Automatic Stereoscopic Presentation of Functions of Two Variables

By BELA JULESZ and JOAN E. MILLER

(Manuscript received September 21, 1961)

Spatial models of functions of two variables are often a valuable research tool. Nomograms and artistic relief drawings in two dimensions are difficult to prepare and still lack the direct impact of a spatial object. It has been demonstrated (see Ref. 2) that objects with a randomly dotted surface permit the determination of binocular parallax and, thus, can be seen in depth even though they are devoid of all other depth cues. This random surface presentation has the advantage that the random brightness points can be evenly and densely placed, whereas the classical contour-line projection at equally spaced heights may leave empty spaces between adjacent contour-lines. A digital computer is used to generate the three-dimensional image of a given $z = f(x, y)$ function and to wrap its surface with points of random brightness. The stereo projections of the function are obtained and, when viewed stereoscopically, give the impression of the three-dimensional object as being viewed along the z-axis. The random surface prevents the accumulation of clusters of uniform regions or periodic patterns which yield ambiguities when fused. Two stereo demonstrations are given of surfaces obtained by this method.

I. INTRODUCTION

Pictorial representations and visual displays are invaluable aids in conveying scientific or technical information. In particular, the problem of presenting three-dimensional data is of interest both from the standpoint of its wide range of applicability and the difficulty involved in the production of such representations.

The methods usually employed to present functions of two variables in the fields of applied mathematics, engineering, cartography, etc., fall into two categories: 1) two-dimensional and 2) three-dimensional displays. The first has the obvious advantage of being suitable for the printed page, thus permitting a wide circulation for the information so presented. The techniques of nomography, orthography, isarithmic (contour line) representations (see Fig. 1) and relief drawings (see Fig. 2) belong to this category and are widely used despite the expense and difficulty in their preparation. However, the greatest objection is perhaps the failure of such displays to match the capabilities of human observers, who are equipped to perceive a three-dimensional object in depth. The second category — that of spatial models or sculpture — answers this objection, but these models are usually much too difficult to execute and much too limited in their applicability.

There is, therefore, a need for a technique which a) eliminates the tedious effort required of draftsmen in producing such displays, b) presents displays complete with the spatial effects inherently belonging to three-dimensional objects and appreciated by human observers, and c) generates displays suitable for the printed page. This first requirement has already been met for two-dimensional representations by the de-

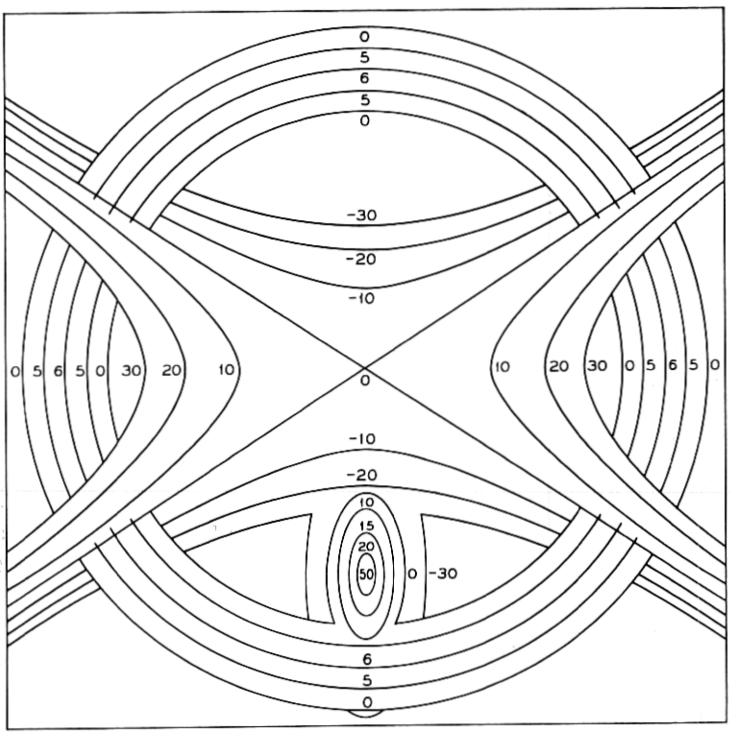

Fig. 1 — Isarithmic (contour-line) drawing (Example 1).

