THE APPLICATION OF SUBSPACE CLUSTERING ALGORITHMS IN DRILL-CORE HYPERSPECTRAL DOMAINING

Kasra Rafiezadeh Shahi, Mahdi Khodadadzadeh, Raimon Tolosana-delgado, Laura Tusa
, and Richard Gloaguen

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

ABSTRACT

Diamond drilling is used in the mining industry to extract drill-cores for characterising mineral deposits. Traditionally, drill-cores are visually analysed by an on-site geologist, subjected to geochemical analyses, and then, few representative samples subjected to additional high-resolution mineralogical studies. However, the choice in samples is frequently subjective and the mineralogical analyses are highly time-consuming. In order to optimize the choice of samples and accelerate the analyses, drill-cores can be partitioned into domains, and then, laboratory analyses can be carried out on selected domains. Nevertheless, in the mining industry, automatic drill-core domaining still remains a challenge. Recently, hyperspectral imaging has become an important technique for the analysis of drill-cores in a non-invasive and non-destructive manner. Several clustering algorithms of hyperspectral data are proposed for automatic drill-core domaining. In this paper, we suggest using advanced subspace clustering algorithms (i.e., sparse subspace clustering algorithm, spectral-spatial sparse subspace clustering algorithm). These algorithms work based on the self-representation property of the hyperspectral data. The clustering methods are tested on two drill-core samples which present different mineralogical and structural features. The subspace clustering algorithms are compared with the result of the K-means clustering algorithm. Our experimental results show that subspace clustering algorithms provide accurate drill-core domains and it is shown that including spatial information significantly improves the clustering results.

***Index Terms*—** Drill-core hyperspectral data, drill-core domaining, subspace clustering, spatial regularization, unsupervised learning

1. INTRODUCTION

In the mining industry, drill-cores are extracted to characterise areas with high underground mineral potential. Traditional drill-core analysis techniques consist of a visual description of drill-cores and geochemical analyses. Additional laboratory measurements are only performed on few representative samples due to the high costs and analysis time. The choice of these areas is important and can be obtained after performing drill-core domaining. Common methods to define the domains comprise a lithological characterisation based on either the visual analysis or the distribution of key chemical elements and elemental ratios [1] [2]. For instance in [2], the authors studied on using geochemical data to extract multi-scale geological boundaries of drill-holes. They proposed an automated method to detect sharp and gradational boundaries by using the results of the continuous wavelet transform method.

Alternatively, hyperspectral data provide valuable information through a large number of narrow spectral bands. These sensors normally cover the spectrum range from the visible to near-infrared, short-wave infrared, and long-wave infrared [3]. Over the past few decades, hyperspectral data are being used in various applications (e.g., medicine [4], vegetation analysis [5]). In the same way, during the exploration of mineral deposits, drill-core hyperspectral data are being increasingly acquired to characterise mineralized systems. It provides valuable mineralogical information that can be used to identify drill-core domains in a fast and accurate manner. Recently, the usage of different machine learning techniques is proposed to have fast and robust analyses of drill-core hyperspectral data [6] [7] [8]. In [6], the performance of different subspace clustering algorithms was evaluated to map minerals in drill-core hyperspectral data. The obtained results show that subspace clustering methods (e.g., sparse subspace clustering (SSC) [9] and low-rank representation clustering (LRR) [10]) are able to accurately map minerals in drill-core hyperspectral data.

Due to the spatial dependency of mineralogical features within drill-cores, including spatial information in clustering algorithms can play an important role in drill-core domaining. However, among the recent studies on drill-core hyperspectral data, only a few are focused on including spatial information in clustering algorithms. For instance, in [8], authors studied on a geo-statistical clustering method to cluster geochemical and hyperspectral drill-core data from an iron ore deposit by considering spatial continuity in drill-core samples. In this paper, in addition to sparse subspace clustering algorithm [9], spectral-spatial sparse subspace clustering algorithm which has been proposed in [11] is applied on drill-core hyperspectral data. The aim of this paper is to investigate the capability of subspace clustering algorithms in drill-core domaining. The rest of the paper is structured as follows: **Section** 2 contains the applied methodology. In **Section** 3, data description, experimental results, and discussions are presented. The conclusions are drawn in **Section** 4.

