Listening to the Earth sing

1 author:

Chris Hayward
Southern Methodist University

Chris Hayward
Southern Methodist University, Department of Geology, Dallas TX

Listening to the Earth Sing

Techniques for auditory monitoring and analysis of seismic data are described. Unlike many other kinds of data, seismograms may be successfully audified with a minimum of processing. The technique works so well because both sound in air and seismic waves in rock follow the same basic physics, that described by the elastic wave equation. Both exploration seismology, which examines with only the upper few miles of the earth, and planetary seismology, which examines the larger structures including the earth core, may make use of auditory display. Previous published work is limited to two papers now nearly 30 years old, which examine the utility of audio display to the discrimination problem of earthquakes and nuclear explosions. The applications are much broader though, including training, quality control, free oscillation display, data discovery, large data set display, event recognition, education, model matching, signal detection, and onset timing. Problems in audifying seismograms arise when the subsonic wide dynamic range signals must be rescaled to the audio without introducing distracting artifacts. Simple processing techniques including interpolation, time compression, automatic gain control, frequency doubling, audio annotation and markers, looping, and stereo are used to create

seven example audio data sets. These seven examples illustrate the use of audio in presenting synthetic seismograms, shallow reflection data, quality control during field recording, noise analysis for earthquake observatories, earthquake analysis for events from various distances, nuclear explosions, and stereo display of seismic array data. The use of audio for seismic quality control, analysis, and interpretation will develop only when audio displays become integrated into the daily tools of seismologists.

1. INTRODUCTION

Seismograms like the ones shown in Figure 1 are graphs of small vibrations of points on the earth plotted against time. They are one of the essential tools of geophysicists studying the interior of the earth. An audio rendering of these can provide unique insight into seismic modeling results, seismic source and wave propagation characteristics, field-recording quality control, and training and education. This rich and unique opportunity to improve the quality of modern seismic interpretation is almost unexplored and yet is inexpensive, simple, and elegant.

In this volume are many examples in which data traditionally graphed or plotted is presented as sound. Stock market prices may be scaled and converted to musical notes or hospital patient status may be converted to analogs of everyday sounds, what Kramer (in this volume) and others have defined as sonification.

However, few people have been successful in directly treating data points as sound waveform samples and simply playing them, a process that Kramer defines as audification. One of the reasons audification fails for arbitrary data such as stock market figures or daily temperatures is that even a slow playback rate of 8,000 samples/second requires many data points to make a sound of any duration. The resultant short sound does not reveal valuable information to the listener. Even for long sequences, the sounds do not resemble natural or familiar noises because arbitrary data points seldom obey the physics of natural sounds transmitted through air.

Seismic data, however, is an almost perfect case for audification. Seismic data sets are large—seldom less than a few thousand samples and as much as several billion samples. A simple audification will play from several seconds to more than an hour. It will sound like a recording of natural environmental sounds because sounds transmitted through air (acoustic waves) have similar physics to seismic vibrations transmitted through the earth (elastic waves). The physics is similar enough that mathematical models that describe sound transmission through gas are successfully used

for seismic modeling.[6] This suggests that doing the opposite, audifying the seismic data, should produce natural sounds such as those heard in the environment.

The advantage of a direct physically consistent auditory display is that as an analog to natural acoustics, it can take advantage of the vast human experience in interpreting noises. For example, a sharp explosion followed by decaying echoes includes information that is interpreted as the size and shape of the echo chamber. A set of echoes followed by the explosion is recognized as something physically ridiculous or artificial.

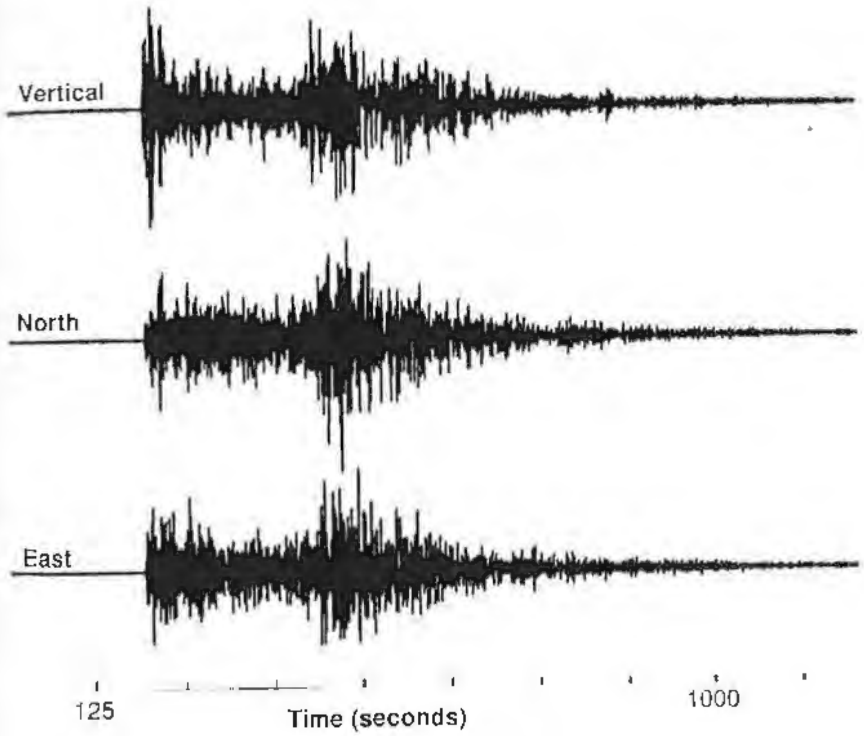

FIGURE 1 Three Component Seismogram of a Nuclear Explosion at NTS Recorded at Lajitas, TX. Each trace records ground velocity in one direction. Data from three GS13 seismometers was recorded at 40 samples/second using a 24-bit digital recorder. The initial impulse represents the arrival of the compression wave. The energy about 4 minutes later is from the shear wave arrivals.

Audified seismograms have been produced as a curiosity by many seismologists. However, no one has demonstrated a common application where auditory presentations are superior to graphical displays or suggested the areas in seismology where such success is likely.

2. BACKGROUND

2.1 COMPARISON OF SOUND WITH SEISMIC WAVES

Most seismic interpretation assumes that the recorded seismograms are elastic waves[1] and, therefore, follow the wave **equation**[3][2]:

This is A misprint. →

$$\frac{\partial^2 y}{\partial t^2} = c^2,$$

where c is constant and $\nabla^2 \equiv \partial^2/\partial x_i^2$. In the equation, c is the velocity that the disturbance travels through the material. This same equation describes the transmission of sound through air. For sound in air c is about 1,100 feet/second. In rock, there are several different kinds of waves that propagate out from the source at different velocities. The fastest compression-rarefractions, or P-waves, travel at 5,000 feet/second in rocks near the surface, and over 40,000 feet/second deep in the earth's mantle. Slower shear wave distortions or S-waves travel at about half the P-wave speed. For sensors and sources near the earth's surface, there are also Rayleigh and Love surface waves that travel at lower speeds.

There are four major differences between the recordings of sound in air and of seismic waves in the earth. Except for unusual circumstances (such as thunder), sounds we experience travel through a simple material, air, with only one velocity. Seismic waves travel at different velocities in each rock type. Sounds have only one type of wave (compression-rarefraction), while seismic waves have several traveling at different velocities. Surface waves that travel along layer boundaries and at the earth's surface are a dominant feature in seismic waves, but are rare in sound. Seismic waves are scaled differently, with dominant frequencies of 40 Hz, much lower than audiable sound.

[1]This is not true in special cases very near the source where the explosion, impact, or earthquake permanently deform or fracture the rock.

[2]For a more general discussion to the elastic equations of motion as used in seismology, see Aki and Richards.[1]

2.2 THE GOALS OF SEISMOLOGY

Geophysicists interpret seismograms to infer the characteristics of the rocks along the energy's travel path or to learn the characteristics of the energy source. Vibrations from seismic sources travel through the earth reflecting and refracting from different layers of rock and, finally, shaking a sensitive displacement, velocity, or acceleration gauge (seismometer or geophone) that records vertical, north-south, east-west, or all three motions.

Studies in the field of seismology can usually be classified by the scale of the phenomena or of the object being studied. Most studies can be classified as either exploration seismology or as planetary seismology.

EXPLORATION SEISMOLOGY. *Exploration seismology* uses a controlled source such as a hammer or small explosive charge, and records reflections and refractions from the rock layers a few feet to several miles below the surface (see Table 1 and Figure 2). Interpretation of these recordings may be used for mineral exploration, characterization of shallow geologic structure for civil engineering studies and geologic hazard estimates, as a way to construct geologic maps of the subsurface or to understand shallow crustal and near-surface processes.

For typical studies, explosions or hammer blows are repeated and recorded on tens to thousands of geophones at the surface to build up a seismic cross section much like a sonar depth section. After each recorded shot, the geophones and energy source are moved forward a short distance.

During interpretation it is assumed that every explosion or hammer blow is identical and every geophone is recording the undistorted ground motion underneath it. Any differences are interpreted as the result of the geologic structure underneath the geophone and explosion. In practice, geophones are sometimes not firmly planted in the ground, which results in a distorted signal. Sometimes hammer blows bounce or are weak. If these problems are unrecognized, they may be mistakenly interpreted as having geologic causes.

PLANETARY SEISMOLOGY. *Planetary seismology* uses recordings of volcanic eruptions, large explosions, and earthquakes (see Table 2). Recordings are interpreted to understand the gross structure of the planet such as the nature of the core, mantel, or crust; the deep structure of continents; and the location and nature of seismic sources (earthquakes, volcanoes, or large explosions).

Planetary seismologists depend on records from seismic observatories located in quiet areas around the globe. Each observatory makes precisely timed high-quality recordings of ground motions that may be associated with recordings from distant stations. These may then be used to develop the location and time of the event.

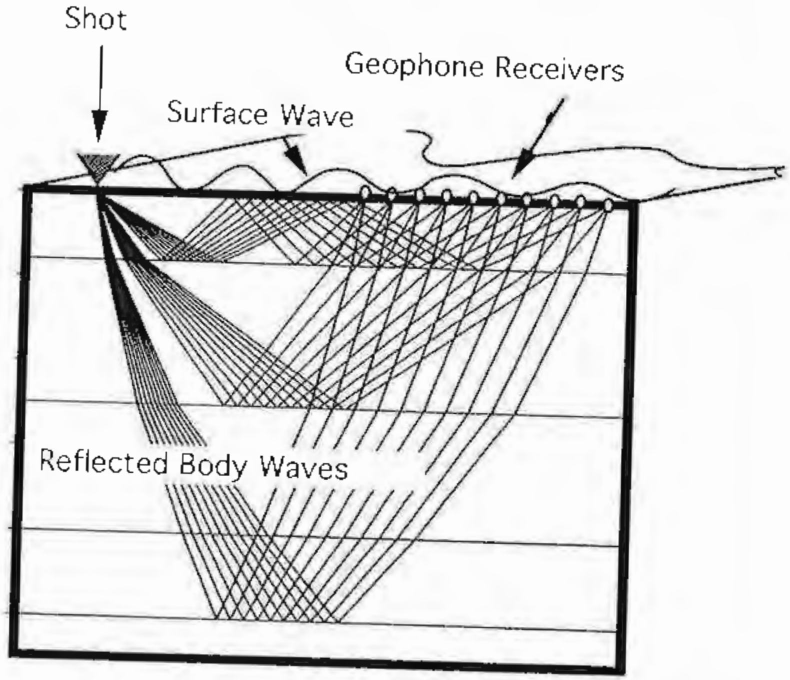

FIGURE 2 Seismic Exploration Data Collection. Seismic energy radiates from an explosion or hammer blow and is then reflected and refracted in the underlying rock. In this example the reflected energy is recorded by ten geophones (sensitive velocity gauges) on the ground surface. In addition to the body wave energy traveling through the rock, surface wave energy travels along the ground (much like ocean waves) and is recorded.

