A TRANSFORMATION FOR EXTRACTING NEW DESCRIPTORS OF SHAPE

Harry Blum
Data Sciences Laboratory
Air Force Cambridge Research Laboratories
Bedford, Massachusetts

Introduction

I have approached the problem of shape by assuming that the current mathematical tools were somehow missing essential elements of the problem. For despite more than two millennia of geometry, no formulation which appears natural for the biological problem has emerged. This is not surprising perhaps when one recognizes that geometry has been born of surveying and has grown in close collaboration with physical science and its mensuration problems. A corollary to this position is that there is some central difference between the biological problem that we are trying to deal with and the physical problem that we have been dealing with. Consequently, such an approach requires a restudy of visual function to assess what such a geometry should indeed try to accomplish. Unfortunately, the problem of exploring function is not easy to do in isolation, since the visual world is extremely rich, and hypotheses about visual shapes and their functional value to an organism may reflect the cultural bias of the experimenter to a large degree. I have chosen to enter the problem from the middle by hypothesizing simple shape processing mechanisms, and then exploring together the geometry and visual function that result. One such mechanism is presented in this paper. Since it leads to a drastic reformulation of a number of notions of visual shape, it may be useful to review briefly some of the notions implicit in our views and experiments.

While local attributes of the visual field (intensity, color, edge, angle, motion) and average global properties such as total luminous flux have been subject to progressive experimentation and understanding, the attributes of shape, global properties in which the particular distribution is precisely relevant, are still largely unkown. Yet, shape has been shown to be a particularly stable organization of our visual stimulus by the degree and variety of distortion that are the stock in trade of artists and cartoonists – even being used to enhance recognition. Why, then, has it been so difficult to tease out the properties

of shape? A primary reason is that the number and variety of shapes are enormous, so that it is impossible within the limits of experiment to explore more than a small sampling of the variables. Consequently, one assumes a theory of shape, at least implicitly, in the selection of experimental stimuli. This reason is compounded by the degree to which geometry and reading, both synthetic processes, imposed on us culturally and with considerable difficulty, have biased our view of shape. As a result, stimuli in even the best animal experiments have been largely rectilinear or other unnatural geometric shapes presented in isolation on a uniform field (for example, see Sutherland, 1960). Except for a limited number of terminated experiments (for example, Attneave and Arnoult, 1956), quantitative work with humans has been subject to similar limitations and greater confusion due to the intimate interaction with cultural artifacts. Consequently, the most primitive problems of shape vision, segmentation of the field into objects, defining locations and recognition of biologically relevant amorphous shapes, are implicitly ignored. It is precisely toward these problems that the process considered here is useful.

If it cannot be assumed that a shape can be isolated and normalized a priori, or that the wide variety of perturbations that a shape may take and still be identified, can be exhaustively enumerated, it is necessary to develop some way of obtaining shape properties without becoming enmeshed in the huge combinatorics of simple congruence views. A number of people have proposed to have the shape interact with itself to accomplish this. The Gestalt school (Koffka, 1935, for example) used field theoretic notions. Deutsch (1955), and Bitterman, Krauskopf and Hochberg (1954), have proposed propagation or diffusion models of interaction. Kazmierczak (1960) has used an electric field analogy and Kirsch, Cahn, Ray and Urban (1958) have used computer rules on a rectangular grid to allow a wide variety of self-interactions, including propagation. However, these models failed to arrive at a set of incisive shape attributes. The present model develops these notions somewhat differently and arrives at a number of distinct attributes which are useful in two ways: (a) in developing a physchology of shape for everyday life, and (b) in focusing attention on some of our biases. The model can be developed in a number of ways. The propagation scheme used has been chosen for its tutorial ease. Another view of it, perhaps having more physiological relevance, is introduced later. A word of warning is important. The model does not intend to deal with the central perceptual problem in which the entire history of the organism, its set, and its total stimulus input must be considered. The output of the transformation merely gives a set of shape attributes on which the organism can perform selection, using criteria of usefulness and relevance similar to those used in other sensory processes.

