

Automatic Detection and Removal of Spiked Points in Hyperspectral Images [†]

Georgi Manchev, Stanislav Penchev , Tsvetelina Georgieva , Eleonora Kirilova and Plamen Daskalov

Department of Automatics and Mechatronics, University of Ruse, 8 Studentska str., 7017 Ruse, Bulgaria; gmanchev@uni-ruse.bg (G.M.); msp@uni-ruse.bg (S.P.); ekirilova@uni-ruse.bg (E.K.); daskalov@uni-ruse.bg (P.D.)

* Correspondence: cgeorgieva@uni-ruse.bg

[†] Presented at the International Conference on Electronics, Engineering Physics and Earth Science (EEPES'24), Kavala, Greece, 19–21 June 2024.

Abstract: This paper presents an approach to eliminate one of the most common defects in hyperspectral images—the appearance of spiked points at some wavelengths. The elimination of this defect was carried out by means of polynomial regression. The Bayes Information Criterion (BIC) was used to determine the correct order of the polynomial. Comparison between polynomial regression and classical filtration with the Savitsky–Golay method shows the advantage of the proposed approach, from the point of view of eliminating the defect in a local area, without changing the typical behavior of the spectral feature in the affected image pixels.

Keywords: hyperspectral images; preprocessing and filtering; defect pixels; automatic detection

1. Introduction

Hyperspectral imaging (HSI) and its processing techniques are tools gaining popularity in many scientific areas. They find applications in the study of nutritional properties, in phenotypic studies of plants, in microbiological analyses, etc. The main advantage of HSI over other analytical approaches is that it is non-destructive and features speed of data collection.

HSIs combine spectrophotometric and visual information. Structurally, they are a three-dimensional array known as a 3-D hypercube ($n \times p \times q$). The first two dimensions ($n \times p$) form the resolution of a visual image, and the third dimension (q) determines the amount of spectral information in each pixel of that image. It can be said that the spectroscopic part provides the link with the chemical characteristics of the samples, while the image adds some details about the spatial information [1].

Often, in order to extract useful information from HS images, they are subjected to various processing procedures, including regression and classification. The qualitative performance of these procedures, however, strongly depends on the pre-processing of the images, in which various types of noise are removed [2]. The other important problem is the presence of incorrect values in the data (known as dead/bad pixels with zero or maximum signal value) or observations that seem inconsistent with the entire data set. These defects complicate the further tasks of object identification and classification, as well as the quantification of specific parameters. The most common reason for the appearance of such defects is a malfunction of the measuring device. It is known that a large part of measurement systems is built on the basis of lines of diode detectors, in combination with adjustable filters [3]. In this situation, a failure in one of them can lead to the appearance of dead pixels, extreme values in the spectral characteristic [4], or local peaks in a certain wavelength.

Dead pixels in an image are represented as missing or null values, their location and size can vary between a single pixel, a group of pixels, or up to an entire row of pixels [5]. Their presence can lead to inaccuracies in the creation of multivariate models. This is



Citation: Manchev, G.; Penchev, S.; Georgieva, T.; Kirilova, E.; Daskalov, P. Automatic Detection and Removal of Spiked Points in Hyperspectral Images. *Eng. Proc.* **2024**, *70*, 32. <https://doi.org/10.3390/engproc2024070032>

Academic Editor: Grigor Mihaylov

Published: 8 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

why locating and handling such defects is an important prerequisite for the success of further data analysis. Different criteria can be used to locate dead pixels, such as thresholding techniques at mean spectra calculated from the data [6], genetic [7], or evolutionary algorithms [8,9].

Spikes are often expressed as the appearance of a sudden and sharp rise or sharp fall in the spectrum. They are often misidentified as informative features of the signal of interest [10]. Spikes can appear as a result of unusual detector behavior, imperfections in electronic circuits, or due to adverse environmental conditions [11].

One of the most common methods of spike detection is by human-eye observation. However, this method requires human care, it is slow, and its performance is poor at low signal-to-noise ratio (SNR) in the image [12]. Its application in hyperspectral images is an even more difficult task because the hyperspectral cube contains hundreds of spectra and not all of them contain spikes.

