Article

Multiband and Lossless Compression of Hyperspectral Images

Raffaele Pizzolante and Bruno Carpentieri *

Dipartimento di Informatica, Università di Salerno, Fisciano (SA) 84084, Italy; rpizzolante@unisa.it

* Correspondence: bc@dia.unisa.it; Tel.: +39-89-969500; Fax: +39-89-969600

Abstract: Hyperspectral images are widely used in several real-life applications. In this paper, we investigate on the compression of hyperspectral images by considering different aspects, including the optimization of the computational complexity in order to allow implementations on limited hardware (*i.e.*, hyperspectral sensors, *etc.*). We present an approach that relies on a three-dimensional predictive structure. Our predictive structure, 3D-MBLP, uses one or more previous bands as references to exploit the redundancies among the third dimension. The achieved results are comparable, and often better, with respect to the other state-of-art lossless compression techniques for hyperspectral images.

Keywords: hyperspectral images; lossless compression; low complexity; 3-D data

1. Introduction

Hyperspectral imaging instruments collect information by exploring the electromagnetic spectrum of a specific geographical area. In contrast to the human eye and traditional camera sensors, which can only perceive visible light (*i.e.*, the wavelengths between 360 to 760 nanometers (nm)), spectral imaging techniques allow to cover a significant portion of wavelengths (*i.e.*, the frequencies of ultraviolet and infrared rays). It is important to note that the spectrum is subdivided into different spectral bands. Therefore, hyperspectral images can be viewed as three-dimensional data (often referred as *datacubes*).

For instance, the *Airborne Visible/Infrared Imaging Spectrometer* (*AVIRIS*) [1] hyperspectral sensor (NASA Jet Propulsion Laboratory (JPL) [2]) measures from 380 to 2500 nm of the electromagnetic spectrum. In particular, the spectrum is subdivided into 224 spectral bands.

From the analysis of hyperspectral data, it is possible to identify and/or classify materials, objects, *etc.* Such capabilities are related to the fact that some objects and materials have a unique signature (a sort of *fingerprint*) in the electromagnetic spectrum, therefore this fingerprint can be used for identification purposes.

Hyperspectral data are widely used in real-life applications including agriculture, mineralogy, physics, surveillance, *etc.* For instance, in geological applications the capabilities of hyperspectral remote sensing are exploited to identify various types of minerals or to search for minerals and oil.

One of the most important parameters to evaluate the precision of a sensor is the spectral resolution, which is the width between two adjacent bands. For instance, by considering the AVIRIS hyperspectral images, the spectral resolution is 10 nm. The spatial resolution is a relevant aspect too. Informally, the spatial resolution denotes how extensive is the geographical area mapped by the sensor into a pixel. It could be difficult to recognize materials and/or objects from a pixel, if a too wide area is mapped into it.

Many hundreds of gigabytes can be produced every day by a single hyperspectral sensor. Therefore, it is necessary to compress these data, in order to transmit and to store them efficiently.

Since such data are often used in delicate tasks and there are high costs involved in the acquisitions, lossless compression is generally required.

This paper focuses on a novel technique for the lossless compression of hyperspectral images. The proposed algorithm is based on the predictive coding model and the proposed predictive structure uses a configurable multiband three-dimensional structure. It is possible to customize our predictor by individuating the number of the previous bands which will be used as references and the wideness of the prediction context. Through appropriate configurations of such parameters, the computational complexity and the memory usage can be optimized depending on the hardware available.

Because of its high configurability, our algorithm is suitable for "*on board*" implementations on hardware with limited capabilities, as for example on an airplane or on a satellite.

The experimental results we have achieved are comparable and often better, with respect to other state of the art approaches. Our scheme provides a good trade-off between computational complexity/memory usage and compression performances.

The rest of the paper is organized as follows: Section 2 briefly reviews previous work on lossless and lossy compression of hyperspectral images, Section 3 outlines the proposed lossless compression approach and Section 4 focuses on the description of experimental results. Finally, Section 5 highlights our conclusion and future work directions.

2. Related Works

Lossless compression of hyperspectral images is generally based on the predictive coding model. The predictive-based approaches have different advantages: they use limited resources in terms of computational power and memory usage and achieve good compression performances. Often, these models are suitable for on board implementations.

Spectral-oriented Least SQuares (SLSQ) [3], Linear Predictor (LP) [3], Fast Lossless (FL) [4], CALIC-3D [5], M-CALIC [5], and RLS [6] are among the state-of-art predictive-based techniques.

The *Consultative Committee for Space Data Systems* (*CCSDS*) has specified the *CCSDS 123 standard*, which outlines a method for lossless compression of multispectral and hyperspectral image data and a format for storing the compressed data [7,8]. The main objective is to establish a Recommended Standard for a multispectral and hyperspectral images, and to specify the compressed data format. In literature, many proposed approaches implement the recommendations of the CCSDS 123 standard for the lossless compression of hyperspectral images, as for instance, the ones described in [9–11].

Other approaches are designed for offline compression, since they use more sophisticated techniques and/or they require the complete availability of the hyperspectral image. These approaches are not suitable for an on board implementation but can achieve better compression performances. Mielikainen, in [12], proposed an approach for the compression of hyperspectral image through Look-Up Table (LUT). LUT predicts each pixel by using all the pixels in the current and in the previous band, by searching the nearest neighbor, in the previous band, which has the same pixel value as the pixel located in the same spatial coordinates as the current pixel. LUT has high compression performances, but it uses more resources in terms of memory and CPU usage.

Other lossless techniques are based on dimensionality reduction through principal component transform [13] or they are based on the clustered differential pulse code modulation [14]. An error-resilient lossless compression technique is proposed in [15].

