SLANT FROM TEXTURE:

Computational methods for recovering surface slant from images of textured scenes

by

Heinz Breu

B.Sc., University of British Columbia, 1978

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies
(Department of Computer Science)

We accept this thesis as conforming

to the required standard

The University of British Columbia

March, 1980

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of COMPUTER SCIENCE

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date APRIL 24, 1980.

Abstract

This thesis examines the problem of computationally recovering or determining the slant of a surface from an image of that surface. Images are restricted to those of planar surfaces produced by orthographic projection. This thesis is concerned only with those cues obtainable from the image texture. These cues arise primarily due to the foreshortening property of orthographic projection.

Texture measures have typically been partitioned into three classes: statistical approaches, micro-structural approaches, and macro-structural approaches. In this thesis, measures from each of these classes are used to develop algorithms capable of detecting surface orientation. It is concluded that these three classes are not distinct and, indeed, are artificially rendered by the prevailing definition of texture. A new definition involving nested structures is suggested.

Acknowledgements

I would like to thank my advisor Alan Mackworth for his many suggestions which I have incorporated in this thesis and for his friendly guidance when my work seemed to be at an impasse. Thanks also to my wife Wendy for all her support while I was in school and especially for her extensive help in the production of the final copy of this thesis.

Table of Contents

Table of Figures

1. INTRODUCTION

"There is more to seeing a pikestaff than is commonly believed ..."
(Boden, '77)

This thesis is about deducing certain properties of a scene from an image of that scene. These properties are the orientations of the surfaces making up the scene and the deduction of these properties is made on the basis of knowledge about the texture present in the image.

1.0 PHYSICAL ASSUMPTIONS

In order to study how perceived texture can enable a reconstruction of the scene, the properties and the events to be analysed must be isolated. One way of accomplishing this isolation is to simplify the world. The following assumptions and restrictions, then, outline our simplification. They will be mentioned again in the thesis as they become appropriate.

1) All surfaces are planar.

2) All images are produced by an orthographic projection.

3) The scenes contain only continuous surfaces.

4) Some preprocessing of the image will always be assumed. The nature of this preprocessing will be mentioned when required.

5) "Nearby" regions of the image that have "similar" textures are really the same region. The term, "similar", will be defined when it is needed. The term, "nearby", is not defined nor is it used by us. This assumption is made largely to allow the following assumption.

6) The images are segmentable into regions, each of which corresponds to a unique surface of the scene. In most cases, we will assume that this segmentation has taken place.

7) Assumptions will be made throughout the thesis as to the nature of the actual texture.

So assumptions 1) - 3) simplify the physical "reality", 4) - 6)

represent work already done which we assume and may call upon, and 7) is our major heuristic, for, without 7), any given image could have been produced from an infinite number of scenes.

1.1 DEFINITIONS

This thesis, then, confronts the task of recovering three-dimensional scene information (namely, surface slant) using texture as a cue from a two-dimensional image produced by an orthographic projection of the scene. To proceed, the terms "texture" and "surface slant" require clarification.

Texture

Texture certainly deserves to be a nebulous concept (there is a pun in there somewhere) and no definition to date has made it very clear. Texture in an image normally consists of intensity variations in the image over space. These variations may be produced by two distinct means:

1) Variations in albedo.

2) Variations in the surface normal.

This thesis assumes the former. This assumption is made as a further simplification of the world, for it is now the case that perceived texture does not depend on the position of the scene's light source(s).

The variations in albedo and the surface markings are referred to

as the actual or scene texture and the variations in the image's intensity as the apparent or image texture. This will be discussed further in Chapters 2 and 4. For now, it is sufficient to think of texture as being a large number of similar (visible) patterns or elements, each of which is small relative to the textured surface, arranged over the surface according to some set of placement rules.

Note that the apparent texture then depends on the scale of the image since different things will become the basic elements as the scale varies. For example, wheat fields viewed aerially from one kilometer have a definite apparent texture quite distinct from the same wheat fields viewed from six centimeters. Image resolution is often associated with scale, resolution decreasing with the image-to-scene ratio. Some Artificial Intelligence paradigms (Kelly, 1971) (Tanimoto and Pavlidis, 1975) vary resolution with the level of attention . Hence the apparent texture also depends on the level of attention. The determination of what constitutes the basic elements for any given textured image is quite difficult and we will largely assume that this has been done for us if we need it (assumption 4). It is felt that this determination problem is one created by an insufficient definition of texture. This opinion is discussed in Chapter 4.2.

Surface Slant

Fortunately, the definition of "surface slant" is somewhat more concrete. We will generally talk about the gradient vector (Huffman, 1971) (Mackworth, 1973) of a surface or minor variations of it. The gradient, (p,q), of a line in a scene is a vector "whose direction is in

the direction of the picture line and whose length is the tangent of the angle the scene line makes with the picture plane." (Mackworth, 1976). See Figure f1.1.1, from (Mackworth, 1976). We will say that the gradient of a surface is the gradient of the set of parallel lines having the steepest inclination to the picture plane. We will refer to the direction of this gradient vector as σ and to the arctan of its length as ϕ. The cartesian co-ordinates, (p,q), and these angles are related by the following equations:

$$p = \tan\phi \cos\sigma$$

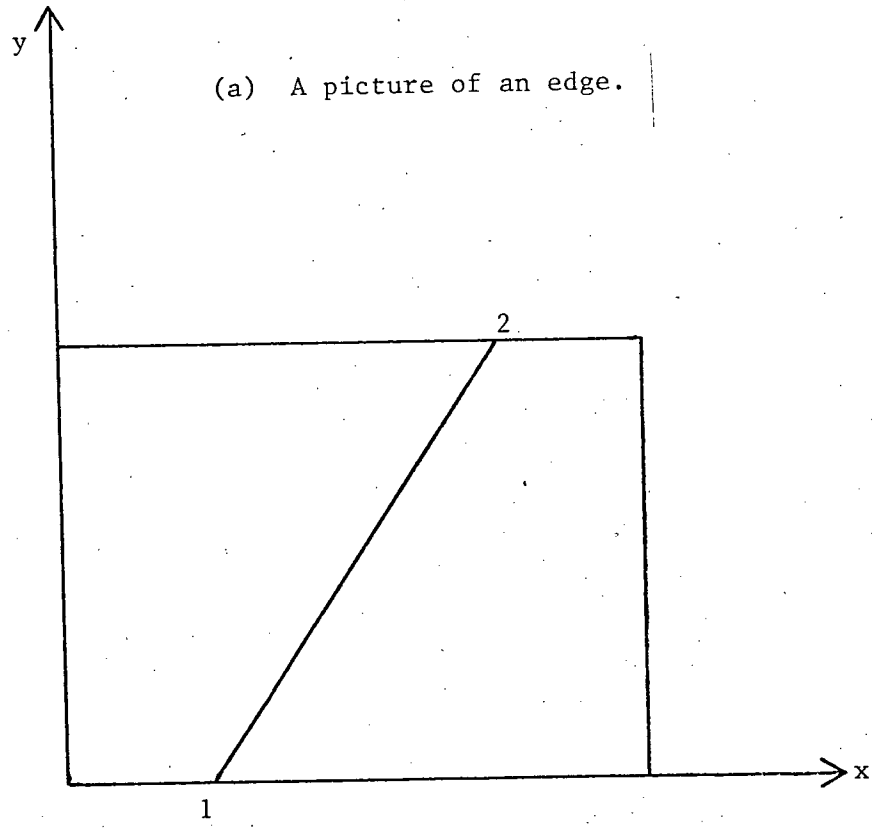
$$q = \tan\phi \sin\sigma$$

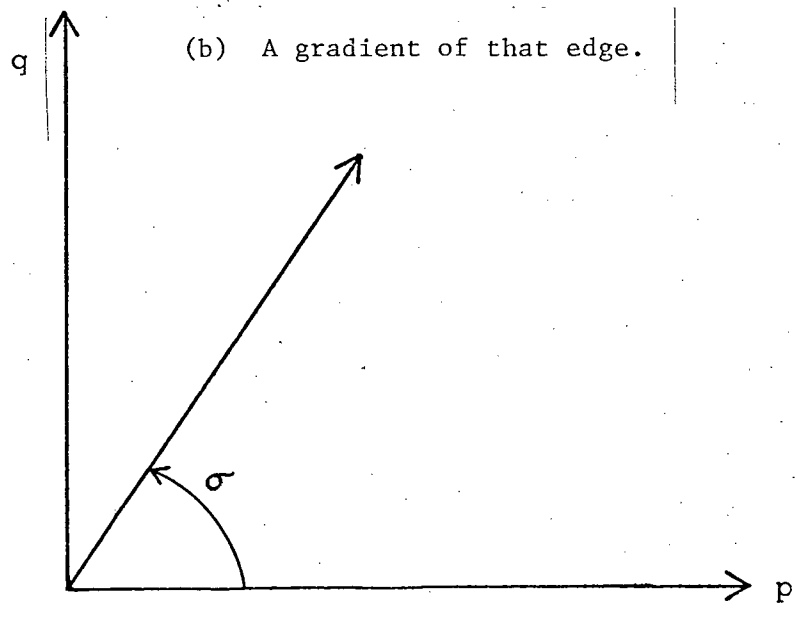
In most cases, we will find it more convenient to determine σ and ϕ than to compute the cartesian co-ordinates of the gradient directly as Kender (see Chapter 2.2) does. We will also call σ the direction of slant and ϕ the amount of slant .

1.2 MOTIVATION

It is not within the scope of this thesis to give motivations for recovering 3-D scenes from 2-D images, so such motivations will be assumed. But why study texture and surface slant? We will use an application to illustrate our motives.

The origami world is a model of the visual world where surfaces, which may stand alone, (Kanade, 1978) are the basic elements rather than polyhedra. That is, the origami world includes as a subset the solid-objects studied by (Huffman, 1971), (Clowes, 1971), (Waltz, 1972) and (Mackworth, 1973) but allows other objects as well. The origami world is





picture and gradient of a line

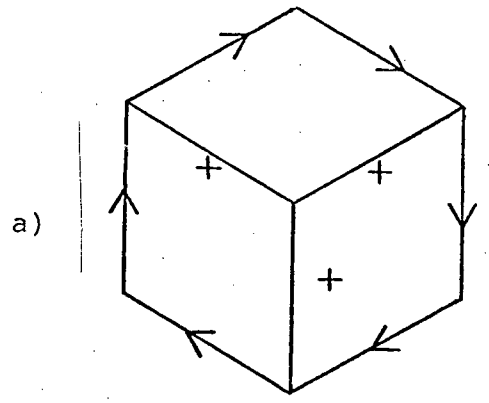
Figure f1.1.1

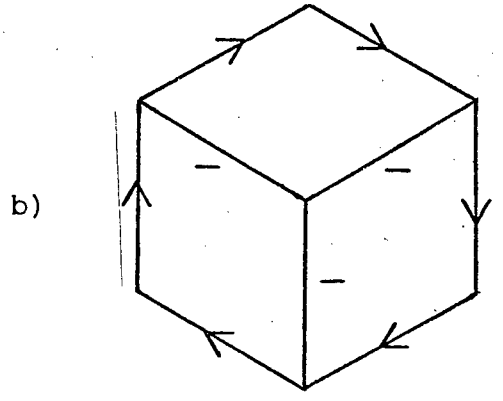
defined on line-drawings and an object is "understood" by assigning one of four labels, + (convex edge), - (concave edge) and, <-, and -> (occluding edge) to each line in the image of the object.

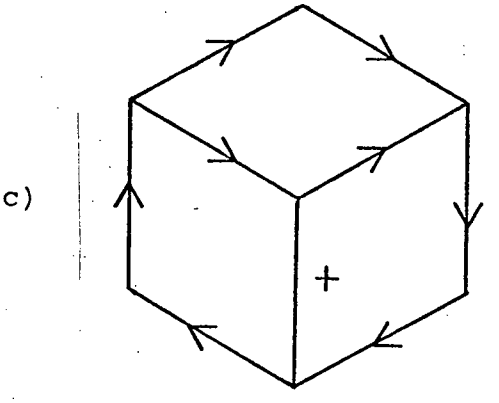
Kanade's system yields the 3 (up to rotation) labellings shown in Figure f1.2.1 for the line drawing shown. We can think of a) as a cube-like configuration, b) as a concave corner and c) as a "roof" placed on a plane.

Kanade uses edge profiles to determine which labelling is "correct" for a given image, but let us see how texture can help us here. Let us first note that the gradients of the surfaces for Figure f1.2.1 a), b) and c) are related as shown in Figure f1.2.2 a), b) and c) respectively. Suppose that all the surfaces have been textured and on this basis we have determined the gradient vector of each surface (it is the purpose of this thesis to demonstrate that this can be done). Then our findings will easily distinguish between Figures f1.2.1 a) and c), for example. Note that we will not be able to distinguish figures f1.2.1 a) and b) on this basis since texture cues will only yield the surface orientation to within Necker reversal. that is, texture will give us two gradients for any given surface, the "real" gradient and its negative or reflection about the origin.

Note that textures are easily carried over into line drawings whereas intensity profiles are not. Hence, it is not necessary to return to the original image, which may be a photograph for example, to determine surface slant from texture. This is especially useful if the photograph is not of sufficient resolution to allow utilization of edge profiles.

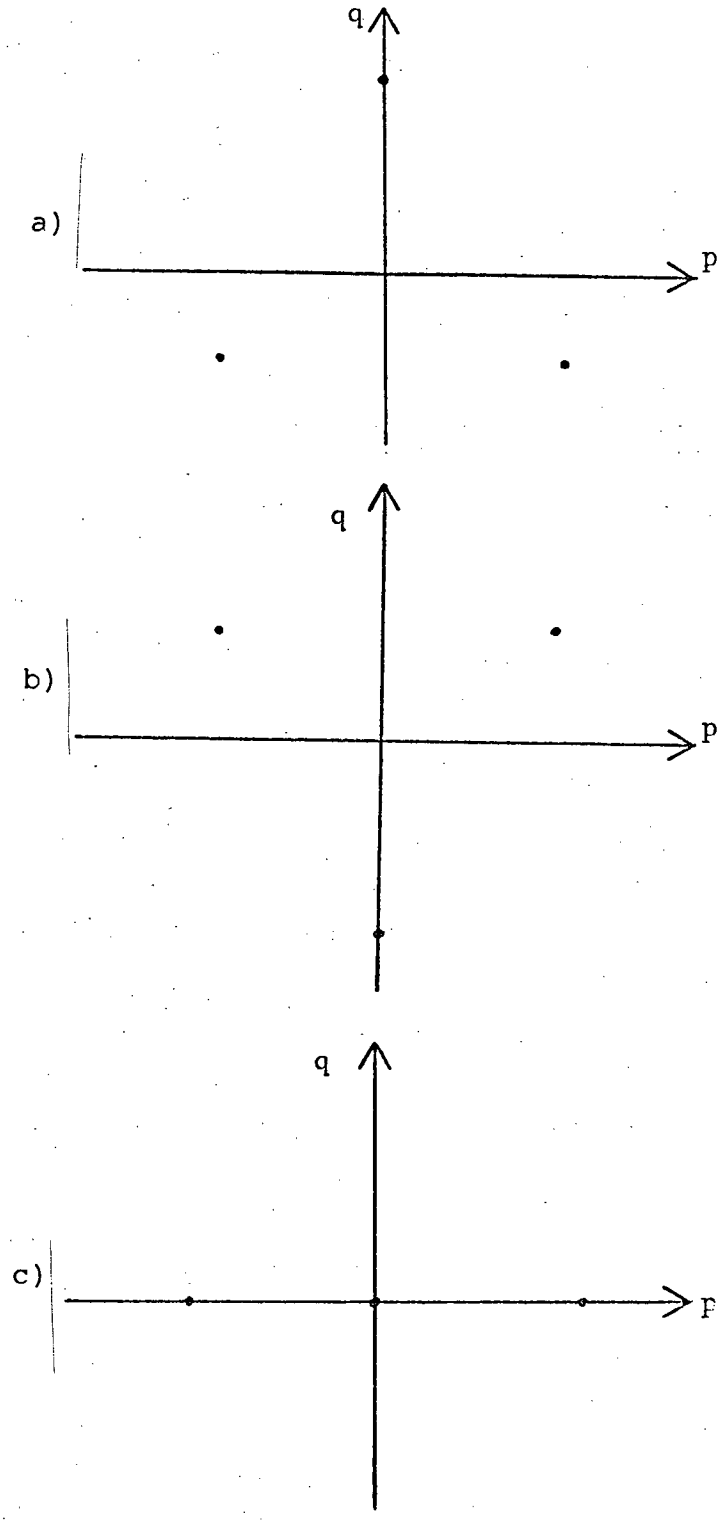






labellings of a "cube"

Figure f1.2.1



gradients of Figure f1.2.1

Figure f1.2.2

1.3 THESIS OUTLINE

It is proposed, then, that texture is a useful cue for the determination of surface slant from orthographic images. Three different "kinds" of textures will be examined and different techniques will be used to analyse them. Programs will be implemented to determine surface slant for two of these texture classes.

Let us now give a summary of the thesis by chapter.

Chapter 1 Introduction

This chapter opens by presenting the major assumptions made this thesis. Section 1.1 discusses concepts crucial to this work. Section 1.2 discusses our motivation with an example of a possible application of this work. Section 1.3 gives an outline of the thesis.

Chapter 2 Related Work

Section 2.1 presents historical attempts at defining texture. These are discussed here mainly as references to later discussions. Section 2.2 introduces the main texture measures used. We will use all but the Fourier transform in our work. Section 2.3 presents applications of these measures that lend credence to assumption 6) by demonstrating texture discrimination. Also discussed is previous work on surface slant detection and depth cues.

Chapter 3 The Solution

Section 3.1 defines "orthographic projection" and proves an important theorem about the nature of texture under this imaging process. Section 3.2 examines the use of statistical measures for determining surface orientation and proves a theorem indicating the limits of this approach. Section 3.3 undertakes a structural analysis of the texture's "basic element" and presents an implemented algorithm for detecting surface slant. Section 3.4 makes use of the Hough transform to analyse the texture's macro-structure. Again, an implemented algorithm is presented.

Chapter 4 An Interpretation of the Results

Section 4.1 summarizes the results of Chapter 3. Section 4.2 then examines a common myth about textures in light of these results.

Chapter 5 Conclusions

This chapter gives a brief summary of what has been accomplished. Several directions for further study are proposed.

1.5 READING PATHS

This thesis will, naturally, be of different interest to different readers. So, let us sketch a few "paths" through the thesis.

Definition of Texture

If the reader is interested in knowing only what we think texture is, much of the work we have done may be glossed over. But read

Sections 1.1 to get acquainted with texture.

2.1 to get a historical perspective.

3.1 to understand the imaging process.

4 and the conclusions to 3.2, 3.3 and 3.4 to get our opinions on texture.

5 the first part, to get our conclusions.

Statistical Approaches

Read 1.1, 2.2, 3.2, 4.1, 4.2, 5.

Micro-structural Approaches

1.1, 2.3, 3.3, 4.1, 4.2, 5.

Macro-structural Approaches

1.1, 2.4, 3.4, 4.1, 4.2, 5.

The reader interested in working with texture in any context is advised to at least browse through the entire thesis. In this endeavour, may I wish the reader fruitful studies.

2. RELATED RESEARCH

2.1 THE DEFINITION OF TEXTURE

Everyone is familiar with visual texture but no one seems to be able to define it adequately. Texture is different things to different people. Certainly, texture depends on one's point of view, or level of attention, as was pointed out in the previous chapter. It appears reasonable, however, to define what one is working with, and several stabs have been made at it.

The lack of successful definitions is due largely to the practice of viewing textured surfaces with the picture plane parallel to the surface and to the confusion caused by considering texture to be composed of micro-structures and macro-structures. This latter problem is discussed in chapter 4. The former results in a conceptual blurring of the distinction between scene and image. That is, definitions have been proposed which describe the actual texture in terms of the image (regions, gray-tones, arrays). Such definitions are acceptable only so long as fronto-parallel views are maintained. If the surface slant is non-zero, however, the question of texture identity and problems in terminology quickly arise. The image-based definitions would say that a non-tilted surface and a tilted surface have different textures. There is no way, in such a definition, of expressing the notion of the same actual or surface texture. Our own definition is scene based, hence we would say that the slanted surface has the same actual texture as the unslanted surface. We would agree, of course, that a different apparent texture is present. It should be realized that all of the following researchers have proposed image-based definitions.

