The University of Wisconsin
Computer Sciences Department
1210 West Dayton Street
Madison, Wisconsin 53706

Received: November 1974

EDGE DETECTION IN COMPLEX SCENES BASED ON GESTALT PRINCIPLES OF PROXIMITY AND SIMILARITY

by

James M. Lester

Computer Sciences Technical Report #229

November 1974

Abstract

An edge detection function is defined on a very broad class of pictures. It departs from previous approaches to edge detection, which are based on differences in intensity or other properties between adjacent areas of a picture, in using the spatial and similarity relations among individual entities, called elements, to determine the locations and probabilities of edges. With auxiliary definitions of similarity, it may be applied to digitized pictures, treating picture points as elements, or to scenes of diversified shapes distributed in any way — although only scenes of disks are discussed in this paper. This generality permits recursive use of the edge detector in the discrimination of visual textures. An algorithm for the efficient computation of the function is described. The program runs at the University of Wisconsin in a larger system, which includes T.V. input and additional programs that cluster edges into boundaries.

Acknowledgments

I am grateful to Leonard Uhr for his valuable advice. This research was supported in part by the National Science Foundation under grant GJ-36312.

I. Introduction

A. Gestalt laws of visual organization

In a classic paper [1] published in 1923 Max Wertheimer described a number of factors or principles of visual perception, some of which appeared to explain the "spontaneous" grouping of distinct objects in a scene. Fig. 1 offers three examples of the type of scenes Wertheimer analyzed: simple shapes (hereafter referred to as elements) against a uniform background.

The first of these illustrates the principle of proximity; the relative nearness of the disks within each of the five columns causes them to be seen as a single coherent group. In Fig. 1(b), the elements are equally spaced in a rectangular fashion, but we clearly see rows, each formed by either gray or black shapes. Here, it is the similarity among members of a group that causes them to cohere. These two organizing forces operate independently in the same scene as shown in Fig. 1(c), where proximity creates columns and similarity creates rows.

B. A paradigm for scene analysis

We regard the Gestalt principles of organization as a loosely formulated paradigm for scene analysis. Within this conception, scenes — at some stage in their processing — consist of a collection of elements related in two spatial dimensions and along an additional dimension of similarity. A major goal in their further processing is the identification of those sub-collections which appear to form natural groups; such collections are characterized by greater proximity and/or similarity among their members relative to the scene as a whole.

The main body of this paper attempts to give substance to and extend this paradigm. We consider the shape of a group — as distinguished from the set of group members — to be primary; we therefore look for group boundaries directly with an edge detection approach. In the exposition of the edge function, restricted cases are first examined, followed by more general cases. An algorithm for the efficient computation of this function is then described. In the final section, we discuss a complete generalization of the function and the role of the edge detection process in a complete perceptual system.

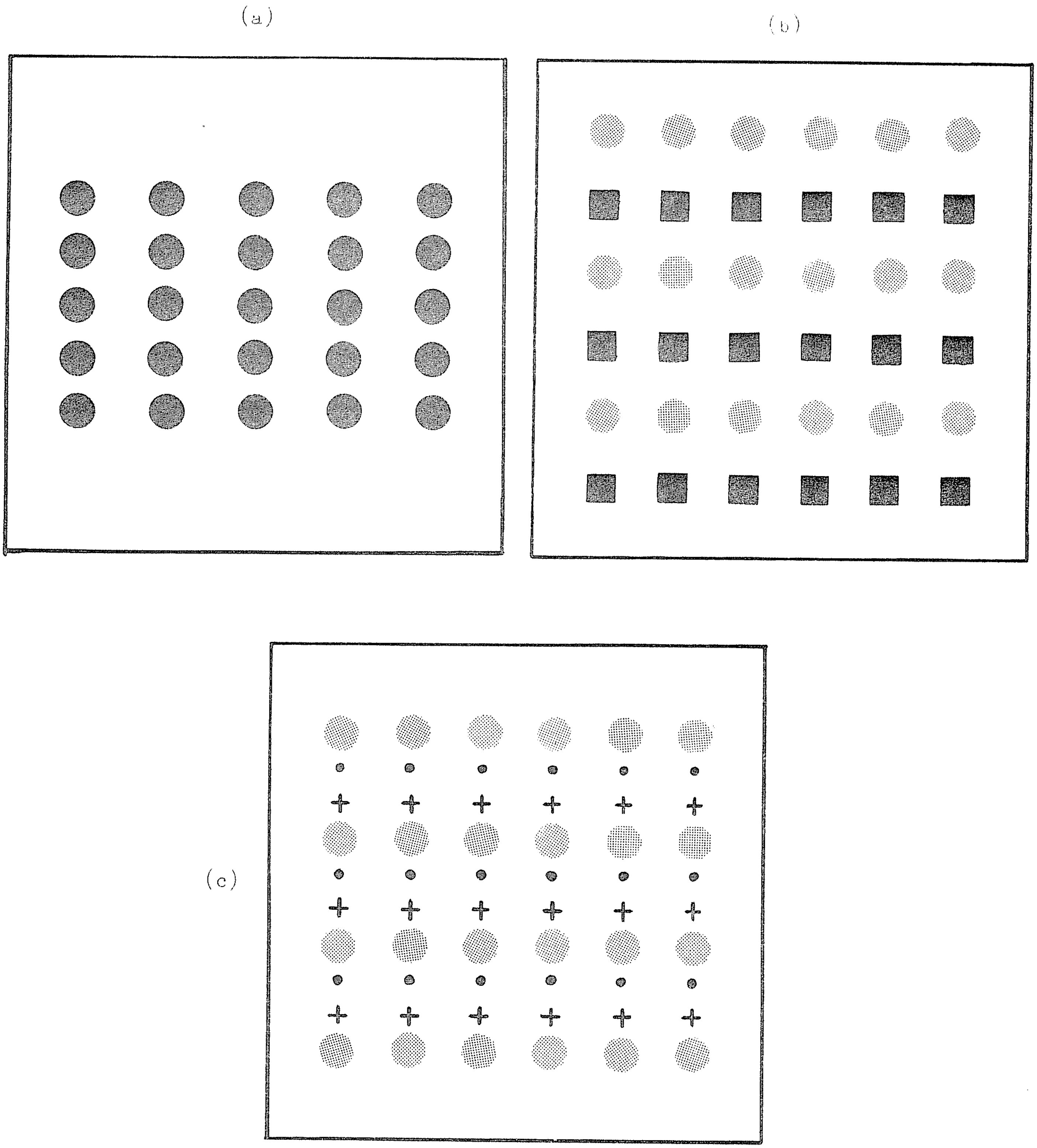


Fig. 1 (a)Column organization due to <u>proximity</u>. (b)Row organization due to <u>similarity</u>. (c)Both columns and rows can be seen, attributable to proximity and similarity respectively.

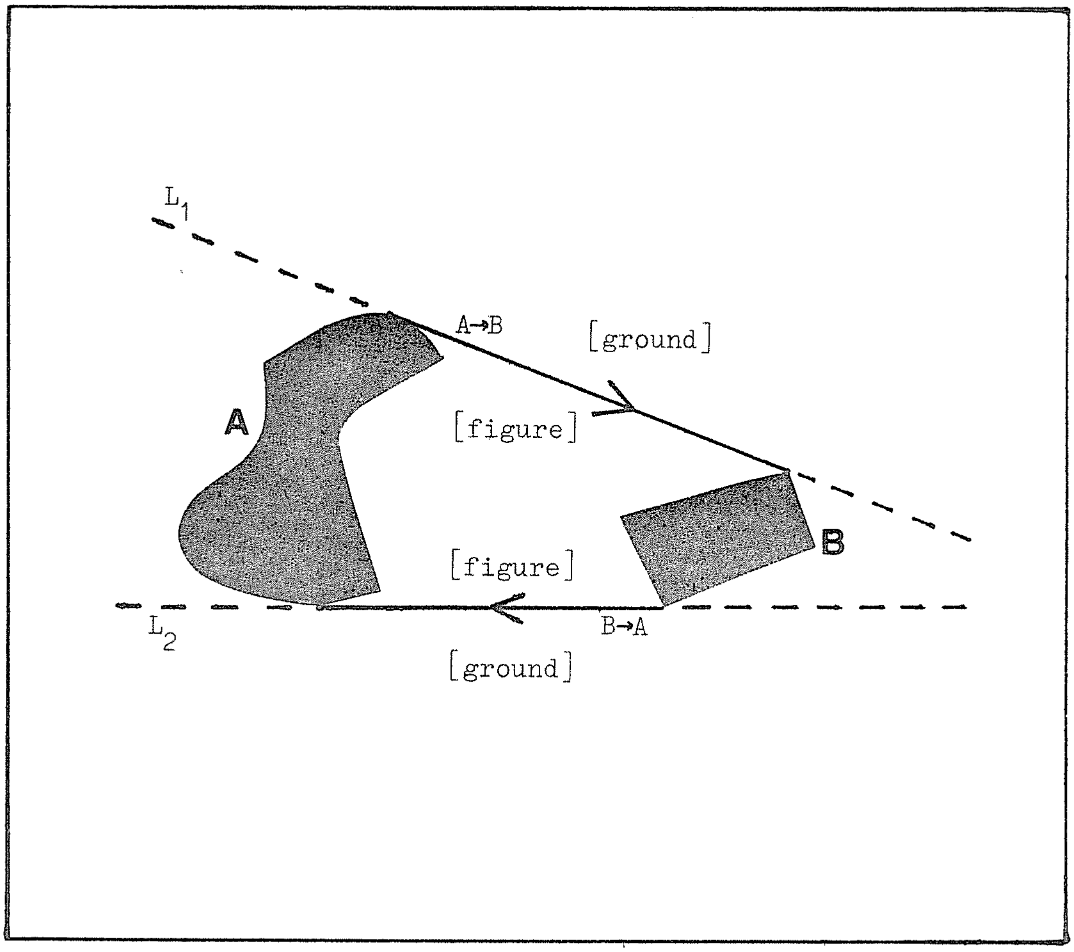


Fig. 2 The two edges spanning a pair of elements, showing "figure" and "ground" sides of each edge.

II. The Edge Detector

A. Definition and interpretation of edges

An edge is a directed line segment spanning two elements. Every pair of elements, A and B, may be spanned by exactly two edges, A→B and B→A, the shapes, sizes and relative positions of the elements determining precisely the lengths, locations and orientations of the edges, as shown in Fig. 2. With reference to this figure, the edge A→B is a segment of L_1, a line tangent to both A and B; the edge begins at the point of intersection of L_1 with A and ends at its intersection with B. Similar remarks apply to B→A and L_2.

An edge thus defines two half planes, one containing all of, the other none of, the elements spanned by the edge. The first of these half planes is on the right facing in the direction of the edge: any element in this region is said to be on the *figure* side of the edge. The *background* side of the edge is to the left facing along the edge.

The definition of edge applies to two elements of any shapes and sizes, provided that neither is entirely enclosed by the other's convex hull. This restriction need not concern us here since the elements to be considered from now on will be convex.

Edges — as discussed in this paper — are more than "dividers" between two regions: each edge partially outlines a shape or figure, setting it off from the background -- the background itself is not directly outlined. A single edge, then, belongs to exactly one figure, not two, and a chain of such edges (in the style of Freeman [2]) can be directly interpreted as the boundary of a region. Further advantageous consequences of this treatment of edges are discussed below in section II.E.

The goal of the edge detection process is, very simply, to estimate a probability for every edge in a scene based on the relations, spatial and otherwise, between elements in the vicinity of the edge. For the remainder of this paper we severely restrict the complexity of these relations in order to focus on the essential factors and problems for an edge probability function.

B. Proximity: scenes of identical elements

In this section we motivate and define PROBd, an edge probability function for scenes of identical elements circular in shape. Some restrictions on the allowed dissimilarity between elements are removed in section II.D; the removal of those affecting shape and size is discussed in III.A.

