

# Lossless Image Compression Using Neural Networks & Arithmetic Coding

Yeo Chai Kiat, \*Soon Ing Yann, Lau Chiew Tong

Nanyang Technological University

School of Applied Science

\*School of Electrical and Electronics Engineering

Nanyang Avenue

S 639798

Republic of Singapore

Email: asckyeo, eiysoon, asclau@ntu.edu.sg

## Abstract

This paper describes a lossless compression technique for compressing multispectral images. It makes use of the pattern recognition ability of neural networks to exploit interband redundancy of multispectral satellite images. The backpropagation neural network used is trained to perform predictive coding of SPOT satellite images using both interband and intraband information. The trained neural network predictor is essentially a 3-layer network, comprising an input layer, a hidden layer and an output layer. Each spectrum of the SPOT image has its own trained network predictor, namely, a Green predictor for the Green band, a Red predictor for the Red band and an Infra-Red (IR) predictor for the IR band. Intra-band prediction is applied to the Green band, while inter-band prediction is used for the Red and IR bands. The Red predictor uses Green and Red band images for its inputs while the IR predictor makes use of images from all 3 bands in its prediction. After the predictive coding, the residue errors are coded using multi-symbol non-adaptive arithmetic coding. Results obtained showed that the neural network based predictive coder is superior in performance compared to lossless JPEG standard.

and has a resolution of 20m per pixel. A single SPOT multispectral scene can take up a full 30 Mbytes of storage space. Storage requirements for such scenes are therefore huge and thus are natural candidates for image compression. Furthermore, given its resolution and in order to provide accuracy for its applications, lossy compression is unsuitable. Moreover, the fact that such images are relatively noise free also render them suitable for lossless compression.

Although the paper uses SPOT multispectral images as test data to prove the worthiness of the algorithm, the technique can be easily applied and extended to other multispectral images including LANDSAT. Most work on lossless compression treats each band individually and ignores the interband redundancies as the relationship between bands cannot be linearly represented or easily modelled. However, intuitively, there should be some interband redundancy too, since objects in all three bands share the same boundaries. Furthermore, from the satellite point of view, the classes of objects on earth is quite restricted, as most objects can be classified into vegetation, water, bare soil or sand and man-made objects. Hence using interband information is certainly helpful in predictive coding [1].

## 1. Introduction

Satellite images, such as SPOT and LANDSAT images, are rich sources of digital image data. They contain information that is useful for many applications such as geographical topology mapping, crop and weather forecasting, road-mapping, oil exploration. A SPOT XS image comprises 3 bands, namely, the Green (G) band, the Red (R) band and the Infra-Red (IR) band

## 2. Neural Network Predictive Coding

The predictive coder used for the lossless compression is based on the 3-layer backpropagation neural network. The backpropagation neural network is extensively used for supervised learning [2], [3]. The network is most suited for the predictive coding of images as it can be trained to recognise the relationship amongst neighbouring samples.

As SPOT images has three spectral bands, G, R and IR, three different nonlinear predictive tools are built for each individual bands using a 3-layer backpropagation neural network. The overall structure for the coder and decoder is shown in Figure 1. Note that each pixel in each band is represented by 8 bits. The order in which the bands are coded are G, R and IR. All the inputs to the neural network are scaled down by 255 to restrict the inputs to lie between 0 and 1. However at the output side, the output is scaled up by 255 and rounded to the nearest integer.

## 2.1 Network Structure

Since the Green band is the first band to be coded, the input layer of the G neural network consists only of 3 previous pixels in the same band. The relationship between the predicted pixel value and its past samples are shown in Figure 2. The output layer consists of only 1 node since at any one time only one prediction is made.

