Multilevel Split Regression Wavelet Analysis for Lossless Compression of Remote Sensing Data

Sara Álvarez-Cortés⬤, Joan Bartrina-Rapesta⬤, and Joan Serra-Sagristà⬤, *Senior Member, IEEE*

Abstract—Spectral redundancy is a key element to be exploited in compression of remote sensing data. Combined with an entropy encoder, it can achieve competitive lossless coding performance. One of the latest techniques to decorrelate the spectral signal is the regression wavelet analysis (RWA). RWA applies a wavelet transform in the spectral domain and estimates the detail coefficients through the approximation coefficients using linear regression. RWA was originally coupled with JPEG 2000. This letter introduces a novel coding approach, where RWA is coupled with the predictor of CCSDS-123.0-B-1 standard and a lightweight contextual arithmetic coder. In addition, we also propose a smart strategy to select the number of RWA decomposition levels that maximize the coding performance. Experimental results indicate that, on average, the obtained coding gains vary between 0.1 and 1.35 bits-per-pixel-per-component compared with the other state-of-the-art coding techniques.

Index Terms—Lossless coding, predictive coding, spectral decorrelation.

I. INTRODUCTION

THE data acquired by the onboard remote sensing sensors is an invaluable tool for governments, rescue teams, and aid organizations to manage infrastructure and natural resources, to appraise climate changes, or to give support when natural disasters strike. The data produced by these sensors are increasing unprecedently in each mission and, therefore, low-complexity and high-performing lossless compression techniques are of paramount importance.

Lossless coding techniques exploit the redundancy in the spatial and spectral dimensions of the scenes, allowing perfect reconstruction. Several transforms, such as discrete wavelet transform (DWT), principal component analysis (PCA) [1], pairwise orthogonal transform (POT) [2] or, the most recent, regression wavelet analysis (RWA) [3], can be employed to take advantage of the spectral redundancy. RWA is a low-complexity transform that reduces the redundancy still remaining after the computation of a wavelet transform along the spectral dimension. In [3], it is shown that RWA followed

The authors are with the Department of Information and Communications Engineering, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain (e-mail: salvarez@deic.uab.cat; joan.bartrina@uab.cat; joan.serra@uab.cat).

by JPEG 2000 [4] (RWA + JPEG 2000) usually achieves higher compression ratios than the widespread state-of-the-art lossless coding techniques, such as PCA + JPEG 2000, POT + JPEG 2000, M-CALIC [5], and CCSDS-123.0-B-1 [6]. In addition, in [7], the performance of RWA + JPEG 2000 at several decomposition levels for different regression models and variants is evaluated, concluding that when using all the regressors and the highest number of decomposition levels, RWA gives rise to the best coding gain on the remote sensing data.

Although JPEG 2000 can achieve high lossless compression ratios, it is too computationally demanding for the remote sensing sensors. In this regard, the consultive committee for space data systems (CCSDS) published in 2012, the CCSDS-123.0-B-1 standard (CCSDS-123 in what follows), intended for lossless coding of multispectral and hyperspectral data and based on prediction and two different entropy encoders. Spatial and spectral redundancy is exploited through an adaptive linear prediction method that makes use of its nearby samples within a small 3-D neighborhood. The prediction is sequentially conducted in a single pass, entailing minor computational complexity.

Despite the affordable computational complexity of CCSDS-123, it allows to achieve high lossless compression ratios. Lately, several contributions [8]–[11] have been presented, which surpass its coding performance at the expense of a higher computational cost. In particular, [10] and [11] use a conventional recursive least-squares technique to predict the current sample and compute the residual, and then, the residual is entropy encoded by an arithmetic encoder. Thus, both contributions are a tradeoff between the coding performance and the computational complexity. In 2017, Bartrina-Rapesta *et al.* [12] presented a compression technique based on the CCSDS-123 predictor and followed by a low-complexity arithmetic coder (CCSDS-123-AC) using a novel and simple context model that only assesses causal adjacent samples and inexpensive low-cost bitwise operations for the symbol probability estimations; on average, it improves CCSDS-123 and M-CALIC by 0.1 and 0.86 bits-per-pixel-per-component (bpppc), respectively.