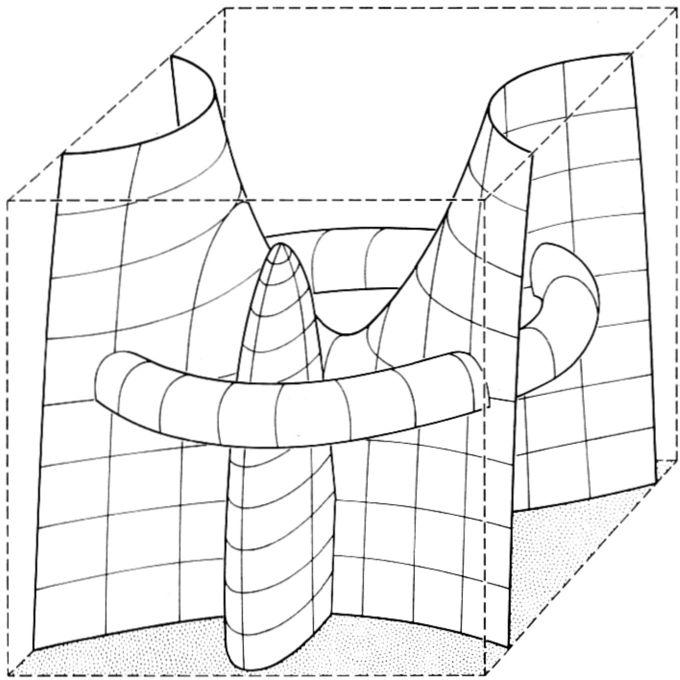

Fig. 2 — Relief drawing (Example 1).

velopment of oscilloscopic displays which automatically project onto the screen of the oscilloscope the object surface defined by one dependent voltage and two independent voltages.[1] The second and third requirement, however, seem of particular interest, and therefore this paper discusses a method employing a computer to make stereoscopic presentations of functions of two variables.

II. METHOD

The technique to be described here may be outlined as follows: the three-dimensional image of a given function $z = f(x,y)$, which is supplied as a table of corresponding x, y and z values, is stored in a digital computer. The computer is programmed to generate a stereo picture pair which, when fused, gives the subjective impression of the three-dimensional object as being viewed along the z-axis perpendicular to the base plane of x and y. This procedure for obtaining the stereo projections of an object can be considered in three parts: 1) defining the function to be presented as a three-dimensional object, 2) "wrapping" the object with a textured surface, and 3) generating a stereo pair by taking proper projections of the object.

In practice, the variables x, y and z must be evenly sampled with a given resolution. Therefore, the object can be defined only by approximation, and the various approximations differ in their fine structure. The classical method is the contour-line approach shown in Fig. 3(a)

