DESCRIPTION OF VISUAL TEXTURE BY COMPUTERS

VISION FLASH 39

by

John Gary Gaschnig

Massacusetts Institute of Technology

Artificial Intelligence Laboratory

Robotics Section

9 MAR 1973

Abstract

Some general properties of textures are discussed for a restricted class of textures. A program is described which inputs a scene using vidisector camera, discerns the texture elements, calculates values for a set of descriptive features for each texture element, and displays the distribution of each feature. The results of the experiments indicate that the descriptive method used may be useful in characterizing more complex textures. This is essentially the content of a Bachelor's thesis completed in June, 1972.

Work reported herein was conducted at the Artificial Intelligence Laboratory, a Massachusetts Institute of Technology research program supported in part by the Advanced Research Projects Agency of the department of Defense and monitored by the Office of Naval Research under Contract Number N00014-70-A-0362-0003.

Vision Flashes are informal papers intended for internal use.

1. Introduction

The perception of and the discrimination between textured objects play a large role in the human visual system, and necessarily so, for our universe is filled with a host of visually textured objects. Humans use visual texture information readily to help distinguish between objects or surfaces, to provide shape, inclination and depth cues, and to tell about the nature and composition of the perceived object or surface.

This paper is concerned with a description of the "type" of a certain class of textures. Relatively little work has been conducted in the study of visual texture. Much of the work that has been done has been concerned with the synthesis of textures [1,3,16]; relatively little effort has been devoted to the analysis of textures. Research in this field may be directed along one of three lines: statistical analysis of local properties, Fourier analysis, and structural or pattern analysis. Statistical methods seem appropriate for textures which can be described solely on the basis of local properties. A fair degree of success has been achieved in synthesizing random-dot textures with Markov processes [7,8,9,13]. Such textures are sometimes used in psychopictoric experiments to measure human texture discrimination [13]. The work of Rosenfeld and Troy [17]

suggests that statistical methods may be of little use in accurately describing or analyzing complex, high-order textures.

Some attempt has been made to use Fourier techniques in the synthesis of textures [1], but less is known about pattern description and picture analysis in the Fourier domain than in picture space; I know of no work that has been done in the analysis of textures using Fourier techniques. Structural description seems to be the most natural of the available domains, because it has obvious analogies to human perception. In both cases, one attempts to describe a scene in terms of recognizable forms contained in the scene, and in terms of the relationships among those forms. This work has been directed toward the formalization of a structural description of textures.

The dependence of current artificial vision systems on line finders to provide the information used to identify and locate objects in a scene is a great handicap when the scene contains visual textures. The mass of lines due to the texture hide the lines representing the true edges of a block, and the system usually fails in its attempt to identify the object. What is desired is the ability to perceive a textured surface as a bounded region of homogeneous texture of a given type. If it were possible to perceive a surface of a block as a homogeneous texture of a

given type, the irrelevant lines could be ignored. If the three visible surfaces of the block formed three homogeneous textures of different type, and if the boundaries between the textured regions could be discerned, it seems likely that the true edges of the block could then be found (Figure 1).

Figure 1: Application of texture discrimination in block recognition

When artificial vision systems attain the ability to perceive textures, the domain of processable scenes will be much enlarged, resulting in a system whose range of input more closely approximates that of humans. When compared with the amount of visual texture information contained in many scenes, line drawings are poor in information content. To attain the goal of sophisticated intelligent artificial vision systems it will be necessary to extract more general types of information, such as texture, from scenes, rather

than relying solely on line-drawings.

2. Texture: Some general properties

2.1 A description of a class of textures

No generally accepted scientific definition of texture is available, even though everyone "knows" what is meant by texture. The following examples may illustrate what people consider as textures: wood grain, animal fur, the surface of a chunk of stone, a cloud formation, a brick wall, a grass lawn, a school of fish, a filled parking lot seen from a tall building, a bouquet of flowers, an aerial view of a city, lizard skin, a tile floor, a random pattern of shapes, a herringbone suit, the upholstery of a couch, an aerial view of a forest, the face of a skyscraper seen from a distance.

A wide class of textures can be described as the placement over a given area according to specific placement rules of many occurences of one or more unit patterns. The texture is then defined by the unit pattern(s) and the placement procedure. The unit pattern may be a simple geometrical figure (e.g., a diamond-shaped figure), or it may be complex itself, consisting of a specific spatial arrangement of geometrical figures. The instances of the unit pattern in the texture need not be identical, but they must be similar enough to each other that they can be

perceived as equivalent. The texture elements are defined to be the forms which are replicated in the texture; they must be individually discernable, i.e., not beyond the resolution of the picture-inputting device. The texture elements are identical to the unit pattern unless the unit pattern is complex, in which case the texture elements are the forms of which the unit pattern is composed. The placement procedure may range from describing a simple spatial periodicity (e.g., a checkerboard) to random placement to a complex pattern whose regularity may be difficult even for humans to see. Often the placement process generates a figure-ground pattern, in which the texture elements are spatially separated in a homogeneous background. In other textures, the texture elements are adjacent (wire mesh).

2.2 Economy of representation

The examples of textures given above all describe visually complex scenes, that is, there are a lot of things to be perceived in the scenes. This is in contrast with scenes like a smooth white cube sitting on a smooth black table, a blank TV picture tube, or a painted bedroom ceiling, which are rather perceptually uninteresting because there is just not much to look at in them. But consider such complex scenes as: the Mona Lisa, a map of the New York City subway system, the cockpit of a Boeing 747, a page of

text in a book. These scenes are visually complex (in terms of the number of visually discernable forms in them), but it is unlikely that they would be classified as textures. The latter examples contain more information than do the textures; more effort is required to understand the whole scene than is required for texture scenes.

The suggestion is that textures are not as perceptually complex as they are visually complex, that is, that the amount of effort needed to extract information from a texture scene or to descibe a texture scene is less than one might expect if one considers only the visual complexity of the scene. Because of the repititive nature of most textures, one can often obtain a good understanding of the texture from the consideration of only a small portion of the entire texture; this cannot be done with a painting.

It has also been suggested [2,18] that there is a strong tendency in the human visual system to represent the outside world as economically as possible, that a representaion of the scene used is the one of minimal complexity. Presumably, a judgement is made as to what is and what is not important in the scene, and unimportant details are subordinated in favor of a generalization of the important contents of the scene.

As an example, picture a sweater that has been thrown onto a table, such that it is not neatly folded but

rather lies in a crumpled heap. If you showed this scene to a person for a moment and then took it away and asked the subject to describe what he saw, he might say "There is a sweater lying on a table." He might also have perceived and remembered some of the details, such as whether the sweater is multi-colored or of a single hue, what the color or colors are, whether it is a cardigan or a pullover. But if he were asked to draw an outline of the sweater as it appears or to describe in detail the folds of the material in that arrangement, or to say whether it was inside-out, or to tell how much of either arm was visible, he might be unable to do so unless he had the opportunity to look at the scene again, concentrating on extracting this detailed information. Unless the viewer is interested in finding specific pieces of information, the tendency is to use one's knowledge about the world and what one expects to see in a scene to make a simple and acceptable generalization about the contents of the scene

2.3 Perceptual equivalence

The texture elements are perceptually equivalent, i.e., they can be interchanged within the scene while leaving the scene perceptually unchanged. This can be illustrated by the example of a brick wall. If one tears down the wall and then rebuilds it in such a manner that no

brick occupies its former location, the new wall will be nevertheless all but indistinguishable from the old, even though the surfaces of no two bricks are ever identical. The elements of a texture need not be identical, but they are similar enough to be grouped together into equivalence classes so that a gross pattern emerges.

2.4 Context and experience sensitivity

The placement procedure is an important source of context information used in determining how a texture is perceived. For example, a texture picture whose texture elements are perceptually equivalent, but in which the spatial density of the elements differs in different sections of the picture will be perceived as two distinct textures. Similarly, if a brick wall is composed of a uniformly mixed assortment of red bricks and orange bricks, it will be perceived as a single texture. But if one looks at two different sections of a brick wall, one of which is composed only of red bricks, the other solely of orange bricks, a distinction between the two sections will be perceived. In both cases the different types of texture elements are similar enough to be grouped together. In the first case this indeed occurs, but in the latter case there is additional evidence (the spatial arrangement) which overrides the grouping tendency and results in a different

perception.

Experience also plays a large part in the human perception of texture. A page of Sanscrit may appear as a texture or as literature depending upon one's familiarity with Sanscrit. People who interpret aerial photographs or photographs of lunar craters use specialized experience in their evaluations. The understanding by humans of many textures requires a high level of knowledge about the world.

3. A model for texture description

A restricted class of textures can be described as follows. The unit pattern consists in a single arbitrary geometrical figure. Stochastically perturbed instances of this unit pattern or texture element are replicated throughout the picture space to form the texture. No information about the placement procedure is included in the description of the texture. Figure 2 illustrates typical textures in this class. Such a texture can be described solely by describing the texture element.

Because of the redundancy of information in the picture space for such textures, it may be efficient to describe such textures by using economies of representation, as described above. What is necessary is to characterize a given texture by a higher level generalization or description, by working in a texture description space

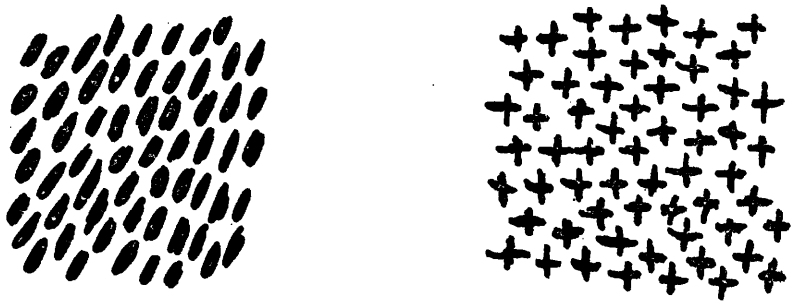


Figure 2: Examples of a restricted class of textures.

rather than in the picture point space. The texture elements must first be individually discerned. This will be done most easily by allowing a contiguous picture region of fairly constant brightness to represent a discernable texture element. A list of descriptive features must be generated for each texture element in the scene. For this purpose, it will be easiest to use textures whose elements can be described by their appearance alone. The frequency distribution of values for each feature for a given texture may be regarded as a description of the texture. If the individual texture elements are similar in appearance, then the histograms of those features which measure a property common to the texture elements will display a localized distribution for the feature values. If the texture elements differ in possession of a given property, then the

corresponding feature histograms will show a more uniform distribution of values. It may be desirable to condense the information contained in the feature histograms. For peaked distributions the mean value of the feature may be considered to characterize the basic texture element for that feature. A uniform distribution may be characterized by the description "uniform." Although it was not attempted in this work, a grouping or clustering operation in feature space may give an accurate description of the basic texture element and of the perturbation from that ideal allowed for actual texture elements. Such a process attempts to create equivalence classes for the samples, all of whose member should be perceptually equivalent.

4. A description of the program

4.1 The Texture Element Finder

A program has been written which discerns the individual texture elements of a scene, calculates values for a set of descriptive features for each region, and displays the frequency distribution of each feature. A scene is inputted to the program using a vidisector camera. The field of vies of the camera consists in a 1024 by 1024 grid of picture points. Associated with each picture point is one of 703 possible intensity values, computed as a function of the logarithm of the actual light intensity from

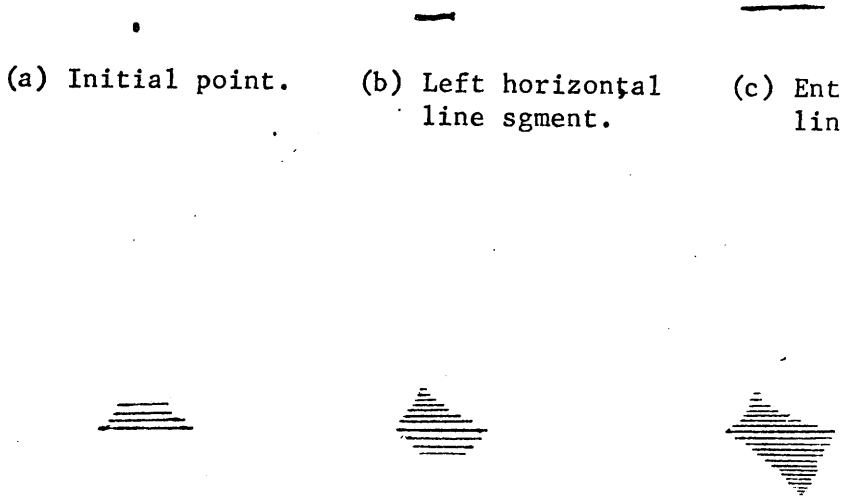
the scene. In the following discussion, "intensity" refers to the value of the light intensity associated with a picture point.

The program can work directly with the vidisector. For the sake of experimental replication and debugging ease, the scene is first stored on disk, and the program then uses the stored scene as input. The texture elements found are picture regions of constant or nearly constant brightness. The procedure for finding the texture elements, explained in detail below, is illustrated in figure 3.

The number of texture elements to be found and the coordinates of a rectangular picture window within which the texture elements are to be found are specified by the user. To find a texture element, a random initial point is generated, and a horizontal line 100 picture points wide and centered on the initial point is scanned. An attempt is made to grow around the initial point a picture region whose points have brightness values similar to that of the initial point. The intensity of the initial point is the reference intensity for the texture element.

The points to the left of the initial point are examined sequentially, and each of these candidate points is accepted into the region defining the texture element if

$$\left| i_r - i_c \right| < d \quad ,$$



(a) Initial point. (b) Left horizontal line sgment. (c) Entire horizontal line segment

(d) Addition of upper horizontal line segments. (e) Addition of lower horizontal lines segments. (e) Completed texture element.

Figure 3: A graphical description of the texture element finder.

where i_r is the reference intensity for the texture element, i_c is the intensity of the candidate point, and d is a user-specified parameter (default value = 16) signifying the maximum deviation from the reference intensity for an allowable point in this texture element. When a point in this left line segment is examined whose intensity is beyond the acceptable range, the left half of the line is ended. The points to the right of the initial point are then

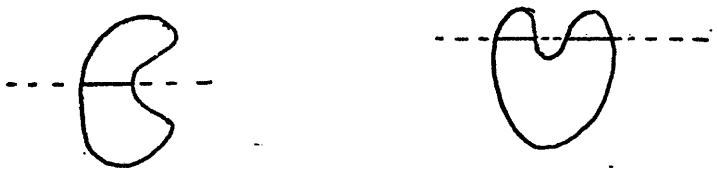
examined in similar fashion, and the right endpoint of the line is determined. The y-coordinate of the line and the x-coordinates of the left and right endpoints of the line are stored.

The line of picture points immediately above the line just considered is examined next. The initial point for the new line is (x_m, y_m+1), where (x_m, y_m) is the midpoint of the line segment just completed. If the intensity of this point is in the acceptable range, the points to its right and left are examined as before, and the coordinates of this line are found and stored. This process of adding line segments above the initial line is continued until the initial point of the next line segment is not accepted, and the top of the texture element is assumed to have been reached. The lines below the initial line are similarly processed, after which the figure is complete. This procedure is repeated for each texture element until the desired number of texture elements has been reached.

The initial random point around which each texture element is grown is examined to make sure it is not within a texture element already found. If the proposed initial point is within another texture element, then another initial point is chosen. This test is performed to insure that no texture element is found more than once. Although the initial point for a texture element is within the

picture window specified by the user, if the picture window is smaller than the entire scene, then a texture element which lies partly within and partly outside the picture window is allowed to grow beyond the picture window boundary.

The maximum size region allowed is a square the length of whose side is 100 picture points. The method of forming the texture elements restricts the class of input figures to those which have no breaks in a horizontal line of the region (figure 4). This restriction is not considered to be important in this initial study. The perimeters of the texture elements are displayed on the PDP-340 display scope.



(a) Allowed (b) Not allowed.

Figure 4: Restriction on class of texture elements.

The element finder works well when there is high contrast between the texture element and its surroundings (approximately binary figure-ground scenes). When the

contrast between texture element and background is not as great, the element finder is much more susceptible to noise introduced by the camera, and to the non-monoticity of intensity gradients. The resulting discerned texture elements do not match the humanly perceived texture elements particularly well.

Varying the range of deviation in brightness from the reference brightness which determines the acceptability of a point for a texture element affects the goodness of match between humanly perceived texture element and constructed texture element. For any figure-background contrast there is a range of the value of acceptable brightness deviation for which the match is best. Greater than this value, the constructed texture element becomes relatively overcomplete (i.e., parts of the scene are included in the texture element which are not part of the real texture element). Below this value, the constructed texture element becomes undercomplete. No automatic procedure for determining the "best" (in the sense of most consistent with human perception) acceptable deviation range has been devised.

For scenes with low contrast between texture element and non-textur element, the element finder often does a relatively poor job. Natural textured objects often have tactile texture in the form of local inhomogeneities in the

smoothness of a surface. Light reflected from such surfaces is diffused more than the incident light, and thus the contrast (i.e., the steepness of the intensity gradients) in the scene is reduced. The amount by which light is so diffused was qualitatively evaluated. One of the experimental scenes, consisting of black figures on white paper was put before the camera and the raster scan picture (i.e., television mode) was magnified, so that a single figure filled about one fourth of the screen. The effective magnification was approximately a factor of 20. The contrast between black figure and white background was sharp. A finely woven checkered dish towel was then placed before the camera. Even with a greater amount of light on the scene and a lower degree of picture enlargement, the checkered squares were difficult to distinguish. The conclusion made was that a better algorithm for finding texture elements will probably be required to obtain good results with natural textures. If the elements to be found are very small grains (e.g., sandpaper), it will probably be necessary to use a different lens on the camera.

There are a number of program options available to the user. Texture elements found which consist of less than a user-given number of points can be rejected. Except for textures with very small grains (e.g., sandpaper), this feature improves the results by eliminating small spurious

regions that are caused by noise added by the camera and which do not represent real texture elements. If the scene is of the figure-ground type and if there is significant difference in the average brightness between figure regions and ground regions, then it is desirable to find only the figure regions while ignoring the ground regions. There is an option that causes only dark regions or only bright regions to be found, instead of both dark and bright alike. The cutoff intensity can be scene dependent, and the mean value of of the intensities of one hundred random points within the picture window is currently used. One can thus choose to concentrate on only the darker or only the brighter sections of the picture, independent of the overall brightness of the scene. An algorithm for eliminating salt-and-pepper noise devised by David Waltz was implemented, but its effect was insignificant for the scenes used.

4.2 The Element Features

Measures or features of the texture elements are calculated and experimental scenes are used in an evaluation of the goodness of each of the features in characterizing the elements. Good measures of shape should be consistent with human perception. This study is an attempt to arrive at a better understanding of textures in human terms. Conceptually, if two figures are similar in shape to a human

observer, then there should be a close agreement in the features used to describe the shapes of the two figures. If humans see two figures as dissimilar, the differences between the two figures should be reflected in the descriptive features. Humans are good at distinguishing between figures which differ in size, brightness, orientation of major axis, eccentricity, smoothness of perimeter, pointedness, symmetry, concavity or convexity, and other features. The features used to describe a figure should contain information such that similar perceptive tasks can be performed. If two figures are perceived to be more similar to each other than to a given third figure, information about this condition should be present in the descriptive features. All of the features need not be mutually orthagonal, but the amount of interaction between features should be minimized. For example, a measurement of eccentricity that is not invariant to rotation will probably not be useful.

The following features are computed for each texture element.

1) AREA

Defined to be the number of points enclosed in the region of the texture element.

2) Mean brightness of the points in the region

3) P^2/A

P = length of the perimeter of the region

A = area of the region

The minimum value of P/A is 4π, for a circle.

Other values for familiar figures are 16 (square),

$8 + 4a/b + 4b/a$ (rectangle of sides a and b),

$36/\sqrt{3} \approx 21$ (equilateral triangle),

$2\pi (a^2+b^2)/ab$ (ellipse of axes a and b).

The computation of the remaining features involves the use of the spatial central moments of the texture element. The central moments are defined as follows.

$$u_{mn} = \sum p(x,y) (\bar{x}-x)^m (\bar{y}-y)^n ,$$

where $\bar{x}$ and $\bar{y}$ are the arithmetic means of x and y, p(x,y) is some function of x and y and in this case is just the constant 1, and the sum is taken over all of the points in the texture element. The moments are invariant to translation. The moments are normalized for size invariance by dividing by $(AREA)^2$. Since the result is often a proper fraction, the moments are multiplied by 100 to produce a rounded integral result. The normalized moments are used in all calculations.

The remaining features are defined as follows:

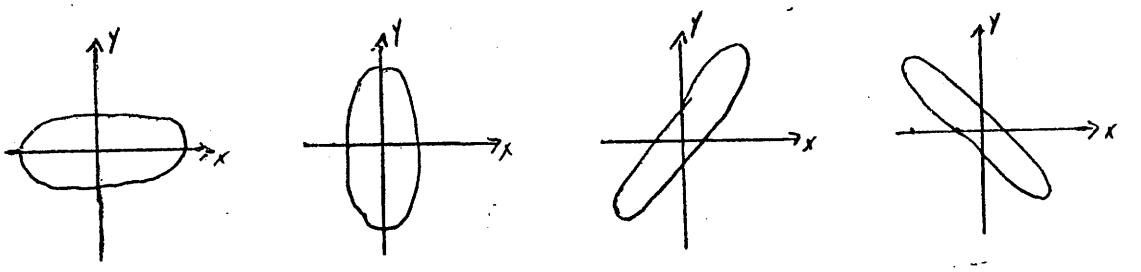
4) $u_{20} + u_{02}$

This sum is identical to $\sum d(x,y)$, where $d(x,y)$ = the distance from the point (x,y) to the point $(\bar{x},\bar{y})$. It is not invariant to rotation.

5) $\sqrt{(u_{20} - u_{02})^2 + 4u_{11}^2}$

This feature can be interpreted as a measure of eccentricity. Figure 5 illustrates the description of this feature. u_{20} is large if the spread of the texture element along the x-axis is large. u_{02} is large if the spread along the y-axis is large. The difference $|u_{20} - u_{02}|$ is large if the spread along one of the axes is small in comparison with the spread along the other axis. If the texture element is symmetric with respect to the origin (x,y), then $u_{20} - u_{02} = 0$. u_{11} is positive and large if the points of the origin-centered texture element lie predominantly in the first and third quadrants, negative and large if the texture element lies predominantly in the second and fourth quadrants, and is equal to zero if the texture element is symmetric with respect to the origin. The function

increases in value as the ratio of lengths of major to minor axis increases. The value of the funtion is invariant to rotation. The proof of this involves showing identity under transformation of coordinates and is straightforward.



(a) u_{20} large (b) u_{02} large (c) u_{11} large $u_{11} > 0$ (d) u_{11} large $u_{11} < 0$

Figure 5: Conditions for which the individual terms of the eccentricity feature are large.

6) Angle of orientation of major axis

tan $2\theta = 2u_{11} / (u_{20} - u_{02})$, where θ is the desired angle.

7,8) XWIDTH,YWIDTH

The width and length of the texture element relative to the major axis. These are the lengths of the sides of the smallest enclosing rectangle,

and are thus equal to or greater than the lengths of the minor and major axes, respectively. YWIDTH is always greater than XWIDTH (the normalized figure is always assumed to be standing upright).

9) Percentage of rectangle fill

The ratio of 100*AREA to XWIDTH*YWIDTH, that is the ratio of 100 times the actual area to the area of the enclosing rectangle.

The implementation of the features was verified by manually inputting the information describing a geometrical figure. The value for each figure was computed manually and compared with the program-generated result.

4.3 The experimental scenes

The experimental scenes chosen are handdrawn artificial textures consisting in the repetition throughout the scene of a particular texture element, drawn in black pen on white paper. This restricted class of scenes was chosen because

1) The texture element finder works best with such scenes;
2) It was desired to test the goodness of the features used in characterizing the texture elements.

In regard to the latter point, if all of the texture elements examined are similar to each other, then the histograms of the features should reflect this similarity. Those features which measure a property for which the texture elements have similar values should yield fairly peaked histograms. For example, if the texture elements are circles, then the P^2/AREA histogram should be strongly peaked. Those features which measure properties for which the texture elements have widely differing values should yield more uniform distributions. For example, the feature measuring the angle of orientation of major axis for circles should give a uniform distribution. If a feature fails this test of goodness, then it is unlikely that it is sufficeiently sensitive to be used in characterizing a wide class of texture elements.

No attempt was made to make the texture elements in the drawings exact duplicates of each other; on the contrary, in order to provide a good test for the features it was desired to make the texture elements similar enough so that a human observer would group them together, but at the same time to have no two texture elements identical. The result is a texture in which the texture elements are stochastically perturbed, yet perceptually equivalent.

There were ten test scenes in all, nine of which

were different and one of which consisted in a composite of two already tested scenes. In six of the scenes the texture consisted in occurrences of a single texture element, all of which were similar in appearance, yet not identical. These scenes were used to measure the individual sensitivity of each feature. A composite of two of these scenes was made by finding texture elements in two different scenes. The feature histogram in this case was the superposition of the two individual histograms. A scene in which two different types of texture elements were present was also used. a scene was used in which the texture elements were not stochastically perturbed occurrences of an ideal model, but rather were rather loosely similar. This scene looks somewhat of a cross between wood grain and a fingerprint. This scene was used to measure how sensitive the features are to higher-level similarities. Finally, a wood block in which the natural wood grain was prominent was used to measure both the goodness of the features and the ability of the texture element finder on scenes which are not binary figure-ground.

For all text scenes, texture elements consisting of less than 25 points were rejected. The acceptable intensity deviation value was set at 16 for all scenes except the natural wood grain, for which best results were obtained when the value was set to 8. For all scenes except the

natural wood grain, the composite scene, and the scene in which there were two texture element types present, 35 texture elements were found. For the latter three scenes, the number of texture elements found was 20, 70, and 70, respectively.

4.4 Experimental results

The experimental scenes, a plot of the texture elements found in each scene, and the histograms of each feature for each scene are contained in the appendix. It is observed that the histograms of some of the features display marked localization for the given scenes, even though the number of samples in each scene is not very large.

The results of the first six scenes afford a qualitative evaluation of the features used.

The values of the AREA feature consistently vary the most. There is not one instance in which the values are peaked around a small range of values. The variation is consistently approximately a factor of two, even though the figures were drawn to be approximately equal. This result suggests that the measurement of area differences of less than a factor of two will be of little use in characterizing texture elements.

Since the scenes used are approximately binary in intensity, the brightness feature is of relatively little

interest. There is a variation in brightness of about five percent among the textures. Except for the wood block scene, the histograms of the brightness feature are not included in the appendix.

Some information has been lost in the current implementation of the major axis angle feature, since the results are limited to the range of -45 to +45 degrees. The results indicate that the feature is fairly sensitive nevertheless. For the dots, triangles, and double-ended arrows, a uniform probability distribution is described. The results for the bricks are peaked close to zero. Most of the values for the ellipses lie in a 20 degree range. This somewhat large range is not so suspicious as it might appear, for the texture elements in the scene apparently have such a spread.

The results for the percentage of fill, XWIDTH, YWIDTH, moment function 1 (i.e., $u_{20} + u_{02}$), eccentricity, and P^2/AREA are generally quite peaked and localized, although in some of the scenes the texture element finder constructed an element which differed radically in appearance from the actual element. These bogus elements are indicated with arrows in the drawings. They are the cause of the occurences of values for some of the features which are widely separated from the main body of values. XWIDTH values are localized and peaked in all six scenes. YWIDTH and

percentage of fill are good for five scenes and somewhat less localized but still peaked in the sixth. The two central moment functions and the P^2/AREA features yield peaked and localized results in five of the scenes. These results suggest that, with the exception of the area and brightness, the features used characterize the texture elemtents fairly well. (Note: Because of the discrete representation of the perimeter of the figures, the perimeter is found to be slightly less than if the figures were continuous. This accounts for the values of P^2/AREA being slightly lower than they would be for continuous figures.)

For the two scenes in which two types of texture elements are present, at least some of the features should be bimodal in distribution. Otherwise, the two texture element types are indistinguishable in feature space, while they are distinguishable by humans. As expected from the previous results, the AREA and brightness feature fail in this respect. Bimodality in distribution is observed for the eccentricity, angle of major axis, and P^2/AREA features. Bimodality is not observed in the distributions of the remaining features for the most part; the two texture element types have similar values for those features. It seems likely that additional good features will be required to accurately discriminate between arbitrary texture

elements.

The results of the tests with the artificial wood-grain/fingerprint and with the natural wood grain indicate that a better texture element finder must be used if natural textures are to be examined. It is difficult to evaluate the goodness of the features in characterizing these textures, because of the significant variation in the performance of the texture element finder. Some degree of peaking is observed for some of the features, and it cannot be ruled out that the features may adequately characterize the wood grain texture elements if the results of the texture element finder were improved.

5. Extensions of the model

The class of textures considered is a very restricted one. This class can be extended to include a broader class of textures. A description of some of these extensions follows.

—Complex unit patterns.

The unit pattern is not a single instance of a single texture element type, but rather is composed of instances of one or more texture element types arranged in a specific pattern. It is the complex unit pattern which is replicated throughout the scene. Figure 6 illustrates this.

—Several unit patterns

Instances of not one but of several distinct unit patterns are replicated to form the texture. Each unit pattern may be simple or complex (Figure 7).

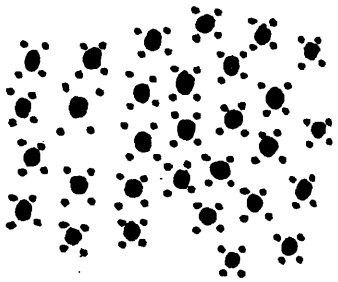


Figure 6: Complex unit pattern.

