SOME INFORMATIONAL ASPECTS OF VISUAL PERCEPTION

FRED ATTINEAVE

Perceptual and Motor Skills Research Laboratory, Human Resource Research Center

In 1952, Charles W. Eriksen presented his research on the idea that information theory can be applied to visual perception. He argued that the way in which information is perceived and processed can be understood through the lens of information theory. This approach allowed for a more quantitative understanding of how visual information is processed by the human brain.

In this paper, Eriksen discussed the concept of stimulus uncertainty and its role in the process of visual perception. He proposed that the uncertainty in a visual stimulus can influence the way it is perceived and processed by the brain.

The implications of this research are significant for various fields, including psychology and neuroscience. It provides a framework for understanding how visual information is processed and how it can be manipulated to improve visual perception.

The Nature of Redundancy in Visual Stimulation: A Demonstration

In this section, Eriksen demonstrates the concept of redundancy in visual stimulation. Redundancy refers to the presence of multiple redundant stimuli that can influence the way in which visual information is processed. This can have implications for tasks that require attention and focus, such as driving or flying.

In conclusion, Eriksen's research on visual perception has laid the foundation for further research in this field. It has opened up new avenues of investigation and has provided a framework for understanding how visual information is processed by the human brain.
redundancy of printed English. We may divide the picture into arbitrarily small elements which "transmit" to a subject (S) in a cumulative sequence, having him guess at the color of each successive element until he is correct. This method of analysis resembles the scanning process used in television and audio-visual systems, and accomplishes the like purpose of transforming two spatial dimensions into a single sequence in time. We are in no way supposing or assuming, however, that perception normally involves any such scanning process. If the picture is divided into say, 50 rows and 80 columns, as indicated, our S will guess at each of 4000 cells as many times as necessary to determine which of the three colors it has. If his error score is significantly less than chance (2/3 X 4000 = 1/3 X 4000 = 4000), it is evident that the picture is to some degree redundant. Actually, he may be expected to guess his way through Fig. 1 with only 15 or 20 errors. It is fairly apparent that the technique described, in its present form, is limited in applicability to simple and somewhat contrived situations. With suitable modifications it may have general usefulness as a research tool, but it is introduced into the present paper for demonstrational purposes only.

Let us follow a hypothetical subject through this procedure in some detail, noting carefully the places on the page he is most likely to make errors, since these are the places in which information is concentrated. To begin, we give him an 80 X 50 sheet of graph paper, telling him that he is to guess whether each cell is white, black, or brown, starting in the lower left corner and proceeding across the first row, then across the second, and so on to the last cell in the upper right corner. Whenever he makes an error, he is allowed to guess a second and, if necessary, a third time until he is correct. He keeps a record of the cells he has been over by filling in black, brown, or white with pencil marks of appropriate color, leaving white ones blank. After a few errors at the beginning, the first row, he will discover that next cell is "always" white, and proceed accordingly. This prediction will be correct as far as Column 20, but if it will be wrong. After a few errors he will learn that the best prediction, as in fact it is at the end of the row. Chances are good that the subject will assume the prediction to be exactly like the first, in which he will guess it with no error; or, if he may make an error or two, the beginning, or at the edge of the table, as before. He is also expected to be entirely correct on Row 3, on subsequent rows through 20. Row 21, however, it is equally likely that he will eventually predict a violation from white to brown on Column where the corner of the table is.

Our subject's behavior to this point demonstrates two principles which, if discussed before we follow through the remainder of his predictions. It is evident that redundant stimulation results from an area of homogeneous color ("color in the usual broad sense, includes brightness"). Or (A color) with homogeneous direction or slope, other words, information is concentrated along contours (i.e., regions where changes abruptly), and is further concentrated at those points on a contour at which its direction changes (at ordinary angles, only, i.e., at angles of 90 degrees). The "scanning" procedure introduces an advantage here, in that a particular cell may be seen as a linear feature first time it is seen. It is obvious that if the starting point of the sequence, and the direction of scan were ordinarily known, a large number of subjects, not even a relatively small number of subjects, would be systematically scanned along such a straight contour.