For the lossy compression of hyperspectral images, the compression algorithms are generally based on 3D frequency transforms: as for examples 3-D Discrete Wavelet Transform (3D-DWT) [16], 3-D Discrete Cosine Transform (3D-DCT) [17], Karhunen–Loève transform (KLT) [18], *etc*. These approaches are easily scalable. On the other hand, they must maintain the entire hyperspectral image at the same time in memory. Locally optimal Partitioned Vector Quantization (LPVQ) [19,20] applies a Partitioned Vector Quantization (PVQ) scheme independently to each pixel of the hyperspectral image.

The variable sizes of the partitions are chosen adaptively and the indices are entropy coded. The codebook is included as part of the coded output. This technique can be used also in lossless mode, but the high costs required in terms of CPU and memory do not allow an on board implementation.

3. Lossless Multiband Compression for Hyperspectral Images (LMBHI)

Hyperspectral images present two typologies of correlations:

- *inter-band correlation*;
- *intra-band correlation*.

In particular, contiguous bands are strongly correlated (inter-band correlation) and the pixels are generally correlated, since, for instance, two adjacent pixels map adjacent areas, possibly composed of the same material, *etc.* (intra-band correlation). Such characterizations are exploited by the compression strategies, in order to optimize the redundancy among the third dimension. The main aim of our approach, which we denoted as *Lossless MultiBand compression for Hyperspectral Images* (*LMBHI*), is to exploit the correlation with a predictive coding model.

In detail, for each pixel, *X*, of the input hyperspectral image, LMBHI performs the prediction of the current pixel, $\hat{X}$, by selecting the appropriate prediction context of *X* (three-dimensional or a bi-dimensional contexts).

All the pixels that belong to the first band are predicted by using a bi-dimensional predictive structure: the 2-D Linearized Median Predictor (2-D LMP) [21], which exploits only the intra-band correlation, since the first band has no reference bands. The other pixels are predicted by using a new three-dimensional predictive approach, which uses a prediction context composed of the neighboring pixels of *X* and its reference pixels in the previous bands.

Once the prediction step is computed, the prediction error *e* (defined in Equation (1)) is modeled and coded.

$$e = \lfloor X - \hat{X} \rfloor \tag{1}$$

3.1. Review of the 2-D Linearized Median Predictor (2D-LMP)

The *2-D Linearized Median Predictor* (*2D-LMP*) [21] uses a prediction context that is composed by three neighboring pixels of *X*, namely, I_A, I_B, and I_C, as shown in Figure 1. In particular, the predictive structure is derived from the well-established *2-D Median Predictor*, which is used in JPEG-LS [22]. The 2-D Median Predictor has the following predictive structure outlined in the Equation (2).

$$\hat{X} = \begin{cases} \max\left(I_A, I_B\right) & if \ I_C \geqslant \min\left(I_A, \ I_B\right) \\ \min\left(I_A, I_B\right) & if \ I_C \leqslant \max\left(I_A, \ I_B\right) \\ I_A + I_B - I_C & otherwise \end{cases} \tag{2}$$

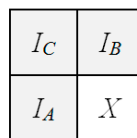| I_C	I_B
I_A	X

Figure 1. The prediction context of the 2D-LMP predictive structure. The gray part is already coded and the white part is not coded yet.

Basically, Median Predictor is in charge of selecting one of the above three options, depending on the context. By combining all the three options, it is possible to obtain the predictive structure of 2D-LMP, defined as in the Equation (3).

$$\hat{X} = \frac{2 \cdot (I_A + I_B)}{3} - \frac{I_C}{3} \tag{3}$$

3.2. 3-D MultiBand Linear Predictor (3D-MBLP)

The Multiband Linear Predictor (3D-MBLP) uses a prediction context by considering two parameters:

- B: number of the previous bands, that are considered for the prediction;
- N: number of the samples for the current and each previous band, which will be used for the creation of the prediction context.

First of all, we define a bi-dimensional enumeration E, graphically represented in Figure 2. The main aim of such an enumeration, is to permit the relative indexing of the pixels with respect to the pixel which is currently under analysis (which has 0 as index in Figure 2).

			16	14		
		11	8	6	9	12
	15	7	3	2	4	10
	13	5	1	0		

Figure 2. The enumeration E we used for the relative indexing with respect to the current pixel, identified with 0 as index.

In order to define the prediction context of 3D-MBLP, we use the following notations:

- $I_{i,j}$: indicates the i-th pixel of the j-th band, according to the enumeration E;
- $I_{0,j}$: denotes the pixel that has the same spatial coordinates of X, of the j-th band, according to the enumeration E.

In the following, we suppose that the current band is the k-th band. In particular, by using our notations, it is possible to observe that X can be also addressed as $I_{0,k}$ In detail, the 3D-MBLP predictor is based on the least squares optimization technique and the prediction is computed by means of the Equation (4).

$$\hat{X} = \sum_{i=1}^{B} \alpha_i \cdot I_{0,\,k-i} \tag{4}$$

The coefficients: $\alpha_0 = [\alpha_1, \ldots, \alpha_B]^T$ are chosen to minimize the energy of the prediction error described by the Equation (5).

$$P = \sum_{i=1}^{N} (I_{i,\,k} - \hat{I}_{i,\,k})^2 \tag{5}$$

P can be rewritten in matrix notation by means of the following equation:

$$P = (C\alpha - X)^T \cdot (C\alpha - X)$$

where

$$C = \begin{bmatrix} I_{1,\,k-1} & \cdots & I_{1,\,k-B} \\ \vdots & \ddots & \vdots \\ I_{N,\,k-1} & \cdots & I_{N,\,k-B} \end{bmatrix}$$

and $X = [I_{1,k}, \ldots, I_{N,k}]^T$.

Subsequently, by taking the derivate of P and by setting it to zero, we obtain the optimal coefficients by means of the Equation (6).

$$\left(C^T C\right) \alpha_0 = \left(C^T X\right) \tag{6}$$

Once the coefficients α_0, which solve the linear system Equation (6), are computed, the prediction, $\hat{X}$, of the current pixel X, can be calculated.

