THE EMERGING ISSUE-3 REVISION OF THE CCSDS-123.0-B LOW-COMPLEXITY LOSSLESS & NEAR-LOSSLESS MULTISPECTRAL & HYPERSPECTRAL IMAGE COMPRESSION STANDARD

Aaron Kiely(1), Mickaël Bruno(2), Aniello Fiengo(3), Miguel Hernández-Cabronero(4), Didier Keymeulen(1), Matthew Klimesh(1), Lucana Santos(3), Englin Wong(5)

(1)NASA Jet Propulsion Laboratory (JPL)
California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
Emails: aaron.b.kiely@jpl.nasa.gov, didier.keymeulen@jpl.nasa.gov, matthew.a.klimesh@jpl.nasa.gov

(2)Centre National d'Études Spatiales (CNES)
Email: mickael.bruno@cnes.fr

(3)European Space Research and Technology Centre, European Space Agency,
2220 AG Noordwijk, The Netherlands
Emails: Aniello.Fiengo@esa.int, Lucana.Santos@esa.int

(4)Dept. of Information and Communications Engineering
Universitat Autònoma de Barcelona
Campus UAB, 08193 Cerdanyola del Vallès, Spain
Email: miguel.hernandez@uab.cat

(5)NASA Goddard Space Flight Center
8800 Greenbelt Road, Greenbelt, MD 20771, US
Email: mark.wong@nasa.gov

Abstract

The CCSDS-123.0-B Low-Complexity Lossless and Near-Lossless Multispectral & Hyperspectral Image Compression standard provides state-of-the-art compression for imaging spectrometer data. Issue 1 of this standard, published in 2012, provides only lossless compression capability. In 2019, Issue 2 was published, adding new features and in particular augmenting the compressor to provide also near-lossless compression, i.e., compression with a user-specified error limit in each spectral band. This presentation will describe the Issue 3 revision currently under development by the CCSDS Data Compression working group. This revision will provide region-of-interest compression capability. That is, a user-provided spatial classification map can be used to specify different data fidelity parameters in different image regions. This would, for example, allow an onboard classification algorithm to identify the most important regions of an image, or regions obscured by clouds, and adjust compression fidelity accordingly in each region to maximize value of imaging spectrometer data returned over constrained space communication links. The revision also aims to add an option that is unrelated to region-of-interest compression but can improve compression effectiveness in some cases.

1 INTRODUCTION

In 2012, the Consultative Committee for Space Data Systems (CCSDS) published the Recommended Standard [1], which describes a lossless compression algorithm for the three-dimensional data sets produced by imaging spectrometers (such spectrometers are often referred to as multispectral and hyperspectral imagers). In 2019, an update [2] was published as Issue 2 (CCSDS 123.0-B-2). Issue 2 adds a number of new options and in particular provides the ability to perform near-lossless compression. This paper assumes the reader has some familiarity with Issue 2.

The Issue 3 revision to CCSDS 123.0 is currently under development by the CCSDS Data Compression working group. This revision will provide region-of-interest (ROI) compression capability, in which a user-provided spatial classification map is used to specify different fidelity parameters in different image regions, on a spatial sample-by-sample basis. This would, for example, allow an onboard classification algorithm to identify the most important regions of an image, or regions obscured by clouds, and adjust compression fidelity accordingly in each region to maximize value of imaging spectrometer data returned over constrained space communication links.

For the particular case of using ROI compression with classification maps that identify cloudy locations, there is a clear practical case for incorporating this capability into Issue 3. In [3] it is mentioned that the European Space Agency (ESA) Copernicus Hyperspectral Imaging Mission for Environment (CHIME) mission (first working launch date 2028) has a requirement to detect clouds and employ compression that compresses cloudy locations with a higher level of loss. Reference [3] presents a VHDL-based hardware solution to meet the needs of CHIME that is consistent with the ROI compression algorithmic changes planned for Issue 3 (although in [3] only the

Block Adaptive choice is used for entropy coding). Reference [3] builds on the earlier [4] which presents a brief description of the "Different Absolute Error (DAE)" method for compressing images with cloudy locations.