In an excellent review of the role of texture in object perception, Pickett defines visual textures to be "two-dimensional arrays of variations" (Pickett, 1970). A more explicit definition, "texture is composed of large numbers of similar 'basic elements' or 'pieces' each of which is small relative to the size of the textured region", is proposed in (Rosenfeld, Lee, and Thomas, 1970). Haralick et al feel that "texture is concerned with the spatial (statistical) distribution of gray tones" (Haralick, Shanmugan and Dinstein, 1973).

Along these lines, Zucker has proposed a general model for texture (Zucker, 1976). This model is based on a Chomsky-style grammar. Zucker proposes an "alphabet" of micro-structures. These micro-structures are then arranged on a surface according to a set of placement rules (syntax). The result is then distorted in some way (transformational component) to yield the final texture. It should be noted that the placement rules are chosen to yield periodic structures.

It is often felt that a suitable description of a texture is one which will enable the re-generation of that texture. Rosenfeld and Lipkin have developed a system allowing texture synthesis (Rosenfeld and Lipkin, 1970). The required texture description is also in the form of (1) the nature of the micro-structures and (2) the appropriate placement rules.

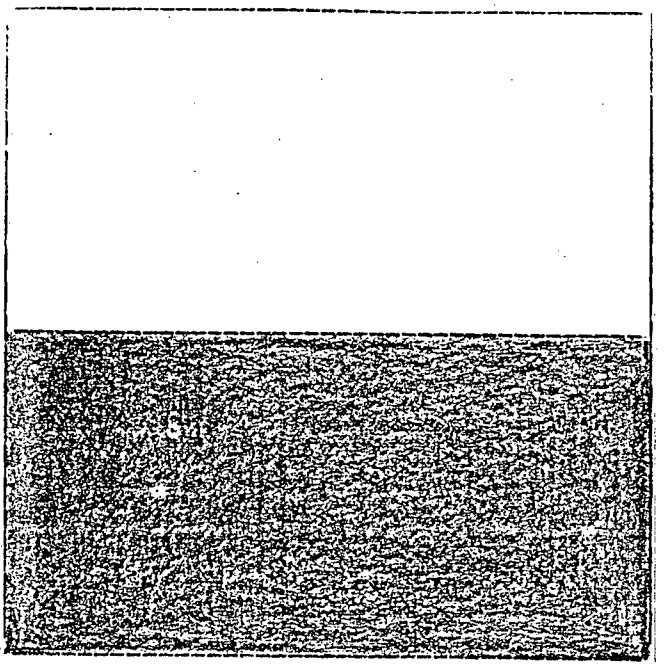
2.2 APPROACHES TO TEXTURE MEASURES

Gray-level Statistics

Gray-level statistics can be classified into 1^{st}, 2^{nd} or, in general, i^{th} order statistics. The order corresponds to the number of pixels or points considered in inter-relation.

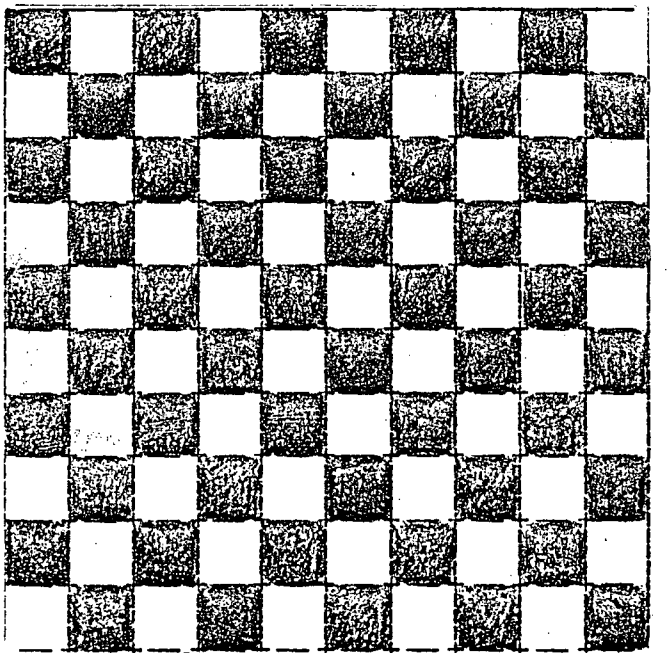
So, 1^{st} order statistics include gray-level average and variance (Hawkins, 1970) for example. Note that these values depend on the individual gray-level values only and not on their arrangement. Nothing can be deduced about micro-structure size or shape from 1^{st} order statistics. For example, figures f2.2.1 and f2.2.2 below yield the same average and variance values. In fact, all first order statistics are invariant under any pixel permutation.

2^{nd} order statistics characterize inter-relations of pairs of pixels or dots. These may be viewed as "dipole statistics" or the probabilities of "needles" of fixed length and orientation thrown on the image having given gray-level values at their tips. It has been conjectured by Julesz (Julesz, 1973) that 2^{nd} order statistical differences are necessary but not sufficient for human visual texture discrimination. That is, a pair of textures may differ in 2^{nd} order statistics and yet not be discriminable by humans. However, Julesz himself has produced counter-examples to his original conjecture; that is, discriminable textures with identical 2^{nd} order statistics were exhibited (Caelli and Julesz, 1977). Julesz modified his psychological model (and conjecture) by introducing pseudo-colinearity detectors that



Divided Square

Figure f2.2.1



Checkerboard

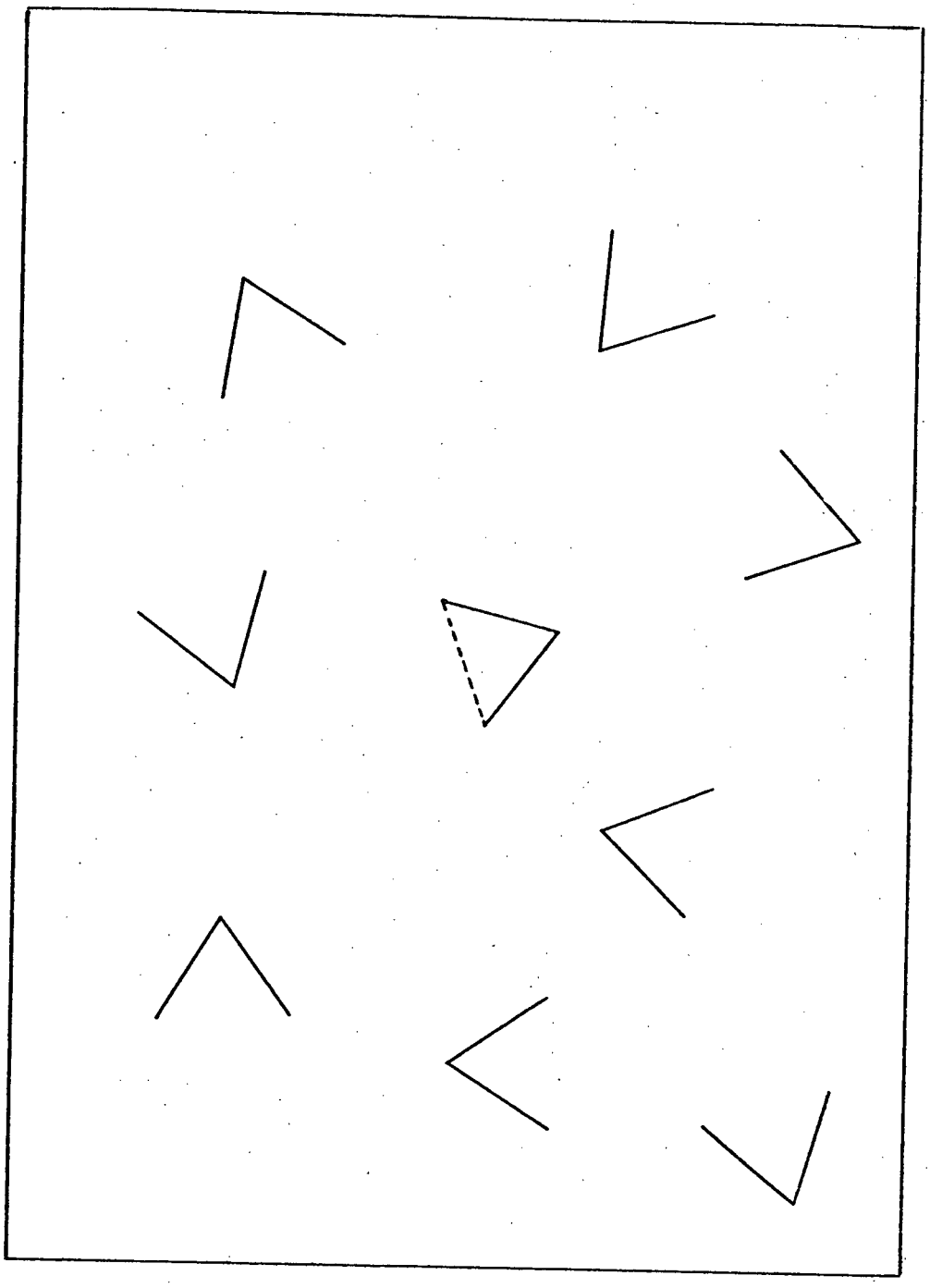
Figure f2.2.2

come into play only when 2nd order statistics are identical. Schatz modified this procedure by considering only those dipoles connecting endpoints of real or local "virtual" lines (Schatz, 1978). An example of a local virtual line is the dotted line in figure f.2.2.3. Schatz's statistics, do yield identical values for some pairs of textures where 2nd order statistics are different. These textures are not subject to human discrimination. Hence, Schatz feels that his method comes closer to providing a sufficient condition for human texture discrimination. Schatz's method assumes that the image has been preprocessed to yield "line drawings" or "primal sketches" (Marr, 1976). Of course, Schatz assumes that the original image is so preprocessable. This is not always the case, as with texture elements composed of curved lines or dots only. In these cases, all lines are virtual lines.

2nd order statistics are computed by Haralick with the help of "gray-tone spatial-dependence" or co-occurrence matrices (Haralick, Shanmugan and Dinstein, 1973). Consider a digitized image and a displacement $\delta = (D_x, D_y)$ defined on it. Then the co-occurrence matrix P_δ is defined such that

$P_\delta(i,j)$ = probability that a point with gray-level j occurs at a displacement δ from a point with gray-level i

Haralick defines 14 different texture statistics that may be computed from these matrices and suggests possible psycho-physical interpretations for some of them (e.g., "coarseness", "entropy"). Note that the displacements, δ, correspond to Julesz's dipoles. Note also that since δ is fixed in orientation, such matrices can be used to measure textural features in a given direction. This is useful for our work in detecting surface slant and also for detecting directionality in a texture. In fact, we will examine "blind" application of this



Virtual line

figure f2.2.3

technique in chapter 3.2. More to the point is our use of dipoles in chapter 3.2 (called "diagonals" in that chapter). Our detection of these dipoles, however, is implemented by a structural analysis . This type of analysis is developed next.

Structural Analysis

(Tomita et al, 1973) have used structural analysis for region detection. The assumptions made here are that the image is made up of clearly outlined atomic elements (such as circles or squares) and that each region to be detected is made up of a homogeneous set of these elements. A certain level of noise (input is via a TV camera) is also assumed and a set of ij^{th} order moments (see pgs. 78 to 79 for definitions), M_{ij}, are chosen as shape descriptors, partially for their resistance to such noise. This process, while used only for texture discrimination here, would be useful to our task as we could determine how the atomic elements have been deformed by the apparent surface slant. Nevertheless, we will not use them in this context, but for region detection only (see Chapter 3.3).

Hawkins describes the use of local "matched filters" for texture discrimination (Hawkins, 1970). This is a structural approach in the sense that structures or shapes (such as edges, lines and wedges) are being searched for, but no attempt is made to isolate possible micro-structures or to describe their shapes. As Hawkins points out, these measures are insensitive to too many image properties. In particular, they are insensitive to changes due to surface slant since no new edges result from tilting the surface.

Schatz's work may also be construed as structural. He assumes that

all real lines in the image have already been found. As well, virtual lines are constructed between real line terminators with a notion of "local". However, Schatz only looks at the overall statistics of these lines rather than their structural properties.

Frequency Domain

Consider a digital image in matrix form, g(x,y), where x and y are integers over the domain (0...k). The image's discrete Fourier transform is given by

$$F(n,m) = (1/k^2) \sum_{x=0}^{k-1} \sum_{y=0}^{k-1} g(x,y) e^{-(2pi)i[(xn+ym)/k]}$$

The power spectrum of F is given by $P(n,m) = |F(n,m)|$. The power spectrum is invariant with respect to translation in the spatial domain but not with respect to rotation. Hawkins points out that textures may differ only in their phase relationships and yet appear very different (Hawkins 1970), indicating a shortcoming of power spectra as texture descriptors. However, Bajcsy (Bajcsy, 1972,1973) uses the directional sensitivity of power spectra to some advantage. The power spectrum is transformed from cartesian coordinates to polar coordinates, $P(r,\phi)$. Then for each direction ϕ and for each frequency r we have the one dimensional functions $P_o(r)$ and $P_r(\phi)$ respectively. She then defines

$$P(r) = 2\sum_{\phi=0}^{k} P\phi(r) \qquad \text{and}$$

$$Q(\phi) = \sum_{r=1}^{w/2} P_r(\phi) \qquad \text{where w is the window size}$$

the basis for her texture descriptors is then the pair $\langle P(r), Q(\phi) \rangle$.

Bajcsy defines many descriptors from this pair. For their definition and further details of the above see (Bajcsy, 1972,1973).

Textural features derived from the Fourier domain have been found to give poorer results than statistical features in some comparative studies (Dyer, Weszka and Rosenfeld, 1975a, 1975b, 1975c). It has been suggested (Rosenfeld,1975) that this may be due to "edge effects" of the discrete Fourier transform. Some work has been done using reflected patterns to minimize this problem (Dyer and Rosenfeld, 1975).

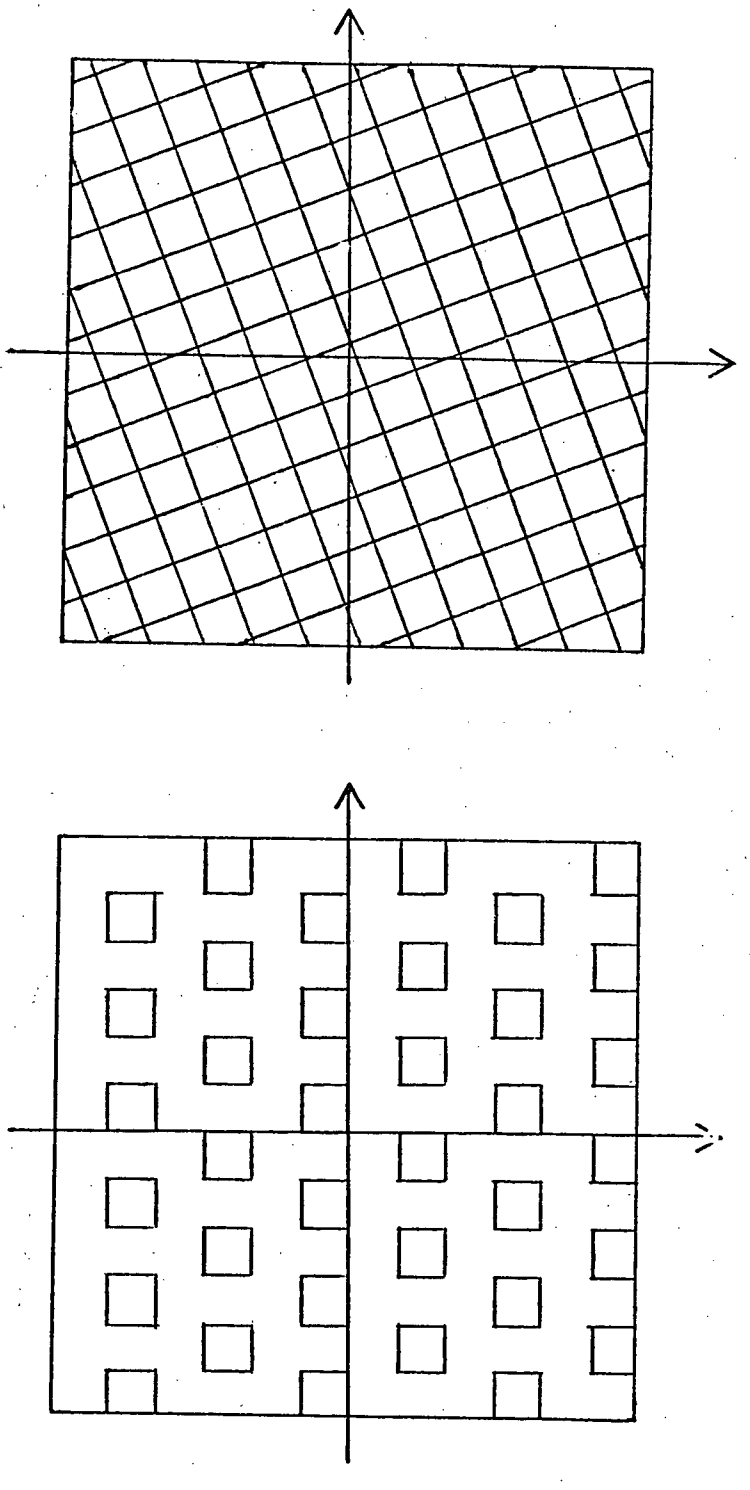
The Fourier-based approach may be useful for our work since directional features show up strongly. However, structural information is disguised by the transform, yet this is quite important information, as we have seen in the previous chapter.

The Fourier transform could be used in the macro-structural analysis presented in Chapter 3.4. Nevertheless, we found the rho-theta Hough transform intuitively more transparent for this application. This transform is developed historically in the following section.

Hough Transforms

The problem of detecting surface slant from texture gradients in the perspective projection case has been tackled by (Kender,1979). Consider the special case where the micro-structures (true shape) are "restricted to be one-dimensional and line-like; they are arranged in a regular mesh-like fashion." Examples of such textures appear in Figure f2.2.4 below.

Kender enlisted help from the techniques of a related discipline, that of line detection in images. To see how Kender uses this technique



"mesh-like" textures

Figure f2.2.4

to some advantage, it may be useful to set the stage by quickly reviewing the technique's development.

This procedure for detecting lines in an image was first proposed by (Hough,1962) and is now commonly referred to as the use of Hough Transforms. Basically, it involves transforming each point in the pattern into a line in a parameter space, in Hough's case, into slope-intercept space. The first really useful development occurred several years later (Duda and Hart,1972). Here, angle-radius (also referred to as theta-rho or rho-theta) space is used. At this stage, a closer look is warranted.