3.3. Modeling and Coding of Prediction Errors

Starting from the consideration that a prediction error can assume positive or negative values. Similarly to [23], we use an invertible mapping function (highlighted in the Equation (7)), in order to have only non-negative values. It is important to note that the mapping function does not alter the redundancy among the errors. For the coding of the mapped prediction errors we use the Arithmetic Coder (AC) scheme.

$$M\left(error\right) = \begin{cases} 2 \cdot |error| & if \ error \geqslant 0 \\ 2 \cdot |error| - 1 & otherwise \end{cases} \tag{7}$$

3.4. Computational Complexity

The main computational costs of our approach are due to the resolution of the linear system Equation (6), used to generate the optimal coefficients, which need the computation of the predicted pixel. By using the normal equation method, the linear system can be solved with $\left(N + \dfrac{B}{3}\right) \cdot B^2$ floating-point operations [24]. Figure 3 shows the trend of the computational complexity of our predictive model, in terms of number of operations (Y-axis) that are required for the solving of the linear system, by using configurations with different parameters (X-axis).
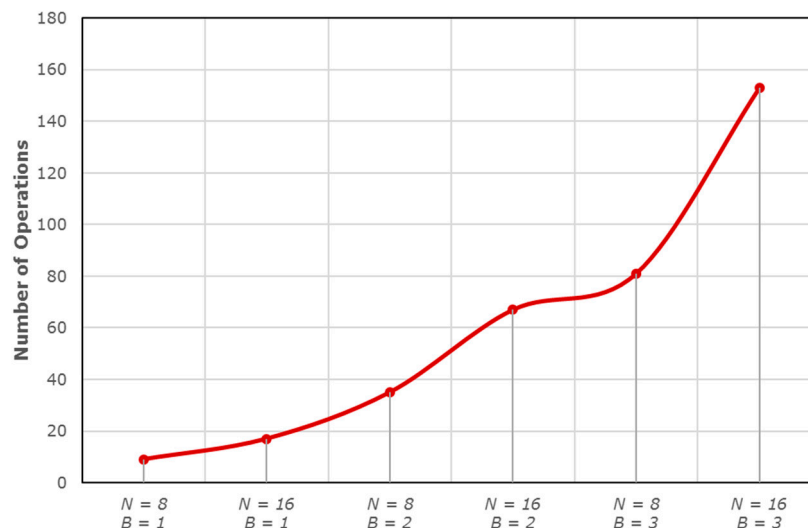

Figure 3. The number of operations (Y-axis) required to solve the linear system Equation (6), by using different parameters (X-axis).

If we use only the previous band as a reference ($B = 1$), only about 20 operations are needed to solve the system. Instead, four or nine times more operations are required if we use two previous bands ($B = 2$) or three previous bands ($B = 3$). A linear system can have three kinds of solutions: no solutions, one solution, and infinity solutions. In the first and the third scenarios, the proposed predictive structure cannot perform the prediction. In these scenarios, it is desirable to use another

low-complexity predictive structure and we have used the *3-D Distances-based Linearized Median Predictor* (*3D-DLMP*) [21].

4. Experimental Results

We performed experiments on two datasets of AVIRIS hyperspectral images: the *1997 AVIRIS Dataset* (Section 4.1) and the *CCSDS Dataset* (Section 4.2).

In our experiments we considered also the PAQ8 algorithm (described in [25]) for the coding of the prediction error. PAQ8 is a state-of-the-art lossless compression algorithm that belongs to the PAQ family of compression algorithm. It is important to note that the PAQ8 family is strictly related to the well-established *Prediction by Partial Matching scheme* (*PPM*) [26]. In general, the PAQ8 algorithm achieve a high degree of compression performances, but the PAQ8 scheme has significant computational complexity. Therefore, such scheme is not fully adequate to be used for on board applications.

The experiments are performed by using a non-optimized *Java-based proof-of-concept* of our approach, which employs few minutes on a medium-end laptop (equipped with an Intel Core i5 4200 M processor and 8 GB of RAM).

4.1. 1997 AVIRIS Dataset

Each one of AVIRIS hyperspectral image of the AVIRIS '97 dataset is subdivided into scenes (the number of scenes is highlighted in Table 1). It is important to note that each scene has 614 columns, 512 lines, and 224 spectral bands. In addition, each pixel is stored by using 16 bits.

Table 1. Description of the dataset used.

Images	Number of Scenes
Lunar Lake	3
Moffett Field	4
Jasper Ridge	6
Cuprite	5
Low Altitude	8

In Tables 2 and 3 we report the results achieved by using $B = 1$ with $N = 8$ and $N = 16$, respectively. Subsequently, in Tables 4 and 5 we report the results achieved, by using $B = 2$. Finally, Tables 6 and 7 report the results achieved by using $B = 3$ and the N parameter equal to 8 and equal to 16, respectively. All the results are reported in terms of *Bits Per Sample* (*BPS*). In each table we report the results achieved by using both the AC and the PAQ8 schemes for the coding of the prediction errors.

Table 2. Achieved results by using the following parameters: $N = 8$, $B = 1$. (*N.P.* indicates that the scene is not present).

Scenes	Lunar Lake		Moffett Field		Jasper Ridge		Cuprite		Low Altitude	
	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8
Scene 01	5.0560	4.9985	5.1463	5.0649	5.0602	4.9801	4.9699	4.9124	5.3784	5.3101
Scene 02	5.0099	4.9532	5.1023	4.9918	5.0524	4.9707	5.0457	4.9916	5.4015	5.3320
Scene 03	4.9963	4.9381	4.9672	4.8426	5.0987	5.0150	4.9930	4.9369	5.3066	5.2429
Scene 04	N.P.		5.1888	5.0984	5.1160	5.0353	5.0380	4.9838	5.3341	5.2687
Scene 05	N.P.		N.P.		5.0532	4.9746	5.0358	4.9821	5.3810	5.3151
Scene 06	N.P.		N.P.		5.0525	4.9726	N.P.		5.3145	5.2536
Scene 07	N.P.		N.P.		N.P.		N.P.		5.3110	5.2511
Scene 08	N.P.		N.P.		N.P.		N.P.		5.3268	5.2712
Average	5.0207	4.9633	5.1012	4.9994	5.0722	4.9914	5.0165	4.9614	5.3442	5.2806

Table 3. Achieved results by using the following parameters: $N = 16$, $B = 1$.