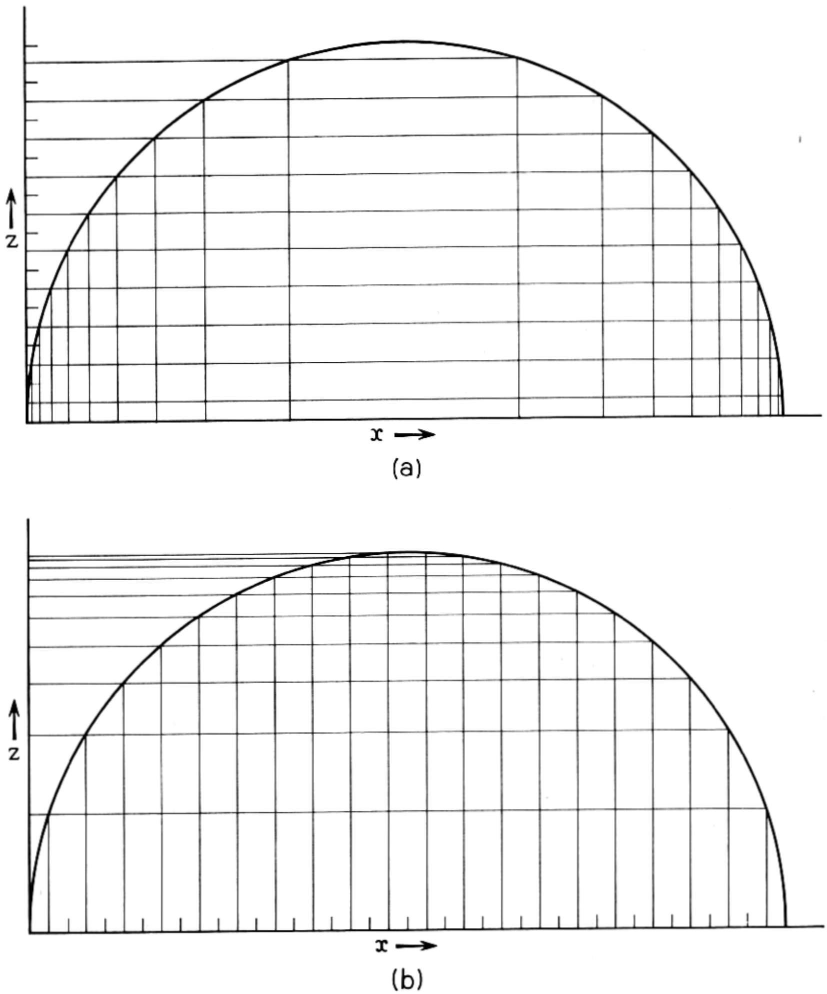

Fig. 3 — (*a*) Surface definition with even z-axis quantization (contour lines); (*b*) surface definition with even x, y plane quantization.

Here the z values are quantized into equal levels, and to any such z level the corresponding x and y values are taken (rounded to their nearest sample). This approximation yields uniformly distributed z values but uneven coverage of the x and y values. If the surface rises sharply toward the observer, the (x,y) points become densely packed, whereas if the surface becomes flat, these points become farther apart, resulting in gaps. Another possible approximation is shown in Fig. 3(b). Here the evenly sampled x and y values are taken and the corresponding z values are determined and rounded to their nearest sample. Hence, the surface to be displayed is defined by a dense covering of points obtained by projection up from the base plane. This second method of approximation is chosen since the object is to be viewed from above, and since the dense covering will result in efficient use of the available stereo picture area. In the case of multiple-valued functions or several functions considered in one display, the projection is made onto the maximum z value, which is the point closest to the observer.

The surface of the object is thus defined but in a rather abstract sense. In order that the object be visible, brightness values must be assigned to every surface point. The use of identical brightness values for all points would yield a surface of homogeneous texture when viewed perpendicular to the base plane. Such a surface would have no patterns, shadows, or brightness changes due to different angles of reflection; that is, it would have neither monocular nor binocular depth cues and thus would be inappropriate for the purpose. Therefore, to obtain depth cues each point must be printed at varying brightness levels. It has been shown[2] that stereo picture pairs comprised of points of random brightness and thus devoid of all cues except binocular parallax can be perceived in depth when fused. Therefore, it is sufficient to assign randomly to each point (x,y,z) a brightness level. The brightness selection on a random basis is a simple procedure, eliminating any consideration for appropriate monocular cues, and has the further advantage of avoiding periodicities and regions of ambiguities. That is, a point domain seen by the left eye may be fused with any periodically repeating domain seen by the right, if such exists, thus producing confusion as to the correct binocular parallax. Therefore, random brightness patterns are used to produce unique point domains which can be fused unambiguously. The question of how many brightness levels to use in the random selection is answered by the requirements of the system of output to be used. However, the use of few levels increases the probability of occurrence of any one level, and clusters of points of equal brightness can produce areas of indeterminate depth on the surface to be viewed. For a photographic output procedure requiring a small number of levels, it would be desirable, therefore, to apply rules to the random selection which would prevent these clusters. Also, a useful monocular cue could be provided by regulating the occurrence of certain brightness levels in a manner dependent upon the z-level of the point domain. Thus, there are many possible refinements to the basic procedure of giving texture to the surface by randomly assigning brightness levels.

