A New Lossless Compression Method for Small Satellite On-Board Imaging

Sofiane Atek and Tanya Vladimirova
Surrey Space Centre
School of Electronics, Computing and Mathematics
University of Surrey
Guildford, Surrey, GU2 7 XH
United Kingdom

Abstract: - Small satellites performing remote sensing missions are equipped with better imaging systems. However, technical limitations of the downlink system create a bottleneck preventing transmission of large amounts of raw data to the ground station. The use of compression on-board a small satellite will allow to offset the downlink problem. In this paper we describe recent research work on lossless image compression using predictive neural networks, Peano Hilbert scan and Integer Wavelet Transforms.

Key-Words: - Lossless Image Compression**,** Neural Networks, Prediction, Peano-Hilbert Scan, Integer Wavelet Transforms.

1 Introduction

Nowadays, due to advances in technology, space equipment developments are more affordable. Small satellites performing remote sensing missions are equipped with better imaging systems. Indeed their remote sensing capabilities are getting closer to those of big satellites. However, technical limitations of the downlink system create a bottleneck preventing transmission of large amounts of raw data to the ground station. The use of compression on-board a small satellite will allow to offset the downlink problem.

A lot of work has been carried out in the field of on-board lossy image compression while less attention has been paid to lossless compression methods. This is because lossless compression methods are considered to have unsatisfactory performance yielding low compression ratios. However, lossless compression allows retrieving images perfectly, for a multipurpose use.

In this paper we describe recent research work on lossless image compression using neural networks for predictive encoding. Previously, similar technique was used for a text compression application [1]. The proposed image compression method is compared to the state-of-the-art in lossless image compression, the CALIC algorithm.

The testing results presented in the paper are based on a traditional set of gray-scale images used for benchmarking of compression methods, which enables comparison against established techniques such as the CALIC algorithm, LOCO-I and the RICE algorithm. Testing experiments based on satellite images are in a process of undertaking.

Our results will show that for some images such as Lena and Goldhill, the ratios obtained for the neural network based prediction method outperformed largely the CALIC algorithm. While for some images such as Baboon and Peppers, the CALIC method did slightly better. All experiments were carried out using Matlab and Microsoft Visual Studio.

The paper is organized as follows. In Section 1 an overview of the proposed method is presented. In Section 3, we describe the neural network predictor used to perform the entropy encoding with an arithmetic encoder. In Section 4, we describe the Integer Wavelet Transform (IWT) method. Experimental results are shown in Section 5. The conclusions are drawn in Section 6.

2 Overview of the Method

Satellite images can get corrupted due to errors that occur during image capture, or bad climatic conditions, such as clouds. A very useful capability is to be able to preview the transmitted data at any time of the transmission. This will help to save transmission time. Such a capability can be provided by the wavelet transforms. The IWT has the property to construct integer outputs for integer inputs and guarantees a perfect reconstruction of the image whereas after using a Discrete Wavelet Transform (DWT) decorrelation stage there can be losses in the data. [2].

A block-diagram of the proposed method is given in Figure 1.

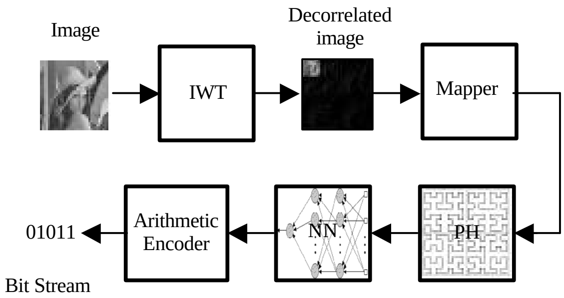

Figure 1: Block diagram of the proposed method

First, the data is fed into an IWT decorrelator. The transformed and mapped image is then scanned with a Peano Hilbert scan (PH) that exploits the redundancies within the two dimensions of the image. The effect of the Peano Hilbert scan on the compression ratios will be illustrated by experimental results in Section 3.1. The mapper can be used either before or after the scanning stage. It has the role to convert each IWT coefficient into a nonnegative integer value.

The predictor, a two-layer, $4x10^6$ by 1 neural network, scans the stream of the mapper output characters, and allocates a probability distribution for the next incoming character.

In proportion to this probability distribution, an arithmetic encoder, starting from an initial interval [0,1), subdivides this range repetitively until obtaining a binary fraction within the final subdivided range. This binary fraction is the shortest number to represent the data and it is known to be optimal within one bit of the data entropy.

