# Comparision of methods for the location and correction of bad pixels present in hyperspectral images

**R.M. Dorrepaal, A.A. Gowen**

*UCD School of Biosystems and Food Engineering, University College Dublin, Belfield, Dubin 4, Ireland*
*ronan.dorrepaal@ucdconnect.ie*

Keywords: Chemometrics, Biomaterials, Bad Pixel Identification, Spike Correction, Signal Processing.

Hyperspectral imaging provides many advantages over traditional point spectroscopy as it allows for large numbers of spectra to be acquired rapidly. Further it introduces a spatial consideration to the acquired spectral data, leading to a greater insight of the materials studied. For this reason, hyperspectral imaging is used ever more by a growing community which spans many academic and industrial fields.

However, the acquisition of large numbers of spectra can yield issues relative to traditional point techniques. Anomalies (known as spikes) in recorded spectra can occur with any spectroscopy technique. The spikes can occur for a multitude of reasons, including malfunction of a detector (especially after some time into the detector's lifetime), cosmic rays or random electrical surges from the detector power source. While these anomalies can occur in the case of traditional point spectra, such spectra are more likely to get the full attention of the experimenter and the anomaly is therefore more likely to be recognised. As hyperspectral images often comprise thousands or even millions of spectra, it is not possible for the experimenter to visually observe all spectra and these anomilies can therefore go unnoticed.

Techniques to find spectral spikes exist (e.g. PCA) [1], as do methods to correct for them (e.g. Median Wavelength Ratio) [1].

The aim of this study is to compare the performance of existing spike location and spike correction methods with novel methods using a ground truth standard. The analysis is performed on NIR hyperspectral images of magnesium oxychloride biomaterials. The hyperspectral data set used is of particular interest as it contains a sharp peak in the NIR range which could be mistaken for a bad pixel and lead to false positive spike detections.

## Instrumentation and Software

Wavelength range of 880–1720nm with spectroscopic resolution of 7 nm and spatial resolution of approximately 700µm (DV Optics Ltd).

Matlab 2016a (Windows 10, 64 bit), with Image Processing Toolbox, Signal Processing Toolbox and Statistics and Machine Learning Toolbox, was used.

Intel Core i7-5500U CPU at 2.4GHz and 8GB RAM and Solid State Disk (SSD).

## Methods

Samples were prepared by dissolving $MgCl_2.6H_2O$ (4M) in deionised water. MgO (1g:1mL deionised water) was added to the solution and was stirred to form a paste. The paste was placed on a glass slide and was allowed to set. After two weeks, NIR imaging was conducted on the samples.

## Bad Pixel Location Algorithms

*1. Ground Truth*

The samples were imaged four times. For each replicate image, the sample was placed in different spatial positions and at different rotational orientations relative to the detector. The pixels were co-registered and overlayed with an affine transformation. This way, when bad pixels were detected, it could be determined whether spectral spikes were due to signal or noise.

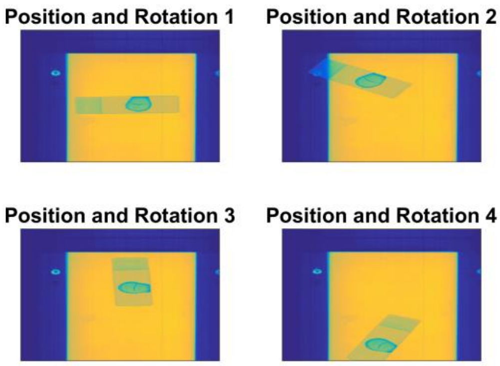


Figure 1. Images of magnesium oxychloride cement sample at different positions and orientations

The following algorithms were compared with this ground truth.

*2. Difference Spectrum*

For each measured spectrum, a difference spectrum was prepared from the subtraction of each measured wavelength from the following measured wavelength. The (a) standard deviation or the (b) maximum value of the difference matrix was then used and thresholded to find bad pixels comprising spikes.

*3. Area Window Mean*

The mean spectrum was found for all possible moving window (3x3 box) positions of image. Each window central pixel spectrum was then compared to the window mean spectrum to look for bad pixels.

*4. Smoothed Spectrum Difference*

A moving spectral window was averaged to produce smoothed spectra. The difference between the original and smoothed spectra was calculated. The maximum difference was thresholded to detect bad pixels comprising spectra with spikes.

*5. Neighbouring Large Difference*

For each measured spectrum, a difference spectrum was prepared from the subtraction of each measured wavelength from the following measured wavelength. A large positive difference value next to a neighbouring large negative difference value was used to determine bad pixels with spike spectra

*6. PCA*

Principle Component Analysis (PCA) [2] can be used as an unsupervised dimensional reduction method. Through PCA, sample outliers can often be distinguished readily.

## Acknowledgement

The authors acknowledge funding from the EU FP7 under the European Research Council Starting Grant programme.

## References

[1] "Replacement of hyperspectral image bad pixels", Burger, J., NIR News, doi: http://dx.doi.org/10.1255/nirn.1151.

[2] "Tutorial: Time series hyperspectral image analysis", Dorrepaal, R., Malegori, C. and Gowen, A., Journal of Near Infrared Spectroscopy, doi: http://dx.doi.org/10.1255/jnirs.1208