Geoinformation for European-wide Integration, Benes (ed.)

Impact of irreversible data compression on spectral distortion of hyper-spectral data

B. Aiazzi, S. Baronti, L. Santurri & M. Selva
Institute of Applied Physics "N. Carrara" IFAC-CNR, Florence, Italy

L. Alparone
Department of Electronis & Telecommunications, Universtity of Florence, Italy

Keywords: Data compression, differential pulse code modulation (DPCM), hyper-spectral data, near-lossless compression, radiometric distortion, spectral distortion

ABSTRACT: Goal of the present work is to investigate and compare different compression methodologies from the view-point of spectral distortion introduced in hyper-spectral pixel vectors. The main result of this analysis is that, for a given compression ratio, near-lossless methods, either MAD- or PMAD-constrained, are more suitable for preserving the spectral discrimination capability among pixel vectors, which is the principal outcome of spectral information. Therefore, whenever a lossless compression is not practicable, the use of near-lossless compression is recommended in such application where spectral quality is a crucial point.

1 INTRODUCTION

It is a widespread belief that the new generation of space-borne imaging spectrometers (e.g. NASA/JPL MODIS and ESA/EnviSat MERIS) will create several problems, on one side for on-board compression and transmission to ground stations, on the other side for an efficient dissemination and utilization of the outcome hyper-spectral data. In fact, the huge amount of data due to moderate ground resolution, but extremely high spectral resolution (around $10nm$), together with the high radiometric resolution (typically 12 bit word-length of the raw, i.e. uncalibrated, data from the digital counter) originates an amount of data of approximately 300 bytes/pixel. Therefore, the use of advanced compression techniques and of suitable analysis/processing procedures for dissemination to users of thematic information is mandatory.

Data compression consists of a decorrelation, aimed at generating a memoryless version of the correlated information source, followed by quantization, which introduces a distortion to allow a reduction in the information rate, and entropy coding. If the decorrelation is achieved by means of an orthonormal transformation, e.g. the discrete cosine transform (DCT), or the discrete wavelet transform (DWT), the variance of quantization errors in the transformed domain is preserved when the data are transformed back into the original domain. Thus, the mean squared error (MSE) can be easily controlled through the step sizes of quantizers. However, quantization errors in the transformed domain, which are likely to be uniformly distributed and upper bounded in modulus by half of the step size, are propagated by the inverse transformation and yield broad-tailed distributions, whose maximum absolute amplitude cannot be generally set a "priori". Therefore lossy compression methods, e.g. those proposed by the Joint Photographic Expert Group, the current standard JPEG and the upcoming standard JPEG 2000 (ISO 2000), are unable to control the reconstruction error but in the MSE sense, which means that, apart from the lossless case, relevant image features may be locally distorted or corrupted by an unpredictable and unquantifiable extent.

Compression algorithms are said to be fully reversible (lossless) when the data that are reconstructed from the compressed bit stream are identical to the original, or lossy otherwise. The difference in performance expressed by the compression ratio (CR) between lossy and lossless algorithms can be of one order of magnitude without a significant visual degradation. For this reason lossy algorithms are extremely interesting and are used in all those application in which a certain distortions may be tolerated. Actually these algorithms are more and more popular and their use is becoming widespread also in such remote sensing applications as those in which it was rightly believed, so far, that the data had to exactly retain their original values for further processing and quantitative evaluations (Vaughn & Wilkinson 1995). This aspect is crucial for transmission from satellite to Earth

receiving stations; in fact, once the data were lossy compressed, they would not be available as they were acquired for the user community. The distortions introduced might influence such research activities as modelling, classification and post-processing in general. As a matter of fact, however, the intrinsic noisiness of sensors prevents from adopting strictly lossless techniques in order to obtain a considerable bandwidth reduction (Roger & Arnold 1996, Aiazzi *et al.* 2001a). In this light, error-bounded near-lossless algorithms (Chen & Ramabadran 1994, Aiazzi *et al.* 2001b) are growing in importance since they are capable to guarantee that at every pixel of the reconstructed image the error is bounded and user-defined.