The probability of an edge is based largely on the distribution of elements on the ground side -- or "outside" -- of the edge. Loosely speaking, if some element is located just outside the edge the probability will be low; conversely, if the area outside the edge is free of nearby elements, the probability may be high. Fig. 3(a) and (b) illustrate the greater negative influence

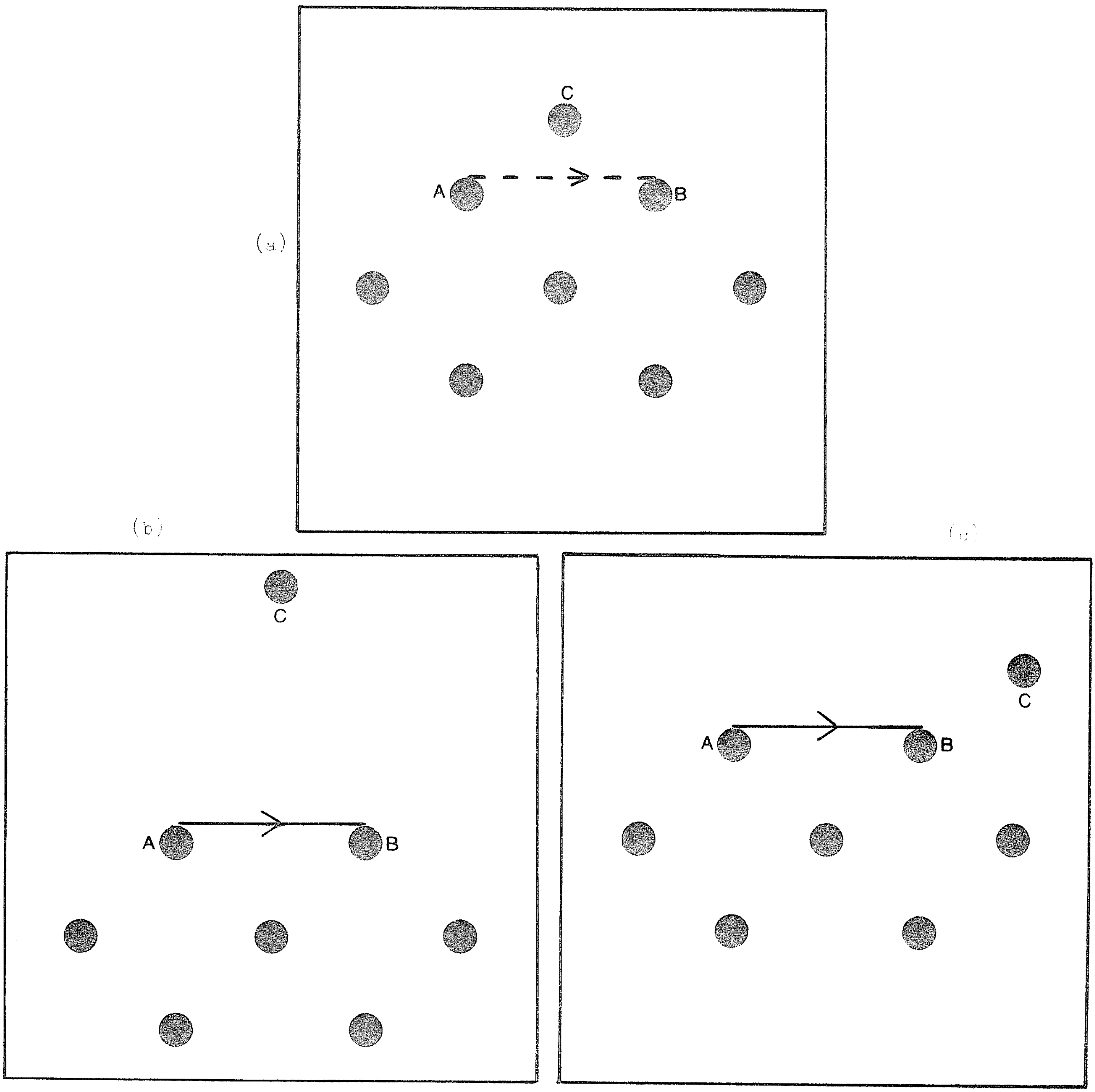
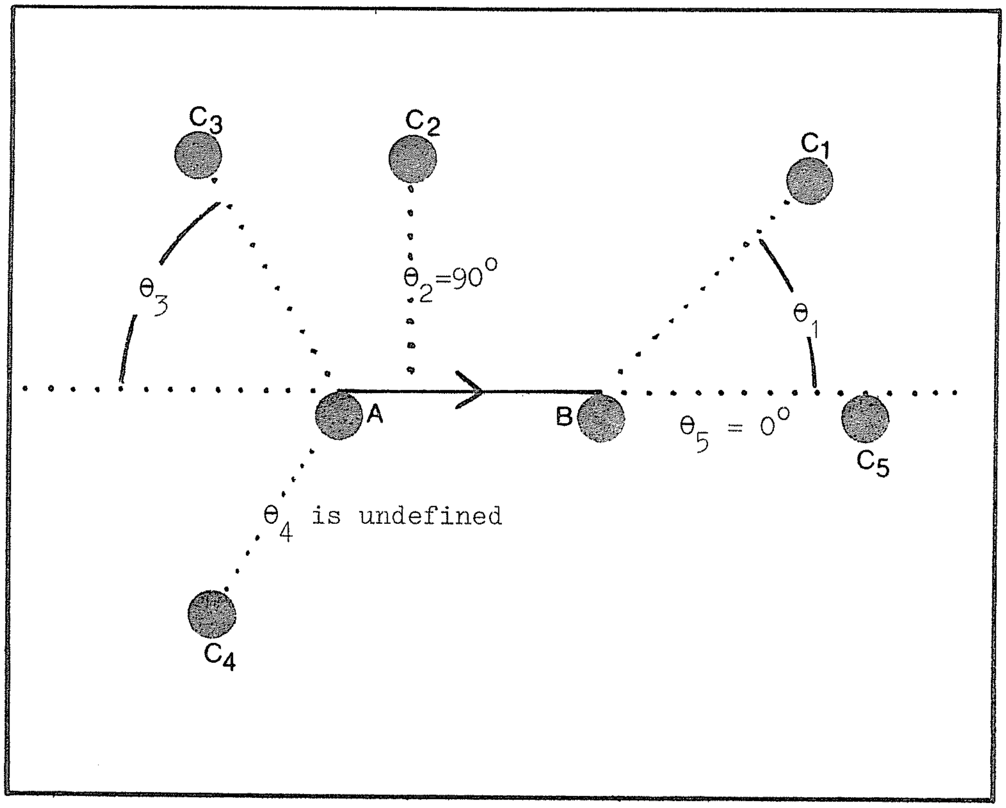


Fig. 3 (a)The proximity of C to the edge and its position directly over it make A→B unlikely. (In this and all subsequent illustrations solid arrows represent relatively strong or probable edges, and broken arrows represent improbable edges.) (b)As C recedes, the probability of A→B increases. (c)As C moves to one side, the probability of A→B increases.

(a)



(b)

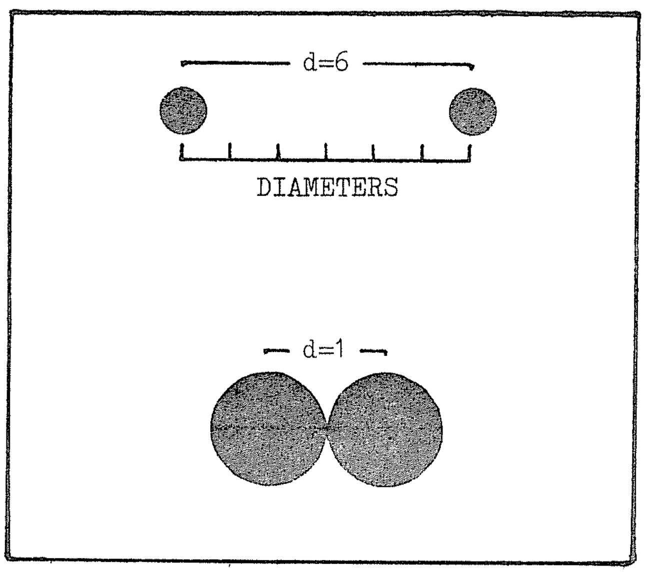


Fig. 4 (a)The angles made by C_i with A→B. Perpendicular to the edge, $\theta=90^\circ$; colinear with it, $\theta=0^\circ$; on the figure side of the edge, θ is undefined. (b)The distance between disks of the same size, according to (1).

of nearby over distant elements on this probability. The effects of elements directly over the edge versus those of elements off to one side are contrasted in Fig. 3(a) and (c).

These two factors, distance and angle, are quantified and related precisely in the definition of STRd, below. In this definition, $\theta_{C,A\rightarrow B}$ is the angle made by the element C with A→B in the sense of Fig. 4(a); d, as shown in Fig. 4(b), is the following scale independent measure of distance:

(1) d(A,B) = the (Euclidean) distance between the centers-of mass of A and B in diameter units

The STRength of the edge A B is based on calculations made independently at both A and B of the EFFect each element C has on the edge. At A we look for that C whose combined proximity to A and angle with A→B make it most damaging to the probability of the edge:

(2) $EFFd_A(C,A\rightarrow B) = 1.0 - d(A,B)*\sin(\theta_{C,A\rightarrow B})/d(A,C)$
$= 0.0$ if above expression is negative

(3) $STRd_A(A\rightarrow B) = \min(EFFd_A(C,A\rightarrow B))$ for all elements C such that θ is defined
$= 1.0$ if there are no C for which θ is defined

Similarly, at C we define:

(4) $EFFd_B(C,A\rightarrow B) = 1.0 - d(A,B)*\sin(\theta_{C,A\rightarrow B})/d(B,C)$
$= 0.0$ if above expression is negative

(5) $STRd_B(A\rightarrow B) = \min(EFFd_B(C,A\rightarrow B))$ for all elements C such that θ is defined
$= 1.0$ if there are no C for which θ is defined

Fig. 5 shows some STRd values for a variety of configurations.

If the probability of A→B were taken to be the average of $STRd_A(A \to B)$ and $STRd_B(A \to B)$ we would have an edge function that worked well in many cases. But Fig. 6 exemplifies the need to consider not only the relative distance distribution of elements above the edge, but the absolute distance between A and B as well. Clearly, the closer are A and B, the more likely is A→B. The edge probability function for the restricted scenes of this section incorporates this idea:

$$(6) \quad PROBd(A \to B) = [(STRd_A(A \to B) + STRd_B(A \to B))/2.0]/d(A,B)$$

C. <u>Experiments with and properties of PROBd</u>

Figs. 7-10 illustrate several desirable features of the edge function PROBd. The first three of these are self-explanatory; the last shows the best edges for a relatively complex scene according to EDGE, a FORTRAN program that computes (6) over the entire scene (described below in II.F).

Since what is finally seen in a picture is not affected by enlargement or reduction of the picture as a whole (except, of course, if new details emerge or disappear in such a change) it is essential that PROBd be invariant with respect to change of scale. This is guaranteed by the definition of d, (1).