This letter presents a novel compression technique based on RWA that obtains higher compression ratios with lower computational cost than the original [3]. In our proposal, the computational cost is reduced by employing a smart strategy that keeps applying further RWA decomposition levels only if it provides coding gain. After the smart-RWA is

conducted, the decorrelated signal is encoded with the low complexity CCSDS-123-AC.

This letter is organized as follows: Section II reviews the RWA transform. Section III introduces our proposed coding technique. Section IV provides experimental results and discussion of the achieved results as compared to the state-of-the-art techniques for data captured by different hyperspectral sensors. Section V brings forward our conclusions.

II. REGRESSION WAVELET ANALYSIS OVERVIEW

RWA [3] is composed by two sequential operations: a DWT—commonly the S-Transform—and a prediction stage—performed through a linear regression.

A. Discrete Wavelet Transforms

The DWT operation comprises a pyramidal multiresolution decomposition along the 1-D spectral dimension. Considering a multicomponent scene with z spectral components and m spatial samples per component, let $\mathbf{V}^0 \in \mathbb{R}^{m \times z}$ be

$$\mathbf{V}^0 = [V^0(1), \dots, V^0(z)], \quad \text{with } V^0(i) = V_i^0 \in \mathbb{R}^{m \times 1}$$

where $i = \{1, \dots, z \cdot 2^{-j}\}$ and $j = \{1, \dots, J\}$, $J = \lceil \log_2(z) \rceil$. The original signal $\mathbf{V}^0$ is decomposed into the half-resolution approximation $\mathbf{V}^1$ and the half-resolution detail $\mathbf{W}^1$ signals. This process is repeated for $\mathbf{V}^{j-1}$, yielding $\mathbf{V}^j \in \mathbb{R}^{m \times (z \cdot 2^{-j})}$ and $\mathbf{W}^j \in \mathbb{R}^{m \times (z \cdot 2^{-j})}$ signals. $\mathbf{V}^j$ contains a reduced version of $\mathbf{V}^{j-1}$ information and $\mathbf{W}^j$ retains the detail information, namely, the difference between $\mathbf{V}^{j-1}$ and $\mathbf{V}^j$.

The one-level DWT decomposition of each $\mathbf{V}^{j-1}$ can be expressed as follows:

$$(\text{DWT})(\mathbf{V}^{j-1}, 1) = (\mathbf{V}^j, \mathbf{W}^j). \tag{1}$$

This decomposition is recursively performed in a pyramid form over $\mathbf{V}^j$, while the detail components $\mathbf{W}^{1 \leq j}$ remain unchanged. By induction, from $\mathbf{V}^0$, the wavelet representation with J levels is described as follows:

$$(\text{DWT})(\mathbf{V}^0, J) = (\mathbf{V}^J, (\mathbf{W}^j)^{1 \leq j \leq J}). \tag{2}$$

B. Ordinary Least-Squares Method

The second sequential operation of the RWA consists of a light linear regression aimed to remove the spectral redundancy that still remains in $\mathbf{W}^{1 \leq j \leq J}$.

At each level j, the regression model estimates each detail component $(\widehat{\mathbf{W}}_i^j \in \mathbb{R}^{m \times (z \cdot 2^{-j})})$ from $\mathbf{V}^j \in \mathbb{R}^{m \times (z \cdot 2^{-j})}$, employing an ordinary least-squares (OLS) method [13].

