



XXIV. SIGNAL PROCESSING*

Academic and Research Staff

Prof. A. G. Bose
Prof. J. D. Bruce

Prof. C. L. Searle
Prof. H. J. Zimmermann

Dr. M. V. Cerrillo
Z. Barat

Graduate Students

J. R. Boddie
R. S. Copithorne

D. R. Cuddy
M. F. Davis

J. T. Doherty
D. G. Richardson

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

Projects related to recording, processing and reproducing acoustic signals are included in our research on signal processing. Both digital and analog techniques are being used to implement linear, time-variant, and nonlinear filters. Similar techniques are being applied to various forms of visual signals.

An important aspect of this work is the attempt to establish quantitative relationships between physical signals and perceptual response. Numerous experiments conducted over the years have indicated qualitatively that small changes in physical signals can cause large perceptual changes. Accordingly, the consideration of sensory mechanisms and perceptual phenomena has important implications for the performance of communication systems.

1. Signal Separation

Our research on signal separation has focused down rather sharply during 1972 to the problem of auditory localization: What linear and nonlinear signal processing does the human employ to localize sound, particularly in the vertical median plane? Our immediate objective is to identify the cues that we use to make these localization judgments; understanding and modeling the process will follow.

Examples of possible cues are monaural spectral coloration (both power spectrum and phase) produced by the pinna, and interaural differences in each of these spectra. Because people utilize different cues under different experimental conditions, it will be necessary to determine the weights of these cues, and how the weights change.

The experiments contemplated or now under way to answer the question of cues and their relative weights can be broadly summarized under three headings.

a. Physical measurements: such things as pinna frequency response versus elevation angle, and detailed comparison of left and right pinna response.

b. Psychophysical measurements: experiments to determine the accuracy with which humans can perform various localization tasks under different experimental conditions, such as restricted frequency range of the stimulus, monaural listening, etc.

c. Simulation by earphone or speaker: starting with a single source (voice, music, noise), we want to create the illusion of a sound source that can be moved around in space at will.

Our research on the parameters of musical instruments which are essential to the recognition process was completed in 1972, and was reported in Quarterly Progress

*This work is supported by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, U.S. Air Force) under Contract DAAB07-71-C-0300, and by the Clarence J. LeBel Fund.

(XXIV. SIGNAL PROCESSING)

Report No. 106 (pp. 205-210). We are continuing our attempts to apply these results to enhance signal separation in musical systems.

2. Signals, Noise, and Systems

Atmospheric radio noise in the VLF and LF ranges is caused principally by electrical disturbances in the Earth's atmosphere. It has long been recognized that this noise has a distinctly non-Gaussian character because of its impulsive or bursty nature. A simple impulse model (e. g., a wave of random height impulses with a Poisson time distribution) will not account, in general, for the complex nature of this process, especially at frequencies above 30 kHz. D. A. Feldman investigated a multiplicative noise model,¹ defined as

$$\text{Observed noise} = a(t) \cdot n(t).$$

In this model $a(t)$ represents a slowly varying envelope process of the noise process $n(t)$. This is a Gaussian process whose power density spectrum is determined by the RF bandwidth of the observation receiver. Feldman has shown that the envelope functions $a(t)$ of noise waveforms observed in disjoint RF channels are statistically related. This result has possible significance in the design of digital communication receivers or navigation receivers (analog estimation receivers), since it reduces a difficult non-Gaussian problem to a time-variant Gaussian problem. Solutions to the last problem have been found in many cases. A receiver using this cochannel information is infinitely adaptable, by definition, to the nonstationary character of $a(t)$ caused by weather, time of day, and time of year.

J. T. Doherty, Jr. will make a digital computer simulation of a joint frequency channel VLF/LF binary digital receiver. The Feldman model for the additive bandlimited noise, which accurately simulates the non-Gaussian nature of low-frequency atmospheric noise, will be used.

3. Sound Recording and Reproduction

There are fundamental problems in sound recording which are common to the problem of measuring the total power radiated from an acoustical source at one point in space without encountering the influence of the normal modes of an enclosure around the source. Experiments will be made with acoustic lenses designed to produce a field at one point whose intensity is an approximate measure of the total radiated acoustic power of a source.

The reproduction of sound is influenced by the normal modes of the vibrating transducer diaphragms. We are studying these modes, using holograms of the diaphragms. At present, only single diaphragms have been studied. We plan to use these techniques to study the interaction of normal modes in closely spaced transducers.