The Issue 3 revision also aims to add additional options in the computation of "sample representative" values, which are used in the prediction calculation in Issue 2. One quantity used in the calculation of sample representative values is the "sample representative damping" value. In Issue 2 the sample representative damping value can vary by band index but otherwise is constant. However, it has been found that with some combinations of compression parameters there can be an appreciable compression effectiveness advantage to letting the sample representative damping take a different value for the first frame ($y = 0$) than with later frames. Thus, it is a goal to provide such a capability in the Issue 3 revision, though how much generality will be provided is yet to be determined.

It should be emphasized that although we are reasonably confident that the final version of the Issue 3 revision will generally agree with the material presented in this paper, it is possible that there will be differences. In addition, some details have yet to be decided on, including ranges of new parameters and the updates to the header structures that will accommodate new options. This paper mostly avoids discussion of updates to the header structures. Importantly though, the Issue 3 revision will be backward-compatible with Issue 2 in the sense that compressed images are compliant with Issue 2 will also be compliant with Issue 3.

2 REGION-OF-INTEREST COMPRESSION

2.1 Error Limits in Issue 2

The fidelity of near-lossless compression in Issue 2 is specified with absolute and/or relative error limits. An image that is compressed with given error limits can be decompressed to an image for which the differences between the original and corresponding reconstructed sample values are guaranteed to conform to the error limits on a sample-by-sample basis. The error limits can vary by spectral band index but the only provision for varying the error limits spatially is by using periodic error limit updating, in which the error limits are updated with a period of a specified number of frames.

2.2 Overview of Region-of-Interest Compression in Issue 3

ROI compression will be optional; the selection of ROI compression will be indicated by a single previously reserved bit in the Essential Subpart of the compressed image header. In an Issue 2 compliant compressed image, this bit will be zero and the compressed image will also be Issue 3 compliant (thus providing backward-compatibility).

A classification map is required for ROI compression. The classification map is a spatial map with the same spatial resolution as the image to be compressed. The idea is that all spatial locations in a given class (i.e., with a given classification index) will be compressed with the same fidelity parameters, but different classes may use different fidelity parameters. Classes might be used to indicate the relative interest of different regions to scientists (perhaps determined by a classifier prior to compression), with regions of higher interest compressed with higher fidelity, thereby attempting to maximize the science value of images returned over a limited communications link. More specifically, if regions obscured by clouds are of little interest, then spatial locations could be classified as either "cloudy" or "not cloudy" and the cloudy class could be assigned significantly lower fidelity.

Notationally, the classification map is an index $c(t)$ (alternatively denoted $c_{x,y}$) that must be defined for each value of the alternate image coordinate index t (or each (x, y) pair of image coordinate indices). Each $c(t)$ must be in the range $0 \leq c(t) < N_c$, where N_c is the user-specified number of classes. The allowed range of N_c is not yet decided on, but a possibility is $1 \leq N_c \leq 256$. (As will be discussed later, $N_c = 1$ is allowed for technical reasons, and is generally *not* the same as not using ROI compression.) The classification map must be known to both the compressor and decompressor. The classification map may be encoded in the header in a straightforward way (each index encoded as a $\lceil \log_2(N_c) \rceil$-bit unsigned binary integer), and there may be an option for instead encoding the classification map periodically in the compressed image body. There may also be an option for including the classification map in user-defined format in the image header (allowing a custom compressor to be used for the classification map). Preliminary performance results for compressing classification maps are described later in Section 2.4. Finally, there may be an option for omitting the classification map from the compressed image entirely, in which case it would have to be communicated separately in order to provide it to the decompressor. Issue 3 will not say anything about how the classification map is generated, since this is a separate activity from compression and decompression (neither the compressor nor decompressor needs to know how the classification map was generated); also, not standardizing a particular method for generating classification maps gives users the flexibility to take advantage of new methods as they emerge.

