N92-12438 70

p. 3

TECHNIQUES FOR CONTAINING ERROR PROPAGATION IN COMPRESSION/DECOMPRESSION SCHEMES

Ben Kobler
Goddard Space Flight Center
Greenbelt, MD 20771

1. The Application

The typical raw Bit Error Rate (BER) for space communications is one bit in 10×10^5 bits. Through error correction of header information, this can be reduced to one bit in 10×10^8 bits; through error correction of the whole data set this can be further reduced to one bit in 10×10^{12} bits. Similarly, the typical raw BER for archive media is one bit in 10×10^6 bits; through error correction this can also be reduced to one bit in 10×10^{12} bits.

The total EOS data volume, however, is at minimum 10×10^{16} bits, to be accumulated over a 15 year period. If BER were to stay at one bit in 10×10^{12} bits, this would result in several uncorrectable errors per day. To avoid this, we must push toward better error correction. However, since we will also be doing data compression to minimize transmission and storage requirements, we have to understand the relationships between error correction and data compression.

2. Key Issue

Data compression has the potential for increasing the risk of data loss. Although data compression reduces the number of bits required for transmission and storage -- and hence the number of bit errors that can be expected -- data compression can also cause bit error propagation, resulting in catastrophic failures. For example, entire images could be rendered useless due to a single bit error. Techniques to detect these errors in compressed data and to minimize the resulting error propagation often involve trade-offs against compression performance.

3. Approaches

There are a number of approaches possible for containing error propagation due to data compression.

1) Data re-transmission - Requests for data re-transmission are only useful, however, when errors are detected, and only when errors are detected early. In space communication retransmission is often impossible; in archive systems re-transmission is often not helpful since the media may already be corrupted.

2) Data interpolation - Data interpolation is also only possible when errors can be detected. In addition, since we often have entire images destroyed, this may require data interpolation between entire time sequenced images -- a difficult technical task, and one that the science community would find difficult to accept.

3) Error containment - Error containment is already done to varying degrees in some data compression algorithms. Vector quantization, for example, sends compressed data in fixed sized blocks, thus limiting error propagation. Some Huffman codes allow quick error detection and re-synchronization, as does the DCT (Discrete Cosine Transform) JPEG (Joint Photographic Experts Group) algorithm which has an appended delimiter pattern and the Rice algorithm which has a fixed line format. Arithmetic codes, however, although efficient in compression performance, do not provide error containment. While this

can be improved via piecewise arithmetic coding, it is done at the expense of reduced compression performance.

4) Error correction - Error correction could be improved so the BER is perhaps only one bit in 10×10^{14} or 10×10^{15} bits. Errors will then occur so infrequently that they would not to be a problem. Improving the BER, however, adds significant additional data bits, thus increasing bandwidth and volume requirements, as well as requiring additional processing power. A related technique, however, to code different information channels with different degrees of error correction depending on their importance, has potential for increasing the effective BER without unduly increasing bandwidth, volume, or processing power requirements. Another technique to look for destruction of specific apriori known information about the data string due to error propagation in data compression also holds promise to allow detection and correction of errors missed through traditional error correction algorithms.

4. Recommendation

The most fruitful techniques will be ones where error containment and error correction are integrated with data compression to provide optimal performance for both. The error containment characteristics of existing compression schemes should be analyzed for their behavior under different data and error conditions. The error tolerance requirements of different data sets need to be understood, so guidelines can then be developed for matching error requirements to suitable compression algorithms. Work should be done to develop new compression algorithms, or modify existing compression algorithms, to improve error containment behavior. Work should also be done to look for ways in which data compression could aid error detection and subsequent error correction.

Participants

The participants in this discussion group were Mayun Chang, Kar-Ming Cheung, P. C. Hariharan, Ben Kobler, Joan S. Langdon, Edward Seiler, and Gregory S. Yovanof. See the appendix for addresses.