At level j, the estimations can be computed as

$$\widehat{\mathbf{W}}_i^j = f_i[\mathbf{V}_i^j] = \beta_{i,0}^j + \beta_{i,1}^j V_1^j + \dots + \beta_{i,k}^j V_k^j \tag{3}$$

where $\boldsymbol{\beta}^j$ corresponds to the regression coefficients at level j, $\boldsymbol{\beta}_i^j \in \mathbb{R}^{(k+1) \times 1}$, $\mathbf{V}_i^j \in \mathbb{R}^{m \times 1}$, and $k = z \cdot 2^{-j}$ is the number of approximation components employed by the OLS method. Note that $\boldsymbol{\beta}^j$ values must be stored as the side information for decoding purposes. The size of the side information amounts to $(z^2/3)(1 - (1/2^{2j})) + z(1 - (1/2^j))$ parameters.

Once the predicted signal $\widehat{\mathbf{W}}^j$ is computed, the residuals are obtained by subtraction:

$$\mathbf{R}^j = \mathbf{W}^j - \widehat{\mathbf{W}}^j. \tag{4}$$

Then, assuming that the highest number of decomposition levels J has been applied, $\mathbf{V}^J$ and $\mathbf{R}^{1 \leq j \leq J}$ are losslessly entropy encoded together. At the decoder side, to recover the signal $\mathbf{V}^{J-1}$, the approximations $\mathbf{V}^J$ and the residuals $\mathbf{R}^J$ are employed. This process is computed in cascade form (for $j = J - 1, \dots, 1$). For a more in-detail explanation of RWA reverse process, see [3].

C. RWA Variants

RWA can be executed on different modes to compute the estimation $\widehat{\mathbf{W}}^j$ of the detail components $\mathbf{W}^j$ based on the approximation components $\mathbf{V}^j$: *Maximum*, *Restricted*, *Fast*, and *Exogenous* variants. The maximum model utilizes all approximation components from $\mathbf{V}^j$ for the prediction of each detail component $\mathbf{W}_i^j$. The restricted model only employs a subset of components from $\mathbf{V}^j$ during the prediction to preserve the component scalability of the original DWT. The Fast variant applies a spatial subsampling of the approximation components only for the regression operation computation. Finally, as scenes captured by the same hyperspectral sensor may have similar statistical relationships among their components, the exogenous variant computes the regression coefficients only once, for a given training scene, and employs these predictions for other scenes captured by the same sensor. This regression coefficients' operation is performed off-line, not onboard, saving significant computational resources and execution time. This is a convenient strategy for those scenes corpora with a very large number of spectral components z.

III. PROPOSED LOW-COMPLEXITY CODING TECHNIQUE

A. RWA With CCSDS-123-AC

Up to now, the RWA output signals $\mathbf{V}^J$ and $\mathbf{R}^{1 \leq j \leq J}$ are losslessly entropy encoded with JPEG 2000 bit plane and entropy coder machinery [3], [7]. Here, we substitute JPEG 2000 by CCSDS-123-AC [12], which is a coding technique based on the predictor of CCSDS-123 followed by a lightweight contextual arithmetic encoder. The performance of this approach is discussed in the following.

All the results reported in this letter have been obtained from 27 hyperspectral scenes[1] captured by three different sensors: the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), the Infrared Atmospheric Sounding Interferometer (IASI) and the Atmospheric Infrared Sounder (AIRS). These instruments store scenes with a bit depth of 16 bpppc. Table I depicts the dimensional information of these scenes together with the sensor's name abbreviations, the RWA regression model or variant, and the CCSDS-123-AC's predictor configurations used per each sensor. The average order-0 entropy is also disclosed.

Regarding the RWA configurations, the maximum regression model is applied for uncalibrated AVIRIS (UA) and

---

[1]The scenes are available at http://cwe.ccsds.org/sls/docs/sls-dc/123.0-B-Info/TestData.