Evidence from other and equally diffuse situations supports both of these conclusions. The concentration of information in contours is illustrated by the striking similarity of appearance of objects in contour and different views. The "triangle" triangle for example, may be either white on black or black on white. Even more impressive and familiar fact that an artist's mask, in which lines are substituted for homogeneous color, may constitute a readily identifiable representation of a thing or thing.

Our experiment relevant to the second conclusion, i.e., that information is further concentrated at points where a change direction most rapidly, is summarized briefly. Eighty subjects were instructed to draw, for each of 16 shapes, a pattern of 10 dots which would resemble the shape as closely as possible, and then to indicate the original outline the exact places where the study has been previously published. The form of a minigraphic note: 1. Relative Importance of Parts of a Contoured Shape. Also, Human Research Center, November 1931.

which the dots represented. A good sample of the results is shown in Fig. 2; in such a figure, the dots are connected by straight line. This conclusion is verified by detailed comparisons of dot frequencies with measured curvatures on both the figure shown and others. Common objects may be represented with great economy, and fairly striking fidelity, by copying the points at which the contours change direction most rapidly, and then connecting these points appropriately with a straight line. Figure 3 was drawn by applying this technique, as mechanically as possible, to a single hemisphere. The information content of a drawing like this may be considered to consist of two components: one describing the positions of the points, the other indicating which points are connected with which others. The first of these components will almost always contain more information than the second, but the exact share will depend upon the precision with which the positions are designated, and will vary from object to object.

A question we would like to ask is this: to what extent can we express the area of a figure by a set of points, connected in such a way that the shape is preserved? This question is of considerable interest, and we attempted to answer it by an experiment of the following kind. We took a set of 100 points, and asked subjects to connect these points in such a way that the shape was preserved. The results of this experiment are shown in Fig. 4. It is evident that the shape is preserved quite well, but that the points are not connected in a straight line.
of the table and the link bottle in Fig. 1. His errors will follow the principles we have just been discussing until he reaches the serrated shoulders of the bottle. (A straight 45° line would be represented in this way because of the grain of the coordinate system, but we shall consider that the bottle is actually serrated, as it is from the subject's point of view.) On the left shoulder there are 12 right angles, but these angles contain considerably less than 13 times the information at a single angle in isolation like the corner of the table. This is true because they fall into a pattern which is repetitive, or redundant, in the everyday sense of the term. They will cease to evolve soon as 5 perceives their regularity and generalizes this extrapolation, precisely like 5's previous extrapolations of color and slope, will have validity only over a limited range and will itself lead to error on Row 38, Column 48. At about the same time that he discovers the regularity of the stair-step pattern (or perhaps a little before), our 5 will also perceive that the ink bottle is symmetrical, i.e., that the right contour is predictable from the left one by means of a simple reversal. As a result he is very unlikely to make any further errors on the right side above Row 32 or 33. Symmetry, therefore, constitutes another form of redundancy.  

It should be fairly evident by now that some reader may be confused by the fact that all subjects have actually been run on the task described. Their scores, as indicated by the number from 1 to 10, were distributed as suggested above, with a single interesting exception: 4 of the 5 assumed on Row 7 that the brown area would be located symmetrically within the field, and guessed "white" on Column 61. By the use of Shannon's formula (17) it was estimated that the entropy contained between 14 (lower limit) and 18 (upper limit) bits of information, in contrast to a possible maximum of 1260 bits. The redundancy was thus estimated to be between 97.5 and 98.7 per cent.

