

Computer Graphics in Acoustics Research

A. MICHAEL NOLL, MEMBER, IEEE

Abstract—Digital computers are being used extensively in acoustics research. An integral, but often unemphasized, part of this research is effective and efficient man-computer communication, particularly in terms of computer-generated graphs and visual displays. This paper explores the many facets and philosophies of computer graphics as an irreplaceable link in man-computer communication. Numerous examples from the diverse areas of acoustics research are described to emphasize the wide range and varied aspects of computer graphics.

A VERY INTRIGUING and essential factor in human development has been man's nearly insatiable need for effective communication that has resulted in the evolution of human speech, and such technological advances as the printed word and the telephone. Perhaps, not surprisingly, man's need for communication has extended even to the products of his own technology. In most cases, the communication from man to the machine was effectively accomplished with a pushbutton or some equally simple mechanism. However, with the birth of the digital computer and its seemingly unlimited power, it became readily apparent that man-to-computer and computer-to-man communication on a much higher level than a simple pushbutton or flashing light would be required. As a result, special programming languages and other input media, which permit the human to communicate with the computer using the languages and skills of his own particular field, were developed. Likewise, the computer utilized special output media to communicate the results of its calculations to man. Any need for computer expertise on the part of the human was thereby greatly reduced, and ideally completely eliminated.

One particular form of man-computer communication that is particularly useful in interpreting the huge volumes of data that computers are adept at handling is computer graphics. Computer graphics includes within its domain the computer production of graphs and pictures as a pure output media for the human, and also pictorial or graphical displays as an interactive media for communication and interplay between man and the computer. This paper explores these different facets of computer graphics, particularly as they relate to research in speech, speech communication systems, hearing, and underwater acoustics. Examples from these areas are used to demonstrate the general principles of computer graphics, both as a pure output medium and as an interactive medium.

COMPUTER GRAPHICS AS AN OUTPUT MEDIUM

Automatic Plotters

Perhaps the most important factor in the very rapid growth of computer graphics has been the availability of automatic plotters that are directly controlled by the digital computer. In its simplest form, the plotter consists of an ink pen that is moved from one point to another on a sheet of paper, thereby drawing a straight line. The required electrical signals for positioning the pen are obtained as output from the digital computer. However, such pen recorders are generally very slow and are not very versatile. These disadvantages are eliminated by the high-speed automatic microfilm plotter.

The high-speed automatic microfilm plotter consists of a cathode-ray tube and a camera, positioned to photograph the face of the tube, as shown in Fig. 1.

Manuscript received August 28, 1967. This paper was presented at the 1967 Conference on Speech Communication and Processing, Cambridge, Mass.

The author is with Bell Telephone Laboratories, Inc., Murray Hill, N. J. 07974

The cathode-ray tube contains a metallic plate with alphabetic and other characters stenciled into it. The electron beam is passed through the selected character in this plate, and the stenciled electron beam is then positioned on the face of the tube. A dot can be chosen and moved from one position on the face of the tube to some other position, thereby producing a straight line on the film in the camera. The film, incidentally, is usually 35-mm microfilm, but facilities are also available for 16-mm film, so that movies for immediate viewing in a standard 16-mm movie projector are also easily obtained. Each axis of the picture area is divided into anywhere from 1024 to 4096 directly addressable locations, depending upon the characteristics of the particular plotter. In conclusion, the automatic microfilm plotter offers the advantages of speed, high quality, and flexibility of output media. Nearly all of the examples described in this paper were produced on a General Dynamics SC-4020 plotter, with a 1024 by 1024 picture area.

Programming Languages

The automatic plotter just described receives its instructions either directly from the main computer, or on magnetic tape obtained as output from the main computer. These instructions tell the plotter either where to position any desired character, or to draw straight lines between some set of points. Although these instructions are very simple, considerable programming effort is required to produce a plot that is both aesthetically pleasing and scientifically meaningful. In other words, these simple instructions are so basic that the programmer is forced into much detailed programming when pictorial output is desired. Nevertheless, the automatic plotter's ability to present vast quantities of data in pictorial form usually overcomes the disadvantages of additional programming effort.

In fact, the results of certain digital simulations could only be presented graphically to be meaningful. As an example, there are 400 short-time spectra and cepstra for 2 seconds of speech, if the calculations are performed every 10 ms.^[1] The user must see a graphical display of all the spectra and cepstra to attach any significance to their time variations. The high-quality plots, shown in Fig. 2, as produced by the automatic plotter are clearly required and are well worth any additional programming effort. The concept of cepstrum pitch determination could not have been evaluated if it were not for computer simulation coupled with computer-generated graphical output.