Scenes	Lunar Lake		Moffett Field		Jasper Ridge		Cuprite		Low Altitude	
	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8
Scene 01	5.0212	4.9665	5.1407	5.0504	5.0547	4.9667	4.9334	4.8791	5.3658	5.2917
Scene 02	4.9767	4.9228	5.1083	4.9796	5.0485	4.9576	5.0194	4.9663	5.3906	5.3151
Scene 03	4.9645	4.9094	4.9654	4.8221	5.0952	5.0020	4.9579	4.9049	5.2892	5.2206
Scene 04	*N.P.*		5.1843	5.0837	5.1158	5.0255	5.0064	4.9540	5.3158	5.2456
Scene 05	*N.P.*		*N.P.*		5.0488	4.9622	5.0002	4.9491	5.3701	5.2981
Scene 06	*N.P.*		*N.P.*		5.0463	4.9604	*N.P.*		5.2963	5.2321
Scene 07	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.2936	5.2303
Scene 08	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.3131	5.2544
Average	4.9875	4.9329	5.0997	4.9840	5.0682	4.9791	4.9835	4.9307	5.3293	5.2610

Table 4. Achieved results by using the following parameters: $N = 8$, $B = 2$.

Scenes	Lunar Lake		Moffett Field		Jasper Ridge		Cuprite		Low Altitude	
	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8
Scene 01	5.0141	4.9707	5.0934	5.0462	5.0140	4.9658	4.9364	4.8929	5.3295	5.2879
Scene 02	4.9689	4.9259	5.0172	4.9525	5.0072	4.9583	5.0326	4.9918	5.3491	5.3065
Scene 03	4.9558	4.9111	4.8890	4.8064	5.0501	5.0005	4.9625	4.9204	5.2648	5.2221
Scene 04	*N.P.*		5.1254	5.0715	5.0654	5.0200	5.0105	4.9693	5.2881	5.2451
Scene 05	*N.P.*		*N.P.*		5.0095	4.9624	5.0012	4.9608	5.3281	5.2866
Scene 06	*N.P.*		*N.P.*		5.0162	4.9639	*N.P.*		5.2707	5.2297
Scene 07	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.2732	5.2326
Scene 08	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.2927	5.2552
Average	4.9796	4.9359	5.0313	4.9692	5.0271	4.9785	4.9886	4.9470	5.2995	5.2582

Table 5. Achieved results by using the following parameters: $N = 16$, $B = 2$.

Scenes	Lunar Lake		Moffett Field		Jasper Ridge		Cuprite		Low Altitude	
	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8
Scene 01	4.9175	4.8732	5.0206	4.9703	4.9397	4.8884	4.8387	4.7945	5.2482	5.2058
Scene 02	4.8743	4.8303	4.9513	4.8806	4.9341	4.8817	4.9483	4.9061	5.2689	5.2254
Scene 03	4.8632	4.8174	4.8160	4.7244	4.9755	4.9227	4.8679	4.8250	5.1799	5.1359
Scene 04	*N.P.*		5.0496	4.9929	4.9940	4.9458	4.9194	4.8771	5.2017	5.1572
Scene 05	*N.P.*		*N.P.*		4.9358	4.8859	4.9047	4.8638	5.2491	5.2066
Scene 06	*N.P.*		*N.P.*		4.9418	4.8863	*N.P.*		5.1859	5.1444
Scene 07	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.1883	5.1470
Scene 08	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.2108	5.1728
Average	4.8850	4.8403	4.9594	4.8921	4.9535	4.9018	4.8958	4.8533	5.2166	5.1744

Table 6. Achieved results by using the following parameters: $N = 8$, $B = 3$.

Scenes	Lunar Lake		Moffett Field		Jasper Ridge		Cuprite		Low Altitude	
	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8
Scene 01	5.1060	5.0699	5.1815	5.1415	5.1037	5.0640	5.0324	4.9965	5.4238	5.3862
Scene 02	5.0621	5.0265	5.0919	5.0362	5.0968	5.0563	5.1363	5.1018	5.4414	5.4026
Scene 03	5.0479	5.0112	4.9649	4.8932	5.1374	5.0962	5.0618	5.0269	5.3610	5.3221
Scene 04	*N.P.*		5.2085	5.1614	5.1523	5.1148	5.1084	5.0737	5.3831	5.3446
Scene 05	*N.P.*		*N.P.*		5.0999	5.0611	5.0963	5.0631	5.4202	5.3825
Scene 06	*N.P.*		*N.P.*		5.1138	5.0697	*N.P.*		5.3656	5.3285
Scene 07	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.3707	5.3341
Scene 08	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.3918	5.3581
Average	5.1117	5.0359	5.1173	5.0581	5.0870	5.0770	5.3947	5.0524	5.3947	5.3573

Table 7. Achieved results by using the following parameters: $N = 16$, $B = 3$.