Each observatory records a large number of earthquakes each day. Seismic analysts perform a winnowing of the events, throwing out events that are uninteresting and saving others for more intense scrutiny.[3] It is assumed that critical events will be recognized. Since analysts always work against time deadlines (tomorrow's data is going to arrive in 24 hours), in practice whether a particular event gets interpreted depends on how much other activity there is and whether the event looks interesting during the first few minutes (or seconds) of study.

[3] Most stations have some degree of automated analysis, sometimes to the point of locating all events. However, at some point an analyst will be required to review the data.

TABLE 1 Characteristics of Exploration Seismic Signals.

Sample Rates Bandwidths	250 to 4000 Hz Usually less than 3 octaves 10–100 Hz common 20–500 Hz high resolution 20–2000 Hz special studies
Dataset Size	12,000 to several billion samples .2 seconds to 5 seconds/record Usually < 4000 samples/record Multiple shots per reflection point
Interpretation	Every peak is significant and waveform details (width or phase) are often also interpreted
Dynamic Range of Recording	8 bits to 24 bits are commonly used

TABLE 2 Characteristics of Planetary Seismic Signals.

Sample Rates	1 to 120 Hz Multiple sample rate streams are common
Bandwidth	1-hour period to 40 Hz (17 octaves)
Data Set Size	Continuous recordings 3–100 Megabytes/day Triggered recordings < 1 Megabytes/day
Interpretation	Precise timing and identification of phase (P, S, etc.) is critical Much interpretation is based on the envelope and spectra of the event
Dynamic Range of Recording	24-bit recording common Most signals easily represented by 16 bits

Seismograms also reveal characteristics about the energy source (typically earthquakes or explosions, but also footsteps, hammer blows, or vibrating machinery). One such application, which is heavily funded, is nuclear test-ban treaty verification. Explosions have particular signatures on seismograms that analysts may be trained to recognize. Even small nuclear explosions may be detected at distant seismic stations (most test nuclear explosions in Nevada are easily detectable in Texas). Being able to recognize the difference between a recording of an earthquake and that of an explosion is the *explosion discrimination problem* of seismology.

2.3 SIGNAL CHARACTERISTICS

Signals recorded for exploration seismology are short and have a limited dynamic range and bandwidth (see Table 1) compared to most sounds. They are more like a recording of a hand clap or gunshot made in a canyon and later analyzed for echoes. The source impulse is short and the interesting portion of the recording is the echo. Equipment limitations set the dynamic range of the recorded signal. Field equipment trades-off fidelity and dynamic range against channel capacity. Geophysicists adjust the recording levels and filters to capture a particular signal of interest and sometimes distort other noise and signals.

Signals recorded for planetary seismology are continuous with wide dynamic range and bandwidth compared to most sounds (see Table 2). The recording problem is like that of attempting to record all sounds that happen in a room, from cannon shots to pin drops, without knowing when such sounds may occur. When they arrive from great distances, earthquake signals become drawn out for many minutes (like distant thunder). Field equipment is designed for broad bandwidth and extreme dynamic ranges.

3. PREVIOUS WORK

Seismic data audification is not new. Many geophysicists with the opportunity and equipment have listened to seismograms. Early recordings made on FM tape recorders were often played at high speed to locate recorded earthquakes. From discussions with experienced seismologists of the 1960s, this seems to have been quite common, although it is not well documented. Time-compressed earthquake recordings appeared on an early Cook Laboratory sound-effects record. Other time-compressed recordings were used to shift seismic recordings up to a frequency band so that speech sonograph equipment could be used to do spectral analysis. Recently, a Mexico City earthquake has been audified by Joe Dellinger at the Hawaii Institute of Geophysics and offered to the Internet.

Even though seismic audification is well known as a curiosity, it has never been pursued broadly as a routine alternative to graphic presentation of all forms of seismic data. Extensive literature searches revealed only two published papers on seismic audification. Both investigated the use of audification to discriminate between earthquakes and explosions. The initial pilot study was by Speeth[8] and a more complete study was conducted by Frantti and Leverault.[5]

Speeth observes that listeners can recognize speakers' voices during telephone conversations, despite wide variances in telephone circuit characteristics, tone, and quality. He hypothesized that if it were possible to discriminate between earthquakes and nuclear explosions from recordings close to the source, discrimination at longer distances and along travel paths through varying geology would be like recognizing two different speakers through differing telephone lines. Speeth's results using a data set of only five explosions and eight earthquakes demonstrated that nine observers with 30 to 60 training events achieved 90% correct recognition. This seemed promising. Unfortunately, all the earthquakes were recorded at one seismic station and all the explosions at another. Also all the earthquakes were farther away than the explosions. This strongly suggests the discrimination results could be artifacts of the data set selection.

Frantti repeated the work in 1965 with 21 subjects and 1500 earthquake-explosion questions. Subjects correctly identified the difference between earthquakes and explosions about 65% of the time. There was a wide variance in individual performance. Two observers averaged about 70%, while one continued to improve with training, never showing the plateau that others had. It seems some people may be able to use their experience and background to develop their understanding of new audio signals. Although a 65% accuracy is poor compared to the performance of trained analysts working with paper seismograms, because Frantti's subjects had little training in geophysics, this may not be a fair comparison.

Several interesting leads and conclusions resulted from the research. The best discrimination was found to be at a time compression of 128, which yielded a P-to-S[4] separation of about 1/4 second on the playback. Listeners found that the events had to be repeated four times in order to understand the sound. During some experiments, when recordings of vertical motion were played back on one channel and recordings of horizontal on

[4] P and S are two phases which travel at slightly different velocities. Their separation in time is a measure of the distance to the explosion or earthquake. Their amplitude ratio can be a measure of how much the event is like an explosion or earthquake.

the other, the discrimination seemed to be better.[5] Discrimination was as good for recordings played forward as for those played backward. Playing earthquakes backward sounds unnatural (something like a piano recording played backward) but observers did no worse than with a more familiar forward recording, suggesting that previous knowledge about natural sounds were not used in discrimination decisions.

Frantti suggested the results could have been improved had listeners been limited to records from a single recording station and been able to compare the unknown record against a catalog of known events. A trained analyst reading paper records in effect does this, but the comparison is with remembered events.

Audifying seismic data in 1965 was cumbersome. The tape playback system had to be built. Once the tapes were mastered, listeners had little control over data playback. In contrast, seismic analysts have a great deal of control over how long and at what scales a particular event is displayed before making a conclusion. Quite frequently they also have a suite of reference displays.

Neither study used trained seismologists as subjects nor compared the success of untrained individuals to untrained individuals doing more classical analysis. Training a seismic analyst is a long process, taking up to a year at a single station. Analysts are generally best when they deal with a few stations that they know well rather than a large collection of stations. The current understanding of auditory seismogram display is so immature that it is difficult to design meaningful statistical tests using human subjects. Instead, it may be more advantageous to extend the understanding of auditory analysis by developing and documenting the successful use of such tools by trained analysts in an operational setting.

Frantti only published the one report. There has been very little published work since. At this point the field looks unexplored.

4. APPLICATIONS OF AUDITORY DISPLAY

For some common seismological tasks, audified seismic data should have distinctive advantages over the traditional plotting of wiggly lines on paper based on the nature of the task itself. Some of these tasks are mundane parts of seismology, but a demonstration that audified seismograms are a superior display will push research into more esoteric subjects.

[5] One classical discriminate for earthquakes and explosions is the ratio of the amplitude of the first arrival (which is on the vertical only) to the later shear arrivals, which also are on the horizontals.

4.1 TRAINING

Some people may learn to identify particular seismic events more quickly by listening to the signal rather than by seeing a wiggly trace. Others may learn better by using a combined graphic and auditory illustration. Just as we learn the alphabet with the ABC song, the audio presentatiou may no longer be necessary once the concepts and images are firmly in place. Since it takes up to a year to train an analyst, an improved educational tool represents a significant savings of money.

4.2 QUALITY CONTROL

For exploratory seismic field recording, audified data could efficiently deliver information when the operator is visually occupied. During field data recording an observer will normally be busy directing crew activity and scheduling shots as well as doing on-the-fly trouble shooting and maintenance. Under such conditions, careful field quality control is sometimes difficult. Particularly insidious are problems that occur slowly or intermittently. Scanning visual displays takes time and concentration away from the primary goal - collecting data. It is useful under such conditions to be able to quality control the data in the background. An audio presentation can alert the operator to problems that may then be addressed with other presentations.

For planetary seismology, quality control includes choosing good sites for seismic observatories and monitoring the background noise levels at existing observatories. This is currently done by comparing spectrograms at noisy and quiet sites. Auditory displays, with their natural ability to summarize the spectral information as a "coloring," provide a quick easy method for monitoring noise levels at new and existing stations. They are particularly well suited for monitoring noise generated by resonant structures, such as telephone lines, buildings, or pipelines.

4.3 FREE OSCILLATION DATA

One interesting application of audio data is for free oscillation data. The earth rings when struck, like any solid object such as a bell. The characteristics of the oscillations depend on where and how the earthquake occurred, the material properties of the surrounding rock, and the location of the seismic recorder. It is not yet clear what could be illustrated in an audio display but, in view of the analogs to bell-like objects, the possibilities are intriguing.

4.4 DATA DISCOVERY

Data discovery includes the processes of examining seismic waveforms for information of interest and forming the initial hypothesis about the causes of the disturbances. Such hypothesis will be tested and explored further with additional specialized processing and displays. This initial look forms the jumping-off point for imaginative new observations and ideas. By enriching the current displays with audio in a simple direct manner with few processing biases, there is less possibility that interesting information will be overlooked.

4.5 LARGE DATA SET DISPLAY

Large data set display is a method of interacting with a large data set to extract the information of interest. Seismic data sets are large, often more than a billion samples for a single experiment. They must be reduced and summarized or excerpted for interpretation. This reduction and selection, besides being time consuming, means that significant features in the data may be missed. Auditory displays summarize the information in a different manner than common plots, allowing a second chance to recognize relationships in the seismic data.

4.6 EVENT RECOGNITION

Event recognition and classification were the problem classically attacked in the 1960s by Speeth and Frantti. There are several motivations to this attack. People seem to recognize events and voices from limited data. It is also an important seismological problem with good funding. Since classifying an event requires a mental comparison with remembered models, providing a larger set of symbols based on both visual and aural cues should result in better classifications. Being able to simultaneously present both auditory and graphic displays may also decrease the time required to make a correct classification.

A long dispersed surface wave train buried in noise is one of the more difficult events to identify on a time-series display. The energy is so spread out over time that the identification is often difficult without using spectrograms or phase-matched filters (if it is even known that the event is present). Such sequences are easily heard and recognized on the audio seismogram.

During seismic storms, a single small area will, in a short time, produce a large number of similar small earthquakes and aftershocks mixed with the normal background noise. Recognizing that an event comes from a seismic storm instead of another location is important during test periods.

One of the most successful audio application may be for identifying noise and equipment problems. Equipment and noise sources sound distinctly different from seismic sources. They tend to be periodic and persistent. When trying to locate these sources, several parameters may need to be adjusted. Being able to listen to the output allows the changes to be made quickly, without having to study plots and frequency diagrams.

4.7 EDUCATION

Education is potentially a very rewarding application. The explanation of how seismic energy travels through the earth can be enhanced with examples and explanations that draw on everyday listening experience, such as echoes or the sounds that objects make when hit. The intuitive grasp of physical acoustics that we accumulate in our life can be carried over to certain aspects of seismology.

4.8 MODEL MATCHING

Much of seismology is devoted to determining reasonable earth structures. Once an initial hypothesis is formed, numerical models are developed to test and further improve the hypothesis. Various methods of "fitting" data to models are available, but it is often more appropriate to build a guided model based on interactive input. The model is iterated and displayed versus the real data while the user makes certain adjustments to model parameters. Model and real data may be summarized with a goodness of fit, but may also be display overlaid. It is natural in such a case to also display the audification in side-by-side comparisons.

