INTER-BAND DISTORTION ALLOCATION IN LOSSY HYPERSPECTRAL DATA COMPRESSION

Leonardo Santurri(1) **Bruno Aiazzi**(1)**, Luciano Alparone**(2)**, Stefano Baronti**(1)**, Cinzia Lastri**(1)

(1)***IFAC–CNR: Institute of Applied Physics "Nello Carrara", Area della Ricerca di Firenze***
10 Via Madonna del Piano, 50019 Sesto F.no (Italy), {B.Aiazzi S.Baronti}@ifac.cnr.it
(2)***DET-UniFI: Department of Electronics and Telecommunications, University of Florence***
3 Via Santa Marta, 50139 Firenze (Italy), Alparone@lci.det.unifi.it

ABSTRACT

The problem of distortion allocation varying with wavelength in lossy compression of hyperspectral imagery is investigated. Distortion is generally measured either as maximum absolute deviation (MAD) for near-lossless methods, e.g. differential pulse code modulation (DPCM), or as mean square error (MSE) for lossy methods (e.g. spectral decorrelation followed by JPEG 2000). Also the absolute angular error, or spectral angle mapper (SAM), is used to quantify spectral distortion. A band add-on (BAO) technique was recently introduced to calculate a modified version of SAM. Spectral bands are iteratively selected in order to increase the angular separation between two pixel spectra by exploiting a mathematical decomposition of SAM. As a consequence, only a subset of the original hyperspectral bands contributes to the new distance metrics, referred to as BAO-SAM, whose operational definition guarantees its monotonicity as the number of bands increases. Two strategies of inter-band distortion allocation are compared: given a target average bit rate, distortion, either MAD or MSE, may be set to be constant varying with wavelength. Otherwise it may be allocated proportionally to the noise level on each band, according to the virtually-lossless protocol. Thus, a different quantization step size depending on the estimated standard deviation of the noise, is used to quantize either prediction residuals (DPCM) or wavelet coefficients (JPEG 2000) of each spectral band, thereby determining band-varying MAD/MSE values. Comparisons with the uncompressed originals show that the average spectral angle mapper (SAM) is minimized by constant distortion allocation. Conversely, the average BAO-SAM is minimized by the noise-adjusted variable spectral distortion allocation according to the virtually lossless protocol.

INTRODUCTION

Technological advances in imaging spectrometry have lead to acquisition of data that exhibit extremely high spatial, spectral, and radiometric resolution. In particular, the increment in spectral resolution has motivated the extension of vector signal/image processing techniques to hyperspectral data, for both data analysis and compression [1]. As a matter of fact, a challenge of satellite hyperspectral imaging is data compression for dissemination to users and especially for transmission to ground station from the orbiting platform. Data compression often performs a decorrelation of the correlated information source, before entropy coding is carried out. To meet the quality issues of hyperspectral image analysis, differential pulse code modulation (DPCM) is usually employed for lossless/near-lossless compression, i.e. the decompressed data have a user-defined maximum absolute error, being zero in the lossless case. DPCM basically consists of a prediction followed by entropy coding of quantized differences between original and predicted values. A unit quantization step size allows reversible compression as a limit case. Several variants exist in DPCM prediction schemes, the most sophisticated being adaptive [2, 3, 4, 5, 6].

When hyperspectral remote sensing acquisition systems, with the instrument on-board a satellite platform, data compression is crucial. If strictly lossless techniques are not employed, a certain amount of information of the data will be lost. In the literature, there are several established distortion measurements, some of which are usually employed also for quality assessment of decompressed data. The problem is that they measure the distortion introduced in the data, but cannot measure the consequence of such distortion; in other words, how the information loss would affect the outcome of an analysis performed on such data. For example, the quantitative evaluation of compression algorithms, and more generally of the results of all methods that process hyperspectral data, is often based on distance metrics that compare two pixel spectra and return a scalar value [7]. The Euclidean minimum distance (EMD) and the spectral angle mapper (SAM) are distance metrics usually adopted for this task and possess distinct mathematical and physical properties; but SAM suffers of some limitation: it is not monotonic and tends to exhibit an asymptotic constant value as the number of bands increases [8]. In a recent paper [9] the characteristics of both metrics are examined and a band add-on (BAO) technique is derived that iteratively selects bands in order to increase the angular separation between two spectra by exploiting a mathematical decomposition of SAM. As a consequence of BAO, only a subset of the original hyperspectral bands contributes to the new distance metric, referred to as BAO-SAM hereafter, whose operational definition guarantees its monotonicity. Thus, BAO-SAM is potentially more useful than SAM for classification, as well as for the evaluation of distortions between pixel spectra, since SAM tends to exhibit an asymptotic constant value when the number of bands increases [8].

