A Lossy Source Coding Algorithm Using The Least Significant Bits

B.Purna Kumari

Abstract— Enhanced radiometric and geometric specifications of the sensors, on-board remote sensing satellites, result in large volumes of data that has to be downlinked. Several compression techniques like JPEG, JPEG 2000, with different encoding, are being employed to compress the data. Here a simple lossy source coding algorithm, using the least significant bits (LSBs), is being proposed to reduce the number of bits to be transmitted. The maximum loss in information of the digital number count will be one-fourth the maximum value of the number of the LSBs. Some of the advantages of the algorithm are good compression with minimal loss in information, conducive for acquiring high bit data, simple and inexpensive compression logic and the algorithm does not introduce artifacts.

Index Terms— source code, data compression, encoding, decoding, least significant bits, quantization, bit reduction, data volume

1 Introduction

The process of reducing the size of data is referred to as data compression. In the context of data transmission, it is called source coding (encoding done at the source of the data before it is stored or transmitted). In signal processing, data compression, source coding or bit-rate reduction involves encoding information using fewer bits than the original representation [1]. In view of the ever increasing data to be stored and transmitted, data compression has become order of the day because it reduces the resources required to store and transmit data. Compression can be either lossy or lossless. Compression reduces bits by removing coding, spatial and temporal redundancy and irrelevant information [2]. No information is lost in lossless compression so that the process is reversible ex: Lempel–Ziv (LZ), Lempel–Ziv-Renau (LZR) etc. Lossy compression involves encoding methods that use inexact approximations and partial data discarding to represent the content. Computational resources are consumed in the compression process and, usually, in the reversal of the process (decompression). The design of data compression schemes involves trade-offs among various factors, including the degree of compression, the amount of distortion introduced (when using lossy data compression), and the computational resources required to compress and decompress the data [1]. Various image compression systems employed on board satellites, development trends and other related issues have been discussed at length by Guoxia Yu et al.,. and Rafael C.Gonzalez et al.,. [2],[3].

There is a growing trend in the development and launch of mini, cube, pico, nano and micro satellites with payloads catering to specific applications. Some of the examples are Indian Space Research Organization's (ISRO) IMS-1, ANUSAT, STUDSAT, Jugnu, SRMSat, SWAYAM, Sathyabamasat, PISAT, PRATHAM, NIUSAT, Youthsat, Microsat and the Doves of Planetlabs. In view of the limited resources and cost, it is desirable to achieve best possible data from the payload by using innovative techniques onboard. Higher radiometric, spatial, spectral and temporal resolutions are some of the desirable specifications of the data from the remote sensing satellites. Higher spatial resolution is achieved by selecting an appropriate orbit, large number of detectors, Instantaneous Field Of View (IFOV) and Time Delay Integration (TDI). A large number of detectors are employed to achieve a desirable swath Ex: 12000 detectors with 0.8m resolutions to get a swath of 9.6 Kms. To further exploit the details provided by high spatial resolution data, it is desirable to collect the data with higher radiometric resolution. High radiometric resolution (quantization levels), geometric resolution (high resolution) and spectral resolution (number of spectral bands) result in higher data volumes. The data may not get transmitted to ground in real time, within the visibility of a ground station. This in turn calls for higher onboard storage for transmission at a later stage, preferably after compression through source code encoding to reduce the bandwidth required. In the case of small and low cost satellites, the on-board compression and the ground decompression logics have to be simple in terms of computational and hardware resources.

ISRO has been employing on-board JPEG/JPEG like compression of data of its satellites from Cartosat-1 onwards. In Resourcesat-1 LISS-IV sensor, data were acquired in 10 bits and the seven most significant bits / selected seven bits were transmitted reducing the data to seven bit.

In this paper, a lossy source coding algorithm is proposed where the original quantization levels can be retained with a marginal loss in information, using the LSBs.

- *Author is currently working as a Scientist for National Remote Sensing Centre, ISRO, Hyderabad,India*
  E-mail : purnakumari_b@nrsc.gov.in

2 Objective

To develop a compression technique with minimal loss in information, with minimal computational resources and achieve higher number of bit reduction.

3 Algorithm

In the proposed algorithm, the least significant bits of data with high quantization levels, of the order of more than 10 bits, are encoded to achieve the compression with minimal loss of information.