In the medical field objective measurements, like MSE, maximum absolute distortion (MAD) and percentage MAD (PMAD) may be integrated with qualitative judgements of skilled experts, e.g. expressed in terms of Receiver Operating Characteristic (ROC) curves (Aiazzi *et al.* 1997b). In remote sensing applications, however, photoanalysis is not the only concern. The data is often post-processed to extract information that may not be immediately available by user inspection. In this perspective, if the MAD error is constrained to be, e.g. one half of the standard deviation of the background noise, assumed to be additive, Gaussian, and independent of the signal, the decoded image will be virtually lossless. This term indicates not only visual indistinguishability from the original, but also that possible outcomes of post-processing are likely to be practically the same as if they were calculated from the original data. The price of compression becomes a small and uniform increment in noisiness.

When higher compression ratios are demanded, a PMAD-constrained approach may be rewarding in terms of scientific quality preservation of the decompressed data (Ryan & Arnold 1997). The rationale is that automatic analysis and processing algorithms may be more sensitive to *relative* errors on pixels, than to *absolute* errors. For best performance, however, relative error-constrained compression requires logarithmic quantization (Aiazzi *et al.* 2001b), which is penalized with respect to linear quantization in the Rate Distortion (RD) sense, with an MSE distortion measure.

When multi-spectral or better hyper-spectral data are being dealt with, *spectral* distortion becomes a primary concern, besides spatial and radiometric distortions. Spectral distortion is a measurement of how a pixel vector (i.e. a vector having as many components as spectral bands) changes because of an irreversible compression of its components. A widely used measurement is the angle between the two vectors. More sophisticated measurements based on information-theoretic criteria have recently proven themselves more effective in discriminating spectral classes (Chang 2000).

2 DISTORTION MEASURES

2.1 Radiometric Distortion

Let $0 \leq g(i,j) \leq g_{fs}$ denote an N-pixel digital image and let $\tilde{g}(i,j)$ be its possibly distorted version achieved by compressing $g(i,j)$ and decompressing the outcome bit stream. Widely used distortion measurements are the following:

Mean absolute error (MAE), or L_1 (norm),

$$MAE \;=\; \frac{1}{N}\sum_i \sum_j |g(i,j) - \tilde{g}(i,j)|; \tag{1}$$

Mean Squared Error (MSE), or L_2^2,

$$MSE \;=\; \frac{1}{N}\sum_i \sum_j [g(i,j) - \tilde{g}(i,j)]^2; \tag{2}$$

Root MSE (RMSE), or L_2,

$$RMSE \;=\; \sqrt{MSE}; \tag{3}$$

Signal to Noise Ratio (SNR)

$$SNR_{(\mathrm{dB})} \;=\; 10 \cdot \log_{10} \frac{\bar{g^2}}{MSE + \frac{1}{12}}; \tag{4}$$

Peak SNR (PSNR)

$$PSNR_{(\mathrm{dB})} \;=\; 10 \cdot \log_{10} \frac{g_{fs}^2}{MSE + \frac{1}{12}}; \tag{5}$$

Maximum absolute distortion (MAD), or *peak error*, or L_∞,

$$MAD \;=\; \max_{i,j}\{|g(i,j) - \tilde{g}(i,j)|\}; \tag{6}$$

Percentage maximum absolute distortion (PMAD)

$$PMAD \;=\; \max_{i,j}\left\{\frac{|g(i,j) - \tilde{g}(i,j)|}{g(i,j)}\right\} \times 100. \tag{7}$$

Both in (4) and in (5) the MSE is incremented by the variance of the integer roundoff error, to handle the limit lossless case, when $MSE = 0$. Thus, SNR and PSNR will be upper bounded by $10 \cdot \log_{10}(12 \cdot \bar{g^2})$ and $10 \cdot \log_{10}(12 \cdot g_{fs}^2)$, respectively.