The object has been defined and its surface has been invested with brightness levels, albeit random and uninformative when viewed monocularly. It remains now to produce a stereo pair in which the point domains are given the proper parallax shift. The calculations for two such pictures follow the simple formulas for projection, which are shown in Fig. 4. The center of projection is considered at a distance H from the base plane of the object. The centers of projection for the stereo picture pair are separated by a base distance B and are positioned symmetrically about the z-axis. The plane of the pair is at a distance F from the centers of projection. The projections for each point (x,y,z) of the surface where $z = f(x,y)$ onto the left and right members of the pair are then given by the relations

$$x_L = \left(x + \frac{B}{2}\right)\cdot\frac{F}{H - z} \qquad \text{and} \qquad x_R = \left(x - \frac{B}{2}\right)\cdot\frac{F}{H - z}$$

$$y_L = y\cdot\frac{F}{H - z} \qquad\qquad y_R = y_L$$

The total parallax for the point (x,y,z) is seen to be $\Delta = BF/(H - z)$ and is shared equally by the two pictures.

It should be pointed out that binocular parallax alone constitutes only the perception of relative depth. Without other depth cues it is not possible to determine absolute depth when fusing the pair obtained by the above projections. That is to say, the perceived z-scaling, which is some monotonic function of $\Delta = \text{const.}/(H - z)$, is obtained by some arbitrary selection for the value H. (In stereoscopic viewing the supplementary depth cues determine the absolute distance of the plane of the stereo pictures from the observer, which is subjectively substituted for H.) If the function $\Delta = \text{const.}/(H - z)$ is used, the parallax shifts will be similar to those experienced by the human optical system and thus will give rise to familiar percepts of z-scaling. Inasmuch as the perceived depth is some monotonic function of the binocular parallax, which is in turn a monotonic function of the height of the surface, it suffices to choose any monotonic function $\Delta = f(z)$. For example, if the range of z is limited, the function $\Delta = z$ gives a good approximation to

