Reprinted from *SearchLites*, **2000**, 6(2), 5-6.

Using the GoldWave Shareware Program for Collection, Storage, and Analysis of SETI Data

Robert A. Lodder
A123 Advanced Science and Technology Commercialization Center, University of Kentucky, Lexington, KY 40506-0286

INTRODUCTION

Often the biggest problem for a SETI station is storage and analysis of data. GoldWave is a shareware program that can help reduce this problem. GoldWave is a digital audio editor for Windows 95/98/NT that uses the standard audio compression manager. The program is usually used by those who need to work with audio for CD editing, Java applications, Web pages, games, radio and TV, etc. GoldWave features real-time amplitude, spectrum, and spectrogram oscilloscopes, large file editing (up to 1GB in size), configurable RAM (fast) or hard disk (large) editing, numerous effects, and support for many sound formats. GoldWave can open and play .au files found in Java applications and on Web pages. It can convert to and from many sound formats including .wav, .voc, .mp3, .au, and binary and text data. The current shareware versions are GoldWave v4.02 for Windows 95/98/NT, v3.03 for Windows 3.1, and v2.14 for Windows 3.1 for old systems (a 275k zip file).

EXPERIMENTAL

The WAV file compression routines used in GoldWave are excellent, requiring only 3.51 Mb of disk space for one hour of SETI data using WAV MSN Audio 8 kHz mono 8200 baud compression (more compression than MP3!). The MSN audio mono 8 kHz 8200 baud compressor, while not as efficient at compression as the Lernout and Hauspie CELP 4.8Kbit/sec 8000 Hz 16-bit compressor (which uses 2.06 Mb of disk space per hour of SETI data), nevertheless preserves ability to detect the smallest signal in our test set (a signal level at 1% of the noise) (see Figure 1). The 30-minute test file was constructed by digitizing AM-mode noise from a microwave receiver and adjusting the gain so that the noise level peaked at full-scale on the A/D. A test signal was digitally mixed at seven different levels (100, 50, 25, 12, 6, 3, and 1% of full scale in the time domain) into the noise. Each signal at 1275.1 Hz was mixed into the file for one minute, starting after five minutes of noise and continuing for seven additional minutes in the noise, after which the file was filled with noise alone.

The Lernout and Hauspie CELP compressor obliterated the 1% signal (see Figure 2) preserved by the MSN

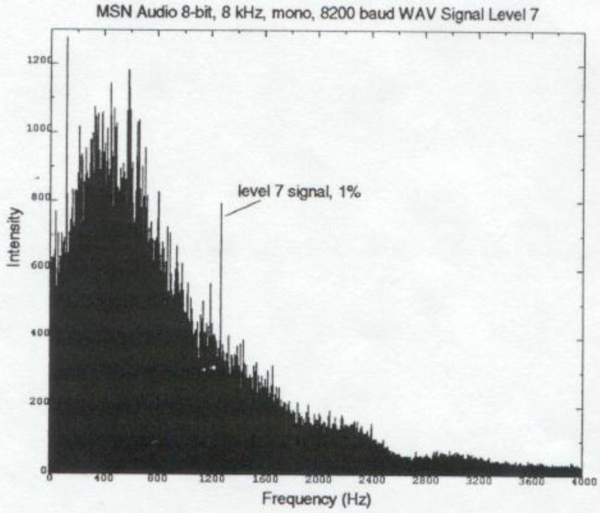

Figure 1. The smallest signal in the test set, present at 1% of the noise level in the time domain, was easily detectable in the compressed data.

compression. In addition, the Lernout and Hauspie CELP compressor introduces other significant distortions into the 4 KHz wide signal, such as forcing the noise intensity to zero at the low frequency end of the spectrum.