When multi-band data are concerned, let $v_l \triangleq g_l(i,j),\ l = 1, \cdots, L$ denote the lth component of the original multi-spectral pixel vector $\mathbf{v}$ and $\tilde{v}_l \triangleq \tilde{g}_l(i,j),\ l = 1, \cdots, L$ its distorted version. Some of the radiometric distortion measurements (1)-(7) may be extended to vector data as: Average RMSE (A-RMSE), or $L_1(L_2)$ (the innermost norm L_2 refers to vector space (l), the outer one to pixel space (i,j)

$$ARMSE \;=\; \frac{1}{N}\sum_{i,j} \sqrt{\sum_l [g_l(i,j) - \tilde{g}_l(i,j)]^2}; \tag{8}$$

Peak RMSE (P-RMSE), or $L_\infty(L_2)$,

$$PRMSE = \max_{i,j} \sqrt{\sum_l [g_l(i,j) - \tilde{g}_l(i,j)]^2}; \quad (9)$$

$$SNR = 10 \cdot \log_{10} \frac{\sum_{i,j,l} g_l^2(i,j)}{\sum_{i,j,l} [g_l(i,j) - \tilde{g}_l(i,j)]^2}; \quad (10)$$

$$PSNR = 10 \cdot \log_{10} \frac{N \cdot L \cdot g_{fs}^2}{\sum_{i,j,l} [g_l(i,j) - \tilde{g}_l(i,j)]^2}; \quad (11)$$

Three-dimensional MAD (MAD^3), or $L_\infty(L_\infty)$,

$$MAD^3 = \max_{i,j,l} \{|g_l(i,j) - \tilde{g}_l(i,j)|\}; \quad (12)$$

Three-dimensional PMAD (PMAD^3)

$$PMAD^3 = \max_{i,j,l} \left\{ \frac{|g_l(i,j) - \tilde{g}_l(i,j)|}{g_l(i,j)} \right\} \times 100. \quad (13)$$

In practice, A-RMSE (8) and P-RMSE (9) are respectively the average and maximum of the Euclidean norm of the distortion vector. SNR (10) is the extension of (4) to the 3-D data cube. PSNR is the maximum SNR, given the full-scales of each vector component. MAD (12) is the maximum over the pixel set of the maximum absolute component of the distortion vector. PMAD (13) is the maximum percentage error over each vector component of the data cube.

2.2 *Spectral Distortion*

Given two spectral vectors $\mathbf{v}$ and $\tilde{\mathbf{v}}$ both having L components, in which $\mathbf{v} = \{v_1, v_2, \cdots, v_L\}$ is the original (hyper)spectral pixel vector $v_l = g_l(i,j)$ while $\tilde{\mathbf{v}} = \{\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_L\}$ is the distorted vector obtained after decompression, $\tilde{v}_l = \tilde{g}_l(i,j)$. Analogously to the *radiometric* distortion measurements, the following *spectral* distortion measurement may be defined.

The spectral angle mapper (SAM) denotes the absolute value of the spectral angle between the pair of vectors:

$$SAM(\mathbf{v}, \tilde{\mathbf{v}}) \triangleq \arccos \left(\frac{< \mathbf{v}, \tilde{\mathbf{v}} >}{||\mathbf{v}||_2 \cdot ||\tilde{\mathbf{v}}||_2} \right) \quad (14)$$

SAM can be measured in either degrees or radians.

Another measurement suitable for hyper-spectral data (i.e. for data with large number of components) is the spectral information divergence (SID) (Chang 2000) derived from information-theoretic Kullback-Leibler distance:

$$SID(\mathbf{v}, \tilde{\mathbf{v}}) = D(\mathbf{v}||\tilde{\mathbf{v}}) + D(\tilde{\mathbf{v}}||\mathbf{v}) \quad (15)$$

with $D(\mathbf{v}||\tilde{\mathbf{v}})$ being the Kullback-Leibler distance (KLD), or entropic divergence, or *discrimination*, defined as

$$D(\mathbf{v}||\tilde{\mathbf{v}}) \triangleq \sum_{l=1}^{L} p_l \log \left(\frac{p_l}{q_l} \right) \quad (16)$$

in which

$$p_l \triangleq \frac{v_l}{||\mathbf{v}||_1} \quad \text{and} \quad q_l \triangleq \frac{\tilde{v}_l}{||\tilde{\mathbf{v}}||_1} \quad (17)$$

In practice SID is equal to the symmetric KLD and can be compactly written as

$$SID(\mathbf{v}, \tilde{\mathbf{v}}) = \sum_{l=1}^{L} (p_l - q_l) \log \left(\frac{p_l}{q_l} \right) \quad (18)$$

which turns out to be symmetric, as one can easily verify. It can be proven as well that SID is always non-negative, being zero iff. $p_l \equiv q_l, \ \forall l$, i.e. $\mathbf{v} \equiv \tilde{\mathbf{v}}$, and is not upper bounded. The measure unit of SID depends on the base of the logarithm: *nat/vector* with natural logarithms and *bit/vector* with logarithms in base two.