For each of the N_c classification index values, absolute and/or relative error limits must be specified. These may depend on the band indices. It is planned that different classes may use different fidelity control methods (among lossless, absolute error limits, relative error limits, and both absolute and relative error limits; also, whether the limits are band-dependent may depend on the class). It is tentatively planned that periodic error limit updating will *not* be allowed when ROI compression is used, in order to avoid the complication of having to periodically update the error limits for all classes (however, it is arguable that there is a use case for combining periodic error limit updating with ROI compression). The details of how the error limits for each class will be specified are not yet determined, but it will be possible to omit this information from the compressed image; this would allow, for example, a mission using the same error limits for an entire mission phase to avoid encoding the error limits with every image.

2.3 Calculation Changes for ROI Compression

Using ROI compression does not make any direct changes to the prediction calculation: the local sums, local differences, weight vectors, and predicted sample values are all calculated in the same way whether or not ROI compression is used. On the other hand, the maximum error value $m_z(t)$ is used for fidelity control and so when ROI compression is used the calculation of $m_z(t)$ must take into account the classification map. This is done in a straightforward way: In Issue 2, depending on the fidelity control method used, $m_z(t)$ may depend on an absolute error limit a_z and/or a relative error limit r_z that in general depend on the band index z. For Issue 3, the calculation is the same except that a_z and r_z now may depend on the class at the spatial location t.

In near-lossless predictive compression in general, the "prediction residuals" (differences between the predicted and actual sample values) are quantized, and the resulting quantizer indices are losslessly encoded in the compressed bitstream. In Issue 2 and the expected Issue 3, there is an intermediate step that reversibly maps the quantizer indices to "mapped quantizer indices" that are nonnegative integers that are smaller when the magnitude of the quantizer indices are smaller. Larger maximum error values generally produce smaller mapped quantizer indices.

Issue 2 provides three choices of method to use for encoding the mapped quantizer indices: "sample-adaptive", "hybrid" and "block-adaptive". No change for Issue 3 is expected for the block-adaptive method. For the other two methods some modifications are planned.

The sample-adaptive and hybrid methods keep simple statistics on the size of the mapped quantizer indices. These statistics are kept separately for different bands. For the sample-adaptive method, these statistics consist of an accumulator $\Sigma_z(t)$ and a counter $\Gamma(t)$ that are adaptively updated during the encoding process. In Issue 2, the ratio $\Sigma_z(t) / \Gamma(t)$ serves as an estimate of the mean mapped quantizer index value in spectral band z. Because the fidelity control parameters are the same for all samples, the maximum error value $m_z(t)$ will either be the same for all samples (in the case fidelity control method is lossless or absolute error limits) or roughly the same for all samples (when relative error limits are used). In either case the ratio $\Sigma_z(t) / \Gamma(t)$ should work reasonably well as an estimate of the expected mapped quantizer index value for all samples in a band. However, when ROI compression is used this may no longer be true: samples for which the ROI class uses a small value of $m_z(t)$ may be expected to have larger mean mapped quantizer index values, and samples for which the ROI class uses a large value of $m_z(t)$ may be expected to have smaller mapped quantizer index values. We have chosen to address this by modifying the calculation and meaning of the accumulator $\Sigma_z(t)$ and modifying the way that $\Sigma_z(t)$ is used in the entropy coding process.

2.3.1 *Sample Adaptive Entropy Coder Changes*

For sample-adaptive entropy coding, in Issue 3 the meaning of $\Sigma_z(t)$ will be different depending on whether or not ROI is used. If ROI is not used then the ratio $\Sigma_z(t) / \Gamma(t)$ provides an estimate of the mean mapped quantizer index value in the spectral band. This ratio determines the variable-length code used to encode the mapped quantizer index $\delta_z(t)$. If ROI is used then the ratio $\Sigma_z(t) / \Gamma(t)$ provides an estimate of the mean mapped quantizer index value in the spectral band if quantization step size 1 were used.

A new "mapped quantizer index scale value", denoted $\lambda_z(t)$, is defined for $t > 0$ as follows. When ROI is not used, $\lambda_z(t) = 1$. When ROI is used and only absolute error limits are used, then $\lambda_z(t) = 2a_z(t)+1$. When ROI is used and only relative error limits are used, then $\lambda_z(t) = 2r_z(t)+1$. When ROI is used and both absolute and relative error limits are used, then $\lambda_z(t) = 2\min(a_z(t), r_z(t))+1$.