Duda and Hart decided to use normal parameterization. Referring to Figure f2.2.5 for the details of the parameters, we note that a straight line can be expressed as

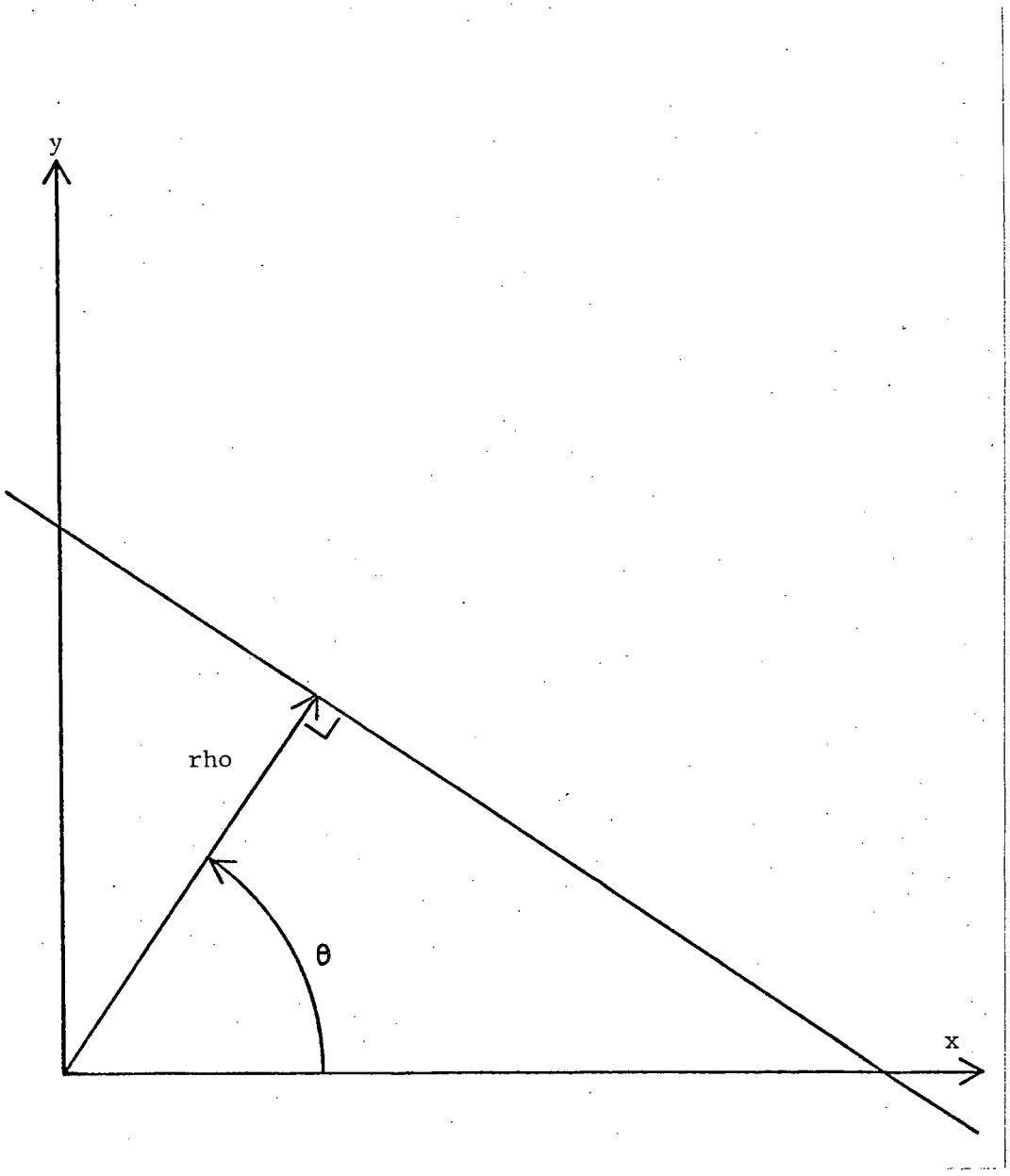
$x\cos(\theta) + y\sin(\theta) = \rho$ (see the Appendix to 2.2 for derivation)

After correlating their digitized image with a differencing operator (a way of detecting intensity changes), Duda and Hart then map each figure point (x_i,y_i) to the sinusoidal curve in (quantized) theta-rho space:

$$\rho = x_i\cos(\theta) + y_i\sin(\theta)$$

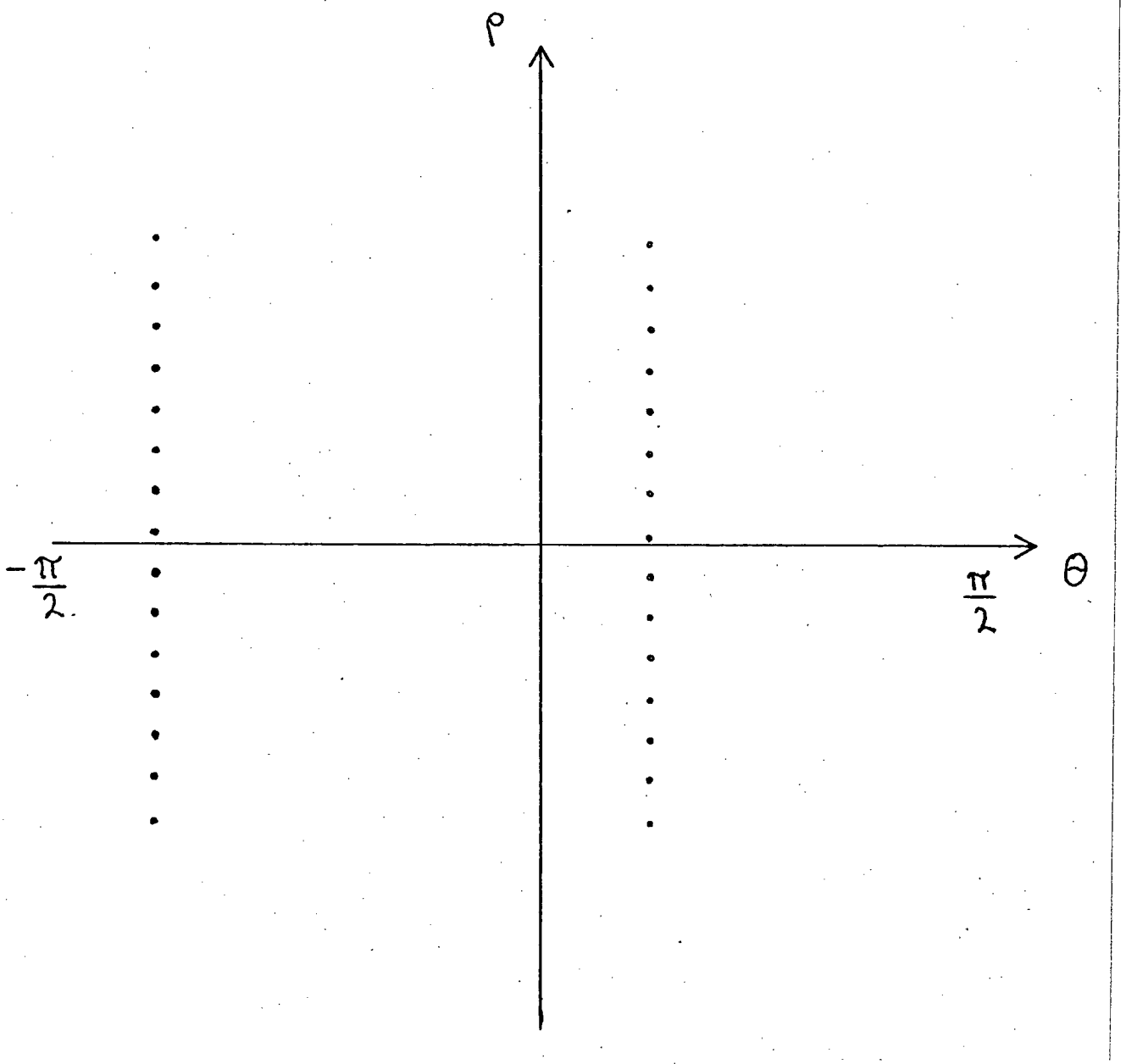
each cell in the space being augmented by 1 every time it is "written" into.

Since collinear points in the image map into lines intersecting at a point corresponding to the line they all lie on, points of maximal accumulation (detected by histogramming) each correspond to a detected line. Figure f2.2.6 below is the theta-rho Hough transform of the "city streets" of Figure f2.2.4.



Rho-theta parameterization of a line

Figure f2.2.5



Hough transform of "city streets"

Figure f2.2.6

The next major step was made by (O'Gorman and Clowes,1973). Following a suggestion by A.K. Mackworth and J. Francis, they decided to make use of local evidence to determine uniquely the value of θ. In effect, they considered each pixel in the image to be an "edgel" (having direction and position but no length). Let us see how this is accomplished. We can estimate the intensity gradient $G=(DX_{xy},DY_{xy})$ at each point (x,y) by correlating the image with the mask

-1/6	0	+1/6
-1/6	0	+1/6
-1/6	0	+1/6

to obtain DX_{xy} and with the mask

+1/6	+1/6	+1/6
0	0	0
-1/6	-1/6	-1/6

to obtain DY_{xy}.

That is, for every point (x,y):

$$DX_{xy} = 1/6\left(\sum_{i=-1}^{+1} I(x+1,y+i) - \sum_{i=-1}^{+1} I(x-1,y+i)\right)$$

$$DY_{xy} = 1/6\left(\sum_{i=-1}^{+1} I(x+i,y+1) - \sum_{i=-1}^{+1} I(x+i,y-1)\right)$$

where I(x,y) is the value of the image's intensity array at point (x,y). Now we can approximate $(\tan(\theta) = DY_{xy}/DX_{xy})$ for each "edgel" since the

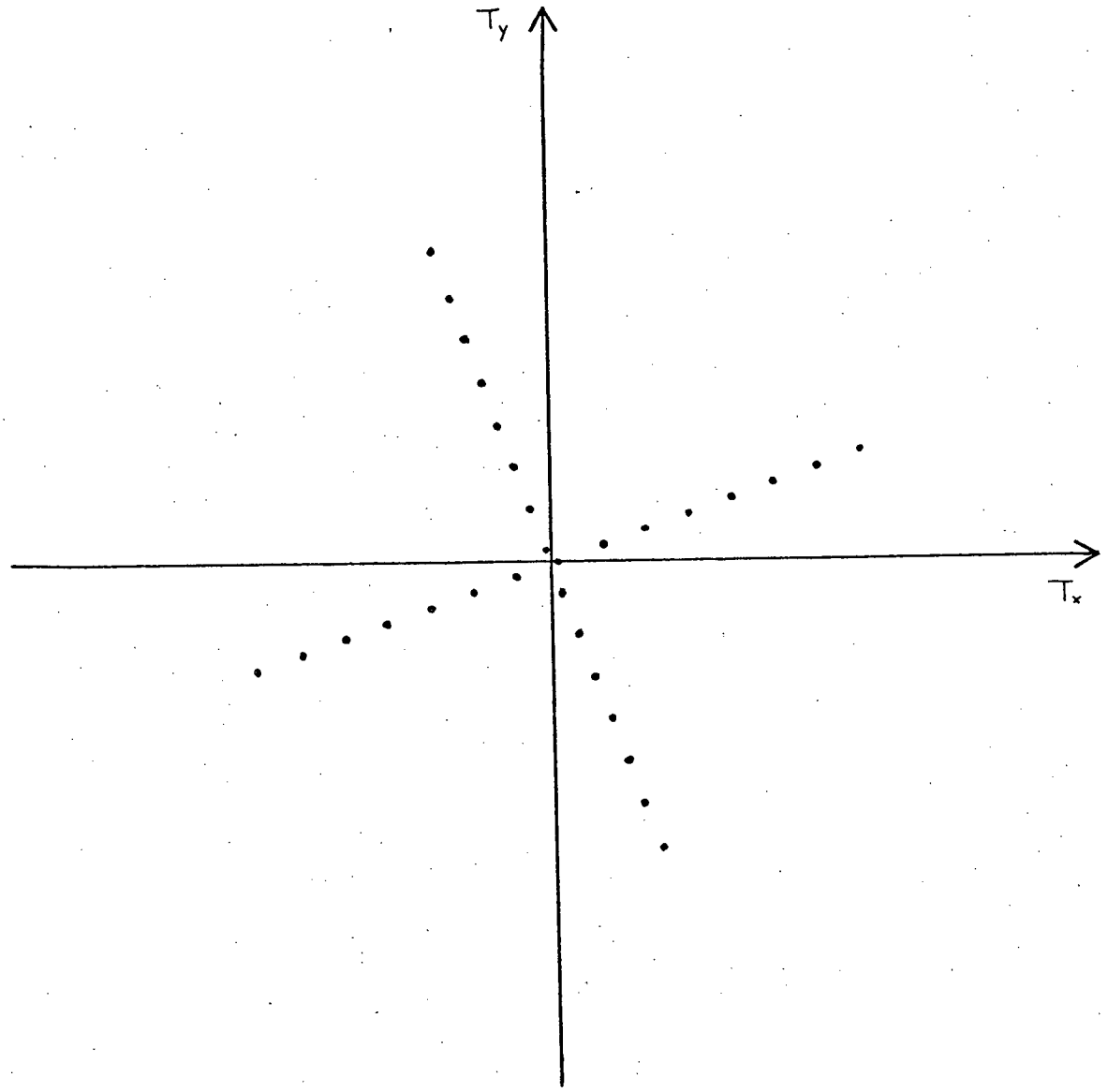
intensity gradient is generally perpendicular to the picture edge. So for each "edgel" we can increment a unique histogram bucket indexed by rho and θ.

Having developed the necessary tools, we are finally ready to look at Kender's work with texture. Kender used the observation that true parallel lines in a perspective image converge to a vanishing point (or points) to note that such lines would map into points in the theta-rho parameter space lying on the sine-curve corresponding to the vanishing point. Sinusoidal curves being rather difficult to detect, Kender proposed two modifications to the transform. The first is to plot the theta-rho transform in polar coordinates. The city streets example of Figure f2.2.4 is plotted below in this space (Figure f2.2.7). Kender noted that parallel lines in the image now map into points lying on a straight line going through the origin and converging lines map into points on a circle through the origin.

It should be noted at this stage that the distance between points corresponding to parallel (in the image) lines is precisely the normal distance between these lines in the image. This feature will be examined in Section 3.4 of the next chapter.

Kender also noted that if we describe a point in the parameter space in terms of the rectangular co-ordinates T_x and T_y then it is not necessary to compute θ explicitly. In fact, no trigonometry at all is required.

Consider an "edgel" defined by a position $P=(x,y)$ and an "edge vector" $E=(E_x,E_y)$. The edge vector is a vector perpendicular to the line segment that it refers to. The value of the edge vector's length



Polar plot of Hough transform of "city streets"

Figure f2.2.7

is left open to interpretation in Kender's paper but may be most conveniently thought of as the "strength" of the line. That is, if the image is correlated with the masks of O'Gorman and Clowes, then the edge vector may be taken to be $E = G = (DX_{xy}, DY_{xy})$. The parametric form of this edgel, $T = (T_x, T_y)$, can then be computed to be:

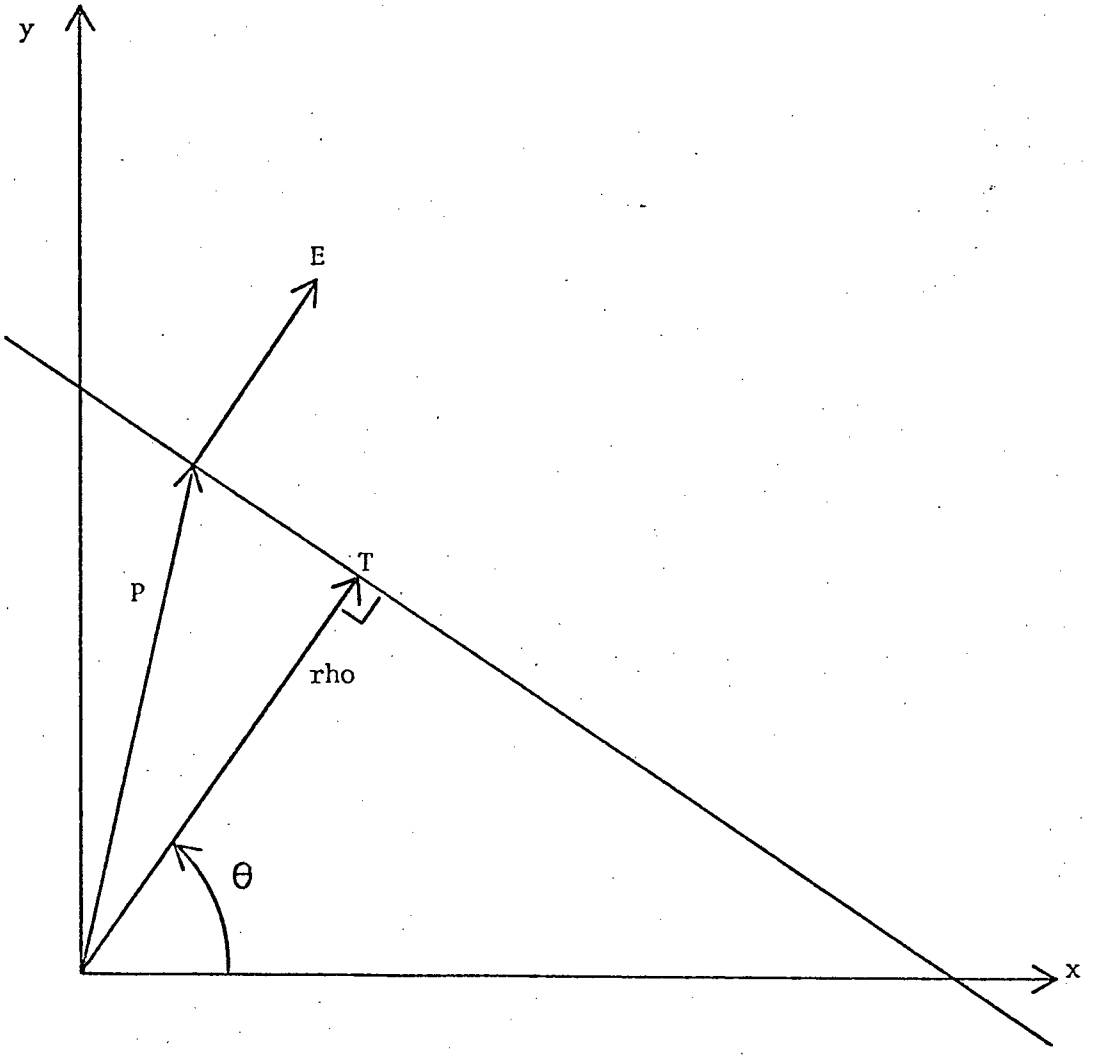
$$T = ((E \cdot P)/|E|^2)E$$

That is, T is in the direction of E, perpendicular to the line segment, and is of length $(E \cdot P)/|E|$ which is just the length of P projected onto the unit vector in the direction of E and T, namely rho (see Figure f2.2.8).

Kender's second modification is simply to map every point, (rho,θ), to $(K/rho,\theta)$ for some constant K. So, T is still in the direction of E, but is now of length K/rho. That is,

$$T = (K/rho)(E/|E|)$$
$$= (K/[(E \cdot P)/|E|])(E/|E|)$$
$$= (K|E|E)/([(E \cdot P)|E|]$$
$$T = (K/(E \cdot P))E$$

Converging lines now map into points lying on a straight line, parallel lines into points lying on a straight line through the origin.



Derivation of T

Figure f2.2.8

Kender then proves the following theorem:

> "Suppose an image contains the perspective projection of a planar surface defined by two or more coplanar sets of parallel line segments. Let the origin of the transform space correspond to the coordinates of the focal point in the image (that is, where the camera is "aimed"). Let R equal the focal distance. Then $T = R\ E\ /\ (E \cdot P)$ transforms edges so that the intersection of lines of accumulation points in the transform space is at (p,q), the gradient vector of the surface."

And so the surface's slant is detected. Notice, however, that this method fails if the projection is orthographic or if the vanishing points are very far away since the intersection point will not be distinquishable from the origin. This method also relies on our distinquishing between parallel lines that converge due to perspective and non-parallel lines. This, as was pointed out earlier, is not always that easy to do.

Nevertheless, Kender has introduced modifications that make the Hough transform computationally more desirable. Chapter 3 will discuss the uses of the Hough transform for orthographic projections.

2.3 APPLICATIONS

Discrimination

The most common application of texture is the discrimination of homogeneous textured regions. This is typically concerned with planar surfaces viewed with the surface normal perpendicular to the picture plane. In this case the image texture is the same as the scene texture. As was mentioned before, this identity is the source of a great deal of confusion in defining texture. These are generally "real-world" applications such as terrain classification from satellite images (Dyer, Weszka and Rosenfeld, 1975a,1975b) and micro-biological images (Hawkins, 1970).

An exception to the above is the work of Tomita *et al* (Tomita, Yachida and Tsuji, 1973) mentioned in the preceding section under "Structural Analysis". The input picture here could be of a textured cube on a differently textured table-top, for example. The image is then segmented into regions on the basis of several textural features. An important aspect of this work is that segmentation occurs in *stages* or *levels* as determined by the "supervisor". For example, the cube image is first segmented into the "cube" region and the "table-top" region on the basis of one feature. The cube region is then further segmented into regions corresponding to the three visible surfaces of the cube. Note that at the top level the cube is seen as being one region. The scene texture is indeed the same over the whole cube. The image texture, however, is different for each region representing a face of the cube. It is a strength of this program that it detects these

levels.

