Progress in the Search for Extraterrestrial Life
ASP Conference Series, Vol. 74, 1995
G. Seth Shostak (ed.)

THE SERENDIP INTERFERENCE REJECTION AND SIGNAL DETECTION SYSTEM

CHARLES DONNELLY, DAN WERTHIMER, STUART BOWYER, and JEFF COBB
Space Sciences Laboratory, Center for EUV Astrophysics
and Department of Astronomy
University of California, Berkeley, California 94720

ABSTRACT The SERENDIP Interference Rejection and ETI Notification system (SIREN) was developed at the University of California, Berkeley for use in the search for extraterrestrial intelligence (SETI). The SIREN system is currently employed in the SERENDIP III ETI search, underway at the Arecibo radio astronomical observatory in Puerto Rico. In this paper we will discuss the SIREN system and its radio frequency interference (RFI) rejection and signal detection algorithms. These algorithms employ dynamically adaptive statistical analysis techniques to reject RFI and identify ETI signals. SIREN senses fixed frequency, drifting frequency, and spread spectrum interference and then selects ETI candidates by identifying statistical anomalies in the RFI-cleaned data set. SIREN conducts massive database searches looking for repeated detections, does beam pattern matching on successive detections, and tracks drifting signals with drift rates up to 10 Hz per second. With the SIREN system, we are able to conduct quality narrowband spectral analysis in the high-interference 424–436 MHz band. In this band we find that 99% of the high powered detections are spurious signals from terrestrial and near-space sources, yet SIREN's selective rejection algorithms tag only 2% of the band as RFI-corrupted.

INTRODUCTION

The University of California, Berkeley SETI project, SERENDIP, is a radio astronomy research effort aimed at detecting narrowband radio emission from extraterrestrial civilizations. SERENDIP is an acronym for the Search for Extraterrestrial Radio Emissions from Nearby Developed Intelligent Populations. SERENDIP's observing model is based on an opportunistic piggyback approach in which SETI observations are conducted simultaneously alongside an observatory's regularly scheduled astronomical observing program.

As a commensal experiment, the SERENDIP science team does not control telescope pointing, and has only limited control over observing frequency. However, given the enormity of the postulated SETI search volume in frequency, sky position, signal strength, and duty cycle, the trade-offs make piggyback observing programs an attractive alternative to the more resource-intensive search efforts. SERENDIP unobtrusively conducts long-term observing programs on the world's most sensitive radio telescopes at a fraction of the cost of comparable SETI efforts that require dedicated telescope resources. SERENDIP III, the latest generation SERENDIP system, is currently installed at the Arecibo observatory in Puerto Rico. To date SERENDIP III has observed over 89% of the sky visible from the Arecibo telescope. For more information on the SERENDIP search program see Bowyer *et al.* (this volume).

THE SERENDIP DATA ANALYSIS SECTION

The SERENDIP Interference Rejection and ETI Notification (SIREN) system is key to SERENDIP's potential for successful detection. SIREN is a data analysis software system designed by the Berkeley science team specifically to support SERENDIP's piggyback operations. SIREN conducts data analysis in two phases: real-time and off-line.

Figure 1 is a schematic illustration of the SIREN system. The figure shows SIREN's seven main processing tasks: post–FFT (fast Fourier Transformation) processing, data logging, health monitoring, dead-time filtering, radio frequency interference (RFI) rejection, reference frame correction, and signal detection. Two of the seven processing tasks, post–FFT processing and data logging, run in real time.

The SIREN real-time system is currently operating at the Arecibo observatory as a post-processor to the SERENDIP III instrument. The SERENDIP III instrument is a four-million-channel, FFT–based spectrum analyzer with 0.6 Hz frequency resolution and a 1.7 second integration period (2.5 MHz instantaneous band coverage). SERENDIP III is now conducting observations in the 424–436 MHz frequency band.