4.9 SIGNAL DETECTION

Signal detection is the identification of segments of continuous recordings that may have seismic earthquake signals on them. These segments are intensively studied to identify the onset time used for event location.

The first problem related to signal detection is the recognition of small quiet signals buried in various types of noise. The converse problem is recognizing when noise segments contain no signals. Automatic seismic signal-detection algorithms can exceed the performance of a trained seismic analyst when background noise conditions are stable. Under varying noise conditions, analysts often do better. People are experienced in picking out specific sounds in a noise background, so it is natural to expect that the signal detection problem is well posed for audio presentations. Since the automatic detector performance exceeds that of human analysts, it suggests the traditional analysts' display could be improved.

Related to the detection of signals is the recognition of multiple signals once a segment has been detected. In some cases several signals, each from a different source, may overlie each other. In other cases a signal may be complex, created by an aftershock sequence (all from the same source). Signals from different sources usually have strong differences in spectral character, so recognizing an overlying signal should be easy on the auditory display which emphasizes spectral differences.

4.10 ONSET TIMING

Once a segment is identified as containing a signal, the event's onset time (the time of the first arrival of energy) must be determined. For impulsive signals, it is quite easy, but the signal emerges slowly out of a background noise for signals whose source is emergent, or for sources at particular distances. Determining the correct onset tends to be a matter of experience and training related to projecting the rise of the signal back into the noise sequence. A combined audio and visual presentation may help, since small changes in the spectral character of the noise may indicate the first arrivals of the signal. Once an analyst is trained, a combined visual and audio presentation might decrease the review time.

5. PROBLEMS IN AUDIFYING SEISMOGRAMS

Ideally, processing techniques should preserve the physics of the input data while making all interesting features audible. This is required if the final display is to take advantage of our native ability and experience in interpreting natural sound. To be audible the sound must fit within human frequency and amplitude limits for comfortable listening (which is usually wider than the commonly available reproduction equipment).

It is obvious from its characteristics (see Table 1) that exploration seismic data requires processing to make events audible. These characteristics present the following problems when audifying raw seismic data:

- **The information is nearly subsonic.**
  Some recorded data includes audio frequencies, but the dominant energy for most reflections is centered near 30 Hz where it is difficult to distinguish small changes in pitch and timbre.
- **The bandwidth is narrow compared to uatural sounds.**
  Seismologists go to great lengths to record broadband signals. Good definition over three octaves is considered excellent. This is far short of the ten-octave range of natural sounds.
- **Both the direct wave and reflected signals are short compared to natural speech and most environmental sounds.**
  Most reflection recording is less than 3 seconds with much of the interest in the first second or two.
- **Classical seismic interpretation (where each waveform peak is interpreted) is nearly impossible in a straight audification.**
  In reflection seismic interpretation, coherent peaks on adjacent traces are interpreted as reflections of a rock layer. The times of these peaks and the interval to the next set of peaks are carefully measured and mapped. This measurement of each wiggle may be more amenable to sonification or visual display.
- **The dynamic range of some signals is extremely large (>100 dB) over a relatively short period.**
  This problem has to be handled by graphical displays too. One method used for plots is to clip waveforms when they exceed a particular amplitude.

Earthquake seismograms (see Table 2) present their own set of problems:

- **The signals are nearly all far infrasonic.**
  Typical teleseisms (distant signals) have their dominant energy below one 1 Hz and only the closest earthquakes and blasts have energy above 10 Hz.
- **Earthquake signals span more than 17 octaves.**
  This presents problems not only for auditory displays, but for almost any other form of raw data display. Researchers have found it difficult to record or deal with displays that include more than a few octaves of signal. Seismographs capable of recording broadband signals have recently become available, but interpretation and display techniques still rely on presenting different filtered bands.
- **The dataset size can be huge.**
  Because stations record continuously for years, the dataset size grows. If instrumentation has not changed, selected waveforms from many years of recording may be appropriate to a particular interpretation problem. Indexing, retrieving, and displaying these selected segments is a problem in itself.
- **The dynamic range of interesting earthquake signals is large.**
  Today, 24-bit recording systems with a dynamic range > 140 dB are common. Fortunately, in many cases 16 bits will adequately represent the waveform.

6. PROCESSING TECHNIQUES

A number of technologies are available to make seismic audifications. Some of them specifically address the problems discussed in Section 5 above. Other techniques are designed to make the audification easier to use.

6.1 PREPROCESSING

Preprocessing consists of adapting the raw seismic field data into a form that can be output to an audio interface. Most data requires a simple amplitude rescaling and DC removal to adjust the range of the recorded data to the range of the output D-to-A converters. Usually high-quality interpolation is also needed to produce an acceptable sample rate without distorting the input waveform.

Tr 80

6.2 TIME-COMPRESSION

Some shallow exploration reflection data has enough high frequency to be audible if played at recorded speed (Track 80, 0:02.7-0:29.9). Most earthquake records are not directly audible since the significant energy is all below 10 Hz. To be heard the seismograms must be processed to move the dominant energy into the audible bands. This process must preserve any of the relations to acoustic physics. For example, a low pianolike sound must remain pianolike when shifted in frequency. It must preserve the characteristic attack and decay of impulsive sounds as well as harmonic relationships. Two tones separated by an octave must maintain an octave separation shifted up in frequency.

Time-compression is playing data faster than it was recorded. It is a simple direct method of making low frequencies audible and it preserves the physics (the result still follows the wave equation). A data stream recorded at 40 samples/second and played at 8000 samples/second has a 200-times compression. This allows 3 hours of recording to be played back in less than a minute, comparable to the amount of time to quickly study 3 hours of plotted data.

6.3 AUTOMATIC GAIN CONTROL (AGC)

Some type of amplitude scaling has to be done to bring the signals in the range of the D-to-A. Many seismic signals have a dynamic range in excess of comfortable listening levels, so signals may be processed through an AGC prior to playback. AGC increases the volume during quiet periods and decreases it for large signals analogous to automatic level controls on

recording equipment. Several different forms of gain control were tried. The most successful method was to divide each sample by the average absolute value of a small window of surrounding samples:

$$(S'_k) = s_k \sum_{i=k-n}^{k+n} |s_i|,$$

where n is half the window size.

Window lengths are typically 200 samples or about 1/40 second for most traces. This gain adjustment is rapid compared to most audio recordings. It changes the nature of the recorded data so that differences in spectral properties of signals are more easily discriminated.

It was also necessary to preprocess traces with AGC before doing frequency doubling. Otherwise the dominant signal swamped all other subtle arrivals.

6.4 FREQUENCY DOUBLING

While time-compressing data does move events from the subaudible to the audible, it brings with it another set of problems. Signals from some earthquakes may only last a minute on the original recording. Speeding up the recording by 200 times will result in a sound that only lasts a fraction of a second, so short that it is difficult to study. It may be recognizable as a single sound but it cannot be taken apart into its components.

Some early experimenters avoided the problem by using the minimum compression to make the seismogram audible. The resultant sounds were dominated by very low frequencies. Others shifted the frequencies into the audio range by multiplying the signal by a carrier to produce an FM signal. Unfortunately, this is an additive shift in frequencies. Thus, if 20 Hz is shifted to 440 Hz, then 40 Hz will be shifted to 480 Hz. This does not preserve the harmonic relationships. It also produces a section that is narrower bandwidth than the original (in terms of octaves) and a sound that is quickly fatiguing.

Close study requires a way to slow the playback without changing the pitch and the harmonic relationships. The operation of frequency doubling shifts any pure tone up one octave. The audio track may then be played at half speed to get the original pitch in a signal with twice the duration. The technique works well for pure sine waves, but complex waveforms have some unexpected results. Fortunately most seismic waveforms are simple enough that their character is preserved.

The development of this method may be illustrated with elementary trigonometry. Assume a pure cosine and sine signal $x = \cos(t)$, $y = \sin(t)$. Using double-angle relations

$$\cos(2t) = \cos^2 t - \sin^2 t,$$
$$\sin(2t) = 2\sin(t)\cos(t),$$

the doubled signals could be produced as

$$x_{\text{doubled}} = x^2 - y^2,$$
$$y_{\text{doubled}} = 2xy.$$

For an arbitrary signal s, we may generate the complex analytic signal,[9] S, whose real component is the original seismogram, and whose imaginary component is the Hilbert transform of s. If s were a pure sine wave, $H(s)$ would be a pure cosine wave. The frequency-doubled analytic signal S is then $S = s^2 - H(s)^2 + i2sH(s)$, where $H(s)$ is the Hilbert transform of s. In a like manner the envelope of a seismic signal is computed as

$$e(s) = \sqrt{s^2 + H(s)^2}.$$

Experience shows that for impulsive signals, the process of frequency doubling makes the signals more bursty, since it, in effect, squares the envelope of the signal. In order to reduce this effect on the frequency-doubled signal, the new envelope is divided out and replaced with the old.

$$S_{\text{corrected}} = \frac{S}{e(S)} e(s).$$

This preserves the amplitude decay of the original signal.

For pure sine and cosine waves these expressions are exact. For other signals, the results are not always as expected. For example, speech signals, doubled and played back at half speed which one might expect to sound simply like a slow talker, sound instead like a badly clipped tape played back at half speed. For the simpler signals in seismology the technique is adequate. Frequency-doubled and half-speed displays may be easily related to the normal displays up to an expansion of about four times. Beyond this, the signals all decay into something that sounds like a large metal sheet being shaken.

One problem with frequency doubling is that it accentuates the dominant energy and suppresses more subtle arrivals. Informal experiments indicated that it improves the display if digital automatic gain control is applied prior to frequency doubling.

6.5 ANNOTATION AND MARKERS

Just as a good graphic needs annotation such as labels and tick marks, so does a good auditory display. In bimodal presentation the graphic and audio must have some common markers to indicate the relation of *time* on the audio track to *position* on the graphic. The obvious solution, the animated cursor on a workstation display, is difficult to mentally synchronize with the correct corresponding sounds. There appears to be a high correlation when random signals are displayed on the video screen, an effect well known to animators and movie arrangers.

Therefore, a system of aural tick marks and annotation was overlaid. This turned out to be more difficult than expected. When the annotation has a markedly different character than the data, it separates perceptually from the data stream and is no longer synchronized. It becomes difficult to tell if an event occurs before or after a particular aural cursor.

Undamped bell tones were used in the initial experiments. They were replaced with drumlike sounds generated by passing an impulse through a narrow bandwidth, first-order Butterworth filter. Families of marker tones were produced by varying the bandwidth while keeping the low-cut frequency constant.

For sections rendered in stereo with multiple markers, each set of markers was panned to a slightly different location in the stereo field. This helped separate the markers from each other and from the data.

6.6 LOOPING

Each selection may be repeatedly played in a short loop. Three or more iterations may be required to understand the sound. Although the individual tracks are about a minute long, it is the selections of only a few seconds within the track that should be looped for study. For some recordings, using an equalizer to boost or cut different frequency ranges while the loop is playing will help extract particular sounds.[6]

6.7 INTERACTIVITY

Initial investigations of seismic audification used a custom interactive program with the ability to snip out pieces of a seismogram for processing and auditory display. Unfortunately, too much interactive flexibility is a problem when working with many sample data sets. It was unusual that an optimum setting for one seismogram could be used on the next. Instead it

[6]The reader might try programming their CD player to create this looping with the examples provided on the accompanying CD. It was avoided on the CD to better utilize the available playing time

was more valuable to play back a wide variety of seismograms with a single processing setup and pick out the obvious similarities and differences. This kept the processing relatively constant through training, providing the listener with a stable auditory environment for cross-sample comparisons.

6.8 STEREO PLACEMENT

Simple stereo processing, wherein a separate signal is played into each ear, was marginally successful. There seems to be more interpretable information in the sounds. In my own experience, repeated listening continues to uncover additional signals and relationships longer on stereo signals than on monaural signals.

Some signals were played back in stereo to convey the direction of a seismic source. The result was only partially successful. While the actual input signals were similar to naturally occurring acoustic phenomena, the selection of two signals for left and right ears was quite arbitrary and not analogous to the stereo effect heard in the environment. More advanced processing or a better choice of inputs may be required for usefnl stereo.