A. G. Bose, J. D. Bruce, C. L. Searle, H. J. Zimmermann

References

1. D. A. Feldmann, "A Model for Low-Frequency Atmospheric Noise," Quarterly Progress Report No. 104, Research Laboratory of Electronics, M. I. T., January 15, 1972, pp. 351-362.

A. LOCALIZATION OF SOUND IN SPACE IN THE VERTICAL MEDIAN PLANE

Clarence J. LeBel Fund

D. R. Cuddy, C. L. Searle, R. Tenney

The problem of how humans localize a sound in space has been studied for at least 50 years. It is now generally accepted that the cues we use for localization in the horizontal plane are interaural time delay for frequencies below approximately 1000 Hz, and interaural amplitude differences for frequencies above 2000 Hz. In contrast, there appears to be no consensus about how humans localize sound in the vertical median plane, for at first glance it would appear that both interaural time delay and interaural amplitude differences would be zero for sound sources in this plane.

Several authors (e. g., Roffler and Butler,¹ Blauert²) have indicated that the pinna plays an important role in vertical localization. Only Batteau and Moshier³ have made detailed measurements of pinna impulse response as a function of vertical angle, and these experimental data are now being systematically analyzed.⁴ We therefore have undertaken direct measurement of the frequency response of human pinna as a function of vertical angle in the median plane.

Previous experimenters have been seriously hampered by lack of suitable microphones for recording sounds received by the pinna, but now high-quality microphones (Thermo Electron Model 526-U) which are small enough to be inserted directly in the ear canal are readily available. Thus our experimental apparatus is very simple. The subject is seated in an anechoic chamber, and a microphone is placed in his ear canal with the diaphragm roughly at the junction of the pinna and the canal. A GR wave analyzer is used to generate a swept sinusoid that is fed to a speaker placed directly in front of the subject. The resulting microphone signal is amplified by a tracking filter, and recorded. The experiment is repeated for the speaker at 15° increments, from 0° (directly in front) to 195° in the vertical plane.

Typical results are shown in Fig. XXIV-1. Free-field measurements with the microphone and speaker in the chamber, using the same arrangement without a subject, indicated a maximum amplitude variation of ± 3 dB from 2-12 kHz. Thus the large variations in amplitude that are evident in Fig. XXIV-1 are probably introduced by the pinna.

We have not yet proved what precise features of these frequency-response curves are used by the subject for vertical localization. Ultimately, this question can only be answered by simulation experiments, but our preliminary observations provide some hints. First, in Fig. XXIV-1, as well as in responses taken on some (but not all) other subjects, there is a systematic change in frequency response as a function of vertical angle in the 135-195° range. Specifically, a high-Q notch at 11 kHz appears in the 150° data,

(XXIV. SIGNAL PROCESSING)

which deepens and moves lower in frequency with increasing vertical angle. Preliminary psychophysical tests indicated that this notch indeed has important perceptual significance.