Pictures with a gradation of tone, or gray scale, can be produced by programming the automatic plotter to repetitively plot dots at specified locations. Fig. 3 shows the amplitude spectrum and square root of the amplitude spectrum thus produced by the author. Each dot was plotted a number of times linearly proportional to the amplitude spectrum or its square root after normalization to some maximum value.

There are, however, many occasions when the computer user does not want to be concerned with the fine details of programming graphical output, but still has a very strong need for effective graphical display. The obvious solution is higher level programming languages and subroutines that fill in the programming details, thereby relieving the user from what might otherwise be considerable and unnecessary programming burden. For example, Kaiser has developed a single subroutine for producing very complicated but aesthetically pleasing graphs on the automatic plotter.^[2] The user need specify only such things as linear or logarithmic scales for the X and Y axis, or the number of separate curves. Grid lines are automatically labeled, titles and labels are automatically centered, and a host of other automatic features are built into the subroutine covering a wide range of plotting versatility. Golden used this subroutine to plot the frequency and phase response of digitally-simulated filters, as shown in Fig. 4.^{[3],[4]} The filter design for computer-simulated vocoders was made considerably easier by such an automatic graphic-output capability.

This, incidentally, brings up another important aspect of computer graphics. In most cases, the scientific user usually thinks in graphic terms, so that a graphic display is most meaningful to him. For example, the filter designer is interested in plots of frequency and phase response and not in tables of numbers. Similarly, the electrical engineer usually thinks in terms of a block diagram, so that the special block diagram compiler written by Kelly, Lochbaum, and Vyssotsky is an especially important programming language.^[5] The electrical engineer, or anyone else for that matter, with a block diagram of any electronic system showing the interconnection of the adders, amplifiers, and other elements can easily learn the block diagram (BLODI) language, and use the computer to simulate the system. A block is available that will produce graphic output showing the waveform at any desired place in the block diagram. Fig. 5 shows the input and output signal waveforms of different channels in a computer-simulated analytic router, as programmed by Schroeder, Flanagan, and Lundry, using the block diagram compiler.^[6]

Computer-Generated Movies

A fair amount of thought and effort has been spent on producing movies using the automatic plotter. Such computer-produced movies are particularly advantageous because, in effect, the scientist or engineer can be his own animator, specifying equations to the computer which then produces the movie on the plotter. A special-purpose language for computer animation, which uses a mosaic of dots and other characters, has been written by Knowlton.^[7] With this language, or with conventional plotting instructions, movies can be produced that should be of considerable assistance in depicting engineering and scientific concepts for educational purposes.^{[8]-[10]}

Computer-generated movies are an entirely new graphic medium, which results in the capability for "moving graphs." Quite obviously, the movies are particularly enlightening when some event changes with time, as, for example, in the selected frames of Fig. 6 from a movie produced by Sondhi.^[11] This movie shows the propagation of underwater sound wavefronts emanating from a point source and based upon ray-path calculations. Different sound velocity profiles can be specified to gain further insight into the nature of the propagation.

Three-Dimensional Graphics

The computer and automatic plotter can calculate and draw two-dimensional perspective projections of any three-dimensional data. But, for many applications, particularly those involving very complicated plots with many hidden portions, a simple perspective plot is unsatisfactory. For these applications, the plotter can be used to produce true three-dimensional plots by drawing separate pictures for the left and right eyes, which, when viewed stereoptically, will fuse and result in a true three-dimensional effect.^[12] The geometry of the situation is shown in Fig. 7. The two

viewing points and their associated projection planes are rotated through an angle β in the $X-Z$ plane, and then through an angle α in the vertical plane, passing through the origin. The three-dimensional straight line P_1P_2 is projected onto the left and right projection planes by conventional perspective projection, i.e., the intersections with the projection plane of lines drawn from the end points of the line P_1P_2 to the viewing point. An analysis of the geometry results in the equations of Fig. 8, which give the coordinates for the left and right projections of some three-dimensional point.

The equations for projecting a point in three-dimensional space were incorporated into a subroutine which would generate separate pictures on microfilm for the left eye and the right eye. This subroutine was written so that the user need only consider in his main program the procedure for generating the three-dimensional data points as three arrays containing the X , Y , and Z axis coordinates of the points. The three-dimensional projection calculations, automatic scaling of the pictures to fill the picture area optimally, and the instructions for the automatic plotter are all done in the subroutine. The user must specify some parameters pertaining to