For other signals, a second audio channel was used to provide an additional annotation. Signals were panned from left to right to represent the seismograms recorded from different sensors on the ground. In this case, stereo position was created by simple amplitude differences. The result is partially successful, but better results might be had with additional stereo cues.[2]

When listening to the examples, there will be a difference in the interpretability through headphones and speakers for the stereo selections. Listening to each channel separately several times will help separate the sounds.

7. THE EXAMPLES

Seven audio examples on CD tracks 78–84 illustrate the problems and rewards of the techniques of seismic data audification.

7.1 MODEL DATA AS A REFERENCE

Model traces are synthetic seismograms created from an arbitrary earth model (see Figure 3) of layers of rock with varying properties (velocity, density, and attenuation). This model assumes a single explosion buried just below the surface. Thirty-nine vertical geophones evenly spaced along a line

from the source position record the seismograms resulting from reflections and refractions of the source explosion.

Models are an appealing starting example because all parameters may be completely controlled, to the extent the assumptions in the modeling process themselves are correct. A model developed from an interpretation of field data may be used to validate the interpretation. If the synthetic seismograms from the model match the field data, then the interpretation could be correct. If not, then model parameters may be adjusted until the resulting synthetic seismograms match the field data. In other cases, results from two different models may be compared to determine if two modeling techniques yield similar answers, or if two different models create recognizable differences in the seismogram. Models also form reasonable

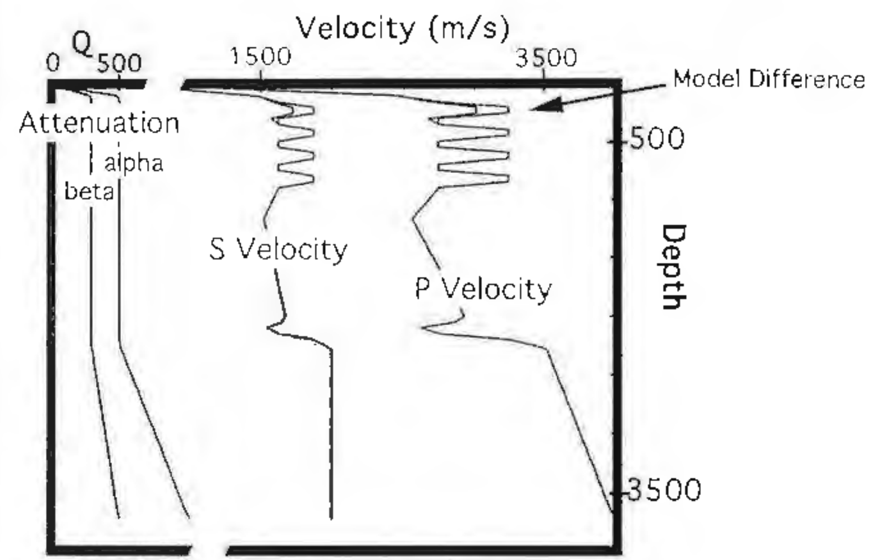

FIGURE 3 Synthetic Seismic Model Description. A complex velocity model was used to describe a possible layer cake geology. Attenuation models (Q) are shown on the far left. The geologic section was assumed to be a loose, dry, low-velocity material at the surface, underlain by four thin interbedded fast and slow layers, a low-velocity zone, and several zones with increasing velocity. The model was not representative of any particular geology, but was selected to represent a section with realistic complexity.

benchmarks for audification. The audification process may be tuned with perfect model data before attempting it with field data.

The first selection (Track 78, 0:01.6-0:07.0) contains synthetic data traces played back at 48 times compression to produce recordings from eight separate geophones every second. The dominant feature is the energy fade away as seismic traces at longer distances are played. Notice that when seismograms of varying distances are plotted at the same scale (see Figure 4), traces near the source are clipped while those at long distances are not visible. Because the audification has a wider dynamic range, it

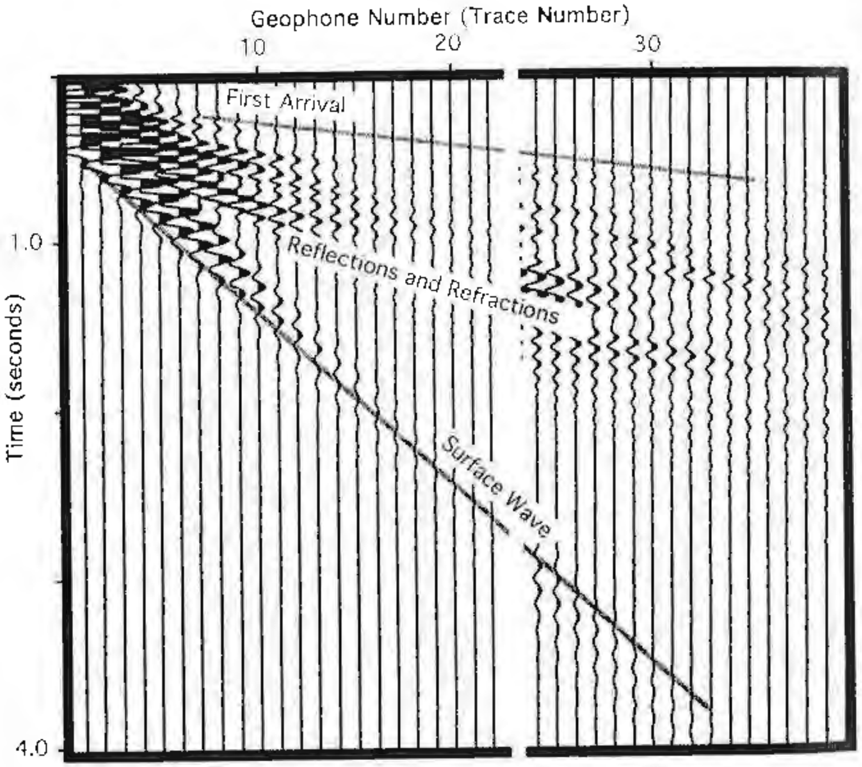

FIGURE 4 Synthetic Seismic Model Section. This synthetic seismic section is created from the velocity model in Figure 3. Each vertical trace is the velocity time series recorded from a single geophone on the ground. Deflections to the right are upward motion. Traces beyond 22 have been plotted with a larger gain to make the deflections apparent.

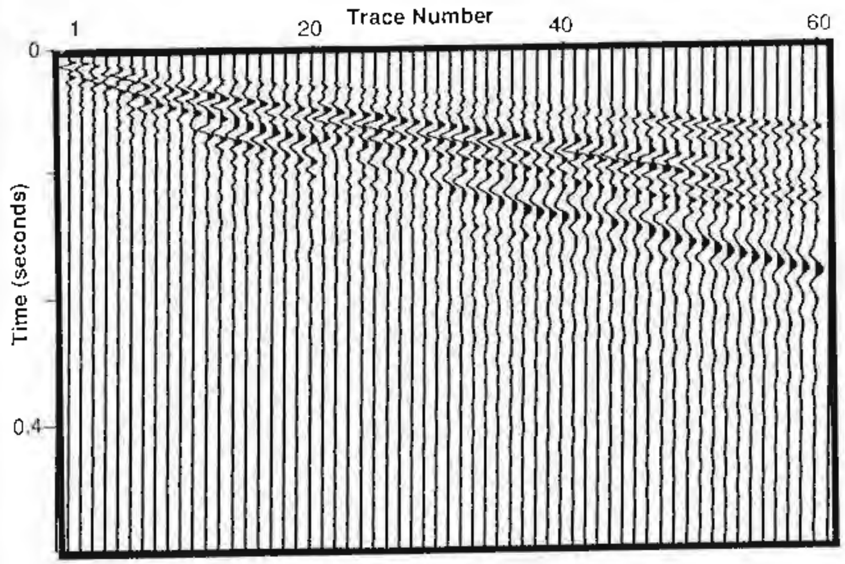

FIGURE 5 Shotgun Seismic Section. This shallow seismic section records the response from a shotgun blast into the earth near trace 1. The recording was done in three groups of 20 traces each. All traces were recorded with the same gain and are plotted with a common scale factor.

presents the closest traces without clipping, while leaving the distant traces audible. Typically, seismograms are normalized to the energy in each trace to remove the effect of energy decay with distance (sometimes called a spherical spreading correction). This may also be done in the auditory domain (Track 78, 0:08.1–0:13.8) so that time and distance changes in the recorded seismic signal produce clear timbral changes in the audification.

The seismograms are automatic-gain-controlled and frequency-doubled to examine the traces in more detail (Track 78, 0:14.9–0:18.6), then played back at half speed (Track 78, 0:19.4–0:25.8) for a time-compression of 24 times. The process may be repeated (Trace 78, 0:28.0–0:34.5) to arrive at a compression of 12 times.

Comparing the seismograms resulting from two different models is usually done by overlaying plots of the two in different colors. A naive method for audio is to play each model into one ear (Track 78, 0:36.1–0:41.8); this is a dichotic presentation. It does not take full advantage of our built-in stereo processing and is more analogous to displaying slightly differing plots on each eye. Differences in the two signals cause an apparent stereo shift from left to right. This is most pronounced at the distant traces.

Imbedded in the signals is a low-frequency surface wave[7] that travels more slowly (see Figure 4) than the other body waves. A second set of drumlike tones has been added to point out this later arrival (Track 78, 0:45.0–1:05.8). Each annotation tone immediately precedes the surface wave and each trace is also marked at the start. The annotation tone is positioned in a separate stereo location by panning it towards one channel away from the other tones and data. This example emphasizes the utility of audio markers to point out particular features.

Tr 79

7.2 SHALLOW SEISMIC FIELD DATA

These same techniques are valid for field data (Track 79) collected as a part of an investigation of the shallow subsurface. A number of seismic recordings were collected along a line of increasing distances from the source, just like the model traces (see Figure 5).

Field recording equipment can only record a few channels simultaneously (20 channels for Track 79, 0:3.5–0:11.5) and it is, therefore, necessary to repeat the explosions after the geophones are moved to build up the complete section. The assumption is that the source explosions are identical and that the geophones are identical so that each trace should sound similar to adjacent traces and the section (Track 79, 0:3.5–0:11.5) should be similar to the model (Track 78). This first example of a shotgun fired into the ground[4] is scaled like the first model section. The differences every 20 traces are a result of slight differences in the three shots fired to create the 60 traces. There are also small differences from trace to trace indicating minor differences in the way individual geophones are attached to the ground.

The second section (Track 79, 0:12.3–0:22.1; see Figure 6) is a composite section formed from six hammer blows recorded into twelve geophones on each blow. Every sixth trace comes from the same hammer blow. Thus traces 1, 7, 13, and so on are from hammer blow 1, traces 2, 8, 14, and so on are from hammer blow 2, and traces 3, 9, 15, and so on are from hammer blow 3. Sharp changes in the sound from trace to trace or regular patterns are artifacts of the field procedure and indicate that the prior assumptions are not completely valid. Extreme variations and prominent patterns indicate inadequate field quality control. Rapid repeating (looping) of each trace emphasizes sharp changes from trace to trace

[7]Surface waves are confined to the uppermost layers of earth and concentrate energy near the upper boundary, much like familiar waves in the ocean. Body waves are those waves that travel through the material, much like familiar sound waves.

