Exploiting noisy hyperspectral bands for water analysis

Daniele Cerra, Peter Gege, Rupert Müller, and Peter Reinartz

German Aerospace Center (DLR), Earth Observation Center, Remote Sensing Technology Institute, Oberpfaffenhofen, 82234 Wessling, Germany; Daniele.cerra@dlr.de, Rupert.mueller@dlr.de, Peter.Reinartz@dlr.de

Abstract. This paper proposes a novel algorithm for the recovery of noisy bands from hyperspectral images. The method, based on spectral unmixing, relies on the spectral behavior of the materials on ground composing each image element. Firstly, reference spectra related to the classes of interest are used to perform spectral unmixing: these exhibit negligible noise influences as they are averaged over areas for which ground truth is available. After the unmixing process, the residual vector is mostly composed by the contributions of uninteresting materials, unwanted atmospheric influences and sensor-induced noise, and is thus ignored in the reconstruction of each spectrum. Finally, the value of a pixel in a given band is predicted as a combination of the noise-free endmembers, resulting in a signal with high signal-to-noise ratio in any spectral band. Experiments show that this method could be used to retrieve spectral information from corrupted bands, such as the ones placed at the edge between Ultraviolet and visible light frequencies, which are usually dominated by atmospheric effects and are thus discarded in practical applications. The proposed algorithm could then be exploited in the study of Coloured dissolved organic matter (CDOM) in natural waters.

Keywords. Spectral unmixing, denoising, image restoration, hyperspectral images, water analysis, CDOM.

1. Introduction

The spectral range characterizing data acquired by state-of-the-art hyperspectral sensors mostly spans the frequencies between 400 nm and 2500 nm. Some sensors, such as the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) acquire data also in the portion of the spectrum which is placed at the edge between Near Ultraviolet (NUV) frequencies and visible light (380-400 nm). Such bands are typically characterized by a low Signal to Noise Ratio (SNR), and as a consequence are usually discarded in a preprocessing step common to most practical applications. For some specific tasks, it would be desirable to keep such spectral information: a typical example is the study of Coloured dissolved organic matter (CDOM) in natural waters. CDOM inhibits phytoplankton productivity by absorbing UV and NUV radiation, affecting in turn remote estimates of clorophyll concentration [1]. As the bands in the NUV-blue portion of the spectrum are usually noisy and therefore difficult to interpret, bands at longer wavelengths characterized by a better SNR are usually preferred to derive empirical indices for its estimation. As an example, in [2] the authors simulate the reflectance at 380 nm using the bands in the 400-500 nm portion of the spectrum. Therefore, a specific denoising methodology enabling a direct use of the information in the mentioned spectral range would represent a great aid for such applications.

Denoising is often carried out in image processing through filtering, usually based on convolutions with sliding windows in the image domain, on operations in the frequency domain, or on estimated noise statistics or degradation functions, if these are known for the image acquisition process [3]. In the case of hyperspectral images, the high dimensionality of the data and the correlation between adjacent bands can be exploited to carry out effective denoising procedures based on dimensionality reduction (DR) algorithms, which project the data onto a subspace where meaningful information is preserved, while noise and some high frequencies are discarded [4].

This paper proposes to use Unmixing-based Denoising (UBD) [5], a methodology resulting from considering any hyperspectral image element as a mixture of a number of materials smaller than the original data dimensionality, which can be effectively described relying on spectral unmixing methods. In the linear mixture model, each spectrum is approximated by a linear combination of the endmembers plus a residual vector. The latter is the sum of the contributions of noise, subtle variations within the reference materials, and errors in the adopted unmixing model [6]. If the set of endmembers is complete and well represents the scene at hand, and if the reference spectra exhibit negligible noise influences, the residual vector will mostly result from atmospheric interferences and instrument-induced noise, and can be ignored in the reconstruction of each single spectrum. Thus, UBD reconstructs each pixel in a given band of a hyperspectral image as a linear combination of the values of the reference spectra in that particular band. As such values would not be reliable for a band with low Signal-to-Noise Ratio, the reference spectra are averaged over an area of interest to greatly reduce the noise influence in the input to the reconstruction. First results on CDOM estimation in open waters are encouraging and suggest that the proposed method could greatly improve the outcomes of such applications.

