AIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202

Centimeter

Inches

MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

1 of 1

LOSSLESS COMPRESSION OF WEIGHT VECTORS FROM AN ADAPTIVE FILTER

M. V. Bredemann, G. R. Elliott, and S. D. Stearns

Sandia National Laboratories
Albuquerque, NM 87185-0972

Abstract--Techniques for lossless waveform compression can be applied to the transmission of weight vectors from an orbiting satellite. The vectors, which are a part of a hybrid analog/digital adaptive filter, are a representation of the radio frequency background seen by the satellite. An approach is used which treats each adaptive weight as a time-varying waveform.

I. INTRODUCTION

Sandia National Laboratories is actively involved in the design of remote sensing instrumentation for use aboard earth-orbiting satellites. Usually the amount of data transmitted to the ground is limited by the data rates on the down-link channels. Data compression techniques are currently being investigated in order to increase the amount of data available for analysis. In general, system requirements dictate that the compression be lossless.

A. The FORTÉ Satellite

The Fast Onboard Recording of Transient Events (FORTÉ) small satellite, a joint project of the Los Alamos National Laboratory and Sandia National Laboratories, will be placed in low earth orbit to monitor transient radio frequency (RF) events such as lightning. It is a part of a series of Department of Energy satellites devoted to increasing the basic understanding of how signals are transmitted to remote orbiting sensors. These small experimental satellites play an important role in a national effort of applied research in the space sciences.

A part of the FORTÉ instrumentation system is a hybrid analog/digital adaptive filter. Its main function is to reduce the effects of narrow-band RF emitters to a point where the background noise spectrum is spectrally flat or white. The filter is called the Adaptive PreWhitener, APW.

FORTÉ orbits the earth with a period of slightly over 100 minutes. During each orbit the APW continually adapts, updating the weight set several times each second. The adaptive filter weight vector is a representation of the noise environment. The Fourier transform of the weight vector is an approximation of the spectrum of the noise. Therefore, an entire 100+ minute set of weight vectors contains an approxi-

mation of the time-varying noise spectrum of a complete pass around the earth. These spectral data have many potential uses.

B. The Adaptive Prewhitener

The APW has been developed at Sandia to provide selective attenuation to interfering narrow band RF signals and pass transient signals of interest with minimal distortion. This filter adaptively converges to variations in the statistical properties of the interfering sources in real time.

Fig 1. Simplified APW block diagram

The APW employs a programmable transversal filter structure operating on signals with a bandwidth of 90 MHz. The delay structure and multipliers are contained on a single Gallium-Arsenide chip. Each multiplier in the transversal filter has an input resolution of 6 bits plus sign. Weight updating is

accomplished by a Digital Signal Processor (DSP) computer system with associated memory. A simplified block diagram is shown in Fig. 1 The adaptive algorithm is based on Widrow's adaptive predictor configuration.[2]

The APW substantially reduces the dynamic range requirements for the FORTÉ digitizing system by removing the higher level narrow band signals prior to digitization. With this improved sensitivity, detection and recording of weak transients are possible. The processing algorithm results in minimum phase distortion; and, generally, the transient signals recorded are affected little by the APW. However, if desired, the APW filter effects can be removed by deconvolving the weight vector with the data. (Within the limits imposed by the numerical problems associated with deconvolution.)

Since the predictor weights approximate the steady state auto-correlation of the input signal, the frequency content of the interfering signals can also be extracted from the current weight values. A global mapping of the narrow-band interference is possible with a sufficiently high weight vector storage rate. In order to minimize memory and data transfer requirements, compression of the ensemble of weight values is desirable. To ensure full recovery of all valuable scientific information, the compression is required to be lossless.

II. DATA COMPRESSION

A. Introduction

Compression of waveform data is significant in many engineering and research areas. It can be used to reduce data storage requirements and/or transmission bandwidth. In the present case, we are using waveform compression to maximize the amount of waveform data sent to earth from the FORTÉ instrumentation system through a channel with a given bandwidth.

Traditional waveform compression methods have produced good results for applications to speech and image data, but these methods are generally lossy. They do not usually allow exact reconstruction of the original data. Lossless compression is, however, a requirement with many kinds of instrumentation and telemetry waveform data, where the user needs to retrieve the digitized waveform data in its exact original form, without any numerical noise added during decompression. In the case of the FORTÉ system, exact reconstruction with error detection is a requirement.

A new lossless waveform compression method was recently proposed [1]. The method is described generally as a two-stage coding scheme. As shown in Fig. 2, the first stage is a special version of the classical linear predictor. Linear prediction is very effective in compressing continuous waveforms. It greatly reduces the amplitude range of narrow-band

data and, in the process, decorrelates the data. After prediction, the residue sequence is approximately a white Gaussian sequence. Lossless prediction can be implemented in a practical and efficient manner, and is suitable to real-time applications. The second stage currently employs arithmetic coding to encode the residues from the first stage.

Fig. 2. Two-stage compression

B. Lossless Compression in Two Stages

In the two-stage process illustrated in Fig. 2 input data to be encoded is partitioned into fixed-length sequences, each having K integer samples. In the first stage, one-step linear prediction is used to convert the input integer sequence, $ix(0:K-1)$, to a prediction error or residue integer sequence, $ir(M:K-1)$. In the second stage, arithmetic coding is used to compress the residue sequence into a final sequence of bytes, $ic(0:N_c-1)$.