Different techniques and algorithms for spike interpolation have been proposed. They are based on the nearest neighbor pixel comparison methodology [10,13] as well as on using wavelets to detect spikes in signals [11,14,15].

Some authors propose an approach based on the calculation of a derivative to remove peaks in a single signal [16]. Their presence causes the signal to change more rapidly than other extrema in the characteristic. The slope of the peaks is usually greater than that of the peaks of interest, and when a derivative is applied to the signal, the large values are associated with the peaks in the original data. Peaks can be separated from other waveforms using a histogram derivative and an appropriate threshold.

This paper proposes an approach to detect and remove spikes in the spectral characteristics of pixels in a hyperspectral image. Defective pixels are detected by comparing the values of the first derivative of the spectral characteristic with a certain threshold. The characteristic is then approximated in a small local region around the spike with polynomials of different orders. The most suitable order of the polynomial is selected using Bayes Information Criterion. A comparison of the performance of the proposed approach with classical filtration with a Savitsky–Golay filter is performed, and the results show promising performance in terms of spike removal while preserving other specific extrema in the spectral characteristic of the defective pixel.

2. Materials and Methods

2.1. Hyperspectral Imaging System

The imaging system consists of a linear translation base driven in one direction by a precision servo motor, four broad spectrum light sources, hyperspectral spectrograph model N17E operating in the wavelength range of 850 nm to 1700 nm coupled with Goldeye CL-008 SWIR camera module (Specim, Spectral Imaging Ltd., Oulu, Finland) with 320 hyperspectral pixels with 256 spectral bands and PC (Figure 1a).

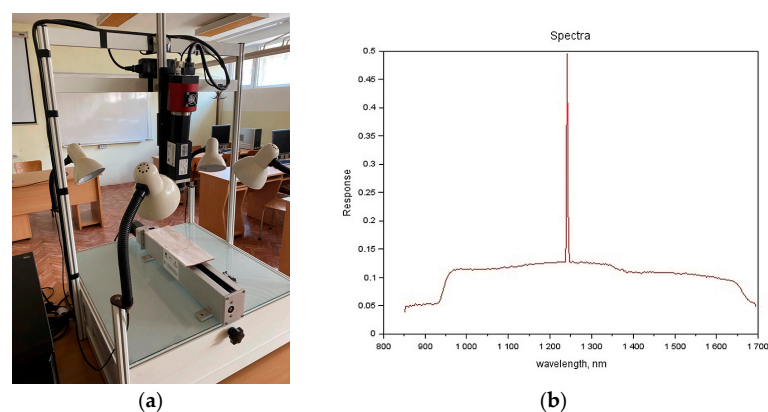


Figure 1. Hyperspectral imaging system (a); An example of a pixel spectrum with spike (b).

2.2. Spiked Point Detection

As mentioned by a number of authors, there is no universal solution for dead pixel detection. In a line scan sensor, the dead pixels appear as a line in the spatial dimension of the HSI at a certain wavelength (Figure 2a). They have an occurrence in each row for the same column index. The spike detection method developed in this paper is based on thresholding in the HSI spatial dimension, applied to the absolute value of the first derivative of the response data for all 320 pixels in a row (Figure 2b).