2. METHODOLOGY

To ease the explanation, let us denote $\mathbf{X}$ as an hyperspectral image which is expressed as $\mathbf{X} \equiv \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_P\}$, where $\mathbf{x}_i \in R^N$, for $i = 1, 2, ..., P$, and $\mathbf{x}_i$ is calculated as the average of data points in each depth, P is the number of pixels in $\mathbf{X}$, and N is the number of spectral bands.
In the subspace clustering theorem, hyperspectral data can be represented in a union of lower dimensional subspace where each subspace corresponds to a cluster. Furthermore, each data point can be written as a linear combination of other data points from the same subspace which is called the *self-expressiveness property* of the data. The linear combination of data points can be formulated as follow:

$$\mathbf{X} = \mathbf{X}\mathbf{Z} \tag{1}$$

where $\mathbf{Z}$ is the sparsest representation coefficient matrix and it is expressed as $\mathbf{Z} \equiv \{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, ..., \mathbf{z}_P\}$, which $\mathbf{Z} \in R^{P \times P}$. In [9], SSC solves the following optimization problem:

$$\begin{aligned} & min \quad ||\mathbf{Z}||_1 \\ S.t., \quad & \mathbf{X} = \mathbf{X}\mathbf{Z}, \quad diag(\mathbf{Z}) = 0 \end{aligned} \tag{2}$$

where $||\mathbf{Z}||_1$ is the $l_1 - norm$ of $\mathbf{Z}$. In the obtained $\mathbf{Z}$, there are nonzero elements whose ideally correspond to data points from the same subspace.
In the adopted spectral-spatial sparse subspace clustering method, the 3D median filter on the calculated sparsest coefficient matrix is used as a spatial term in the standard SSC formula [11] [12]. We refer to this adopted method as sparse subspace clustering algorithm with spatial regularization (SSC-med). In [12], the spatial information is included in the sparse optimization problem (2) as follow:

$$\begin{aligned} & min \quad ||\mathbf{Z}||_1 + \frac{\alpha}{2}||\mathbf{Z} - \bar{\mathbf{Z}}||_F^2 \\ S.t., \quad & \mathbf{X} = \mathbf{X}\mathbf{Z}, \quad diag(\mathbf{Z}) = 0 \end{aligned} \tag{3}$$

where the $||\mathbf{Z} - \bar{\mathbf{Z}}||_F^2$ is the spatial term, and $\bar{\mathbf{Z}}$ is the coefficient matrix after applying 3×3 median filter on $\mathbf{Z}$. In (3), α is the trade-off parameter between spectral and spatial information.
After solving the sparse optimization problem in (2) and (3), the obtained $\mathbf{Z}$ is used to calculate the symmetric similarity matrix $\mathbf{W}$.

$$\mathbf{W} = |\mathbf{Z}| + |\mathbf{Z}|^T \tag{4}$$

where $|\mathbf{Z}|^T$ is the transposed matrix of $\mathbf{Z}$. Note that $\mathbf{W}$ is needed to be symmetrised to make sure that all data points from the same subspace are connected to each other. Finally, the spectral clustering algorithm is applied on $\mathbf{W}$ to produce the final clustering results.

3. EXPERIMENTAL RESULTS

3.1. Data acquisition

A SisuRock drill-core scanner equipped with an AisaFENIX VNIR-SWIR hyperspectral sensor that was employed to capture hyperspectral images of drill-core samples as it is presented in Fig 1. The FENIX sensor is covering the spectral range of 380 - 2500 nm in 450 bands. The obtained pixel size is 1.5 mm/pixel. Furthermore, the drill-core samples were analysed visually and segmented into domains based on the present veins and matrix composition. The RGB images of two obtained drill-core samples and generated reference data by a geologist are presented in Fig 2.

Fig. 1: The SisuRock drill-core scanner with AisaFENIX VNIR-SWIR hyperspectral sensor