Both SAM (14) and SID (18) may be either averaged on pixel vectors, or the maximum may be taken instead, as more representative of spectral quality.

3 NEAR-LOSSLESS COMPRESSION OF HYPER-SPECTRAL DATA

Differential Pulse Code Modulation (DPCM) schemes, either causal, like JPEG in lossless mode (Rao & Hwang 1996) and JPEG-LS (Weinberger *et al.* 2000), or non-causal (Aiazzi *et al.* 1997), are usually employed for error-free image compression. DPCM basically consists of a prediction followed by entropy coding of the outcome prediction errors. Whenever multi-spectral images are to be compressed, advantage may be taken from the spectral correlation of the data for designing a prediction that is both *spatial* and *spectral*, from a causal neighborhood of pixels (Wang *et al.* 1995, Roger & Cavenor 1996, Wu & Memon 2000). Causal means that only previously scanned pixels on the current and previously encoded bands may be utilized for predicting the current pixel value. This strategy is as more effective as the data is more spectrally correlated, as in the case of hyper-spectral data (Aiazzi *et al.* 1999).

The simplest way to design a predictor, once a causal neighborhood is set, is to take a linear combination, or regression, of the values of such a neighborhood, with coefficients optimized in order to yield minimum mean square prediction error (MSPE) over the whole image (Rao & Hwang 1996). In case of 3-D images, the causal neighborhood may also contain pixels belonging to the previously encoded bands, in order to exploit inter-band correlation. Such a prediction, however, is optimum only for stationary signals. To overcome this drawback, two variations have been proposed: *adaptive* DPCM (ADPCM) (Rao & Hwang 1996), in which the coefficients of predictors are continuously recalculated from the incoming

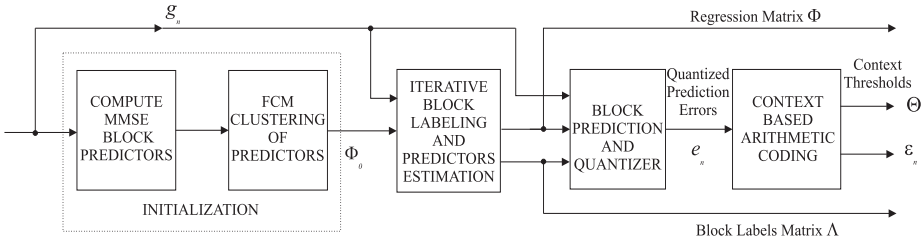


Figure 1. Flowchart of the classified 2-D/3-D DPCM encoder for near-lossless compression of remotely sensed optical data, based on a relaxation-labelled prediction.

new data, and *classified* DPCM (Aiazzi *et al.* 2002b), in which several classes are preliminarily recognized, an optimized MMSE predictor is calculated for each class, and such predictors are enabled to attain the best space-varying prediction.

The DPCM encoder (Aiazzi *et al.* 2001b), shown in Fig. 1, is of the latter type, being based on a classified linear-regression prediction followed by context-based arithmetic coding (Aiazzi *et al.* 2002a) of the outcome residuals. Image bands are partitioned into blocks, typically 8×8, and an MMSE linear predictor is calculated for each block. Given a preset number of classes, a clustering algorithm produces an initial guess of as many classified predictors to be fed to an iterative labelling procedure which classifies pixel blocks simultaneously refining the associated predictors. All the predictors are transmitted along with the label of each block. In order to achieve a *near-lossless* compression, prediction errors are quantized with an odd valued step size and then arranged into activity classes, which are entropy coded by means of arithmetic coding (Witten *et al.* 1987).

4 EXPERIMENTAL RESULTS

The data set includes a sequence of hyper-spectral images collected in 1997 by the *Airborne Visible InfraRed Imaging Spectrometer* (AVIRIS), operated by NASA/JPL, on the *Cuprite Mine* test site, in Nevada. The sequence is constituted by 224 bands recorded at different wave-lengths in the range $380 \div 2500\ nm$, with a average spectral separation between two bands of $10\ nm$. The size of each image is 614×2048 pixels. The raw sequence was acquired by the $12\ bit$ analog-to-digital converter (ADC) with which the sensor was equipped in 1995, in place of the former 10-bit ADC. The raw data from the digital counter have been radiometrically calibrated by multiplying by a gain and adding an offset (both varying with wave-lengths), and are expressed as radiance values, rounded to integers, and packed in a 16-bit word-length, including a sign bit. Thus, the full-scale is $g_{fs} = 32767$ for all bands. Band # 48 ($808\ nm$) is shown in Fig. 2. The second spectrometer, covering the near-infrared

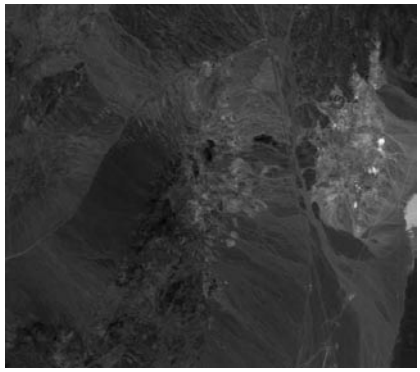
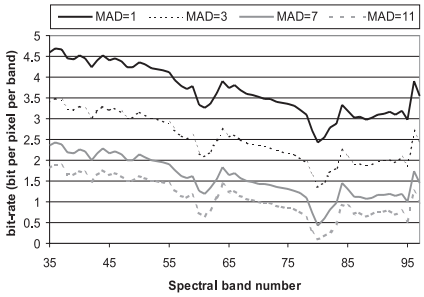


Figure 2. Band # 48 ($808\ nm$ wavelength) of NASA/JPL AVIRIS *Cuprite Mine*. The detail shown is 614×512 pixels with a word-length of 16 bits.

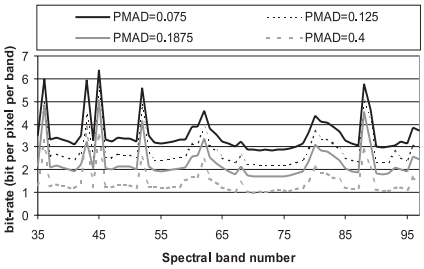
(NIR) spectrum, was analyzed in a recent work by the authors (Aiazzi *et al.* 2001a). It was found that the noise affecting AVIRIS data is somewhat correlated spectrally and along track, and less across track, due to the "wisk-broom" scan mechanism, as well as to post-processing. In this light, the claimed achievement of the theoretical compression limits, based on a noise whiteness assumption (Roger & Arnold 1994, Roger & Cavenor 1996), should be revised.

Bands 35 to 97 covering the NIR wave-lengths have been compressed in both MAD-constrained mode (linear quantisation) and PMAD constrained mode (logarithmic quantisation). The work parameters of the RLPE algorithm are non-crucial (Aiazzi *et al.* 2002b) and have been chosen so as to balance coding performances with encoding time.

The outcome bit-rates varying with band number, together with the related distortion parameters are shown in Fig. 3. As it appears the bit-rate plots follow similar trends varying with the amount of distortion, but quite different trends for the two types of distortion (i.e. either MAD or PMAD). For example, around the water vapour absorption wave-lengths ($\approx$ Band 80) the MAD-bounded plots exhibit pronounced valleys, that can be explained because the intrinsic

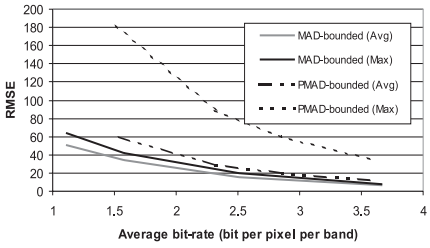


(a)

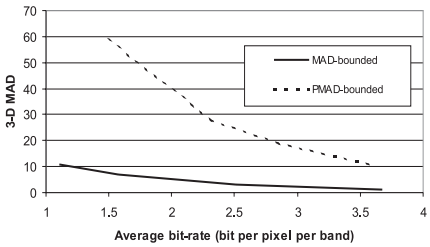


(b)

Figure 3. Bit-rates produced by 3-D RLPE on the data produced by the second spectrometer (NIR) of AVIRIS *Cuprite Mine '97*: (a) linear quantisation to yield user-defined MAD values; (b) logarithmic quantisation to yield user-defined PMAD values.