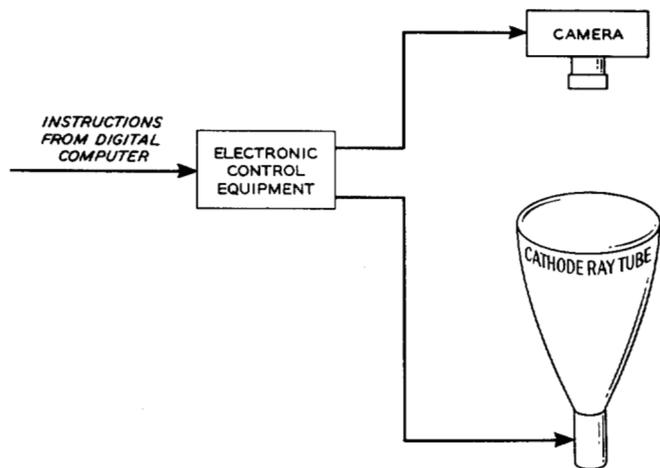


Fig. 1. Block diagram of high-speed automatic plotter, consisting of a camera and cathode-ray tube, both under the control of a digital computer.

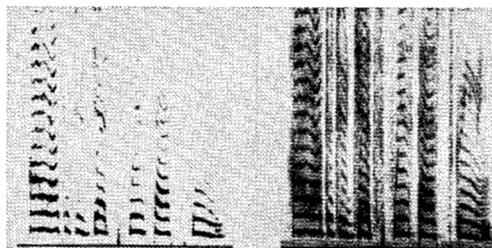


Fig. 3. Computer-generated spectrograms of the amplitude spectrum (left) and square root of the amplitude spectrum (right) for the sentence "High altitude jets whiz past screaming" spoken by a female. Frequency from 0 to 4 kHz is plotted vertically, while time is plotted along the horizontal axis.

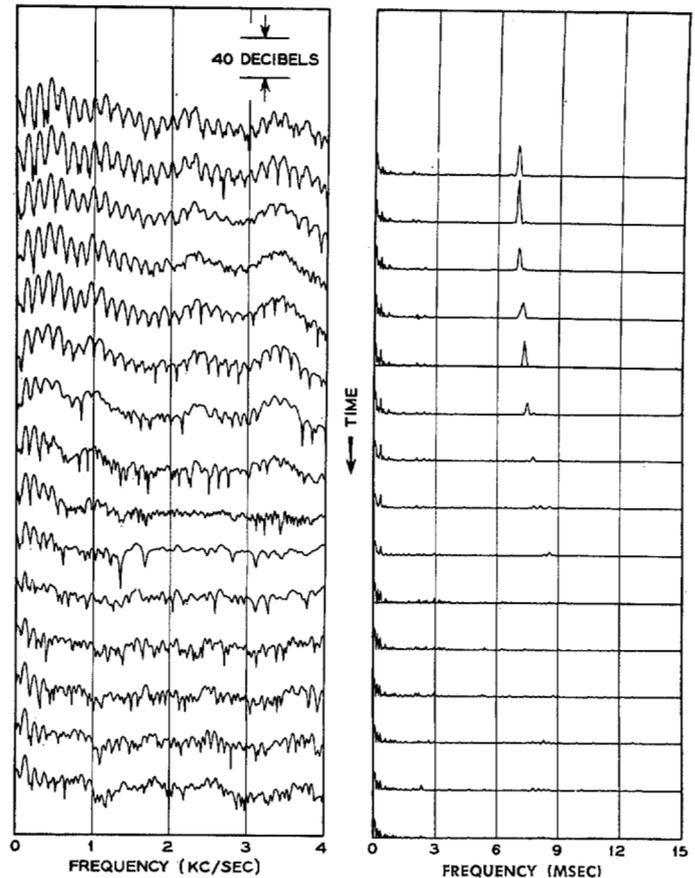


Fig. 2. Computer-generated logarithm spectra (left) and cepstra (right) for a male talker recorded with a condenser microphone. The 40-ms long Hamming time window moved in jumps of 10 ms.

