

EQUIPMENT REVIEW

SIGNAL™/RTSD™ SOUND ANALYSIS SYSTEM

Computer analysis of animal sounds is becoming increasingly commonplace as computers become faster and cheaper (Stoddard 1990, Williams and Slater 1991, McGregor and Ranft 1994, Wilkinson 1994). The SIGNAL (v. 3. 1) and RTSD (Real Time Spectrogram Direct to Disk v. 1.0) software programs have been in continuous development for over a decade and provide powerful tools for sound analysis and synthesis. The developer of SIGNAL, Kim Beeman, has worked in close contact with bioacousticians throughout SIGNAL's evolution, and the program reflects their feedback and "wish lists."

Technical Information

Both SIGNAL and RTSD run in DOS. Macintosh or Windows versions are not available. Minimum requirements are an IBM-AT compatible computer with math coprocessor, 640 kB RAM, EGA graphics, and DOS 3.3 or higher. A fast 586 CPU with 8-16 Mb RAM will facilitate computationally intensive analyses in SIGNAL. In desktop versions SIGNAL and RTSD use instrumentation quality analog I/O boards from Data Translation. Several different versions with maximal sampling speeds ranging from 50-250 kHz (12 bit) or 100 kHz (16 bit) are available. The sampling rate can be selected in software by the user. Desktop RTSD requires a separate DSP board in addition to the I/O board. For laptop users, a two-channel 16 bit PCMCIA card with a maximal sampling rate up to 50 kHz per channel serves as the I/O device for SIGNAL. RTSD is not currently available for laptops. Desktop users may also require a separate anti-aliasing filter (Stoddard 1990) for each I/O channel. Two anti-aliasing filters are built into the PCMCIA card for laptops. Included with the SIGNAL or RTSD system is an interface panel with connectors for signal input and output.

System Overview

RTSD provides a scrolling, real-time spectrogram and waveform of the signal source. RTSD can write the input stream directly to the computer's hard disk, up to the unused capacity of the hard drive, or store the input in memory. For example, one can acquire about 70 minutes of sound sampled at 25 kHz with 200 Mb free on the hard drive. In extended memory mode, 5 minutes of sound (at 25 kHz) can be retained in 15 Mb of RAM. Files this long can be processed by SIGNAL (see below). RTSD is especially useful as an acquisition "front end," allowing the user to scan a recording while listening to it. The program

can be stopped at any point, and the signal entered up to that point can be examined by scrolling back and forth. The user has control over the scroll rate, input sensitivity and gain, dynamic range of the spectrogram, sampling rate and frequency resolution (FFT length). Cursors can be used to demarcate stop and end points, and the portion of the signal between the cursors can be written to a file, sent to a printer, or played back through a loudspeaker. Cursors can also be used to make time and frequency measurements off the display, and the values can be stored in a separate "log file" for later statistical analysis. Some users may find RTSD entirely satisfactory for all their sound analysis needs as it can produce high quality spectrograms and makes signal measurement possible. SIGNAL, however, offers extremely powerful capabilities for the analysis, measurement and recording of data, synthesis, and playback, and I turn to it after I have acquired my sound files in RTSD.

SIGNAL stores data in three different types of buffers: time (T), frequency (F) and "spectrogram" (FT) buffers (Figure 1). In addition, an acquisition buffer is used for acquisition (if the user does not use RTSD for this operation) and/or playback of time buffers. Up to four channels can be acquired (eight with an optional interface panel), and two channels played back simultaneously. The user can configure the number and sizes of each buffer type within the limits imposed by the amount of random access memory in the computer (SIGNAL is available in 8 or 16 Mb versions). As in RTSD, the user can choose the sampling frequency desired. The acquisition buffer is a continuous ring buffer that "remembers" the last x points sampled. Hitting <Enter> terminates acquisition and enters the sampled data into the T buffer designated by the user at the beginning of acquisition. A very useful feature is that the program will report the headroom in the sampled data, and provides a warning if the I/O converter range (± 10 volts) was exceeded.

Time buffers store the sampled waveform of the signal (a series of voltages at each sampling point). A signal can be entered into a T buffer from the acquisition buffer as described above, or can be read in from a disk file. In addition to storing a sampled waveform, a T buffer can also store numeric data in the form of measurements made on other buffers (see below). SIGNAL and RTSD both read and write T files using the same SIGNAL format. In addition, SIGNAL can read and write T files in .WAV, .AIFF, and binary formats for interchange with other programs.