TABLE I

SUMMARY OF DATA EMPLOYED FOR THE EXPERIMENTAL RESULTS. (LEFT TO RIGHT) SENSOR NAME TOGETHER WITH THE DIMENSIONS OF THEIR RECORDED SCENES, SENSOR'S NAMES ABBREVIATIONS, USED SCENES, SCENES AVERAGE ORDER-0 ENTROPY, RWA REGRESSION MODEL OR VARIANT, AND MODE AND LOCAL SUM USED BY THE CCSDS-123-AC PREDICTOR. z IS THE NUMBER OF SPECTRAL COMPONENTS, y IS THE HEIGHT, AND x IS THE WIDTH

Sensors & Dimensions	Abbreviation	Scenes	Order-0 entropy	RWA Model or Variant	Predictor Mode	Predictor local sum
Uncalibrated AVIRIS $z=224, y=512, x=680$	UA	Yellowstone, sc: 00, 03, 10, 11, 18	12.16	Maximum	Neighbor Oriented	Full Mode
Calibrated AVIRIS $z=224, y=512, x=680$	CA	Yellowstone, sc: 00, 03, 10, 11, 18	9.77	Maximum	Neighbor Oriented	Full Mode
Uncalibrated IASI Level 0 $z=8359$ $y=1528$ $x=60$	UI	L0 1: 20091007093900Z L0 2: 20091007143900Z L0 3: 20100319050300S6 L0 4: 20120718075700Z	8.12	Exogenous	Neighbor Oriented	Full Mode
Calibrated IASI Level 1 $z=8461$ $y=1530$ $x=60$	CI	L1 1: 20130816230553Z L1 2: 20130817004753Z L1 3: 20130817041457Z L1 4: 20130817055657Z L1 5: 20130817073857Z	12.89	Exogenous	Neighbor Oriented	Full Mode
AIRS Gran $z=1501, y=135, x=90$	AG	sc: 9, 16, 60, 126, 129, 151, 182, 193	11.39	Exogenous	Neighbor Oriented	Reduced Mode

calibrated AVIRIS (CA) scenes and exogenous variant for uncalibrated IASI (UI), calibrated IASI (CI), and AIRS Gran (AG) scenes (see Table I). Though the maximum model attains better prediction, it also needs an important amount of side information to recover the original signal. For instance, UI, CI, and AG sensors need, respectively, 0.918, 0.939, and 1.249 bpppc for the side information if the maximum variant is employed. The expensive storage of the side information yields to a more competitive result when using the exogenous variant. The scenes, UI-1, CI-1, and AG-129, have been employed only for the training procedure. Results are, therefore, not provided for these three scenes.

Leaning on the results of [12] and after conducting an extensive evaluation, experimental results are produced for the CCSDS-123 predictor's parameters, as reported in Table I (columns 6 and 7).

Table II reports the average bit rates for RWA at 1, 5, and the highest number of decomposition levels followed by the JPEG 2000 standard (RWA + JPEG 2000) and CCSDS-123-AC (RWA + CCSDS-123-AC). RWA + CCSDS-123-AC coding gain values with respect to the RWA + JPEG 2000 are also disclosed in Table II (column 6). It is worth noting that, after RWA computation, two very different types of data must be processed, the approximations $\mathbf{V}^j$ and the residuals $\mathbf{R}^{1 \leq j}$. As in JPEG 2000 each component is encoded independently, the approximations and residuals can be considered a single signal to be encoded. However, with CCSDS-123 and after computing the RWA, this consideration does not hold, since, to estimate data from a certain component, the CCSDS-123 predictor can employ information from other components or even form decomposition levels. This leads to an incorrect prediction that may penalize the encoding performance. Consequently, for CCSDS-123-AC, each RWA decomposed signal (approximation and details) is separately encoded. In this case, the best results per each sensor are enhanced in bold. From these results, we can see that: 1) RWA + CCSDS-123-AC almost always outperforms RWA + JPEG 2000 at all levels; 2) the best coding performance for RWA + JPEG 2000 is always obtained when