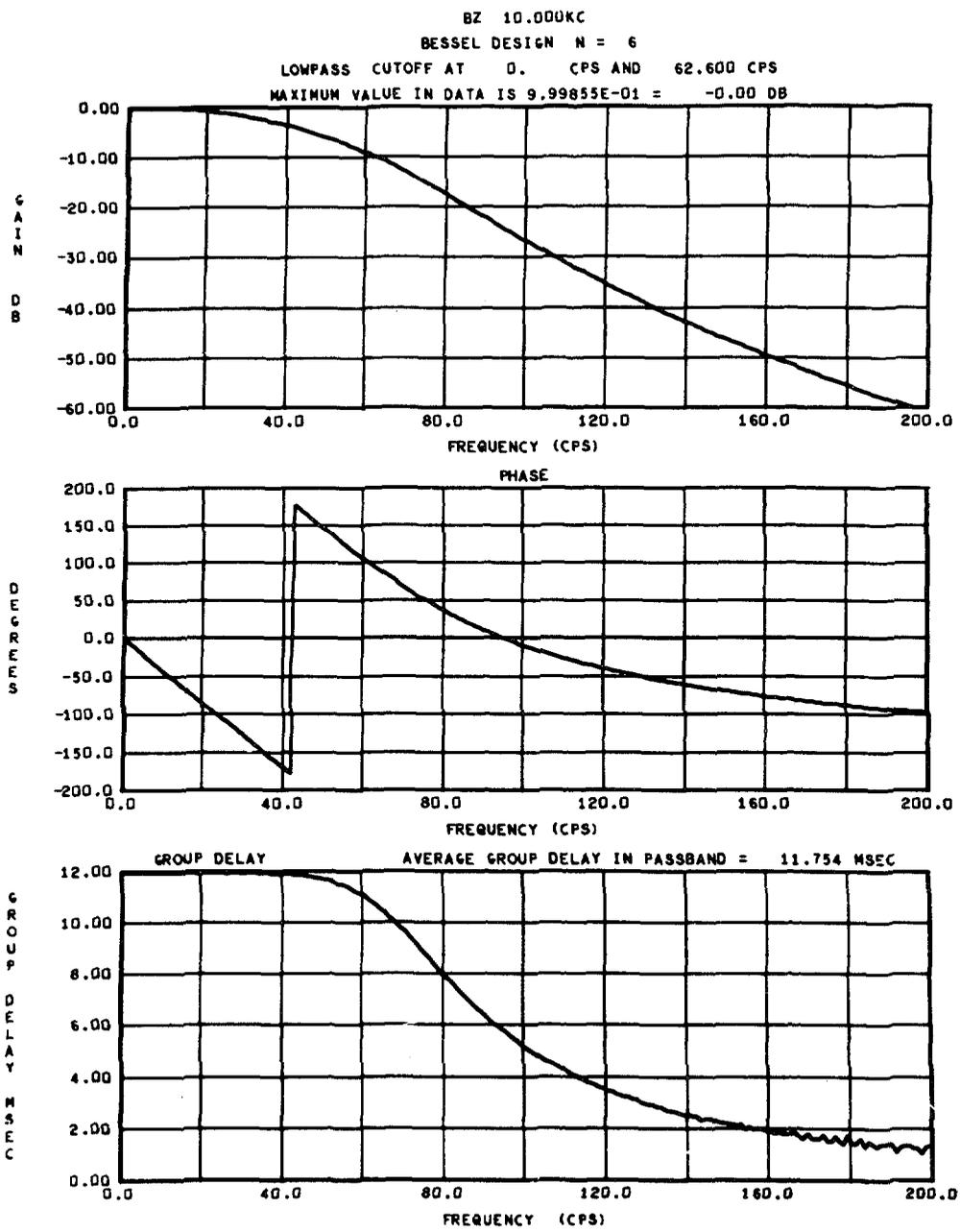


Fig. 4. Computer-generated graph showing the response of a digital filter for a computer-simulated speech bandwidth compression system.