F and FT buffers contain frequency domain representations of a T buffer, or part thereof. Both buffer types can be written to, and read from disk files. An F buffer contains a single power spectrum calculated by the discrete Fourier transform (DFT). The user has control of the parameters necessary to derive an F buffer: frequency resolution (transform length), window (rectangular, Hanning, Hamming), and magnitude scaling. Transform length has an upper limit of 32 kpoints, but buffers longer than this can be analysed using a separate command (XFL) that calculates the average of successive transforms. Signals shorter than the desired transform length can be zero-padded so that the desired frequency resolution can be obtained on short signals.

An FT buffer contains the familiar spectrographic representation of a T buffer calculated by a series of DFTs at short increments. In addition to the parameters associated with calculation of F buffers, the user can select the number of transforms (steps), or the transform increment, calculated in the

FT buffer. Increasing the number of steps increases computation time, but produces a more detailed spectrogram.

SIGNAL has extensive graphics capabilities. Figure 1 shows a "standard" view of a T buffer on top with its associated FT buffer on the bottom. The user controls the dynamic range displayed on the spectrogram. A status line below the FT graph shows the settings used in producing the FT buffer, and below that is a caption that contains information about the recording. I find the ability to put captions on buffers invaluable. A screen cursor can be invoked to make measurements on any buffer displayed on the screen. Cursor values are displayed on the screen and can also be saved into "user variables" or as successive points in T buffers (for example, one T buffer for amplitudes, another for frequencies). The values can also be written to a disk file. The cursor can also be used to select regions of a graph on both the frequency and time axes to "zoom" in on the selected region. For example, either the T or FT buffer in Figure 1 can be overwritten with an expanded section of part of the entire T buffer. In addition, an F buffer can be calculated beginning at a selected cursor position with the result replacing the T graph. The original T buffer is not overwritten in this process, and the entire buffer can be recalled and graphed.

The default graphics capabilities will probably suffice for most users. However, SIGNAL gives the user wide control over the content of the graphics screen. Figure 2 displays one example of four different graphs. Graph (a) is a spectrogram of a white-crowned sparrow song calculated from the T buffer in graph (c). The spectrogram status line present in Figure 1 has been suppressed here. The cursors were used to select the second of the paired syllables in graph (a), a new FT buffer calculated, and the result graphed in (b). Note the time and frequency ranges differ from graph (a). Graph (d) is a smoothed F buffer (16 kpoint FFT) of the initial whistle of the song. The placement and scaling of each graph on the screen is at the user's discretion, as is the use of cross hatching (in graph d) versus ticks in the other three graphs. The content and precise position of each label is also under user control. SIGNAL can print output at up to 600 dpi on a laser printer to produce publication quality graphics.

SIGNAL contains extensive capabilities for signal analysis. In lieu of the user manually making measurements with the screen cursor, many measurements (RMS level, maximum, minimum, etc.) on T, F, and FT buffers can be made automatically using the ME(asure) command. The results are stored in user variables that can be manipulated or saved to a file. Amplitude changes can be tracked either from a waveform or a spectrogram to produce an amplitude envelope (Figure 3). Spectral contours, the measurement of frequency variation with time can be produced by four methods: 1) Manually with a mouse, 2) zero crossing analysis, 3) spectrogram contour detection (Figure 3), or 4) Hilbert transform. It is also possible to track the spectral contours and amplitudes of individual harmonics within spectrally complex sounds.

The amplitude envelope and spectral contour can be measured, or edited and modified to synthesize signals based on the sound they were originally derived from. For example, Figure 4 shows the original note, and a synthesized version after 1000 Hz was added to the spectral contour in Fig. 3d. The modified spectral contour was fed through a digital voltage controlled oscillator to produce the synthetic note in Fig. 4. The decomposition and