TABLE II

AVERAGE BIT RATES (IN bpppc) WHEN ENCODING DIFFERENT SCENES CORPORA (SEE TABLE I) WITH JPEG 2000 AND CCSDS-123-AC AFTER RWA AT 1, 5, AND THE HIGHEST NUMBER OF DECOMPOSITION LEVELS. P_R CORRESPONDS TO THE NUMBER OF PREVIOUS COMPONENTS THAT THE CCSDS-123 PREDICTOR USES WHEN ESTIMATING THE RWA RESIDUAL COMPONENTS. THE BEST RESULTS PER EACH SENSOR ARE ENHANCED IN BOLD

	RWA levels	RWA + JPEG 2000	RWA + CCSDS-123-AC		Coding gains
			$P_R = 0$	$P_R = 3$	
UA	1	7.86	5.79	5.79	2.07
	5	5.87	5.71	**5.68**	0.19
	8	5.83	5.72	5.69	0.14
CA	1	5.27	3.64	3.69	1.63
	5	3.56	**3.52**	3.63	0.04
	8	3.52	3.53	3.63	-0.01
UI	1	3.86	2.66	2.77	1.20
	5	2.54	**2.51**	2.75	0.03
	14	2.44	**2.51**	2.77	-0.07
CI	1	8.17	6.54	6.41	1.76
	5	6.60	6.38	6.24	0.36
	14	6.46	6.37	**6.23**	0.23
AG	1	5.62	4.15	4.23	1.47
	5	4.25	**3.99**	4.16	0.26
	11	4.18	4.01	4.17	0.17

the highest number of decomposition levels is applied; and 3) RWA + CCSDS-123-AC at the highest decomposition level does not always yield the lowest bit rates.

B. Multilevel Split RWA

As seen previously, fixing the number of RWA levels is not the best strategy when CCSDS-123-AC is harnessed. In order to obtain the best coding performance, we automatically select the adequate number of decomposition levels for RWA + CCSDS-123-AC, and we name this proposal as multilevel split RWA (MLS-RWA) + CCSDS-123-AC.

Fig. 1 renders our proposed coding technique when applying RWA for any regression model or variant. The main insight of
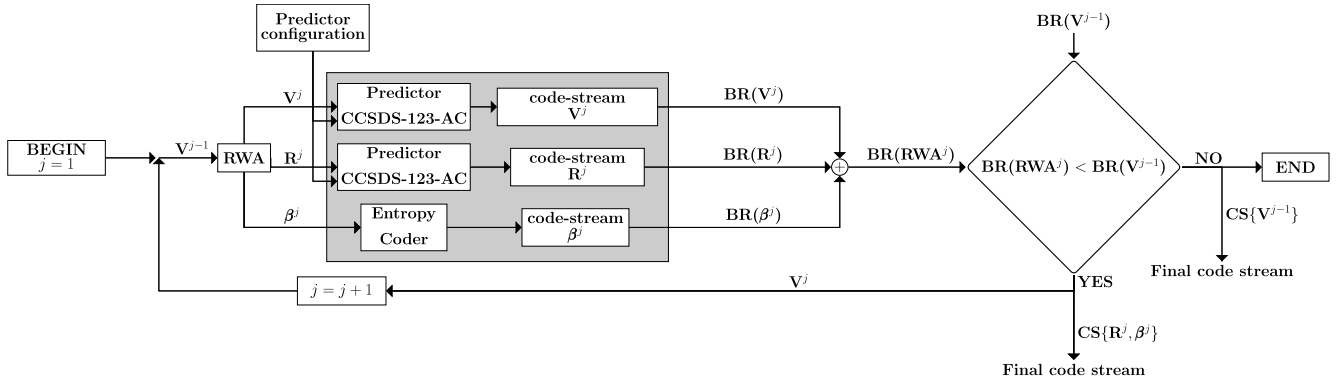
Fig. 1. Proposed coding scheme.