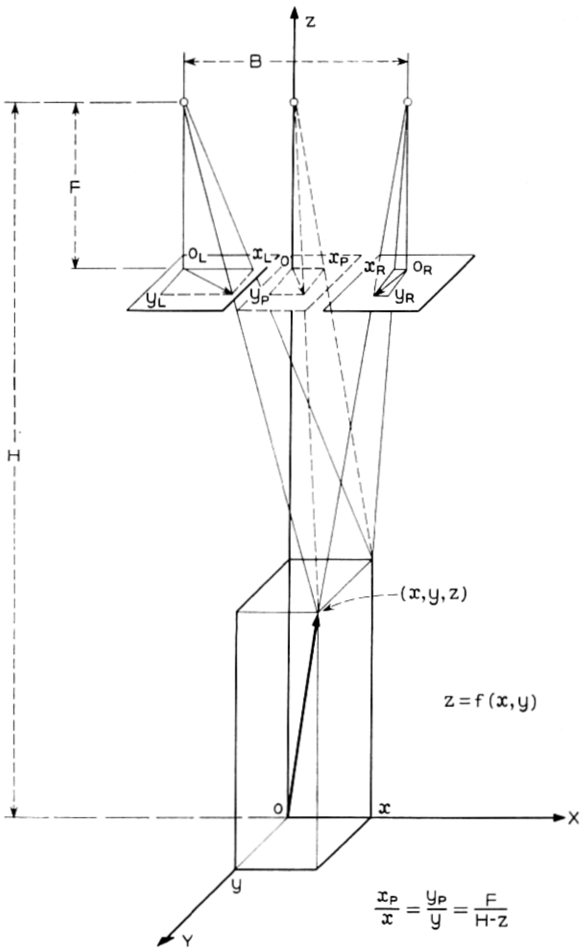

Fig 4 — Projection of an object onto a stereo pair.

the projection of Fig. 4. This function merely provides a different subective z-scaling. If a numerical z-scale is provided, which can be perceived in depth together with the surface to be presented, and if both are generated according to the same projection rules, then the problem of a correctly labelled stereoscopic projection is solved.

Consequently, the parallax shift can be computed, having selected a function $\Delta = f(z)$, which gives a new position to each point, and an identical brightness level can be assigned at random to the corresponding points in the left and right fields. The stereo pair then results by use of a suitable output medium, that is, a video transducer in which the x,y positions correspond to the deflection, and in which the brightness values correspond to the intensity of the beam in a cathode-ray tube display. Inasmuch as a digital system is used, the video transducer generates a sampled display. Since the projection can produce expansions and contractions of point domains on the surface or parallax shifts which are not integral multiples of the sampling intervals, over-sampling is required. This, however, results in great strain on the storage capacity of available computers and on the resolution requirements of video transducers. Therefore, it is necessary to make compromises by trading resolution in object definition for resolution in depth. In the present state of technology, however, there are devices available which will satisfy this requirement.

III. INSTRUMENTATION

The above steps were carried out by quantizing the base plane into 10,000 points with the scale on the x and y axes running from 1 to 100. An IBM 7090 computer was used to generate an array of corresponding values $z = f(x,y)$ and to assign a random number designating the brightness for each of the points. For simplicity, the function $\Delta = z$ was chosen for the parallax shift instead of the geometric projection and was applied in the x-direction only. That is, the coordinates of the point x,y in the left and right pictures were

$$x_L = x + z/2 \qquad\qquad x_R = x - z/2$$
$$\text{and}$$
$$y_L = y \qquad\qquad y_R = y.$$

This corresponds to projecting the stereo pair with a cylindrical lens, the axis of which runs parallel to the y-direction. This position and brightness information was then written on digital magnetic tape and put into a General Dynamics S-C 4020 microfilm printer, which served as the output device for the stereo pair. Different brightness levels were achieved by randomly employing each of the sixty-four type characters available on the microfilm printer. The variation in density of each of the characters gave sufficient variation in brightness level and provided an efficient means for plotting brightness information. The grid size of the microfilm output was 1024 x 1024 and provided, therefore, an oversampling of ten to one for the chosen picture size. This oversampling gave enough stereo resolution for most applications. In order that the type characters did not overlap and totally obscure each other, the maximum shift permitted between two points was taken to be six micro-positions. This limited the total parallax shift to twelve units. That is, the maximum angle of rise between two points on the surface which could be displayed was approximately 85°. The total range of the surface was further restricted by the locations of the peaks and valleys relative to the side boundaries of the grid. For the examples to be shown, a scaling on the z-values of about 60 levels running from -30 to $+30$ was chosen.

An alternative semi-automatic method of output using an optical system was also investigated. This technique resulted in the production of a solid model of the surface to be displayed, which was then photographed by a stereo camera to obtain the desired picture pair. The model was prepared in layers by printing the points belonging to each of the quantized z-levels on transparent glass slides as black and white dots. The slides were then stacked together in register to form a solid cube, where the width of the glass plates determined the scale factor for the z-axis. The prints for each level were obtained by writing the picture information on magnetic tape in digital form as computer output and by using a digital-to-analog converter and a slow-speed television monitor to produce oscilloscope displays, which were then photographed.[3,4,5] A secure mounting of the stack of glass slides in which the entire stack remained transparent was achieved by making an air-tight seal between each plate with a polyester resin having an index of refraction sufficiently near that of the glass.* This technique results, therefore, in a stereo pair in which the parallax shift corresponds to the geometric projection, and furthermore, gives rise to a solid model which is a desirable by-product.