Definition Using Time

Consider a continuous isotropic plane (an idealization of an active granular material or a rudimentary neural net) that has the following properties at each point: (a) excitation – each point can have a value of either 0 or 1; (b) propagation – each excited point excites an adjacent point with a delay proportional to the distance; and (c) refractory or dead time – once fired, an excited point is not affected by a second firing for some arbitrary interval of time. A visual stimulus from which the contours or edges have been extracted, impinges on such a plane at some fixed time and excites the plane at those points. This excitation spreads uniformly in all directions but in such a way that the waves generated do not flow through each other. Figure 1 shows this process occurring for an input pattern of two points. The waves start as two expanding circles and merge into a single contour when these circles intersect. (One can visualize the contours as the front of a grass fire.) Figure 2 shows the generated wave for a number of other simple shapes. This wave passes by each point of the space once and only once. For most points of the space the wave flows by in a good fashion, each point on the wave front generating a new point on a new wave front. This is not the case, however, for those places where corners appear in the wave, or where waves collide with each other in a frontal or circular manner. The wave front undergoes cancellation at these places. These places are shown by the dotted lines in Figures 1 and 2. The two dots of Figure 1 give rise to a line which starts with an infinite velocity and becomes asymptotic to the space velocity. From Figure 2, it can be seen that such points occur only on the inside of an angle or an arc. The corner appears only after the center of curvature is reached. In the triangle three corners start propagating and disappear at the center of the largest inscribed circle. In the case of the ellipse, these points start at the shortest radius of curvature and disappear at the center of the largest inscribed circle. Since both of these shapes are convex, they generate no such

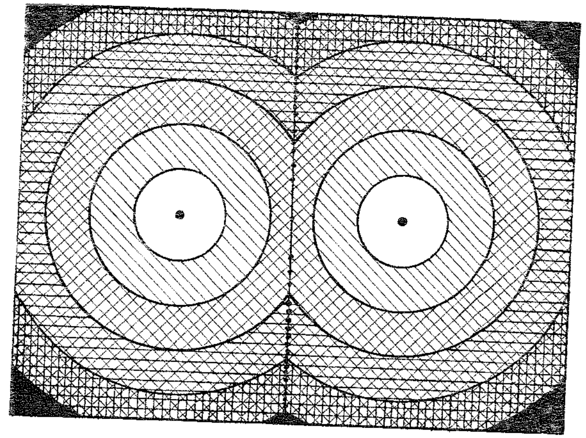

Fig. 1. Wave fronts generated by a two-point grass fire excitation. (Note the absence of flow-through associated with propagation of waves in a pond. The mid-line represents the locus of corners in this process.)

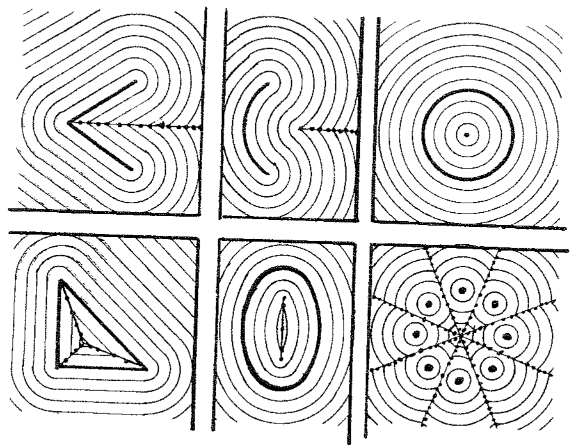

Fig. 2. Wave fronts and "corners" generated by some simple excitations (dotted lines represent locus of "corners").

points outside of themselves. A straight line and a point, since they are convex and have no inside, have no such points unless a special definition is used to include them. The circle is a special case whose description is a single point. The series of dots in a circle is included to show a closure property of the process. By setting a velocity threshold (this is equivalent to a smoothness criterion of the wave front), one sees that it is possible to have these points disappear and reappear at the center. In this way, a circle and a series of dots in a circle can be equated. Clearly, under these conditions, the process is not dimension preserving. Let us examine the process more carefully since these points along with the times or distances associated with them form an alternate description of a shape.

Figure 3 consists of 2 circular arcs and the tangents connecting them. A corner does not immediately appear in the wave front since it is a smooth figure. A corner appears at time t_1, propagates to the right and disappears at time t_2. Each stage of these expanding waves represents a figure parallel to the original one. The appearance of a single corner occurs at the minimum radius of curvature of the figure. The disappearance of a corner represents a largest circle that can be drawn in the figure. All the corner points have a velocity associated with them that is determined by the angle formed — the sharper the angle, the faster the velocity. Since the parallel figures generated in Figure 3 all have the same angle, the velocity along this central line is a constant. An accelerating point would result from curvature toward this locus; a decelerating point would result from curvature away from this locus. The locus of corners will be called the "medial axis" (MA). When the times of corner occurrence on the medial axis is included, it will be referred to as the "medial axis function" (MAF).[1] As long as the MA is a straight line, the figure is symmetrical. Clearly an arbitrary parameter, such as parallel distance, could be substituted for time.

