lawrence livermore national laboratory

UCRL-53282

Sound and Computer Information Presentation

Sara Bly

(Ph.D. Thesis)

March 1, 1982

Work performed under the auspices of the U.S. Department of Energy by the UCLLNL under contract No. W-7405-ENG-48.

university of california • davis

DISCLAIMER

DISCLAIMER

Work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

UCRL--53282

DE82 015782

UCRL-53282
Distribution Category UC-32

Sound and Computer Information Presentation

Sara Bly

Computing Science Group

University of California, Davis

Davis, California

Manuscript Date: March 1, 1982

LAWRENCE LIVERMORE LABORATORY
University of California • Livermore, California • 94550

Available from: National Technical Information Service • U.S. Department of Commerce
5285 Port Royal Road • Springfield, VA 22161 • $10.00 per copy • (Microfiche $3.50)

ABSTRACT

This thesis examines the use of sound to present data. Computer graphics currently offers a vast array of techniques for communicating data to analysts. Graphics is limited, however, by the number of dimensions that can be perceived at one time, by the types of data that lend themselves to visual representation, and by the necessary eye focus on the output. Sound offers an enhancement and an alternative to graphic tools.

Multivariate, logarithmic, and time-varying data provide examples for aural representation. For each of these three types of data, the thesis suggests a method of encoding the information into sound and presents various applications. Data values were mapped to sound characteristics such as pitch and volume so that information was presented as sets or sequences of notes. In all cases, the resulting sounds conveyed information in a manner consistent with prior knowledge of the data.

Experiments showed that sound does convey information accurately and that sound can enhance graphic presentations. Subjects were tested on their ability to distinguish between two sources of test items. In the first phase of the experiments, subjects discriminated between two 6-dimensional data sets represented in sound. In the second phase of the experiment, 75 subjects were selected and assigned to one of three groups. The first group of 25 heard test items, the second group saw test items, and the third group both heard and saw the test items. The average percentage correct was 64.5% for the sound-only group, 62% for the graphics-only group, and 69% for the sound and graphics group. In the third phase, additional experiments focused on the mapping between data values and sound characteristics and on the training methods.

As the use of computers spreads, the need for methods of conveying information from the computer to humans grows. Computer-generated sound offers an alternative that has not been widely utilized. This thesis suggests areas of future exploration in both applications and techniques for aural data representation. The results of the work described in this thesis and the many questions to be studied indicate the broad range of use for sound.

KEYWORDS: Human/computer interface   Computer graphics   Data analysis techniques   Information presentation   Human factors   Interactive systems   Sound synthesis   Computer sound

ACKNOWLEDGEMENTS

Many people helped me turn the idea of using sound for information presentation into a reality. I could not have done this work without their imagination, knowledge, and support.

My committee, Steve Levine, Charles Wetherell, and Herschel Loomis, were particularly helpful throughout *several* years. I thank them not only for all the ideas, the technical information, and the critical reviews, but also for the continual prodding.

Others also gave generously of their time and expertise. Stanley Grotch provided data, graphical representations, and an encouraging interest in a new technique. His applications gave my ideas special value. Jeff Brandt forced me to write down my thoughts for an experiment. Dick Mensing set up tests for the statistical analysis and patiently explained them to me. He equally kindly showed me that I did not always get the *desired* results. Norm Badler, Jim McGraw, Frank Moses, Joyce Moulden, and Mary Zosel read, commented, and criticized; the final version owes much to their efforts.

I thank those who helped with the implementation of the experiments, recordings, and paper production. Jim Carlson kept the Imsai microcomputer running so that even an earthquake did not cause a setback. Numerous friends participated in the experiments, lending extra support by their willingness and prompt responses. Dick Rufer and Don Patterson helped record the sounds so I did not have to rely entirely on visual explanations of aural effects. Kelly O'Hair and Pete Keller taught me the intricacies of plotting programs and text formatting. Their help was supplemented by the valuable work of John Beatty, Mark Blair, and Hank Moll.

My work was done while I was a member of the Computation Department at Lawrence Livermore National Laboratory. D-Division personnel were especially generous in providing the equipment and time necessary for producing the sounds and running the experiments. Both D-Division and the Engineering Research Division provided environments in which I could continue my work (even if the sounds coming from the lab did seem strange at times).

I am continually grateful to my family who has always supported my goals and to Aunt Ida who saw that my mother got to college years ago. My special thanks goes to George who was there day after day, available for brainstorming, editing, criticizing, talking, cooking -- whatever would keep me going.

Thank you all!

Sara Bly
Livermore, California
March 1, 1982

CONTENTS

INTRODUCTION

With the advent of computers, the ability to perform data calculations has increased tremendously. Much of the data analysis, however, still depends on humans to find patterns and extract information. These analysts are faced with the problems of an ever-increasing number of numeric results and with the desire to comprehend these results easily and quickly. A need exists to understand large amounts of data and to present this computer data to humans in a readily-understood form.

In the past few years, computer graphics has revolutionized the ability to obtain information from computers. Output far more various than printed words and numbers is now available. From bar graphs to three-dimensional color displays, computer graphics involves vision as an active aid in data interpretation. Plots enhance analysis, a computer paintbrush expands art, designer plans emerge almost as they are envisioned, and simulations help explore a range of alternatives.

If visual feedback improves the use of computer calculations, why not use other senses as well? Our perception of the world around us comes from hearing, touching, smelling, and tasting, as well as from seeing. Although experiments have been done with touch [32], most computer research for output other than graphics has been with sound. Sound technology already exists for mass reproduction and playback. Additionally, sound has a well-defined structure that taste, smell, and touch lack. A vocabulary exists to talk about sounds, and the relationships among sounds can be categorized. Sound is a natural choice for further study of ways to convey computer data.

This thesis explores the use of sound output to examine data from computers. It includes the problems of information presentation, the existing methods of sound generation, and the alternatives for using sound. An experiment shows that listeners can classify data values encoded into sound. Chapter 1 outlines the current ways of communicating information from the computer to a human and looks at the associated problems. Chapters 2 and 3 depart from the discussion of presenting information in sound to provide background material. Techniques for computer-synthesized sound are described in Chapter 2. Facilities and desired tools are recorded in Chapter 3. Chapter 4 refocuses on information presentation and suggests a variety of ways to use sound. Chapter 5 presents the results from an experiment in which subjects used sound output for data discrimination. Chapter 6 summarizes the conclusions and outlines further areas of exploration.

A recording provides examples of several methods of using sound to present computer-generated information. Appendix A describes the recording in detail. Just as graphic illustrations help in clarifying text, the recording is most useful in illustrating the material in Chapters 4 and 5. The reader will benefit by listening to the recording while reading.

1: INFORMATION PRESENTATION

Information presentation is important. The more rapidly we can obtain information, the more quickly we can question and expand our knowledge. Techniques for presenting computer-generated information are particularly interesting. A computer is extremely fast and provides quantities of data at a rate far too rapid for human comprehension. However, the computer's strength in providing information does not compensate for its weakness in presenting that information to humans in an easily understandable manner. Tools for computer data presentation are crucial for making adequate use of computer information.

This chapter concentrates on methods of presenting computer information. Computer graphics offers a variety of ways to deal with the flow of information; however, there is data which can not be absorbed by graphics alone. It is appropriate to expand the methods for computer information presentation. Consider the ways in which humans process information, the computer output which supports these skills, the limitations of this information flow, and some alternatives using sound.

1.1: Human Information Processing

In order to increase the understanding of information presentation, it is necessary to understand how humans deal with information. In the work that follows, a few general skills are significant for gaining information from data, particularly from large sets of data. Four such skills are the ability to note specific features in the data, the ability to recognize patterns or groupings in the data, the ability to assimilate overall structure and properties of the data, and the ability to perceive the data in more than one way. This section elaborates on each of these skills.

The most common means of dealing with large amounts of data is to note those events which have particular significance or stand out in some way from the remaining information. A summary or outline of a subject is often sufficient to point out the important topics. An experiment in which only the occurrence of a particularly large value is important may record only a few of the many facts observed. A graph quickly indicates one or two points that lie far away from the others. These features (or errors!) of the data often become the focus of further work.

Many times, however, it is not sufficient to observe only a few facts regarding the data. The significant information may lie in the patterns or groupings of the data. Even though an isolated value is important, the events leading to the noted occurrence are often more enlightening. For example, restricting experiment output to large data values might hide the fact that a particularly low value always preceded a particularly high value. Data usually is classified by noting relationships and similarities which correspond to known information. This categorization can be helpful in simplifying a complex set of information.

When nothing is known in advance about the data, even patterns or groupings are difficult to determine. It is important to be able to comprehend the entire body of data in some way. All the information should be considered before structuring the data into patterns or specific events. A two-dimensional projection of n-dimensional data may hide a natural division of the data in another plane. Being mindful of the overall structure of the data may prevent isolated events or groupings from falsely dominating any conclusions.

Most important is the ability to consider data from a variety of viewpoints. What may appear as ordinary in one context may be quite significant in another. What may have no meaning from one viewpoint may take on useful structure from another. For example, a graph of eight-dimensional data may appear as random points when projected onto the first two axes but may show distinct separation into sets in a different projection. Often data plotted linearly appears scattered but shows definite patterns when plotted logarithmically. Figures 1.1

and 1.2 illustrate two equally important views of one such data set when the x-axis is a linear scale and when the x-axis is a logarithmic scale.[1] Analysis involving several methods of examining data increases the understanding of the data, its properties and structure.

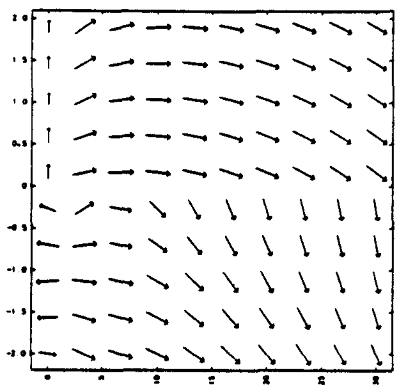

Figure 1.1: Linear Axis

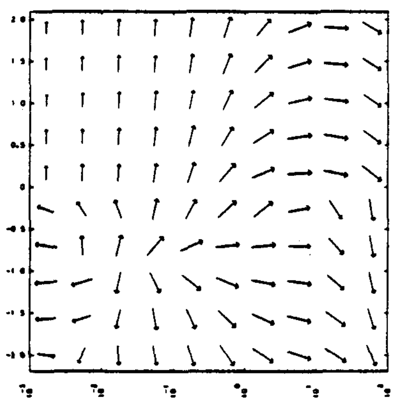

Figure 1.2: Logarithmic Axis

These human skills for analyzing data should be supported by computer methods for presenting the data. Many such computer tools do exist. Current forms of computer output and suggestions for alternatives should be examined by considering how well they contribute to human understanding of data information.

1.2: Visual Computer Data Presentation

A computer can be especially powerful in aiding the human ability to comprehend and manipulate large bodies of information. Given algorithms for extracting, grouping, and summarizing information, the computer offers a wide variety of tools for presenting the results. Currently, this digital output is primarily visual; in particular, text and graphics are extensively used to communicate from a computer

---

1. Figures 1.1 and 1.2 were provided by the Applied Technology Group of the Engineering Research Division at Lawrence Livermore National Laboratory.

to a human user.[2]

Text output falls into two main categories: messages and computed results. Messages may be merely status information or more in-depth reports of program states. Calculational results may be as few as a final value or as complex as tables of data. Figures 1.3 and 1.4 are examples of text information for three sets of four-dimensional data.

RESULTS OF PRINCIPAL COMPONENT ANALYSIS

ten highest eigenvalues

4.228	0.243	0.078	0.024

corresponding eigenvectors

1 x1	0.4219	0.8992	-0.9734	0.4186
2 x2	-0.0987	1.0000	1.0000	-0.4242
3 x3	1.0000	-0.2374	0.1275	-0.6367
4 x4	0.4182	-0.1034	0.9129	1.0000
%var.expln=	92.462	5.307	1.710	0.521
cum.%expln=	92.462	97.769	99.479	100.000

Figure 1.3: Analysis Output

Transformed Values

x1	x2	x3	x4
3.2898e+00	7.7330e+00	-1.1035e+00	-4.1251e-02
3.2547e+00	7.0532e+00	-1.4066e+00	8.7141e-02
3.0506e+00	7.0971e+00	-1.0266e+00	-1.7758e-02
3.2183e+00	6.8597e+00	-1.0040e+00	-1.4453e-01
3.2377e+00	7.7431e+00	-9.0611e-01	-1.2553e-01
3.7605e+00	8.3109e+00	-7.7466e-01	-7.8384e-02
3.1305e+00	7.1731e+00	-6.2545e-01	-1.0613e-01
3.3574e+00	7.5193e+00	-1.0934e+00	-1.0436e-01
3.0537e+00	8.5035e+00	-1.0220e+00	-7.9740e-02
3.3030e+00	7.1398e+00	-1.3673e+00	-1.1895e-01
3.4966e+00	8.1790e+00	-1.1627e+00	-6.4183e-02
3.3731e+00	7.3158e+00	-8.8592e-01	-2.5175e-01
3.1707e+00	6.9738e+00	-1.4027e+00	-5.4719e-02
2.8598e+00	6.5952e+00	-9.5424e-01	-7.3019e-02
3.3357e+00	8.9100e+00	-1.3104e+00	1.8700e-01
3.8377e+00	9.1281e+00	-5.9219e-01	-3.5580e-02
3.3605e+00	8.4059e+00	-8.2568e-01	1.7831e-01
3.3316e+00	7.7227e+00	-1.0122e+00	5.8749e-02
3.8551e+00	8.4910e+00	-1.2580e+00	-8.3590e-03
3.4020e+00	7.9989e+00	-6.9941e-01	-1.3219e-01
3.7262e+00	7.8316e+00	-1.4572e+00	-6.4250e-02
3.4537e+00	7.8686e+00	-7.0812e-01	1.0235e-02
2.6690e+00	7.4784e+00	-5.6774e-01	-3.8306e-02
3.7350e+00	7.4306e+00	-9.9133e-01	1.5259e-01
3.6731e+00	7.2446e+00	-8.4787e-01	-4.4275e-01
3.4969e+00	7.0956e+00	-1.4806e+00	1.6682e-03
...	...	...	...

Figure 1.4: Data Vectors

Notice that although computer text output can provide comprehensive and complete information, such output provides little aid to the reader in extracting information. Text is most useful for describing particular features of data rather than patterns or overall structures.

Graphical output is an especially effective means of presenting data [24, 49]. A picture can be visualized at a glance and can contain more information in a limited space than text alone. Therefore, graphics usually offers a faster and broader view of the information.

---

2. In general, all computer output is visual with two notable exceptions. First, speech and music outputs are rapidly gaining in use. Second, computer output can take the form of actions controlling peripheral equipment.

Plots, bar graphs, pseudo-color images, contour maps, and movies are but a few of the graphic tools in use. Because graphics provides the major medium for computer information presentation, it is worth considering some of the methods in more detail.

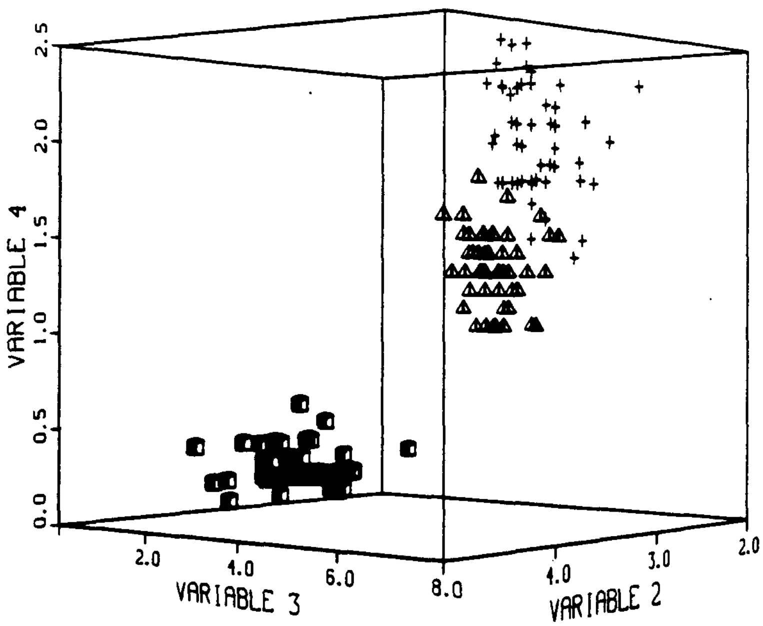

Figure 1.5: Three-dimensional Perspective Plot

Plots appear constantly in computer graphics applications. Plots include an x-y representation of points, a bar graph of several groups of data, or a three-dimensional perspective of an object. A graphical plot is a two- or three-dimensional representation of data in which the data variable values are mapped to locations on the visual display. Figure 1.5 presents a graphical view of the raw data referenced in Figures 1.3 and 1.4. In this instance, variables 2, 3, and 4 of the four-dimensional data sets are plotted in a two-dimensional rendering of three-dimensions. Notice that one set is clearly separated from the other two. Analysis methods reduce n-dimensional data to fewer components by mathematically determining linear transformations of the original variables. The transformed variables are such that the first has maximum variance, the second has maximum variance subject to being uncorrelated with the first, and so on [24]. An alternative view of the data shown before in Figure 1.5 results from plotting the transformed data after a principal components analysis. In Figure 1.6, the first two principal components are plotted. Notice again that Set 1 is easily distinguished from Sets 2 and 3. These are only two examples of graphical plots; other plots for data representation include probability plots, logarithmic plots, histograms, and curves.

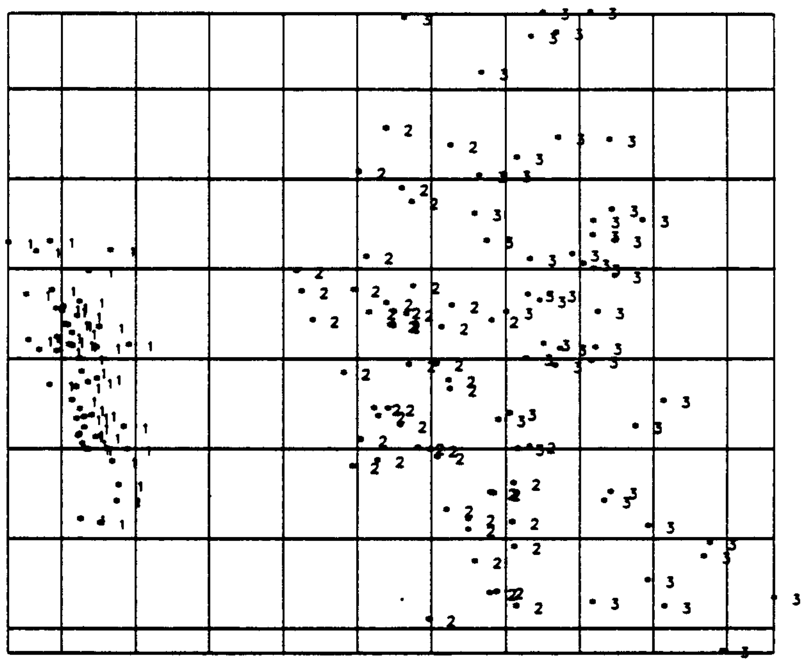

Figure 1.6: Two-dimensional Plot

Several especially interesting methods of graphical representation attempt to present all dimensions of the data to aid the human view of the gestalt in the subsequent analysis. Two examples are Andrews' function plots [4] and Chernoff's FACES [13]. Both of these imaginative methods accept multi-dimensional data and present a view of that data to the human user. Figures 1.7 and 1.8 plot ten samples from each of the three sets displayed in Figures 1.5 and 1.6. For each data sample, Andrews plots a curve which is based on a function whose coefficients are the values of the data sample variables.

(Eq. 1.1)

$$f_v(t) = v_1/\sqrt{2} + v_2 \sin t + v_3 \cos t + v_4 \sin 2t + v_5 \cos 2t + \dots$$

$$\text{where } -\pi < t < \pi$$

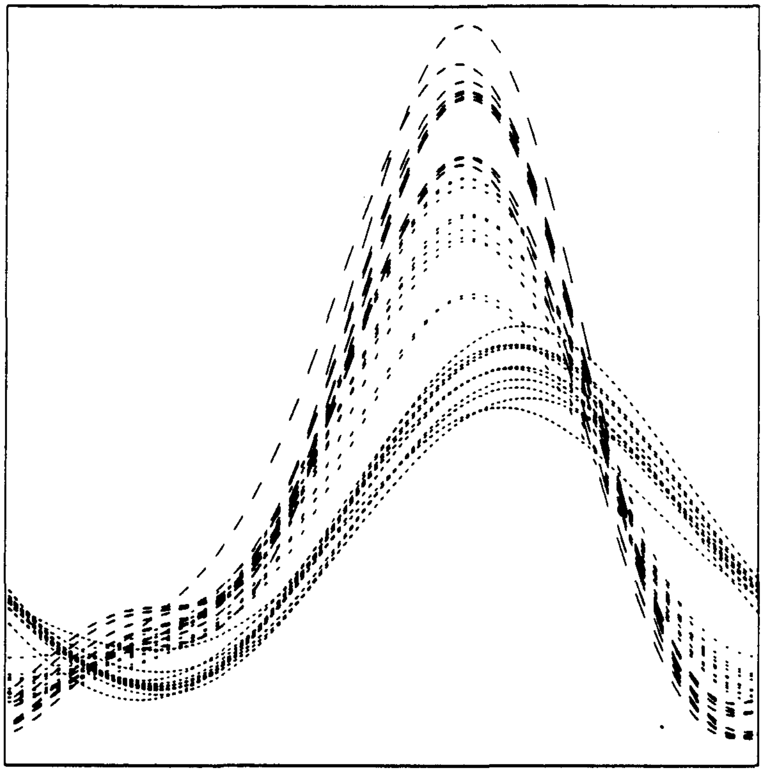

Figure 1.7: Andrews' Functions

Chernoff depends on the ability of the human to assimilate the face as a chunk of information and thus group similar samples. In both of these figures, one is able to separate the data clearly into the three known sets.

Image analysis and simulations are other outstanding examples of computer graphical output. Given an image divided into small cells, each with an associated data value, image analysis uses grey scaling or color to emphasize areas (and thus value levels) which might otherwise be hidden. This is an objective interpretation of numerical information inherent in the data but the method enhances subtle changes. Figure 1.9[3] shows a picture before and after image

---

3. Figure 1.9 was provided by the Signal and Image Processing Research Group of the Engineering Research Division at Lawrence Livermore National Laboratory.

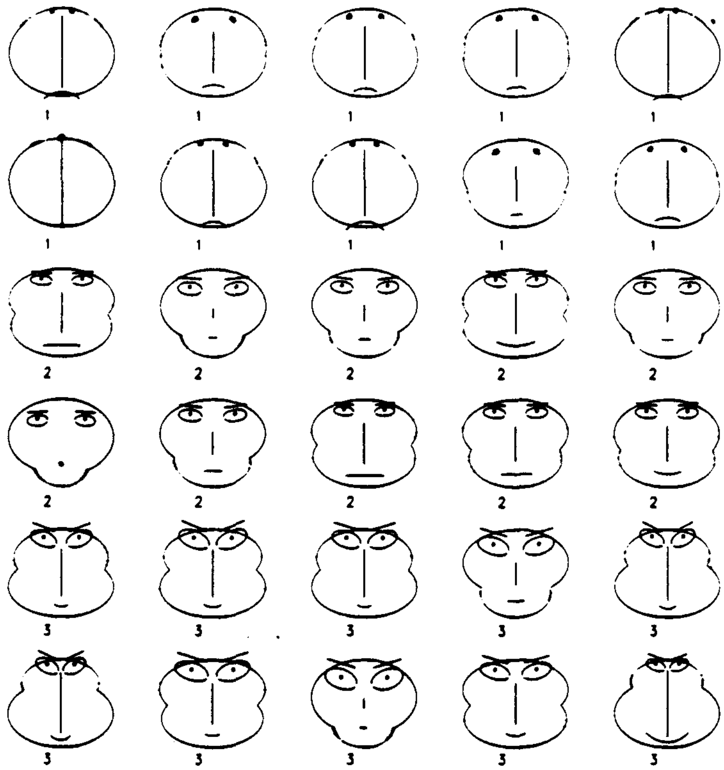

Figure 1.8: Chernoff's FACES

enhancement. Adjusting the mean and variance of picture cell values in each small region of the first picture creates more distinct edges in the second. Image analysis not only offers a varied perspective of the data but allows specific occurrences and patterns in the data to be observed as well as providing an overall view of the data.