Immediate discrimination of more abstract (and perhaps artificial) textures has been studied by Julesz (Julesz et al, 1973) (Caelli and Julesz, 1978) from a psychological point of view and by Schatz (Schatz, 1978) from a computational point of view. Only absolute discrimination (are the two textures different or not?) is considered by all of these researchers, although recently (Caelli and Julesz ,1978) have mentioned "ease of discrimination". No interpretation (such as apparent texture from true texture) was placed on differences in apparent texture.

In the above cases, discrimination was performed by generating a histogram of the discriminating feature, setting thresholds from the histogram, and segregating textures as to how the feature applied over the texture related to the threshold. For example, all atomic regions with area greater than the threshold are grouped into one region (Tomita et al, 1973). We will demonstrate this in Chapter 3.3.

Surface slant

Gibson has studied the use of texture gradient (the change, due to perspective, in some textural features such as coarseness) in human "space perception" (Gibson, 1950). Gibson has shown that, in the case of perspective images, discontinuities in texture gradient give rise to the perception of edges, such as caused by corners or discontinuities in the surface ("occluding edges"), and that the texture gradient could be used to determine the orientation or shape of a surface. Computationally, many problems arise. Among them: it is difficult to establish if the texture gradient over a region in an image is caused by perspective or the curve of the surface. Gibson also observed that the

monocular perception of depth depends on texture gradient to some degree.

We have already mentioned the work by Kender in this field, and we now note that, he also, exploits texture gradient in perspective images by computing the vanishing point. This thesis does not exploit texture gradient, as will be discussed later.

Both of Gibson's and Kender's theories do not determine surface slant in orthographic images. It is the purpose of this thesis to present a theory that succeeds in this respect.

Depth cue

Computations based on texture gradient have actually been used as a depth cue (Bajcsy and Lieberman, 1976) in perspective images. The assumption is made that all surfaces are planar. "Real-world" scenes (such as a grassy field and an ocean view) were considered and images of these were used as input to the program. Features derived from the Fourier domain were used and relative distances were correctly deduced. It was conjectured in this work that absolute distances could be calculated if the program contained a camera model with information such as the lens focal length. Some problems were encountered when the planar surface assumption was violated (the grassy field rolled a bit). This work, may have benefited if discrimination of surface slant had been incorporated into the program.

2.4 SUMMARY

Texture is usually defined as many basic elements (micro-structures) on a surface. Each micro-structure may or may not have a structure associated with it. As will be further explained in Chapter 3.1 and 3.3, texture will be viewed as micro-structures placed on a surface according to a macro-structure (placement rules). The word "texel" is used to refer to a projected image of a micro-structure. The shortcomings of this approach will be discussed in Chapter 4.4.

Many texture measures have been developed and they fall into three broad classes: statistical, micro-structural and macro-structural. All three classes will be used in the next chapter, but, as we shall see in Chapter 3 and discuss in Chapter 4, these class distinctions are somewhat ambiguous.

It will be assumed that texture discrimination is possible in the remainder of this dissertation. Furthermore, work concerned with determining surface slant from images of textured scenes has been restricted to perspective images. The rest of this thesis will extend this work by examining texture in orthographic projections.

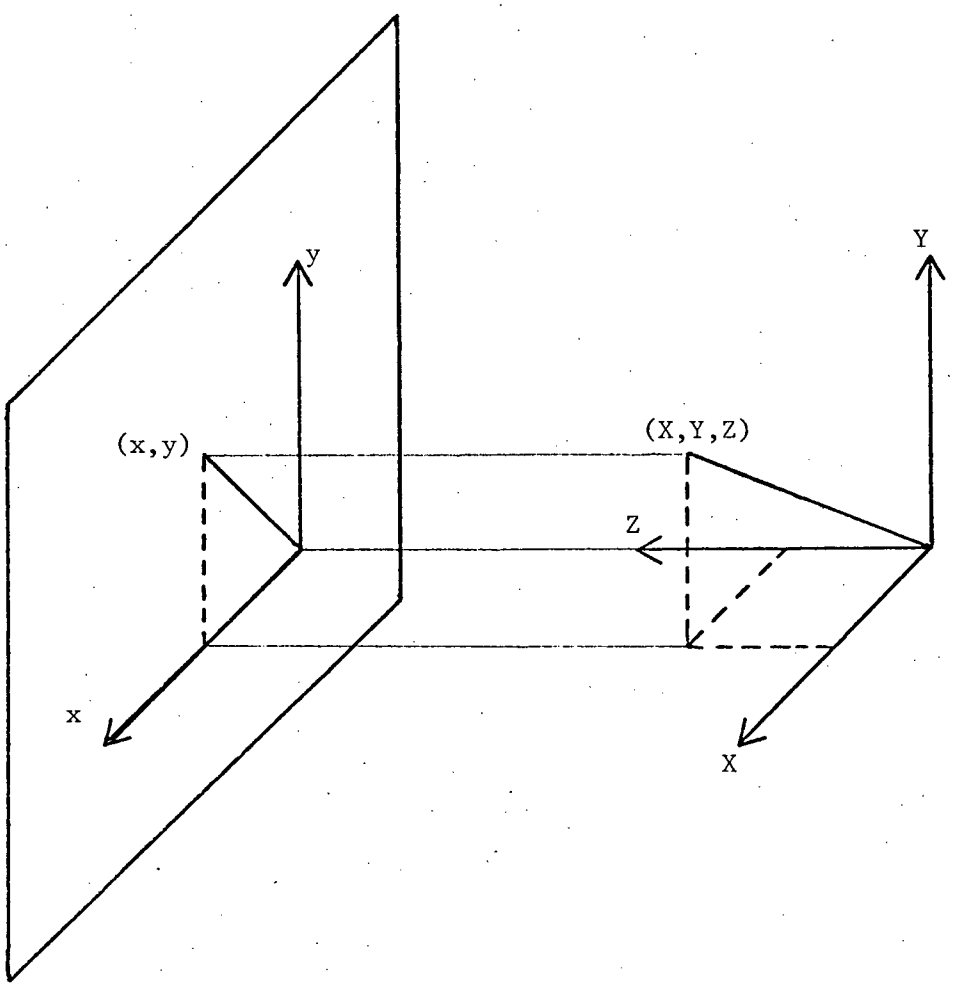
3. APPROACHES TO COMPUTATIONAL DETERMINATION OF SURFACE SLANT FROM TEXTURE MEASURES

3.1 INTRODUCTION

It is demonstrated in this chapter that information about a surface's orientation can be extracted from the apparent texture of the surface's orthographic image. This demonstration is constructive in that we will actually specify algorithms to accomplish this task (and in some cases our computer will accomplish it). With this end in mind, let us now look at the orthographic picture taking process.

Let us suppose that we have represented a given object in terms of the Cartesian co-ordinates of all of the points on its surface. This object, and the space that it is in will be referred to as the scene. Now let us position a picture plane in such a way that the x- and y-axes of the picture plane are parallel to the X- and Y-axes of the scene's frame of reference and the picture plane origin lies on the Z-axis of the scene's reference frame. We will say that a point, (x,y), on the picture plane is an orthographic projection of a point (X,Y,Z), in the scene if and only if $x = kX$ and $y = kY$ for some constrant k. This process is illustrated for a line in figure f3.1.1.

Suppose now that our scene consists of a planar surface with lines and dots on it. A surface is considered to be "infinite" in all directions (the extent of any lines and dots on it, however, is finite). It is easy to see that the surface, unless it is parallel with the picture plane, intersects the picture plane along a straight line. Without loss of generality, we can think of this intersection line as being parallel to the x-axis (if it is not, we need simply rotate the



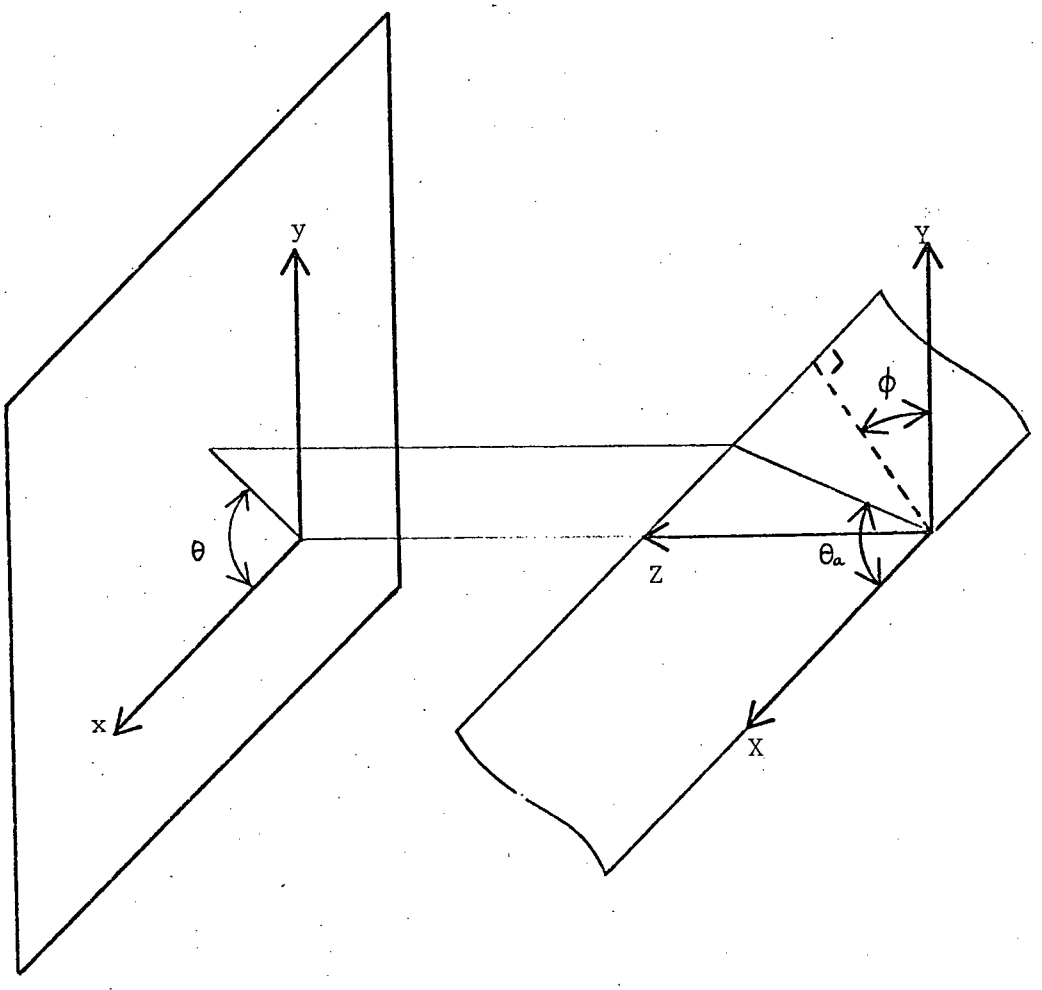
Orthographic projection of line

Figure f3.1.1

frame of reference about the z-axis until it is). We will also assume for simplicity that the surface passes through the reference frame origin and that any line on the surface also passes through the origin unless otherwise mentioned. These last two assumptions are made simply to allow referring to absolute co-ordinates, (x,y), rather than being forced to use relative co-ordinates (delta x, delta y).

We now want to talk about how the surface is slanted with respect to the picture plane. To make this task somewhat easier, we define the following terms. We define the amount of slant, ϕ, of a surface as the angle between the X-Y plane (or the picture plane) and the surface's gradient vector. If the frame of reference is rotated as we assume it to be, then ϕ is the angle between the Y-axis and that line on the surface which is normal to the X-axis. For any line (passing through the origin) we define θ_a (a for actual) to be the angle between the line and the X-axis. We also define θ to be the angle between the projected image of the line and the x-axis. Note that $\theta \leq \theta_a$ in general. These concepts are illustrated in figure f3.1.2.

It is the goal of this chapter to recover ϕ from information obtained only from the image (and the assumptions we make about it). In general, for a given image and a given surface, the surface will not intersect the picture plane at a line parallel to the x-axis, contrary to our assumption. In fact, the intersection line will be at an angle σ_n (which may be 0) from the x-axis of the picture plane. Recall that σ is the direction of the surface's gradient vector, hence the notation σ_n to indicate that the intersection line is normal to the direction of the gradient. It is also our goal to recover σ from the same information. For the purpose of illustration, however, let us stick to the $\sigma_n = 0$



Surface slant notation

Figure f3.1.2

assumption.

The feature of this projection that is most useful to our work is the effect of "foreshortening" of lengths. To see what this means, let us prove the following theorem.

Theorem 3.1.1

Assume that all of the lines on the surface are of length L and that these lines take on all orientations ($-\pi/2 \leq \theta_a \leq \pi/2$). Let the surface be tilted by an angle ϕ. Let L' be the length of a projected image of a line. Then a polar plot of L' versus θ is an ellipse centred on the origin with a major axis a = L and a minor axis b = $L\cos(\phi)$.

Proof:

Referring to Figure f3.1.3, it is noted that

$$X = OA = L\cos\theta_a$$

$$x = O'A' = L'\cos\theta$$

and that X = x since OA = O'A' so

$$L\cos\theta_a = L'\cos\theta \qquad (1)$$

It is also observed that

$$\begin{aligned} Y = BP = CD &= OE \\ &= OD\cos\phi \\ &= AP\cos\phi \\ &= (OP\sin\theta_a)\cos\phi \\ &= L\sin\theta_a\cos\phi \end{aligned}$$

$$\begin{aligned} y &= A'P' \\ &= L'\sin\theta \end{aligned}$$

and Y = y since Y = BP = A'P' = y so

$$L\sin\theta_a\cos\phi = L'\sin\theta \qquad (2)$$

Then, dividing (2) by (1) we get

$$\tan\theta = \tan\theta_a\cos\phi \qquad (3)$$

To see the ellipse in L', we now want to express L' as a function of a, b, and θ. Since a = L and b = $L\cos\phi$, we want to express L' as a function of L, $\cos\phi$, and θ. Now

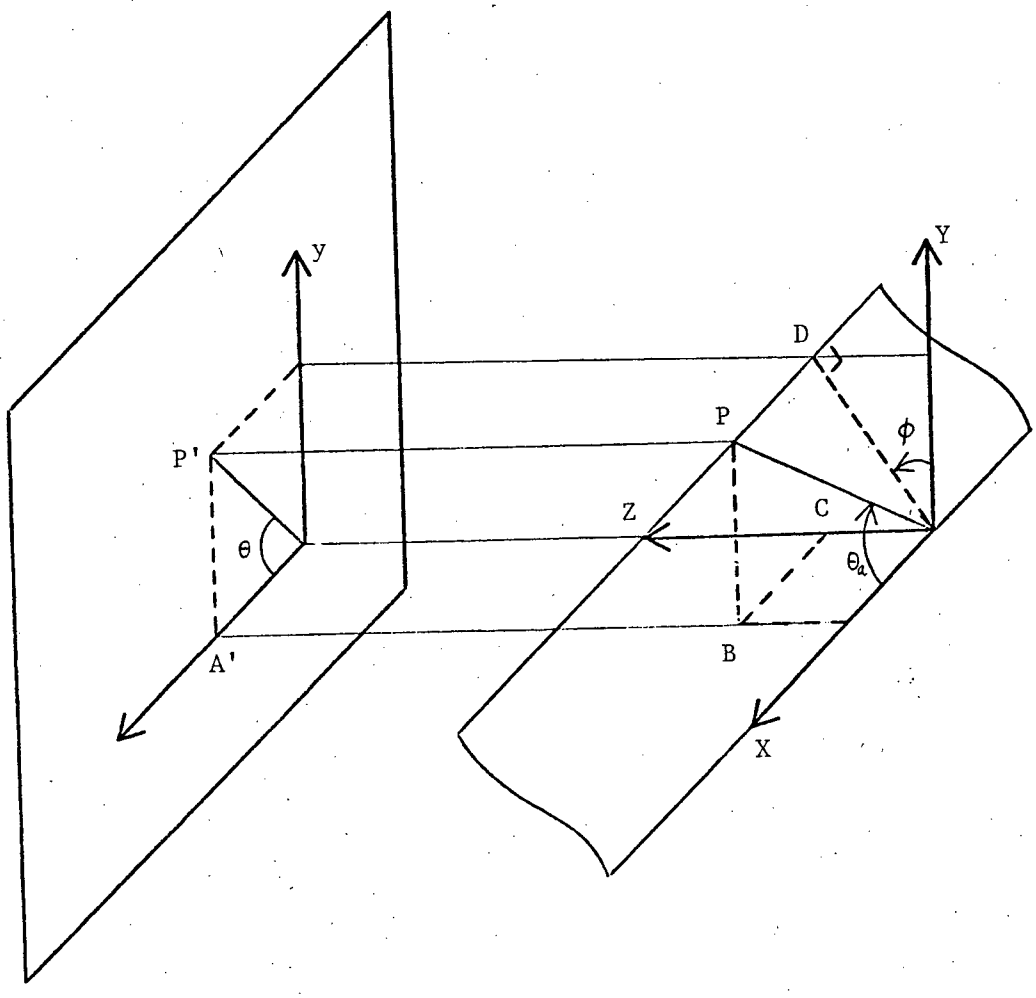


Figure for theorem f3.1.1

Figure f3.1.3

$$L'^2 = \frac{L^2\cos^2\theta_a}{\cos^2\theta} \qquad \text{by (1)}$$

$$= \frac{L^2}{(1+\tan^2\theta_a)\cos^2\theta} \qquad \text{(trigonometric equality)}$$

$$= \frac{L^2}{(1+\frac{\tan^2\theta}{\cos^2\phi})\cos^2\theta} \qquad \text{by(3)}$$

$$= \frac{L^2\cos^2\phi}{\cos^2\theta\cos^2\phi + \sin^2\theta}$$

then multiplying by L^2/L^2 we get

$$= \frac{(L^2)(L^2\cos^2\phi)}{L^2\sin^2\theta + (L^2\cos^2\phi)\cos^2\theta}$$

$$= \frac{a^2b^2}{a^2\sin^2\theta + b^2\cos^2\theta}$$

which is the equation of an ellipse.

Q.E.D.

It is now fairly easy to see that any given line in the scene will be maximally foreshortened in the direction of the surface slant (the direction of the gradient vector) and not foreshortened at all in a direction normal to it. Furthermore, theorem 3.1.1 allows us to calculate the amount of slant, ϕ, as:

$$\phi = \cos^{-1}(b/a)$$

There is, nevertheless, a defect in the method of determining ϕ from foreshortening. From the theorem just proven, lines in the scene are foreshortened by a factor $\cos\phi$ in the direction of the surface's gradient. But due to the invariance of the function, cos, over the sign of its argument, it is possible to determine only the <u>absolute value</u> of

ϕ from the foreshortening. That is, since $\cos\phi = \cos(-\phi)$, we cannot distinguish between the image of a scene and the image of its Necker reversal.

With these observations in hand, let us now look at texture.

3.2 STATISTICAL APPROACHES

Second order statistics have been used by Julesz, Haralick, and Schatz (see Chapter 2 for details). These statistics are examined in the hope that they may be useful for the detection of surface slant and the results obtained are presented in this section. Input is in the form of a 64X64 matrix with two possible gray-levels, 0 (blank) and 1 (a mark). 2X2 co-occurrence matrices (see Section 2.2 for definitions) are calculated for displacement vectors $\delta = (x,y)$ where $0 \leq x \leq 10$, $0 \leq y \leq 10$.