The mapped quantizer index scale value is meant to be the value that a mapped quantizer index should be multiplied by in order to approximate what the mapped quantizer index would be if the sample was being compressed losslessly. Accordingly, the update equation for the accumulator, which is (60) in [2], is modified to be the following:

$$\Sigma_z(t) = \begin{cases} \Sigma_z(t-1) + \lambda_z(t)\cdot\delta_z(t-1), & \Gamma(t-1) < 2^{\gamma^*} - 1 \\ \left\lfloor \frac{\Sigma_z(t-1)+\lambda_z(t)\cdot\delta_z(t-1)+1}{2} \right\rfloor, & \Gamma(t-1) = 2^{\gamma^*} - 1 \end{cases}. \tag{1}$$

Note that the only change from (60) in [2] to (1) is that both occurrences of $\delta_z(t\text{-}1)$ are replaced by $\lambda_z(t) \cdot \delta_z(t\text{-}1)$.

The determination of the variable length code parameter $k_z(t)$ is also modified from 5.4.3.2.4.3 of [2]; the new determination is that $k_z(t) = 0$ if $2\Gamma(t)\cdot\lambda_z(t) > \Sigma_z(t) + \left\lfloor \frac{49}{2^7}\Gamma(t)\cdot\lambda_z(t) \right\rfloor$; otherwise $k_z(t)$ is the largest positive integer $k_z(t) \le D - 2$, such that

$$\Gamma(t)\cdot\lambda_z(t)\cdot 2^{k_z(t)} \le \Sigma_z(t) + \left\lfloor \frac{49}{2^7}\Gamma(t)\cdot\lambda_z(t)\right\rfloor. \quad (2)$$

One can see from the preceding two paragraphs that lossy compression using ROI and just one class is *not* the same as not using ROI compression, since the accumulator $\Sigma_z(t)$ is calculated and used differently. These differences can result in a minor difference in the compressed image body size. A reason for allowing lossy compression with $N_c = 1$ is that it lets an implementation produce the effect of non-ROI-compression without treating such compression differently from ROI compression.

2.3.2 Hybrid Entropy Coder Changes

For the hybrid entropy coder, in Issue 3 there will be changes that are analogous to the changes for the sample-adaptive entropy coder. The changes are similar enough to the changes for the sample-adaptive entropy coder that we will not provide details here.

Given the meaning of the mapped quantizer index scale value $\lambda_z(t)$ as "the value that a mapped quantizer index should be multiplied by in order to approximate what the mapped quantizer index would be if the sample was being compressed losslessly", it might seem better to define it as equal to $2m_z(t)+1$ instead of the actual definition above in terms of $a_z(t)$ and $r_z(t)$; this would be more accurate when relative error limits are used. However, this would not work for the hybrid entropy coder because the value $m_z(t)$ might not be known to the decoder at the time it needs it: When using the hybrid entropy coder, the decoding of the mapped quantizer indices in a band occurs in reverse order and must complete before the actual predicted sample values can be computed (since these are determined in forward order). The mapped quantizer index scale values for a band must be known in order to decode the mapped quantizer indices, so this implies that the mapped quantizer index scale value *cannot* depend on the predicted sample values. It *would* be possible to define $\lambda_z(t)$ as equal to $2m_z(t)+1$ for the sample-adaptive entropy coder, but we have made the choice to define $\lambda_z(t)$ in the same way for both the sample-adaptive and hybrid entropy coders.

2.4 Classification Map Compression Performance

Several lossless compression approaches are being examined for compressing the classification maps introduced by the upcoming Issue 3. One approach, presented in [6], is an algorithm specifically designed for this type of data. Contrary to most traditional image compressors, it does not make the assumption (generally false for classification maps) that sample values change smoothly at small spatial scales. Its reported performance is compared here to that of the following approaches:

- *No coding*: the shortest fixed-length code sufficient for the number of classes in the map. This would be analogous to not compressing the class indices.
- *Issue 2*: the CCSDS 123.0-B-2 standard using the hybrid entropy coder. Analogous to compressing the classification map as a regular spectral band, but with spatial-only prediction.
- *Issue 1 – No prediction*: a non-compliant modification of the CCSDS 123.0-B-1 to use the sample adaptive entropy coder but skipping the prediction stage. Analogous to compressing the classification map as a regular spectral band, but without any type of prediction.
- *Huffman*: optimal prefix code generated for the global probability distribution of the classification map.
- *Deflate*: LZ77 followed by Huffman.
- *RLE*: run-length encoding.
- *PAAC*: a prediction filter with pre-fixed weights, followed by a binary adaptive arithmetic coder.
- *JPEG LS, JPEG XL*: direct application of these well-known standards.