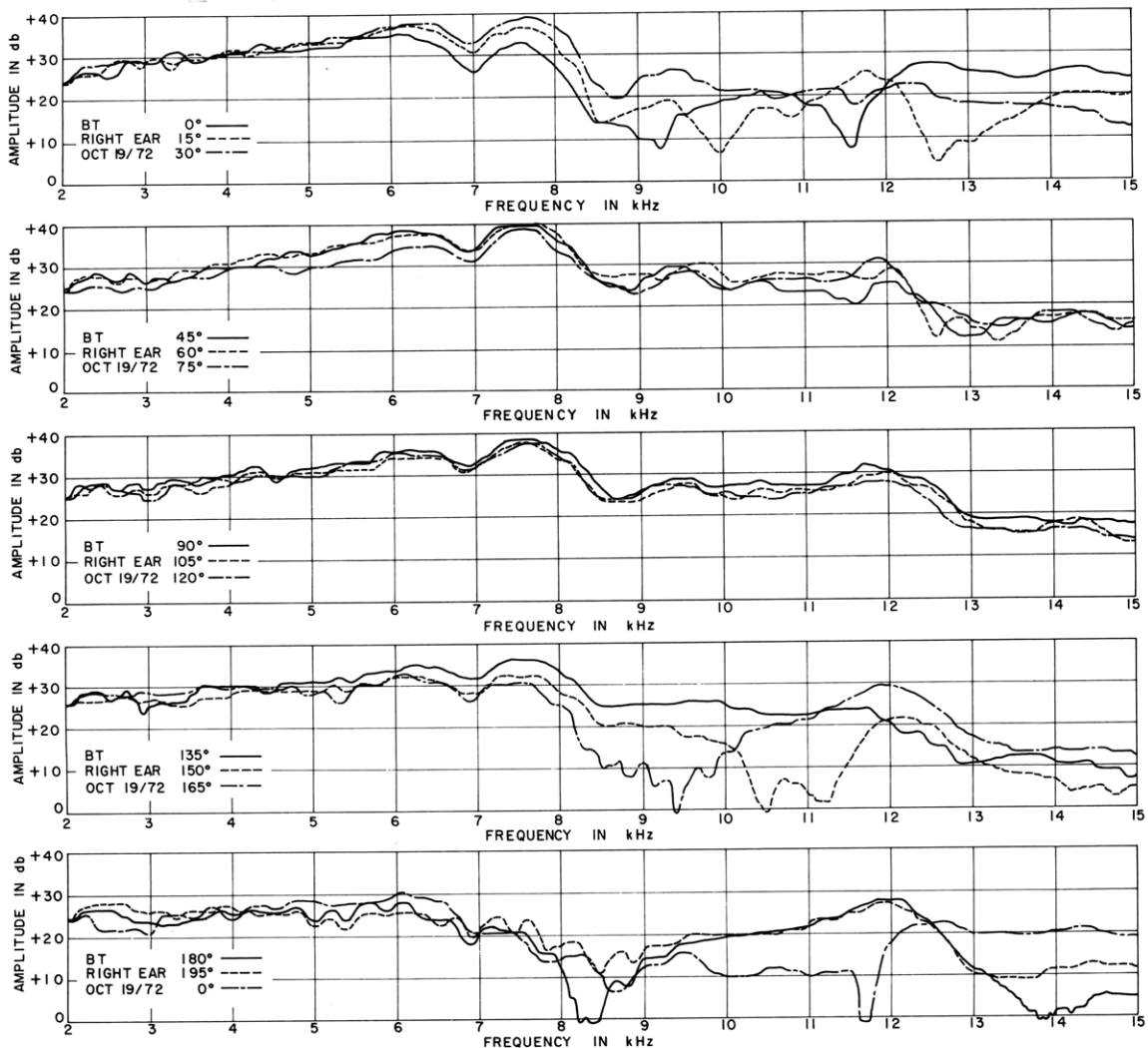


Fig. XXIV-1. Pinna frequency response. Horizontal scale: frequency in kHz. Vertical scale: relative response magnitude, 40 dB from bottom to top.

Second, we have been concerned about the problem of a priori information: Even if the individual pinna responses as a function of angle are known, is it possible to determine the angle of arrival without knowing the shape of the incoming signal spectrum? For example, Roffler and Butler state, "it seems unlikely that a pattern of intensity ratios (developed by the pinna) alone would allow humans to determine vertical signal angle."¹

Our measurements indicate one possible explanation: There appears to be a systematic difference between the left- and right-ear responses, as shown in Fig. XXIV-2. Here we have plotted notch frequency against vertical angle, as directly perceived by the

subject in a monaural listening test. Note that for this subject, the left-ear notch is consistently a few hundred cycles higher than the right-ear notch. We speculate that