3 The Neural Network Predictor

When modelling a particular data distribution we derive the joint distribution $P(y \mid x)$, where y is the following bit and x the previous sequence. According to the maximum entropy principle we have [1]:

$$P(x,y) = 1/Z \prod_i a_i^{f_i(x,y)} \qquad (1)$$

where $f_i(x,y)$ is an arbitrary "feature" function equal to x_i in this case, and a_i are parameters that are found by generalized iterative scaling (GIS) [1]. Z is a normalization constant to make the probabilities sum to 1.

When we take the log of the expression (1) and employ the property $P(y|x)=P(x,y)/P(x)$, we obtain the following expression:

$$P(y \mid x) = 1/Z' \exp(\sum_i \log(a_i)\ x_i) \qquad (2)$$

where Z' is set so that

$$P(y=0 \mid x) + P(y=1 \mid x) = 1 \qquad (3)$$

When simplifying (2) we obtain :

$$P(y \mid x) = g(\sum_i w_i\ x_i) \qquad (4)$$

Where $w_i = \log(a_i)$ and $g(x) = 1/(1+e^{-x})$.

This is a two-layer perceptron neural network with N input units x_i and one output $P(y=1|x)$. We train on-line the network by adjusting the weights w_i so that we minimize the error $E=y-P(y)$, between the real output y and the predicted output $P(y)$. We update the weights using the expressions:

$$w_i = w_i + \Delta w_i \qquad (5)$$

$$\Delta w_i = h\ x_i\ E \qquad (6)$$

where η is the learning rate.

After the prediction step an arithmetic encoder begins with a range [0, 1), and divides the range into two sub-ranges for each input bit. The use of an arithmetic encoder is justified by the fact that it performs better than the other known classic compressors. It has been shown that the arithmetic encoder is optimal within one bit of the entropy [1]. In our experiment we use a two-layer neural

network with a size N= $4x10^6$ by 1. The N inputs are divided into 6 sets.
The weight update is carried out according to the following expression:

$$w_i = w_i + (R_S + R_L/\text{sigma}^2)\, x_i E \qquad (7)$$

where R_S and R_L are the short and long term learning rates.
The variance of the training data in context x_i is given by

$$\text{sigma}^2 = \frac{(C_0 + d)(C_1 + d)}{(C_0 + C_1 + 2d)} \qquad (8)$$

where C_0 and C_1 are the counts of y=0 and y=1 in context x_i. The value d is a parameter between 0 and 1 so that it avoids a division by 0.
R_S is between 0 and 0.5, in our case R_S=0.02. For stationary data, 0 is optimal. R_L should be between 1/6 and 1, here R_L=0.38.

3.1 Effect of Scan on the Neural Network Predictor

To investigate the effect of scan on the compression ratios of the NN-based method, we applied different scans such as the spiral scan and the Peano-Hilbert scan before the NN predictor block. We also tested a Burrow Wheeler Transform; used as a decorrelation stage (see figure 2). The results are listed in Table 1.

Images	NN+PH	NN+Spir	NN+BWT	NN	CALIC
Lena	4.560	2.962	3.997	2.924	2.062
Barbara	3.054	3.039	2.789	2.910	1.545
Goldhill	3.550	3.483	3.344	3.504	1.565
F16	4.701	4.572	4.761	4.773	1.861
Baboon	1.169	1.135	1.168	1.179	1.392
Peppers	1.469	1.441	1.412	1.427	1.778
Splash	1.795	1.837	1.769	1.790	2.172
House	1.676	1.672	1.708	1.735	2.073
Average	2.747	2.518	2.619	2.530	1.806

Table 1: Compression ratios after applying the different scans and the Burrow Wheeler transform

In Table 1, "NN" stands for neural network predictor method using a raster scan, "BWT" stands for the Burrow Wheeler transform, "Spir" stands for the spiral scan and "PH" for the Peano Hilbert scan.

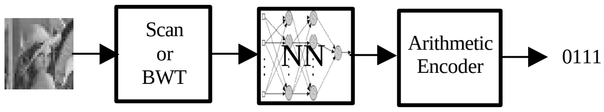

Figure 2: Block diagram of the experiment.

Table 1 shows that the PH scan brought an improvement of 8.5% to the NN based method. The Spiral scan degraded the ratio by 4.9%. The Burrow Wheeler Transform affected the compression ratio of Lena, but did not outperform NN for the other images.

4 Integer Wavelet Transform Based Compression

IWT is a wavelet transform that constructs an integer wavelet coefficient output for an integer input. The oldest IWT is an integer version of the Haar Transform, which is called the S transform. The filters used are [3, 4]:

$$d[n] = x[2n+1] - x[2n] \qquad (9)$$

$$s[n] = x[2n] + \left\lfloor \frac{d[n]}{2} \right\rfloor \qquad (10)$$

Where x is the input vector, s the low pass output vector and d the high pass vector.
Said and Pearlman came out with an improved method that performs the S transform plus a Prediction step in order to generate a new set of highpass coefficients. This transform is called the S+P transform [5].