To cover the entire 12 MHz band, SERENDIP III steps across the passband in 2.5 MHz increments by mixing the intermediate frequency (IF) signal from the telescope with a signal generated by a programmable local oscillator. In this way SERENDIP III scans the entire IF signal available from the telescope's line-feed receiver system. The SERENDIP III hardware system is discussed in detail in Werthimer *et al.* (this volume). SIREN's post–FFT processing activities are powered by an Intel 80960-CA 33 MHz processor running highly optimized assembly language programs that take advantage of the 80960-CA super-scaler architecture. Each critical routine streams multiple instructions per clock cycle to the processor core. The total number of instructions executed each second thus exceeds the processor's clock speed by nearly a factor of two.

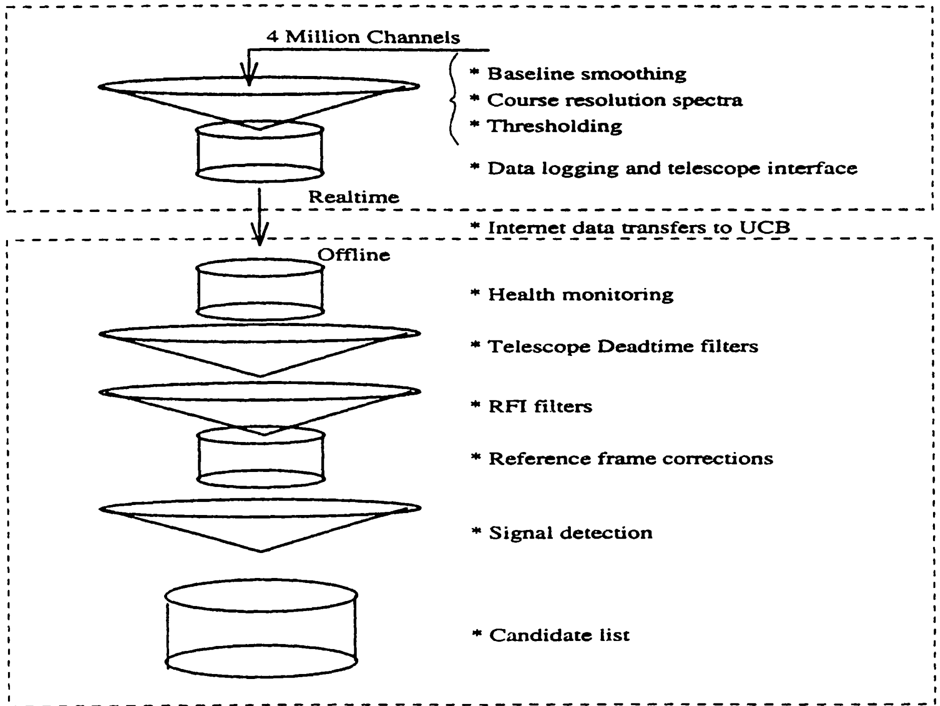


Figure 1. Schematic illustration of the SIREN data analysis. Two main data processing tasks are conducted in real-time at the observatory. The remaining data analysis is conducted off-line at UC Berkeley. The conic icons represent data reduction filters. The pipe icons represent data conversion or health monitoring filters. To date SIREN has processed over 40 trillion radio channels, and has tagged 164 coordinates for further investigation.

The SIREN real-time system processes four million spectral powers every 1.7 second integration period. The major real-time tasks include baseline smoothing, computing coarse resolution spectra, and detecting and logging power peaks in any of the four million spectral channels. Baseline smoothing is conducted by dynamically sliding an eight thousand channel local-mean boxcar across the spectrum. The local mean establishes a statistical basis, with 1% rms error, for signal thresholding. To gain sensitivity to modulated signals, a sequence of progressively coarser resolution spectra are computed. Spectrum $n+1$ in the sequence is computed from its predecessor, spectrum n, by adding the power in adjacent bins to form a new spectrum whose frequency resolution is given by 0.6×2^n Hz. All channels are compared with a threshold based on the mean spectral power and recorded along with time, pointing coordinates, detection frequency, and signal power if they exceed the detection level.