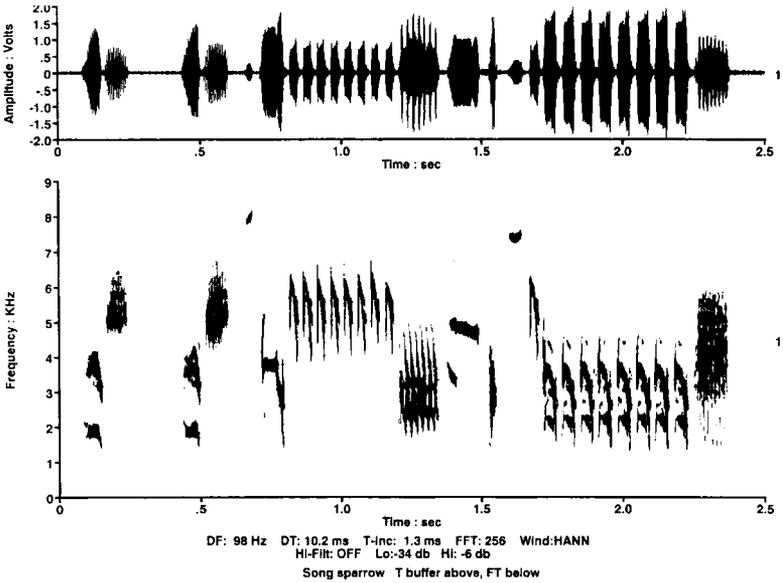


Figure 1. SIGNAL FT buffer (“spectrogram”) below with T buffer above. The status line below the FT buffer lists the transform parameters and graphics dynamic range (Lo, Hi). Below the status line is a user-entered caption. The caption is stored with the associated buffer in disk files. All the Figures in this paper were produced at 300 dpi on a Hewlett Packard laser printer.

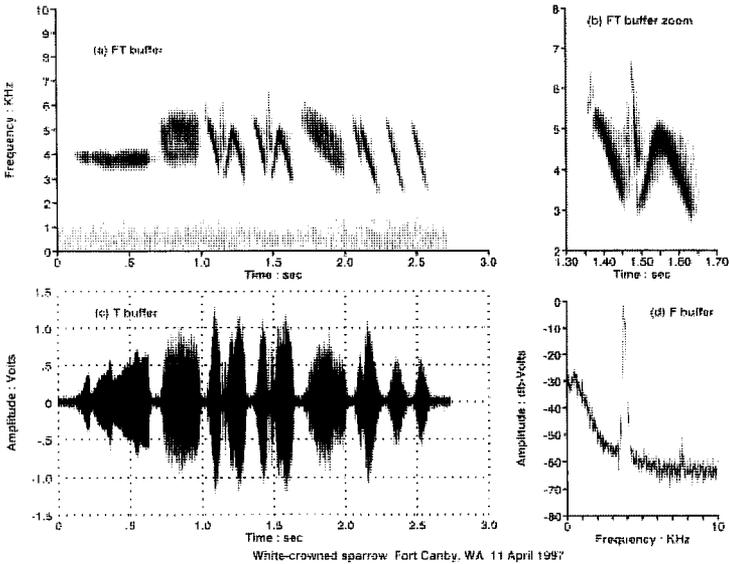


Figure 2. Demonstration of some of the graphics capabilities in SIGNAL. See text for description.

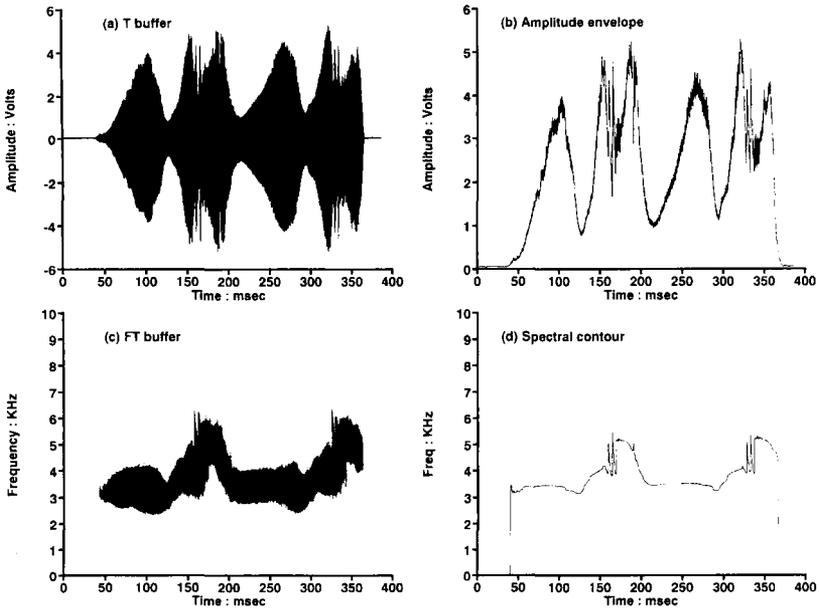


Figure 3. Examples of two analytical techniques in SIGNAL. Graph (b) is the amplitude envelope extracted from the full rectified waveform in graph (a). A user selectable decay constant controls the detail preserved in the waveform. The spectral contour in graph (d) is created from the FT buffer in (c) by measuring the frequency with maximum amplitude at each time step in the FT buffer. The resulting T buffer (d) is expanded by interpolating points to match the original duration and sampling rate of the FT buffer.