BANDWIDTH COMPRESSOR WITHOUT BP 1/2 BESSEL FILTERS
SAMPLE NO. IN THIS RUN = 1 IN THIS RECORD =

2 RECORD(S)/LINE

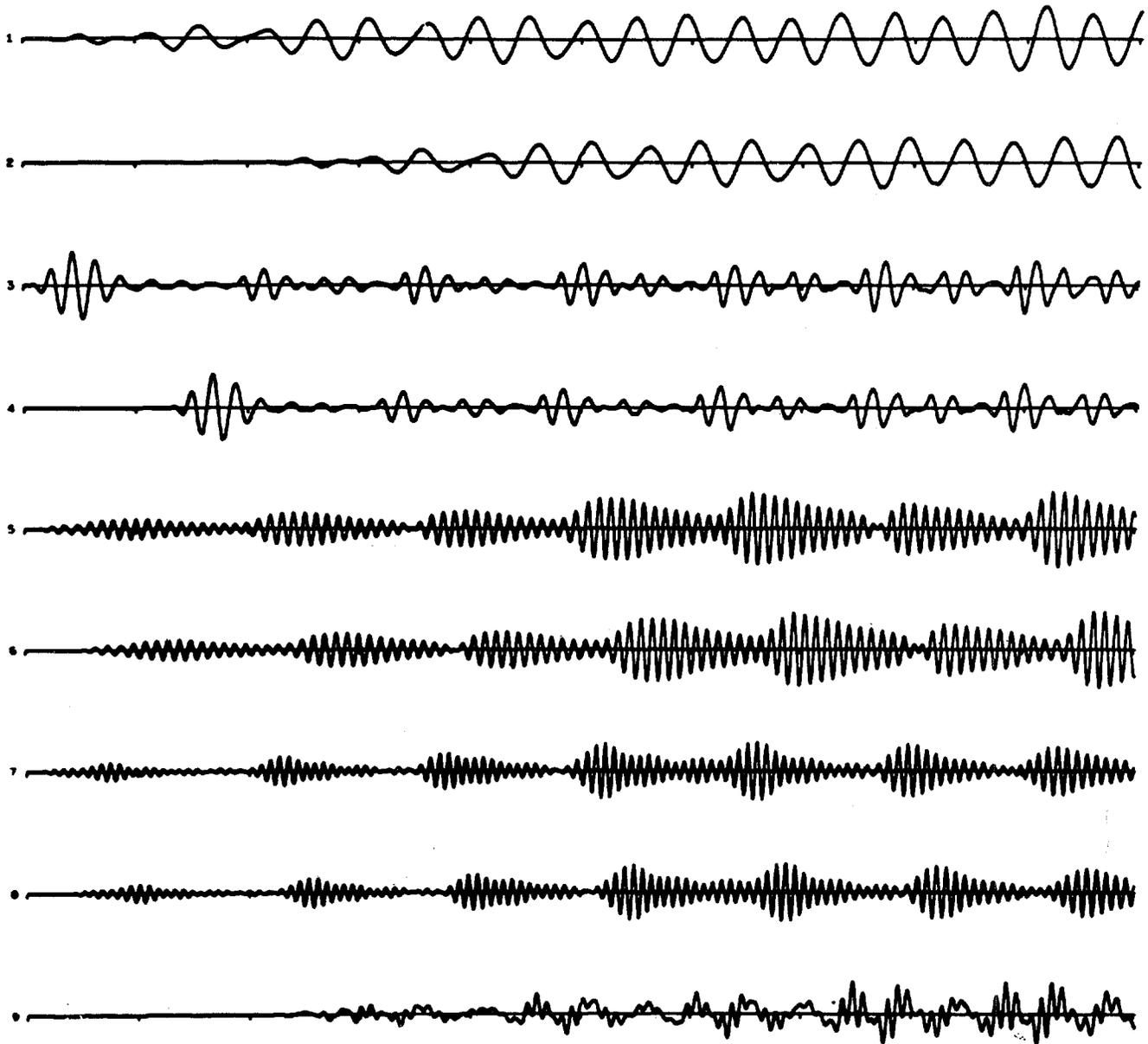


Fig. 5. Computer-generated plot of the waveforms at different portions of a speech bandwidth compression system simulated on the computer, using a block diagram compiler programming language.

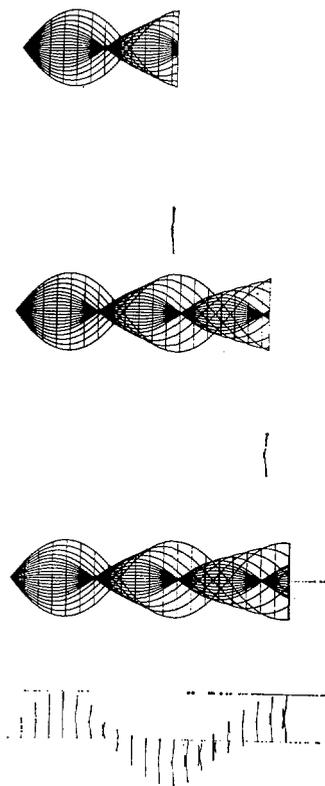


Fig. 6. Selected frames from a computer-generated movie showing the propagation of underwater sound from a point source. Some of the ray paths are shown in the upper portion of each frame, while the corresponding wave front is shown in the lower portion.