Data are transferred from Arecibo to Berkeley across the Internet where off-line data analysis activities begin. Instrument health is monitored by checking for detection of a weak artificial test signal that is periodically injected into the IF. Telescope dead-time filters then remove data acquired during telescope activities that are incompatible with data analysis. Telescope slew movement is critical to SERENDIP's RFI detection scheme, but extra rapid slew rates preclude acquisition of accurate positioning information. SIREN detects both extra rapid and retarded telescope movement, and flags all data acquired during these periods.

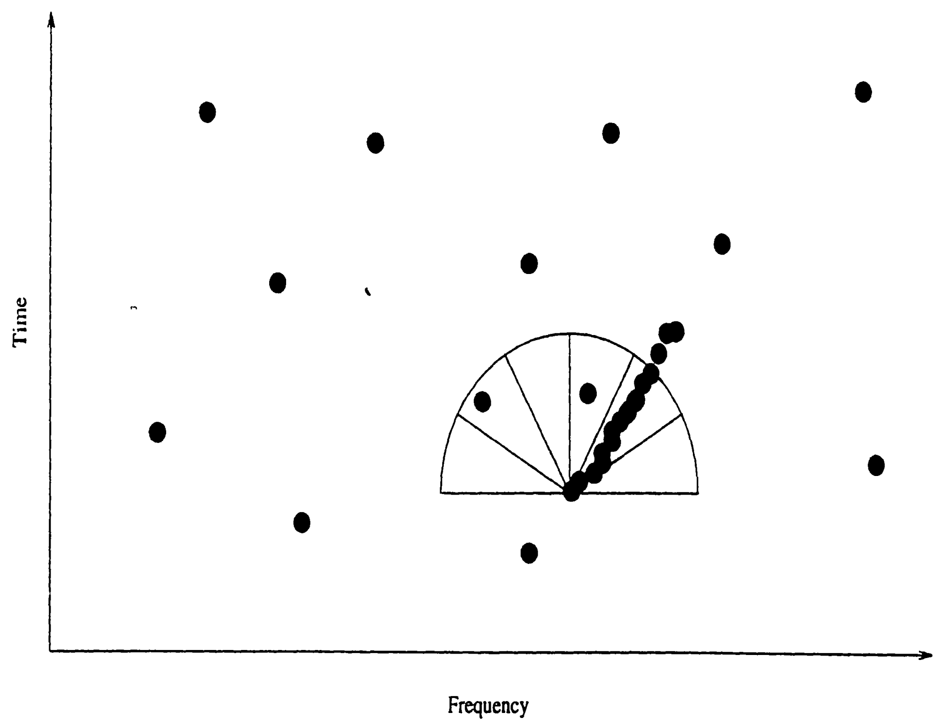


Figure 2. Illustration of SIREN's drifting RFI detection algorithm. A sectored ellipse of frequency-time space around each detection is analyzed for signals that exhibit time-coherency and persistence over multiple pointings of the telescope's beam.

SERENDIP's RFI rejection algorithms incorporate dynamically adaptive statistical analysis routines that detect spurious signals from terrestrial and near-space sources. Four cluster analysis tests are conducted on each input data file spanning several hours of observation. Signals that (1) are detected over broad areas of the spectrum in one or more integration periods (spread spectrum interference), (2) persist in receiver frequency through multiple telescope beams, (3) leak in at the baseband or (4) drift in frequency but persist over multiple telescope beams are identified and rejected.