The transformation from the original pattern to the medial axis function description is unique and hence invertible. It can be seen that the original pattern can be described by the envelope of circles of proper radius (MAF time divided by the space velocity) associated with each point. This can be done more elegantly by the following inversion procedure. Excite the space along the MA at a time defined by the negative of the MAF. Thus, the later points of the function are pre-excited and allowed to expand by the proper amount before the earlier

---

[1] This locus has also been called the "skeleton" and the function on the skeleton, the "skeletal function."

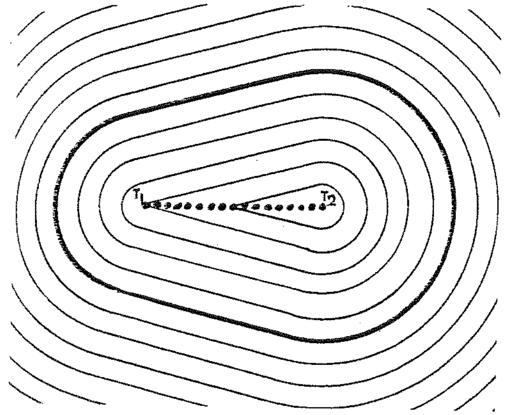

Fig. 3. Generation of new contour description (medial axis function) centered on enclosed space. (t_1 is appearance of corner, t_2 is disappearance. The locus of points and their times are required.)

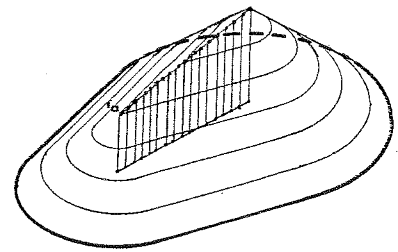

Fig. 4. A three-dimensional static alternative to the two-dimensional kinetic view of the process. (The MAF is the ridge formed where the union of cones on the input contour intersect each other.)

points of the function are excited. This procedure uses a reciprocity property like that of electromagnetic theory.

Alternate Definitions

There are a number of alternate ways of looking at the transformation, each of which sheds a somewhat different light on the process. Consequently, it is worth describing them. Figure 4 shows a formulation proposed by Kotelly (1963) in applying an earlier version of this process to the problem of pattern masking (metacontrast). Here, vertical distance is substituted for time. Instead of a propagating wave, one obtains a representation as a surface which is the envelope of cones whose apexes sit on the original contour. The MAF is then the locus of discontinuities in this surface. Still another way of looking at the process is from the viewpoint of field theory. The wave can be considered as a "nearest distance" field. The MAF is then the locus of discontinuities in the derivative of that field. Such a field departs sharply and in important ways from the conventional fields of physics and psychology, in which the value at a point is an integral of a distance function to the total input. Still one more, and I feel particularly interesting definition of it, is gotten from an equidistance viewpoint. Consider the (minimum) distance from each point of the plane to the stimulus shape. For most points this distance is unique to one point on the shape. For points on the MA, however, this distance is not unique to one point on the stimulus. Consequently, the MA is the locus of points equidistant from the pattern and in this sense represents a line of symmetry of the pattern. The MAF can be considered a symmetrical central description of the space whose boundary is the stimulus contour. It has been particularly satisfying to me to find that the process can be expressed in so simple a geometric notion. If one considers contours as coming from a world of objects, then these original contours are extremities of objects or transition points between objects. For even so simple a notion as location of an object, the contour or perimeter description is particularly poor.

Properties

Let us see what kind of categorizations are natural to such a process by considering a number of smooth, simply connected closed curves. The MAF of the circle and the ellipse were shown in Figure 2. Figure 5 shows the MAFs for a number of more undulating curves. The "pinch" in the upper left figure generates a double set of MA points which propagate in opposite directions. The pinches in the lower left "3-bulb" do the same. The outward moving points disappear, indicating closure; the inward points form a triple branch indicative of the three-sided structure of the figure. That the "3-star" on the lower right has such a triple structure is again shown by this three branch. At this point, it can be seen heuristically that by use of the directed graph of the MAF, it is possible to specify a generic 3-bulb, for example. This leads to a geometry whose character lies between topology, which is too general for shape categorization, and the "hard" congruent geometries such as Euclidean and projective geometry. Figure 6 shows such a geometry at work in describing an anthropomorphic outline and a gross distortion of it. It should be pointed out that since these figures are not convex, they have MAF points which lie outside themselves. The outside MAFs are generated only by the points of the contours that are not on the convex hull of the contours.