4.1 The Lifting Scheme

The Lifting Scheme (LS) is a method to construct biorthogonal wavelets [6], which are required Those latter to obtain a perfect reconstruction of the image [7].
First the input data are split into even and odd samples. The even signal is convolved with a low pass filter, and is added to the odd signal. The roles of the signals are inverted and then we can apply M lifting steps. In the output of the process

we will have one high pass signal $d^{(n)}$ and a low pass signal $s^{(n)}$.

We call the lifting steps that use the low pass signal "prediction steps", and those, which use the high pass signal, the "update steps" [2]. To reconstruct the original signal, we simply use the same structure but just in the reverse order.

It was shown in [2] that LS insures a perfect reconstruction of the signal. One important property is that we can replace the linear filters by nonlinear ones and preserve the Perfect Reconstruction (PR) properties. We generate the IWT coefficients by introducing a rounding operation.

The Lifting steps are summarized in what follows [8]:

1. The input x(n) is split into even and odd indexed samples:

$$s^{(0)}[n] = x[2n] \qquad (11)$$

$$d^{(0)}[n] = x[2n+1] \qquad (12)$$

2. M alternating lifting steps are applied:

$$d^{(i)}[n] = d^{(i-1)}[n] - \left\lfloor (\sum_k p^{(i)}[k] s^{(i-1)}[n-k]) + \frac{1}{2} \right\rfloor \qquad (13)$$

$$s^{(i)}[n] = s^{(i-1)}[n] + \left\lfloor (\sum_k u^{(i)}[k] d^{(i)}[n-k]) + \frac{1}{2} \right\rfloor \qquad (14)$$

for $i = 1....M$

3. Finally, a scaling factor is applied:

$$d[n] = K d^{(M)}[n] \qquad (15)$$

$$s[n] = s^{(M)}[n] / K \qquad (16)$$

In the literature [3, 8, 9, 10, 11] many transforms were used to perform the IWT. In equation (17) and (18) the (2,2) IWT transform is presented. Here the notation $(N, \tilde{N})$ represents a transform with N and $\tilde{N}$ vanishing moments in the analysis and synthesis high pass filters respectively [11]:

$$d[n] = x[2n+1] - \left\lfloor \frac{1}{2}(x[2n] + x[2n+2]) + \frac{1}{2} \right\rfloor \qquad (17)$$

$$s[n] = x[2n] + \left\lfloor \frac{1}{4}(d[n-1] + d[n]) + \frac{1}{2} \right\rfloor \qquad (18)$$

The choice of this filter is justified by the fact that it requires less calculations and thus it is ideal for an onboard implementation.

4.2 The Two-Dimensional Integer Wavelet Transform [12]

To apply the integer wavelet transform on images, we have to compute two steps. First a one-level forward IWT is applied on the horizontal dimension. The result of this will be two integer output sub-images. We then use another one-level forward IWT in the vertical dimension. The first operation in the horizontal direction produces one low frequency image L and a high frequency image H. The second operation produces four subimages: one low-frequency subimage LL and three high-frequency subimages LH, HL and HH. The next level of the IWT will process the LL sub-image using the same procedure.

4.3 The Mapper

The mapper has the role to convert each wavelet coefficient y_i to a nonnegative integer value z_i. To make sure that more probable symbols are encoded with shorter codewords, the preprocessed symbol z_i should satisfy:

$$p_0 \geq p_1 \geq p_2 \geq ...p_j \geq ...p_{(2^n-1)} \qquad (19)$$

Where p_j is the probability that z_i equals the integer j, for $i = \{0,1,2,...,2^n - 1\}$.

The prediction error mapping function used in section 5 is based on the mapper used in the Rice Algorithm [13], because it is easy to implement in a low complex program. The mapper function is described as follows

$$z_i = \begin{cases} 2y_i & 0 \le y_i \le q \\ 2|y_i| & -1-q \le y_i \le 0 \\ q+|y_i| & otherwise \end{cases} \quad (20)$$

where θ = y_{min}. This mapping function has the property that $p_i < p_j$, $\forall |y_i| > |y_j|$.

5 Experimental Results

After decomposing the images using the (2,2) filter for 1-level of decomposition we applied the NN+PH method and the CALIC algorithm independently on each sub-image. We noticed that the NN+PH method performed better than CALIC in high frequency sub-images. We increased the level of decomposition and compared the results. The best performances were obtained for n=2. This value gives a good tradeoff between the ratio and the calculation complexity.

Figure 3: Lena decomposition using a 2-level decomposition (2,2) integer wavelet filter.