that many of the gestalt principles of perceptual organization pertain essentially to information distribution. A good Gestalt is a figure with some degree of internal redundancy. Subject A's performance suggests that the grouping laws of similarity, continuity, and common fate all apply to conditions which reduce the information available to the subject; at least when it is clear enough after the preceding censure, and we shall presently observe that similar information in an efficient manner would necessitate the solving of something of the sort. Mussett's studies have very close to the potential of this extrapolation, precisely like 5's previous extrapolations of color and slope, will have validity only over a limited range and will itself lead to error on Row 38, Column 48. At about the same time that he discovers the regularity of the stair-step pattern (or perhaps a little before), our 5 will also perceive that the ink bottle is symmetrical, i.e., that the right contour is predictable from the left one by means of a simple reversal. As a result he is very unlikely to make any further errors on the right side above Row 32 or 33. Symmetry, therefore, constitutes another form of redundancy.  

It should be fairly evident by now that some reader may be confused by the fact that all subjects have actually been run on the task described. Their scores, as indicated by the number from 1 to 10, were distributed as suggested above, with a single interesting exception: 4 of the 5 assumed on Row 7 that the brown area would be located symmetrically within the field, and guessed "white" on Column 61. By the use of Shannon's formula (17) it was estimated that the entropy contained between 14 (lower limit) and 18 (upper limit) bits of information, in contrast to a possible maximum of 1260 bits. The redundancy was thus estimated to be between 97.5 and 98.7 per cent.

The abstraction of statistical parameters  

Although Fig. 1 presents a situation much simpler, or more redundant, than the visual situations which ordinarily confront us, the reader need merely look around the room in which he is sitting to find that the principles illustrated apply to the real world. Further, it may be argued on neurological grounds that the human brain could not possibly utilize all the information provided by states of stimulation which were not highly redundant. According to Polyn- skiy's (14) estimate, the retina contains not less than four million cones. At any given instant each of these cones may be in either of two states: firing or not firing. The retina as a whole might be in any one of about 2*400 or 10^12000 states, each representing a different configuration of visual stimulation. Now, if by some unspecified mechanism each of these states were to evoked a different unitary response, and if a unitary response consists merely of the firing of a single unique neuron, then 10^12000 of such response-neurons would be required. The fantastic magnitude of this figure becomes somewhat apparent when one calculates that only about 1074 neurons could be packed into a cubic light year. The fact that the number of patterns of response-neurons might plausibly equal the number of retinal configurations simplifies matters only if there are certain one-to-one connections between cones and response-neurons, in which case the response is to some degree merely a copy of the stimulus.

We may nevertheless ask: would how an observer respond to a situation in which the retinal receptors were stimulated quite independently of one another? This situation would be in prac-

ice very difficult to achieve (even more difficult than its diametric opposite, the
tion of the random field suggested above might look at some particular limit. Perhaps the most striking thing about the figure is the subjective impression of homogeneity that it gives: the half of the figure seems, at least in general way, very much like the right half. This is remarkable because there have been previously associated homogeneity with redundancy, and Fig. 4 was constructed to be completely nonredundant. Now, in psychological terms, it is far clearer than the characteristic way in which the figure appears homogeneous is what Gibson (6) would call a percept. In physical terms, and the factors may be specified: (a) probability that any cell will be black rather than white, and (b) size of cells. Both of these factors probably contribute to perceived texture, which is undoubtedly a functional variable, though the latter is somewhat the more important. The figure is viewed from a sufficient distance, these two parameters being identifiable with (a) the central tendency, and (b) the dispersion, of a continuous brightness distribution in dimensions. It appears, then, that when some portion of the visual field contains a quantity of information grossly in excess of the observer's perceptual capacity; it treats those components of information which do not have redundancies between them as a statistician treats "error variance," averaging out particular blocks and abstracting certain statistical homogeneities. Thus an averaging process was involved in drawing the cat. It was said earlier that the points of the drawing corresponded places of maximum curvature on contour of the cat, but this was strictly correct; if the principles had not been followed rigorously, it would not have been necessary to represent the body individual hairs by points. In observing a cat, however, one does not ordinarily perceive its hairs as individual entities; instead one perceives that the cat is furry. Furriness is a kind of texture, and the statistical parameters which characterize it presumably involve average shapes and directions, as well as number, of elements. The perceived course of a cat (e.g., the contour from with the points of Fig. 3 were taken) is the result of an orthogonal averaging process in which texture is eliminated and in some way as if a photograph of the object were blurred and then printed on the contrast paper (cf. Kovesky, 15 and Colburn, 5). The sense in which a surface of a particular texture may be said to provide redundant stimulation has perhaps been adequately described. This sort of redundancy might be demonstrated by (a) guessing-game technique, with a considerable modification in the level of prediction required, i.e., by increasing the unit area to be predicted and requiring the subject to select from a multidimensional array of samples the texture (i.e., the statistical parameters) he believes the next unit will have. In view of Gibson's (6) convincing argument that a physical edge, or a contour, is as likely to be represented as an object by an abrupt texture change as by an abrupt color change, I have considered it important to show how texture may be substituted for color without materially altering the principles derived from Fig. 1.