Since the contour and the MAF description are equivalent and invertible there is no information lost or gained in the transformation. It is clear, then, that this categorization could have been done on the contour itself. What then is the virtue of the new description? I believe

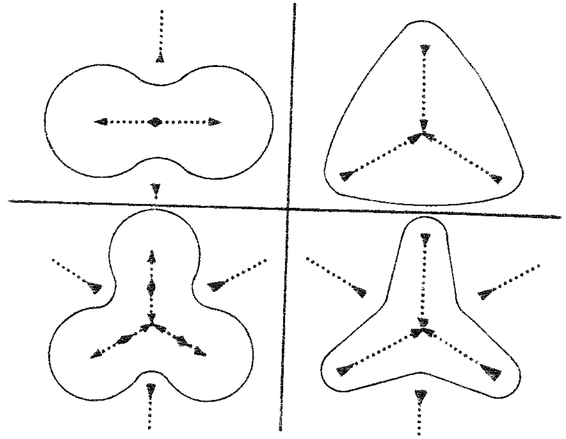

Fig. 5. Representation of simple shape properties by directed graph structure of the MAF.

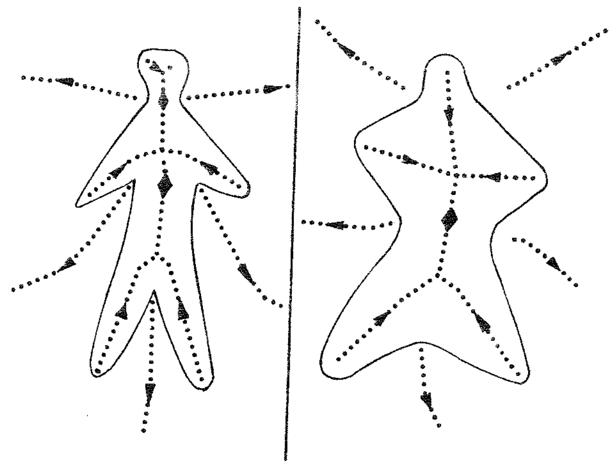

Fig. 6. Two anthropomorphs and their MAFs.

it to be in defining natural shape properties on the new primitive descriptors. It could be argued that the properties I have discussed in the last paragraph could be gotten by such simple contour properties as counting the inflection points and the total curvature of the shape.

Figure 7 shows a case where simple equivalent descriptors cannot be found on the boundary. The upper and lower figures can be arranged to have the same number of inflexion points and the same total curvature; yet they are quite different visual figures. What is different about the figures has to do with the distance across the enclosed space. Carrying along such a distance function to all sorts of other points across the contour as one moves along the contour is a formidable task indeed. To try to deal with such properties from a boundary description such as the intrinsic co-ordinates, for example, implies that one cannot describe a swan's neck until one has gone completely around the swan. This point is illustrated in a different way in Figure 8. The upper figure shows a rectangle and its MAF. The heavy dots indicate points which represent parallel lines and so occur at the same time (infinite velocity). The middle figure shows this rectangle distorted perspectively. It can be seen that the velocity of the central element, a local property, becomes the cue to perspective angle if one assumes one is looking at a rectangle. The lowest figure has been included to emphasize that all is not roses with this theory. For extreme distortions, the MAF structure stays the same but the perspective cue changes. Note that in going from the top to the bottom figure, one passes through a degenerate state when the figure can be circumscribed about a circle. It is an interesting conjecture, that with proper constraints, the MAF structure cannot change without going through such a degenerate state.

An interesting question of continuity is raised in Figure 9. In both parts of this figure the outside shape is continuously transformed to the circle, but by different paths. The intermediate shapes lead to distinctly different MAs. In the upper figure, the MA is continuously deformed from that of the original figure to that of the circle, the central dot. In the lower figure, however, the limiting MA is a radius of the circle. The velocity of the lower MAF, however, becomes asymptotic to the space velocity. Consequently, a viable notion of

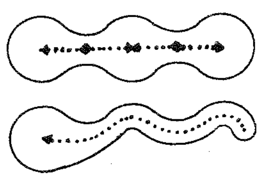

Fig. 7. The width measure of the shape as a primitive property.

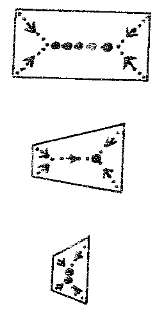

Fig. 8. The MAF for a rectangle "full face" and in perspective.

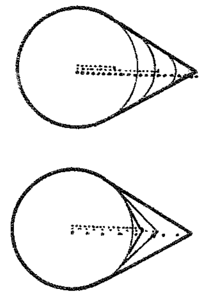

Fig. 9. Dependency of MAF limit on path of curve perturbation.

continuity must include the velocity dimension. This dimension is also important for the closure property introduced in the circle of dots in Figure 2 and in the completing of masked figures discussed later.