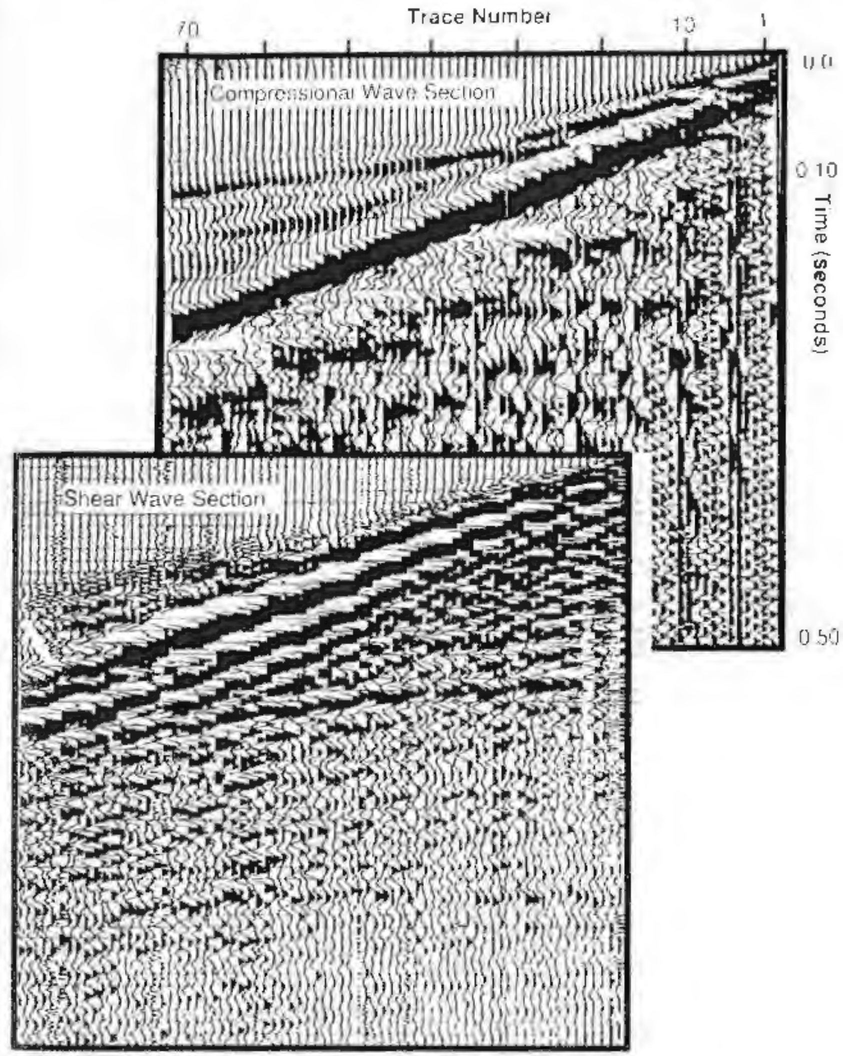

FIGURE 6 P and S section from Caldwell Ranch. At Caldwell Ranch 72 channels were recorded in 6 groups of 12. Traces are numbered from right to left. Data was acquired with 10 Hz geophones for the compressional section and 50 Hz geophones for the shear section. Channel gains were optimized for reflected energy expected at 0.20 seconds. Earlier arrivals were clipped. This section was processed with AGC before being plotted. On the compressional section several of the first few traces are ringy.

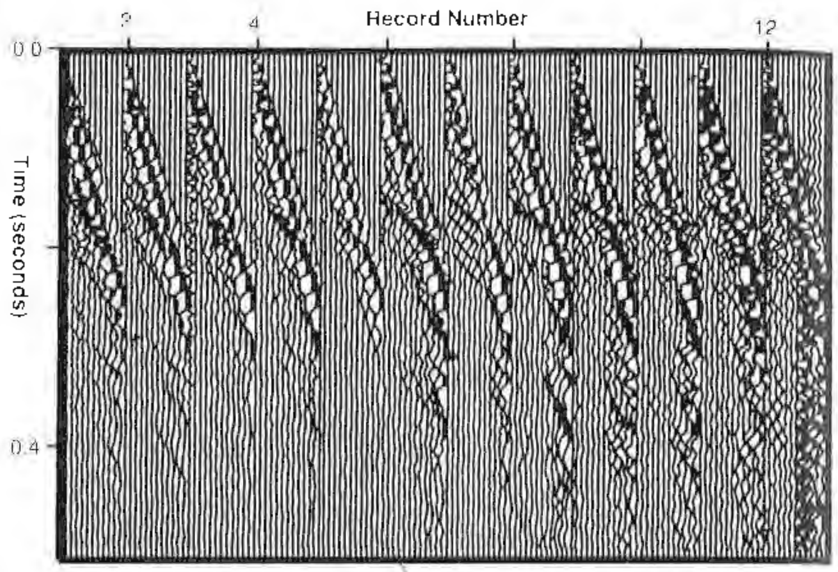

FIGURE 7 Twelve Consecutive Shot Records. Each set of 12 traces was acquired with a different hammer blow. Channel gains were adjusted in the field to try to keep the recorded energy constant. With perfect acquisition and flat-lying bedrock, each record should be identical to the adjacent ones. With this particular equipment, the field operator can view only one record (12 traces) at a time.

(Track 79, 0:22.7–0:37.7). Applying AGC to the section (Track 79, 0:38.8–0:48.6—left channel) better emphasizes some of the weaker signals. In this case, a second set of shear wave seismograms is used to measure depth to the water table. Interpreting these joint experiments requires recognizing arrivals common to the two sections and arrivals that appear to be similar but show slightly different velocities and arrival times on the two sections. On plots these two may be overlaid in color. For audio, each is played into one ear (Track 79, 0:38.8–0:48.6).

7.3 SHALLOW SEISMIC REFLECTIONS

Common depth point recording (CDP) is a standard exploration method that records a hammer blow into a line of evenly spaced geophones (in these examples, 12 geophones). The hammer source point and the geophones are shifted a measured distance in-line and a second recording made. This is repeated hundreds of times until the line of geophones may be moved several miles from its initial position. Again sudden changes are indications of field problems. It is likely the seismic observer was unable to find time to do careful quality control. In this case the audio channel forms another important method to warn of problems.

A set of four sample records illustrate some field problems that may be encountered (Track 80, 0:02.7–0:29.9). This high-frequency seismic section is unusual in that much of it is audible with no time compression or special processing. These records illustrate significant quality control problems not detected on the visual field display during the field collection phase (field crew of seven people, newly trained operator, sledge hammer source, four to six shots per minute). Tr 80

A set of 12 consecutive records (Track 80, 0:32.0–0:56.8; see Figure 7), played at 8 times compression, illustrates the case of quality control. Traces are played in stereo with a simple amplitude pan relative to their ground position. Ideally the listener would hear the sound originate at the corresponding geophone position. The audio evidence of field problems could even be radioed to each crew member, allowing each individual to monitor his contribution to recording quality. For small (two- to four-person) operations it may even allow elimination of one crew member.

7.4 NOISE AT EARTHQUAKE SEISMIC STATIONS

For seismic observatories recording earthquakes, a key quality control issue is finding quiet locations. Seismically noisy locations are generally poor candidates for recording small distant earthquakes. Typically a noise survey is made at several candidate sites by recording background noise for days or weeks. These recordings (see Figure 8) are then analyzed to produce frequency spectra (see Figure 9) of the background noise.

An audio selection of background noise from four different stations (Track 81, 0:01.8–0:18.0) demonstrates that each station has a distinctive character. Because this data is dominated by 1/20 to 5 Hz frequencies, it requires much higher time-compression than exploration data– 200 times in this case. The distinctive sound suggests that site noise conditions may be recognized by listening for a few minutes to a compressed recording. Tr 81

Following the four noise examples are two extended selections (Track 81, 0:19.4–1:05.4) from two of the sites represented in the short recording. These two stations are immediately recognizable from the prior 5-second samples. The stations have distinctly different and recognizable noise spectra, although this is not immediately apparent on the time-series (see Figures 10 and 11). Some individuals may learn to recognize characteristic sounds quicker than plots of characteristic spectra. Tr 81

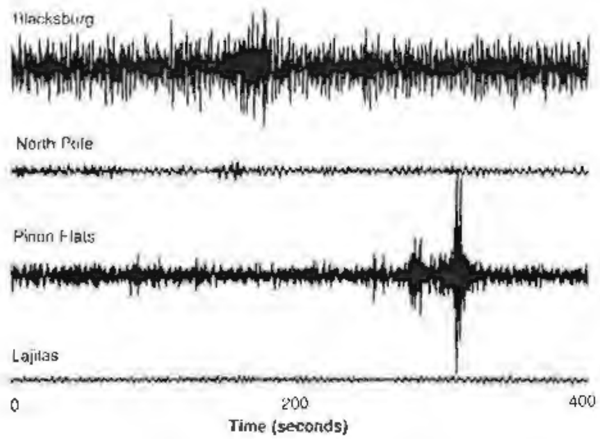

FIGURE 8 Background Noise at Four Seismic Observatories. All seismograms were plotted with a common scale factor. Each seismogram is of the vertical GS13 recorded at 40 samples/second into a 24-bit digitizer. Segments were selected from random times and may not be representative of long-term noise conditions.

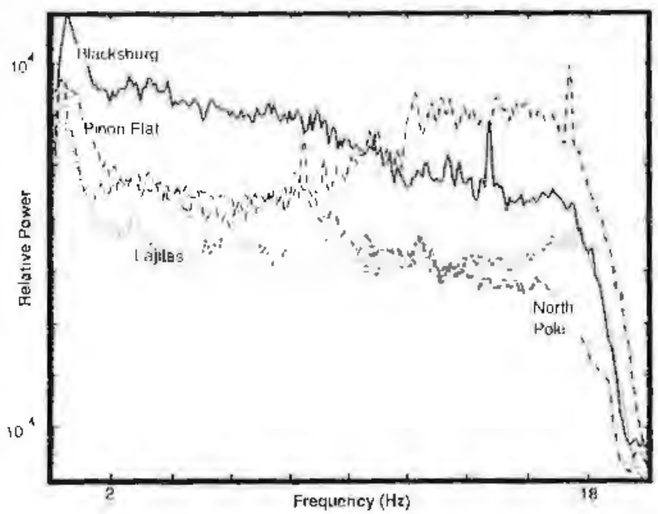

FIGURE 9 Background Noise Spectra at Four Seismic Observatories. Four hundred seconds of data from each of four stations was used to calculate the average noise spectra. Strong peaks at Blacksburg and Piñon Flats are indications of resonant man-made structures.

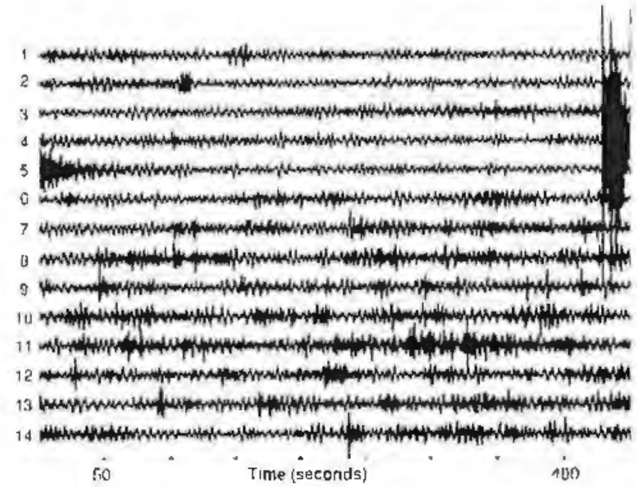

FIGURE 10 Seismogram from North Pole, Alaska Near Fairbanks. Each horizontal segment is 450 seconds of data and is contiguous with the segment below. A small earthquake is recorded on segments 4 and 5. A microseismic storm or other source of noise occurs from segments 8 through 14.

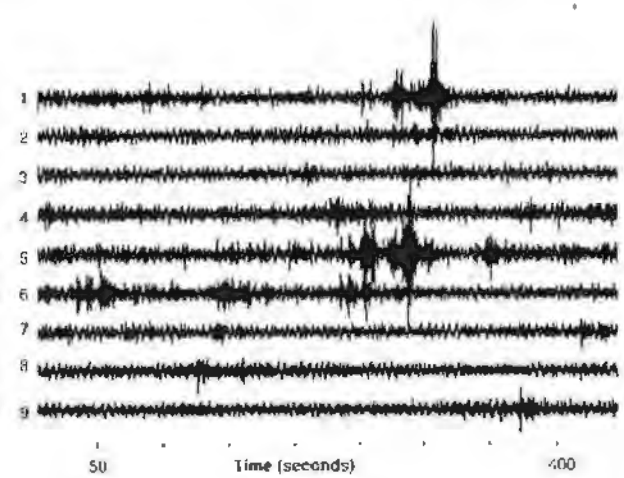

FIGURE 11 Seismogram from Piñon Flats, California. Some of the bursts on the seismogram are small earthquakes although the seismogram is also contaminated with noise.