TABLE III

AVERAGE BIT RATES (IN bpppc) FOR PCA + JPEG 2000, RWA + JPEG 2000, M-CALIC, CCSDS-123, CCSDS-123-AC, AND MLS-RWA + CCSDS-123-AC FOR THE SCENES OF TABLE I. THE BEST CODING PERFORMANCES OF OUR PROPOSAL ARE ENHANCED BOLD. THE CODING GAINS WITH RESPECT TO OUR BEST PROPOSAL'S OUTCOMES ARE ALSO INCLUDED IN PARENTHESES. P_R CORRESPONDS TO THE NUMBER OF PREVIOUS COMPONENTS EMPLOYED BY CCSDS-123-AC'S PREDICTOR STAGE TO ESTIMATE THE RWA RESIDUAL COMPONENTS

	MLS-RWA levels	PCA+JPEG2000	RWA+JPEG2000	M-CALIC	CCSDS-123	CCSDS-123-AC	MLS-RWA+CCSDS-123-AC	
							$P_R = 0$	$P_R = 3$
UA	4	5.81 (0.14)	5.83 (0.16)	6.06 (0.39)	5.98 (0.31)	5.92 (0.25)	5.69	**5.67**
CA	4	3.69 (0.17)	3.52 (0.00)	3.86 (0.34)	3.72 (0.20)	3.76 (0.24)	**3.52**	3.62
UI	6	—	2.44 (-0.07)	2.88 (0.37)	2.88 (0.37)	2.82 (0.31)	**2.51**	2.75
CI	7	—	6.46 (0.23)	6.88 (0.65)	6.59 (0.36)	6.53 (0.30)	6.36	**6.23**
AG	4 (sc:16) 5 (sc:126,193) 6 (sc:9,60,151,182)	7.69 (3.70)	4.18 (0.19)	4.34 (0.35)	4.33 (0.34)	4.29 (0.30)	**3.99**	4.14
	Average	5.73 (1.35)	4.48 (0.10)	4.80 (0.42)	4.70 (0.32)	4.66 (0.28)	4.38	

MLS-RWA is to process a new RWA decomposition level over V^j only if the sum of bit rates of approximations, residuals, and regression coefficients does not exceed the bit rate required to losslessly compress the approximations of the lower level, i.e., BR (V^j) + BR (R^j) + BR (β^j) < BR (V^{j-1}). Iteratively, further RWA decomposition levels are applied until no coding gain exists. The code streams from the last applied level are discarded and the process stops. In Fig. 1, CS and BR refer, respectively, to the code stream term and bit rate operation. BR (V^0) has been previously obtained after encoding V^0 with CCSDS-123-AC. We use the Lempel–Ziv–Markov chain algorithm (LZMA) [14] as an entropy coder for encoding the regression coefficients. However, another lossless coding technique could be employed. All the residual and regression coefficients code streams and the approximation coefficients code stream at the highest decomposition level, which satisfy the condition, are stored together into the final code stream. For exogenous variant, the regression coefficients would have been fixed beforehand and would not be stored into the final code stream.

MLS-RWA decreases the computational cost by reducing the number of RWA levels to apply and by replacing JPEG 2000 with CCSDS-123-AC lower complexity encoder. By applying a lower number of RWA decomposition levels, we decrease not only the number of floating-point operations but also the number of regression coefficients that have to be computed onboard and stored as side information.

In this letter, neither the spatial subsampling nor a spectral components selection are proposed to alleviate the complexity of the regression operation block as explained in [3] and [7], although MLS-RWA could apply them too. Also, MLS-RWA could be combined with any other coding system.