Such distortion metrics as Mean Square Error (MSE), Maximum Absolute Deviation (MAD), i.e. peak error, average and maximum SAM, and spectral information divergence (SID) [10], are usually adopted to verify the efficiency of compression algorithms. As the number of bands increases, SAM partially loses discrimination capability, as it tends to saturate. The BAO approach could be adopted to overcome this problem with two main objectives. Providing a deeper insight in the spectral distortion affecting hyperspectral data. Assessing the feasibility of BAO-SAM bounded compression algorithms. In fact, a data compression method of general validity should not be specialized to a specific application, e.g. spectral anomaly detection [11], land-cover classification, or spectral unmixing [12] and material detection [13]. That is, a compression method should not be optimised and assessed on a single application. Rather, one should try to optimise a hyperspectral data compression method in terms of a distortion metrics, e.g. MAD, as it happens with "near-lossless" compression methods, but also SAM. Then one should demonstrate how the detection/classification/unmixing accuracy depends on the distortion metrics, for which the compression algorithm is optimised.

A distortion measurement derived from the BAO approach and called BAO-SAM has been applied for quality evaluation of compressed hyperspectral data. Compression with different MAD values and comparisons with the undistorted originals show that BAO-SAM is useful for characterizing the spectral distortion of compressed hyperspectral data. Although a compression algorithm capable of producing reconstructed vectors with upper-bounded user-defined BAO-SAM is unfeasible, it is possible to *indirectly* reduce the BAO-SAM by allocating more distortion to those bands that are less relevant, on an average, for its computation. This is implicitly achieved by adopting a virtually-lossless compression strategy.

BAO-BOUNDED COMPRESSION

In a recent paper [9], the characteristics of both EMD and SAM are examined and a band add-on (BAO) technique is derived that iteratively selects bands in order to increase the angular separation between two spectra by exploiting a mathematical decomposition of SAM. As a consequence of BAO, only a subset of the original hyperspectral bands contributes to the new distance metric, referred to as BAO-SAM hereafter, whose operational definition guarantees its monotonicity.

BAO-SAM is potentially more useful than SAM for classification, as well as for the evaluation of distortions between pixel spectra, and an interesting characteristic is that BAO-SAM can be used both for quality assessment of compressed data and for operatively discriminating materials.

In particular we can see that compression algorithm runs on board, where data are available in radiance units. After being transmitted on ground, they are converted into reflectance units, by means of the subsequent formula:

$$\rho(\lambda) = \frac{R(\lambda) \cdot \pi}{I(\lambda) \cdot T(\lambda)} \tag{1}$$

where $\rho(\lambda)$ is the reflectance of the pixel being acquired, $I(\lambda)$ is the solar irradiance on ground, $T(\lambda)$ is the atmospheric transmittance, and finally $R(\lambda)$ is at-sensor radiance: all these quantities are function of the wavelength λ. It is clear that a distortion introduced on radiance data would be amplified or attenuated depending on the values that are assumed by the product $I(\lambda) \cdot T(\lambda)$.

The initial idea was of developing a BAO-bounded compression algorithm, but this approach is not practicable. What is possible is to allocate distortion among bands, in such a way that the BAO-SAM originated from compression is as small as possible, so that the BAO-SAM deriving from spectral differences is mostly preserved, even when the data are converted to reflectance units.

Band Add-On Spectral Angle Mapper

SAM measures the angle $\theta(\mathbf{x},\mathbf{y})$, where **x** and **y** are N-dimensional spectral vectors having real-valued nonnegative components, $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$, respectively,

$$\theta(\mathbf{x},\mathbf{y}) \;=\; \arccos\left(\frac{\langle \mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}\right), \quad 0 \leq \theta \leq \frac{\pi}{2}. \tag{2}$$

in which $\langle \mathbf{x},\mathbf{y}\rangle$ is the scalar product between **x** and **y**

$$\langle \mathbf{x},\mathbf{y}\rangle \triangleq \sum_{i=1}^{N} x_i \cdot y_i \tag{3}$$

and $\|\cdot\|$ represents the Euclidean norm, i.e. $\|\mathbf{x}\|^2 = \langle \mathbf{x},\mathbf{x}\rangle$.