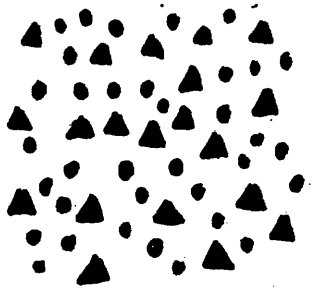


Figure 7: Several unit patterns.

—Consideration of placement information

The manner in which the unit patterns are placed in the scene is perceptually significant. Two textures in which the unit patterns are identical but in which different placement procedures are used can be distinguished. Placement variables are the density of the instances of unit patterns per unit picture area, the description of placement pattern regularity (or non-regularity), a type of placement process (figure-ground or texture element adjacency), and so

on.

There are many simple tricks which can be used to deduce placement information from a scene. Inter-center distance between texture elements, collinearity of texture element centers, detection of an area in the scene composed only of one type of texture element. Higher-order detection processes are needed to complement the low-order measures.

—A unit pattern not specified by appearance alone

Texture elements need not be absolutely similar in appearance (in the autocorrelation sense) for one to perceive them as being similar. Wood grain provides an example of a higher-order similarity in appearance by which texture elements can be grouped into perceptual equivalence classes.

Grouping of objects in scenes often occurs due to similarities not only in appearance but also in meaning or function. Bricks in a wall, blades of grass in a lawn, fish in a school look similar to each other and also have a common meaning or classification on grounds other than perceptual.

In some cases, the common function of the objects in a scene may be more prominent than the visual similarities between the objects. As an example, consider a filled parking lot seen from a height. There are many different

kinds of cars present displaying marked differences in their outward appearances, from VW's to Cadillacs. Yet a human has no trouble making the generalization that they are similar because they are all automobiles and that this particular arrangement of automobiles can be labelled "a filled parking lot." Similarly, an aerial view of a city exposes 50 story office buildings and three story town houses, yet one immediately groups these forms together as buildings and applies the label "aerial view of a city" to the scene. If one looks at the contents of a tool box, similarities in function are certainly of more use than similarities in appearance in grouping the objects in the box. It is clear that high-level knowledge about the world is used in the grouping of similar objects in many scenes.

The method of describing textures used in this work appears to be expandable to describing complex textures. The texture element finder can be used to find occurences of different texture element types. The histograms of some of the features should show different localizations, corresponding to different ranges of feature values for different texture element types. A grouping or clustering operation should yield a distinct equivalence class for each texture element type.

Placement rules for a texture should be a part of

the description of the texture. A method for deducing placement patterns of elements in a picture is unknown at present. Presumably, the information necessary to perform such an activity must be obtained by examining the locations of the elements in the picture. The present texture element finder provides this information.

The greatest difficulty in using measures of appearance to characterize texture elements is that much information is lost in the transformation from picture space to feature space. Humans are more likely to characterize (at least verbally) a given figure not with a feature vector, but rather in terms of familiar object (e.g., triangle, dumbbell). Such an object identifier would make short work of the simple textures used in this work, because the similarity between the texture elements is more apparent at a higher level (e.g., "all look like triangles" vs. "noticeable peak in P^2/AREA").

6. Perception of texture

This work is directed toward the ability to perceive a texture as contiguous, that is to automatically determine that some section of a scene is a texture of a particular description, and toward the ability to discriminate between different textures in the same scene. For example, if the scene was figure 8 we would like to say that there are two

distinct textures present, one occupying the left half of the picture, the other occupying the right half, and to be able to determine where the boundary between the textures is. As noted in the introduction, this has direct applicability to our vision system.

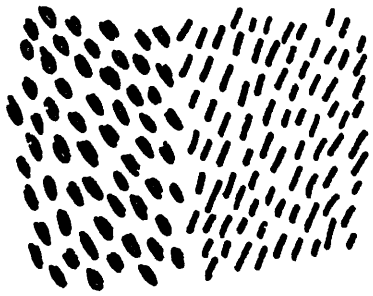


Figure 8: Adjacent distinct textures.

Since little work has been done on the structural discrimination of textures, it is difficult to credibly propose a method for doing so. For such simple textures as those used in this study, a grouping operation in feature space is a likely candidate for use in a simple texture discriminator. Placement information must also be incorporated in the discriminator.

When one considers the great versitality of human perception, it seems clear that a texture discriminator whose domain includes anything but very simple textures must

be a very complex process indeed. Consider the texture illustrated in figure 9. Humans have no trouble finding the boundary line between the two halves of the scene, even though it is a boundary between two textures which are essentially identical. Pictures of biological and man-made camouflage, in which an object presents an appearance essentially identical to its surroundings, are examples in which the limit of human ability for detection of boundaries between textures is approached. It is difficult to imagine how the detection of the moth on the tree bark in the familiar example of this type could be described. Clearly, A good texture discriminator requires high-level perceptual abilities about which we now have very little understanding.

Figure 9: Adjacent identical textures.

7. Conclusions

For the experimental scenes used, the feature histograms seem to offer a credible description of the texture elements. There is a marked histogram localization in many instances. a comparison of the corresponding features of different scenes shows a separation in the probability distributions for at least some of the features when different textures are compared. A possible test of the goodness of the description is a grouping operation in feature space. If the description is good, such a test should yield a clustering together of all of the perceptually equivalent texture elements as instances of an ideal model of the texture elements. Such a test is as yet unperformed.

The results of the natural wood grain scene offer hope that some version of this method may be useful in characterizing natural textures. Natural textures are seldom restricted to binary figure-ground scenes in which texture elements are perceptually similar. This work can be considered at best a starting point for additional development.

Additional features should be sought and tested. In particular, none of the features used in this study measure symmetry, pointedness (possession of sharp corners), concavity or convexity.

Natural textures generally have great variations in shading, and tactilely textured objects diffuse light reflected from their surfaces. Because of this, better ideas about finding texture elements will be required before good results can be obtained with natural textures.

The description method should be extended to be applicable to complex textures. In particular, very little is known about extracting placement rule information from a scene. Future work should address these issues.

Bibliography

1. Anderson, G., Frequency-Domain Image Representation, Sc.D. Thesis, Dept. of E.E., M.I.T., 1970.

2. Attneave, F., "Multistability in Perception", Scientific American, Vol. 225, No. 6, December, 1971.

3. Butt, Johnston, Kolmus, Lee, Rosenfeld, Troy, Wells, Studies in Visual Texture Manipulation and Synthesis, Report 64-94, Computer Science Center, U. of Maryland, June, 1969.

4. Brodatz, P., Textures, New York: Dover Press, 1966.

5. Gibson, J., The Perception of the Visual World, Boston: Houghton Mifflin, 1950.

6. Hu,M.K., "Visual Pattern Recognition by Moment Invariants", IRE Transactions on Information Theory, IT-8, No. 2, February, 1962.

7. Julesz, B., "Cluster Formation at Various Perceptual Levels", in Watanabe, S., Methodologies of Pattern Recognition, New York: Academic Press, 1969.

8. Julesz, B., "Texture and Visual Perception", Scientific American, Vol. 212, No. 2, Feb., 1965.

9. Julesz, B., "Visual Pattern Discrimination", IRE Transactions on Information Theory, IT-8, Feb., 1962.

10. Krakauer, L.J., Computer Analysis of Visual Properties of Curved Objects, Ph.D. Thesis, Dept. of E.E., M.I.T., June, 1970.

11. Leeuwenberg, E.L., Structural Information of Visual Systems, Mouton & Co., The Hague, 1968.

12. Lipkin, B., "Psychopictorics and Pattern Recognition", SPIE Seminar Proceedings, Pattern Recognition Studies, Vol. 18, June 9-10, 1969.

13. Pickett, R., "The Perception of Visual Texture", Journal of American Psychology, Vol. 68, No.1, 1964.

14. Pickett, R., "Visual Analyses of Texture in the Detection and Recognition of Objects", in Lipkin and Rosenfeld, Picture Processing and Psychopictorics, New York: Academic Press, 1970.

15. Rosenfeld, A., "Figure Extraction", in Automatic Interpretation and Classification of Images, Grasselli, A. (ed.),Academic Press, New York, 1970.

16. Rosenfeld, A., Lipkin, B., "Texture Synthesis", in Picture Processing and Psychopictorics, Academic Press, New York, 1970.

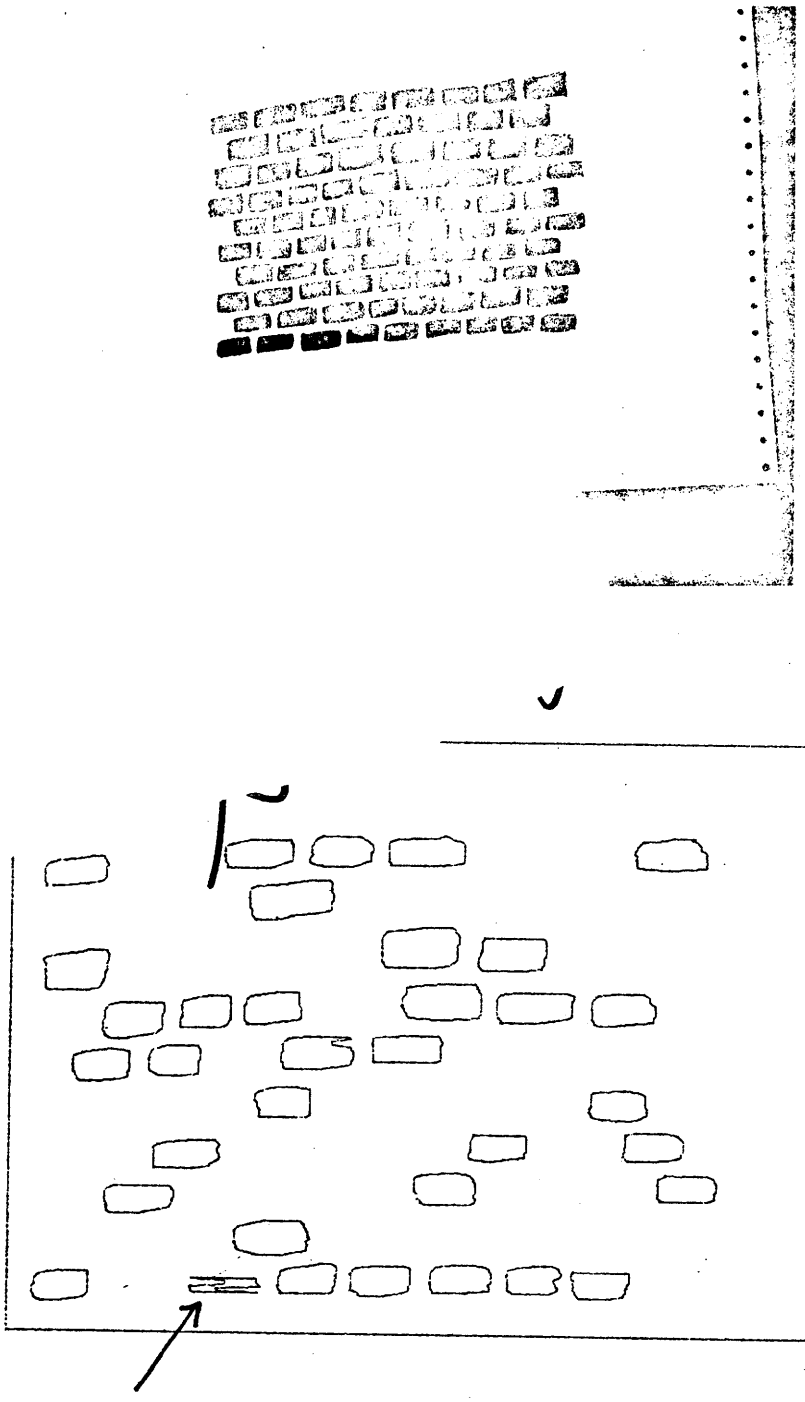
17. Rosenfeld, A., Troy, E., "Visual Texture Analysis", I.E.E.E. Symposium on Feature Extraction and Selection in Pattern Recognition, Oct. 5-7, 1970.

18. Simon, H., Sciences of the Artificial, M.I.T. Press, Cambridge, Mass., 1969.

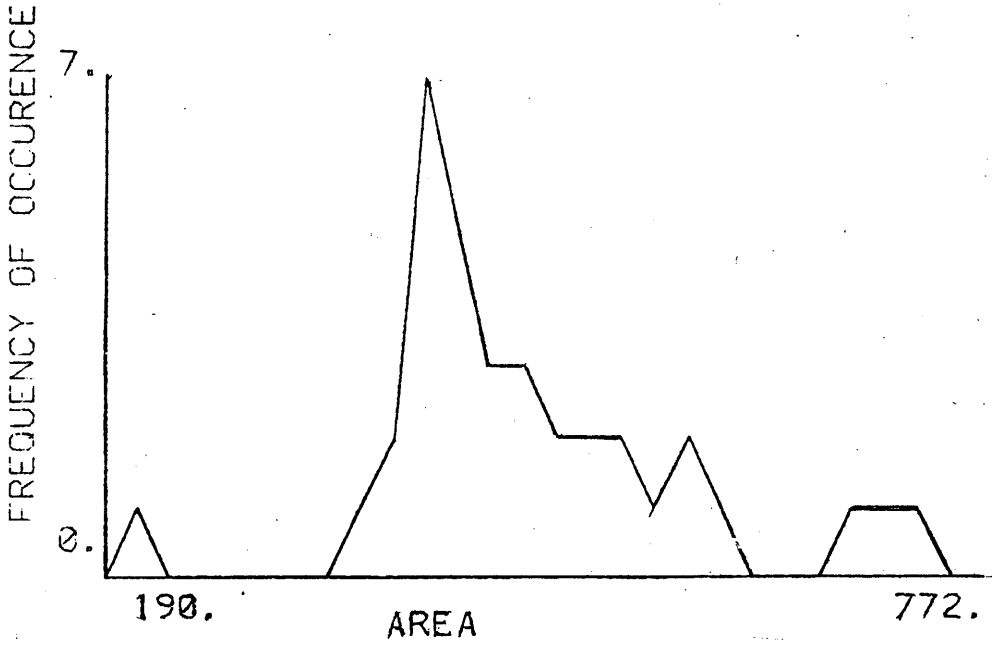
Appendix: Results in graphical form

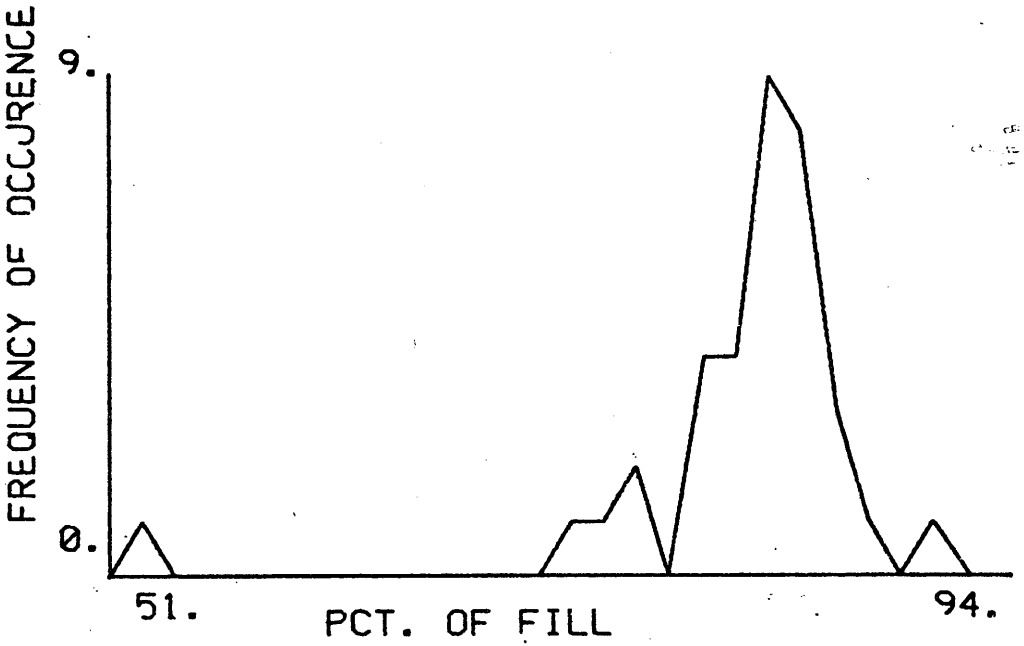
For each of the ten scenes used, a photograph of the scene, the result of the texture element finder, and the eight feature histograms are shown in the figures. For the natural wood grain, the brightness histogram is also shown.

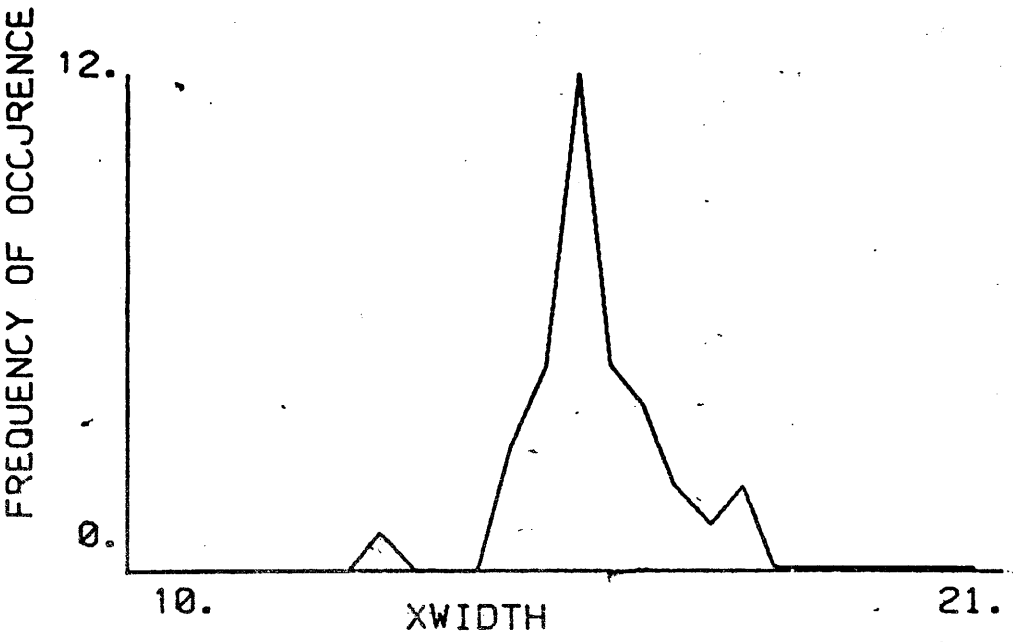
Each histogram has 27 positions, of which the first and last, for the sake of ease of interpretation, are not used to represent the distribution of the feature values. The remaining 25 positions represent the frequency of occurrence of feature values, from the minimum to the maximum value for that feature. If the range of values is less than 25, each position corresponds to a single integral value of the feature, and the distribution is centered in the histogram. The x-axis limits printed below the graph correspond to the actual minimum and maximum values of the feature values. If the range of frequency values is greater than 25, each histogram position may correspond to several integral values of the feature, such that the 25 positions together cover exactly the spread of feature values.

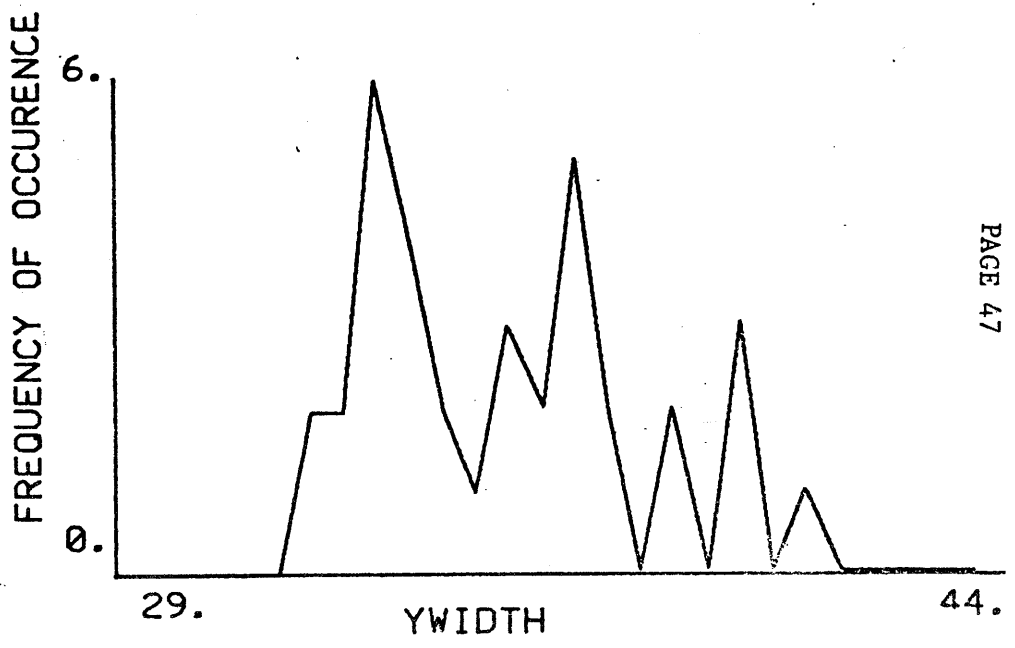


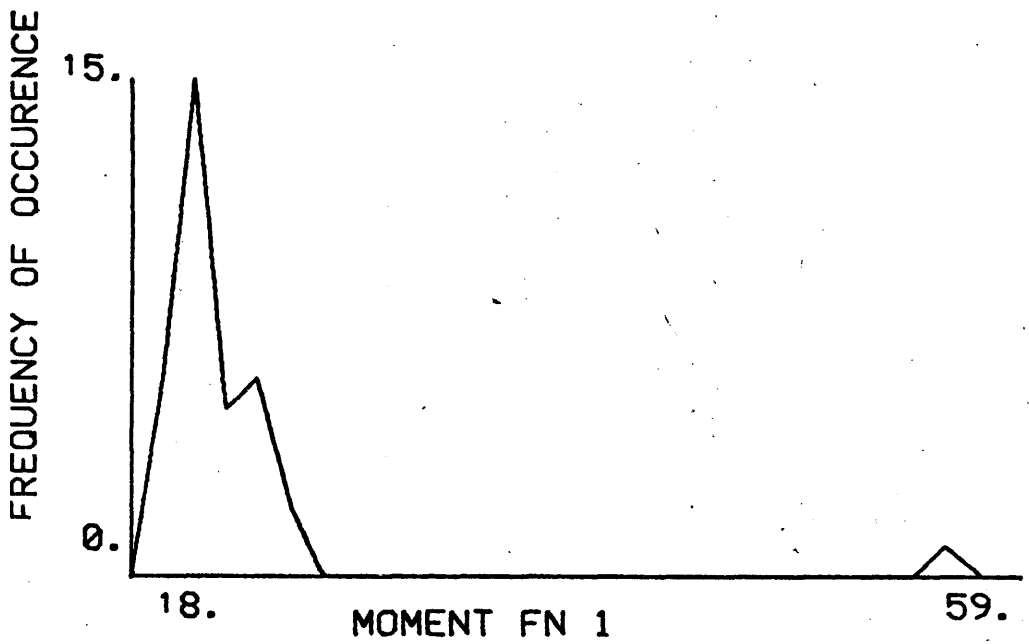
Arrow indicates constructed texture element which does not match humanly perceived texture element.

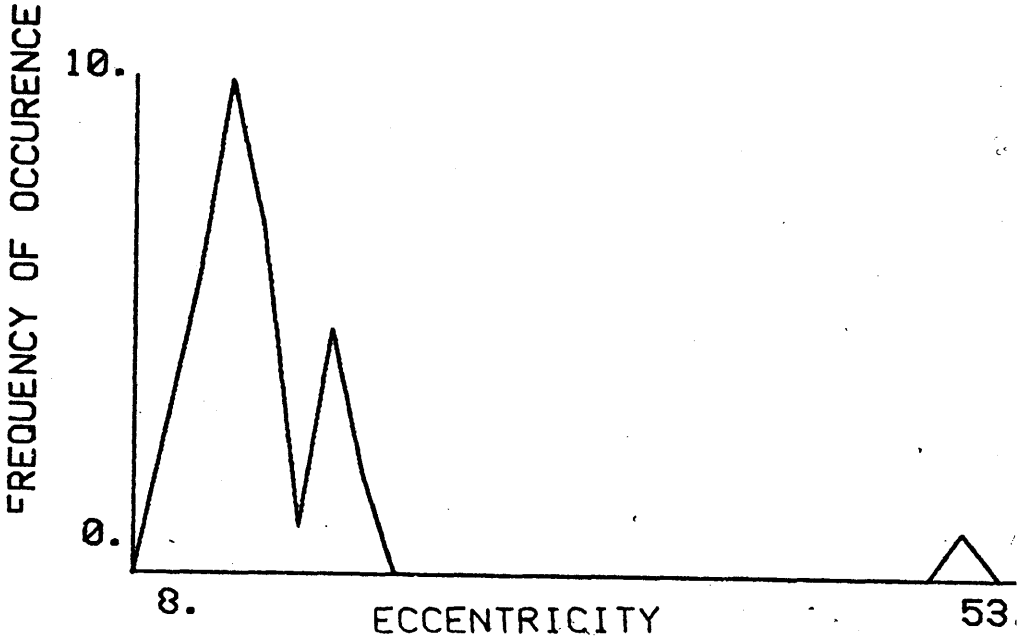

FREQUENCY OF OCCURENCE
7.
0.
190.
AREA
772.

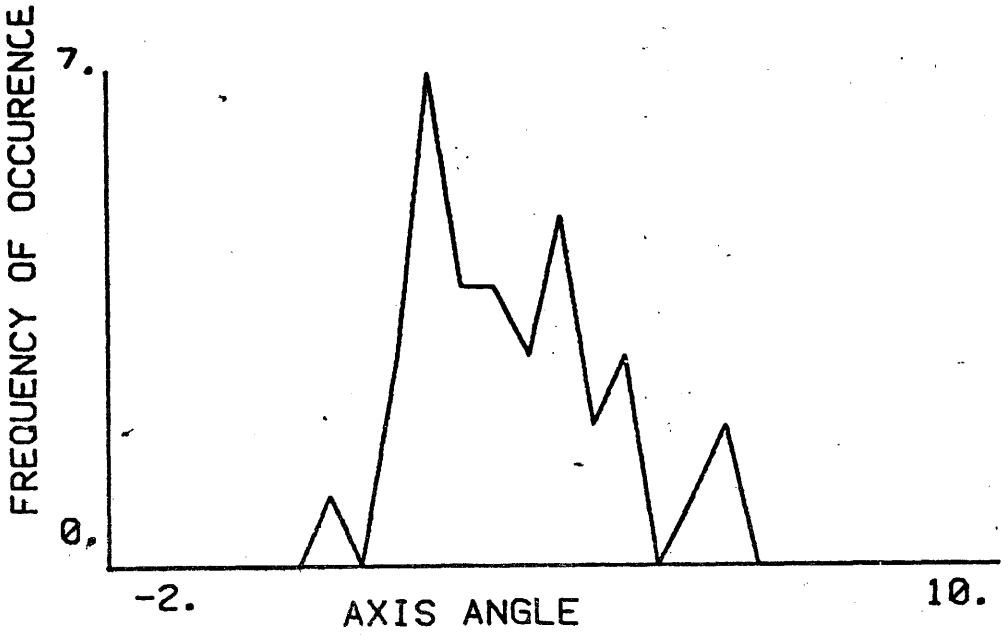

FREQUENCY OF OCCURENCE
9.
0.
51.
PCT. OF FILL
94.

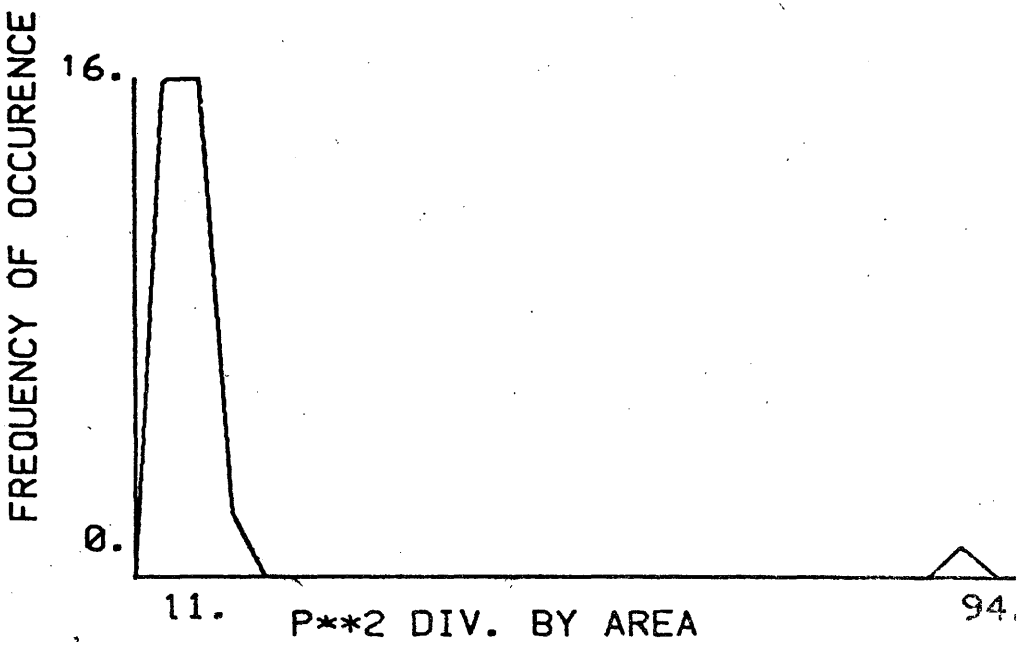

FREQUENCY OF OCCURENCE
12.
0.
10.
XWIDTH
21.

