THE EFFECT OF DATA REDUCTION ON IMAGE INTERPRETATION

by

N A McEVOY and M H B THOMAS

LOGICA Ltd, 64 Newman Street,
London
England

ABSTRACT

With the improvement in the spatial and spectral resolution of sensors, the problems of handling very large quantities of data are becoming increasingly acute. A partial solution to this problem is to compress the data. However, the process of compression and reconstruction generates errors in the pixel intensities and this will have an effect on user interpretation algorithms. The particular compression algorithm we have studied is the Rice Machine 2 with particular reference to a multispectral scanner obtaining imagery over the sea. The effect of data compression on an atmospheric correction/chlorophyll concentration algorithm is considered. We conclude that data compression offers substantial improvement in coverage at the price of small but acceptable errors in pixel intensities.

1 INTRODUCTION

When a satellite is out of range of a ground station any data acquired by on-board instruments must be recorded and then replayed when the ground station is in view. Clearly, the quantity of data that can be acquired in this way is limited by the capacity of the tape recorder. In order to increase the coverage are, it is necessary to use data compression techniques so that imagery from a given area can be stored in a fewer number of bits.

We have given particular consideration to an Ocean Colour Monitor (OCM) instrument which was originally scheduled to fly on ERS-1 but is now planned for a later mission.

2 THE RICE MACHINE 2

The particular data compression technique we have studied is known as Rice Machine 2 (RM2). This has been described elsewhere (Rice 1985 ; McEvoy 1981) but a brief description will be given here.

Cascaded Hadamard transformations are applied to the image data, which is processed in blocks of 64 square. Four levels of the transform are performed and the transformed data is entropy coded in blocks of 16. The Hadamard transformations produce unimodal data which is symmetrically distributed around zero. The basic principle is that the most commonly occurring values are given the shortest codes. This type of compression which can be achieved depends on the randomness or "activity" of the data. Using CZCS data from the North Sea we have obtained compression ratios of typically 2.8.

Our objective has been a compression ratio of 4 and to obtain this it is necessary to reduce the bit quantisation of the transformed data. This naturally introduces errors in the data when it is reconstructed and we have found that the r.m.s error is 0.7, compared to a maximum value of 255 for 8-bit data.

3 COMPARISON WITH OTHER SYSTEM ERRORS

This reconstruction error needs to be compared with other system errors. In particular :

1) Detector noise. The detector themselves are noisy with a finite signal-to-noise ratio. For the CZCS, this is of the order of 125 to 1.

2) Calibration errors. The data acquired by the instrument consists of digital numbers which have to be converted to radiances through a linear relationship. For the OCM, the absolute calibration is expected to be accurate to 2% and the interband calibration to 1%.

3) Atmospheric correction. Before the imagery can be processed to calculate chlorophyll concentrations, it must be corrected for atmospheric effects. At present, this is not a calculation which can be performed to any high degree of accuracy and there is considerable uncertainty concerning some of the critical parameters eg the aerosol optical thickness. It is unlikely that the atmospheric contriburion to the total radiance is known to better than 2%.

By comparison, the error due to compression and reconstruction is comapatively small and can be regarded as an acceptable price to be paid for the improvement in coverage.

4 DETERMINATION

We have used a simple power law expression to determine chlorophyll concentrations. This is an algorithm which is widely used amongst experimenters. Explicitly, the chlorophyll concentration, C, is given by :

$$C = a\left(\frac{L_1}{L_3}\right)^b$$

where L_1 , L_3 are atmospherically corrected radiances in CZCS bands 1 and 3

a , b = adjustable constants.

The constants a and b are normally determined by least-squares fitting to sea-truth measurements and their values can vary considerably.

Data compression causes errors in L_1 and L_3 , which then cause errors in the value of C. Theory shows that the percentage error in C is directly proportional to the percentage errors in L_1 and L_3 and to the value of b but independent of a. In our experiments, we have used values of b ranging from -2 up to -8.286, which represents the range used by experimenters. Errors in the chlorophyll concentration were found to be almost directly proportional to b as predicted by theory. Using b=-8.286 (the worst case) errors were found to be 22%rms.

This figure should be compared with the errors in chlorophyll concentration due to the other system errors (see section 3). In particular a 2% error in the atmospheric correction will lead to an error in C of approximately 60%, for b=-8.286.

We should also consider the errors which are found in surface measurements. Because of the variation in concentration from point to point, a surface measurement may be quite unrepresentative of the average concentration within a CZCS pixel (800m square). Errors from this source can be up to 50% (Charlton, 1979).

5 COVERAGE IMPROVEMENT

We have considered a satellite carrying an instrument package similar to ERS-1, namely :

OCM	- 2.8 Mbit/s
SAR (wave mode)	- 600 kbits/s
Altimeter	- 10 kbits/s
Wind scattermeter	- 1 kbit/s

These are the instruments whose data can be fed to the tape recorder. A compression factor of 4 reduces the data rate from the OCM to 700 kbits/s.

We have examined various combinations of instruments sending data to the tape recorder. For example, with all these instruments switched on, a 5 Gbit tape recorder would be full after 24.2 minutes, but data compression would increase this to 63.5 minutes, so that 2.6 times as much data can be stored and played back. If only the OCM by itself were switched on, coverage is obviously improved by a factor of 4.

6 CONCLUSIONS

Data compression presents a way of greatly increasing the coverage area for the case where data is being recorded and played back to a ground station. The penalty wich ispaid is a degree of error in the digital imagery. However, these errors are considerably smaller than the system errors and the uncertainty in user algorithms.

7 ACKNOWLEDGEMENTS

This work was supported by the European Space Agency under contract n° 4835/81/NL/HP (SC).

Thanks are also due to M. J-C Degavre of ESTEC and to DR S M Singh of Dundee Unviersity for their help and encouragement.

REFERENCES

1 Charlton J A, 1979, Proc Remote Sensing Society Conference, Dundee.

2 McEvoy N A, 1981 Proc Remote Sensing Society Conference, London.

3 Rice R F, 1975, SPIE Seminar Proceedings on Efficient Transmission of Pictorial Information.