SAM is a *nonadditive* distance metrics [14], which means that, if the components of **x** and **y** are partitioned as $\mathbf{x} = [\mathbf{x}_a, \mathbf{x}_b]$ and $\mathbf{y} = [\mathbf{y}_a, \mathbf{y}_b]$, where $\mathbf{x}_a, \mathbf{y}_a \in \mathbb{R}^{+a}$ and $\mathbf{x}_b, \mathbf{y}_b \in \mathbb{R}^{+b}$, with $a + b = N$, then $\theta(\mathbf{x},\mathbf{y}) \neq \theta(\mathbf{x}_a,\mathbf{y}_a) + \theta(\mathbf{x}_b,\mathbf{y}_b)$.

According with [9], we can rewrite (2) in terms of the angle between $\mathbf{x}_a$ and $\mathbf{y}_a$, namely θ_a,

$$\theta_a(\mathbf{x},\mathbf{y}) \triangleq \theta(\mathbf{x}_a,\mathbf{y}_a) \;=\; \arccos\left(\frac{\langle \mathbf{x}_a,\mathbf{y}_a\rangle}{\|\mathbf{x}_a\|\|\mathbf{y}_a\|}\right) \tag{4}$$

in such a way that

$$\begin{aligned}\cos(\theta) \;&=\; \frac{\langle \mathbf{x}_a,\mathbf{y}_a\rangle + \langle \mathbf{x}_b,\mathbf{y}_b\rangle}{\sqrt{\|\mathbf{x}_a\|^2 + \|\mathbf{x}_b\|^2}\sqrt{\|\mathbf{y}_a\|^2 + \|\mathbf{y}_b\|^2}} \\ &=\; \cos(\theta_a) \cdot \frac{1 + \dfrac{\langle \mathbf{x}_b,\mathbf{y}_b\rangle}{\langle \mathbf{x}_a,\mathbf{y}_a\rangle}}{\sqrt{1 + \dfrac{\|\mathbf{x}_b\|^2}{\|\mathbf{x}_a\|^2}}\sqrt{1 + \dfrac{\|\mathbf{y}_b\|^2}{\|\mathbf{y}_a\|^2}}}.\end{aligned} \tag{5}$$

The mathematical decomposition of SAM reported in (5) has been used to derive a novel angular distance between two vectors that has the desirable property of being monotonic as the number of vector components increases, unlike SAM [15]. To this purpose, an algorithm, denoted as band add-on (BAO), has been developed in [9] and consists of starting with $a = 2$ and iteratively adding one vector component at a time in such a way that the overall spanned angle is maximized. The couple of startup components are usually chosen to maximize $\cos(\theta_2)$, among all possible $N(N-1)/2$ combinations. This choice simplifies the following procedure without significantly degrading the optimality of the results [16]. Starting from (5), let us note that the second factor of the product on right side, which will be referred to as β in the following, can be either ≥ 1 or < 1. Thus, if $\beta \geq 1$, then $\theta \leq \theta_a$; if $\beta < 1$, then $\theta > \theta_a$. Upon these premises, the BAO algorithm works as follows:

1. Select a couple of starting components, e.g. such that the spanned "subangle"is maximized. Set $a = 2$.

2. Calculate β for each of the remaining $N - a$ components.

3. Find the component yielding the lowest β, provided that $\beta < 1$, and add it to the set of the selected components. If no component yields $\beta < 1$, end; otherwise set $a = a + 1$ and go to step 2).

In this way, given the two spectra **x** and **y**, one finds the subset of components, whose number is a, such that $\theta(\mathbf{x}_a,\mathbf{y}_a)$ is maximum.

Our interest in exploiting the measure given by BAO-SAM was raised from the property of BAO-SAM of being monotonic, while SAM usually decreases, as the number of bands increases.

