A New Lossless Compression Algorithm For Satellite Earth Science Multi-Spectral Imagers

Irina Gladkova[a] and Srikanth Gottipati[a] and Michael Grossberg[a]

[a]CCNY, NOAA/CREST, 138th Street and Convent Avenue, New York, NY 10031, USA

ABSTRACT

Multispectral imaging is becoming an increasingly important tool for monitoring the earth and its environment from space borne and airborne platforms. Multispectral imaging data consists of visible and IR measurements from a scene across space and spectrum. Growing data rates resulting from faster scanning and finer spatial and spectral resolution makes compression an increasingly critical tool to reduce data volume for transmission and archiving. Examples of multispectral sensors we consider include the NASA 36 band MODIS imager, Meteosat 2nd generation 12 band SEVIRI imager, GOES R series 16 band ABI imager, current generation GOES 5 band imager, and Japan's 5 band MTSAT imager. Conventional lossless compression algorithms are not able to reach satisfactory compression ratios nor are they near the upper limits for lossless compression on imager data as estimated from the Shannon entropy. We introduce a new lossless compression algorithm developed for the NOAA-NESDIS satellite based Earth science multispectral imagers. The algorithm is based on capturing spectral correlations using spectral prediction, and spatial correlations with a linear transform encoder. The algorithm as presented has been designed to work with NOAA's scientific data and so is purely lossless but lossy modes can be supported. The compression algorithm also structures the data in a way that makes it easy to incorporate robust error correction using FEC coding methods as TPC and LDPC for satellite use. This research was funded by NOAA-NESDIS for its Earth observing satellite program and NOAA goals.

1. INTRODUCTION

The MODerate resolution Imaging Spectroradiometer (MODIS) is a key instrument aboard the Terra (EOS AM) and Aqua (EOS PM) polar satellites. Terra's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. The MODIS instrument provides high radiometric sensitivity (12 bit) in 36 spectral bands ranging in wavelength from 0.4 μm to 14.4 μm. These 36 distinct spectral bands are divided into four separate Focal Plane Assemblies (FPA): Visible (VIS), Near Infrared (NIR), Short- and Mid-Wave Infrared (SWIR/MWIR), and Long-Wave Infrared (LWIR). Each FPA focuses light onto a certain section of detector pixels, which are relatively large, ranging from 135 μm to 540 μm square. The large number and variety of detector pixels are what make the wide variety of MODIS data possible. When light hits a detector pixel, it will generate a distinct signal depending on the type of light it is sensitive to. The signals that the pixels generate are what scientists process and study to learn about Earth's land surfaces, water surfaces, and atmosphere.

In this paper we outline a new algorithm for lossless imager data compression. The algorithm can be thought of as a pre-processing step to boost the compression of a purely spatial lossless 2D imaging algorithm such as JPEG2000. The algorithm uses a non-linear statistical method based on histogram specification to remap the intensities in neighboring spectral bands in order to predict intensities. The method presented takes advantage of destriping methods we have introduced in paper.[1] Striping is a well known impairment that affects the radiometric measurements of MODIS. It is due to the anomalous behavior of the input/output transfer function of the single detectors in the FPA. There are 10 detector elements along track for each of the 1 km bands, 20 for each of the 500 m bands, and 40 for the 250 m bands. A more detailed overview of the MODIS data attributes for various band types are shown in Table 1.

The paper first shows some entropy estimated compression ratios and then argues that the state of the art compression software lag with respect to these achievable compression ratios. We then present our algorithm by outlining both the compression as well as decompression part of the lossless compression algorithm.

Further author information: (Send correspondence to I. Gladkova)
E-mail: gladkova@cs.ccny.cuny.edu, Telephone: 1 212 650 6261

Table 1. MODIS data set attributes

Resolution	No of sensors	No of rows	No of columns	No of channels	Memory (bits)	Memory (%)
250m	40	2080*4	1400*4	2	1490944000	%39
500m	20	2080*2	1400*2	5	931840000	%24
1km day	10	2080	1400	14	652288000	%17
1km night	10	2080	1400	17	792064000	%20