**PERCEPTION AS ECONOMIC DESCRIPTION**

It is sometimes said that the objective science is to describe nature economically. We have reason to believe, however, that some such process of parsimonious description has its beginnings in a fairly naive perceptual level, in scientists and their fellow organisms alike; thus the difficulty, mentioned earlier, of distinguishing between perception and inductive reasoning. It appears likely that a major function of the perceptual machinery is to strip away some of the redundancy of stimulation, to describe, or encode incoming information in a form more economical than that in which it impinges on receptors.

If this point of view is sound, we should be able to generate plausible hypotheses as to the nature of specific perceptual processes by considering rational operations which one might deliberately employ to reduce redundancy. The approach suggested, as it applies to the perception of a static visual field, is equivalent to that of a communications engineer who wishes to design a system for transmitting pictures of real things over a practically noise-free channel with the utmost economy of channel time and bandwidth, but in a manner designed to meet standards such as human observers are likely to have. Some of the reduction principles which he might usefully employ in such a system are listed below. It will be found that these principles serve to summarize and integrate ideas which have been developed somewhat informally in the foregoing sections, as well as to introduce new considerations. The principles may be grouped according to the forms of redundancy with which they are concerned: thus I-4 deal with varieties of continuous regularity; S and 6 with discontinuous regularity, or recurrence; 7-9 with proximity; and 10 with situations involving interaction.

1. An area of homogeneous color may be described by specifying the color and the boundaries of the area over which it is homogeneous. (It is assumed that limits of error tolerance on relevant dimensions have been agreed upon, e.g., that there is some definite number of
colors from which the receiving mechanism may be directed to choose.

Likewise, an area of homogeneous texture may be described by specifying the statistical parameters which characterize the texture and the boundaries of that area, even if all the parameters are relatively invariant. Thus, if Fig. 4 represented a part of the upholstered surface of a chair, we would not necessarily have to simply instruct the receiving mechanism to reproduce the texture by filling in cells of a certain size at any one of a large number of random points. It is true that this process might result in the complete loss of 19,600 bits of information; the essential point is that we are dealing here with a class of stimuli from which such a huge information loss is perceptually tolerable.

An area over which either color or texture varies according to some regular function may be described by specifying the function and the boundaries of the area over which it obtains (cf. Gibson's [6] texture gradient). This principle actually implies both 1 and 3 as special cases.

Likewise, if some segment of an area boundary is either continuous or other maintains a constant direction or varies according to some other regular function, it may be described by specifying the function and the loci of its limiting points. Figure 3 illustrates a special case of the process.

If two or more identical stimulus patterns (these might be either successive portions of a continuous or discrete objects) appear at different places in the same field, all may be described by specifying one and specifying the positions of the others and the fact that they are identical (cf. similarly as a grouping law).

If two or more patterns are similar but not identical, it may be economical to proceed as in 3, in addition specifying either (a) how subsequent patterns differ from the first, or else (b) how each pattern differs from some common base pattern which includes the communalities of the group (cf. the 'schema-wifthing' idea discussed by Woodworth [20]; also Hebb's [7] treatment of perceptual schema).