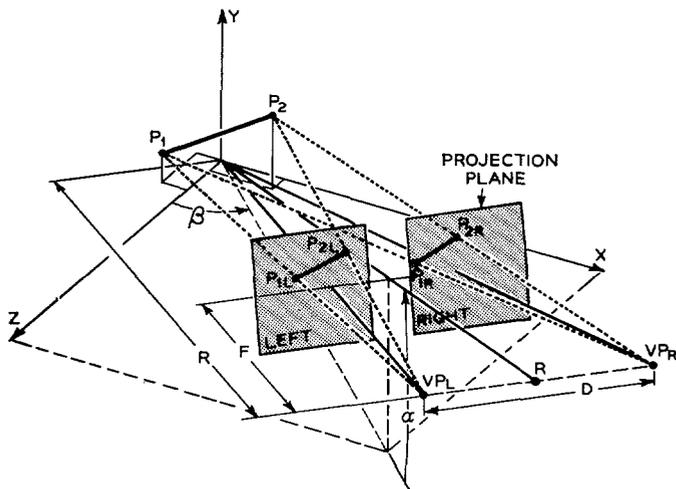
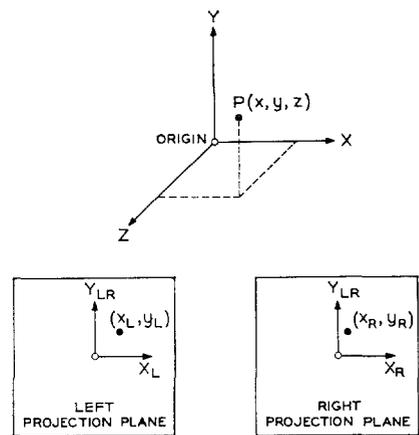


Fig. 7. The geometry of three-dimensional stereographic projection. The two viewing points VP_L and VP_R correspond to the left and right eyes.

the viewing geometry, but otherwise the subroutine eliminates the bulk of the programming effort, so that three-dimensional pictures can be simply and routinely obtained. As an example, Fig. 9 shows speech spectra in three-dimensions, as produced by Golden using the three-dimensional plotting subroutine.

A three-dimensional movie is produced by repetitively calling the three-dimensional subroutine, except that the left and right images for each three-dimensional



- α = ANGLE OF INCLINATION OF VIEWING POINTS WITH HORIZONTAL PLANE
- β = ANGLE OF ROTATION OF VIEWING POINTS ABOUT VERTICAL AXIS
- D = DISTANCE BETWEEN LEFT AND RIGHT VIEWING POINTS
- R = DISTANCE FROM VIEWING POINTS TO ORIGIN
- F = DISTANCE FROM VIEWING POINT TO PROJECTION PLANE

$$X_L = \left[\frac{D}{2} + (x \cos \beta - z \sin \beta) \right] G$$

$$X_R = \left[\frac{D}{2} - (x \cos \beta - z \sin \beta) \right] G$$

$$Y_L = Y_R = \left[y \cos \alpha - (z \cos \beta + x \sin \beta) \sin \alpha \right] G$$

$$\text{WHERE } G = \frac{F}{R - [y \sin \alpha + (z \cos \beta + x \sin \beta) \cos \alpha]}$$

Fig. 8. The basic equations for the stereographic projection of a three-dimensional point into the left and the right projection planes.

pair are now drawn on a single frame of film.^[13] Fig. 10 shows some selected frames of the simulated motion of the basilar membrane in the human ear,^[14] from a three-dimensional movie produced by Lumms and Sondhi. The quiescent membrane is represented by a line in the shape of a three-dimensional spiral that approximates the actual shape of the human basilar membrane. This spiral shape also results in relatively small displacements of the line compared to its uncoiled length, so that a very good feel for the traveling-wave motion of the membrane is obtained. For the simulation, the program accepted a digitized version of the sound waveform at the eardrum and computed the responses, in accordance with Flanagan's model, of a number of different locations on the membrane.^[15] This three-dimensional movie and a newer two-dimensional version, depicting the quiescent membrane as a straight line, have been of considerable assistance in visualizing the motion of the basilar membrane in response to many different acoustic signals.

COMPUTER GRAPHICS AS AN INTERACTIVE MEDIUM

In all of the preceding examples, the man-computer communication channel had a comparatively very long time delay of the order of a few hours between the man's input to the computer and the computer's output

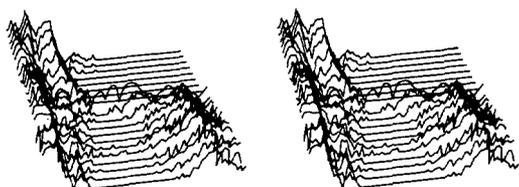


Fig. 9. Computer-generated three-dimensional spectrogram of speech. Increasing frequency is plotted to the right, and each spectral slice is for a 40-ms speech interval. (To view the spectrogram in three dimensions, place a sheet of paper on edge between the stereo pair. Position your head so that each eye sees only one image. The pictures should then seem to converge and appear three dimensional.)