Table 2. Estimated compression ratios from entropy estimation

File Name	EV_250m	EV_500m	EV_1km_day	EV_1km_night
MOD01.A2006174.1005.005.2006214124917	3.81	4.04	4.14	5.38
MOD01.A2006175.0915.005.2006214125006	2.84	3.40	3.19	5.31
MOD01.A2006176.0955.005.2006214124831	3.42	3.64	7.23	5.06
MOD01.A2006177.0900.005.2006215131020	3.22	3.46	4.80	5.18
MOD01.A2006179.0850.005.2006214124349	2.75	3.01	3.74	4.66
MOD01.A2006180.0930.005.2006214124332	3.49	3.73	7.09	5.62
MOD01.A2006181.1010.005.2006214124336	3.50	3.77	3.22	5.55

2. COMPRESSION LIMITS AND CURRENT STATE OF THE ART.

We briefly review prior estimates of the maximum compression ratios achievable for MODIS imager data. These estimates have been based on Shannon's information theory which equates the optimal compression rate with the entropy of the images. We have in a previous paper[2] developed an algorithm to estimate compression ratios based on the entropy estimators developed Kozachenko.[3] For our sample data sets, the Table 2 shows the estimated compression ratios that are achievable.

In a paper by the authors,[4] we have run the data sets on the state of the art compression algorithms for image compression for identifying which compression algorithm perform better for MODIS data sets. JPEG2000

Table 3. JPEG2000 compression ratios for MODIS data

File Name	EV_250m	EV_500m	EV_1km_day	EV_1km_night
MOD01.A2006174.1005.005.2006214124917	3.06	2.71	5.19	3.29
MOD01.A2006175.0915.005.2006214125006	2.67	2.48	4.01	3.15
MOD01.A2006176.0955.005.2006214124831	2.75	2.53	5.85	3.25
MOD01.A2006177.0900.005.2006215131020	2.68	2.48	5.43	3.17
MOD01.A2006179.0850.005.2006214124349	2.45	2.25	4.75	2.92
MOD01.A2006180.0930.005.2006214124332	2.74	2.54	5.90	3.34
MOD01.A2006181.1010.005.2006214124336	2.94	2.60	4.18	3.19

consistently outperforms other software that we tried. Table 3 shows the means of the compression ratios for the same 7 MODIS granules defined in the Table 2 using the Jasper open source library. The compression ratios are computed based on 16-bit baseline.*

However, it could be claimed that the algorithms which are primarily developed for 2D images do not achieve the estimated compression ratios and that there is significant gap which needs to be bridged by developing more sophisticated algorithms which take both the spatial and spectral information into account.

3. COMPRESSION APPROACH

State of the art compression techniques take advantage of the spatial correlation found in natural images. However, compression could be bettered by taking the spectral correlation into account. In our algorithm we predict a channel from the previous one to exploit spectral dependencies. The prediction is done in a different domain, defined by histogram specification. After prediction, the difference is compressed by 2D compressor, so our approach can also be considered as pre-processing. However, there are 'more' of spatial correlations than spectral, since the image is 2030 by 1354 spatially and only 36 channels spectrally. We would therefore like to exploit these spatial dependencies by preprocessing images so that they look more correlated with each other and at the same time make the images smoother to get rid of sensor bias.

Before we describe the algorithm we present two pre-processing steps: destriping and offset removal. We use destriping to remove the sensor bias as the images we get from the satellite are very stripy. This introduces discontinuities in the image and standard 2D compression algorithms fail to perform well. In a paper by authors we have examined two statistical methods that can be used to destripe images: histogram equalization and histogram specification. In Figure 1, we show a typical destriped image using one of these algorithms.

Another eccentricity that is present in MODIS images is that there are jumps in gray values for the same sensor. Figure 2 shows a typical set which has these jumps within sensors. We need to remove these offsets so that we have a smoother image per sensor.
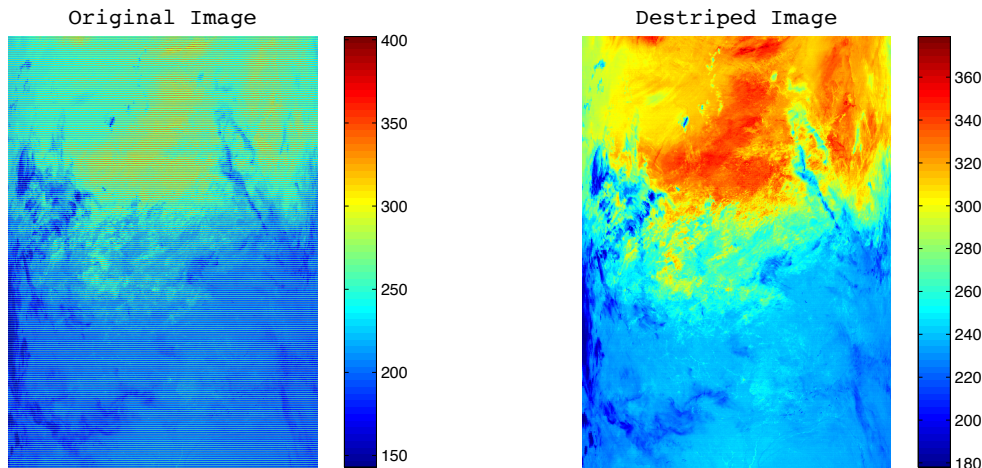

Figure 1. Destriping using histogram specification

4. COMPRESSION ALGORITHM

The imager granules are divided into mulitpectral data blocks. Each block consists of a some number, K, of channels. These channels are homogeneous in the sense that the size of the images in that block (number of rows and columns) are the same. Currently the algorithm is applied to the blocks independently.

---

*The data has a dynamic range that is 12-bit. However, it is typically stored and transmitted as 16-bit integers.
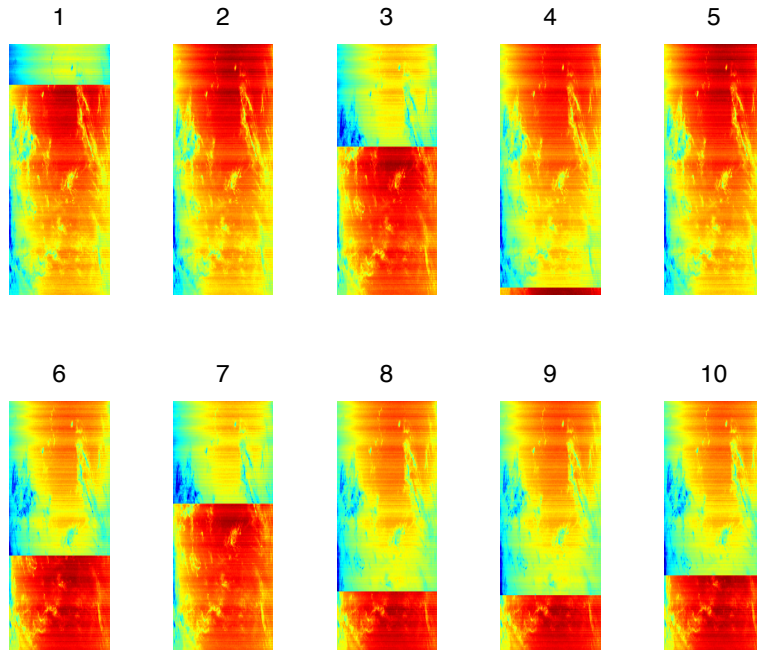
Figure 2. An example of an image which has offsets: per sensor images are shown in the figure.
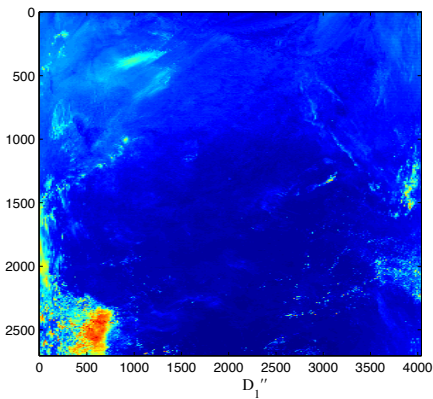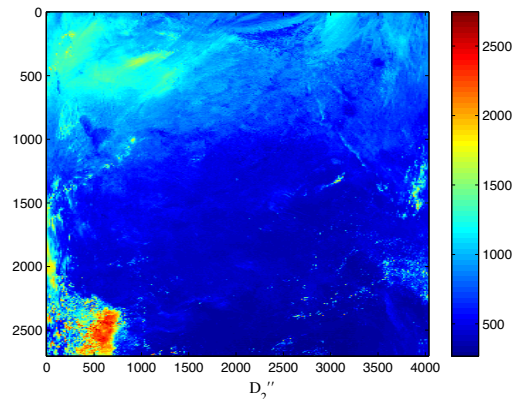

Figure 3. Destriped image D''_1.

Figure 4. Destriped image D''_2.

In the initialization phase of the algorithm, an initial channel D_1 (e.g. channel 1 of the block) is processed to remove a constant offsets. These are large discontinuities that result from a constant being added to the digital counts of a given sensor after some point in the acquisition of the granule (cf. Figure 2). If we denote the result of offset removal as D'_1, then let D''_1 denote image after D'_1 has been destriped losslessly using compressed histogram equalization (cf. Figure 1). Finally, D''_1 is compressed with conventional 2D compression algorithm (e.g. JPEG2000). In our implementation we used JPEG2000 but any 2D compressor can be used for this step.

After the initial image is compressed, we present an inductive algorithm that processes image D_{k+1} assuming that the kth image D_k has already been processed (e.g. $k = 1$). In particular, offsets are removed to yield D'_k, and losslessly destriped to give D''_k, in a previous step as shown in Figure 3. The same processes are performed on D_{k+1} to obtain an image D'_{k+1} with offsets removed and a losslessly destriped image D''_{k+1} (cf. Figure 4). The
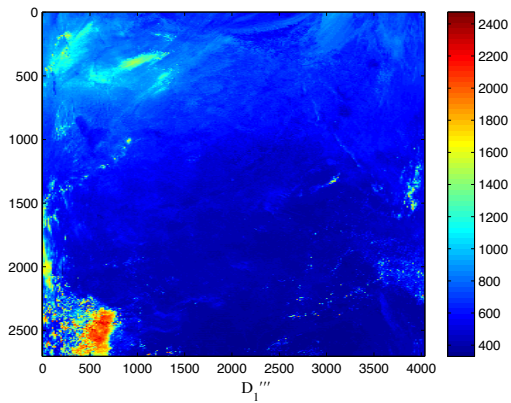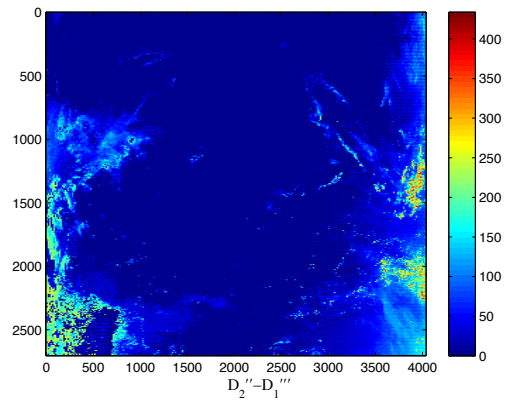
Figure 5. Intensity remapped image D_1'''.
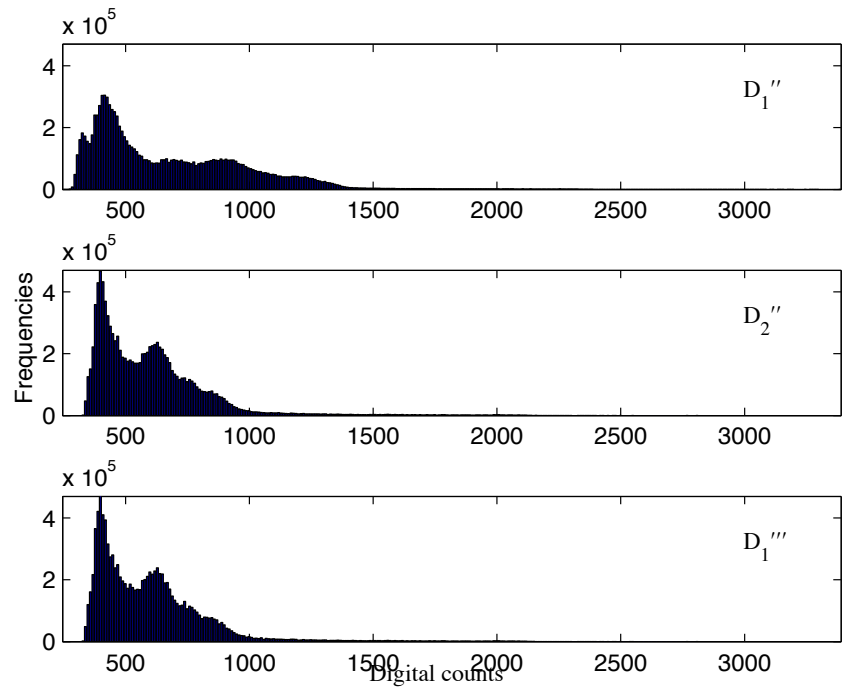