3.2. Clustering results

The clustering algorithms (i.e., SSC, SSC-med) are tested using default parameters suggested in [9] [12]. In order to have a quantitative validation of obtained clustering results, adjusted rand index (ARI) is used [13]. In ARI, the similarity between generated reference data and obtained clustering results is measured. The results of applied clustering algorithms are presented in Fig 2. As it can be observed, the employed subspace clustering algorithms perform well to characterise present veins and alteration halos comparing to the traditional K-means algorithm for both analysed samples. In **Sample 1**, the standard SSC and SSC-med could accurately cluster the small mineralogical changes in comparison to the traditional K-means. For instance, subspace clustering algorithms are able to identify the transition from a low to a high quartz vein density. SSC is not locally able to distinct between quartz veins and quartz-gypsum-sulphide ones comparing to SSC-med. A reason for the mislabeling, can be the slight variation in the alteration halo intensity surrounding the quartz veins. The second sample includes a stronger variability in the vein distribution and orientation. These strong variabilities lead SSC-based methods to have considerably different performance. SSC provides non-smooth results within small intervals of the drill-core. In addition, SSC is not robust against minor changes in vein orientations and abundances. While, SSC-med shows good results by comparing to the reference data, and it is robust against negligible mineralogical changes. According to these changes, K-means performs well, nonetheless it is not able to distinguish between fine quartz veins with high density and coarse chlorite-chalcopyrite veins. As it is also discussed in **Sample 1**, the obtained results can be caused by the similarity of the composition of the alteration halo from the two vein types. Although, the mineral composition and their economic importance are different, accurate results for drill-core domaining which as obtained by SSC-med are required. Regarding the obtained accuracy results of applied clustering algorithms, in **Sample 1**, SSC obtains more accurate results comparing to the traditional K-means by 4%. Whilst the most accurate clustering result belongs to SSC-med (60.98%). In **Sample 2**, again SSC-med provides the highest accurate clustering result in contrary with the other studied methods. Whereas, in this sample, traditional K-means performs well in comparison to the standard SSC by 6%. The complete accuracy results are presented in Table1.

4. CONCLUSIONS AND REMARKS

In this paper, in order to automate and accelerate the analysis of drill-core hyperspectral data, the performance of advanced unsupervised machine learning methods is investigated. The assumption of the self-expressiveness of drill-core hyperspectral data encouraged us to investigate the use of subspace clustering algorithms in drill-core domaining. The obtained clustering results show that subspace clustering algorithms provide reliable drill-core domains in comparison to the results of the traditional K-means clustering algorithm. Furthermore, the exploitation of spatial information in the standard SSC significantly improves the clustering results in different types of drill-core samples, particularly where spatially and compositionally heterogeneous features such as veins with various orientations are present. As future work, we will explore different strategies of exploiting spatial information in subspace clustering algorithms. We will also examine the application of subspace clustering methods for generating 3D geological models of mineralized systems.

5. REFERENCES

[1] S.M. Gandhi and B.C. Sarkar, "Chapter 8 - drilling," in *Essentials of Mineral Exploration and Evaluation*, S.M. Gandhi and B.C. Sarkar, Eds., pp. 199 – 234. Elsevier, 2016.

[2] E.J. Hill and Y. Uvarova, "Identifying the nature of lithogeochemical boundaries in drill holes," *Journal of Geochemical Exploration*, vol. 184, pp. 167 – 178, 2018.

[3] A.F.H. Goetz, G. Vane, J.E. Solomon, and B.N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, 1985.

[4] G. Lu, L. Halig, D. Wang, X. Qin, Z.G. Chen, and B. Fei, "Spectral-spatial classification for noninvasive cancer detection using hyperspectral imaging," *Journal of Biomedical Optics*, vol. 19, no. 10, pp. 106004, 2014.

[5] J.L. Hatfield and J.H. Prueger, "Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices," *Remote Sensing*, vol. 2, no. 2, pp. 562–578, 2010.

[6] M. Khodadadzadeh, C. Contreras, L. Tusa, and R. Gloaguen, "Subspace clustering algorithms for mineral mapping," in *Proc. SPIE 10789, Image and Signal Processing for Remote Sensing XXIV*, 2018, vol. 10789.

[7] Y. Guo, J. Gao, and F. Li, "Spatial subspace clustering for drill hole spectral data," *Journal of Applied Remote Sensing*, vol. 8, pp. 1 – 20 – 20, 2014.

[8] F. Fouedjio, E.J. Hill, and C. Laukamp, "Geostatistical clustering as an aid for ore body domaining: case study at the rocklea dome channel iron ore deposit, western australia," *Applied Earth Science*, vol. 127, no. 1, pp. 15–29, 2018.

[9] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2765–2781, Nov 2013.

[10] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 171–184, Jan 2013.

[11] H. Zhang, H. Zhai, L. Zhang, and P. Li, "Spectralspatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3672–3684, June 2016.