Figure 10 illustrates the process considered as a co-ordinate transformation. The upper and middle figures depict symmetrical contours and so have a straight-line MA. The new description consists of the location of the MA point as a function of time (or length of the normal) to the MA. Acceleration represents curvature to the axis; deceleration, away from the axis. An explicit formulation can be given for curvature in terms of velocity and acceleration of the MAF point. The condition that no branching occurs at a point on the MA generated by two contours is that the center of curvature of the points mapping into it do not lie between the MA and those points. (The condition of two contours is inserted since a third contour may enter for curves with their centers of curvature on the side opposite the MA such as is shown in the case of the 3-bulb of Figure 5. The lowest figure shows a curved MAF. In this case the contours are described by two components, the MAF curvature and the acceleration curvature — one curve

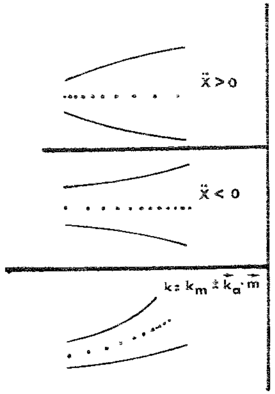

Fig. 10. The MAF as a co-ordinate transformation.

being the sum and the other being the difference when proper components of these curvatures have been taken. Consequently, a curve has an infinite number of descriptions since it can be described from a wide variety of MAs, and the location of the MAs is determined by what else is in the input field. Figure 11 illustrates this. The upper left figure is a semiellipse, its MA being its evolute continued by the perpendicular bisector of its end points. The upper right figure shows the change in form of description gotten by completing the ellipse. The lower left figure shows a descriptor appearing outside the elliptic arc when space is enclosed there, and the altered description obtained by putting a barrier line underneath the elliptic arc. These illustrate clearly the barrier effect of the process. Only the closest contour can affect the MAF. This guarantees that a description of a shape represents the bounding contour alone. Conventional field theories alter the relevant description by modifications due to outside contours both far and near. As such, these and other complete interaction processes, such as two-dimensional autocorrelation, are certain to become unwieldy as the complexity of the input goes up. I feel that this type of segmentation property is essential to a realistic theory of shape.

Figure 12 shows an abstract generalized input. Except for special cases, which complicate the description, but not the properties, the transformation maps each point on the contour into a 2 MAF points, and each MAF point results from 2 contour points. (To make the statement general, it is necessary to consider ray paths rather than contours and MAFs.) Gross properties of the MAF are equivalent to gross properties of the pattern. Each connected space is represented

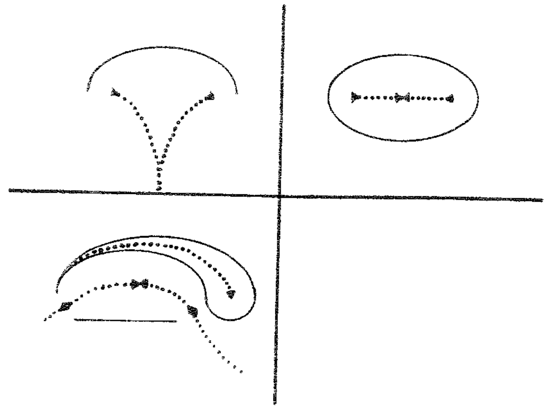

Fig. 11 . Nonsuperposability of partial descriptions and barrier effect.

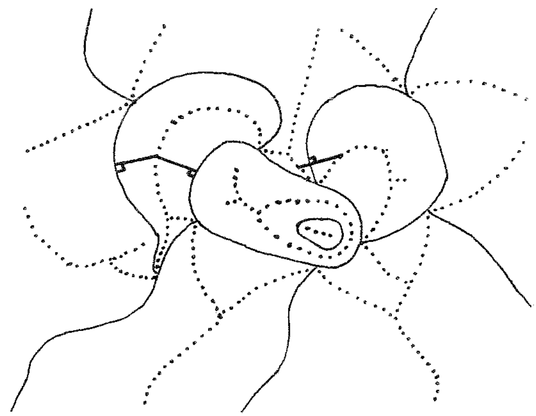

Fig. 12. A generalized input showing MAF as limit of space seen from contour.

by a separate graph piece. A connected space with no holes in it is represented by a "tree," each hole in a space being represented by a loop. (Topological properties of the space are thus available as primitive properties of the graph.) The order and connection of the expanding spaces is represented by the structure of the directed graph. Since the MA represents discontinuities in the wave, the discontinuities of the MAF (appearance, disappearance, impulse, and branch points) can be considered as second-order discontinuities of the process. These can be categorized by a number pair in which the first number is the number of MAF points immediately prior to the discontinuity and the second, the number immediately after the discontinuity. That each of these represents a particular natural figure property is shown by Figure 13. A 0,0 represents a circle; a 0,1 a circular arc of less than 180 degrees; a 1,0 a circular arc of greater than 180 degrees; a 2,0 a double closure such as in an ellipse; a 0,2 a pinch in the contour; a 1,1 a case where the center of curvature lies on the MAF, etc. Simple constraints of geometry, such as the closest distance between three points being between two of them, impose a number of constraints on the possibility of such pairs, which leads to forbidden pairs and pairs possible only at t = 0. Figure 14 illustrates the equivalent matrix of the graphical description. Graph discontinuity properties go into the diagonal, continuous properties into the off diagonal. Each term in the matrix is a vector, in general. Finding good vectors cannot in general be done theoretically but must be decided by new experiments on visual shape. This may be a fruitful area of exploration.