Figure 6. Difference $D_2'' - D_1'''$.

Figure 7. Histograms of D_1'' (top), D_2'' (middle), and D_1''' (bottom).

histograms of D_k'', and D_{k+1}'', are computed (Figure 7). These histograms are used to build an intensity lookup table (LUT) so that after D_k'' is remapped using the LUT, the resulting image, D_k''' (Figure 5) is a predictor for D_{k+1}''. Note because we have assumed that the whole image D_k has already been stored in a prior step, the prediction process that creates D_k''' may be lossy. The difference $D_{k+1}'' - D_k'''$ is computed and compressed as shown in Figure 8.

The compressed data that must be stored to reconstruct losslessly is:

1. A list of values as well as the beginning and end locations where offsets have been applied to the digital counts.

2. Lossless destriping lookup tables for each sensor (or equivalently Histograms for each of the sensor in each image)

3. Prediction lookup tables (or equivalently histograms for the images D_k'')

4. 2D compressed initial image D_1''

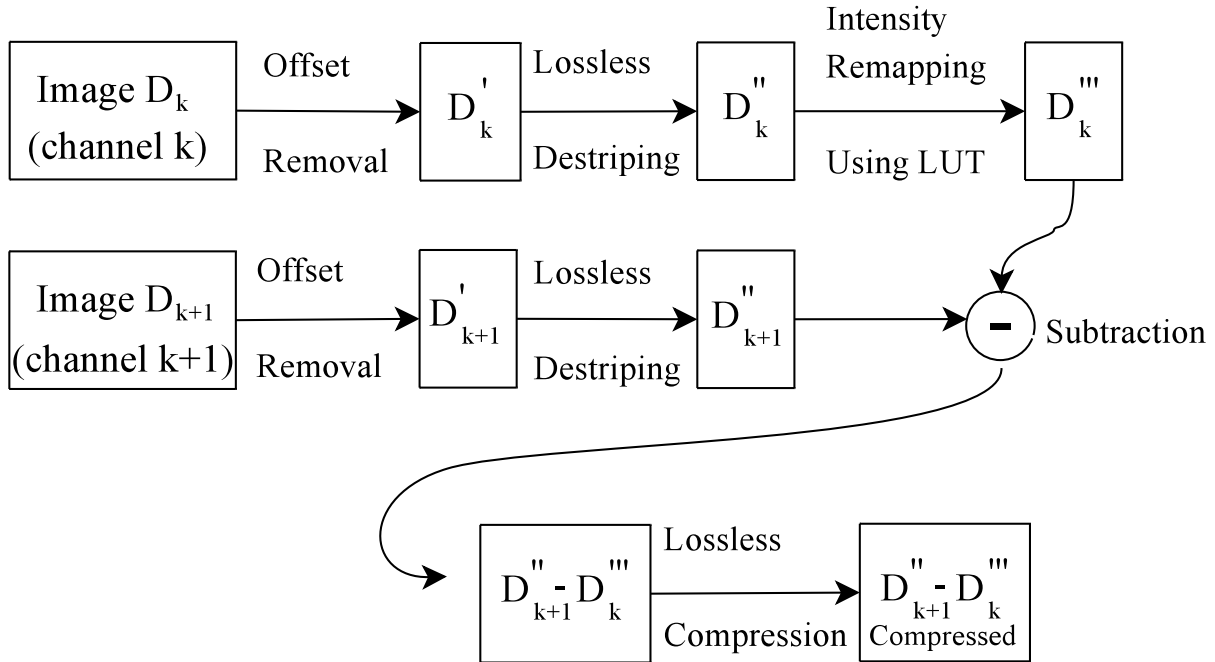5. 2D compressed difference images $D_{k+1}'' - D_k'''$