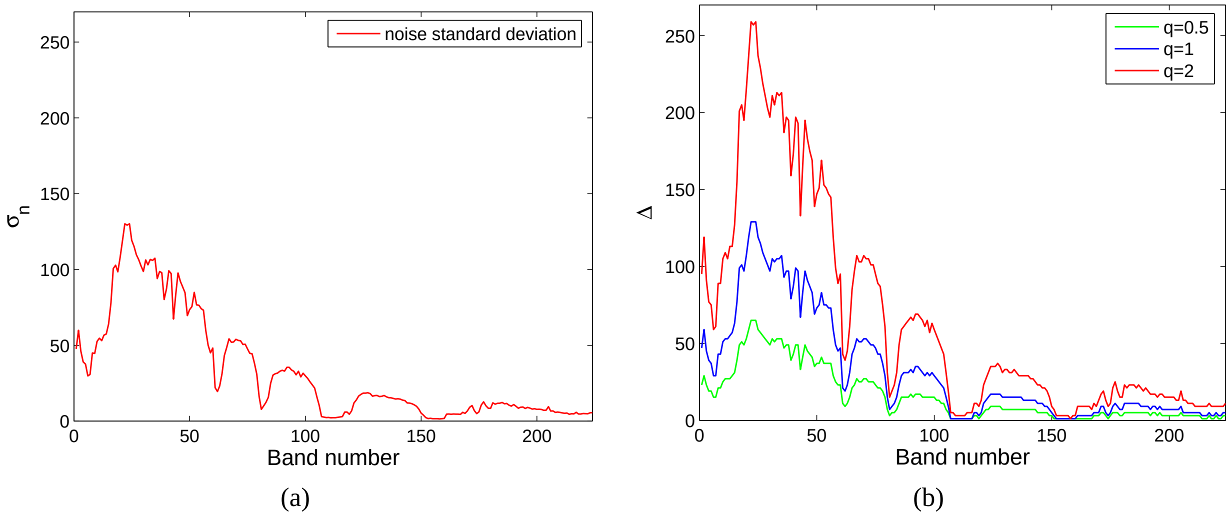


Figure 1: (a) measured noise standard deviation and (b) quantization step size for virtually lossless compression varying with band number.

VIRTUALLY LOSSLESS COMPRESSION

The term *virtually-lossless* indicates that the distortion introduced by compression should appear as an additional amount of noise, being uncorrelated and having space-invariant first order statistics such that the overall probability density function (PDF) of the noise corrupting the decompressed data, i.e. intrinsic noise plus compression-induced noise, closely matches the noise PDF of the original data. This requirement is trivially fulfilled if compression is lossless, but may also hold if the difference between uncompressed and decompressed data exhibits a peaked and narrow PDF without tails, as it happens for near-lossless techniques, whenever the user defined MAD is sufficiently smaller than the standard deviation σ_n of the background noise. Both MAD and σ_n are intended to be expressed in either physical units, for calibrated data, or as digital counts otherwise. Therefore, noise modeling and estimation from the uncompressed data becomes a major task to accomplish a virtually-lossless compression [3]. The underlying assumption is that the dependence of the noise on the signal is null, or weak. However, signal independence of the noise may not strictly hold for hyperspectral images, especially for a data set which is not definitely raw, namely postprocessed. This further uncertainty in the noise model may be overcome by imposing a margin on the relationship between target MAD and RMS value of background noise.

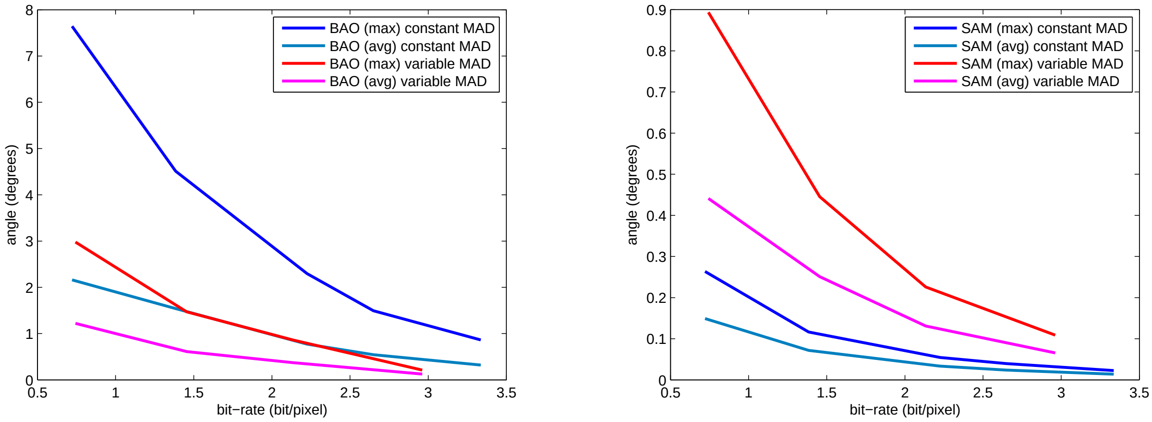
Quality evaluation of compressed remote sensing images, and specifically of hyperspectral data, cannot rely on PSNR distortion measurements only. For example we can notice that the wavelet-based JPEG2000 algorithm achieves the effect of progressively “denoising” an image as the target compression ratio increases. This fact is not surprising, since it has been demonstrated that suppression of small wavelet coefficients, which happens because of quantization, yields a powerful method for image denoising, established also in the field of astrophysical image processing [17]. Therefore, the data may become little useful once they have been compressed by means of an otherwise advanced L_2-bounded method like JPEG2000.