When the spatial loci of a number of points are to be described in an arbitrary order, and the locations are arranged in clusters or proximity groups (as in Fig. 5), it may be economical to describe the positions of the groups rather than the individual points. The order of points within each group is immaterial, and the information required is that the groups are distinctly separate, the distances between adjacent points in each group are minimized, and transmitted with each point serving as origin for the one following it. This procedure may reduce some saving if the points are scattered, as in Fig. 5, but it is most clearly applicable when the points are coming-out in some obvious sequence. In the latter case, a further economy may be achieved by the use of special coordinates such as distance from group boundaries, grouping through the two preceding points (or from an arc through preceding points, etc.; cf. 4 [20]).

Certain areas and objects may be described in a relatively simple way, procedures of the sort suggested above, if they are first subjected to some systematic distortion or transformation. Consider the case of a complex, symmetrical, two-dimensional pattern viewed from an angle such that its retinal or photographic image is not symmetrical. It will be economical to transmit a description of the pattern as if it were in the frontal plane, and thus symmetrical (eliminating the redundancy of symmetry by means of (a)), together with a description of the transformation which relates the frontal aspect described to the oblique aspect in which the pattern is viewed (cf. Gibson's [6] discussion of perspective transformations; also the "Thompsonian coordinates" of D'Arcy Thompson [19]).

Koffka (16) and other gestalt psychologists have held that many objects have some "preferred" aspect, and that this aspect has the characteristics of a "good gestalt." The present principle supports this view on functional grounds, since the perceptual transformation of a figure to an aspect in which similarities among parts are maximized may be interpreted as the initial step in an efficient information-digesting process. It should be clearly recognized, however, that an overall economy is achieved only if the amount of information required to describe the transformation is less than the amount of information saved by virtue of the transformation; that a transformation must be relatively simple to be considered useful, at least by the present criterion.

Let us indicate briefly how these considerations may be integrated with others of a more general nature. Interdependencies among sensory events may exist either in space or in time, or they may cut across both space and time. In studying the redundancy of spoken English (11), for example, one is dealing with interdependencies which may be considered purely temporal. The present discussion has been restricted, quite arbitrarily, to relationships in space: to the problem of redundancy and information distribution which may obtain in the
visual field at a particular instant, and which a computer of conceivable com- plexity might evaluate from a photo- graph. The extension of the visual field in time, which I propose to discuss in a subsequent paper, introduces new varieties of redundancy among the temporal continuation or recurrence of spat- tial-configurational which may be non- redundant at any instant considered in isolation. Any individual learns a great deal, over his life span, about what- ever-it-contains. Thus, if an eye is dis- closed in a situation like that illustrated by Fig. 1, the observers can predict that a mouth, nose, eyes, etc. are also pres- ent, and approximately where they are. This sort of redundancy is spatial-temporal in its basis; predictions are not possible merely on the basis of the present visual field, but depend also upon previous fields which have con- tained faces. Principle 6 also suggests the approach to economical description which might be extended to such cases. Further, as Brunswik (2, 3, 4) has pointed out in some detail, ecological principles of very broad generality may be derived from experience. 7 For ex- ample, the frequency with which an observer has encountered symmetrical objects in his past may certainly affect the point at which, in predicting suc- cessive cells of Fig. 1, he "assumes" that the ink bottle is symmetrical. Likewise in terms of economical encod- ing: each of the varieties of spatial redundancy suggested above will itself occur with some determinate frequency over any given set of fields (e.g., the

1 Brunswik (2), Welch (3), and the Ames group at Princeton (W) have always advanced views concerning the role of expectation in perception which have much in common with the author and with my own position in the matter. It appears to me, however, that they have in general tended to underestimate (as the gestalt psychologists have sometimes over- estimated) the importance of lawful relation- ships which may exist within the static and isolated visual field.


Refereced: