ON-BOARD LOSSLESS HYPERSPECTRAL DATA COMPRESSION: LUT-BASED OR CLASSIFIED SPECTRAL PREDICTION?

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

This paper presents a novel algorithm suitable for the lossless compression of hyperspectral imagery. The algorithm generalises two previous algorithms, in which the concept nearest neighbour (NN) prediction implemented through lookup tables (LUTs) was introduced. Here, the set of LUTs, two or more, say M, on each band are allowed to span more than one previous band, say N bands, and the decision among one of the NM possible prediction values is based on the closeness of the value contained in the LUT to an advanced prediction, spanning N previous bands as well, provided by a top-performing scheme recently developed by the authors and featuring a classified spectral prediction. Experimental results carried out on the AVIRIS '97 data-set show improvements up to 15% over the baseline LUT-NN algorithm. However, preliminary results carried out on raw data show that all LUT-based methods are not suitable for on-board compression, since they take advantage uniquely of the sparseness of data histograms, which is originated by the on-ground calibration procedure.

INTRODUCTION

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 consists of a decorrelation of the correlated information source, possibly followed by quantization, to reduce transmission rates, and entropy coding of residues. To meet the quality issues of hyperspectral image analysis, differential pulse code modulation (DPCM) is usually employed for lossless/near-lossless compression [1]: 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 to be achieved as a limit case. The simplest way to design a predictor, once a causal neighbourhood, i.e. a set of pixels that have been previously scanned, is fixed, is to take a linear combination of the values of its samples. The coefficients of predictor can be calculated so as to yield minimum mean squared error (MMSE) over the whole image. Such a prediction is optimal only for stationary signals. Therefore, two strategies have been proposed to overcome this drawback and obtain an *adaptive* prediction. Adaptive DPCM (ADPCM), in which the coefficients of predictors are continuously recalculated from the incoming new data is traditionally used for one-dimensional signals [2]. A more recent DPCM typology suitable for digital images is *classified* DPCM, in which a number of statistical classes of samples are preliminarily recognized, an optimized predictor is calculated for each class, and such predictors are switched, either hardly [3] or softly [4], to attain the best space-varying prediction. The two strategies of hard/soft classified prediction will be referred to as *adaptive selection/combination of adaptive predictors* (ASAP/ACAP).

Whenever multiband 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 neighbourhood of pixels [1]. Causal means that only previously scanned pixels on the current and previously encoded bands may be used for predicting the current pixel value. This strategy is more and more effective as spectral correlation increases, as it happens with hyperspectral data.

State-of-the-art schemes feature a three-dimensional (3D) prediction, i.e. jointly spatial and spectral, whose support comprises pixels belonging to the current and few previously scanned bands, one [5] or two at most [6, 7, 8]. The ACAP encoder [4] has been extended by the authors to 3D data [9], same as the 3D ASAP encoder [6], by simply changing the two-dimensional (2D) neighbourhood (spatial) [3] into a 3D one spanning up to three previous bands.

A different approach specific to hyperspectral images is the extension of 3D CALIC [10], originally conceived for color images, having few spectral bands, to image data having a greater number of highly correlated bands. The method, referred to as M-CALIC [8], significantly outperforms 3D CALIC, to which it largely sticks, with a moderately increased computational complexity and absence of setup parameters crucial for performances.

The non-stationarity characteristics of hyperspectral data in both spatial and spectral domains, together with computational constraints, make the jointly spatial and spectral prediction to take negligible extra advantage from a number of previous bands greater than two [6, 7, 8]. However, if the original hyperspectral pixel vectors are classified into spatially homogeneous classes, whose map must be transmitted as side information, then a purely spectral prediction may be profitably carried out on pixel spectra belonging to each class by means of a different set of linear spectral predictors, as many as the bands, of length up to twenty, i.e. spanning up to twenty previous bands. This original approach was recently introduced for lossless compression [11], and provided the best compression ratios in the literature. Unfortunately, the computational effort for clustering hyperspectral pixel vectors makes the method unsuitable for applications requiring low complexity. In fact, computing power is limited for satellite on-board compression, and coding benefits must be traded off with computational complexity [12, 8, 5]. Furthermore, since the cost of overhead (classification map and spectral predictors) is independent of the target compression ratio, the method [11] seems to be not recommendable for near-lossless compression, which might be achieved in principle.

The rationale of a purely spectral prediction performed on pre-classified data has lead to the development of the up-to-date crisp and fuzzy adaptive spectral predictors, referred to as *Spectral-Relaxation Labeled Prediction* (S-RLP) and *Spectral-Fuzzy Matching Pursuits* (S-FMP) [13].

MIELIKAINEN'S LUT-NN

In a recently published paper [14], Mielikainen introduced a very simple spectral prediction given by the value taken on the current band by its nearest neighbour (NN), i.e. the spatially closest pixel, previously encountered along the scan path, having the same value as the current pixel on the previous band. Such a prediction, which is computationally very simple and can be effectively implemented by means of a dynamically updated lookup table (LUT), outperforms most of state-of-the-art methods on a large test set of AVIRIS data, (16-bit radiance format). The reason of its success relies on that prediction is forced to yield a value that does appear in the histogram of the current band, unlike conventional prediction strategies, which are likely to produce values that do not appear in the current band. On the other hand, the depth of spectral prediction is limited to one and this may lead to not fully exploit the spectral correlation of the data set. This method will be referred in the following as LUT-NN.

HUANG'S LAIS-LUT

The rationale of prediction based on LUTs has been later extended by Huang and Sriraja in [15] by exploiting two LUTs, respectively containing the closest neighbour and the second closest neighbour of the current pixel on the previous band. To yield the current pixel prediction, the choice between the two values contained in the two LUTs, indexed by the radiance level of the current pixel in the previous band, is based on the similarity to a reference prediction, which takes into account of the cross gain between the current and the previous band, as indicated by its acronym *locally average interband scaling* (LAIS). This strategy, which will be referred as LAIS-LUT in the remainder of this paper, yields significant coding benefits (3-4%) over [14] at a moderate extra cost. The reason of its success is twofold: changes in gain across wave-lengths are considered in the final prediction; furthermore, two distinct LUTs allow a wider choice of NNs to be available for selecting the best prediction.

PROPOSED S-RLP-LUT AND S-FMP-LUT

The idea of LUT-NN and especially the development of LAIS-LUT have lead to the generalization proposed in this paper. Here, prediction may span an arbitrary number of bands, say N. The number of LUTs on each band is arbitrary as well, say M. The decision based on LAIS, as in [15], was replaced by a more fitting reference prediction. In fact the advantages of more than one band for LUT-based prediction and of more than two LUTs for each band quickly vanish as M and N increase. Thus, the decision among one of the NM possible prediction values is based on the closeness of the value contained in the LUTs to an advanced prediction, spanning N previous bands as well, provided by the top-performing schemes S-RLP and S-FMP [13], featuring a crisp and fuzzy adaptive spectral prediction, respectively. The drawback that the optimal predictions provided by S-RLP and S-FMP do not provide a radiance value contained in the histogram of the current band to predict is mitigated by the multiple LUTs, which make a wide choice of NNs to be available for prediction. The proposed methods will be denoted in the next sections as S-RLP-LUT and S-FMP-LUT, depending on whether the reference prediction is crisp or fuzzy.

EXPERIMENTAL RESULTS AND COMPARISONS

The first data set comprises four hyperspectral images collected in 1997 by the *Airborne Visible InfraRed Imaging Spectrometer* (AVIRIS), operated by NASA/JPL. The four test sites are *Cuprite Mine* and *Lunar Lake* in Nevada, *Moffett Field* and *Jasper Ridge* in California. All images comprise 224 bands recorded at different wave-lengths in the range 380 to 2500 nm, with a nominal spectral separation of 10 nm between two adjacent bands. Each image is constituted by a variable number of scenes of size 512 lines by 614 columns. All data that have been considered for compression are in *radiance* units, 16-bit format.

Besides LUT-NN, LAIS-LUT, the proposed S-RLP-LUT and S-FMP-LUT and the up-to-date S-RLP and S-FMP, the other methods compared are the clustered DPCM (C-DPCM) [11], M-CALIC [8], and the Spectral-oriented Least SQuares (SLSQ) encoder [5]. Table 1 reports lossless bit rates produced by the nine algorithms on the whole image test set. All scenes of each image have been compressed. The results clearly show that the LUT strategy coupled with the adaptive spectral predictions provided by S-RLP and S-FMP outperforms any other scheme in the literature, either based on LUTs or not, by an extent of 15% up to almost 25%. Such results are really surprising, especially because they concern lossless compression, in which the attainable bit rate is lower bounded by the entropy of the source. An explanation of that may be that conventional DPCM schemes, i.e. non based on LUTs, implicitly assume a model of source that is not verified by hyperspectral radiance data.

To check this hypothesis, a further set of coding experiments was carried out on an additional data set. A sample 512 $\times$ 512 scene acquired in 92 spectral bands (VNIR+SWIR) by the *Multispectral Infrared and Visible Imaging Spectrometer* (MIVIS) on the urban area of Viareggio, in Italy, was available in both raw and calibrated versions. The former are 12-bit digital counts; only the dark current has been subtracted. The latter are 16-bit radiance units, obtained by multiplying the raw data acquired at each wave-length by a suitable real-valued gain, constant throughout the scene, and rounding the outcome to integer.

Table 1: Bit rates (bit/pel/band on disk) for lossless compression of AVIRIS 1997 test hyperspectral images. S-RLP-LUT and S-FMP-LUT use $N = 20$ previous bands and $M = 4$ LUTs per band. Lowest rates in boldface.

	Cuprite	Jasper	Lunar	Moffett	Average
SLSQ	4.94	4.95	4.95	4.98	4.96
M-CALIC	4.89	4.97	4.80	4.65	4.83
C-DPCM	4.68	4.62	4.75	4.62	4.67
S-RLP	4.69	4.65	4.69	4.67	4.67
S-FMP	4.66	4.63	4.66	4.63	4.64
LUT-NN	4.66	4.95	4.71	5.05	4.84
LAIS-LUT	4.47	4.68	4.53	4.76	4.61
S-RLP-LUT	3.94	4.07	3.97	4.11	4.02
S-FMP-LUT	**3.91**	**4.05**	**3.94**	**4.07**	**3.99**

Table 2: Bit rates (bit/pel/band on disk) for lossless compression of a MIVIS hyperspectral image in both raw and calibrated versions. Lowest rates in boldface.

	MIVIS raw	MIVIS calibrated
S-RLP	4.38	6.93
S-FMP	**4.34**	6.91
LUT-NN	5.32	6.25
LAIS-LUT	5.06	5.94
S-RLP-LUT ($N = 12$, $M = 4$)	4.56	5.16
S-FMP-LUT ($N = 12$, $M = 4$)	4.44	**5.12**

The lossless coding results reported in Table 2 reveal that on calibrated data all LUT-based methods more or less outperform the crisp and fuzzy adaptive spectral predictors S-RLP and S-FMP. On raw data, however, even the most sophisticated LUT-based methods proposed by the authors cannot attain the results given by any of S-RLP and S-FMP. Furthermore, calibrated data are more difficult to be compressed by any type of algorithm, notwithstanding they differ from the corresponding raw data essentially by a gain factor varying with the wave-length.

DISCUSSION

A thorough analysis of the characteristics of all LUT-based algorithms reveals that this strategy takes advantage uniquely of the sparseness of radiance histograms in each band [16]. Actually, LUT-based prediction is equivalent to a dynamic remapping of radiance values, such that the remapped histograms have no holes [17] and spectral correlation, on which compression performances rely, is preserved. Preliminary results of compression of raw, i.e. uncalibrated, 12-bit data produced by an on-board instrument reveal that no advantage whatsoever is given by LUTs and that the methods described in [13] yield the lowest bit rates on such data. Therefore, it seems that the LUT strategy, though simple and attractive, is not recommended for on-board lossless compression. Furthermore, the nature of LUT prediction make it unsuitable for near-lossless compression. It is expected that the advantages of LUTs vanish as the amount of error increases.

CONCLUDING REMARKS

This paper has introduced a generalization of two existing LUT-based lossless compression algorithms for hyperspectral data. Experimental results carried out on the same AVIRIS dataset as in [14, 15, 13] show improvements of about 15% over [14], of 10% over [15] and of 13% over [13]. The computational complexity is about 10% higher than that of [13], but several times higher than that of [14] and [15]. However, most of computing time is due to the advanced reference prediction [13], which is responsible for about 5% of the bit rate decrement. The remaining decrement over [14] and [15] depends on the use of multiple LUTs on more previous bands.

Furthermore, this work has demonstrated that LUT-based prediction strategies are attractive for lossless compression of *calibrated* hyperspectral data, i.e. spectral radiance data, where all conventional DPCM schemes fail in providing bit rates close to the entropy of the data. In fact the entropy of prediction errors is *inflated* by the on-ground calibration procedure aimed at producing spectral radiance data. Further concepts and ideas can be incorporated in the baseline LUT-NN to improve its performance [18]. However, the goal is attaining the bit rates that can be obtained on the raw data by compressing the calibrated data.

As a last remark, we wish to remind that an algorithm suitable for satellite on-board compression, either lossless or near-lossless, should have several favorable characteristics, such as low complexity, low power and storage requirements, capability of working in *Band Interleaved by Line* (BIL) format, and quasi-constant bit rate varying with scenes. However, it should be evaluated on raw, i.e. *uncalibrated* data, as they are produced by the instrument, not on the widespread calibrated data, which instead are used for applications. Therefore, a performance ranking of compression methods may be different on raw data.

References

[1] R. E. Roger and M. C. Cavenor, “Lossless compression of AVIRIS images,” *IEEE Trans. Image Processing*, vol. 5, no. 5, pp. 713–719, May 1996.

[2] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, Englewood Cliffs, NJ, 1984.

[3] B. Aiazzi, L. Alparone, and S. Baronti, “Near-lossless image compression by relaxation-labelled prediction,” *Signal Processing*, vol. 82, no. 11, pp. 1619–1631, Nov. 2002.

[4] B. Aiazzi, L. Alparone, and S. Baronti, “Fuzzy logic-based matching pursuits for lossless predictive coding of still images,” *IEEE Trans. Fuzzy Systems*, vol. 10, no. 4, pp. 473–483, Aug. 2002.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] B. Aiazzi, L. Alparone, S. Baronti, and L. Santurri, “Near-lossless compression of multi/hyperspectral images based on a fuzzy-matching-pursuits interband prediction,” in *Image and Signal Processing for Remote Sensing VII*, S. B. Serpico, Ed., 2002, vol. 4541 of *Proc. SPIE*, pp. 252–263.

[10] X. Wu and N. Memon, “Context-based lossless interband compression–Extending CALIC,” *IEEE Trans. Image Processing*, vol. 9, no. 6, pp. 994–1001, Jun. 2000.

[11] 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.

[12] B. Aiazzi, L. Alparone, S. Baronti, A. Bertoli, C. Lastri, F. Lotti, E. Magli, G. Olmo, and B. Penna, “Advanced methods for onboard lossless compression of hyperspectral data,” in *Mathematics of Data/Image Coding, Compression, and Encryption VI, with Applications*, M. S. Schmalz, Ed., Dec. 2003, vol. 5208 of *Proc. SPIE*.

[13] B. Aiazzi, L. Alparone, S. Baronti, and C. Lastri, “Crisp and fuzzy adaptive spectral predictions for lossless and near-lossless compression of hyperspectral imagery,” *IEEE Geosci. Remote Sensing Lett.*, vol. 4, no. 4, pp. 532–536, Oct. 2007.

[14] J. Mielikainen, “Lossless compression of hyperspectral images using lookup tables,” *IEEE Signal Processing Lett.*, vol. 13, no. 3, pp. 157–160, Mar. 2006.

[15] B. Huang and Y. Sriraja, “Lossless compression of hyperspectral imagery via lookup tables with predictor selection,” in *Image and Signal Processing for Remote Sensing XII*, L. Bruzzone, Ed., 2006, vol. 6365 of Proc. of SPIE, pp. 63650l–1.

[16] P. J. S. G. Ferreira and A. J. Pinho, “Why does histogram packing improve lossless compression rates,” *IEEE Signal Processing Lett.*, vol. 9, no. 8, pp. 259–261, Aug. 2002.

[17] A. J. Pinho, “An on-line pre-processing technique for improving the lossless compression of images with sparse histograms,” *IEEE Signal Processing Lett.*, vol. 9, no. 1, pp. 5–7, Jan. 2002.

[18] J. Mielikainen and P. Toivanen, “Lossless compression of hyperspectral images using a quantized index to lookup tables,” *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 474–478, Jul. 2008.