On the contrary, near-lossless methods, like JPEG-LS and RLPE seem to be more suitable than JPEG2000 for locally preserving even small areas of variable coarseness. The main reason of that is the *quantization noise-shaping* effect achieved by L_∞-bounded image encoders, like those based on DPCM.

The rationale of virtually-lossless compression can be summarized by the following protocol: measure the noise RMS, σ_n; if $\sigma_n < 1$, lossless compression is mandatory. Otherwise, if $1 \leq \sigma_n < 3$, near-lossless compression with MAD=1 (hence, $\Delta = 3$) might be attempted. For $3 \leq \sigma_n < 5$, compression with MAD=1 is recommended, to avoid wasting bits encoding the noise. In the general case, the relationship between MAD and σ_n, also including a margin of approximately one dB, is:

$$\mathrm{MAD} = \lfloor \max\{0, (\sigma_n - 1)/2\} \rfloor \tag{6}$$

In the case of scientific data, the signal may have been previously quantized based on different requirement; afterwards a check on the noise is made to decide whether lossless compression is really necessary, or near-lossless compression



(a) (b)

Figure 2: Spectral (angular) distortion between original and compressed spectra varying with bit rate for both near-lossless (MAD constant with wavelength) and virtually-lossless (MAD varying with wavelength) compression strategies: (a) BAO-SAM; (b) SAM.

could be used instead without penalty, being "de facto" *virtually-lossless*. Depending on the application context and the type of data, the relationship (6) may also be relaxed, e.g. by imposing that the ratio MSE(noise)/MSE(compression) is greater than, say, 3 dB, instead of the 10÷11 dB, given by (6).

The key to achieve a compression preserving the scientific quality of the data for remote-sensing is represented by the following twofold recommendation:

1. Absence of *tails* in the PDF of the error between uncompressed and decompressed image, in order to maximize the ratio $\sqrt{\mathrm{MSE}}/\mathrm{MAD}$, i.e. RMSE/MAD, or equivalently to minimize MAD for a given RMSE.

2. MSE lower by one order of magnitude than the variance of background noise σ_n^2.

Near-lossless methods are capable to fulfill such requirements, provided that the quantization step size Δ is chosen as an odd integer such that $\Delta \approx \sigma_n$. If the data are intrinsically little noisy, the protocol may lead to the direct use of lossless compression, i.e. $\Delta = 1$, to obtain what has been denoted as *virtually-lossless* compression.