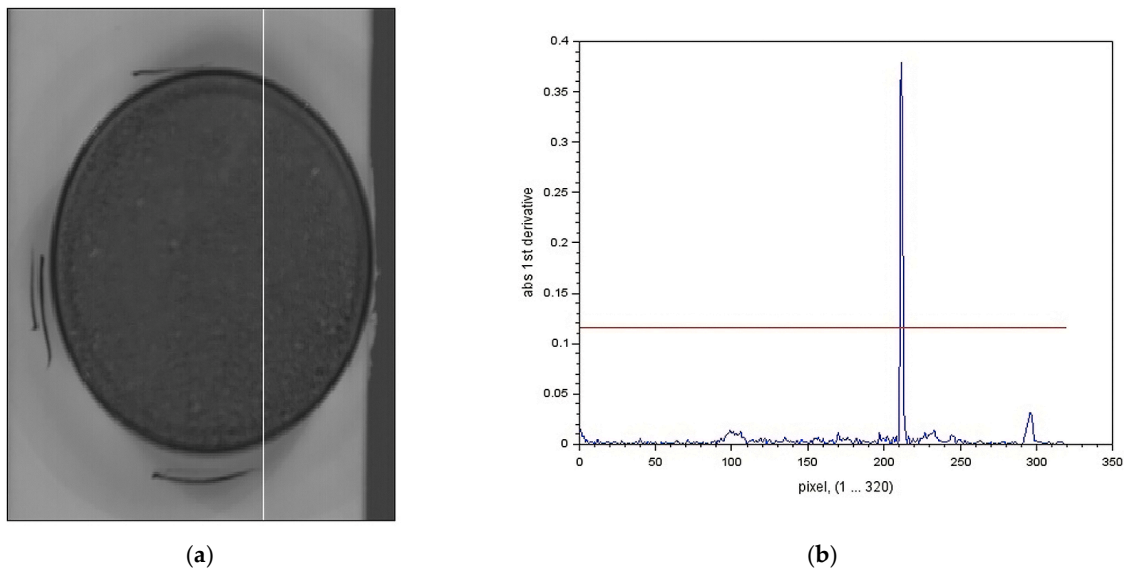


Figure 2. Image with dead pixels (the white vertical line) at 1241 nm (a); The absolute value of the first derivative of a corrupt spectral characteristic and a threshold value (the red line) (b).

The check is performed for each spectral band:

$$\left(\frac{\max(M'_{row})}{\text{mean}(M'_{row})} \right) > Th, \quad (1)$$

where M'_{row} is the absolute value of the first derivative and Th is a certain threshold.

Pixels, where the ratio of the maximum and average absolute value of the first derivative exceeds the preset threshold, are defined as “dead”.

2.3. Polynomial Regression and Bayes Information Criterion

Since spikes appear as an abrupt change in a single spectral band, the recovery of its value is performed by applying a polynomial interpolation in the spectral dimension of the HSI.

Unlike filtering methods, where a change is applied to multiple points (bands), in this case, corrections are made locally only over the defective pixel.

Polynomials with orders up to 5 are calculated, where the data set includes 5 points before and 5 points after the damaged band excluding its value.

A Bayes Information Criterion (BIC) [17] is applied to select the best-fitting model, where the minimum value corresponds to the degree of the model. The pixel recovery value is calculated by interpolation with the model specified by BIC. An example of the use of BIC is presented in Figure 3.

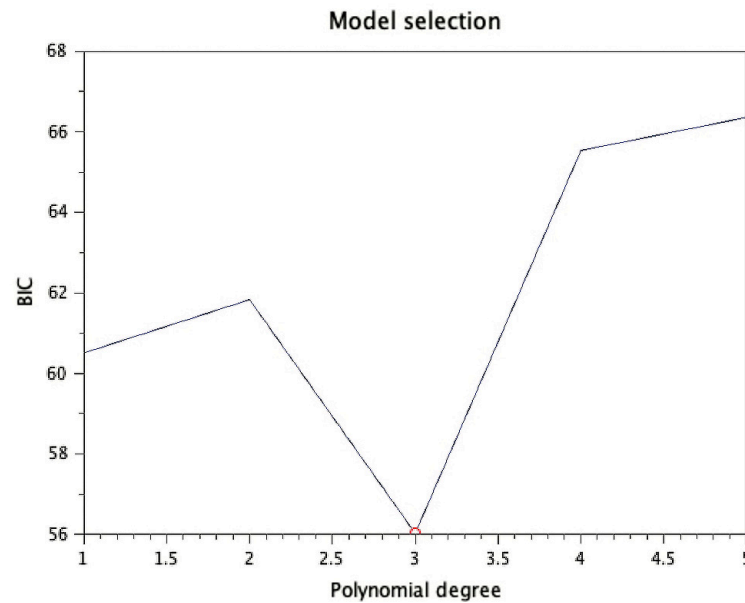


Figure 3. BIC values for fitting polynomials of order 1 to 5. The smallest value of BIC (presented by the red half circle) corresponds to a 3rd-order polynomial.