Table 1: The acquired adjusted rand index of applied clustering algorithms in order to evaluate their performance

Sample's number/Clustering algorithms	**K-means**	**SSC**	**SSC-med**
Sample 1	39.21%	43.33%	60.98%
Sample 2	38.87%	32.71 %	62.28%

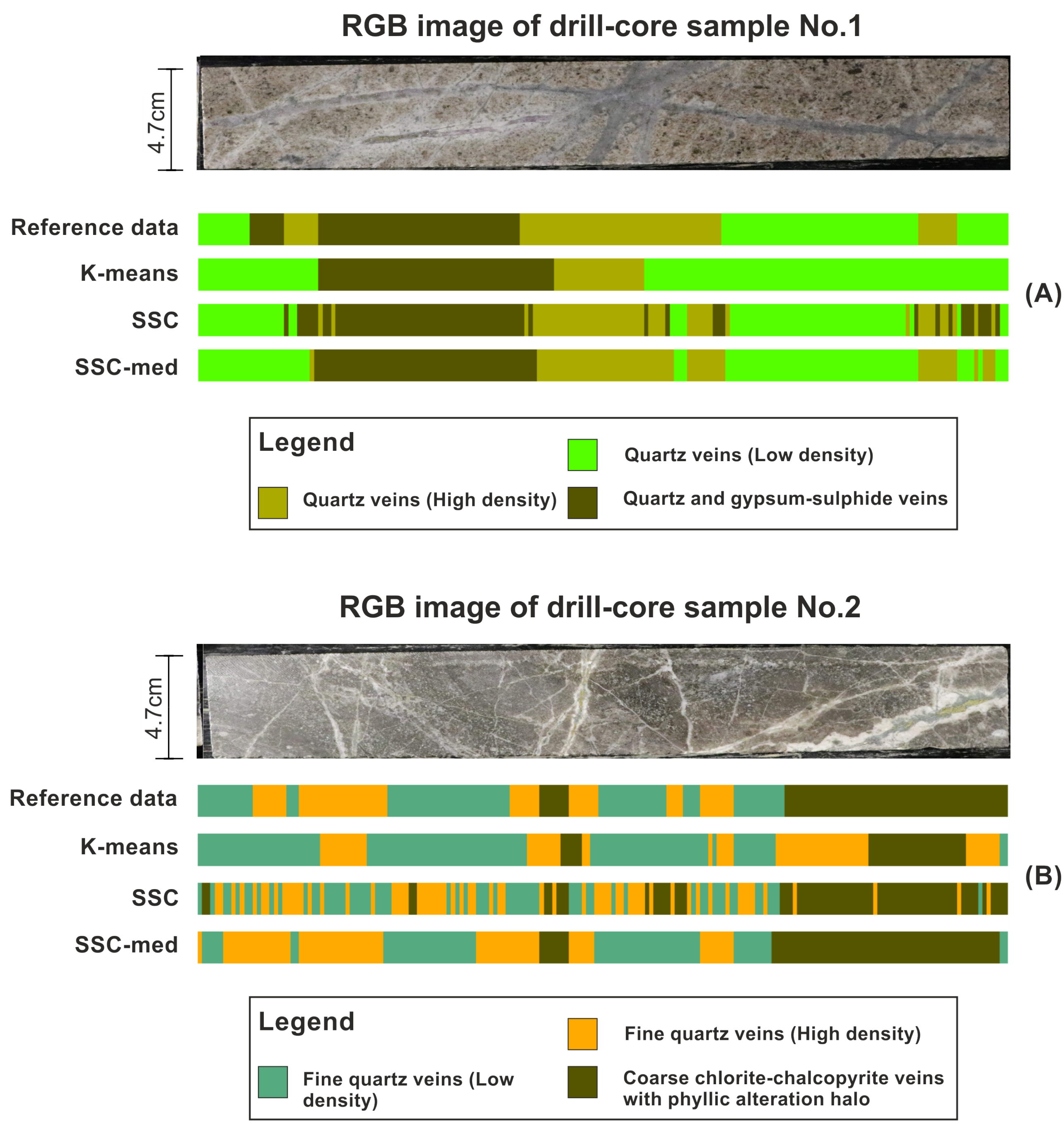

Fig. 2: The RGB of drill-cores ((A) Sample 1, (B) Sample 2), their generated reference data, and the clustering results of applied traditional K-means, SSC, SSC-med for each sample are presented

[12] C. Hinojosa, J. Bacca, and H. Arguello, "Coded aperture design for compressive spectral subspace clustering," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 6, pp. 1589–1600, Dec 2018.

[13] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, Dec 1985.