(a)



(b)

Figure 4. Radiometric distortions versus bit-rate for compressed AVIRIS *Cuprite Mine '97* data: (a) RMSE; (b) MAD^3 (3-D MAD).

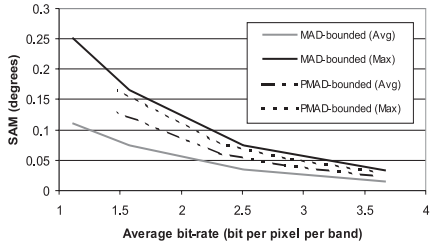
SNR of the data gets lower; thus the linear quantizer dramatically abates the *noisy* prediction errors. On the other hand the PMAD-bounded encoder tends to quantise the noisy residuals more finely when the signal is lower. Therefore bit-rate peaks are generated instead of valleys. More generally speaking, bit-rate peaks from the PMAD-bounded encoder are associated with low responses from the spectrometer. This explains why the bit-rate plots of Fig. 3(b) never fall below one bit per pixel per band.

Some of the rediometric distortion measures defined in Subsect. 2.1 have been calculated on the distorted hyper-spectral pixel vectors achieved by decompressing the bit-streams generated by the near-lossless encoder, both MAD- and PMAD-bounded.

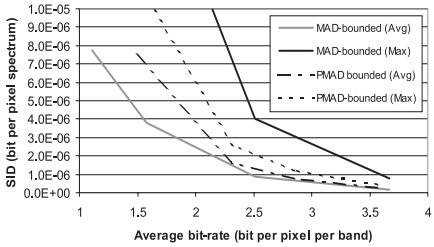
RMSEs of the vector data, both *average* RMSE (8) and *peak*, i.e. maximum, RMSE (9) are plotted in Fig. 4(a) as a function of the bit-rate from the encoder. The MAD-bounded encoder obviously minimises both the radiometric distortions: average (A-RMSE) and maximum (P-RMSE) Euclidean norm of the pixel error vector. A further advantage is that A-RMSE and P-RMSE are very close to each other for all bit-rates. The PMAD-bounded encoder is somewhat poorer: A-RMSE is comparable with that of the former, but P-RMSE is far larger, due to the high-signal components that are coarsely quantised in order to minimise PMAD. Trivially, the MAD of the data cube (3-D MAD) (12) is exactly equal to the desired value (see Fig. 3(a)), whereas the PMAD, being unconstrained, is higher. Symmetric results, not reported here, have been found by measuring PMAD on MAD-bounded and PMAD-bounded decoded data.

As far as *radiometric* distortion is concerned, results are not surprising. Radiometric distortions measured on vectors are straightforwardly derived from those measured on scalar pixel values. The introduction of such *spectral* measurements as SAM (14) and SID (18) may encompass the rationale of *distortion*, traditionally established in the signal/image processing community.

Fig. 5 shows spectral distortion measures calculated between original and decompressed hyper-spectral pixel vectors. The PMAD-bounded algorithm yields plots (maximum and average SAM in Fig. 5(a)) that lie in the middle between the corresponding ones produced by the MAD-bounded algorithm and are very close to each other too. Since the *maximum* SAM is a better clue of spectral quality of the decoded data than the *average SAM* may be, a likely conclusion would be that PMAD-bounded compression optimizes the *spectral* quality of the data, while MAD-bounded is superior in terms of *radiometric* quality.



(a)



(b)

Figure 5. Spectral distortions versus bit-rate for compressed AVIRIS *Cuprite Mine '97* data: (a) spectral angle mapper (SAM); (b) spectral information divergence (SID).

The considerations expressed for SAM are emphasized by the plots of Fig. 5(b) reporting average and maximum SID. As it appears, the latter is capable of discriminating spectral quality more finely than SAM does, as previously noticed by Chan (2000) in the case of multi-spectral classification.