Regular dots

Consider a surface textured with dots (with, theoretically, no size) arranged on the vertices of an (invisible) square grid. The surface is then tilted back by $\phi = \pi/4$ about the horizontal axis of the grid (which is also the picture plane horizontal). This rotated surface is then 'imaged'. Analysis of the second order statistics for this image indicates that $P_\delta(1,1)$ will have high values for δs in the horizontal direction having lengths equal to multiples of the interpoint distance. High values for $P_\delta(1,1)$ will also occur for δs in the vertical direction and having lengths equal to $\cos(\phi)$ times the interpoint distance. For δs of other directions, $P_\delta(1,1)$ will yield high values for those lengths of δ that correspond to distances between

points having the same direction of δ between them (see Figure f3.2.1). The plot of $P\delta(1,1)$ for $\delta = k(1,0)$ where k=1,2,3,...,10 peaks at δs corresponding to multiples of the interpoint distance. The plot of $P\delta(1,1)$ for $\delta = k(0,1)$ where k is as above, while somewhat muddled, gave recognizable peaks at $\cos(\phi)$ times the interpoint distance as was expected. Thus ϕ could be determined to be:

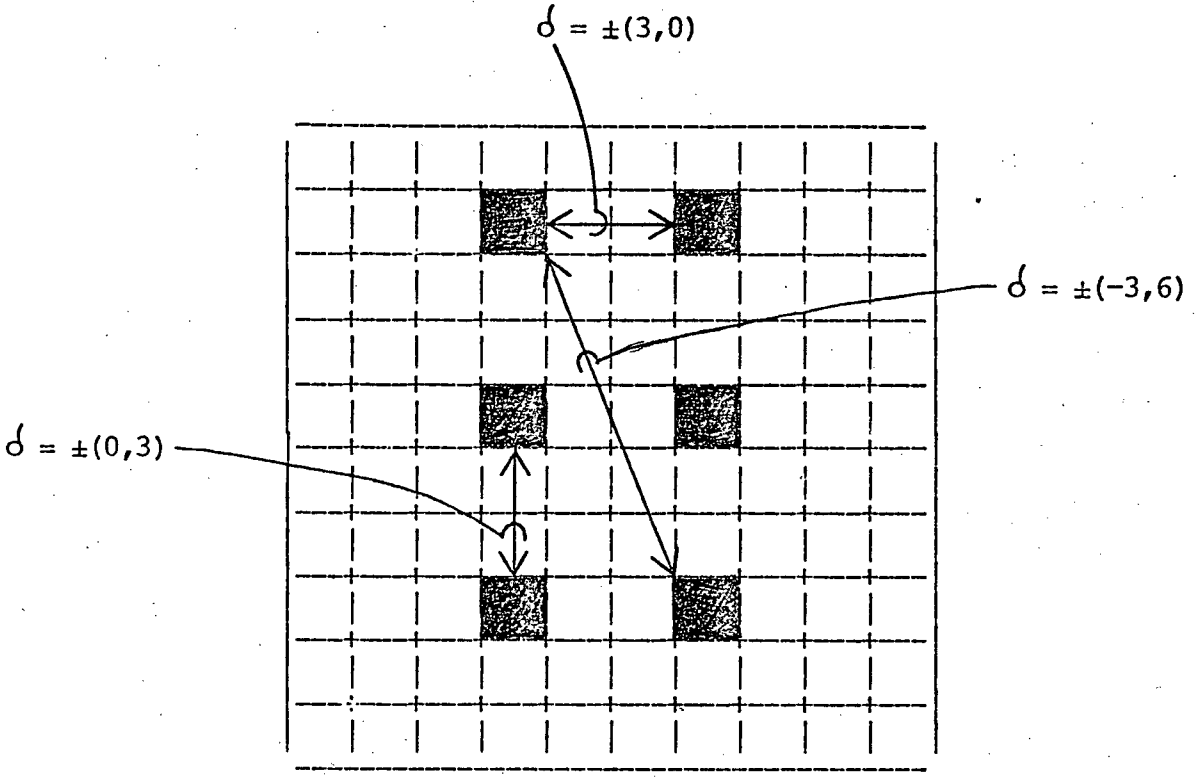
$$\phi = \cos^{-1} \frac{\text{length}(\delta) \text{ at } 1^{st} \text{ peak of } P\delta(1,1) \ \delta = k(1,0)}{\text{length}(\delta) \text{ at } 1^{st} \text{ peak of } P\delta(1,1) \ \delta = k(0,0)}$$

Note that if this technique were to be used on an image of a surface with unkown slant, we would have to be sure to "pick" the δs such that they fall on adjacent points, otherwise we could not determine the size of the squares of the grid. The problem would be alleviated if the grid were "filled in", that is, if we had a grid of lines as well as vertices. In this case we could assume that a polar plot of length(δ) for smallest peaks would give a square and we could then fit our data accordingly.

Note that we are actually looking at a grid and that "connecting the dots" as in (Schatz,1977,1978) would not change our result. In fact, the dominant characteristic here seems to be the structure of the grid, rather than the statistics such a grid presents.

Circles

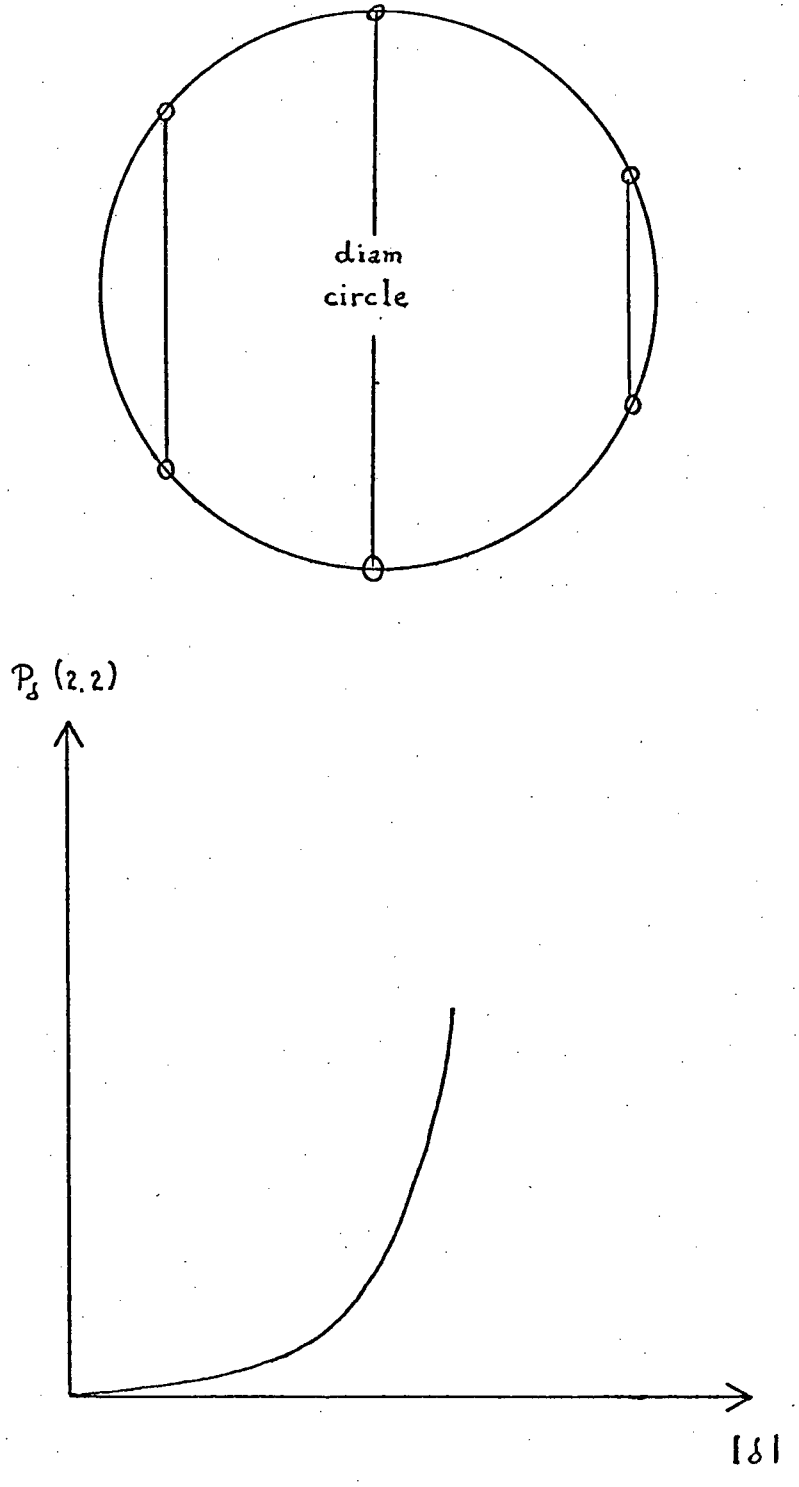
The surface here consists of rings arranged on the vertices of a square grid. The circles are small enough that no overlap occurs. Again, the surface is slanted by $\phi = \pi/4$. My collegues (actually barely willing fellow graduate students) and I had no difficulty seeing the surface slant ourselves (human perception), however, an analysis of 2nd



Examples of deltas

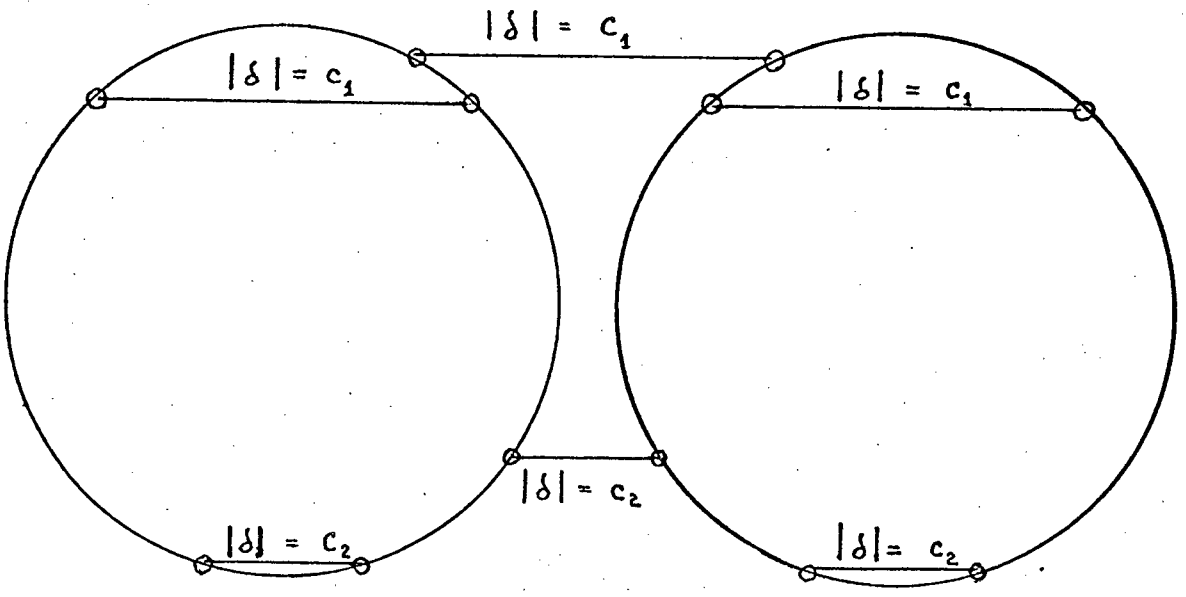
Figure f3.2.1

order statistics for the resulting digitized image proved very messy. The statistics of just a single circle are shown in Figure f3.2.2 below. If there were no other contributors to $P_{\delta}(1,1)$ when $|\delta| \leq$ diam circle and diam circle is sufficiently large for the plot of $P_{\delta}(1,1)$ to be recognizable, this method would pick out the "diameter" of the circle in any direction. Since circles are deformed into ellipses by the surface slant, we could measure the major(a) and minor(b) axis and set $\phi = \cos^{-1}(b/a)$. We could also say (correctly) that the surface is tilted about the major axis. Unfortunately, neither of our conditions are met, in general. First off, since the minimum intercircle distance is less than diam circle for most textures (and in particular for our test case) involving "circles" (e.g. Corkboard, concrete walkways), these intercircle pairs of points also contribute to $P_{\delta}(1,1)$ in the region of interest (see Figure f3.2.3 below). Digitization attributes (the top, bottom, and sides of the circles are flattened) also crept in to deform our experimental plots (see Figure f3.2.4). The implementation verified this analysis and circle diameters did not seem to be detectable without prior knowledge of the arrangement. Again we find that structural properties dominate. This makes sense to some extent since our micro-structures (circles) are highly structured objects indeed. Perhaps not so obvious is the effect of the arrangement, or macro-structure, on the second order statistics. It is the macro-structure that is muddying the plot of $P_{\delta}(1,1)$ by introducing the intercircle distances into the intracircle part of the plot. Yet it was precisely the macro-structure of the regular dot texture that allowed us to say anything at all about the surface slant for that pattern (as dots do not have a useful micro-structure). In this case at least, a structural analysis of some sort appears to be necessary.



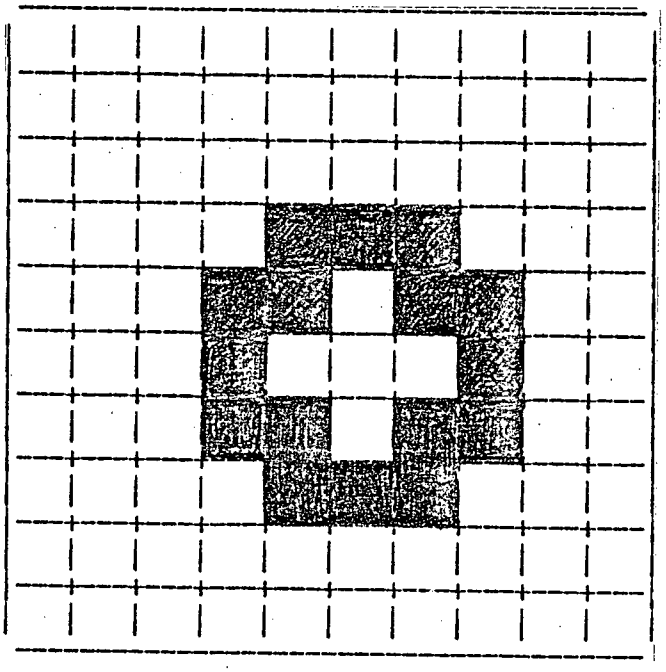
Second order statistics of a circle

Figure f3.2.2



Intercircle deltas

Figure f3.2.3



Digitization of a circle

Figure f3.2.4

Other Micro-structures

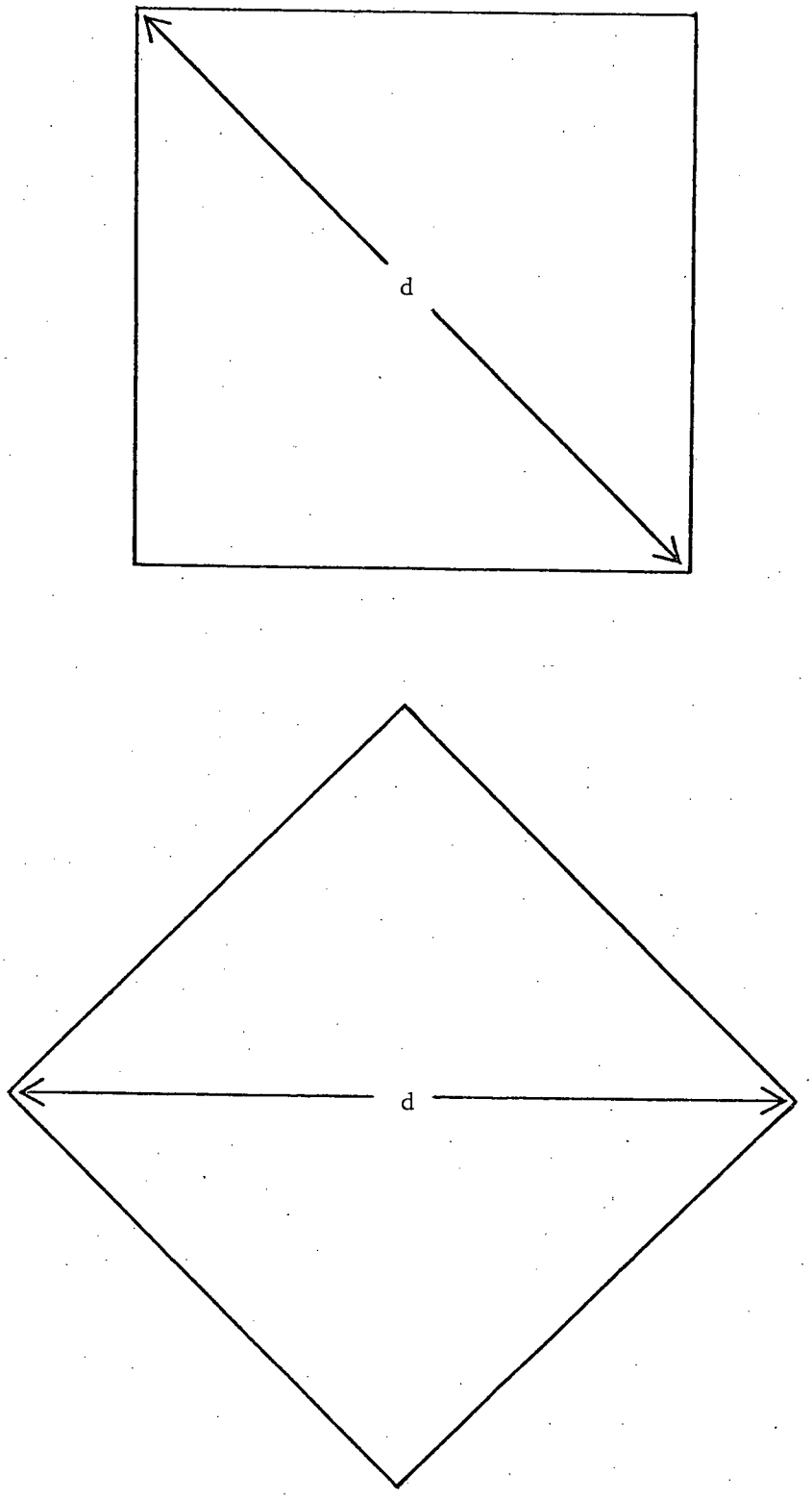
Circles were chosen in the previous example for their simplicity. They are rotationally invariant under $P_{\delta}(1,1)$ and directional measurements on them are easily described due to the simple polar equation of a circle ($r = c$).

More complex regular shapes (e.g. squares) are not rotationally invariant under $P_{\delta}(1,1)$ in that Figure f3.2.5a will yield a lower value (0) for $P_{\delta}(1,1)$ where $\delta = (d,0)$ than Figure f3.2.5b. It is now the case that inter-texel distances depend not only the macro-structure of the texture but also on the rotation of the micro-structures (which, it is true, may be construed to be a part of the macro-structure) as shown in Figure f3.2.6 below.

It is clear that second order statistics quickly become unintelligible in this case. Intuitively, however, squares, circles and other regular shapes are ideally suited for detection of surface slant as circles are deformed into ellipses and squares into parallelograms in a clearly specified way.