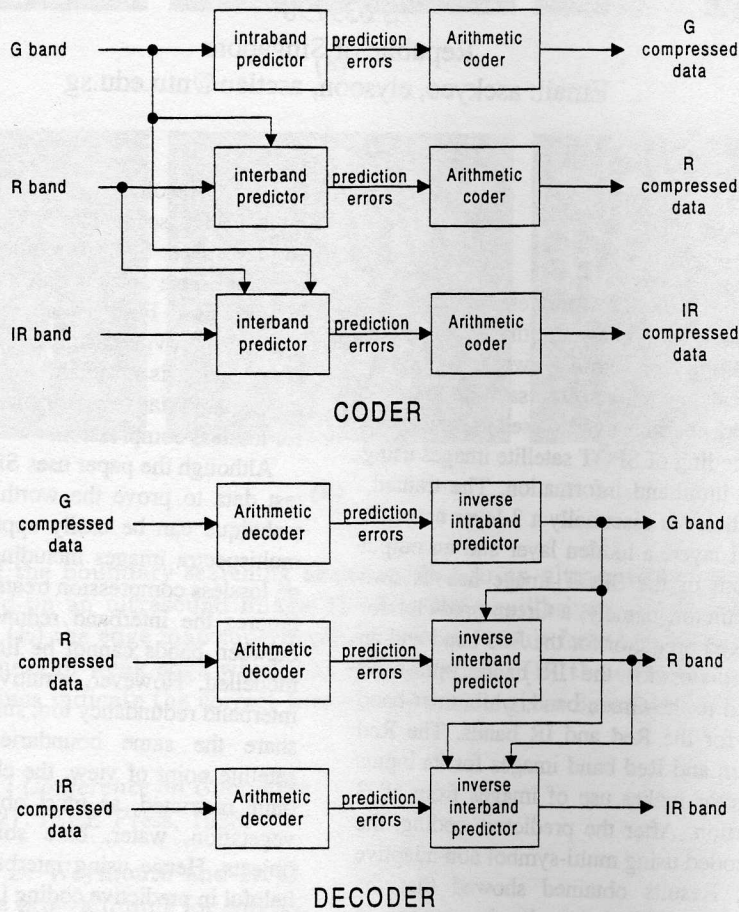


Figure 1 Coder Decoder Structure.

The SPOT images can be broadly classified into 2 categories, namely, rural and urban. As rural and urban images differ in the amount of complexity, a different set of predictor has to be developed for each category of images.

However the number of nodes in the hidden layers cannot be determined through inspection alone. Hence the number of nodes in the hidden layers were decided using a technique called network growing. The process started by using only one hidden node and then the number of hidden nodes is increased incrementally until the network performance saturates. The point at which saturation begins will be used to decide on the number of nodes. Network performance is determined by the error

function. In this paper, the error function used is the log error,  $\log E$  [4] function with  $E$  defined as follows:

$$E = \sum_j [t_j \cdot \ln(x^l_j) + (1 - t_j) \cdot \ln(1 - x^l_j)]$$

where  $t$  is the actual pixel value  
 $x$  is the output from neuron  $j$   
 $l$  is the layer number

Input 1	Input 2
Input 3	Output

Figure 2 Input and Output Samples for Green Band Predictor

### Rural Predictor

For the green band, the optimum number of hidden nodes is 18.

The next band to be coded is the Red band. The prediction for the Red band is based on the 4 samples from the Green band and 3 past samples from the Red band. The arrangement is shown in Figure 3. Similarly network growing is used to determine the size of the hidden layer. The optimum number of nodes is 10.

The last band to be coded is the Infra-red band. The

### Urban Predictor

The urban predictor is identical to the rural predictor in structure except for the differing number of hidden nodes used in the various bands. The number of hidden nodes for the Green, Red and Infra-Red bands are 12, 8 and 21 respectively.

## 2.2 Network Training

The weights of the backpropagation neural network between the input layer and the hidden layer are initialised to zeros, while the weights between the output layer and the hidden layer are initialised to small random values with an absolute value of less than 0.3.

The network was then trained using the backpropagation rule ([5], [6], [7]) with variable momentum factor and learning rate. A larger momentum and learning rate were used initially to speed up convergence and to prevent the network from being trapped in a local minima. As training proceeds, the momentum factor and learning rate were slowly reduced to ensure stability and accuracy.

## 2.3 Predictive Coding

Predictive coding is the standard technique for lossless

Green Band		Red Band	
Input 1	Input 2	Input 5	Input 6
Input 3	Input 4	Input 7	Output

Figure 3 Input and Output Samples for Red Band Predictor

Green Band		Red Band		Infra-Red	
Input 1	Input 2	Input 5	Input 6	Input 9	Input 10
Input 3	Input 4	Input 7	Input 8	Input 11	Output

Figure 4 Input and Output Samples for Infra-Red Band Predictor

prediction for the Infra-red band is based on 4 samples from the Green band, 4 samples from the Red band and 3 past samples from the Infrared band (See Figure 4). The optimum number of hidden nodes is 12.

compression [8]. However methods of performing the prediction may differ. This section presents a brief overview of the method.

Let the image sample to be coded be  $P$  and the predicted image sample be  $P'$ .

Therefore,  $P' = NN(\text{past samples})$

where NN is the neural network function and P is the output of the neural network.

The prediction error, E, is given by

$$E = P' - P$$

The above predictive coding does not apply to the top row and left column of the image. For the top row and left column, 1D DPCM is used instead. The prediction error is simply current pixel minus previous pixel in the respective directions.

### 3. JPEG Lossless Extension

The JPEG standards for lossless extension [9] are used as a benchmark to compare the performance of the neural network predictor. A total of seven predictors are given as shown in Figure 5. From Figure 5, it can be seen that pixel x is the predicted result while its neighbouring pixels, a, b and c are used in the respective prediction models.

c	b
a	x

- Technique 1:  $x = a$
- Technique 2:  $x = b$
- Technique 3:  $x = c$
- Technique 4:  $x = a + b - c$
- Technique 5:  $x = a + ((b - c) / 2)$
- Technique 6:  $x = b + ((a - c) / 2)$
- Technique 7:  $x = (a + b) / 2$

Figure 5 JPEG Prediction Models

### 4. Arithmetic Coding

All the prediction errors are coded using P-based, multi-symbol non-adaptive arithmetic coding. Arithmetic coding is an entropy encoding technique that encodes data by creating a code that represents a fractional value on the number line between zero and one [10]. It works by recursively partitioning an interval of the number line (starting from an interval from zero to one, inclusive of zero) and retaining one of the interval as the new interval depending on the event or symbol to be encoded. For P-based method, the coding is based on the probability of occurrence of the symbols. Hence a frequency table has to be built to count the number of occurrence of each symbol. Each spectra band is given its own frequency table for higher coding efficiency.