IV. RESULTS

Example 1, generated and displayed automatically, is shown in Fig. 5, and can be perceived in depth when viewed stereoscopically. Viewing may be facilitated by use of Fresnel lenses accompanying the article cited in Ref. 2. The following three surfaces are presented:

1) the hyperbolic paraboloid,

$$\left(\frac{x - 50}{30}\right)^2 - \left(\frac{y - 50}{20}\right)^2 = \frac{z}{30},$$

with saddle point at (50, 50, 0),

* The slide mounting techniques were developed by R. A. Payne of Bell Telephone Laboratories.

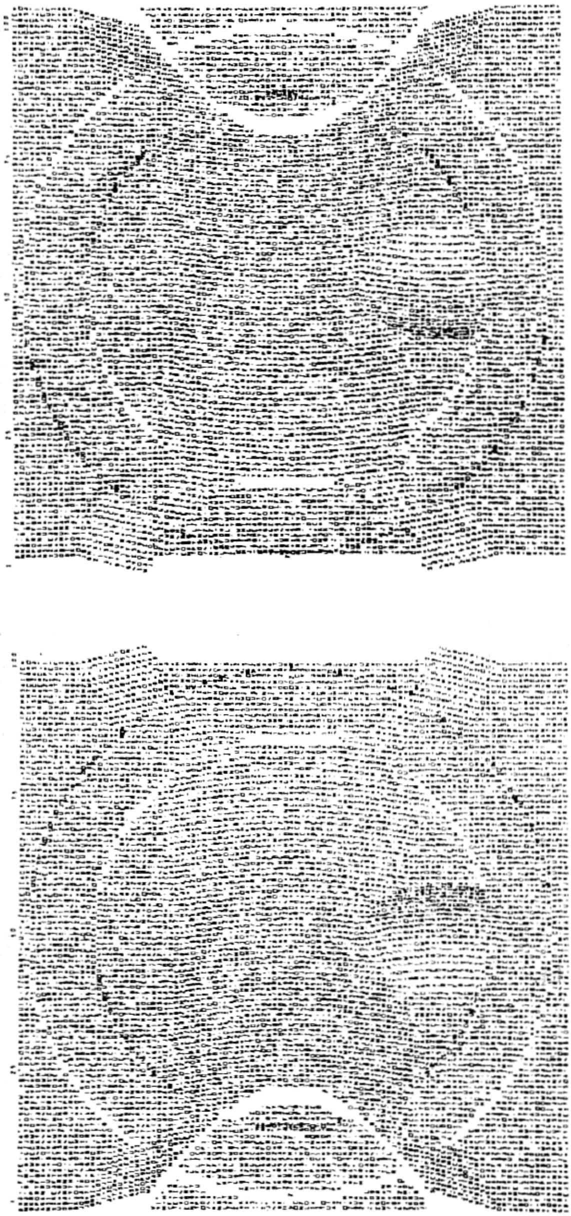

Fig. 5 — Automatic stereoscopic presentation of a function (Example 1).

2) the elliptical paraboloid,

$$\left(\frac{x-50}{10}\right)^2+\left(\frac{y-79}{20}\right)^2=\frac{-(z-20)}{50},$$

with vertex at (50, 79, 20), and

3) the torus,

$$\left(\sqrt{(x-50)^2+(y-50)^2}-42\right)^2+z^2=6^2,$$

centered at (50, 50, 0) and having a radius of 6. (Conventional two-dimensional displays of Example 1 were given in Figs. 1 and 2.)