Figure 8. Block diagram of the iterative step in the Compression algorithm. At each step a destriped and offset corrected (kth) image which has previously been compressed is used to predict the values of the next (($k + 1$)st) image. This is accomplished using a histogram specification technique to build a (non-linear) intensity LUT. The difference between the predicted values and the actual values are stored and compressed.

Due to the preprocessing and the underlying spectral correlations between measurements in D_k, and D_{k+1} the $D_{k+1}'' - D_k'''$ is on average small (cf. Figure 6). As a result the $D_{k+1}'' - D_k'''$ typically compresses well. It is possible that for some images $D_{k+1}'' - D_k'''$ may not compress as well as D_{k+1}'' alone because the channels k and $k + 1$ may sometimes not be correlated. One possible example where we expect this would be if k is a visible channel and $k + 1$ is a water vapor channel. This type of issue is easily handled by simply checking if SizeCompressed($D_{k+1}'' - D_k'''$) > SizeCompressed(D_{k+1}''). In such cases we can break the data block into two

Image D_k
(channel k)

Offset Removal

D'_k

Lossless Destriping

D''_k

Intensity Remapping Using LUT

D'''_k

$D''_{k+1} - D'''_k$ Compressed

Lossless 2D Decompression

$D''_{k+1} - D'''_k$

$+$ Addition

D''_{k+1}

Lossless Restriping

D'_{k+1}
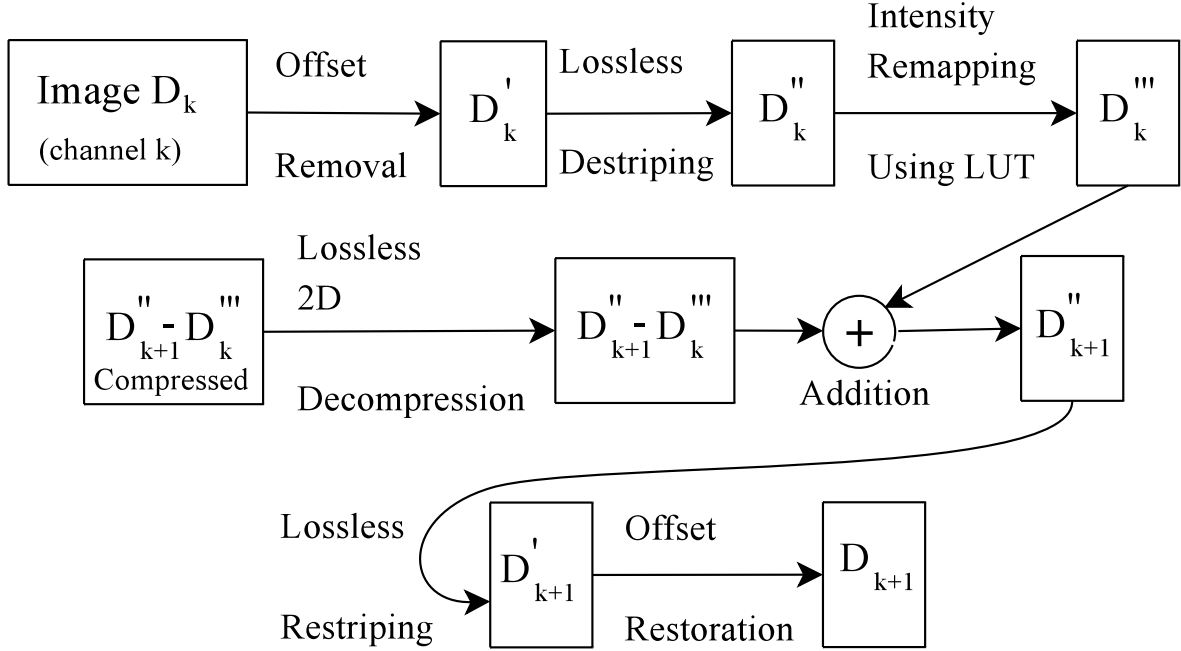
Offset Restoration

D_{k+1}

Figure 9. Block diagram of the iterative step in the Decompression algorithm. As in compression, at each step a destriped and offset corrected (kth) image which has previously been decompressed is used to predict the values of the next (($k+1$)st) image.

sequences, the second sequence starting at $k + 1$. This way we can insure that the algorithm will always improve on the base lossless 2D compression algorithm.

