A Background Correction Algorithm for Hyperspectral Images

Said Sadeg, Jean Cauzid, Cécile Fabre
Université de Lorraine, CNRS, GeoRessources
F-54000 Nancy, France
{firstname.lastname}@univ-lorraine.fr

Yingying Song
Chengdu University of Technology
Chengdu, China
yingying.song.ys@icloud.com

David Brie, El-Hadi Djermoune
Université de Lorraine, CNRS, CRAN
F-54000 Nancy, France
{firstname.lastname}@univ-lorraine.fr

***Abstract*—This paper introduces a method for background (or baseline) correction in hyperspectral images. The method is based on the optimization of a criterion incorporating a non-quadratic robust loss (data fidelity term) and both spatial and spectral regularization terms to enforce baseline smoothness. Unlike the classical approach based on a pixel-by-pixel baseline correction, the proposed algorithm exploits jointly the spatial and spectral information. The effectiveness of the the proposed algorithm is demonstrated using simulated and geological hyperspectral images.**

***Index Terms*—Hyperspectral image processing, background estimation, spatial and spectral regularization**

I. Introduction

Hyperspectral imaging has received considerable interest in the last years as it is of major importance in many fields of application such as medical imaging [1], remote sensing [2], chemical characterization of materials [3], [4], and food quality control [5], [6]. This method allows to obtain more detailed information than those provided by conventional imaging techniques based only on three color channels or, at least, a few spectral bands. In contrast, each pixel of a hyperspectral image contains several spectral bands that can range from X-rays to infrared.

The spectral-spatial information provided by hyperspectral imaging systems is often embedded in the so-called background spectrum related to the acquisition system or the imaged material. The latter often appears as a sample-independent smooth curve in the spectrum corresponding to a given pixel. Several methods have been proposed to fit and substract this baseline from raw spectra. Some of them are now discussed. Asymmetric Least Squares (ALS) [7] seeks to iteratively estimate the baseline using weights that penalize rapid changes. An improvement to this method, named Adaptive Iteratively Reweighted Penalized Least Squares (airPLS), uses an exponential weighting of the negative residuals [8]. Another improvement (Improved Asymmetrically Reweighted Penalized Least Squares, iarPLS) was recently proposed by Ye *et al.* [9]. It uses an adjusted weighting formula to improve the baseline estimation when fitting small peaks in noisy data. Other approaches include Peaked Signal's Asymmetric Least Squares Algorithm (PSALSA) [10], Iterative Reweighted Spline Quantile Regression [11], and Statistics-sensitive Non-linear Iterative Peak-clipping (SNIP) [12], [13] based on peak clipping. The method proposed in [14] approximates the baseline with a polynomial function whose coefficients are estimated by minimizing a non-quadratic cost function. Another way to fit a baseline is to use deep learning techniques. For instance, Schmidt *et al.* [15] proposed a convolutional neural network approach to estimate the baseline. A trained adversial neural network was also presented in [16].

All of these methods aim to fit a baseline to a single spectrum associated with a pixel in a hyperspectral image. When processing a hyperspectral image, it is obviously useful to exploit neighboring pixels to improve baseline estimation. It is worth to mention that the algorithm presented in [17] is able to perform baseline correction from multiple spectra collected several times for the same sample. The objective of this paper is to propose an algorithm, tailored for hyperspectral images, and in which baseline estimation also accounts for the similarity of immediate neighboring pixels. We also propose an implementation scheme to avoid direct inversion of the underlying large matrices.

The remainder of this paper is organized as follows. In Section II, we present a regularized non-quadratic strategy to estimate the baseline of one-dimensional spectra. This strategy is then extended in Section III to deal with hyperspectral images with spectral and spatial regularizations. We present the proposed method and discuss its numerical implementation. Experiments using simulated and real hyperspectral images are detailed in Section IV. Finally, the conclusions are drawn in Section V.

II. Single spectrum background estimation

Let $\mathbf{y} \in \mathbb{R}^p$ be a p-points spectrum. Each point corresponds to a spectral band noted y_i, $i = 1, 2, \ldots, p$. It is assumed that $\mathbf{y}$ is the superposition of a smooth baseline $\mathbf{x} \in \mathbb{R}^p$, the useful spectrum composed of peaks whose positions and intensities depend on the considered sample, and noise. When the intensity of all useful peaks is greater than the baseline, it has already been reported that estimating the latter using an asymmetric least squares-type method leads to better results [14]. Baseline correction can be formulated as the minimization of the following cost function:

$$J(\mathbf{x}) = \sum_{i=1}^{p} \varphi(y_i - x_i) + \alpha ||\mathbf{D}_p \mathbf{x}||_2^2, \tag{1}$$

where $\varphi(x)$ is a non-quadratic and asymmetric function which strongly penalizes negative and small values of x and whose cost is lower for large positive values. This property allows for robust estimation as the minimizer is less sensitive to outliers and/or large peaks in data than the squared error loss. In this paper, we use the asymmetric Huber function defined as follows:

$$\varphi(x) = \begin{cases} x^2, & \text{if } x < s, \\ 2sx - s^2, & \text{otherwise.} \end{cases} \quad (2)$$

where $s \geq 0$ is a tuning parameter. Notation $||\cdot||_2$ stands for the Euclidean norm and $\mathbf{D}_p \in \mathbb{R}^{(p-1)\times p}$ is a finite difference matrix of order 1:

$$\mathbf{D}_p(i,j) = \begin{cases} -1, & \text{if } j = i, \\ +1, & \text{if } j = i+1, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The parameter $\alpha > 0$ allows one to control the smoothness of the baseline along the spectral dimension. As the objective function in (1) is not quadratic in $\mathbf{x}$, its minimizer cannot be expressed analytically. The method proposed in [18] consists in using a half-quadratic minimization which is more efficient than gradient methods. The half-quadratic minimization consists of introducing auxiliary variables $\mathbf{z} = [z_1, \ldots, z_p]^\top$ in a new criterion $K(\mathbf{x}, \mathbf{z})$ that admits the same minimum as the original criterion $J(\mathbf{x})$:

$$\begin{aligned} K(\mathbf{x}, \mathbf{z}) = & \frac{1}{c} \sum_{i=1}^{p} \tfrac{1}{2} \left[(y_i - x_i - z_i)^2 + u(z_i) \right] \\ & + \alpha ||\mathbf{D}_p \mathbf{x}||_2^2, \end{aligned} \quad (4)$$

where $u(z) = \sup_x (c\varphi(x) - (x-z)^2/2)$ and $c = 1/2$ [14]. The criterion $K(\mathbf{x}, \mathbf{z})$ is half-quadratic because it is quadratic in $\mathbf{x}$ and convex in $\mathbf{z}$. The minimizer $(\hat{\mathbf{x}}, \hat{\mathbf{z}})$ is calculated using an alternating approach: for a given $\mathbf{x}$, we seek for the minimizer over $\mathbf{z}$, and with this value the criterion is then minimized with respect to $\mathbf{x}$. Let $(\mathbf{x}^{(k-1)}, \mathbf{z}^{(k-1)})$ be the solution at iteration $(k-1)$, and $(\mathbf{x}^{(0)}, \mathbf{z}^{(0)})$ as the initial point. At iteration k we calculate:

$$\begin{aligned} \mathbf{z}^{(k)} &= \arg\min_{\mathbf{z}} K(\mathbf{x}^{(k-1)}, \mathbf{z}) \\ &\Longrightarrow z_i^{(k)} = -(y_i - x_i^{(k-1)}) + c\varphi'(y_i - x_i^{(k-1)}), \quad (5) \\ \mathbf{x}^{(k)} &= \arg\min_{\mathbf{x}} K(\mathbf{x}, \mathbf{z}^{(k-1)}) \\ &\Longrightarrow \mathbf{x}^{(k)} = (\mathbf{I}_p + \alpha \mathbf{D}_p^\top \mathbf{D}_p)^{-1} (\mathbf{y} + \mathbf{z}^{(k-1)}), \quad (6) \end{aligned}$$

with $\varphi'(x) = 2\min(x, s)$ and $\mathbf{I}_p \in \mathbb{R}^{p\times p}$ is the identy matrix. We assume that the convergence is reached when $||\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}||_2 / ||\mathbf{x}^{(k)}||_2 < \epsilon$, where $\epsilon = 10^{-6}$ in this paper.