The reduction required for reproducing the stereo pictures here has made the resolution of individual type characters very difficult. For this reason, a presentation in depth of the numerical z-scale was omitted. However, a very effective display can be achieved, including the z-scale, by using a larger picture size.

In Fig. 6 the same surface is displayed by the optical method. Two

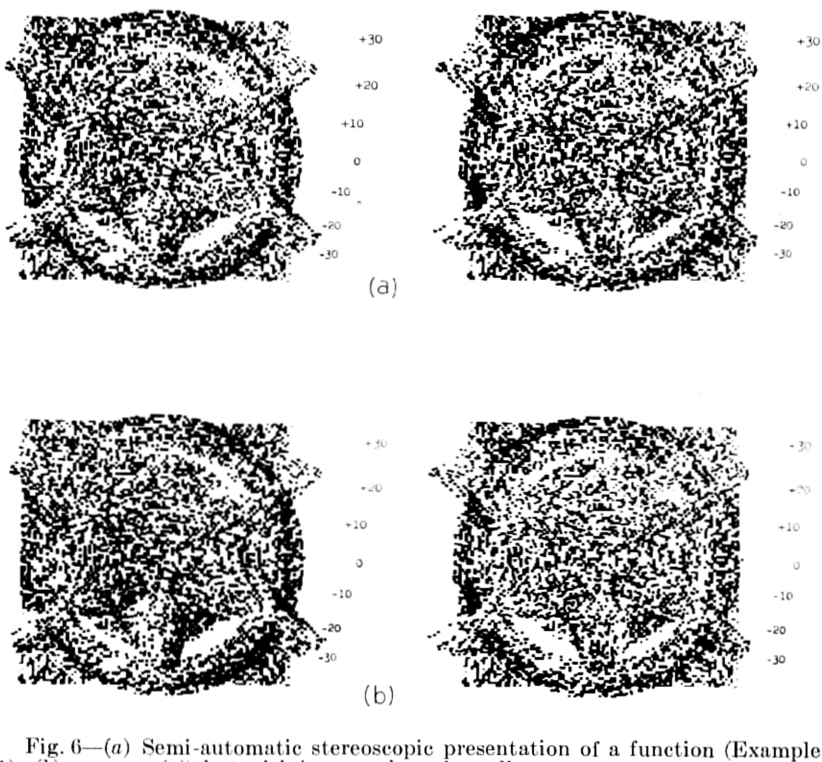

Fig. 6—(a) Semi-automatic stereoscopic presentation of a function (Example 1); (b) same as (a), but with increased z-axis scaling.

views are presented to show that variable scaling on the z-axis is possible. The amount of depth is determined by both the width of the glass plates and the base distance between the two lenses of the stereo camera (or distance between two positions of a single-lens camera). Fig. 6 (b) was produced with greater base distance, and consequently the surface has greater stretching in the z-direction. The numerals on the right edge of the displays indicating the z-levels were applied to the appropriate slides by hand and were not generated as picture material. It will also be pointed out that in this method, the points of the two pictures are more clustered and less uniform in distribution. This demonstrates the expanding and contracting of point domains produced by the transformation of projection and provides a helpful monocular cue. All displays are far more evenly filled with brightness elements, however, than if the contour-line method had been used.

Example 2 shown in Fig. 7 is that of a spiral given by the parametric equations

$$x = \rho \cos \theta + 50.5$$

$$y = \rho \sin \theta + 50.5$$

$$z = \frac{10}{\pi} \theta - 30$$

with

$$0 \leqq \theta \leqq 6\pi$$

$$\rho = 50 - \frac{15}{2\pi} \theta.$$

This presentation is another display from the microfilm printer, illustrating the procedure in its completely automatic form. Approximately one minute of time is required to generate and display the stereo information.