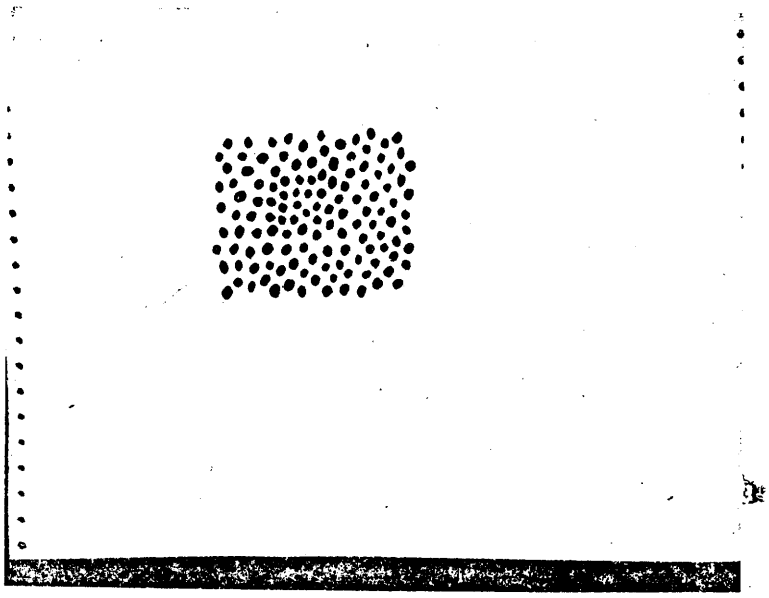

FREQUENCY OF OCCURENCE
6.
0.
29.
YWIDTH
44.

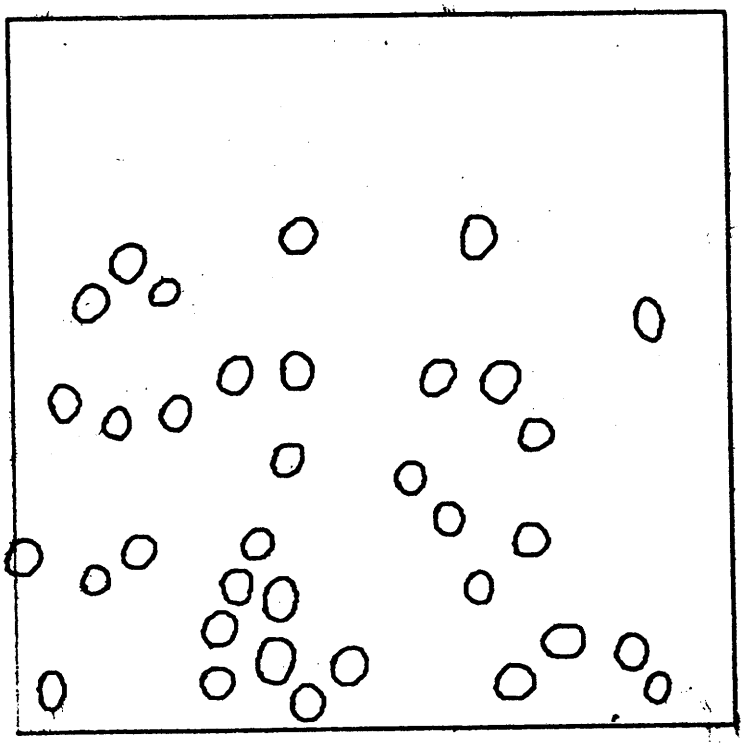

FREQUENCY OF OCCURENCE
15.
0.
18.
MOMENT FN 1
59.

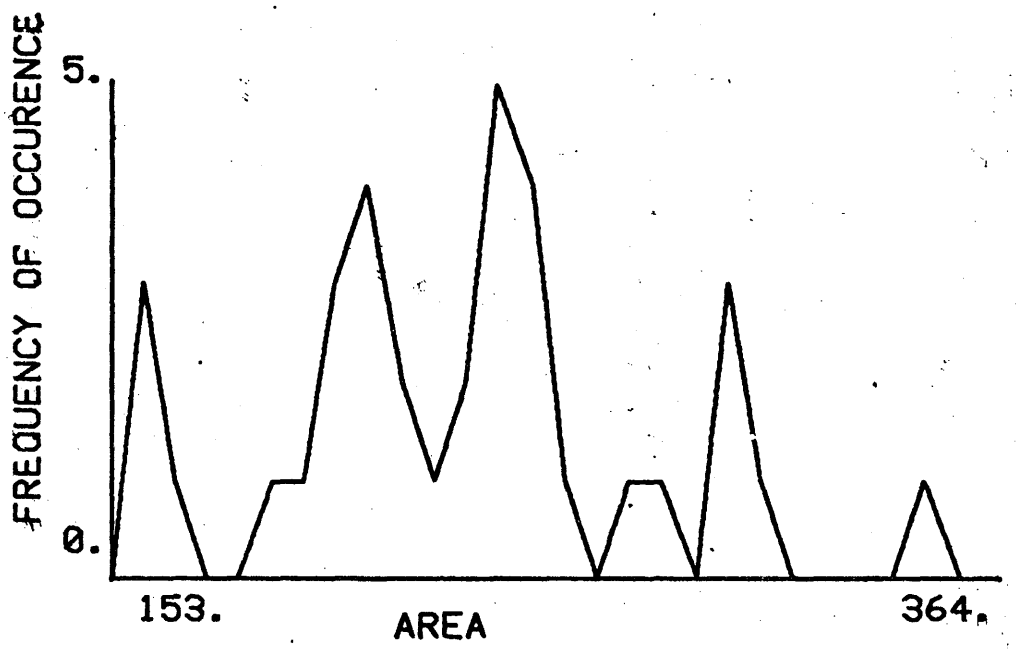

FREQUENCY OF OCCURENCE
10.
0.
8.
ECCENTRICITY
53.

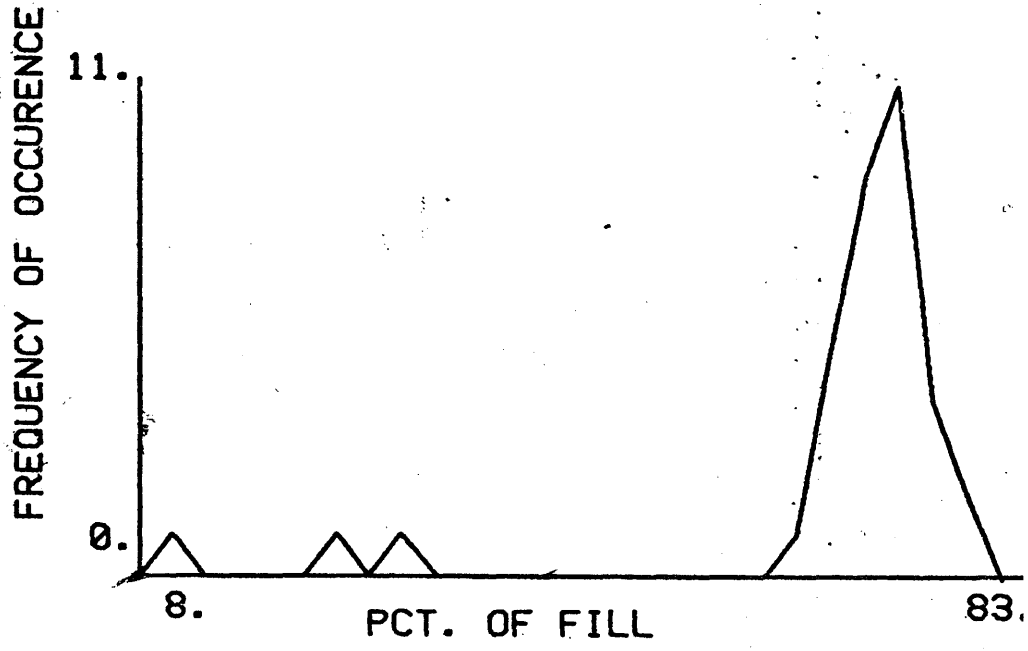

FREQUENCY OF OCCURENCE
7.
0.
-2.
AXIS ANGLE
10.

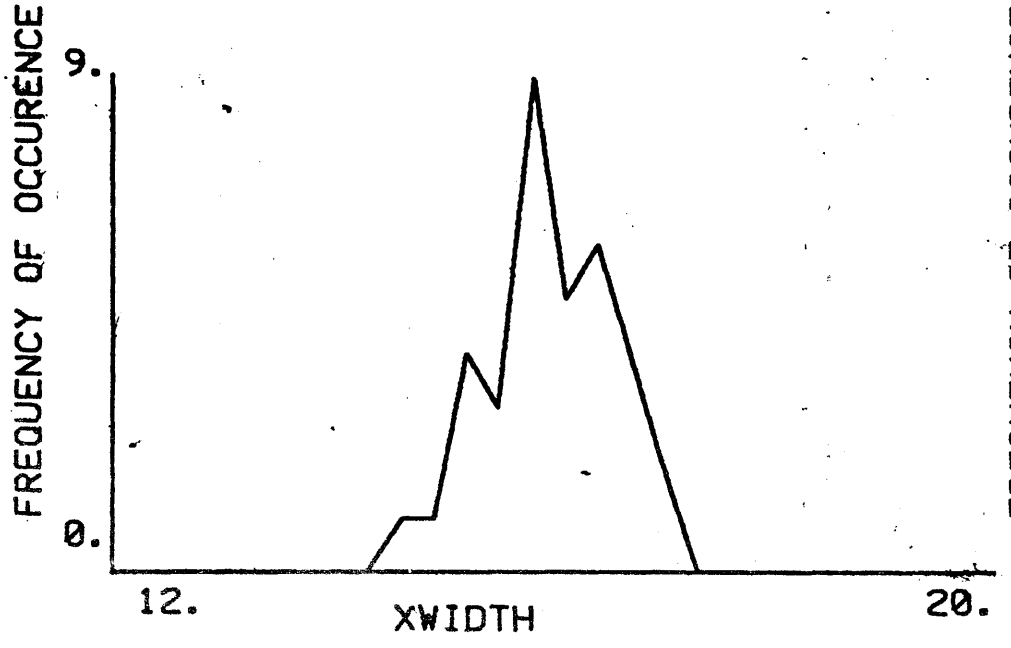

FREQUENCY OF OCCURENCE
16.
0.
11.
P**2 DIV. BY AREA
94.

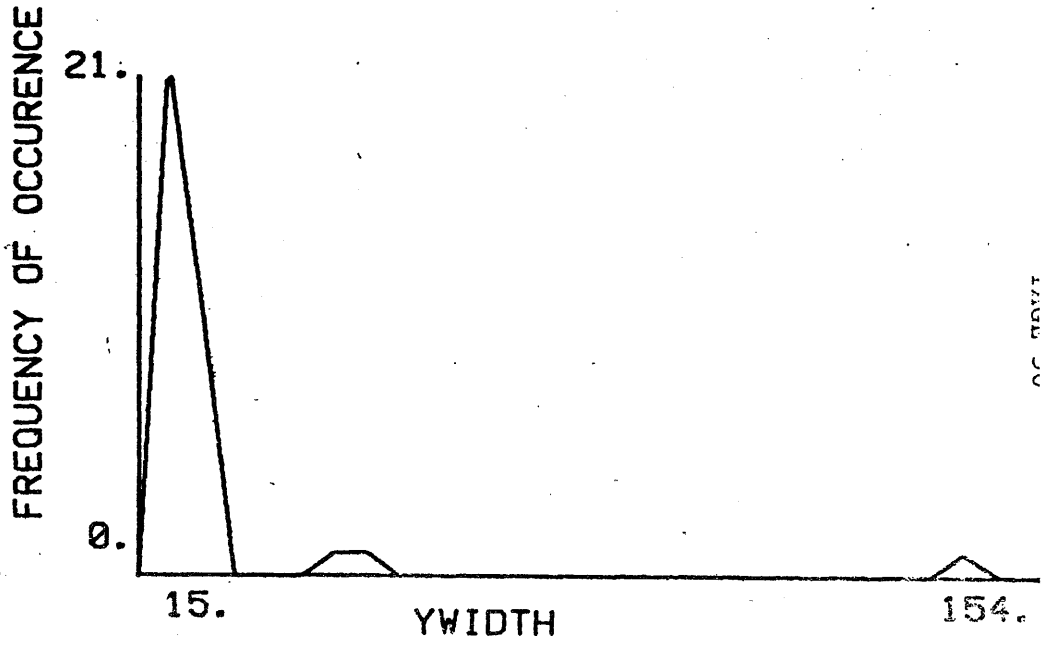


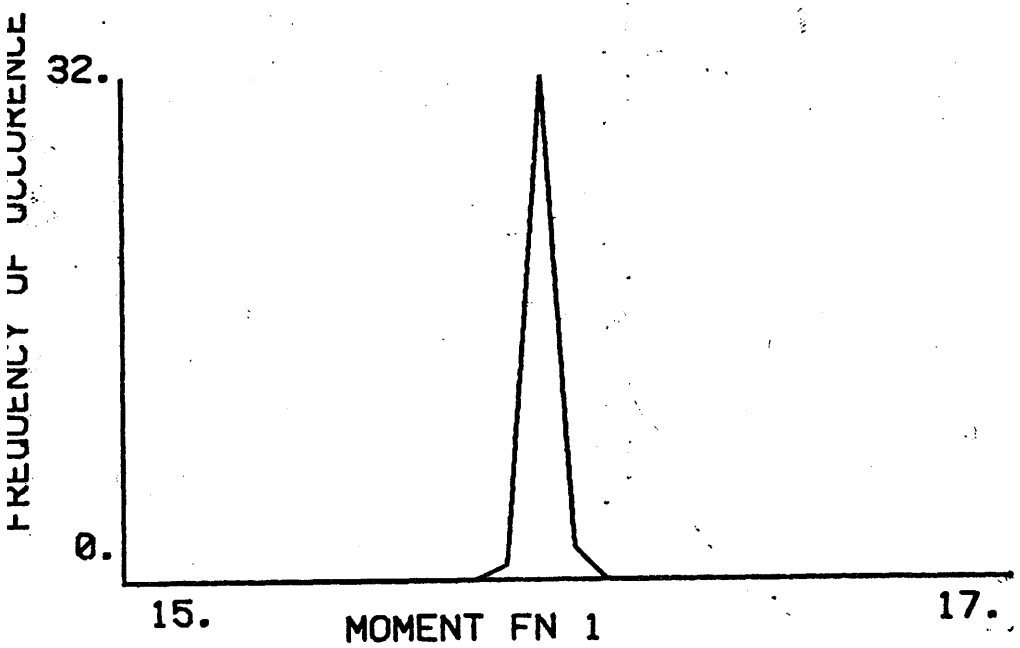


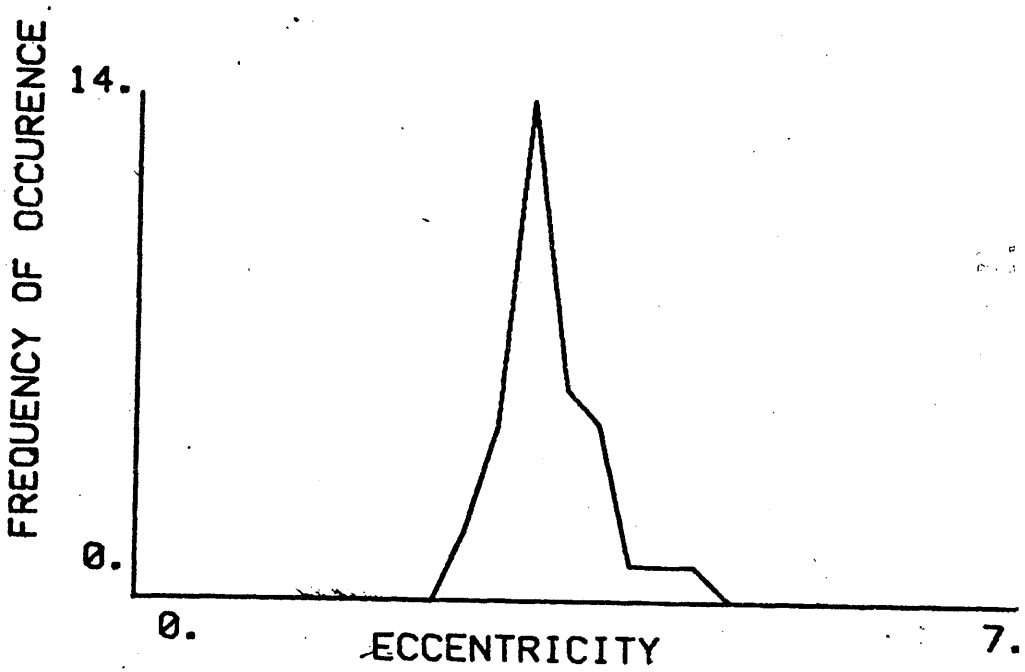

FREQUENCY OF OCCURENCE
5.
0.
153.
AREA
364.

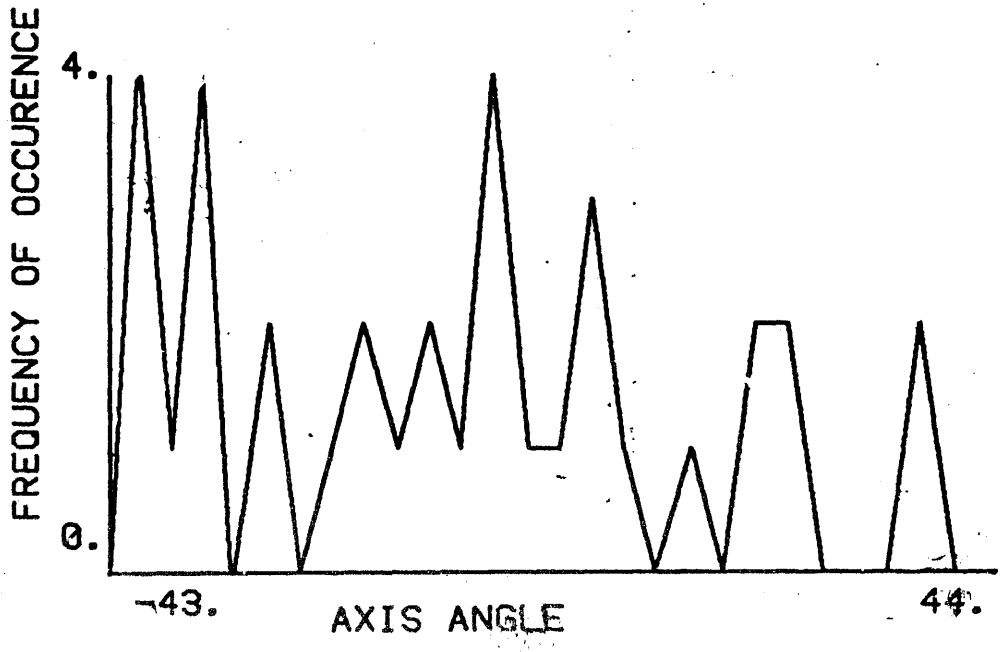

FREQUENCY OF OCCURENCE
11.
0.
8.
PCT. OF FILL
83.

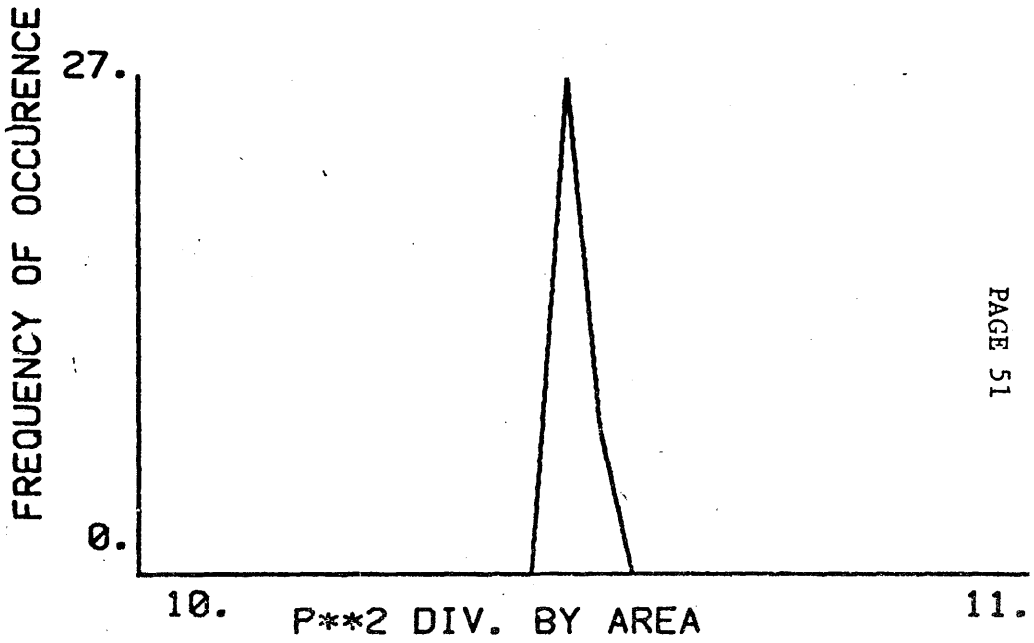

FREQUENCY OF OCCURENCE
9.
0.
12.
XWIDTH
20.

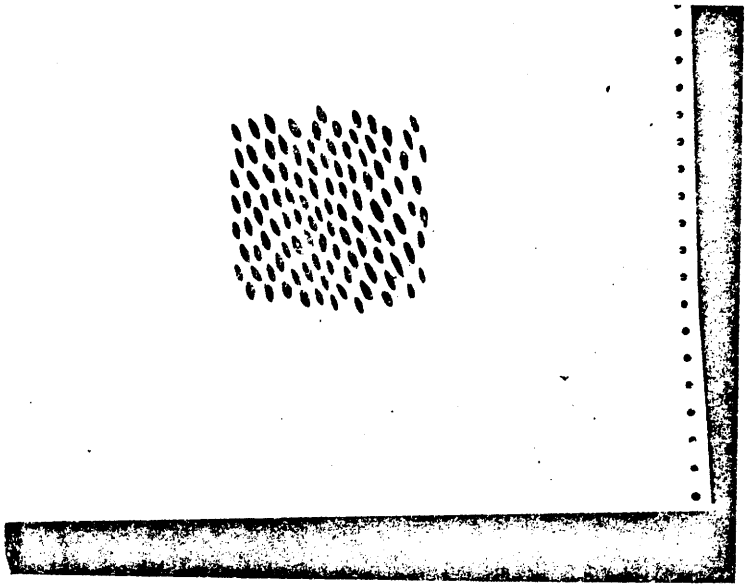

FREQUENCY OF OCCURENCE
21.
0.
15.
YWIDTH
154.

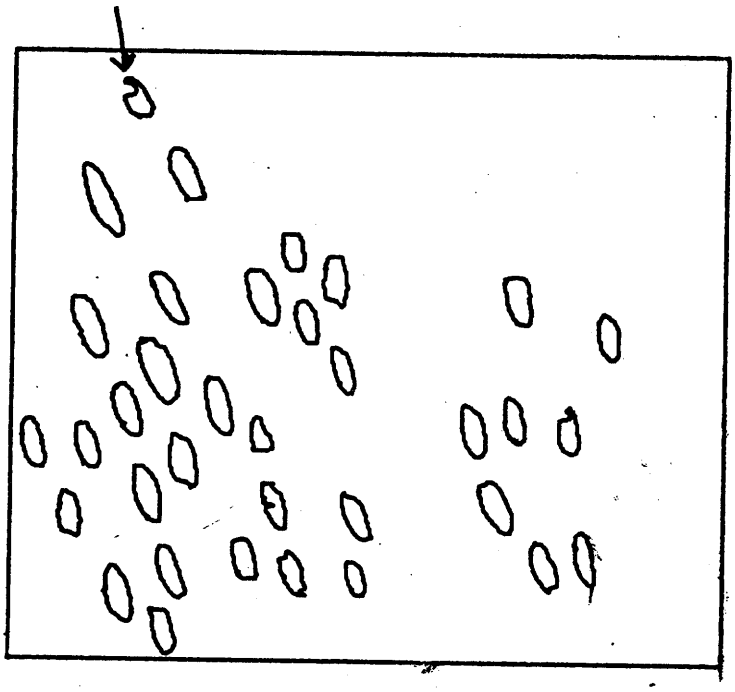

FREQUENCY OF OCCURENCE
32.
0.
15.
MOMENT FN 1
17.

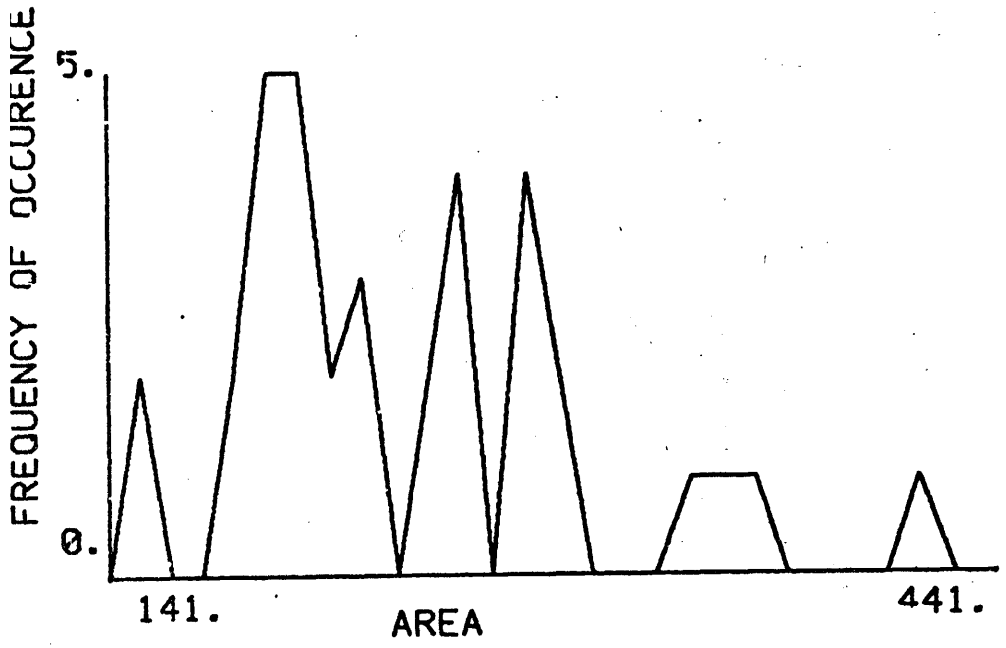

FREQUENCY OF OCCURENCE
14.
0.
0.
ECCENTRICITY
7.

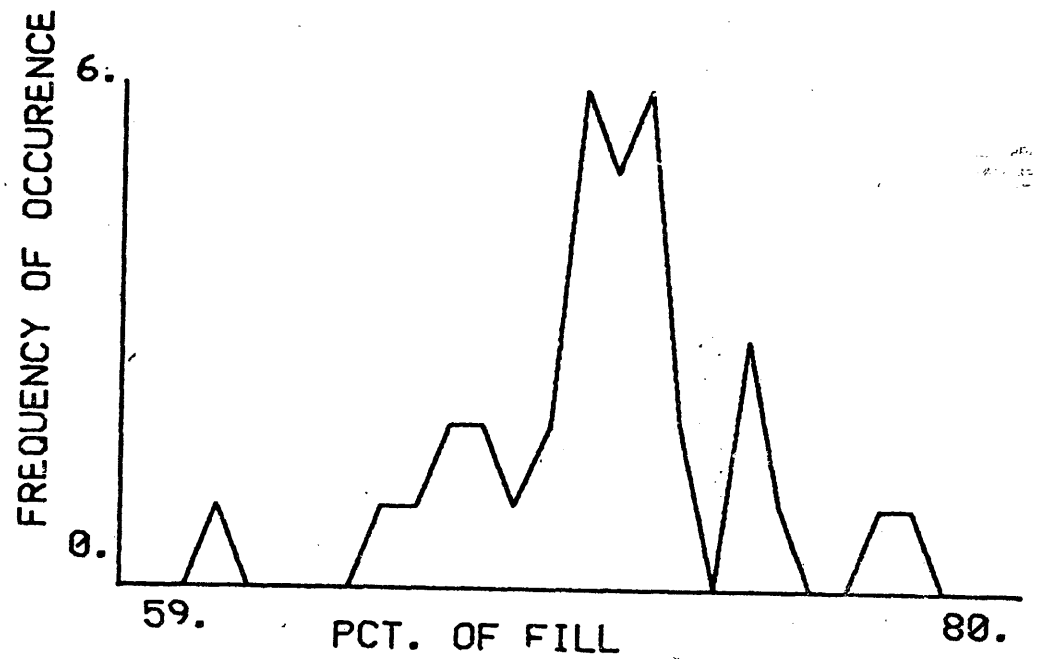

FREQUENCY OF OCCURENCE
4.
0.
-43.
AXIS ANGLE
44.

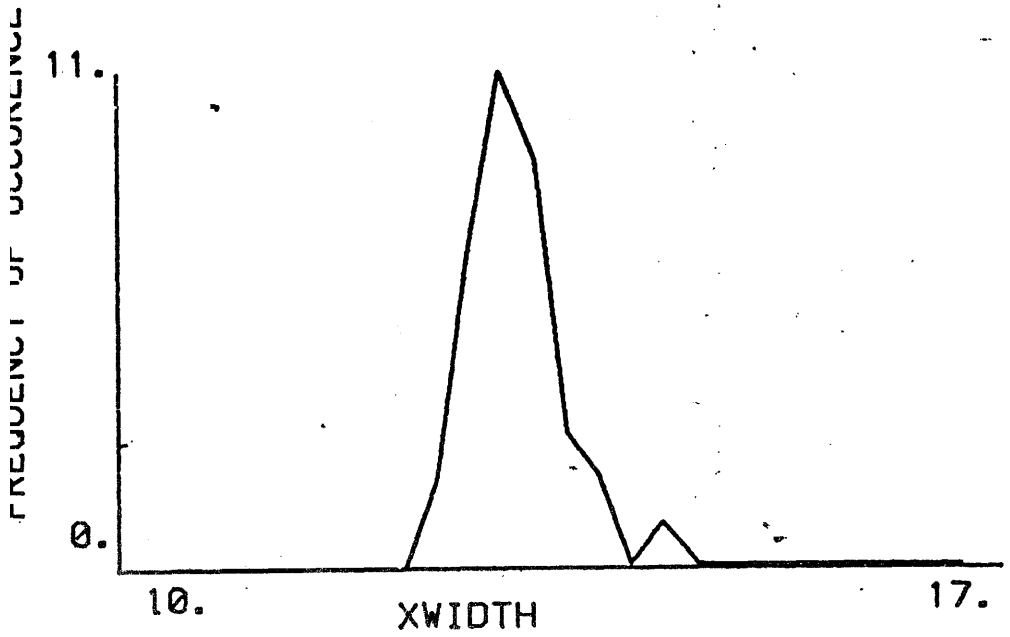

FREQUENCY OF OCCURENCE
27.
0.
10.
P**2 DIV. BY AREA
11.