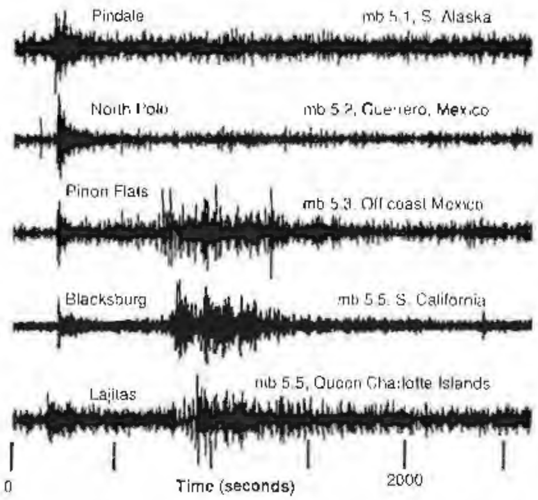

FIGURE 12 Seismograms from Earthquakes 27 degrees away from the seismic observatory. The initial impulse is the compressional wave arrival. The second package of energy (seen best on the Blacksburg station) is the surface wave arrivals. Each seismogram was normalized with a separate gain factor. The separation of the compressional wave arrival and surface wave arrival is nearly constant for a given distance.

Tr 82

7.5 SELECTED EARTHQUAKES

Recordings of selected earthquakes, all about the same distance (27 degrees) from the recording station, illustrate the characteristic features of audio earthquake seismograms. Five different earthquakes recorded at five different stations (Track 82, 0:03.7–0:20.5; see Figure 12) have been time-compressed 800 times. On plots the primary control of the appearance of a particular recording is the distance from the event. At most distances an audio seismogram of an earthquake has three parts: two sharp sounds (reminiscent of a heavy door with a latch closing) followed much later by a drawn out, low-frequency surface wave that usually shifts from higher frequencies.

A selection of earthquakes recorded at from 12 to 101 degrees (Track 82, 0:23.3–0:52.8; see Figure 13) illustrates the progression as the station is farther and farther away. At long distances part of the energy is subaudible, even at 800 times compression. Doubling the compression to 1600 times (Track 82, 0:54.0–1:05.7) demonstrates that all three parts are present even at far distances. These low-frequency surface waves, while audible, may not always be easily visible on a plot without special filtering. Because relative distances may be recognized on either audio or graphic seismograms, the display choice may be optimized for the operator preference.

Tr 83

7.6 NUCLEAR EXPLOSIONS

Two past nuclear explosions at White Sands, New Mexico, recorded in Lajitas, Texas, are played back at 200 times compression (Track 83, 0:01.4–0:13.5; see Figure 1) . These relatively close events (13 degrees) do not have a strong surface wave which is typical of explosions. They also reverberate compared to previous recordings. The reverberant structure is an indication of an unusual geology along the travel path from the test site to Lajitas.

Stereo recordings (Track 83, 0:15.0–0:21.7) were created by playing vertical ground velocity into the center channel, north-south velocity into the left channel, and east-west velocity into the right channel. Shear waves and surface waves have strong horizontal components, and they should demonstrate a stereo preference. The apparent "roominess" of the sound was unexpected and is yet unexplained. Sounds within the explosion seem to shift from channel to channel in a complex way; the patterns are more

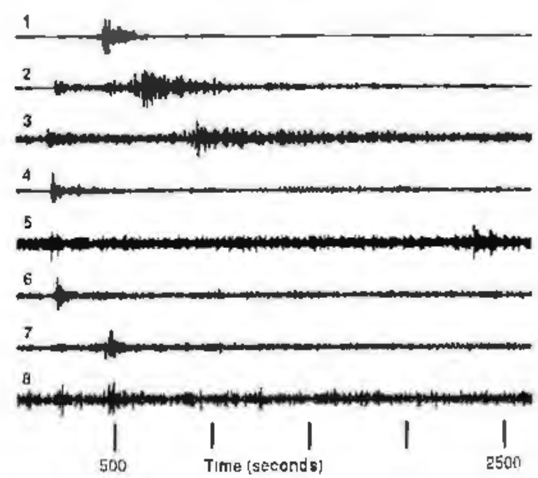

FIGURE 13 Eight Earthquakes from 12 to 101 degrees away. Eight earthquakes illustrate the progressive separation of the initial and later arrivals with increasing distance. The low-frequency dispersed surface wave is easily visible on seismograms 4 and 7 (for a listing of the events and exact distances, see the CD annotations).

apparent when the explosions are frequency doubled and slowed down (Track 83, 0:21.7–0:25.7).

7.7 ARRAY DATA

A particularly rich application for stereo presentation is recorded array data. Seismic arrays usually include 10–30 seismometers spaced in a regular pattern, such as concentric rings. Recordings of such arrays may be beamformed during processing to enhance signals in particular directions or to estimate the direction to the origin of a particularly interesting earthquake. If an audio seismogram from each element of an array were played back simultaneously, a situation similar to a noisy cocktail party or choral choir would result. The goal is to be able to pick out individual voices to find those elements that are unusually quiet or noisy. Another possibility is to beamform the array to enhance particular directions and then position those signals correctly in the stereo field. Ideally, the output of a head motion sensor attached to the display user could be used to control the beamform direction. Turn your head left and signals to the west would be enhanced.

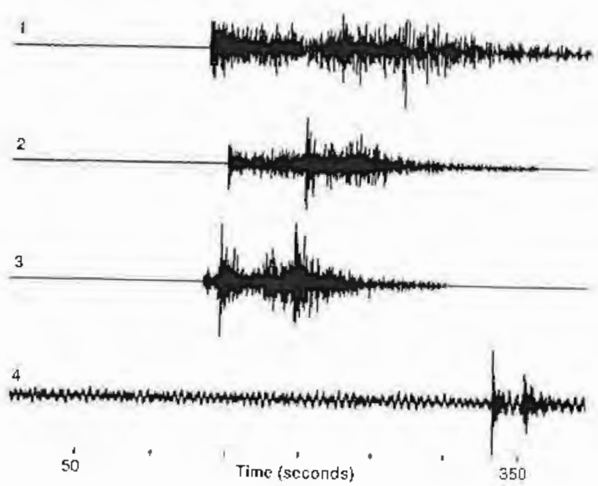

FIGURE 14 Center Element for a Stereo Experiment. The four displayed seismograms are for one of the vertical channels used in forming the stereo audification. The events are listed in the CD annotation. Each event is plotted with an independent scale factor.

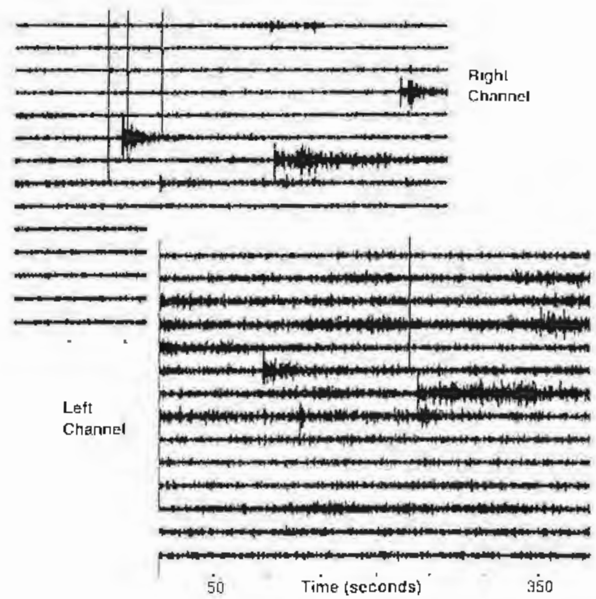

FIGURE 15 Two Channels from the 100-Km Array Audification. The right channel contains a number of data glitches (three are obvious on the plot). The left channel contains one. Seismic events on the two channels arrive at slightly different times. The event on segment 4 of the right channel has no corresponding signal on the left channel. This was either a very local signal, or a signal that is obscured by high-noise conditions on the left channel.

For example, on audio track 84, stereo imaging was employed as follows: Three elements, spaced in a triangle 2 kilometers on a side, were selected. One was played in the center channel, one on the left channel, and one on the right channel (Track 84, 0:01.5–0:30.1; see Figure 14). The stereo separation is prominent for all these events although it is difficult to associate a preferred direction for a particular earthquake. A single earthquake will include sounds with different delays between the channels because different wave types travel at different velocities. This is not something that is common to acoustic events.

The first three events are all dominated by low frequencies. In acoustic signals, low frequencies are difficult to locate and, therefore, we have less success in accurately locating these signals. The final earthquake is a high-frequency signal. The stereo audio clearly suggests two different directions although both sounds originate from the same signal.

The final example (Track 84, 0:33.3–1:09.9; see Figure 15) is an audification of a three-element array with each element separated by 100 kilometers. The energy arrives with significant delays between stations. The separation is great enough that a waveform recorded at one station may look significantly different from those at the other two. The immediately obvious characteristics are the distinctive noise patterns related to each channel. Channel noise conditions appear unrelated to each other. Glitches and pops occur at unrelated times on the left and right channel, but never the center channel. The earthquake at the end of the recording appears to occur almost simultaneously on all channels.

8. CONCLUSIONS

These examples of audification suggest its value for doing diagnostic quality control in field reflection seismology and in earthquake seismology. A method for moving the frequencies of seismic waves from subaudible to audible has been demonstrated. Repeated listening to a short waveform helps to better understand the signal and to increase the effective change in character from sound to sound. Simple processing yields the most natural interpretable sounds. Stereo presentations seem to include more extractable information and have potential, although the simplified method used here is inadequate for most signals.

Annotation is particularly important. Good annotation allows the listener to match events on a graphic display with a correlated audio. In multiple annotations, tones must be separated from each other by slight frequency differences or positions in stereo. Annotation tones must be similar to the signal or the listener will perceive two unrelated streams and it will be difficult to relate the timing between the two. Good annotation avoids the problem of finding terminology to describe sounds to a reader. Examples may be played with audible markers just prior to the sound of interest in a manner analogous to adding arrows to a graph.

The wiggle plot results in a poor representation of frequency information, but the auditory display gives a good idea of the overall characteristics and evolution of the spectra. Just as it is possible to process and plot the seismogram to extract spectral information and plot the spectrogram, the seismogram data could be used to control a sound generator (i.e., sonified) to present each wiggle. However, this approach does not fully utilize the natural advantage of each style of presentation. Obvious events (those with high signal-to-noise ratios) are clear in both the auditory and graphic displays. More subtle events may be initially recognized in only one domain, but may then be processed to enhance them in the other domain.

The natural advantages of each seem to encourage the use of graphics to display time domain variations and the use of audio for frequency domain information.

9. FUTURE EXPERIMENTS

A number of experiments in several subject areas seem worthwhile. A natural extension of Speeth's earlier work would be to bury known seismic signals in background noise and then to compare audio and visual identification. Of particular interest, however, is the real-world performance of analysts using their standard visual analysis tools with and without audio.

A second rich area for exploration is the use of stereo auditory displays. Single stations could include three-component motion detectors so that processing information from these detectors would determine direction. Even more exciting would be the use of stereo ,pmotpromg with array data, when head position could be used to beamform the output of the array. Alternatively, the raw output of the array could be processed to set up the station locations in spatial locations. The use of stereo might allow identification of compression and shear arrivals. Shear energy will predominate the horizontal channels that could be placed in a distinct location in audio space.