BIC is a statistical method used for model selection considering both its goodness of fit to the data and their complexity. A lower BIC value indicates better model fitting.

A common form of BIC equation is:

$$BIC = -2 \times \ln(L) + k \times \ln(n), \quad (2)$$

where $\ln(L)$ is the log-likelihood of the model, k is the number of free parameters to be estimated in the model, and n is the number of data points in the sample.

The Log-likelihood of the model represents how well the model fits the data, and its higher value indicates better model fitting. For the case with polynomial regression models, it is based on the sum of square residuals SSR and variance of error σ^2 shown in Equation (3).

$$\ln(L) = -n \times \frac{\log(\sigma^2)}{2} - \frac{SSR}{(\sigma^2)} \quad (3)$$

The number of free parameters for a polynomial model of degree 'd' is shown in (4), and it includes the intercept and the coefficients for each degree of the polynomial.

$$K = (d + 1), \quad (4)$$

3. Results and Discussion

The procedure for the detection and removal of spiked points in a hyperspectral image includes the following main steps. Every pixel spectrum is checked for the existence of spikes. If the spectrum derivative satisfies the condition in Equation (1), then this spectrum should be corrected. The position of the maximum value of the derivative matches the position of the spike in the spectrum.

A polynomial interpolation with equations in order from 1 to 5 is performed in the local neighborhood of the spike, considering its ten neighbors (five before and five after the spike point). The best-fitting curve is chosen by using the minimum value of the Bayes Information Criterion, and the new value of the spike point is calculated from the best-fitting equation.

Figure 4 illustrates the result of the procedure when dealing with the image shown in Figure 2a.

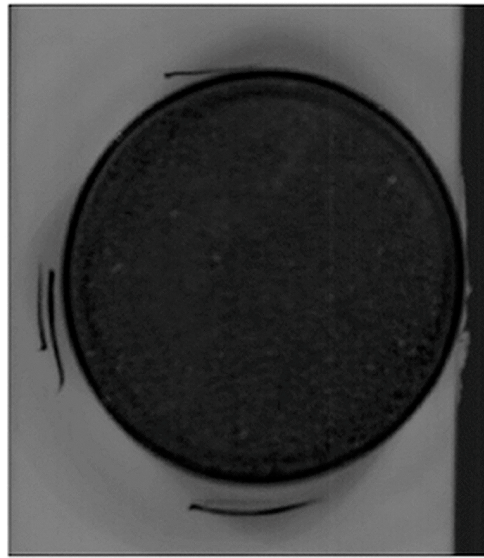


Figure 4. Fixed image (without dead pixels) at 1241 nm.

The performance of the proposed approach is compared with the classical Savitsky–Golay filtering procedure. The results are shown in Figures 5 and 6. Figure 5 represents the whole original spectrum with spike, the polynomial interpolation of the local spike area, and the curve, obtained from Savitsky–Golay filtering. It can be seen that as a result of the Savitsky–Golay filtering, not only is the spike eliminated, but all other small extrema in the spectrum are also smoothed. This is not a good result because some of these extrema could bring essential information about some of the physical or chemical properties of the object in the hyperspectral image. Figure 6 shows the same curves, but only in the local region of the spike. The same result is observed here: the Savitsky–Golay curve is smooth, which means that not only the spike but all other small extrema are also smoothed. On the other hand, the polynomial regression approach makes correction only over the value of the spike. All other neighboring values are not affected.