3.1 Pre-requisite

Most of the electro-optical imaging systems use multiple Charge Coupled Devices (CCDs) per band. For instance, in the case of Cartosat-2S PAN sensor, there are two detectors, each with 8000 CCD elements to image over a swath of 10Kms. Since all the detectors will not have the same response, it is necessary to adjust the electronic gain in each channel so that the output voltage (the digital number in digital transmission systems) is same for equal radiance input. This is carried out by elaborate ground calibration. The ground calibrated data is stored in the form of a look-up-table (LUT) and applied either on-board or during pre-processing [5]. In the case of Indian Remote Sensing (IRS) satellites, like Resourcesat-2 and Cartosat-2S, there is a provision to uplink and store the detector non-uniformity correction (NUC) data. So, there shall be a provision to correct the source code for correcting the non-uniform response before applying the encoding algorithm.

Secondly, since the performance of the detectors is likely to deteriorate with time, there shall be provision to acquire the data in compression bypass mode to re-work the LUT.

3.2 Encoding process

Let N be the number of original quantization bits.
Let M be the number of bits used to transmit the data after source coding.
Let P = N-M
Do not disturb the M-1 number of MSBs of the original data
Read out the LSBs beyond M-1
Replace the Mth LSB with a

- value 0 if the combined value of the P+1 number of LSBs in the original data is less than half the value of max value of P+1 bits (all 1s).
- value 1 if the combined value of the P+1 number of LSBs in the original data is greater than half the value of max value of P+1 bits (all 1s)

3.3 Decoding process

Reconstruct the original N bit data by replacing the LSB

- P+1 bits with a value 1/4 the max value of P+1 bits (all 1s) if the LSB is 0
- P+1 bits with a value 3/4 the max value of P+1 bits (all 1s) if the LSB is 1

3.4 Example

Let the original quantization be 11 bits
Let the number of bits used to transmit the data after source coding be 8.
Number of bits reduced = 3
Do not disturb 7 MSBs of the 11 bit original data
Read out the LSBs beyond 7th bit, i.e., the 4 LSBs
Replace the 8th bit with a

- value 0 if the combined value of the 4 LSBs in the original data is less than or equal to 8
- value 1 if the combined value of the 4 LSBs in the original data is greater than 8

During decoding, the 8th bit is replaced with 4 bits corresponding to a value 4 if it were 0 and with bits corresponding to a value 12 if it were 1.

4 Observations and Results

The algorithm was tested for images acquired over different terrains. Fig 1 depicts an part if an image captured by ISRO's Cartosat-2S satellite, multi-spectral sensor. The sensor provides data with a spatial resolution better than 2m in blue, green, red and near infra-red bands with 11 bit quantization. JPEG 2000 like compression is implemented on board these satellites to achieve data compression. Fig 2 depicts the simulated image generated using the proposed algorithm wherein the 4 LSBs have been encoded.

The original and encoded multi-spectral images were subjected to unsupervised classification (ISODATA) using 10 classes (Fig 3). As can be seen in Fig 3, there is no difference in the classification results. The data were tested for fifty classes and no change was observed in classification. There could be a miniscule change in the classification results only when a large number of classes of the order of more than hundred, are used.

Fig 4 and 5 depict the original, encoded images collected by Cartosat-2S Panchromatic camera (11 bit). Fig 6 depicts the graph of one-line of original, encoded images and the difference in DN value. It can be observed that the maximum loss of information is +/- 4 digital counts out of 0 to 2047 counts per pixel (Fig 7).

Fig 1. Original image captured by Cartosat-2S Mx (256 Lines x 256 Pixels)

Fig 2. Cartosat-2S Mx Image with the proposed algorithm implemented

Input image classification results

Row	Histogram	Color	Red	Green	Blue	Opacity	Class_Names
0	0		0	0	0	0	Unclassified
1	6956		0.21	0.2	0.37	1	Class 1
2	7878		0.3	0.27	0.56	1	Class 2
3	6593		0.36	0.35	0.52	1	Class 3
4	5657		0.41	0.42	0.44	1	Class 4
5	5401		0.47	0.47	0.43	1	Class 5
6	4922		0.51	0.54	0.43	1	Class 6
7	4386		0.56	0.6	0.47	1	Class 7
8	4653		0.66	0.67	0.52	1	Class 8
9	5512		0.82	0.83	0.61	1	Class 9
10	3594		1	1	0.78	1	Class 10

Output image classification results

Row	Histogram	Color	Red	Green	Blue	Opacity	Class_Names
0	0		0	0	0	0	Unclassified
1	6956		0.21	0.2	0.37	1	Class 1
2	7878		0.3	0.27	0.56	1	Class 2
3	6593		0.36	0.35	0.52	1	Class 3
4	5657		0.41	0.42	0.44	1	Class 4
5	5401		0.47	0.47	0.43	1	Class 5
6	4922		0.51	0.54	0.43	1	Class 6
7	4386		0.56	0.6	0.47	1	Class 7
8	4653		0.66	0.67	0.52	1	Class 8
9	5512		0.82	0.83	0.61	1	Class 9
10	3594		1	1	0.78	1	Class 10

Fig 3. Isodata unsupervised classification results of the input and output images.