Decompression Algorithm

As we described earlier, compression algorithm runs inductively through the images within the each data block. Initially the first image is decompressed to obtain D''_1. The lossless destriping is inverted on D''_1 to obtain a restriped D'_1. The offsets are then restored to recover D_1. The inductive recovery of the remaining images in the data block is shown in Figure 9. Note that although offset removal and lossless destriping are shown in the diagram for symmetry with the compression diagram, the image D''_k is already computed as a byproduct of the recovery of D_k. To recover D_{k+1} we first decompress the stored difference $D''_{k+1} - D'''_k$. We apply our prediction LUT which has been stored in the compressed file to obtain D'''_k from D''_k. Then, by adding D'''_k to the decompressed difference $D''_{k+1} - D'''_k$ we obtain D''_{k+1}. As in the initial step we know can restripe this image to compute D'_{k+1}. Restoring the offset information which has been stored in the compressed file we finally recover D_{k+1} losslessly.

We should note also that although we have not been able to fully analyze the complexity of our algorithm, histogram specification, equalization, and mapping by LUT are among the fastest image processing operations. Hardware based programmable LUT are common and often part integrated directly into imaging hardware. The pre-processing step we have presented does not represent a significant increase in processing time over the base 2D lossless compression.

5. RESULTS

In this section we will present tables with compression ratios for the 7 MODIS granules defined in the Table 2 as well as average bits per sample needed to store compressed MODIS granules. The first column of Table 4 shows the results of compression of the granules with JPEG2000 using the Jasper open source library. The compression

ratios are computed based on 16-bit baseline.[†] The middle column shows the results of compression of the granule using our prediction algorithm in conjunction with spatial compression to obtain a lossless algorithm which exploits both spectral and spatial redundancy. Despite variation in the granules, our non-linear prediction consistently outperforms JPEG2000 along. For comparison, the third column of table Table 4 shows entropy estimates of optimal compression for the same granules using the methods described in.[3] While these estimates show there is still considerable room for improvement, they show that the new algorithm makes a significant step in optimally compressing the data.

Table 4. Compression ratios for 7 granules using JPEG 2000 (J), the prediction algorithm (P), and approximate achievable ratios from entropy estimation (E)

J	P	E
3.23	3.62	4.17
2.87	3.14	3.35
3.06	3.5	4.12
2.97	3.37	3.79
2.70	2.98	3.23
3.07	3.54	4.25
3.04	3.27	3.80

Table 5. Compression ratios using the prediction algorithm (P) and JPEG 2000 (J) for the 2nd channel in 250m bands on 7 granules

P	3.50	2.57	3.32	3.15	2.71	3.36	2.95
J	2.83	2.38	2.55	2.48	2.30	2.53	2.73

Table 6. Bits per Sample(Bps) using the prediction algorithm (P) and JPEG 2000 (J) for the 2nd channel in 250m bands on 7 granules

P	4.57	6.22	4.81	5.08	5.91	4.77	5.43
J	5.65	6.72	6.28	6.45	6.95	6.32	5.86

As an important example, we present results for the compression benefit to using the first 250m visible channel to predict the second in Table 5. Again prediction gives consistent and significant gains over pure 2D compression algorithms. The is of particular importance because the 250m bands represent a type of "worst case scenario" with respect to spectral dependencies. In particular this portion of the MODIS data has very

---

[†]The data has a dynamic range that is 12-bit. However, it is typically stored and transmitted as 16-bit integers.