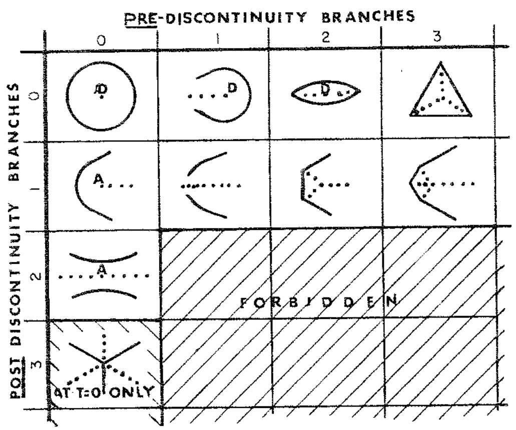

Fig. 13. Shape equivalents of discontinuities on MAF.

Compound Shapes

Thus far we have been considering shapes merely as spaces on a plane. Figure 15 contains an illustrative example for showing an extension of this process for shapes which are masked by others. In the

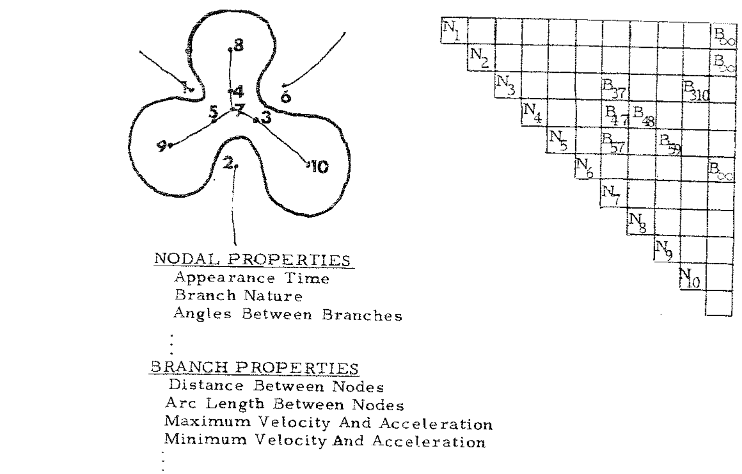

Fig. 14. MAF description allows analysis by matrix of graph.

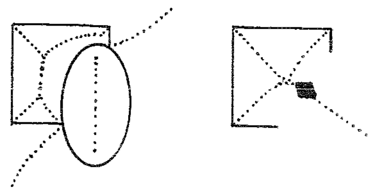

Fig. 15. Extension of MAF to masked shapes.

left-hand figure, an ellipse and partially masked square are shown, the masking considerably complicating the MAF of the square. If the MAF of the ellipse is recognized and its contour removed after its recognition, as shown in the right figure, the representation is considerably simplified. (This can be done from the MAF function by presenting it in negative time so that the ellipse is formed at t = 0. If further propagation is stopped, the input can be made refractory at the locations of the ellipse.) The remaining square contour can easily be recognized. In this way, one can attempt to cancel sequentially the entire stimulus contour and equate this cancellation with an acceptable interpretation of it. Clearly, alternate sequences are possible. A number of properties of such a sequential process are particularly interesting. (a) Masking cues are primary outputs of the process. (b) There is an inherent figure-ground effect that comes from matching MAF descriptions rather than the contour itself. (c) There is an asymmetry of the process with respect to additive and subtractive noise. Additive noise, such as a masking object, must be explicitly removed. Subtractive noise, such as a missing contour, is implicitly filled in. Figure 16 shows a stylized animal behind a tree. If the tree is recognized by any cues, such as the green of its leaves and brown of its trunk, its contour can be removed. This facilitates making a single object of the animal. Such an object completion property may be an extremely important one to vision; a cat must see through grass, a monkey through leaves.

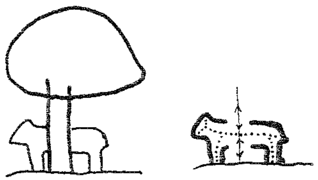

Fig. 16. Making a contiguous world of masked objects.