Movies or simulations are significant output for time-varying data as well as a means of providing different views of static data. By observing the change of information from one state to another, one better understands the structure of that information. For example, Figure 1.10 shows nine successive rotations of two sets of data. By observing more than one view of the data, it can be seen that one set of data moves diagonally through the other set. Imagine how the view of that three-dimensional data would improve given a constantly rotating figure. The relative positions of all the data points would

Figure 1.9: Image Enhancement

become much clearer. An ability to animate the data provides a powerful analysis tool.

Other graphic techniques for presenting data include color and stereo. In all cases, color can be used to highlight particular information or to provide another dimension for discriminating among types of data. Stereo output is easily generated by creating two images, one shown to each eye from a viewpoint appropriate to that eye. The resulting depth perception provides a good representation of three-dimensional data when no other clues are present.

Graphics, and specifically the power of computer graphics, gets information to users more quickly than text output alone. Visual methods aid those skills necessary for extracting, grouping, and summarizing data information. Many graphical techniques are available so that various perspectives of the data are offered. Analysts using the power of computer graphics have the advantage of exploring their data by many methods.

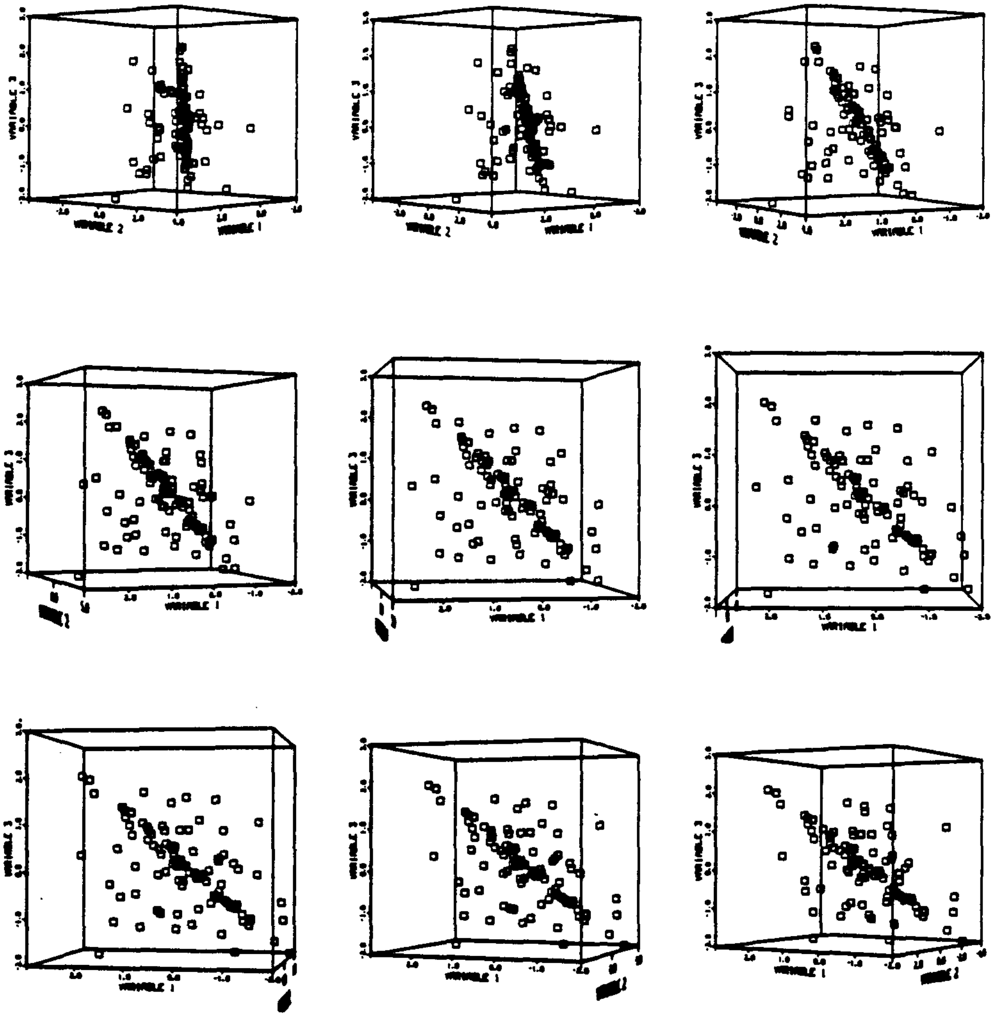

Figure 1.10: Rotating Data

1.3: Limitations

Despite the multitude of output from a computer, there is still information not being presented in a meaningful or useful manner. The limitations may be technical limitations based on the properties of graphical output or they may be perceptual limitations based on the characteristics of vision. Large data bases and multivariate data bases are particularly difficult to present. Current technical methods are restricted in the number of dimensions available, limited by the cost and speed of output, and confined primarily to visual attention on the output. Perceptually, not all data problems lend themselves to graphical representation. Understanding the limits of current information presentation methods can be helpful in seeking alternate techniques.

Restrictions on the number of output dimensions available limits the presentation of multivariate information. . By reducing an n-dimensional problem to an m-dimensional problem ($m<n$), one may lose information. Despite the work of Andrews, Chernoff, and others, n-dimensional output techniques do not include many satisfactory representations. In complex simulations, with many events occurring simultaneously, static approaches are not even applicable. Increasing the bandwidth between computer and human is necessary.

The widespread use of graphics is still costly, particularly for flexibility in working with large amounts of data.[4] A display is updated at most 60 times a second. The resolution and update speed place an upper limit on the actual number of bits of information which can be presented graphically (ignoring for the moment the number of bits of information which can be comprehended). An easy availability of output media for large amounts of data is needed.

Visual output requires that the observer's eyes be on the graphic output in order to obtain the information being presented. In a simulation, attention to one event may cause another simultaneous event to go unnoticed. The ability to provide more than one form of sensory input would be helpful.

The perceptual limitations are as significant as the technical constraints. Our vision is accustomed to a three-dimensional space that consists of distances related linearly. We have no model for visualizing an n-dimensional space or non-Euclidean geometries. Data

---

4. Despite an impressive array of hardware at ever decreasing costs, a typical 512x512x1 display costs approximately \$10,000. Such a display allows 262,144 bits of information, assuming each bit of information is simply in one of two states (on or off). The number of pieces of data will be greatly reduced if the display is used for more complex representations of the data. The cost will increase if color is added to differentiate among data points.

which does not fit into the constraints imposed by visual familiarity is difficult to present and comprehend graphically.

1.4: Alternatives Using Sound

Computer-generated sound offers another option for information presentation. Sound is already a familiar means of obtaining everyday information (consider an alarm clock, the news, or a well-known melody). Furthermore, audio output may help overcome some of the graphical limitations described above [72]. Sound provides both an enhancement to graphic output and an alternative to data presentation.

Motivation

Technically, sound offers a much different medium than graphics. The dimensionality of sounds, though not as familiar as x-y axes, consists of several components (such as pitch, volume, and duration). Sound equipment is available and relatively inexpensive.[5] The data rate can be extremely high since thousands of samples a second constitute a sound. And, unlike visual output, sound can be heard regardless of the physical orientation of the listener's attention.

Perceptually, sound offers a structure that is already familiar (as smell, taste and touch do not). We have a language for discussing sounds. It has meaning to say that one sound has higher pitch or lower volume than another. Because perception of sound is different than visual perception, sound can offer a different intuitive view of the information it presents. For example, musical octaves are familiar to most people and provide a natural expression of logarithmic variance.

---

5. For many micro-computers, sound boards are available in the $300-$500 range and will output frequencies from 50 Hz to 20,000 Hz with volume and waveshape control.

Sound output is not without limitations as well. First is the necessity for an environment in which distracting noises can be minimized. Second, sounds are transient, unlike graphical output which can be static. Third, the static attribute of graphics also makes reference points, such as grids or time lines, readily available.

The advantages and limitations of both sight and sound serve as a reminder that a variety of tools is desirable for presenting information. Furthermore, graphics and sound combine naturally and can be presented simultaneously. The possibility of utilizing sound with graphics will strengthen the best capabilities of each.

Current Research Efforts

Research has begun to consider the use of computer sound for information presentation. Natural sounds have been recorded and used successfully for data discrimination [84]. More recently, Mathews [58], Yeung [99], and Wilson [97] have tried encoding data in sound. Their work shows positive results.

Mathews used both graphics and sound to present up to five dimensions of data. Two variables provided a visual x-y scatter plot, while frequency, timbre and amplitude modulation provided a corresponding note of three dimensions for each point. A user interactively selected a sequence of points to hear. Using real data which had been analyzed by other means as well, Mathews showed that auditory representations did reveal structure in the data.

Yeung maps multivariate data vectors to properties of sound such as pitch, volume, direction, timbre and duration. He drives the sound output directly and thus utilizes complex waveshape functions such as damping. Furthermore, Yeung feels that it is possible to use at least nine dimensions in sound and perhaps as many as twenty. For instance, Yeung repeats a note with a rest period between repetitions. This rest period is one of his dimensions. Four analysts heard 40 samples from four data sets and were subsequently able to achieve 90% to 100% correct classification when hearing a data item again.

In a third approach, Wilson relies heavily on the musical aspects of sound. She suggests a unique representation of data by playing sequential tones for each data vector. Instead of a single note corresponding to an ordered set of values, an ordered set of notes (or melody) results. Wilson also suggests that tables of data may be encoded into a set of notes so that patterns in the table data are readily apparent. Wilson emphasizes the value of interactively using sound to explore data sets. She reports positive results when comparing her methods to traditional methods of exploratory data analysis.

The research of Mathews, Yeung, and Wilson represents only a beginning of the potential use for sound and information presentation. All three concentrate on data which is represented by n-dimensional vectors. In addition, Wilson uses sound for problems involving contingency tables. They do not suggest other types of data which might be advantageously represented by sound. Only Mathews combines sound with a graphic presentation. There are no formal studies to verify that sound does compare favorably with traditional methods of data discrimination. Mathews, Yeung, and Wilson have provided good initial findings on which to broaden the scope and understanding of sound and information presentation.

2: SOUND

This chapter provides a brief review of techniques for computer generation of sound in order to consider sound as an alternative for presenting information. The material is not critical to the understanding of later chapters but provides background information. Much of the discussion is based on material from *The Physics of Musical Sounds* [88] and *The Technology of Computer Music* [57]. This chapter first describes the nature of sound. There is a discussion of computer sound generation and a brief explanation of the perception of sound. The chapter concludes by reviewing a few of the current research efforts in computer-generated sound.

2.1: Description

A sound is a wave which causes pressure variations to be transmitted to, and thus heard by, the ear. Physically a sound wave is characterized by its frequency and amplitude as a function of time. The resulting waves are interpreted by the ear in terms of pitch, timbre, and loudness. The relationships between the physical and physiological attributes of sound are not well understood and are a significant area of research.

A wave which is a continuous function of time can be represented mathematically as an infinite sum of sine functions.[6] Furthermore, two or more waves passing through a given point simultaneously have no

---

6. A theorem of real analysis demonstrates that any continuous function may be built from a countable sum of appropriately chosen basis functions. Fourier analysis uses trigonometric functions.

effect on each other; the result in sound is the algebraic sum of the component sounds. Therefore any sound or combination of sounds heard by the ear may be expressed as an infinite sum of various sine waves.

Since sound is a continuous function of time, the following equation represents any sound wave:

(Eq. 2.1)

$$f(t) = \sum_{k=1}^{\infty} A_k \sin(2\pi f_k t + \varphi_k).$$

A_k = amplitude of the k^{th} component
f_k = frequency of the k^{th} component
φ_k = phase shift of the k^{th} component

Frequency determines the number of cycles in a given unit of time. Figure 2.1 illustrates these components. Since the sine function repeats every 2π radians, the phase component shifts one sine wave relative to another. Adding any combination of these will yield a continuous function of time.

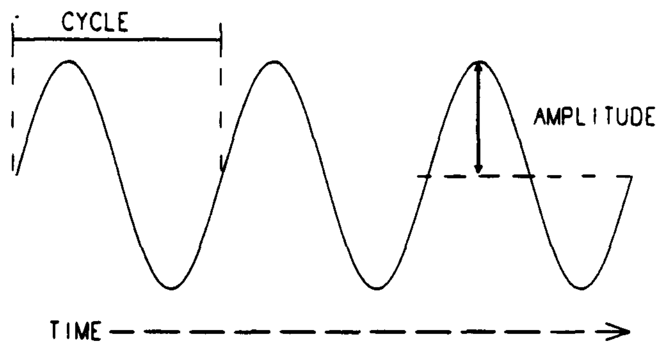

Figure 2.1: Waveform

Pure and musical tones are characterized by periodic functions, $f(t) = f(t+p)$. Not only does the function repeat itself at a regular interval but also the function components are organized in some way. Other sounds may consist of periodic functions with dissonant frequencies or random non-periodic functions with no perception of regularity. Because periodic functions are amenable to digital computation, current sound generation techniques have usually assumed periodicity in the underlying formulas. This does not exclude other possibilities, particularly when producing sound from a set of data values rather than from a function equation.

A simple sine function (i.e. $y = A \cdot \sin(2\pi ft)$) produces a pure tone.[7] That is, the output tone has the frequency and amplitude of the sine function. Any other sound consists of several sine waves, each with a particular amplitude and frequency. When higher and lower frequency sine waves are included (called *partials*), the tone becomes fuller and more complex. These resident frequencies combine by adding and subtracting to produce a wide spectrum of perceived frequencies. Musical tones are a subset of such functions in which the partials are integral multiples (i.e. *harmonics*) of the fundamental frequency. When inharmonic partials (those which are not small integral multiples) are present, the pitch is musically ill-defined, much like a bell or gong.

The equation for musical tones is derived from the general equation (Equation 2.1) by restricting component sine waves to be integral multiples of the fixed fundamental frequency f. That is, instead of adding sine waves of various frequencies, f_k, each k^{th} component is a multiple of the fundamental ($k \cdot f$).

(Eq. 2.2)

$$f(t) = \sum_{k=1}^{\infty} A_k \cdot \sin(2\pi kft + \varphi_k)$$

In addition to the sine wave, other simple functions which produce musical tones are the square, triangle, and sawtooth waves. A *square* waveform is theoretically an infinite sum of sine functions, each of which is an odd harmonic of the fundamental frequency. In Equation 2.3 below, a square wave is a sum of sine waves with frequencies that are odd multiples ($2k-1$) of the fundamental frequency f with proportionately smaller amplitudes. Figure 2.2 shows 300 odd harmonics and the sum as it approaches a square wave. Because the square wave consists only of odd harmonics, it is like the wave set up

---

7. A pure tone is a musical term for the tone produced by an idealized vibrating string.

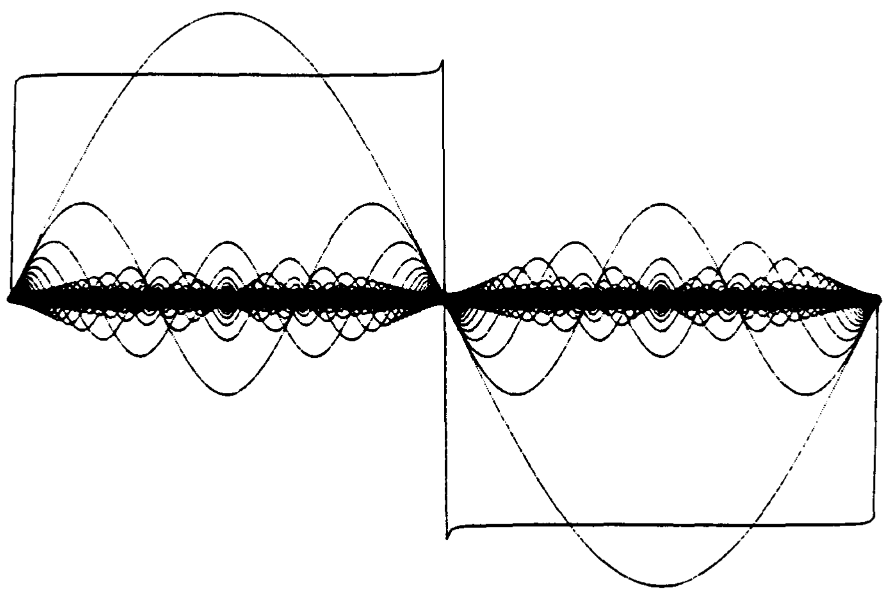

Figure 2.2: Square Wave

by a closed-end pipe. Three odd harmonics are shown in the pipe of Figure 2.3. Clarinets and organs are instruments which are examples of variations of closed-end pipes. Thus a square wave produces a sound which is hollow and somewhat woody, much like a clarinet [1].

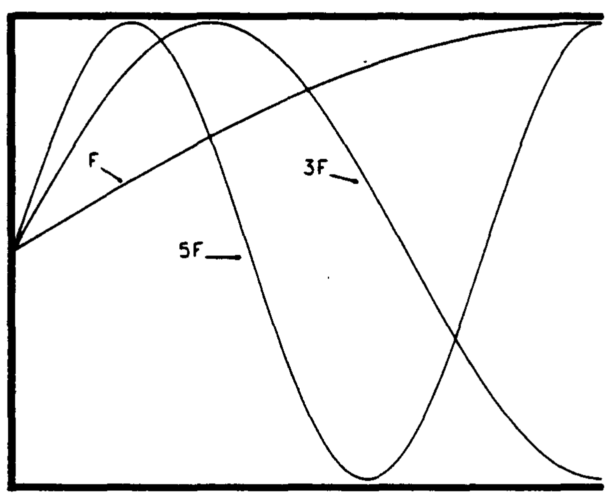

Figure 2.3: Closed-end Pipe

The *sawtooth* wave consists of frequencies inversely proportional to their amplitudes. Equation 2.4 shows that for each integral multiple of the frequency f, the amplitude A is proportionately smaller. The sound is more like that of a stringed instrument or English horn [1]. A triangle waveform also falls in this set of functions which contain a complete overtone series, but the triangle waveform has a different distribution than the sawtooth.

Square wave

(Eq. 2.3)
$$f(t) = \sum_{k=1}^{\infty} (A/(2k-1)) \cdot \sin[2\pi(2k-1)ft]$$

Sawtooth wave

(Eq. 2.4)
$$f(t) = \sum_{k=1}^{\infty} (A/k) \cdot \sin(2\pi kft)$$

Table 2.1: Equations for Musical Tones

2.2: Techniques of Sound Synthesis

The current results in sound synthesis techniques are those which are a side-effect of the research in computer speech and music. Our experiences with sound fall primarily into four categories: speech, music, auditory cues such as bells or alarms, and effects of natural events. All of these forms convey information to the listener, but speech and music have been the main areas for human creation and control. It is not surprising that computer-generated sound has concentrated also on the fields of music and speech.

The goal of speech synthesis is duplication of spoken language. The sounds are generally of short duration with frequencies in a fairly narrow bandwidth. In addition, the phonemes of speech form a relatively limited set of data for reproduction. Because speech synthesis is strictly directed toward particular types of sounds, the techniques are not applicable for general use.

Musical sounds include a wide range of frequencies and a wide variety of waveforms. Musicians are interested in new sounds, as well as those produced by traditional musical instruments, so that exploration of new techniques is important. Much of the effort has focused on instrumental tones which are periodic, contain a complex combination of harmonics, and have definite attack, steady-state, and decay phases. The research in computer music has created a wide base of techniques which can be extended to other areas of sound research.

Synthesis of sound concerns itself with producing waveforms which, when converted to an analog signal, can be heard by the human ear. Techniques include combining various frequency and amplitude components to produce a composite waveform, reducing a complex waveform by removing some of its components, and modulating the waveform by some (usually non-linear) formula. These techniques are realizable by signal processing systems ranging from an all analog electronic synthesizer to a computer with a digital-to-analog converter and amplifier/loudspeaker. Flexibility makes digital computers the basis for the rest of this discussion.

In his survey of signal processing aspects of computer music, Moorer [62] distinguishes three methods for synthesizing sound: direct synthesis, analysis-based synthesis, and *musique concrete*. In direct synthesis, the computer constructs wave samples using only mathematical functions. In both analysis-based synthesis and *musique concrete*, the waveform samples come from digitized natural sounds. Analysis-based synthesis includes functions which modify the digitized samples. *Musique concrete* essentially replays the digitized sounds without change. These latter two methods are data driven and thus not as general for wide application of sound production; they are not considered further in this paper.

Direct Synthesis

Direct synthesis endeavors to use a mathematical model to calculate a sound waveform. Equation 2.1 (described earlier) is an example of a mathematical model in which various frequency and amplitude components are added to produce a desired waveform. This method of direct synthesis is called *additive synthesis*. Another frequently used method, called *subtractive synthesis*, models a filtering function and reduces a wide spectrum of frequencies to a desired waveform.[8] In either method, results are not based on analysis of existing sounds but on construction from function parameters. Because additive synthesis is most commonly used to produce a wide variety of general sounds, it will be explained in some detail.

To eliminate unnecessary computation, direct synthesis of a waveform is often table-driven. Mathews [57] particularly exploited this concept by using modules to produce tables for later computational use. Typically, a module produces a table for a specific function such as the square or sawtooth function. A resulting table consists of equally spaced amplitude values for one cycle of the waveform.

TABLE(1) = 0	TABLE(501) = 500
TABLE(2) = 1	TABLE(502) = -500
TABLE(3) = 2	TABLE(503) = -499
TABLE(4) = 3	TABLE(504) = -498
.	.
.	.
.	.
TABLE(499) = 498	TABLE(1000) = -2
TABLE(500) = 499	TABLE(1001) = -1

Table 2.2: Sawtooth Function

Table 2.2 shows the sawtooth function discussed earlier with 1001 table entries and a peak amplitude value of 500. For synthesis, the time for developing tables is part of the preprocessing and therefore does not affect the run-time cost.

---

8. White noise is often the basis for subtractive synthesis. It is a random sound composed of many different frequencies uniformly distributed over a wide range and not harmonically related [92].

These table values represent one cycle of a periodic wave. The number of times the table values are output per second is the number of cycles per second or frequency. Typically, a computer sound system will output a specific number of samples or table values per second, usually 20,000 to 30,000. This number is the sampling rate, S. Thus to obtain a desired frequency (number of cycles per second), the table length is dependent on that frequency and on the sampling rate. That is, if the sampling rate is S samples per second and the desired frequency is C cycles per second, then the table must contain exactly S/C entries to describe one cycle of the desired waveform. Obviously this is the fastest output method and certainly allows real-time generation given that memory access rates exceed the sampling rate.[9]

In the situation above, a different table would have to exist for each possible frequency. In practice, a single table is used by simply choosing from that table the necessary number of entries for a particular frequency. If C is the desired frequency in cycles per second, then $NUM = S/C$ is the number of output samples per cycle. If E is the number of entries in a table describing one cycle of a given waveform, the output consists of incrementing an appropriate amount (E/NUM) through the table. Combining output from several tables representing different waveshapes builds more complex waveshapes.