### 5. Results

Arithmetic coding is used to compress the prediction errors generated by both the neural network predictor and the JPEG prediction models. Four test images were used, two belonging to the urban category and the remaining two belonging to the rural category. The best JPEG prediction model was selected and the resulting compression ratios achieved by the neural network predictor and the JPEG predictor were tabulated in Table 1.

The overall results in Table 1 show that the neural network predictor exhibits more than 10% improvement over the JPEG predictor. The performance of the neural network (NN) coder over the JPEG predictor for the Green band is less obvious as all the coders are using only intraband coding. However the improvement in performance using the NN coder for the Red band is very significant. This is due to the high interband correlation

Bands	Urban 1		Urban 2		Rural 1		Rural 2	
	JPEG	NN	JPEG	NN	JPEG	NN	JPEG	NN
Green	1.85	1.86	2.01	2.03	3.00	3.04	3.72	3.74
Red	1.86	2.56	1.91	2.66	2.89	3.61	3.70	5.13
Infra-Red	1.71	1.74	1.72	1.76	2.11	2.13	2.40	2.45
Overall	1.80	2.05	1.87	2.09	2.61	2.93	3.14	3.45

Table 1 Compression Ratios

that exists between the Green and Red bands. The results thus testify that the pattern recognition capability of the nonlinear neural network predictor can optimally exploit these interband redundancies. The Infrared band on the other hand has a very much higher entropy than the other two bands and hence can only be compressed to a smaller extent compared to the other bands.

The neural network based predictor has proven to perform better than the JPEG lossless predictor for the SPOT satellite images. The only overhead incurred is that prior training needs to be performed. However, once the neural network is trained, the predictor is as easy to use as the JPEG lossless predictor.

## 6. Conclusion

This paper illustrates the ability of the backpropagation neural network to recognise the nonlinear correlation between the different bands of a multispectral images such as the SPOT XS images. Using such a predictor for lossless compression together with arithmetic coding yields very good results compared to lossless JPEG prediction techniques with arithmetic coding. This is due to the fact that such a predictor can optimally exploit the interband redundancies to maximise the compression ratios. Tests conducted showed that there is a more than 10% improvement in compression ratios when using the NN predictor. The NN predictor needs only to be trained once to obtain all the necessary weights required for carrying out the prediction. Once trained, the NN predictor can efficiently perform predictive coding of the satellite images just like its JPEG prediction models.

## 7. References

- [1] C. K. Yeo, I. Y. Soon, "Neural Networks based Predictive Coder," *Proceedings of International Conference on Signal Processing, Circuits & Systems*, IEEE SICSPCS95, Jul 95, Singapore, pp. 299-303.
- [2] Don R. Hush and Bill G. Horne, "Progress in Supervised Neural Networks," *IEEE Signal Processing Magazine*, Jan 1993.
- [3] R. P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE Acoustics, Speech and Signal Processing Magazine*, 4(2):4-22, Apr 1987.
- [4] Patterson Dan W., *Artificial Neural Networks: Theory and Applications*, Prentice Hall, 1996.
- [5] J. Makhoul, A. El-Jaroud and R. Schwartz, "Formation of Disconnected Decision Regions with a Single Hidden Layer," *Proceedings of the International Joint Conference on Neural Networks*, vol 1, pp. 455-460, 1989.
- [6] Russell C. Eberhart and Roy W. Dobbins, *Neural Network PC Tools A Practical Guide*, Academic Press Inc, 1990.
- [7] J. L. McClelland and D. E. Rumelhart, *Explorations in Parallel Distributed Processing, A Handbook of Models, Programs and Exercises*, MIT Press, Cambridge, MA, 1986.
- [8] I. Y. Soon, "Lossless Compression of SPOT XS Images," *Proceedings of International Conference in Image Processing, ICIP'92*, Sep 92, Singapore, pp. 11-15.
- [9] JPEG Technical Specification, Revision 8, ISO/IEC JTC1/SC2/WG10.
- [10] Pennebaker, W. B., J. L. Mitchell, G. G. Langdan Jr. And R. B. Arps, "An Overview of the Q-coder Adaptive Binary Coder", *IBM J. Res. Develop.*, vol. 32, no. 6, 1988.

## Acknowledgements

The authors would like to thank Ms Wendy Huang Yingyi for her excellent efforts in performing the training of the neural networks as well as the collection and collation of training results.