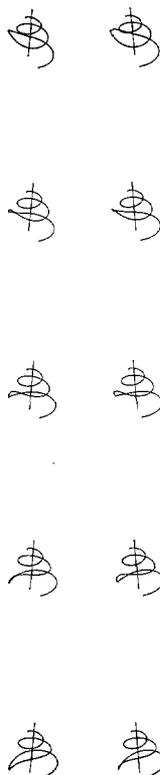


Fig. 10. Selected frames from the computer-generated three-dimensional movie showing the simulated motion of the basilar membrane responding to the sound "o-o." (To view this figure in three dimensions, place a sheet of paper on edge between the stereo pairs. Position your head so that each eye sees only one set of images. The pictures should then seem to converge and appear three dimensional.)

to the man. Even so, the communication channel was reasonably efficient, since the man communicated with the computer in terms of special programming languages, while the computer communicated graphically with the man. There are, however, many occasions when one desires an almost immediate response from the computer, so that a sort of dialogue between man and the computer is possible. The computer performs those calculations and operations for which it is especially well suited, while the man performs those decisions and evaluations which cannot be programmed because of lack of detailed knowledge about the decision and evaluation processes. For these applications, fast and efficient man-computer communication and interaction

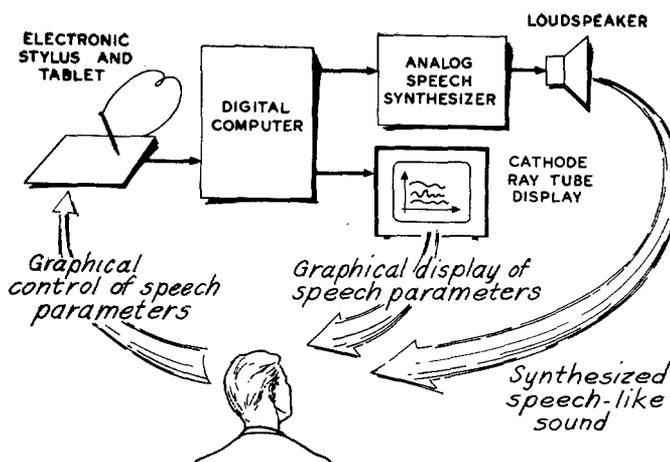


Fig. 11. Block diagram of the man-computer communication channels in a system used to synthesize speech-like sounds.

is an absolute necessity. The man must think about his problem in the terminology of his particular field, and, accordingly, the man-computer communication link must be in terms of this terminology. Here, man is simply unable to react fast enough to the computer, unless high-level forms of communication are used.

Two-way graphical communication between man and the computer is a very important and central theme of these extremely fast or "real-time" interactive facilities. The user watches a large cathode-ray tube on which a computer-generated display is presented. Through the use of either a light pen or an electronic stylus and tablet, the user instructs the computer to manipulate the display. The computer can then act upon the graphic display, perform certain previously specified operations, and perhaps modify the display. Knobs and a typewriter are sometimes used as supplementary input to the computer in addition to the graphical facilities. The interaction is completely dynamic; the computer reacts immediately to the instructions from the man, and generates results nearly instantaneously.

As an example, Denes has programmed a computer to produce a graphic display showing the time variation of certain speech parameters, such as pitch, the first three formant frequencies, and the bandwidths of the first two formants.^[16] The user can graphically vary these parameters by using an electronic stylus and tablet, which are also connected to the computer as shown in Fig. 11. When the user is ready, the computer uses the graphically specified parameters to produce electrical signals for controlling an electronic speech synthesizer. The user hears the synthesized speech-like sounds, and can then make appropriate changes in the graphically displayed parameters to investigate such things as the effect of pitch variations on vocal stress. The man-computer communication is completely in terms of the time variation of speech parameters, which are readily known and understood by all researchers in the speech field. Absolutely no programming experience is required for a speech researcher to use the system.

SUMMARY AND CONCLUSION

Recapitulating, the digital computer and automatic plotter can be used to obtain graphical output in the conventional form of graphs and plots, but can also produce plots in three dimensions, and motion picture films in two or three dimensions. Through special programming languages and general purpose subroutines, the immense power of insight from graphical output from the computer is available, with an absolute minimum of programming expertise. Graphical displays and associated equipment are also available, so that the user can interact and communicate graphically with the computer and almost immediately see (or hear) the results of the interaction. Considerable research effort is presently being expended to develop even faster and more efficient forms of man-computer communication, and many of these efforts lean even more heavily on computer graphics to tighten the man-computer loop.

Although some rudimentary, but still very impressive, work has already been done in obtaining graphs and movies in color as graphical output from the computer, in the future colored graphical output and displays will be obtained as routinely as all the previously described black and white examples. The added "dimension" of color can be augmented by the capability for manipulating and presenting displays of four and even higher dimensional data. The final data display could be accomplished by successive projections of the data, until a final three-dimensional projection was obtained.^[17] Such techniques would be extremely useful for multidimensional representations of the results of experiments and other experimental or calculated data.