Problems arise because the table increment must be an integer value in order to index into a table of discrete values. Three methods are used to determine the proper sample output [57]: truncation, rounding, and interpolation. Truncation and rounding are the simplest (and therefore the fastest for a computer), but they lead to obvious distortion in the output. Because output samples differ from actual waveform values, the resulting waveform approximation also differs from the actual waveform. The dotted line in Figure 2.4 represents a sine wave defined by 100 values. The solid line shows the waveshape which results when an increment of 13.7 is used to pick values from the table using truncation. When this difference is audible, the ear

---

9. Given a sampling rate of 25,000, the memory access time must be at most $1/25{,}000 = 40\ \mu sec$. This time is easily met by almost all modern mini-computers.

hears distortion. For the same length table, interpolation has been shown to give better results.[10]

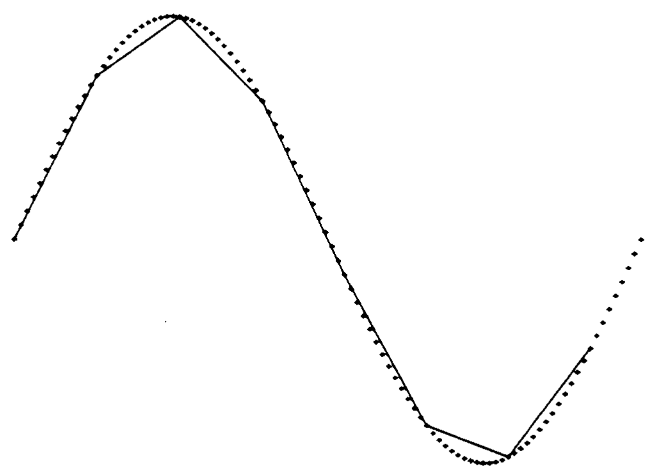

Figure 2.4: Table Truncation

Generating more complex natural sounds, particularly for music, requires significantly more complex waveforms. These sounds are characterized not only by their individual sinusoidal combinations but also by their attack and decay periods. Chowning [16] has developed another model, digital frequency modulation, which has proved quite successful in music synthesis. His work includes consideration of the attack and decay periods, harmonics, and modulations of basic waveshapes for realizing various instrumental tone qualities.

Chowning's equation

(Eq. 2.5) $$f(t) = A(t) \cdot \sin[2\pi f_c t + I(t)\sin(2\pi f_m t)]$$

represents a frequency-modulated waveform. The fundamental (or carrier) frequency f_c is sinusoidal but modulated by a second sinusoid

---

10. Mathews [57] states that in general, rounding is about twice as accurate as truncation. While doubling the table length doubles the accuracy for truncation or rounding, it quadruples the accuracy for interpolation.

of frequency f_m. The modulation index, I, is the amplitude of the side frequencies. The amplitude is also a function of time. For many tones, this synthesis formula yields a complex wave spectrum without the necessity of adding a multitude of component waveforms or of obtaining a spectrally rich waveform and passing it through a time-varying filter. Changing the ratio of the carrier frequency to the modulating frequency controls the spectral partials; changing the modulation index controls the contributions of the partials by determining their amplitudes.

Problems

The first problem of sound synthesis is that, unlike the continuous sinusoidal waveform models described in Section 2.1, the corresponding digital approximations are finite discrete sequences. A waveform is well-represented by a discrete sequence if interpolation between terms yields a smooth curve which has a small error relative to the desired continuous function. Thus to represent a waveform digitally, it is necessary to choose carefully the minimum number of samples. The three primary problems in representing a continuous function digitally are aliasing, quantization, and windowing [96].

Aliasing (or foldover) is caused by an insufficient number of samples for accurate reproduction of the waveform.[11] It introduces incorrect frequencies which then combine with existing frequencies to produce additional sums and differences. When a curve is drawn through the wave by interpolating between samples, a very different wave is generated. Using the Sampling Theorem of Fourier analysis, it can be shown that a frequency, F, must be sampled at a rate of at least 2F to reproduce the waveform. If the frequency of the wave is F and sampling rate is R such that R < 2F, then the waveform generated has a frequency F1 = R - F. F1 will add unexpected and, very likely, inharmonic frequencies to the waveform.[12] Figure 2.5 shows a waveshape of frequency 9 (F=9). If sixteen samples (R=16) of this waveshape are

---

11. A common example of aliasing in video is the wheel which turns backwards when its rotational frequency exceeds 12 revolutions a second in a standard 24 frames per second movie.

taken, the resulting curve has frequency 7 (R-F=F1=7). Note that although 64 samples produce a more accurate reproduction of the original waveshape, 18 samples are sufficient to produce a frequency of 9.

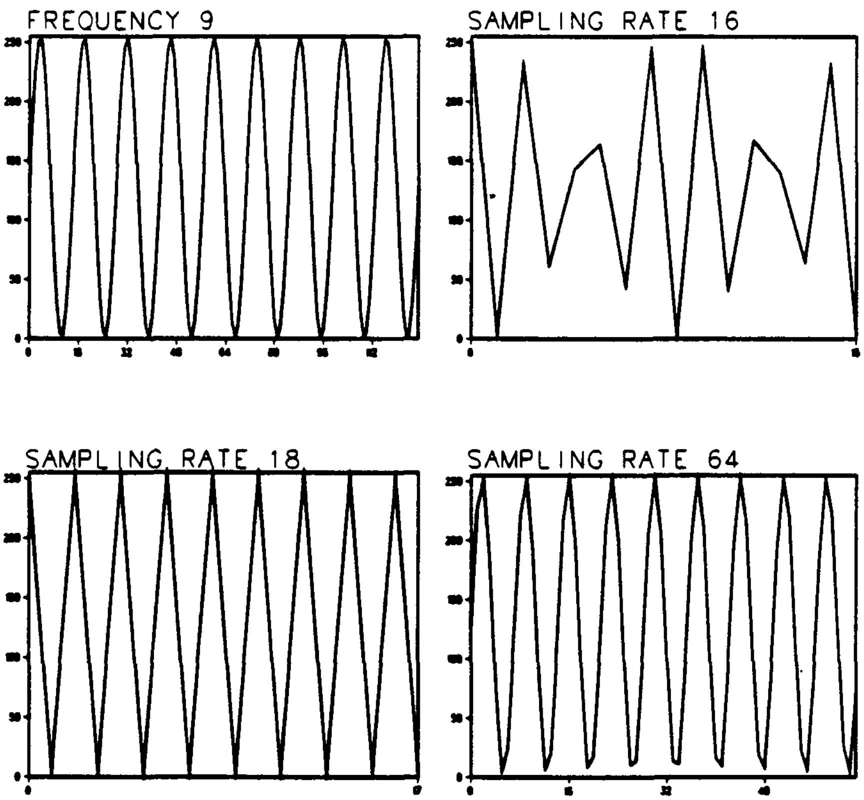

Figure 2.5: Sampling

Quantization introduces noise distortion and is caused not by the sampling rate but by the accuracy of the sampled values. For each sample value S(t), the error S(t) - f(t) depends on the computer word size. This error results in a signal-to-noise ratio of approximately 6 times the number of bits used per value.[13] Most audio equipment has

---

12. Hearing ranges from 20 Hz to 18,000 Hz so that a sampling rate of 36,000 Hz is desirable; most music software uses about 25,000 samples per second.

13. The signal-to-noise ratio is doubled for each additional bit of accuracy. Expressed in decibels, this is an additive component of 6 (i.e. 20log 2 = 6).

a signal-to-noise ratio of 66-72dB so that 11 or 12 bits is a reasonable degree of digital accuracy.

Windowing occurs when a sample set consists of less than a complete cycle. Because it arises when real data drives the synthesis, windowing is not a serious problem for most sound synthesis. If a set of data exists with an unknown cycle time, the number of samples needed can be determined by Fourier analysis.

A second major problem in sound synthesis is computation time. Even a simple tone with high frequency components requires about 25,000 samples each second. Any musical sound consists of several instruments or voices, each requiring the generation of 25,000 samples every second. For real-time sound generation, time is especially critical. Even the fastest hardware is limited in the number of operations which can be performed in 40μsec. With the increase of micro-computers and distributed processing, most sound synthesis systems now drive the actual sample generation with an independent processor.

A third major difficulty of sound synthesis arises from the lack of understanding of the relationships between physical and physiological aspects of hearing. In generating an arbitrary sound from a mathematical model, it is difficult to predict the perceived effects. If synthesized sound is to be used to convey meaningful information, it is particularly important to know what effects the equation parameters have on hearing.

2.3: Psycho-acoustics

Having described the generation of sound, it is important to note some facts regarding the human sensations produced by sound. Various combinations of sound characteristics have different effects on the perceived sounds. Both music and psychology have concentrated research in this area, for example by comparing human responses to a variety of differences in sound characteristics and by analyzing instrumental sounds. The perception problem is, in some ways, more serious than the problem of producing sound since one would like to predict in advance the effects of the produced sound.

Studies suggest that the human ear can discriminate between any two of 400,000 different sounds presented in rapid succession and that most people can remember and correctly identify 49 different sounds at one time [68]. In fact, reaction time to auditory stimuli is faster than to visual stimuli [8]. However, the perception of the sound one hears is dependent on the context of that sound -- that is, on the combination of parameters describing the particular sound and on the notes preceding and following the sound. To represent data using sound, it is important to understand how the various sound characteristics are perceived.

Most notable, perhaps, are the interrelationships among characteristics of sound. Sounds in the middle range of frequencies seem more varied, and are therefore easier to separate, than high range sounds. Not only are the frequency changes more difficult to perceive but the volume of a high note appears softer than for a corresponding note at lower frequency. Duration also affects the perception of a note. For example, a shorter duration note appears to be lower in pitch and a louder note appears to lengthen the duration [50, 98].

Timbre, "the multidimensional wastebasket category for everything that's not pitch nor loudness" [50], affects perceived pitch, loudness, and duration! Varying the number and relative intensities of existing overtones makes the same fundamental frequency sound differently. The attack and decay periods heavily influence sound, particularly those of short duration. Listening to a trumpet and a violin is an example of the importance of timbre. Two notes, each at the same frequency and amplitude, sound completely different.

Fully utilizing sound output for presenting data will require an understanding of the relationships and effects of sound parameters. The observations discussed above are only a few of the facts regarding perception of sound. Further studies must consider the effects on data representation.

2.4: Sound Synthesis Research

Music and speech are the current focus for research in sound synthesis. Digital music supports music analysis, music performance, and music composition. Researchers are exploring the perception of sounds using the computer to generate a variety of waveforms under controlled conditions not previously possible with conventional instruments. Speech is being used to further the study of human/machine communication.

Two examples of computer music research are the efforts at Bell Laboratories and at Stanford University. Max Mathews is considered by many to be the father of computer music. His book, *The Technology of Computer Music* [57], is a basic guide for the understanding of digital music. His software FORTRAN program, MUSIC V, has been extensively used for studies of timbre. Work at Stanford involves a frequency modulation equation as the synthesis technique to produce more complex and natural musical sounds. Chowning [15] has achieved the synthesis necessary for a perceived sound location which varies dynamically. Both of these studies are especially relevant to the perceptions of sound and the further use of sound.

Given the methods and equipment for computer-generated sound, it is appropriate that its use has extended to areas other than music and speech. The tools for producing sounds are available and the need for additional methods of presenting information exists. Initial studies indicate that data values can be encoded into sounds so that data relationships are preserved. The next chapters expand the use of sound for information presentation by explaining tools for using sound to represent data, applications for a wide variety of data types, and results of experiments in which subjects classified data based on sound representations.

3: FACILITIES AND TOOLS

Research in the use of sound for information presentation requires both hardware and software capabilities. The system must support sound generation that can be based on data values and that allows control of a number of sound parameters (dimensions). Four requirements provide a minimum capability for the variety of sounds needed for the work described in this thesis:

1: Capability for at least two simultaneous notes or voices,

2: Control of envelope and waveshape for timbre variations,

3: Control of pitch and volume, and

4: Control of note duration.

A sampling rate of 20,000 samples per second allows a maximum frequency range of 10,000 Hz; this encompasses the normal hearing range. To be generally useful, all components must be easy to use and inexpensive. Fortunately, the computer music field provides extensive resources.

Two additional resources are helpful in examining the potential of sound output for data: a computer graphics capability and traditional data analysis tools. Since most computer information presentation is visual, graphics can provide a basis for comparison and for experiments in combined visual and aural output. Traditional computer tools provide the capability for generation and analysis of complex data sets. Although neither of these functions are required for presenting computer information in sound, they add valuable support for sound experiments.

The research in this thesis was carried out on an interactive four-voice system and on a large time-sharing system with a great deal of compute power and batch output. The equipment often determined specific features of the studies so a description of the facilities and tools is helpful in understanding later chapters. Additionally, these two very different facilities provide good examples of useful computer sound systems.

3.1: Interactive System

The interactive system simultaneously produced color graphics and sound output.[14] The facility did not provide tools for creating nor manipulating large data sets, so all data was preprocessed on other computer systems. Sound output varied with respect to pitch, volume, duration, envelope, and waveshape for four notes. Sound could not be preserved except by recording the analog output in the usual manner on cassette tape.

Hardware

Most projects described in Chapters 5 and 6 ran on a Varian (V73) minicomputer with an Imsai (8080) microcomputer and four Solid State Music synthesizer boards. All data computations ran entirely on the Varian while the synthesizer boards generated the actual sound output. The Imsai provided an easy interface between the Varian and the music synthesizer boards; the music boards could not interface directly to the Varian bus. In addition, an Aydin 5214 frame buffer display system provided graphic output with capabilites for interactive input.

A synthesizer board computed the necessary wave samples which drove an analog output signal. Input for each music synthesizer board included

----

14. This facility is a special purpose system for interactive battlefield simulations and no longer includes the sound output capability. The sound equipment is being moved to another computer for continued work on information presentation.

a 128 byte waveform definition, a 16 byte envelope definition, 15 levels of volume, and a frequency range of 15 Hz to 25K Hz. The 128 byte waveshape determined one cycle of the periodic wave; the synthesizer board automatically repeated the wave as necessary for the desired frequency and duration. The frequency was determined by a frequency value ranging from 0 to 255 combined with an octave value ranging from 0 to 7.[15]

(Eq. 3.1)

$$\text{Frequency output} = \frac{20{,}000{,}000}{(256 - fv)\cdot(128)\cdot(2^{8-ov})}$$

fv = board frequency value
ov = board octave value

Not only were the synthesizer boards easy to use by merely specifying the necessary inputs, the total cost of the four boards was less than $600.

Each synthesizer board output an analog signal which had to be amplified. The four signals (from the four synthesizer boards) were combined in a preamp mixer and fed into a stereo system comparable to a home sound system. The sound output peripherals consisted of the preamp mixer with four stereo inputs (Numark DM 1100), an amplifier (Technics SU-8055), two speakers (Electrovoice Interface, Line Series 2), and a headphone (Koss Pro 4AA). This peripheral equipment cost was also less than $600.

Software

Two types of software provided valuable support: routines for generating the sound output and programs to control dynamically various sound characteristics. Access to the synthesizer boards was through a set of subroutines which specified each of the sound

----

15. For data sounds, a frequency expressed in hertz would be more appropriate. The configuration of the sound boards was probably due to the fact that they were intended for music output.

attributes (such as volume and frequency). The two most useful programs interactively controlled the music synthesizer boards and interactively manipulated the mapping between data sets and sound characteristics.

A FORTRAN-callable set of routines interfaced computer programs to the sound boards. These routines allowed the various synthesizer inputs to be set and controlled the on and off states of each board. To generate sound output, a program set up a waveshape, envelope, frequency, and volume for each board to be played. The program then had to send the signals to turn the board on or off. The time interval between turning the board on and turning the board off determined the note duration.

The sound control program graphically displayed a control panel for the four synthesizer boards. The frequency, octave, volume, and envelope duration controls corresponded directly to the synthesizer board inputs. A drawing area allowed the user to modify predefined envelopes and waveshapes and store them for listening or for input to subsequent software development. At any time, any of the four boards could be turned on or off.

The control program was particularly valuable for comparing the effects of various sound characteristics, for waveshape definition, and for diagnostics. Changing a sound parameter such as volume immediately changed the output sound. By turning on and off two sound boards with different settings, the different effects could be compared. Since waveshape determines the timbre of the sound, the ability to try interactively a variety of waveshapes was extremely useful. Finally the control program provided a diagnostic tool for checking all aspects of the sound boards.

The mapping program presented data as discrete sounds. Given a data sample as an n-tuple

$$S = (v_1, v_2, v_3, \ldots, v_n),$$

each v_i determined the level of a sound parameter (such as volume) for a specified board. Thus, each data sample S was encoded into a single sound. The mapping control program permitted dynamic manipulation of the mapping between v_i variables and sound characteristics.

The mapping program allowed up to four sets of data with up to 15 variables per sample as input. This interactive program was particularly helpful in determining the effect of various sound characteristics and the significance of individual data variables by providing a means for setting up a data-to-sound mapping. Varying only one sound characteristic at a time quickly illustrated how well that characteristic discriminated among data values. Conversely, holding only one sound parameter constant indicated how important that characterisitic was in the overall note definition. One easily obtained a good indication of whether a variable was helpful in discriminating among sets by mapping one data variable (v_i) to all sound parameters and then listening to the samples in two or more sets.

3.2: Batch Sound System

For batch output of sound, I used an extremely powerful timesharing computer network with a high precision output film recorder [45]. The sound is generated by drawing the soundtrack onto film. Thus, the system simultaneously generates both graphics and sound. Since the computer network is used for many types of computation and analysis, a wide selection of data creation and manipulation tools are available. For sound output, up to sixteen notes may be generated at a time, thereby allowing control of tones which can be used as overtones for timbre control. Unfortunately, the waveshape itself is not under program control. Because feedback is not immediate, the system is most useful for generating precise sound (and graphics) that can be preserved for playback at a later time.

Hardware

The calculations and soundtrack definitions are produced from programs executing on a Control Data Corporation 7600 mainframe computer. The results go to a magnetic tape which is read by an Information International, Inc. FR80 film recorder. This film recorder draws the soundtrack on 35mm film. These soundtracks are generated a frame at a time by drawing horizontal lines across the width of the soundtrack.

To hear the sound requires a 35mm playback facility or conversion to the more commonly used 16mm film.

Software

Software support includes FORTRAN-callable routines for generating the soundtracks and executable programs for manipulating large data sets. These latter programs were used for sound output on the timesharing system and for preprocessing data for input to the interactive sound system.

Software routines allow the frequency and amplitude to be set for up to 16 tones. The waveshapes for the tones are added to create a final output wave. Frequencies range from 10 Hz to 5000 Hz. The amplitude of the combined tones has a fixed maximum. Note duration was controlled by specifying an integral number of frames for each tone.[16]

The most useful data manipulation tools were those for creating random distributions, for traditional discriminant analysis of data sets, and for graphical representation of multivariate data. A random distribution generator creates complex data sets that differ in well-defined ways. Principal component, linear regression, and quadratic regression analysis programs are available as a basis for determining the difficulty of discriminating data sets. Finally, programs for graphical data presentation include two- and three-dimensional scatter plots and Chernoff's FACES.

---

16. Film is shown at 24 frames per second so the shortest note duration is 1/24 second (42 msec).

4: PRESENTING DATA IN SOUND

The past chapters suggest that computer sound can be used to present data. Two questions now must be addressed: first, what data is to be presented and second, how is this data to be encoded in sound? The literature does not yet provide a well-defined set of problems which best adapt themselves to sound presentation nor a well-defined method of mapping data to sound. My preliminary studies examine a variety of data types and techniques for using sound to represent data.

A general consideration of what data is to be presented in sound includes several data problems that are difficult to present graphically. In addition to multivariate data sets, logarithmic data and time-varying data are two such areas. Although logarithmic plots are familiar, the graphical relationships are nevertheless hard to grasp. The eye reponds to postion in a linear way; in sound, pitch and loudness are logarithmically related to frequency and amplitude.[17] Since sound is created by waves which are a function of time, sound offers interesting possibilities for time-varying data as well.

Having identified data analysis problems for audio output, the next task is to discover a straightforward method of relating data to sounds. Just as color, location, size, and shape are characteristics of visual presentations, several sound parameters are fundamental to audio presentations. The most obvious are pitch, volume, and duration. Another significant characteristic is timbre. The timbre of a note is affected by attack, decay, harmonics, and the waveshape itself. The relationships among notes also contribute to the sound

---

17. The intensity of color is logarithmically related to energy; however, humans have a more well-organized stucture for considering pitch relationships than for considering color relationships.

output. Consider varying sequences of notes (for example, melodies) and combinations of notes (for example, chords). Given these several characteristics of sound, the question then is how best to use them to represent data.

To explore the uses of sound for information presentation, both the data type and the data encoding in sound must be considered. The three data types outlined above provide a broad and varied basis for examination. Although multivariate, logarithmic, and time-varying data are not independent classes, specific features of each data type suggest various representations in sound. For each of the three classes of data, this chapter proposes a method for representing the information in sound and discusses my observations after presentation of that data in sound.

4.1: Multivariate Data

New methods for examining multivariate data are welcome because there is no *universal* technique for finding patterns and structures within a set of data or for assigning an unknown sample to one of several possible sets. Traditional techniques often throw away some of the data information so multivariate data lends itself to the possibility of increasing the information presented. Multivariate data is readily available from a variety of disciplines. A typical problem of exploratory analysis is to determine what, if any, factors distinguish one set of data from another. Specifically, given two or more sets of data, can an unknown sample be assigned to one of the sets? These problems are interesting because there is no predetermined algorithm for separating the sets of data. A specific data discrimination problem will often be interesting because there is no criterion for choosing the best method of assigning an unknown sample to a set.

The next two sections describe one method for data-to-sound encoding and two applications of real data. The encoding technique is the basis for the encoding used for logarithmic and time-varying data so it is described in great detail. This encoding method is also the one used for the experiments in Chapter 5. The positive results of applying sound representations to two groups of real data provide the motivation for later experiments.

Encoding Data In Sound

A set of data consists of individual data samples. If the data is n-dimensional, then each sample is an n-dimensional data vector or n-tuple. Thus, a data sample vector consists of n variable values. Each of the n variables maps to one of the n characteristics of sound. For a particular data sample, the specific n values associated with it produced a corresponding discrete sound (note).

Seven characteristics of an individual note were chosen for representing data values. These were pitch, volume, duration, fundamental waveshape, attack envelope, 5^{th} harmonic, and 9^{th} harmonic. Pitch, volume, and duration are necessary attributes of a note and easily varied. Initial attempts at producing a variety of distinguishable sounds indicated that timbre is one of the most noticeable characteristics. Thus, the fundamental waveshape, attack envelope, and harmonics were included as significant parameters. These seven parameters, then, were each varied over a specific range in a predetermined manner. (Listen to the parameter variations in Appendix A). Since four tones could be played simultaneously on the available equipment and seven sound parameters were chosen to vary for each note, 28 dimensions were technically possible.[18]

The actual variation of each of these seven characteristics is described in more detail. My intent was to vary each characteristic over a range which provided noticeable extremes, was easily audible throughout, and was computationally straightforward. An integer code, from 1 to m, corresponded to each variation throughout the range of a characteristic. The number of output levels available for each of the note characteristics varied widely. For example, only fifteen increments of volume were available but 48 values of pitch were used. Unfortunately, this affects the precision of the data value being encoded.

---

18. I did not attempt to determine an upper limit on the number of dimensions that are useful or a hierarchy of the sound characteristics. In fact, only 4 to 6 dimensions seemed to add information. Most users of the system said that pitch and duration were most valuable in distinguishing one note from another.

Several sound characteristics were not used and these deserve attention in the future. Certainly location is an easily detectable attribute of sounds. Harmony, dissonance, chords, and melodies are so familiar in music that they should be considered as potential parameters for presenting data. I chose the characteristics used because they vary easily and provide a wide range of distinct sounds.