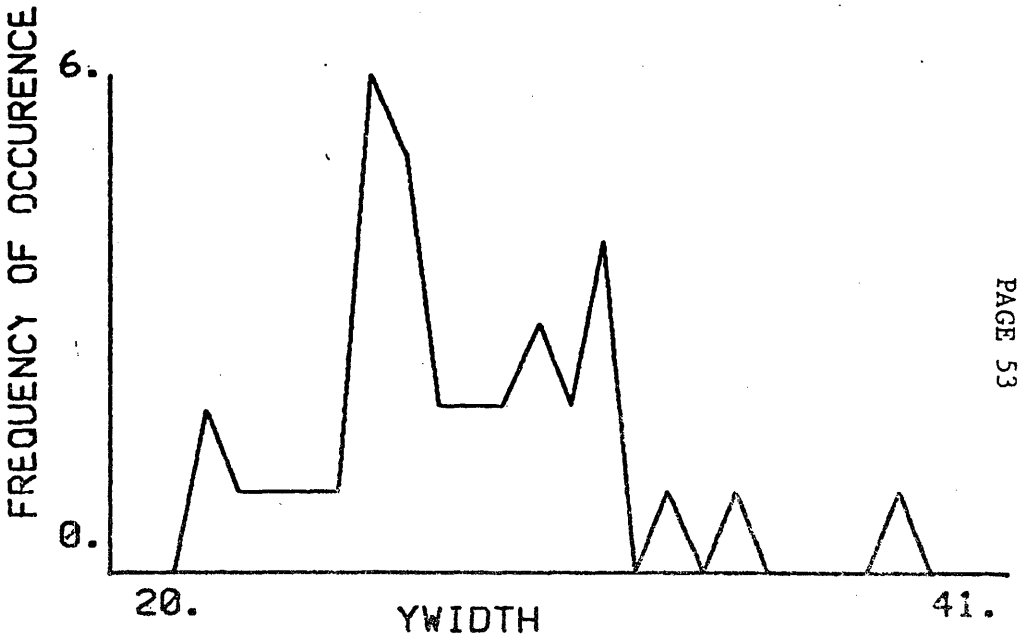


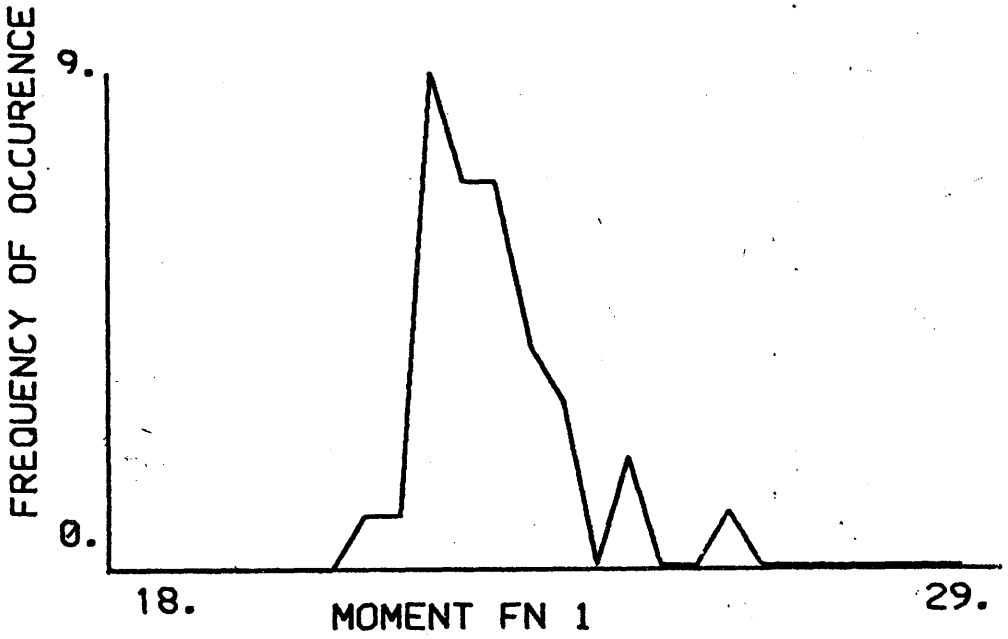


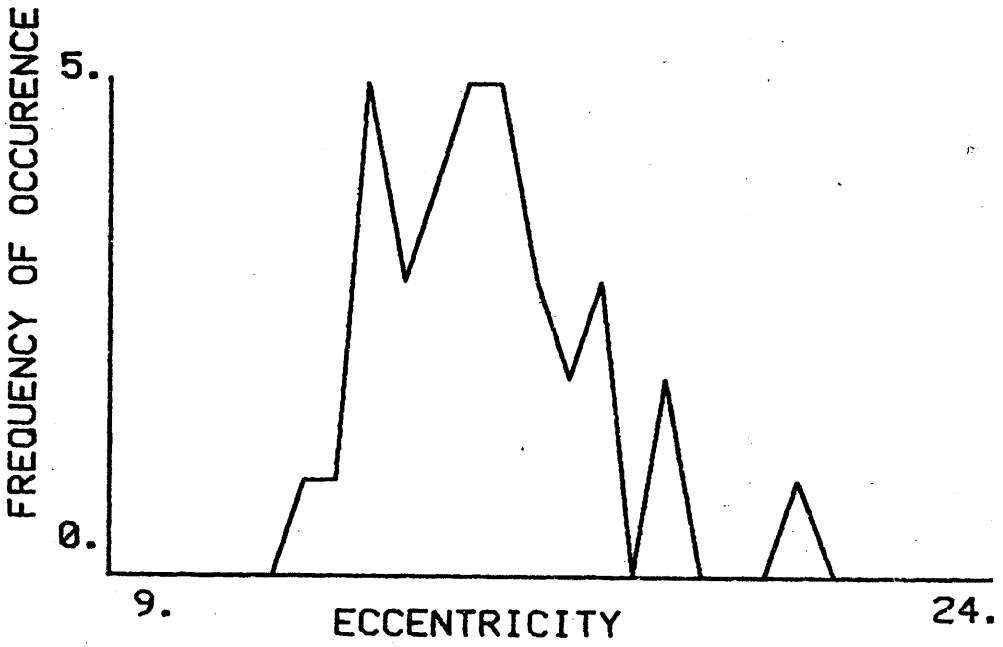

FREQUENCY OF OCCURENCE
5.
0.
141.
AREA
441.

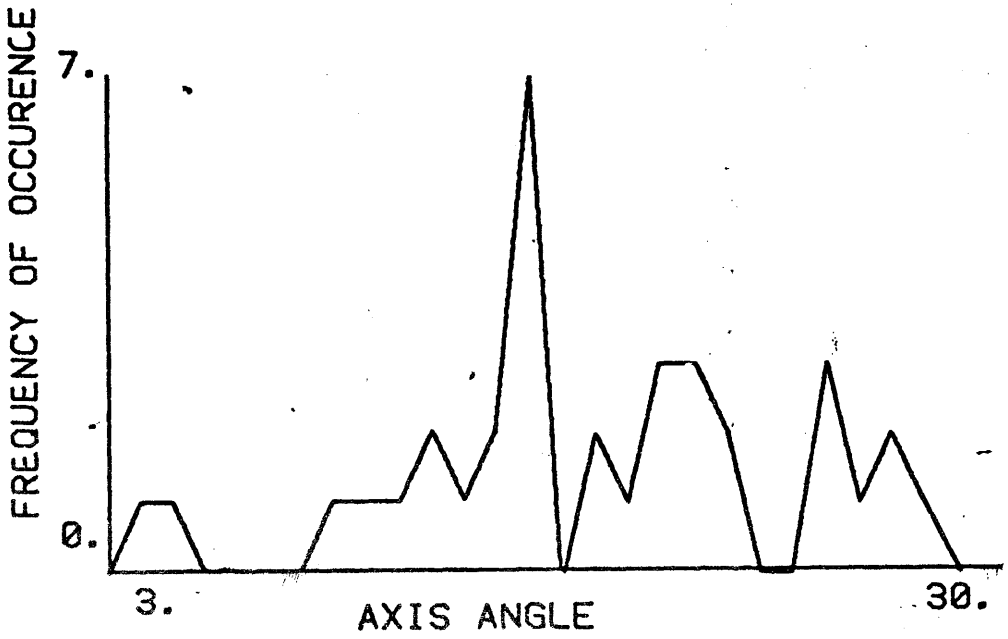

FREQUENCY OF OCCURENCE
6.
0.
59.
PCT. OF FILL
80.

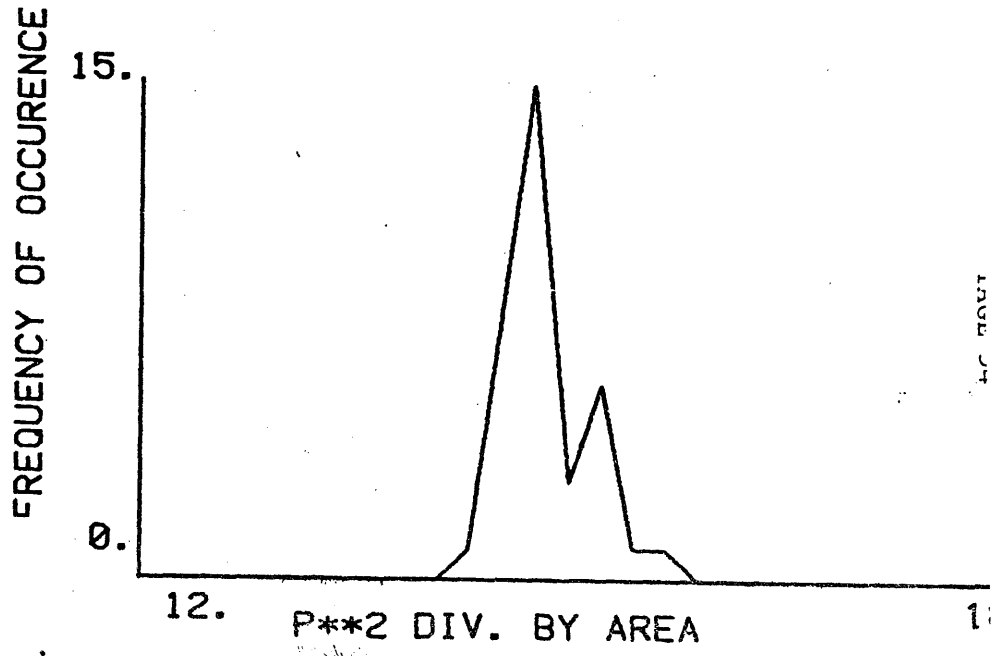

11.
0.
10.
XWIDTH
17.

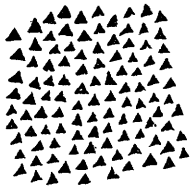

FREQUENCY OF OCCURENCE
6.
0.
20.
YWIDTH
41.

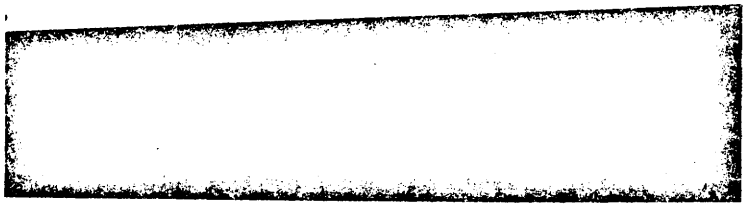

FREQUENCY OF OCCURENCE
9.
0.
18.
MOMENT FN 1
29.

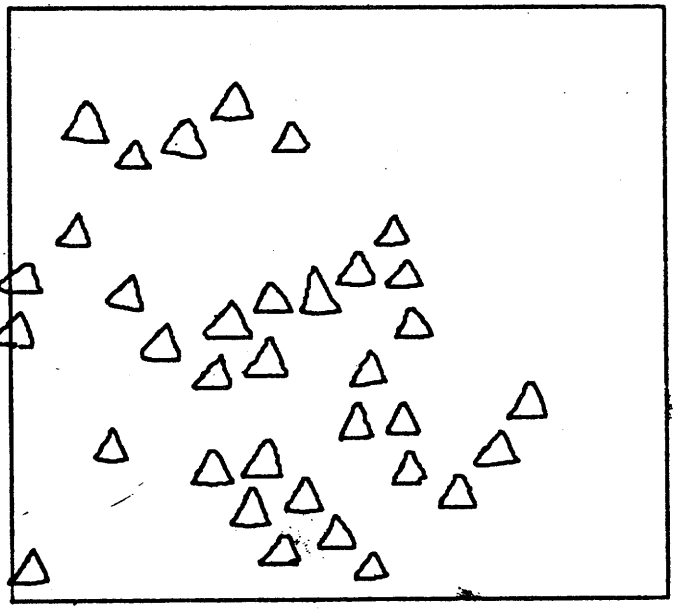

FREQUENCY OF OCCURENCE
5.
0.
9.
ECCENTRICITY
24.

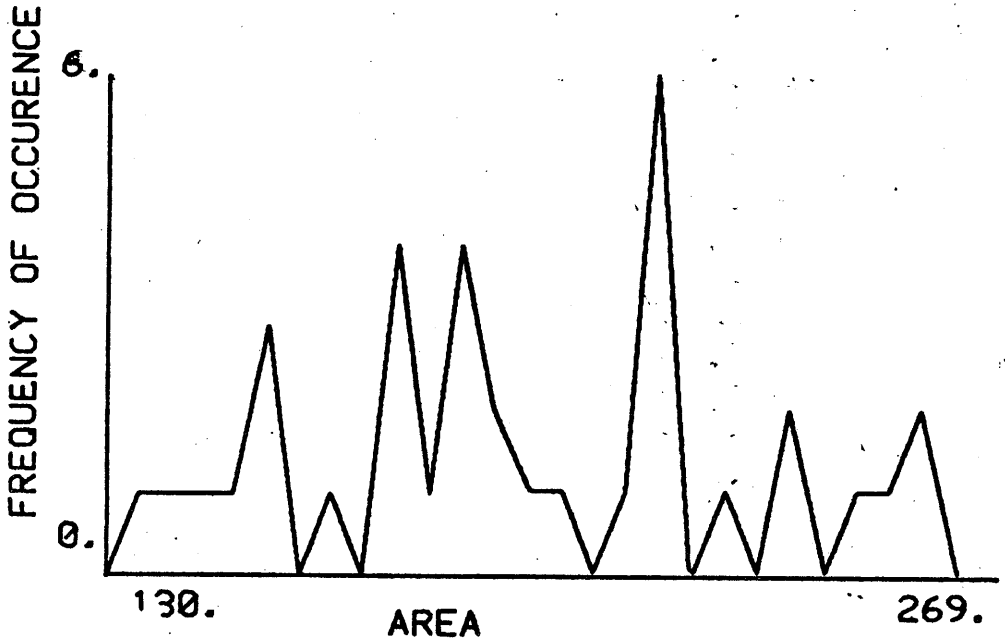

FREQUENCY OF OCCURENCE
7.
0.
3.
AXIS ANGLE
30.

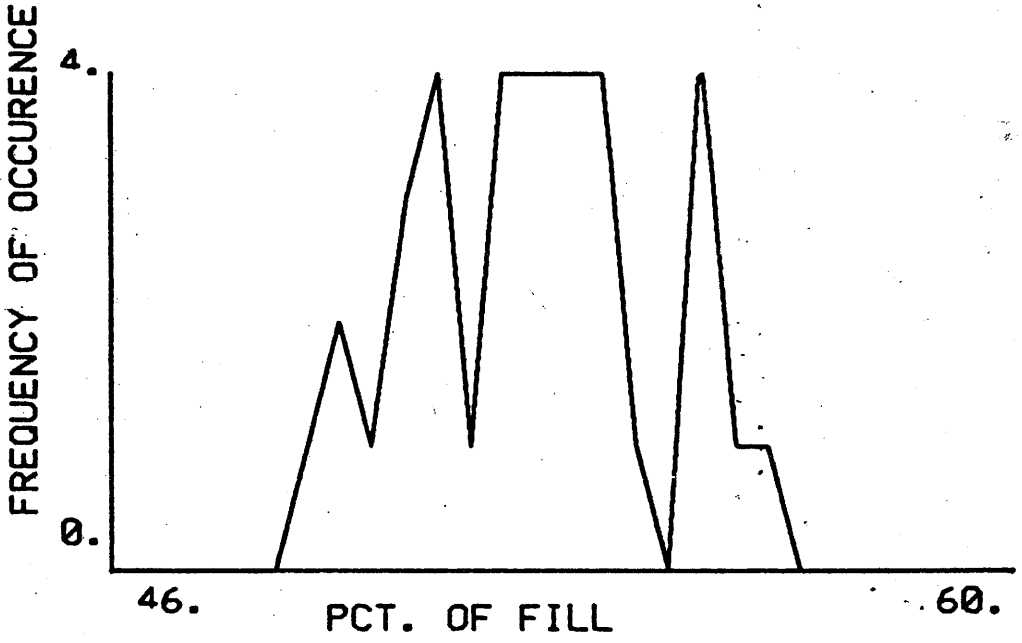

FREQUENCY OF OCCURENCE
15.
0.
12.
P**2 DIV. BY AREA

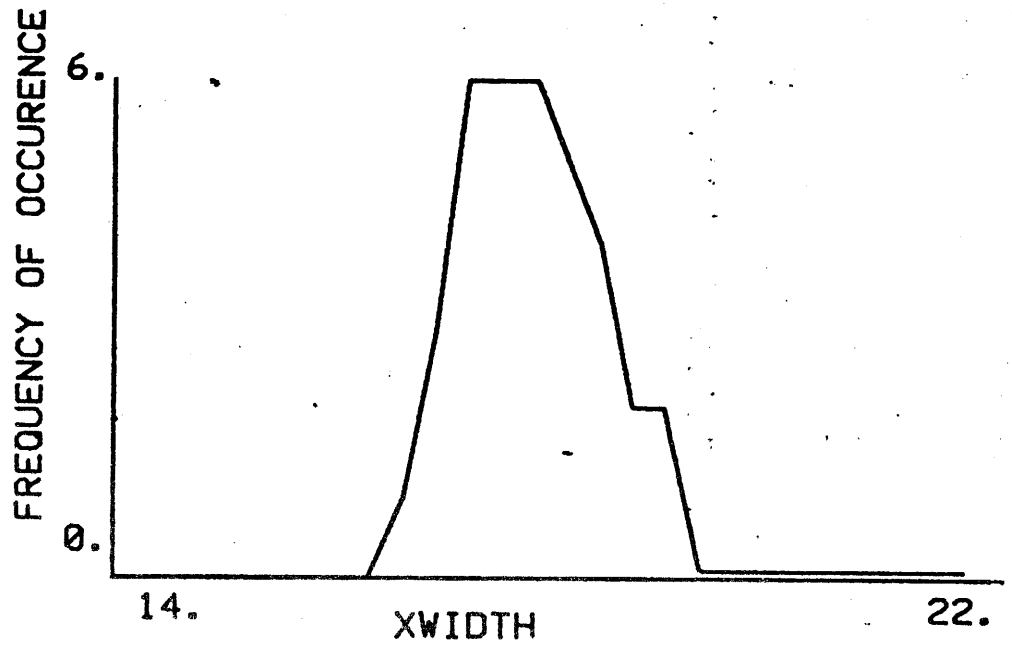


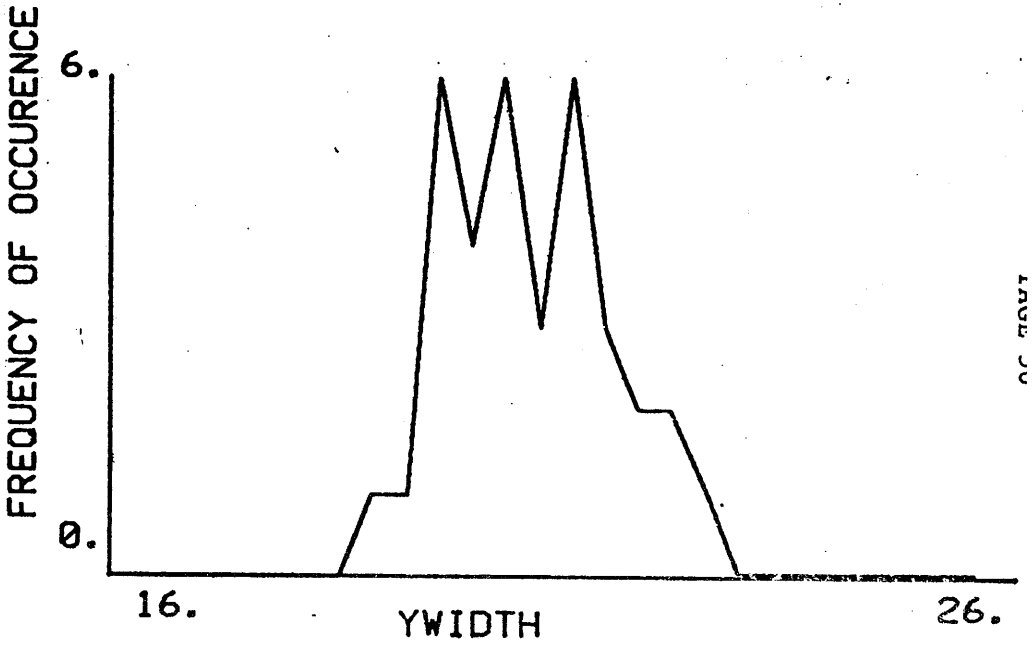


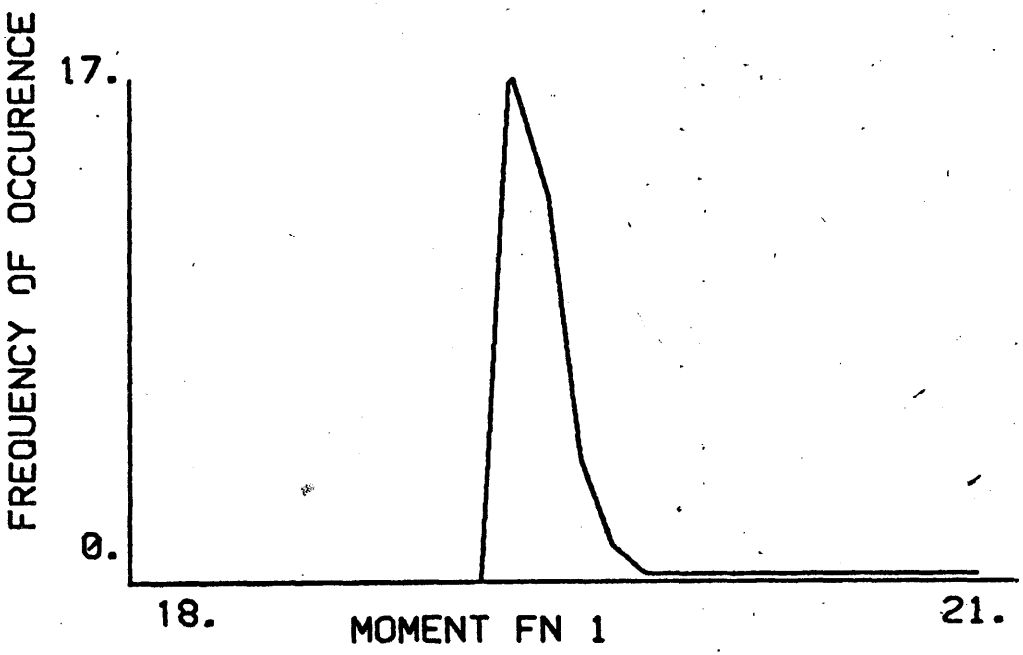


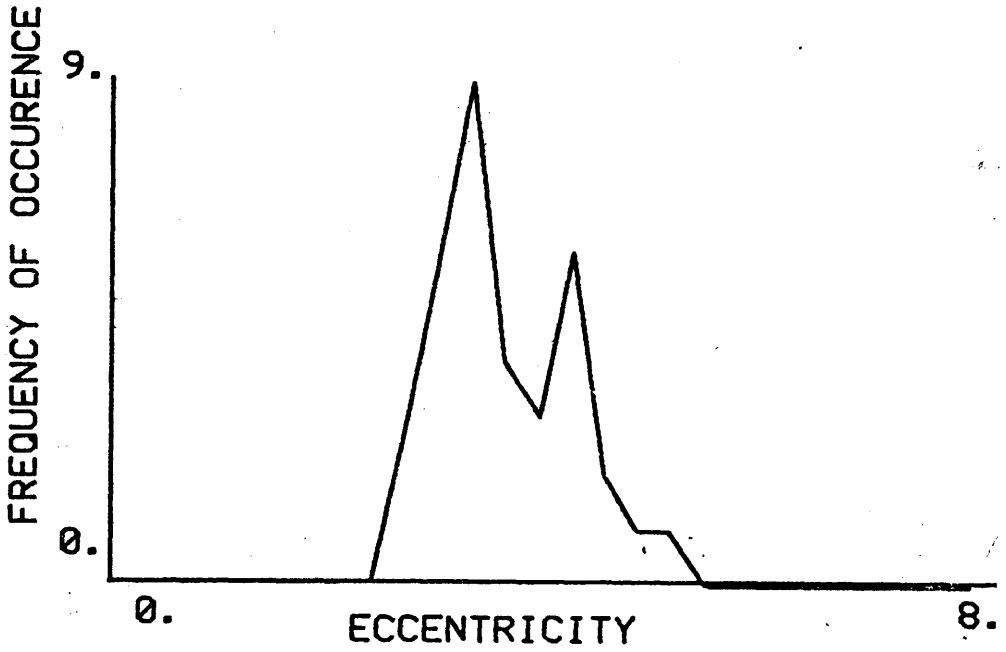

FREQUENCY OF OCCURENCE
6.
0.
130.
AREA
269.

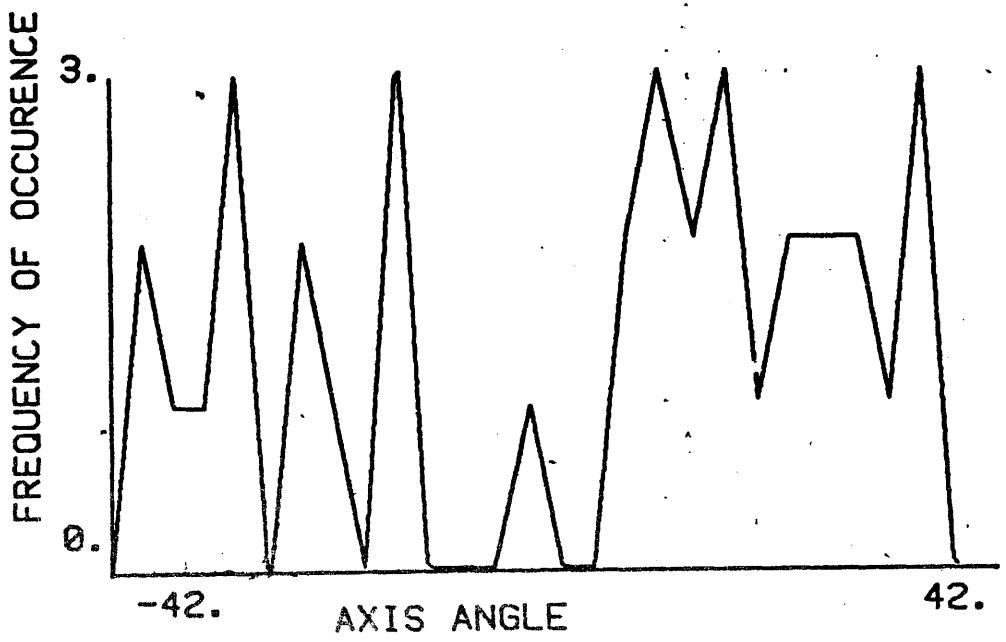

FREQUENCY OF OCCURENCE
4.
0.
46.
PCT. OF FILL
60.

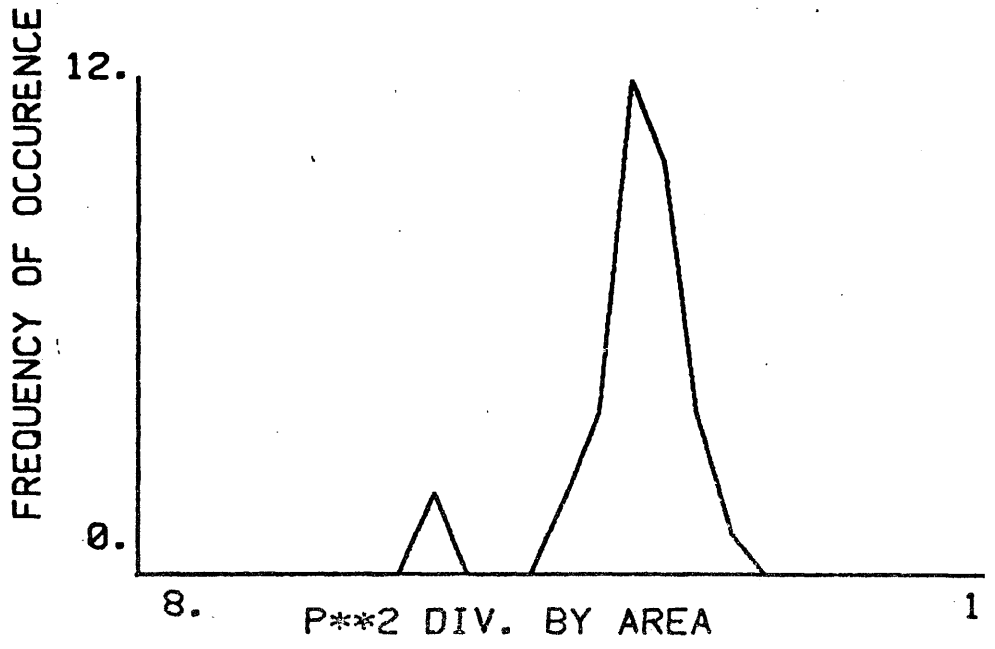

FREQUENCY OF OCCURENCE
6.
0.
14.
XWIDTH
22.

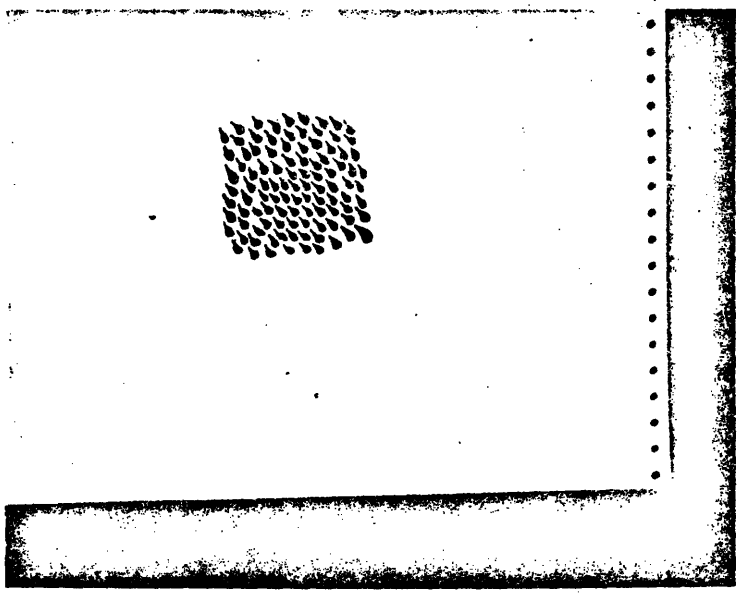

FREQUENCY OF OCCURENCE
6.
0.
16.
YWIDTH
26.

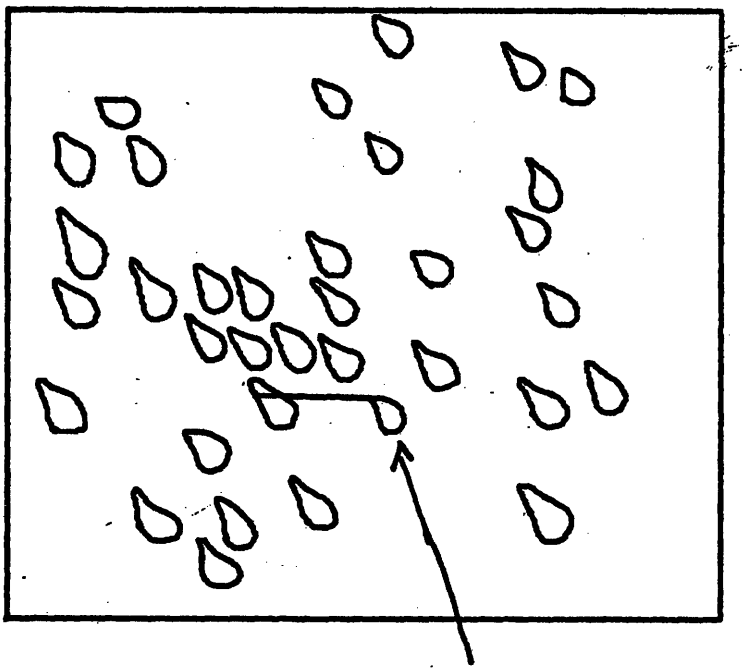

FREQUENCY OF OCCURENCE
17.
0.
18.
MOMENT FN 1
21.

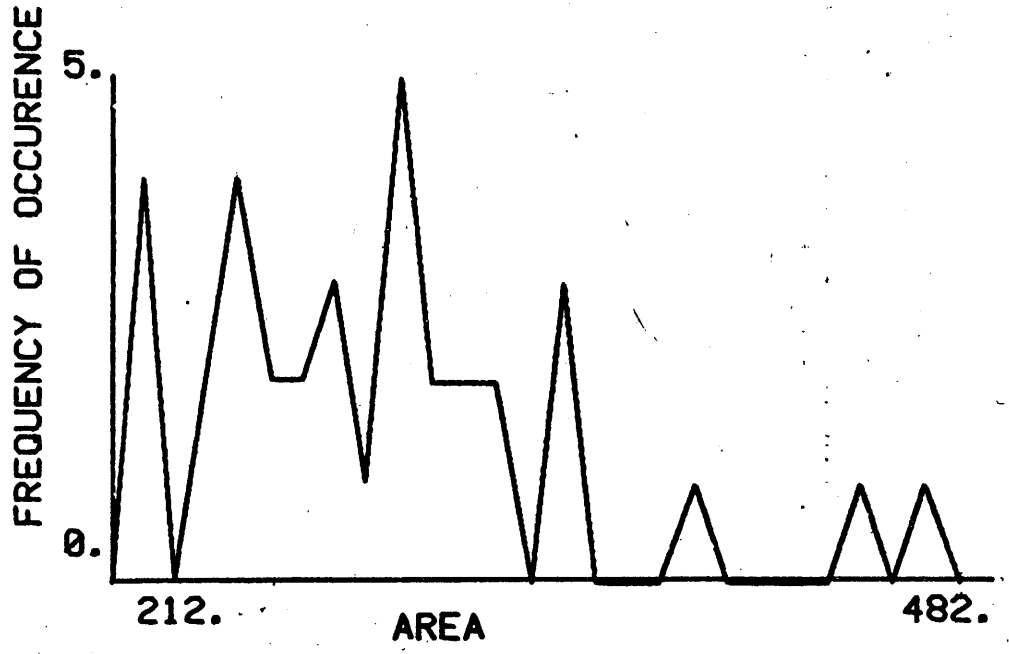

FREQUENCY OF OCCURENCE
9.
0.
0.
ECCENTRICITY
8.

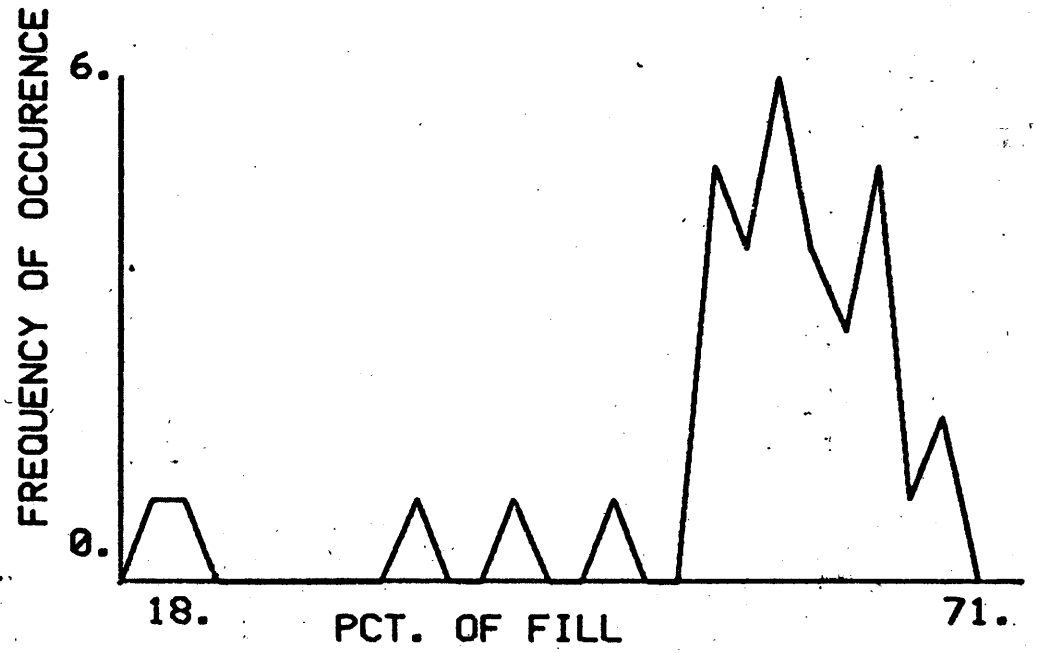

FREQUENCY OF OCCURENCE
3.
0.
-42.
AXIS ANGLE
42.

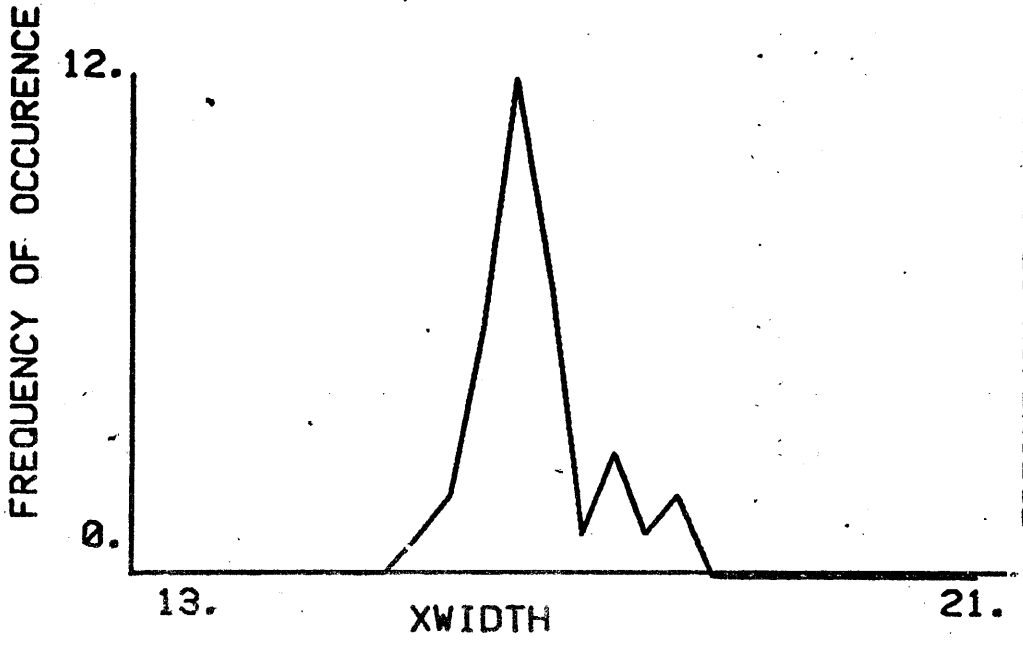

FREQUENCY OF OCCURENCE
12.
0.
8.
P**2 DIV. BY AREA
1

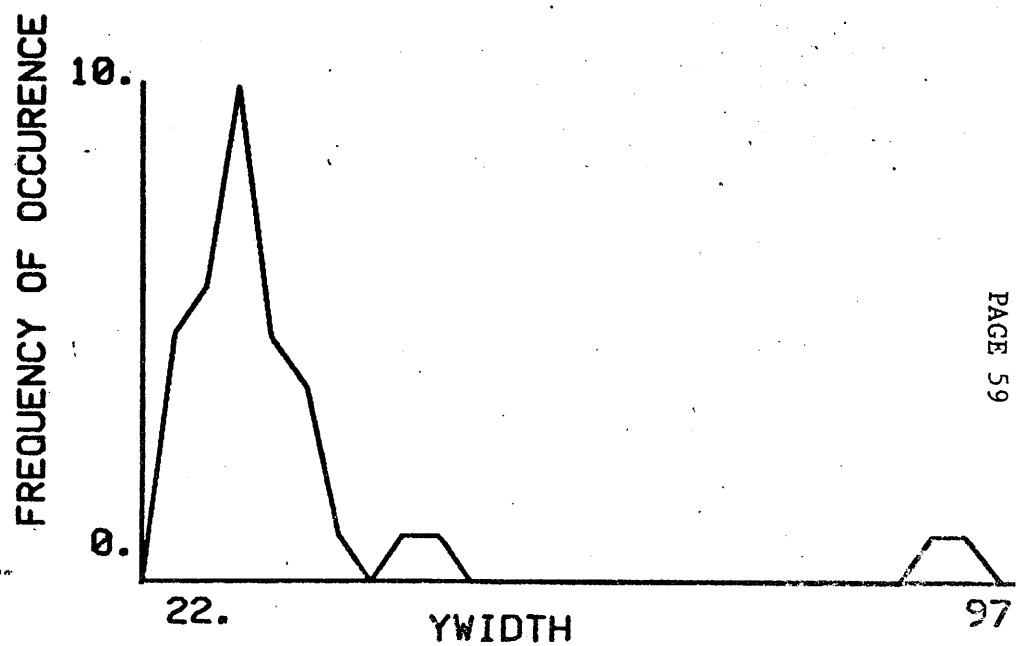


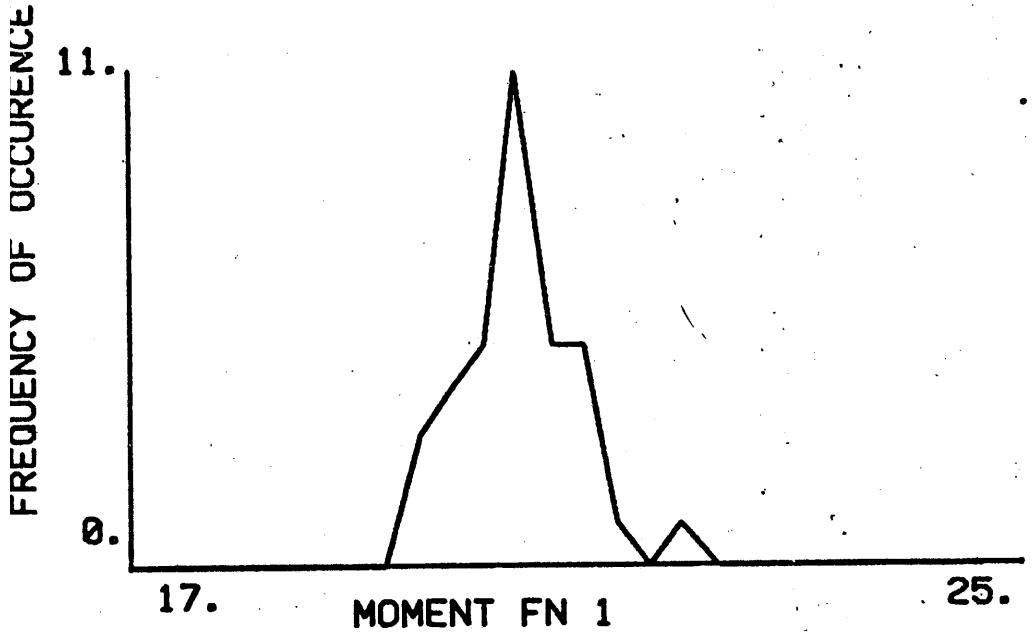


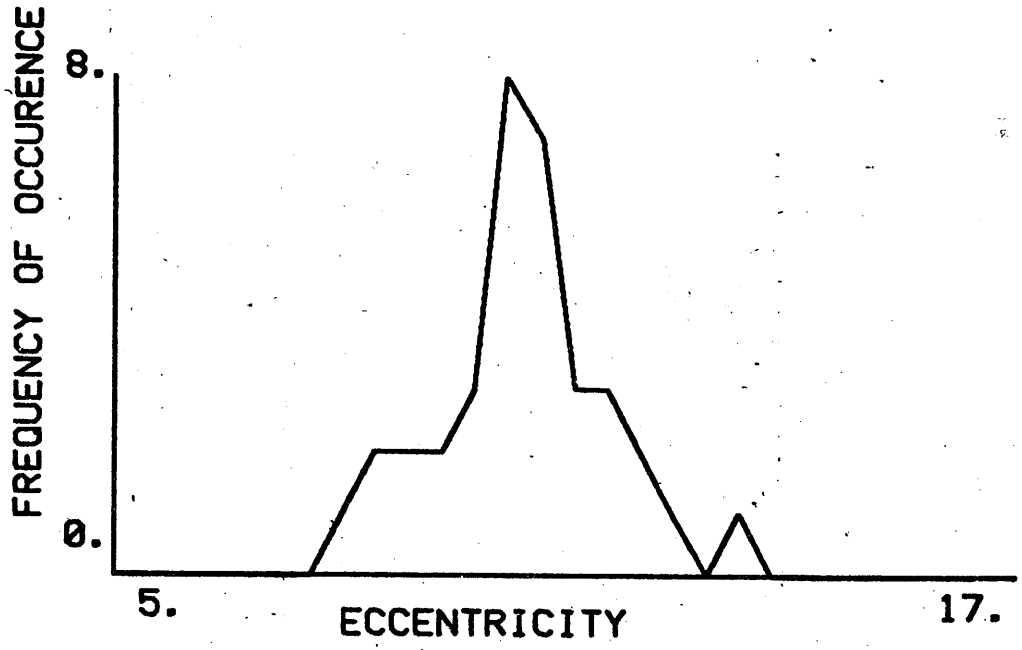

FREQUENCY OF OCCURENCE
5.
0.
212.
AREA
482.

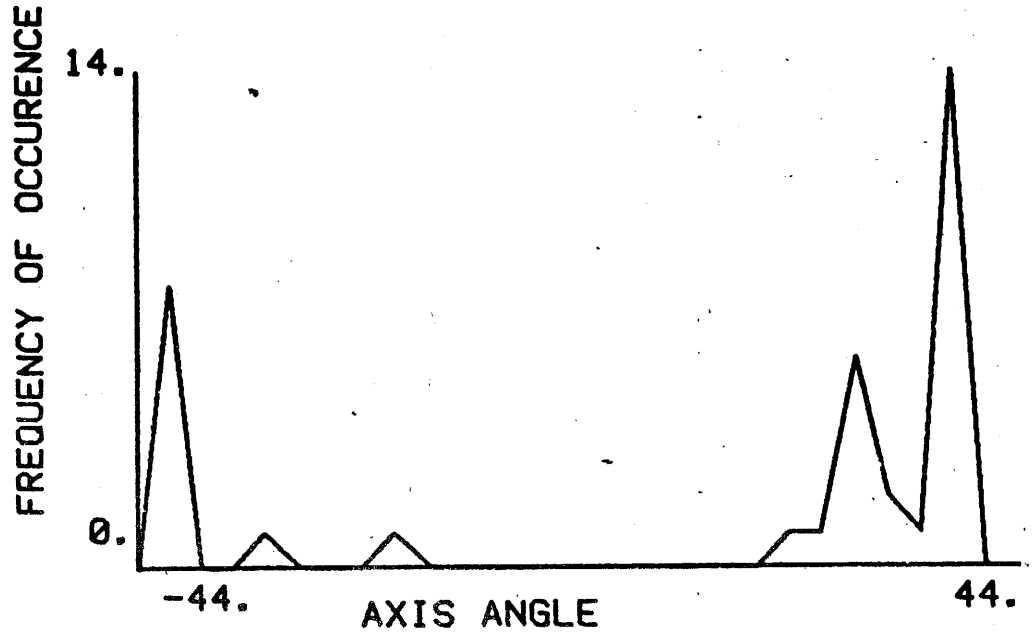

FREQUENCY OF OCCURENCE
6.
0.
18.
PCT. OF FILL
71.

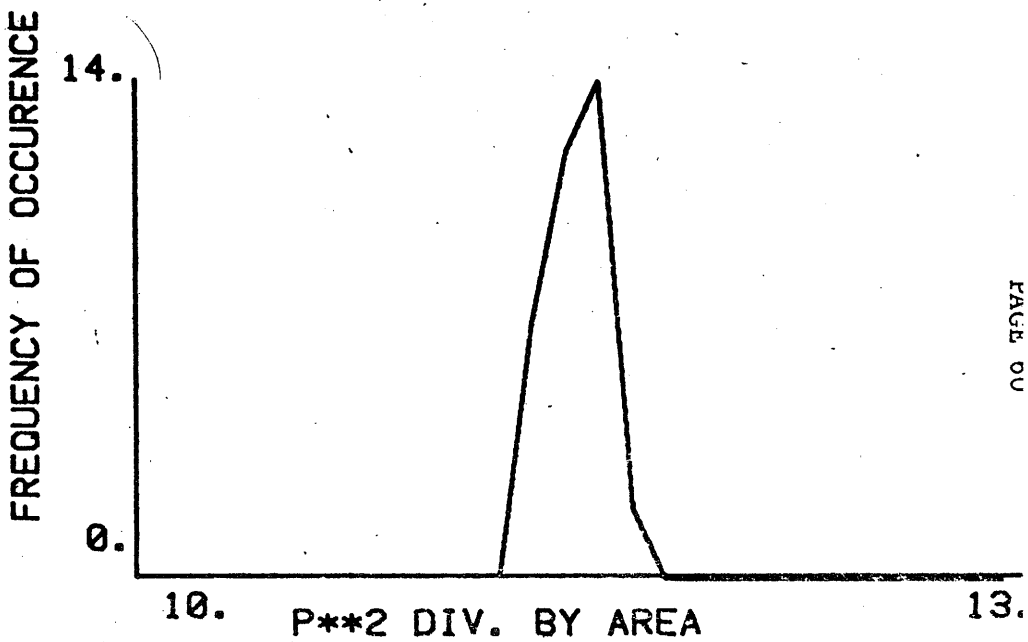

FREQUENCY OF OCCURENCE
12.
0.
13.
XWIDTH
21.

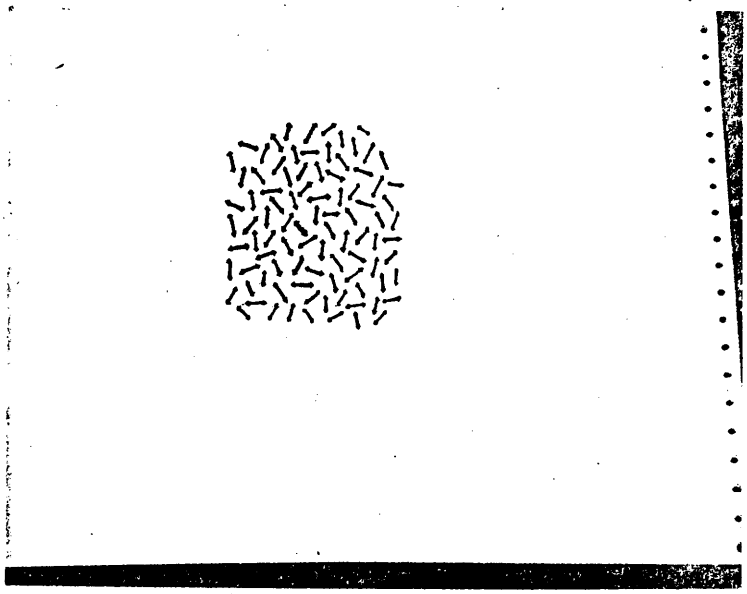

FREQUENCY OF OCCURENCE
10.
0.
22.
YWIDTH
97

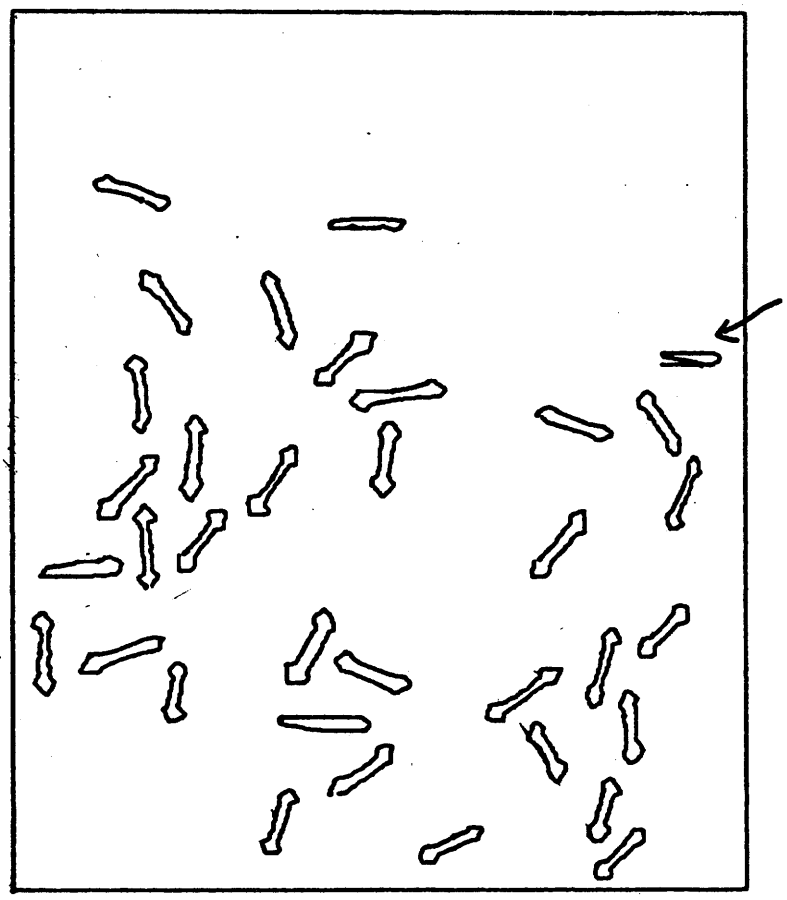

FREQUENCY OF OCCURENCE
11.
0.
17.
MOMENT FN 1
25.

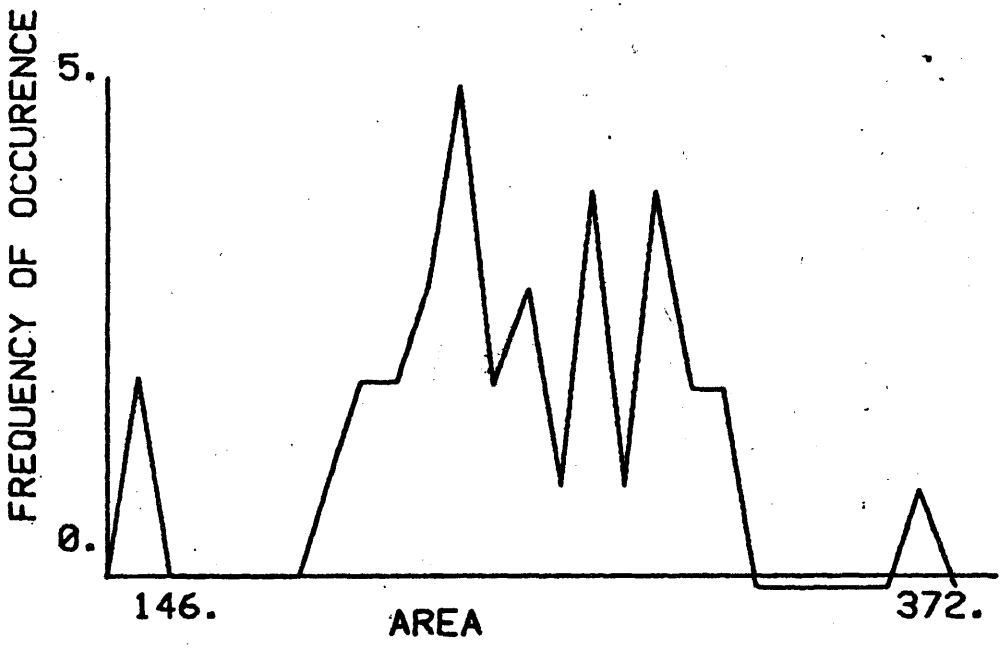

FREQUENCY OF OCCURENCE
8.
0.
5.
ECCENTRICITY
17.

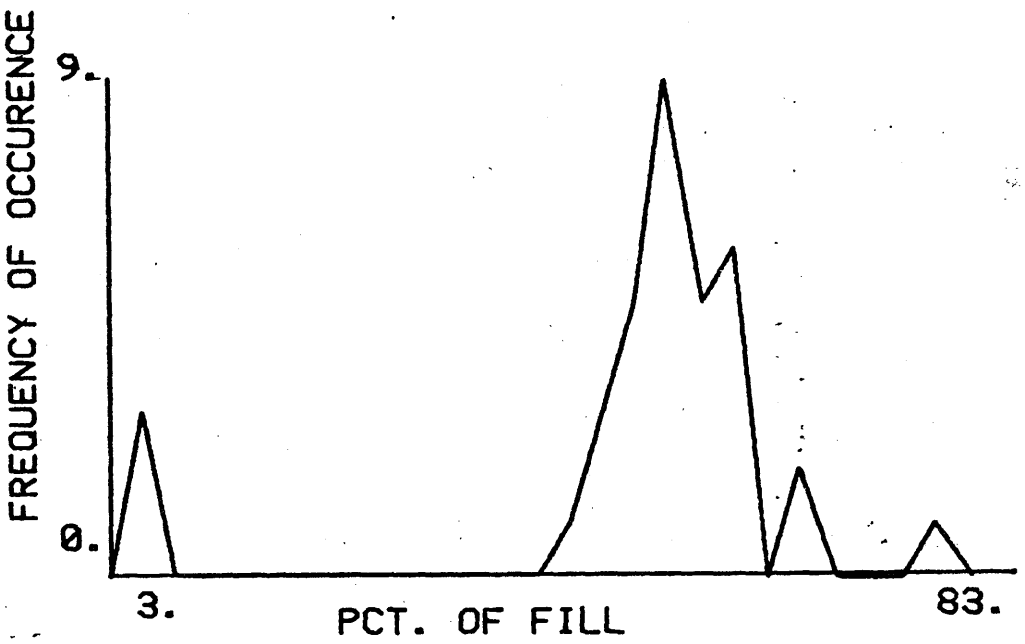

FREQUENCY OF OCCURENCE
14.
0.
-44.
AXIS ANGLE
44.

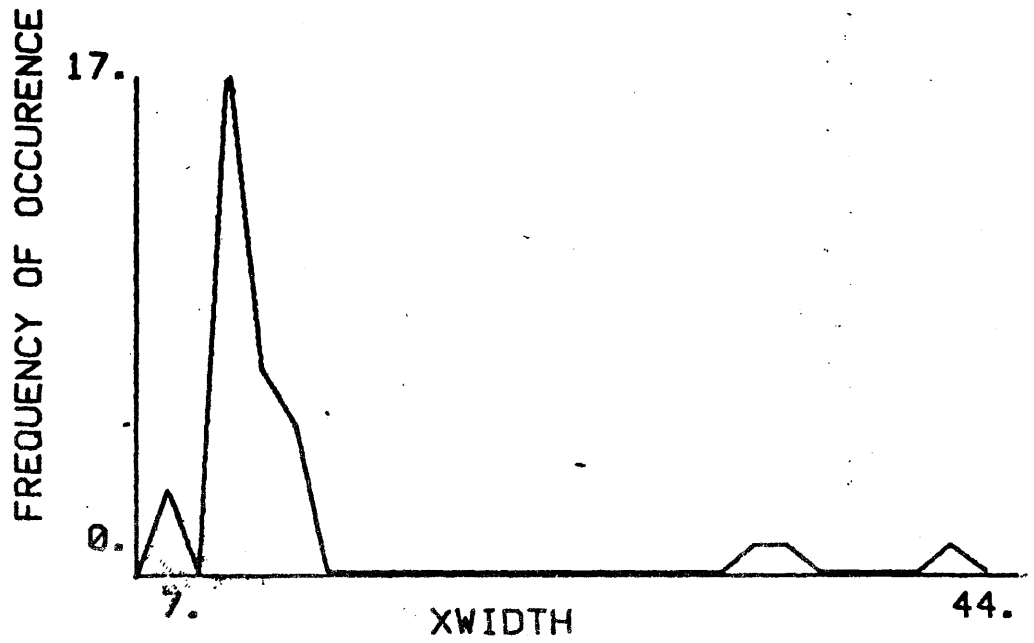

FREQUENCY OF OCCURENCE
14.
0.
10.
P**2 DIV. BY AREA
13.

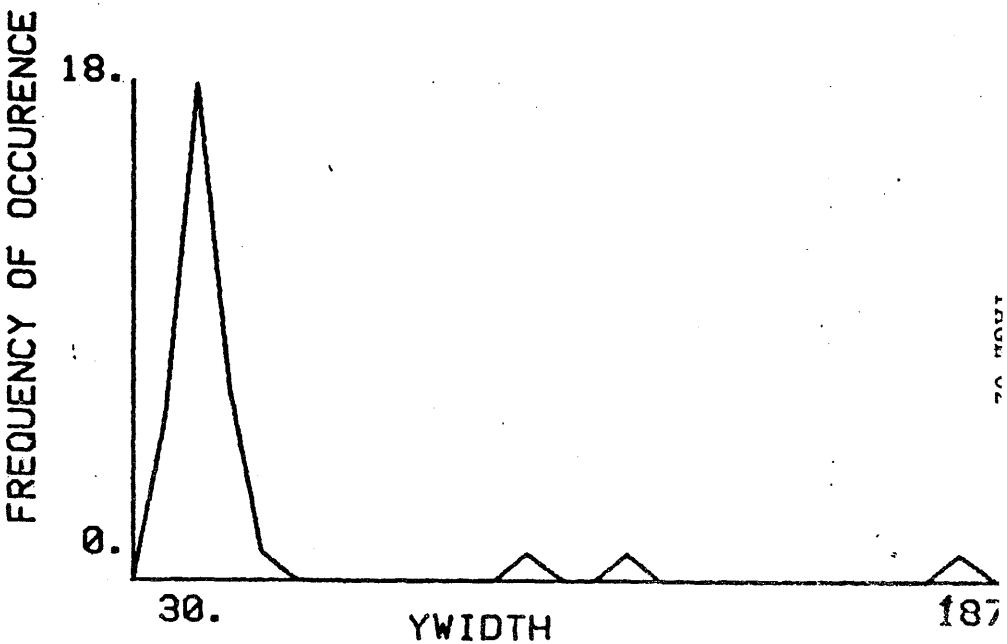


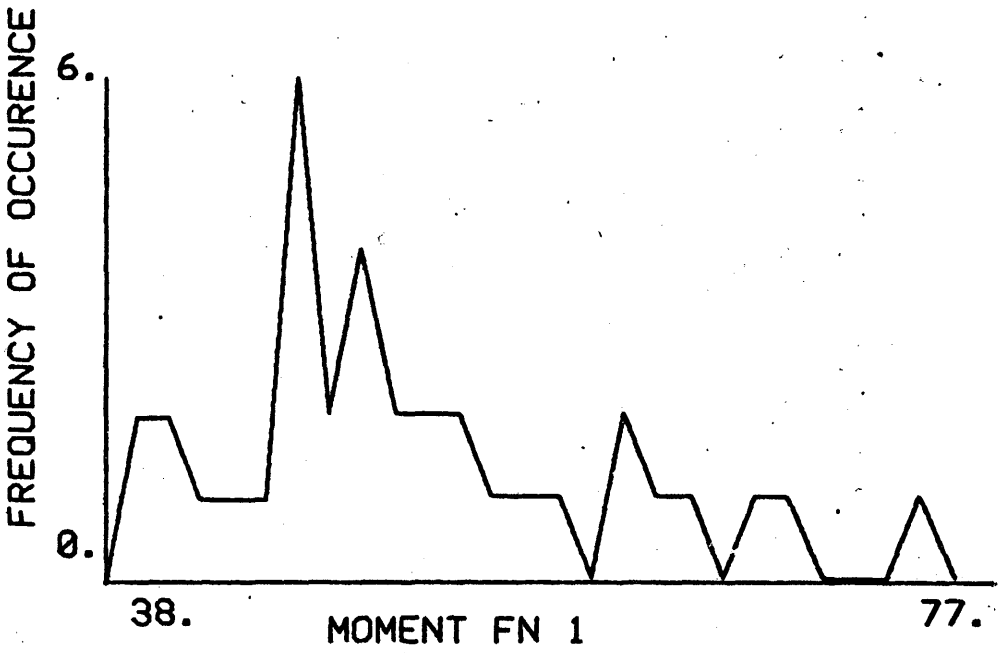


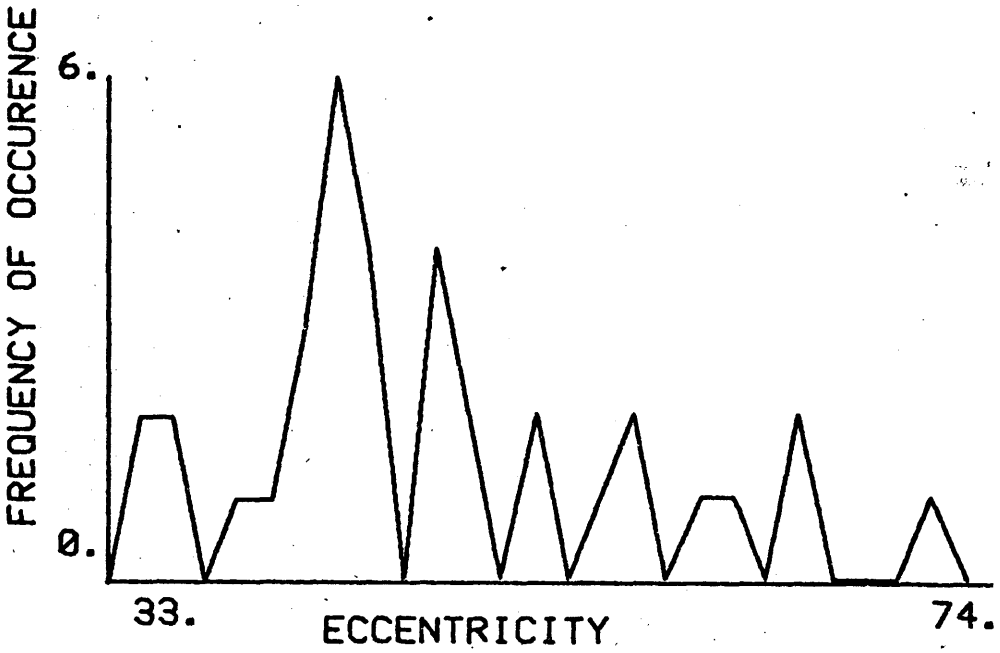

FREQUENCY OF OCCURENCE
5.
0.
146.
AREA
372.

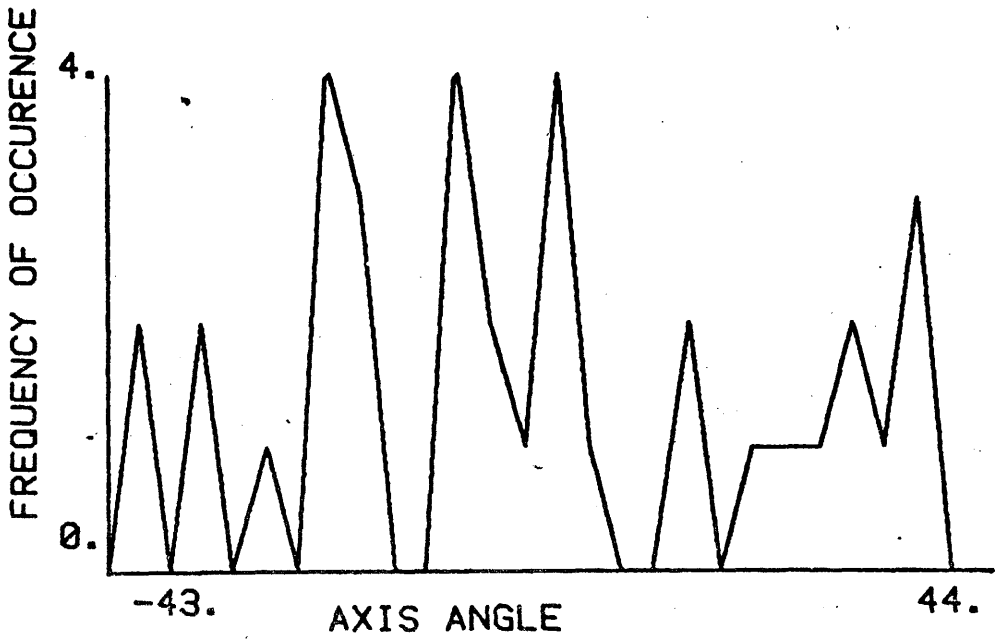

FREQUENCY OF OCCURENCE
9.
0.
3.
PCT. OF FILL
83.

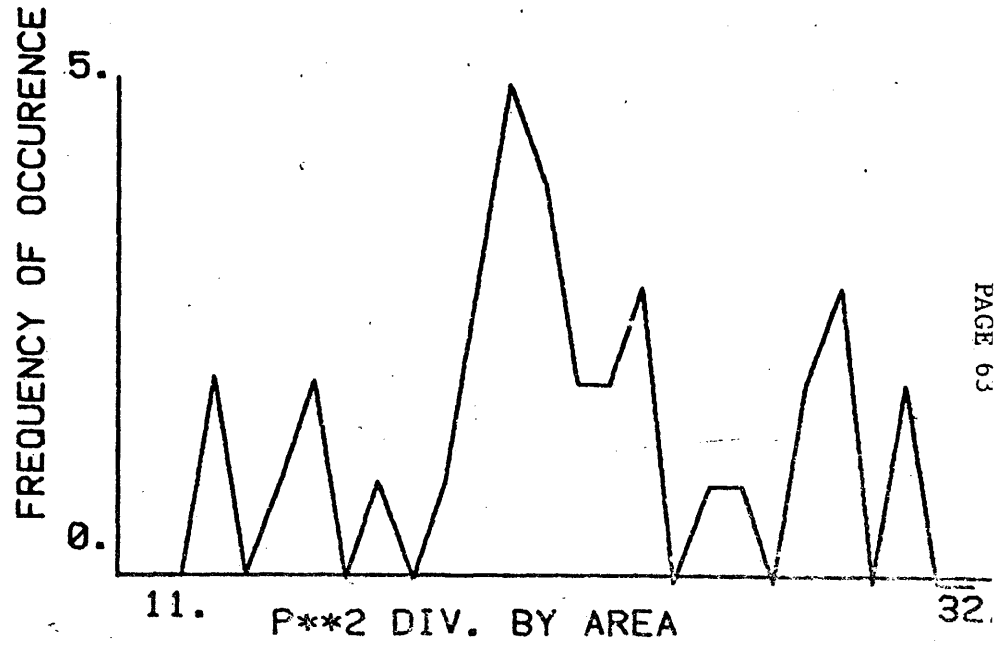

FREQUENCY OF OCCURENCE
17.
0.
7.
XWIDTH
44.

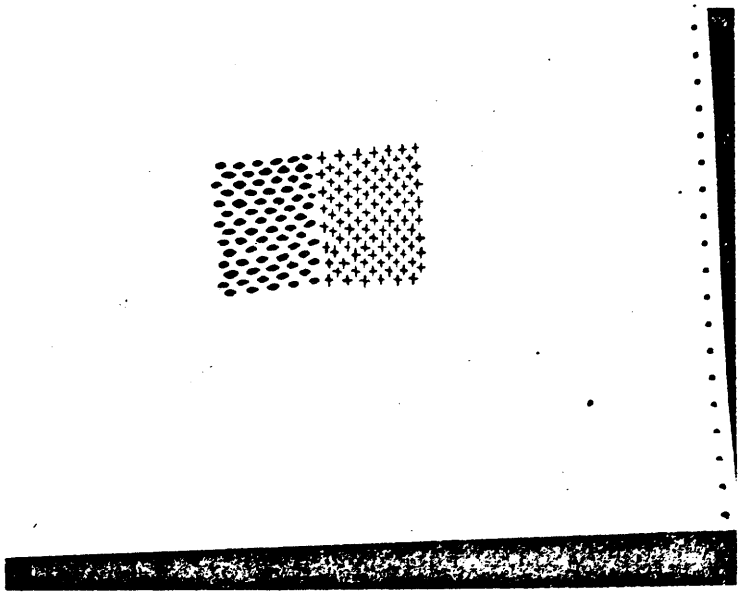

FREQUENCY OF OCCURENCE
18.
0.
30.
YWIDTH
187

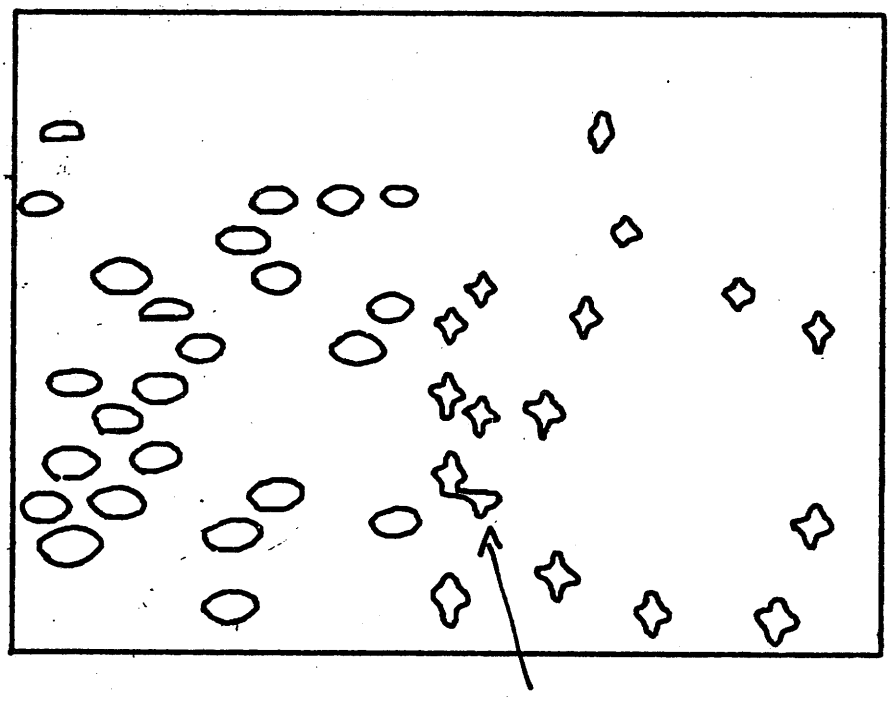

FREQUENCY OF OCCURENCE
6.
0.
38.
MOMENT FN 1
77.

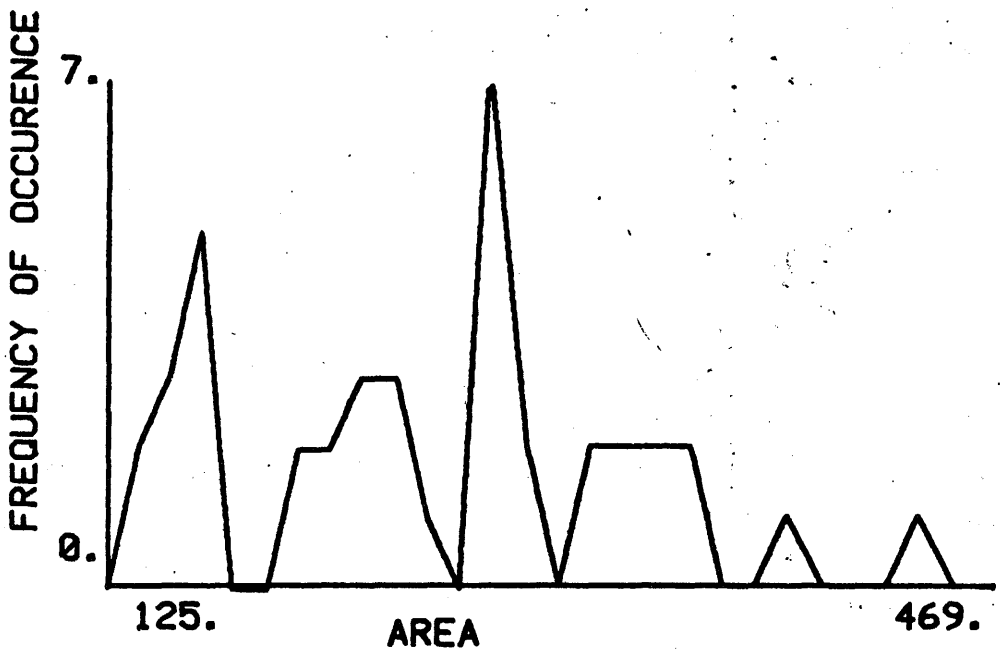

FREQUENCY OF OCCURENCE
6.
0.
33.
ECCENTRICITY
74.

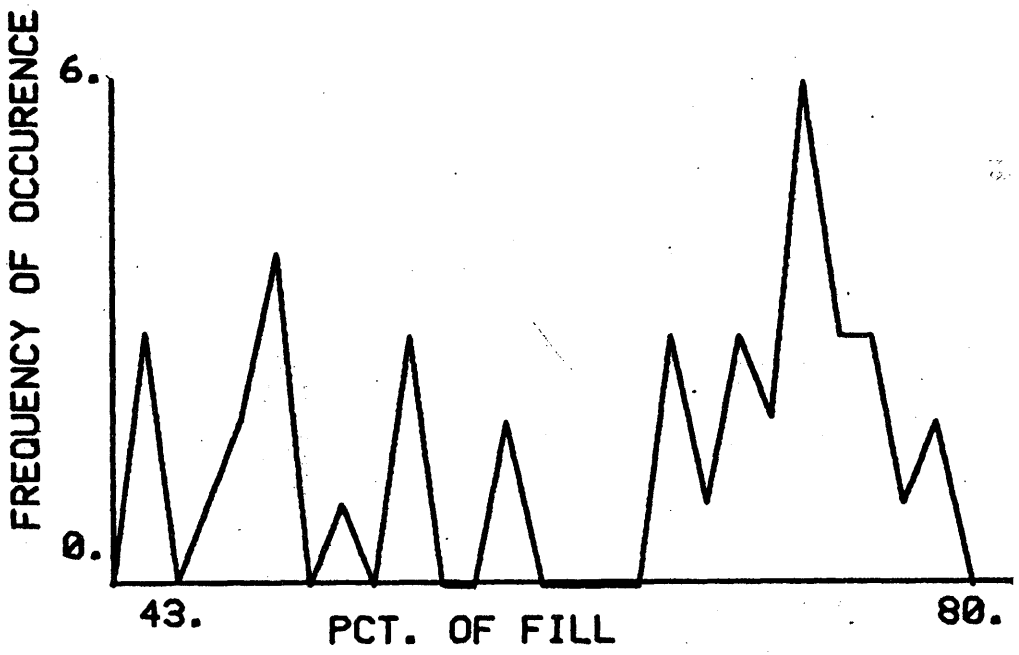

FREQUENCY OF OCCURENCE
4.
0.
-43.
AXIS ANGLE
44.

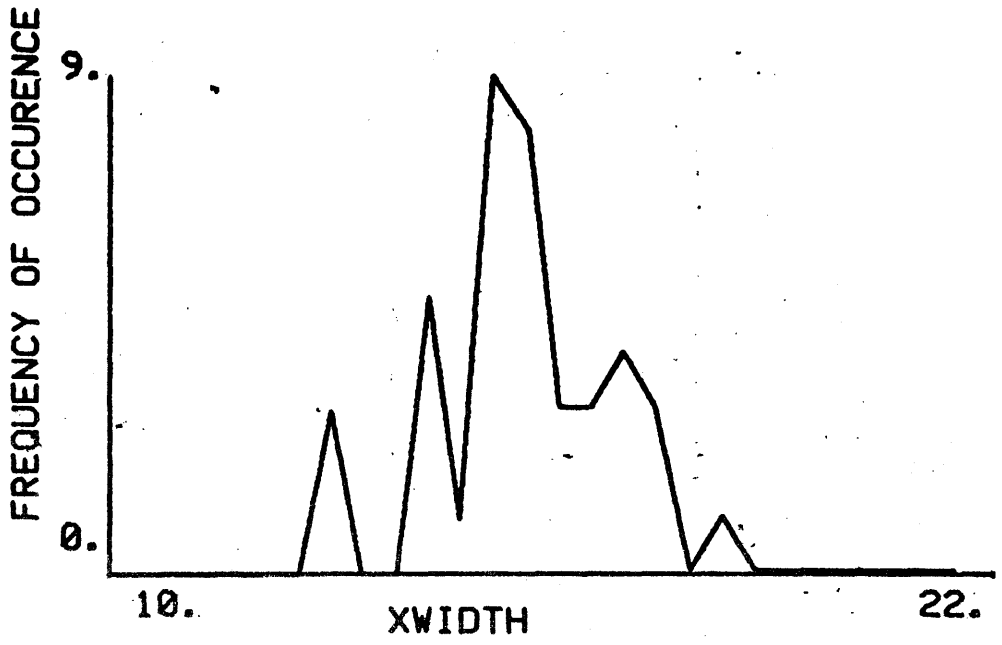

FREQUENCY OF OCCURENCE
5.
0.
11.
P**2 DIV. BY AREA
32.

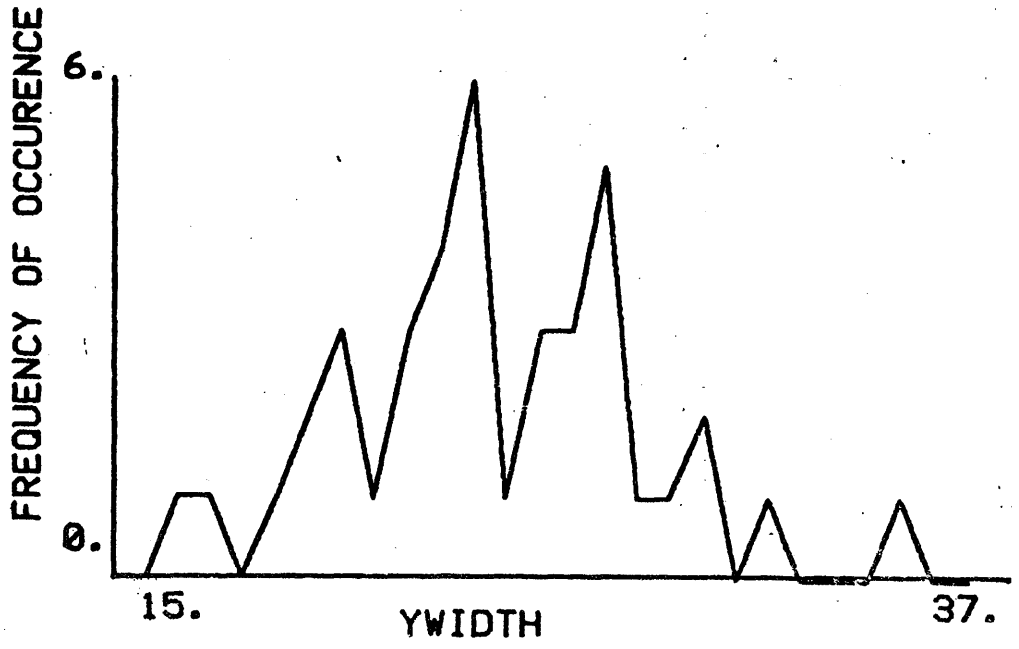


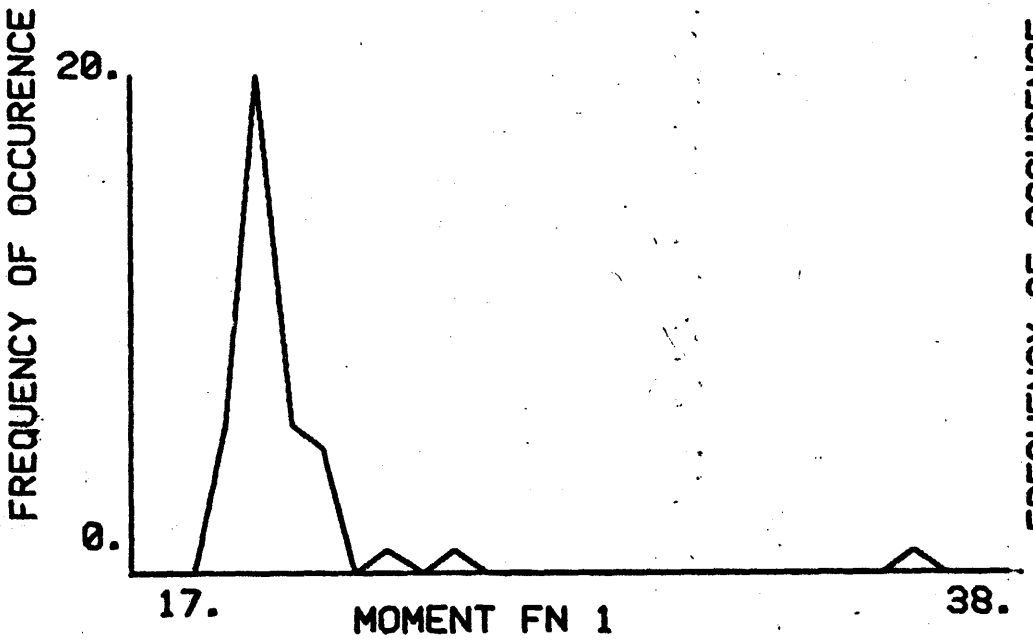


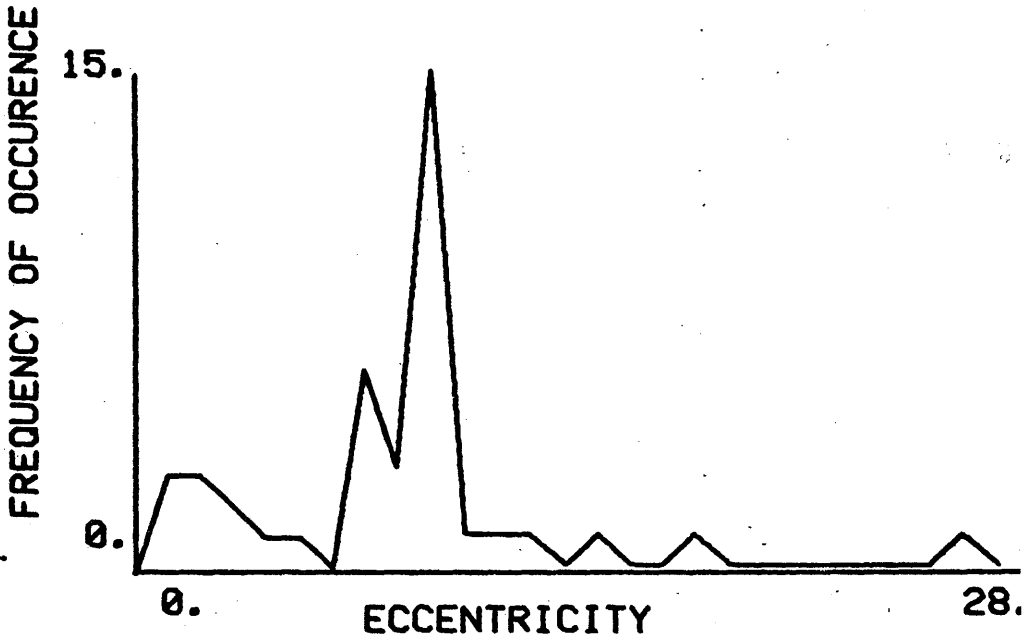

FREQUENCY OF OCCURENCE
7.
0.
125.
AREA
469.

