COMPUTER RECOGNITION OF PRISMATIC SOLIDS

ARNOLD KOONS GRIFFITH

August 1970

PROJECT MAC

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Cambridge Massachusetts 02139

COMPUTER RECOGNITION OF PRISMATIC SOLIDS*

Abstract

An investigation is made into the problem of constructing a model of the appearance to an optical input device of scenes consisting of plane-faced geometric solids. The goal is to study algorithms which find the real straight edges in the scenes, taking into account smooth variations in intensity over faces of the solids, blurring of edges and noise. A general mathematical analysis is made of optimal methods for identifying the edge lines in figures, given a raster of intensities covering the entire field of view. There is given in addition a suboptimal statistical decision procedure, based on the model, for the identification of a line within a narrow band on the field of view given an array of intensities from within the band. A computer program has been written and extensively tested which implements this procedure and extracts lines from real scenes. Other programs were written which judge the completeness of extracted sets of lines, and propose and test for additional lines which had escaped initial detection. The performance of these programs is discussed in relation to the theory derived from the model, and with regard to their use of global information in detecting and proposing lines.

*This report reproduces a thesis of the same title submitted to the Department of Mathematics, Massachusetts Institute of Technology, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ACKNOWLEDGEMENTS

I would like to express my considerable appreciation for the advice and encouragement of Professors Seymour A. Papert and Marvin L. Minsky. For additional encouragement and editorial assistance, I am also indebted to Miss Patricia A. Martino.

Thanks are also extended to Professors Manuel Blum, Adolfo Guzman, and Edward Fredkin; to Dr. Thomas O. Binford, and to Miss Carole Naymie. I am further indebted to Project MAC, M.I.T., for financial support.

Arnold K. Griffith

Work reported herein was supported in part by Project MAC, an M.I.T. research project sponsored by the Advanced Research Projects Agency, Department of Defense, under Office of Naval Research Contract N00014-20-A-0362-0002. Reproduction of this report, in whole or in part, is permitted for any purpose of the United States Government.

Government contractors may obtain copies from:

Defense Documentation Center, Document Service Center, Cameron Station, Alexandria, VA. 22314

Other U.S. citizens and organizations may obtain copies from:

Clearinghouse for Federal Scientific and Technical Information (CFSTI), Sills Building, 5285 Port Royal Road, Springfield, Va. 22151

TABLE OF CONTENTS

CHAPTER 1

TOWARD A THEORY OF THE OPTIMAL USE OF INTENSITY INFORMATION IN THE DETECTION OF LINES IN A VISUAL FIELD

I. INTRODUCTION AND DEFINITIONS

We are interested in studying procedures which identify the images of straight lines in some two-dimensional (picture-like) projection of an array of real objects, given some large set of intensity values from points on the two-dimensional projection plane. We shall confine our attention to arrays of plane-faced solids. Applying a line-detection procedure to the image of an array of such objects yields a "description" of the objects which is, in effect, a line drawing of the objects from a particular point of view. Such a description is complete in the sense that it

contains all information obtainable from that point of view which is relevant to determining the shape, location in space, and so forth of the objects. Information of the latter sort, which may be regarded as constituting a more detailed description of the array of objects, may be derived from one or several "line drawings" of a particular array of objects, by methods surveyed in the last chapter of this paper.

In the first chapters of this paper, we shall investigate the problem of constructing line detection procedures which make optimal use of the set of intensity values to which they are applied, by being maximally accurate, in various senses, about the lines they claim to exist. Such investigations require a formalism, definitions and so forth, which will presently be developed in this section. In the meantime we shall informally discuss several considerations relevant to the development of such a theory, with the idea of motivating some of the formal presentation of the rest of the chapter.

First, it appears that a theory of optimal edge finding procedures would be intractably difficult if we were to consider procedures which decide on the locations at which to measure each successive intensity value. This leads into the realm of

sequential decision procedures; and we should not attempt this until we first have a reasonable understanding of the simpler non-sequential case. One may get an idea of the inherent complexity of optimal sequential decision procedures by consulting, e. g., Bellman (Bellman). For the moment, then, we shall assume that the intensities are obtained from the two-dimensional projection of the array of objects in a predetermined pattern, such as at the vertices of a grid of squares, and that the line finding procedure acts on exactly these intensities. We shall return to the problem of sequential analysis of scenes in chapter four, where the matter is approached at the heuristic level.

A second assumption to be made concerning the set of intensities upon which a line finding procedure acts concerns the relation between the actual intensity at a point in the two dimensional view of the scene and the intensity available to the line-finding procedure, which would be obtained by a real optical input device. In general such a device is subject to various forms of optical blurring and distortion, as well as to random noise. If there is nothing known about the noise, then

the set of intensity values obtained by such a device contains no information about the array of objects at hand, in particular about the locations of the projections of their edge lines. It is a reasonable assumption that the obtained values of intensity at any point be normally distributed about the actual intensity, and that the standard deviation of such a distribution be a function of the intensity at the point. It is further reasonable to assume that any spatial distortion of the intensities be accountable for principally as some sort of blurring phenomenon. These forms of distortion and noise are present in any real optical or electro-optical system; and certain devices, such as an image-dissector, may be designed so that these are the principal effects.

A third restriction involves the relation between the size of the objects in the scenes under consideration, or more properly the size of the images of these objects in a two dimensional projection, and the spatial density of points at which intensities are measured. If the intensities are obtained in a simple grid pattern with a spacing of one unit on the image plane, it would clearly be impossible to detect edge lines whose images, for example, were shorter than one unit. In judging the optimality of an edge detector, it would surely be unfair to

penalize it for not detecting so short a line. A solution to the problem might be to consider only arrays of objects whose images are larger than a certain size relative to the spacing of the points at which intensities are measured. An alternative approach, which will actually be employed, is as follows: Choose a suitable geometry of points in the image space at which to measure intensities. Then define a simple predicate on exact values of intensity which would be obtained at these points in the absence of all noise, distortion and imperfections in the objects being examined. Then use this criterion for the existence of a line to judge the performance of a line detection procedure which measures intensities at the same points, but with a real optical input device. By this means, one is essentially determining how well the procedure "sees through" the distortion and noise; and does not penalize it for not "seeing" lines that it could not possibly detect from intensities measured at points of the given grid even in the absence of noise.

The concepts so far discussed may be summarized in the following definitions.

Definition: an intensity function, I(x, y), is a real-valued function whose values and derivatives are defined

on a square in 2-space, except for a set of measure 0.

Definition: a description space, D, is an arbitrary set of elements $\{d_i\}$.

Definition: a sampling set, S, is a finite set of points $\{(x_i, y_i), \ldots\}$.

Definition: a noise-free sample, I(S), of an intensity function I(x, y), is the set of values of $I(x_i, y_i)$ for the points (x_i, y_i) of S.

Definition: a noisy sample, J(S), is defined as above, but with the values of intensity modified by some distortion and noise functions.

Definition: an identification function, F(I), is a mapping from a set of intensity functions onto some description space.

$F: \{I\} \rightarrow D$

Definition: a noise-free sample identification function, G(I(S)), is a mapping from a set of noise-free samples onto a description space.

$G: \{I(S)\} \rightarrow D$

Definition: a noisy-sample identification function, H(J(S)), is a mapping from a set of noisy samples to a description space.

$H: \{J(S)\} \rightarrow D$

Definition: a sample identification function G, together with a space of noise-free samples $\{I(S)\}$ and a description space D are said to be descriptively complete with respect to an identification function F if

$F(I)=d_i \Leftrightarrow G(I(S))=d_i$

This definition describes the situation where all that one wishes to say about a scene may be inferred from the values of intensity at a predetermined finite set of locations on the visual field.

Definition: an intensity function I*(x, y) on a domain A is said to be equivalent under the blurring function B(x, y) to a function I(x, y) if:

$$I^*(x, y)= \int_A I(u, v)B(u-x, v-y)du\ dv$$

Definition: an intensity function I**(x, y) is said to be Gaussian-noise modulated with respect to an intensity function I*(x, y) if successive values of $I^{**}(x_0, y_0)$ distribute normally about $I^*(x_0, y_0)$ with

a standard deviation which is dependent solely on the value of $I^*(x_0, y_0)$.

Definition: a transducer which goes from an intensity signal $I(x, y)$ to a Gaussian-noise modulated function $I^{**}(x, y)$ of a function $I^*(x, y)$ which is equivalent under blurring to $I(x, y)$, is called a transducer subject to fundamental distortion.

Definition: a scene system with identifiers consists of:

$\{I_i(x, y)\}$, a set of intensity functions on a common domain;

$\{(x_i, y_i)\} = S$, a sampling set;

$\{I_i(S)\}$, a set of noise-free samples, one for each intensity function I_i;

$\{J_i(S)\}$, a set of noisy samples, an infinite number of $J_i(S)$ for each $I_i(S)$. The relation between a set of $J_i(S)$ and the corresponding $I_i(S)$ is entirely accountable for by fundamental distortion.

$\{d_i\} = D$, a description space;

$F: \{I_i(x, y)\} \rightarrow D$, an identification function;

$G: \{I_i(S)\} \rightarrow D$, a noise-free sample identification function;

$H: \{J_i(S)\} \rightarrow D$, a noisy-sample identification function.

Finally, for completeness:

Definition: a scene system with identifiers is said to be descriptively complete if the relevant parts conform to the descriptive completeness criterion above.

The various restrictions set forth in the earlier parts of this section may be simply summarized by saying that we will concern ourselves exclusively with scene systems with identifiers which are descriptively complete.

II. COMPLETE SYSTEMS OF STRAIGHT-LINE SCENES

A criterion for what constitutes a line in a noise-free sample will now be given. This definition will provide a basis for a definition of a line in an intensity function which will be appropriate for ensuring completeness.

Prior to making this definition, however, it is necessary to digress slightly to consider a plausible restriction on the geometry of the sampling set. An important desideratum of any line finder is that it not be especially sensitive to lines in a particular part of the field of view, or at any particular orientation in the field of view. It may be that of two procedures, one will find more lines with a given amount of effort, but at the

expense of uniformity in this sense. One would still not prefer the "better" line finder, on account of its inhomogeneous or anisotropis behavior. One thus wishes to guarantee that any line-finding procedure that is a candidate for being considered optimal is in addition uniform in this sense. A necessary condition seems to be to have a sampling set which is distributed uniformlu over the entire field of view. Without proving this last assertion, we shall henceforth restrict our attention to systems whose sampling sets consist of points on all vertices of some grid of squares covering the entire field of view.

A straight line in a noise-free sample over a sampling region as described above may be defined in terms of the notions of discontinuity and anomaly. We define three such notions for, e.g., a set of colinear intensities i_1, i_2, i_3 and i_4, using constants K, K' and K'':

1) A discontinuity of intensity occurs at a point between the points at which i_2 and i_3 were obtained if:

$$|i_2 - i_3| \geq K(|i_1 - i_2| + |i_3 - i_4|).$$

2) An anomalous intensity occurs at the point where i_3 was obtained if:

$$|i_2 - (i_1 - i_3)/2| \geq K'.$$

3) A discontinuity in shading occurs at a point between the points at which i_2 and i_3 were obtained if:

$|(i_1 - i_2) - (i_3 - i_4)| \geq K''$.

A line may be defined as two or more such points, all of the same type and all adjacent and colinear.

It is now easy to specify a class of systems with identifiers which are descriptively complete in their identification of lines in a scene. One may simply choose a scene system with a description space appropriate for describing lines, and define a noise-free sample identification function G in such a way that it appropriately identifies all the lines in a noise-free sample according to the above definition of a line. One may then define the identification function F as follows:

$F(I)=d_i \iff G(I(S))=d_i$.

The completeness follows a fortiori.

In whatever system we shall henceforth be considering, we shall assume that the description space is the power set of the set of all lines definable as above on noise-free samples. We shall further assume that the noise=free sample identification function of the system is as defined in the previous paragraph. The choice of an optimal noisy-sample identification function for this type of system is the subject of the remainder of this paper.

III. OPTIMIZATION CRITERIA

The variables to be considered in discussing optimal object recognition algorithms are number of intensities read, time taken to read each intensity value, length of computation time, and rate of error in identifying lines. We will not be concerned with computation time, as this would require more knowledge about "mimimal" versions of the algorithms concerned than is available today. On the other hand, the number of points at which intensities are taken, and the length of time taken in getting the intensities can be related in a natural way to the accuracy of identifying and locating lines, or other features of a scene. Intuitively, there must exist a trade-off between the measurement cost of a certain amount of intensity information; and accuracy, measured in terms of number of lines correctly identified. Optimization may consist in minimizing some type of error rate, with a fixed overhead in effort. One may, alternatively, wish to set a certain error rate criterion, and minimize the amount of effort necessary to achieve it. It is easy to see that this latter

form of optimization may be reduced to the former: Suppose there exists a procedure P(E) which, using E units of effort, minimizes the error rate. If it is desired to minimize effort for a fixed error rate, one may simply consider all possible instances of the procedure P(E) with error rate below the desired one, and choose the procedure P(E*) where E* is a minimum among all such amounts of effort. This procedure uses the least possible amount of effort while guaranteeing the error rate to be below a certain value. We may thus without loss of generality restrict our attention to procedures which make a minimum of errors for a given amount of overhead.

Further discussion of error rate requires a more detailed discussion of various types of classification errors and the probabilities with which they occur. We shall base this discussion on the notion of false positive error, the assertion by a line finder that a line exists at a particular place on the visual field when no such line exists in the scene under consideration, and upon the converse notion of true negative error. We will begin with some definitions regarding line identification functions.

Definition: In a system of the type under consideration, for some noise-free sample $I_i(S)$ and a corresponding

noisy pattern $J_j(S)$, we call the set of lines $H(J_j(S))-G(I_i(S))$ the set of <u>false positive errors</u> in the identification of the scene from the noisy sample $J_j(S)$.

Definition: For similar conditions, the set $G(I_i(S))-H(J_j(S))$ is called the set of <u>true negative errors</u> in the identification of the scene from the noisy sample $J_j(S)$.

Definition: Let A and B be two sets, then the <u>difference cardinality</u>, N(A, B), is the cardinality of the set A-B.

It follows immediately that for any $J_j(S)$ and corresponding $I_i(S)$, the number of false positive errors is given by:

$$N(H(J_j(S)), G(I_i(S))),$$

and the number of true negative errors by:

$$N(G(I_i(S)), H(J_j(S))).$$

With the aid of the function N defined above, we may indicate the performance of a noisy sample identification function in terms of the number of errors of each type it will make in describing any particular noisy sample. Before being able to indicate the <u>over-all</u> performance of a noisy sample identification function, one must as well be able to specify the relative

probabilities of occurrence of the various pairs of noise-free and noisy samples. For the present we need only the first of the following definitions; however they will all ultimately be necessary.

Definition: $P(I_i(S), J_j(S))$ is the joint probability of occurrence of the noise-free sample $I_i(S)$ and the noisy sample $J_j(S)$.

Definition: $P(J_j(S)/I_i(S))$ is the conditional probability that the noisy sample $J_j(S)$ will occur given the occurrence of noise-free sample $I_i(S)$.

Definition: $P(I_i(S)/J_j(S))$ is defined similarly.

Definition: $P(I_i(S))$, $P(J_j(S))$ are the a priori probabilities of occurrence of the noise-free sample $I_i(S)$ and the noisy sample $J_j(S)$.

We may now discuss the expected rate of error of a classification system of the type under consideration.

Lemma: The false-positive error rate, or expected number of false positive errors made by a complete scene system is given by:

$$\sum_{i,j} P(I_i(S), J_j(S))N(H(J_j(S)), G(I_i(S)))$$

Proof: Whenever the noisy sample $J_j(S)$ occurs in conjunction with the scene representable by

the noise-free sample $I_i(S)$, the number of false positive errors is:

$N(H(J_j(S)), G(I_i(S)))$.

This occurs with probability

$P(I_i(S), J_j(S))$. QED

Lemma: The true negative error rate in a similar situation is given by:

$$\sum_{i,j} P(I_i(S), J_j(S))N(G(I_i(S)), H(J_j(S))),$$

the proof is similar.

With the aid of the following definitions, we may define two additional error-rate measures:

Definition: $M(A, B)$, the total number of errors is defined by:

$M(A, B)=N(A, B)+N(B, A)$.

Definition: $O(A, B)$, the identity predicate, is defined by:

$O(A, B)=1$ if $A=B$,

$O(A, B)=0$ if $A\neq B$.

Lemma: The correct recognition rate, or expected rate at which a scene is perfectly identified is given by the following:

$$\sum_{i,j} P(I_i(S), J_j(S))O(G(J_j(S)), H(I_i(S))),$$

the proof is as above.

Theorem: The total error rate, or expected number of errors made by a complete scene system is given by:

$$\sum_{i,j} P(I_i(S), J_j(S))M(G(J_j(S)), H(I_i(S))),$$

the proof is as above.

As a corrolary to the above theorems, we make the following definitions.

Definition: $a(J_j(S), d_j) = \sum_i P(I_i(S)/J_j(S))N(d_j, G(I_i(S)))$ is the expected rate of false positive errors on occasions when the noisy sample $J_j(S)$ occurs, provided the description d_j is assigned to it.

Definition: $b(J_j(S), d_j) = \sum_i P(I_i(S)/J_j(S))N(G(I_i(S)), d_j)$ is the expected error rate of true negative errors under similar conditions.

IV. OPTIMIZATION PROCEDURES

With the aid of the definitions in the previous section, we are in a position to discuss four optimization procedures. The first involves the maximization of the probability of correctly recognizing a scene in its entirety. The second involves minimizing the sum of the false positive and true negative error rates. The last two involve the maximization of one of the error

rates with the other error rate held constant. Since these latter two procedures are mathematically nearly equivalent, we shall consider only the case of holding the false positive error rate constant, and minimizing the true negative error rate.

Theorem: The scene recognition procedure which maximizes the rate of correct recognition of scenes in their entirety is given by the function $H(J_j)=d_k$ such that

$$\sum_i O(G(I_i(S)), d_k)P(I_i(S), J_j(S))$$

is maximized.

Proof: If description d_k is assigned to the noisy sample $J_j(S)$, then a perfect recognition of the scene will occur whenever $I_i(S)$, the corresponding state of the real world, has the property that $G(I_i(S))=d_k$. This occurs with a probability $P(I_i(S)/J_j(S))$ for a particular $I_i(S)$ having this property. Hence the expected rate of correct identification for the noisy sample $J_j(S)$ is:

$$\sum_{G(I_i(S))=d_k} P(I_i(S)/J_j(S)).$$

This is equivalent to:

$$\sum_i O(G(I_i(S)), d_k)P(I_i(S)/J_j(S))$$

or:

$$1/P(J_j(S))\sum_i O(G(I_i(S)), d_k)P(I_i(S), J_j(S)).$$

For a given $J_j(S)$, the rate of correct recognition is obtained by maximizing this sum. This optimization for a particular $J_j(S)$ is independent in its effect on optimizing the recognition of other $J_k(S)$. Hence this noisy-sample-wise optimization procedure provides an optimum recognition procedure for the set of all $J_j(S)$.

Theorem: The scene recognition procedure which minimizes the rate of false positive errors plus true negative errors is given by the function $H(J_j(S))=d_k$ with the property that:

$$\sum_i M(G(I_i(S)), d_k)P(I_i(S), J_j(S))$$

is minimized.

Proof: Similar to that of the preceding theorem.

The last case is a bit more complex.

Theorem: Suppose it is desired to keep the false positive error rate at or below a constant value α, while minimizing the true negative error rate. Then there exists a $\underline{c}$, depending only on α, with the property that the desired optimality may be achieved by assigning to $J_j(S)$ the description d_k such that

$c \cdot a(J_j(S), d_k) + (1-c)\, b(J_j(S), d_k)$

is minimal for all d_i.

Proof: 1) There must exist a description function having false positive error rate at or below α, and minimal true negative error rate. This follows from the fact that, e.g., $H(J_j(S)) \equiv 0$ gives a false positive error rate of 0. Hence the set of decision functions with false positive error rates below α is non-empty. There must exist one with minimal true negative error rate. Let it be denoted by $H(J_j(S))$, and its true negative error rate by β.

2) Consider a decision function $H^*_x(J_j(S))$ which assigns to $J_j(S)$ the description d_k such that $x \cdot a(J_j(S)), d_k) + (1-x)\, b(J_j(S), d_k)$ is minimal for all d_i. We shall show that for a suitable value $\underline{c}$ of $\underline{x}$, the function $H^*_c(J_j(S))$ has a false positive error rate at or below α, and a true negative error rate of β. From this it will follow that the function $H^*_{\underline{c}}(J_j(S))$, which is the function referred to in the theorem, has the desired optimality.

3) Let:

$a(j)=a(J_j(S), H(J_j(S)))$,

$b(j)=b(J_j(S), H(J_j(S)))$; and also

$a^*_x(j)=a(J_j(S), H^*_x(J_j(S)))$,

$b^*_x(j)=b(J_j(S), H^*_x(J_j(S)))$.

From the definition of $H^*_x(J_j(S))$, we know that for all j,

$$x \cdot a(j) + (1-x)\, b(j) \geq x \cdot a^*_x(j) + (1-x)\, b^*_x(j).$$

Multiplying by $P(J_j(S))$ and summing on j, we have:

$$x\sum_j a(j)P(J_j(S)) + (1-x)\sum_j b(j)P(J_j(S)) \geq$$

$$x\sum_j a^*_x(j)P(J_j(S)) + (1-x)\sum_j b^*_x(j)P(J_j(S)).$$

But, by definition,

$$\sum_j a(j)P(J_j(S)) = \alpha, \text{ and}$$

$$\sum_j b(j)P(J_j(S)) = \beta.$$

Also, letting

$$\sum_j a^*(j)P(J_j(S)) = \alpha^*_x, \text{ and}$$

$$\sum_j b^*_x(j)P(J_j(S)) = \beta^*_x,$$

we have:

$$x\alpha + (1-x)\beta \geq x\alpha^*_x + \beta^*_x.$$

4) It is not difficult to see that $\underline{x}$ may be chosen so as to make $\alpha^*_{\underline{x}} = \alpha$. This follows by assuming that $\alpha^*_{\underline{x}}$ is a continuous function of $\underline{x}$; and noting that $\alpha^*_1 = 0$, since H^*_1 minimizes false positive errors ignoring true negative errors, and $H(J_j(S)) \equiv \emptyset$ achieves a false positive error rate of 0. By similar reasoning, $\alpha^*_{\underline{x}}$ increases without bound as $\underline{x}$ decreases from 1. Thus for some value $\underline{c}$ of $\underline{x}$ the value of $\alpha^*_{\underline{c}} = \alpha$. This approach may be applied to the discrete case, where at least α^*_c will be very close to α. The expression of the previous step will then reduce to:

$c\alpha + (1-c)\beta \geq c\alpha + (1-c)\beta^*_c$.

But β is minimal, and $(1-c) \geq 0$, thus:

$\beta = \beta^*_c$. QED

CHAPTER 2

THE OPTIMAL USE OF REGIONAL INTENSITY INFORMATION IN THE DETECTION OF LINES

I. THE REGIONAL QUALITY OF ISOLATED LINES

In this section we shall direct our attention to situations where there exist no systematic spatial relationships among the lines appearing in scenes under consideration. We shall briefly discuss the extent to which lines in this case are "regional" in the sense that all intensity information relevant to the detection of a line is contained on the visual field within a certain short distance from the line.

Definition: Let S be a sampling region consisting of points uniformly distributed over a field of view. Let F(A, B, S) be a predicate on intensity values at points

of S, which attempts to decide whether or not there is a line segment through points A and B. Let α and β respectively be the false positive and true negative error rates of the function F(A, B, S). Then the quantity $\delta_F(\epsilon)$ is defined as the smallest distance such that there exists a predicate F*(A, B) which decides on the existence of a line through A and B only on the basis of intensities of S lying within this distance of AB, and which has false positive and true negative error rates respectively below $\alpha+\epsilon$, and $\beta+\epsilon$.

For some small value of ϵ, the value of $\delta_F(\epsilon)$ may take on a wide variety of values in different sorts of situations. For example $\delta_F(\epsilon)$ would be large in the situation where scenes consist of rectangular parallelapipeds. In this case, lines parallel to a proposed line, or incident with an end, constitute useful evidence for the existence of the proposed line. The spatial remoteness of useful evidence in this case might make $\delta_F(\epsilon)$ not appreciably smaller than the radius of the field of view for small values of ϵ. However, $\delta_F(\epsilon)$ may be relatively small in the case of scenes consisting of single discontinuities of intensity across the field,

or of sets of straight lines drawn at random across a sheet of paper. A hypothetical situation of particular interest is that in which scenes consist of lines which, within a relatively large radius, behave like lines given rise to by edges of plane-faced solids; but which are randomly distributed across the field of view. In cases such as this, two effects limit the predictive value of intensity information remote from a proposed line.

The first effect has to do with the unpredictibility of intensity values away from an edge. In the absence of optical distortion and noise, intensities on a line normal to an edge might appear as in Figure 2.1. The determination of the existence of an

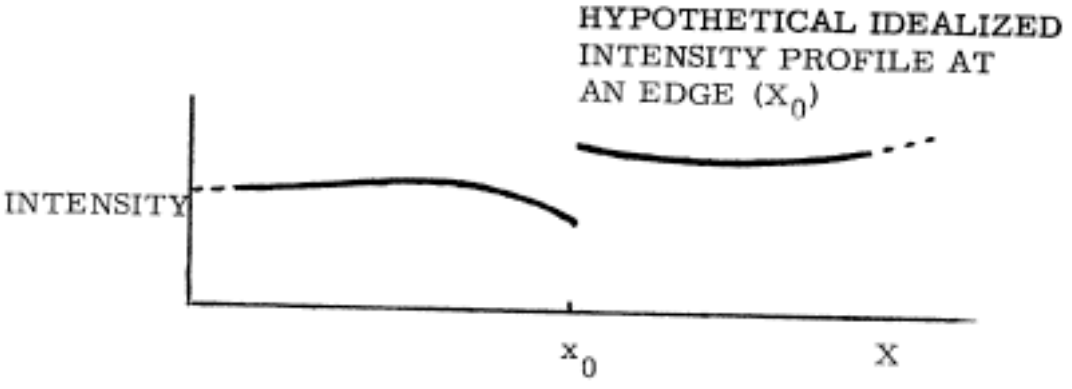


Figure 2.1

intensity discontinuity at x_0 is dependent upon finding the right and left limits of the intensity function at the point x_0. The intensities

along the normal away from x_0 contain information about the left and right limits at x_0. However the amount of information contained in these intensities falls off as one goes farther from the edge. This is due to possible unpredictible non-uniformities in the illumination and reflectivity of the objects in the scene.

The second effect has to do with the nature of the blurring inherent in an optical system. An edge subject to such blurring appears as in Figure 2.2. An abrupt step function is smeared out

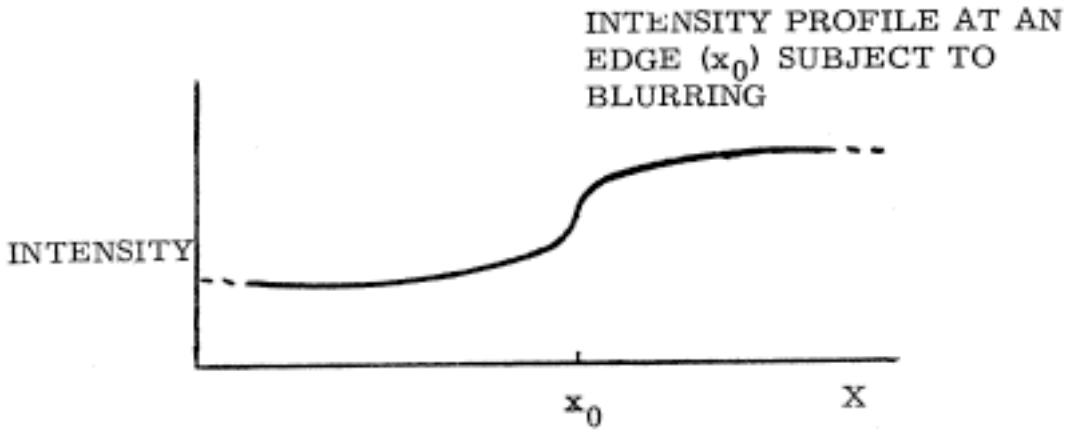


Figure 2.2

a small distance from the edge and causes extreme values of the various derivatives of the intensity function in a small neighborhood of the edge. It is the extreme values of the derivatives that characterize and identify an edge. However, the effect of the existence of an edge on the derivatives falls off drastically as one goes away from the edge. Hence, derivative information on a set of intensities remote from a possible edge can give little

information about the presence or absence of the edge.

The spatial extent of this effect is determined by the critical resolution radius defined as follows.

Definition: The critical resolution radius is the radius of a circle whose intensity equals (I+J)/2 around the image of a point of intensity I on a background of intensity J.

This radius is related to such factors as the resolving power of the optics of the system. Empirical evidence suggests that for relatively small values of ϵ, the value of $\delta_F(\epsilon)$ is on the order of magnitude of a few times this radius, in the hypothetical case mentioned of randomly distributed parallelopiped-edge-like lines.

In situations where the existence of a line is indeed virtually determinable entirely from intensity information contained within a few critical radii of the proposed line, we can test for the presence of a line in a given location by examining an area which could only possibly contain a few other lines resolvable from the given line. In general there would seldom be any line but the central line in an examined region. This suggests the development of a theory of the optimal use of intensity information within a narrow rectangular region, based on a model which assumes the

possible existence of at most one line in the region.

II. A REGIONAL VERSION OF THE GENERAL THEORY

A regional version of the theory presented in the previous chapter will now be discussed. The theorems presented give rise to a procedure which is optimal in deciding on the presence or absence of lines within a narrow rectangular region.

For the theory to be discussed, we adapt the following from chapter 1:

Let $\{I_i(x, y)\}$ be a set of real valued functions defined over a narrow rectangle R.

Let $D = \{0, 1\}$ be a two element description space.

Let S, a sampling set, be a finite subset of R.

Let $F(I_i(R))$ be an identification function mapping onto D, with the inverse image of the element $\underline{1}$ consisting of exactly the intensity functions given rise to by lines down the center of R.

Let $G(I_i(S))$ be a noise free sample identification function such that it and the function $F(I_i(R))$ conform to the descriptive completeness criterion. In this case, descriptive completeness requires that any line identified on R by $F(I_i(R))$ be identifiable by $G(I_i(S))$ from intensities on S.