The net number of pixels that are affected will be much less than the total number of pixels (whenever the difference is zero) and the deviation is both positive and negative from the original value. There will not be any artifacts since all the pixels undergo the transformation - by a small number of DN count change. There will be a change, only in the unit's place of the DN, even if five LSBs are encoded. This is a better option than reducing the quantization levels since the finer details discerning the features can be captured with a minimal loss of information (ex : can achieve a range of 0-2047 DN values using the same 8 bits instead of 0-527 values).

In the case of images acquired over uniform terrains, no striping was seen when the algorithm is applied after applying the detector non-uniformity correction, since all the identical pixels undergo the same change in DN count.

In the case of images with varying features, encoding effects like local block effect were not observed since the maximum deviation in DN number of neighbouring pixels, from the original values, will be half the value of the maximum value of the LSBs encoded – eg., eight in the case of a eleven bit image where four LSBs are encoded.

Fig 4. Original image captured by Cartosat-2S PAN sensor

Fig 5. Cartosat-2S PAN Image with the proposed algorithm implemented

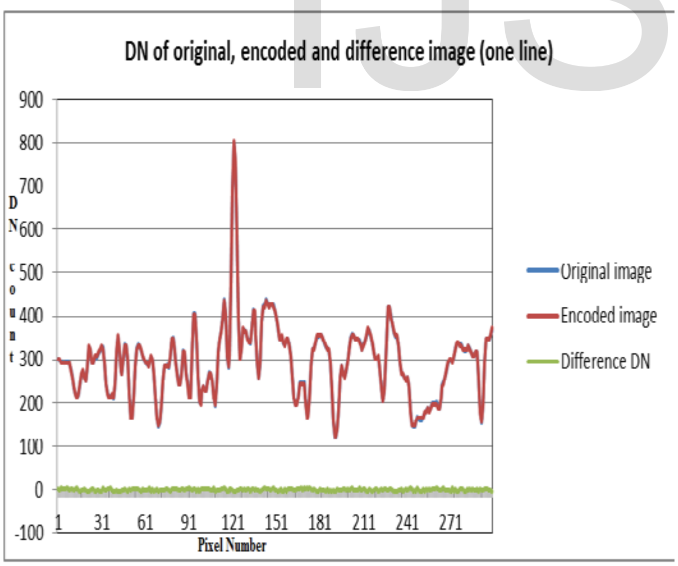

Fig 6. Graph depicts one-line of original, encoded Panchromatic images and the difference in DN value

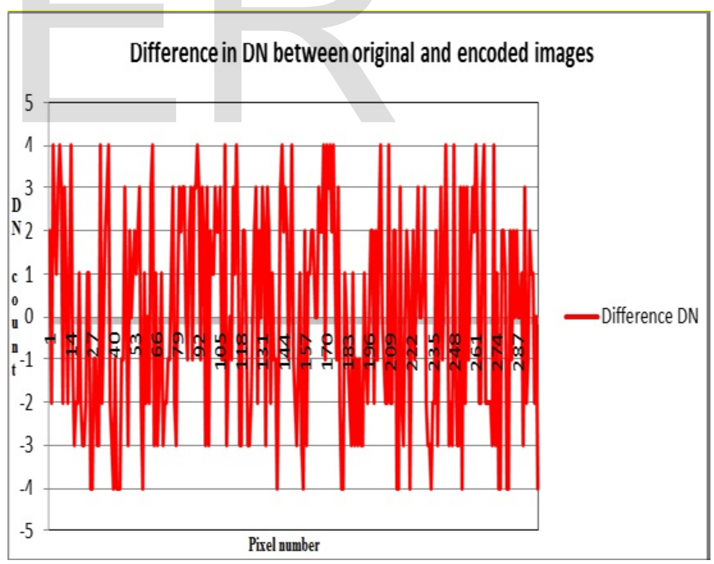

Fig 7. Graph depicts difference in DN values between original and encoded Panchromatic images (one-line of the images)

ISSN 2229-5518

5 MEASURE OF PERFORMANCE

A compression algorithm can be evaluated in a number of different ways. We could measure the relative complexity of the algorithm, the memory required to implement the algorithm, how fast the algorithm performs on a given machine, the amount of compression, and how closely the reconstruction resembles the original [4]. Some of the parameters considered for evaluation, acceptable limits and study results are tabulated in Table1. These are some of the parameters used for evaluating the performance of the compression logic used for ISRO's, Cartosat-1) [6].