Psychology and Physiology

Although the transformation has been developed to deal with the visual problems of everyday life, it is interesting to see how it checks with the anomalies of vision, the illusions. In Figure 17, I have shown a number of standard illusions and the MAFs associated with these. In the right-hand figures, I have accented the faster velocity MAFs. If one assumes that the visual process tries to line up MAFs, (these represent the spaces or objects) rather than the contours themselves, the illusions have a natural meaning. The figures on the left are self-explanatory, I feel. The MAF opens a new range of processes for explaining the standard illusions. I feel that it goes much deeper than that, however. The process allows us to take a much broader view of shape vision. We have been looking at vision as a recognition process, rather than as a segmentation process. The closure properties suggest that the appearance of a dotted figure as a continuous contoured one (Figure 2) is an illusion. The process suggests that the whole of our sensitivity to circles, parallels, and symmetric figures may be an illusion. If one asks what life and death problem such sensitivities have emerged to solve, one finds no ready answer. It is suggested here that these are by-products of an inverting process of the visual co-ordinates useful for coping with the problems of localization, segmentation, and organization of a world of visual objects.

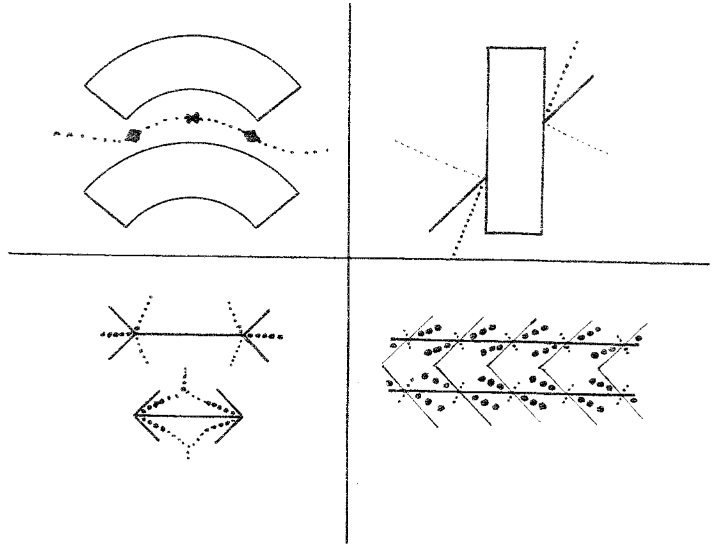

Fig. 17. The MAF as a source of new explanations for optical illusions.

A word must be said about how such a process relates to the results of Hubel and Wiesel (1962, 1965) on cortical cell responses. This has been abstracted and stylized in Figure 18. The upper-left figure shows the receptive fields of the "simple cells" that would be fired by a particular contour. The upper right shows the receptive fields of their "complex cells." Whereas Hubel and Wiesel have looked at these responses as detectors of lines at particular angles, I am suggesting that they can be looked at as generators of parallel shapes. The lower left shows the kinds of corners that could be formed from the interaction of the output of such cells. The lower right shows a selection of proper "symmetrical corners" obtained by having a cell that responds to a corner at a particular location, but <u>not</u> in an adjacent location. Such a sequence can perform the transformation without the barrier property. This requires a further condition. Unfortunately no one has looked for such a property, as far as I know.

General Comments and Conclusion

A word should also be said about the mathematical implication of the transformation. (It should be clear that I speak for mathematicians

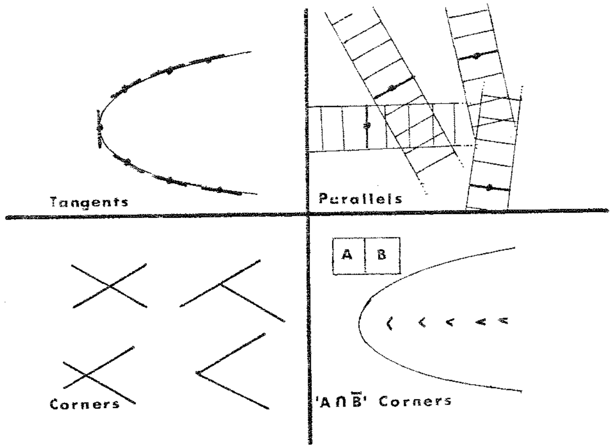

Fig. 18. MAF extraction by idealized properties of visual cortical cell receptive fields.

here without their permission.) The general Cartesian view of geometry metricizes a space (builds a co-ordinate system) and then describes the curve in that metric in some functional form. Unfortunately, the transformation from geometry to analysis is highly discontinuous and gives preference to shapes which have easy functional rather than geometric meaning. The approach introduced here suggests that the co-ordinate system be built on the curve and space be explored from the curve. The MA is then the boundary of the space (in some Riemann sense) that must exist if one is to have a good single valued co-ordinate system. The process indicates that the curve is then describable from the boundary as the envelope of circles on that boundary. I have talked hitherto about a Euclidean orthogonal process, but there is no need to impose such a limitation. This transformation can be done with non-Euclidean processes, as the spherical visual mapping may well demand, or it can be done in higher dimensions, as a three- or four-dimensional world may demand.