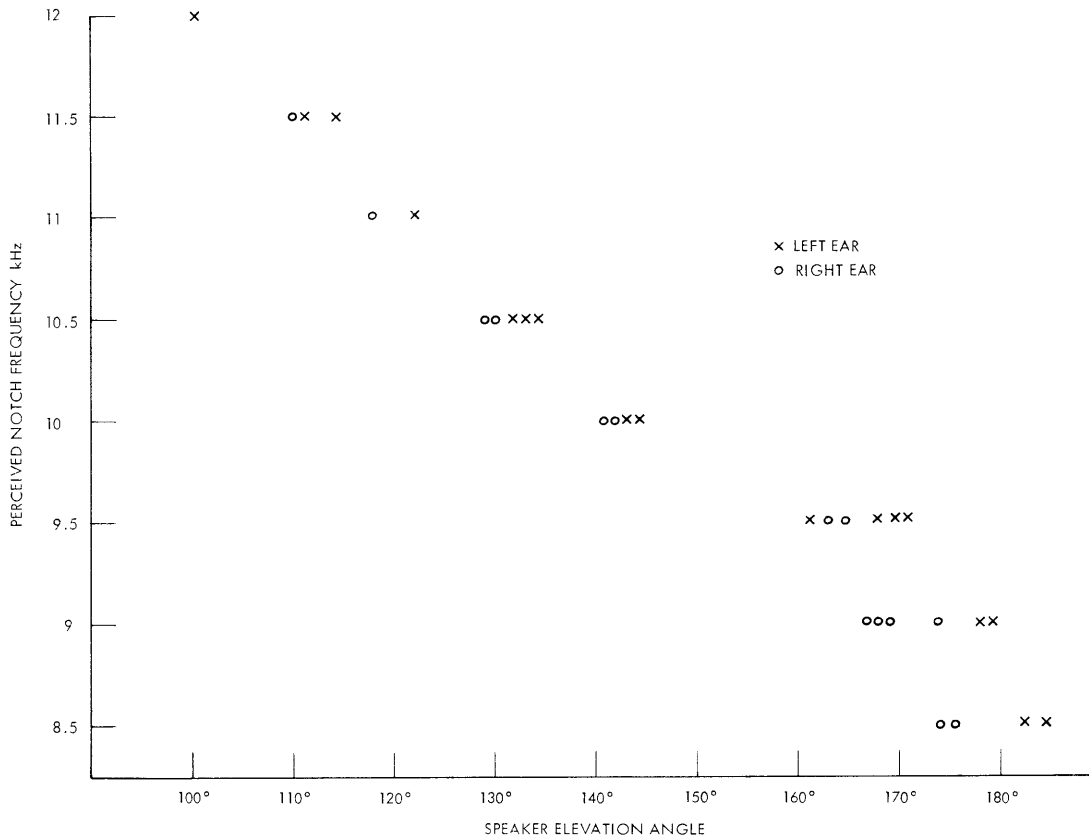


Fig. XXIV-2. Frequency in kHz of perceived amplitude minima vs loudspeaker elevation angle. Subject: CLS.

subjects can exploit these systematic differences (by taking the magnitude of the amplitude differences, for example) to eliminate the effect of incoming signal spectrum, and localize unambiguously on the basis of pinna cues.

References

1. Suzanne K. Roffler and R. A. Butler, "Factors that Influence the Localization of Sound in the Vertical Plane," *J. Acoust. Soc. Am.* **43**, 1255-1259 (1968).
2. J. Blauert, "Sound Localization in the Median Plane," *Acustica* **22**, 205-213 (1969/70).
3. S. L. Moshier, "Synthesis of Auditory Localization Perceptions by Means of Pinna Transfer Functions," *Listening, Inc.*, July 1, 1969.
4. Professor O. McNeil, Walla Walla College, College Place, Washington, Private communication, 1972.

(XXIV. SIGNAL PROCESSING)

B. SIMULATION OF AUDITORY SIGNALS FOR LOCALIZATION

Clarence J. LeBel Fund

J. R. Boddie

A system is being assembled which will process an arbitrary audio-band sound source so that a subject listening to the output through stereo headphones will perceive the sound in some determined position in free space. A minicomputer will control analog elements, such as filters, attenuators, and time delays, to simulate the effects of interaural time differences, intensity differences, and the frequency-response transformations of the pinna. The pinna responses will be approximated with 1/3 octave resolution by passing the signal through a General Radio 1/3 octave bandpass multifilter and digitally controlling the attenuation of the channel outputs. Since an interaural intensity difference caused by head shadowing is frequency-dependent, it can be simulated in the same way as pinna responses.

In order to simulate the effects of interaural time differences, a system has been constructed which variably delays one audio channel with respect to another. Two M-31 analog "bucket-brigade" delay lines, made by Phillips, Inc., were used. In such units, the delay time is inversely proportional to the frequency of the sampling clock. A divide-by N counting circuit was designed to divide the frequency of a high-frequency clock by an 8-bit number. The result is a digitally variable frequency clock that is used as the sampling clock for the bucket-brigade units. Additional logic receives a sign bit that decides which channel is to be delayed with respect to the other. The system has a 15-kHz bandwidth, and delays from 0 to 440 μ s in 8- μ s steps. Longer delays can be achieved by cascading more bucket-brigade units, and the delay-per-step figure can be decreased by increasing the frequency of the fixed clock.