2. Unmixing-based Denoising

Given a hyperspectral image element *m* with *p* bands, and a training dataset containing *n* samples from each of *k* classes, with $k < p$, the Unmixing-based Denoising (UBD) is a simple procedure which can be described as follows. Firstly, a set of reference spectral signatures is defined as $A = \{x_1 \ldots x_i \ldots x_n\}$, where x_i is the average of the *n* spectra belonging to class *i*. Considering the mean value for a given reference spectrum reduces the presence of noise to a minimum, if each class is spectrally homogeneous. It must be remarked that no assumption on the purity of the reference spectra is made. Then, any unmixing procedure can be employed to decompose the signal in a combination of the reference spectra. If we assume this to be linear, we have:

$$m = \sum_{i=1}^{k} x_i s_i + r \tag{1}$$

where s_i is the fraction or abundance of the reference spectrum *i* in *m*, and *r* the residual vector. The latter is mostly composed by the contributions related to materials not present in *A*, subtle variations of one or more materials in *A*, atmospheric interferences, and instrument-induced noise. If the spectra in *A* are noise-free and represent well the classes of interest, we expect the last two terms to be predominant in the residual vector for bands with low SNR, and we can derive a reconstruction $\hat{m}$ for a spectrum *m* as:

$$\hat{m} = \sum_{i=1}^{k} x_i s_i \tag{2}$$

ignoring *r*, and along with it most of the noise affecting *m*. The described procedure is based on the assumption that if the contributions to the radiation reflected from a resolution cell are known, the value of noisy bands in that area can be derived by a combination of the average values characterizing each component in that spectral range. The proposed method is supervised, as it needs as input a set of spectra that well characterize the scene, and is carried out independently for each pixel.

As a certain homogeneity of the classes of interest is assumed, the method is expected to perform better on natural scenes where man-made objects (usually having a higher variability) are not prevalent.

3. Results

In the following experiments we choose inversion through Non-negative Least Squares (NNLS) as unmixing algorithm [6]. It is of interest to remark that NNLS naturally enforces sparsity in its solution, as several components in it are set to zero: this intuitively well agrees with the characteristics of a hyperspectral pixel, which is usually composed by a limited number of materials [7]. Unconstrained Least Squares and first attempts at using sparse reconstruction tools did not yield satisfactory results, while in recent years the fully-constrained least squares method, which enforces not only non-negativity but also the sum-to-one constraint on the estimated abundances, has been debated by the community and is therefore not considered in these experiments [6].

We apply the proposed method to a hyperspectral scene acquired by the HySpex sensor over the lake Starnberg in Germany. The Signal-to-noise Ratio for these data in the spectral range up to 450 nm is low, making difficult to extract reliable information from the related spectral bands, especially for water areas. We select 25 spectra from averaged homogeneous areas in the scene and use this set of reference spectra as input for UBD. Results for the denoising of a sample band are reported in Fig. 1, while Fig. 2 shows the difference image between the original and the reconstructed band.

(a) (b)

Figure 1. band 6 from the Starnberger lake dataset (434 nm), before and after denoising.

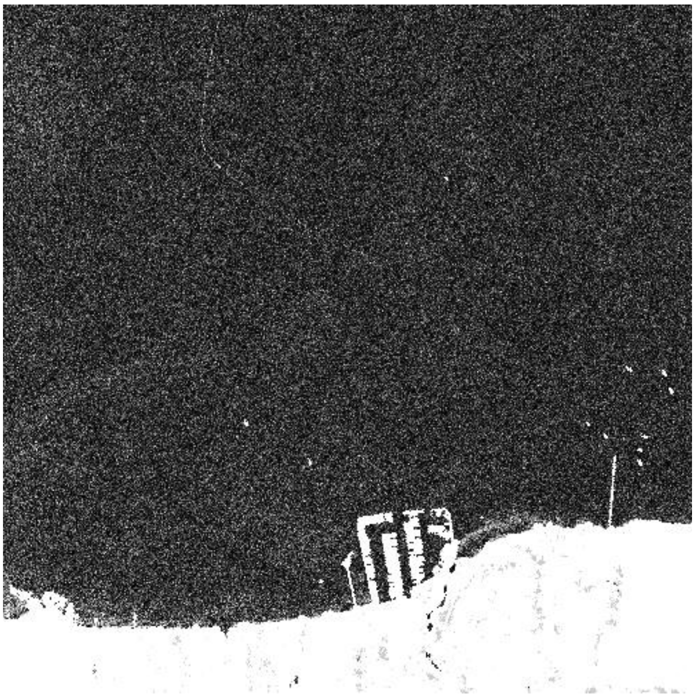

Figure 2. Absolute difference between the two images in Fig. 1. It is hard to spot features in the image, suggesting that noise is removed while meaningful information is preserved.

