Synthesis and Processing of Audible Notification and Warning Signals

3354 (A-7)

Eric P. Moss and Robert C. Maher
Department of Electrical Engineering
University of Nebraska-Lincoln
Lincoln, NE, USA 68588-0511

Presented at the 93rd Convention 1992 October 1–4 San Francisco

Synthesis and Processing of Audible Notification and Warning Signals:
Design of an Audible Signal Testbed

Eric P. Moss and Robert C. Maher
Department of Electrical Engineering
University of Nebraska--Lincoln
209N WSEC, Lincoln, NE 68588-0511

ABSTRACT

A brief listing of the limitations of audio information transmission and interaction between the audio and visual sensory channels is presented, followed by a description of an automated audible warning signal testbed accounting for those parameters. This station, using a microcomputer and a real-time digital signal processing board, allows real-time processing of two analog input channels, plus synthesis of a third channel. The synthesized channel is composed of "scripts"--timed sequences of pure tones, arbitrary waveforms, noise and combinations of the above. The synthesized signals may be panned between two speakers and mixed with the two live analog channels to create the expected ambient sound field in the target application.

0 INTRODUCTION

While the audio sensory channel has been used since antiquity as a medium for information transmission, few would argue that it has been used to its fullest capability in any given situation. This seems particularly true given recent revelations regarding human error disasters in which audible alarms were intentionally disabled because of their annoying nature (Sorkin, 1988).

Because lives are at stake in situations involving air transport, nuclear power generation and the like, the proper design of audible alarms and information displays is crucial. There seem to exist few standards regarding audible displays, however (Veitengruber, 1977). This may be at least partially due to the lack of a testbed for quickly devising possible audible warnings and testing their effectiveness.

With the availability of powerful personal computers and digital signal processing boards which can provide significant real-time capabilities, this problem can now be investigated. This paper deals with the design of such a system, from a brief analysis of what is known about the audio channel to how a testbed can be designed with those capabilities in mind.

1 THE AUDITORY CHANNEL

The human ear is generally capable of detecting pure tones with frequencies between 20 Hz and 20 kHz. Its response is far from flat, with greatest sensitivity in the 4 kHz range, and is also level-dependent. These effects appear in the form of the subjective parameters of loudness and pitch (Fletcher and Munson, 1933). Whereas frequency and intensity values are directly measurable, pitch and loudness are perceptual values dependent upon both frequency and intensity. For example, more intense tones are perceived to be lower in frequency than tones of the same frequency but of lower intensity. Also, as the ear's response is not flat, tones at the same intensity but with different frequencies will not necessarily be perceived as having the same intensity. To further complicate matters, these effects vary from person to person. The present paper will therefore concern itself with the measurable parameters of frequency and intensity rather than the subjective assignments of pitch and loudness.

The ear-brain system is relatively insensitive to phase, at least between steady-state tones (Licklider and Webster, 1950). However, time delays between the ears allow the localization of sound sources, as do differences in intensity of a signal panned stereophonically (King and Laird, 1930). The above vagaries hint at the capabilities of the audible channel, in that information might be encoded in frequency, intensity and time sequences of the above, which additionally can be placed in different apparent locations.

Although the ear can detect many thousand frequencies, absolute identification of frequencies is typically limited to approximately five (Pollack, 1953). The same goes for loudness. Although humans can generally detect relative differences in intensity down to 1 dB, absolute detection is limited to about five values spaced between whisper-quiet and deafeningly loud (Garner, 1953). In the case of directionality, the minimum audible angle is approximately two degrees between sources located straight ahead, degrading significantly as the sources move progressively to one side of the head (Mills, 1958). The limit of absolute direction identification is around seven, limited for the most part to the horizontal half-plane in front of the observer.

Aside from the absolute acoustic variables of frequency, intensity and apparent location, there exist relative temporal parameters such as repetition rate (which appears as a pitch at higher rates), time delay between signals, length of signal and the like. When combined, these variables can produce a constellation of about 150 absolutely identifiable signals (Pollack, 1953). Unfortunately, such presentation requires considerable concentration on the part of the observer, who must listen to the signals for long periods (tens of seconds), sequentially deciphering each variable's value. As some parameters are not independent (e.g. frequency and intensity) and the subject's memory is not perfect, interference between parameters occurs. The resulting rule-of-thumb is that when the number of acoustic variables is increased, the precision with which each variable may be deciphered in a reasonable amount of time is decreased.

The apparent limit of this progression is that approximately seven variables may be given binary (e.g. high/low or on/off) values with the expectation of consistent identification under ideal conditions of long, quiet analysis intervals. Laboratory conditions are rarely available in the field, however, the most notable difference being noise. Noise masks overt audible signals and audible feedback which tells the operator that a response has been recorded. Noise also masks the operator's "inner speech" that is often commonly used during complex tasks (Poulton, 1977). "Inner speech" is a more scientific term for silently "talking to one's self" in order to improve concentration on a given task. That is why it is often said that intense noise makes it difficult to "hear one's self think."

The above information indicates that not only must an information transmission system be designed to maximize the perceived difference between signals so as to maximize the total number of types of information transmittable, but it must also account for interaction between those signals and corrupting signals in the listening environment. The first requirement is further complicated by the preferences of the operators. For example, pilots prefer that guidance information "sound like the airplane" in order to minimize the amount of decoding needed to initiate a proper response. This indicates that warning and notification signals referring to one subsystem must sound "enough alike" to show their relatedness, thus minimizing the deciphering necessary, but "enough different" to prevent disastrous misinterpretation of the intent. A pertinent example would be that "flaps up" should sound more like "flaps down" than like "landing gear up," but should not sound too much like "flaps down," either.

2 PROPOSITION

Given these complications, the question arises as to whether a warning/information system could be designed which is "better" than those presently available. The affirmation of this possibility requires detailed information regarding both the nature of the information to be presented and the environment with which it must integrate. While a complete set of rules for audible information transmission is more attractive than trial and error schemes, such a set of rules seems unlikely to surface at the moment. A more practical short-term solution would be to devise a method for quickly testing signals in a simulation of the target environment. Such a scheme makes trial and error less time-consuming, and would also be helpful in empirically defining any rules which might exist. Such rules could then be incorporated in the testbed itself. It is therefore proposed to design a testbed which allows the following:

1. manipulation of the intensity and apparent spatial positioning of two "live" channels of sound which can be used to simulate the expected ambient sound field in the target application.

2. synthesis of signals containing one or more of the following-- pure tones, arbitrary waveforms, broadband noise and clicks.

3. manipulation of the parameters of each synthesized signal-- intensity, frequency, duration, start time and apparent location.

4. variation of apparent location and volume over time, as well as graphical display of this information.

5. recording of the script to a disk for archival purposes.

3 TESTBED DESCRIPTION

Hardware

The current testbed uses a 33 MHz 80386 PC running MS-DOS and an Ariel model DSP-56 plug-in board based upon the Motorola DSP56001 digital signal processor. This board supports two channels of 16 bit A/D and D/A at sampling rates of up to 100 kHz. The chip features 24 bit precision fixed point arithmetic running at 10 Mflops, with a 56 bit accumulator to ease the task of data scaling. The DSP can be loaded with an assembly program of up to 512 24-bit words on-chip using a Motorola cross-assembler. The board allows interfacing to the PC controller's program, presently coded in C, providing the means to upload and download data while the DSP is processing.

Software

Assuming that a full 20 kHz bandwidth is desired, the DSP can perform the computational equivalent of 100 multiply/accumulates (MACs) per sample when sampling two channels at 44 kHz. This limits the real-time capability of the system, so that complex signals must first be created off-line, stored in look-up tables on-chip, and finally accessed and replayed when the appropriate cues are received from the host-controller. While much of this calculation could be accomplished at greater speed on the DSP, it is often preferable, practically speaking, to perform as much as possible on the PC, as the code upkeep is far more straight-forward when coded in a high-level language. Further, the DSP has limited on-chip storage capacity for programs, and speed is greatly reduced when program code accesses must occur off-chip. Therefore, this system apportions as much of the design and calculation of the actual signals on the PC--and as much of the real-time processing on the DSP, as is possible. Below is a list of the tasks and how they are assigned.

PC:
graphical display
design of signals
calculation of waveform samples
time base generation and tracking
downloading of mix/volume factors and timing cues to DSP

DSP:
digitizing of analog channels
storage of PC-designed waveforms in look-up tables
access of those tables when cued by the PC
scaling of those values using mix/volume factors

Specifically, as the DSP has a 256 point sampled sine wave stored in on-chip ROM, replication of the effort in the PC is unwarranted. Further, it was deemed more important to dedicate as much of the DSP RAM as possible to arbitrary waveforms. Therefore, the total of the PC's duty in sine generation is to provide an address increment for resampling the pre-stored sine wave.

Digital noise generation has not been given a rigorous treatment in the system prototype as of yet. Presently, noise should be handled as an analog signal passed through one of the live channels. Clicks, or more properly, finite magnitude impulses, are available as complete signals in themselves, or as part of arbitrary waveforms. Their amplitude is user-selectable, with total signal magnitude clamped at unity.

There exist two live input channels which are digitized on the DSP board. These can come from any analog source--a CD player, noise generator, etc. The testbed allows separate, digital manipulation of each input's gain on a linear scale from off (i.e. zero gain) to full-scale (i.e. unity gain). Each input can also be panned left-to-right between two output channels, e.g. loudspeakers. This information is displayed in a box which represents volume (on a linear vertical scale) versus mix (on a linear horizontal scale).

By first selecting which channel to manipulate and then clicking the PC mouse inside the box (see Figure 1), the user can quickly set both the volume and mix of that channel, leaving a dot deposited on the screen to record that choice. By clicking within the vertical margin to the left of the box, independent modification of the gain may be accomplished; similarly so with the horizontal margin and the mix. The selection process is timed on a ten second interval which records the selections for optional disk storage and playback, and which connects the selections with a line to hint at the time progression of the data.

A second data field appears to the right of the mix/volume display. Its horizontal axis represents a ten second interval, its vertical axis represents either the gain of each channel or the mix, depending upon which was selected. As selection of data with a mouse can be imprecise, a means of editing that data is given via this second field. Clicking at two desired endpoints in the field sets those points' values and linearly interpolates the samples between them, thus affording some rudimentary control over precision.

If further control over precision is desired, there exists to the right of the time display box a column of data readouts and corresponding control buttons. By clicking in the topmost pair of buttons, a highlighted datum appears in the editing field, and quickly scans left or right through the samples. Once a particular sample is chosen, the other buttons allow editing of that sample's volume and mix, the results being simultaneously updated in all displays.

Below the time display reside two bars. Clicking in the upper of the two results in a vertical line being drawn in the time display which represents the beginning of a playback selection. Clicking in the lower of the two, to the right of the previous point, draws another marker and highlights the selected data in both displays. Selecting GO begins the timer and replays that selection of mix and volume for the given channel.

Below the volume/mix display box are a set of buttons which allow storage of the "scripts" under user-selected names in user-chosen directories. Both storage and retrieval follow a similar pattern: select the directory box, type the directory path, select the BROWSE button to display the script files (which are assumed to have the extension .SCR), select the desired name or type it into the file selection box, and click on OPEN or STORE. Other buttons allow selection of some quick default types of volume/mix scripts--circular, linear and square, which can be sized under mouse control.

Once the scripts for the two live channels have been set, the test signal can be designed. Clicking on any of the following signal template buttons opens a menu for setting its parameters: noise, pure tone, arbitrary waveform and finite magnitude impulse, i.e. click. There are presently no noise parameters which may be set. Upon completion of FIR and/or IIR filter modules, the bandwidth of the noise will be user-adjustable. The pure tone parameters include frequency and sample rate. The same holds for the arbitrary waveform, with the addition that its 256 samples may be generated in an editing field similar to that of the mix and volume editing field. The click's only parameter is amplitude.

Once set, each of these signal templates' offspring may be assigned one or more starting times and durations, and may be manipulated with respect to mix and volume parameters over its duration. In this way, compound signals may be developed by creating multiple sinewaves of differing frequencies and amplitudes and assigning them the same duration and start times. To pan such an aggregate in the sound field, a single template of mix and volume data may be created and assigned to each component in the group. Initially, this is a bit tedious, but it provides useful flexibility and a consistent programming model from which to work. Eventually, an extensible system menu will be implemented, i.e. a system whereby these aggregates may be added to the basic menu by the user.

This system does not now take into account the actual operational details of the ear-brain system: the measurable quantities of intensity and frequency are used rather than the subjective quantities of loudness and pitch. In addition, no limitations are built into the testbed regarding the absolute discrimination capabilities of humans, such as for the minimum audible angle. As processing power is becoming closer to being "free," excess precision in setting a relative angle between two apparent sources does not seriously degrade the speed of the system, especially since those parameters are dealt with at the PC level. The same attitude was applied to the precision with which other parameters were treated.

4 AN EXAMPLE SESSION

Following is an example session of how this system might be used. The question under investigation is "Does masking of a single 1 kHz tone by 90 dB broadband noise mostly (e.g. 2/3) from the right of the test subject vary with the apparent left-right position of that tone?" To set up the test conditions, one would first determine the total SPL expected in either channel from the noise plus the test tone. This is done so that the DSP accumulator will not overflow. As a quick approximation, one might guess that the noise would require approximately one-fourth-scale on the processor and that the tone would take up approximately one half-scale, leaving some headroom.

The investigator would connect the noise generator to one of the live analog inputs of the board. Then, the mix/volume data of that channel would be set with the mouse so that 2/3 of the signal was routed to the right channel and its volume was set to .25 of full-scale. Then the user would adjust the analog volume controls on the power amplifier until 90 dB was measured with a hand-held meter.

Following this, the user would click on the sine wave template and fill in the parameters, in this case, 1 kHz. This signal could be left on the entire time, formed into bursts by setting up a script which modulated the intensity between full-scale and off. The bursts could also be formed by creating several separate gated sinewaves and placing them at the desired intervals during the 10 second script. Either way, the user would complete the script by setting the apparent location and volume of those bursts. As a further note, the burst pattern could be in the form of several bursts at the same location with varying intensities or the same intensities at varying locations.

Once the data was taken, the investigator might then wish to examine how detection of a stationary 90 dB, 1 kHz tone is affected by 90 dB bursts of noise from different locations. This could be accomplished in two ways. First, the investigator could write two scripts, one for the live channel which simply switches the pre-existing noise on and off and changes its mix and volume values. The other script, for the sine wave, would involve a continuous wave in a single apparent location. The other scheme for accomplishing this would be to leave the live channels unused and write several scripts for the synthesized channel, each of which used different sets of noise bursts (when implemented) in conjunction with a synthesized pure tone. The live analog channels could then be used to add an actual recording from the target application sound field, e.g. an aircraft cockpit.

5 RESULTS

This system is incomplete from the standpoint of not including every acoustic parameter which may be of interest. It therefore needs to have added the capability of generating arbitrary waveforms and cataloguing them for quick retrieval, plus digital filtering. It is also recommended that the capability of arbitrary-length sessions be added to allow for easier vigilance testing. To improve the graphical interface as well as improve the ease of software maintenance, an object-oriented language should be used instead of C. This would allow for easy spawning and manipulation of signals from the signal palettes, easier graphics updates when data changes, and safer addition of editing facilities in the future.

6 CONCLUSION

This paper has described a testbed for designing and testing audible warning and notification signals using a PC and an Ariel digital signal processing board. The system allows quick graphics-based, mouse-driven creation and editing of pure tones, noise, arbitrary waveforms and aggregates of these signals. The system also allows digital manipulation of the signals' intensity and apparent placement in acoustic space. These signals may then be mixed with two similarly manipulated live analog channels. The testbed also provides for storage of the scripts developed to test each type of signal.