Pitch

Because many people lose the ability to hear high frequencies, the fundamental frequency was limited to the range 130 Hz to 2000 Hz. Although musical ability is not common to all analysts, the musical scale provides a familiar set of sounds. Frequency increments were based on the piano scale. Thus, pitch varied over 48 values, twelve notes in four octaves.

Volume

The volume range was controlled by the output of the equipment. It is questionable how valuable volume is as a discriminant among sounds. Small differences are often difficult to detect, and the perceived intensity is strongly related not only to volume but also to the other characteristics of the note. Nevertheless, I used twelve of the fifteen values of volume available for variation, from very soft to very loud.

Duration

The length of a note varied from 50 msec to 1050 msec. This range was chosen arbitrarily with the intent that no note be so short as to be inaudible or so long as to cause the user to stop listening. Duration varied in 5 msec steps so there were 201 levels of duration.

Waveshape

A buzz and a pure tone are quite distinguishable, so the waveshape variation was based on a sine wave and a random waveshape.[19] The sound synthesis equipment allowed a waveform of 128 values. Thus, a sine wave was composed of 128 values ranging between -1.0 and 1.0. To begin altering the sine wave to a waveform that would produce a buzzing sound, one of the 128 values was randomly selected and randomly changed to a value between -1.0 and 1.0. For each subsequent change, one of the 128 values not already selected was randomly changed to a value between -1.0 and 1.0. In this manner a set of 128 waveshapes, varying from a pure sine function to a random waveshape, were defined (see Figure 4.1).

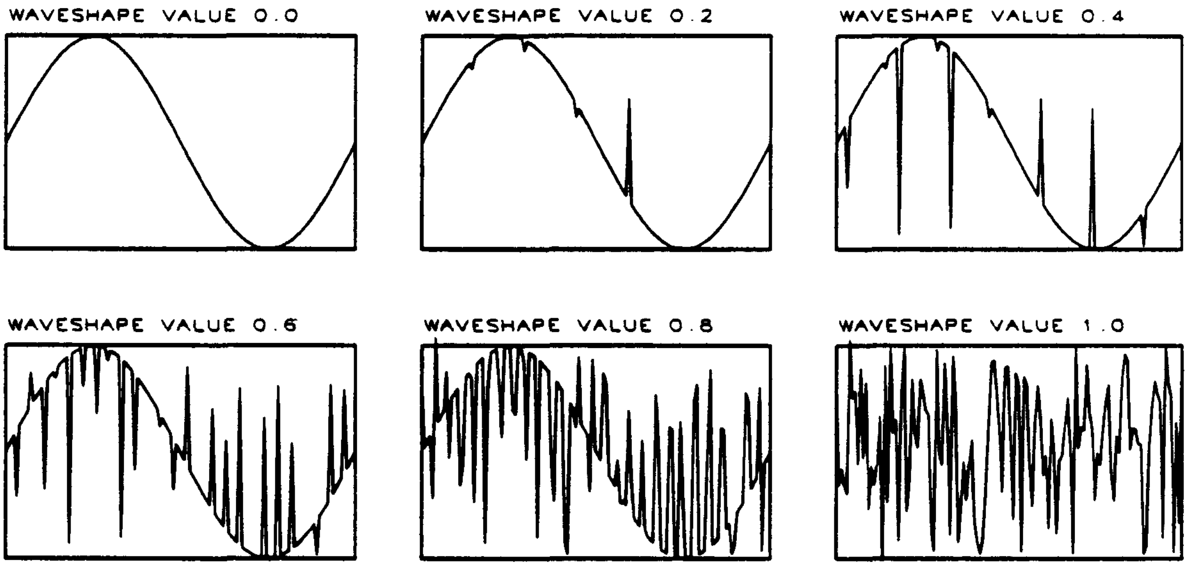

Figure 4.1: Waveshape Variations

---

19. In an earlier attempt in creating a range of waveshapes, I varied the waveshape among computationally simple but aurally different functions. As noted earlier, a simple sine function produces a pure tone, a square wave function produces a sum of odd harmonics, and a triangle function produces a complete overtone series. Since these functions are straightforward but produce very different sounds, interpolation between any two seemed like an obvious method for varying the waveshape. In fact, it was found that more complex waveforms offer greater aural differences.

Attack

The attack envelope affects the shape of the note and thus its overall timbre. A long attack makes the start of a note sound softer; a sharp attack causes the note to begin crisply, much like a pop from a gun. To have affected the end of a note instead, decay might just as well have been chosen. In all cases, the envelope affected the entire note regardless of total duration, volume or waveshape. Fifteen values were chosen which ranged from a long attack to a constant amplitude envelope (see Figure 4.2).

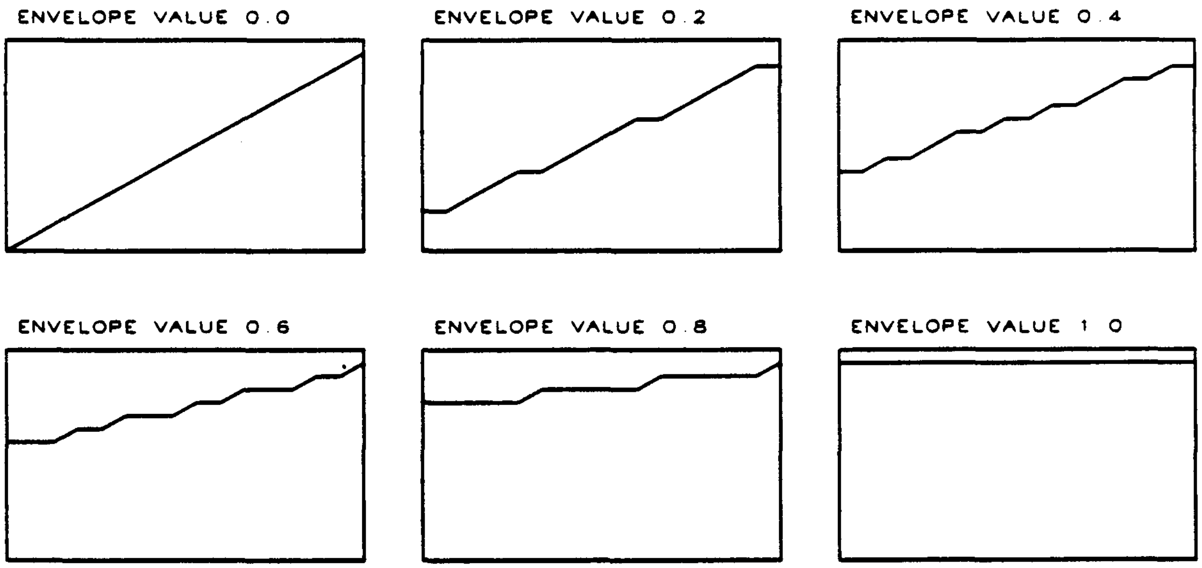

Figure 4.2: Attack Variations

Harmonics

Odd harmonics, the hollow-like sounds, seemed more noticeable than even harmonics, so the fifth and ninth harmonics were added as overtones (Figure 4.3). Each harmonic waveshape varied over a range between a pure sine function and a random shape, just as the fundamental waveshape varied.[20] The resulting overtone was

---

20. Harmonics might have been heard more effectively as a sound parameter if the amplitude of the harmonic waveshape added to the fundamental was varied instead of varying the harmonic waveshape itself.

algebraically added to the desired fundamental waveshape to provide a final waveshape for output. If no variable value was mapped to overtone, then no overtone was included. As with fundamental waveshape, there were 128 variations of overtone waveshape.

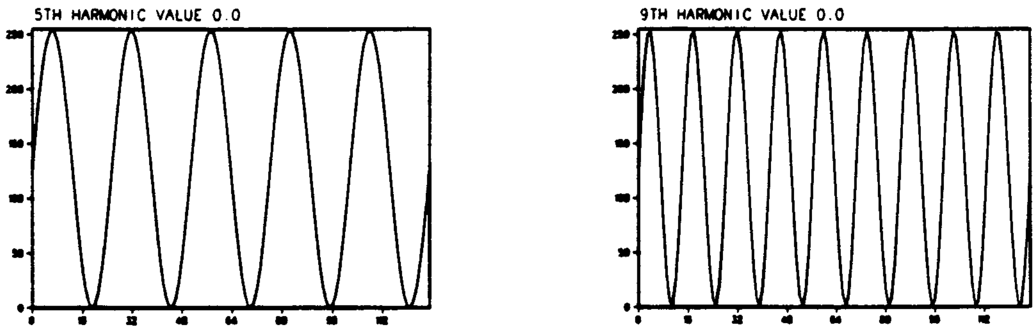

Figure 4.3: Harmonic Waveshapes

Example

A data set was encoded into sound by first, normalizing the data values to a [0.0, 1.0] range; second, determining a mapping of data variables to the sound parameters; and third, producing the note for a specified data sample. The range of a particular dimension throughout all sets in the problem was found and then mapped to [0.0, 1.0]. Thus, all data values were between 0.0 and 1.0. A different sound characteristic was chosen to correspond to each of the n dimensions. Finally, the values of a particular data sample were used to determine the levels of the appropriate sound characteristics. A given data sample then corresponded to a well-defined note.

For example, consider a four-dimensional data set including these three samples:

$$s_1 = (3.0, 5.0, 150.0, 3.0)$$
$$s_2 = (7.4, 2.0, 152.0, -1.2)$$
$$s_3 = (14.0, 8.0, 153.0, -4.0).$$

Assume that in the complete data set variable 1 ranges between 3.0 and 14.0, variable 2 between 0.0 and 10.0, variable 3 between 150.0 and 155.0, and variable 4 between -4.0 and 10.0. After normalizing these values to a [0.0, 1.0] range, the transformed samples become

$$S_1 = (\ 0.0,\ 0.5,\ 0.0,\ 0.5\ )$$
$$S_2 = (\ 0.4,\ 0.2,\ 0.4,\ 0.2\ )$$
$$S_3 = (\ 1.0,\ 0.8,\ 0.6,\ 0.0\ ).$$

The next step is to choose a data-to-sound mapping. Assume that variable 1 is represented by pitch, variable 2 by duration, variable 3 by volume, and variable 4 by waveshape. Since 48 levels of pitch are used, a data value of 0.0 will correspond to a pitch level of 1 (130 Hz) and a data value of 1.0 will correspond to a level of 48 (2000 Hz). Likewise, values for the other three dimensions will map to levels of duration, volume, and waveshape.

Each sample now produces a note. S1 would be represented as a very soft, low note of medium duration with a slightly buzzy sound. S2 would have a mid-range pitch and volume in a less buzzy note of short duration. S3 would produce a very high, long note with a pure sound at mid-range volume.

A change in the data-to-sound mapping will change the resulting notes. If variable 1 is mapped to waveshape, variable 2 to pitch, variable 3 to duration, and variable 4 to volume, then S_1 would be very pure and short compared to S_3 which would sound buzzy and longer.

Applications

Through the interest of Dr. Stanley Grotch [37], several multivariate data sets were available for sound encoding. Several different applications of data show that the sound encoding is a useful discriminant. Two of these applications are iris data from three different plant species and spectra data from four different materials. Listeners were able to discriminate among the sets with much the same success as with traditional discrimination methods.

Iris Data

Perhaps the most widely recognized data in the field of discriminant analysis is the Fisher iris data [28]. Using data collected by Dr. E. Anderson, Fisher examined the differences among three species of flowers, *Iris setosa*, *Iris versicolor*, and *Iris virginica*, with 50 plants of each species. For each plant, sepal length, sepal width, petal length and petal width were recorded. The problem is to use the measurements of an arbitrary plant to classify it as one of the three species.

The four variables (sepal length and width, petal length and width) were mapped to pitch, volume, duration, and fundamental waveshape. The resulting notes yield results that appear similar to traditional methods of discriminant analysis. The first set, *Iris setosa*, is quite distinct and easily identified. The remaining two sets are distinguishable in most instances but seem to have some overlap. (Listen to the examples in Section 4 of Appendix A). Most casual observers who attempted to place an unknown note into one of the three sets only missed an occasional one or two which belonged to sets 2 or 3. In fact, it is possible in four-space to separate the three sets so that all but one of the 150 samples are correctly classified.

<u>γ-ray spectra</u>

Recently, scientists at Lawrence Livermore National Laboratory measured the energy emissions of four different materials. Each spectrum consisted of seventeen channels of counts versus energy. Thirty γ-ray spectra were obtained for each material. This data was preprocessed using the statistical technique of principal components. Grotch [36] encoded the resulting transformed data into several graphical outputs, including Chernoff's FACES. Figure 4.4 shows twenty samples from each of the four data sets. Next he randomly selected another twenty from the remaining samples in the four groups of spectra data. These became the unknowns to be classified and are shown in Figure 4.5. In an informal study, a number of individuals of various ages and skill levels were able to identify the correct group for an average of 94% of the unknowns.

The same set of data was encoded into sound. (Listen to the examples in Section 5 of Appendix A). Five individuals, familiar with the sound testing procedures, were able to correctly identify an average of 94% of the unknowns. Table 4.1 shows the raw data for both exercises. The score shows the number of correctly identified samples over the number of samples given as unknowns.[21]

4.2: Logarithmic Data

Logarithmic data seems particularly appropriate for sound encoding since frequency and pitch are logarithmically related. A perceived linear pitch difference naturally corresponds to an exponential frequency difference. I examined two types of problems involving logarithmic data for sound representation. One type was discrete n-tuples much like the multivariate data described in Section 4.1 but which contained at least one dimension that varied exponentially. Another type was two-dimensional data typically represented graphically on a log/linear or log/log plot.

Encoding Data In Sound

In both examples of logarithmic data, frequency represented the dimension which varied exponentially. In the case of the n-tuples, the data samples were encoded in the same manner as the multivariate samples. However, frequency was always used to represent the logarithmic value with a mapping that preserved the power function. Instead of using the 48 levels of the piano scale, the frequency was mapped directly from the data value.

(Eq. 4.1) $\text{Output frequency} = a \cdot 2^{b \cdot \text{Data Value}}$

Equation 4.1 calculates the frequency for a given data value. The range of the data values determined constants a and b so that the frequency range was audible.

---

21. I used the transformed spectra data in a slightly different manner. First, the data was normalized differently. Second, each subject was given a *random* selection from each data set as training and a different additional random group of unknowns from all four sets.

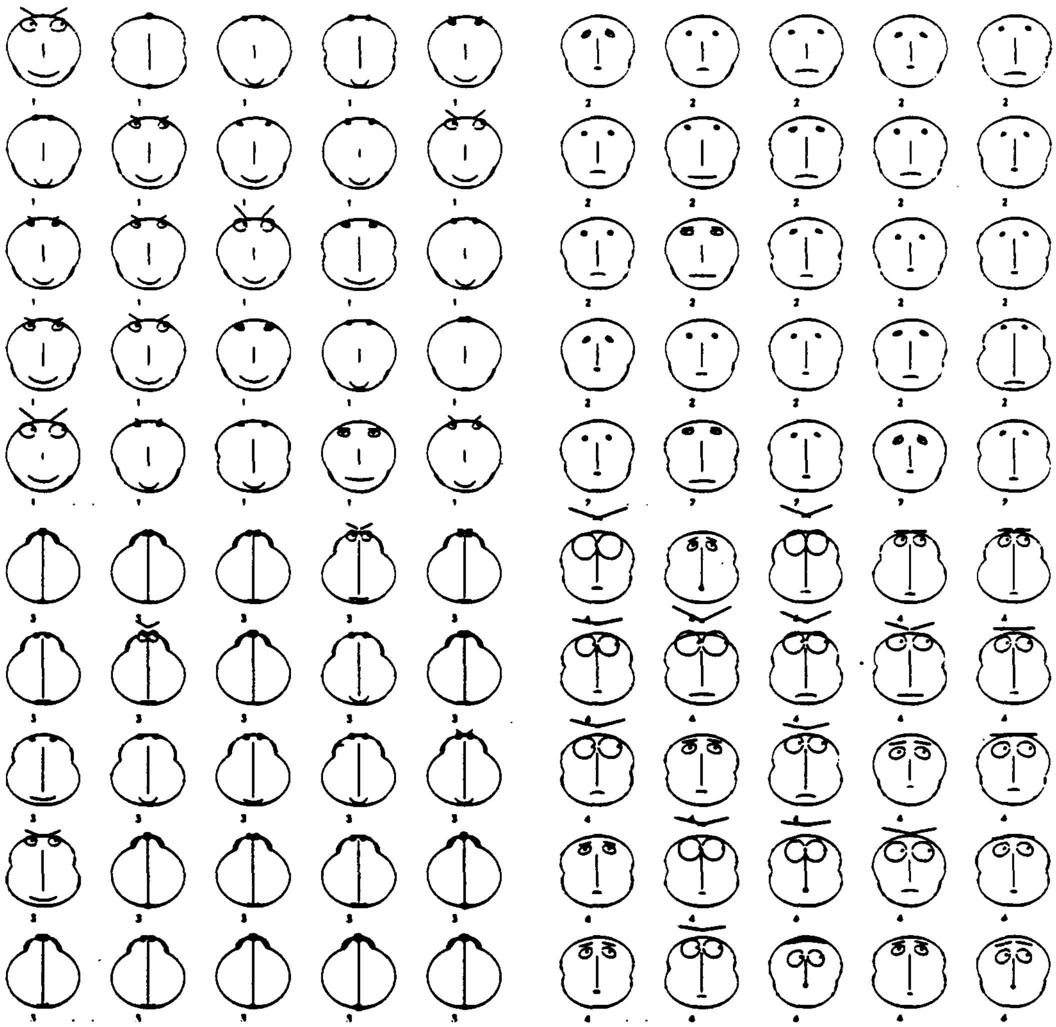

Figure 4.4: Training

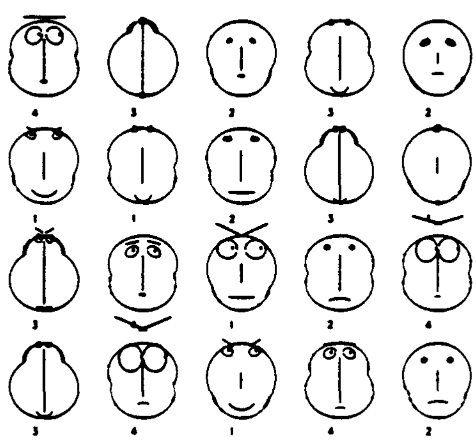

Figure 4.5: Test

FACES

Subject	Score	% correct
1	20/20	100%
2	18/20	90%
3	19/20	95%
4	19/20	95%
5	18/20	90%
6	20/20	100%
7	17/20	85%
8	20/20	100%
9	19/20	95%
10	19/20	95%
11	19/20	95%
12	17/20	85%
13	18/20	90%
14	18/20	90%
15	18/20	90%
16	19/20	95%
17	19/20	95%
18	19/20	95%
19	19/20	95%
20	19/20	95%
21	18/20	90%
22	20/20	100%
23	19/20	95%
24	18/20	90%
25	20/20	100%
26	20/20	100%

Average % correct: 94%

Sound

Subject	Score	% correct
1	18/20	95%
2	19/20	95%
3	27/30	90%
4	29/30	97%
5	30/30	100%

Average % correct: 94%

Table 4.1: Informal γ-ray Spectra Study

In the case of the two-dimensional plots, I defined two modes of audio output. A *chirp* designated a sound plot in which the x-value was encoded as time and the y-value provided the frequency. Thus, as the y-value increased with increasing x, the sound increased in frequency. A *warble* designated a sound plot in which the slope (y/x) provided the frequency base. For example, a linear plot has a constant warble. In general, a log/linear plot was represented by a sequence of notes which varied only in frequency. These representations used constant sound characteristics with the exception of the changing frequency. A two-dimensional plot is then a *sequence* of notes varying in time with the increasing x-value.

Applications

Although I did very little experimentation with either of these two methods, I was encouraged by the resulting sounds. Seismic information provided the data for the multivariate logarithmic data trials. Computer-generated data produced log/linear plots for the two-dimensional trials. In all cases, the sequences of notes differed as expected.

Seismic data

Sound and graphics are useful in illustrating a time history of the earthquakes which occurred in California in the year 1979 [52]. The information for each earthquake is longitude, latitude, depth, magnitude, and start time. The longitude and latitude were used to update a display map with a dot each time an event happened. As the dot appeared, an accompanying sound represented the magnitude. The pitch of a sound was such that a low note (rumble!) indicated a high magnitude quake. A magnitude of 0.0 mapped to 5120 Hz; 8.0 mapped to 20.8 Hz. Magnitude also controlled the volume and duration of the note.

At most, four events (four notes) could be heard simultaneously. Patterns and extremely large events were easily observable. Variables such as depth were not used but could add more information if included in the graphical or sound representation. The trials were positive indications that sound be used to highlight features that are most relevant to seismologists.

Two-dimensional plots

Differences in plots of logarithmic data often appear to be visually the same, when in fact, the relative differences are quite significant. For illustrating this point, plots consisted of scatter points and the corresponding best fit line of the form

$$y = Ae^{Bx}.$$

By varying the values of A and B, different plots were obtained which appeared visually quite similar.

The work consisted of comparing four different aural outputs for each of the plots: a chirp of each scatter plot, a chirp of each best fit line, a warble of each scatter plot, and a warble of each best fit line. For the warble of a scatter plot, the slope at a point (x2,y2) was

$$\frac{(\ \ln(y2) - \ln(y1)\ )}{(\ x2 - x1\ )}$$

where x1 was x2 less the increment in x and y1 was $Ae^{b(x1)}$. The resulting warble for a scatter plot varied up or down in frequency as the scatter plot points deviated above or below the best fit line. The slope of a best fit line was taken to be B (in $y = Ae^{Bx}$). The warble was constant for a given line but varied from one plot to another.

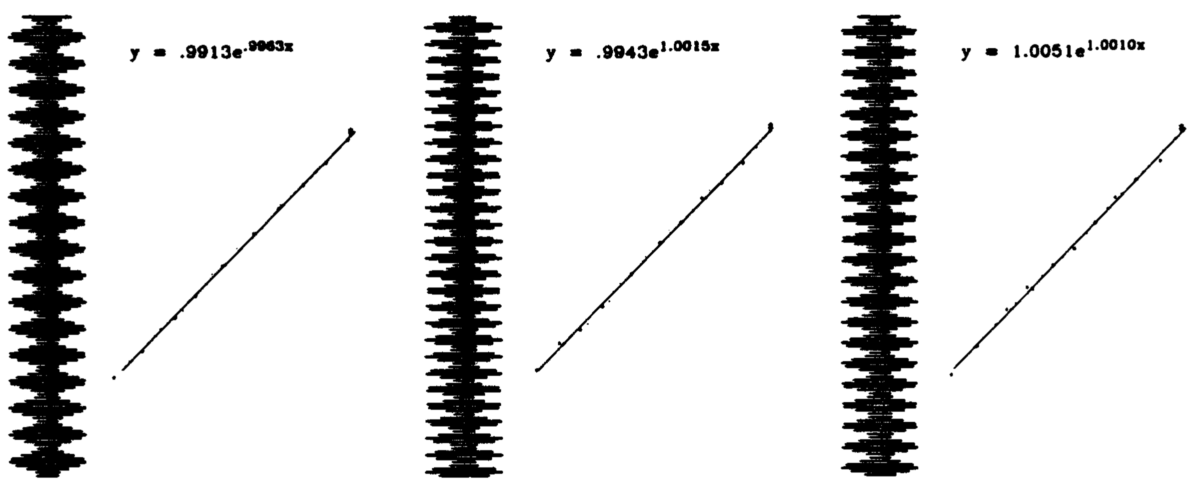

Figure 4.6: Logarithmic Plots

Plots of the data and the corresponding output waveforms are shown in Figure 4.6. Although the two-dimensional plots look alike, the audio output varies significantly as seen by the three different waveshapes. The chirps, either for the best fit lines or for the scatter plots themselves, were distinguishable. As might be expected, the warbles for the scatter plots varied so much within a single plot that it is difficult to make any distinctions from one plot to the next. The most interesting are the warbles for the best fit lines. They are most illustrative of the differences among the actual best fit lines.