IV. DISCUSSION AND RESULTS

Table III reports the lossless coding performance in average bit rate for PCA + JPEG 2000, RWA + JPEG 2000, CCSDS-123, M-CALIC, CCSDS-123-AC coding techniques, which are competitive coding techniques, and our approach MLS-RWA + CCSDS-123-AC. For CCSDS-123 and CCSDS-123-AC coding techniques, 3 previous components are used in the prediction process. The entropy encoder used for CCSDS-123 is the sample adaptive, which is the entropy coder that attains higher lossless compression ratios. For CCSDS-123-AC and when encoding approximation coefficients with our approach, the number of prediction components is set to 3 too, because, as with the original image components, the approximation components are still highly correlated in the spectral domain. For MLS-RWA + CCSDS-123-AC, the best performance of the predictor is achieved with different configurations depending on the sensor, as higher or lower statistical relationships can still remain within the residual components after RWA computation. Therefore, and only for our proposal, we report the outcomes when 0 and 3 previous components

are employed for predicting the residuals. The best coding performance is enhanced in bold. Our proposal's best coding gains with respect to each coding technique are reported within parenthesis. The last row reports the average of each coding technique; MLS-RWA + CCSDS-123-AC average value is computed only using the best results.

Comparing with coding techniques based on transforms, such as PCA + JPEG 2000 and RWA + JPEG 2000, we can observe that our proposal yields significant lower bit rates for all the sensors, except for CA and UI scenes when applying RWA + JPEG 2000. For CA scenes, no improvement is achieved, while an average loss of only 0.07 bpppc is produced for UI sensor. As mentioned earlier, MLS-RWA + CCSDS-123-AC entails also a decrement of the computational cost. For CA and UI, our proposal only computes 6 and 14 RWA decomposition levels, letting aside 130 and 14 components for which neither a Haar-DWT nor a regression operation to estimate the details need to be performed. On average, code stream size reductions of, respectively, 23.56% and 2.23% between our proposal and PCA + JPEG 2000 and RWA + JPEG 2000 are achieved. Results for PCA + JPEG 2000 are not provided for UI and CI scenes due to their large number of spectral components, which makes the PCA computation unfeasible, as the covariance matrix and eigenvalues calculation, the matrix factorization, and the amount of side information become prohibitive. For AG sensor, PCA's average side information requires 3.71 bpppc, yielding a final bit rate of 7.69 bpppc.

M-CALIC and CCSDS-123 achieve quite competitive results at low computational complexity too; for instance, CCSDS-123 is, on average, about 0.1 bpppc worse than CCSDS-123-AC. However, our low complexity proposal MLS-RWA + CCSDS-123-AC improves, on average, by about 8.75%, 6.80%, and 6.00% as compared to M-CALIC, CCSDS-123, and CCSDS-123-AC, respectively.

Comparing with the recently published coding techniques based on the conventional recursive least squares (and adaptations thereof) for UA and CA scenes, the best results reported in [10] are, respectively, 5.57 and 3.29 bpppc, whereas the best results reported in [11] are, respectively, 5.55 and 3.31 bpppc. These results are, respectively, about 0.1 and 0.2 bpppc better than those of MLS-RWA + CCSDS-123-AC, which is reasonable due to the tradeoff between coding performance and computational complexity of [10] and [11].

V. CONCLUSION

This letter introduces a lossless coding technique that provides superior coding performance than the state-of-the-art coding techniques at low computational complexity. It is based on RWA followed by a recently presented lightweight contextual arithmetic encoder prepended by the CCSDS-123 prediction (CCSDS-123-AC). RWA removes the spectral redundancy by computing a DWT and, in sequence, a light regression operation to estimate the detail coefficients from the approximation coefficients.

Different to what happens when coding RWA code streams with JPEG 2000, applying the highest number of RWA

decomposition levels with CCSDS-123-AC does not imply the best coding performance. To provide lower bit rates, we developed a smart strategy to properly select the number of RWA decomposition levels that afford the highest coding gain, named MLS-RWA. By decreasing the number of levels and by replacing JPEG 2000 by a light arithmetic coder, our proposal also reduces the computational cost of the original RWA + JPEG 2000 approach.