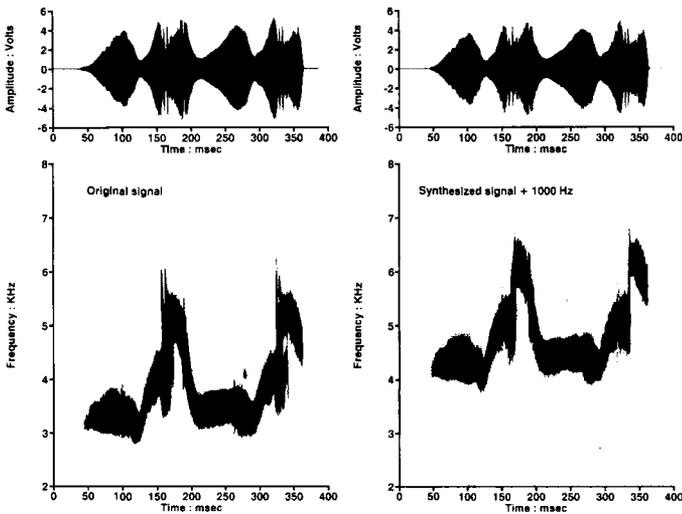


Figure 4. An example of tonal synthesis. 1000 was added to the spectral contour in Fig. 3d to shift the note higher in frequency without changing temporal properties.

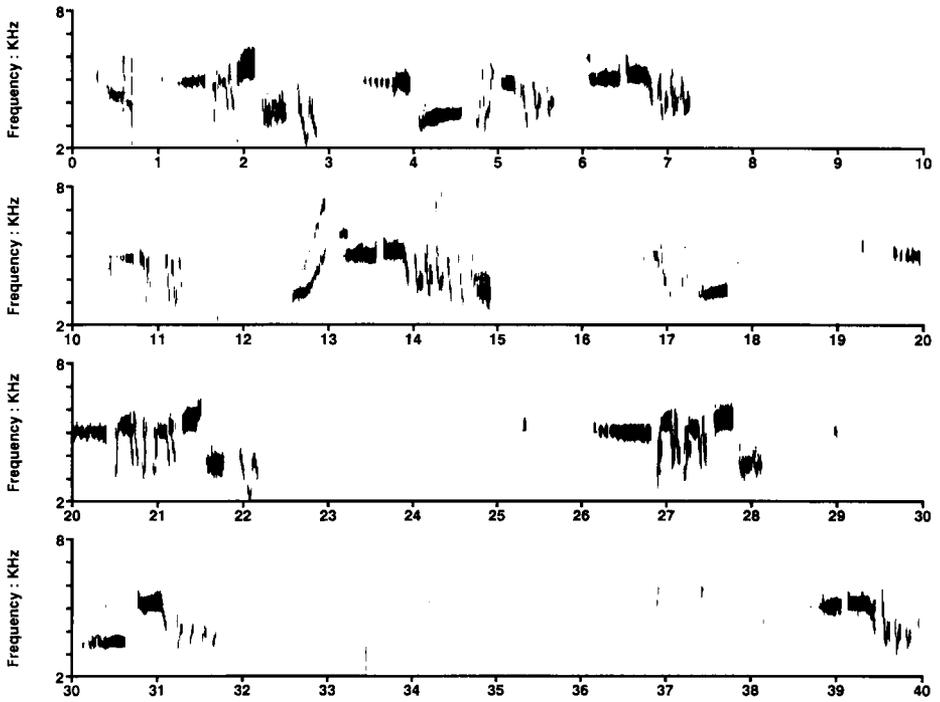


Figure 5. Example output from the STRIP command which produces a continuous spectrogram. The user controls all parameters of the output, as with any FT buffer in SIGNAL.

synthesis of sounds are implemented in the SYNIN and SYNOUT commands. The example shown here is of a tonal sound, but SIGNAL contains more extensive synthesis options for polytonal, harmonic, and pulse repetition sounds.

Cross-correlation of FT, T, and F buffers can also be performed to provide an index of sound similarity (Clark et al. 1987, Khanna et al. 1997). Smoothing and digital filtering are useful to the average user, while advanced users will find convolution, cepstral analysis, Hilbert analysis, and complex spectrum analytical techniques.

The outstanding feature of SIGNAL is its programmability. All of the graphical and analytical techniques mentioned above can be incorporated into programs. While Macintosh/Windows generation users may complain about the DOS interface, some of us who cut their computing teeth on key punches and teletypes find it natural and faster to type in commands, which can be incorporated into SIGNAL programs or macros. Beginning users can carry out SIGNAL's basic functions through a series of menus. As a user gains experience, they will almost certainly want to use "program mode." The ability to program complicated, or even simple, operations separates SIGNAL

from other sound analysis packages that I am aware of. In a lab with multiple users, each user can save a "startup" file on disk, which when executed at the beginning of a work session will configure SIGNAL for that user's specific needs (number of T, F, FT buffers, transform lengths, etc.). There are enormous advantages in having a program guide you through the repetitive operations involved in making many measurements on a large data set of recorded sounds. The ascii data files so produced are then read into a statistical program for further analysis. In addition, I find SIGNAL programs useful in preparing playback tapes for tutoring birds in the laboratory, or in field and laboratory experiments on the perception of sounds. SIGNAL can normalize the amplitudes of stimuli, and then present them with order, number of repetitions and intervals precisely controlled.