Spread spectrum interference is identified by the rejection test

$$\frac{\Sigma d^2}{x} > S$$

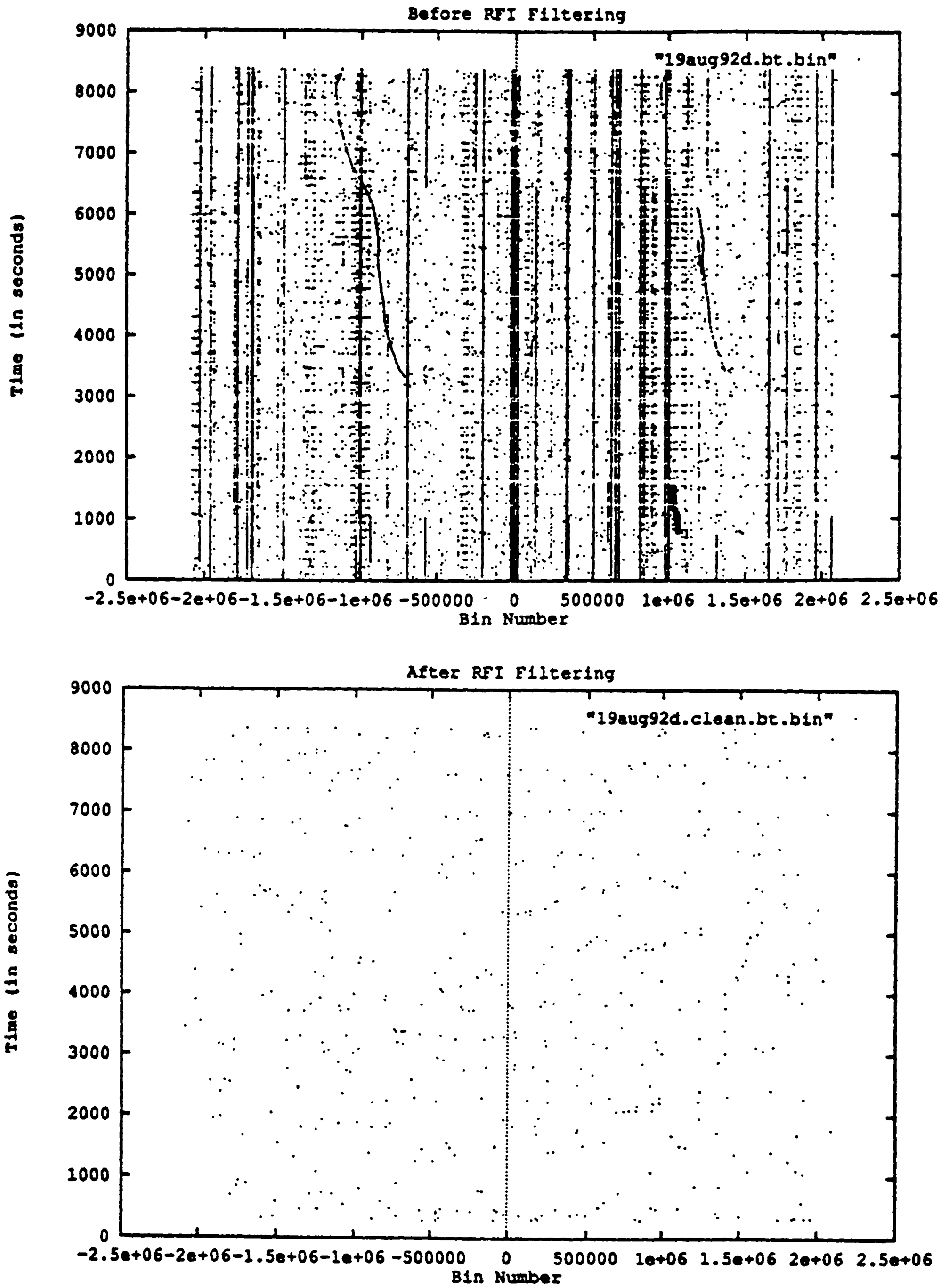


Figure 3. Real data from a period of severe interference. Each dot represents a high power signal detection by the SERENDIP instrument. After RFI filtering the data are nearly Poisson, as expected from a white noise only observing environment.