Table 7. Compression ratios using the prediction algorithm (P) and JPEG 2000 (J) for 500m bands on 7 granules

	2	3	4	5
P	3.52	2.73	3.12	3.05
J	3.19	2.51	2.54	2.22
P	3.42	2.49	2.52	2.81
J	2.85	2.29	2.31	2.12
P	3.53	2.58	3.01	2.98
J	2.90	2.37	2.36	2.09
P	3.41	2.50	2.87	2.79
J	2.82	2.33	2.32	2.07
P	2.91	2.28	2.37	2.53
J	2.42	2.16	2.18	2.02
P	3.57	2.66	3.07	3.00
J	2.88	2.38	2.37	2.10
P	3.17	2.49	2.88	2.90
J	3.00	2.42	2.45	2.19

high spatial resolution and very low spectral resolution. Thus one expects only modest gains over state of the art conventional image compressors. In light of such expectation the compression gain we have achieved is remarkable. The same results are shown in Table 6 in terms of bits per sample.

The 500m bands make up a significant portion of the MODIS data partially as a result of their relatively high spatial resolution. Hence the performance of compression of these can have a significant impact on compression. Table 7 and Table 8 show the improvement here based on our the prediction based algorithm. The tables are expressed both in terms of compression ratios and bit per sample for the 4 channels where prediction was used, You will note that the compression improvement is significant. Here only relationships between neighboring channels is being exploited leaving opportunities for future work to exploit relationships between non-adjacent channels.

6. CONCLUSION

The algorithm we have presented is able to significantly improve compression of imager data by exploiting both spatial and spectral correlations. We take into account sensor specific characteristics. Our non-linear spectral prediction is based on intensity statistics. By carefully removing sensor artifacts and applying our intensity prediction we are able to preserve spatial correlations in a manner so that conventional compression algorithms can exploit them. Thus our method can be thought of as a lossless preprocess to 2D spatial algorithms such as JPEG2000. We also insure that our preprocess will improve the base algorithm by testing at each stage of the inductive decompression. For MODIS the algorithm shows significant improvement for the 250 meter, 500 meter and 1 kilometer day channels. This is important because these channels make up the majority of the data size and thus they have the most significant impact on overall compression. The remaining 1 kilometer night channels are emissive rather than reflective and thus it is not surprising that they have different statistical properties. Given the discrepancy between the optimal compression estimates from the computed entropy and the compression ratios using conventional compression, it is likely that it should be possible to compress these channels as well but this is the subject of future work.

Table 8. Bits per Sample(Bps) using the prediction algorithm (P) and JPEG 2000 (J) for 500m bands on 7 granules

	2	3	4	5
P	4.55	5.87	5.13	5.24
J	5.02	6.38	6.31	7.21
P	4.68	6.43	6.35	5.69
J	5.62	7.00	6.94	7.55
P	4.53	6.20	5.31	5.37
J	5.52	6.74	6.79	7.67
P	4.70	6.40	5.58	5.73
J	5.67	6.86	6.88	7.73
P	5.51	7.01	6.75	6.32
J	6.62	7.41	7.35	7.93
P	4.49	6.02	5.21	5.34
J	5.55	6.72	6.74	7.61
P	5.04	6.43	5.56	5.51
J	5.33	6.62	6.52	7.31

7. ACKNOWLEDGMENTS

REFERENCES

1. M. Grossberg, S. Gottipati, I. Gladkova, and T. Schmit, "The impact of striping artifacts on compression," in *Proceedings of SPIE – Satellite Data Compression, Communications, and Archiving III*, 2007.
2. M. Grossberg, S. Gottipati, I. Gladkova, M. Goldberg, and L. Roytman, "An analysis of optimal compression for the advanced baseline imager based on entropy and noise estimation," in *Proceedings of SPIE – Satellite Data Compression, Communications, and Archiving II*, W. Heymann, C. W. Wang, and T. J. Schmit, eds., **6300**, September 2006.
3. L. F. Kozachenko and N. N. Leonenko, "Sample estimate of the entropy of a random vector," *Problems of Information Transmission* **23**(2), pp. 95–101, 1987.
4. S. Gottipati, J. Goddard, M. Grossberg, and I. Gladkova, "A comparative study of lossless compression algorithms on modis data," *Proceedings of SPIE – Satellite Data Compression, Communications, and Archiving III* , 2007.