4.3: Time-varying Data

Intuitively sound seems particularly appropriate for use with time-varying data despite the success in using sound to discriminate among sets of discrete multivariate data. Given an application in which events vary over time, sound can help highlight changes and relationships.

Encoding Data In Sound

The basic method of encoding the data was that used for the multivariate data work. The values of all variables at a given time step comprised a data sample. The duration was constant from time step to time step. Because time-varying data often involves two or more simultaneous events, a different note represented each event at every time step. However, it is difficult to track simultaneously several different notes if only pitch and volume are varying. Here the waveforms themselves appear to have the most potential for distinguishing among events. If two very distinct waveforms are chosen, then each note can be more easily followed and other sound characteristics noted.

Applications

Professor Sam Parry of the Naval Postgraduate School [67] suggested a time-varying application. Computer battlefield simulations which run from start to finish without human interaction provide information about the state of the battle at each time step. To an analyst interested in the results of the simulated battle, this information is often an overwhelming collection of statistics. Nevertheless, it is important to note the battle characteristics which yield various results. Thus the information at each time step encoded into sound results in a song for each battle. Listening to the songs provides a quick view of the battle in progress and draws attention to critical points during the battle.

For simplicity, only four variables of a two-sided battlefield simulation were used [81]. At every time step, the simulation recorded the number of units at the front and the number of units in transit to the front for sides A and B. Side A of the battle was encoded into a pure tone (a sine wave). Side B of the battle was encoded into a note with very noisy timbre (a randomly defined wave). The number of units at the front mapped into pitch and the number of units in transit mapped into volume. Thus, a high and loud note represents a very strong battlefield position. Because the Side A note did not have the overtones of the Side B note, it was necessary to adjust the volume so that the Side A volume range was somewhat higher than the Side B volume range.

It was possible to listen to both notes simultaneously and therefore follow the progress of each side in the battle. Battles which have similar outcomes do not necessarily produce similar songs. (Listen to the examples in Section 6 of Appendix A).

4.4: Observations

Listeners were enthusiastic and successful in discriminating among different data sets throughout the variety of data applications and sound encoding techniques. The multivariate data applications particularly appealed to users. Those who attempted to distinguish among data sets based on sound representations quickly became adept at the procedures and improved their skills rapidly.

In the battle songs, listeners had no trouble distinguishing one song from another. However, listeners had some difficulty tracking one side relative to the other. This difficulty in hearing the two sides independently did not impair the ability to recognize battles in which significant events occurred. It did make it hard to identify the characteristics (greater front line forces, for example) of the significant events.

The encoding techniques were satisfactory for initial applications of sound presentation. Several combinations of sound characteristics were also tried and subsequently abolished. Combining two or more notes for one discrete data sample seemed an obvious possibility. When eight or more variables were presented, two notes were played simultaneously each with its own frequency, volume, waveshape, and overtones. The resulting chord represented a single data sample but the sound became noisy and much more difficult to classify. Fewer variables might also utilize two or more notes if each variable controls several sound characteristics. For example, one variable might control frequency, volume, and waveshape of one note while a second variable controls frequency, volume, and waveshape of another. Again, the presence of more than one note heard at a time seemed distracting rather than helpful. This may be due to the limited resolution of the sound board waveshapes, the method of varying the timbre, and the serial nature of turning one note on after another.

The interactive system was definitively preferable to the timesharing system with batch output. The turn-around time of the batch system simply did not allow enough experimentation. The sound output could not be varied dynamically so that a wide range of comparisons among notes and sound characteristics was difficult. The inability to vary the waveshape on the batch system (each of the 16 notes was based on a sine wave) was also a hindrance. With either system, I found that intense listening could become tiring after a period of time (30 to 45 minutes).

All the applications of sound encoding for data representation described here indicate the value of sound as a means of information presentation. Each of the three data types offers potential for further study. These positive indications support the desirability of a formal experiment to substantiate the validity of sound representations. Chapter 5 describes the subsequent experiments and their results.

5: EXPERIMENTS

Having explored types of data appropriate for sound representation and ways of encoding .that data into sound, I wanted to demonstrate more concretely the feasibility of using sound to present data information. I ran a set of experiments for two reasons: first, to determine that sound does convey accurate information about the data and second, to examine the potential of using sound to convey more information than the usual means of data analysis. The problems of multivariate data provided a good basis for the experiments because it was possible to construct a straightforward task for volunteers and because other methods of analysis exist for comparison.

Eight separate experiments covered a variety of data set differences and a few alternatives for sound presentation. Phase 1 consisted of three experiments in which one data set was translated, scaled, or contained greater variable correlation relative to a second data set. The positive results of Phase 1 led to a more formal set of experiments in Phase 2. Three experiments made up Phase 2; two data sets were presented with sound only, with graphics only, and with the combination of sound and graphics. Finally, Phase 3 included two additional experiments based on the sound-only experiment of Phase 2. The first changed the mapping of data values to sound characteristics; the second changed the training methods for sound comparisons. Each subject participated in only one of the eight experiments.

For each one of the experiments, two sets of data differed from each other in a well-defined way. The task for each subject was to determine whether each unknown test sample belonged to Set 1 or to Set 2. Since discriminant analysis offers a variety of computational tools, I was able to compare the results with traditional methods of discrimination. The procedures for each phase of the experiment were the same. The following section describes the general procedure for an experiment, followed by detailed explanations of each phase and its results.

5.1: Procedures

The experiment data bases were generated in well-defined ways. Using a multivariate normal random deviate generator, I created two sets of data for each experiment. Each set of data was six-dimensional and contained fifty samples. Following the procedures described for multivariate data in Chapter 4 (Section 4.1), the data for both sets of an experiment was encoded into sound. The sound characteristics corresponding to the six dimensions were pitch, volume, duration, attack, fundamental waveshape, and a fifth harmonic waveshape.

For a particular trial, a subject was told that the task involved listening to sounds which were grouped into two sets. The objective, after listening to a few sounds from each set, was to decide for each subsequent sound whether it belonged in Set 1 or in Set 2. First the subject spent about five minutes learning the mechanics of the facility -- adjusting volume and earphones and learning to use a cursor for interactive input. Left alone, the subject was then presented with ten samples *randomly* selected from Set 1 and ten from Set 2. These were considered training samples and could be replayed as many times as desired before proceeding to the testing. However, once the training was completed, it could not be repeated later. Note that because the training samples were chosen randomly, they varied from subject to subject.

When satisfied with the training, a subject indicated testing was to begin. Samples were played one at a time with no Set 1 or Set 2 identification. The subject's task was to determine for each sample whether it belonged in Set 1 or in Set 2. The subject could repeat the current test sound as many times as desired before making a choice. After the choice was made, the correct set for that sample was displayed. Via this feedback, the subject was further trained. Finally, a group of samples was played, one at a time, with the subject choosing Set 1 or Set 2 for each but receiving no feedback on the correct answer.

For a given trial, the goal of the experiment was for the subject to identify correctly those samples which were in Set 1 and those which were in Set 2. Note that the subjects were given no information about the ways in which the two sets differed. The subjects were not familiar with the sound output; they knew nothing of the data set relationships; and they had no previous experience in using sound for categorizing data in this way. In order to separate samples into the two groups, each subject had to rely on patterns derived solely from listening to the training samples.

For each pair of data sets, results were produced from traditional methods for discriminating sets of data. These results provided a known standard for comparison. One program calculated the principal components of the data, the variance accounted for by considering a subset of the variables, and a plot of any two dimensions of the transformed data (from the principal components). A second discriminant analysis program (linear or quadratic regression) indicated the percentage of unknowns correctly identified given a known training set.

Table 5.1 summarizes the eight experiments. For each trial of an experiment, the variables were 1) the subjects participating, 2) the training samples presented, and 3) the test items presented. The recorded results were 1) the number of test items from Set 1 and the number from Set 2, 2) which samples from each set were incorrectly identified by the subject, and 3) where these samples occurred in the testing sequence. Thus a phase of the experiment consisted of n subjects, each of whom was tested on m items. A subject received twenty training items, ten drawn randomly from each of the two sets of data. The m test items were then drawn *randomly* from the remaining eighty items in the combined sets.

5.2: Phase 1

To establish a basis for the fact that sound does convey information, six-dimensional data bases were generated in which Set 2 variables were translated relative to Set 1, scaled relative to Set 1, and had weak correlation (whereas Set 1 had strong correlation) among the variables. I chose this basis for defining two sets because they

Experiments	# subjects	# training per set	# test items
Phase 1			
Translation (3 trials)	7	10	40 per trial
Scaling (3 trials)	7	10	40 per trial
Correlation	4	10	60
Phase 2			
Graphics only	25	10	40
Sound only	25	10	40
Graphics with sound	25	10	40
Mapping change	20	10	20
Training change	10	10	40

Table 5.1: Experiment Characteristics

could be separated easily by varying amounts (although any two data sets were not necessarily distinct). For simplicity, all variables were equivalent; that is, each dimension in Set 2 was translated, scaled, or correlated by the same amount. Figure 5.1 shows an example of each in which only two of the six dimensions are used for an x-y plot. (Listen to the data set differences in Section 2 of Appendix A).

Translation

Seven subjects used data bases in which Set 2 was translated relative to Set 1. Given two sets of multivariate normal random data with a standard deviation of 1, all sample values in the second set were increased by n. Each subject participated in three different trials: a) Set 2 translated by $n = 3$, b) Set 2 translated by $n = 1$, and c) Set 2 translated by $n = 0.5$. In general, the sounds in Set 2 were higher, louder, longer, and more buzzy than the sounds in Set 1.

The average percentage of correct responses (for twenty samples with and twenty without feedback) was 92% for translation by 3, 70% for translation by 1, and 53% for translation by 0.5. The computer

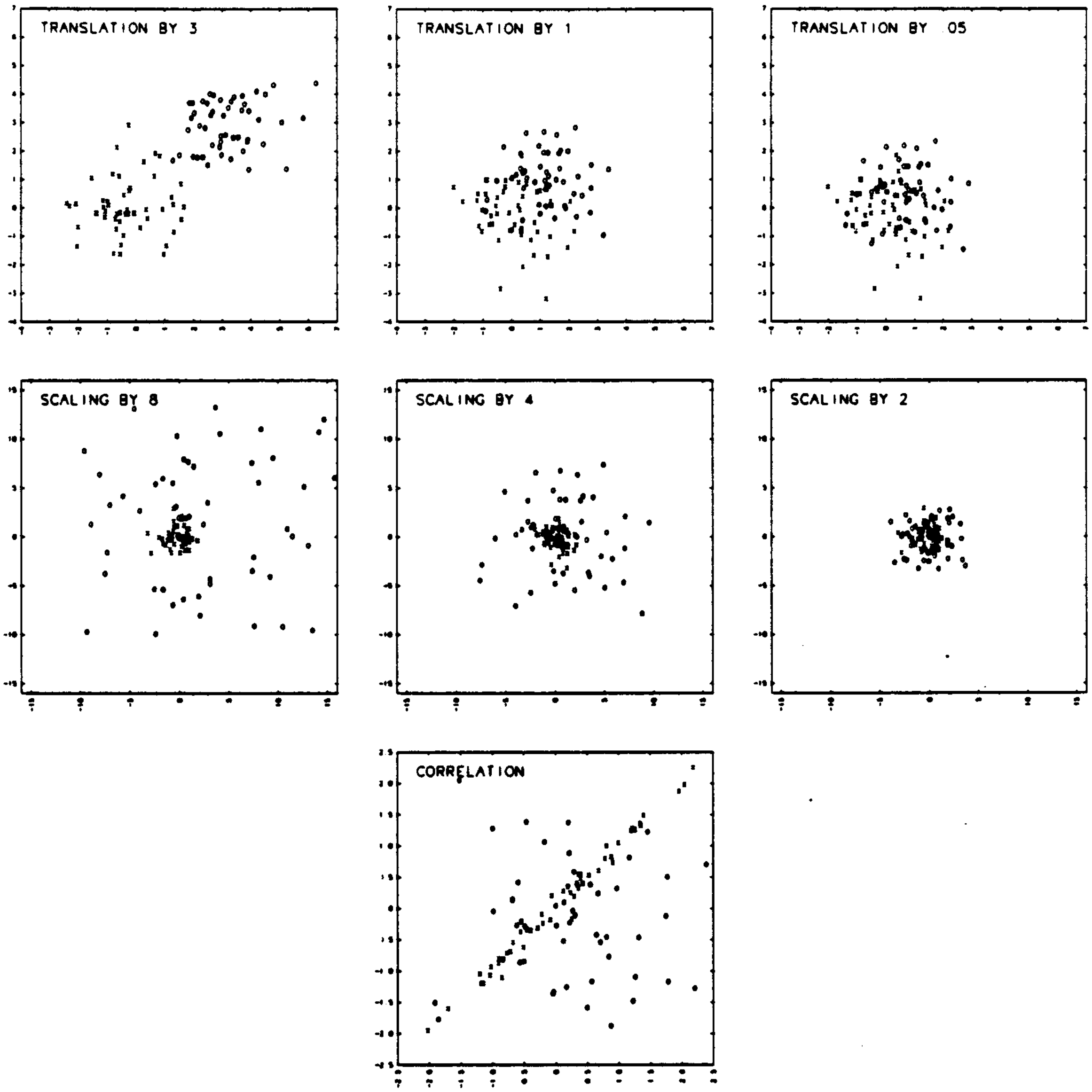

Figure 5.1: Phase 1 Data

discriminant analysis program was given 15 samples identified as Set 1 and 15 identified as Set 2 for known training samples. The program correctly identified 100% of the 40 unknowns for translation by 3, 95% of 20 unknowns for translation by 1, and 70% of 70 unknowns for translation by 0.5. As hoped, the subjects performed in a consistent (though not nearly so successful a) manner as the traditional method. As the separation of the data sets decreased, the ability to distinguish the sets by sound diminished.

Scaling

Seven subjects used data bases in which Set 2 was scaled relative to Set 1. Given a set of multivariate normal random data, for all samples in Set 2, each value was multiplied by n. Each subject participated in two or three different trials: a) Set 2 scaled by $n = 8$, b) Set 2 scaled by $n = 4$, and c) Set 2 scaled by $n = 2$. In general, the sounds in Set 2 were higher or lower, louder or softer, longer or shorter, and more or less buzzy than sounds in Set 1. The Set 1 sounds had mid-range characteristics and all sounded very much alike.

The average percentage of correct responses (for twenty samples with and twenty without feedback) was 69% for scaling by 8, 74% for scaling by 4, and 55% for scaling by 2 (only 2 subjects). For five of the subjects, the third trial was again scaled by 8 instead of 2. In this case, the average was 76.5%. The discriminant analysis program, given 30 knowns, correctly identified 96% of the 70 unknowns for scaling by 8, 90% of 70 unknowns for scaling by 4, and 56% of 70 unknowns for scaling by 2.

In this set of trials, there are two likely implications of the results. One is that scaled data is more difficult to distinguish than translated data, particularly since the subjects had no prior information about the data or the patterns to expect. This is somewhat intuitive, recognizing that Set 1 lies within Set 2 and thus does not have quite so clear a separation. Secondly, there seems to be an effect of training, particularly from scaling by 8 to scaling by 4. The five who did a set scaled by 8, then by 4, and then by 8 again seemed to improve consistently indicating that the differences between scaling by 8 and scaling by 4 were not so great as the differences between first attempting the experiment and attempting the experiment after some practice.

Correlation

Four subjects used data bases in which the variables of Set 1 were strongly correlated. In generating the usual data bases, the covariance matrix for the multivariate normal random generator contained 1's on the diagonal with zeroes elsewhere. For this last part of Phase 1 experiments, the data base for Set 1 was created using a covariance matrix in which all values off the diagonal were .99. Thus, the variables in Set 2 had no strong correlation, while those in set 1 did. In general, the sounds in Set 1 were high, loud, long, and buzzy or low, soft, short, and pure or mid-range pitch, volume, duration and timbre. Set 2 sounds might have any combination of characteristics.

The average percentage of correct responses was 60% for 60 unknowns (all with feedback). The discriminant analysis program, given 30 knowns, correctly identified 96% of the 70 unknowns. This relationship among data seems particularly difficult to distinguish using sound, at least without further training.

	#subjects	#test items	average % correct	analysis program % correct
Translation				
by 3	7	40	92%	100%
by 1	7	40	70%	95%
by 0.5	7	40	53%	70%
Scaling				
by 8	7	40	69%	96%
by 4	7	40	74%	90%
by 2	2	40	55%	56%
repeat by 8	5	40	76.5%	--
Correlation				
.99	4	60	60%	96%

Table 5.2: Experiment Results

Results

The results were consistent with my expectations and are summarized in Table 5.2. Although there were not enough subjects in each group to confirm that the results were significant, the responses are positive. When the two sets are widely separated (and thus easier to discriminate), the number of correct responses appears much better than chance. When the two sets overlap, the successful performance decreases. I was sufficiently convinced that using sound was feasible to try a more complicated experiment.

5.3: Phase 2

During the second phase of the experiment, the intent was to discover if, for some data, sound could add information to other methods. It was necessary to provide a data base whose characteristics were dependent on the multivariate nature of the data so that discrimination by usual methods was more difficult. The experiment then consisted of three different means of presenting the data information -- visually, aurally, and with a combination of graphics and sound.

Data Sets

Once again, the six-dimensional data was obtained from a multivariate normal random deviate generator. A set of 100 samples was generated and then separated into two sets such that a sample, s = (x1, x2, x3, x4, x5, x6), belonged to Set 2 if and only if

$$x2^2 + x3^2 + x4^2 + x5^2 + x6^2 \leq 1.5^2$$

OR

$$x1^2 + x3^2 + x4^2 + x5^2 + x6^2 \leq 1.5^2$$

OR

$$x1^2 + x2^2 + x4^2 + x5^2 + x6^2 \leq 1.5^2$$

Only samples in which all 6 variables had positive values were included. Thus, at least five of the six variables in each sample of Set 2 had value less than 1.5 and at most one of the variables x1, x2, or x3 could have value greater than 1.5. Although these two sets are completely distinct, a projection plot onto any two axes has overlap. Note that the two sets are completely distinct *only* in six-space. In a complete positive and negative space, Set 2 lies somewhat within Set

1.[22] Figure 5.2 shows several plots of two of the six dimensions.

For sound output, x1 was mapped to waveshape, x2 to overtone waveshape, x3 to pitch, x4 to envelope, x5 to duration, and x6 to volume. This mapping was determined arbitrarily by listening to a variety of mappings and subjectively picking that which seemed to distinguish Set 1 from Set 2 most clearly. The sounds in Set 2 were generally shorter and softer than Set 1 sounds because variables x4, x5, and x6 were always less than 1.5. If a Set 2 note was buzzy, then it was low in pitch. When a Set 2 note was high in pitch, it could not be very buzzy.

To establish the visual portion of the experiment, two-dimensional plots were generated for each pair, (xi,xj), of the 100 samples in Sets 1 and 2. (Examples are given in Figure 5.2). I chose, as the graph for visual information, the plot in which Set 1 points seemed most distinct from Set 2 points. As a result, the horizontal axis represented variable x1 and the vertical axis represented variable x5. For graphic experiments, the identification of each point as belonging to Set 1 or to Set 2 was removed and the entire plot was always displayed. Figure 5.3 shows the plot of variable x1 versus variable x5 for the graphic display, and Figure 5.4 shows the plot as presented to the subjects. A particular sample for training or testing was indicated by highlighting the appropriate point.

Seventy-five subjects participated in Phase 2, 25 in each of three groups. Each trial consisted of ten training samples from Set 1 and ten from Set 2. The subject then had forty test samples to identify, twenty samples with feedback and twenty without feedback. The first group of 25 heard each sample, the second group both heard and saw each sample, and the third group only saw each sample.

---

22. Variables x1, x2, and x3 of Set 2 lie within three cylinders, each of radius 1.5 about the first three axes. Variables x4, x5, and x6 of Set 2 lie within a sphere of radius 1.5 about the origin of the last three axes. A Set 1 sample may overlap into either of those two three-space containers but not both.

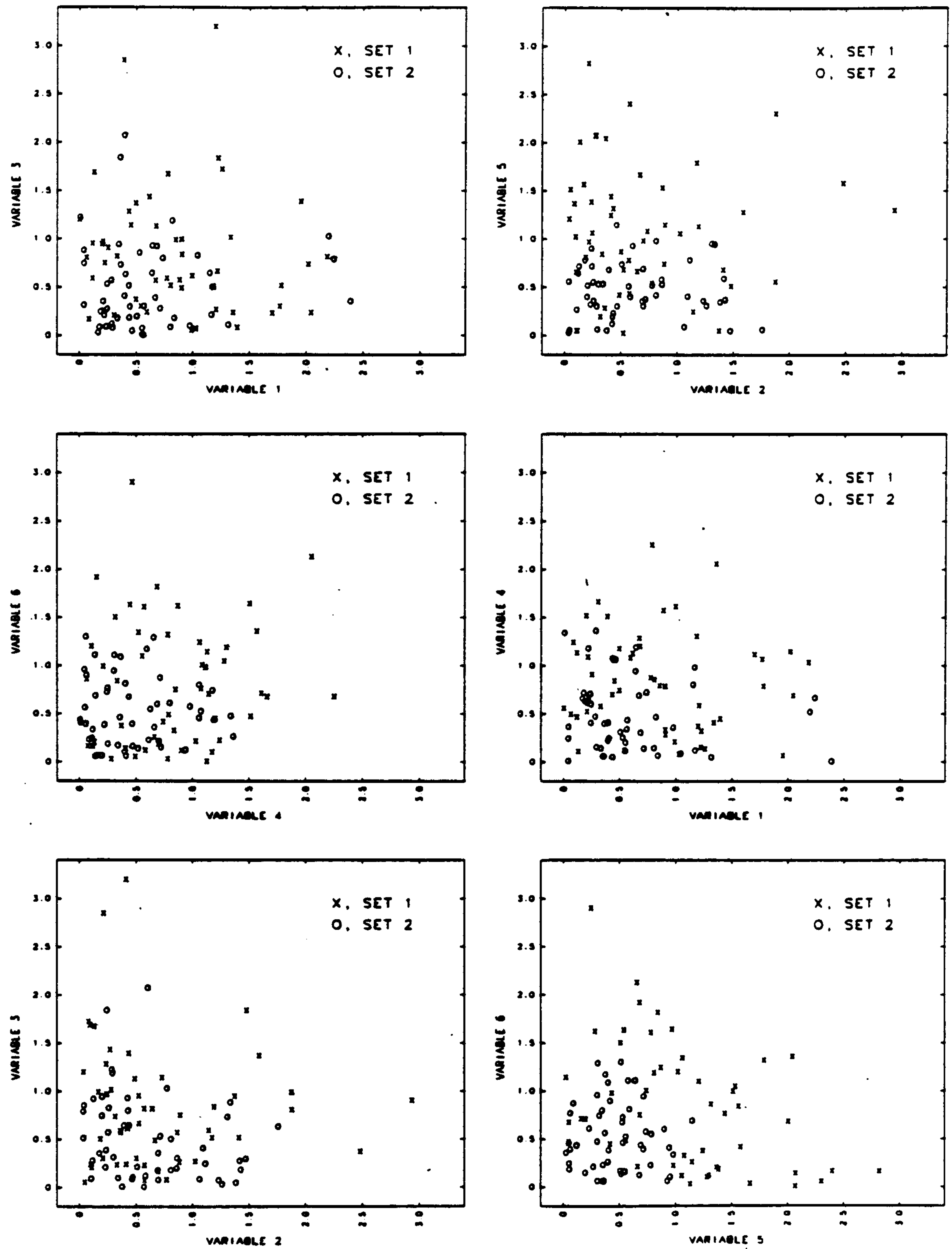

Figure 5.2: Phase 2 Data

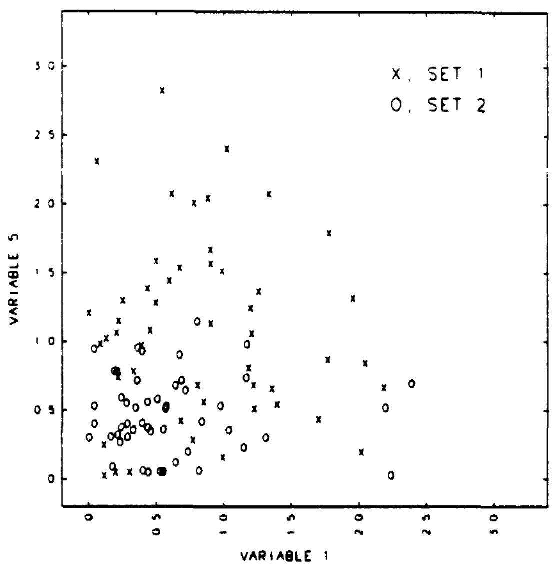

Figure 5.3: Phase 2 Graphics Data

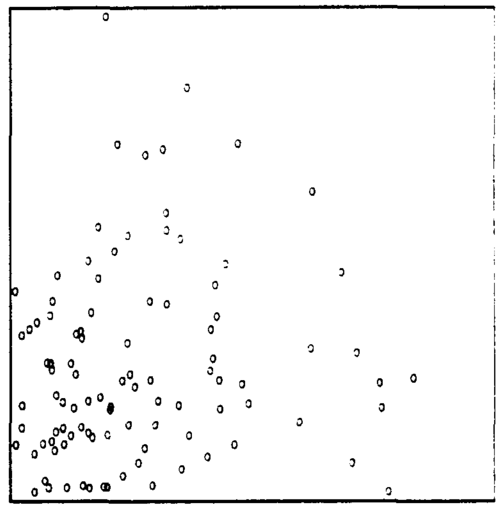

Figure 5.4: Graphics-only Display

Results

Figure 5.5 shows the raw data which has been rank ordered so that subject 1 is the subject who had the fewest number of correctly identified samples and subject 25 is the subject who had the greatest number of correctly identified samples.[23] The average percentage of samples correctly identified was 64.5% for the sound-only presentation, 62% for the graphics-only presentation, and 69% for the presentation combining sound and graphics. Discriminant analysis using quadratic regression offered much better results. Of 60 unknowns, 87% were correctly identified after a training set of 40 knowns.