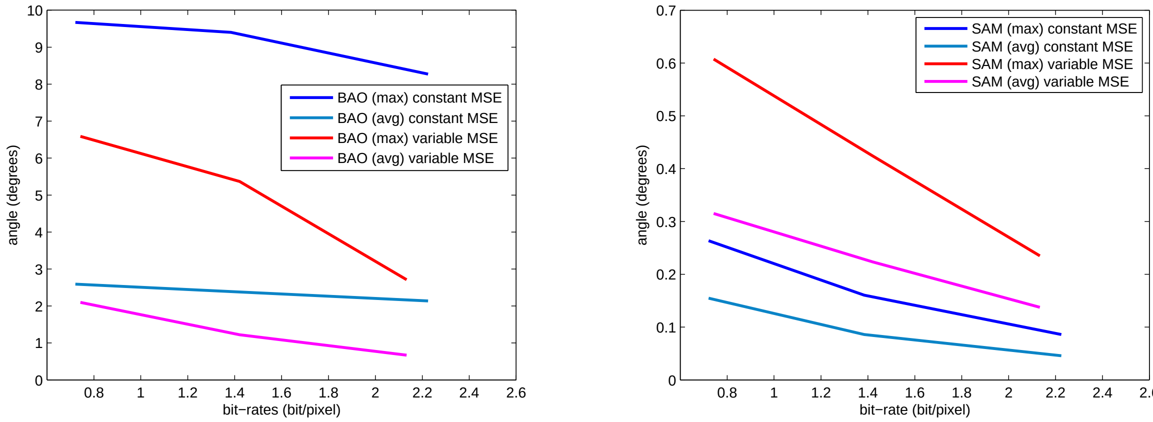
CODING RESULTS

The data set used for carrying out the experiments is composed of a sequence collected in 1997 by the *Airborne Visible InfraRed Imaging Spectrometer* (AVIRIS), operated by NASA/JPL on *Cuprite Mine*, NV test site.

The hyperspectral data have been compressed by means of the SRLP algorithm [18], which is a MAD-bounded near-lossless algorithm. The noise standard deviation σ_n of the test sequence *Cuprite Mine*, was measured by means of the scatterplot-based method [19]. We can observe in Fig. 1(a) that, for some bands, near-lossless compression with $\mathrm{MAD} = \delta = 1$ (i.e. quantization step size $\Delta = 2\delta + 1 = 3$) would yield an RMS distortion $\epsilon = \sqrt{2/3} \approx 0.82$, slightly greater than the noise RMS value of such bands, which would have the effect of increasing by a factor greater than $\sqrt{2}$, after decompression. Equivalently, the intrinsic SNR of the uncompressed image would be decremented by 3 dB after compression/decompression. In this specific case, *virtually-lossless* compression should better coincide with *lossless* compression. Near-lossless compression of such bands with MAD equal to one is unable to retain the quality of the data, because the compression-induced MSE is not one order of magnitude lower than σ_n^2, as it would be recommended for virtually-lossless compression.

To vary the compression ratio, a scale factor q is introduced, such that

$$\Delta_n = \mathrm{round}[q \cdot \sigma_n]. \tag{7}$$



(a) (b)

Figure 3: Spectral (angular) distortion between original and compressed spectra varying with bit rate for compression with MSE constant and varying with wavelength: (a) BAO-SAM; (b) SAM.

If $q \leq 1$ a strictly virtually-lossless compression is achieved, since the compression-induced distortion is over 10% lower than the intrinsic noisiness of the data. Otherwise, if $q > 1$, compression is loosely virtually-lossless, even though distortion is properly allocated among the spectral bands. Fig. 1(b) shows quantization step sizes for three different values of q. To compare virtually-lossless compression with near-lossless compression (a unique quantization step size for the whole data cube), the equivalent step size yielding approximately the same compression ratio is the integer roundoff of the geometric mean of the step sizes (7).

Fig. 2 shows that for a given bit rate, or compression ratio, the virtually-lossless compression yields far lower BAO-SAM than the near-lossless one does. Conversely, the near-lossless compression strategy is favored in terms of SAM (average and maximum). Therefore if one aims at minimizing the BAO-SAM between original and lossy compressed spectra, the virtually-lossless strategy is undoubtedly preferable.

An analogous experiment has been performed using JPEG2000. To achieve a 3D decorrelation also for JPEG2000, a spectral decorrelation (a single spectral predictor with 3 coefficients, adaptively calculated for each band) has been adopted prior to JPEG2000. This strategy allows a significant bit-rate reduction to be achieved in the lossless case. JPEG2000 has been run by setting the *rate* parameter for every band equal to the rate obtained by SRLP on the same band. That is equivalent to obtain a (roughly) *constant* MSE on every band, and an MSE *variable* from band to band. The results are reported in Fig. 3.