With the RTSD the user can easily acquire long continuous recordings. While it is possible to visually scroll through the recording displayed on the terminal, it is often desirable to have a hardcopy record of long recordings. Previously, only dedicated real-time spectrum analysers that produced photographic (Hopkins et al. 1974) or dot-matrix printer output, or the Kay DSP Sonagraph™ (with additional memory) have been available to analyze recordings over several minutes in length. The STRIP command in SIGNAL 3.1 can read in long files and produce continuous spectrograms "wrapped" onto successive lines and onto successive pages. Figure 5 shows 40s of continuous practice singing by an immature white-crowned sparrow.

A recent addition to the SIGNAL family is the Event Detector, a module that runs in SIGNAL, and detects "events" according to user-defined criteria. The detector can be configured to detect the onset and offset of individual notes in a string of notes comprising a call, or to count the entire call itself as one event. The Event Detector should be useful in analysing long sequences of sounds. An entire 60 min tape could be saved to hard disk using RTSD. The disk file can then be processed by Event Detector with the time onset and offset of events stored in a list file. The list file of events can then be processed and examined one by one using a SIGNAL command file. The successful use of Event Detector requires a good signal-to-noise ratio within the frequency bandwidth of the desired events. Out of band noise can be digitally filtered and will not degrade performance. The Event Detector is an expensive (\$2000) option, but I could imagine it paying for itself in the analysis of large datasets.

Summary

SIGNAL/RTSD comprise an extremely sophisticated and powerful package for the analysis of sound and vibration. All the programs' functions are fully described and illustrated with examples in the 700+ page User's Manual. To help learn SIGNAL, there are on line demonstrations of the program's commands, on line tutorials of different signal processing concepts, and an extensive on line help menu. My only real complaint is that SIGNAL and RTSD require separate DOS environments, so the computer needs to be re-booted when switching between programs. I feel this is a small inconvenience considering the power and features these programs provide. It is safe to say that almost any analytical or synthetic tool required in the study of animal

sound is either already implemented or can be performed within the open, programmable architecture in SIGNAL.

REFERENCES

- Clark, C.W., Marler, P. & Beeman, K. (1987). Quantitative analysis of animal vocal phonology: An application to swamp sparrow song. *Ethology*, **76**, 101-115.
- Hopkins, C.D., Rossetto, M. & Lutjen, A. (1974). A continuous sound spectrum analyzer for animal sounds. *Z. Tierpsychol.*, **34**, 313-320.
- Khanna, H., Gaunt, S.L.L. & McCallum, D.A. (1997). Digital spectrographic cross-correlation: tests of sensitivity. *Bioacoustics*, **7**, 209-234.
- McGregor, P.K. & Ranft, R.D. (1994). Equipment for sound analysis and playback: a survey. *Bioacoustics*, **6**, 83-86.
- Riede, T. (1995). SIGNALIZE: Signal analysis for speech and sound. *Bioacoustics*, **6**, 327-331.
- Stoddard, P.K. (1990). Audio computers – theory of operation and guidelines for selection of systems and components. *Bioacoustics*, **2**, 217-239.
- Wilkinson, G.S. (1994). Canary 1.1: Sound analysis software for Macintosh computers. *Bioacoustics*, **5**, 227-241.
- Williams, J.M. & Slater, P.J.B. (1991). Computer analysis of bird sounds: a guide to current methods. *Bioacoustics*, **3**, 121-128.

SIGNAL costs \$2500 and the RTSD \$1595 not including hardware. Inquiries in the U.S. can be directed to: Engineering Design, 43 Newton Street, Belmont, MA 02178 (phone 617-484-3520, FAX 617-484-6559, email 70700.701@compuserve.com). In Europe: Noldus Information Technology, Costerweg 5, 6702 AA Wageningen, Netherlands; (phone +31-317-497677, FAX +31-317-424496, email info@noldus.nl, http://www.noldus.com/product/signal/prod_sig.htm). Either source can handle inquiries from other countries.

Douglas A. Nelson
Borror Laboratory of Bioacoustics
The Ohio State University
USA