where d is the number of bins (0.6 Hz/bin) between adjacent events and x is the number of hits in the spectrum. For SERENDIP III, S, as a thresh-

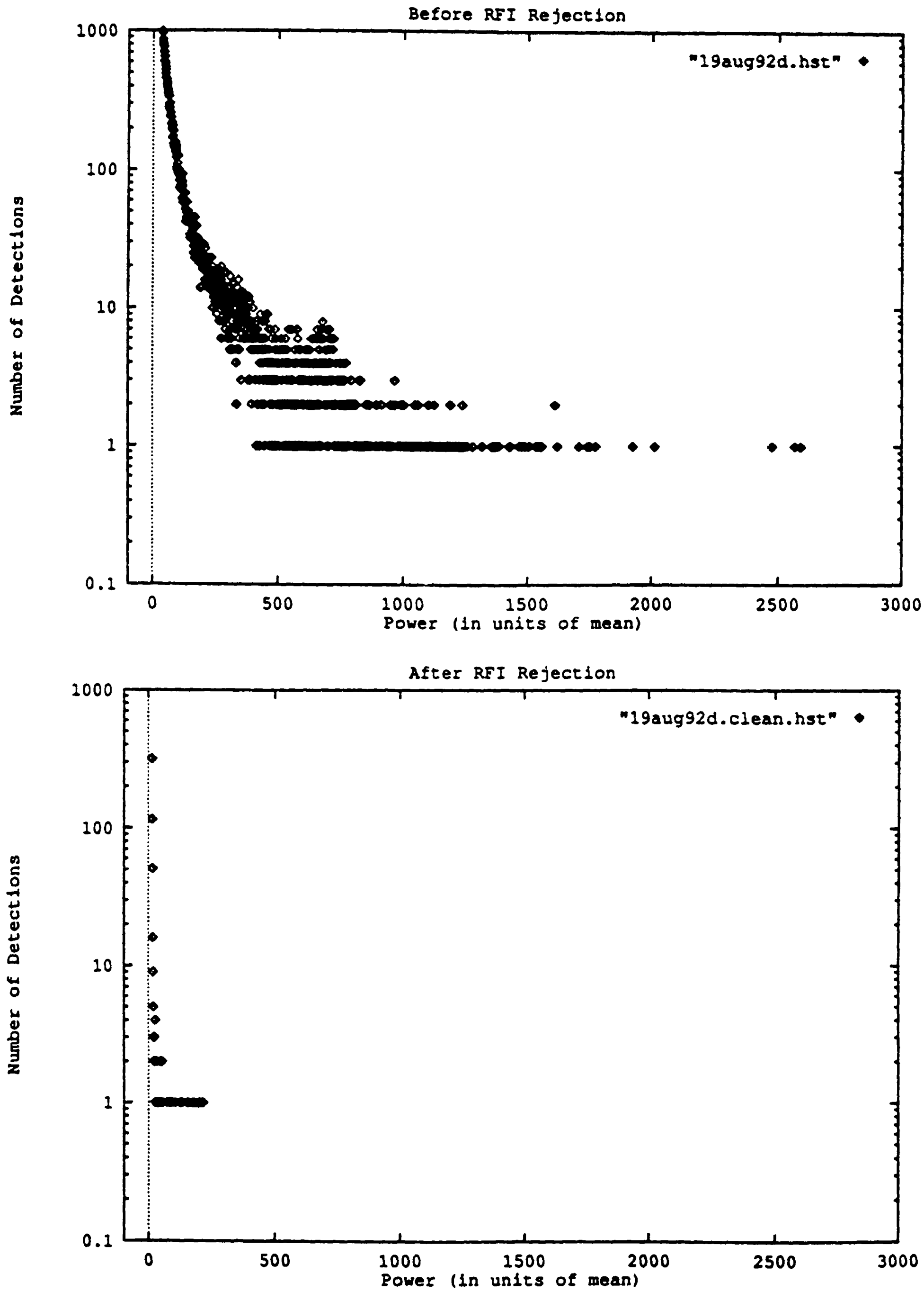


Figure 4. Signal power distribution before and after RFI filtering. Note that after application of the SIREN RFI rejection algorithms, the data approach the expected exponential power distribution of astrophysical background sources.

old value, is set at $S = 10^8$. The spread spectrum test is invoked only for spectra having 50 or more detections.