Table 2 lists the ratios obtained after applying a 2-level IWT using a (2,2) filter:

Images	NN+PH+IWT	NN+IWT	CALIC+IWT
Lena	2.294	2.205	1.880
Barbara	2.132	2.081	1.821
Goldhill	2.295	2.286	1.990
F16	2.566	2.522	2.287
Baboon	1.770	1.752	1.403
Peppers	2.141	2.110	1.851
Splash	2.677	2.646	2.222
House	2.330	2.920	2.024
Average	2.276	2.236	1.935

Table 2: Compression ratios after applying a (2,2) IWT for 2-levels of decomposition.

From Table 2 we can notice that the use of IWT in the NN+PH compressor improved the ratios of Baboon and Peppers in Table 1, whereas it degraded the performance of NN+PH in general. In the contrary it increased the ratio when applied on the CALIC algorithm. The results in Tables 1 and 2 show that the proposed method (NN+PH+IWT) outperformed the CALIC algorithm.

6 Conclusions

This paper presents a new image compression scheme for small satellite on-board imaging. This compressor is based on a neural network prediction and an arithmetic encoder. The latter is optimal within one bit of the entropy. The advantage of this method is that the neural network can learn and predict in one pass. It gives the best prediction when short and long learning rates are combined to balance, between using the whole input history and the most recent inputs in order to adapt to the change of the image statistics. This method gave an average ratio of 2.530, which is good compared to the CALIC algorithm (1.806). A Peano Hilbert scan has been introduced to exploit the redundancies along two dimensions. The experiment led to an average ratio of 2.747, which is 8.5% better than the NN compression ratio using a raster scan.

Since the idea of multiresolution was attractive for our application, an (2,2) IWT has been used as a preprocessing step. A 1-level of decomposition showed that NN+PH performs better on high frequency bands than low frequency ones. The next experiment showed that for a 2-level decomposition we obtain a good tradeoff ratio-calculations complexity. After fixing the value n=2, the experiments showed that the IWT+NN+PH algorithm outperformed the CALIC compressor by 13.1%.

References:

[1] M. V. Mahoney "Fast Text Compression with Neural Networks" Florida Institute of Technology, presented at FLAIRS, May 23, 2000, Orlando, Florida.

[2] J. Reichel, G. Menegaz, M. J. Nadenau ,M. Kunt, " Integer Wavelet Transform for Embedded Lossy to Lossless Image Compression", IEEE Transactions on Image Processing, vol. 10,No.3, March 2001.

[3] A.R. Calderbank, I.Daubechies, W.Sweldens, Boon-Lock Yeo "Lossless Image Compression Using Integer to Integer Wavelet Transforms", Technical Report, Departement of Mathematics, Princeton University, 1996.

[4] V.K Heer and E.E Reinfeldor " A Comparison of Reversible Methods for Data Compression", in Medical Imaging IV, pp 354-365. Proc- SPIE 1233,1990.

[5] A. Said and W.A Pearlman, "An Image Multiresolution Representation for Lossless and Lossy Image Compression", IEEE Transactions on Image Processing, vols.5 pp.1303-1310, September 1996.

[6] Wim Sweldens "Wavelets and the Lifting Scheme: A 5 minute tour", Proceedings of the 1995 ICIAM Meeting Hamburg,1995.

[7] Alexander Stoffel, " Remarks on the Unsubsampled Wavelet Transform and the Lifting Scheme", Signal Processing journal, volume 69, issue 2, pp. 177-182, 1998.

[8] Fang Sheng, Ali Bilgin, Philip J. Sementilliy, Michael W. Marcellin "Lossy And Lossless Image Compression Using Reversible Integer Wavelet ransforms" International Conference on Image Processing October 4-7, 1998 Chicago, Illinois, USA.

[9] Nasir Memon, Xiaolin Wu, Boon-Lock Yeo, "Entropy Coding Techniques for Lossless Image Compression with Reversible Integer Wavelet Transforms." 1998 International Conference on Image Processing October 4-7, 1998, Chicago, Illinois, USA.

[10] Michael D.Adams and Faouzi Kossentini "Low-Complexity Reversible Integer-to-Integer Wavelet Transforms for Image Coding". Data Compression Conference, DCC 1999, 29 - 31 March, 1999, Snowbird, Utah, USA. IEEE Computer Society, 1999.

[11] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet Transforms that Map Integers to Integers," Applied and Computational Harmonic Analysis, vol. 5, no. 3, pp. 332-- 369, July 1998.

[12] Eric E Majani "Algorithm for Lossy/Progressive/Lossless Image Compression", Nasa Tech Brief Vol. 21,No. 12, Item #146,1997.

[13] "Lossless Data Compression. Recommendation for space data systems Standards", CCSDS 121.0-G. Green Book. Issue 1. Washington, D.C.: CCSDS, May 1997.