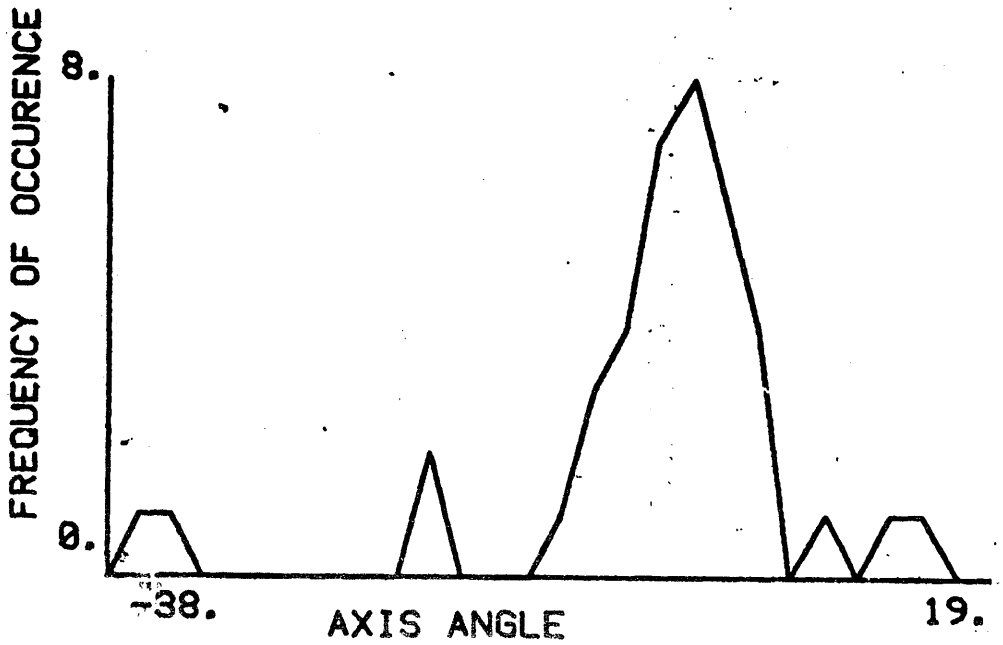

FREQUENCY OF OCCURENCE
6.
0.
43.
PCT. OF FILL
80.

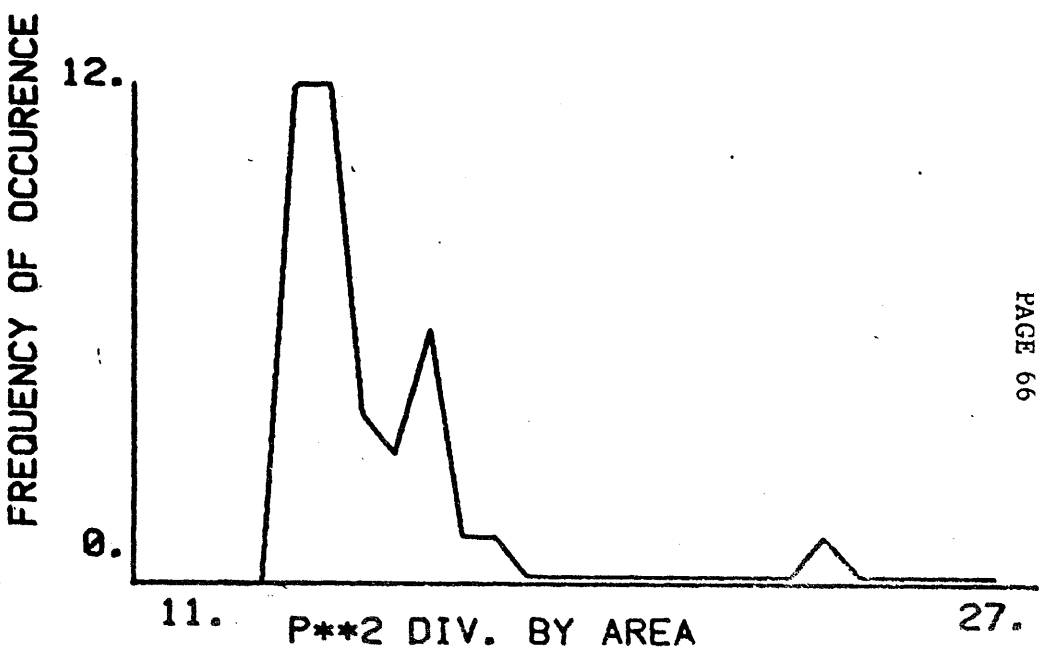

FREQUENCY OF OCCURENCE
9.
0.
10.
XWIDTH
22.


FREQUENCY OF OCCURENCE
6.
0.
15.
YWIDTH
37.

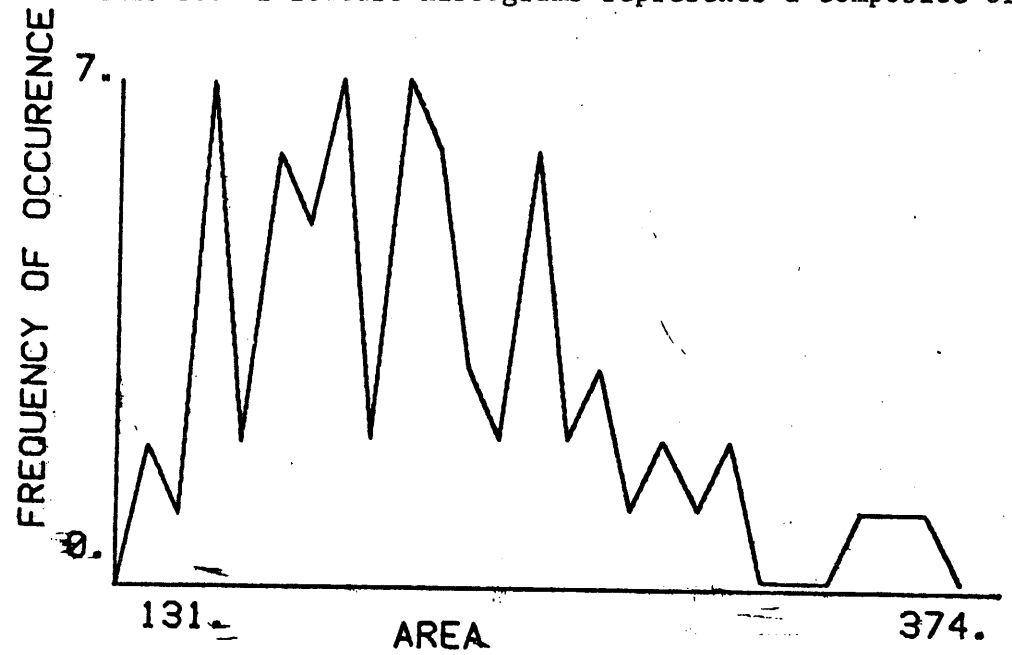

FREQUENCY OF OCCURENCE
20.
0.
17.
MOMENT FN 1
38.

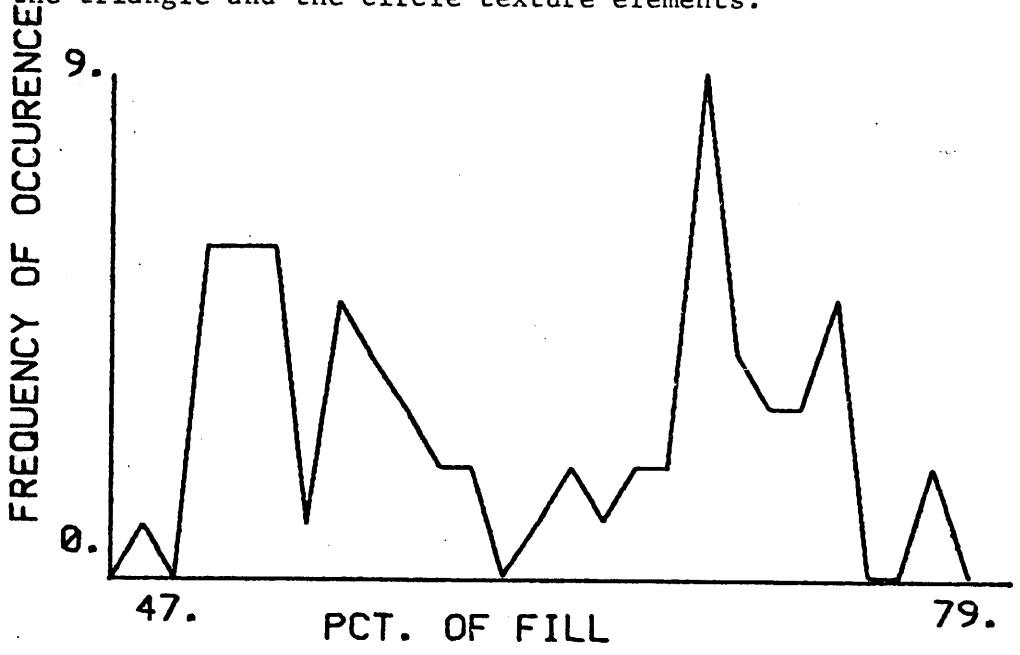

FREQUENCY OF OCCURENCE
15.
0.
0.
ECCENTRICITY
28.

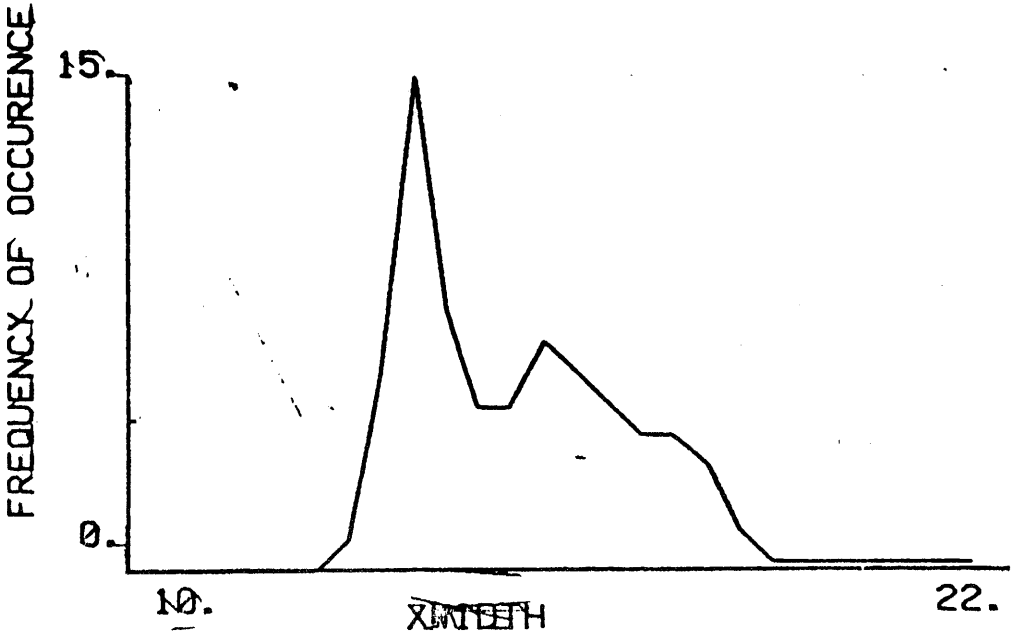

FREQUENCY OF OCCURENCE
8.
0.
-38.
AXIS ANGLE
19.

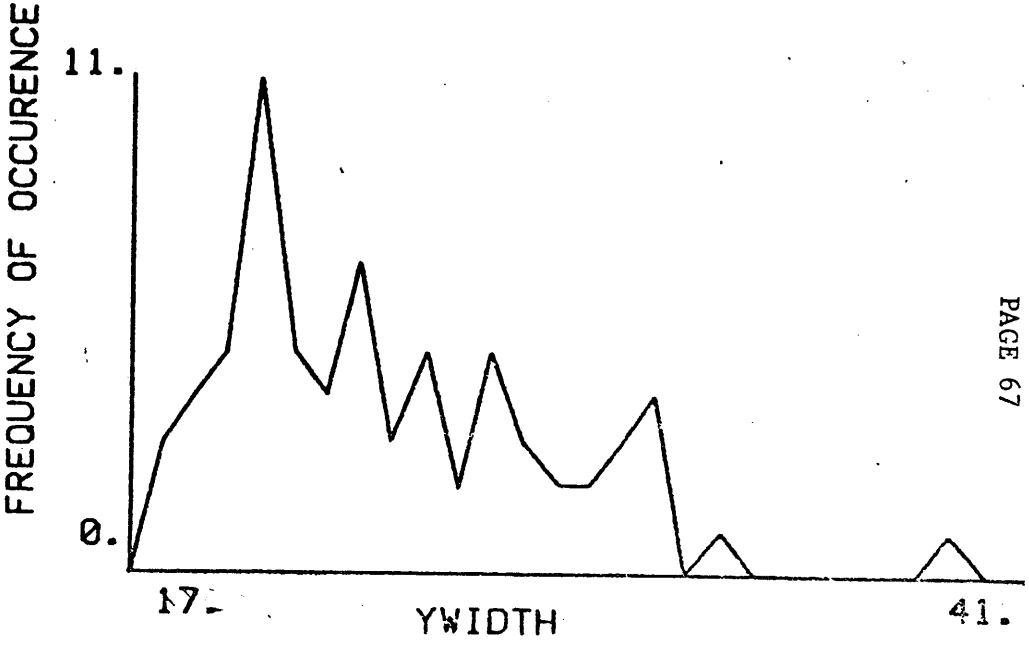

FREQUENCY OF OCCURENCE
12.
0.
11.
P**2 DIV. BY AREA
27.

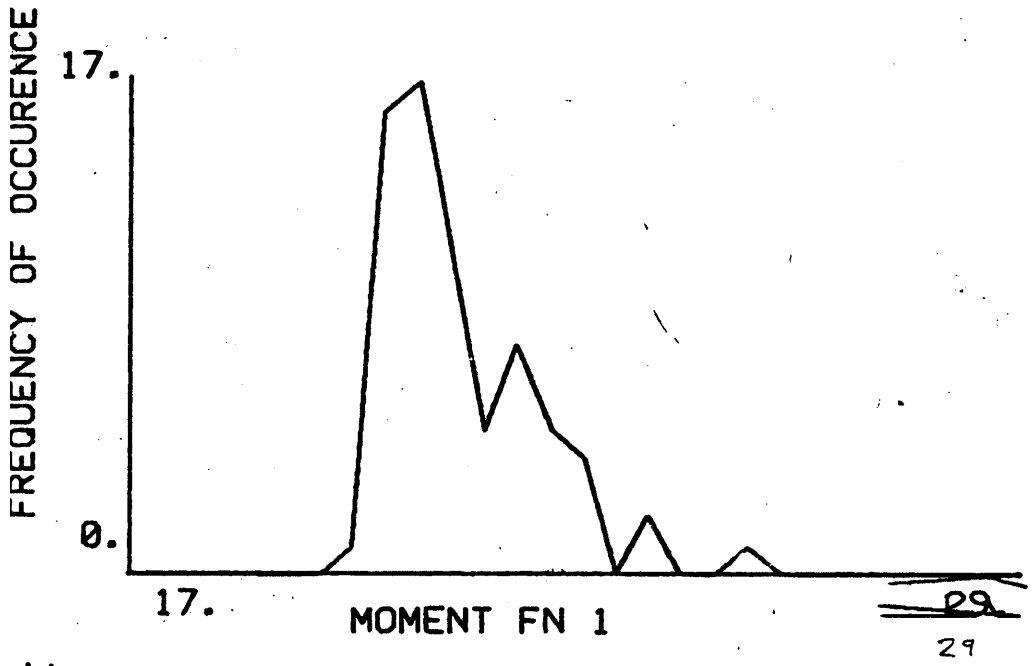
This set of feature histograms represents a composite of the triangle and the circle texture elements.

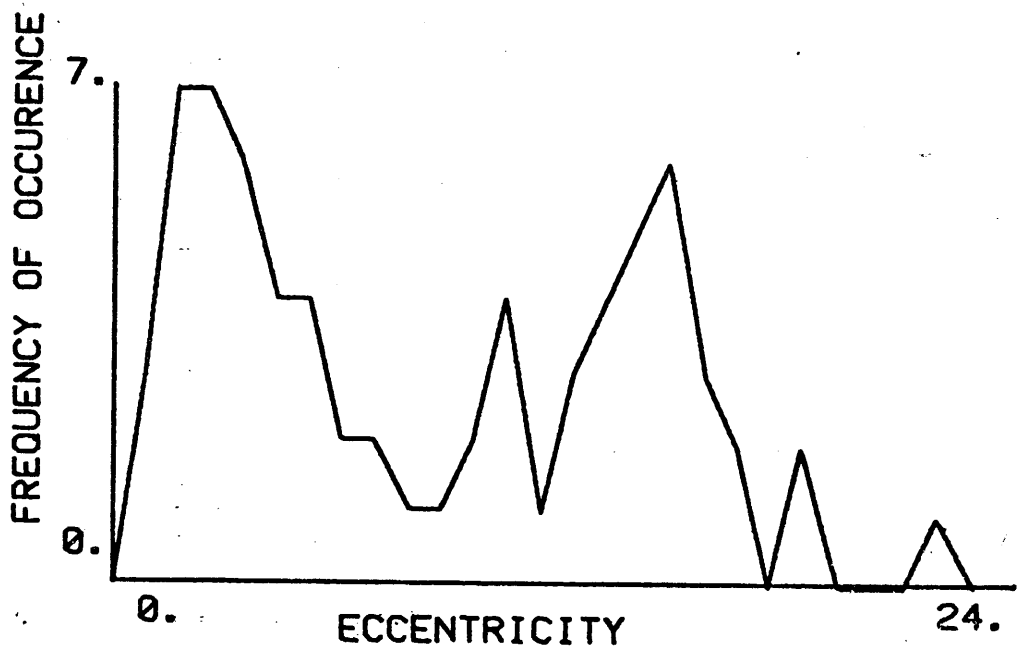


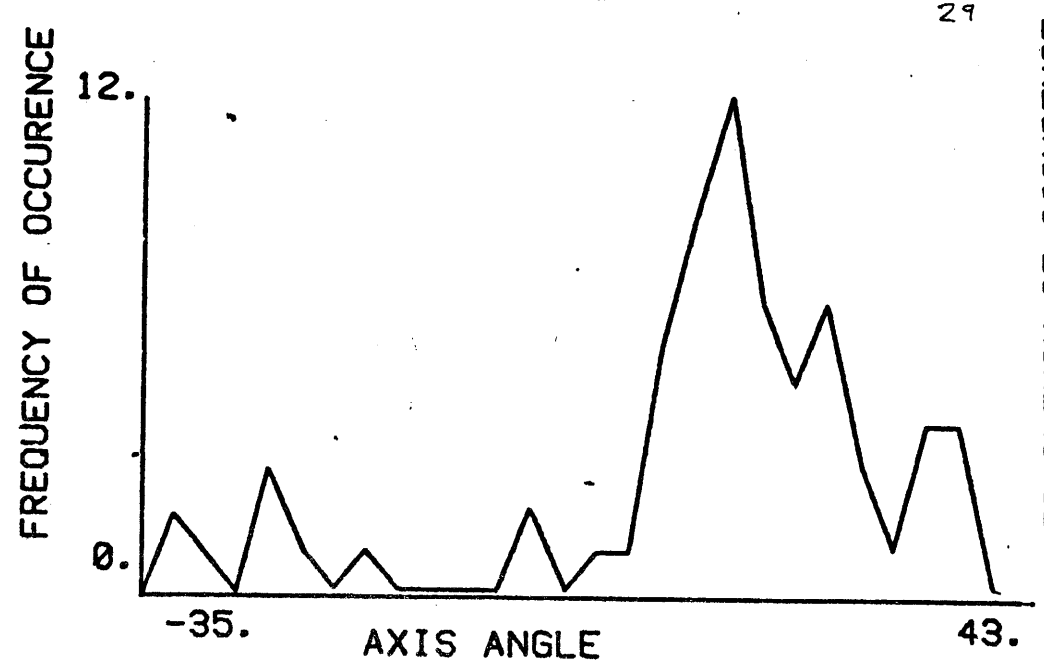


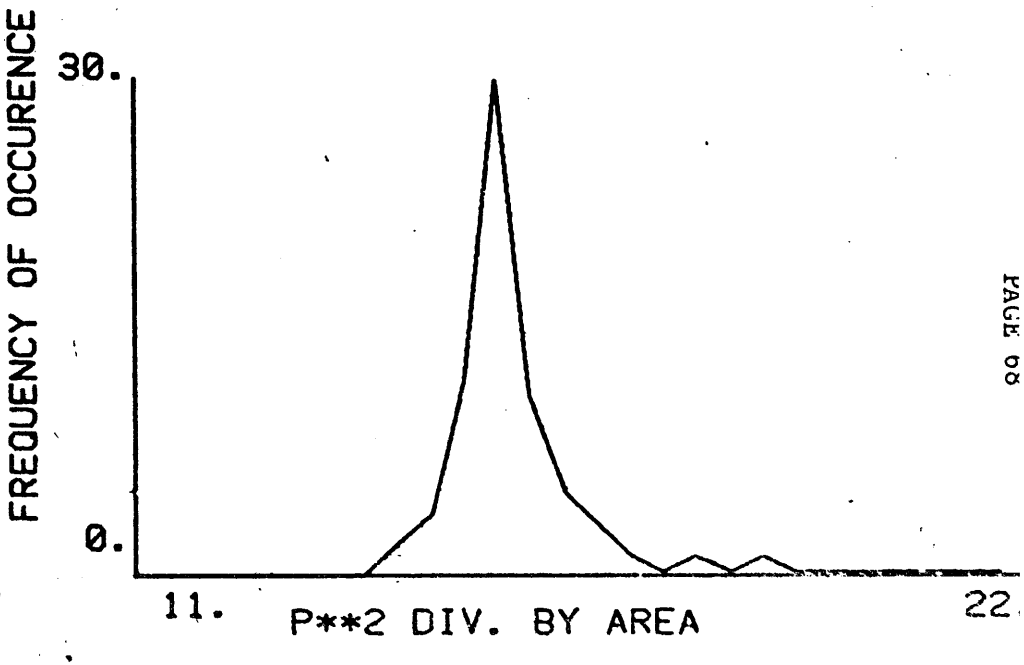


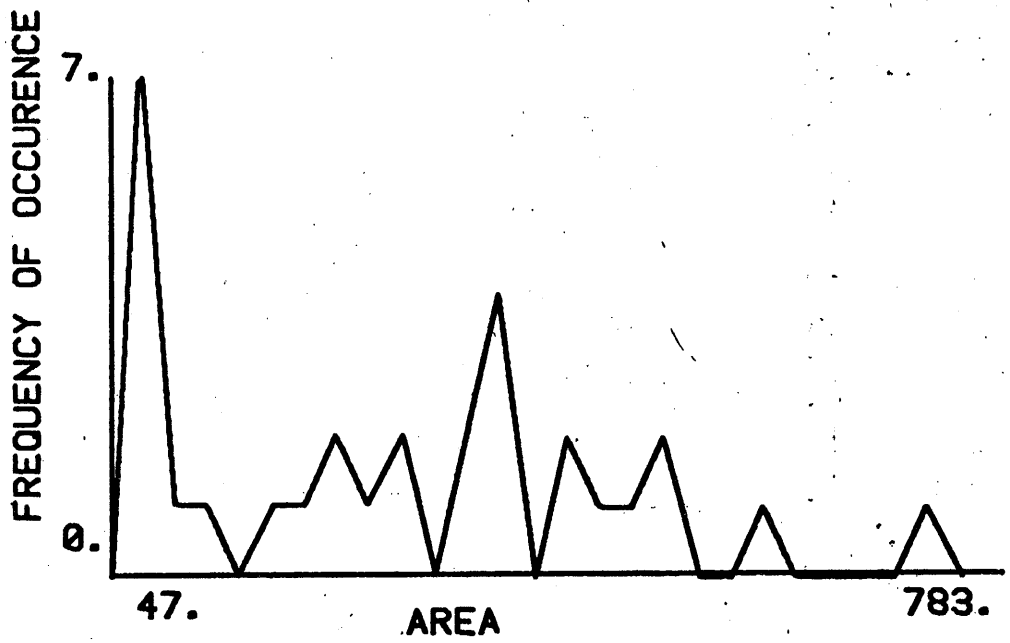


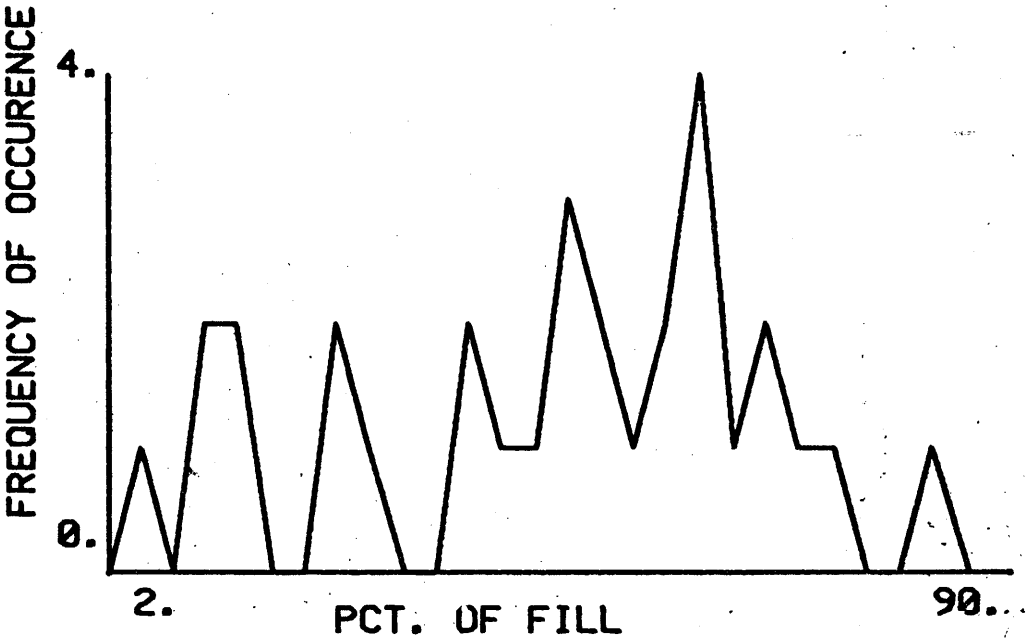

FREQUENCY OF OCCURENCE
17.
0.
17.
MOMENT FN 1
29
29

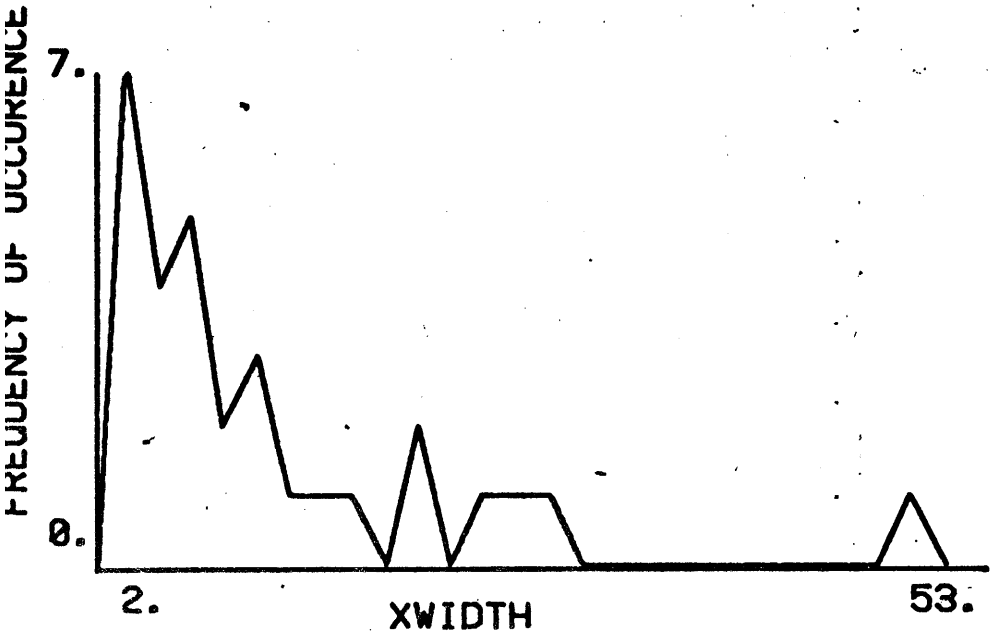

FREQUENCY OF OCCURENCE
7.
0.
0.
ECCENTRICITY
24.


FREQUENCY OF OCCURENCE
12.
0.
-35.
AXIS ANGLE
43.

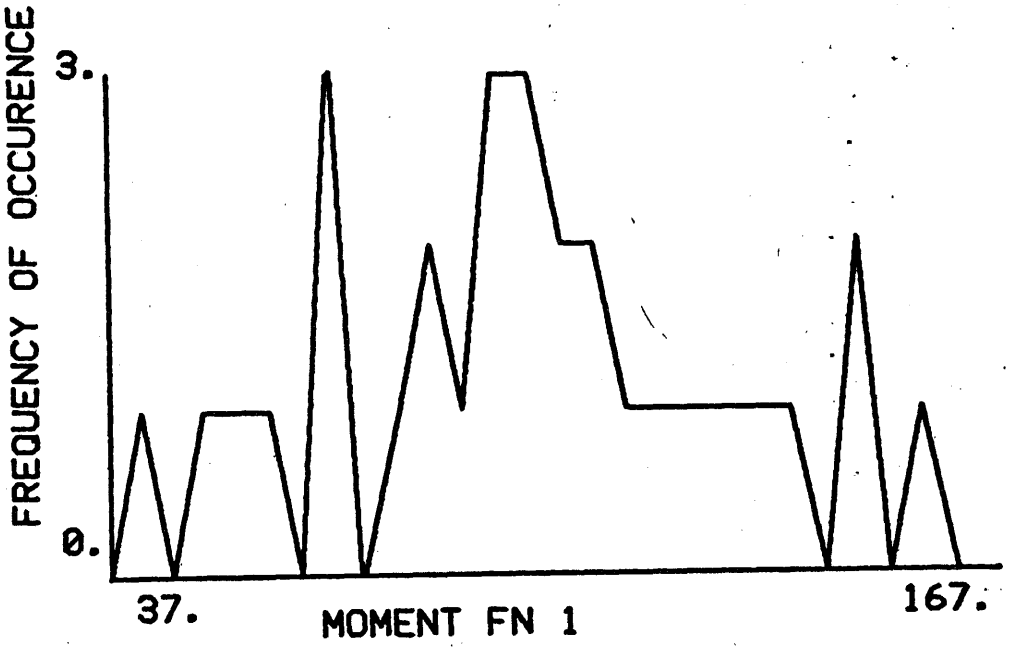

FREQUENCY OF OCCURENCE
30.
0.
11.
P**2 DIV. BY AREA
22.