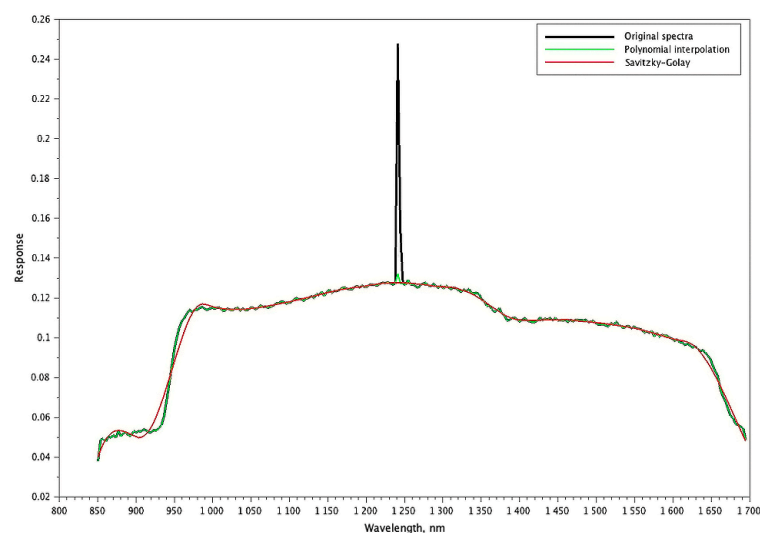


Figure 5. Representation of the Savitsky–Golay filtering and polynomial interpolation.

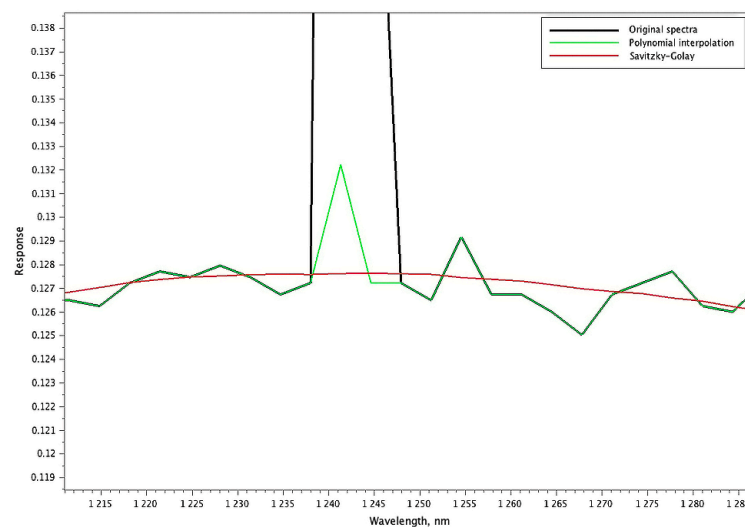


Figure 6. The local spike area with Savitzky–Golay filtering and polynomial interpolation.

In this regard, the approach presented in this paper performs better because it operates only in a small region of the spectrum (where the spike is), thus preserving all other local extrema.

4. Conclusions

An approach for the elimination of spiked points in a hyperspectral image is presented. It is based on a polynomial regression in a local neighborhood around the spike in the affected pixel spectrum.

An automatic procedure for the detection of spikes is developed. It uses the absolute value of the first derivative spectrum as well as an appropriate threshold as a criterion for spike detection. The procedure allows only the part with abnormal behavior to be taken into account, instead of the whole spectrum.

The polynomial regression is performed with equations of a different order. For that reason, the Bayes Information Criterion (BIC) is used to determine the best-fitting model that should be used for spectrum correction.

The performance of the method is compared with the classical Savitzky–Golay filtration. It is concluded that the approach presented here performs better because, unlike the Savitzky–Golay filtration, it operates in only a small region of the spectrum, thus preserving all other local extrema, without changing the typical behavior of the spectral feature in the affected image pixels.

Author Contributions: Conceptualization, S.P. and G.M.; methodology, S.P.; software, G.M.; validation, T.G.; S.P. and E.K.; formal analysis, S.P.; investigation, G.M.; resources, G.M.; data curation, P.D.; writing—original draft preparation, S.P. and E.K.; writing—review and editing, S.P.; visualization, G.M. and E.K.; supervision, P.D.; project administration, T.G.; funding acquisition, T.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created.