We might also investigate the consequences of uniform shrinkage or expansion of all elements. In the simple case of identical disks, altering the size by a factor of q (q cannot be so large as

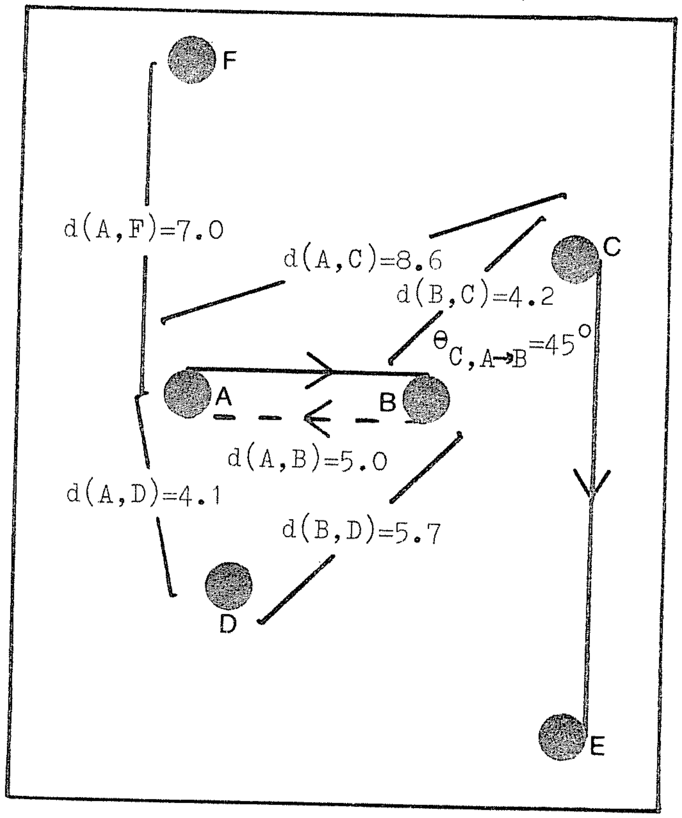


Fig. 5 Values of STRd for the three edges drawn above:

$STRd_A(A \to B) = EFFd_A(F, A \to B) = 1.0 - 5.0*\sin(90^o)/7.0 = .29$

$STRd_B(A \to B) = EFFd_B(C, A \to B) = 1.0 - 5.0*\sin(45^o)/4.2 = .16$

$STRd_B(B \to A) = EFFd_B(D, B \to A) = 1.0 - 5.0*\sin(90^o)/5.7 = .12$

$STRd_A(B \to A) = EFFd_A(D, B \to A) = 0.0$

$STRd_C(C \to E) = 1.0$

$STRd_E(C \to E) = 1.0$

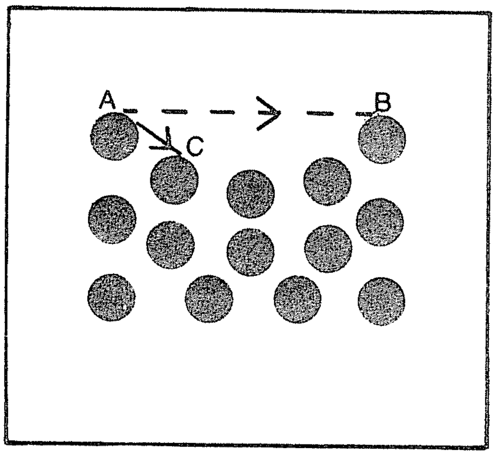


Fig. 6 STR values for edges that are part of the convex hull boundary of a shape, like A→B, are greater than those for edges in a concavity, like A→C. Using the absolute length of an edge in the calculation of edge probability, (6), allows concavities to emerge properly.

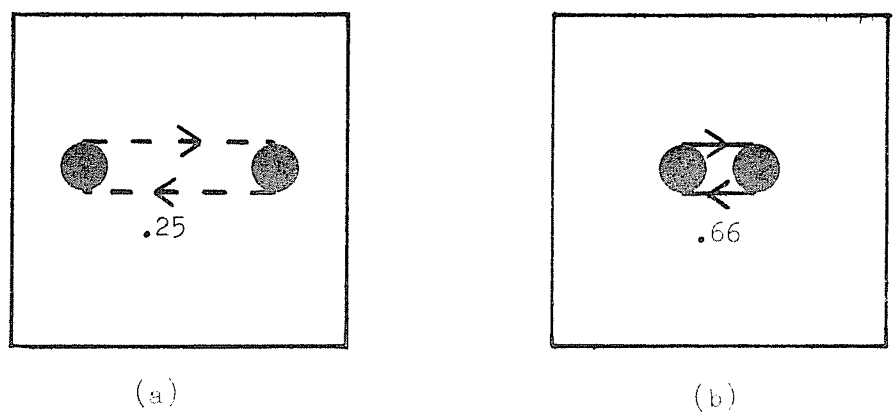


Fig. 7 Edges become more probable as elements approach one another.

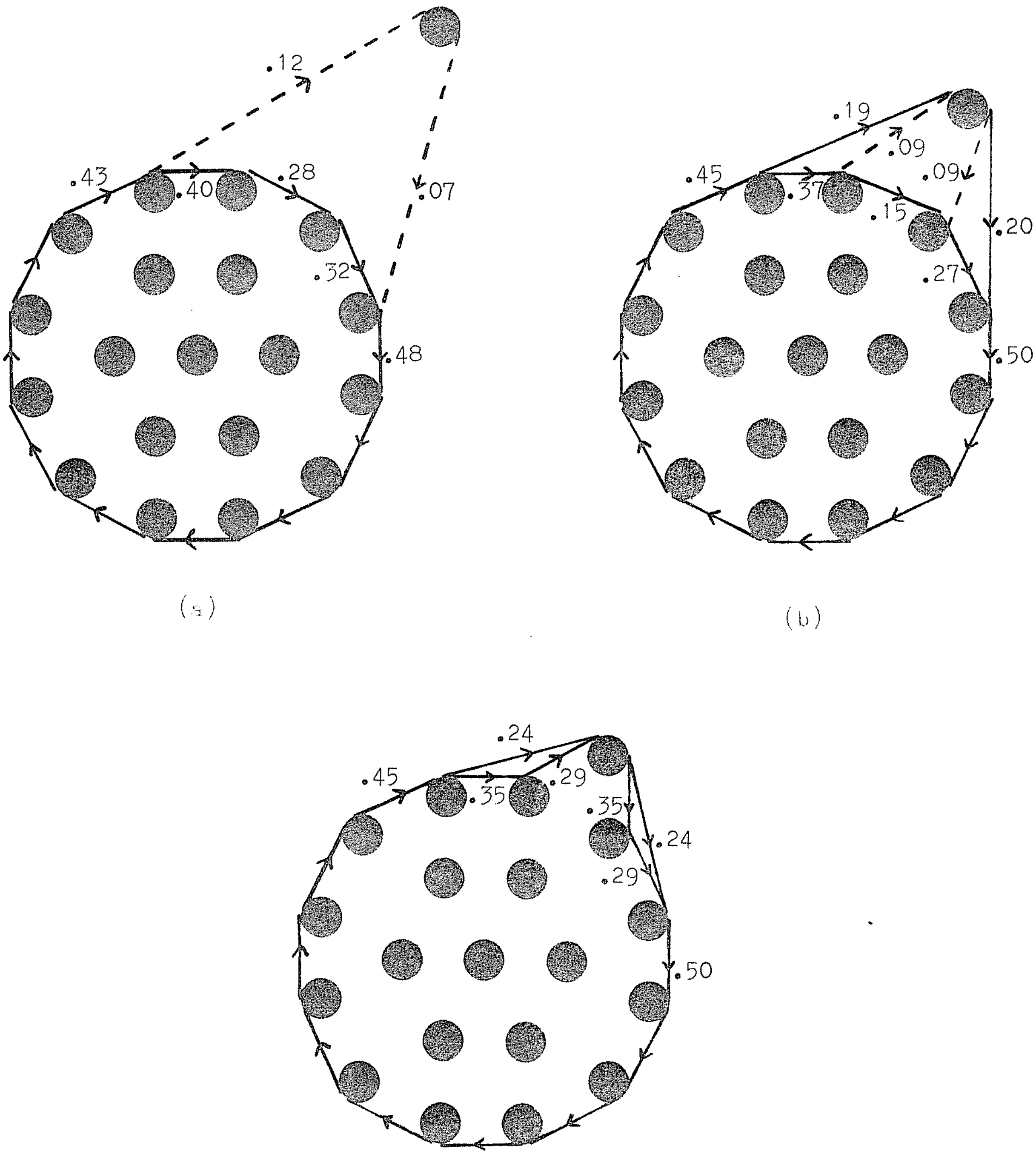


Fig. 8 "Planet and satellite" of (a) becomes "circle with protuberance" in (c), as edges on the planet boundary nearest the satellite weaken, and edges between the two become stronger.

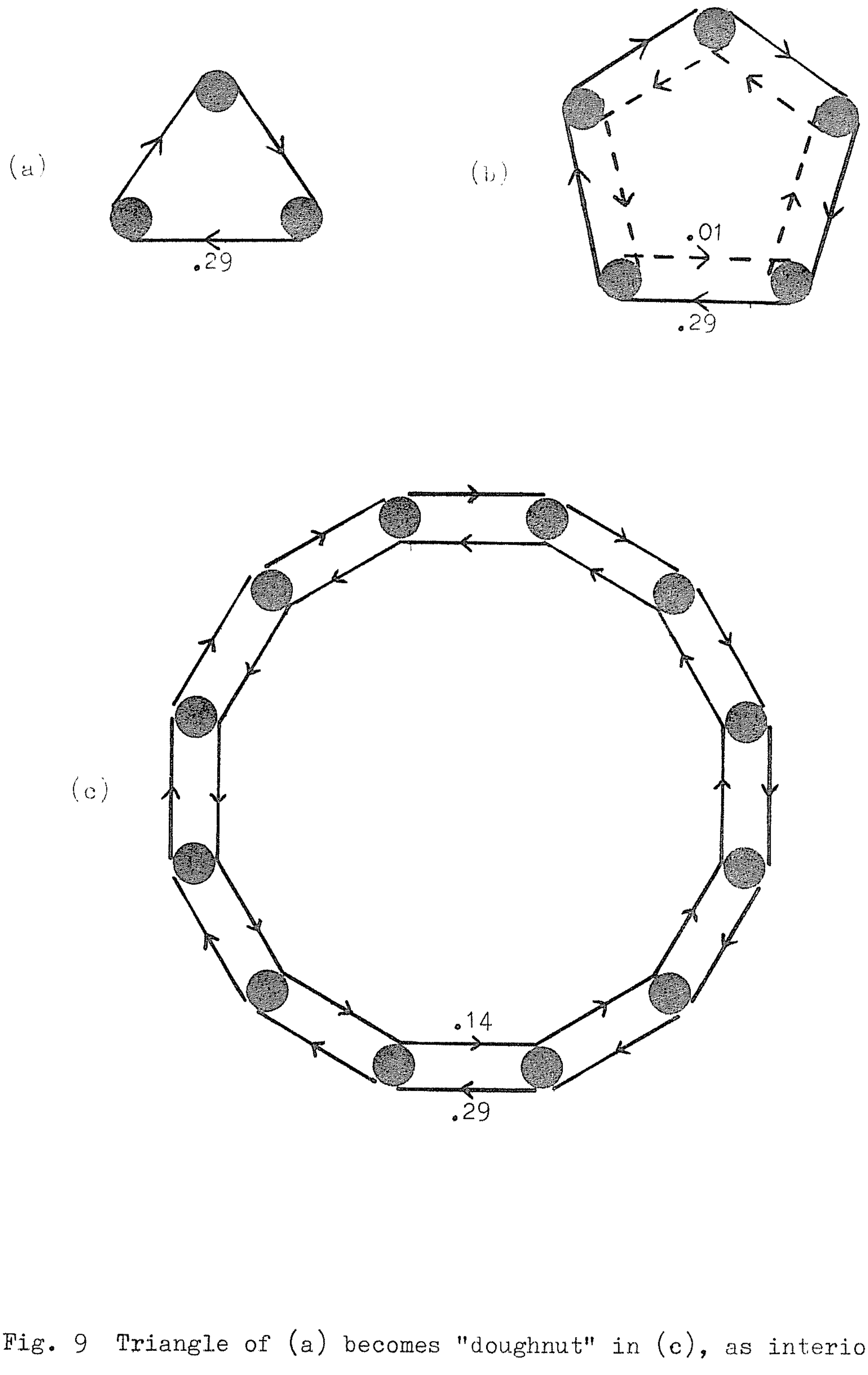


Fig. 9 Triangle of (a) becomes "doughnut" in (c), as interior edges appear.