Extensive experimental results over 27 hyperspectral scenes from 5 corpus sensors have been performed, indicating that MLS-RWA followed by CCSDS-123-AC outperforms CCSDS-123.0, M-CALIC, PCA + JPEG 2000, CCSDS-123-AC, and RWA + JPEG 2000, and provides significant average coding gains of 0.32, 0.42, 1.35, 0.28, and 0.10 bpppc, respectively, for lossless compression. Our approach yields very competitive results even when the RWA is computed only for a single decomposition level. In this case, an average coding gain of at least 1.74 bpppc over RWA + JPEG 2000 is obtained.

REFERENCES

[1] I. T. Jolliffe, *Principal Component Analysis*. Berlin, Germany: Springer-Verlag, 2002, p. 487.

[2] I. Blanes and J. Serra-Sagristà, "Pairwise orthogonal transform for spectral image coding," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 961–972, Oct. 2010, doi: 10.1109/TGRS.2010.2071880.

[3] N. Amrani, J. Serra-Sagristà, V. Laparra, M. W. Marcellin, and J. Malo, "Regression wavelet analysis for lossless coding of remote-sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5616–5627, Jun. 2016, doi: 10.1109/TGRS.2016.2569485.

[4] D. Taubman and M. W. Marcellin, *JPEG 2000: Image Compression Fundamentals, Practices and Standards*. New York, NY, USA: Springer, 2001.

[5] E. Magli, G. Olmo, and E. Quacchio, "Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 1, pp. 21–25, Jan. 2004, doi: 10.1109/LGRS.2003.822312.

[6] Consultative Committee for Space Data Systems. (May 2012). Lossless multispectral & hyperspectral image compression CCSDS 123.0-B-1. Blue Book. [Online]. Available: https://public.ccsds.org/Pubs/123x0b1ec1.pdf

[7] S. Álvarez-Cortés, N. Amrani, and J. Serra-Sagristà, "Low complexity regression wavelet analysis variants for hyperspectral data lossless compression," *Int. J. Remote Sens.*, vol. 39, no. 7, pp. 1971–2000, Sep. 2017, doi: 10.1080/01431161.2017.1375617.

[8] J. Song, Z. Zhang, and X. Chen, "Lossless compression of hyperspectral imagery via RLS filter," *Electron. Lett.*, vol. 49, no. 16, pp. 992–994, Aug. 2013, doi: 10.1049/el.2013.1315.

[9] W. Jiaji, K. Wanqiu, M. Jarno, and H. Bormin, "Lossless compression of hyperspectral imagery via clustered differential pulse code modulation with removal of local spectral outliers," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2194–2198, Dec. 2015, doi: 10.1109/LSP.2015.2443913.

[10] G. Fang and G. Shuxu, "Lossless compression of hyperspectral images using conventional recursive least-squares predictor with adaptive prediction bands," *J. Appl. Remote Sens.*, vol. 10, no. 1, p. 015010, Feb. 2016, doi: 10.1117/1.JRS.10.015010.

[11] A. C. Karaca and M. K. Güllü, "Lossless hyperspectral image compression using bimodal conventional recursive least-squares," *Remote Sens. Lett.*, vol. 9, no. 1, pp. 31–40, Jan. 2018, doi: 10.1080/2150704X.2017.1375612.

[12] J. Bartrina-Rapesta, I. Blanes, F. Aulí-Llinàs, J. Serra-Sagristà, V. Sanchez, and M. W. Marcellin, "A lightweight contextual arithmetic coder for on-board remote sensing data compression," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4825–4835, May 2017, doi: 10.1109/TGRS.2017.2701837.

[13] J. Nocedal and S. J. Wright, "Least-squares problems," in *Numerical Optimization*. New York, NY, USA: Springer, 2006, pp. 245–269.

[14] I. Pavlov. (2007). *LZMA SDK (Software Development Kit)*. [Online]. Available: https://www.7-zip.org/sdk.html