Scenes	Lunar Lake		Moffett Field		Jasper Ridge		Cuprite		Low Altitude	
	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8	AC	PAQ8
Scene 01	4.9243	4.8842	5.0206	4.9758	4.9413	4.8968	4.8494	4.8094	5.2546	5.2153
Scene 02	4.8822	4.8422	4.9383	4.8754	4.9353	4.8898	4.9667	4.9282	5.2731	5.2327
Scene 03	4.8704	4.8291	4.8052	4.7226	4.9739	4.9279	4.8823	4.8434	5.1887	5.1476
Scene 04	*N.P.*		5.0445	4.9931	4.9919	4.9501	4.9320	4.8932	5.2093	5.1681
Scene 05	*N.P.*		*N.P.*		4.9379	4.8944	4.9149	4.8776	5.2528	5.2133
Scene 06	*N.P.*		*N.P.*		4.9510	4.9012	*N.P.*		5.1935	5.1545
Scene 07	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.1980	5.1594
Scene 08	*N.P.*		*N.P.*		*N.P.*		*N.P.*		5.2214	5.1854
Average	4.8923	4.8518	4.9522	4.8917	4.9552	4.9100	4.9091	4.8704	5.2239	5.1845

In Tables 8 and 9 the average results on all the tested hyperspectral images are reported. In detail, the first column indicates the N parameter and from the second to the fourth columns the average results for $B = 1$, $B = 2$, and $B = 3$, respectively.

Table 8. Average Results on the 1997 AVIRIS Images (AC).

N	$B = 1$	$B = 2$	$B = 3$
8	5.1110	5.0652	5.2211
16	5.0936	4.9821	4.9865

Table 9. Average Results on the 1997 AVIRIS Images (PAQ8).

N	$B = 1$	$B = 2$	$B = 3$
8	5.0392	5.0178	5.1161
16	5.0175	4.9324	4.9417

As it is possible to observe from Figures 4 and 5 which graphically represent the average results, the best results are achieved when the following parameters are used: $N = 16$ and $B = 2$ (Figures 4b and 5b). The worst results are obtained by using the following parameters: $N = 8$ and $B = 3$.
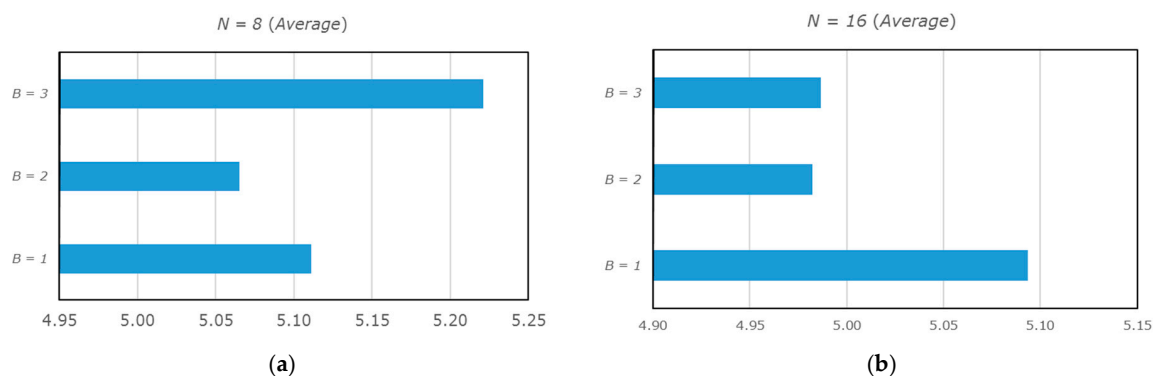
Figure 4. Graphical representation of the average results by using the AC scheme for the coding of the prediction errors. $N = 8$ (**a**) and $N = 16$ (**b**).
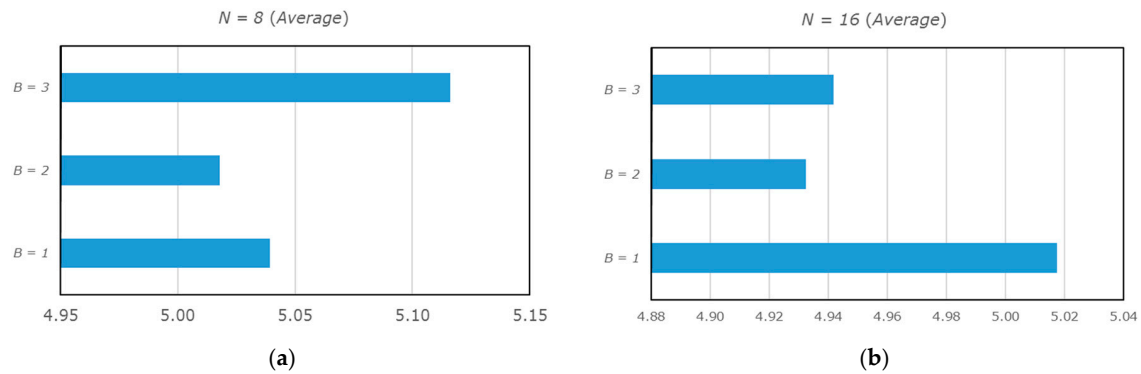
Figure 5. Graphical representation of the average results by using the PAQ8 scheme for the coding of the prediction errors. $N = 8$ (**a**) and $N = 16$ (**b**).

Comparison with other Approaches

In order to compare the experimental results achieved by our approach, we consider the Compression Ratio (C.R.) as a measure for the compression performances. In detail, Table 10 reports the results achieved by considering several parameters on all the hyperspectral images of the used dataset. More precisely, the results are reported in terms of C.R. and they are compared with other state of the art lossless compression schemes.

Table 10. Compression results, in terms of compression ratio (C.R.) achieved by LMBHI (by using various parameter configurations), compared to other lossless compression methods.