Figure 19 shows my first physical embodiment of the process. It uses a movie projector and camera with high contrast film. These are symmetrically driven apart from the lens in such a way as to keep a one to one magnification, but to increase the circle of confusion (defocussing). The lower right inset shows how this dilates the contour. Corner detection is done separately by a subsequent process. I am presently building a closed loop electronic system to do both the wave

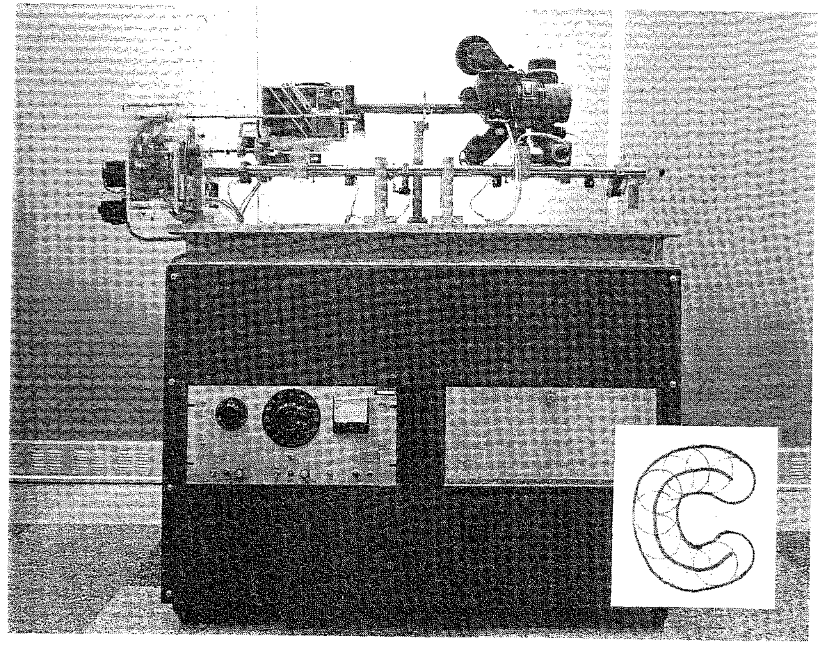

Fig. 19. An early opticomechanical device for obtaining grassfire propagation using photographic defocussing.

generation and corner detection.

In conclusion, I would like to emphasize that our view of vision has been constrained for too long by problems of physical geometry. Successful methods there have looked at shape as a superposable process where the total collection of elements is equal to the sum of the parts. I have tried to depart from this by looking for processes in which the interaction is precisely the item of interest, without getting into a conceptual space so rich in possibilities that all things are encompassable in it.

References

Attneave, F., and Arnoult, M. D. The quantitative study of shape and pattern perception. Psychol. Bull., 1956, 53, 452-471.

Bitterman, M. E., Krauskopf, J., and Hochberg, J. Threshold for visual form: A diffusion model. Amer. J. Psychol., 1954, 67, 205-219.

Blum, H. An associative machine for dealing with the visual field and some of its biological implications. In E. E. Bernard and M. R. Kare (Eds.), Biological prototypes and synthetic systems. Vol. 1. New York: Plenum, 1962.

Deutsch, J. A. A theory of shape recognition. *Brit. J. Psychol.*, 1955, *46*, 30-37.

Hubel, D. H., and Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 1962, *160*, 106-164.

Hubel, D. H., and Wiesel, T. N. Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.*, 1965, *28*, 229-289.

Kazmierczak, H. The potential field as an aid to character recognition. In *Proceedings international conference on information processing*. London: Butterworth, 1960.

Kirsch, R. A., Cahn, L., Ray, C., and Urban, G. H. Experiments in processing pictorial information with a digital computer. *East. joint Comp. Conf.*, 1958, *12*, 221-229.

Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt Brace, 1935.

Kotelly, J. C. A mathematical model of Blum's theory of pattern recognition. *AFCRL-TR-63-164*, USAF Camb. Res. Labs., 1963.

Sutherland, N. S. *The methods and findings of experiments on the visual discrimination of shape by animals*. Exp. Psychol. Monogr. No. 1. Cambridge, Eng.: Heffer, 1960.