III. Hyperspectral image background estimation

A. Algorithm

In this section, we present our proposed method to estimate the baseline of an entire data cube. Instead of considering spectra as independent from each other, we stack those spectra in a tensor $\mathcal{Y} \in \mathbb{R}^{m\times n\times p}$ along the third dimension while the first two represent the spatial dimensions. We then add a penalty to the cost function, accounting for the similarity of the baseline over neighboring pixels, to estimate the baseline $\mathcal{X} \in \mathbb{R}^{m\times n\times p}$ for the image $\mathcal{Y}$. Let us denote $\mathcal{A} \bullet_d \mathbf{B}$ the p-mode product of a tensor $\mathcal{A}$ and a matrix $\mathbf{B}$ [19]. In this product, summation is performed on the second index of the matrix $\mathbf{B}$, e.g., if $\mathcal{A} \in \mathbb{R}^{m\times n\times p}$ and $\mathbf{B} \in \mathbb{R}^{q\times m}$, then $\mathcal{A} \bullet_1 \mathbf{B} \in \mathbb{R}^{q\times n\times p}$ with $[\mathcal{A} \bullet_1 \mathbf{B}]_{ijk} = \sum_{\ell=1}^{m} \mathcal{A}_{\ell jk} B_{i\ell}$. The new cost function can be expressed as

$$\begin{aligned} J(\mathcal{X}) = \phi(\mathcal{Y} - \mathcal{X}) + \alpha ||\mathcal{X} \bullet_3 \mathbf{D}_p||_F^2 + \\ \beta \left(||\mathcal{X} \bullet_1 \mathbf{D}_m||_F^2 + ||\mathcal{X} \bullet_2 \mathbf{D}_n||_F^2 \right), \end{aligned} \quad (7)$$

where $||\cdot||_F$ stands for the Frobenius norm and β is the parameter that controls the spatial smoothness of the baseline. With a slight abuse of notation $\phi(\mathcal{Y} - \mathcal{X}) = \sum_{i,j,k} \varphi(\mathcal{Y}_{ijk} - \mathcal{X}_{ijk})$. The last term on the RHS of (7) penalizes large deviations of the baselines among connected pixels. It is also worth to notice that this penalty is separable along the two spatial dimensions. This choice is motivated by the need to design a fast algorithm able to process large data cubes.

Now, we need to vectorize equation (7) in order to build an algorithm estimating the entries of tensor $\mathcal{X}$. Let $\mathbf{x} = \text{vec}\,(\mathcal{X})$ and $\mathbf{y} = \text{vec}\,(\mathcal{Y})$, where $\text{vec}\,(\cdot)$ denotes the vectorization operator applied on a matrix or a tensor. We denote $\mathbf{X}_{(d)}$ the mode-d matricization (or flattening) of the tensor $\mathcal{X}$; this operation arranges the mode-d fibers to be the columns of the resulting matrix [19]. The following property will help us to express the matricization operation by means of Kronecker products [20]:

$$\begin{aligned} & \mathcal{Y} = \mathcal{X} \bullet_1 \mathbf{A}_1 \cdots \bullet_N \mathbf{A}_N \Longleftrightarrow \\ & \mathbf{Y}_{(d)} = \mathbf{A}_d \mathbf{X}_{(d)} \left(\mathbf{A}_N \otimes \cdots \mathbf{A}_{d+1} \otimes \mathbf{A}_{d-1} \otimes \cdots \mathbf{A}_1 \right)^\top, \end{aligned} \quad (8)$$

where $\otimes$ stands for the Kronecker product. The vector form of the first term on the RHS of (7) can be easily expressed as

$$\text{vec}\,(\mathcal{Y} - \mathcal{X}) = \mathbf{y} - \mathbf{x}. \quad (9)$$

Moreover, using the flattening property in (8) and the fact that $\text{vec}\,(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A})\text{vec}\,(\mathbf{B})$, we obtain:

$$\begin{aligned} \text{vec}\,(\mathcal{X} \bullet_3 \mathbf{D}_p) &= \text{vec}\,(\mathcal{X} \bullet_1 \mathbf{I}_m \bullet_2 \mathbf{I}_n \bullet_3 \mathbf{D}_p) \\ &= \text{vec}\left(\mathbf{I}_m \mathbf{X}_{(1)} (\mathbf{D}_p \otimes \mathbf{I}_n)^\top \right) \\ &= (\mathbf{D}_p \otimes \mathbf{I}_n \otimes \mathbf{I}_m) \text{vec}\left(\mathbf{X}_{(1)} \right) \\ &= (\mathbf{D}_p \otimes \mathbf{I}_n \otimes \mathbf{I}_m)\mathbf{x}, \end{aligned} \quad (10)$$