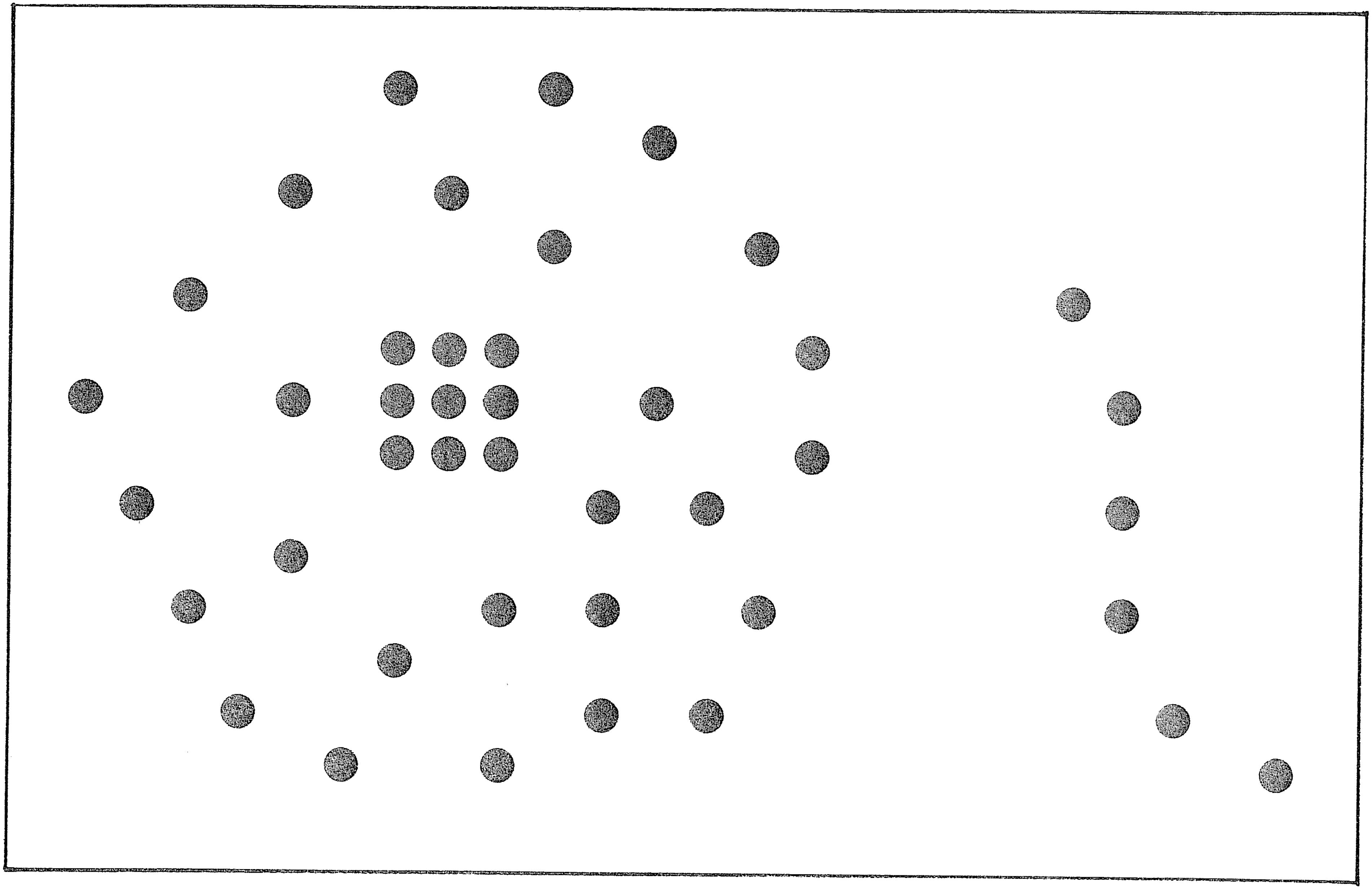


Fig. 10 (a)

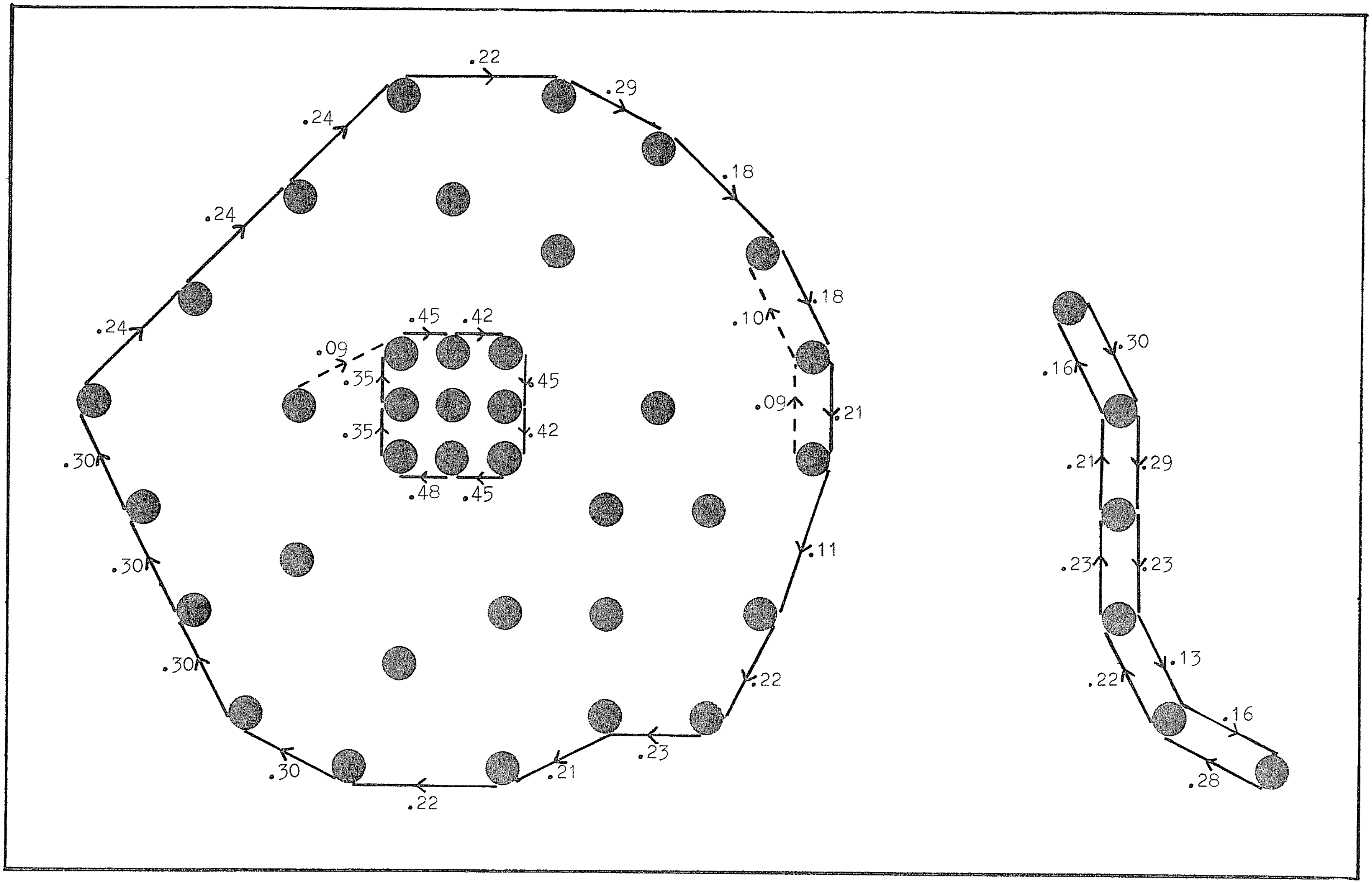


Fig. 10 (b)The 38 most probable edges for the scene of (a) as determined by EDGE. These include all edges along the three most obvious boundaries as well as several "interior" edges which are weaker.

to cause "overlapping" of nearby elements) causes inter-element distances to change by a factor of $1/q$. Angles and relative distances are unchanged, however, so that STRd is not affected; the only difference in PROBd, therefore, is in multiplication by $1/q$. The simplest interpretation of this change is that the same edges, boundaries and shapes will be formed regardless of the choice of diameter — but the very perception of sub-collections of elements *as shapes* will be diminished or enhanced. The reader may convince himself of the reality of this phenomenon by studying Fig. 11. We point out again that we are only considering scenes of identical disks: as we will see, the results of shrinking or expanding dissimilar elements are a good deal more striking.

D. *Similarity: scenes of dissimilar elements*

We now move on to a more general class of elements: they are still circular and the same size but may differ in any other ways, e.g., brightness, color or texture. The retension of restrictions on shape is not primarily for the purpose of making easier the estimation of similarity between elements (although it does), but rather to preserve the simple notions of distance and angle of II.B while developing the proper form of the interaction of distance with similarity.

The most obvious treatment of similarity is simply as a "third dimension" alongside the two spatial dimensions. We would begin by defining a composite distance function:

(7) $ds(A,B) = SQRT(d(A,B)^2+SIM(A,B)^2)$ where SIM is a numerical measure of the similarity of A and B

Substituting ds for d in (2), (4) and (6), however, yields a new probability function with the undesirable property of permitting a tradeoff between proximity and similarity. In Fig. 12, for example, no matter how close together they are positioned, the dissimilar rectangular fields maintain their integrity; furthermore, if two elements from the different groups are closer to one another than to their nearest neighbor in their respective groups, they will form a group. In the first instance, similarity edges are formed; in the second, proximity edges. In Fig. 1(c) there are similarity edges running horizontally and proximity edges running vertically.

These examples could persuade us to consider the opposite extreme, that is, a completely separate function, "PROBs," based entirely on SIM and to be combined (perhaps averaged) with PROBd in a total probability function. Fig. 13, however, plainly shows this approach to be wrong. The correct interaction of proximity and similarity lies between these extremes.

First, we note the effect of individual elements on an edge. We have already seen that, if elements are identical, a proximity edge is formed to the extent that the area immediately outside the edge is free of other elements. With dissimilar elements we have the opposite situation: to the extent that elements outside a po-

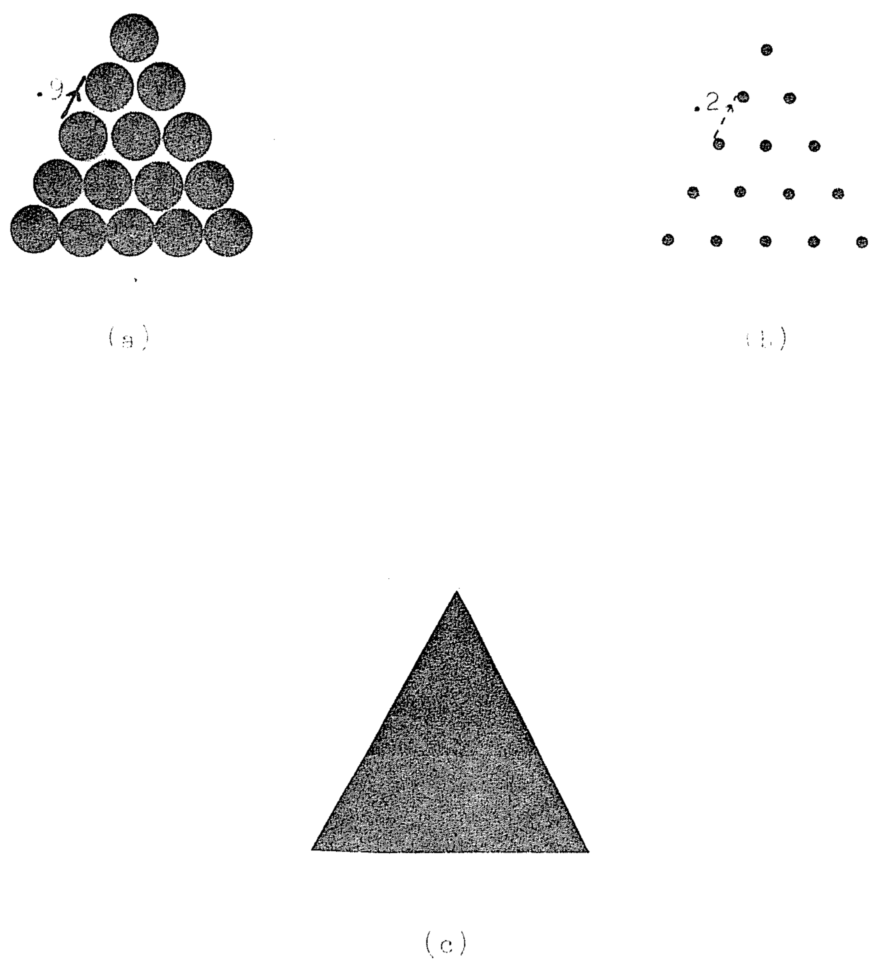


Fig. 11 The effect of uniform change in element size. Both (a) and (b) may be perceived as the triangle of (c), but this interpretation is more compelling for (a). The difference is reflected in the values of PROBd, which are higher for (a).