7 REFERENCES

Fletcher, Harvey and Munson, W. A. Loudness, its definition, measurement and calculation. *The Journal of the Acoustical Society of America.* Vol. 5. 1933. 82-108.

Garner, W. R. An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology.* Vol. 46. 1953. 323-332.

King, Walter G. and Laird, Donald A. The Effect of Noise Intensity and Pattern on Locating Sounds. *The Journal of the Acoustical Society of America.* July, 1930. 99-102.

Licklider, J. C. R. and Webster, J. C. The Discriminability of Interaural Phase Relations in Two-Component Tones. *The Journal of the Acoustical Society of America.* Vol. 22 (2). March, 1950. 191-195.

Mills, A. W. On the Minimum Audible Angle. *The Journal of the Acoustical Society of America.* Vol. 30 (4). April, 1958. 237-246.

Pollack, Irwin. The information of elementary auditory displays, II. *The Journal of the Acoustical Society of America.* Vol. 25. 1953. 765-769.

Poulton, E. C. Continuous intense noise masks auditory feedback and inner speech. *Psychological Bulletin.* Vol. 84 (5). 1977. 977-1001.

Sorkin, Robert D. Why are people turning off our alarms? *The Journal of the Acoustical Society of America.* Vol. 84 (3). September, 1988. 1107-1108.

Veitengruber, J. E., Boucek, G. P. Jr. and Smith, W. D. Aircraft Alerting Systems Criteria Study Volume 1: Collation and Analysis of Aircraft Alerting Systems Data. FAA-RD-76-222,1. May, 1977.

Veitengruber, J. E., Boucek, G. P. Jr. and Smith, W. D. Aircraft Alerting Systems Criteria Study Volume 2: Human Factors Guidelines for Aircraft Alerting Systems. FAA-RD-76-222,2. May, 1977.

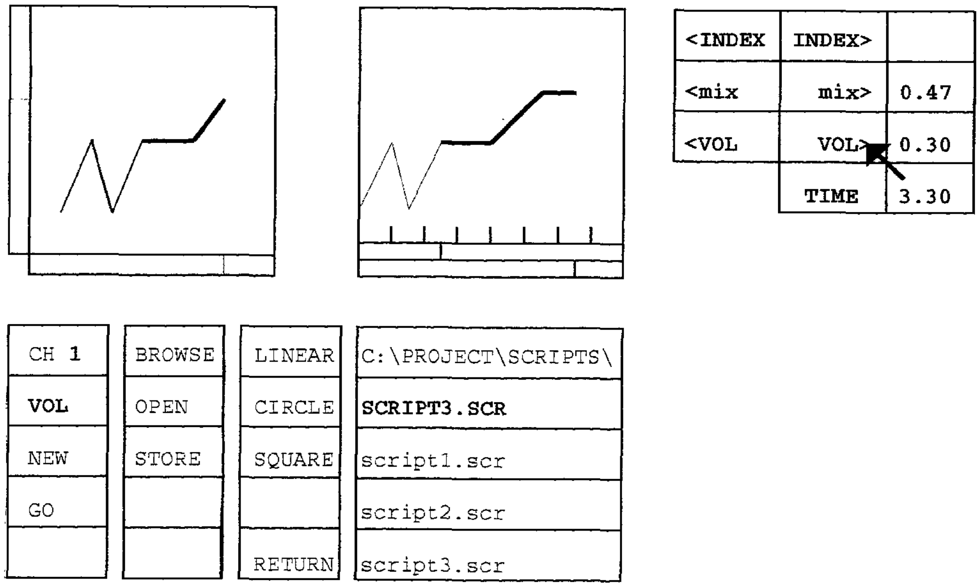

Figure 1. The Mix/Volume Script development screen. Notice that SCRIPT3.SCR is the active script file, that Channel 1 is being developed, and that the intensity of time index 3.30 sec is about to be incremented.