TABLE 1
Algorithm evaluation parameters, limits and study results

Parameter	Limits	Study results
Root Mean Square Error (RMSE)	Local – should not exceed 2 counts in a block of 128 Global- 2 counts with a variation of maximum 0.5 (as obtained from literature) The desired number from the sensor characterization is 1 count with maximum variation of 0.5	The local RMSE – column wise, was observed to be between 0.8 to 1.006 Global RMSE was 1 in the images taken up for evaluation
Histogram comparison	97 % within +2 counts difference	100% within ± 1/4th the maximum value of the LSBs encoded.
Absolute error statistics (global)	99% within 10 counts	100% within ± 1/4th the maximum value of the LSBs encoded.
Acceptability with respect of land use applications	Classification accuracies, statistical measures with 95% confidence interval for measure of dispersion	Unsupervised classification results are not affected even when large number of classes of the order of 50, are considered.

Table 2 gives the details of deviation in DN of each pixel when four LSBs of a 11 bit data set are encoded.

TABLE 2
Deviation in the DN count from the original value

Original value of the 4 LSBs	Deviation in the modified pixel from the original value
15	-3
14	-2
13	-1
12	0
11	1
10	2
9	3
8	-4
7	-3
6	-2
5	-1
4	0
3	1
2	2
1	3
0	4

The maximum deviation in the DN number of each pixel is 0 to +/- 4 in a 11 bit encoded data (2048 levels)

Table 3 gives the details of the maximum deviation in DN value of each pixel of the original data of 10, 11, 12 and 13 bits and transmitted in coded eight bits.

TABLE 3
Bits encoded Vs Deviation in the DN count

No of Bits of original data	Max DN value	Assumed number of Bits transmitted	Number of Bits reduced	Max value of LSBs encoded	Maximum deviation in DN value of each pixel
10	1023	8	2	7	-2 to +2
11	2047	8	3	15	-4 to +4
12	4095	8	4	31	-8 to +8
13	8191	8	5	63	-16 to +16

In general

- The deviation in the DN number of each pixel will be 1/4th the maximum value of the number of LSBs.
- The bit reduction increases with the increase in the quantization levels, if the number of bits transmitted are at kept at a certain chosen value (say 8). At the same time

the deviation from original also increases but confines to the unit place value up to 4 LSBs.

6 Conclusion

The proposed algorithm offers a means of simple source coding with minimal loss of information, simple compression and decompression logic.

- The loss of information in each pixel is 1/4th of the maximum value of LSBs encoded.
- No artifacts are introduced as in the case of other compression techniques like DPCM [7].
- This can be implemented in the payloads providing data with high quantization (more than 11 bits) and for small, low cost satellites like Nano, Pico and mini satellites.
- The data can be further compressed by using other proven algorithms.
- By encoding the encoded LSBs in the manner explained, the loss in information can be reduced to ±1 DN count. To make the process lossless, the lost information can be stored and transmitted separately and restored on ground.

Acknowledgment

The author is grateful to the management of NRSC and colleagues and management at SDSC, SHAR, LPSC for giving an opportunity to work on this topic as a part of the Structured Training Programme.

References

[1] Wade, Graham (1994). Signal coding and processing (2 ed.). Cambridge University Press. p. 34. ISBN 978-0-521-42336-6. Retrieved 2011-12-22.

[2] Image compression systems on board satellites – Guoxia Yu, Tanya Vladimirova, Martin N. Sweeting – ScienceDirect – Acta Astronautica 64 (2009) 988-1005

[3] Rafael C.Gonzalez, Richard E. Woods, 2008, *Digital Image Processing*, 3rd edition, Prentice Hall

[4] Khalid Sayood, University of Nebraska, Introduction Data Compression, 2006 by Elsevier Inc.

[5] Fundamentals of Remote Sensing, third edition, George Joseph and C Jeganathan.

[6] Cartosat-1 mission document (classified).

[7] Pushpakar A,Dube N,Dhar D, Ramakrishnan R, 2013, Resourcesat-2 image restoration for differential pulse code modulation compression artifacts, IEEE Second International Conference On Image Information Processing.