We shall make the following assumptions as well:

1) That the intensity gathering device is subject to fundamental distortion, as defined in the previous chapter. In particular, we assume that this distortion governs the relationship between noisy samples $J_i(S)$, and corresponding noise free samples $I_i(S)$.

2) That at most one line can occur in the set S.

Some of these definitions and assumptions will not be explicitly used until the next chapter. In particular, the theorem about to be proved makes rather little explicit use of them.

The optimization of the system in question will, as before, be a matter of making optimal use of a series of a priori and conditional probabilities in the minimization of certain error rates. In particular, we will assume that the a priori probabilities $P(I_i(S)$ of each noise-free sample $I_i(S)$, together with the conditional probabilities $P(J_j(S)/I_i(S))$ are known; or alternatively that the joint probabilities $P(J_j(S), I_i(S))$, the product of the previous two, are known. Again we may wish to hold the true negative error rate to some value and minimize the false positive error rate; hold the false positive error rate to a certain value and minimize the true negative error rate; or minimize the sum of the two.

A surprising result emerges from the theorem about to be discussed: that these three types of optimization are brought about by essentially the same procedure.

We need the following definition before stating the main theorem of this section.

Definition:

$$Q(J(S)) = \sum_{G(I_i(S))=1} P(I_i(S)/J(S)) \qquad (2.1)$$

This quantity is the probability, given the noisy sample J(S), that there is a line down the center of the region R.

Theorem: Suppose it is desired to keep the false positive error rate below a certain value α, and to minimize the true negative error rate. Then there exists a β such that the following decision function H(J(S)) satisfies the conditions:

$$H(J(S))=1 \Longleftrightarrow Q(J(S)) \geq \beta \qquad (2.2)$$

Where β is the minimum value such that:

$$\sum_{Q(J_i(S))\geq\beta} (1-Q(J_i(S))P(J_i(S)) \leq \alpha. \qquad (2.3)$$

Proof: 1) Suppose a noisy pattern J(S) occurs, and that H(J(S))=1, i. e. a line is claimed to exist. Then (1-Q(J(S))) is the relative probability of a false

positive error whenever this pattern J(S) occurs, and (1-Q(J(S))P(J(S)) is the absolute probability of a false positive error involving the pattern J(S)). Thus:

$$\sum_{H(J_i(S))=1} (1-Q(J_i(S))P(J_i(S))$$

is the over-all false positive error rate of the function H(J(S)).

2) It is clear that the value $\underline{1}$ in place of β in equation (2.3) satisfies it, and that such values are bounded below by 0. Hence there exists a least value satisfying the equation, which we shall call β. Also, in the continuous case, for this value of the equality of equation (2.3) will hold. One may consider this to be the case as well in the discrete case, if the number of discrete values is large.

3) It remains to show that this procedure gives a minimum number of true negative errors. This error rate is given by:

$$\sum_{H(J_i(S))=0} Q(J_i(S))P(J_i(S)) \qquad (2.4)$$

Suppose we consider another decision function H*(J(S)) satisfying the condition of having false positive

error rate equal to or less than α, i.e.

$$\sum_{H^*(J_i(S))=1} (1-Q(J_i(S))P(J_i(S)) \leq \alpha \qquad (2.5)$$

We wish to show that its true negative error rate is not less than that of H(J(S)), given by expression (2.4), i.e. that

$$\sum_{H^*(J_i(S))=0} Q(J_i(S))P(J_i(S)) \geq \sum_{H(J_i(S))=0} Q(J_i(S))P(J_i(S)) \qquad (2.6)$$

4) Let R be the set of i's s.t. $H(J_i(S))=1$, and $H^*(J_i(S))=0$,

and let T be the set of i's s.t. $H(J_i(S))=0$, and $H^*(J_i(S))=1$,

and let U be the set of i's s.t. $H(J_i(S))=0$, and $H^*(J_i(S))=0$,

and let V be the set of i's s.t. $H(J_i(S))=1$, and $H^*(J_i(S))=1$.

Then

$$\sum_{H^*(J_i(S))=0} Q(J_i(S))P(J_i(S)) =$$

$$\sum_{U} Q(J_i(S))P(J_i(S)) + \sum_{R} Q(J_i(S))P(J_i(S)),$$

and

$$\sum_{H(J_i(S))=0} Q(J_i(S))P(J_i(S)) =$$

$$\sum_{U} Q(J_i(S))P(J_i(S)) + \sum_{T} Q(J_i(S))P(J_i(S)),$$

So to prove (2.6), we need to show:

$$\sum_{R} Q(J_i(S))P(J_i(S)) \geq \sum_{T} Q(J_i(S))P(J_i(S)) \qquad (2.7)$$

If $i \epsilon R$, then $Q(J_i(S)) \geq \beta$, and if $i \epsilon T$, $Q(J_i(S)) < \beta$, so

$$\sum_{R} Q(J_i(S))P(J_i(S)) \geq \beta \sum_{R} P(J_i(S)) \qquad (2.8)$$

and

$$\sum_{T} Q(J_i(S))P(J_i(S)) \leq \beta \sum_{T} P(J_i(S)). \qquad (2.9)$$

Assuming equality in expression (2.3), we have

from (2.3), (2.5) and the definition of R, T, and V:

$$\sum_{R \cup V} (1 - Q(J_i(S))P(J_i(S)) = \alpha \geq \sum_{T \cup V} (1 - Q(J_i(S))P(J_i(S)),$$

or, expanding and eliminating the sum on V:

$$\sum_{R} P(J_i(S)) - \sum_{R} Q(J_i(S))P(J_i(S)) \geq$$

$$\sum_{T} P(J_i(S)) - \sum_{T} Q(J_i(S))P(J_i(S)). \qquad (2.10)$$

Combining (2.8), (2.9) and (2.10):

$$1/\beta \sum_{R} Q(J_i(S))P(J_i(S)) - \sum_{R} Q(J_i(S))P(J_i(S)) \geq$$

$$\sum_{R} P(J_i(S)) - \sum_{R} Q(J_i(S))P(J_i(S)) \geq$$

$$\sum_{T} P(J_i(S)) - \sum_{T} Q(J_i(S))P(J_i(S)) \geq$$

$$1/\beta \sum_{T} Q(J_i(S))P(J_i(S)) - \sum_{T} Q(J_i(S))P(J_i(S)),$$

or:

$$(1/\beta - 1) \sum_{R} Q(J_i(S))P(J_i(S)) \geq (1/\beta - 1) \sum_{T} Q(J_i(S))P(J_i(S)).$$

Formula (2.7) follows from the fact that $(1/\beta - 1)$ is positive. QED

The case of minimizing the false positive error rate with a fixed true negative error rate is almost identical:

Theorem: Suppose it is desired to keep the true negative error rate below a certain value α, and to minimize the false positive error rate. Then there exists a β such that the following decision function satisfies the conditions:

$H(J(S))=1 \Leftrightarrow Q(J(S)) \geq \beta$.

Proof: As for the previous theorem.

It follows from these two theorems that whatever the value

of β used, the procedure of thresholding $Q(J_i(S))$ is better than all other procedures with the same false positive error rate, and better than all other procedures with the same true negative error rate. In this respect it is similar to a saddle point solution of classical 2 by 2 game theory.

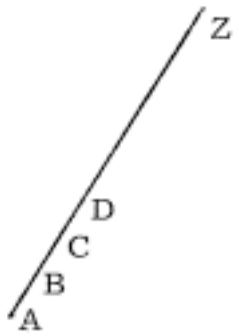
It is not difficult to see that this procedure can, for a suitable choice of β, be used to minimize the sum of the two error rates. Consider an arbitrary choice procedure H* which gives a false positive error rate $\underline{a}$ and true negative error rate $\underline{b}$. Then there exists a β such that thresholding $Q(J(S))$ at β gives false positive error rate $\underline{a}$ and true negative error rate b^*, where $b^* \leq b$. This will provide a decision procedure whose total error rate is $a + b^* \leq a + b$. Hence thresholding of $Q(J(S))$ with this β gives a minimal sum of error rates.

III. THE RELATIONSHIP BETWEEN GLOBAL AND REGIONAL LINE PREDICATES

This section discusses the relation between the global line finding procedure described in chapter one, and the regional procedure of the present chapter. It is clear that the procedure of chapter one is computationally unfeasible.

However a systematic application of the regional predicate to a large number of regions on the field of view may be computationally practical. The latter procedure, in certain cases of interest, can not achieve the optimality of the global procedure. However it does trade a certain amount of accuracy for computational feasibility.

It has been pointed out that in certain cases where lines on the visual field bear no systematic relationship to each other, a regional line predicate of the sort described in the definition at the beginning of this chapter seems to be potentially as accurate as a global one. It is tempting to formalize this notion into some sort of theorem, stating conditions under which a systematic application of a regional predicate of this sort to a suitably large number of regions on the field is equivalent to a global procedure. Unfortunately, an exact correspondence between these two procedures is difficult to make. One problem is illustrated in Figure 2.3. It is reasonable to assume that in the system described in chapter one, if an element of the description space contains a segment AB, it would not also contain a subsegment of that line. More informally stated, the scene descriptions would not contain redundant segments.



A REGIONAL LINE PREDICATE APPLIED TO MANY REGIONS ON THE FIELD WOULD REPORT LINES AB, AC, . . . , AZ, BC, . . . , BZ, ETC. A GLOBAL PROCEDURE WOULD OMIT THE SUB-SEGMENTS AND REPORT ONLY AZ.

Figure 2.3

On the other hand, the systematic application of a regional predicate of the sort mentioned would not only give positive identification for some complete line in the scene, but for all examined subsegments as well. Another difficulty is that the regional predicate might report the existence of two nearly coincident lines at very slightly different inclinations, where there is actually only one line. These difficulties will be encountered in chapter four, which describes an actual computer program embodying a regional line predicate. In the program, the problems are handled at the heuristic level. However, from a formal point of view, obvious efforts to solve them appear artificial. We will thus confine the remainder of this discussion to the informal level.

It is convenient to informally divide the evidence relating to the existence of a particular line into three categories:

1) Global evidence, namely evidence from areas spatially remote from a possible line;

2) Regional evidence, namely evidence from an area within a small radius of a possible line;

3) Local evidence, namely evidence contained within a neighborhood of some point on a possible line.

It has been pointed out that there are certain problems associated with a formal identification of a systematic use of a regional predicate with a global predicate. This identification appeared plausible only in cases where there is no relationship among the lines in the figures. However, the situation of particular interest is that in which the scenes consist of sets of plane-faced solids. In this case there is definitely a relationship among the lines in the figures; and a regional predicate cannot possibly achieve the optimality of a global one. Figures 2.4 and 2.5 illustrate the possible failures of a regional line predicate, which might be avoided by the use of global evidence as well as regional evidence.

Figure 2.4 illustrates three lines radiating from a common

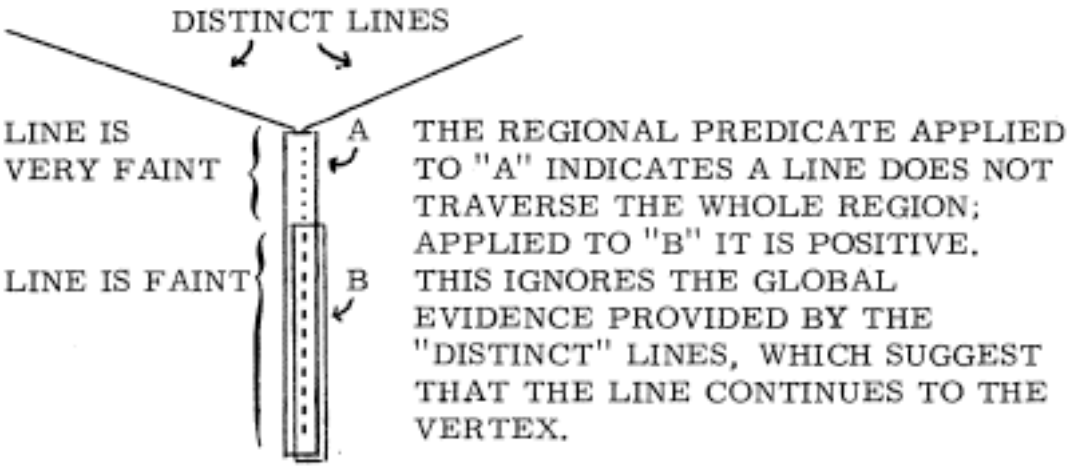
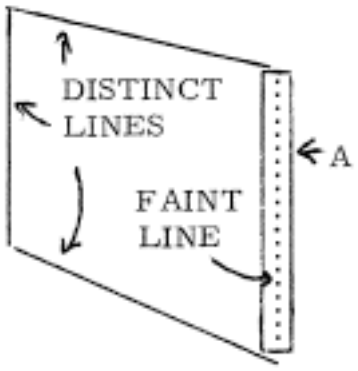


Figure 2.4

vertex. The upper two lines are quite distinct, whereas the lower line is faint, especially at the top. A regional predicate applied to region A may claim that a line does not traverse the entire region. However, when applied to B, there is sufficient evidence to claim the existence of a line. The fact that the line in region B "points to" a close vertex is ignored by the regional predicate. Global evidence indicates that the line continues to the vertex. The latter would likely be the conclusion of a global procedure, as discussed in chapter one, which would assign low plausibility (a priori probability) to scenes with isolated line ends, and high

plausibility to closed line figures. Figure 2.5 illustrates a



THE REGIONAL PREDICATE APPLIED TO "A" MIGHT CLAIM THAT THERE IS NO LINE. HOWEVER, GLOBAL EVIDENCE STRONGLY SUGGESTS THERE IS.

Figure 2.5

situation similar to that in Figure 2.4. In this case a line is totally ignored by the regional predicate, despite considerable global evidence for its existence.

In chapter four, a program will be described which uses a regional predicate in the analysis of scenes consisting of prismatic solids. It is convenient to here describe briefly two features of this program, as the current context justifies their use. Specifically, the two features are designed to overcome the non-globality of the regional predicate as illustrated by the previous two examples. The sort of problem illustrated in Figure 2.4 is overcome by the use of what will be termed links, which tentatively extend a line to a possible vertex. The sort of problem illustrated in Figure 2.5 is handled by the procedure of

proposing lines. The procedure in effect assigns a high a priori probability to lines suggested by lines already identified. This tends to augment regional evidence for the proposed line.

A final fact should be pointed out in this context concerning the relationship between the use of purely local evidence and the use of a regional predicate. This is analogous to the relationship between the systematic use of a regional line finder and the use of a global one. In the latter case, it is necessary that the lines in the scenes be spatially independent in order that the two procedures be equally effective. In order for purely local decisions as to the existence of a line to be as effective as regional decisions, it would, by analogy, be necessary that the points along a particular line be spatially independent. This is, of course, totally contrary to the nature of a line. It is thus quite impossible that a systematic application of a simple local decision procedure to a suitably large number of neighborhoods covering a region could be as sensitive in detecting lines as a regional predicate. By a "local decision procedure" is meant a procedure which decides entirely on the basis of intensity information from within a neighborhood of a point, whether or not a line traverses the neighborhood. The output of such a procedure would be a set of

points on the visual field, which would be incorporated into lines by finding maximal sets of adjacent colinear points.

This two stage procedure is a special case of a set of procedures which process intensities in a neighborhood, and then process the results for a set of neighborhoods contained within a region. Although the particular two stage procedure mentioned in the previous paragraph can never be as sensitive as an optimal regional line finder, there exist two stage procedures of this sort which are. This is not surprising, since all regional predicates fall into this class, if one regards the first stage to be the identity function, and the second stage to be the regional predicate itself. There also exist non-trivial examples of members of this class which are as sensitive as an optimal regional procedure. The procedure described in the next two chapters involves such a two stage predicate. For the first stage, four values are obtained from the intensities in a neighborhood. Sets of these four-tuples, obtained from all the neighborhoods within a region, are processed by the second stage; and a decision is made as to the existence of a line in the region. The procedure may be regarded intuitively as extracting all necessary local evidence in the first stage, and employing all necessary regional evidence in the second stage.

CHAPTER 3

COMPUTATIONAL APPROACHES TO THE REGIONAL THEORY

I. INTRODUCTION

The object of this chapter is to give a computationally feasible method for computing values of the function given by Formula 2.1 of the last chapter:

$$Q(J(S)) = \sum_{G(I_i(S))=1} P(I_i(S)/J(S)).$$

It will be recalled that this is a function of the finite set J(S) of intensities obtained by a particular sort of optical input device from within some region on the field of view. The thresholding of this value provides an optimal regional line detection procedure.

From Bayes' rule it follows that

$$P(I_i(S)/J(S)) = \frac{P(J(S), I_i(S))}{P(J(S))}$$

so that

$$Q(J(S)) = \frac{\sum_{G(I_i(S))=1} P(J(S), I_i(S))}{P(J(S))}$$

Using the relation

$$P(J(S)) = \sum_{G(I_i(S))=1} P(J(S), I_i(S)) + \sum_{G(I_i(S))=0} P(J(S), I_i(S)).$$

and

$$P(J(S), I_i(S)) = P(J(S)/I_i(S))P(I_i(S)),$$

we have:

$$Q(J(S)) = \frac{\sum_{G(I_i(S))=1} P(J(S))/I_i(S))P(I_i(S))}{\sum_{G(I_i(S))=1} P(J(S)/I_i(S))P(I_i(S)) + \sum_{G(I_i(S))=0} P(J(S), I_i(S))P(I_i(S))}$$

Sections II and III of this chapter will be concerned with the calculation of the values of $P(J(S)/I_i(S))$. Sections IV and V propose a model for the relative probabilities of the various noise-free patterns $I_i(S)$. On the basis of the values of $P(I_i(S))$ given by the model, it is possible to express the sums in the above expression as integrals, and to give a relatively simple explicit formula for $Q(J(S))$. In the last section of this chapter, this formula is applied to some simple cases of image-dissector output.

II. NOISE-FREE PATTERNS AND BLURRING DISTORTION

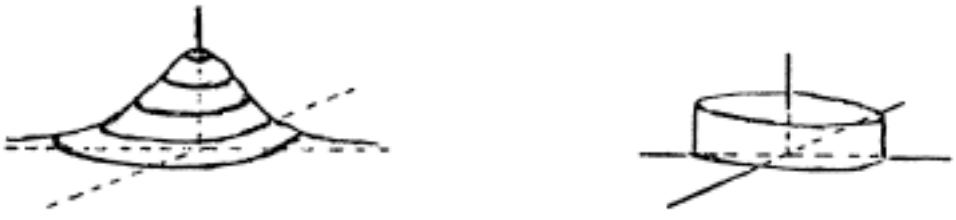
The value of the conditional probability $P(J_i/I_j)$ is a function of the relation between the two sets of intensities J_i and I_j. This relationship can be decomposed into:

1) a determinate relation between the noise-free sample I_j and a distorted but noise-free version I^*_j; and

2) a simple probabilistic relation between this I^*_j and the noisy pattern J_i.

The former will be discussed in this section; the latter in the next section.

According to the definition of fundamental distortion, the relation between a noise-free pattern I_j and a distorted version I^*_j may be expressed in terms of a convolution of the former with some function f(x, y). Possible examples of f(x, y) are given in Figure 3.1. Theoretically, from the way an



HILL FUNCTION PILLBOX FUNCTION

Figure 3.1

image-dissector is designed, this function should be cylindrically shaped as in the right hand illustration of Figure 3.1. In fact, it is more like a gaussian "hill" as shown in the left hand illustration. An intensity function I(x, y) is acted upon by this convolution function f(x, y) to produce a blurred intensity function I*(x, y) as follows:

$$I^*(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u, v)I(x-u, y-v)du\ dv.$$

If we had an edge which is reasonably uniform for a relatively long distance compared with the "width" of the convolution function, we can describe the convolution effect, at least locally, by means of a single integral instead of a double integral. In particular, suppose for some local region the intensity function I(x, y) represents a line parallel to the y-axis. Then locally we may write I(x, y) = I(x) and the corresponding blurred but noise-free intensity function I*(x, y) may be expressed in this region as:

$$I^*(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u, v)I(x-u)du\ dv$$

$$= \int_{-\infty}^{\infty} I(x-u)\left[\int_{-\infty}^{\infty} f(u, v)dv\right]du.$$

Since the inner integral is a function only of u, we may indicate it with a single function name:

$$g(u) = \int_{-\infty}^{\infty} f(u, v)dv,$$

and we have at a reasonably extensive and uniform edge parallel to the y-axis:

$$I^*(x, y) = I^*(x) = \int_{-\infty}^{\infty} I(x-u)g(u)du. \qquad (3.1)$$

We may obtain g(u) empirically by scanning at right angles across a very narrow line in the visual field. Let the intensity along a line normal to a vertical line in the visual field be given by:

$I(x) = b$ if $x < -a,\ x > a$

$I(x) = c$ $\quad -a \leq x \leq a.$

Then the distorted version of the intensity locally around the line is given by:

$$I^*(x) = \int_{-\infty}^{\infty} I(x-u)g(u)du = c\int_{-x-a}^{-x+a} g(u)du + b.$$

$$\text{But } \lim_{a \to 0}\left[c \int_{-x-a}^{-x+a} g(u)du \right] = c \cdot a \cdot g(-x), \quad \text{so}$$

$$I^*(x) \equiv c \cdot a \cdot g(-x) + d,$$

for very narrow lines.

Figure 3.2 shows plots of intensity taken along normals

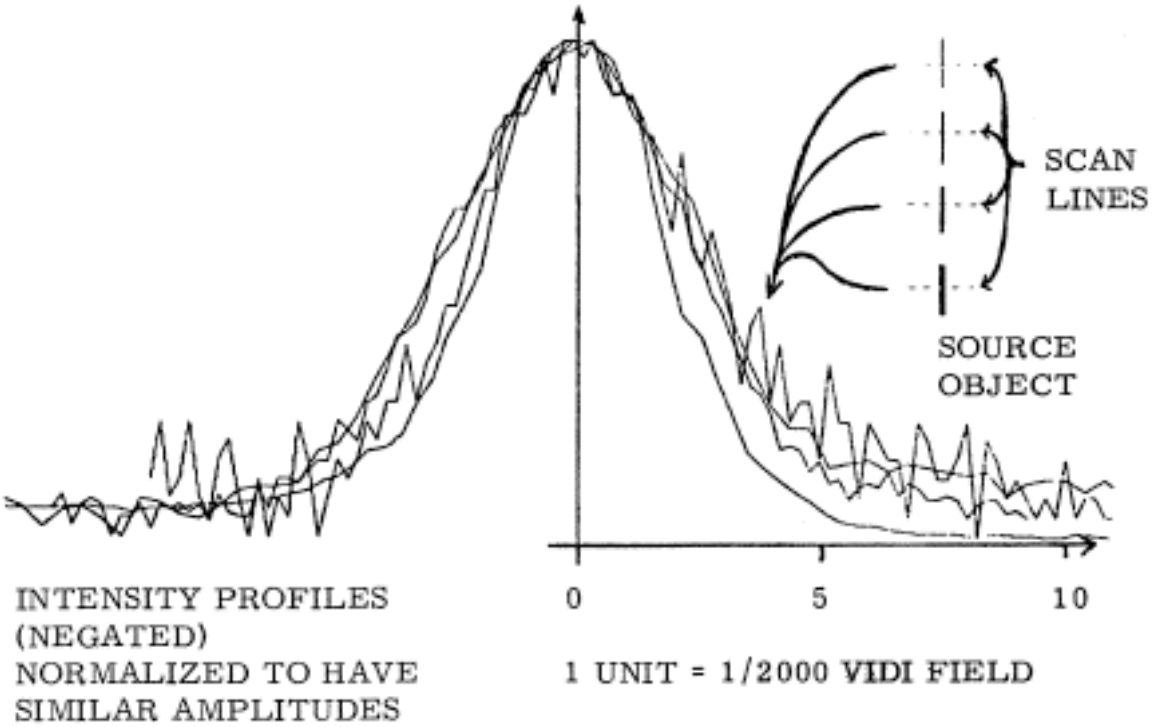
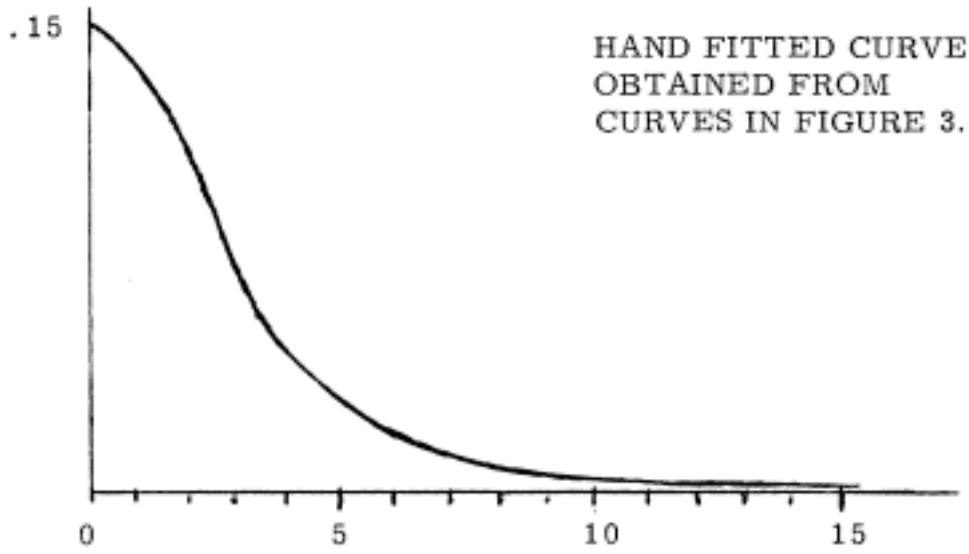


Figure 3.2

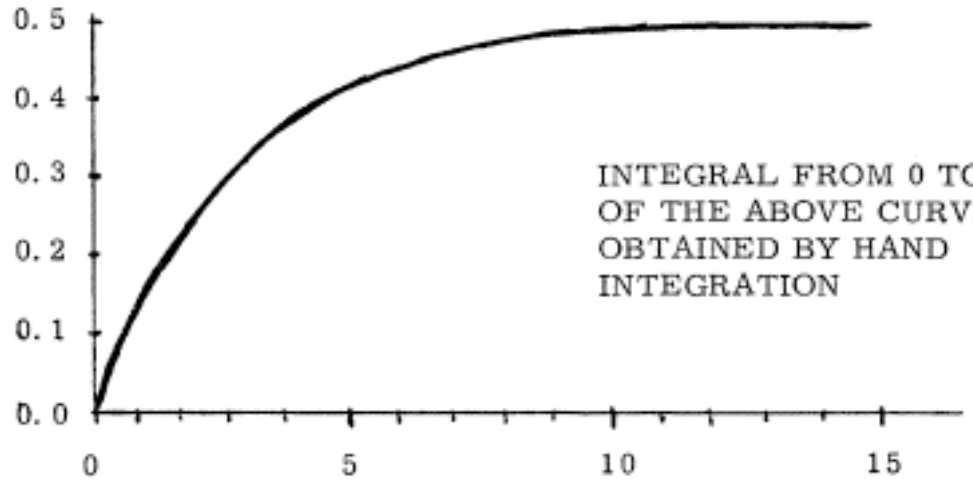
to various vertical lines. According to the preceding paragraph, these curves should be of the form $a \cdot g(x) + b$.

They are shown superimposed and normalized to have similar amplitudes. The lines are all quite narrow relative to the width of the hill plotted, and represent lines of a wide variety of intensities. The values are all proportional to absolute intensity, having been suitably transformed from vidissector output which is logarithmic. The degree of variation of un-normalized amplitude was greater than ten to one; and may be observed from the relative noisiness of the curves. All were recorded at a fixed noise level, and the amount of local fluctuation of a given line in the illustration is a measure of the degree to which the hill in question was enlarged in the normalization process. It should be noted that the shape of the curves is remarkably consistent. This indicates a good agreement between the theory just discussed and the actual behavior of the blurring process in the vidissector. Figure 3.3 illustrates a curve obtained from the data in Figure 3.2 by hand fitting. Again, this curve should be of the form $a \cdot g(x) + b$, where $g(x)$ is the one-dimensional blurring function. Also illustrated is a curve of the integral of this function, obtained by hand integration. This latter curve will be compared with a similar curve obtained by two different methods



HAND FITTED CURVE
OBTAINED FROM
CURVES IN FIGURE 3.2

1 UNIT = 1/2000 VIDISSECTOR FIELD



INTEGRAL FROM 0 TO X
OF THE ABOVE CURVE,
OBTAINED BY HAND
INTEGRATION

1 UNIT = 1/2000 VIDISSECTOR FIELD

APPROXIMATIONS TO g(x) AND
ITS INTEGRAL MAY BE OBTAINED
EMPIRICALLY FROM THE DATA
IN FIGURE 3.2

Figure 3.3

which will now be described.