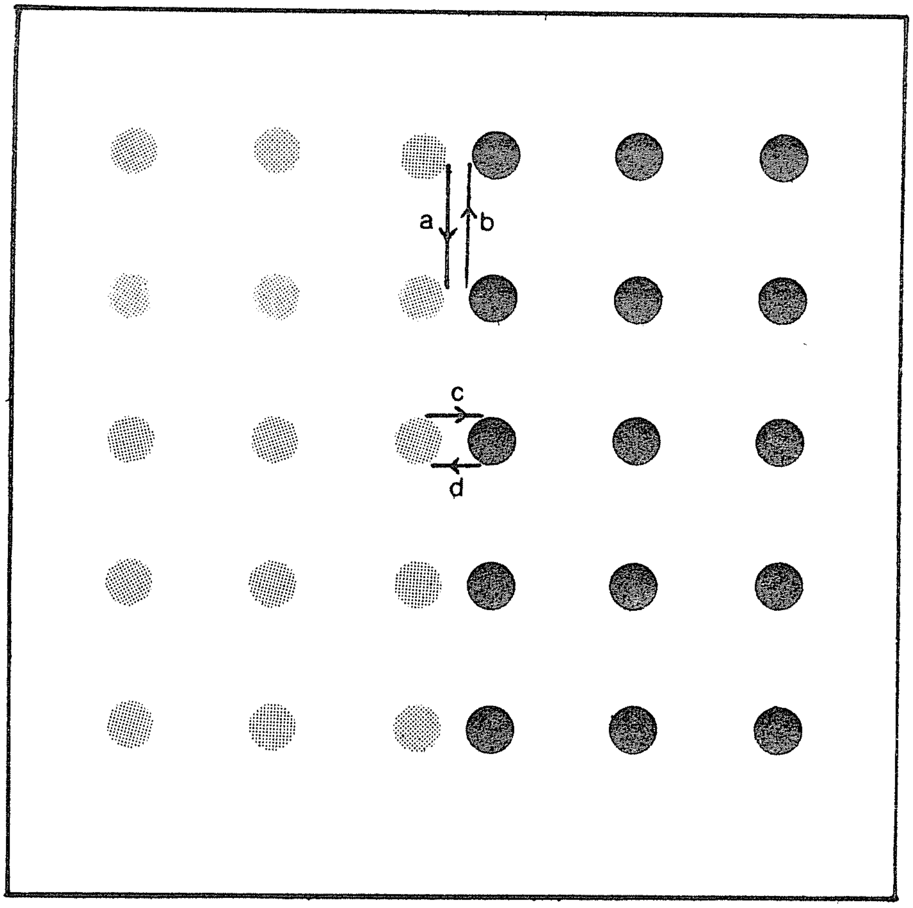


Fig. 12 "Similarity edges" such as (a) and (b) separate the two rectangular fields; these are constant regardless of how close together the two fields are. (c) and (d) are "proximity edges" formed when the fields are close and bounding pairs of dissimilar elements.

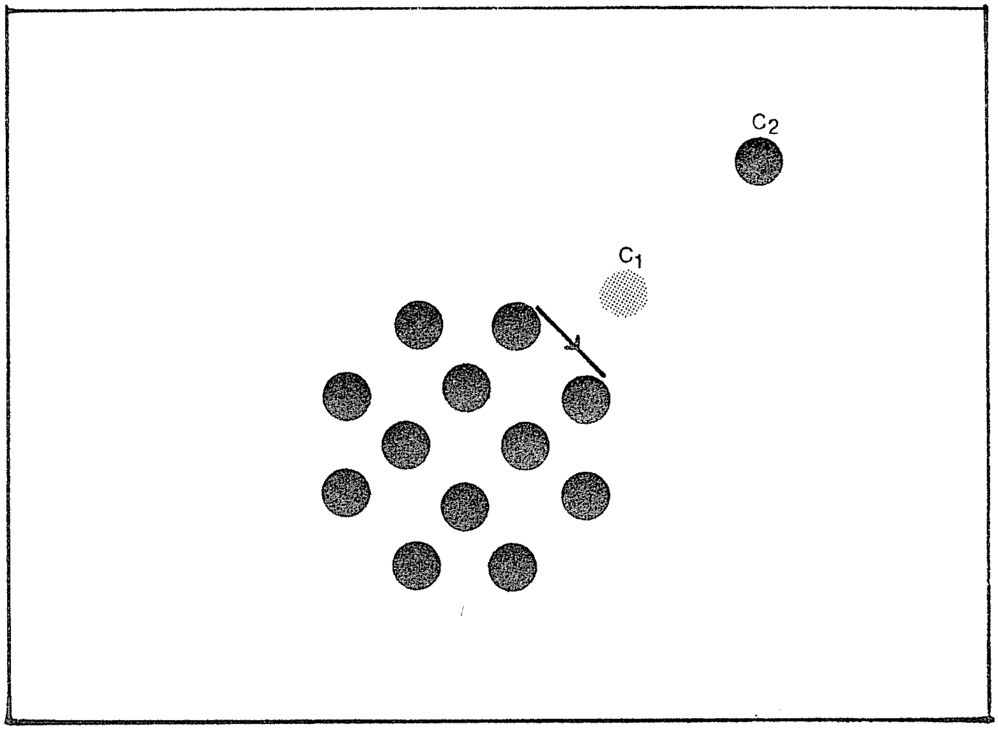


Fig. 13 If the probability of the edge drawn here were based on two separate components, PROBd and PROBs (a hypothetical function modelled on PROBd), it would be zero, with C_1 the "most damaging element" in the calculation of PROBd, and C_2 playing the same role for PROBs. Since the edge is actually strong, a <u>single</u> element must determine both similarity and proximity contributions.

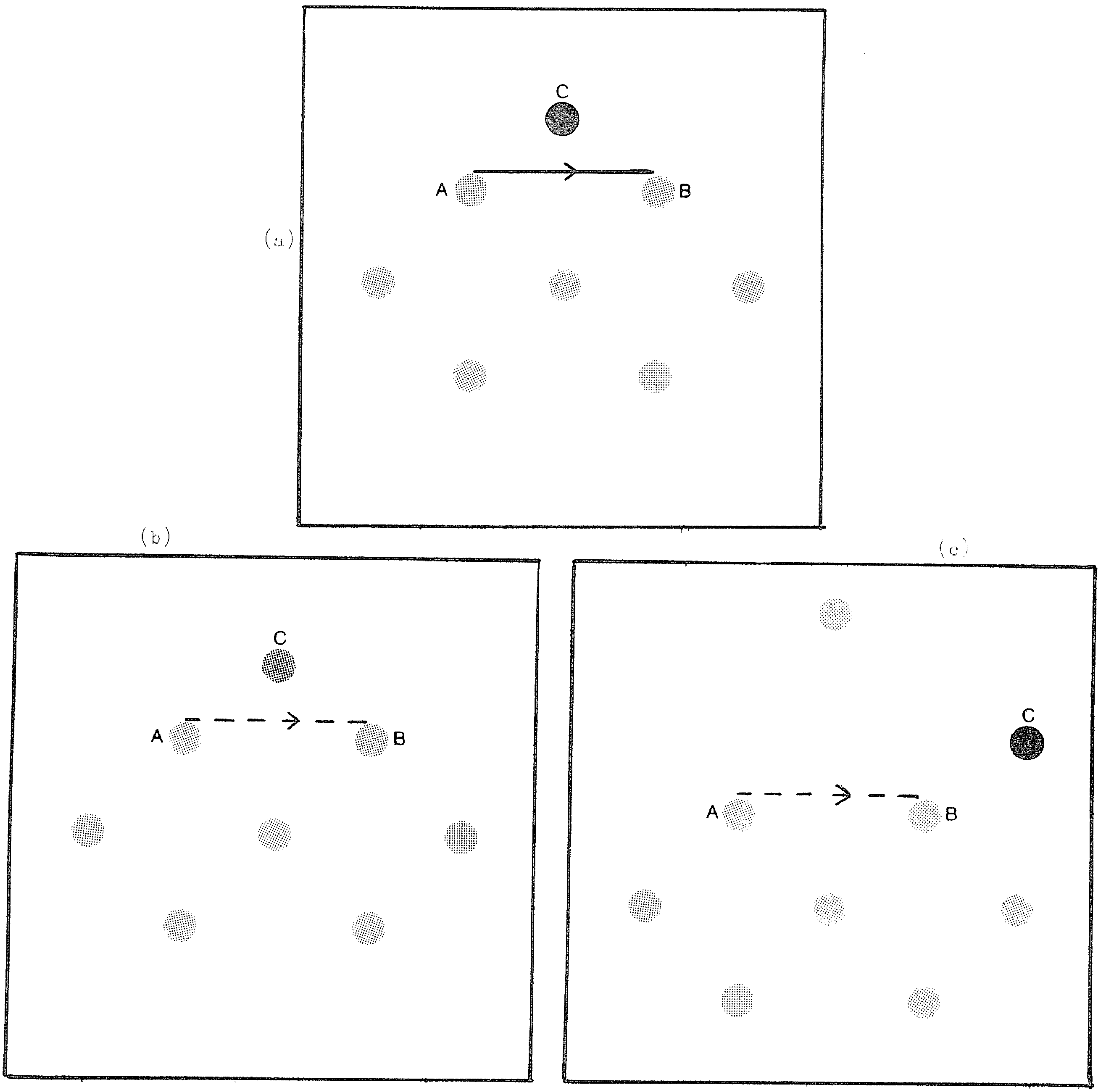


Fig. 14 (a)The dissimilarity of C to A and B and its position directly over the edge make A→B very probable. (b)As contrast diminishes, the probability of A→B decreases. (c)As C moves to one side, the probability of A→B decreases (actually, what decreases is the similarity contribution to the edge — an additional element is introduced in order to balance the increased proximity contribution brought about by the move).

tential edge are dissimilar to the elements spanned there is a similarity edge; as these dissimilar elements make increasingly large angles with the edge, their contribution to a strong edge decreases. In analogy with Fig. 3 and (2) and (3) we have Fig. 14 and:

$$(8)\quad \mathrm{EFFs}_A(C,A\rightarrow B) = [1.0-(\mathrm{SIM}(A,B)+1.0)]*\sin(\theta_{C,A\rightarrow B})/(\mathrm{SIM}(A,C)+1.0)$$
$$= 0.0 \quad \text{if above expression is negative}$$

$$(9)\quad \mathrm{STR}_A(A\rightarrow B) = \min(\mathrm{EFFd}_A(C,A\rightarrow B)+\mathrm{EFFs}_A(C,A\rightarrow B))/2.0 \quad \text{for all } C \text{ such that } \theta \text{ is defined}$$
$$= 0.5 \quad \text{if there are no } C \text{ for which } \theta \text{ is defined}$$

The strength of A B at A is still based on the single most damaging element outside the edge -- but the damage is computed independently for proximity and similarity. We go on to define at B:

$$(10)\quad \mathrm{EFFs}_B(C,A\rightarrow B) = [1.0-(\mathrm{SIM}(A,B)+1.0)]*\sin(\theta_{C,A\ B})/(\mathrm{SIM}(B,C)+1.0)$$
$$= 0.0 \quad \text{if above expression is negative}$$

$$(11)\quad \mathrm{STR}_B(A\rightarrow B) = \min(\mathrm{EFFd}_B(C,A\rightarrow B)+\mathrm{EFFs}_B(C,A\rightarrow B))/2.0 \quad \text{for all } C \text{ such that } \theta \text{ is defined}$$
$$= 0.5 \quad \text{if there are no } C \text{ for which } \theta \text{ is defined}$$

Finally, on the analogy of (6), and using the composite distance function ds:

$$(12)\quad \mathrm{PROB}(A\rightarrow B) = [(\mathrm{STR}_A(A\rightarrow B)+\mathrm{STR}_B(A\ B))/2.0]/\mathrm{ds}(A,B)$$

In the next section we look at some of the more important properties of PROB, continuing to use as examples scenes of circular elements of the same size that vary only in brightness. We merely observe here that PROB does allow for both proximity and similarity edges, and will handle cases like Figs. 12 and 13.