Rejection algorithms (2) and (3) compute the expected number of detections under pure white noise observing conditions and conduct Poisson analysis on the distribution of events over receiver frequency and baseband frequency bins. Bins with excessive detections are further analyzed for signal persistence over multiple telescope beams. Persisting signals are rejected as RFI.

Data surviving the first three rejection criteria are further analyzed for RFI that drifts rapidly in frequency and is therefore not rejected by algorithms (2) and (3). Figure 2 illustrates SIREN's drifting frequency RFI detection algorithm. Event density in frequency-time space is calculated for each data set. A sectored ellipse in frequency-time space around each detection is analyzed for sectors having an unusually high number of detections. Sectors exceeding the count threshold are further analyzed for the presence of drifting signals that persist through multiple pointings of the telescope beam.

Figure 3 shows a period of particularly severe interference. Note that after application of SIREN's RFI filters, data approach the expected Poisson noise distribution. Of particular note, SIREN's drifting frequency RFI detection algorithm extracted drifting interference detected between 3,000 and 8,000 seconds into the observing run.

Figure 4 shows the power distribution of a data set before and after RFI filtering. Note that after RFI rejection, the data approach an exponential distribution in power as expected from Gaussian background noise.

Signal detection algorithms search for candidate ETI signals in the last phase of the data analysis process. The "clean" data set is frequency corrected to the geocentric and barycentric reference frames, and analyzed for coherent signals. SIREN's signal detection algorithms are sensitive to signals with linear frequency drifts over short time scales, point source signals matching the telescope's characteristic beam profile, and signals that are detected on different days. SERENDIP's entire database of accumulated observations is periodically searched as new data are acquired and added. All signals identified in the signal detection phase are cataloged for further study and possible reobservation.

To date SIREN has processed some 40 trillion spectral bins, logged information on 160 million events detected above the signal threshold, and "selected" 164 candidate detections. Pure Poisson processes over spatial and frequency bins for the SERENDIP accumulative data set and selection criteria predict approximately 100 multiple detection candidates. However, SERENDIP's data analysis system has detected subtle clustering in frequency. The effects of event clustering explain additional detections.

The SIREN system runs continuously as SERENDIP III operates at Arecibo. Modifications to support the next generation SERENDIP system, SERENDIP IV, are currently under way.