5 CONCLUDING REMARKS

This paper has demonstrated the potential usefulness of *near-lossless* compression, i.e. with bounded pixel error, either *absolute* or *relative*. Unlike lossless compression achieving typical CRs around *two*, near-lossless compression can be adjusted to allow a *virtually lossless* compression with CR always larger than *five* for hyper-spectral data. The main result of this analysis is that, for a given CR, near-lossless methods, either MAD- or PMAD-constrained, are more suitable for preserving the spectral discrimination capability among pixel vectors, which is the principal outcome of spectral information. Therefore, whenever a lossless compression is not practicable, the use of near-lossless compression is recommended in such application where spectral quality is a crucial point. Furthermore, since the maximum reconstruction error is defined by the user before compression, whenever higher CRs are required, the loss of performance expected on application tasks can be accurately modelled and predicted.

ACKNOWLEDGEMENTS

The authors are grateful to F. Lotti of IFAC-CNR for valuable suggestions and useful discussion on multi-spectral data quality assessment.

This work has been carried out in part under grants from the *Italian Space Agency* (ASI).

L. Alparone wishes to acknowledge the financial support of the Italian Ministry of Education, University and Research (MIUR), COFIN 2001.

REFERENCES

Aiazzi, B., L. Alparone, S. Baronti & F. Lotti, 1997a. Lossless image compression by quantization feedback in a content-driven enhanced Laplacian pyramid. *IEEE Transactions on Image Processing*, 6(6):831–843.

Aiazzi, B., L. Alparone, S. Baronti, G. Chirò, F. Lotti & M. Moroni, 1997b. A pyramid-based error-bounded encoder: An evaluation on X-ray chest images. *Signal Processing*, 59(2):173–187.

Aiazzi, B., P. Alba, L. Alparone & S. Baronti, 1999. Lossless compression of multi/hyper-spectral imagery based on a 3-D fuzzy prediction. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2287–2294.

Aiazzi, B., L. Alparone, A. Barducci, S. Baronti & I. Pippi, 2001a. Information-theoretic assessment of sampled hyperspectral imagers. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7):1447–1458.

Aiazzi, B., L. Alparone & S. Baronti, 2001b. Near-lossless compression of 3-D optical data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(11):2547–2557.

Aiazzi, B., L. Alparone & S. Baronti, 2002a. Context modeling for near-lossless image coding. *IEEE Signal Processing Letters*, 9(3):77–80.

Aiazzi, B., L. Alparone & S. Baronti, 2002b. Near-lossless image compression by relaxation-labelled prediction. *Signal Processing*, 82(7) (in press).

Chang, Chein-I, 2000. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46(5):1927–1932.

Chen, K. & T. V. Ramabadran, 1994. Near-lossless compression of medical images through entropy-coded DPCM. *IEEE Transactions on Medical Imaging*, 13(3):538–548.

ISO/IEC JTC 1/SC 29/WG1, 2000. *ISO/IEC FCD 15444-1: Information technology–JPEG 2000 image coding system: Core coding system*, [WG 1 N 1646].

Rao, K. K. & J. J. Hwang, 1996. *Techniques and Standards for Image, Video, and Audio Coding*, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Roger, R. E. & J. F. Arnold, 1994. Reversible image compression bounded by noise. *IEEE Transactions on Geoscience and Remote Sensing*, 32(1):19–24.

Roger, R. E & M. C. Cavenor, 1996. Lossless compression of AVIRIS images. *IEEE Transactions on Image Processing*, 5(5):713–719.

Ryan, M. J. & J. F. Arnold, 1997. Lossy compression of hyperspectral data using vector quantization. *Remote Sensing of Environment*, 61(9):419–436.

Vaughn, V. D & T. S. Wilkinson. 1995. System considerations for multispectral image compression design. *IEEE Signal Processing Magazine*, 12(1):19–31.

Weinberger, M. J., G. Seroussi & G. Sapiro, 2000. The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS. *IEEE Transactions on Image Processing*, 9(8):1309–1324.

Witten, I.H., R. M. Neal & J. G. Cleary, 1997. Arithmetic coding for data compression. *Communications of the ACM*, 30(6):520–540.

Wang, J., K. Zhang & S. Tang, 1995. Spectral and spatial decorrelation of Landsat-TM data for lossless compression. *IEEE Transactions on Geoscience and Remote Sensing*, 33(5):1277–1285.

Wu, X. L. & N. Memon, 2000. Context-based lossless interband compression–Extending CALIC. *IEEE Transactions on Image Processing*, 9(6):994–1001.