Methods/Images	Lunar Lake	Moffett Field	Jasper Ridge	Cuprite	Low Altitude	Average
3D-MBLP + PAQ8 ($N = 16, B = 2$)	3.31	3.27	3.26	3.30	3.09	*3.25*
3D-MBLP + PAQ8 ($N = 8, B = 2$)	3.24	3.22	3.21	3.23	3.04	*3.19*
3D-MBLP + PAQ8 ($N = 8, B = 1$)	3.22	3.20	3.21	3.23	3.03	*3.18*
3D-MBLP + AC ($N = 16, B = 2$)	3.27	3.23	3.23	3.27	3.07	*3.21*
3D-MBLP + AC ($N = 8, B = 2$)	3.21	3.18	3.18	3.21	3.02	*3.16*
3D-MBLP + AC ($N = 8, B = 1$)	3.18	3.14	3.16	3.19	2.99	*3.13*
LPVQ	3.31	3.01	3.12	3.27	2.97	*3.14*
SLSQ	3.15	3.14	3.15	3.15	2.98	*3.11*
JPEG-2000	2.98	2.99	2.96	2.98	2.82	*2.95*
LP	3.05	2.88	2.94	3.03	2.76	*2.93*
JPEG-LS	2.87	2.90	2.87	2.87	2.74	*2.85*
Diff. JPEG2000	2.94	2.83	2.82	2.92	2.69	*2.84*
Diff. JPEG-LS	2.93	2.84	2.81	2.91	2.7	*2.84*
M-CALIC	3.19	3.27	3.06	3.14	N.D.	*N.D.*
CALIC-3D	3.06	3.08	3.09	3.25	N.D.	*N.D.*
LUT	3.44	3.23	3.40	3.17	N.D.	*N.D.*

From the experimental results, it should be observed that LMBHI gets its best results by using two previous bands as references (*i.e.* when the following parameter is used: $B = 2$), LMBHI outperforms, in average, all the other state of the art approaches.

On the other hand, when only the previous band is used (*i.e.*, when $B = 1$), LMBHI outperforms all the compared state of the art techniques, with the exception of LPVQ. But, LPVQ is not suited for on board implementation.

In this latter case, our approach achieves better results with respect to LPVQ on three of the five hyperspectral images: Moffett Field, Jasper Ridge, and Low Altitude, but LPVQ gains on Cuprite and especially on Lunar Lake. In addition, LUT obtains better results of our approach on two of four compared hyperspectral images: Lunar Lake and Jasper Ridge.

The high flexibility and adaptability of our approach makes it considerable for on board implementations. In fact, the coding parameters can be customized depending on the hardware available.

4.2. CCSDS Dataset

In this section we focus on the experimental results we have achieved by considering the CCSDS Dataset, which is composed by five calibrated and seven uncalibrated hyperspectral images. This dataset is provided by Consultative Committee for Space Data Systems (CCSDS) Multispectral and Hyperspectral Data Compression [27].

In Table 11, we shortly describe the dataset by reporting the number of scenes (second column) and the number of samples per line (third column) for the calibrated and the uncalibrated images (first column). The samples of the calibrated and the uncalibrated images are stored by using 16 bits (16-bit signed integer for the calibrated and 16-bit unsigned for the uncalibrated), except for the *Hawaii* and *Maine* images in which the samples are stored by using 12 bits (unsigned) [27]. Each image is composed by 512 lines.

Table 11. Description of the CCSDS dataset.

Images	Num. of Scenes (*Denotation of the Scenes*)	Samples
Calibrated		
Yellowstone	5 scenes (*3, 10, 11, 18*)	677
Uncalibrated		
Yellowstone	5 scenes (*3, 10, 11, 18*)	680
Hawaii	1 scene (*1*)	614
Maine	1 scene (*10*)	680

In Table 12, we report our results in terms of bits-per sample (BPS). The results refer to the calibrated hyperspectral images (first column), by using several configurations for our approach (columns from the second to the fourth). Analogously to Table 12, in Table 13 we report our experimental results for the uncalibrated images. In each table, we report the results by using the AC and the PAQ8 schemes.

The best results are achieved when the following configuration is used: $B = 2$ and $N = 16$.

Table 12. Achieved results for the calibrated images of the CCSDS dataset.

Images/Configurations	$N = 8$, $B = 1$		$N = 8$, $B = 2$		$N = 16$, $B = 2$	
	AC	PAQ8	AC	PAQ8	AC	PAQ8
Yellowstone 0	4.1881	4.0898	4.0435	3.9917	3.9783	3.9198
Yellowstone 3	4.0831	3.9658	3.9449	3.8701	3.8795	3.7972
Yellowstone 10	3.5470	3.3974	3.4712	3.3488	3.3925	3.2592
Yellowstone 11	3.7878	3.7128	3.7035	3.6554	3.6311	3.5751
Yellowstone 18	4.1495	4.0350	3.9773	3.9154	3.9082	3.8403
Average	3.9511	3.8401	3.8280	3.7563	3.7579	3.6783

Table 13. Achieved results for the uncalibrated images of the CCSDS dataset.

Images/Configurations	N = 8 , B = 1		N = 8 , B = 2		N = 16 , B = 2	
	Calibrated (16-bit)					
	AC	PAQ8	AC	PAQ8	AC	PAQ8
Yellowstone 0	6.7889	6.6034	6.4564	6.4012	6.4085	6.3411
Yellowstone 3	6.6854	6.4961	6.3553	6.2757	6.3106	6.2183
Yellowstone 10	6.0635	5.8620	5.8397	5.7096	5.7754	5.6309
Yellowstone 11	6.3023	6.1701	6.0889	6.0418	6.0329	5.9723
Yellowstone 18	6.7909	6.5840	6.4235	6.3504	6.3715	6.2865
Average	6.5262	6.3431	6.2328	6.1557	6.1798	6.0898
	Calibrated (12-bit)					
Hawaii	2.9533	2.9041	2.8748	2.8434	2.7965	2.7616
Maine	3.0746	2.9835	2.9528	2.9030	2.9005	2.8413
Average	3.0140	2.9438	2.9138	2.8732	2.8485	2.8015

Comparison with Other Approaches

We have compared our results on the CCSDS dataset with other state-of-art approaches. Table 14 reports the results achieved by considering several values for the B and N parameters on the calibrated hyperspectral images of the CCSDS dataset, by using the AC scheme as well as the PAQ8 scheme for the coding of the prediction errors.

Table 14. Comparison with other lossless compression methods (calibrated images). The results are reported in bits-per-sample (BPS).

Methods/Scenes (Yellowstone Calibrated)	0	3	10	11	18	Average
3D-MBLP + PAQ8 ($N = 16, B = 2$)	3.92	3.80	3.26	3.58	3.84	*3.68*
3D-MBLP + PAQ8 ($N = 8, B = 2$)	3.99	3.87	3.35	3.66	3.92	*3.76*
3D-MBLP + PAQ8 ($N = 8, B = 1$)	4.09	3.97	3.40	3.71	4.04	*3.84*
3D-MBLP + AC ($N = 16, B = 2$)	3.98	3.88	3.39	3.63	3.91	*3.76*
3D-MBLP + AC ($N = 8, B = 2$)	4.04	3.95	3.47	3.70	3.98	*3.83*
3D-MBLP + AC ($N = 8, B = 1$)	4.18	4.08	3.55	3.79	4.15	*3.95*
FL	3.96	3.83	3.40	3.63	3.94	*3.75*
FL#	3.91	3.79	3.37	3.59	3.90	*3.71*
LUT#	4.82	4.62	3.96	4.34	4.84	*4.52*
LAIS-LUT#	4.48	4.31	3.71	4.02	4.48	*4.20*
TSP-W1	3.94	3.81	3.37	3.60	3.92	*3.73*
TSP-W2	3.99	3.86	3.42	3.67	3.97	*3.78*

Tables 15 and 16 report the comparison between the proposed approach and other approaches for the 16-bit uncalibrated and 12-bit uncalibrated hyperspectral images of the CCSDS dataset.

From Tables 12 and 13 it comes clear that the best results are achieved when the value of N is equal to 16 and B is equal to 2.

By looking at Table 14, it should be observed that our approach, when the following configuration is used: $N = 16$ and $B = 2$, achieves results that are comparable but slightly worse with respect to FL and FL# [27]. Our approach outperforms all the other techniques when the PAQ8 scheme is used for the coding of prediction errors with $N = 16$ and $B = 2$. However, in such configuration the computational complexity of our approach is not suitable for on board implementations.

Table 15. Comparison with other lossless compression methods (16-bit uncalibrated images). The results are reported in bits-per-sample (BPS).

Methods/Scenes (Yellowstone Uncalibrated)	0	3	10	11	18	Average
3D-MBLP + PAQ8 *(N = 16, B = 2)*	6.34	6.22	5.63	5.97	6.29	*6.09*
3D-MBLP + PAQ8 *(N = 8, B = 2)*	6.40	6.28	5.71	6.04	6.35	*6.16*
3D-MBLP + PAQ8 *(N = 8, B = 1)*	6.60	6.50	5.86	6.17	6.58	6.34
3D-MBLP + AC *(N = 16, B = 2)*	6.41	6.31	5.78	6.03	6.37	6.18
3D-MBLP + AC *(N = 8, B = 2)*	6.46	6.36	5.84	6.09	6.42	6.23
3D-MBLP + AC *(N = 8, B = 1)*	6.79	6.69	6.06	6.30	6.79	6.53
FL	6.23	6.10	5.65	5.86	6.32	*6.03*
FL#	6.20	6.07	5.60	5.81	6.26	*5.99*
LUT#	7.14	6.91	6.26	6.69	7.20	*6.84*
LAIS-LUT#	6.78	6.60	6.00	6.30	6.82	*6.50*
TSP-W1	6.23	6.09	5.59	5.83	6.28	*6.01*
TSP-W2	6.27	6.13	5.64	5.88	6.32	*6.05*

Table 16. Comparison with other lossless compression methods (12-bit uncalibrated images). The results are reported in bits-per-sample (BPS).

Methods/Scenes (Yellowstone Uncalibrated)	Hawaii	Maine	Average
3D-MBLP + PAQ8 *(N = 16, B = 2)*	2.76	2.84	2.80
3D-MBLP + PAQ8 *(N = 8, B = 2)*	2.84	2.90	*2.87*
3D-MBLP + PAQ8 *(N = 8, B = 1)*	2.90	2.98	*2.94*
3D-MBLP + AC *(N = 16, B = 2)*	2.80	2.90	*2.85*
3D-MBLP + AC *(N = 8, B = 2)*	2.87	2.95	*2.91*
3D-MBLP + AC *(N = 8, B = 1)*	2.95	3.07	*3.01*
FL	2.64	2.72	*2.68*
FL#	2.58	2.68	*2.63*
LUT#	3.26	3.45	*3.35*
LAIS-LUT#	3.05	3.19	*3.12*
TSP-W1	2.61	2.71	*2.66*
TSP-W2	2.62	2.74	*2.68*

5. Conclusions and Future Works

In this paper, we have investigated on the lossless compression of hyperspectral images by introducing a multiband three-dimensional predictive structure, we named as 3D-MBLP.

Because of its configurability, it is possible to implement the algorithm on different typologies of sensors, by using appropriate configuration for each type of sensors. Moreover, the proposed approach can be also easily scaled for future generation sensors which will have better hardware capabilities. The experimental results we achieved are comparable and often outperform the other state of the art lossless compression techniques.

In future works, we will include a pre-processing stage before the compression of the hyperspectral image, which substantially reorders the bands by considering their correlation. This will possibly improve the compression performance as in [28,29].

References

1. AVIRIS NASA Page. Available online: http://aviris.jpl.nasa.gov/ (accessed on 1 December 2015).
2. Jet Propulsion Laboratory (JPL) Page. Available online: http://www.jpl.nasa.gov/ (accessed on 1 December 2015).
3. Rizzo, F.; Carpentieri, B.; Motta, G.; Storer, J.A. Low-complexity lossless compression of hyperspectral imagery via linear prediction. *IEEE Signal Process. Lett.* **2005**, *12*, 138–141. [CrossRef]
4. Klimesh, M. Low-complexity lossless compression of hyperspectral imagery via adaptive filtering. *IPN Prog. Report* **2005**, *42*, 1–10.
5. Magli, E.; Olmo, G.; Quacchio, E. Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC. *Geosci. Remote Sens. Lett.* **2004**, *1*, 21–25. [CrossRef]
6. Song, J.; Zhang, Z.; Chen, X. Lossless compression of hyperspectral imagery via RLS filter. *Electron. Lett.* **2013**, *49*, 992–994. [CrossRef]
7. Shah, D.; Bera, K.; Sanjay, J. Software Implementation of CCSDS Recommended Hyperspectral Lossless Image Compression. *Int. J. Image Gr. Signal Process.* **2015**, *4*, 35–41. [CrossRef]
8. Consultative Committee for Space Data Systems (CCSDS), Lossless Multispectral & Hyperspectral Image, Lossless Multispectral & Hyperspectral Image Compression. Available online: http://public.ccsds.org/publications/archive/123x0b1ec1.pdf (accessed on 5 February 2016).
9. Snchez, J.E.; Auge, E.; Santal, J.; Blanes, I.; Serra-Sagrist, J.; Kiely, A.B. Review and implementation of the emerging CCSDS recommended standard for multispectral and hyperspectral lossless image coding. In Proceedings of the 2011 First International Conference on Data Compression, Communications and Processing (CCP), Palinuro, Italy, 21–24 June 2011; pp. 222–228.
10. Keymeulen, D.; Aranki, N.; Hopson, B.; Kiely, A.; Klimesh, M.; Benkrid, K. GPU lossless hyperspectral data compression system for space applications. In Proceedings of the Aerospace Conference, Big Sky, MT, USA, 3–10 March 2012; pp. 1–9.
11. Keymeulen, D.; Aranki, N.; Bakhshi, A.; Luong, H.; Sarture, C.; Dolman, D. Airborne demonstration of FPGA implementation of Fast Lossless hyperspectral data compression system. In Proceedings of the NASA/ESA Conference on Adaptive Hardware and Systems (AHS), Leicester, UK, 14–17 July 2014; pp. 278–284.
12. Mielikainen, J. Lossless compression of hyperspectral images using lookup tables. *IEEE Signal Process. Lett.* **2006**, *13*, 157–160. [CrossRef]
13. Pickering, M.; Ryan, M. Efficient spatial-spectral compression of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1536–1539. [CrossRef]
14. Wu, J.; Kong, W.; Mielikainen, J.; Huang, B. Lossless Compression of Hyperspectral Imagery via Clustered Differential Pulse Code Modulation with Removal of Local Spectral Outliers. *IEEE Signal Process. Lett.* **2015**, *22*, 2194–2198. [CrossRef]
15. Abrando, A.; Barni, M.; Magli, E.; Nencini, F. Error-Resilient and low-complexity on-board lossless compression of hyperspectral images by means of distributed source coding. *IEEE Trans. Geosci.* **2010**, *48*, 1892–1904. [CrossRef]
16. Lim, S.; Sohn, K.; Lee, C. Compression for hyperspectral images using three dimensional wavelet transform. In Proceedings of the IGARSS, Sydney, Australia, 9–13 July 2001; pp. 109–111.
17. Markman, D.; Malah, D. Hyperspectral image coding using 3D transforms. In Proceedings of the IEEE ICIP, Thessaloniki, Greece, 7–10 October 2001; pp. 114–117.
18. Penna, B.; Tillo, T.; Magli, E.; Olmo, G. Transform coding techniques for lossy hyperspectral data compression. *IEEE Trans. Geosci.* **2007**, *45*, 1408–1421. [CrossRef]

19. Carpentieri, B.; Storer, J.A.; Motta, G.; Rizzo, F. Compression of hyperspectral imagery. In Proceedings of the IEEE Data Compression Conference (DCC 03), Snowbird, UT, USA, 25–27 March 2003; pp. 317–324.

20. Motta, G.; Rizzo, F.; Storer, J.A. *Hyperspectral Data Compression*; Springer Science: Berlin, Germany, 2006.

21. Pizzolante, R.; Carpentieri, B. Lossless, low-complexity, compression of three-dimensional volumetric medical images via linear prediction. In Proceedings of the 18th International Conference on Digital Signal. Processing (DSP), Fira, Greece, 1–3 July 2013; pp. 1–6.

22. Carpentieri, B.; Weinberger, M.; Seroussi, G. Lossless compression of continuous tone images. *Proc. IEEE* **2000**, *88*, 1797–1809. [CrossRef]

23. Motta, G.; Storer, J.A.; Carpentieri, B. Lossless image coding via adaptive linear prediction and classifications. *Proc. IEEE* **2000**, *88*, 1790–1796. [CrossRef]

24. Golub, G.H.; van Loan, C.F. *Matrix Computations*; The Johns Hopkins University Press: Baltimore, MD, USA, 1996.

25. Knoll, B.; de Freitas, N. A Machine Learning Perspective on Predictive Coding with PAQ8. In Proceedings of the Data Compression Conference (DCC), Snowbird, UT, USA, 10–12 April 2012; pp. 377–386.

26. Salomon, D.; Motta, G. *Handbook of Data Compression*; Springer: Berlin, Germany, 2010.

27. Kiely, A.B.; Klimesh, M. Exploiting calibration-induced artifacts in lossless compression of hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2672–2678. [CrossRef]

28. Carpentieri, B. Hyperpectral images: Compression, visualization and band ordering. *Proc. IPCV* **2011**, *2*, 1023–1029.

29. Pizzolante, R.; Carpentieri, B. Visualization, band ordering and compression of hyperspectral images. *Algorithms* **2012**, *5*, 76–97. [CrossRef]