To determine if the UBD process can help improving the results of practical applications we estimate the CDOM content from the described image. We adopt [8] as bio-optical model for shallow waters and its derivation through inverse modeling, and use its software implementation in WASI (WAter color SImulation) [9]. The concentration of CDOM is computed for the original image with the first 10 bands removed (420-450 nm) and again for the same dataset with the first 10 bands denoised through UBD included. Results for estimated CDOM absorption are reported in Fig. 3. On the left of the UBD-processed image artificial image patterns (triangular structures) are evident. These are caused by the large errors of atmosphere correction in the denoised bands below 450 nm: if the inversion problem has no unique solution, the fit result depends on the choice of the initial values, and the measurements do not fit well the model. The triangles are caused by the algorithm to initialize the fit parameters using previously processed pixels. These problems can be solved by recalibrating the original data. Apart from these inversion problems caused by erroneous spectra, the denoised image reveals finer structures of CDOM concentration and less noise.

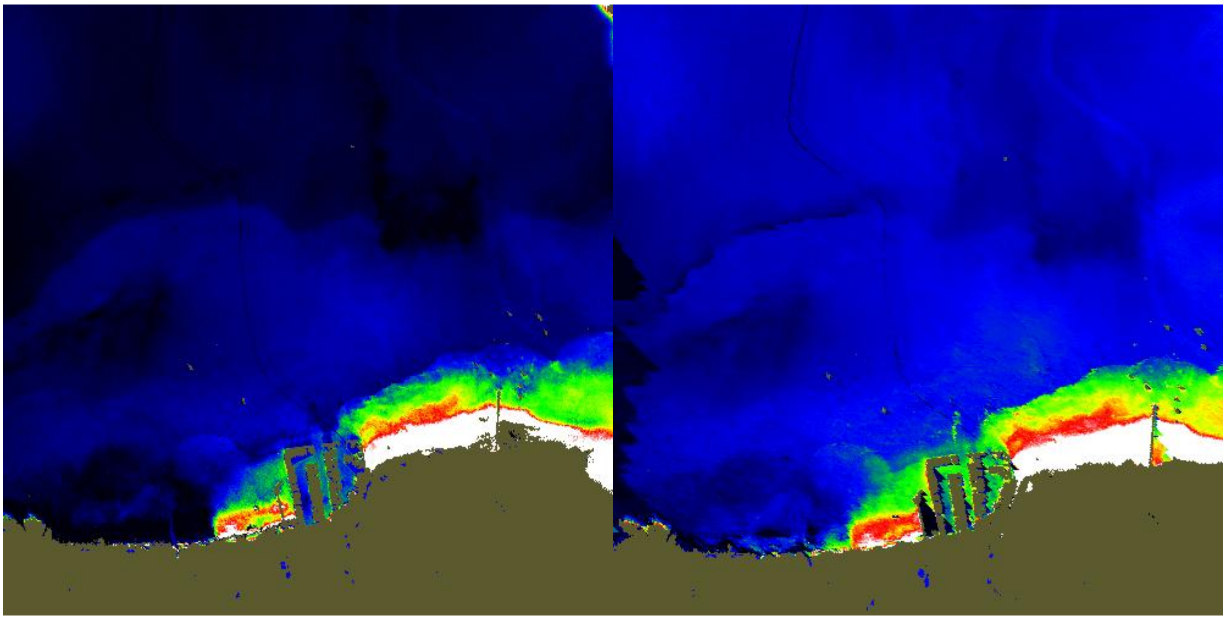

Figure 3. CDOM concentration estimation before and after denoising of noisy bands in the blue portion of the spectrum with UBD (Scale: 0 - 1 mg/l).

It is also interesting to consider that the number of iterations in the inversion problem has drastically decreased after denoising. Fig. 4 shows the pixels for which more than 200 iterations were necessary. These are much less after the application of UBD. Also the number of pixels for which no solution was found is much lower in the denoised image.