Quite admittedly, all of the preceding has been a rather extensive trip through the many different forms of graphic output from the digital computer. The numerous examples of computer graphics, as used in acoustics research at Bell Telephone Laboratories, have hopefully exemplified the many aspects of computer graphics in the context of the broader area of man-computer communication. However, the real essence of this paper is the desirability of an even closer partnership between man and the computer, in which efficient man-computer communication is so very essential.

REFERENCES

- [1] A. M. Noll, "Cepstrum pitch determination," *J. Acoust. Soc. Am.*, vol. 41, pp. 293-309, February 1967.
- [2] J. F. Kaiser, "Graphs should be computer drawn," *Proc. Symp. on the Human Use of Computing Machines* (Bell Telephone Labs., June 1966), pp. 9-14.
- [3] R. M. Golden, "Digital computer simulation of sampled-data communication systems using the block diagram compiler: BLO-DIB," *Bell Sys. Tech. J.*, vol. 45, pp. 345-358, March 1966.
- [4] R. M. Golden, "Digital computer simulation of a sampled data voice-excited vocoder," *J. Acoust. Soc. Am.*, vol. 35, pp. 1358-1366, September 1963.
- [5] J. L. Kelly, Jr., C. L. Lochbaum, and V. A. Vyssotsky, "A block diagram compiler," *Bell Sys. Tech. J.*, vol. 40, pp. 669-676, May 1961.
- [6] M. R. Schroeder, J. L. Flanagan, and E. A. Lundry, "Bandwidth compression of speech by analytic-signal rooting," *Proc. IEEE*, vol. 55, pp. 396-401, March 1967.
- [7] K. C. Knowlton, "A computer technique for producing animated movies," *AFIPS Conf. Proc.*, vol. 25, pp. 67-87, 1964.
- [8] E. E. Zajac, "Computer animation: a new scientific and educational tool," *J. SMPTE*, vol. 74, pp. 1006-1008, November 1965.
- [9] E. E. Zajac, "Film animation by computer," *New Scientist*, vol. 29, pp. 346-349, February 10, 1966.
- [10] W. H. Huggins, "Film animation by computer," *Mech. Engrg.*, vol. 89, pp. 26-29, February 1967.
- [11] M. M. Sondhi, "Computer movies of wavefront motion," *J. Acoust. Soc. Am.*, vol. 42, p. 1210, November 1967.
- [12] A. M. Noll, "Stereographic projections by digital computer," *Computers and Automation*, vol. 14, pp. 32-34, May 1965.
- [13] A. M. Noll, "Computer-generated three-dimensional movies," *Computers and Automation*, vol. 14, pp. 20-23, November 1965.
- [14] R. C. Lummis, A. M. Noll, and M. M. Sondhi, "A 3-D glimpse of the hearing process," unpublished memorandum meant to accompany film entitled, "Simulated basilar membrane motion (3D)," available on loan from Film Library, Bell Telephone Labs., Murray Hill, N. J. 07974.
- [15] J. L. Flanagan, "Models for approximating basilar membrane displacement—part II. Effects of middle-ear transmission and some relations between subjective and physiological behavior," *Bell Sys. Tech. J.*, vol. 41, pp. 959-1009, May 1962.
- [16] P. B. Denes, "Real-time speech research," *Proc. Symp. on the Human Use of Computing Machines* (Bell Telephone Labs., June 1966), pp. 15-23.
- [17] A. M. Noll, "A computer technique for displaying n -dimensional hyperobjects," *Commun. ACM*, vol. 10, pp. 469-473, August 1967.



A. Michael Noll (S'59-M'68) was born in Newark, N. J., on August 29, 1939. He received the B.S.E.E. degree from Newark College of Engineering, Newark, N. J., in 1961, and the M.E.E. degree from New York University, New York, N. Y., in 1963.

He joined the Bell Telephone Laboratories as a Member of the Technical Staff in 1961. He was initially concerned with the assessment of telephone quality and, in particular, the subjective effects of peak clipping and sidetone. In 1965, he was transferred to the Acoustics Research Department, where he was concerned with computer simulations and investigations of short-time spectrum analysis and the cepstrum method for vocal pitch determination. He is currently in the Speech and Communication Research Department at Bell Telephone Laboratories, Inc., Murray Hill, N. J. His present interests include computer-generated three-dimensional displays of data, the application of computer technology to the visual arts, and psychological investigations of human reactions to pseudo-random patterns. He is the author of various publications.

Mr. Noll is a member of Tau Beta Pi, Eta Kappa Nu, the Acoustical Society of America, the American Society of Aesthetics, the Association for Computing Machinery, and the Audio Engineering Society. He is a licensed Professional Engineer in the State of New Jersey.