If we compare Fig. 3(a) with Fig. 2(a) it is apparent that the tailed error distribution originated by JPEG2000 heavily affects the BAO-SAM measure, especially the maximum. For JPEG2000 compression, the maximum BAO-SAM is $2°$ higher for constant MSE than for constant MAD allocation and $4°$ higher for variable MSE than for variable MAD allocation. It is worth noting that for BAO-SAM the variable MAD case was more favorable than constant MAD. Also the BAO-SAM average values are slightly higher for JPEG2000. Generally, the BAO-SAM distortion of data compressed by means of JPEG2000 follows the same trends highlighted by DPCM compression: the BAO-SAM is lower if distortion is allocated band by band following the noise variance, but in the JPEG2000 case BAO-SAM is always higher at the same bit-rates. Also for SAM the trends are similar in the two cases (Fig. 3(b) and Fig. 2(b)). The sole difference is that the SAM originated from a variable MSE distortion allocation is lower for JPEG2000 than for SRLP, even if SAM is always lower with a constant distortion allocation.

CONCLUDING REMARKS

This work has shown that the BAO-SAM measure can be useful in the evaluation of the distortion introduced by near-lossless compression. BAO-SAM is more sensitive than SAM with the further advantage that it is monotonic with the number of bands. These properties may be valuable for quality assessment of hyperspectral data as they are reconstructed after lossy compression. Although a compression algorithm capable of producing reconstructed vectors with upper-bounded user-defined BAO-SAM is unfeasible, it is possible to *indirectly* reduce the BAO-SAM by allocating more distortion in those bands that are less relevant, on an average, for its computation.

References

[1] D. A. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 17–28, Jan. 2002.

[2] B. Aiazzi, P. Alba, L. Alparone, and S. Baronti, "Lossless compression of multi/hyper-spectral imagery based on a 3-D fuzzy prediction," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 5, pp. 2287–2294, Sep. 1999.

[3] B. Aiazzi, L. Alparone, and S. Baronti, "Near-lossless compression of 3-D optical data," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, no. 11, pp. 2547–2557, Nov. 2001.

[4] J. Mielikainen and P. Toivanen, "Clustered DPCM for the lossless compression of hyperspectral images," *IEEE Trans. Geosci. Remote Sensing*, vol. 41, no. 12, pp. 2943–2946, Dec. 2003.

[5] E. Magli, G. Olmo, and E. Quacchio, "Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC," *IEEE Geosci. Remote Sensing Lett.*, vol. 1, no. 1, pp. 21–25, Jan. 2004.

[6] F. Rizzo, B. Carpentieri, G. Motta, and J. A. Storer, "Low-complexity lossless compression of hyperspectral imagery via linear prediction," *IEEE Signal Processing Lett.*, vol. 12, no. 2, pp. 138–141, Feb. 2005.

[7] B. Aiazzi, L. Alparone, S. Baronti, C. Lastri, and L. Santurri, "Near-lossless compression of hyperspectral imagery through crisp/fuzzy adaptive DPCM," in *Hyperspectral Data Compression*, G. Motta, F. Rizzo, and J. A. Storer, Eds., pp. 147–177. Springer, Berlin, Heidelberg, New York, 2006.

[8] G. Mercier and M. Lennon, "Joint classification and compression of hyperspectral images," in *Hyperspectral Data Compression*, G. Motta, F. Rizzo, and J. A. Storer, Eds., pp. 179–196. Springer, Berlin, Heidelberg, New York, 2006.

[9] N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 7, pp. 1552–1565, July 2004.

[10] Chein-I. Chang, "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis," *IEEE Trans. Inform. Theory*, vol. 46, no. 5, pp. 1927–1932, Aug. 2000.

[11] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 58–69, Jan. 2002.

[12] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[13] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[14] R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 713–718, Mar. 1992.

[15] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, MA, 2nd edition, 1990.

[16] N. Keshava and P. Boettcher, "On the relationships between physical phenomena, distance metrics, and best bands algorithms in hyperspectral processing," in *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*, Michael R. Descour Sylvia S. Shen, Ed., 2001, vol. 4381 of *Proc. SPIE*, pp. 55–67.

[17] J. L. Starck and F. Murtagh, "Image restoration with noise suppression using the wavelet transform," *Astron. Astrophys.*, vol. 288, pp. 342–350, 1994.

[18] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, C. Lastri, and L. Santurri, "Near-lossless compression of hyperspectral data through classified spectral prediction," 2005, vol. 5236 of *Proc. of SPIE*, pp. 105–115.

[19] B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, and I. Pippi, "Estimating noise and information of multispectral imagery," *J. Optical Engin.*, vol. 41, no. 3, pp. 656–668, Mar. 2002.