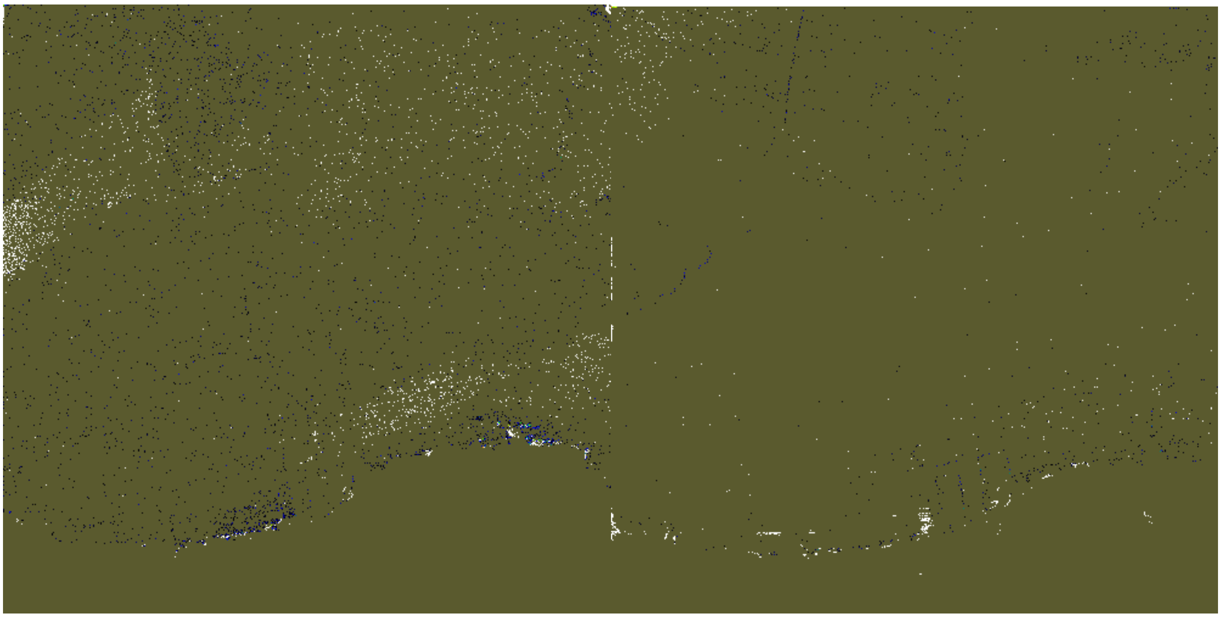

Figure 4. Number of iterations for the solution of the inversion problem before and after denoising. Only pixels are shown for which more than 200 iterations were necessary. The white pixels indicate that no convenient fit was found, and inversion was broken up artificially after 1000 iterations. The number of such cases has decreased significantly for the denoised image.

4. Conclusions

Unmixing-based Denoising (UBD) is a supervised methodology for the recovery of bands characterized by a low Signal-to-Noise Ratio (SNR) in a hyperspectral scene. UBD reconstructs any pixel in a given band as a linear combination of reference spectra belonging to materials present in the scene, which have negligible noise influences as they are averaged over areas for which ground truth is available. As the residual vector from the unmixing process is mostly composed by contributions of uninteresting materials, unwanted atmospheric influences and sensor-induced noise, this is ignored in the reconstruction process. Experiments on real data suggest that this method could be used to retrieve bands which are usually discarded in practical applications: this would help in exploiting important and often overlooked information within hyperspectral scenes, resulting in a better characterization for areas in which the information of such bands is important. An example is represented by applications to natural waters, where the coloured Dissolved Organic Matter (CDOM) could better estimated by employing directly the spectral information at the edge between visible and Near Ultraviolet frequencies, which is typically noisy [1]. A drawback of the method is that it requires as input a reasonably large amount of pixels to derive each reference spectrum, in order to have a meaningful mean value robust to noise and local variations. Future work includes a fully automatic workflow to iteratively select the best reference spectra for the denoising of specific spectral ranges.

References

[1] Kutser, T., Tranvik, L., and Pierson, D. C., “Variations in colored dissolved organic matter between boreal lakes studied by satellite remote sensing,” Journal of Applied Remote Sensing, vol. 3, no. 1, pp. 33 538–33 538–10, 2009.

[2] Kutser, T., Pierson, D. C., Kallio, K. Y., Reinart, A., and Sobek, S., “Mapping lake cdom by satellite remote sensing,” Remote Sensing of Environment, vol. 94, no. 4, pp. 535 – 540, 2005.

[3] Gonzalez, R. C.and Woods, R. E., Digital Image Processing (3rd Edition), 3rd ed. Prentice Hall, Aug. 2007.

[4] Renard, N., Bourennane, S., and Blanc-Talon, J., “Denoising and dimensionality reduction using multilinear tools for hyperspectral images,” IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 2, pp. 138–142, april 2008.

[5] Cerra, D., Müller, R., and Reinartz, P., Noise Reduction in Hyperspectral Images through Spectral Unmixing”, IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 1, January 2014.

[6] Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot J., “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, no. 2, pp. 354–379, 2012.

[7] Bieniarz, J., Müller, R., Zhu, X., and Reinartz, P., “Sparse approximation, coherence and use of derivatives in hyperspectral unmixing,” in 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing WHISPERS 2012, June 2012.

[8] Albert, A., Mobley, C.D., 2003. An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters. Optics Express 11, 2873-2890.

[9] Gege, P., Albert, A., “A tool for inverse modeling of spectral measurements in deep and shallow waters”, L.L. Richardson and E.F. LeDrew Eds, Kluwer book series: Remote Sensing and Digital Image Processing, pp. 81-109, 2006.