Brightness is the similarity variable of choice throughout this presentation for two reasons: it is relatively easy to quantify (unlike texture) and relatively easy to have printed (unlike color). Defining SIM functions over more general classes of elements is decidedly non-trivial and -- under somewhat different names (e.g., "similarity grouping" [3]) and in less well-defined contexts -- has been investigated by a number of psychologists, notably Julesz [4] and Beck [5]. Any SIM function, however, must meet the following requirements:

(13) $SIM(A,B) \geq 0.0$

(14) $SIM(A,B) = 0.0$ if and only if A and B are identical

(The SIM scale begins at zero, rather than one as does the proximity scale; this accounts for the use of SIM+1.0 in (8) and (10).) It goes almost without saying that SIM should be insensitive to the absolute sizes of elements and should reflect greater dissimilarity with larger values.

E. Experiments with and properties of PROB

If PROBd is correct for the highly restricted scenes to which it may be applied, PROB should give the same results for the same scenes. Actually, as can be seen from inspection of (7), (8), (9) and (14), the probabilities according to PROB are exactly half (see, however, the description of IMPSIM in section II.F) those of PROBd. Thus, for scenes of non-contrasting elements edge probabilities are in the same proportion, although lower, and, very naturally, added contrast between elements may increase values of PROB as edges significant for reasons of proximity are rendered even more apparent. PROB, then, can be considered a generalization of PROBd.

PROB retains the important property of scale independence in this generalization since SIM and ds, like d, are invariant under scale changes. Several other interesting features of PROB are exemplified in Figs. 15 and 16, counterparts of Figs. 7 and 8.

It will be recalled from section II.D that a uniform shrinkage or expansion of elements led to an overall decrease or increase in edge probability. This remains true for scenes of contrasting elements; however, edge probabilities no longer change at the same rate. It is easily seen that STR values are unaffected by this type of transformation -- changes in PROB are due solely to changes in ds. It follows that probabilities of edges spanning dissimilar elements will be more resistant to change than those of edges spanning elements that are relatively similar. We could describe this effect by saying that shrinkage diminishes the effects of similarity (in comparison to proximity), whereas expansion enhances them. Fig. 17 is an example of greatly altered perception under this transformation.

Most edge detectors (for example, those of Roberts [6], Rosenfeld and Thurston [7], or Hueckel [8]) are designed to operate on digitized pictures exclusively; PROB, in contrast, is defined on a much broader class of scenes which includes digitized pictures. The latter, from our point of view, consist of elements (the "picture points") that are evenly distributed (rectangularly or hexagonally) and differ in intensity (and, possibly, color). Apart from its greater generality, PROB is distinguished from other edge detectors in not being confined to delineating areas that contrast strongly with neighboring areas in average brightness: it is also able to handle <u>structural</u> differences of the kind that characterize different visual textures, as shown in Fig. 18. We believe that PROB is a significant first step towards the automatic analysis of texture; in section III.B we outline the role of this edge function in a complete system capable of such an analysis.

Fig. 19 shows even more sharply that a uniform approach to the detection of edges <u>cannot be based on contrast alone</u>: as the scene as a whole becomes more homogeneous with respect to average brightness (when viewed from a sufficiently great distance Fig. 19(b) contains no inner circular region -- that is, in order to see this region properly the elements must be individually visible) the perception of the central area is <u>facilitated</u>. Accordingly, the values of PROB along the boundary of this area are higher in (b) than in (a) — but edge detectors based on (average) contrast between areas of the scene will produce the opposite result.

This example also highlights another distinctive -- and advantageous — feature of the overall treatment of edges in this paper:

Fig. 15 Edges beoome more probable as elements become more similar. (Compare Fig. 7)

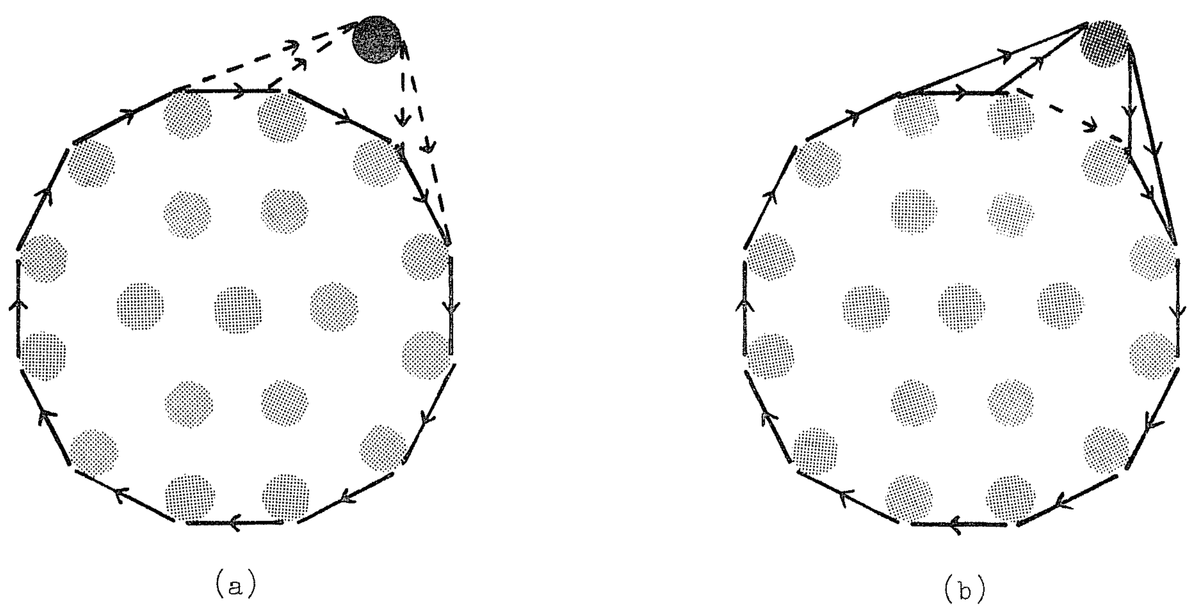


Fig. 16 "Planet and satellite" of (a) becomes "circle with protuberance" in (b). (Compare Fig. 8)

(a)

(b)

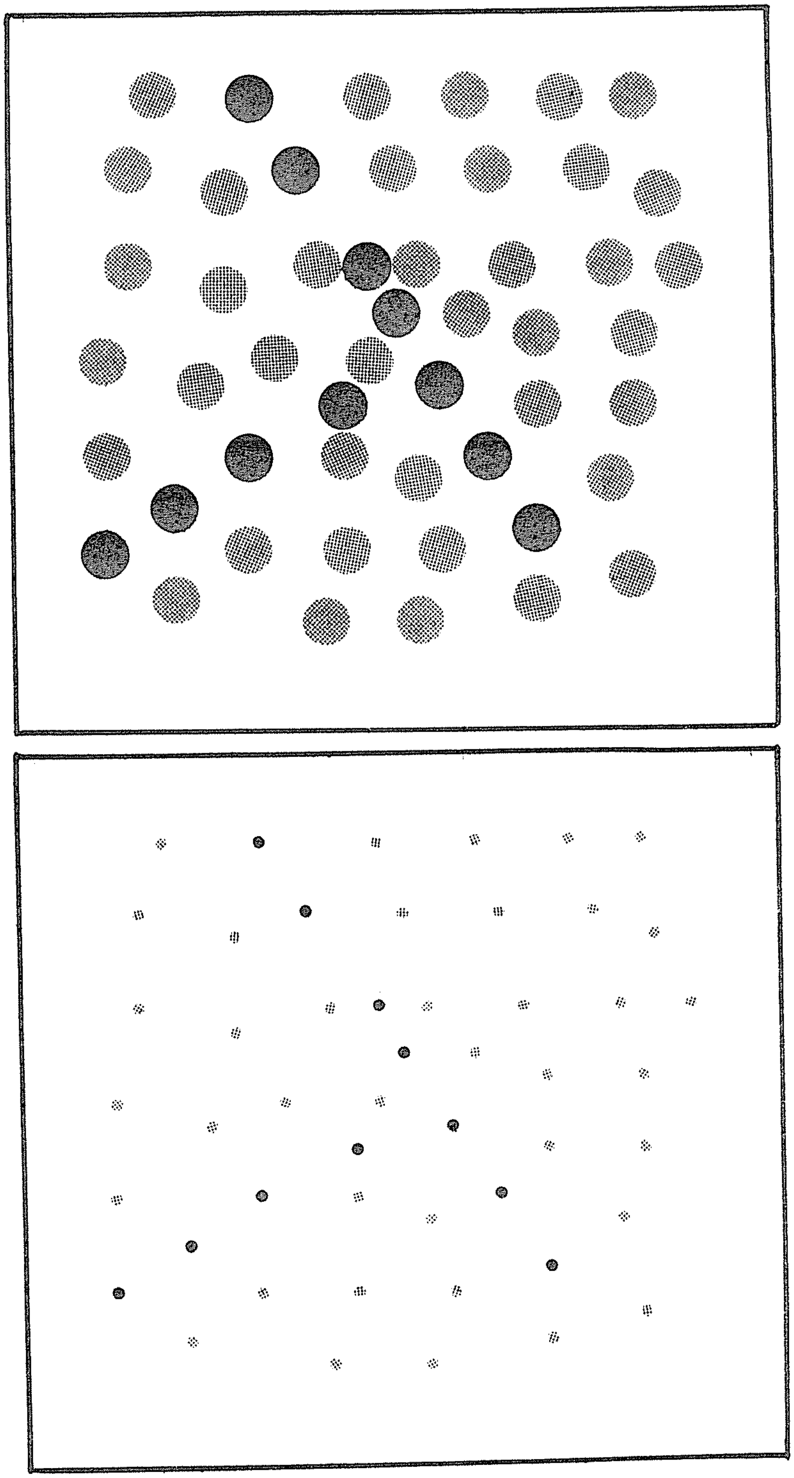
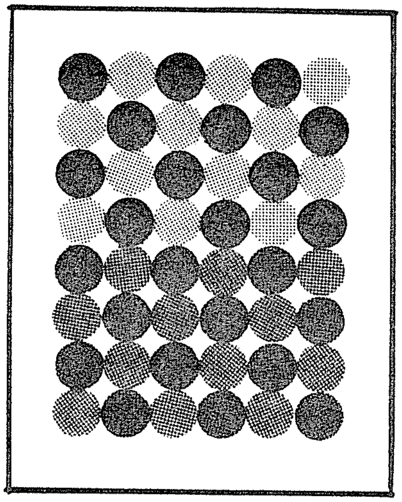


Fig. 17 (a)"Expanded" scene: "T" organization, due to similarity edges, predominates. (b)"Shrunken" scene (the same brightness and distribution as (a), but smaller elements): "7" organization, due to proximity edges, predominates.