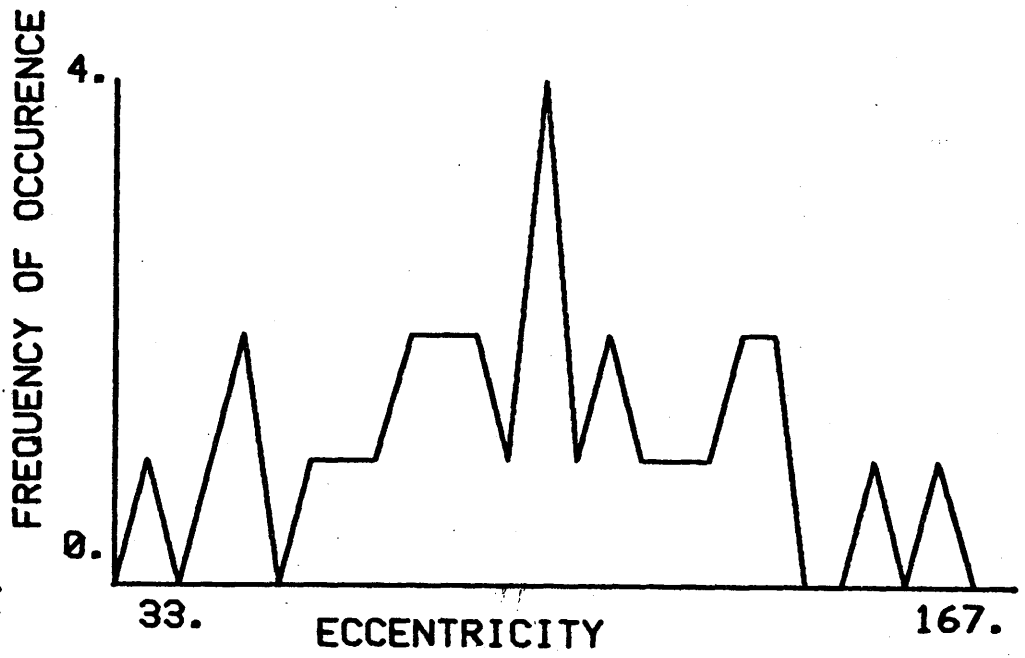

FREQUENCY OF OCCURENCE
7.
0.
47.
AREA
783.

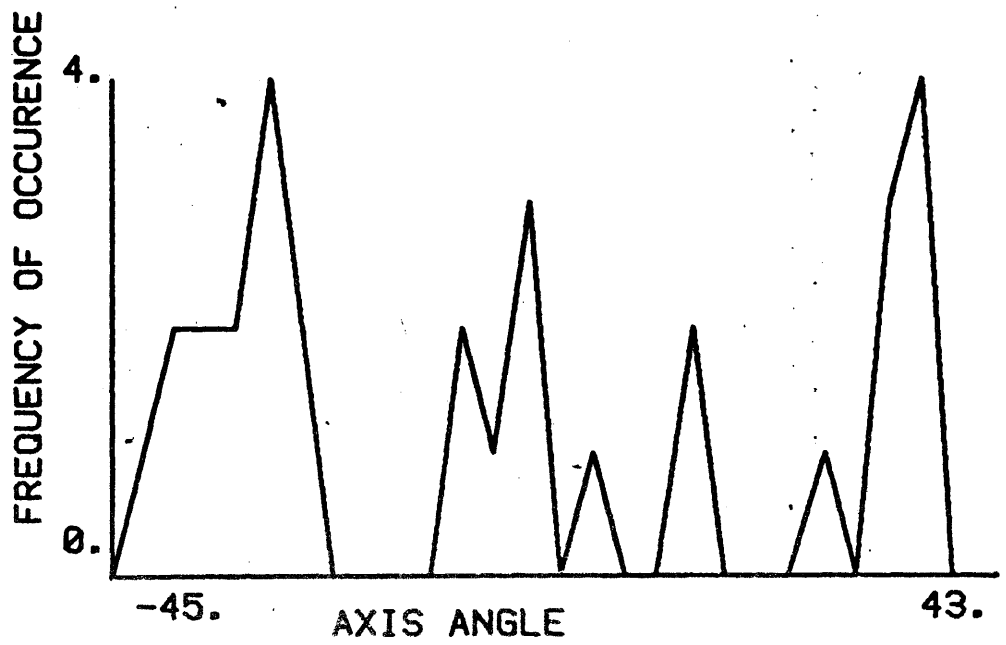

FREQUENCY OF OCCURENCE
4.
0.
2.
PCT. OF FILL
90.

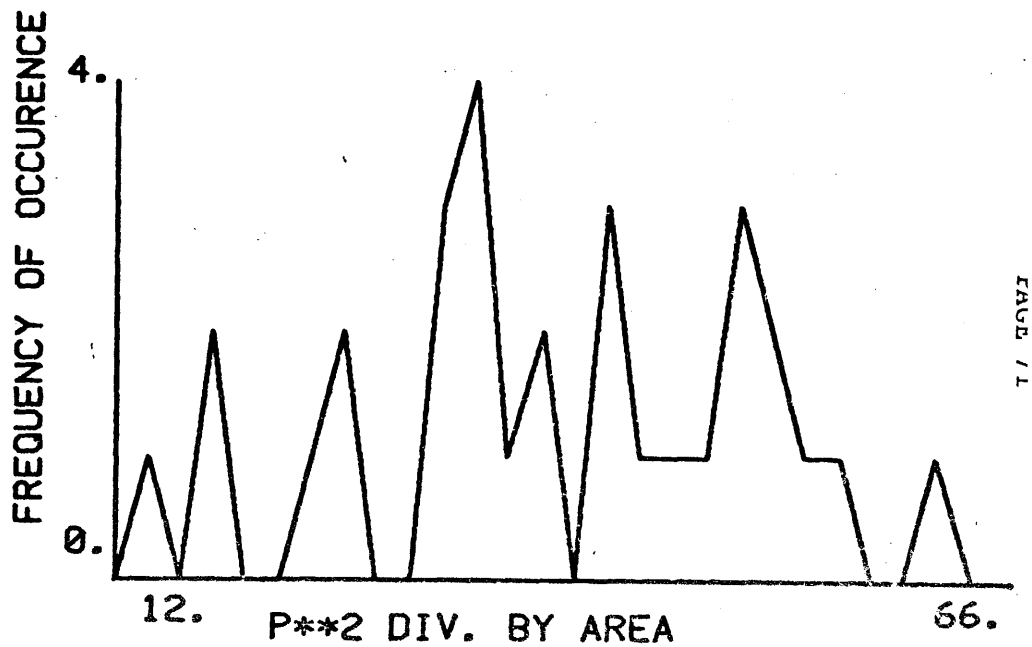

FREQUENCY OF OCCURENCE
7.
0.
2.
XWIDTH
53.

FREQUENCY OF OCCURENCE
5.
0.
16.
YWIDTH
204.

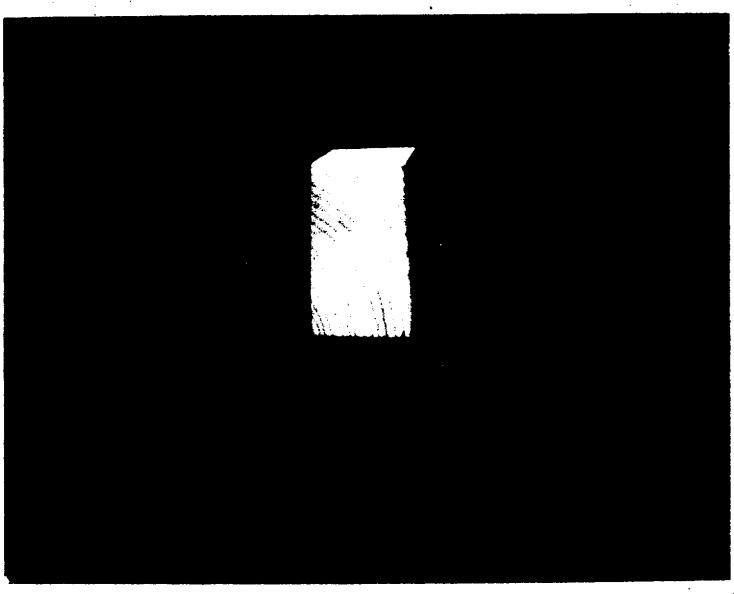

FREQUENCY OF OCCURENCE
3.
0.
37.
MOMENT FN 1
167.

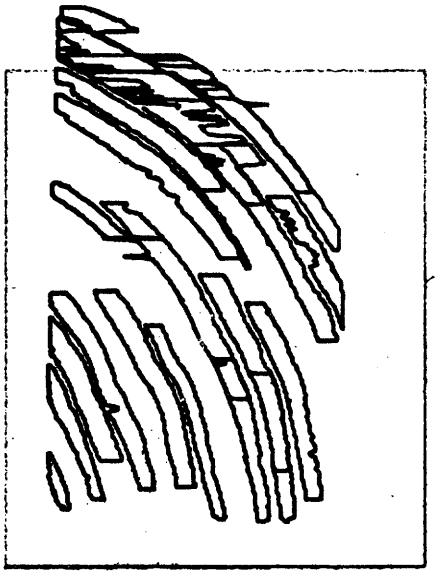

FREQUENCY OF OCCURENCE
4.
0.
33.
ECCENTRICITY
167.

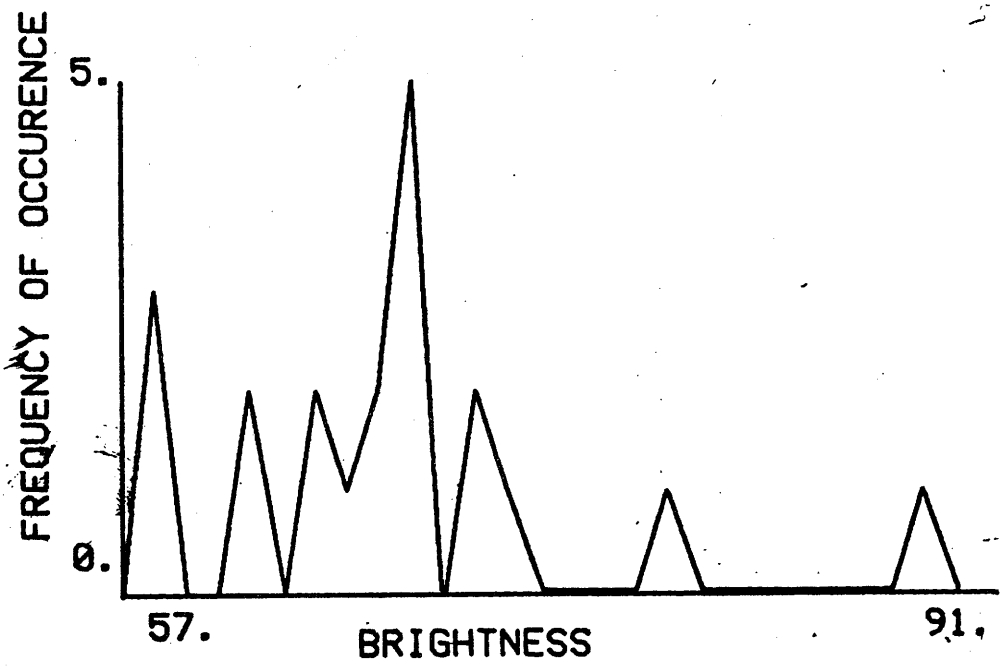

FREQUENCY OF OCCURENCE
4.
0.
-45.
AXIS ANGLE
43.

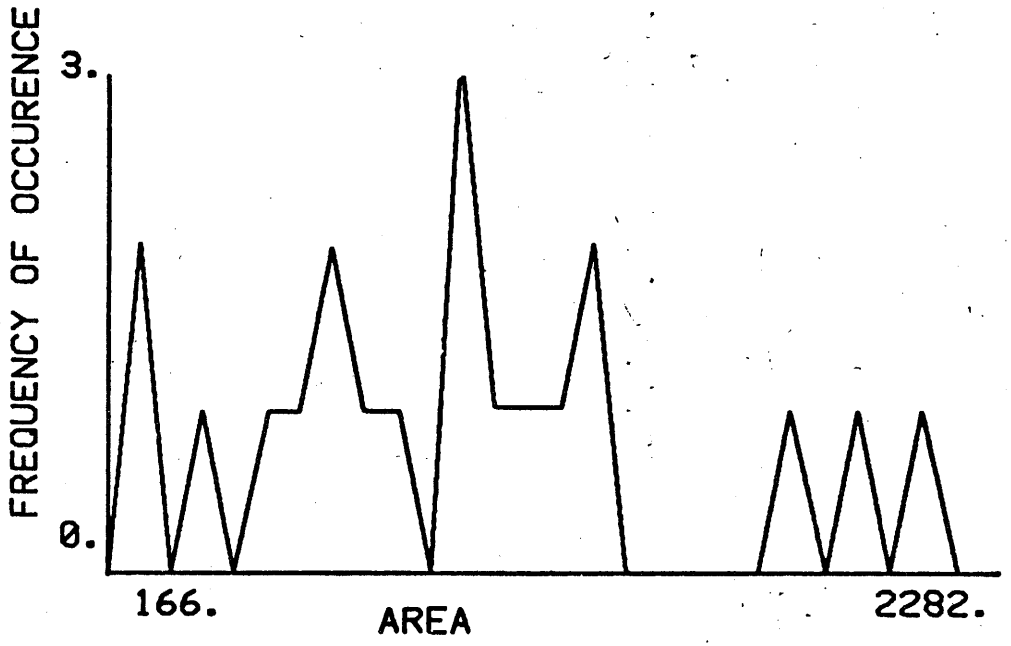

FREQUENCY OF OCCURENCE
4.
0.
12.
P**2 DIV. BY AREA
66.

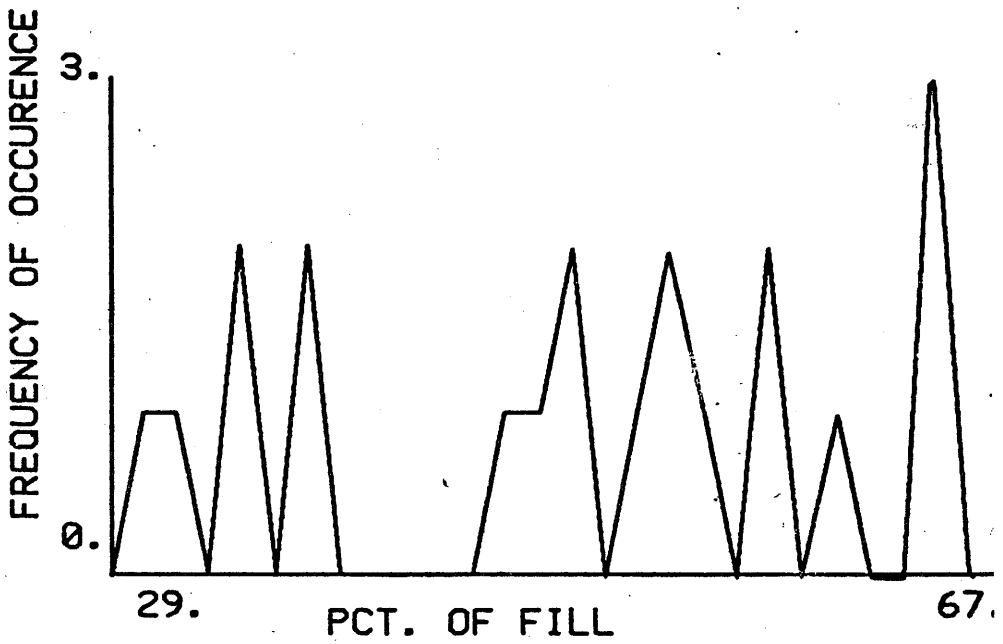
FAGE 1

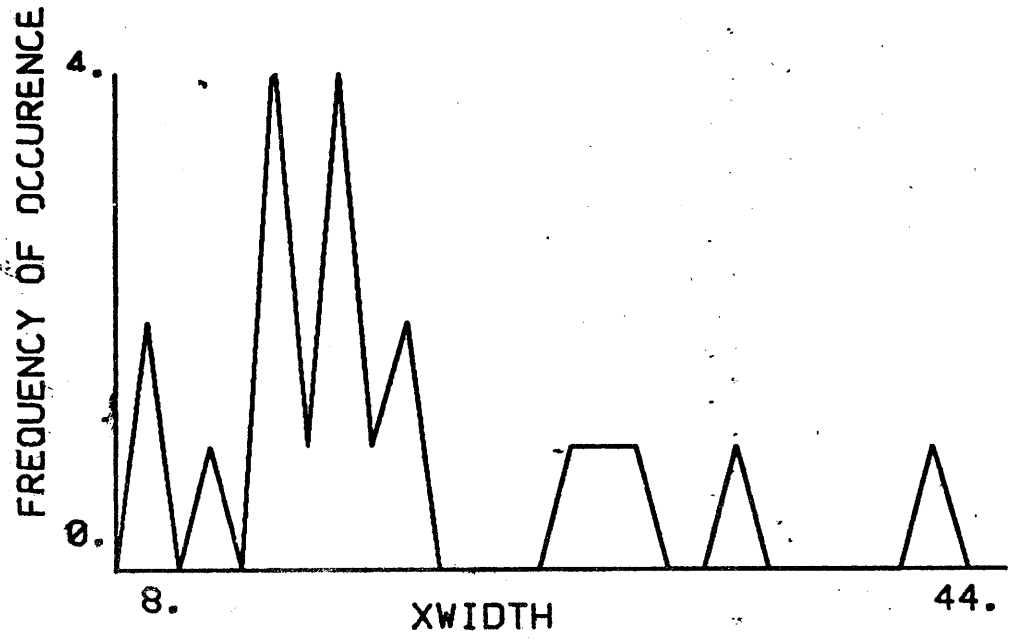


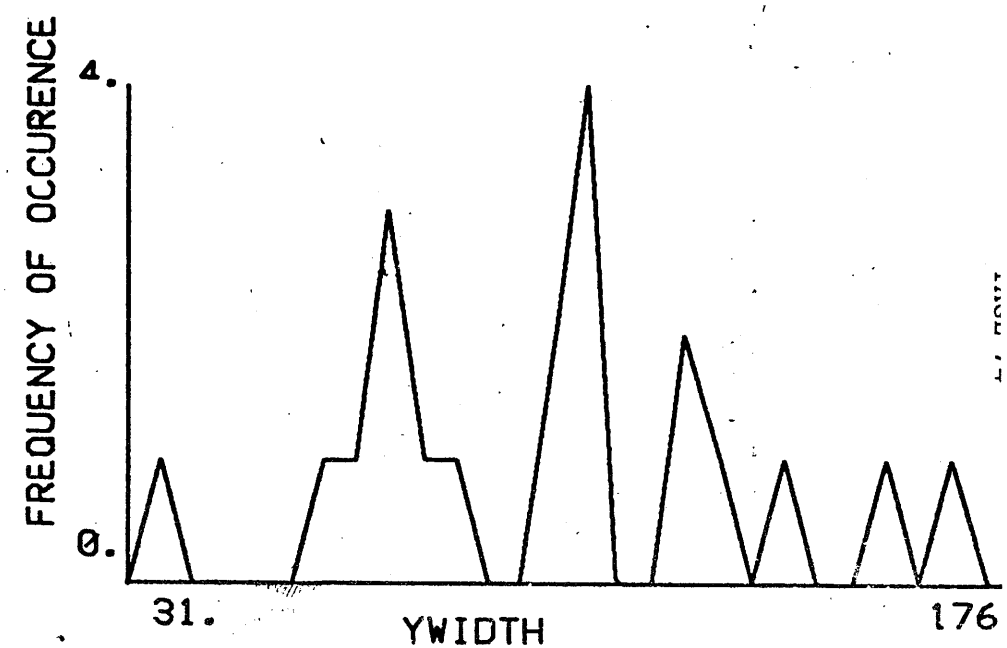


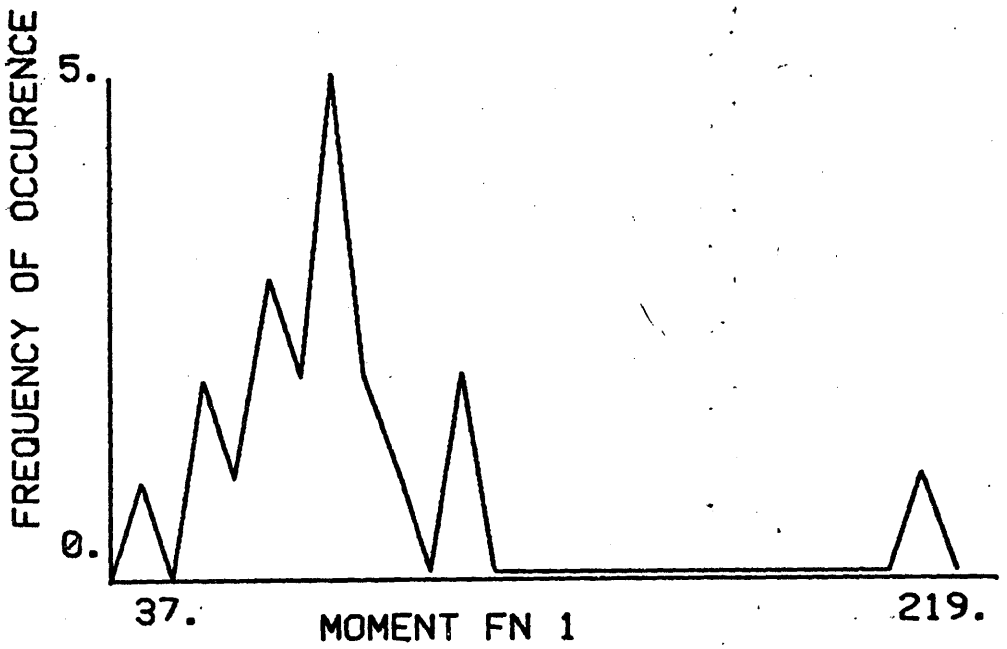

FREQUENCY OF OCCURENCE
5.
0.
57.
BRIGHTNESS
91.

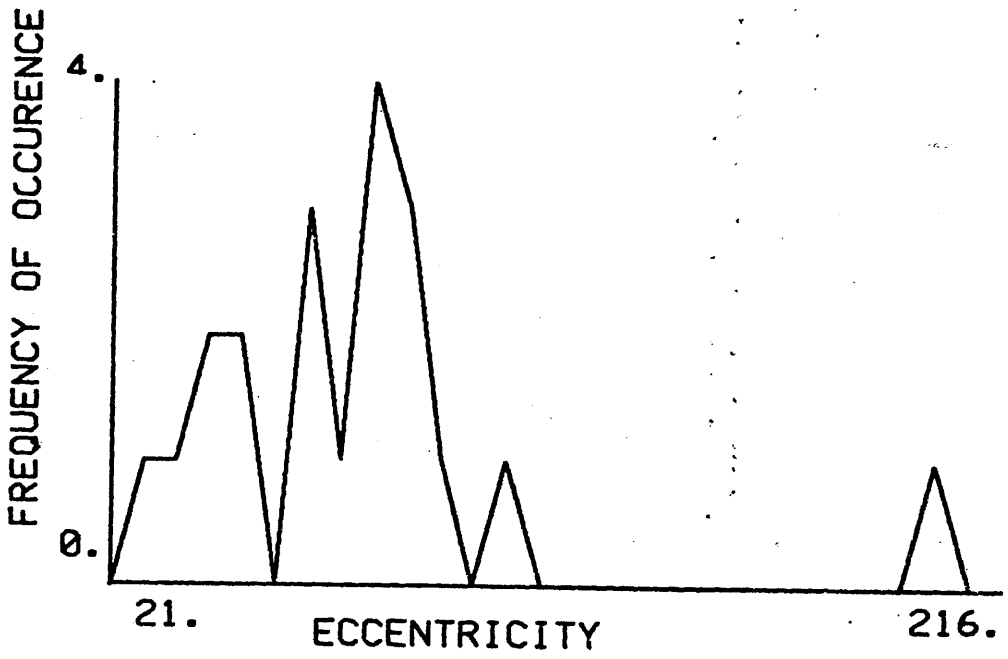

FREQUENCY OF OCCURENCE
3.
0.
166.
AREA
2282.

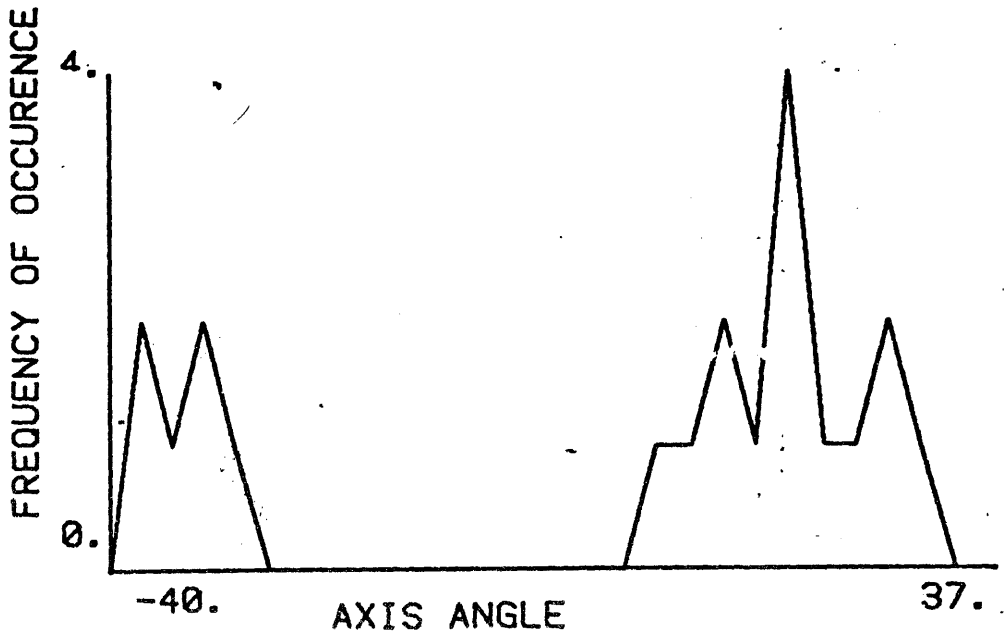

FREQUENCY OF OCCURENCE
3.
0.
29.
PCT. OF FILL
67.

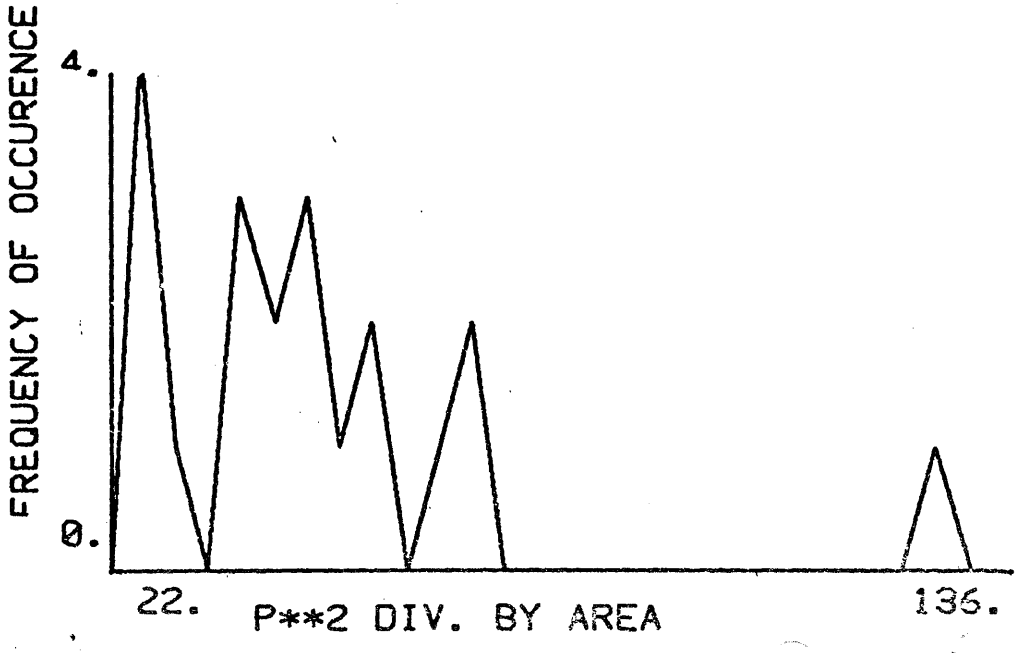

FREQUENCY OF OCCURENCE
4.
0.
8.
XWIDTH
44.


FREQUENCY OF OCCURENCE
4.
0.
31.
YWIDTH
176


FREQUENCY OF OCCURENCE
5.
0.
37.
MOMENT FN 1
219.


FREQUENCY OF OCCURENCE
4.
0.
21.
ECCENTRICITY
216.


FREQUENCY OF OCCURENCE
4.
0.
-40.
AXIS ANGLE
37.


FREQUENCY OF OCCURENCE
4.
0.
22.
P**2 DIV. BY AREA
136.