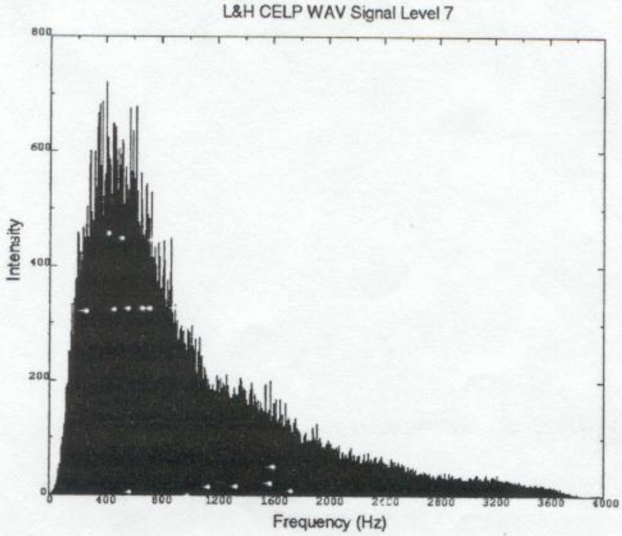

Figure 2. Further compression using the L&H CELP codec obliterates the smallest signal in the test set.

RESULTS AND DISCUSSION

It appears that at least some pattern recognition algorithms can be applied directly to the MSN WAV-compressed data to detect unusual signals. The two most significant advantages to such a procedure for analysis of SETI data are reduced disk space and memory requirements, and increased speed of analysis. One hour of SETI data can require 60 to 70 Mb of storage space when using 4 kHz bandwidth, 8 bit data. The MSN-compressed format requires only 3.5 Mb for the same hour of data. Because the data array can be made a factor of 20 smaller, the analysis can be made a factor of 20 faster (at least when the same pattern recognition algorithms can be used on compressed and uncompressed data). Fortunately, at least one nonparametric subcluster detection algorithm (1) appears to be usable on both compressed and uncompressed data.

The subcluster detection algorithm referenced above was applied to the 30-minute test file containing seven different signal levels, and produced the graph in Figure 3. The graph assumes the position of each byte in the file is proportional to the elapsed time in the recorded data. The algorithm was apparently able to detect the unusual patterns of at least four signals without any knowledge of the MSN WAV file format or header structures. It should be noted that analysis of the uncompressed data reveals all seven test signals. The result of pattern recognition testing on the compressed data suggests that a better understanding of the MSN WAV compression algorithm will increase the effectiveness of automated detection of unusual signals on such compressed data. Nevertheless, subcluster detection can already be used as a tool to flag data blocks that should be decompressed for more thorough analysis. Further testing is now underway.

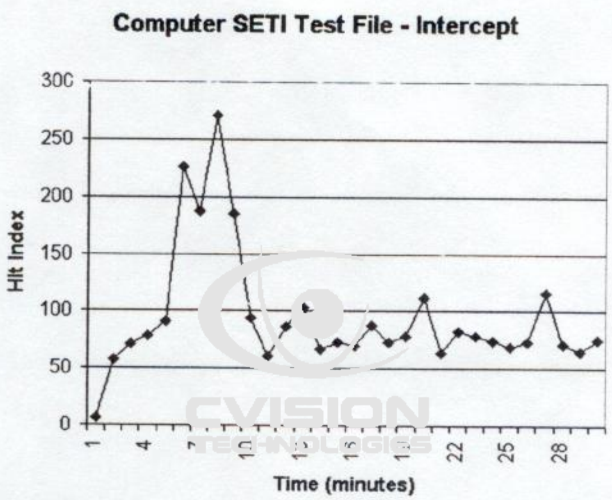

Figure 3. The first 4 of 7 test signals are detectable in the MSN-compressed data.

CONCLUSION

Analysis and storage of radiotelescope data can be a major problem for Project Argus SETI stations. GoldWave is a shareware program with compression algorithms that can help with data collection, analysis and storage. At the University of Kentucky, many students use GoldWave to screen radiotelescope data on the web server at http :// hendrix. pharm. uky. edu / seti / volunteers. html and seldom experience any problems despite use on all kinds of computers. New users should set temporary storage to RAM if possible in the Options / File menu as this speeds the slowest functions. It is also better to use the GoldWave clipboard instead of the Windows clipboard for large file manipulations. GoldWave displays both time domain and frequency domain signals simultaneously in different windows. Users can toggle through different display formats by clicking on the displays. GoldWave offers useful functionality to small SETI stations at relatively low cost.

REFERENCES

1. R. A. Lodder and G. M. Hieftje, "Detection of Subpopulations in Near-Infrared Reflectance Analysis", *Appl. Spectrosc.*, **1988**, *42*, 1500-1512