ACKNOWLEDGEMENTS

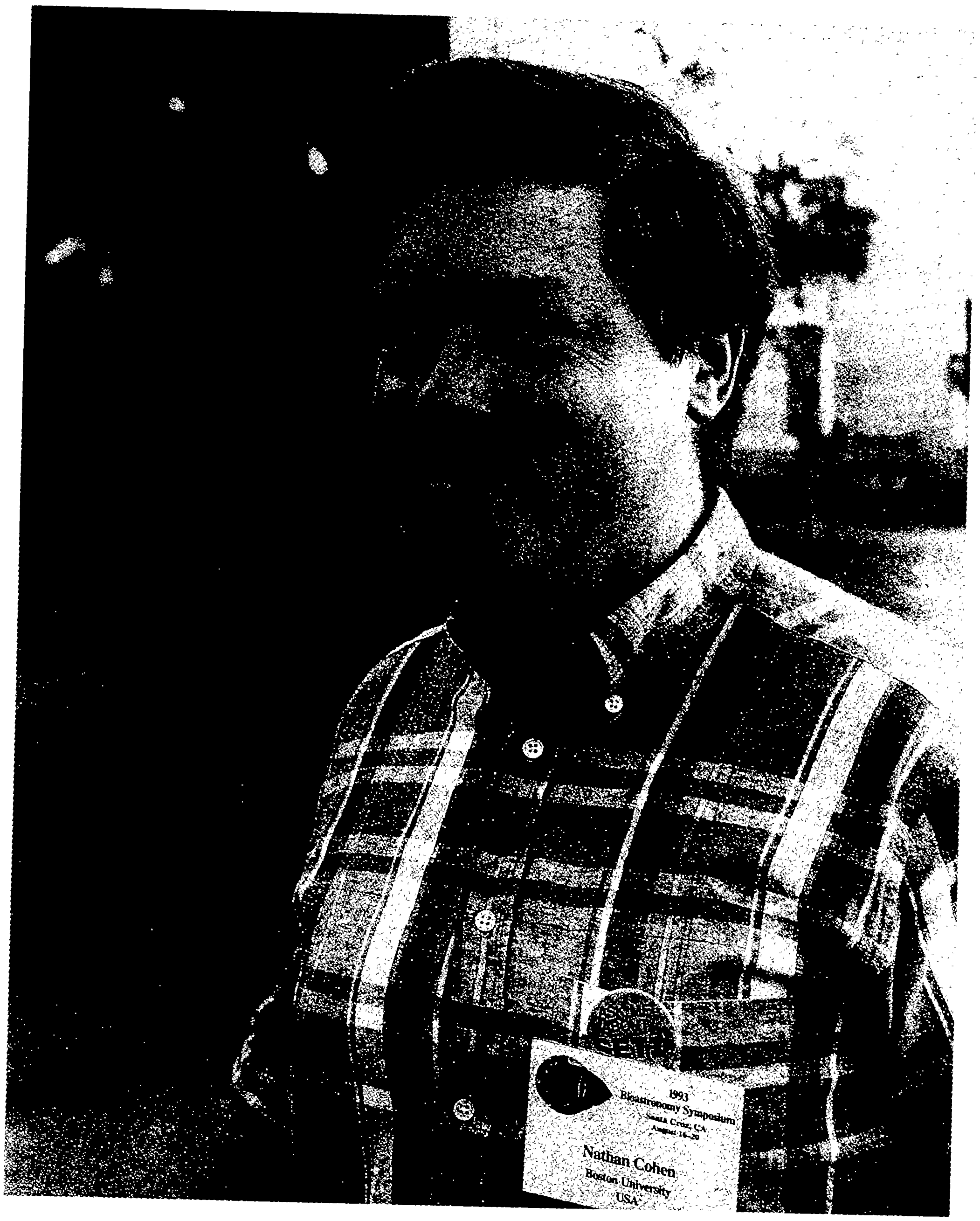
The authors would like to acknowledge the excellent support from the Arecibo telescope operators and staff, and the diligent efforts of the many volunteers who have worked on SERENDIP over the years. In particular we would like to thank Tanwei Fang and Kim Huysken for their recent efforts. The SERENDIP project is supported by NASA grant NAGW-526 and a significant contribution from Mrs. Elisabeth Bowyer. SERENDIP also receives donations from corporations and private individuals. Microtec Research Inc. donated a compiler and other development tools for the 80960-CA processor which were used in the development of the SIREN system. Stuart Bowyer acknowledges support from the John Simon Guggenheim Memorial Foundation.

REFERENCES

Werthimer, D., Ng, D., Donnelly, C., and Bowyer, S. "The SEREN- DIP III and IV Instrumentation," this volume

Bowyer, S., Werthimer, D., and Donnelly, C. "Forty Trillion Signals from SERENDIP: The Berkeley SETI Program," this volume

Steve Levin takes a whack at Woody Sullivan's creampuff pitch.



Chip Cohen has several good lines.