In general, for continuous waveforms, the residue sequence has the following important properties: It is essentially decorrelated, its amplitude distribution is approximately Gaussian with zero mean, and its variance is small when the sequence $ix(0:K-1)$ consists of narrow-band or other predictable components. The predictor in Fig. 2 differs from the usual linear predictor in one important way. Since we require exact recovery of the input sequence $ix(0:K-1)$, the predictor coefficients, $ib(0:K-1)$, are integers, and also the prediction and recovery algorithms are designed to reverse each other exactly, with no roundoff errors in the process. Each input sequence, $ix(0:K-1)$, is assumed to have exactly K samples, each in the range $\pm (2^{30}-1)$.

Table 1: Compressed frame contents.

Data	Contents
Keyword	$M, ix(0:m-1), ib(0:M-1), N_r, ir_L$, and other parameters
Interval sequence	$ir_f(M:K-1)$
Offset sequence	$ir_o(M:K-1)$

The format of a compressed data sequence is shown in

Table 1. The residue sequence is included in two separate parts, $ir_I(M:K-1)$ and $ir_O(M:K-1)$. Both sequences begin at M rather than 0 because the first M data samples, $ix(0:M-1)$, are included along with the integer predictor coefficients, $ib(0:M-1)$, and other parameters in a "keyword" at the beginning of the compressed data frame, where each data sample is coded as a signed binary integer. These are used as initial values for inverse linear prediction in the recovery procedure. Each residue, $ir(k)$, is recovered from

$$ir(k) = ir_L + 2^{Nr} ir_I(k) + i_O(k) \quad (1)$$

in which ir_L is the minimum residue value, $ir_I(k)$ is the number of an interval of length, 2^{Nr} and $i_O(k)$ is an offset within the interval composed of N_r bits. Arithmetic coding is used for $ir_I(M:K-1)$, and ordinary binary coding is used for $ir_O(M:K-1)$, on the basis that the former has an approximately Gaussian amplitude distribution, whereas the latter has an approximately uniform amplitude distribution.

Decompression at the receiver is carried out by a decoder which also has two stages. First, the keyword is decoded to determine the parameters for the two stages. Then, the first stage decodes the residue interval sequence $ir_I(M:K-1)$ using arithmetic decoding, and combines this sequence with the offset sequence, $ir_O(M:K-1)$, to reconstruct the original residue sequence, $ir(M:K-1)$ in accordance with (1). The second stage is the inverse linear predictor, that is, the all-pole filter which performs an exact recovery of the original sequence $ix(0:K-1)$ using the initial values, $ix(0:M-1)$ from the keyword, plus the recovered residue sequence.

III. COMPRESSING THE FORTÉ APW WEIGHT VECTORS

A. Data Organization

In an adaptive predictor such as the APW each weight in the weight vector $w_0, \dots, w_{L-1}$ changes slowly with time. The interfering signals typically vary gradually as the satellite orbits the earth. Considering that the lossless compression algorithm involves prediction, we felt that the best results would be obtained by compressing each weight as a time-varying function. We establish a generalized representation of each weight at a given time, w_{lp} , where l is the index related to the weight number, and p is the time index associated with the weight update interval. This can be represented as a matrix, $W(0:L-1, 0:P-1)$. Weights are presented to the compression algorithm for a fixed l , with p varying over the range of 0 to $P-1$. For the APW there are 126 weights: $L=126$. Data are typically stored for 1024 updates: $P=1024$.

B. Application of the Programs

A standard software library of compression programs was developed as result of the work described in [1]. The programs are intended primarily for developmental activities, and provide a good basis for evaluating the usefulness of the approach for this application. Actual APW hardware was used, set up in a development laboratory at Sandia. A complex set of signal generators was set up to simulate a bad-case scenario for noise environment, and the weights recorded for 1024 updates. Off-line, at one of the development workstations, the compression and decompression algorithms were run after reordering the data as described above.

The compression routines ran without difficulty, reducing each string of 1024 6 bit integers into 100-300 bytes.

C. Results

Both compression and decompression algorithms ran reliably, resulting in the exact same input and output data, verifying the lossless nature of the compression/decompression technique. In the test runs of 1024 updates the ratio of the number of bytes in the input data set (which, of course is equal to the number in the output data) to the number of bytes in the compressed data set was 5.91. In the final design of the satellite communication system control words will need to be inserted which will include at least some error checking abilities, decreasing the realizable compression ratio slightly.

IV. COMMENTS

A. Coding Considerations for the Onboard Hardware

The APW contains a very capable DSP microcomputer system and several megabytes of memory. The compression can be accomplished in one of two ways. The weights can be stored in a buffer for several updates and the compression software run in batch mode, much as was done in the laboratory simulation. This has the advantage that the existing software will need little modification other than adapting to the DSP chip. The disadvantages are that more memory is required and more of the computer processing time will be expended.

A second approach would be to compress each weight as it is available, running in a near-real time mode. This requires less memory for weight storage; however 126 copies of the compression software will be required or the routines will have to be made reentrant. Either way will require more program memory, using up much of the savings in buffer storage. The batch real-time trade-off has not been analyzed.

V. CONCLUSIONS

Lossless compression of weights from an adaptive predictor such as the APW has the potential for greatly reducing the data flow over the transmission down-link. The approach presented in [1] functions well on actual data sets from the APW hardware, resulting in compression ratios of nearly six to one and exact reproduction after decompression. Some work remains to be done in adapting the algorithms for the satellite onboard restraints.

REFERENCES

- [1] S. D. Stearns, Li-Zhe Tan, N. Magotra, "Lossless Compression of Waveform Data for Efficient Storage and Transmission," *IEEE Trans. Geoscience and Remote Sensing*, vol. 31 no. 3, pp 645-654, May 1993.
- [2] B. Widrow, et al, "Adaptive Noise Cancelling: Principles and Applications," *Proc. of the IEEE*, vol. 63, no 12, pp 1692-1716, December 1975.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DATE

FILMED

9 / 19 / 94

END