Preliminary coding performance results are shown in Table 1 for the dataset introduced in [6]. This dataset comprises 5 maps of size 614x512, derived from an actual AVIRIS scene. These maps contain 4, 7, 9, 17 and 32 classes, respectively, to account for different user requirements when choosing the granularity of the ROI classes. Compressed data rates are expressed in bits per sample (bps) individually for each of these maps, and on average for all five.

Table 1. Compressed data rates in bits per sample (bps) of different algorithms on class maps with different number of classes. The best result for each number of classes is highlighted in bold font.

	Number of classes					
Codec	**4**	**7**	**9**	**17**	**32**	**Average**
Xie et al. [6]	**0.176**	**0.279**	**0.487**	**0.915**	**2.415**	**0.834**
No coding	2.000	3.000	4.000	5.000	5.000	3.800
Issue 2	0.359	0.776	1.506	2.807	4.775	2.045
Issue 1 – No prediction	2.808	2.995	3.789	4.243	5.511	3.869

Huffman	1.650	2.173	2.809	3.761	4.917	3.062
Deflate	0.391	0.559	0.854	1.397	2.845	1.209
PAAC	0.329	0.672	1.230	2.216	4.224	1.734
RLE	0.874	1.337	1.927	3.114	5.635	2.577
JPEG LS	0.355	0.588	0.911	1.688	3.896	1.488
JPEG XL	0.239	0.328	0.583	1.107	2.777	1.007

The method described in [6] yields the best results for all tested number of classes, with improvements between 0.05 and 0.36 bits per sample (between 14.98% and 35.79% of the compressed bitrate) over the JPEG XL standard, which ranks seconds in all test cases. Of the remaining image compressors, Issue 2, PAAC and JPEG LS have a performance comparable to that of JPEG XL for 4-7 classes, but performance drops for more classes, likely due to the aforementioned assumption of local sample smoothness. The remaining methods, all of them general entropy coders, are unable to close the performance gap with [6] for any number of classes. This is as expected, since these methods lack information about the 2D structure that image coders exploit. Notwithstanding, their performance is not generally so severely compromised for larger number of classes. As a result, Deflate exhibits some of the best average results, only worse than [6] and JPEG XL.

3 SAMPLE REPRESENTATIVE DAMPING PARAMETER

One of the innovations incorporated into Issue 2 is the use of “sample representatives”. The idea behind sample representatives is that there is some noise in the image samples values from imaging spectrometers, and if one knew the “true” value of the signal that produced a given sample, using that true value in the prediction of later samples may give a slightly better prediction than using the actual sample. Of course, the true value is not known, but it turns out that when compressing image data from some imaging spectrometers, using a “sample representative” that combines information from the sample value with the predicted sample value can on average produce a more accurate prediction of later sample values than using only the sample value. The use of sample representatives can indeed provide an appreciable improvement in compression effectiveness in some cases; see [5] for some examples.

The sample representative damping parameter ϕ_z, combined with the sample representative offset ψ_z and the sample representative resolution parameter Θ, determine how the sample representative is calculated from the predicted sample value. Roughly, the sample representative for a sample is calculated as a value between an approximation to the original sample and the predicted sample value. The sample representative is set to the value $\phi_z/2^\Theta$ of the way from the approximation to the original sample to the predicted sample value.

In Issue 2, the sample representative damping parameter ϕ_z can vary depending on the spectral band, but within a band it is a constant. However, it has been found that for some situations, there can be an appreciable advantage in allowing ϕ_z to vary with the image line (y coordinate).