Another approach to finding g(x) is based on the fact that the peak value of an intensity scan normal to a line of non-negligible width $\underline{a}$ is related in a simple way to the integral

$$\int_{-a/2}^{a/2} g(x)\ dx,$$

namely, if the actual intensity along a normal to a narrow vertical line is given by:

$$I(x) = c \quad \text{if} \quad x < b, \quad x > b + a$$

$$I(x) = c + d \quad b \leq x \leq b + a,$$

then I*(x), the distorted intensity function is given by:

$$I^*(x) = \int_{-\infty}^{\infty} I(x-u)g(u)du = \int_{-\infty}^{\infty} (c + I_1(x-u))g(u)du,$$

where $I_1(x) = 0$ except in the interval (b, b+a), where $I_1(x) = d$. One may write the latter integral as:

$$c\int_{-\infty}^{\infty} g(u)du + \int_{-\infty}^{\infty} I_1(x-u)g(u)du,$$

But the former integral is a constant, and the latter integral

is equal to:

$$d\int_{b-x}^{b+a-x} g(u)du,$$

and has an extremum at $x = 1/2(a + b)$, which is a minimum or a maximum depending on the sign of d. At this extremum, the integral has the value:

$$d\int_{-a/2}^{a/2} g(u)du.$$

Thus a scan made perpendicular to a line of width a and intensity c+d on a background of intensity c has a maximum (or minimum) whose value relative to the intensity somewhat away from the line is:

$$d\int_{-a/2}^{a/2} g(u)du.$$

Suppose a scan is taken across a pair of narrow triangles as in the left of Figure 3. 4. In the figure, kLn(1/intensity) is plotted for successive scans. In the right hand side of the figure, the values of the maxima of each scan minus the background value are plotted for each scan as

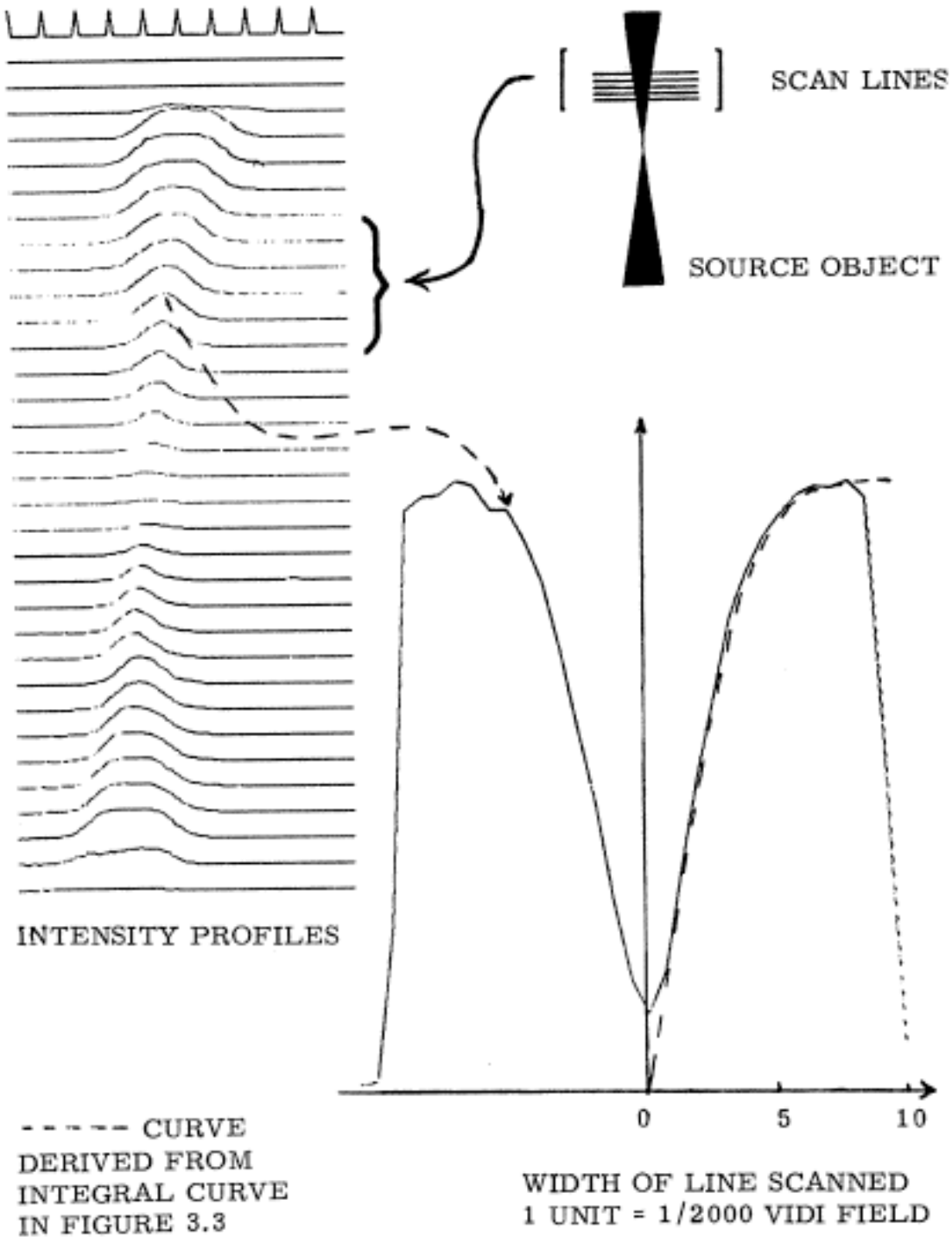

SCAN LINES
SOURCE OBJECT
INTENSITY PROFILES
0
5
10
- - - - - CURVE DERIVED FROM INTEGRAL CURVE IN FIGURE 3.3
WIDTH OF LINE SCANNED
1 UNIT = 1/2000 VIDI FIELD

Figure 3.4

a function of the spatial location of the scan along the set of triangles. Since the width of the triangle at a particular scan is a linear function of the position of the scan along the triangle, the plot just mentioned may be considered to be a function of the width of the line being scanned. Hence the x-axis in the plot is in units commensurate with this distance measure, and in the same scale as is used in Figures 3.2 and 3.3 relative to the physical dimensions of the vidissector. The values plotted are thus those of $kLn(1/m(x)) - kLn(1/c)$, where $m(x)$ is the actual minimum (since the target is dark) of intensity along a scan of a line of width $\underline{x}$, and $\underline{c}$ is the background intensity. If the blurring function behaves according to the theory proposed, we should have:

$$c - m(x) = d\int_{-x/2}^{x/2} g(u)du = 2d\int_{0}^{x/2} g(u)du \qquad (3.2)$$

where $\underline{d}$ is the absolute value of the intensity difference between the background and the target. Using the curve of the integral of the function in Figure 3.3, and doing a certain amount of arithmetic involving logarithms, we can obtain a curve of predicted maxima in accordance with the premise expressed

in (3.2), and with the assertion that the curves in Figure 3.3 represent g(x) and its integral. This curve is indicated in Figure 3.4 as the dashed line superimposed on the plotted curve in the figure at the right. The closeness of these two curves indicates good consistency between these two approaches to finding g(x); and provides further evidence that the curves in Figure 3.3 represent g(x) and its integral.

A third approach to the computation of g(x) is based on the fact that at a wide edge, the value of the intensity along a line normal to the edge is related in a simple way to the function:

$$h(y) = \int_{-\infty}^{y} g(x)dx.$$

In particular, if the intensity function for a vertical edge is given by:

$$I(x) = a \quad x \leq 0$$

$$I(x) = b \quad x > 0,$$

then the distorted value of intensity, I*(x), is given by:

$$I^*(x) = \int_{-\infty}^{\infty} I(x-u)g(u)du = a + b\int_{-x}^{\infty} g(u)du.$$

This equality is derived by an argument similar to that given above. In Figure 3.5 we have the values of kLn(1/intensity)

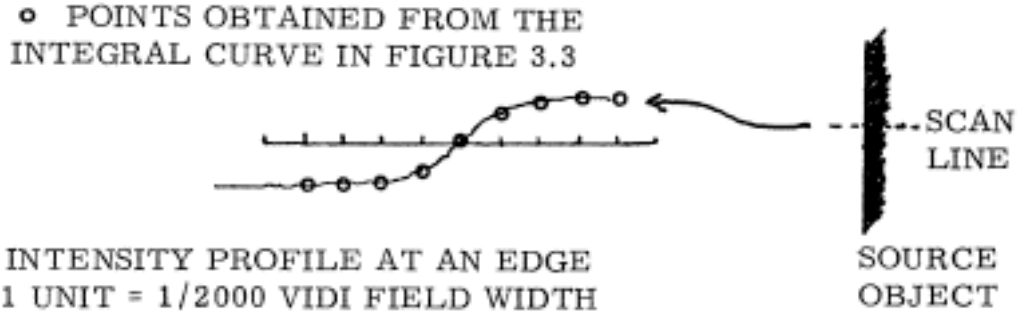


Figure 3.5

for the intensities along a line normal to an edge, subject to the blurring of the vidissector. The illustrated points along the curve represent values obtained from the integral curve in Figure 3.3, suitably scaled. For clarity, the curve represented by these points has been omitted. The close correspondence between the points and the curve further verifies that the curves of Figure 3.3 are a good approximation to g(x) and its integral.

III. NOISE AND THE COMPUTATION OF $P(J_i/I_j)$

A given noise-free sample I_i turns into a distorted sample I_i^* as shown in the previous section by the convolution of I_i with some blurring function. The conditional probability

$P(J_j/I_i)$ may easily be expressed as the probability that simple noise on the sample I_i^* would give rise to the configuration of intensities J_j.

It has been claimed that a large number of intensities read at a given point under a given lighting situation by the vidissector have approximately a normal distribution. According to the way in which the vidissector is constructed, the absolute noise level is a constant for the logarithmically scaled intensities delivered. Actually there are four possible noise levels; however all work in the remainder of this paper will involve only one, the most accurate. Hence we shall assume that there is a single constant noise level. The intensities read at a particular point for a particular set of lighting conditions follow the distribution:

$$\frac{1}{a\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x}{a}\right]^2} dx.$$

where $\underline{a}$ is a constant. From this may be computed the probability of a given configuration of intensities $v_1, v_2, \ldots, v_m$ at points with actual intensities $u_1, u_2, \ldots, u_m$. Since the noise-induced deviations are

statistically independent, this probability is:

$$\prod_{i=1}^{m} \frac{1}{a\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{u_i - v_i}{a}\right]^2} dv_1 dv_2 \ldots dv_m,$$

or, letting

$$D_{i,j} = \sum_k (u_k - v_k)^2, \tag{3.3}$$

with :

$$I_i^* = (u_1, u_2, \ldots, u_m)$$

$$J_j = (v_1, v_2, \ldots, v_m),$$

we have:

$$P(J_j/I_i) = P(J_j/I_i^*) = \frac{1}{(a\sqrt{2\pi})^m} e^{-\frac{1}{2}\left[\frac{D_{ij}}{a^2}\right]} dv_1 dv_2 \ldots dv_m. \tag{3.4}$$

The value of a for the signal to noise ratio has been reported (Horn 1969) in terms of the percentage error of a logarithmic intensity value corresponding to an error of one standard deviation of the errors at that intensity. Again, this percentage is a constant for all intensities, and is reported to have the value 1.6%. Since the logarithmic intensities are given by:

$$100_8 \text{Log}_2(1/I),$$

the absolute error corresponding to one standard deviation

is:

$$a = 64.\,\mathrm{Log}_2(I - .016I) - 64.\,\mathrm{Log}_2(I) = 1.5 \qquad (3.5)$$

This is the value of $\underline{a}$ which will be used in (3.4).

The procedure for computing the conditional probability $P(J_j(S)/I_i(S))$ is to determine the blurred version $I_i^*(S)$ of the set of intensities $I_i(S)$, and to calculate the value of $P(J_j(S)/I_i^*(S))$ by (3.4). The relationship between $I_i(S)$ and $I_i^*(S)$ is in general governed by the function $f(u, v)$ referred to in the previous section. However we shall consider only cases where (3.1) applies, and will need only the function $g(x)$ to determine this relationship.

IV. A MODEL FOR THE SET $\{I_i^*\}$ AND COMPUTATIONAL CONSEQUENCES

In this section we shall give models for the sets $\{I_i^*(S)\}$ of noise-free distorted intensity configurations taken over particular rectangular regions S. Our attention will be confined to sampling regions S which have a particular width, and whose lengths are integral multiples $\underline{n}$ of the width. We may consider such a region to be composed of $\underline{n}$ square

sub-components, denote it by S_n, and use the definition:

$$I^*_i(S_n) = I^{*n}_i.$$

Rules will be given for assigning a priori probability values $P(I^{*n}_i)$ to the elements of $\{I^{*n}_i\}$. It will be recalled that there is a determinate 1:1 correspondence between elements I^{*n}_i and elements $I_i(S_n)$. Thus the analysis in section I of this chapter remains valid with I^{*n}_i in place of $I_i(S)$. It follows that the value of $Q(J(S))$ of (2.1) is determined by (3.4) and by the values of $P(I^{*n}_i)$. The latter are the subject of this section.

Consider, for a fixed $\underline{n}$, the set $\{I^{*n}_i\}$ of all noise-free distorted samples over rectangular regions composed of $\underline{n}$ square sub-components. We shall divide this set into four subsets as follows:

Let CL be the subset of $\{I^{*n}_i\}$ consisting of sets of intensities from lines in the visual field exactly centered in the rectangular region.

Let CE be the subset of $\{I^{*n}_i\}$ consisting of sets of intensities from edges in the visual field exactly centered in the rectangular region.

Let CH be the subset of $\{I^{*n}_i\}$ consisting of sets of

intensities from a homogeneous area in the rectangular region.

Let CS be the subset of $\{I^{*n}_i\}$ consisting of sets of intensities from lines or edges which are not centered in the rectangular region.

We shall first consider a model of CL, and base it on the following assumptions:

1) That the absolute intensities of the samples of CL are of no consequence, so that in each square sub-component of a sample I^{*n}_i the sum of the intensities may be assumed to be 0.

2) That on any section across some j-th sub-component of a member I^{*n}_i of CL, the intensity function has the form:

$a_{i,j}g(x) + b_{i,j}$

where $a_{i,j}$ will be termed the relative amplitude, and $b_{i,j}$ is chosen to conform with assumption 1).

3) That the relative amplitude $a_{i,j}$ of an intensity peak (or valley) is uniform within the j-th sub-component of I^{*n}_i. A sample I^{*n}_i may thus be described in terms of

an n-tuple of relative amplitudes ($a_{i,1}, \ldots, a_{i,n}$), e.g. as $I^*(a_{i,1}, \ldots, a_{i,n})$.

4) That each sample $I^*_i{}^n$ was given rise to from some line, highlight, etc., in the real world with an "idealized" relative peak (or valley) amplitude a_i, which is uniform along the length of the line.

5) That this idealized relative amplitude is purturbed randomly along the length of the line so that there is a minor fluctuation of relative amplitude $a_{i,j}$ among the various sub-components of a sample $I^*_i{}^n$, with a normal distribution with variance σ^2_n.

The first of these assumptions ignores a certain amount of information about a line which is useful in distinguishing it from a non-line. Broadly speaking, it prevents a distinction between a series of peaks in a line whose absolute intensity is constant or regularly varying; and a series of peaks in a line whose absolute intensities vary randomly or discontinuously. This is not undesirable in the case of a line lying in the real world partly in shadow. However, it would be desirable to eliminate as lines a series of "lined up" intensity peaks whose absolute intensity varied randomly, as this would likely be due

to noise.

The second assumption is justified on the basis of the discussion in the first section, and by the fact that we are considering noise-free but distorted samples.

The third assumption is justified on the basis that the "ridges" of intensity of samples in CL are rather uniform along their whole lengths, and would thus be quite uniform on square sub-components. This assumption is particularly valid if a square sub-component is only sampled along a line traversing it perpendicular to the orientation of the rectangle. This is approximately what is done in the program which will be discussed in the next chapter.

The "idealized" constant relative amplitude of assumption 4) may be thought of as the relative amplitude which would be obtained in the absence of physical defects of the object being viewed or on the vidissector photocathode, and in the absence of spurious local reflections and anomalies of lighting. The purturbations of assumption 5) may be thought of as a result of these defects etc., e. g. a nick on the edge of a cube, a slight shadow or a bad spot on the photocathode of the vidissector. The idea of an idealized constant value of

the relative amplitude along an edge is further justified by the fact that intensities used will be logarithmic. Any ridge of intensity, for example a highlight, is a relative contrast phenomenon, and may vary in absolute intensity, but not in relative intensity. A logarithmic intensity function would tend to register amplitude of highlight the same regardless of the absolute intensity. The same is true when the argument is applied to intensity discontinuities at an edge.

It follows from these assumptions that if we let F(a) be the a priori probability of an edge in the real world with "idealized" intensity $\underline{a}$, then the probability of existence of a particular $I^*_i{}^n = I^*(a_{i,1}, \ldots, a_{i,n})$ given rise to by an edge with idealized relative amplitude $\underline{a}$ and parameterized by amplitudes $a_{i,1}, \ldots, a_{i,n}$ is given by:

$$F(a)\prod_{j=1}^{n} \frac{1}{\sigma_n\sqrt{2\pi}}\, e^{-\frac{1}{2}\left[\frac{a-a_{ij}}{\sigma_n}\right]^2} da_{i,1} \cdots da_{i,n}.$$

The probability that this $I^*_i{}^n$ would occur from a source with any idealized amplitude is given by the sum over $\underline{a}$ of the above expression. This sum is simply the a priori probability $P(I^*_i{}^n) = P(I^*(a_{i,1}, \ldots, a_{i,n}))$.

This sum may then be evaluated provided the distribution

F(a) is known. Assuming it is normal with standard deviation ρ_n, we have:

$$F(a) = \frac{P(CL)}{\rho_n\sqrt{2\pi}}\, e^{-\frac{1}{2}\left[\frac{a}{\rho_n}\right]^2} da,$$

where P(CL) is the a priori probability of existence of a member of CL. We then have:

$$P(I^*(a_{i,1}, \ldots, a_{i,n})) =$$

$$\frac{P(CL)}{\sqrt{2\pi}^{n+1}\sigma_n^n\rho_n}\sum_a e^{-\frac{1}{2}\left[\frac{a}{\rho_n}\right]^2-\frac{1}{2}\left[\sum\frac{(a-a_{ij})^2}{\sigma_n^2}\right]} da\, da_{i,1} \cdots da_{i,n} \quad (3.6)$$

This sum, including the differential da may be expressed as as integral, with the exponent slightly rewritten:

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{a^2}{\rho_n^2}+\frac{na^2 - 2a\sum a_{ij} + \sum a_{ij}^2}{\sigma_n^2}\right]} da.$$

The exponent may be further rewritten as:

$$-\frac{1}{2}\left[a^2\left[\frac{1}{\rho_n^2}+\frac{n}{\sigma_n^2}\right] - 2a\frac{\sum a_{ij}}{\sigma_n^2} + \frac{\sum a_{ij}^2}{\sigma_n^2}\right]$$

Letting:

$$A = \frac{1}{\sigma_n^2}+\frac{n}{\rho_n^2},$$

$$B = \frac{\sum a_{ij}}{\sigma_n^2}$$

and completeing the square, the exponent becomes, in

terms of A and B:

$$-\frac{1}{2}\left[A(a^2-2a(B/A)+(B/A)^2) - B^2/A + \sum_i a_{ij}^2/\sigma_n^2\right].$$

The integral becomes:

$$\int_{-\infty}^{\infty} e^{-A/2(a - B/A)^2 - 1/2(\sum_i a^2/\sigma_n^2 - B^2/A)} da.$$

The second term in the exponent is independent of the integration variable a, and may be taken outside the integral sign. The remainimg definite integral is well known to have the value $\sqrt{\pi/A}$. Thus the integral is equivalent to:

$$\sqrt{\frac{\pi}{A}}\, e^{-\frac{1}{2}\left[\frac{\sum_i a_{ij}^2}{\sigma_n^2} - \frac{B^2}{A}\right]}$$

The result of substituting in this expression the values of A and B, of substituting the resulting expression into (3.6), and of doing some straightforward algebraic mainpulation is:

$$P(I^*(a_{i,1}, \cdots a_{i,n}) =$$

$$\frac{P(CL)}{\sqrt{2\pi}\sigma_n^n}\frac{\sigma_n}{\sqrt{\sigma_n^2+n\rho_n^2}}\, e^{-\frac{1}{2}\left[\frac{\sum_i a_{ij}^2}{\sigma_n^2+n\rho_n^2} + \frac{n\rho_n^2\sum_i(a_{ij}-\bar{a}_{ij})^2}{\sigma^2(\sigma_n^2+n\rho_n^2)}\right]} da_{i,1}\cdots da_{i,n} \quad (3.7)$$

An analysis of the CE category may be made in a nearly identical fashion. In this case, the function g(x) of assumption 2) is replaced with the function h(x), its integral. As was pointed out in Section II, this function gives the intensity along a normal to an edge, and is the obvious correspond to g(x), which gives the intensity along a normal to a line. A formula for the a priori probability of a member of class CL may be derived from an analysis nearly identical to that given above. The resulting formula differs from (3. 7) only in that different values appear for the variances σ_n and ρ_n; and that P(CE) is used in place of P(CL).

Again, an analysis of the CH category may be made along nearly identical lines. In this case, the function g(x) is replaced with l(x) = ax + b, the function which gives the intensity profile across a homogeneous region. Again the resulting formula for a priori probabilities of members of CH differs from (3. 7) only in the values of σ_n and ρ_n; and in the use of P(CH) in place of P(CL).

At this point it is convenient to recapitulate some of the foregoing by rewriting the formula for Q(J(S)) appearing at the end of Section I in terms of the specific notation introduced in

later sections, and some additional new notation. First, it will be recalled that the notation $I^*(a_{i,1}, \ldots, a_{i,n})$ was introduced to refer to noise-free distorted samples over a region consisting of $\underline{n}$ square sub-components, and containing a line centered in the region. A value $a_{i,j}$ refers to the relative amplitude of the intensity profile of the i-th sample, taken across the j-th sub-component. We may similarly adopt the notation $I^*(b_{i,1}, \ldots, b_{i,n})$ to refer to noise-free distorted samples from the class CE, and $I^*(c_{i,1}, \ldots, c_{i,n})$ to refer to samples from CH. We shall reserve the notation $I^*_i{}^n$ to refer only to distorted noise-free samples from the class CS. We may introduce similar notation for noisy samples by a definition similar to that used at the beginning of Section IV:

$$J_i^n = J_i(S_n).$$

By this notation is meant the i-th possible noisy sample on a region consisting of $\underline{n}$ square sub-components.

The sets CL, CE, CH and CS provide us with an alternative for expressing the function $G(I_i(S))$ used in the formula for Q(J(S)) in Section I. If we make the following

definitions:

$$CL(J_j^n) = \sum_i P(J_j^n/I^*(a_{i,1}, \ldots, a_{i,n}))P(I^*(a_{i,1}, \ldots, a_{i,n})), \quad (3.8)$$

$$CE(J_j^n) = \sum_i P(J_j^n/I^*(b_{i,1}, \ldots, b_{i,n}))P(I^*(b_{i,1}, \ldots, b_{i,n})), \quad (3.9)$$

$$CH(J_j^n) = \sum_i P(J_j^n/I^*(c_{i,1}, \ldots, c_{i,n}))P(I^*(c_{i,1}, \ldots, c_{i,n})), \quad (3.10)$$

$$CS(J_j^n) = \sum_i P(J_j^n/I^{*n}_i)P(I^{*n}_i). \quad (3.11)$$

Then it follows from the 1:1 correspondence of distorted noise-free elements, and non-distorted noise-free elements that:

$$\sum_{G(I_i(S_n))=1} P(J_j^n/I_i(S_n))P(I_i(S_n)) = CL(J_j^n) + CE(J_j^n),$$

and

$$\sum_{G(I_i(S_n))=0} P(J_j^n/I_i(S_n))P(I_i(S_n)) = CH(J_j^n) + CS(J_j^n).$$

Consequently we may rewrite the formula at the end of Section I as:

$$Q(J_j^n) = \frac{CL(J_j^n) + CE(J_j^n)}{CL(J_j^n) + CE(J_j^n) + CH(J_j^n) + CS(J_j^n)} \quad (3.12)$$

We may now turn to an analysis of the class CS. It turns out not to be necessary to make an explicit model as was done for the classes CE, CL and CH. To see this, recall Formula (3.12) of the previous paragraph, and Formulas (3.8) - (3.11), which define its components. An element in these latter sums, e.g.,

$$P(J_j^n/I^*(a_{i,1}, \ldots, a_{i,n}))P(I^*(a_{i,1}, \ldots, a_{i,n}))$$

from the sum CL, may be thought of as a measure of similarity between J_j^n and the distorted noise-free sample $I^*(a_{i,1}, \ldots, a_{i,n})$, weighted by the plausibility of the latter. Thus the value of CS, for example, is roughly proportional to the plausibility or frequency of occurrence of members of CS, and to the similarity of J_j^n to members of CS.

If J_j^n occurs over a rectangle S_n on the visual field which is centered on the image of a line, the value of $CS(J_j^n)$ is quite small relative to the other members of the denominator of $Q(J_j^n)$, owing to the relative infrequency of members of CS as compared with members of CH, and to the fact that such a J_j^n is as similar to a pattern of intensities from a homogeneous region as it is to a pattern of intensities

containing a skewed line. Since the value of $CS(J_j^n)$ is thus small, it could be omitted in this case, or at least approximated with the sum of values of

$$CL^k(J_j^n) + CE^k(J_j^n)$$

over a set of regions S_n^k which are nearly coincident with S_n, but are slightly skewed with respect to it. These values may be obtained "for free" for any region S_n by a line-finder which has to be able to find lines in all positions in the visual field; since it would necessarily compute $CL^k(J_j^n)$ and $CE^k(J_j^n)$ in the process of testing the other regions S_n^k for the presence of lines.

If S_n is a region in the visual field which did not contain a line, a noisy sample J_j^n over S_n would, by a similar argument, have the property that $CS(J_j^n)$ makes a rather small contribution to the denominator of $Q(J_j^n)$, and could be similarly approximated, or indeed omitted.

It is only for the case that the region S_n contains the image of a skew line that the value of $CS(J_j^n)$ becomes significant. However, this effect occurs only for rectangles approximately coincident with a rectangle S_n centered on a line in the visual field. Let $Q^*(J_j^n)$ be the same as $Q(J_j^n)$ but with

the value of $CS(J_j^n)$ omitted from the denominator:

$$Q^*(J_j^n) = \frac{CL(J_j^n) + CE(J_j^n)}{CL(J_j^n) + CE(J_j^n) + CH(J_j^n)} . \qquad (3.13)$$

It has been empirically checked that $Q^*(J_j^n)$ obtains a maximum value locally approximately on the rectangle S_n centered on the line. This is quite plausible, since the value of $Q^*(J_j^n)$ should be large for rectangles approximately coincident with S_n, and if it didn't have a maximum on S_n, it would either be offset parallel to S_n, which is unlikely because of the symmetry of the intensity pattern; or there would be several maxima, which seems not to happen. Hence the omission of $CS(J_j^n)$ from the denominator of $Q(J_j^n)$ has the effect of giving erroneously high values on regions nearly centered on a true line, but not erroneously high values on the maximum value region, namely a region centered on the line. However, these values are not so augmented as to create false maxima for regions not quite centered on the line.

In summary, the set of noise-free distorted samples is partitioned into four classes CL, CH, CE and CS. Our threshold function $Q(J_j^n)$ is expressed, in (3.12), in terms of

four sums $CL(J_j^n)$, $CH(J_j^n)$, $CE(J_j^n)$ and $CS(J_j^n)$ over these four classes. Explicit models of CL, CH and CE allow a simple explicit expression of the a priori probabilities of members of these sets. These probabilities, together with the results of Sections II and III, determine the values of the sums $CL(J_j^n)$, $CH(J_j^n)$ and $CE(J_j^n)$. It has been shown that $CS(J_j^n)$ may be computed indirectly, or even omitted from the denominator of $Q(J_j^n)$. The latter is permissible, resulting in Formula (3.13), in the case where the regional predicate is applied systematically over the whole field of view.

V. COMPUTATIONS OF $CL(J_j^n)$, $CE(J_j^n)$ AND $CH(J_j^n)$ BY MEANS OF INTEGRALS

In the previous section, it was pointed out that the threshold function $Q(J_j^n)$ could be approximated with the function $Q^*(J_j^n)$ given by (3.13). This function is determined entirely by the values of sums $CL(J_j^n)$, $CE(J_j^n)$ and $CH(J_j^n)$, as defined by (3.8), (3.9) and (3.10). In this section, we shall show how to express these sums as integrals, and to reduce them to relatively simple closed form.

It might seem a good strategy to evaluate the

expression, for example $CL(J_j^n)$ as given by (3.8), by substituting values given by (3.4) and (3.7) respectively for the two terms within the summation sign, expressing the resulting sum as an integral, and finding a closed form for the value of the integral. This approach appears to run into difficulties, and a slightly different approach will be used instead.

Our first step is to express, e.g., $CL(J_j^n)$ in a more convenient form. Since the sum in (3.8) is taken over all n-tuples, $CL(J_j^n)$ may be expressed as:

$$\sum_{a_{i,1}} \sum_{a_{i,2}} \cdots \sum_{a_{i,n}} P(J_j^n / I^*(a_{i,1}, \ldots, a_{i,n}))P(I^*(a_{i,1}, \ldots, a_{i,n})),$$

where the sum over $a_{i,k}$ means that $a_{i,k}$ in the summand takes on all positive integral values. We may express $P(I^*(a_{i,1}, \ldots, a_{i,n}))$ as

$$\sum_{a} P(I^*(a_{i,1}, \ldots, a_{i,n})/a)P(a),$$

where P(a) is the a priori probability of a line with idealized amplitude $\underline{a}$; and $P(I^*(a_{i,1}, \ldots, a_{i,n})/a)$ is the relative probability of occurrence of a noise-free distorted

sample with sub region amplitudes $a_{i,1}, \ldots a_{1,n}$, given a line with idealized amplitude $\underline{a}$. The conditional probability $P(I^*(a_{i,1}, \ldots, a_{i,n})/a)$ may be expressed, as in Section IV, as a product of probabilities:

$$P(I^*(a_{i,1}, \ldots, a_{i,n})/a) = \prod_{k=1}^{n} P(a_{i,k}/a),$$

where $P(a_{i,k}/a)$ is understood as the probability, given a line with idealized amplitude $\underline{a}$, that the k-th sub-component of a noise-free sample will have a relative amplitude $a_{i,k}$. The above probability may be expressed as a product, since the elements in the product are statistically independent. Likewise, by statistical independence, $P(J_j^n/I^*(a_{i,1}, \ldots, a_{i,n}))$ may be expressed as a product:

$$P(J_j^n/I^*(a_{i,1}, \ldots, a_{i,n})) = \prod_{k=1}^{n} P(J_j^n(k)/a_{i,k}),$$

where $J_j^n(k)$ is understood to mean the set of intensities of the noisy sample J_j^n which lie within the k-th square sub-component of the region. Substituting and rearranging, we have:

$$CL(J_j^n) =$$
$$\sum_a P(a) \sum_{a_{i,1}} \cdots \sum_{a_{i,n}} \prod_{k=1}^{n} P(J_j^n(k)/a_{i,k}) P(a_{i,k}/a).$$

It is not very difficult to see that the product over $\underline{k}$ and the summations over the $a_{i,k}$'s may be interchanged, so that:

$$CL(J_j^n) =$$

$$\sum_a P(a) \prod_{k=1}^{n} \sum_{a_{i,k}} P(J_j^n(k)/a_{i,k}) P(a_{i,k}/a). \qquad (3.14)$$

Similar expressions exist for $CE(J_j^n)$ and $CH(J_j^n)$.

