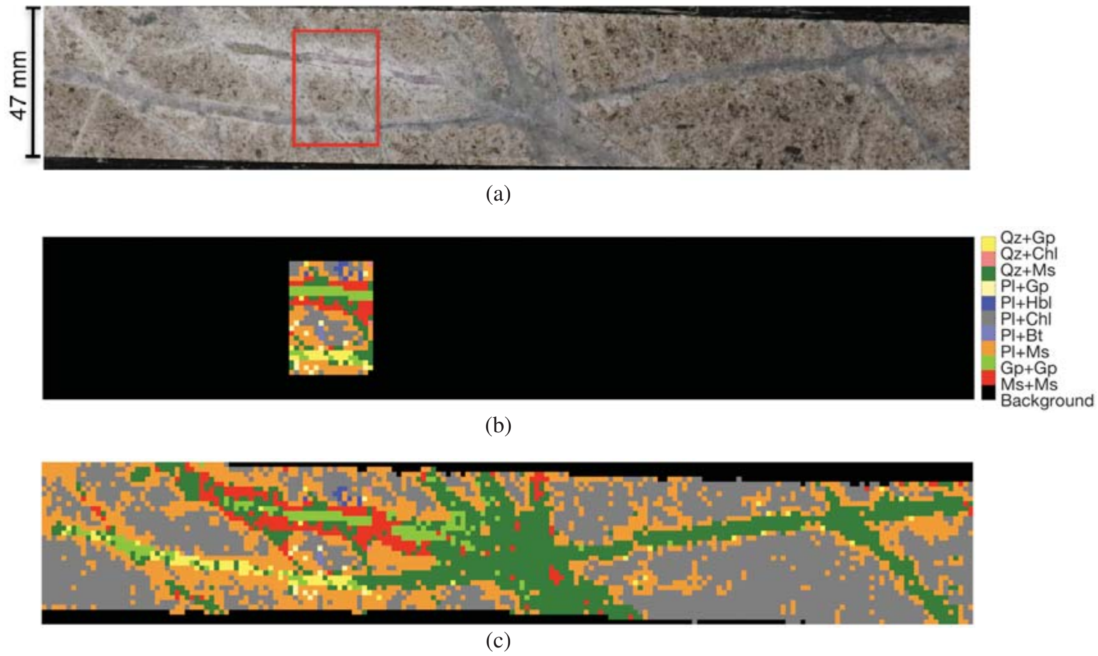
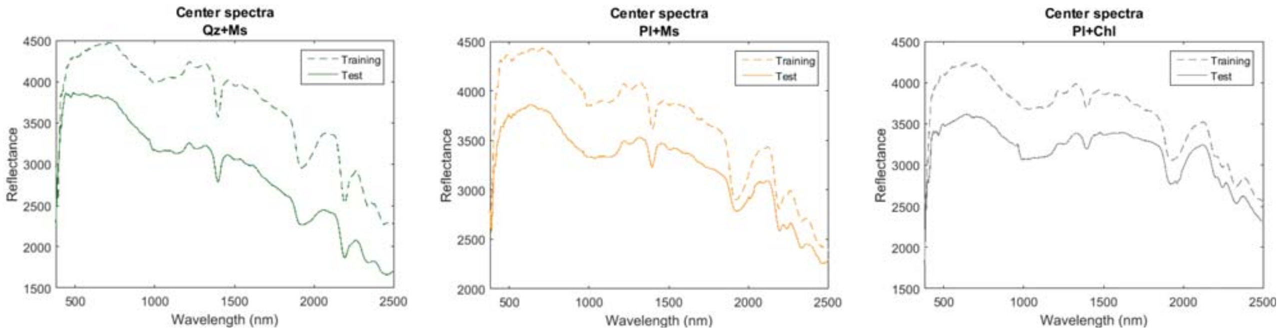


Fig. 4: (a) RGB image of the drill core sample. The red rectangle represents the area where the MLA was performed. (b) MLA reference map. (c) Mineral map obtained with the proposed approach using PCA as the feature extraction technique. Mineral abbreviations after [14].



(a) (b) (c)

Fig. 5: Training and test center spectra of classes: (a) *Quartz+Muscovite*, (b) *Plagioclase+Muscovite*, and (c) *Plagioclase+Chlorite*.

Photogrammetry and Remote Sensing, vol. 128, pp. 27–39, 2017.

[6] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines," *Ore Geology Reviews*, vol. 71, pp. 804–818, 2015.

[7] S. Schneider, R. J. Murphy, and A. Melkumyan, "Evaluating the performance of a new classifier - the GP-OAD: A comparison with existing methods for classifying rock type and mineralogy from hyperspectral imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 98, pp. 145–156, 2014.

[8] P. J. Sylvester, "Use of the Mineral Liberation Analyzer (MLA) for Mineralogical Studies of Sediments and Sedimentary Rocks," *Mineralogical Association of Canada Short Course 42*, , no. 42, pp. 1–16, 2012.

[9] R. Fandrich, Y. Gu, D. Burrows, and K. Moeller, "Modern SEM-based mineral liberation analysis," *International Journal of Mineral Processing*, vol. 84, pp. 1–4, 2017.

[10] Y. Gu, "Automated Scanning Electron Microscope Based Mineral Liberation Analysis An Introduction to JKMRC/FEI Mineral Liberation Analyser," *Journal of Minerals and Materials Characterization and Engineering*, vol. 02, no. 01, pp. 33–41, 2003.

[11] K. Laakso, M. Middleton, T. Heinig, R. Bärs, and P. Lintinen, "Assessing the ability to combine hyperspectral imaging (HSI) data with Mineral Liberation Analyzer (MLA) data to characterize phosphate rocks," *International Journal of Applied Earth Observation and Geoinformation*, vol. 69, no. December 2017, pp. 1–12, 2018.

[12] V. Kuosmanen, H. Arkimaa, and M. Tiainen, "Hyperspectral close-range LWIR imaging spectrometry - 3 case studies," *Geological Survey of Finland, Special Paper 58*, pp. 117–144, 2015.

[13] S. Jakob, R. Zimmermann, and R. Gloaguen, "The Need for Accurate Geometric and Radiometric Corrections of Drone-Borne Hyperspectral Data for Mineral Exploration: MEPHySTo-A Toolbox for Pre-Processing Drone-Borne Hyperspectral Data," *Remote Sensing*, vol. 9, no. 1, 2017.

[14] D. L. Whitney and B. W. Evans, "Abbreviations for names of rock-forming minerals," *American Mineralogist*, vol. 95, no. 1, pp. 185–187, 2010.