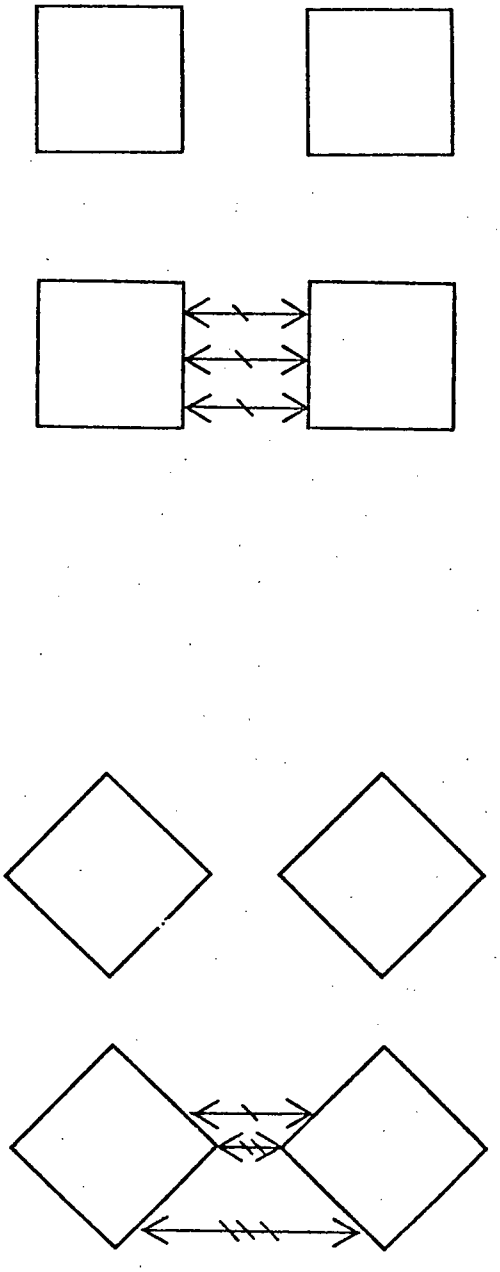
Uniform Dots

Consider, now, a surface textured with dots randomly distributed over the plane (e.g., a high-contrast stucco wall). Suppose that the surface is again tilted by an angle ϕ away from the picture plane in a direction σ. One would expect the dots to be "compressed" in the image of this surface, and this is indeed the case. But this observation is misleading. Suppose the digitized image, I, has a resolution of N*N and that the dots are now uniformly distributed in the



Effect of rotation on co-occurrence matrices

Figure f3.2.5



Effect of rotation on inter-texel distances

Figure f3.2.6

discrete case. The average density (# dark pixels / N*N) is higher in the image I than in an image of the untilted surface. But, can we use the knowledge of this compression to detect the surface slant from I? Equivalently, is there a directional dependence detectable in the statistics of the projected image? Suprisingly, the answer is "no".

Theorem 3.2.1: In an orthographic image of a surface textured by dots positioned randomly and uniformly there is no directional dependence in the co-occurence matrices.

Proof:

Consider a square section of the surface of dimensions a x a aligned with the scene axes X and Y. Suppose the probability density function (p.d.f) of the dot distribution on the surface is

$$f(x,y) = g(x)g(y)$$

where $g(s) = 1/a$ for $0 \leq s \leq a$.

In the image, the p.d.f of the dot distribution will be

$$f'(x,y) = g(x)h(y)$$

where g(x) is as above and

$$h(y) = 1/a\cos\phi \text{ for } 0 \leq y \leq a\cos\phi$$

The p.d.f. F' can be regarded as a constant, uniform p.d.f. for independently sampled points over the rectangle in the image of dimensions a x $a\cos\phi$. If $f'(x,y)$ is re-expressed in polar co-ordinates r,θ then it is clearly a function of r alone. Hence there can be no directional dependence in the co-occurence matrices.

Q.E.D.

As it seems reasonable to assume that directionality is necessary for the detection of surface slant, it is concluded that it is not

possible to determine the slant of this surface from second order statistics of the texture. It should be pointed out, however, that the amount of slant, ϕ, can be determined from the image if the parameters of the probabilty density function are known.

The theorem indicates that a plot of $P_{\phi}(1,1)$ would be flat for such an image. This has been experimentally verified using the implementation mentioned.

Conclusions

The results obtained from the implementation indicate that second order statistics on structural (in the micro sense) textures quickly become muddled and unintelligible. This is not too suprising, as it seems appropriate to use structural techniques on structural textures and statistical techniques on "statistical" textures.

However, we have seen that for truly random texture, second order statistics do not help us at all! Furthermore, second order statistics seem to apply best to regular dots. But here, the dots are arranged in a pattern, or structure. Again, structural techniques are appropriate, this time for analysis of the macro-structure of the texture. In this case the micro-structures are simply dots and their analysis need not concern us here.

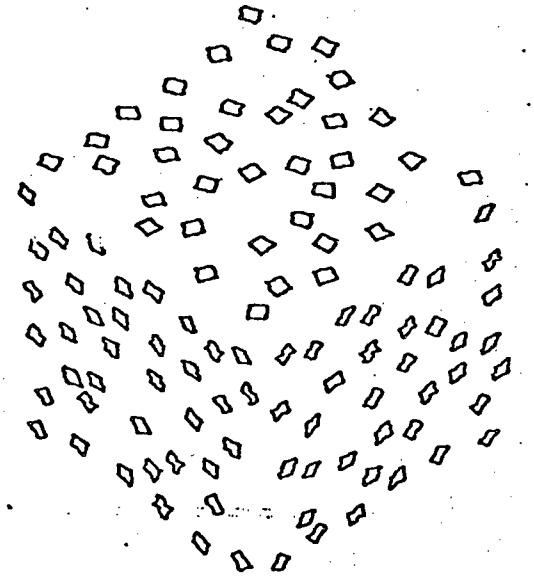
The shortcomings of statistical analysis strongly suggest the use of structural analysis, and this will be examined in the next two sections.

3.3 STRUCTURAL APPROACHES: THE MICRO-STRUCTURE

Consider a surface textured by placing copies of a given micro-structure over it. We will use the term micro-structure to refer to an abstract definition of shape and size. These are the patterns out of which to construct a texture. We will use the term texel to refer to a texture element in the image. Thus, the image of a copy of a micro-structure that has been placed on a surface is a texel. We will assume that the micro-structure is not altered in any way (other than rotation and translation) and we will not concern ourselves with the details of the macro-structure (the pattern in which the micro-structures are placed). The purpose of this section, then, is to examine the problem of surface slant detection from texel analysis.

Consider Figure f3.3.1 below. Intuitively we would say that the figure represents a cube (or corner), each surface of which has been textured with square micro-structures. Furthermore, we would partition the image into regions corresponding to the three visible surfaces of the cube. We would also say that the texels in any given region are squashed in the direction of the surface's gradient vector. We make these observations even though we do not know, a priori, that the micro-structure is a square, neither do we know its size, and we do so in the presence of a fair amount of noise (all the micro-structures would have mapped into parallelograms in a noiseless image).

Any shape-from-texture system should be just as versatile. The following algorithms attempt to embody some of the intuitive notions mentioned above. It should be mentioned, however, that they do not detect the shape of the true micro-structure. For example, for Figure f3.3.1, they do not recognize the texels as slanted squares.



"cube" textured with squares

Figure f3.3.1

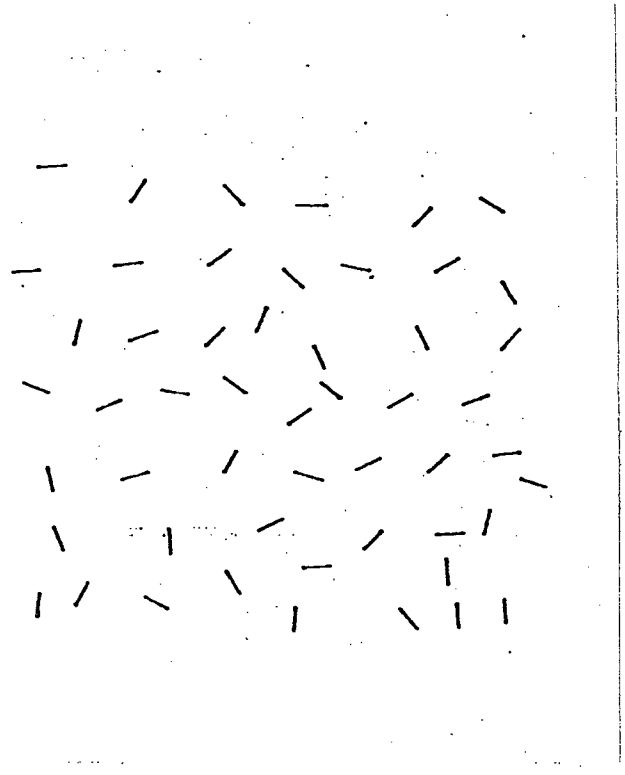
The Single Surface

Before we actually tackle an image like Figure f3.3.1, let us first look at some simpler cases and work our way up. Consider the image of Figure f3.3.2. The scene represented here is a single surface textured with randomly rotated straight line segments of a given length. This surface has been rotated (slanted) by an angle ϕ about a line at an angle σ_n from the X-axis. That is, its gradient vector is of length $\tan(\phi)$ and has direction σ. Before setting about analysing this image, let us set down a few assumptions (in addition to our usual ones). These assumptions will be retained for the duration of this section (3.3).

- -there is only one micro-structure used.
- -the surface is textured by placing copies of the micro-structure rotated to make all orientations equally likely over it. Note that no assumptions are made as to where these copies are placed. For example, the placement may be random or orderly.
- -the image (of the scene) is pre-processable to yield a list of texels (sufficiently well represented, by a list of vertex coordinates, for example).
- -the image has been so pre-processed.

One Diagonal

For each texel (line segment, in this case) let us measure its (Euclidean) length, L, and its angle of orientation, θ, from the x-axis ($-\pi/2 \leq \theta \leq \pi/2$). Let us refer to this set of (L,θ) pairs as the L-θ-list. Since each micro-structure copy is randomly rotated before being

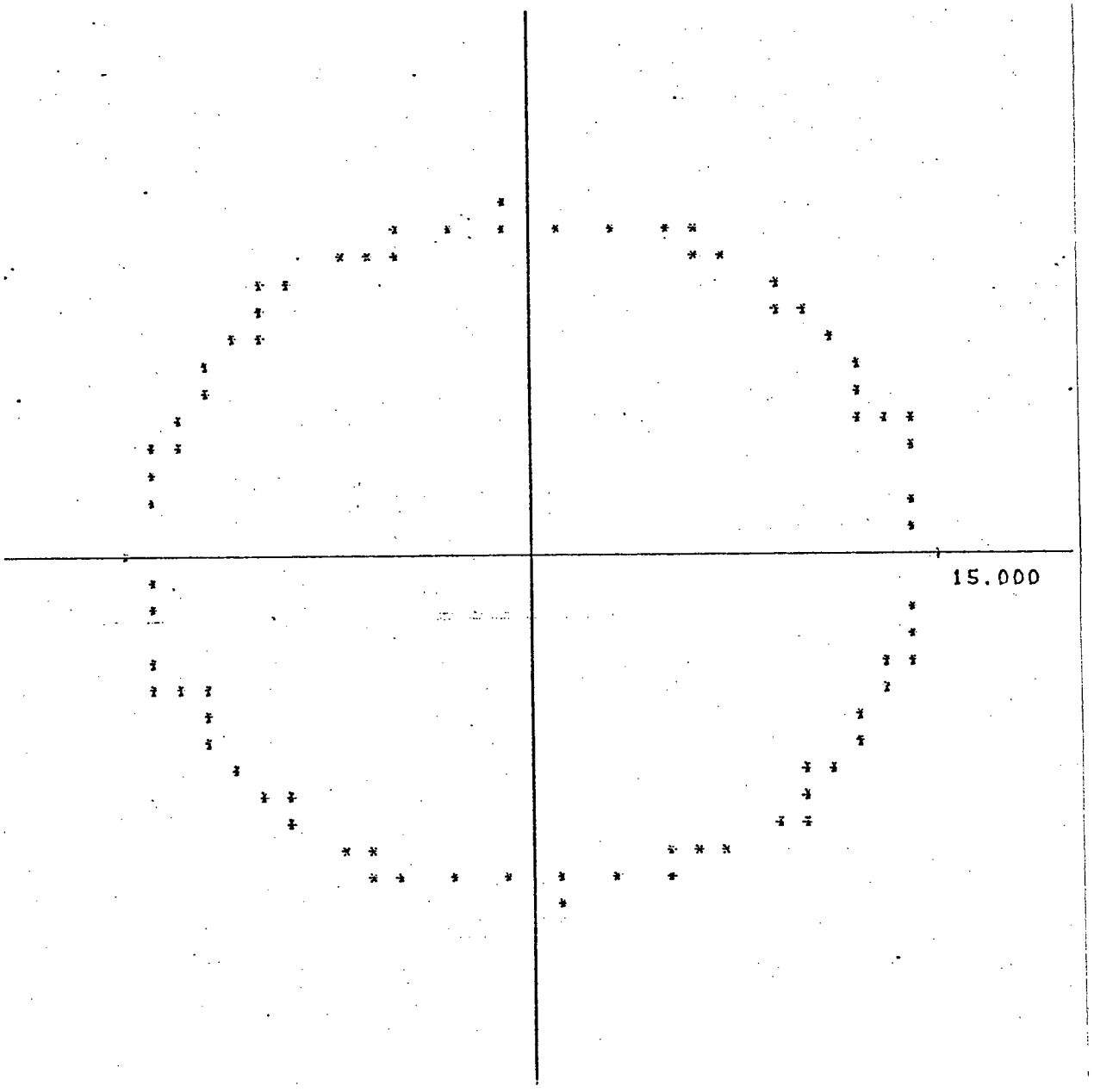


Tilted surface textured with line segments

Figure f3.3.2

placed on the surface, we can now appeal to theorem 3.1.1 to see that the measured data will lie roughly on an ellipse. The image, of course, is not a perfect orthographic projection, but it is a good approximation to one within a noise factor (caused by digitization). The data, therefore, actually represents a noisy ellipse. A polar plot of this data taken of Figure f3.3.2 is shown in Figure f3.3.3. Our task is now reduced to finding the ellipse factors a, b, and σ_n from the data, L-θ-list. Since we have determined that those line segments making an angle $\theta = \sigma_n$ with the x-axis are not foreshortened (theorem 3.1.1 again), we can reasonably estimate σ_n to be that θ corresponding to the largest L in L-θ-list. This is clearly a guess, the point may be very noisy and therefore in error, but it is hoped that it is accurate enough (and it does, after all, save us some work). A more noise resistant approach would be to treat σ as a parameter in the least-squares fitting process. We would, however, like the function that we are trying to fit the data to be linear with respect to its parameters (see the appendix to this section). This is not the case for the function that we have decided to use. To make σ_n a linear parameter then, it will be necessary to factor out σ_n and then find a similar function which does depend linearly on all three parameters, a, b, and σ_n.

Let us now think of the data, L-θ-list, as a function of θ (each pair is a mapping of some θ to some L). Now assume that the probability density function of the variation in L is a symmetric function about a mean of 0. We can now determine the factors a and b by fitting the data to an ellipse with a least-squares approximation (Conte and de Boor, 1965; pg. 241-246). The details may be found in the appendix to this section. We are now nearly done. Appealing yet again to theorem 3.1.1, we see that



Polar plot of L-θ-list taken of Figure f3.3.2

Figure f3.3.3

$$\phi = \cos^{-1}(\text{minor-axis/major-axis})$$

where minor-axis = MIN(a,b) and major-axis = MAX(a,b).

Also, σ_n has already been determined. Thus the gradient vector (see Chapter 1) of the surface can be determined to be:

$$p = -\tan\phi \sin\sigma_n \qquad \text{or} \qquad p = +\tan\phi \sin\sigma_n$$
$$q = +\tan\phi \cos\sigma_n \qquad\qquad q = -\tan\phi \cos\sigma_n$$

This algorithm would be of little interest if line-segments were the only micro-structures that we could handle. However, if our micro-structure had some consistently identifiable lines, then we could work with these lines only. That is, if lines of some given length in the micro-structure were marked in a way (coloured, highlighted, topologically unigue, etc.) that allows determination of the corresponding length in any given texel, then we could treat this line as if it were a single line-segment micro-structure. One way of ensuring that we have an identifiable line is to require our micro-structure to have a sufficiently long diagonal (s.l.d.). We shall say that d is a sufficiently long diagonal (for the image in question) if for all other diagonals d_i (we will use the term diagonal to refer to both true diagonals and edges) in the micro-structure,

$$\text{length}(d) = \text{length}(d_i)$$
$$\text{or } \text{length}(d) * \cos(\phi) > \text{length}(d_i)$$

where ϕ is the angle of tilt (with respect to the picture plane) of the surface. The idea here is to make sure that we have a diagonal that will not be foreshortened "too much". That is, we would like its image to be identifiable by its length. Note that a s.l.d. is either unique or, if it is not, the same length as all other s.l.d.s. The identifiability of a sufficiently long diagonal depends on the angle of

tilt, ϕ. Hence, a s.l.d. is a weaker concept than a uniquely identifiable length (u.i.l.) since this dependence is not required. Indeed, other usable u.i.l.s such as topologically unique lines do not suffer from this restriction. Consider the class of micro-structures that have precisely one sufficiently long diagonal. Examples of such micro-structures (for some images) appear in Figure f3.3.4 below. Note that the micro-structures need not be convex, polygonal, closed, nor even contain any line segments at all. To go any further, it is necessary to first prove the following theorem for texels containing one or more sufficiently long diagonals. This theorem gives the reason that the diagonal is detectable.

Theorem 3.3.1:
The longest diagonal of a texel corresponds to one of the sufficiently long diagonals of the micro-structure.

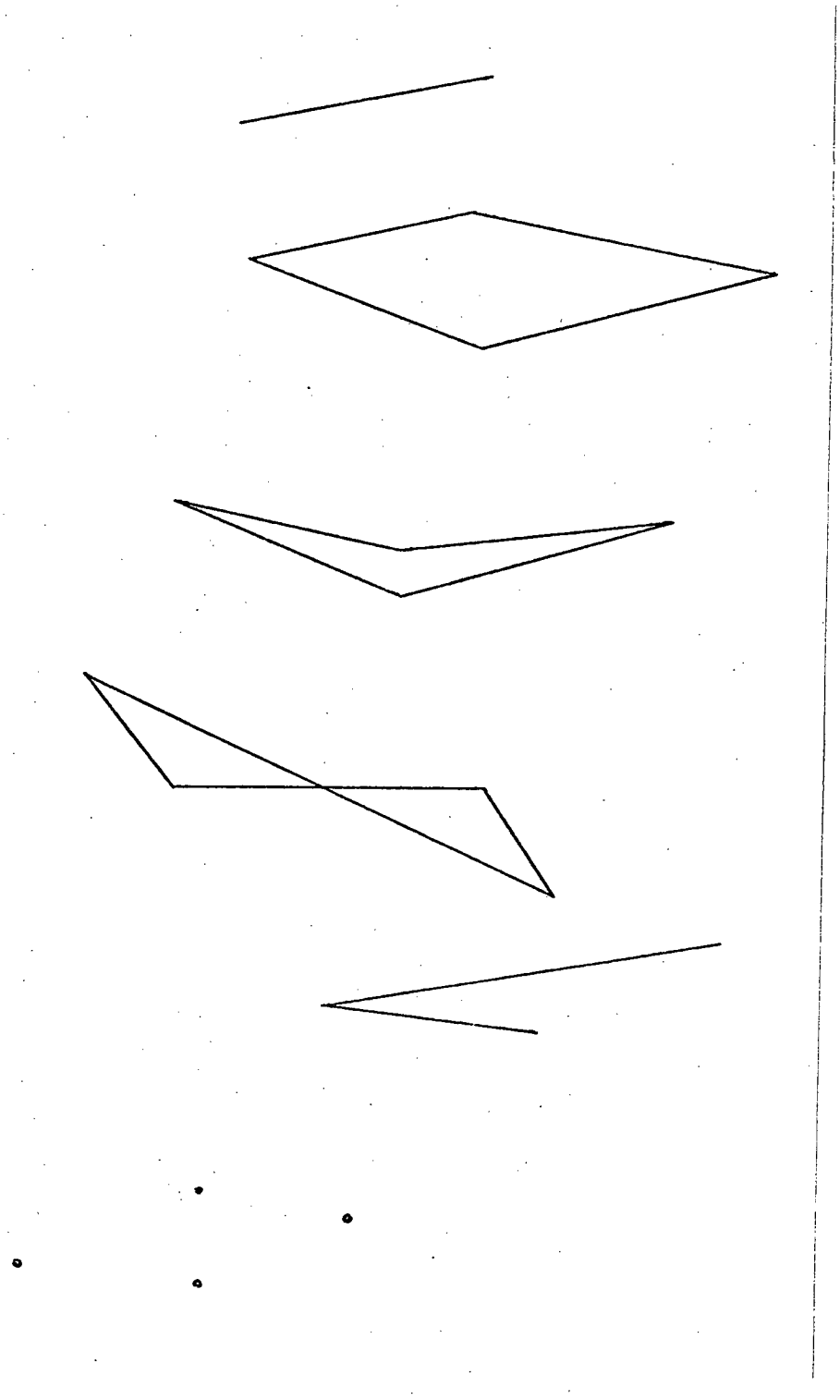
Proof:

Without loss of generality, let the surface be tilted about the X-axis by an angle ϕ (Figure f3.1.1). Then any diagonal of length L and direction θ in the micro-structure will be mapped into a diagonal of length

$$L'^2 = L^2\sin^2\theta\cos^2\phi + L^2\cos^2\theta$$
$$= L^2(\sin^2\theta\cos^2\phi + \cos^2\theta)$$

Thus L' is longest when $\theta = 0$ or $\theta = \pi$ (L'= L) and is shortest when $\theta = \pi/2$ or $\theta = 3\pi/2$ ($L' = L\cos\phi$). But since a sufficently long diagonal is such that $L\cos\phi > L_i$ for all other L_i, we see that L' at its shortest is still longer than L_i' at its longest. The theorem follows immediately.