(a)



(b)

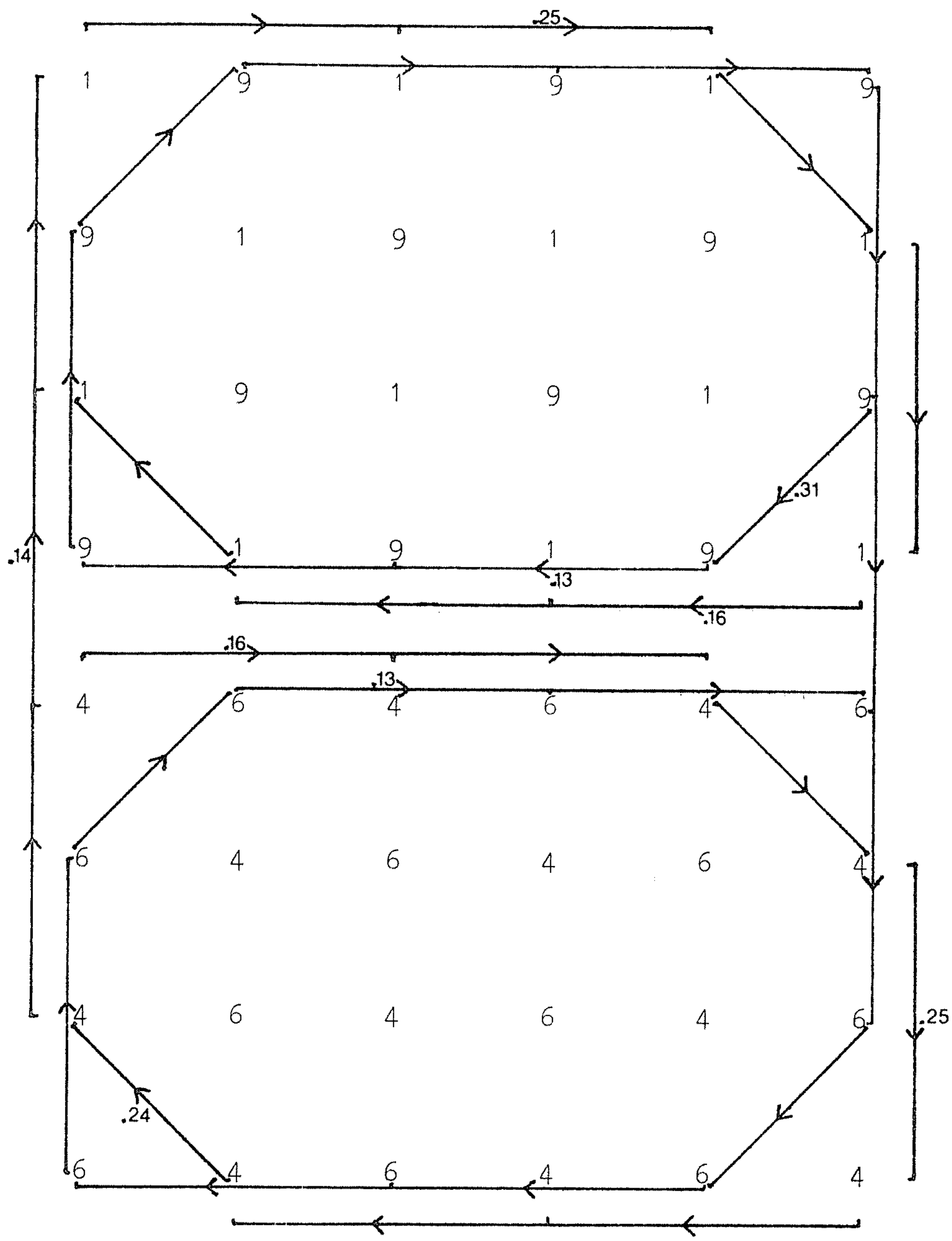


Fig. 18 (b)The intensity of elements in (a) and edges (not shown in actual position because of overlap) found by EDGE (using for SIM the absolute value of the difference in intensity). The edges suggest that each textured region really consists of two regions whose boundaries coincide (except at corners).

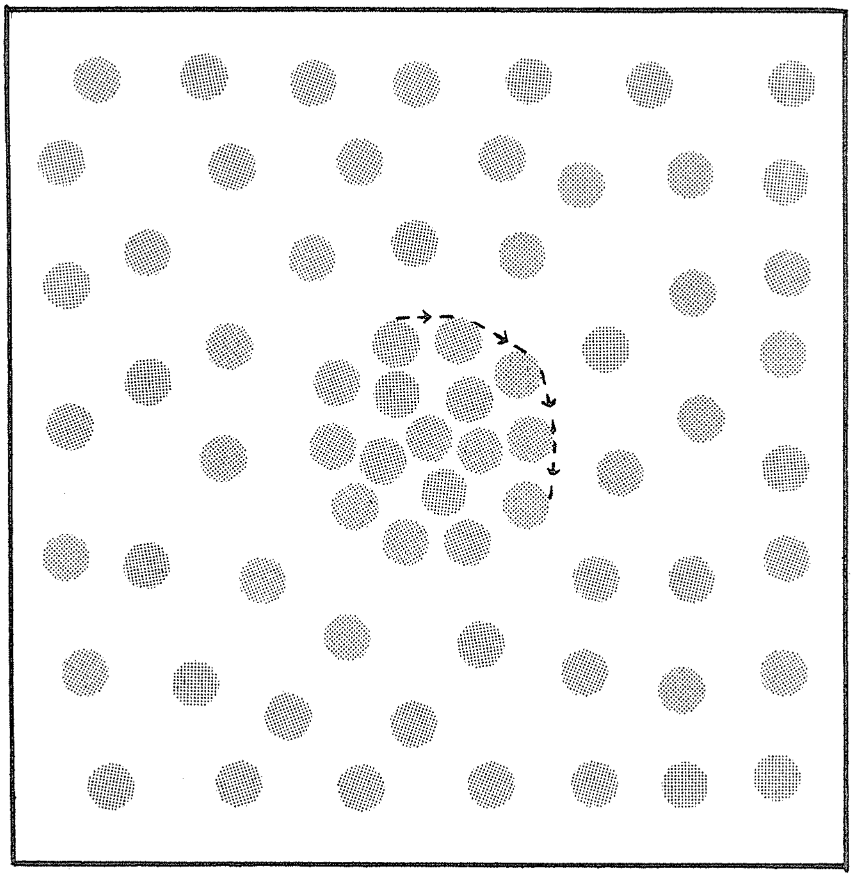
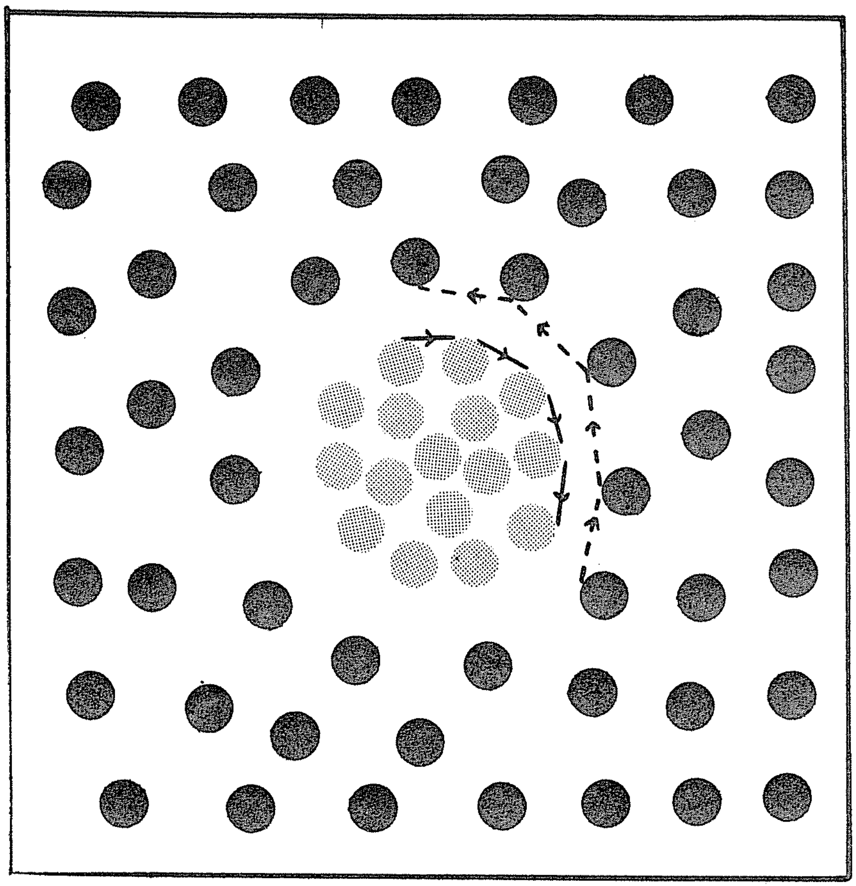


Fig. 19 (a)



(b)As elements in the area of lower density become darker, and average intensities in the two regions approach one another, two perceptual changes occur: the circular area becomes easier to see, and a "hole" appears around the circle. PROB correctly reflects these changes with larger values for edges outlining the circle (resulting from an added similarity contribution) and the emergence of similarity edges around the hole.

that is, the attachment of each edge to a single shape (as discussed in section II.A). In Fig. 19(a), edges outline only the circle and enclosing square; as the background darkens in (b), edges appear around the "hole" as well. These edges explain differing global perceptions of the two scenes: in (a), we see two "positive" (to use the artist's term) shapes, namely the circle and square — the "nut," or square with a hole in the center, exists only as a "negative" complement; in (b), the "nut" becomes directly visible. Because they are both positive in (b) the circle and surrounding hole are independent in both size and shape, as evidenced here by their different sizes (and very slightly different shapes). Although most figure-ground distinctions are probably determined by "higher level" perceptual mechanisms — that PROB, in other words, can explain only certain basic cases of this phenomenon — we believe that our treatment of edges permits a very direct and "low level" interpretation of the figure-ground phenomenon generally.

F. EDGE: an efficient method of estimating PROB

PROB may be calculated very simply by brute force: for every pair of elements, A and B, compute probabilities for A→B and B→A by looking at all other elements, C, in the scene. For scenes with few elements this strategy is satisfactory. EDGE, however, was designed to handle quickly scenes of more than 10,000 elements — which is made possible by the quasi-local nature of PROB.

This local character has two aspects: first, the most probable edges span elements that are close together and similar; second, those elements found to be most injurious (in the sense of (9)) to an edge are close and similar to the elements spanned by it. EDGE takes advantage of this by constructing, for each element A, a "neighborhood" consisting of the nearest (in terms of ds) elements in every direction. In practice "every" becomes sixteen, and it is this small neighborhood that determines which edges will be computed and which C's will figure in the computations.

To be precise, if N[A] is the neighborhood of A, STR_A will be computed for edges A→e and e→A, for all e in N[A]. For each edge, a maximum of eight values of EFF must be computed at A, one for every C in N[A] outside the edge (see Fig. 20). Only thirty-two edges that span A are possible, but this number is further reduced by immediately discarding both A e and e A for any e such that A is not in N[e]. PROB itself is calculated for A B when STR_A(A→B) and STR_B(A→B) have both been calculated.

A good deal of parallelism is possible with this approach. For example, in the construction of neighborhoods all directions can be dealt with simultaneously; the relevant EFF values for each edge can be computed independently; STR_A(A→e), for all e in N[A], can be computed at the same time; and, most importantly, every element can be treated simultaneously as an "A."

It should be noted that the above strategy for computing PROB may not be accurate: there are pathological cases in which the "worst C" is not in N[A]. The restriction to sixteen directions also causes errors, and of a more serious type, as quite good edges may not be considered at all in certain situations.