In order to support the inference that sound was useful in discriminating between data sets, several hypotheses about the experiment results were checked statisically (see Appendix C). The

---

23. The curve representing random choice was determined by a psuedo experiment in which the response for each test item was a random draw of Set 1 or Set 2. This psuedo-experiment randomly drew 40 test items for each of 100 trials.

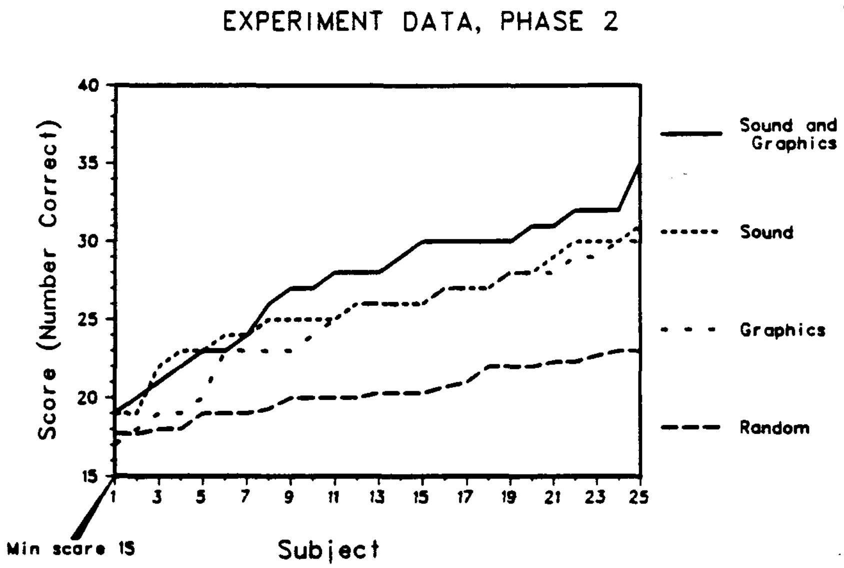

Figure 5.5: Results

experiment results are positive, but it was worth questioning whether the results could have occurred by chance and whether the results did differ among the three experiments. The calculations show at a 99.9% confidence level that the participants were not responding by guessing. The second hypothesis was that the performance in each of the three groups did not differ from group to group. This hypothesis was rejected at a 95% confidence level. The performance scores of participants in the group receiving the combined graphics and sound presentation were higher than those in the graphics-only presentation. The sound-only results were not significantly different than the graphics-only results. Overall these results verify that sound indeed provides information about multivariate data when it is presented to human analysts.

The results support two observations. First, it seems that poor performance was equivalent in any one of the three methods. Second, a combination of sound and graphics provided an easier discrimination task for subjects who did well than either sound alone or graphics alone. Furthermore, sound alone seems to be as good a method for discrimination as the two-dimensional graphics. This seems particularly interesting knowing that subjects are usually familiar with x-y plots but were totally unfamiliar with using sound for data representation.

As expected, the experiments suggested several unanswered questions which deserve further study.

1: How does one determine the mapping from data variables to sound parameters?

2: How many sound parameters (and thus how many variables) can be utilized?

3: Which sound parameters convey the most information?

4: How does the interaction among sound parameters affect the information presented?

5: Does a learning curve exist? What is the best performance that can be expected?

6: How does one provide a reference basis for comparing sounds to a given standard or to known data?

7: What types of data are best suited for sound encoding?

8: How can sound and graphics be used together for information presentation?

Although it is possible that sound may provide more information about some data than traditional methods, such comparisons were not made. In particular, no attempt was made to utilize the various *graphical* methods for presenting multivariate data. The potential for combining sound with graphics is exciting. However, further understanding of the use of sound itself is critical. Thus, the first questions to address are those concerning sound alone.

5.4: Phase 3

Additional experiments provided initial observations in two areas suggested above. One experiment used a different mapping between data values and sound parameters. The other expanded the training for subjects. In both cases, I used the same data and procedures as for the sound-only experiment in Phase 2. The mapping change provided no noticeable change in responses; the training change increased the subjects' scores.

Mapping Change

Twenty subjects (who had not participated in Phase 1 or Phase 2 experiments) were tested with a different mapping of variables to sound parameters. The mapping from data values to sound characteristics determines the nature of the sounds. If one or two data variables provide more of the data set variance than other variables, then it would be wise to incorporate that knowledge in defining the mapping. The task of data discrimination based on sound presentation would be more straightforward if a procedure could be found for selecting the optimal mapping from data values to sound.

A new mapping was chosen based on the output of a traditional discriminant analysis program. The Phase 2 data sets described in Section 5.3 were processed using a principal components analysis. This offered some basis for ranking the importance of the contribution of each variable as x2, x1, x3, x6, x4, and x5. Likewise, listening to sound variations offers some basis for ranking the characteristics in importance as pitch, duration, waveshape, volume, envelope, and overtone. Thus, for a change of mapping, x2 was mapped to pitch, x1 to duration, x3 to waveshape, x6 to volume, x4 to envelope, and x5 to overtone.

The experiment procedures were like those of Phase 2 except that only twenty subjects participated and only twenty samples (all with feedback) were used as test items. Of the twenty subjects who participated in the experiment with a change of mapping, the average number of correctly identified samples was 62%. The results do not indicate a significant difference with this new mapping.

Training Change

A subject's success might increase with additional training since each participant in the original Phase 2 experiment was exposed to both unfamiliar equipment and unfamiliar tasks with only 15 to 20 minutes learning and testing. Also, it appeared that a subject performed better with constant feedback (based on the difference in the number of samples correctly identified when feedback was given versus the number of samples correctly identified when no feedback was given). Thus, some method of providing a continuous reference base could be helpful in improving performance.

Another short experiment was run to indicate whether increased training and available reference sounds would increase the number of correctly identified samples. I used the same experiment data as used in the Phase 2 sound-only group. Fifteen subjects were randomly selected from the 25 who participated in the sound-only portion of the original Phase 2 experiment. Therefore, the participants in this experiment had previous experience with the sound presentation. Of these, ten subjects repeated the sound-only experiment with the same data base and mapping but with modified training.

Training

The following changes to the experiment procedure were added to improve the training.

1: A participant could listen to each of the six characteristics of sound varied one at a time while the other five were held constant. This demonstrated the range of each parameter for the listener. (Listen to the first section in Appendix A).

2: When the ten training samples of each of the two sets were played, a subject could listen to any subset in any order, rather than always playing the twenty training samples sequentially. This prevented any grouping of notes based on a sequence pattern.

3: A subject could refer to the training set at any time during the experiment. This gave the subject a reference point for review, much as an x-y plot provides a constant reference on the axes for relative positioning.

4: Since the data sets used in the Phase 1 experiment were ones which varied in separation, they offered a convenient base of data sets for providing experience in identifying data samples as sound. I ran short sequences of the translation, scaling, and correlation sets with follow-up explanations of that data. This exercise provided each subject with a variety of experiences in using sound for discriminating sets of data.

The testing procedure was the same as that of the Phase 2 experiment.

Results

Subjectively, the most noticeable result was the attitude of the participants. In general, they felt successful and enjoyed the exercise more than they had the first time. Most were enthusiastic about their ability to separate sets of data via sound and were willing to participate in further trials.

The ability to refer to the training sets at any time during the experiment was particularly useful. Since sound is transient by nature, subjects were able to refresh their memories on the differences between Set 1 and Set 2 sounds. In fact, often a subject repeatedly cycled between listening to a test sample and listening to one or more training samples.

The average number of correct responses was 29.5 of the 40 test items (73.8%) compared to a 25.8 average (64.5%) originally. One can conclude that there was a significant improvement in the scores. Unfortunately, there was a flaw in the experment. Training items could be repeated as test items in this particular experiment (although the subjects believed there was no overlap between the training set and the test items). Thus it is possible that the improved scores resulted in part from the occasional overlap of training and test items. Consequently, any statistical analysis is meaningless. In observing the participants, I do not believe that the overlap greatly influenced their behavior. I believe that the improved training did in fact improve performance.

5.5: Observations

The experiments confirmed the hypothesis that sound is useful in data presentation. Subjects were successful in identifying sounds which belonged in one of two sets. Their performances varied consistently with the difficulty of the task as determined by the difference in the two sets. It is particularly important that the performance improved as the subjects became more familiar with the procedures and with a variety of data set differences encoded into sound. The results in using a combination of sound and graphics are especially positive and indicate a technique for improving current methods of presenting information.

The subjects were not requested to make judgements about the nature of the data. Generally data samples were discussed in terms of single sound characteristics; higher or lower, louder or softer, longer or shorter. The results of Phase 1 show that translation differences were easier to recognize than other methods of separating data sets. That is, when one group of sounds was consistently higher or louder or longer than another, subjects discovered the pattern readily. The correlation of sound characteristics was difficult to perceive (listen to the data in Section 2 of Appendix A), although a visual correlation of variables is recognizable (look again at the same correlation data in Figures 1.10 and 5.1). With practice, subjects did improve in hearing the variable correlation in sound. No attempt was made with any of the data to have subjects derive any absolute information about that data. Even in a relative sense, subjects tended to think of only one dimension at a time ("This note is louder than that note", or "The first note was shorter than the second note"), although I am convinced they heard and discriminated on the gestalt.

6: SUMMARY

Computer-generated sound is one solution to the existing need for additional methods of information presentation. The increasing use of computers provides an abundance of data output. To be useful, this data must be presented in ways which are easily grasped by analysts. Computer graphics offers a variety of ways to display information but is limited to one sensory input and in the number of dimensions available for encoding. Sound is useful in daily life and already has a well-understood language. Data can be encoded into sounds so that information is preserved. This sound encoding provides a new tool for exploratory data analysis. The possibilities for future research are vast.

The next two sections summarize the conclusions of the work described in previous chapters and outline several topics for further exploration.

6.1: Conclusions

This thesis demonstrates that sound is useful for computer information presentation. The work describes successful attempts in encoding different types of data into sound. A formal experiment verifies that sound representation of data is useful. Finally, sound adds information when presented in combination with graphics.

Several types of data are appropriate for sound representation. The studies described in Chapter 4 provide a variety of data types which are particularly natural for encoding into sound. Listeners were able to discover differences and patterns in logarithmic, time-varying, and discrete multivariate data. Earlier work of Yeung, Mathews, and Wilson indicates the value of using sound to present multivariate data samples and patterns in table data. The studies presented here expand that base to include other data types which are appropriate for sound.

Sound is a useful discriminant in classifying data samples. The experiments explained in Chapter 5 validate the hypothesis that sound can be used for information presentation. In the Phase 2 experiments, subjects did classify data based on sounds and sound encoding was as good for discrimination as a two-dimensional graphics plot. These results provide formal verification of the ideas suggested by the initial studies in Chapter 4 and by the observations of Yeung, Mathews, and Wilson.

A combined presentation of graphics and sound is more useful than two-dimensional graphics alone for classifying six-dimensional data samples. Subjects performed better when discriminating between two sets of six-dimensional data given information in sound combined with a two-dimensional plot than when given graphics alone.

6.2: Future Exploration

The results thus far indicate that sound offers an exciting new means for computer information presentation. The work that has been done is only a small beginning. A better understanding of both applications and techniques is needed to build a solid foundation for the use of sound. Given the methods described already, many data problems could be examined and encoded. At the same time, these applications and techniques are only a few of the possible implementations.

This section reviews types of data that are particularly applicable for sound encoding and describes methods that deserve further attention for audio data presentation. A value of any technique is its ability to aid an application. Recognition of a broad range of applications and techniques will confirm that sound is an effective tool for human/machine communication.

Applications

The applications for sound presentation involve not only the type of data to be presented but also the audience and the interface to that audience. I have discussed briefly a few types of problems for which sound could be useful. However, one of the greatest difficulties I have faced is the acceptance that sounds are a valid means of presenting data. Particularly when exploring *scientific* data, analysts are accustomed to visual representations. It is obvious that sound presentation of data broadens the base of information available to those who cannot see or who are in tasks which already require visual attention. However, it is not so intuitive that sound can adequately present scientific information for general use. The burden is on the research in sound presentation to show applications in which sound does convey useful information and to find ways of introducing the sound presentation to listeners.

One way of demonstrating the feasibility of sound for data use is to emphasize applications in which the sound encoding provides *more* information about the data than traditional methods alone. I have already discussed multivariate data in great detail. I believe that the area is well worth pursuing. Now I would like to focus more attention on audio cues, logarithmic data, and time-varying data. The work described in Chapter 4 already shows the value of sound representation for these areas of data presentation. I would like to suggest applications of these data types which are particularly suited for exploring sound encoding.

Audio Cues

One of the most straightforward applications of presenting information in sound is that of audio cues. I define an audio cue as a single sound of short duration which is used to focus the listener's attention. Computer terminals have bells which signal requests for user input or task completion. It would be a simple matter to extend this idea to more complex situations.

The use of audio cues seems most appropriate for interactive applications with graphic output when the user is involved and visual attention is captured. One such example is any application in which the contents of the display are very complex and rapidly changing. An audio cue would be especially helpful every time critical events occurred (such as explosions in a battlefield simulation). The audio cue is essential when the application is such that the display does not contain all the events at any one time. An audio cue can signal those events which occur off the current display and would not be seen. Similarly, an audio cue might be used to track a single event such as temperature in an application of weather monitoring. These are not technically difficult concepts but their potential helpfulness has not been fully utilized.

Robert Lee used audio cues in a film portraying a two-step laser isotope separation process [46]. A gas of mixed U-235 and U-238 was exposed to a laser that excited the U-235. A beep corresponded to each isotope excitation. Subsequently, a second laser ionized those isotopes in the excited state so that they were attracted to a negative plate. For each ionization, a tone was heard which lasted until the ionized isotope reached the negative plate. If more than one isotope was ionized, then more than one tone was heard. Graphics displayed the motion of all isotopes. An isotope raised to an excited state was enlarged; an ionized isotope immediately fell toward the negative plate. However, the many events occurring on the display often distracted the observers' attention away from other critical events. The audio cues were valuable in drawing the observer's attention to the isotope excitation and the subsequent ionization. Furthermore, the observer heard sequences of events without having to scan the display for rapidly changing situations.

Logarithmic Data

The examples of logarithmic data encoded into sound were encouraging. Although the sound encoding did not necessarily use more information than the usual logarithmic plots, it produced more noticeable differences and relationships in the data. An analyst obtains information more quickly or finds information that was not previously obvious.

Often analysts are searching for patterns in the behavior of logarithmic data. Such tasks almost always are aided by a different perspective of the data. Since sound has inherent logarithmic characteristics, I believe it adds new insight to the data.

Time-varying Data

Like logarithmic data, time-varying data deserves additional follow-up to the work described earlier. In the battlefield examples with similar outcomes, songs often varied noticeably indicating that very different events led to the same conclusions. Sound allows a relatively short capsule replay of information which varied over time. In such replays, it is easy to notice critical events within a song or vast differences among songs.

Whereas the audio cues are most suggestive of interactive data analysis, sound for time-varying data has potential for reviewing batch data analysis. I would like to explore further applications in which the number of interacting variables is at least four or five and for which the analyst involvement is one of looking back at intermediate actions and final results. I believe that encoding the events in sound provides a helpful tool for quick review of the entire problem.

Techniques

Whatever the available applications for using sound, the techniques are critical for insuring that the sound adequately describes the data. The next sections describe five areas relating to the techniques for using sound to present information. Two are concerned with the mechanics of presenting the information in sound: one, the mapping between data variables and sound parameters and two, the need for a constant reference base of sounds. The third area touches on the need to understand better our perception of sound. The other two areas suggest additional characteristics of sound for encoding data. Timbre and stereo/location are two aspects of sound which could be utilized to provide stronger recognition of data.

Variable-To-Sound Mapping

Several considerations arise when encoding a data set into sound. In particular, one must determine which data variables map to which sound characteristics. My methods were sufficient to establish sound as a useful means of presenting data. More sophisticated methods are needed for widespread use.

One issue to address is how many characteristics (dimensions) of sound can be utilized. Yeung [99] suggests that up to twenty are possible. My own experience indicates that probably only four or five will actually add information. Discovering guidelines for the amount of information that can be encoded into sound will help determine how other factors of sound encoding can be optimized.

A second consideration is the data itself. In many cases, some of the data variables may not be independent. This fact may or may not be known at the time of initial analysis. For example, the knowledge that two variables vary inversely would be useful in determining which sound characteristics to assign to those variables.

A third consideration is the relationship among sound parameters. Some aspects of sound are more significant than others. Additionally, some aspects of sound work together. Because perceived volume decreases as pitch increases, it is not clear that two independent variables can be meaningfully mapped into pitch and volume. One solution is to base the volume on the value of the pitch. Using sound for information presentation increases the need for a good understanding of psycho-acoustics.

It would be ideal if a general algorithm existed which would find the mapping between a data set and sound characteristics. It is not likely that one such procedure will satisfy all applications. Rather, further exploration will determine guidelines for satisfactory use of sound to encode data.

Sound Reference Base

One of the most significant changes reported by most subjects who participated in the Phase 3 experiment with additional training was the ability to refer to the training sets at any time. Just as we often make visual decisions by comparing two or more objects, discriminating among sounds was easier when a comparision was available. Other reference bases should be explored.

One suggestion by an experiment subject was that the range of each sound parameter being used in the information presentation be constantly available. For example, if pitch, volume, duration and waveshape are varying, then provide access to the low, median and high values for each. That is, at any time, a user could listen to the low volume or the median duration or the high value waveshape as a reference for the current sound being heard.

The reference base becomes more difficult when the data is time varying. Because there is no pause between notes, there is no opportunity for reference to other sets and standards of sound. It is conceivable that a constant note could accompany the data sound and thus provide a base much like an axis runs past plots. However, such an approach immediately complicates the output sound.

Resolution In Sound

The resolution of sounds affects the actual information that analysts derive from sound presentations. Graphics offers very good spatial resolution that provides absolute distances for data value separation. Color differences in hue, saturation, and intensity are not so easily quantified. Two colors which are very close may look alike, and it is difficult to determine anything other than relative data information (i.e., one color or data item is different than another color or data item). Sound resolution offers similar problems.

First, how different do two sounds have to be in order to be classified by a listener as different? Many timbre differences are noticeable only if the duration of the note exceeds some minimum time. For example, two different attacks will not be effective if the duration of the notes is less than the length of the attacks.

Second, how do differences among sounds translate into information about the data? Subjects can distinguish between data sounds and correctly classify a sound with others of similar characteristics. In a data plot, one can determine absolute object sizes and distances among samples. It is not clear that such absolute information is available from sound. Graphically, one would expect also to notice facts about the various sets such as their relative masses or shapes. Sound may provide similar information (in the Phase 1 scaling experiment, subjects did recognize a set which was within another set), but experience is needed to understand better the methods.

Timbre

The effect of timbre is certainly one of the most significant characteristics of sound. For example, note how easily one can distinguish between a trumpet and a clarinet at the same pitch and volume. This ability to discriminate based on the shape of the sound needs further study to realize its potential.

The difference between the sound of a clarinet and that of a trumpet can be shown by examining their respective waveshapes [87]. In a clarinet, the even harmonics tend to be suppressed. The resulting sound tends to be like that of a square wave. In a trumpet, the higher the frequency harmonic, the later it appears in the tone. That is, higher-numbered harmonics do not rise to their steady-state values as quickly as lower-numbered harmonics.

The encoding methods described in Chapters 4 and 5 used the waveshape in a very straightforward and computationally simple manner. Beginning with a pure sine wave defined by 128 values, random values were chosen and substituted for random positions in the sine function until the shape of the sine was lost and a buzz resulted. This method indeed provided a varying sound but did not address the full range of possible discriminations based on timbre. At least two approaches are worth consideration. The first is that of altering the waveshape itself by changing discrete values. A second is to create the waveshape by adding components. More control of the waveshape allows more noticeable note differentiation.

Varying from a sine wave to a random buzz is one example of the first approach. Use of various functions, such as sine, square, or sawtooth (refer to Section 2.1), is another. One possible alternative is to move from a given waveshape (say, that of a clarinet) to some other (say, that of a trumpet). Such interpolation needs careful thought, especially to insure a smooth and even transition.

The second approach suggests creating a waveshape and directly controlling the overtones for each data value. This offers an additional capability, particularly to users with some musical training. Building a waveshape of chosen overtones controls the harmony or dissonance of the sound. The amount of dissonance becomes an additional sound parameter. If two or more notes are being used simultaneously, harmonic variations may be a more meaningful way to vary their respective pitches than by an absolute mapping of variable values directly to pitch.

Stereo

Stereo sound can be achieved quite easily given at least two output sound signals and headphones or dual speakers. Chowning [15] has gone a step further by examining the spatial orientation of sound. Thus another aspect of sound offers itself for use in presenting data.

For a single note encoding, stereo might be used to represent one of the variable values. For example, in sending one note to two speakers, the stereo value might determine the amplitudes of each signal going to its respective speaker.

Location variation is perhaps more straightforward though less easily implemented. The value of the variable mapped to location determines the location of the sound. This capability has a direct correspondence to graphical plotting. It seems particularly applicable to three-dimensional data since location is considered in a three-dimensional space. The comparision of three-dimensional plots with three-dimensional sounds is intriguing.

BIBLIOGRAPHY

[1] Ahmed, N. and K.R. Rao, *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag, New York, 1975.