So, for the particular case of precisely one sufficiently long diagonal, we know that the longest diagonal of each texel is the image



Examples of micro-structures with exactly one sufficiently long diagonal

Figure f3.3.4

of the sufficiently long diagonal of the micro-structure. This means that we can effectively ignore all of the points of any texel except those belonging to the longest diagonal, and in doing so, ignore all of the diagonals of the micro-structure except the sufficiently long one. Since the micro-structure is randomly rotated before being placed on the surface, so is the sufficiently long diagonal. If we then measure and record the length, L, and the direction, θ, of each texel's longest diagonal, we will have a list, L-θ-list, satisfying all of the assumptions of theorem 3.1.1. We can then extract ϕ and σ_n from L-θ-list in exactly the way we did for the single line-segment case.

To summarize briefly, an algorithm has been developed which extracts surface slant information from images of surfaces textured with

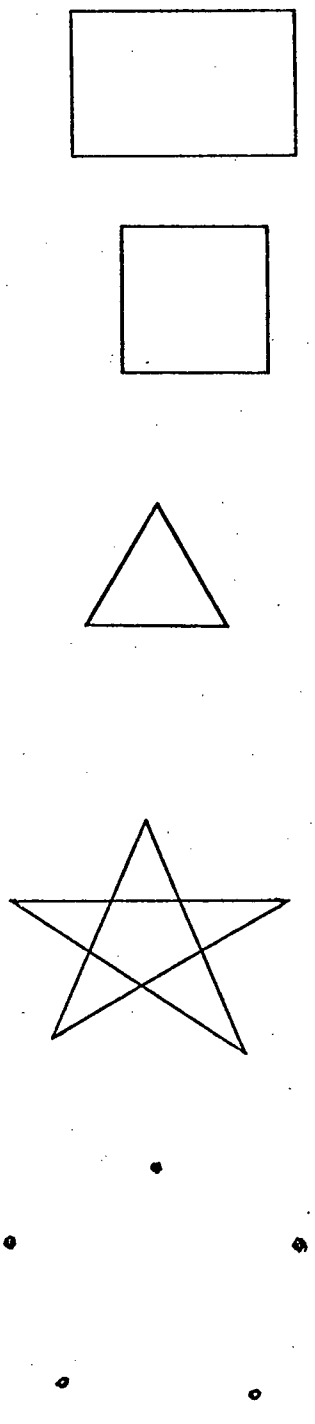
1) "Dipoles" or micro-structures consisting of just a single line-segment
2) "Thick dipoles" or micro-structures containing precisely one sufficiently long diagonal.

Furthermore, we note that 1) is just a special case of 2).

To allow treatment of a more natural class of micro-structures, we now relax our restriction on them to include micro-structures containing one or more sufficiently long diagonal(s).

One or More Diagonals

Some examples of such micro-structures are shown in Figure f3.3.5. The situation is now somewhat more difficult. The difficulty arises primarily because the longest diagonal of a texel may now be the image of any one of the sufficiently long diagonals of the micro-structure.



Examples of micro-structures containing one or more sufficiently long diagonals

Figure f3.3.5

This affects us in the following way.

Theorem 3.3.2:
Let d_1 and d_2 be two of the sufficiently long diagonals of the micro-structure. Let the angle[1] between them (in the micro-structure) be μ. If d is the longest diagonal of a texel, then the angle between d and the major axis of the ellipse is less than $\mu/2$.

Proof:
See Figure f3.3.6. Since d_1 and d_2 are distinct, $\mu < \pi$. Without loss of generality, let the surface be tilted about the x-axis. That is, the major axis of the ellipse is the X-axis. Let the angle between d_1 and the X-axis be θ_a. Also without loss of generality, let θ_a be less than or equal to $\mu/2$ (if this is not true for d_1, it will be for d_2). Then d must be the image of d_1 since the image of d_1, d_1', will be at least as long as the image of d_2, d_2'. Let θ_2 be the angle between d_2 and the X-axis. Then

$$|d_1'| = x^2 + y^2$$
$$= L^2\cos^2\theta_a + L^2\sin^2\theta_a\cos^2\phi$$
$$|d_2'| = L^2\cos^2\theta_2 + L^2\sin^2\theta_2\cos^2\phi$$

Hence $|d_2'|^2 \leq |d_1'|^2$ since $\theta_a \leq \theta_2$

Note that $|d_1'| = |d_2'|$ only when $\theta_a = \theta_2$ in which case it can be assumed that d is the image of d_1 without loss of generality.

Referring back to equation (3) of Theorem 3.3.1, we know that

$$\tan\theta = \tan\theta_a\cos\phi$$

So $\tan\theta < \tan(\mu/2)\cos\phi$ since $\theta_a \leq \mu/2 < \pi/2$

$\tan\theta \leq \tan(\mu/2)$ since $0 \leq \phi < \pi/2$

$\theta < \mu/2$ since $0 < \theta, \mu/2 < \pi/2$ Q.E.D.

[1] d_1 and d_2 are considered to be lines rather than vectors. The "angle between them" is considered to be the major angle of their intersection (unless otherwise noted).

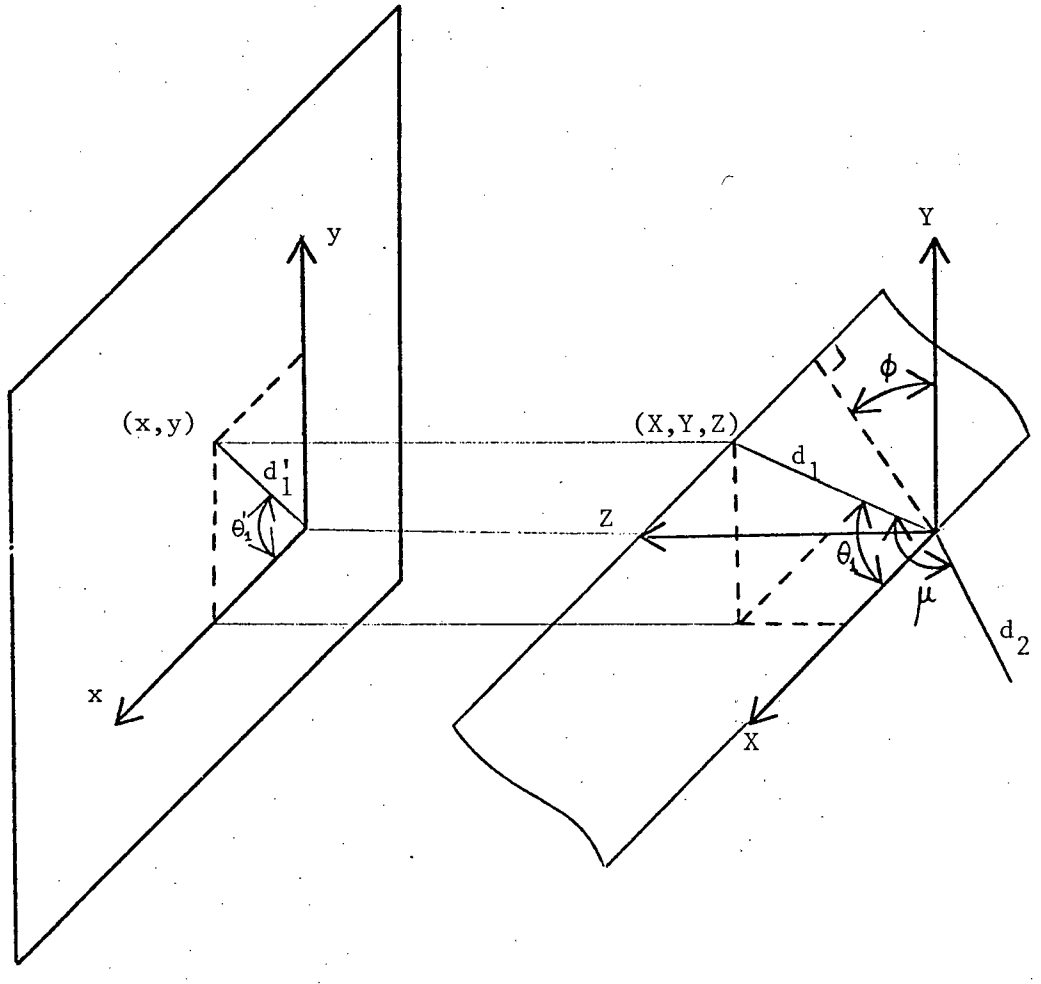


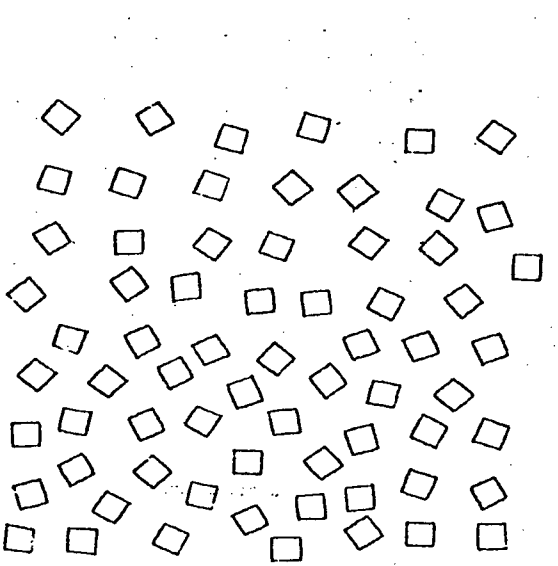
Figure for theorem f3.3.2

Figure f3.3.6

This is easily generalized to more than two diagonals. Note also that $\mu/2$ is not the tightest bound, but it is our purpose merely to show that such a restriction exists.

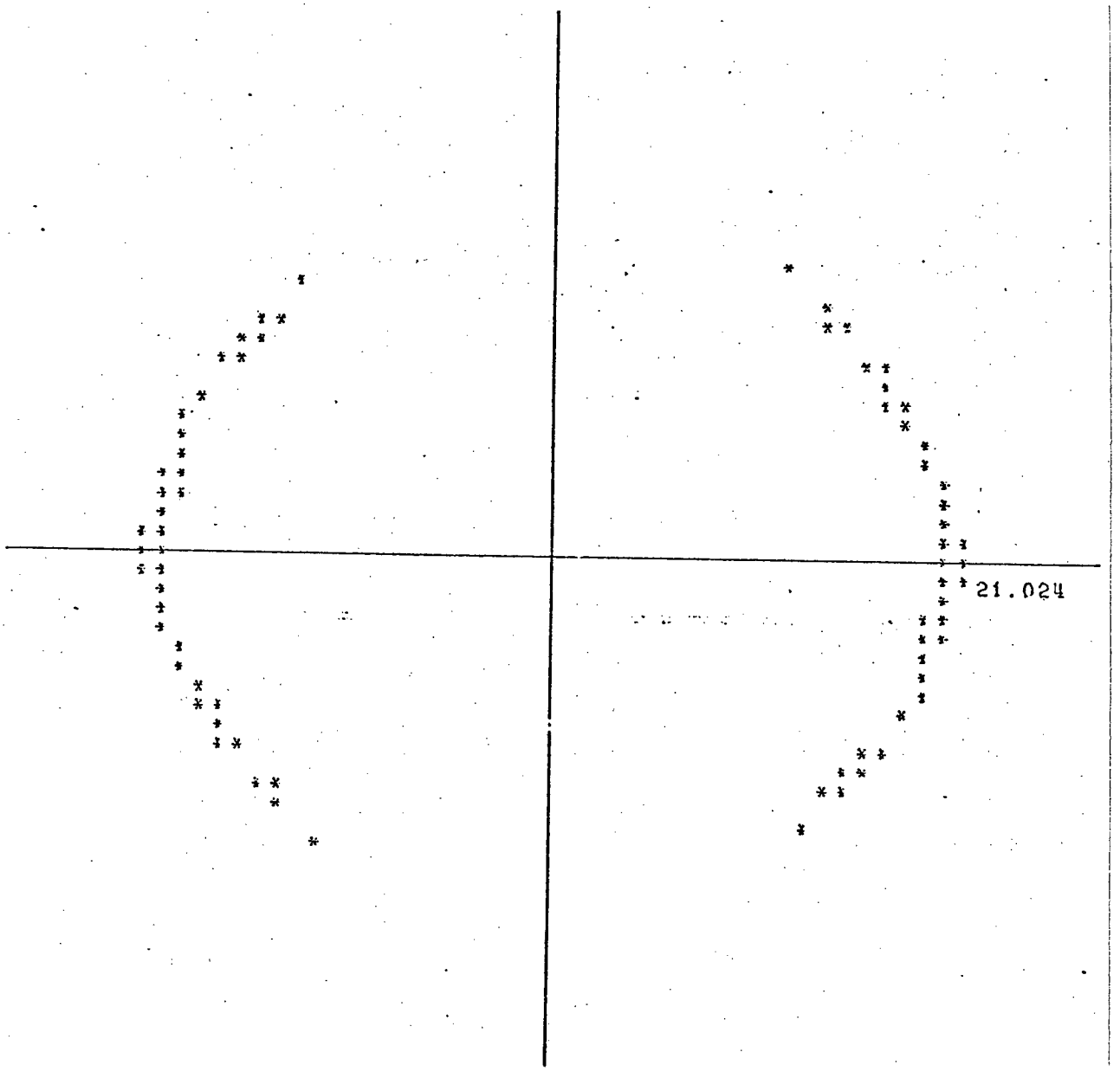
From theorems 3.3.1 and 3.1.1 we know that all of the data points in L-θ-list still lie on an ellipse, but theorem 3.3.2 shows that not all of the ellipse will be fairly represented. In fact, there may be no points at all at the minor axis of the ellipse.

Figure f3.3.7 shows an image of a surface textured with squares (which have two sufficiently long diagonals). A polar plot of the L-θ-list information for this image is shown in Figure f3.3.8 illustrating the theorem. We will, however, ignore this gap in our data and hope that enough data are available to accurately fit an ellipse to what we do have. It should now be clear why we have decided to use a least-squares, trigonometric approximation, which is computationally a bit expensive (especially when using LISP), rather than a computationally inexpensive smoothing of data by averaging, for example. The least-squares method makes statistical assumptions only about the noise. A smoothing-by-averaging method, on the other hand, also assumes that the probability of any given point on the ellipse being represented is equal to that of any other point. Theorem 3.3.2 shows us that this is not a wise assumption to make. It should be noted that as the number of sufficiently long diagonals of the micro-structure increases, the amount of ellipse represented by the L-θ-list decreases. This leads to the paradoxical case of the micro-structure with infinitely many diagonals, a circle. For this case, the algorithm falls apart completely. The only point in the L-θ-list for this image is the point at the end of the



Tilted surface textured with squares

Figure f3.3.7



Polar-plot of L-θ-list taken of Figure f3.3.7

Figure f3.3.8

major axis in the range ($-\pi/2$ to $\pi/2$) and yet we can easily determine the surface slant simply by fitting an ellipse to any given texel (which will be an ellipse) and analysing this fit in the same way as we have been doing. The algorithm can be made to handle this case simply by doing exactly that when the L-θ-list represents only one point on the ellipse. Nevertheless, it is unpleasant to have to deal with a special case in this manner.

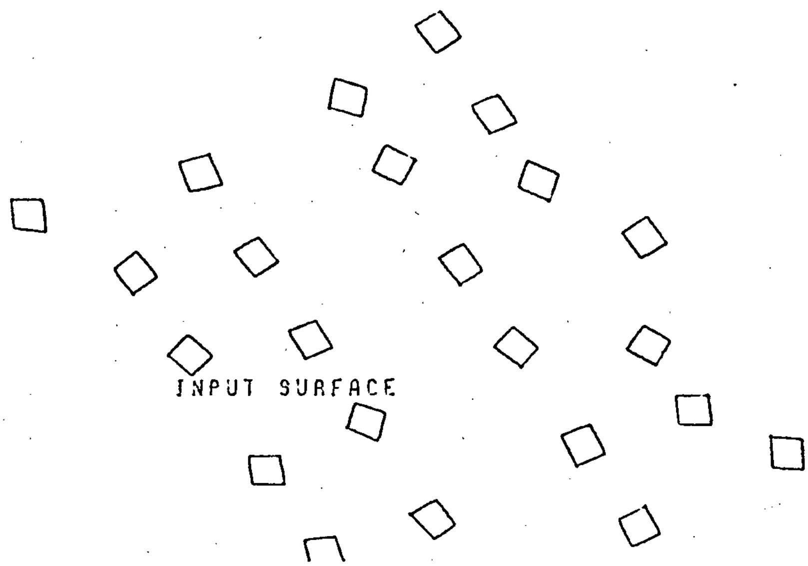
Single surface results

Pages 71-72 show two sample runs of the algorithm on some representative synthesized images. Given also are the actual values of ϕ and σ_n for comparison.