Combined audio and visual displays could be useful in the routine analysis of seismic data. If an analyst is given audio tools, the question is "under what conditions are they useful and when are they distracting?" If the tools are to some extent "adjustable," what sorts of adjustments are most useful to the analyst?

The use of audio suggests new possibilities for seismic attribute displays in exploration data, including direct hydrocarbon indicator displays of various types, audio attribute displays, and amplitude vs. offset displays. Of particular interest are displays for modeling and quick look, when the audio tag may indicate whether the visual display is useful.

The audio seismogram is still in its infancy. The tools to produce it are currently cumbersome and not part of the daily tool set available to seismic analysts. The real proof of the utility of the tool will come only when auditory display is a natural part of seismic analysis. This will require making audio tools accessible to working seismologists in a natural well-designed interface.

ACKNOWLEDGMENTS

The author gratefully acknowledges the unending patience and encouragement of Gregory Kramer who would not let this work die. Without his continued efforts and substantive input this paper would still be an uncompleted collection of notes and outlines. Thank you, Greg. The initial material development was a direct result of the First International Conference on Auditory Display held at the Santa Fe Institute. The comments and suggestions from participants helped define the problems and method of attack.

REFERENCES

1. Aki, K., and P. G. Richards. *Quantitative Seismology: Theory and Methods.* W. H. Freeman, 1980.
2. Begault, Durand R., and Elizabeth Wenzel. "Techniques and Applications for Binaural Sound Manipulation in Human-Machine Interfaces." *Intl. J. Aviation Psych.* **2(1)** (1992): 1–22.
3. Bullen, K. E., and Bruce A. Bolt. *An Introduction to the Theory of Seismology,* 4th ed. Cambridge: Press Syndicate of the University of Cambridge, 1985.
4. Clark, R. D. "Gunning for Data." *The Leading Edge* **3(11)** (1985): 20.
5. Frantti, G. E., and L. A. Leverault. "Auditory Discrimination of Seismic Signals from Earthquakes and Explosions." *Bull. Seismol. Soc. Am.* **55(1)** (1965): 1–26.
6. Hilterman, Fred J. "Three-Dimensional Seismic Modeling." *Geophysics* **35(6)** (1970): 1020–1037.
7. SMU Staff. "Development of an Intelligent Seismic Facility and Preparation for Participation in the Conference on Disarmament Group of Scientific Experts Technical Tests." Semi-Annual Technical Report, Southern Methodist University, Defense Advanced Research Projects Agency Contract #MDA 972-88-K-0001, 1991.
8. Speeth, Sheridan Dauster. "Seismometer Sounds." *J. Acous. Soc. Am.* **33(7)** (1961): 909–916.
9. Taner, M. T., F. Koehler, and R. Sheriff. "Complex Seismic Trace Analysis." *Geophysics* **44(6)** (1979): 1041–1063.

TRACKS 78–84

Chris Hayward

"Listening to the Earth Sing"

All audio examples other than Track 78, the synthetic model seismograms, are actual seismic field recordings made to determine the depth and nature of the bedrock or are recordings from well-known seismic observatories.

Raw data was previewed on a Sun SPARCstation 2 and interesting examples were audified. Sample audio segments were reviewed using the built-in Sun 8000-Hz single-channel μlaw output device. Forty-five minutes of sound was selected for scrutiny and critical listening. These interesting segments were transferred to an SGI Indigo Extreme where they were converted from the original seismic data to 44.1-KHz 16-bit stereo AIFF files. The individual examples were edited to meet the time available on the CD and then annotated to produce the DAT used for mastering.

Track 78: Model Traces. A computer program simulates the recordings from a shallow buried explosion in a multi-layered earth. Thirty-nine separate seismometers record the seismic activity at increasing distances from the explosion. When created each seismogram is about six second long is sampled at 166 Hz and contains frequencies as low as 5 Hz.

Track 78: 0:01.6–0:07.0. Thirty-nine simulated seismograms are played 48 times faster than recorded (time-compressed 48 times) resulting in 8 audio seismograms played each second. Each fifth and tenth seismogram begins with a short wood-block tone.

The energy fades away as the seismometers get further from the source. This decay in amplitude so dominates the recording that it is difficult to compare other aspects of the changing nature of the seismograms. Although the power in each seismogram varies by over 600,000, almost all seismograms are audible without doing special variable scaling as would be necessary for plotted data.

Track 78: 0:08.1–0:13.8. The model is repeated with the seismograms individually scaled to remove the effects of energy decay with distance. Each six-second seismogram has been divided by the maximum value in the seismogram.

The gradual shift in dominant in tone from high to low as seismograms move farther from the explosion illustrate the loss in high frequencies with increasing distance from the source. Later seismograms have an echo. This echo occurs at the more distant seismograms where energy traveling at two different velocities becomes well separated in time. On seismogram 30 it is also possible to distinguish the delay between the annotation tone and the initial arrival of the seismic impulse.

Track 78: 0:14.9–0:18.6. The previous normalized seismograms are further modified with a fast automatic gain control (AGC) that adjusts the amplitude each 1/40 second to keep the energy constant even within the individual seismogram. Then the seismograms are frequency-doubled. Frequency doubling preserves the harmonic relationships between fundamentals and harmonics but shifts more of the energy into higher frequencies where small differences are more easily distinguished.

Only the first 20 seismograms are played as an example. The seismograms have a distinctly springy sound. It is sometimes difficult to distinguish the start of individual seismograms.

Track 78: 0:19.4–0:25.8. Frequency-doubled seismograms when played at half speed have the same frequency spectrum as the original version. Annotation tones have been added to the start of each seismogram in addition to the tones at every fifth and tenth seismogram. The tones are overlaid with a slightly different stereo balance than the seismic data to help separate them.

Track 78: 0:28.0–0:34.5. Frequency doubling followed by half-speed playback is repeated to yield audio seismograms with 12-times compression.

Each seismogram is a sliding tone which moves quickly from high to low frequency. The first few seismograms have a constant superimposed sharp metallic tone. This may be an artifact of the modeling program. Later seismograms have a more complex extended shift in frequency. This shift from high to low is a consequence of the rapid attenuation of high frequencies in earth materials. Energy returning from deeper reflections is of lower frequency than that which only travels a short distance through the shallow layers.

Track 78: 0:36.1–0:41.8. One channel is the previous frequency-doubled model traces played at 48-times compression. The other channel is the result of a slightly different model. It takes careful comparison of the plots of the two models to distinguish differences between them.

In the audio display, the differences are not apparent on the first few traces, but become progressively more distinct at later times, finally becoming a bouncing from channel to channel. While it is possible to tell that the two models are different, it is difficult to describe the differences.

Track 78: 0:45.0–1:05.8. The model frequency-doubled twice and played at 12-times compression. There are annotation tones at the start of each seismogram as well as every fifth and tenth seismogram. In addition a tone is added just before the low-frequency surface wave. The annotation tones are separated from each other with slightly different frequencies and balance.

In this case not only do the traces have every fifth and tenth trace marked, but there is also a mark on each trace at the expected time of the surface wave. With the annotation, the surface wave is clearly distinguishable as a short low-frequency tone that becomes progressively delayed from the start of the seismogram. Although the multiple annotation tones are at first listening complex, on repeated listening they are separated sufficiently that surface wave annotation may be distinguished from tones at the start of the seismogram.

Track 79: Shallow Seismic Exploration. These examples use field seismic records acquired with geometry's similar to the model. Each seismogram is at an increasing distance from the explosion or hammer blow. Often in the field, equipment limitations require that a recording of one shot into many geophones or seismometers be simulated with multiple shots into the available geophones. In these examples 20 or 12 geophones are recorded at a time. If all the shots or geophones are not identical, artifacts in the recorded data may obscure information used to interpret geology. In some cases, these artifacts may be misinterpreted as geologic structures.

Recordings with many geophones, such as these, are made during the initial investigation of an area to determine basic seismic and noise characteristics region prior to starting a more detailed seismic investigation.

Track 79: 0:03.5–0:11.5. This seismic section was acquired with 20 RefTek 16-bit A/D recorders sampling at 250 Hz. All gains were set identically to avoid clipping even the strong near signal. The source was a Betsy Seisgun, an industrial shotgun that points down into the ground. Geophones are separated by 1 meter. Sixty records or traces in three sets have been recorded in increasing distances.

Traces are played back at 16-times compression or 8 traces per second. Every fifth, tenth, and twentieth trace is prefixed with an annotation tone. A single scaling factor was used for all traces.

Listen for the overall change in the tone after each twentieth seismogram and the general change in frequency and amplitude with distance. The abrupt changes in tone every twentieth trace are a result of slight differences in the characteristics of the shotgun source from shot to shot. Traces 21 through 40 have a higher background noise (it sounds like traffic noise) relative to the other two shots. This is probably wind noise, and if recognized in the field may have been avoided.

While every fifth, tenth, and twentieth traces are marked at their start, it is difficult to match the marker with the exact trace. It is also difficult to recognize the progressive delay between the annotation marker and the dominant energy although it is evident that the markers and arrivals do not exactly coincide.

Track 79: 0:12.3–0:22.1. This section of 72 traces was recorded with six different sledge hammer blows into twelve geophones. The data was recorded with an EG&G twelve-channel 10-bit recorder sampling at 2000 Hz. The limited dynamic range of the recorder made it necessary to individually adjust the gain of each channel. Gains were adjusted to best record the seismic reflections and, therefore, the strong arrivals are clipped.

The section is played back at a compression of four, or about eight traces per second. Each geophone is played with the six consecutive hammer blows, then the next geophone is played. Every fifth, tenth, and twentieth trace are marked. No other processing has been applied.

The dependence on geophone coupling is evident in the sharp changes from each group of six, while the shot dependence appears as a repeated six-beat pattern. The first three shots have a pronounced ring evident on the first three beats of the first three measures of six (traces 1–18). At more distant offsets from the source the ringiness is no longer distinguishable because most of the high frequencies have been attenuated. A difference in background noise that sounds like something walking through tall grass or heavey-breathing is associated with some shots and is similar to that heard in the previous Betsy Seisgun section. Recognizing the changes from sensor to sensor is more difficult. It is most obvious at the far offsets where there is a distinct change in tone every six beats.

Track 79: 0:22.7–0:37.7. Rapidly repeating each individual trace three or more times emphasizes the differences between the individual traces as well as allowing the traces to be studied more closely. Each trace is annotated once at the start as well as every fifth and tenth trace. The ringiness of the first four shots is clearly audible in the first 18 traces although at longer distances these high frequencies are nearly gone. However, at the farther offset a difference in background noise (sounds a little like walking through tall grass or breathing) associated with some shots similar to that heard in the previous Betsy Shotgun section. Listening for the changes from sensor to sensor is more difficult. It is most obvious at the far offsets where there is a distinct change in tone every six beats.

Track 79: 0:38.8–0:48.6. The same section is taken, but normalized with automatic gain control and played through the left channel. The right channel has the corresponding shear wave section (horizontal hammer blow). In this case annotation has been deleted to avoid complicating the selection.

The breathing sound on the right channel corresponds to different noise conditions at the time of each shot. Theoretically, the direct arrival of the shear waves (right channel) should occur after the compression wave, but the compression Rayleigh (left channel) surface wave should occur after the shear Love surface wave (right channel).

Track 79: 0:51.3–1:03.7. The same compression and shear sections are slowed down by 4 and frequency shifted to maintain the section in the audible range. Strong clipping of the compression section (left channel) is obvious (it was clipped in the field). The initial arrival occurs at the same time on each trace, while the later surface waves generally arrive first on the right channel.

Track 80: Shallow Seismic Reflections. Once the initial investigations are complete, the next step is often to profile an area to build up a geologic cross section. This involves recording a number of records as the line is scooted along the profile. In this case each set of 12 traces represents one shot. Data is recorded from 14-Hz horizontal geophones sampled at 2000 Hz with a 10-bit EG&G recorder.

Individual traces are placed in the stereo field to represent their position along the ground.