V. SUMMARY

A method for automatically presenting three-dimensional information in depth has been described. The advantages are threefold in that (*i*) such presentations make possible displays which are very difficult if not impossible to obtain by other means, (*ii*) they carry the spatial impact enjoyed by human observers, (*iii*) and they are suitable for the printed page. The technique has been outlined in three steps: definition of surface, texturing of the surface with brightness elements, and generation

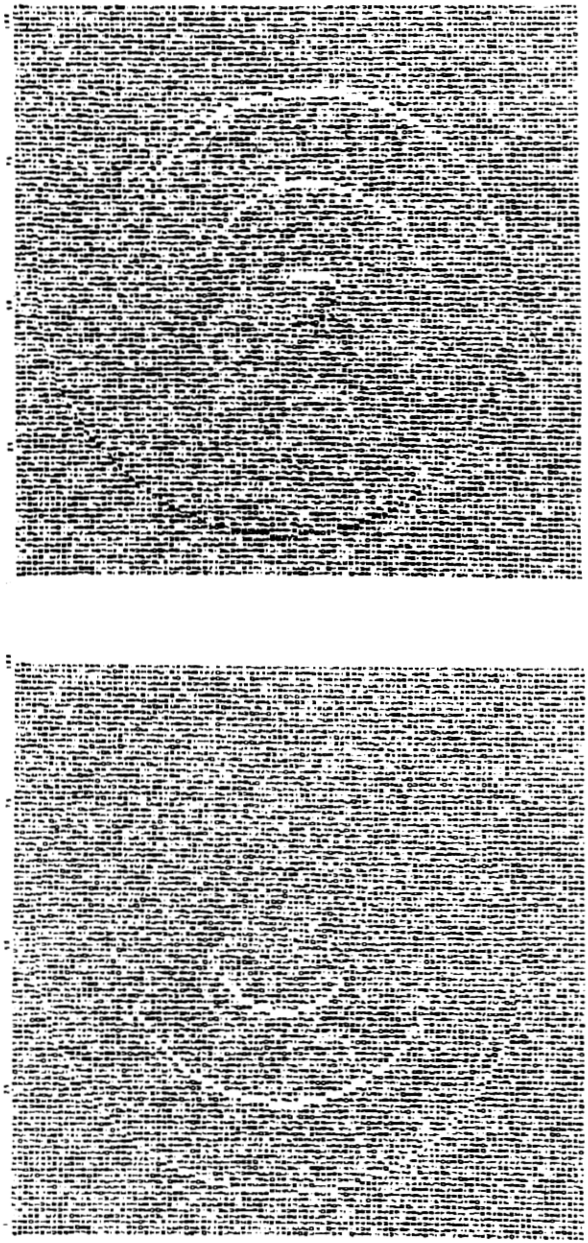

Fig. 7 — Automatic stereoscopic presentation of a function (Example 2).

of stereo projections of the surface. By use of digital computers and the special-purpose output devices now available, this procedure can be carried out in a completely automatic fashion, thus making possible a simple and effective demonstration of three-dimensional data.

REFERENCES

1. MacKay, D. M., A Simple Multi-Dimensional CRT Display Unit, Electronic Engineering, **32,** London, June, 1960, pp. 344–347.
2. Julesz, B., Binocular Depth Perception of Computer-Generated Patterns, B.S.T.J., **39,** Sept., 1960, pp. 1125–1161.
3. David, E. E., Jr., Mathews, M. V., McDonald, H. S., Experiments with Speech Using Digital Computer Simulation, I.R.E. Wescon Convention Record-Audio, Aug., 1958, p. 3.
4. Graham, R. E., Kelly, J. L., Jr. A Computer Simulation Chain for Research on Picture Coding, I.R.E. Wescon Convention Record — Computer Applications, Aug., 1958, pp. 41–46.
5. Julesz, B., A Method of Coding Television Signals Based on Edge Detection, B.S.T.J., **38,** July, 1959, pp. 1001–1020.