An expression for $P(J_j^n(k)/a_{i,k})$ may be obtained by adapting (3.4) to the present context, using the value of $\underline{a}$ from (3.5):

$$P(J_j^n(k)/a_{i,k}) = \frac{1}{(1.5\sqrt{2\pi})^m} e^{-\frac{1}{2}\frac{D_{ijk}}{(1.5)^2}} dv_{j,k}^1 \ldots dv_{j,k}^m, \qquad (3.15)$$

where

$$D_{i,j,k} = \sum_{t=1}^{m} (u_{i,k}^t - v_{j,k}^t)^2, \qquad (3.16)$$

which is similar to (3.3) in that it refers to some i-th noise-free distorted sample and a j-th noisy sample, but has a third subscript to indicate it involves only the k-th sub-component. The values $u_{i,k}^t$ are the $\underline{m}$ intensity values within the k-th sub-component of the i-th noise-free distorted sample containing a line; and are taken over points which may be denoted by $p_{i,k}^t$. Since the amplitude in this region is

$a_{i,k}$, we have by assumption 2) of Section IV that the intensities at points $p^t_{i,k}$ are given by:

$$u^t_{i,k} = a_{i,k} g(x^t_{i,k}) + b_{i,k},$$

where $x^t_{i,k}$ is the perpendicular distance between $p^t_{i,k}$, and a line traversing the center of the region in the direction of the expected line. If the geometry of the points $p^t_{i,k}$ is fixed for all samples and sub-regions, then the value of $g(x^t_{i,k})$ is independent of the sample and sub region, and we may express such an intensity value by $g(x^t)$. By assumption 1) of Section IV, we may choose $b_{i,k}$ so that the sum of $u^t_{i,k}$ is zero, i.e., by choosing $b_{i,k} = \overline{g}$. Hence defining values $(U^1, \ldots, U^m)$ by:

$$U^t = g(x^t) - \sum_t g(x^t)/m, \tag{3.17}$$

we may express the $\underline{m}$ intensities in the k-th sub-component of the i-th distorted noise-free sample by:

$$(u^1_{i,k}, \ldots, u^m_{i,k}) =$$

$$(a_{i,k}U^1, \ldots, a_{i,k}U^m) \tag{3.18}$$

where:

$$\sum_t U^t = 0. \tag{3.19}$$

To be consistent with assumption 1), we should also assume that:

$$\sum_t v^t_{j,k} = 0. \tag{3.20}$$

The value of $D_{i,j,k}$ of (3.15) may be expressed in terms of the values just defined as:

$$D_{i,j,k} = \sum_t (v^t_{j,k} - a_{i,k}U^t)^2. \tag{3.21}$$

It is easy to show that:

$$D_{i,j,k} = b(a_{i,k} - a^*_{j,k})^2 + R_{j,k}, \tag{3.22}$$

where:

$$a^*_{j,k} = \frac{\sum_t U^t v^t_{j,k}}{\sum_t (U^t)^2} \tag{3.23}$$

$$b = \sum_t (U^t)^2 \tag{3.24}$$

$$R_{j,k} = \sum_t (v^t_{j,k} - a^*_{j,k}U^t)^2. \tag{3.25}$$

Combining we have, omitting the differentials $dv^t_{j,k}$:

$$P(J^n_j(k)/a_{i,k}) = \frac{1}{(1.5\sqrt{2\pi})^m} e^{-\frac{1}{2}\left[\frac{b(a_{ik} - a^*_{jk})^2 + R_{jk}}{(1.5)^2}\right]}. \tag{3.26}$$

We may express $P(a_{i,k}/a)$ as before by assuming a normal distribution on the various values of $a_{i,k}$, with

standard deviation σ_n and mean $\underline{a}$:

$$P(a_{i,k}/a) = \frac{1}{\sqrt{2\pi}\,\sigma_n} e^{-\frac{1}{2}\left[\frac{a - a_{ik}}{\sigma_n}\right]^2} da_{i,k}. \qquad (3.27)$$

Combining (3.14), (3.26) and (3.27) we have, again omitting differentials $dv^t_{j,k}$:

$$CL(J^n_j) = \sum_a P(a) \prod_{k=1}^{n} \sum_{a_{i,k}} \frac{1}{\sqrt{2\pi}\sigma_n (1.5\sqrt{2\pi})^m} e^{-\frac{1}{2}\left[\frac{b(a_{ik} - a^*_{jk})^2 + R_{jk}}{(1.5)^2} + \frac{(a - a_{ik})^2}{\sigma_n^2}\right]} da_{ik} da \qquad (3.28)$$

with b, $a^*_{j,k}$ and $R_{j,k}$ given by (3.23), (3.24) and (3.25). The inner sum may be put in the form:

$$A \sum_{a_{i,k}} e^{-Ba^2_{i,k} + Ca_{i,k} + D} da_{i,k},$$

which admits of the approximation used in the preceding section:

$$A \sum_x e^{-Bx^2 + Cx + D} dx \approx$$

$$A \int_{-\infty}^{\infty} e^{-Bx^2 + Cx + D} dx = \frac{A\sqrt{\pi}}{\sqrt{B}} e^{\frac{C^2}{4B} + D}.$$

Applying this approximation to the inner sum in (3.28), we

have after considerable algebra:

$$CL(J_j^n) \approx \sum_a P(a) \prod_{k=1}^{n} \frac{1.5}{(\sqrt{2\pi}\,1.5)^m \sqrt{(1.5)^2 + b\sigma_n^2}}$$

$$e^{-\frac{1}{2}\left[\frac{R_{jk}}{(1.5)^2} + \frac{b(a^*_{jk} - a)^2}{(1.5)^2 + b\sigma_n^2}\right]} da,$$

where again, the diffentials $dv^t_{j,k}$ have been omitted. This may alternatively be expressed:

$$CL(J_j^n) \approx \sum_a P(a) \left[\frac{1.5}{(\sqrt{2\pi}\,1.5)^m \sqrt{(1.5)^2 + b\sigma_n^2}}\right]^n$$

$$e^{-\frac{1}{2}\left[\frac{\sum_k R_{jk}}{(1.5)^2} + \frac{b\sum_k (a^*_{jk} - a)^2}{(1.5)^2 + b\sigma_n^2}\right]} da,$$

and expressing P(a) as before by:

$$P(a) = \frac{P(CL)}{\sqrt{2\pi}\,\rho_n} e^{-\frac{1}{2}\left[\frac{a}{\rho_n}\right]^2}$$

this latter expression for $CL(J_j^n)$ may be put in the form:

$$A \sum_a e^{-Ba^2 + Ca + D} da,$$

and may be approximated as before to yield:

$$CL(J_j^n) \approx P(CL)\left[\frac{1.5}{(\sqrt{2\pi}1.5)^n\sqrt{(1.5)^2+b\sigma_n^2}}\right]^n \frac{\sqrt{b\sigma_n^2+(1.5)^2}}{\sqrt{b\sigma_n^2+nb\rho_n^2+(1.5)^2}}$$
$$e^{-\frac{1}{2}\left[\frac{\sum_k R_{jk}}{(1.5)^2}+\frac{b\sum_k a_{jk}^{*2}}{(1.5)^2+b\sigma_n^2+nb\rho_n^2}+\frac{nb^2\rho_n^2\sum_k (a_{jk}^{*}-\overline{a_{jk}^{*}})^2}{(b\sigma_n^2+(1.5)^2)(b\sigma_n^2+nb\rho_n^2+(1.5)^2)}\right]}, \quad (3.29)$$

with differentials $dv_{j,k}^t$ omitted. Similar expressions exist for $CE(J_j^n)$ and $CH(J_j^n)$.

VI. APPLICATION OF THE THEORY TO SOME ACTUAL VIDISSECTOR OUTPUT

In this Section we shall apply some computational procedures suggested by the foregoing theory to intensity values obtained from the vidissector. These intensity values will be taken along single lines normal to various lines and edges in the visual field. The results of the computation will serve to validate the theory presented, and justify some assumptions and approximations in its computational implementation which will be used by the program described in the next chapter.

In applying the foregoing theory of a regional line predicate, it is first necessary to choose a suitable physical width for a rectangular region. It appears that, for the

vidissector in current use, this width should be between 25 and 50 units, where one unit is 1/2000 of the full field width. This range of values is chosen to conform with the assertion that, in the terminology of chapter two, the majority of regional evidence for the existence of a line is contained within a neighborhood of the line whose radius is on the order of a few times the "critical resolution radius" of the vidissector. This latter quantity, as defined in Section I of chapter two, may be easily obtained by examining Figure 3.3. It may be seen that the intensity in the upper curve falls to one half of its maximum value at a radius of three units from the center of the curve. This value is approximately the critical resolution radius as previously defined. For present purposes, we took "a few times" to mean specifically eight times, which yielded a region width of 50 units. Later results, reported in the next chapter, indicated that this value may have been conservatively large, and somewhat better results seem to have been obtained with a region width of 25 units. In any case, the value of 50 units turned out to be adequate for present purposes.

It seems appropriate to begin an empirical investigation of the properties of the regional line predicate here developed with a consideration of the case where the region to which it is applied consists of a single square sub-region. This admits of a simplification in the formula for $Q^*(J_i^n)$ as given by (3.13) and (3.29). In particular, the term involving $(a^*_{j,k} - \overline{a^*_{j,k}})$ in (3.29) becomes zero, since $\overline{a^*_{j,k}}$ is the average of only one term. Thus this term is eliminated from the exponent in (3.29); and likewise similar terms are eliminated from corresponding expressions for $CE(J_j^n)$ and $CH(J_j^n)$.

Another simplification of the exponent in the expression for $CL(J_j^n)$ given by (3.29), and likewise for corresponding expressions for $CE(J_j^n)$ and $CH(J_j^n)$, follows from a consideration of the relative magnitudes of the values $(1.5)^2$ and $b\sigma_n^2$. The latter is the value of $\underline{b}$ given by (3.24) times the variance of values of idealized relative amplitude of the various lines in scenes in the real world. Since the distribution of these values has a mean of zero, it follows from elementary properties of the normal curve that about 1/3 of all lines have idealized relative amplitude $\underline{a}$ with the

property that $|a| \geq \sigma_n$. Since the variance of the sub-region amplitudes $a_{i,j}$ is small relative to the variance of the values of $\underline{a}$, the same holds true of these amplitudes, i.e., about 1/3 of all sub-components of noise-free samples containing lines have amplitudes $a_{i,j}$ with $|a_{i,j}| \geq \sigma_n$. If the intensities within a sub-region of a noise-free distorted sample are given by $(u^1_{i,k}, \ldots, u^m_{i,k})$ and their <u>spread</u> is defined as:

$$\sqrt{\sum_t (u^t_{i,k} - \bar{u}_{i,k})^2},$$

with

$$\bar{u}_{i,k} = \sum_t u^t_{i,k} / m$$

then it is easy to see from (3.18) that this spread has the value $a_{i,j}\sqrt{b}$, where $\underline{b}$ is as defined in (3.24). It follows from what was said previously that about 1/3 of all sub-components of noise-free samples have a spread of greater than $\sqrt{b}\sigma_n$. More informally stated: $\sqrt{b}\sigma_n$ is a not uncommonly large value of the spread of intensities in sub-components of noise-free samples containing lines. In the case of <u>noisy</u> samples containing lines, empirical investigations suggest that a spread of 30 is not uncommon; and it is easy to see from the fact that noise level is relatively much smaller than

30, that this value is also not an uncommonly large value of spread for noise-free samples containing lines. Consequently:

$$b\sigma_1^2 >> (1.5)^2. \tag{3.30}$$

Now by simple algebra the exponent in (3.29) may be written, excluding the third term which is zero, and using the relations in (3.23), (3.24) and (3.25):

$$-\frac{1}{2}\left[\frac{\sum_t (v_{j,1}^t)^2}{(1.5)^2} - \frac{\left[\sum_t U^t v_{j,1}^t\right]^2 (\sigma_1^2 + n\rho_1^2)}{(1.5)^2((1.5)^2 + b\sigma_1^2 + nb\rho_1^2)}\right].$$

It follows from (3.30) that the coefficient of the second sum may be approximated by $(1/b(1.5)^2)$, so that the exponent of (3.29) is almost exactly equal to:

$$-\frac{1}{2}\left[\frac{\sum_t (v_{j,1}^t)}{(1.5)} - \frac{(\sum_t U^t v_{j,1}^t)^2}{b(1.5)}\right]. \tag{3.31}$$

A similar sort of simplification of the coefficient of $\underline{e}$ in (3.29) may be made from a consideration of the relative magnitudes of $nb\rho_n^2$ and $b\sigma_n^2$. The value of σ_n^2 represents the variance of a quantity $\underline{a}$, idealized relative intensity; and ρ_n^2 is, according to assumption 5) of Section IV, the variance of small purturbations of the latter. It is reasonable to assume that $\sigma_n^2 >> \rho_n^2$, and that even $b\sigma_n^2 >> nb\rho_n^2$. By this approximation

it follows that the coefficient in (3.29) is close to:

$$\frac{P(CL)}{\sqrt{b}\,\sigma_n(\sqrt{2\pi}1.5)^m} . \tag{3.32}$$

If this and the approximation in (3.21) are applied to (3.29), and similar approximations are applied to corresponding formulas for $CE(J_j^1)$ and $CH(J_j^1)$; Formula (3.13) reduces to:

$$Q^*(J_j^1) \approx$$

$$\frac{\frac{P(CL)}{\sqrt{b}\,\sigma_1} e^{\frac{1}{2}\frac{SL(j)}{(1.5)^2}} + \frac{P(CE)}{\sqrt{b'}\,\sigma_1'} e^{\frac{1}{2}\frac{SE(j)}{(1.5)^2}}}{\frac{P(CL)}{\sqrt{b}\,\sigma_1} e^{\frac{1}{2}\frac{SL(j)}{(1.5)^2}} + \frac{P(CE)}{\sqrt{b'}\,\sigma_1'} e^{\frac{1}{2}\frac{SE(j)}{(1.5)^2}} + \frac{P(CH)}{\sqrt{b''}\,\sigma_1''} e^{\frac{1}{2}\frac{SH(j)}{(1.5)^2}}} \tag{3.33}$$

Where: σ_1, σ_1', and σ_1'' are respectively the idealized relative amplitude variances of elements of CL, CE and CH;

$$b = \sum_t (U^t)^2, \tag{3.34}$$

and $(U^1, \ldots, U^m)$ is some paradigm intensity pattern for component regions of members of CL;

$$b' = \sum_t (U'^t)^2, \tag{3.35}$$

and $(U'^1, \ldots, U'^m)$ is some paradigm intensity profile for component regions of members of CE;

$$b'' = \sum_t (U''^t)^2, \tag{3.36}$$

and $(U''^1, \ldots, U''^m)$ is some paradigm intensity profile for component regions of members of CH;

$$SL(j) = (\sum_t U^t v^t_{j,1})^2/b \tag{3.37}$$

$$SE(j) = (\sum_t U'^t v^t_{j,1})^2/b' \tag{3.38}$$

$$SH(j) = (\sum_t U''^t v^t_{j,1})^2/b'' \tag{3.39}$$

and $(v^1_{j,1}, \ldots, v^m_{j,1})$ are the intensities in J^1_j.

It is necessary to assign values to the remaining constants in (3.33). First we shall assume that lines and edges are equally likely, so that P(CL) = P(CE). Also in the examples which follow, it will become clear that an edge or line occurs in one sample out of 75, so that:

$$1/75P(CH) = P(CE) = P(CH).$$

Secondly, it follows from empirical investigation of the range of spread of homogeneous samples with various gradients, that five is a reasonable value for $b''\sigma''_1$. It was already pointed out that 30 is a reasonable value for $b\sigma_1$, and we shall assume this is the case for $b'\sigma'_1$ as well. Thus:

$$Q^*(J^1_j) \frac{e^{\frac{SL(j)}{2(1.5)^2}} + e^{\frac{SE(j)}{2(1.5)^2}}}{e^{\frac{SL(j)}{2(1.5)^2}} + e^{\frac{SE(j)}{2(1.5)^2}} + 450\, e^{\frac{SH(j)}{2(1.5)^2}}} \tag{3.40}$$

For values greater than . 5 this function is nearly monotone with:

$$\text{Max}(SL(j)/2(1.5)^2,\ SE(j)/2(1.5)^2) - SH(j)/2(1.5)^2 - \text{Ln}(450.),$$

or equivalently, with the function $Q^{**}(J_j^1)$ given by:

$$Q^{**}(J_j^1) = \text{Max}(SL(j), SE(j)) - SH(j) - 27.5 \qquad (3.41)$$

This follows from an observation about the results given later in this section, namely that SL(j) and SE(j) are rarely of the same order of magnitude. This is not surprising, in view of the fact that one expresses the similarity of the obtained pattern to a noise free edge, and the other expresses the similarity to a noise-free line. For values of $Q^{*}(J_j^1)$ greater than . 5, the pattern must be at least somewhat similar to either an edge or a line. It could not be simultaneously similar to both. Consequently the formula (3. 40) above, which is of the form

$$\frac{A + B}{A + B + C}$$

has the property that when its value is above . 5, A and B are considerably different in value, and both positive.

It is not difficult to show that in such a case if, for example

$$\text{Max}(A,\ B) \geq 10\ \text{Min}(A,\ B),$$

then the approximation:

$$\frac{\text{Max}(A,\ B)}{\text{Max}(A,\ B) + C} \approx \frac{A + B}{A + B + C}$$

is good to better than 6 percent. The fact that $Q^*(J_j^1)$ is monotone with $Q^{**}(J_j^1)$ follows from the fact that

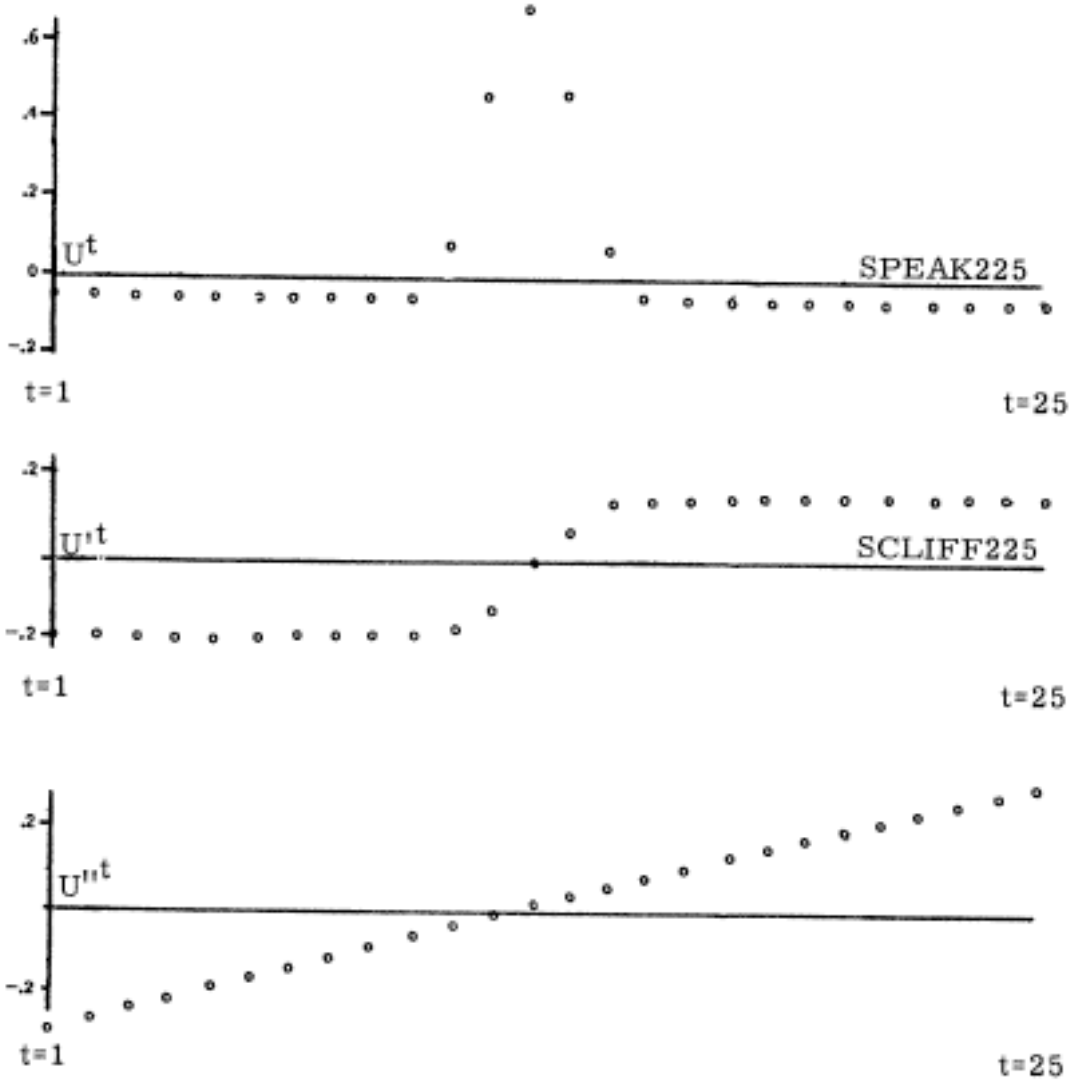
$$\frac{\text{Max}(e^a,\ e^b)}{\text{Max}(e^a,\ e^b) + e^c}$$

is monotone with Max(a, b) - c, if a, b and c are positive.

It is desirable to apply the predicate to a variety of regions which are homogeneous, contain off-center lines, or contain centered lines or edges. We have chosen data from a series of areas on the visual field which are 200 units wide and contain the images of vertical edges and lines approximately centered laterally. The predicate will be applied to 75 successive overlapping regions 50 units wide, which are offset relative to each other by two units and thus exactly cover the 200-unit-wide area. The procedure will be applied to vertically oriented regions only, since only vertically oriented lines exist in the data samples. The fact that 75

regions are tested and only one contains a centered line or edge justifies the relation P(CH) = 75P(CE) used above.

Finally, before applying the predicate to real data, which will be done by computing the function $Q^{**}(J_j^1)$, it is necessary to give values for the sets $\{U^t\}$, $\{U'^t\}$ and $\{U''^t\}$ used in formulas (3.34) - (3.39). It will be recalled that these are paradigm sets of intensity values at points in some fixed geometry within a square sub-region. They correspond respectively to a line centered in the region, an edge centered in the region, and a homogeneous region. According to the fixed geometry which we shall use, the points will be equally spaced along a line traversing the square sub-component through its center and oriented normal to the expected line. These values may be ordered in an obvious manner according to their positions along the line mentioned. They are shown graphed as a function of this ordering in Figure 3.6. The values of b, b' and b'' for these sets of intensities are all unity. The intensities from the real world will come from single scans of 100 points taken normal to a vertical edge or line in the real world. The points will be separated laterally by two



PARADIGM INTENSITY PROFILES

USED FOR FIGURE 3.7

Figure 3.6

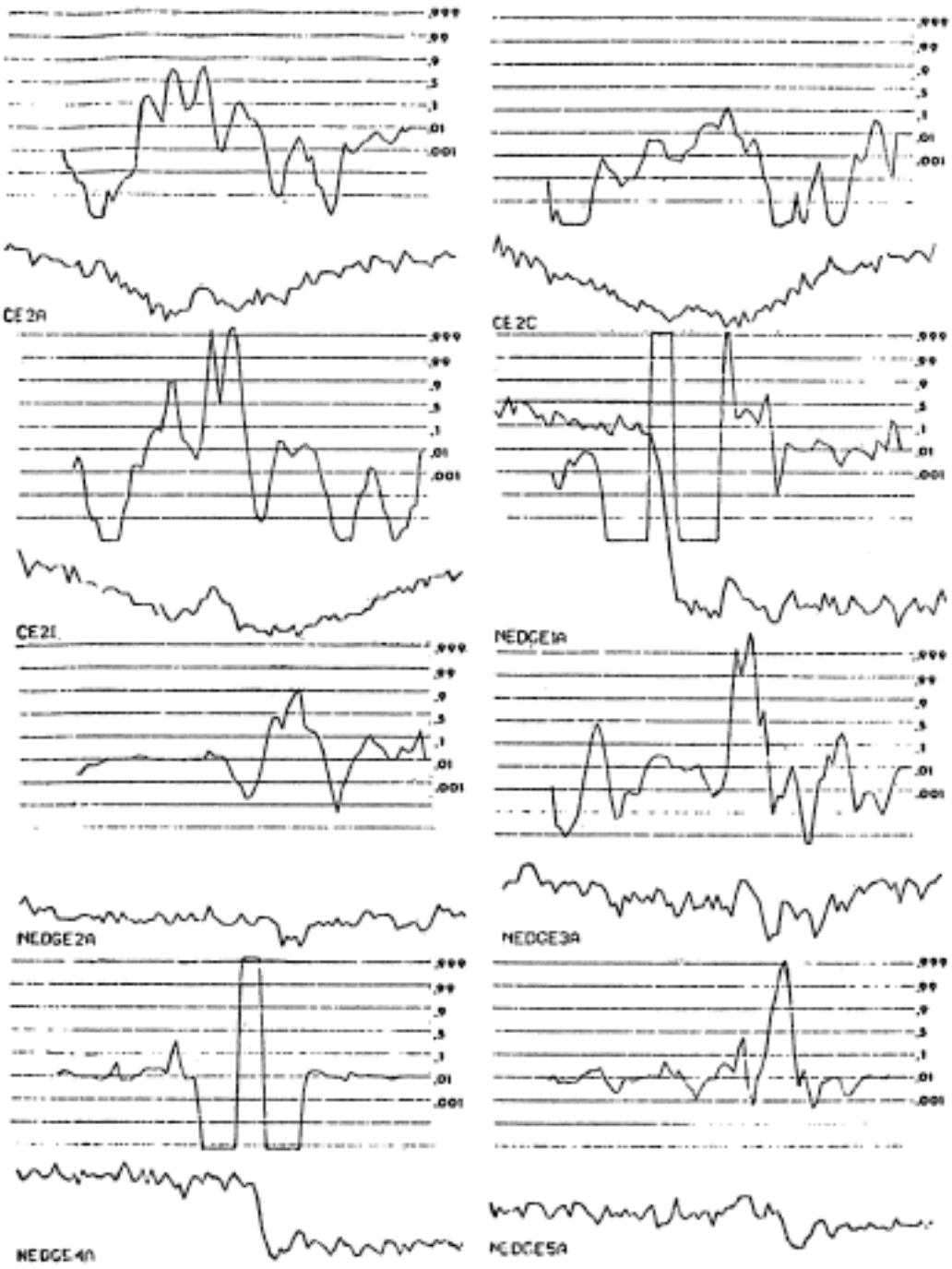
units, so that the total width of the area scanned will be 200 units. If it is desired to apply $Q^{**}(J_j^1)$ to some region along this scan, out of the 100 intensities $I_1, \ldots, I_{100}$, some set $I_{i+1}, \ldots, I_{i+25}$ fall within the region. Then it is trivial to select intensities from among these to be used as the values $v_{j,1}^1, \ldots, v_{j,1}^m$ for use in formulas (3.37), (3.38) and (3.39). They are simply $I_{i+1}, \ldots I_{i+25}$ respectively. Thus we will be applying (3.41) to 75 successive 25-tuples from the 100 intensities in a particular scan.

It is important to digress here to point out the advantages of the sort of geometry of points within a sub-region which is used above. This advantage will be realized in the next chapter when it is desired to apply the regional predicate exhaustively over the whole field. In particular it will be desired to compute the predicate for lines of all orientations through any point P on the visual field. Consider those situations where the regions are oriented within $\pm$ 45 degrees of the vertical axis of the vidissector, and all pass through some point P; and consider only the square sub-region containing the point P, which will be termed the p-th

sub-region. We shall in chapter IV take intensities relevant to regions so oriented, along lines separated by approximately the width $\underline{s}$ of a region, and oriented parallel to the x-axis of the vidissector. We may thus assume that there will be a unique scan passing within s/2 of the point P; and that the intensities $I_1, \ldots, I_m$ along its length within a distance s/2 of P are exactly the ones from which the $v^t_{j,p}$ will be selected. The procedure for selecting the values $v^t_{j,p}$ from among this set is parameterized by the angle Ø which the region makes with the vertical. If the spacing of points in the fixed geometry of points is $\underline{d}$, then one wishes to choose points from among the $I_1, \ldots, I_m$ such that their actual spacing is approximately $d/\mathrm{Cos}(\emptyset)$. If the spacing of the points along the scan line is itself $\underline{d}$, which will usually be the case in practice, then this matching of obtained intensities with those in the prototype amounts to "stretching" the former by some factor between 1 and $\sqrt{2}$. Thus the various values of, e.g., $SE(J^n_{j,p}) = \sum_t v^t_{j,p} U^t$, which are used in the computation of $Q^*(J^n_j)$ for various regions containing P, differ only in the amount of "stretching" of one spatial axis relative to the other in the matching of intensity values of the scan against

those of the set $\{U^t\}$. Denoting these various values of $CE(J^n_{j,p})$ by $CE(J^n_{j,p}, \emptyset)$, it turns out empirically that the range of these values is rather small; and in particular for $|\emptyset| \leq 45°$, the values are within something like 20 percent of the closer of $CE(J^n_{j,p}, 10°)$, $CE(J^n_{j,p}, 30°)$. One may thus compute these two values only, and for any region containing the intensities from which they were computed, use one or the other in the calculation of $Q^*(J^n_j)$ for that region. In fact we shall in chapter IV actually compute these two values by stretching the spatial dimension of the paradigm, rather than vice versa. This amounts to using two peaks of different widths as paradigm profiles for lines, and two edges of different severity of slope for paradigms for edges. It should be clear that if the fixed geometry of the points were two-dimensional, e.g. on a square grid, no such simple accommodation for the various angles might be possible; and doubtless the amount of computation involving the intensities in each neighborhood would be far more than the calculation of two values.

To return to the application of Q** to real data, our first example, in the upper left of Figure 3.7, comes from a horizontal scan across the forward edge of a cube. The



VALUES OF Q* FOR VARIOUS INTENSITY PROFILES

Figure 3.7

scan covers approximately one eighth of the apparent width of the cube. The cube is a solid metal block which has been painted with white paint so as to have a very uniform surface and sharp edge. In each of the eight figures, the logarithms of the reciprocals of intensity are given as the bottom of the two curves. The absolute range of values is 11 units, where 64 units corresponds to a factor of two in intensity. It is easy to infer from the physical properties mentioned that the slope of the intensity curve on either side of the edge is due to a systematic non-uniformity in illumination, rather than to a non-uniformity of reflectance properties. At the edge itself there appears to be a sort of "negative highlight"(a local peak). This appears also at the centers of the scans in the upper right and upper-middle left examples of Figure 3.7, which were taken from the same cube; and is apparently a property of this particular edge. The 100 intensities in the example give rise to 75 values of Q** which are graphed above the intensity curve. A value of Q** is derived from 25 successive intensities consisting of the one just below it and twelve on either side. Thus a peak in the value of Q** should exactly coincide with a peak or discontinuity in the intensity profile. The horizontal lines in each of the eight figures are placed at distances along the Q** axis which

correspond to values of Q* which are indicated on the lines.

A reasonable threshold value for Q* appears from an analysis of this example, and those appearing later in this section, to be about 0.1. A cutoff level of approximately this magnitude will be used by the program described in the next chapter. This cutoff value is chosen on the basis of its giving about 70% false positives. There is no simple exact relationship between the Q* cutoff and the false positive error rate; however it follows from the first theorem of chapter two that if β is the cutoff value for Q*, then $1-\beta$ is an upper bound of the false positive error rate. This may be seen by noting that for the threshold procedure, the sum in (2.3) gives the false positive error rate. Every element in the sum is some probability $P(J_j(S))$ times a value less than $1-\beta$. Since the values of $P(J(S))$ sum to one, the sum in (2.3), i.e., the false positive error rate, is less than $1-\beta$. Thus thresholding Q* at 0.1 should give a false positive error rate less than 0.9. However, the extent to which the false positive error rate is actually less than 0.9 is an empirical matter. Evidently the various probabilities and so forth are such as to give an actual value of 0.7 in this case.

An obvious question now arises: how subtle may an edge or line be relative to the noise level in order to be just detectable, e.g., with the threshold of 0.1. An example of such an edge is given in Figure 3.7 at the upper right. It is possible to make a general quantitative description of such a line or edge by means of a value derived from the cutoff value of Q**. Suppose the cutoff value is 0.1; then we have:

$$\frac{\text{Max}\left[e^{\frac{SL(j)}{2(1.5)^2}}, e^{\frac{SE(j)}{2(1.5)^2}}\right]}{\text{Max}\left[e^{\frac{SL(j)}{2(1.5)^2}}, e^{\frac{SE(j)}{2(1.5)^2}}\right] + 450\ e^{\frac{SH(j)}{2(1.5)^2}}} \approx 0.1,$$

or:

$$9.\ \text{Max}\left[e^{\frac{SL(j)}{2(1.5)^2}}, e^{\frac{SE(j)}{2(1.5)^2}}\right] \approx 450.\ e^{\frac{SH(j)}{2(1.5)^2}},$$

or

$$\text{Max}(SL(j)/2(1.5)^2, SE(j)/2(1.5)^2) + \text{Ln}(9) \approx$$

$$SH(j)/2(1.5)^2 + \text{LN}(450.),$$

or:

$$\text{Max}(SL(j),\ SE(j)) - SH(j) - 27.5 \approx 2(1.5)^2 \text{Ln}(9) = 9.9$$

or:

$$\text{Max}(SL(j), SE(j)) - SE(j) \approx 17.6 \qquad (3.42)$$

Since SL(j), SE(j) and SH(j) are positive, the minimum value of Max(SL(j), SE(j)) subject to (3. 42) occurs when SH(j) = 0. This is the case when the net intensity gradient for the sample is zero; a fact which can be deduced from (3. 39) and a knowledge of some of the arithmetic involved with simple least-squares analysis. In this case for a marginally detectable line or edge, SL(j) or SE(j) has the value 17. 6. If the intensity profile follows the form, e. g. , of the upper curve in Figure 3. 6; then it is not difficult to see, by the use of (3. 37), what its actual peak to valley range would be: Suppose the 25 intensities in the noisy intensity profile are given exactly by $aU^1, \ldots, aU^{25}$. Then SL(j) for these values is, according to (3. 37):

$$\frac{\left[a \sum_t (U^t)^2\right]^2}{\sum_t (U^t)^2} .$$

However, the denominator term, which is simply $\underline{b}$ as defined in (3. 34), was chosen to be unity for the set of U^t's in Figure 3. 6; so the value of SL(j) turns out to be simply a^2. This, it was agreed, is 17. 6; so the value of $\underline{a}$ is 4. 2. Thus a marginally detectable intensity profile of this shape

is similar to that at the top of Figure 3. 6, but multiplied by a factor of 4. 2; and thus has a peak to valley range of about three units. In the case that SE(j) is a maximum and the phenomenon is edge-like, one may assume the approximate shape of the intensity profile is that given by the middle curve of Figure 3. 6. By a similar argument, the marginally detectable sample of this sort has a range, in this case net intensity difference, of about 1. 7 units. The central feature of the scan in the upper right of Figure 3. 7 has a value of Q* at threshold; and the amplitude of the peak at the center may be compared with the former of the two theoretical values just referred to. The intensity range for the whole scan is 17 units, from which it is easy to see that the peak to valley height of the small peak in the center is about four units. This compares well with the theoretical value of three units.

Clearly the amplitudes of marginally detectable lines and edges must be considered in relation to the noise level. Noise may in general be due both to time noise, random fluctuations in intensity measured at a point under constant illumination; and to space noise, which consists in random fluctuations in intensity as one scans across a surface. A

combination of the two may be obtained from an analysis of successive differences in intensity of the above samples in areas away from the known features. This analysis yields a standard error of successive differences of about two in all cases. Since the variance of successive differences is twice the variance of individual values about a mean, then the standard error of the latter should be about $2/\sqrt{2}$, or about 1.6. This is approximately the value of the standard error involving time noise only, as reported in (3.5). Consequently, the space noise is in this case negligible, which is not surprising in view of the deliberate choice of smooth-surfaced objects. In any case, the marginally detectable amplitudes of 3 units and 1.7 units for lines and edges respectively should be thought of in relation to an underlying noise level whose standard error is 1.5. For example, using the 25 point geometry previously described, one may state that a marginally detectable peak has a peak to valley range of twice the noise level, and a marginally detectable edge has a net intensity difference of about 1.2 times the noise level.

Six additional examples are given in Figure 3.7. The left-hand upper-middle scan was taken from the same

cube and under the same lighting conditions as the top two in the illustration. The other five were obtained from a scan across the vertical forward edge of a second cube under various lighting conditions. The lower-middle pair, marked NEDGE2A and NEDGE3A, were taken under identical lighting conditions, but at different heights along the vertical edge. This cube is uniformly painted, but the edges are slightly rounded, accounting for the pronounced highlight (dip in intensity, since the values are actually Log(1/intensity)) apparent in some of the profiles, particularly the one marked NEDGE2A.

There is a certain weakness in the use of so great a region width as 50 units, namely that the intensity profiles across a region may have non-negligible gradients toward the edges of the region, appearing, e. g., like the profiles in Figure 3. 8. For this reason, it is useful to generalize the thresholding procedure so far described by computing values $SL_1(j)$, ... , $SL_n(j)$, $SE_1(j)$, ... , $SE_m(j)$ for a set of $\underline{n}$ peaked profiles and $\underline{m}$ edge profiles, and using a maximum of all these values in place of the maximum over two elements as in 3. 41. Denoting this generalized version

of Q** by Q'**, we may express it in two slightly different forms:

$$Q'^{**}(J_j^1) =$$

$$\max_{i,k} (SL_i(j), SE_k(j)) - SH(j) - 27.5 =$$

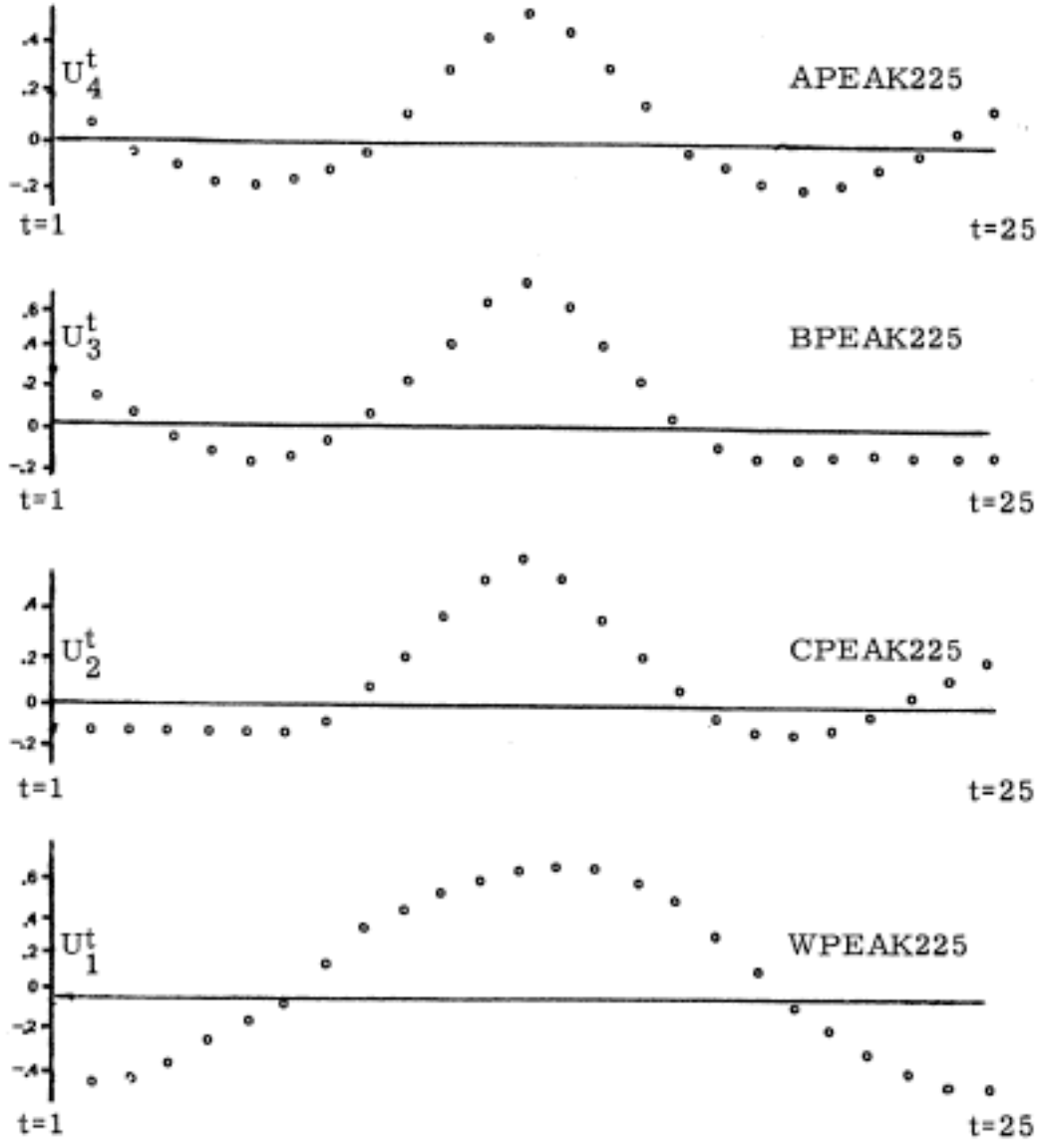
$$\max_{i,k} (SL_i(j) - SE(j), SE_k(j) - SE(j)) - 27.5. \qquad (3.43)$$

This generalized formula will be used in the program described in the next chapter. In the remainder of this chapter we will discuss its application to the previously analyzed data of Figure 3.7.

For present purposes, we shall consider only a possible multiplicity of paradigm peak profiles. Four additional paradigm profiles besides the one in Figure 3.6 are illustrated in Figure 3.8, and denoted by WPEAK, CPEAK, BPEAK, and APEAK respectively. It is instructive to compute individually the values whose maximum gives the value of Q'** according to the second expression in (3.43). In particular we shall consider six values in the following order:

$SL_1(j)$-SH(j), corresponding to the paradigm WPEAK;

$SL_2(j)$-SH(j), corresponding to the paradigm CPEAK;



PARADIGM INTENSITY PROFILES

USED FOR FIGURES 3.9 AND 3.10

Figure 3.8

$SL_3(j)$-SH(j), corresponding to the paradigm BPEAK;

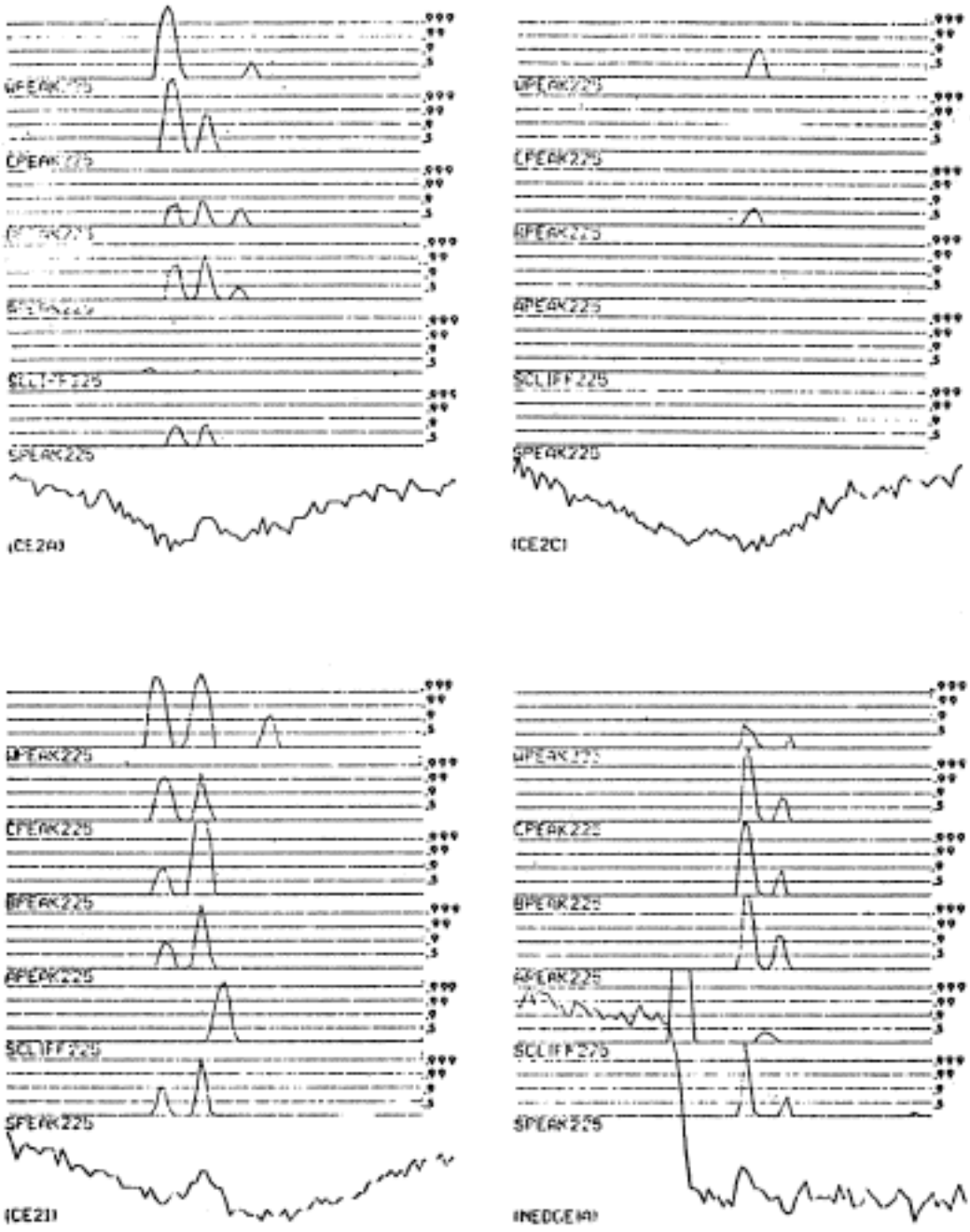
$SL_4(j)$-SH(j), corresponding to the paradigm APEAK;

SE(j)-SH(j), corresponding to the edge paradigm of Figure 3. 6;

$SL_5(j)$-SH(j), corresponding to the peak paradigm of Figure 3. 6.

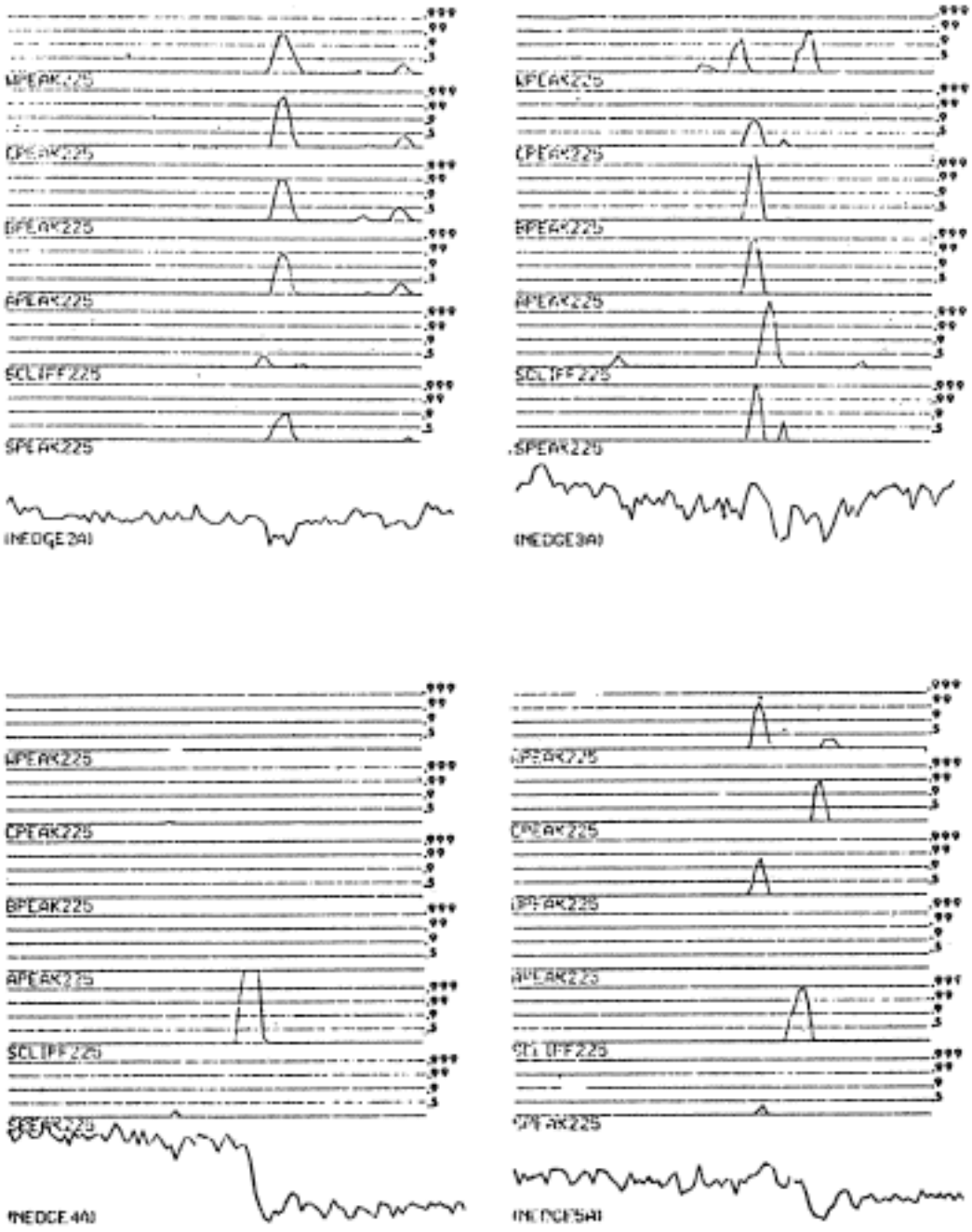
In Figures 3. 9 and 3. 10, we have these six "components" of Q'** individually graphed above the corresponding intensity profile, in the above order top to bottom, in a manner similar to that used in Figure 3. 7.

One may make a series of observations concerning the various curves in Figures 3. 9 and 3. 10. Comparing the lowermost two component curves in the upper left of Figure 3. 9, it is easy to see that the pair of peaks in the upper left illustration of 3. 7 were due to the peak-detector profile entirely. This follows from the fact that the peaks appear only in the lower of the pair of curves from 3. 9 just mentioned. Regarding the same profile, one may see that the major peak from among the component curves is the uppermost. This corresponds to the profile WPEAK, which is a wide peak. Evidently the feature represented by the maximum of this



VALUES OF Q'* FOR VARIOUS INTENSITY PROFILES

Figure 3.9



VALUES OF Q'* FOR VARIOUS INTENSITY PROFILES

Figure 3.10

curve is the inverted wide peak to the left of the center of the intensity scan. Another observation concerns the lower right example in 3. 9. It is reassuring to note that although the discontinuity present in the intensity profile is quite large, it was "noticed" by none of the peak-detector profiles. This is evidenced by the fact that in the top four component curves, and in the bottom one, the curves are not above threshold in the vicinity of the extreme intensity discontinuity. On the other hand, the next-to-bottom curve, corresponding to an edge, has an extreme peak at this location. Lastly, one may make the general observation that the maximum of all six component values of Q'** may be thresholded at a considerably higher level than Q**, for example at a value corresponding to a Q'* value of . 99, and still detect all edges, but at a considerably better signal-noise ratio. Observation of the figures shows that at this level a few false positives remain, e. g., the peak to the right of the large discontinuity in the lower right example in Figure 3. 9. An examination of the intensity curve itself would appear to justify the contention that there really was "something there" after all, perhaps a spot on the cube, or an insensitive spot on the vidissector photocathode.

CHAPTER 4

A COMPUTER PROGRAM FOR FINDING LINES

I. INTRODUCTION

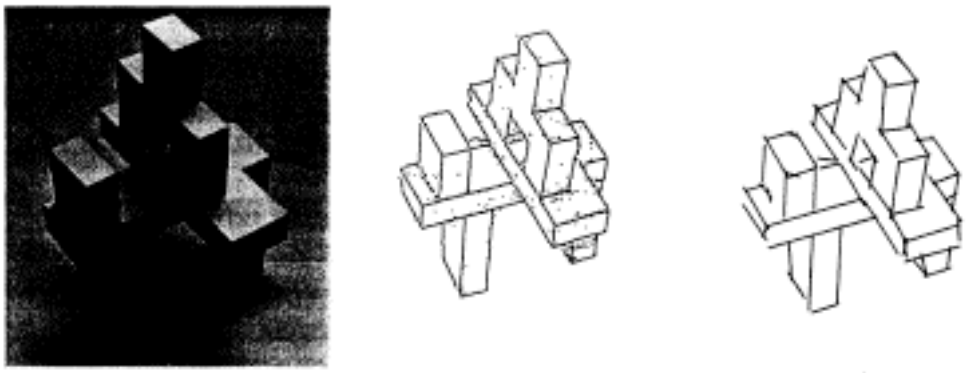


Figure 4.1

Figure 4.1 illustrates the first stages in the analysis of scenes by the program to be described in this chapter. On the left is a photograph of a set of objects placed before a random access optical input device connected to a digital computer. Programs in the computer have access to

intensities in the field of view at arbitrary locations on a 20000_8 by 20000_8 grid. Intensities are obtained by the program along 50 vertical and 50 horizontal scans of 500 intensities each. The intensities are converted into "feature points", as illustrated in the center illustration of 4.1. Feature points are locations where the scans are adjudged, by spatially local processing, to intersect with an edge of an object in the real world. Often noise in the real world or within the vidissector will produce an intensity configuration along part of a scan which is edge-like and gives rise to a "false positive" feature point. The right-hand portion of Figure 4.1 illustrates the output of a program which extracts lines from an array of feature points and which is designed to ignore these extraneous "false positive" points by detecting chains of lined-up feature points. A final stage of the analysis, not illustrated in Figure 4.1, is the proposing and verification of lines not located by the above procedure. Lines are proposed on the basis of lines already located, in places where figures appear to be incomplete.

II. A DESCRIPTION OF SCENES ANALYZED

In this chapter we will describe an object recognition program based on the theory so far developed. This section will be devoted to a discussion of the restrictions imposed by the theory on the scenes analyzed, and on the optical input device used. Various scenes conforming to the restrictions will be illustrated.

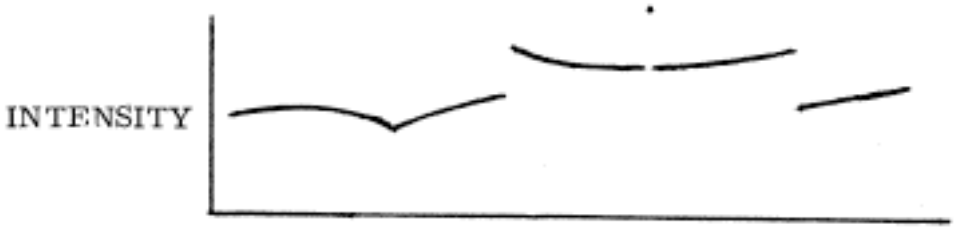
The theory developed in the preceeding chapters is applicable to certain types of scenes, and to optical input devices satisfying certain conditions. Stated informally, the restrictions on the scenes are:

1.1) The scene consists entirely of relatively homogeneous regions bordered by edge lines.

1.2) The edge lines are all straight.

1.3) The objects of the scene may be recognized entirely from the edge lines.

The first of these restrictions is fundamental. It is virtually equivalent to the assumption that an intensity profile across a scene consists of a series of smooth curves bounded by discontinuities as in Figure 4.2.



HYPOTHETICAL INTENSITY PROFILE ACROSS A SCENE IN THE ABSENCE OF BLURRING AND NOISE

Figure 4.2

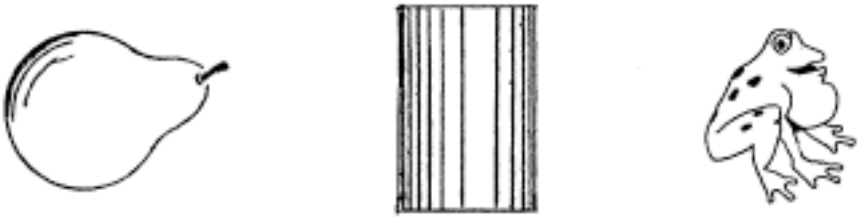
The second of these assumptions is essential if one is applying an edge test only to narrow rectangular regions, as is the case in the program under consideration. Of course one could develop a theory, analogous to the foregoing, for amalgamating local feature point information over curved bands, and thereby detect edge lines of arbitrary shape. However we have not done this; it almost certainly would require computation of an order of magnitude above that used in the current program.

The third restriction is a matter of convenience. Actually the program may be presented with scenes not conforming to this restriction without any impairment of its behavior. For example, it might be presented with an edge-on view of a cylinder, and successfully recognize the outline.

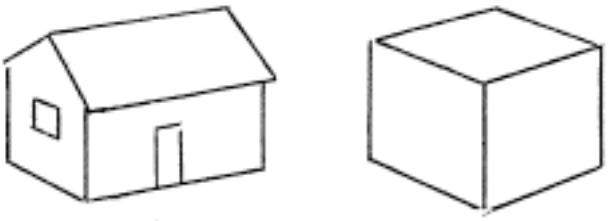
The roundness, though detectable by the vidissector from the intensity gradient across the surface, would not be detected by the present program.

These restrictions are summarized in Figure 4.3.

NOT APPROPRIATE:



APPROPRIATE:



ATTENTION WILL BE LIMITED TO A CERTAIN CLASS OF SCENES

Figure 4.3

The theory requires certain restrictions on the instrument which obtains the intensities from the real world. These are, in summary:

2.1) The extent to which the input device blurs the image received must be uniform over the whole field.

2.2) The extent to which the intensities are modulated by noise must be uniform over the field, and independent

of relative intensity.

Stated more specifically, the first assumption says that if I(x, y) represents the intensity of the scene over the field of view, D; then the intensity received by the optical input device in the absence of noise is:

$$I^*(x, y) = \int_D I(x-u, y-v) f(u, v)\, du\, dv$$

for some blurring function f(x, y).

These assumptions are summarized in Figure 4. 4:

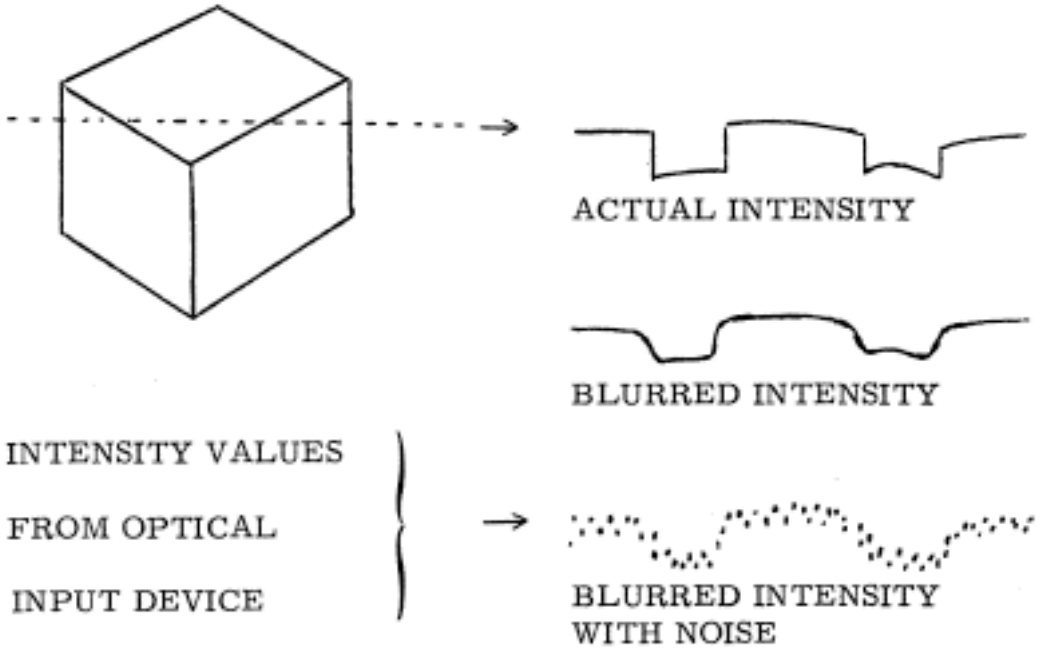
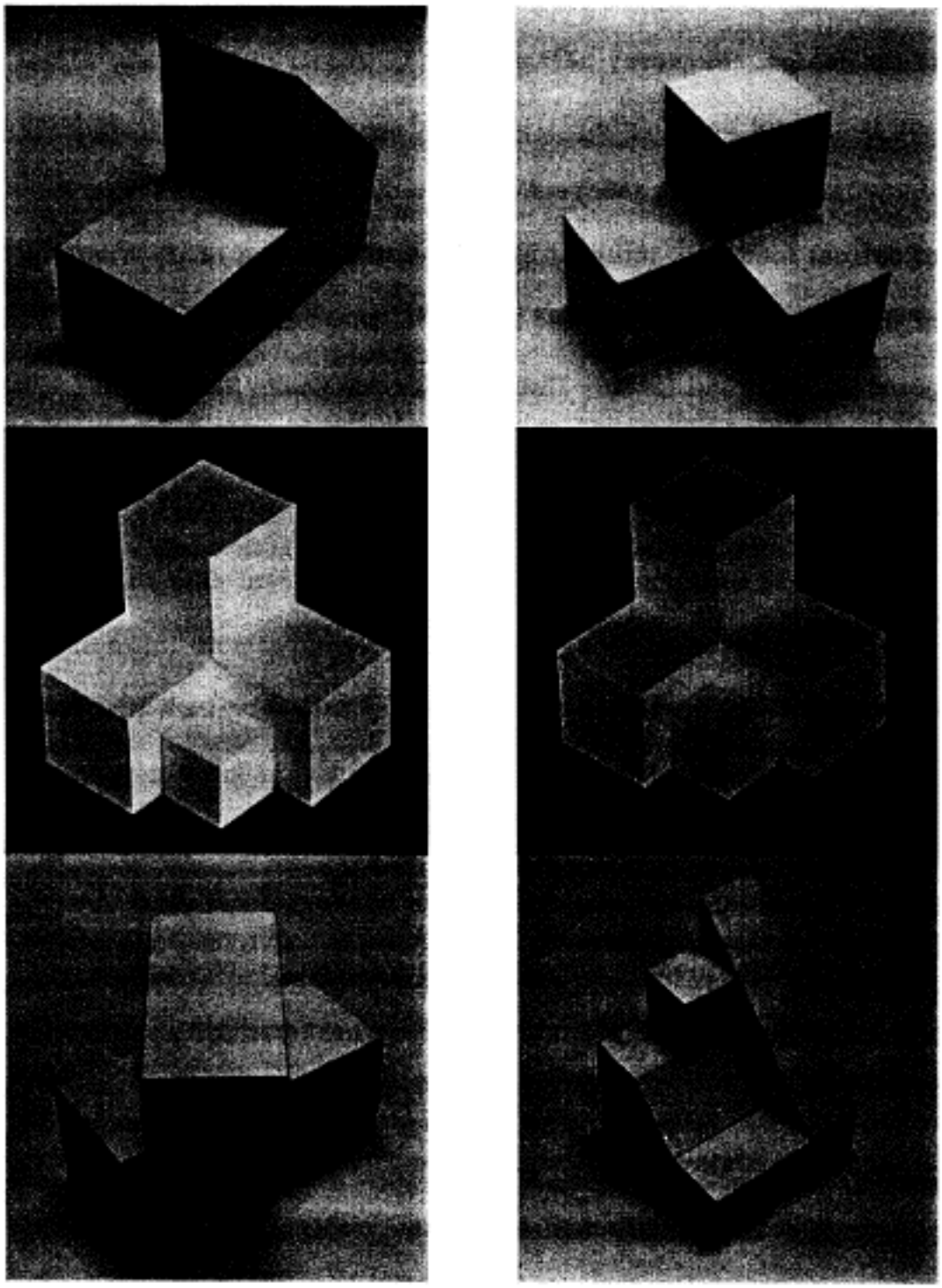


Figure 4. 4

Some scenes analyzed by the program are shown in Figure 4. 5.

It is easy to see that criteria 1. 2 and 1. 3 are satisfied by scenes of this sort. However, it is not readily apparent that criterion 1. 1 is satisfied. This follows from the fact that real lines such as highlights and cracks between cubes, in the absence of blurring and noise are never of zero width as stipulated in the definition of a line. However it is easy to see that they may be of nonzero width so long as they are sufficiently narrow to be indistinguishable from zero-width lines under the blurring function. This blurring function was discussed in Section II of chapter three. From its properties it is easy to see that blurred versions of lines of less than about five units width, where one unit is 1/2000 of the vidissector field width, are virtually indistinguishable from each other and from blurred versions of lines of negligible width. Criterion 1. 1 has thus been satisfied for these figures by keeping the apparent widths of the edge lines below the aforementioned five units.

Criterion 2. 2 has been met by the optical input device used, an Information International vidissector, by the nature



EXAMPLES OF SCENES ANALYZED

Figure 4.5

of its design. However certain care is required to satisfy 2.1, and it can not be completely satisfied by the instrument used.

The blurring inherent in the vidissector is a function of optical focus, of internal electronic focus, and of the size of the internal aperture used to collect intensity data from a point in the visual field. The latter two matters have been discussed at length by Horn (Horn 1969), and in the literature on the instrument itself. Suffice it to say here that there is a more or less uniform degree of blurring inherent as a result of these two factors, though a certain amount of non-uniformity is detectable by careful measurements. It is fortunate that the order of magnitude of these effects is somewhat greater than the optical blurring of the system when it is focused to minimize this blurring. By keeping the objects in the field of view in a sufficiently flat plane, it is possible to have the non-optical blurring effects dominate the optical effects even for portions of the field which are relatively out of focus. This, to a first order of approximation, insures a uniformity of over-all blurring, despite the fact that the purely optical blurring is quite non-uniform. The criteria for "a sufficiently flat plane"

for the image is given by laws of optics. From the well known relation that:

$$1/F = 1/a + 1/b$$

Where:

a = object distance

b = image distance

F = focal length of the lens,

it follows that:

$$da/db = - a^2/b^2.$$

Hence for a given displacement Δa of the object from perfect focus, the image displaces by an amount $-(\Delta a)(b^2/a^2)$. This quantity, divided by the focal ratio f, gives the amount an image is blurred by displacing the object a distance Δa from perfect focus. Thus the allowable scene depth for a particular maximum allowable amount of blurring depends on the square of the scene-distance image-distance ratio; and inversely on the f-ratio. If a particular object or arrangement of objects increases in absolute size, and its distance from the lens increases proportionately (and if the object distance is large relative to the lens-image distance), its image remains approximately constant in size, but the depth of the object

remaining in good focus increases linearly relative to the size of the object. It is thus possible, for a given arrangement of objects to keep the blurring over the whole field down below any pre-assigned limit by making the f-ratio sufficiently large, and by making the objects sufficiently large. In practice, due to the low sensitivity of the vidissector it was often impossible to insure adequate depth of field by the use of a suitably small aperture alone. In particular, in the case of the objects illustrated in Figure 4. 5, it was necessary to make an effort to make the objects as large as possible to satisfy criterion 2. 1.

III. THE DETERMINATION OF FEATURE POINTS FROM INTENSITY INFORMATION

In this section we will be concerned with a discussion of a procedure for extracting information from a set of intensities taken over a visual field, as a first step in the determination of the locations of lines in the field. The relevant theory, developed in the previous chapter, will be reviewed; and an algorithm for reducing intensity information to "feature points" will be discussed.

In applying the line predicate to various regions on the visual field, we shall obtain intensities along scan lines

perpendicular to the axes of the vidissector. In particular we shall use intensities obtained from scans parallel to the x-axis in applying the line predicate to long rectangular regions oriented within $\pm$ 45° of the y-axis. Similarly, whenever a line predicate is applied to a region within $\pm$ 45° of the x-axis, intensities from scans parallel to the y-axis will be used. We may thus henceforth without loss of generality assume that we are dealing with regions oriented within $\pm$ 45° of the y-axis, and with intensities obtained from scans taken parallel to the x-axis. A scan shall consist of 500 intensities taken at intervals of 2 units along a line, where, as in the previous chapter, one unit is 1/2000 of the full field width. The line is centered laterally in the field. Fifty scans will be taken at intervals of 20 units, with the 25-th scan half way up the field of view. The set of scans will thus encompass a square area exactly centered in the field of view, whose edge length is 1000 units, or one half the edge length of the entire field of view. The reason for not using the entire field is simply that there are various physical problems at the edges, one of which involves an aperture which prevents obtaining intensity values at the corners.

The central square used is approximately the largest useful square region in the field of view. For purposes of visualizing the following explanations, it may be assumed that an entire set of 2500 intensities is obtained at once, prior to any further processing. This is not entirely true, but may as well be considered to be the case.

We will want to be able to apply the line predicate to any of approximately fifteen million rectangular regions in the field of view. This figure is obtained from the fact that we are considering regions extending between pairs of scans, and the latter number $50 \cdot 49/2 = 1250$; we are considering regions at any of fifty orientations; and we are considering regions with any of 500 different intercepts. In fact we shall apply a modified version of the predicate to all of a subset of these, namely those which traverse the field entirely, which number $50 \cdot 500 = 25,000$. This will be discussed in Section IV. Also the predicate will be applied to shorter subregions of those among the latter which appear to contain a line somewhere along their lengths. Further the predicate will be applied to regions proposed on the basis of lines already found. This will

be discussed in Sections V, VI and VII.

The application of Q* to a region on the field of view consists in computing the value given by (3.13):

$$Q^*(J_j^n) = \frac{CL(J_j^n) + CE(J_j^n)}{CL(J_j^n) + CE(J_j^n) + CH(J_j^n)},$$

where $CL(J_j^n)$ is given by (3.29), and $CE(J_j^n)$ and $CH(J_j^n)$ are given by similar formulas. A region whose maximum minus minimum y-distance is $\underline{s}$ units intersects with $\|s/20\|$ scan lines, where $\|i\|$ means the closest integer to $\underline{i}$. Instead of assuming that a region is of a particular fixed width measured perpendicularly to its length, we shall assume it has a fixed x-direction width of 50 units. Consequently the intensities along a particular scan which lie within a rectangular region will be exactly 25 in number, since the separation of intensity points along a scan is two units. We shall also assume that the length of a sub-region, instead of being a fixed quantity as measured along the axis of a region, has a fixed y-direction length of 20 units. Also sub-regions will be taken to be vertically centered on a scan line. This arrangement is

diagrammed in Figure 4. 6.

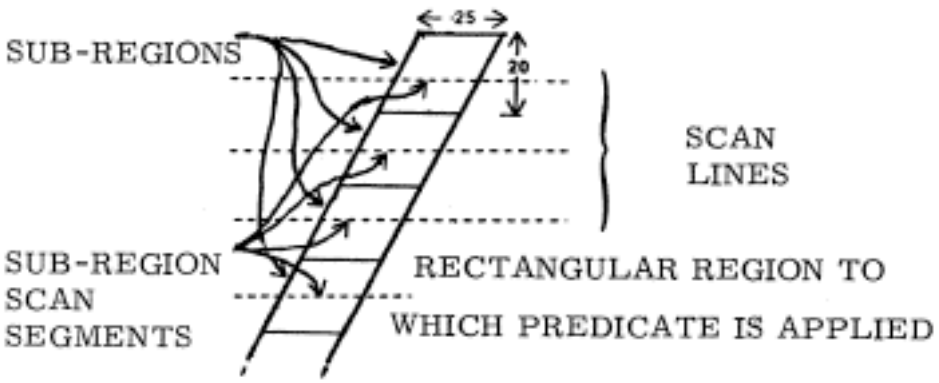


Figure 4. 6

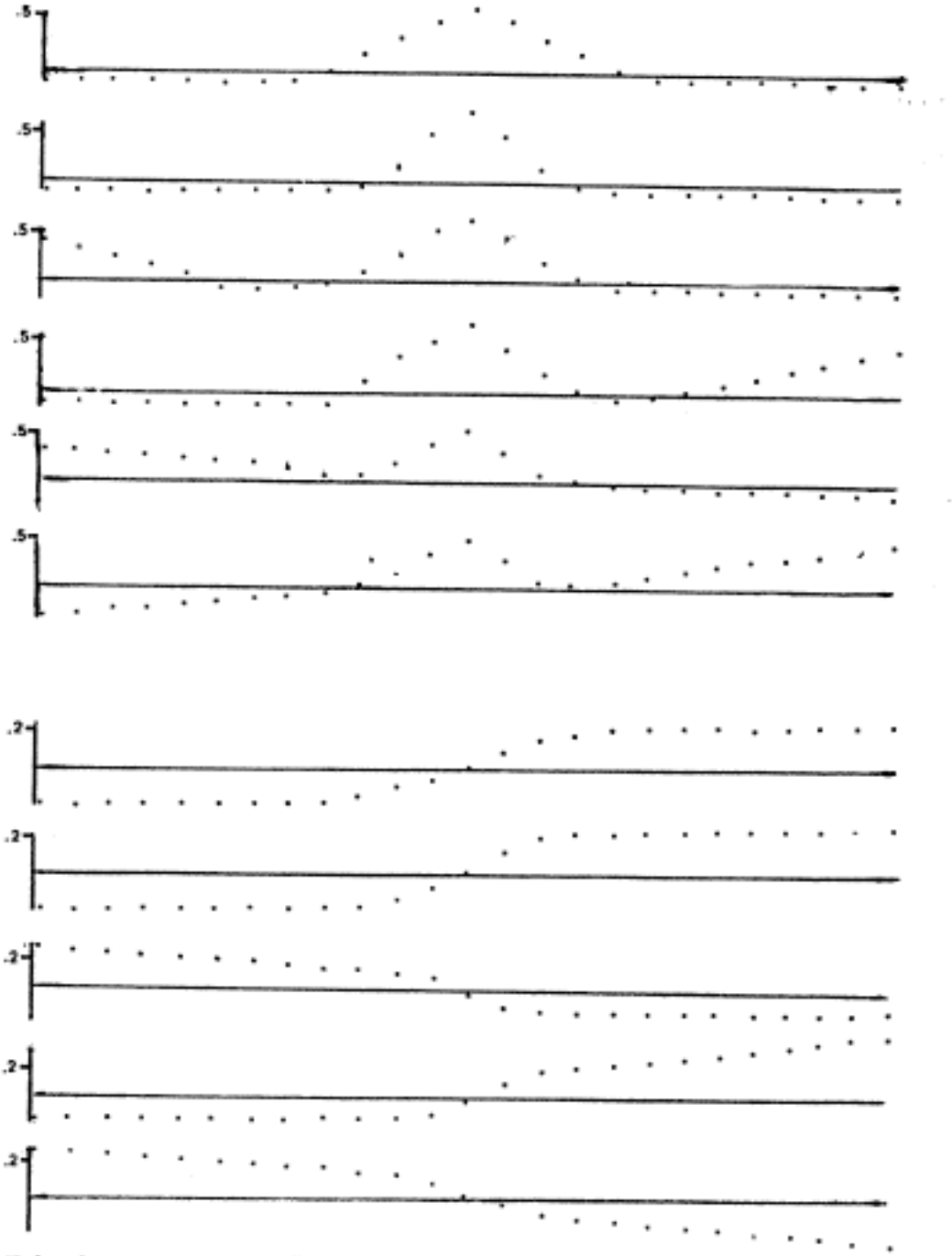
We may thus identify a segment of a scan enclosed within a region with a particular sub-region. The advantages of this sort of geometry were discussed at the end of the previous chapter. We shall use a specific term for the segment of a scan which we are associating with a particular sub-region, namely a sub-region scan segment. In summary, a region of length a consists of $\|a/20\| = m$ sub-regions; and the value of Q* is computed from the m sub-region scan segments by calculating values of CL, CE, and CH according to formulas of the form given by (3. 29).

Evidently the intensities of a particular sub-region scan segment enter into the computation of Q* for several thousand regions enclosing the segment. Examining (3. 29)

it is apparent that, e. g., in computing CL, it is necessary to compute a value $R_{j,k}$ from the intensities in this segment, and also a value $a^*_{j,k}$; and that CL is determined entirely from these values. It is not hard to see that these values are the same for a particular segment, when the enclosing regions are all at the same inclination. Further, it was pointed out in some detail in the last chapter that these values for a particular segment do not vary a great deal with orientation of the enclosing rectangular region assuming, as we are, that the regions remain within $\pm$ 45° of a normal to the scan lines. Thus, as before, only values of $R_{j,k}$, etc., for two region orientations, 10° and 30°, need be computed. This is done, as in the previous chapter, by using two paradigm profiles of different widths. Finally, we may recall that at the end of the previous chapter the procedure of computing a single value of R, R', etc., for a particular segment and orientation was generalized to the computation of several, corresponding to different possibilities of the shape of the intensity profiles toward the edge of the region. Consequently, for a given sub-region scan segment, we will compute exactly the values $R_{j,k,1}, \cdots, R_{j,k,6}$, $a^*_{j,k,1}, \cdots, a^*_{j,k,6}, R'_{j,k,1}, \cdots, R'_{j,k,5}, a'^*_{j,k,1}, \cdots,$

$a'^*_{j,k,5}$, and also $R''_{j,k}$ and $a''^*_{j,k}$. The six R's and a*'s are computed as in (3.25) and (3.23) from six different peaked profiles $(U_1^1, \ldots U_1^{25}), \ldots, (U_6^1, \ldots, U_6^{25})$ of various widths and with various combinations of gradients towards the ends. Similarly the R''s and a'*'s are computed from five different cliff-like profiles. Clearly, no such multiplicity is necessary in the case of R'' and a''*, which corresponds to a homogeneous region. These paradigm profiles, 25-tuples of values $(U_1^1, \ldots, U_1^{25})$, etc., from which the R's, a*'s, R''s, a'*'s, and R'' and a''* are computed, are illustrated in Figure 4.7.

Since values of $R_{j,k,1}$, $a^*_{j,k,1}$, $\ldots$, $R'_{j,k,5}$, $a'^*_{j,k,5}$, $R''_{j,k}$, and $a''^*_{j,k}$ will be required for every sub-region scan segment, they are all computed. Since there are 475·50 sub-region scan segments, and 24 values computed per segment, this is a matter of 570,000 intermediate values, an order of magnitude more than the 25,000 intensity values from which the values were computed. However, the derived set of values admits of considerable compactification. First, for a particular sub-region scan segment, it is necessary to retain only the minimum of all the R's and R''s, together with the corresponding value of a* or a'*. Secondly, consider the



PARADIGM INTENSITY PROFILES USED IN COMPUTING Q'**

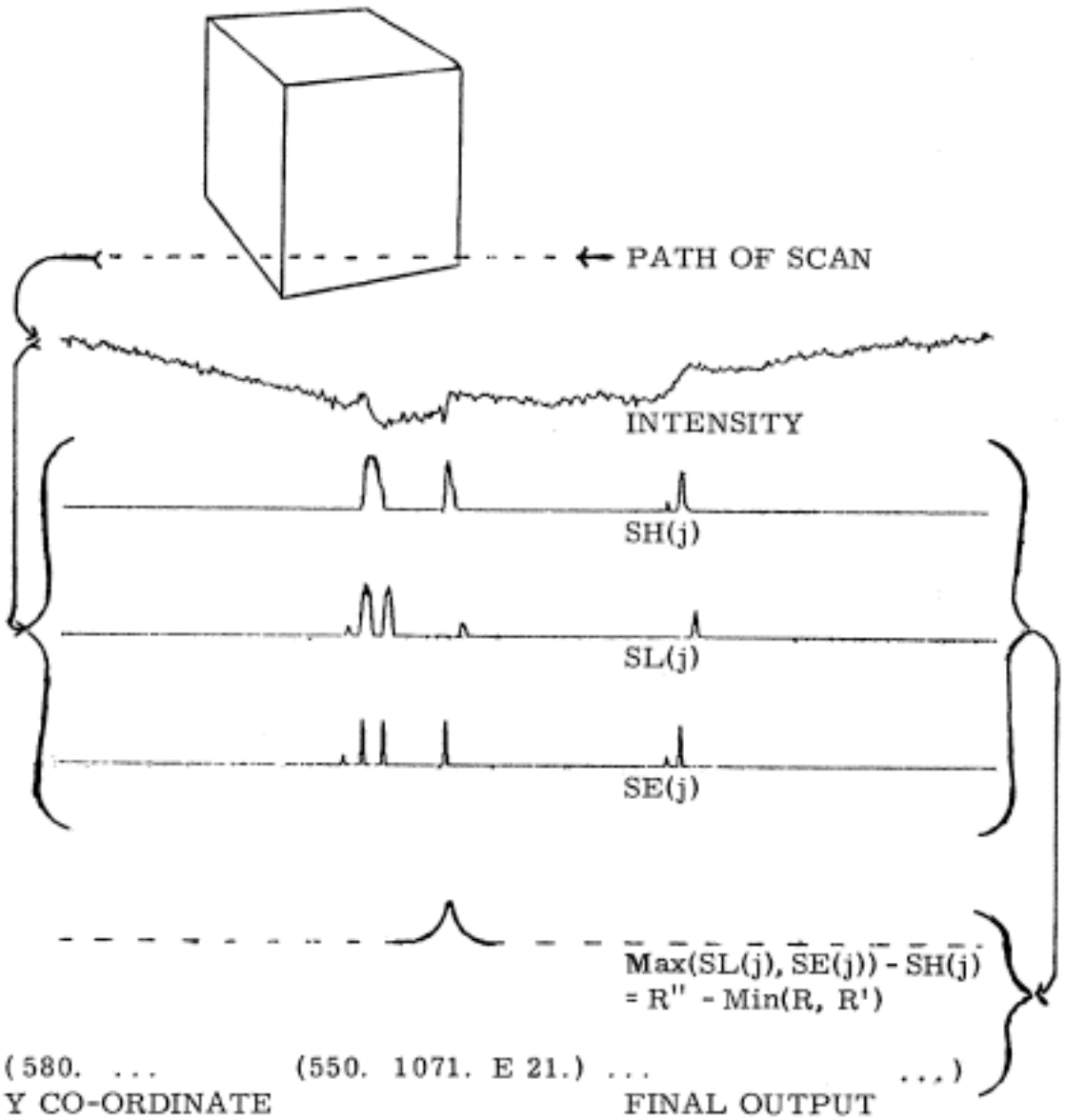
Figure 4.7

values of R'' minus this minimal R or R' for the 475 successive sub-region scan segments along a particular scan. Then there exists a threshold for these differences such that most values lie below the threshold and indicate with virtual certainty that a line does not traverse a region containing that segment. Such values may be omitted, thereby indicating that they fall within this category. Also, only the local maxima of the values above threshold need be considered. Consequently an entire collection of 475 sets of values may be reduced to somewhere around ten sets. These latter sets need only consist of 4 values: the value of R'' minus the minimal R or R' which is locally maximal; the value of a* or a'* corresponding to the minimal value of R or R'; an indicator as to which R or R' was minimal; and the x co-ordinate of the point in question. Such a four-tuple, together with the y co-ordinate of the scan from which it was derived, will be termed a feature point. We will not justify in detail here why the 570,000 values can be reduced to approximately 500 feature points. However we shall point out that the subsequent use made of derived information in the computation of Q* and similar values uses no more than the reduced feature point data; that many of the

values can be considered to be in a single null category; and that the remaining values have sufficient regularity to be specifiable in terms of their local maxima only.

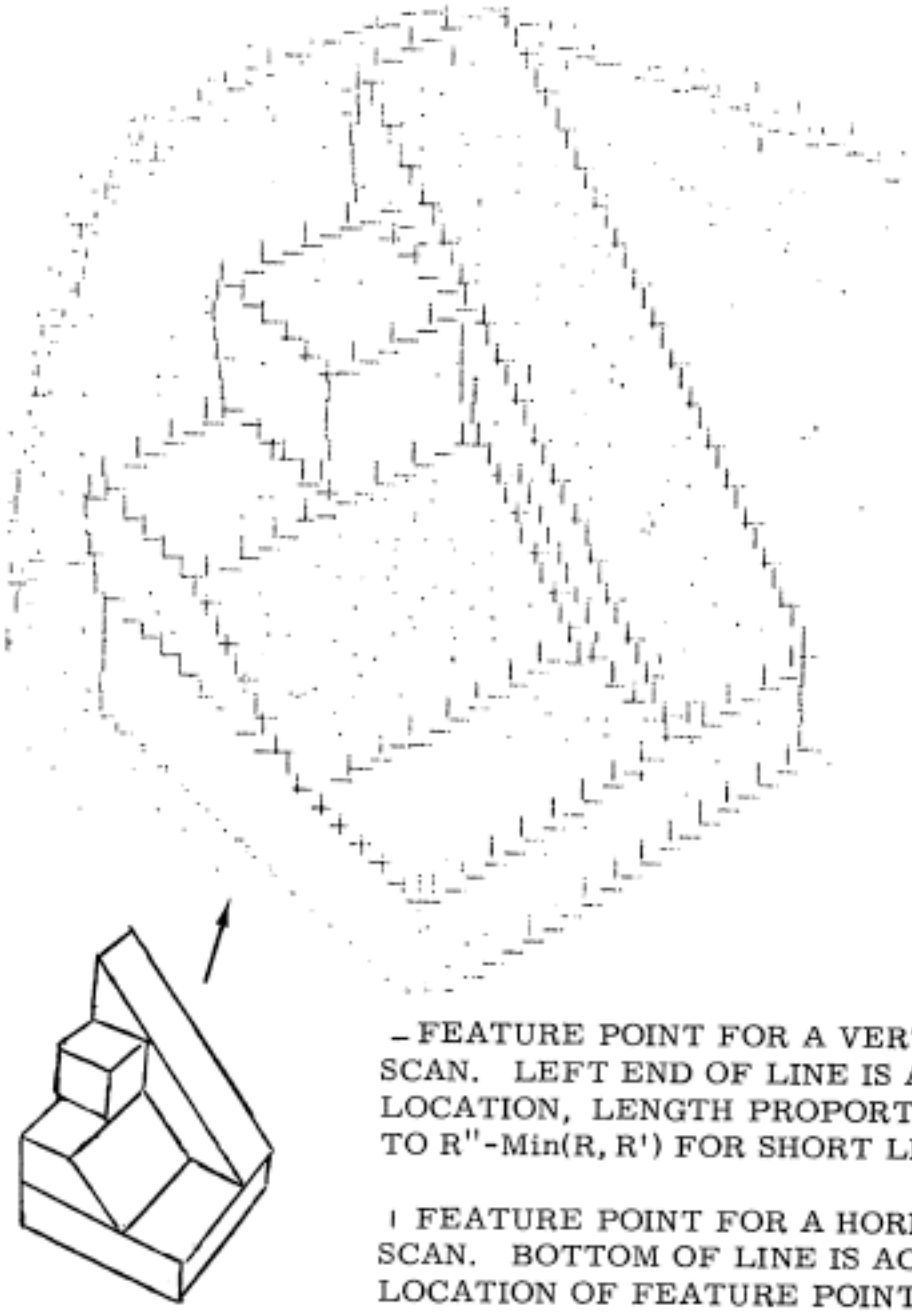
The program (GETFP X S) obtains intensities for a single scan and performs the calculations sketched above. The second argument is 1 or 0 depending on whether a scan is to be made vertically or horizontally respectively. The value of X is the appropriate ordinate or abcissa on a scale from 0 to 500. Five hundred intensities are taken across the field along the line defined by the arguments and are processed according to the algorithm sketched above. The output is a series of 4-tuples as described above. A schematic diagram of this process appears in Figure 4. 8.

In the line finding program which has been developed, the sub-program GETFP is called 50 times for vertical scans and 50 times for horizontal scans. This procedure is illustrated in Figure 4. 9. In this figure, the short lines indicate the locations of maxima previously described. Their lengths are proportional to the corresponding values of $R'' - \text{Min}(R, R')$ for small values of this difference. For values of this difference above a certain threshold, a constant length of line is used



SIMPLIFIED VERSION OF THE PROCESS OF EXTRACTING FEATURE POINTS FOR A SINGLE SCAN ACROSS THE FIELD

Figure 4.8



– FEATURE POINT FOR A VERTICAL SCAN. LEFT END OF LINE IS ACTUAL LOCATION, LENGTH PROPORTIONAL TO R''-Min(R, R') FOR SHORT LINES.

| FEATURE POINT FOR A HORIZONTAL SCAN. BOTTOM OF LINE IS ACTUAL LOCATION OF FEATURE POINT.

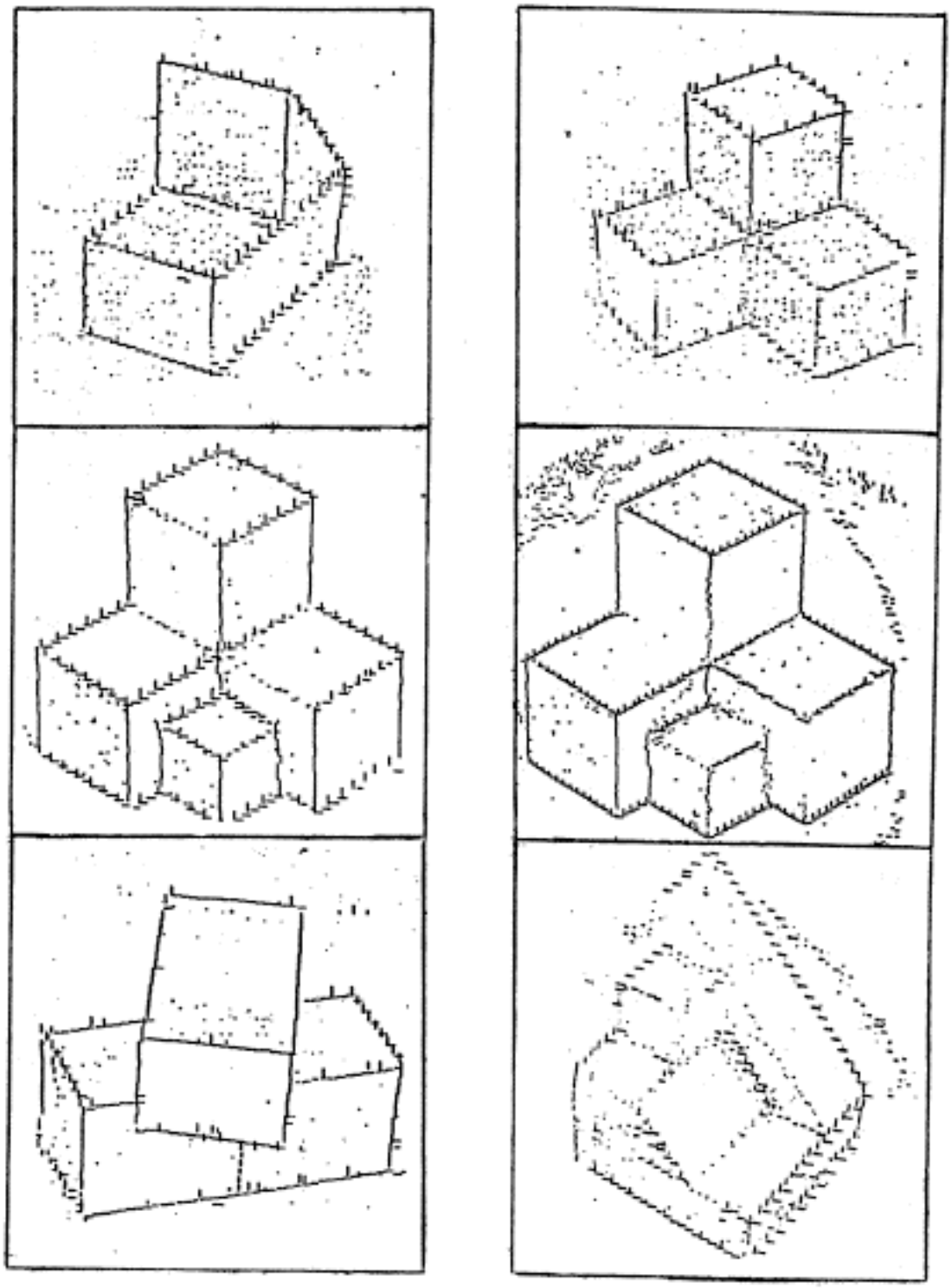
FEATURE POINTS FROM 50 VERTICAL AND 50 HORIZONTAL SCANS

Figure 4. 9

corresponding to this threshold value. These points will be termed feature points. Rasters of feature points corresponding to the illustrations in Figure 4. 5 are given in Figure 4. 10.

IV. A GLOBAL PROCEDURE FOR EXTRACTING LINES FROM ARRAYS OF FEATURE POINTS

If one were to examine the feature point rasters of Figure 4. 10, one would be able to pick out almost all the lines in the figures they represent. The perception of some of the lines may depend on complex global perception processes, for example, the use of obvious lines to provide clues as to the existence of less obvious ones. A computer approximation to this process will be the subject of subsequent sections. On the other hand, some of the lines are obvious without either reference to other parts of the figure or a priori knowledge of what constitutes a plausible figure. This is because the feature points are both strong (represented in 4.10 by long lines), and they line up very exactly. This section describes a program, LINES, which takes advantage of the obviousness of certain lines to find them in an array of feature points, by a rather simple and not overly lengthy procedure.



FEATURE POINTS EXTRACTED FROM THE SCENES IN FIG. 4.5

Figure 4.10

The problem of locating lines by accumulating locally obtained information has been approached by line following procedures. (Roberts) This often runs into the problem that the line follower gets lost, since, in effect, it is making a series of decisions (whether or not to continue the line) on the basis of local considerations. This problem becomes greater as the number of lines in the scene increases. Line followers are also subject to the problem of never having "attached" themselves to certain lines in the first place. It is thus desirable to use a more exhaustive and global procedure, for example, covering the feature point raster with a very large number of narrow rectangles, and applying a thresholded predicate to the feature points within each rectangle. The apparent tremendous cost of such a procedure can be greatly reduced by observing that if a certain narrow rectangle contains evidence of a line, then so does a rectangle with the same orientation but extending across the entire field. The converse of this statement is not true, since a rectangle across the whole field may contain a large amount of line evidence in the form of scattered noise and spurious values due to feature points along lines intersecting the rectangular region at some angle.

However, the testing of narrow rectangles across the whole field does provide a sort of screening procedure. Any relatively strong lines will fall within at least one of these bands, and thus be detected. On the other hand, it is necessary to further analyze the contents of one of these bands which has a large amount of line evidence in it in order to determine both if there really is a line within the band; and if so, where it is.

The procedure for assessing the total amount of evidence for the existence of a line within a band is based on the idea of projecting a two dimensional array into a one dimensional line. For example, if an array of points (x_i, y_i) is given, and it is desired to determine the locations (x-intercepts) of vertical bands containing a large number of points, then it is merely necessary to histogram the x-co-ordinates of all the points. The resulting histogram has large values for x-intercept values of bands containing a large number of points. By this procedure, all vertical bands are investigated in parallel. The procedure is simply modified for arbitrary orientations, by histograming numbers of points versus a quantity $x_i + ay_i$, where a depends on the orientation under consideration. This corresponds to projecting the raster of feature points onto

the x-axis at some angle.

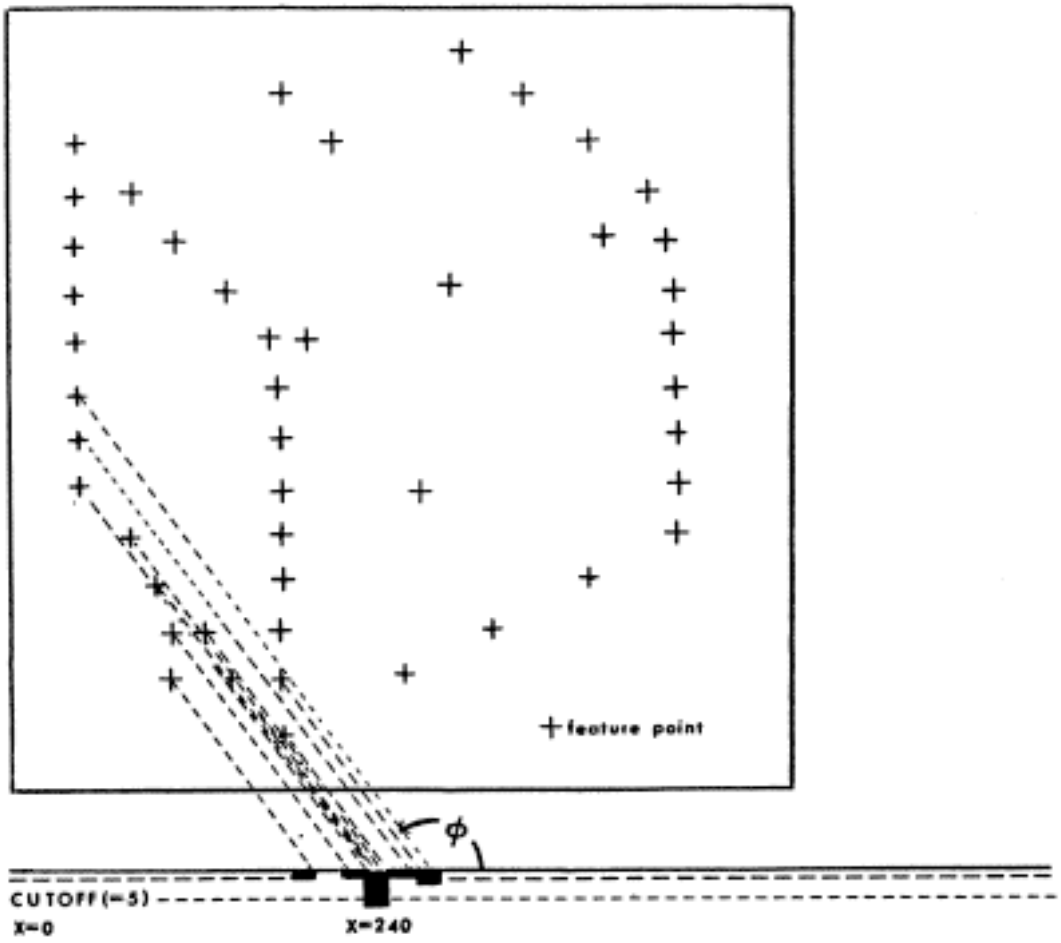
The program (PROJ X SL) carries out this projection procedure on an array of feature points X, projecting at an angle SL. The main program calling PROJ uses 100 values of SL. For 50 values, the feature points obtained from horizontal scans (parallel to the x-axis) are used. The values of SL in this case correspond to orientations between $\pm$ 45 degrees with the vertical (the y-axis). The remaining 50 projections use the feature points from the vertical scans, and involve angles between $\pm$ 45 degrees with the horizontal. The values histogramed are not simply numbers of points, but sums of integer weights attached to the various feature points. These integers have values from 1 to 4, and depend monotonically on the value of R'' - Min(R, R'), the first feature point value, for a particular point. Thus, for a particular band, it is not simply the quantity of feature points within it that is considered, but something approximately equal to the sum of the values of R'' - Min(R, R') for points within it. This gives extra weight to distinct lines, and may be shown to approximate Q* in some sense. The value returned by (PROJ X SL) is the set of x-intercepts of centerlines of bands of slope SL which have a large amount of line evidence

in them. The algorithm employed by PROJ is diagrammed in Figure 4.11.

The set of outputs of (PROJ X SL) must be analyzed to extract the actual lines in the scene. One problem which must be taken into account arises from the fact that a particular set of feature points may give rise to an intercept value in the output of several applications of PROJ, for successive values of SL. This is a result of the fact that a given set of feature points which fall in a line along the visual field are contained within several rectangles of almost identical slopes. Another problem is that the output of PROJ contains a large number of false-positive values due to spurious effects described previously.

The program LINES, which calls PROJ for various values of the arguments, takes these factors into account by a complex process of:

1) retrieving the feature points in sets from areas suggested by the output of PROJ,
2) eliminating sets which do not actually constitute lines,
3) eliminating spurious points from sets which do represent lines,



OUTPUT:
(. . . . . (240. 6.) . . . . .)

SIMPLIFIED DIAGRAM OF (PROJ X SL) APPLIED TO
X, A SIMULATED FEATURE POINT RASTER, AND
SL, A SLOPE CORRESPONDING TO ϕ IN THE DIAGRAM.

Figure 4.11

4) sorting by the apparent strength of lines,

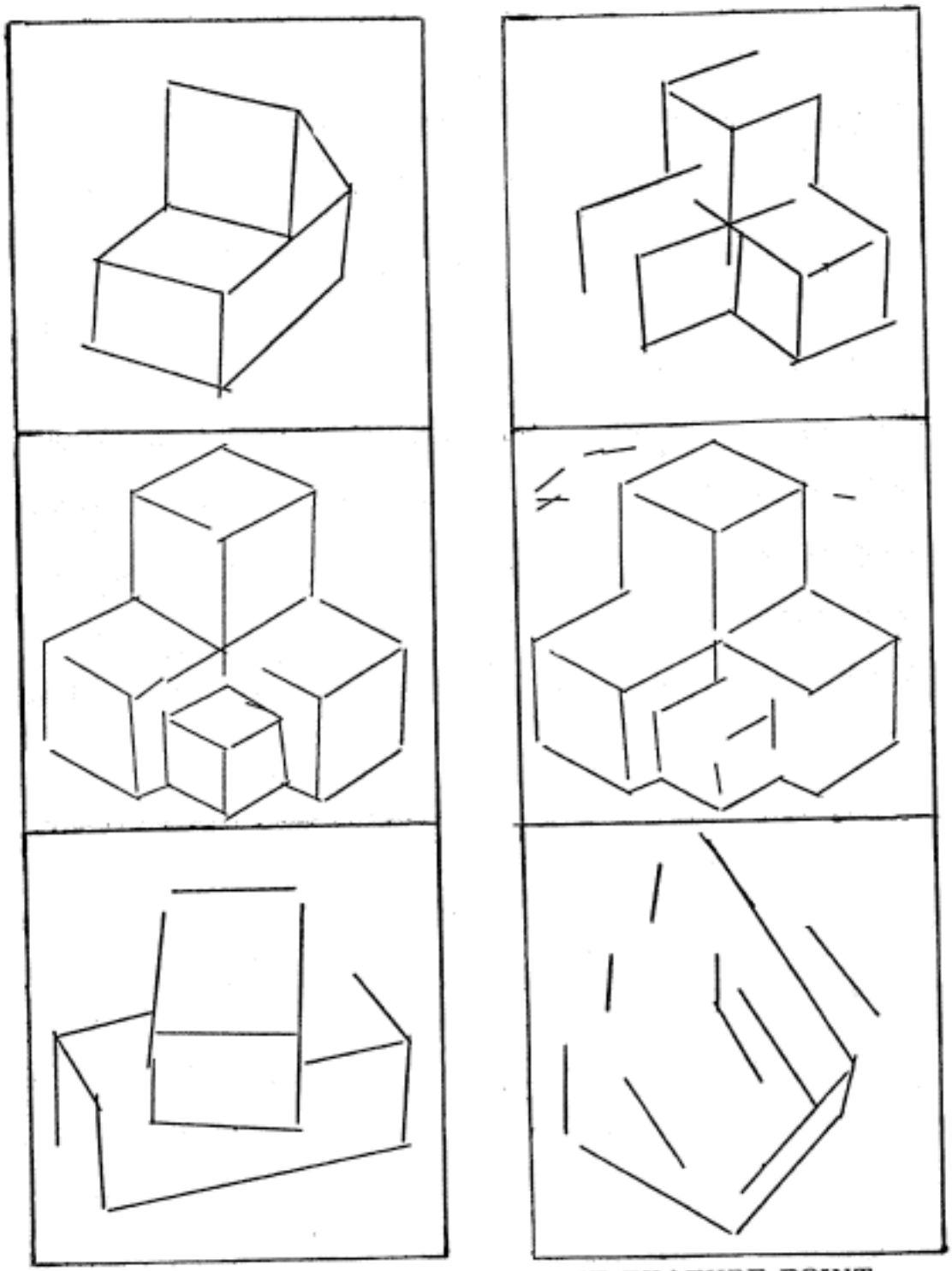
5) eliminating redundant sets of points by considering the set in the order imposed by step 4 and eliminating sets which are highly redundant with (more distinct) sets which had previously been considered,

6) fitting lines to the resulting sets of points.

The output of (LINES X) is a set of lines retrieved from the set of feature points X. LINES is applied once to the feature points from the vertical scan; and once to the feature points from the horizontal scan. The result of applying LINES to the sets of feature points illustrated in Figure 4.10 is shown in Figure 4.12.

V. LINKING LINES TO FORM PARTIAL FIGURES

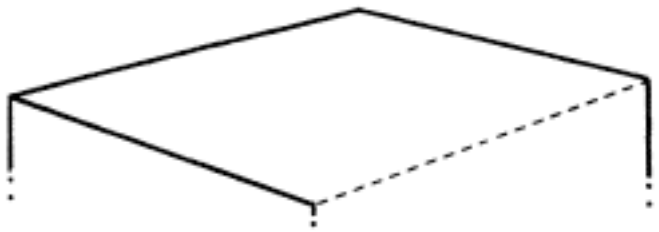
Having located a certain number of lines in a scene by the methods of the previous section, a logical next step is to join them together and form figures. This may be viewed as a terminal step, provided all lines have been located by the projection procedure. This would not, in general, be the case; and the joining procedure described in this section is actually a first step in the location of the remaining lines of the figure.



OUTPUT OF LINES, APPLIED TO THE FEATURE POINT RASTERS IN FIGURE 4.10

Figure 4.12

We shall assume that the output of LINES, when applied to a set of feature points, does not include all the lines of interest in the original scene. This is, in fact, the case for 3 out of 4 scenes in Figure 4.12. It is a reasonable hypothesis that the lines found, together with some assumptions about the nature of the scenes examined, provide some clues as to the locations of the remaining lines. It appears not to be the lines individually which suggest the locations of other lines, but rather collections of lines which form a figure which is incomplete relative to a model of how the actual figures appear. An example is shown in Figure 4.13.



THE SOLID LINES HAVE BEEN IDENTIFIED,
THE DOTTED LINE IS SUGGESTED BY THEM.

Figure 4.13

It would be desirable to join the lines into groups which would suggest the possible location of other lines. A line verifier could then be applied to the locations suggested. This procedure requires that all locations of missing lines ultimately be proposed. Otherwise, some lines would never be found by

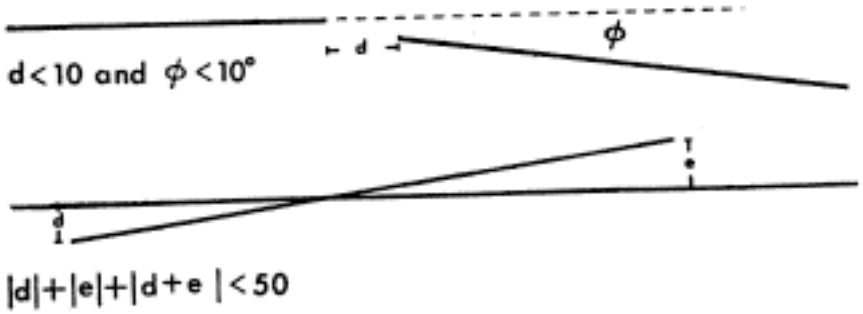
either procedure. It is thus important that the whole mechanism for line proposing be liberal in proposing lines. Any excess lines proposed will almost certainly be rejected by a suitably powerful verifier.

Prior to the application of the linking procedure, a merging procedure is applied to the lines. This is necessary to eliminate redundant lines from two sources:

1) Lines which are at approximately at a 45-degree angle with the vertical, and appear in the output of LINES applied to both the vertical and the horizontal scans.

2) Single lines which appear, for one reason or another, as a pair of contiguous lines in the output of LINES.

The criterion for merging lines is given in Figure 4.14.



CRITERIA FOR MERGING LINES

Figure 4.14

If two lines are to be merged by this criterion, the pair is replaced by a single line which best fits the endpoints of the two original lines.

The joining procedure consists in applying directed links from one line to another. There are four possible types:

1) Regular links, denoted by the symbol T. If any part of line B lies within a neighborhood of radius 10 units around A, then A is given a directed T-link to B. This is illustrated in Figure 4.15.

2) Extension links, denoted by the symbol E. If any part of line B lies in a particular region near the end of A, or if the end of line B lies within another more distant region extending away from the end of line A, then an E-link from A to B is established. These regions are illustrated in Figure 4.15. The motivation for this type of link is that possibly a segment of A extending to B, or a longer segment extending to the end of B, has been omitted. The reason for postulating a rather long extension to the end of another line is that a line's end, lining up with a given line, provides strong evidence that the original line may actually extend to that end.

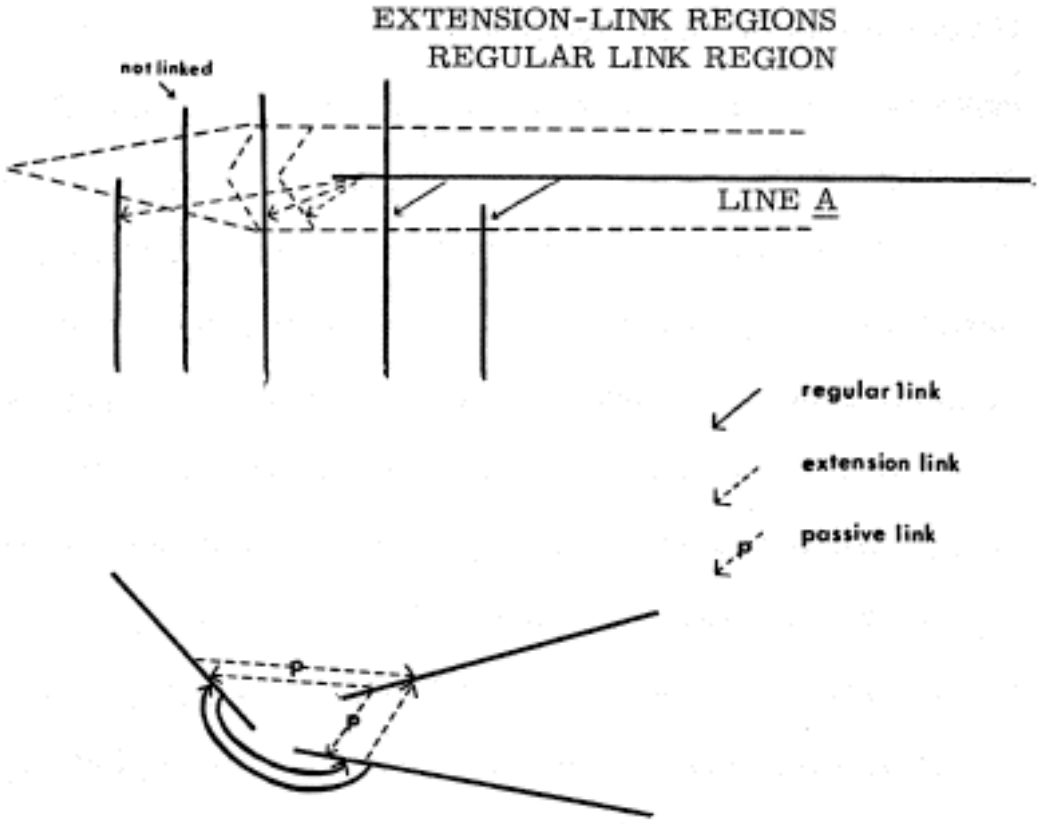


Figure 4.15

3) Passive regular links, denoted by PT. If line B links to line A by a T-link, but A does not link to B by either type of link, then a PT-link is established from A to B.

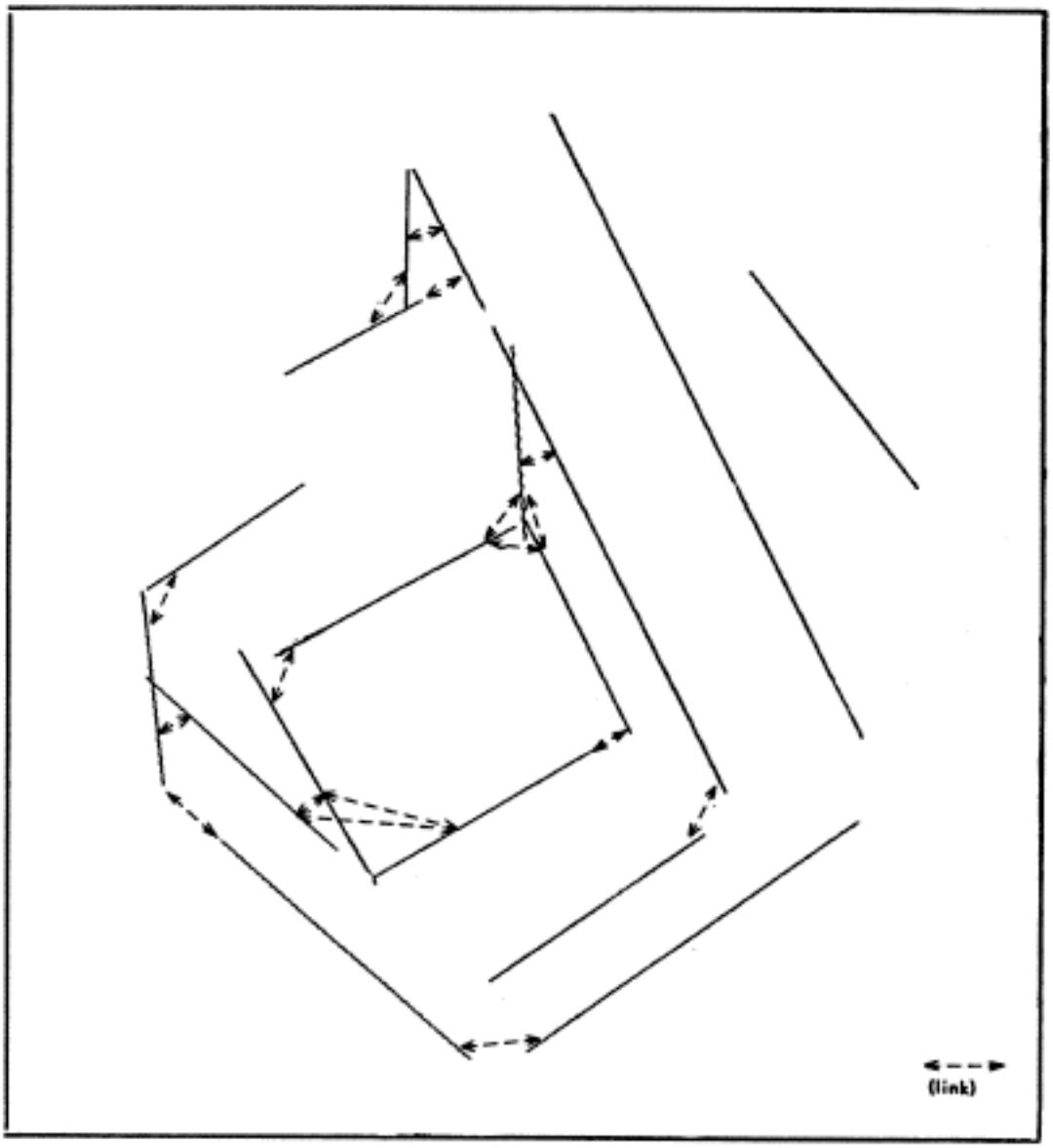
4) Passive extension link, denoted by PE. This is similar to a PT-link, but is applied if B links to A by an E-link.

An illustration of the links applied to one of the sets of lines from Figure 4.12 is given in Figure 4.16. It should be emphasized that, in accordance with the necessary liberal policy underlying the proposing of lines, the lines are over-linked. These links are not intended to be those applied when all the lines have been located.

VI. HEURISTIC LINE PROPOSING

This section describes the program (PROPOS X), which proposes additional lines from a possibly incomplete set X of lines from a scene. Each line is considered in turn, and additional lines suggested by it, and by lines linked to it, are proposed. A verification program, to be explained in the last section, is applied to each proposed line. If the verifier claims that the line actually exists, it is added to the list of lines, together with links between it and other lines; and it may be used subsequently in proposing further lines. When no more lines can be proposed from the (possibly augmented) set of lines, PROPOS returns the entire set.

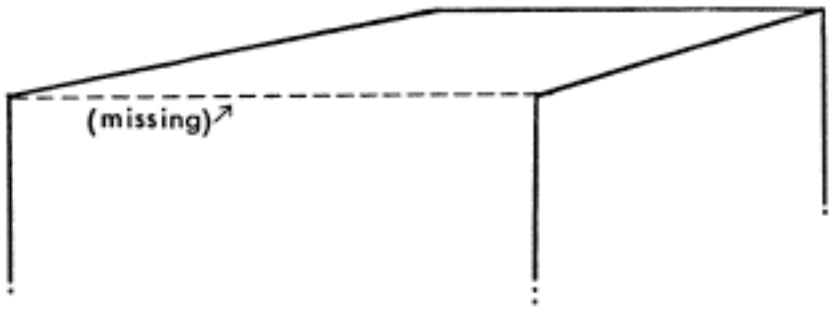
The proposing procedure is based on the premise that if a line is missing from a set S, then at least one line having a



LINKS BETWEEN MEMBERS OF THE OUTPUT OF LINES

Figure 4.16

common end-point with the missing line is in the set X, or can be successfully proposed from X. This is illustrated in Figure 4.17. It is further assumed that vertices have no more



IT IS ASSUMED THAT IF A LINE OF THE ORIGINAL FIGURE WAS NOT AMONG THOSE OUTPUT BY LINES, THEN AT LEAST ONE LINE CONNECTING TO IT WAS.

Figure 4.17

than three or possibly four lines radiating from them. It follows that lines should be proposed from the ends of lines already found, provided that there are not already many lines incident with a particular end.

It is in general not entirely clear where a line generated by LINES actually terminates. This is partially due to the rough "mesh" of the scanning grid, and partly due to the fact that the feature point generator gets confused in the neighborhood of a vertex. In order to propose lines from the ends of known

lines, it is necessary to make a reasonable estimate as to exactly where the ends lie. The first step in proposing lines from a given line is to dispose of the ends in the following manner:

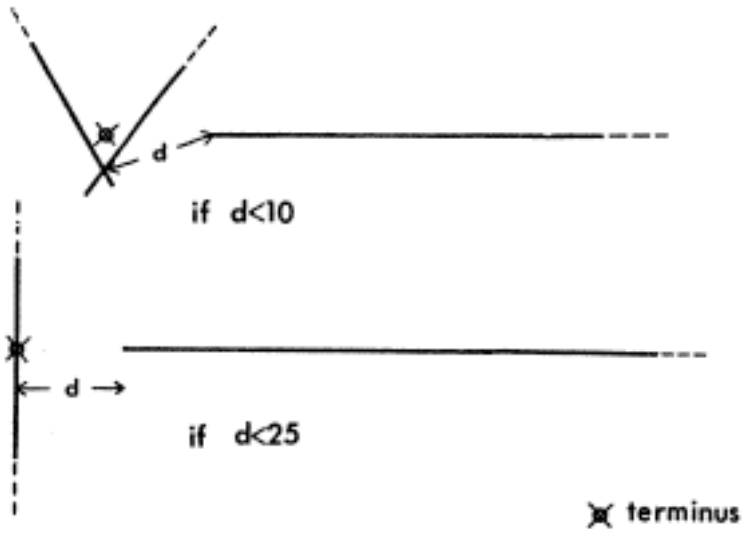
1) If two lines intersect within ten units of the end of a given line, the point on the line (possibly extended) closest to the intersection is taken as the end or "terminus".

2) If a terminus is not determinable by criterion 1, but a line intersects the given line within 25 units of the end, this intersection is taken as the terminus.

3) If no terminus can be found by 1 or 2, that end of the line is said to be of "indeterminate terminus."

The determination of termini is diagrammed in Figure 4.18.

Another assumption made by PROPOS, which is independent of the first, is that the opposite sides of faces of the figures are approximately parallel. If it is desired to propose a line extending, e.g., downward from the left terminus of a horizontal line, the direction of the proposed line is then the same as that of another "secondary" line extending downward from the given line. In particular, this



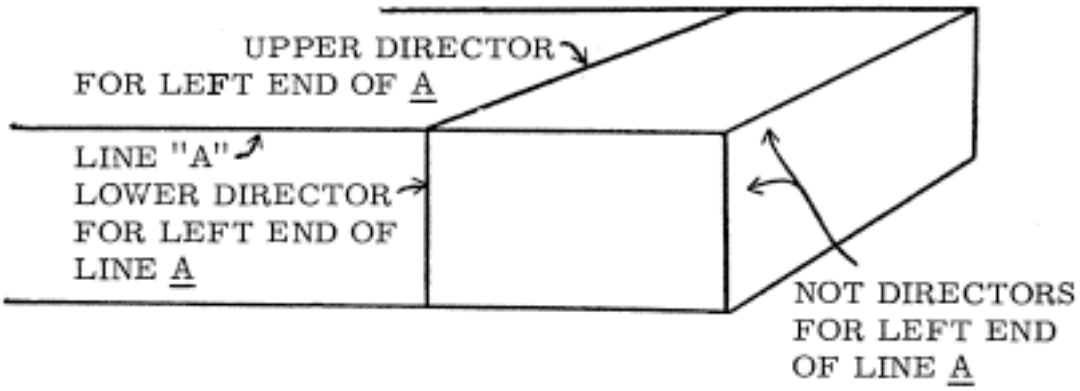
THE "TERMINI" OF THE HORIZONTAL LINES ARE AS INDICATED.
LINES PROPOSED FROM THE HORIZONTAL LINES EXTEND FROM THE TERMINI.

Figure 4.18

"secondary" line should be the first line "to the right" of the left terminus, as that is the obvious candidate for an opposite side to the proposed one. This determination of these secondary lines, termed "directors" is diagrammed in Figure 4.19.

Having determined the termini and directors for a particular line, the procedure for proposing lines from a particular end is governed by four cases:

1) The end of the line is of undetermined terminus;



DIRECTORS ARE LINES WHOSE DIRECTION SUGGESTS THE DIRECTION OF A MISSING LINE.

Figure 4.19

2) Only one line intersects with the given line at the selected end;

3) Two lines lying on the same side of the given line intersect the selected end; or

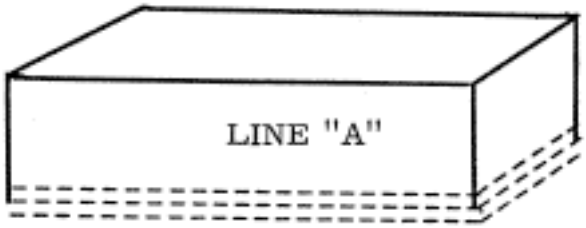
4) Something else occurs at the selected end.

In the first case lines are proposed from either side of the original line for which a director has been found. Three parallel lines are proposed, 10 units apart, parallel to the director, and with the end of the middle line coinciding with the indicated end of the given line. The other ends of the proposed lines are determined in this case, and in the cases

which follow, by making the proposed line no longer than the director, and truncating it so that it does not intersect any other lines in the set. In case 2), lines are proposed on the opposite side from the intersecting line, provided a director exists on that side. Three lines are proposed ten degrees apart, all extending from the terminus. Their lengths are determined as above, and the middle line is again parallel to the director. If no lines are found on the side away from the intersecting line, then lines are similarly proposed on the same side as the intersecting line, subject to the additional provision that the proposed lines all lie between the intersecting line and the director. If this were not the case, it would be impossible that the proposed line and the director be opposite sides of a face. In case 3), lines are proposed similarly to the first set described for case 2). In case 4), no lines are proposed. These cases are diagrammed in Figure 4.20.

In summary, PROPOS goes down the list of linked lines proposing in the manner described above, adding lines which it finds to the end of the list. This procedure is applied to the resulting list, and so on until no lines on the current list result

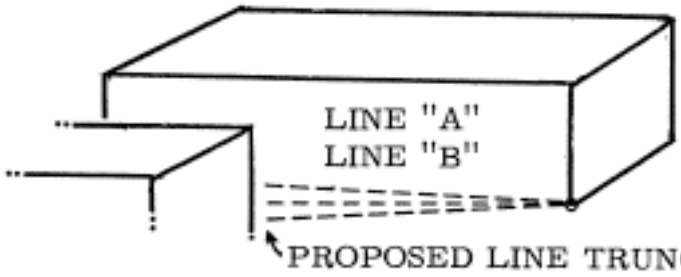
1) LOWER TERMINUS OF "A" IS UNDEFINED,
PARALLEL LINES PROPOSED:



DIRECTORS FOR
LOWER END OF "A"

--- PROPOSED
LINES

2) ONE LINE INTERSECTS LOWER END OF "A",
RADIAL LINES PROPOSED:

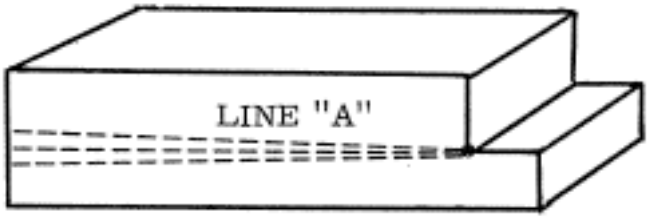


o TERMINUS

---- PROPOSED
LINES

PROPOSED LINE TRUNCATED BECAUSE OF "B"

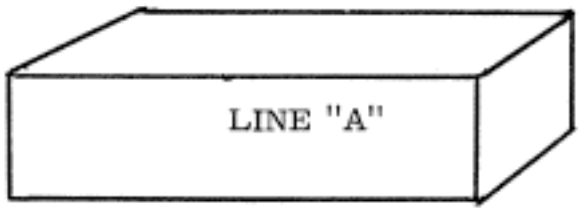
3) TWO LINES AT LOWER END OF "A" ON ONE SIDE,
RADIAL LINES PROPOSED ON OTHER SIDE:



o TERMINUS

-·-- PROPOSED
LINES

4) LINES ON EITHER SIDE OF LOWER END OF "A",
NO LINES PROPOSED:



PROPOSING OF LINES FROM AN END OF A GIVEN ONE
IS GOVERNED BY FOUR CASES.

Figure 4.20

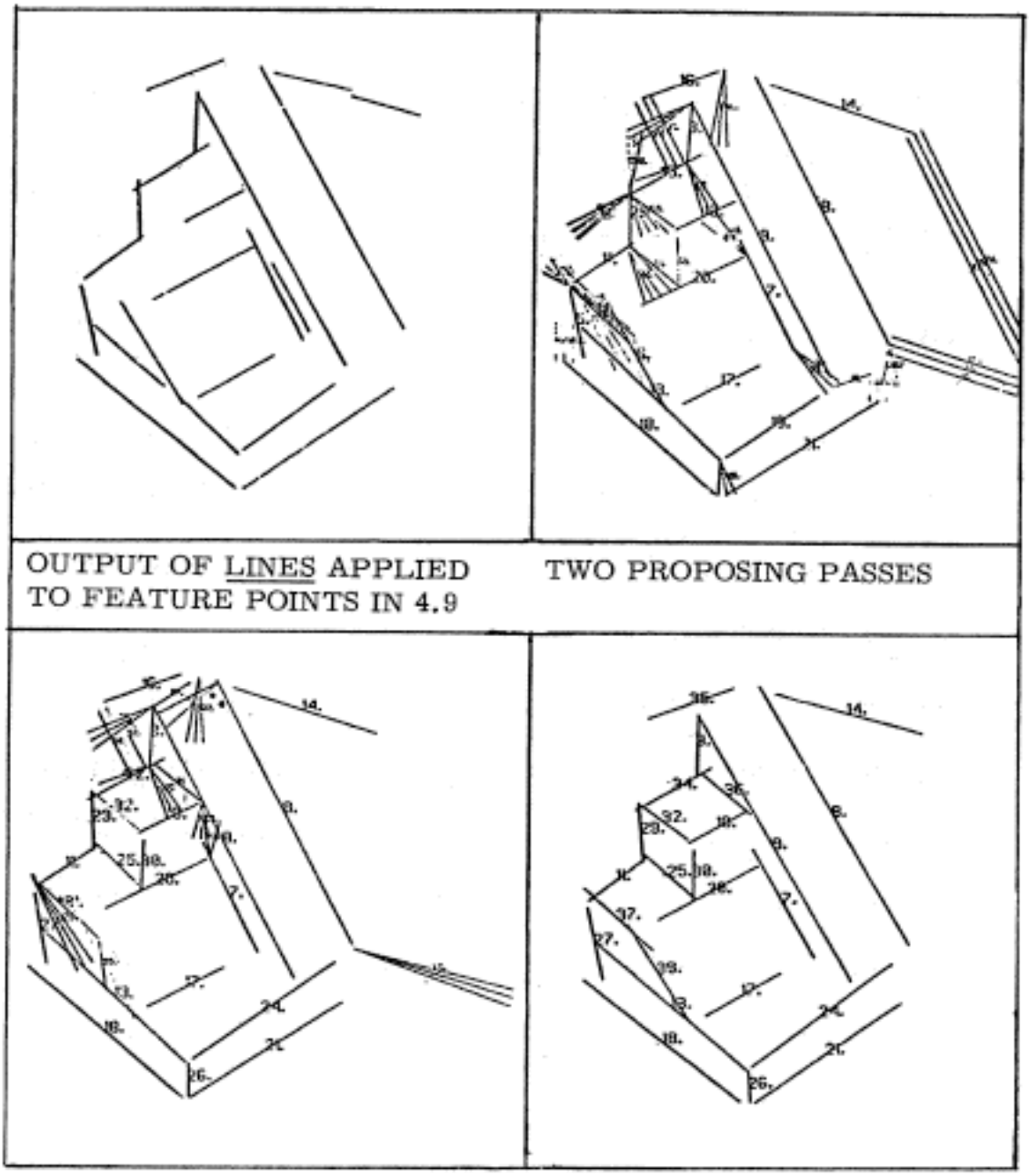
in the proposing of any further lines. At this point, the final set of lines is returned by the program. The process of proposing lines is diagrammed in Figure 4.21.

VII. A LINE VERIFIER

Line verification consists in applying the function $Q^*(J_j^n)$ as given by (3.13) and (3.29), or some variant of it, to the intensities J_j^n, or some values derived from them, lying in a region suggested by the line proposer. The region is assumed to be divided into $\underline{n}$ square or nearly square sub-regions, and Q^* or its variant is computed from the values $R_{j,1}, \cdots, R_{j,n}, a^*_{j,1}, \cdots$, etc., according to Formulas (3.13) and (3.29), or formulas derived from them. In most of the remainder of this section we shall be concerned with a generalization of $Q^*(J_j^n)$ which has been suggested by the discussion at the end of chapter three. This generalization consists in assuming that the values of $R_{j,k}$ and $R'_{j,k}$ are given by:

$$R_{j,k} = \text{Min}(R_{j,k,1}, \cdots, R_{j,k,6}), \tag{4.1}$$

$$R'_{j,k} = \text{Min}(R'_{j,k,1}, \cdots, R'_{j,k,5}), \tag{4.2}$$



OUTPUT OF LINES APPLIED TO FEATURE POINTS IN 4.9

TWO PROPOSING PASSES

THREE ADDITIONAL PROPOSING PASSES

FINAL SET OF LINES

DIAGRAM OF THE PROCESS OF LINE PROPOSING.
SMALL NUMBERS INDICATE THE ORDER OF PROPOSING.

Figure 4.21

where the $R_{j,k,t}$'s and $R'_{j,k,u}$'s are computed from the 25 intensities of the k-th sub-region scan segment according to (3.25), and from the paradigm intensity profiles illustrated in Figure 4.7. According to the discussion in Section III, the values in the feature point rasters are derived from the values given by (4.1) and (4.2), and the present line verifier draws upon exactly these values. In this section we shall review the relevant theory in detail, describe the present line verifier, and suggest possible improvements and generalizations of it.

The present line verifier, in computing a variant of $Q^*(J_j^n)$ for a region, obtains the values of, e.g., $R_{j,k}$ for CL(j) of (3.29) from the feature point rasters such as are partially depicted in Figure 4.10. In principle, if the region encloses $\underline{n}$ 25 point sub-region scan segments of the original scans, then the values $R_{j,k}$, etc., should be computed from exactly these scan segments. These values were computed for all possible sub-region scan segments, in the generalized manner given by (4.1) and (4.2), by the feature point generation procedure. However, most of the values of the R's and R''s

were discarded during the compactification process. In fact it was only locally maximal values of E, given by:

$$E_{j,k} = R''_{j,k} - \text{Min}(R_{j,k}, R'_{j,k}) \tag{4.3}$$

where $R_{j,k}$ and $R'_{j,k}$ are given by (4.1) and (4.2), that were retained. However, it turns out that we may obtain a reasonable estimate of the value of $E_{j,k}$ for the intensities of a particular sub-region scan segment along a particular scan from one of these maxima nearby along the same scan. In particular, if $E_{j,k}$ is a desired value of E for a sub-region scan segment along some scan, then let $\hat{E}_{j,k}$ be the value of E which is locally maximal along the same scan and is computed from a sub-region scan segment closest to the one in question. We assume that there exists a maximum within something like 15 units of the segment in question, else we assign $E_{j,k}$ the value 0. If $\hat{E}_{j,k}$ so defined actually exists, then the following relationship holds:

$$E_{j,k} \approx \hat{E}_{j,k} - ax^2 \tag{4.4}$$

where $\underline{a}$ is some constant, and $\underline{x}$ is the distance between the sub-region scan segments from which $E_{j,k}$ and $\hat{E}_{j,k}$ were computed. This relationship follows from the behavior of the E curve in the neighborhood of a maximum, which may be

observed in Figures 3.9 and 3.10. It will be seen later that the $\underline{n}$ values of $E_{j,k}$ are all that are necessary to compute the present variant of Q* used in the line verifier.

This present variant of Q* omits the second and third terms in the exponent of, e.g., the value of CL as given by (3.29), and thus does not need the values of $a^*_{j,k}$, $a'^*_{j,k}$ and $a''^*_{j,k}$. The middle term may be ignored because it turns out to be always very small relative to the other two. To see this, recall that line verification in the present set of programs is applied only to lines, or possible lines, which escaped detection by the LINES program. They are thus lines which have a net relative amplitude whose magnitude is small relative to the average magnitude of relative amplitudes for lines in general. One may observe that the denominator of the middle term is dominated by its third term, so the term is approximately $(\sum_K a^{*2}_{j,k}/n)/\rho^2_n$. This is approximately the square of the net amplitude divided by its variance. Since the net amplitude in situations under consideration is small relative to the average magnitude of the net amplitudes, and the average net amplitude over lines in general is zero; then the net amplitude

is small relative to its standard deviation. The second term of (3. 29) is thus in this case very small, being approximately the square of the quotient of the net amplitude with the standard deviation.

The omission of the third term of the exponent of (3. 29) was based on the empirical observation that for the majority of the scenes analyzed, this value seemed not to be related in a reliable way to the existence of lines. A series of experiments was begun to properly test whether any such relation existed. The results appeared promising but not so far conclusive. Consequently the third term was omitted pending further investigation.

The resulting variant of $Q^*(J_j^n)$ is quite similar to the one discussed at the end of the previous chapter. It is given by (3. 13), where the various values, e. g., of CL are given by (3. 29) but with only the first term in the exponent. This was the case in the situation described in the end of chapter three. Also, this time we expect a scan of 500 points to intersect with about five lines in the visual field, so the relation between P(CL), P(CE) and P(CH) remains approximately the same as in the situation at the end of chapter three. Finally, we may as before

assume that one of the terms in the numerator is considerably larger than the other. Thus by an argument similar to that used in deriving (3. 41) and (3. 43) we arrive at the function $Q'^{**}(J_j^n)$ defined by:

$$Q'^{**}(J_j^n) =$$

$$\sum_{k=1}^{n} \text{Max}(SL_k(j)\text{-}SH_k(j), SE_k(j)\text{-}SH(j)) + K, \qquad (4.5)$$

where, e. g., $SL_k(j)$ is like SL(j) given by (3. 37) with subscript k in place of 1. As before, this function is approximately monotone in $Q^*(J_j^n)$ and hence thresholding it provides an approximately sub-optimal regional decision predicate. One may see from (3. 37), (3. 38), (3. 39), (3. 23), (3. 24), (3. 25) and a certain amount of algebra that:

$$SL_k(j) - SH_k(j) = R''_{j,k} - R_{j,k},$$

and

$$SE_k(j) - SH_k(j) = R''_{j,k} - R'_{j,k}.$$

Using (4. 3) we may rewrite (4. 5) as:

$$Q'^{**}(J_j^n) = \sum_{k=1}^{n} E_{j,k} + K, \qquad (4.6)$$

with the values of $E_{j,k}$ obtainable via (4. 4) from values $\hat{E}_{j,k}$, which are available in the feature point raster. In

summary we have:

$$Q'^{**}(J_j^n) = \sum_{E_{j,k} \neq 0} \hat{E}_{j,k} - ax_k^2 + K, \tag{4.7}$$

recalling that if a suitably close feature point for obtaining a value $\hat{E}_{j,k}$ did not exist, then the corresponding $E_{k,j}$ was assigned the value zero.

In practice it was found to give better results to deviate slightly from a strict thresholding of (4. 7) in the verification of lines. Prior to applying (4. 7) to a region it was observed which of the $E_{j,k}$ were given nonzero values. In the case that the locations of the nonzero $E_{j,k}$'s were sparse or concentrated in one place in the region, the thresholding procedure was omitted, and a negative answer was returned as to the existence of a line. This is approximately the same as observing the pattern of the locations of feature points within the region, and denying the existence of a line if they are all clumped together or are sparsely scattered within the region. These two situations are usually due to the existence of feature points from lines that cross the region at non-negligible angles. Another variation from (4. 7) which was found to be more convenient, and in some cases more accurate, was to

threshold the sums of the $\hat{E}_{j,k}$'s and the sums of the x_k^2's separately. Finally, instead of taking x_k to be the deviation of the $\hat{E}_{j,k}$ feature point from the center of the region, it was taken to be the deviation of the $\hat{E}_{j,k}$ feature point from a best-fit line to the various $\hat{E}_{j,k}$ feature points.

The line verifier (VERIFY P1 P2) takes as arguments the end-points P1 and P2 of the proposed line and outputs either NIL, if the line is thought not to exist by the above criterion; or the end-points of the best fit line to the feature points if the line is considered to exist.

In observing the line verifier in action in Figure 4.21, occasional failures are apparent. Although no lines are claimed to exist which do not exist in the figure, several actually existing in the figure were proposed but claimed, on the basis of the feature point raster information, not to exist by the verifier. An examination of the feature point raster indicates the presence of feature points in these regions. The problem was principally that the region intersected with too few scans, on the order of three to five; giving too few values for the threshold function (4.7) to work with. Evidently a finer scan would eliminate this problem, as well as some other approaches to

its improvement which are about to be discussed.

A major source of improvement for the line verifier is for it to obtain a new set of intensities within the region to which it is applied, at considerably higher density than that of the original scan. This would eliminate the difficulty described in the previous paragraph, would allow more accurate values of the R's, E's or whatever, to be obtained, and would eliminate the necessity of the approximation to the $E^{n}_{j,k}$'s in (4. 4). Another improvement would result from the utilization of the third term of (3. 29) and similar expressions. As was stated previously, investigations in this area are still under way. It seems plausible that the model in Section IV of chapter three may have to be modified in assumptions 4) and 5) to admit of a regular linear variation of "idealized" relative amplitude along a noise-free ridge-like or cliff-like noise-free sample.

CHAPTER 5

RELATED WORK

I. OTHER APPROACHES TO LINE FINDING

An early line-finding program was written by Roberts (Roberts). This program analyzed photographs, of good contrast and focus, of scenes whose content was principally defined by straight edge-lines. Feature points were obtained from a predicate on four adjacent intensities of a 2 by 2 square on the visual field. Feature points were organized into lines by a sort of line-following procedure. Higher level heuristics handled the resulting lines.

A comparison may be made between the behavior of Roberts' four-point feature point predicate, and the present feature point gathering procedure. This analysis is a

generalization of a procedure due to Binford and Herskovits (Herskovits).

An upper bound on the performance of the former may be inferred from theoretical considerations. Data on the actual performance is not available. The feature point predicate is applied to intensities $x_{i,j}$, $x_{i+1,j}$, $x_{i,j+1}$, and $x_{i+1,j+1}$ at four adjacent points in a square pattern on a 256 by 256 point grid. The thresholded value is:

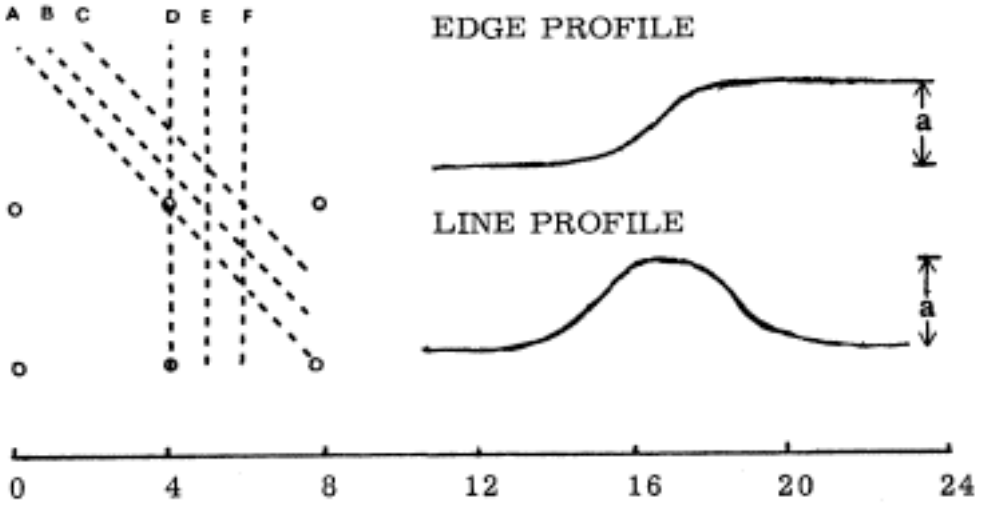
$$z_{i,j} = \sqrt{(y_{i,j}-y_{i+1,j+1})^2 + (y_{i+1,j}-y_{i,j+1})^2},$$

where

$$y_{i,j} = \sqrt{x_{i,j}}.$$

Assume that the noise is gaussian, and that the noise level, as measured by the standard error about a mean of values of $\underline{y}$ taken at a single point under constant lighting conditions, is $\underline{s}$ for intensities in the vicinity of the average of the four y's in question. It is not difficult to show that $\underline{z}$ as a function of the random variables $y_{i,j}$, $y_{i+1,j}$, $y_{i,j+1}$ and $y_{i+1,j+i}$ has a median of about 2s, and an upper interquartile point value of about 3s. In the context of the set of figures we are considering, one would expect about one feature point per 100 quadruples of

intensity. Thus for about 50% false positives, one would wish to set the threshold on $\underline{z}$ high enough that $\underline{z}$ falls above threshold by chance once in 100 times. For a normal distribution, this is about 2.5 standard deviations, which corresponds to 3 times the semi-interquartile range. Thus $\underline{z}$ should cut off at about 5s. Now it is not difficult to see what values of $\underline{z}$ would result for various lines and edges of amplitude $\underline{a}$ oriented in various ways with respect to the grid of squares. In Figure 5.1 we have computed such values for 12 cases. In the diagram, we are assuming that the dimensions of the blurring function are such as to give the line and edge profiles of the width shown. This width appears to be about an optimum compromise. If considerably lower then more edges would have a relatively high z-value, but more lines would have a zero z-value. As the width increases, both lines and edges tend toward having a zero z-value. From the diagram it may be estimated that the average value of $\underline{z}$ for edges is about .8a, and for lines about .4a. Thus the amplitudes of lines and edges respectively which are at threshold are about 6.3s and 12.6s. This may be compared with the values 1.2s and 2s obtained for Q* at the end of chapter three. The difference is due to two factors. First,



1 UNIT = 1/2000 VIDI FIELD WIDTH

LOCATION	APPROX. VALUE OF Z-OPERATOR LINE	EDGE
A	$a/\sqrt{2}$	a
B	a/2	a
C	0	a
D	$a/\sqrt{2}$	$a/\sqrt{2}$
E	$a\sqrt{2}$	$a\sqrt{2}$
F	0	$a\sqrt{2}$

VALUES OF ROBERTS' Z-OPERATOR FOR LINES AND EDGES OF VARIOUS ORIENTATIONS.

Figure 5.1

the Roberts predicate has a support of only four points, whereas the Q* tested had a support of 25 points. One would expect the signal to noise ratios to differ by a factor of $\sqrt{25/4} = 2.5$. The actual difference is about 5, and the remaining factor of two is probably due to the fact that the Roberts predicate is isotropic, whereas Q* is directional.

The Topologist, based on the work of Binford, Sussman (Sussman) and Herskovits (Herskovits), represents another approach to the analysis of simple scenes. In this case, the scene is divided into a set of small square regions which are individually tested for homogeneity. Adjacent homogeneous regions are clustered into maximal homogeneous regions, which are presumably the areas of the scene which lie between edge lines. This procedure has the properties that edge lines need not be straight, and that arbitrarily complex predicates for homogeneity may be used.

Recently, a method of spatial frequency filtering and matching has been applied to the problem of location of feature points (Hueckel). This procedure is roughly a two dimensional version of the feature point analysis presented here. For a circular neighborhood, an estimate is given of the directional

and amplitude parameters of a possible edge through the neighborhood. This procedure seems to be based on a rather simple model of the nature of intensity profiles at edges. Also, it seems to be relatively insensitive to the presence of lines.

An approach bearing some mathematical similarities to the feature point detection procedure reported here, was employed by Saunders (Saunders) in an analysis of filmed images of scope tracings. In this case it was necessary to locate the intersections with parallel scan lines of the image of a narrow trace along the film.

Worthy of mention here, though not specifically related to line finding, is the work in statistical decision theory of Chow (Chow). The theorem in the first chapter is essentially an adaptation and generalization of Chow's results in the theory of optimal error reject trade-off.

II. RELATION OF THE PRESENT WORK TO THE DEVELOPMENT OF A VISUALLY ORIENTED REAL-TIME OBJECT MANIPULATOR

The program here described is part of a larger effort to develop the visual-perception aspects of a real-time object manipulator. (MAC Progress Report 1968) For the present,

effort has centered on programs to obtain descriptions of scenes composed of cubes and other prismatic solids. By a description of an object is meant a list of the spatial co-ordinates of its vertices, together with a list of pairs of vertices defining edges of the object. By a scene description is meant a list of object descriptions, one for each object in the scene.

There is an obvious disparity between the output of the present program and a corresponding scene description. The former is, or can easily be reduced to, a set of vertices and a set of vertex pairs representing edge lines. A description requires that these vertices and vertex pairs be grouped into sets corresponding to the various objects in the scene. One may regard the output of the current program as a set of lines dividing the field into regions. The problem then becomes one of grouping regions which correspond to faces of a single object, and matching regions which may be parts of a single face which has an occluding object in front of it which divides it into two regions. This problem has been investigated in detail, and a program called SEE (Guzman) performs the described grouping and matching to reduce the output of the present program into a

proper two dimensional scene description.

Several approaches to the problem of determining the locations of objects in three dimensions have been investigated in greater or lesser degrees of detail. One suggestion involves the use of two light sources located close to the point of view of the optical input device. One would be a point source, whose intensity falls off as the square of the distance from the point of view. The other would be a source placed "at infinity" by means of lenses, whose intensity would fall off negligibly with intensity. If illumination is rapidly alternated between these two sources, and intensities obtained for a particular point under both conditions, the relation between the two intensities would yield the distance from the point of view to the point in the field. Another approach to determining the distance from the point of view to a point in the scene involves the use of optical focusing. (Horn 1968) By this procedure, a point in the scene is brought into perfect focus and the position of the lens of the optical input device is transmitted to the computer. After a suitable calibration procedure, the absolute distances of objects may thus be obtained automatically. A third approach (Horn) involves the use of intensity gradient to infer the shape of the contour of a

curved surface. By this technique, relative distance information may be obtained for points on curved surfaces. A final approach worked out by the present author, involving an analysis of the scene from two different points of view, relates directly to the current work. It is assumed that the present program has examined a suitable scene from two different points in space, and that the resulting two sets of lines have been processed by SEE. Central to this approach is a subroutine called WHERE which is based on results in projective geometry, and is described in a recent Project MAC Artificial Intelligence Memo (Minsky). This program takes as input the two dimensional co-ordinates of a point from each scene. In many cases, WHERE can determine that the two points are not images of the same point in space. If the points are from the same point in space, WHERE outputs the three-dimensional co-ordinates of that point. One might apply this program in a straightforward manner to a point in one scene together successively with all points of the other. This would greatly reduce the possible "matches" of features in the two scenes, but would, in general, not suffice to determine the three-dimensional locations of all the vertices. The problem of resolving the remaining ambiguities, as well as

developing a more efficient point-matching procedure seems to be quite complex. It seems likely that the organization that SEE imposes in the lines would provide a basis for using contextual clues to match the vertices. The problem is currently under investigation.

A final problem worthy of mention is that of the determination of the stability of a configuration of prismatic solids, given a three-dimensional scene description of it. (Blum) The relevance of this problem derives from the fact that it is one of the goals of the object manipulator project to be able to construct structures from rectangular blocks. In order to direct the object manipulation device, a program must first determine the order in which the blocks are to be set down. It may be possible to construct a sub-configuration of the desired configuration which has the property that adding any additional block results in an unstable configuration. To forestall such a possibility, it is necessary to be able to detect instability in searching for an appropriate order in which to build the structure.

BIBLIOGRAPHY

Bellman, R. 1957. Dynamic Programming. Princeton, N. J.: Princeton University Press.

Blum, M., Griffith, A. and Neumann, B. 1970. A Stability Test for Configurations of Blocks. M. I. T. Project MAC Artificial Intelligence Memo 188.

Chow, C. K. 1969. On Optimum Recognition Error and Reject Tradeoff. M. I. T. Project MAC Artificial Intelligence Memo 175.

Guzman, A. 1968. Computer Recognition of Three Dimensional Objects in a Visual Scene. M. I. T. Project MAC Technical Report: MAC-TR-59 (Thesis).

Guzman-Arenas, A. 1967. Some Aspects of Pattern Recognition by Computer. M. I. T. Project MAC Technical Report: MAC-TR-37 (Thesis).

Herskovits, A. 1970. On Boundary Detection. M. I. T. Project MAC Artificial Intelligence Memo 183.

Horn, B. K. P. 1968. Focusing. M. I. T. Project MAC Artificial Intelligence Memo 160.

Horn, B. K. P. 1969. The Image Dissector "Eyes". M. I. T. Project MAC Artificial Intelligence Memo 178.

Hueckel, M. 1969. An Operator Which Locates Edges in Digitized Pictures. Stanford Artificial Intelligence Project Memo: AIM 105.

Project MAC 1968. Project MAC Progress Report V.

Minsky, M. L. 1967. Stereo and Perspective Calculations. M. I. T. Project MAC Artificial Intelligence Memo 143.

Roberts, L. G. 1965. Machine Perception of Three-Dimensional Solids. In Optical and Electro-Optical Information Processing, ed. Tippett, J. T., et al., pp. 159-197. Cambridge, Massachusetts: M. I. T. Press.

Saunders, R. (undated). Radar Film Reading. Memo by Information International Incorporated, Maynard, Mass.

Sussman, G., and Guzman, A. 1966. A Quick Look at Some of Our Programs. M. I. T. Project MAC Artificial Intelligence Memo 102.

TECHNICAL REPORTS BY THE AUTHOR

Griffith, A. 1962. A Lisp Manual. University of Florida, Quantum Chemistry Group, Preprint 31.

Griffith, A. 1966. A New Machine-Learning Technique Applied to the Game of Checkers. M. I. T. Project MAC Artificial Intelligence Memo 94.

Griffith, A. 1967. POLYSEG. M. I. T. Project MAC Artificial Intelligence Memo 131.

Griffith, A. 1967. Appendix B. In Stereo and Perspective Calculations. Minsky, M. L., M. I. T. Artificial Intelligence Memo 143.

Blum, M., Griffith, A., and Neumann, B. 1970. A Stability Test for Configurations of Blocks. M. I. T. Project MAC Artificial Intelligence Memo 188.