[2] Allan, J.J. and A.M. Chiu, "An Effectiveness Study of a CAD System Augmented by Audio Feedback", *Computers and Graphics*, Vol. 2, p. 231, 1977.

[3] Allen, J., "Synthesis of Speech from Unrestricted Text", *Proceedings of the IEEE*, Vol. 64, No. 4, 1976.

[4] Andrews, D.F., "Plots of High-dimensional Data", *Biometrics*, Vol. 28, p. 125, 1972.

Andrews' plots are a particularly interesting method of presentng multi-dimensional data. The example in Chapter 1 is based on this article.

[5] Andrews, D.F., "Graphical Techniques for High Dimensional Data", *Discriminant Analysis and Applications*, T. Cacoullos, editor, Academic Press, New York, 1973.

[6] Ashton, A.C., *Electronics, Music, and Computers*, Ph.D. dissertation, University of Utah, Salt Lake City, UTEC-CSc-71-117, 1971.

[7] Babbitt, M., "An Introduction to the R.C.A. Synthesizer", *The Journal of Music Theory*, Vol. 8, No. 2, p. 251, 1964.

[8] Barbour, A., "An Investigation of Dual Sensory Stimulus Presentation of Complex Noise-like Sounds" Pennsylvania State University, Technical Memorandum File No. TM 79-155, 1979.

[9] Beauchamp, J., and J. Melby, editors, *Proceedings of the Second Annual Music Computation Conference*, 1975.

[10] Beniger, J.R. and D.L. Robyn, "Quantitative Graphics in Statistics: A Brief History", *The American Statistician*, Vol. 32, No. 1, 1978.

[11] Beranek, L.L., "Digital Synthesis of Speech and Music", *IEEE Transactions on Audio and Electroacoustics*, Vol. AU-18, No. 4, p. 426, 1970.

[12] Blesser, B.A., K. Baeder, and R. Zaorski, "A Real-Time Digital Computer for Simulating Audio Systems", *Journal of the Audio Engineering Society*, Vol. 23, No. 9, p. 698, 1975.

[13] Chernoff, H., "The Use of Faces to Represent Points in k-Dimensional Space Graphically", *Journal of the American Statistical Association*, Vol. 68, No. 342, p. 361, 1973.

Chernoff's FACES are a unique method for representing multivariate data. His paper includes several examples which illustrate the ability of humans to differentiate easily among faces.

[14] Chernoff, H. and M.H. Rizvi, "Effect on Classification Error of Random Permutation of Features in Representing Multivariate Data by Faces", *Journal of the American Statistical Association*, Vol. 70, No. 3, p. 548, 1975.

[15] Chowning, J.M., "The Simulation of Moving Sound Sources", *Journal of the Audio Engineering Society*, Vol. 19, No. 1, p. 2, 1971.

This article is especially relevant for studying the perception of sounds. In particular, Chowning discusses the synthesis of sound which has the perceived effect of motion.

[16] Chowning, J.M., "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation", *Journal of the Audio Engineering Society*, Vol. 21, No. 7, p. 526, 1973.

Chowning has developed a frequency modulation equation for sound synthesis which proves quite successful for duplicating natural sounds, as well as for creating new ones. His equation is the basis for several digital sound systems which run in real-time for a limited number of voices. One of his equations appears in Chapter 2 under direct synthesis.

[17] Chowning, J.M. and R.K. Clark, "A Program for the Real-Time Generation of Musical Sounds", *Journal of the Audio Engineering Society*, Vol. 14, No. 1, p. 21. 1966.

Although the system does not run in real-time in the sense of immediate audio feedback, the paper includes an outline of the actual FORTRAN program.

[18] Chowning, J.M., J.M. Grey, L. Rush, and J.A. Moorer, "Computer Simulation of Music Instrument Tones in Reverberant Environments", Dept. of Music, Stanford University, Report No. STAN-M-1, 1974.

[19] Colquhoun, W.P., "Evaluation of Auditory, Visual, and Dual-Mode Displays for Prolonged Sonar Monitoring in Repeated Sessions", *Human Factors*, Vol. 17, No. 5, p. 425, 1975.

[20] Divilbiss, J.L., "The Real-Time Generation of Music with a Digital Computer", *The Journal of Music Theory*, Vol. 8, No. 1, p. 99, 1964.

Divilbiss uses a hybrid system and includes a nice discussion of Mathews' work.

[21] Dorf, R.H., *Electronic Musical Instruments*, RADIOFILE, New York, 1968.

[22] Dworak, P., A.C. Parker, and R. Blum, "The Design and Implementation of a Real-Time Sound Generation System", *4th Annual Symposium on Computer Architecture*, p. 153, 1977.

[23] Everitt, B.S., *Cluster Analysis*, Heinemann Educational Books Ltd, London, 1974.

[24] Everitt, B.S., *Graphical Techniques for Multivariate Data*, Heinemann Educational Books Ltd, London, 1978.

This book is an excellent summary of several graphical methods, all clearly explained. In addition, there are some helpful comments on exploratory data analysis in general.

[25] Everitt, B.S. and P. Nicholls, "Visual Techniques for Representing Multivariate Data", *The Statistician*, Vol. 24, No. 1, p. 37, 1975.

The paper contains good descriptions, examples and comparisons of three techniques: Non-linear Mapping, Andrews Function plotting method, and FACES.

[26] Farnsworth, P.R., *The Social Psychology of Music*, The Iowa State University Press, 1969.

[27] Fienberg, S.E. "Graphical Methods in Statistics", *The American Statistician*, Vol. 33, No. 4, p. 165, 1979.

[28] Fisher, R.A., "The Use of Multiple Measurements in Taxonomic Problems", *Annals of Eugenics*, Vol. 7, p. 179, 1936.

[29] Fisherkeller, M.A., J.H. Friedman, and J.W. Tukey, "PRIM-9 An Interactive Multidimensional Data Display and Analysis System", Stanford Linear Accelerator Center, Stanford, Calif., SLAC-PUB-1408, 1974.

[30] Freedman, M.D., "A Digital Computer for the Electronic Music Studio", *Journal of the Audio Engineering Society*, Vol. 15, No. 1, p. 43, 1967.

A good article which includes some timing statistics.

[31] Gabura, J. and G. Ciamaga, "Computer Control of Sound Apparatus for Electronic Music", *Journal of the Audio Engineering Society*, Vol. 16, No. 1, p. 49, 1968.

[32] Geyer, K.E. and K.R. Wilson, "Computing With Feeling", *Proceedings of the Conference On Computer Graphics, Pattern Recognition, and Data Sructure*, IEEE Catalog No. 75CH0981-1C, p. 343, 1975.

The authors discuss touch communication systems for increasing perception of spatial awareness.

[33] Gnanadesikan, R., *Methods for Statistical Data Analysis of Multivariate Observations*, John Wiley and Sons, 1977.

This contains a good discussion of the multivariate problems of classification and clustering. There are several complete examples.

[34] Gold, B., "Digital Speech Networks", *Proceedings of the IEEE*, Vol. 65, No. 12, 1977.

[35] Greer, W.H., "Monaural Sensitivity to Dispersion in Impulses and Speech", University of Utah, Salt Lake City, UTEC-CSc-75-142, 1975.

[36] Grotch, S.L., "Interactive Multivariate Graphics: Shapes that Inform", *Energy and Technology Review*, Lawrence Livermore National Laboratory, Livermore, Calif., 1981.

[37] Grotch, S.L., Lawrence Livermore National Laboratory, Livermore, Calif., private communications, 1980-1982.

[38] Hiller, L.A., "Electronic Music at the University of Illinois", *The Journal of Music Theory*, Vol. 7, No. 1, p. 99. 1963.

Hiller includes some interesting information on waveforms.

[39] Hiller, L.A., "Musical Applications of Electronic Digital Computers", *Gravesaner Blatter*, Vol. 7, No. 27-28, p. 62, 1965.

[40] Hiller, L. and P. Ruiz, "Synthesizing Musical Sounds by Solving the Wave Equation for Vibrating Objects: Part 1", *Journal of the Audio Engineering Society*, Vol. 19, No. 6, p. 462, 1971.

[41] Hofmann, M.A. and N.W. Heimstra, "Tracking Performance with Visual, Auditory, or Electrocutaneous Displays", *Human Factors*, Vol. 14, No. 2, p. 131, 1972.

[42] Hutchins, B.A., "Application of a Real-Time Hadamard Transform Network to Sound Synthesis", *Journal of the Audio Engineering Society*, Vol. 23, No. 7, p. 558, 1975.

[43] Hutchins, B.A., "Experimental Electronic Music Devices Employing Walsh Functions", *Journal of the Audio Engineering Society*, Vol. 21, No. 8, p. 640, 1973.

[44] Julesz, B., *Foundations of Cyclopean Perception*, The University of Chicago Press, Chicago, Illinois, 1971.

[45] Lee, R.M., "Recording Motion Picture Sound Tracks Using a Computer Output Film Recorder", *Computer Graphics and Image Processing*, Vol. 5, p. 41, 1976.

[46] Lee, R.M. and R.A. Riviello, "Laser Separation of Uranium Isotopes", Lawrence Livermore National Laboratory, Livermore, Calif., film, 1975.

[47] Lincoln, H.B., editor, *The Computer and Music*, Cornell University Press, Ithaca, New York, 1970.

[48] Lincoln, H.B., "Use of Computer in Music Research", *Computers and the Humanities*, Vol. 8, No. 5-6, p. 285, 1974.

[49] Lohrding, R.K., M.M. Johnson, and D.E. Whiteman, "Computer Graphics for Extracting Information from Data", *Proceedings of the Eleventh Interface Symposium on Computer Science and Statistics*, 1978.

[50] McAdams, S. and A. Bregman, "Hearing Musical Streams", *Computer Music Journal*, Vol. 3, No. 4, 1979.

[51] Mantock, J.M. and K. Fukunaga, "A Two-dimensional Display for Multiclass Multivariate Data", *Pattern Recognition in Practice*, E.S. Gelsema and L.N. Kanal, editors, North-Holland, p. 361, 1980.

Mantock and Fukunaga present a method for two-dimensional display of n-dimensional data in which the two-dimensional coordinates relate the two largest pairwise log-likelihood ratios.

[52] Marks, S.M. and F.W. Lester, "Catalog of Earthquakes Along the San Andreas Fault System in Central California, January-March 1977", United States Department of the Interior, U.S. Geological Survey, Menlo Park, Calif., Open-File Report 80-1233, 1980.

[53] Martin, J., *Design of Man-Computer Dialogues*, Prentice Hall, New Jersey, 1973.

[54] Mathews, M.V., "An Acoustic Compiler for Music and Psychological Stimuli", *The Bell System Technical Journal*, Vol. 40, No. 3, p. 677, 1961.

[55] Mathews, M.V., "The Digital Computer as a Musical Instrument", *Science*, Vol. 142, p. 553, 1963.

[56] Mathews, M.V., "The Electronic Sound Studio of the 1970's", *Proceedings of the Stockholm Meeting*, 1970.

[57] Mathews, M.V., *The Technology of Computer Music*, The M.I.T. Press, Cambridge, Mass., 1969.

The book is an excellent introduction, concentrating on table-driven sound generation with a detailed explanation of Mathews' sound synthesis package, MUSIC V. Mathews provides an introductory section on sound, which includes some discussion of digital sound synthesis problems. The information in his earlier papers [54, 55, 61] is contained in this book.

[58] Mathews, M.V., Bell Laboratories, Murray Hill, New Jersey, private communication, 1977.

[59] Mathews, M.V., F.R. Moore, and J.C. Risset, "Computers and Future Music", *Science*, Vol. 183, No. 4122, p. 263, 1974.

[60] Mathews, M.V. and F.R. Moore, "GROOVE -- A Program to Compose, Store, and Edit Functions of Time", *Communications of the ACM*, Vol. 13, No. 12, p. 715, 1970.

[61] Mathews, M.V., J.R. Pierce, and N. Guttman, "Musical Sounds from Digital Computers", *Gravesaner Blatter*, Vol. 6, No. 23-24, p. 119, 1962.

A very basic article with good definitions.

[62] Moorer, J.A., "Signal Processing Aspects of Computer Music--A Survey", *Computer Music Journal*, Vol. 1, No. 1, p. 4, 1977.

This is an excellent and extensive article written on digital sound synthesis. Moorer covers analysis, several methods and applications of synthesis, and current hardware capabilities. Some of the material on synthesis techniques in Chapter 2 is based on Moorer's survey.

[63] Morse, A., "Some Principles for the Effective Display of Data", *Computers and Graphics*, 1979.

Morse presents a good discussion on visuals that provides a basis for analogous questions concerning audio representations.

[64] Neisser, U., *Cognitive Psychology*, Prentice Hall, New Jersey, 1967.

[65] Nelson, G., "MPL: A Program Library for Musical Data Processing", *Creative Computing*, Vol. 3, No. 2, p. 76, 1977.

MPL uses Chowning's frequency modulation.

[66] Oppenheim, A.V., R.W. Schafer, and T.B. Stockham, "Nonlinear Filtering of Multiplied and Convolved Signals", *Proceedings of the IEEE*, Vol. 56, No. 8, p. 1264, 1968.

[67] Parry, S., Naval Postgraduate School, Monterey, Calif., private communication, 1981.

[68] Pierce, J.R. and E.E. David, *Man's World of Sound*, Doubleday, New York, 1958.

The book contains valuable summaries of hearing characteristics.

[69] Rabiner, L.R. and R.W. Schafer, "Digital Techniques for Computer Voice Response: Implementations and Applications", *Proceedings of the IEEE*, Vol. 64, No. 4, p. 416, 1976.

[70] Rader, G.M., "A Method for Composing Simple Traditional Music by Computer", *Communications of the ACM*, Vol. 17, No. 11, p. 631, 1974.

[71] Revesz, G., *Introduction to the Psychology of Music*, University of Oklahoma Press, 1954.

This is an excellent reference about the perception of sounds including the relationships among amplitude, frequency and partials with respect to intensity, pitch, and timbre. Notes on synaesthesia are included.

[72] Riganati, J.P. and M.L. Griffith, "Interactive Audio-Graphics for Speech and Image Characterization", *Data Structures, Computer Graphics, and Pattern Recognition*, Academic Press, Inc., 1977.

The authors make a strong point of the potential for using sound in information presentation.

[73] Risset, J.C., "Synthesis of Sounds by Computer and Problems Concerning Timbre", *Music and Technology*, Proceeding of the Stockholm Meeting, 1970.

[74] Roberts, A., "An All-FORTRAN Music-Generating Computer Program", *Journal of the Audio Engineering Society*, Vol. 14, No. 1, p. 17, 1966.

Roberts uses Mathews' system, MUSIC IV, in a program to generate music for later playback.

[75] Roederer, J., *Introduction to the Physics and Physchophysics of Music*, Springer-Verlag, New York, 1975.

[76] Rosenstein, M., "Computer-Generated Pure Tones and Noise and Its Application to Psychoacoustics", *Journal of the Audio Engineering Society*, Vol. 21, No. 2, p. 121, 1973.

[77] Rozenberg, M., "Microcomputer-controlled Sound Processing Using Walsh Functions", *Computer Music Journal*, Vol. 3, No. 1, 1979.

[78] Saunders, S., "Real-time Digital FM Audio Synthesis", *Proceedings of the Second Annual Music Computation Conference*, 1975.

This is a helpful programming guide for an efficient implementation of Chowning's FM synthesis.

[79] Schaeffer, M., "The Electronic Music Studio of the University of Toronto", *The Journal of Music Theory*, Vol. 7, No. 1, p. 73, 1963.

[80] Schottstaedt, B., "The Simulation of Natural Instrument Tones Using FM with a Complex Modulating Wave", *Computer Music Journal*, Vol. 1, No. 4, 1978.

[81] Smith, G.C., "CODE SECECH -- Simulation of the Impact of Second Echelon Attacks on Outcome of (NATO) Central Battle", Lawrence Livermore National Laboratory, Livermore, Calif., internal report D-80-1, 1980.

[82] Smith, L., "SCORE -- A Musician's Approach to Computer Music", *Journal of the Audio Engineering Society*, Vol. 20, No. 1, p. 7, 1972.

[83] Snell, J., "Design of a Digital Oscillator Which Will Generate Up to 256 Low Distortion Sine Waves in Real Time", *Computer Music Journal*, Vol. 1, No. 2, 1977.

[84] Speeth, S.D., "Seismometer Sounds", *The Journal of the Acoustical Society of America*, Vol. 33, No. 7, 1961.

[85] Stahlin, R., "Optimizing Computer Graphics", *Computer Graphics World*, Vol. 3, No. 8, 1980.

[86] Stockham, T.G., T.M. Cannon, and R.B. Ingebretsen, "Blind Deconvolution Through Digital Signal Processing", *Proceedings of the IEEE*, Vol. 63, No. 4, p. 678, 1975.

[87] Strawn, J. (editorial notes) with J.A. Moorer, and J. Grey, "Lexicon of Analyzed Tones", *Computer Music Journal*, Vol. 1, No. 3, 1977.

The article contains a helpful explanation of timbre.

[88] Taylor, C.A., *The Physics of Musical Sounds*, American Elsevier, New York, 1965.

This book is an excellent reference book for a very thorough and technical discussion of sound.

[89] Tenney, J.C., "The Physical Correlates of Timbre", *Gravesaner Blatter*, Vol. 7, No. 26, p. 106, 1965.

Tenney includes a discussion of timbre which is relevant to psychoacoustics.

[90] Tenney, J.C., "Sound Generation by Means of a Digital Computer", *The Journal of Music Theory*, Vol. 7, No. 1, p. 24, 1963.

[91] Tobias, J.V., editor, *Foundations of Modern Auditory Theory*, Volume II, Academic Press, New York, 1972.

[92] Tremaine, H.M., *Audio Cyclopedia*, Howard W. Sams, Indianapolis, Indiana, 1973.

[93] Truax, B., "Organizational Techniques for C:M Ratios in Frequency Modulation", *Computer Music Journal*, Vol. 1, No. 4, 1978.

The paper contains good information on FM parameters.

[94] Tukey, J., *Exploratory Data Analysis*, Addison-Wesley, Menlo Park, Calif., 1977.

The book contains a wonderful description of a variety of techniques for data analysis.

[95] von Foerster, H. and J. Beauchamp, *Music By Computers*, Wiley, New York, 1969.

Although this is a good overview of the subject, the primary focus is on music composition and music analysis.

[96] Weaver, J., *Discrete and Continuous Fourier Analysis, Volumes 1-3*, Lawrence Livermore National Laboratory, Livermore, Calif., internal report UCIR 998, 1977.

These three volumes provide an excellent course in Fourier analysis and transforms. Chapter 1 introduces the necessary mathematical background before the discussion plunges into Fourier series. The book contains many examples with a particular focus on the application of the mathematics.

[97] Wilson, S.R., "Sound and Exploratory Data Analysis", The Research School of Social Sciences, The Australian National University, Canberra, Australia, draft, 1980.

This paper suggests a musical representation of multivariate data and offers a thorough description of the method used. It is a unique approach and is particularly interesting in comparison with other techniques.

[98] Winckel, F., *Music, Sound and Sensation*, Dover, New York, 1967.

[99] Yeung, E.S., "Pattern Recognition by Audio Representation", *Analytical Chemistry*, Vol. 52, No. 7, p. 1120, 1980.

To examine multivariate data, each measurement of a *k*-dimensional data vector is translated into a property of sound. Yeung reports excellent results when the method was applied to pattern recognition. This paper is a leader in the area of using sound for data discrimination.

APPENDIX A: RECORDING

The accompanying recording contains examples of the sounds used to present data. Listening to this recording will help understanding of the procedures and ideas described in this paper.

The sections labeled Phase 2 Experiment, Fisher Iris Data, and Spectra Data are arranged so that the listener can hear the training samples and then try to select the correct set for each test item. These test items are identified in the text for checking the results of a self-test or for reading while listening to the notes.

1: Parameter Variations

Each variable value was mapped to the (0.0, 1.0) range. Each sound parameter was varied over some integral number of levels. A variable value can be easily mapped to a corresponding level of a sound parameter. Thus, a value of 0.0 causes the lowest pitch (130 Hz), the softest volume, or the shortest duration (55 msec). A value of 1.0 maps to the highest pitch (2000 Hz), the loudest volume, or the longest duration (1.05 second). For attack, a value of 0.0 will cause a slow attack envelope and a value of 1.0 will cause a high, sharp attack envelope. 0.0 maps to a pure sine waveshape for either the fundamental waveshape or the overtone waveshape. 1.0 maps to a random waveshape for either the fundamental or the overtone. Section 4.1 contains a complete description of the parameter variations.

On the recording, each sound parameter varies for data values 0.0, 0.1, 0.2, ..., 1.0 while the other five parameters are held constant. The constant values for pitch, volume, and duration correspond to data values of 0.5. When not being varied, the attack is constant (1.0), the waveshape is pure (0.0), and the overtone is not used. The first section on the recording illustrates these parameter variations.

LISTEN to eleven variations for each of the six sound parameters: pitch, volume, duration, attack, fundamental waveshape, and overtone waveshape.

2: Normalization

Pairs of data sets were used to verify that information was presented meaningfully and consistently to listeners. First, Set 2 data was translated relative to Set 1 data. Second, Set 2 data was scaled relative to Set 1 data. Third, Set 1 variables had a .99 correlation with each other while Set 2 variables had no correlation. Section 5.2 contains a complete description of the normalization (Phase 1) experiments and Figure 5.1 illustrates the data set differences.

The data was six-dimensional with each dimension varying in the same way. Thus, the mapping between data variables and sound parameters was not significant.

Translation

In the first pair of data sets, Set 2 was translated by 3 standard deviations. Note that Set 2 sounds are sharper, louder, longer, higher, and buzzier. In the second pair, Set 2 was translated only 1 standard deviation. Set 2 data and Set 1 data more frequently overlap but Set 2 still has a noticeably sharp attack and long duration.

LISTEN to the translation pairs.

Scaling

In the first pair of scaled data sets, Set 2 was scaled by 8 standard deviations. Note that Set 1 samples are very mid-range and all sound very much alike. Set 2 samples seem to jump all around. They may be high and short, soft and low, or any combination whatsoever of parameters. In the second pair, Set 2 was scaled by 4 standard deviations. Ten values will be played from each set for each of two scalings.

LISTEN to the scaled pairs.

Correlation

Set 1 variables had a .99 correlation while the variables in Set 2 had no correlation. Thus, Set 1 sounds have corresponding characteristics. A sample in Set 1 which is low in pitch is also soft, short, very pure, and has a long attack. A Set 1 sample which is loud is also long, buzzy, high-pitched, and has a sharp attack with buzzy overtone. A good example of the difference in Set 2 is the fourth training sample which has low pitch (a low data value) but loud volume and long duration (high data values). Ten values will be played from each set of the correlation pair.

LISTEN to the correlation pair.

3: Phase 2 Experiment

These sounds are a subset of those which were used as part of the training and testing for Phase 2 of the experiments described in Section 5.3. The six-dimensional experiment data was mapped into six parameters of sound. These parameters were pitch, volume, duration, attack, fundamental waveshape, and overtone waveshape. Each participant was given ten samples from Set 1 and ten samples from Set 2 as training. Subsequently, forty test samples were played, one at a time. The participant repeated the test sample as often as desired before indicating whether the sample belonged in Set 1 or in Set 2.

Training Sets

Ten samples were randomly selected from Set 1 and ten from Set 2 for training. Variable 1 was mapped to waveshape, 2 to overtone, 3 to pitch,. 4 to attack, 5 to duration, and 6 to volume. At most, one of pitch, waveshape, or overtone has a high value in Set 2. (Refer to Section 5.3 for more details of the data sets). Note that Set 2 samples are in general soft, low, short, and pure sounding.