Example: An image from the NASA Earth Surface Mineral Dust Source Investigation (EMIT) mission has dimensions $N_X = 1280$, $N_Y = 32$, $N_Z = 328$, and 14-bit samples. Compression is performed with column-oriented local sums and number of spectral bands used for prediction (P) is 3. Lossless compression is selected.

- For wide local sums, choosing $\phi_z/2^\Theta = 3/8$ works reasonably well (call this the baseline).
- Changing to narrow local sums, and keeping other parameters the same, results in a compressed size that is more than 10% larger than in the baseline case, even though narrow column-oriented local sums are the same as wide column-oriented local sums when y > 0. As mentioned in [2], one might want to use narrow local sums in a hardware implementation to facilitate pipelining. (Setting $\phi_z/2^\Theta$ to a smaller value yields some improvement but is still somewhat worse than baseline.)
- Still with the narrow local sums, but now setting $\phi_z = 0$ for $y = 0$, keeping $\phi_z/2^\Theta = 3/8$ elsewhere, and making no other changes produces a compressed size that is very close to the size achieved in the baseline case.

For the above example, investigation of the reason for narrow local sums producing such a significant increase in size revealed that the cause appears to that for $y = 0$, predicted values are bad enough to produce instability and oscillatory behavior in the sample representatives. This makes the predicted sample values quite poor and thus incurs a significant bit cost. The effect propagates for several y values; e.g., there is still significant oscillation in sample representative values at $y = 4$.

More experimentation is needed, but it is planned to incorporate some form of allowing ϕ_z to vary with y in Issue 3. A minimal change could be to just include an option to use $\phi_z = 0$ only when $y = 0$, and otherwise use another specified value for ϕ_z. A more flexible possibility would be to allow different values for ϕ_z in the four cases $y = 0$, $y = 1$, $y = 2$, and $y > 2$.

4 CONCLUSION

The Issue 3 revision to CCSDS 123.0 is currently under development by the CCSDS Data Compression working group. Issue 3 will provide ROI compression capability, in which a user-provided spatial classification map is used to specify different fidelity parameters in different image regions, on a spatial sample-by-sample basis. The Issue 3 revision also aims to add additional options in the computation of “sample representative” values, with the aim of providing an improvement to compression effectiveness in certain cases.

ACKNOWLEDGEMENTS

The research conducted by A. Kiely, D. Keymeulen, and M. Klimesh was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Research by M. Hernández-Cabronero was partially funded by projects PID2021-125258OB-I00, SGR2021-00643 and DTN/TVO/ET/2024-01059 by the Spanish Government, Catalan Government and CNES, respectively.

REFERENCES

[1] *Lossless Multispectral & Hyperspectral Image Compression*. Recommendation for Space Data System Standards, CCSDS 123.0-B-1. Blue Book. Issue 1. Washington, D.C.: CCSDS, May 2012.

[2] *Low-Complexity Lossless and Near-Lossless Multispectral and Hyperspectral Image Compression*. Recommendation for Space Data System Standards, CCSDS 123.0-B-2. Blue Book. Issue 2. Washington, D.C.: CCSDS, Feb. 2019.

[3] A. Sánchez, D. Ventura, Y. Barrios, L. Berrojo, C. Carrasco, F. Veljkovic, P. Rodríguez, and R. Sarmiento, “A smart compression approach based on the CCSDS 123.0-B-2 standard for CHIME,” *8th International Workshop on On-Board Payload Data Compression (OBPDC)*, September 2022, 8 pages.

[4] D. Lebedeff, M. Foulon, R. Camarero, R. Vitulli, and Y. Bobichon, “On-board cloud detection and selective spatial/spectral compression based on CCSDS 123.0-B-2 for hyperspectral missions,” *7th International Workshop on On-Board Payload Data Compression (OBPDC)*, September 2020, 8 pages.

[5] *Low-Complexity Lossless and Near-Lossless Multispectral and Hyperspectral Image Compression.* Informational Report, CCSDS 120.2-G-2. Green Book. Issue 2. Washington, D.C.: CCSDS, Dec. 2022.

[6] H. Xie and M. Klimesh, “Lossless Compression of Classification Map Images, IPN Progress Report, pp. 42-169, 2007.