Track 80: 0:02.7–0:29.9. In this unusual high-frequency seismic section, four records (1, 4, 5, 12) selected from a larger set are audible without frequency shifting or time compression. Each record begins with a marker tone.

All records should sound identical under ideal conditions. Instead the third record is noticeably weaker, and the fourth is stronger but has obvious 60-Hz noise pickup on several traces. The second and third records also contain a small chirp caused by grass rubbing against the geophone. These significant problems were not recognized in the field during recording but are immediately identifiable here.

Track 80: 0:32.0–0:56.8. The set of all 12 consecutive records is played back at 8-times speed. Each record begins with marker tone and traces are panned across the channels to represent their ground placement. At this increased speed, the previous two chirps will be sharp clicks. There are also at least 8 noise spikes, 4 traces with high 60 Hz noise, several records with differences in the shot, and several interesting background tonal noises (on records 3 and 7).

Track 80: 0:58.4–1:08.9. Frequency-doubling the first 5 records of the preceding brings out the background noise and suppresses the dominant thump of the initial arrival. In this manner the traces may be examined to verify that all background white noise is similar.

The dominant frequency is rather high initially and then falls off towards trace 12 on each record. In Record 1 a difference is shown suggesting that there may have been a problem with the first couple of channels. The differences from shot to shot are most obvious when the first few traces of each shot are compared. The second record has generally lower frequency.

Track 81: Seismic Stations. Examples of seismic stations used for earthquake recording have been extracted from the US UN GSE seismic network. A number of high-quality 24-bit ADC sampling at 40 and 10 Hz have been used to collect the data. The wide dynamic range here ensures these stations do not clip, even during large events. For the purposes of audification, most records must be normalized to make the noise audible. Data is played back at 8000/40-times compression, allowing 24 hours of data to be scanned in 7.2 minutes, faster than many stations can scroll the data on a CRT display.

Track 81: 0:01.8–0:18.0. Samples of background noise for four seismic observatories—Blacksburg, Virginia; North Pole, Alaska (near Fairbanks); Piñon Flats, California; and Lajitas, Texas—are played back at 200-times compression. These were all observatories used in a three-month test of a world-wide system for monitoring seismic signals to detect nuclear tests.

All segments have been normalized to the same scale. Only 5 seconds of audio is sufficient to identify each of the stations.

Blacksburg has a strong low-frequency component audible as a low rumbling of 100 Hz and less. Superimposed on this noise are small surface wave (Lg) arrivals (sounding like failed attempts to whistle) from distant mining explosions. A strong tone at 15 Hz is shifted to 3000 Hz and forms an annoying high frequency whistle. It is helpful to vary the settings on a graphic equalizer while looping this sequence.

North Pole has an overall quieter background. A small frequency peak at 12 Hz on the spectra is shifted to 2400 Hz and in the audio display is a set of two repeated whistlelike tones. These are probably a natural harmonic of a manmade structure, such as power poles vibrated into resonance by wind. From the audio section it is obvious that these tones are not present during the whole time but change. They are quite distinctive and would be immediately recognizable. The quiet staticlike sound is a part of the natural noise background and may be distant earth tremors related to volcanic activity.

Piñon Flats has a number of tonals that sound like continuous high-pitched whistles on the recording. These are likely a result of one or more manmade structures. From the recording it is evident that this is more of a problem at Piñon Flats than Alaska. There is also a distant surface wave arrival similar to that recorded at Blacksburg.

Lajitas is almost silent in contrast to Piñon Flats. At the very end of the Lajitas selection the careful listener will hear a small surface wave, again like that of Blacksburg. Lajitas shows none of the tonals present on other stations.

Track 81: 0:19.4–0:43.0. An extended record of 100 minutes of recorded data from North Pole (labeled as Fairbanks on the recording) includes a near local earthquake and two more distant earthquakes. The near earthquake reverberates, like a gunshot in an empty gymnasium. Three smaller distant earthquakes may be recognized by their similarity to a heavy door latching as it closes. These are recognizable underneath the large number of crackling and staticlike sounds that may be a microseismic storm.

Track 81: 0:45.7–1:05.4. The distinctly different record from Piñon Flats is difficult to listen to because of the continuous high-pitched tonals but listening to the selection through an equalizer with the highs above 3000-Hz cut helps. There are three regional earthquakes (each less than 1500 km away), each sounding like the double thump of a heavy door latching and closing. Times between the two thumps are related to the distance so these three events can be ranked by distance during three or four listening.

Along with the impulses are two surface (Lg) waves. These begin with low frequencies and generally end with high frequencies. They are reminiscent of what might be used to simulate wind blowing across the desert or a roadrunner dashing off across the desert. One local event is heard as a sharp pop or click, almost like a production problem.

Track 82: Selected Earthquakes. Recorded earthquake signals vary depending on the size, location, distance, depth, and mechanism of the quake. The geology between the quake and observatory and the recording instrument response also influences the appearance of the record.

Five earthquakes recorded from 27 degrees[6] away at five different stations are played back at 800-times compression. The traces were originally recorded at 10 Hz with 24-bit digitizers.

The recordings are:

Track 82: 0:03.7–0:06.1. 5.1 mb on March 9, 1990 in Southern Alaska recorded in Pinedale, Wyoming (31 degrees).

Track 82: 0:06.1–0:09.9. 5.2 mb on March 13,1990 in Guerrero, Mexico recorded near Fairbanks, Alaska (27 degrees).

Track 82: 0:09.9–0:13.2. 5.3 mb on July 10,1990 off Coast of Mexico recorded at Piñon Flats, California (27 degrees).

Track 82: 0:13.2–0:16.8. 5.5 mb on February 28, 1990 in Southern California recorded at Blacksburg, Virginia (30 degrees).

Track 82: 0:16.8–0:20.5. 5.5 mb on February 3, 1990 in Queen Charlotte Islands recorded at Lajitas, Texas (29 degrees).

[6]A degree is approximately 111 km.

Listen for the separation of the initial compression and shear arrival (it sounds a little like a heavy door closing) and the separation between the initial arrival and the surface wave (a low-frequency boing). A careful listener will recognize the stations at Blacksburg, Piñon Flats, and Fairbanks from the characteristic noise. While these events are all similar distances, there is still a wide variety depending on the exact station and event combination.

Track 82: 0:23.3–0:52.8. Eight earthquakes from 12 degrees to 101 degrees away are played back at 800-times compression. All events are preceded by an annotation tone. All events were recorded at Latjias, Texas. See Table 2.

The additional compression is required to make the low-frequency surface wave at 1/20 Hz audible. At 101 degrees the low-frequency surface wave is nearly inaudible. Three distinct arrivals are audible for each record, although at the near distances the first two occur close together. Some stations may be identified by the noise underneath the event.

Track 82: 0:54.0–1:05.7. Three of the preceding earthquakes, the first one at 12 degrees and the last two near 100 degrees, are compressed by 1600 to emphasize the low frequencies. The long duration and dispersion are indicators of extreme distance to the earthquake. These surface waves were not associated with the original even during the first stage of visual analysis. In fact, the surface waves were not immediately apparent on the plots. With audio displays as with video, the wrong form of display can result in missing important information.

TABLE 2 Eight earthquakes, recorded at Latjias, Texas, at 800-times compression; see Track 82: 0:23.3–0:52.8.

Num	Time	Date	Distance (degrees)	Magnitude (mb)
1	23:45:46.6	May 11, 1990	12.31	5.2
2	00:56:34.0	Jul. 10, 1990	20.72	5.3
3	10:01:02.0	Feb. 3, 1990	29.35	5.1
4	15:23:14.0	Aug. 14, 1990	56.65	5.8
5	18:10:17.0	Jul. 4, 1990	61.32	5.1
6	07:34:20.0	May 5, 1990	90.59	5.3
7	03:31:36.0	Jul. 10, 1990	99.45	6.0
8	02:41:48.0	Feb. 17, 1990	101.7	6.1

Track 83: Nuclear Explosions. Seismic data recorded from nuclear explosions have much in common with earthquakes but have a smaller surface wave and shear wave in relation to an equivalent sized earthquake. In this dataset, two U.S. underground nuclear tests were recorded at Lajitas.

Track 83: 0:01.4–0:13.5. Two nuclear explosions (June 13, 1990 and March 10, 1990) of slightly different magnitudes recorded in Lajitas at 40 Hz and played back at 200-times compression.

The reverberations suggest that the travel path from the explosion to Lajitas includes a large number of nearly perfect reflectors, much like a light pipe. The absence of a surface wave is one of the diagnostic features of explosions.

Track 83: 0:15.0–0:21.7. Stereo recordings were created by playing the recorded vertical channel into both channels, the recorded north channel into the left channel and the recorded east channel into the right.

The roominess of the resultant signal is startling. Careful listening will reveal that the sounds making up the seismic signal are placed differently in the stereo space. On repeated listening one can distinguish other features besides the primary event. The background noise is different on the two channels. Near the end of the main signal is a faint signal from an unrelated earthquake.

Track 83: 0:21.7 0:25.7. Time compressing by 400 times was done to listen for longer period surface waves (none are immediately audible). The second event almost masked in the previous recording is clearer. The shift to lower frequencies as the reverberation dies away is obvious and the characteristic two-thump signature is clear.

Track 84: Array Data. The GERESS (GErman Regional Experimental Seismic Station) seismic array includes 25 elements distributed over a 2-Km aperture. Three elements on the outer ring were formed into a stereo image. The first element is played on both channels, the east element into the left channel and the west element into the right channel. The data is played back at 200-times recorded speed. It was expected that the stereo placement of each event should give an indication as to the distance and the direction to the event.

Track 84: 0:01.5–0:03.9. Earthquake 1. Oct. 28, 1991 at 00:31, Yugoslavia 4.0 mb, 773 km distance, 127° azimuth.

Track 84: 0:03.9–0:06.3. Earthquake 2. Oct. 31, 1991 at 09:30, Northern Italy 4.7 mb, 507 km distance, 214° azimuth.

Track 84: 0:06.3–0:08.3. Earthquake 3. Nov. 20, 1991 at 01:53, Switzerland 4.8 mb, 338 km distance, 235° azimuth.

Track 84: 0:08.3–0:10.8. Earthquake 4. Jan. 8, 1992 at 08:48, Austria 1.6 mb, 179 km distance, 149° azimuth.

The same data is frequency-doubled and played with a 100-times compression.

Track 84: 0:13.2–0:17.5. Earthquake 1.

Track 84: 0:17.5–0:21.0. Earthquake 2.

Track 84: 0:20.3–0:25.4. Earthquake 3.

Track 84: 0:25.4–0:30.1. Earthquake 4.

Frequency doubling and expanding are used to increase detail. The quakes are not in order from closest to most distant. Relative distances may be determined by comparing these quakes with those on track 82. The final quake does not have the reverberant nature of the preceding three. The third quake appears to be complex having a double arrival which is duplicated in the later S wave arrival. Background noise on the three channels is different. Portions of each event are stronger on one channel or another. The last quake with its extended high-frequency range is particularly distinctive, appearing to shift from one channel to the other. This should not be the case, as the direction of arrival for both phases should be the same although the velocity will to be different.

Track 84: 0:33.3–1:09.9. Elements of a 100-Km array are recorded at 20 Hz and played back at 400-times compression. The selection represents about 200 minutes of recording. The Lajitas station is played back on both channels, Marathon, Texas (to the east) on the left channel and Fort Shafter, Texas (to the west) on the right channel.

The noise is distinctive, with several clicks and pops that denote data glitches appearing in only the left or right channel. The right channel has the largest problem with glitches. There are a number of events of various sizes, all close to the same distance. None localize very well. There is a large event near the end of the section and during the coda, a number of smaller events. The background noise at the Marathon and Fort Shafter is down just prior to the main event.

Acknowledgments. The earthquake data was collected by the Southern Methodist Geophysics staff as part of Defense Advanced Research Projects Agency contract numers MDA 972-89-C-0054 and MDA 972-88-K0001. Exploration seismic data were collected by Tom Goforth at Baylor University.