Acknowledgments: The report reflects the results of the work on project No. 24-FEEA-05, financed by the “Scientific Research” fund of the University of Ruse, Bulgaria. The research was conducted within the framework of European Union-Next Generation EU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project N^o BG-RRP-2.013–0001-C01.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liu, D.; Zeng, X.-A.; Sun, D.-W. Recent Developments and Applications of Hyperspectral Imaging for Quality Evaluation of Agricultural Products: A Review. *Crit. Rev. Food Sci. Nutr.* **2015**, *55*, 1744–1757. [\[CrossRef\]](#)
2. Xiao, J.; Ji, Y.; Wei, X. Hyperspectral Image Denoising with Spectrum Alignment. In Proceedings of the MM 2023—31st ACM International Conference on Multimedia, Ottawa, ON, Canada, 29 October–3 November 2023; pp. 5495–5503. [\[CrossRef\]](#)
3. Amigo, J.M. Practical issues of hyperspectral imaging analysis of solid dosage forms. *Anal. Bioanal. Chem.* **2010**, *398*, 93–109. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Prats-Montalbán, J.M.; de Juan, A.; Ferrer, A. Multivariate image analysis: A review with applications. *Chem. Int. Lab. Syst.* **2011**, *107*, 1–23. [\[CrossRef\]](#)
5. Maji, S.K.; Aetesam, H. Lp-norm-based successive denoising approach for hyperspectral images. *Rem. Sens. Lett.* **2023**, *14*, 334–345. [\[CrossRef\]](#)
6. Burger, J.; Geladi, P. Hyperspectral NIR image regression part 1: Calibration and correction. *J. Chem.* **2005**, *19*, 355–363. [\[CrossRef\]](#)
7. Karakaplan, M.; Avcu, F.M. Deconvolution of Gaussian peaks with mixed real and discrete-integer optimization based on evolutionary computing. *J. Chem.* **2020**, *34*, e3229. [\[CrossRef\]](#)
8. Walczak, B. Outlier detection in multivariate calibration. *Chem. Int. Lab. Syst.* **1995**, *28*, 259–272. [\[CrossRef\]](#)
9. Amigo, J.M.; Santos, C. Preprocessing of hyperspectral and multispectral images. *Data Handl. Sci. Technol.* **2020**, *32*, 37–53. [\[CrossRef\]](#)
10. Cozzolino, D.; Williams, P.J.; Hoffman, L.C. An overview of pre-processing methods available for hyperspectral imaging applications. *Microchem. J.* **2023**, *193*, 109129. [\[CrossRef\]](#)
11. Sharan, T.S.; Sharma, S.; Sharma, N. Denoising and Spike Removal from Raman Spectra using Double Density Dual-Tree Complex Wavelet Transform. *J. Appl. Spectrosc.* **2021**, *88*, 117–124. [\[CrossRef\]](#)
12. Nenadic, Z.; Burdick, J.W. Spike detection using the continuous wavelet transform. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 74–87. [\[CrossRef\]](#)
13. Behrend, C.J.; Tarnowski, C.P.; Morris, M.D. Identification of outliers in hyperspectral Raman image data by nearest neighbor comparison. *J. Appl. Spectrosc.* **2002**, *56*, 1458–1461. [\[CrossRef\]](#)
14. Koshino, K.; Zuo, H.Y.; Saito, N.; Suzuki, S. Improved spike noise removal in the scanning laser microscopic image of diamond abrasive grain using wavelet transforms. *Opt. Commun.* **2004**, *239*, 67–78. [\[CrossRef\]](#)
15. Du, P.; Kibbe, W.A.; Lin, S.M. Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics* **2006**, *22*, 2059–2065. [\[CrossRef\]](#)
16. Feuerstein, D.; Parker, K.H.; Boutelle, M.G. Practical methods for noise removal: Applications to spikes, nonstationary quasi-periodic noise, and baseline drift. *Anal. Chem.* **2009**, *81*, 4987–4994. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Schwarz, G.E. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.