For reference to the actual data (Appendix B), the training sample sequence numbers are listed. For example, the first training sample in Set 1 is sample 43, (0.989, 0.050, 0.052, 0.209, 1.511, 0.992). Given that variables 1, 5, and 6 were mapped to waveshape, duration, and volume, the note is noticeably loud, long, and buzzy. Variable 3 is quite low as is the corresponding pitch of sample 43.

Set 1: 43, 49, 28, 6, 33, 37, 14, 24, 46, 21

Set 2: 50, 18, 27, 9, 4, 30, 48, 19, 10, 11

LISTEN to the ten training notes for Set 1 and the ten for Set 2.

Testing

LISTEN to twenty of the test samples.

The correct identification for the samples is as follows.

#1 is Set 2, sample 32. It is very short, soft, and low in pitch.

#2 is Set 1, sample 39. This note is noticeably longer, louder, and higher in pitch.

#3 is Set 1, sample 40. The mid-range pitch and buzz indicate that it is not in Set 2.

#4 is Set 2, sample 12. Although it is a higher note, it is soft with a long attack.

#5 is Set 2, sample 25. The note is low in pitch and short.

#6 is Set 1, sample 27. The loud volume and longer duration are indicative of Set 1.

#7 is Set 2, sample 2. The pitch is high and slightly long, but all else remains low as expected in Set 2.

#8 is Set 1, sample 45. Many higher values are represented by a sharper attack and buzzy note.

#9 is Set 2, sample 42. Although the note is buzzy, it is short and softer with a slow attack.

#10 is Set 2, sample 38. The note is very low and short.

#11 is Set 1, sample 20. The note is loud with a sharp attack despite being low and short.

#12 is Set 2, sample 34. It is low and otherwise a bit mid-range.

#13 is Set 1, sample 15. The note is too high and long to be in Set 2.

#14 is Set 1, sample 3. The note is harder to place in either Set 1 or Set 2 by listening. It is a bit too long and loud for Set 2, given other mid-range values.

#15 is Set 1, sample 30. The note is low but also long with a sharp attack.

#16 is Set 2, sample 41. The note is only mid-range in pitch and soft and short.

#17 is Set 1, sample 23. Set 1 characteristics are the sharp attack, the high volume, and the duration.

#18 is Set 2, sample 23. The note is louder than usual for Set 2, but it has a slow attack and a low pitch.

#19 is Set 1, sample 18. The note is short but has the louder volume and sharper attack of Set 1. The short duration makes it difficult to hear the buzziness.

#20 is Set 2, sample 44. A note with no outstanding features is characteristic of Set 2 (particularly since the attack is noticeably slow).

4: Fischer Iris Data

R. A. Fisher used measurements from fifty plants for each of three species, *Iris setosa*, *Iris versicolor*, and *Iris virginica*, for discriminant analysis by linear functions during the 1930s. This data consists of four variables for each flower: sepal length, sepal width, petal length, and petal width. Since several studies in discriminant analysis refer to this data, it is included here. Section 4.1 also describes the iris data.

Training

There are only four variables, so no overtone was used and the attack was a constant sharp attack. By listening to the notes, the sounds for each set can be generalized. The notes in Set 1 are extremely low pitched, short, and loud. Set 2 notes and Set 3 notes are higher and longer, but Set 2 notes are midrange in pitch while Set 3 notes are generally higher pitched, longer, and slightly buzzy.

LISTEN to ten training samples for each of the three data sets.

Testing

LISTEN to ten of the test samples.

The correct identification for the samples is as follows.

#1 belongs to Set 1. The note is soft, low, and short.

#2 belongs to Set 3. Note the high pitch and long duration of a very pure tone.

#3 belongs to Set 3. This note is slightly longer and less pure but still high.

#4 belongs to Set 2. Note the mid-range of all parameters.

#5 belongs to Set 1. As usual, a Set 1 note is low, short, and soft. Note that this example is especially buzzy.

#6 belongs to Set 3. The note is long and high pitched.

#7 belongs to Set 2. It's a pure-sounding tone with very mid-range characteristics.

#8 belongs to Set 1. The especially low pitch is clearly indicative of Set 1.

#9 belongs to Set 3. Note that this is longer and louder than #6.

#10 belongs to Set 2. This note is *very* pure, but mid-range otherwise.

5: Spectra Data

Actual γ-ray spectra data consisted of four sets of 17-dimensional data. After transforming the original data with a principal component analysis, the resulting first six principal components were used as the six variables for each sample.

Training

Like the iris data, the spectra data has no well-defined algorithm for determining the different sets. The notes in Set 1 are low in pitch, loud, somewhat buzzy, and mid-duration. Set 2 notes are higher, louder, and with a sharp attack as are Set 3 notes. However, Set 3 notes differ from Set 2 notes by being shorter and a bit more buzzy. Set 4 notes are noticeably higher in pitch than any of the others.

LISTEN to ten training samples for each of the four data sets.

Testing

LISTEN to ten of the test samples.

The correct identification for the samples is as follows.

#1 belongs to Set 2. Note the sharp attack and midrange pitch.

#2 belongs to Set 4. Note the high pitch.

#3 belongs to Set 4. Note the high pitch and longer duration.

#4 belongs to Set 2. Note the longer duration with some buzziness and overtone.

#5 belongs to Set 3. Set 3 notes usually are midrange in all parameters and shorter than Set 2 notes.

#6 belongs to Set 2. Note again the longer duration of Set 2 notes.

#7 belongs to Set 4. The Set 4 notes are easily distinguished by the high pitch.

#8 belongs to Set 3. Note the shorter duration.

#9 belongs to Set 3. The short duration separates Set 3 from Set 2.

#10 belongs to Set 1. The lower pitch and longer duration make Set 1 easy to identify.

6: Battle Songs

Three songs were created from the time data output of three battle simulations. For each of two sides, units could be in any of three states -- in reserve, in transit to the front, or in combat at the front. Losses occurred from user scheduled attacks or from mutual attrition at the front. For a given battle at each time step, the number of units at the Side A front lines, the number of units at the Side B front lines, the number of Side A units in transit, and the number of Side B units in transit were recorded. Two notes were played, a pure note with sharp attack for Side A and a buzzy note with overtone and long attack for Side B. For each note, the number of units at the front determined the corresponding pitch for that time step. The number of units in transit determined the volume. Thus the pitch of a note rose as the front line units increased, and the volume increased as the number of units moving toward the front line increased. Section 4.3 contains additional information on battle songs.

Battles With The Same Starting Input

Three battles were run with constant starting parameters. That is, at the start of each battle, the number of units for Sides A and B did not vary from one battle to the next. Also, the times of major attacks were specified. Probabilistic random-number-generated events determined specific unit movements and losses.

LISTEN to Side A of the first battle. Note the sharp decrease in pitch which signifies a decrease in the number of units at the front.

LISTEN to Side A of the second battle. Even though the battle began with the same parameters, the front units increase to a greater number than in the first battle.

LISTEN to both sides of the first battle. Despite Side A's loss of units, Side A regains strength at the end with more front units than Side B has.

LISTEN to both sides of the second battle. Note the strong comeback of Side A front units while the number of units at the front of Side B stays fairly constant in the second half of the battle.

LISTEN to Side A of the third battle. Although the volume indicates an increase in the number of units in transit about halfway through the battle, Side A never makes a strong comeback.

LISTEN to Side B of the third battle. Note the increasing number of units at the front for Side B.

LISTEN to both sides of the third battle. Once again, the front units of Side A suddenly decrease. This time, however, Side B grows in strength and ends the battle with more units at the front than Side A.

Battles With Varying Start Parameters

In this set of three battles, the start parameters for each of the three battles differed.

LISTEN to Side A of the first battle. The number of front units for Side A increases with only a little fluctuation.

LISTEN to Side B of the first battle. Despite some increase in the number of units in transit for Side B, the number of units at the front does not increase.

LISTEN to both sides of the first battle. Note that Side A always has more forces at the front than Side B. In fact, the number of Side B front units stays constant or drops.

LISTEN to both sides of the second battle. Side A has a sudden decrease in the number of units at the front and never regains enough strength to overcome Side B.

LISTEN to both sides of the third battle. Side A loses a few units at the front but then recovers to end with greater force than Side B. Note that the number of front units for Side B wobbles and that Side A has a fairly large number of units in transit.

APPENDIX B: PHASE 2 DATA

The following data is the complete set of samples used for the experiment, Phase 2. This is the original data before it was normalized to the (0,1) range. Sample i,j is the j^{th} sample of the i^{th} set. The *Minimum Sum* is the minimum of the three values

$$x2^2 + x3^2 + x4^2 + x5^2 + x6^2$$

$$x1^2 + x3^2 + x4^2 + x5^2 + x6^2$$

$$x1^2 + x2^2 + x4^2 + x5^2 + x6^2.$$

Note that the minimum sum in Set 1 is always greater than 2.25 and the minimum sum in Set 2 is always less than 2.25.

Sample	x1	x2	x3	x4	x5	x6	Minimum Sum
1, 1	1.700	0.577	0.229	1.114	0.432	0.978	2.769
1, 2	0.680	0.488	1.126	1.200	0.418	0.448	2.516
1, 3	0.222	0.892	0.748	0.665	1.145	0.258	2.429
1, 4	0.906	0.169	0.991	0.334	1.562	0.838	4.103
1, 5	0.904	0.669	0.487	0.783	1.665	0.031	4.071
1, 6	2.184	0.645	0.815	1.031	0.663	0.212	2.628
1, 7	0.886	0.366	0.570	1.571	2.040	1.357	8.930
1, 8	0.117	1.143	0.590	0.465	0.246	2.900	9.049
1, 9	0.435	0.236	1.280	0.698	1.382	0.182	2.675
1,10	1.258	0.085	1.722	0.133	1.364	0.202	3.509
1,11	1.394	0.775	0.080	0.446	0.541	1.631	3.759
1,12	0.807	1.411	0.515	0.851	0.681	0.749	2.665
1,13	1.355	0.099	0.235	2.054	0.656	2.127	9.238
1,14	0.498	2.482	0.370	0.743	1.580	0.415	3.606

1,15	1.197	0.411	3.198	0.370	1.241	0.373	3.418
1,16	1.955	0.433	1.390	0.067	1.314	0.858	4.587
1,17	0.330	0.572	0.815	0.574	0.778	1.606	3.950
1,18	2.019	0.317	0.737	1.144	0.193	0.705	2.487
1,19	0.996	0.426	0.613	1.612	0.157	0.711	3.686
1,20	0.775	0.359	0.589	0.874	0.281	1.618	3.937
1,21	1.335	0.280	1.014	0.407	2.072	0.138	5.584
1,22	0.251	2.937	0.902	0.905	1.296	0.118	3.389
1,23	1.182	0.186	0.503	1.302	0.807	1.188	4.045
1,24	0.782	0.134	1.671	2.251	2.005	0.678	10.176
1,25	1.206	1.025	0.267	0.584	1.056	0.116	2.592
1,26	0.674	0.867	0.567	1.283	1.531	1.046	5.860
1,27	0.207	0.244	0.969	0.523	1.060	1.342	3.300
1,28	0.907	1.190	0.834	0.281	1.129	0.029	2.873
1,29	0.453	0.733	1.138	0.840	1.078	0.324	2.715
1,30	0.086	0.699	0.167	1.241	0.980	0.222	2.585
1,31	2.046	0.333	0.235	0.689	0.840	1.816	4.644
1,32	0.064	1.879	0.804	0.494	2.304	0.052	6.206
1,33	1.028	0.576	0.070	0.077	2.401	0.162	6.134
1,34	1.219	0.527	0.663	0.154	0.679	1.917	4.877
1,35	0.853	1.876	0.985	0.792	0.558	0.488	2.875
1,36	0.390	0.215	2.847	1.509	0.967	1.642	6.107
1,37	0.307	0.112	0.205	1.662	0.048	0.674	3.273
1,38	0.217	0.885	0.259	1.087	0.739	1.005	2.852
1,39	0.130	0.096	1.686	0.108	1.020	1.198	2.513
1,40	1.224	1.481	1.835	0.318	0.510	1.500	6.303
1,41	0.200	1.373	0.943	1.517	0.047	0.472	3.456
1,42	0.616	0.271	1.431	1.127	2.070	0.004	6.008
1,43	0.989	0.050	0.052	0.209	1.511	0.992	3.316
1,44	0.596	0.413	0.240	1.077	1.439	0.760	4.036
1,45	0.497	1.590	1.365	1.175	1.278	0.101	5.134
1,46	0.117	0.524	0.950	1.131	0.022	1.142	2.872
1,47	1.781	1.173	0.515	0.786	1.789	1.318	7.197
1,48	0.543	0.209	0.300	0.120	2.819	0.160	8.120
1,49	0.002	0.039	1.199	0.557	1.204	1.097	2.965
1,50	1.769	0.503	0.301	1.065	0.866	1.242	3.770
2, 1	0.395	1.092	0.407	0.220	0.404	0.384	0.681
2, 2	0.396	0.605	2.070	0.409	0.927	0.060	1.553
2, 3	0.530	0.042	0.852	0.250	0.053	0.185	0.382
2, 4	0.362	1.308	0.729	0.397	0.951	0.104	1.735

2, 5	0.236	0.107	0.091	0.705	0.266	0.211	0.632
2, 6	1.042	0.259	0.825	0.086	0.355	0.231	0.934
2, 7	0.281	0.253	0.570	0.470	0.550	0.161	0.692
2, 8	0.292	0.699	0.076	0.153	0.303	0.064	0.210
2, 9	0.347	0.204	0.941	0.059	0.515	1.299	2.118
2,10	0.554	1.232	0.074	0.198	0.358	0.065	0.484
2,11	0.715	0.118	0.277	0.138	0.641	1.108	1.748
2,12	0.358	0.242	1.839	0.057	0.714	0.391	0.853
2,13	1.151	0.435	0.642	0.801	0.229	0.610	1.668
2,14	0.507	0.862	0.198	0.309	0.578	1.107	1.951
2,15	0.167	1.261	0.031	0.658	0.306	1.289	2.217
2,16	1.170	0.812	0.500	0.119	0.979	0.334	1.994
2,17	0.645	0.424	0.926	1.189	0.118	0.434	2.212
2,18	0.040	1.335	0.880	0.010	0.944	0.406	1.832
2,19	0.286	0.585	0.120	1.361	0.398	0.263	2.176
2,20	0.686	0.127	0.919	0.303	0.715	0.943	1.979
2,21	0.219	0.230	0.209	1.177	0.320	0.743	2.131
2,22	1.313	0.469	0.108	0.046	0.301	0.957	1.240
2,23	0.805	0.463	0.085	0.143	1.145	0.687	2.025
2,24	0.462	1.383	0.047	1.059	0.344	0.798	2.092
2,25	0.570	0.866	0.301	0.434	0.526	0.674	1.335
2,26	0.838	0.812	0.178	0.064	0.417	0.897	1.674
2,27	0.178	1.062	0.085	0.716	0.088	0.871	1.318
2,28	0.671	0.237	0.388	0.689	0.899	0.600	1.850
2,29	2.388	0.698	0.355	0.006	0.691	0.437	1.282
2,30	0.551	0.377	0.007	0.112	0.051	0.246	0.218
2,31	0.242	1.417	0.275	0.405	0.590	0.810	1.302
2,32	0.404	1.758	0.631	0.247	0.060	0.767	1.214
2,33	0.443	1.472	0.297	1.058	0.046	0.455	1.613
2,34	0.188	1.112	0.245	0.628	0.782	0.547	1.400
2,35	0.007	0.286	1.223	1.339	0.300	0.475	2.190
2,36	0.640	0.393	0.642	0.941	0.678	0.123	1.924
2,37	0.737	0.427	0.796	0.723	0.194	0.149	1.308
2,38	0.565	0.571	0.005	0.340	0.507	0.168	0.720
2,39	0.976	0.342	0.098	0.356	0.532	0.461	0.749
2,40	1.165	0.510	0.212	0.980	0.736	0.574	2.137
2,41	0.821	0.291	1.185	0.465	0.061	0.392	1.132
2,42	2.195	0.776	1.027	0.518	0.517	0.135	2.211
2,43	0.438	1.427	0.183	0.050	0.370	0.563	0.682
2,44	0.243	0.717	0.531	0.597	0.376	1.170	2.208
2,45	0.042	0.199	0.745	0.364	0.399	1.086	1.512

2,46	0.209	0.178	0.353	0.613	0.778	0.226	1.107
2,47	0.041	0.301	0.314	0.243	0.530	0.726	0.959
2,48	0.333	0.692	0.176	0.138	0.354	0.058	0.290
2,49	0.434	0.036	0.513	1.075	0.557	0.524	1.930
2,50	2.239	0.036	0.790	0.663	0.025	0.357	1.193

APPENDIX C: ANALYSIS

This appendix contains more detailed information concerning the results of Phase 2 of the experiment described in Chapter 5 (see Section 5.3). The purpose of Phase 2 of the experiment was to determine whether sound might enhance other methods of data discrimination. Three groups of subjects were tested for their ability to distinguish between two sources of test items. The data analysis of the results is used to verify that the subjects' responses were better than guessing and to conclude that sound enhances graphics in data discrimination.

1: Experimental Data

For Phase 2 of the experiment, 75 subjects were selected and randomly assigned to one of three groups. Group 1 saw two-dimensional representations of the data, group 2 heard six-dimensional representations of the data, and group 3 both saw two-dimensional representations and heard six-dimensional representations of the data.

The experiment task was for a subject to identify correctly a test item as belonging to one of two data sets. Each subject was presented with 40 test items, chosen at random from the two sets, and asked to identify the source of each test item. The number of test items correctly placed in one of the two sets was recorded as the individual's score. The raw scores (ordered from low to high) are summarized in Table C.1.

Index	Group 1	Group 2	Group 3
1	17	19	19
2	18	19	20
3	19	22	21
4	19	23	22
5	20	23	23
6	23	24	23
7	23	24	24
8	23	25	26
9	23	25	27
10	24	25	27
11	25	25	28
12	26	26	28
13	26	26	28
14	26	26	29
15	26	26	30
16	27	27	30
17	27	27	30
18	27	27	30
19	28	28	30
20	28	28	31
21	28	29	31
22	29	30	32
23	29	30	32
24	30	30	32
25	30	31	35
Mean	24.8	25.8	27.5
Standard Deviation	3.75	3.07	4.16

Table C.1: Phase 2 Raw Scores

2: Data Analysis

The data analysis attempts to answer the following two questions:

1: Were the subjects' responses better than would be expected from guessing?

2: Did some test groups perform better than others?

Definitions

The following variables were identified for the statistical methods used in the data analysis.

N is the number of test items presented to each subject; N = 40 for each subject.

N(i,j) is the number of test items correctly identified by the j^{th} subject in the i^{th} group.

i = 1, 2, or 3 where

1 is the group exposed to graphics only,

2 is the group exposed to sound only, and

3 is the group exposed to both graphics and sound.

j = 1,2,...,25.

X(i,j) is the proportion correctly identified by the j^{th} subject in the i^{th} group.

$$X(i,j) = \frac{N(i,j)}{N}$$

P(i,j) is the probability of a correct response by the j^{th} subject in the i^{th} group, assumed to be constant for all test items.

P(i,·) is the *average* probability of a correct response among subjects in the i^{th} group.

Question 1: Were Responses Better Than Chance?

To decide if performances were better than would be observed if subjects identified the set of a test item by chance alone, the hypothesis of chance selection,

$$H_{01}: P(i,j) = 0.5 \text{ for all } i, j,$$

was tested. The following analysis uses a chi-square statistic to show that one can reject H_{01} and conclude that the responses were not by chance alone.

Under the assumption that a subject has a constant probability p of correctly identifying the set of a test item, the variable $N(i,j)$ is a binomial random variable with parameters $N = 40$ and $p = P(i,j)$. Under the hypothesis of chance identification, the expected number of correct responses is $pN = 0.5 \cdot 40 = 20$.

To test the hypothesis H_{01}, an appropriate test statistic is a chi-square statistic with 75 degrees of freedom (25 subjects in each of three groups).

$$\chi^2_{75} = \sum_{i=1}^{3} \sum_{j=1}^{25} \frac{(N(i,j) - pN)^2}{p(1-p)N}.$$

Using the data in Table C.1,

$$\chi^2 = 386.0.$$

Since the 99.9th percentile of the chi-square distribution with parameter $\nu = 75$ is

$$\chi^2_{75,.999} = 119.0,$$

one can reject H_{01} and conclude that the responses were not chance selections.

Question 2: Did One Group Perform Better Than Any Other Group?

The hypothesis to be tested,

$$H_{02}: P(1,\cdot) = P(2,\cdot) = P(3,\cdot),$$

is that the average performance (i.e. probability of correct identification) is the same for all three groups. The following analysis uses an F-distribution statistic and the Newman-Keuls statistic. First, the variation between groups is compared with the variation of individuals within a group to show that the group performances are different. Second, the average differences among groups are compared. The statistic shows that group 3 (sound and graphics) performed better than group 1 (graphics only) at a .95 confidence level.

To use the method of analysis of variance for testing hypothesis H_{02}, it is necessary to assume that the variance of the responses is constant for all i, j. Since the response, N(i,j), is a binomial random variable for which the variance is related to the mean, the variances can be different among the groups. Thus, a transformed variable was used to approximate more closely the constant variance assumption for the analysis of variance method. The appropriate transformed variable is

$$Y(i,j) = \sin^{-1}\sqrt{X(i,j)} \ .$$

The data in Table C.1 was transformed and analyzed. The analysis of variance table, Table C.2, summarizes the variation in the Y(i,j) attributable to the two sources of variation: 1) the variation in the average probability of correct identification, P(i,·), between groups; 2) the variation in the probability of correct identification, P(i,j),

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Groups	2	229.3	114.6
Subjects/group	72	2350.3	32.6
Total	74	2579.6	

Table C.2: Analysis of Variance Table

between individuals within a group.

The test statistic for testing H_{02} is

$$F = \frac{\text{Group Mean Squares}}{\text{Subject Mean Squares}}$$

Using the data in Table C.2,

$$F = 3.51.$$

Since the 95th percentile of the F-distribution with parameters $\nu_1 = 2$ and $\nu_2 = 72$ is

$$F_{2,72}(0.95) = 3.13,$$

one can reject H_{02} and conclude that group performances were different.

Given that the hypothesis of equal performance is rejected, it is appropriate to identify groups which performed better than others. The Newman-Keuls statistic was used to identify such groups. This statistic is based on the range between the respective group averages. Using the transformed data, $Y(i,j)$, the group averages are

Group	Index	Mean, $\overline{Y}(i,\cdot)$
Graphics only	i=1	52.12
Sound only	2	53.54
Graphics and Sound	3	56.33

Thus comparing graphics only versus sound only involves a range of two groups, comparing graphics only versus graphics and sound involves a range of three groups, and comparing sound only versus graphics and sound involves a range of two groups. The observed differences are

$$|\overline{Y}(1,\cdot) - \overline{Y}(2,\cdot)| = 1.42$$

$$|\overline{Y}(1,\cdot) - \overline{Y}(3,\cdot)| = 4.21$$

$$|\overline{Y}(2,\cdot) - \overline{Y}(3,\cdot)| = 2.79$$

The Newman-Keuls test compares the difference $|\overline{Y}(i,\cdot) - \overline{Y}(k,\cdot)|$ for each pair of groups against

$$q_{r,72}(0.95)\sqrt{\frac{\text{Subject Mean Square}}{25}}$$

where $q_{r,72}(0.95)$ is the 95th percentile of the normalized difference between groups, and

25 is the number of subjects per group,
72 is the degrees of freedom for the Subject Mean Square, and
$r = 2$ or 3 is the number of groups per range.

For a range $r = 2$, the value is 3.222 and for a range $r = 3$, the value is 3.873. Since $4.21 > 3.873$, one can conclude that the performance of the graphics and sound group ($i=3$) was better than that of the graphics only group ($i=1$). It cannot be concluded that the graphics and sound group was better than the sound-only group nor that the sound-only group was better than the graphics-only group.