Proceeding in the same way for $(\mathcal{X} \bullet_1 \mathbf{D}_m)$ and $(\mathcal{X} \bullet_2 \mathbf{D}_n)$, we get:

$$\text{vec}\,(\mathcal{X} \bullet_1 \mathbf{D}_m) = (\mathbf{I}_p \otimes \mathbf{I}_n \otimes \mathbf{D}_m)\mathbf{x}, \quad (11)$$

$$\text{vec}\,(\mathcal{X} \bullet_2 \mathbf{D}_n) = (\mathbf{I}_p \otimes \mathbf{D}_n \otimes \mathbf{I}_m)\mathbf{x}. \quad (12)$$

Finally, equation (7) can be rewritten as follows:

$$\begin{aligned} J(\mathbf{x}) = \phi(\mathbf{y} - \mathbf{x}) + \alpha ||(\mathbf{D}_p \otimes \mathbf{I}_{mn})\mathbf{x}||_2^2 \\ + \beta ||(\mathbf{I}_{np} \otimes \mathbf{D}_m)\mathbf{x}||_2^2 + \beta ||(\mathbf{I}_p \otimes \mathbf{D}_n \otimes \mathbf{I}_m)\mathbf{x}||_2^2. \end{aligned} \quad (13)$$

Algorithm 1: Hyperspectral image baseline correction algorithm

Input : $\mathcal{Y} \in \mathbb{C}^{m\times n\times p}$, s, α, β, ϵ

Initialization: $\mathbf{y} \leftarrow \text{vec}\,(\mathcal{Y})$, $\mathbf{x}^{(0)} \leftarrow \mathbf{0}$, $\mathbf{z}^{(0)} \leftarrow \mathbf{0}$, $k \leftarrow 0$, $\mathbf{A}$: `sparse`;

repeat

  $k \leftarrow k + 1$;

  $z_i^{(k)} \leftarrow -(y_i - x_i^{(k-1)}) + c\varphi'(y_i - x_i^{(k-1)})$, for $i = 1, \ldots, mnp$;

  $\mathbf{x}^{(k)} \leftarrow \texttt{pcg}(\mathbf{A}, \mathbf{y} + \mathbf{z}^{(k-1)})$;

until $||\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}||_2 < \epsilon$;

$\mathcal{X} \leftarrow \texttt{unfold}(\mathbf{x})$;

Output: $\mathcal{X} \in \mathbb{R}^{m\times n\times p}$

Therefore, the minimizer of $J(\mathbf{x})$ can be computed similarly to the one-dimensional case. The hyperspectral counterpart of the iteration in (6) is:

$$\mathbf{x}^{(k)} = \mathbf{A}^{-1}(\mathbf{y} + \mathbf{z}^{(k-1)}), \tag{14}$$

except that here $\mathbf{z}^{(k)} \in \mathbb{R}^{mnp}$, and

$$\mathbf{A} = \mathbf{I}_{mnp} + \alpha(\mathbf{D}_p^\top \mathbf{D}_p) \otimes \mathbf{I}_{mn} + \beta \mathbf{I}_{mp} \otimes (\mathbf{D}_m^\top \mathbf{D}_m) + \beta \mathbf{I}_p \otimes (\mathbf{D}_n^\top \mathbf{D}_n) \otimes \mathbf{I}_m. \tag{15}$$

B. Implementation

As most of iterative algorithms for baseline correction, the proposed scheme involves the inversion matrix $\mathbf{A}$ in (15). This is a very challenging issue, both from memory requirement and calculation time points of view, because the dimension of the matrix is $mnp \times mnp$ which can be very large for large data cubes. Fortunately, as $\mathbf{A}$ is a sparse matrix containing only a small amount of non-zero entries, efficient iterative techniques have been developed to solve a system of linear equations of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ without explicitly inverting matrix $\mathbf{A}$. For instance, we can cite the conjugate gradient and Lanczos algorithms [21], [22] as well as the Wiedemann algorithm [23]. In this paper, only the non-zero entries of $\mathbf{A}$ are stored in memory[1] and we use the preconditioned conjugate gradient method (PCG), as sketched in Algorithm 1. The Matlab code is available on request from the authors.

IV. Experimental results

In this section, we will discuss the results obtained by applying the algorithm on both synthetic and real geological hyperspectral images.

A. Simulated data

In this first experiment, our objective is to assess the quality of the estimated background in a fully controlled setting; i.e. when the true baseline is perfectly known. The hyperspectral image is of size $10 \times 10 \times 1000$. Each pixel is a spectrum composed of a sum of narrow Gaussian peaks embedded in a wide Gaussian baseline and an additive zero-mean white noise. The baseline is subject to random shifts from one pixel to another; the amplitude of the shifts obeys a uniform law in the interval $[-\frac{1}{4}, \frac{1}{4}]$. A spectrum extracted from the resulting image is shown in Figure 1(a). During our experiments we have tested different signal-to-noise ratios (SNRs) to compare the influence that have the spatial regularization on the baseline fitting. We use the root-mean-square error (RMSE) metric to assess the fitting quality; it is defined by $\text{RMSE} = ||\hat{\mathcal{X}} - \bar{\mathcal{X}}||_F/\sqrt{mnp}$, where $\hat{\mathcal{X}}$ and $\bar{\mathcal{X}}$ are the estimated and mean baseline over all pixels, respectively. The parameter α is set to 1500. As s depends on the noise level, its value is fixed at each SNR so as to obtain the best fit. As can be seen in Figure 1(b), the RMSE decreases as the SNR increases. The RMSE obtained with a spatial regularization ($\beta = 0.01$) is smaller than the one with $\beta = 0$ even at small SNR.

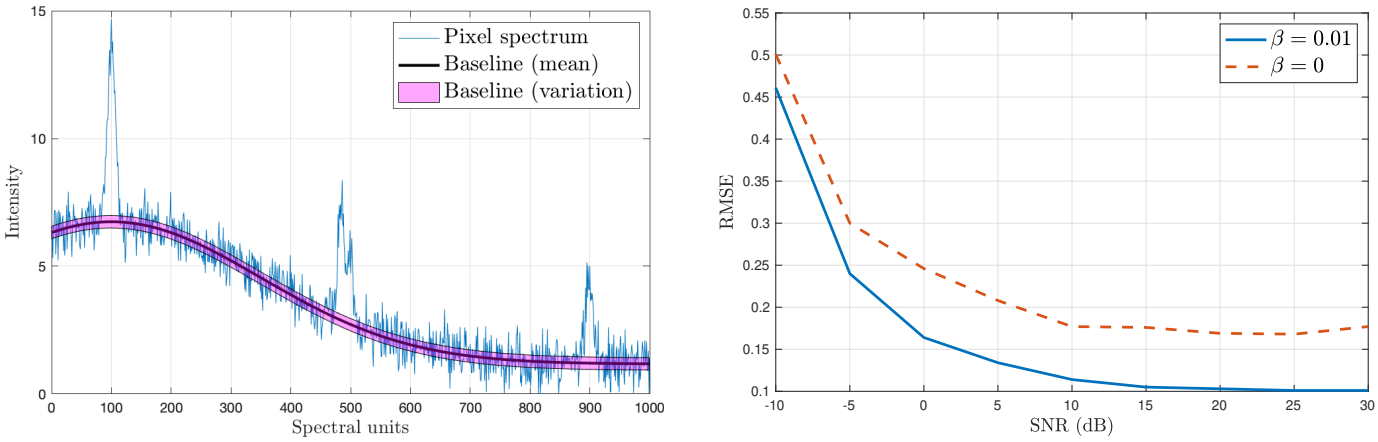
B. Application on real geological data

The geological data comes from a μ-XRF hyperspectral image obtained on a thin section of a solid sample from the Saint-Melany Sn-W deposit in French Massif Central. This sample is composed of wolframite and cassiterite macro-crystals surrounded by quartz. During our experiment we imaged a subsection from this sample centered on a cassiterite crystal with quartz surrounding it. Data acquisition was done with a Bruker Tornado M4 equipped with a rhodium (Rh) tube working under a voltage of 50 kV and an intensity of 300 mA. Detection was done with two 30 mm^2 XFlash SDD detectors. The pixel size was set to 20 μm with a dwell time of 20 ms in each pixel.

Due to the low SNR, we performed spatial binning to group pixel patches of size 10×10 in one pixel of a new hyperspectral image. Hence, the original image of size $550 \times 550 \times 4096$ was reduced to $55 \times 55 \times 4096$ containing a total of 3025 spectra. The proposed algorithm is running under Matlab on a computer equipped with Intel Xeon 3.20 GHz with 32 GB of RAM. Different value of β have been tested to see the effect of the spatial regularization on the baseline fitting. The other parameters were fixed: $\alpha = 1500$ and $s = 2.5$. The results are compared with those obtained using SNIP algorithm which is applied sequentially on each spectrum in the image. The half-window length in SNIP was set at 50.

As shown in Figure 2(a), when we do not apply spatial regularization to the hyperspectral image, we observe that the baseline follows too much the variations of the spectrum even in areas without peaks and picks up parts of the signal especially between 5 and 10 keV. When we set $\beta = 0.007$, we note that the baseline is smoother and passes through the lower parts of the peaks, especially at the fluorescence peaks, although it remains high at the scattering peaks from 18 to 23 keV. However, with a strong spatial penalty, the baseline does not average out the noise in areas where there are no spectral peaks. Using SNIP algorithm, we can see that the baseline is lower than the one obtained with the proposed algorithm with $\beta = 0.007$, particularly in the range 9–15 keV where we see that the former is clearly underestimated. When

[1]The number of non-zero entries in $\mathbf{A}$ is linear in mnp.



(a) Example of a pixel spectrum with its baseline for SNR = 5 dB

(b) RMSE as a function of SNR

Fig. 1. Results of baseline fit on the simulated hyperspectral image

we compare all baselines together in Figure 2(b), we observe that, with spatial regularization, the curves become closer and closer as β increases. A $\beta = 0.5$ ends with a baseline almost constant on the entire image, which is not expected as density influences the baseline shape and quartz has a much lower density compared to cassiterite. Hence, a $\beta = 0.007$ was chosen for the following steps.

Once the background of the image has been corrected, we perform K-means clustering to check whether the fluorescence information is preserved. In Figure 3, we show the clustering on the original data without extracting the baseline and with baseline extraction. We choose to cluster the data into three groups: the central cluster represents the cassiterite crystal, the second cluster is composed of quartz surrounding the crystal, and we added a third cluster for the other minerals present in small quantities in the sample and pixels in which more than one crystal was present. We note that the clustering applied to the raw image and the one applied to the corrected one are almost identical with no lost information from data peaks. Even if the calculated baseline sometimes enters peaks, that subtraction remains negligible compared to the preserved signal. Most subtraction of peaks from the baseline occurs on scattering peaks (18 to 23 keV), which holds no elementary information. At the same time, in Figure 3(c), we can observe the baseline of each cluster. Cluster "other" corresponds to pixels included or at the edge of the cassiterite crystal (Figures 3(a) and 3(b)). It is then expected that some cassiterite signal will be held into these pixels, which is visible with baselines "cassiterite" and "other" being similar (Figure 3(c)). On the contrary, the "quartz" baseline is really different with a much higher elevation between 5 and 23 keV, which is expected from the lower quartz density. The ability of the algorithm to maintain this feature when the spatial regularization is not too high is essential in its applicability to geological samples.

V. Conclusion

We proposed a baseline correction algorithm for hyperspectral images capable to enforce smoothness along spatial and spectral dimensions. The complexity of the algorithm is linear in data length: provided that the data can be stored in memory, it is able to process large hyperspectral images. During our experiments, we tested several parameters and configurations to assess the influence of the additional spatial regularization on synthetic and real data. When compared to a classical method that does not use spatial information, we observed an improvement in the smoothness of the baseline. However, the choice of spectral and spatial parameters remains a significant challenge that must be carefully manipulated to prevent large fluctuations in the baseline. Obtaining the density information from the baseline of spectra opens avenues to propagate the surface information from X-Ray fluorescence to the third spatial dimension accessible from X-ray computed tomography, which, in first approximation, depends on sample density.

References

[1] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," *Journal of Biomedical Optics*, vol. 19, no. 1, p. 010901, 2014.

[2] W. L. Wolfe, *Introduction to imaging spectrometers*. SPIE Press, 1997.

[3] A. Thumm, M. Riddell, B. Nanayakkara, J. Harrington, and R. Meder, "Near infrared hyperspectral imaging applied to mapping chemical composition in wood samples," *Journal of Near Infrared Spectroscopy*, vol. 18, no. 6, pp. 507–515, 2010.

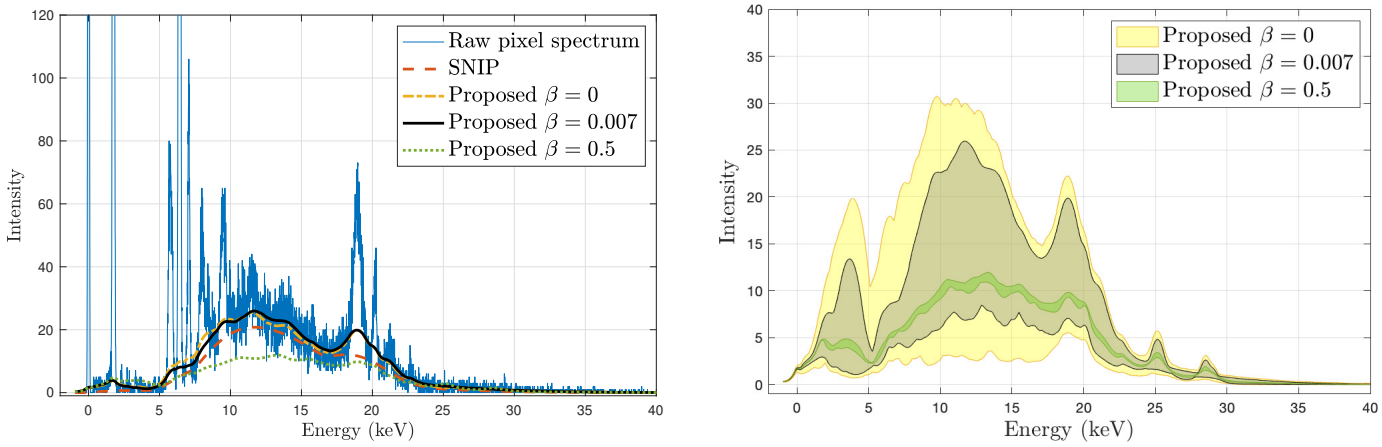
[4] J. J. Kim, F. T. Ling, D. A. Plattenberger, A. F. Clarens, A. Lanzirotti, M. Newville, and C. A. Peters, "Smart mineral mapping: Synchrotron-based machine learning approach for 2D characterization with coupled micro XRF-XRD," *Chemom. Intell. Lab. Syst.*, vol. 156, 2021.

[5] D. F. Barbin, G. ElMasry, D.-W. Sun, and P. Allen, "Non-destructive determination of chemical composition in intact and minced pork using near-infrared hyperspectral imaging," *Food Chemistry*, vol. 138, no. 2-3, pp. 1162–1171, 2013.

[6] H. Huang, L. Liu, and M. O. Ngadi, "Recent developments in hyperspectral imaging for assessment of food quality and safety," *Sensors*, vol. 14, no. 4, pp. 7248–7276, 2014.

[7] P. Eilers and H. Boelens, "Baseline correction with asymmetric least squares smoothing," Leiden University Medical Centre Report, Tech. Rep., 11 2005.

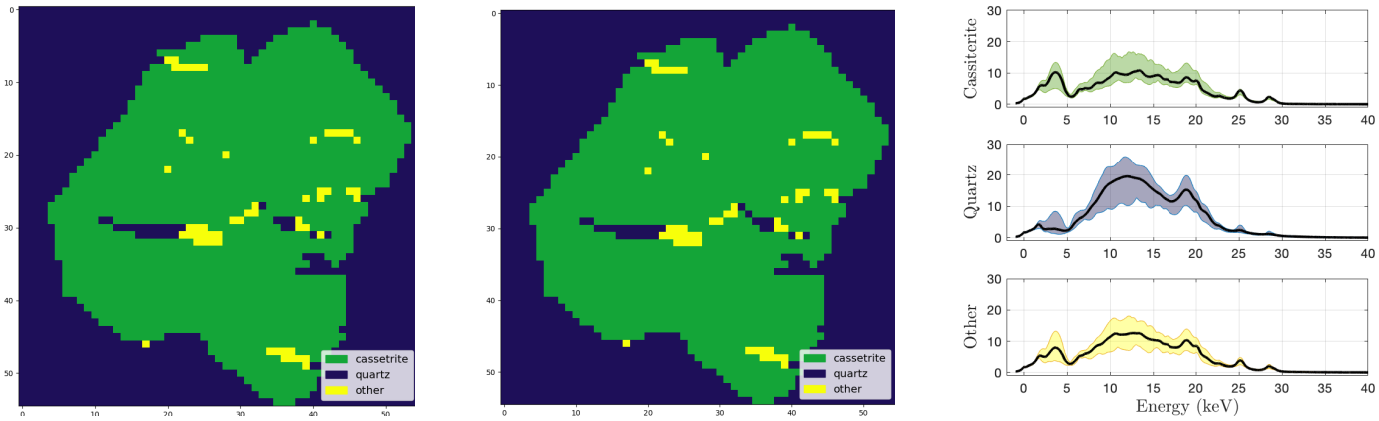
[8] Z. M. Zhang, S. Chen, and Y. Z. Liang, "Baseline correction using adaptive iteratively reweighted penalized least squares," *Analyst*, vol. 135, no. 5, pp. 1138–1146, 2010.



(a) Baseline at one pixel

(b) Baseline distribution over all image pixels

Fig. 2. Baseline fit on the geological hyperspectral image. (a) Example on a 10×10 cluster with various spatial regularization factors and SNIP. (b) Variability of the baseline on the entire image with the same spatial regularization factors as in (a)



(a) Original image

(b) Cleaned image with $\beta = 0.007$

(c) Baseline distribution in the 3 clusters

Fig. 3. Clustering results on (a) original data and (b) cleaned data. Green = cassiterite; Dark blue = quartz; Yellow = other crystals

[9] J. Ye, Z. Tian, H. Fei, and Y. Li, "Baseline correction method based on improved asymmetrically reweighted penalized least squares for the Raman spectrum," *Applied Optics*, vol. 59, no. 34, pp. 10 933–10 943, 2020.

[10] S. Oller-Moreno, A. Pardo, J. M. Jimenez-Soto, J. Samitier, and S. Marco, "Adaptive asymmetric least squares baseline estimation for analytical instruments," in *IEEE 11th International Multi-Conference on Systems, Signals Devices (SSD14)*, 2014, pp. 1–5.

[11] Q. Han, S. Peng, Q. Xie, Y. Wu, and G. Zhang, "Iterative reweighted quantile regression using augmented Lagrangian optimization for baseline correction," in *2018 5th International Conference on Information Science and Control Engineering (ICISCE)*, Los Alamitos, CA, USA, Jul. 2018, pp. 280–284. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/ICISCE.2018.00066

[12] C. G. Ryan, E. Clayton, W. L. Griffin, S. H. Sie, and D. R. Cousens, "SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications," *Nuclear Instruments and Methods in Physics Research B*, vol. 34, no. 3, pp. 396–402, Sep. 1988. [Online]. Available: https://ui.adsabs.harvard.edu/abs/1988NIMPB..34..396R

[13] M. Morháč, J. Kliman, V. Matoušek, M. Veselský, and I. Turzo, "Background elimination methods for multidimensional coincidence γ-ray spectra," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 401, no. 1, pp. 113–132, 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900297010231

[14] V. Mazet, C. Carteret, D. Brie, J. Idier, and B. Humbert, "Background removal from spectra by designing and minimising a non-quadratic cost function," *Chemom. Intell. Lab. Syst.*, vol. 76, no. 2, pp. 121–133, 2005.

[15] M. N. Schmidt, T. S. Alstrom, M. Svendstorp, and J. Larsen, "Peak detection and baseline correction using a convolutional neural network," in *IEEE ICASSP*, 2019, pp. 2757–2761.

[16] Y. Liu, "Adversarial nets for baseline correction in spectra processing," *Chemom. Intell. Lab. Syst.*, vol. 213, pp. 455–500, 2021.

[17] J. Peng, A. J. S. Peng, J. Wei, C. Li, and J. Tan, "Asymmetric least squares for multiple spectra baseline correction," *Analytica Chimica Acta*, vol. 683, no. 1, pp. 63–68, 2021.

[18] Y. Song, "Extensiometry signal denoising," Master's thesis, Université de Technologie de Troyes, France, 2015, (in French).

[19] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[20] T. G. Kolda, "Multilinear operators for higher-order decompositions," Sandia National Laboratories, Albuquerque, NM, Livermore, CA, Tech. Rep. SAND2006-2081, 2006.

[21] D. Coppersmith, A. Odlyzko, and R. Schroeppel, "Discrete logarithms in GF(p)," *Algorithmica*, vol. 1, pp. 1–15, 1985.

[22] A. M. Odlyzko, "Discrete logarithms in finite fields and their cryptographic significance," in *Advances in Cryptology*, T. Beth, N. Cot, and I. Ingemarsson, Eds. Berlin, Heidelberg: Springer-Verlag, 1985, pp. 224–314.

[23] D. H. Wiedemann, "Solving sparse linear equations over finite fields," *IEEE Trans. Information Theory*, vol. 32, pp. 54–62, 1986.