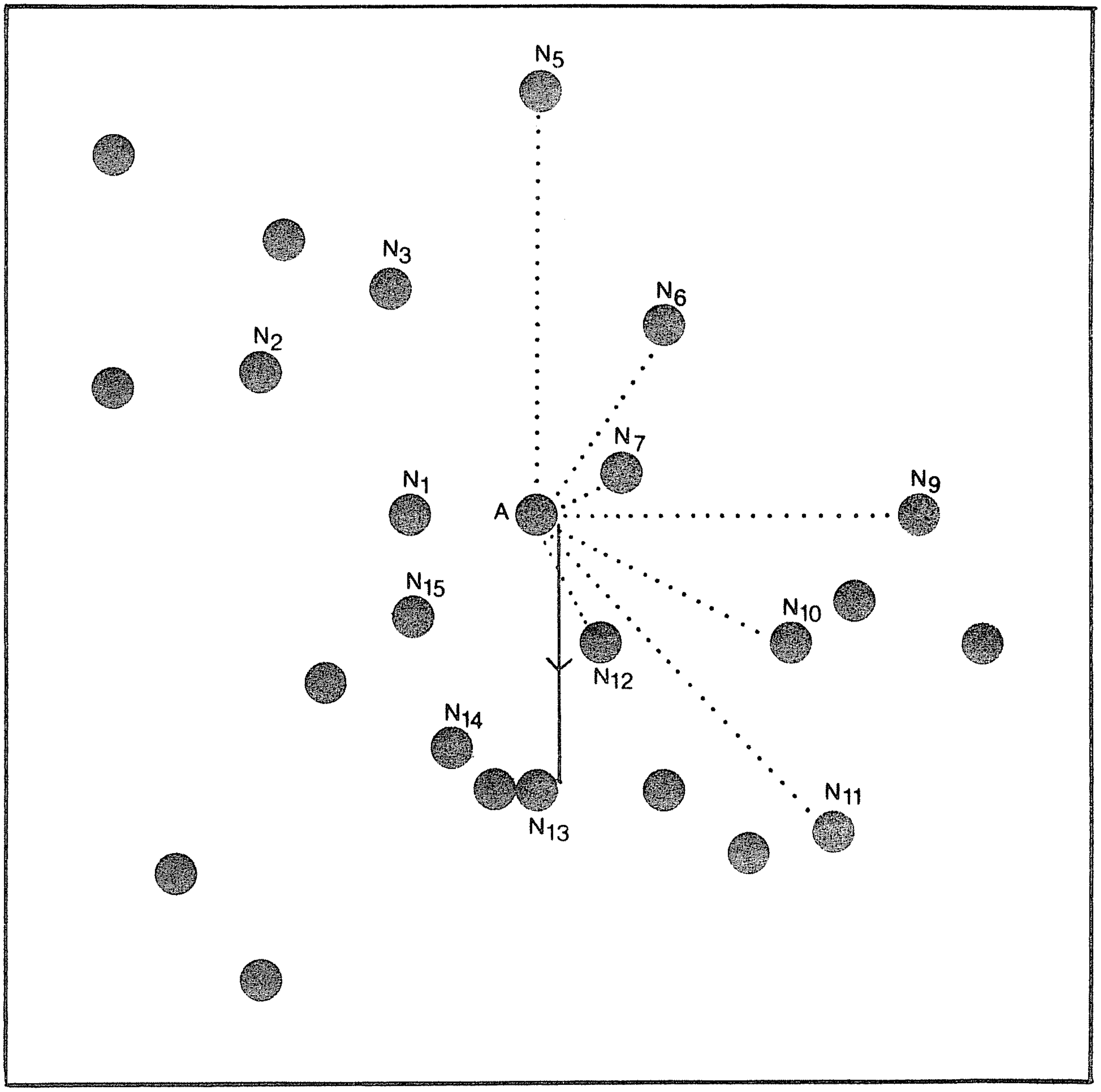


Fig. 20 The neighborhood of A, consisting of the nearest (in terms of ds) elements in each of 16 sectors (some sectors, such as 4 and 16, may be empty). In the calculation of $STR_A(A \rightarrow N_{13})$, EFF is calculated for only seven elements, indicated above by dotted lines.

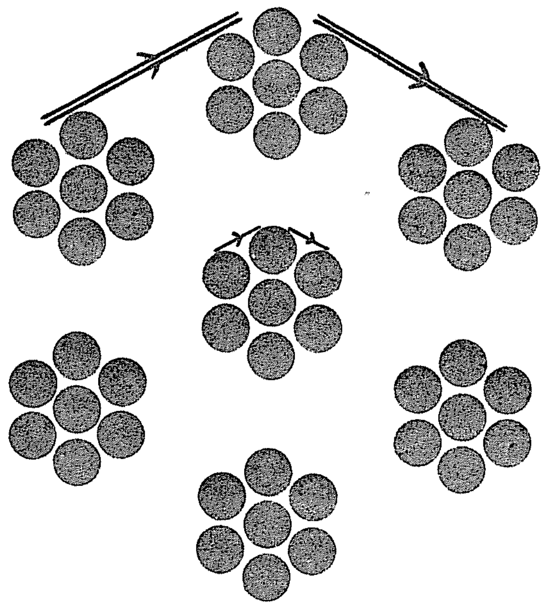


Fig. 21 Edge detection at two levels in a hierarchical structure. Edges and boundaries at one level (single arrows) are used in the formation of new elements at the next. These elements are used in turn for further edge detection (double arrows) at this higher level.

EDGE is an experimental program and is parameterized accordingly. Perhaps the most important instance of this is the use of a scaling factor in (9) (and, of course, in (11)):

$$(9)' \quad STR_A(A\rightarrow B) = \min(EFFd_A(C,A\rightarrow B)+IMPSIM*EFFs_A(C,A\rightarrow B))/(1.0+IMPSIM)$$
$$\text{for all C such that } \theta \text{ is defined}$$
$$= 1.0/(1.0+IMPSIM) \quad \text{if there are no C for which } \theta \text{ is defined}$$

Adjusting this parameter changes the relative importance of similarity with respect to proximity. A value of ten, for example, will bias the program towards the perception of similarity edges. The introduction of IMPSIM in this equation is not a hedge — on the contrary, we believe that the variation in behavior brought about by changing IMPSIM has its counterpart in human perception: that "set," "expectation," "experience" — in general, the circumstances in which the perceiving takes place — may emphasize similarity factors at the expense of proximity factors, or vice versa. Although PROB assumes a single function for similarity, any such function that captured adequately the differences between elements in fairly unrestricted scenes would no doubt break down into several components, for example brightness, size, and color. We believe that varying the contributions of these components within the total similarity function would produce a further interesting parallel to the flexible perceptions of humans.

III. Discussion and Conclusion

A. Generalizing the edge probability function

We have described an edge detector that is both very simple and very general. It can be applied to digitized pictures, or to scenes of randomly arranged colored disks. Put another way, this is an edge detector for a very general class of pictures, which includes the above examples: in full, it comprises scenes of disks, all the same size but allowed to differ in any other way, arbitrarily placed.

The complete generalization — to scenes of elements that are not identical in shape and size, and not required to be circular — has been thoroughly carried out, but not yet programmed and, therefore, only sketchily tested. Although it is beyond the scope of this paper to describe this extension, we briefly note that it involves new measures of distance and angle for complex shapes, and nothing more: that is, d and Θ are redefined, but (7) through (12) remain valid.

B. Edge detection in a complete visual system

Edge detection, in our view, is the first stage in the segmentation of a scene into its meaningful parts. The most important step in segmentation, after edge detection, is boundary formation, the chaining together of edges into longer boundaries. A system for boundary formation — called BNDRY — based on minimal spanning tree (Zahn [9]) clustering of edges, using their probability, length, orientation and location, has also been designed and coded.

The tandem EDGE-BNDRY was developed as a non-semantic component of a complete perceptual system, in particular, one capable of discriminating visual textures. In terms of existing programs, it could function as a pre-processor for a program like that described by Grape [10] in which recognition is based on matching object prototypes to possibly incomplete boundaries. However, we have in mind its use in a recursive system, where new elements may be formed at successive "levels" and then used for edge detection and boundary formation. A layered structure of this kind has been described by Uhr [11]. The disk-elements of our previous examples could be thought of as occurring at any level in this structure, as exemplified by Fig. 21: here, the largest disk has a two level "dotted" texture.

The hierarchical analysis of texture has been discussed by, for example, Pickett [12]; Tomita et al. [13] have designed a hierarchical system using a regional approach; Rosenfeld and Thurston [7] have suggested using an edge detector as the basis for such a system. Our proposed system, we feel, combines the best features in the latter two. Edge detection appears to have a greater potential for parallel implementation (and, therefore, for greater speed of computation) than region finding; moreover, it is inherently more general since important boundaries in real pictures are not always closed, and regions that correspond to such boundaries do not exist. Traditional edge detection, on the other hand, is tied to the concept of digitized picture as an approximation to a "picture function" defined on the real plane and sees edge detection as an attempt to approximate some form of the gradient of these functions (this point of view is summarized in [14]). We substitute a uniformly finite and hierarchical concept of picture or scene and a general edge detection process that, ideally, can be applied to any collection of elements and, consequently, "bootstrapped" naturally in the analysis of a scene.

C. Future research

One minor obstacle in the way of completing the proposed system is the extension of EDGE to handle more general cases of spatial interaction between elements along the lines sketched in section III.A. The major obstacles appear to be the design of a subsystem to form elements at each level and the definition of similarity over a set of elements varying in color, texture, shape and size, as well as brightness. Progress towards these goals depends heavily on psychological experimentation, in particular continuation of the "similarity grouping" line of investigation [3] [4] [5] and research on textural features or descriptors [15]. Conversely, these empirical studies might benefit from the explicit use of a theoretical model like the EDGE-BNDRY system, with which hypotheses about similarity could be linked automatically to the perception of form.

References

[1] M. Wertheimer, "Untersuchungen zur Lehre von der Gestalt," Psychol. Forsch., vol. 4, 1923, pp. 301-350; English translation, "Investigations on the Gestalt theory," in A Source Book of Gestalt Psychology, W.D. Ellis, Ed. New York: Harcourt, 1938, pp. 71-88.

[2] H. Freeman, "Boundary encoding and processing," in Picture Processing and Psychopictorics, B.S. Lipkin and A. Rosenfeld, Eds. New York: Academic Press, 1970, pp. 241-266.

[3] F. Attneave and R.K. Olson, "What variables produce similarity grouping?," The American J. of Psychology, vol. 83, pp. 1-21, 1970.

[4] B. Julesz, "Cluster formation at various perceptual levels," in Methodologies of Pattern Recognition, S. Watanabe, Ed. New York: Academic Press, 1969, pp. 297-315.

[5] J. Beck, "Effect of orientation and of shape similarity on perceptual grouping," Perception and Psychophysics, vol. 1, pp. 300-302, 1966.

[6] L.G. Roberts, "Machine perception of three-dimensional solids," in Optical and Electro-Optical Information Processing, J.T. Tippet, et al., Eds. Cambridge, Mass.: MIT Press, 1965, pp. 159-197.

[7] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," IEEE Trans. on Computers, vol. C-20, no. 5, pp. 562-569, May, 1971.

[8] M.H. Hueckel, "An operator which locates edges in digitized pictures," JACM, vol. 18, no. 1, pp. 113-125, Jan., 1971.

[9] C.T. Zahn, "Graph-theoretical methods for detecting and describing Gestalt clusters," IEEE Trans. on Computers, vol. C-20, no. 5, pp. 68-86, May, 1971.

[10] G.R. Grape, "Model based (intermediate-level) computer vision," Stanford Artificial Intelligence Memo AIM-201, May, 1973.

[11] L. Uhr, "Layered 'recognition cone' networks that preprocess, classify and describe," IEEE Trans. on Computers, vol. C-21, no.7, pp. 758-768, July, 1972.

[12] R.M. Pickett, "Visual analyses of texture in the detection and recognition of objects," in Picture Processing and Psychopictorics, B.S. Lipkin and A. Rosenfeld, Eds. New York: Academic Press, 1970, pp. 289-308.

[13] F. Tomita, et al., "Detection of homogeneous regions by structural analysis," Third International Joint Conference on Artificial Intelligence (advance papers of the conference), pp. 564-571, Aug., 1973.

[14] R. Duda and P. Hart, Pattern Classification and Scene Analysis. New York: John Wiley & Sons, 1973, pp. 268-272.

[15] R.M. Pickett, "On enhancing visual perceptions of solar flare precursors," Air Force Cambridge Research Laboratories Report AFCRL-TR-73-0210, April, 1973.