More than One Surface

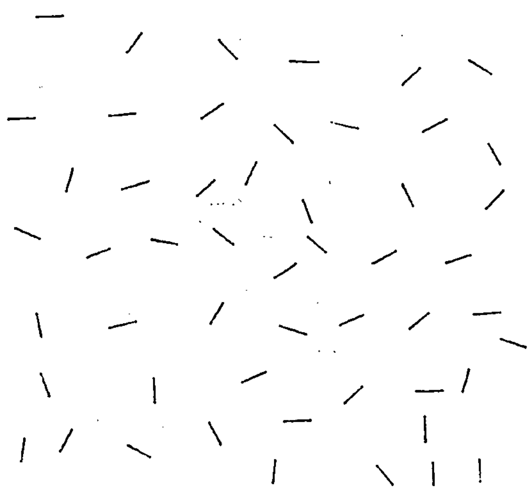
In the beginning of this section, it was hinted that several surfaces (actually the image regions corresponding to these surfaces) could be segmented on the basis of their apparent texture. As was mentioned in Chapter 1, such segmentation is assumed. Nevertheless, to see if it is possible, this problem is examined for this texture class. Since the L-θ-list data of any given region corresponding to a surface lie on an ellipse unique to the surface, we may reasonably suggest that this data may be used for the region detection task. Figure f3.3.9 shows an image of a cube made up of surfaces like Figure f3.3.2. A polar plot of the L-θ-list data for this image is shown in Figure f3.3.10 which can be seen as three superimposed ellipses. For this image, it is conceivable to detect the ellipses in its L-θ-list. One way to detect these ellipses is with a Hough transform (Hough,1962) that

```
* (DISKIN grad)
>    30 FORMS READ FROM: grad
>
* (gradient (image>
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
```



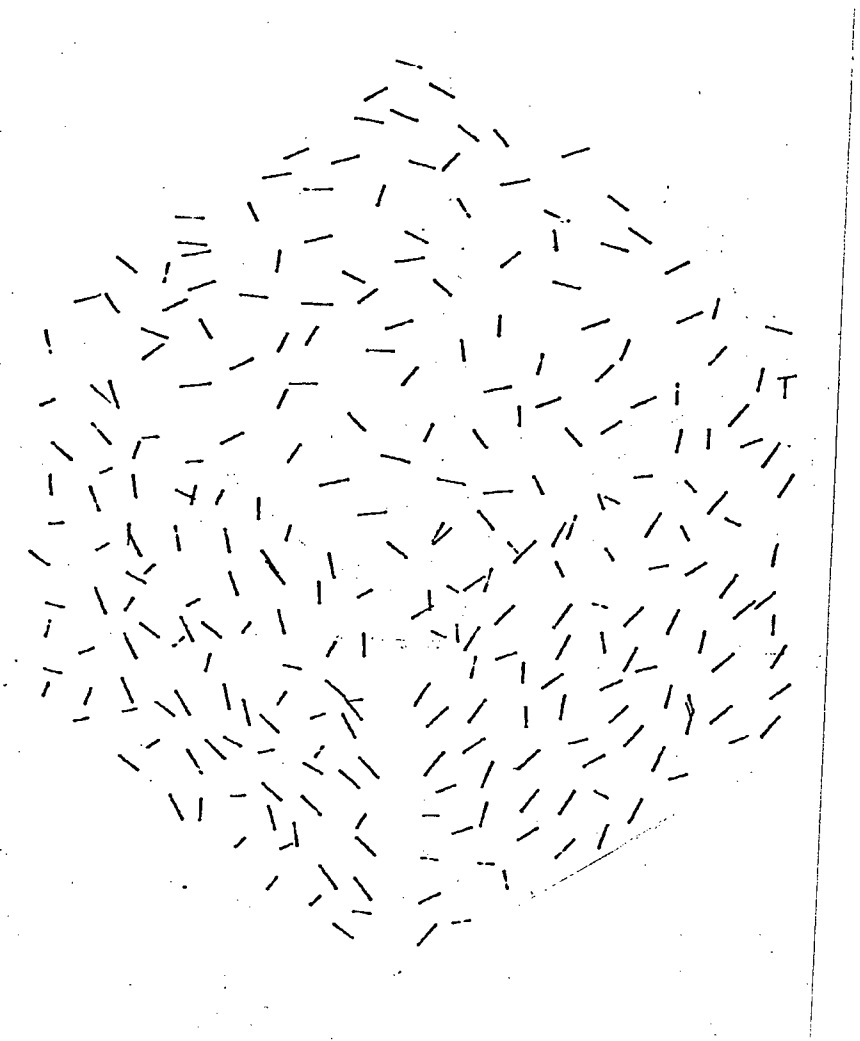

```
>    Input surface ok?
* T
>    This surface is tilted by an angle phi = ±0.474918
>    about a line an angle sigma_n = -.444419 from the X-axis
>    Therefore, its gradient vector is given by
>    p = -0.221058      or     p = 0.221058
>    q = 0.464222              q = 0.464222
> The actual value of sigma_n = -.523598
> The actual value of phi = ±.523598
*
*
```

```
* (gradient (image))
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
*
```



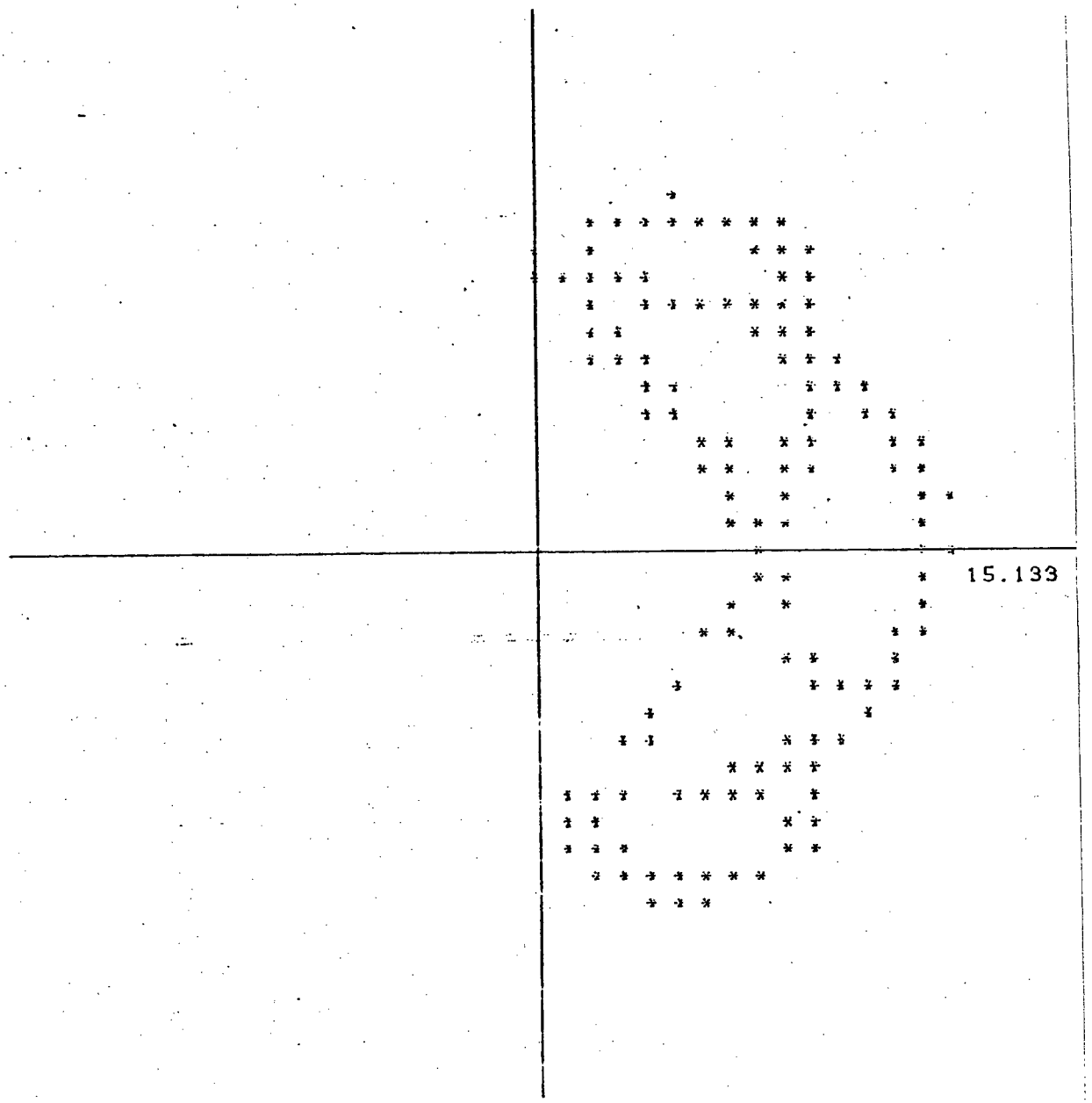
Input surface

```
>    Input surface ok?
* T
>    This surface is tilted by an angle phi = ±0.767956
>    about a line an angle sigma_n = 0.571337 from the X-axis
>    Therefore, its gradient vector is given by
>    p = 0.522215        or        p = -0.522215
>    q = 0.812335                  q = -0.812335
> The actual value of sigma_n =  0.523598
> The actual value of phi =  ±0.785398
*
*
```



"cube" textured with line segments

Figure f3.3.9



Polar of the L-θ-list taken of Figure f3.3.9

Figure f3.3.10

maps ellipses into points in some transform space. One need then simply look for accumulation points in the transform space. However, two problems immediately become apparent.

The first is one of computational complexity. If we assume that we have no information about the ellipses, except that they all have their origins at the co-ordinate origin, then the transform space must necessarily be three-dimensional (since there are three remaining parameters).

The second problem is much more fatal and persistent. If the ellipse were noiseless, this Hough transform technique would work well. As it is, we have found, experimentally, that the transform is very sensitive to noise, and gives unexpected results.

This second problem persists even if we try to eliminate the first. We can do this by noting that, since the same micro-structure is used over the entire image, the length of the major axis will be the same (and equal to the length of the micro-structure's sufficiently long diagonal) for each ellipse. We can estimate this length to be the maximum L value found in the L-θ-list. Hence, if the ellipses are parameterized in terms of the lengths of the major- and minor-axis and the rotation of the major-axis, σ_n, a two-dimensional transform space can now be used. Let us further assume that σ has somehow been correctly determined as well. The transform space is now one-dimensional and, if all were well in the world, it would now be a simple histogram problem.

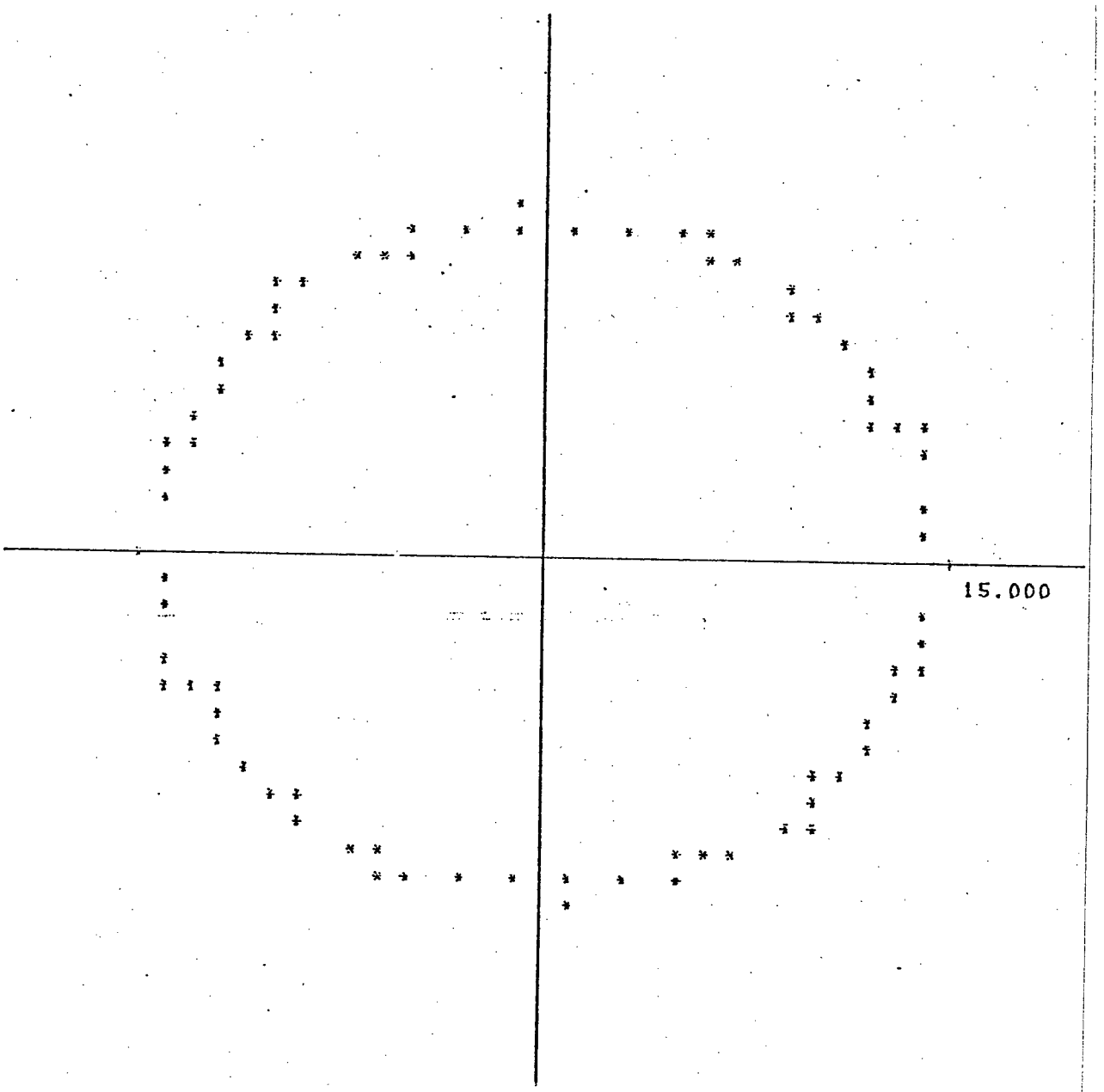
Nevertheless, it is easy to see why the second problem is still present by looking at the following analysis. The transform will

consist of evaluating the minor-axis length for each σ_n. Let us assume that we only have one ellipse, then let us see if we can detect it. Consider those points near the ends of the major axis. Note that the lengths of all of these will be less than the length of the major axis (as they should be) but due to the selection of the major axis length and the noise in the points themselves, some lengths will be much smaller. It can be seen that, for these points, very small values of the minor-axis will be obtained. Note that these are the correct values for these points; they do indeed lie on very eccentric ellipses and not on the ellipse we would like to detect. Figure f3.3.11 illustrates this. We could try to correct this problem by finding a better estimate for the major-axis length and by smoothing the data (we will have to be careful that the smoothing algorithm does not try to smooth the ellipse into a circle as averaging smoothers will). We did not have much success with any of these solutions, however, so we decided that another approach is clearly in order.

"We all have our moments"anonymous

Fortunately, texture discrimination is an old problem and much work has already been done on it (see Chapter 2). As was mentioned in the last chapter, however, most of the work has been concerned with distinguishing different textures on the same surface (such as wheat fields from grass fields in an aerial image). The task considered in this thesis is better expressed as distinguishing different surfaces that have the same (actual) texture on them.

Before proceeding, let us re-emphasize that, although we are using the term "surface", we are actually practicing region detection at this



Ellipse detected for noisy points

Figure f3.3.11

stage of the analysis. Since the surface is part of the scene (and not part of the image) it will be maintained that a surface has not been detected until its three-dimensional attributes (such as surface slant) can be described.

Several researchers (Tomita et al, 1973) (Wang et al, 1979) have found that the second order moments of inertia of the texels form a good feature space for surface detection of this kind. We can see why this should be so by considering the case where the micro-structures are all circles. Then the texels will all be ellipses whose major axis is in the direction of the line that the surface has been tilted about and whose minor axis is in the direction of the surface's gradient vector. In such a case, the minimum angular inertia of all the texels on a given region will be that taken about the major axis. The domain of micro-structures for this section may be thought of as polygons approximating circles. So, rather than getting unique points in the moment space for each class of texels on a region, slightly scattered clusters in the space are obtained. If a texel is described as a list of points $((x_1,y_1)\ (x_2,y_2)\ \ldots\ (x_n,y_n))$ which are the co-ordinates of the vertices within the texel, then the moments of these vertices are given by:

$$I_{xx} = (1/n) \sum_{i=1}^{n} (y_i - y_m)^2$$

$$I_{yy} = (1/n) \sum_{i=1}^{n} (x_i - x_m)^2$$

$$I_{xy} = (1/n) \sum_{i=1}^{n} (x_i - x_m)(y_i - y_m)$$

where (x_m,y_m) are the co-ordinates of the vertices' centre of mass. For a more general definition see (Wang et al,1979). In Chapter 2, in the

section on structural analysis, we referred to the ij^{th} order moments, M_{ij}. These are given by

$$M_{ij} = 1/n \sum_{k=1}^{n} (x_k - x_m)^i (y_k y_m)^j$$

Thus $M_{11} = I_{xy}$

$M_{02} = I_{xx}$

$M_{20} = I_{yy}$

So, assume that we these moments have been computed for every texel in the image. The task now is to pick out the regions in the image corresponding to the surfaces. Rather than constructing one-dimensional histograms of each moment (Tomita et al,1973), or computing the eccentricity and direction of each texel (Wang et al, 1979), let us simply view the set of computed moments as points in the three-space, I_{xx} X I_{yy} X I_{xy}, and then look for clusters in this space.

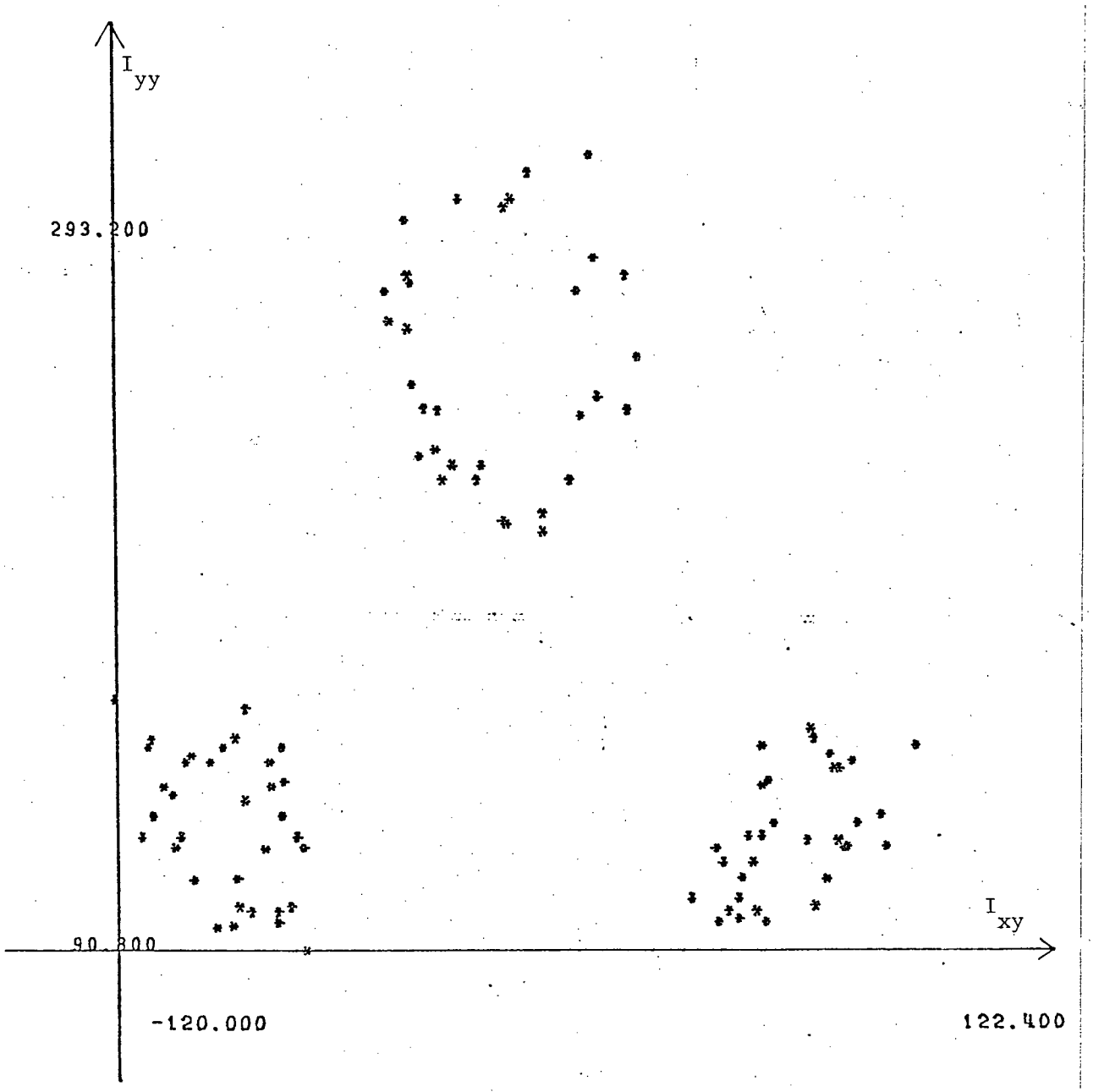
Note that, in so doing, we are making all of the usual clustering assumptions. One of these is that the feature space will indeed produce clusters separated by some threshold. In this case, the metric (distance function) is simply Euclidean distance, and the two works previously referenced give us reason to believe that the feature space is useful. We are also relying heavily on our previous assumptions and our notion of texture. For example, the scene could conceivably be composed of as many surfaces as texels with one texel to a surface. This case clearly contradicts the definition of texture outlined in Chapters 1 and 2. And of course, the assumption is still made that only one micro-structure is used. Note that if the scene consists of two distinct (in space) but identically slanted surfaces, the clustering algorithm will only see one surface. To ensure that segmentation occurs

in a meaningful way, it will be necessary to go back to the image and verify that the detected cluster indeed defines a single region. If not, further processing will have to take place. The discussion of these problems is not within the scope of this paper and this implementation does not tackle them. The problem does exist, however, and any complete scene analysis system will need to apply itself to it.

Keeping in mind that these assumptions have been made, it is now possible to set about implementing a cluster analysis algorithm. Any algorithm, such as one of those described in (Duda and Hart, 1973) or (Tou and Gonzalesz, 1974), would do here, since we have found the clusters to be widely separated. In the implementation, a simple algorithm known as the maximin-distance algorithm (Tou and Gonzalesz, 1974) is used. This served well for purposes of demonstration. For a large scene analysis system, however, a more dependable graph-theoretic algorithm, such as one of those described in (Zahn, 1971) is recommended.

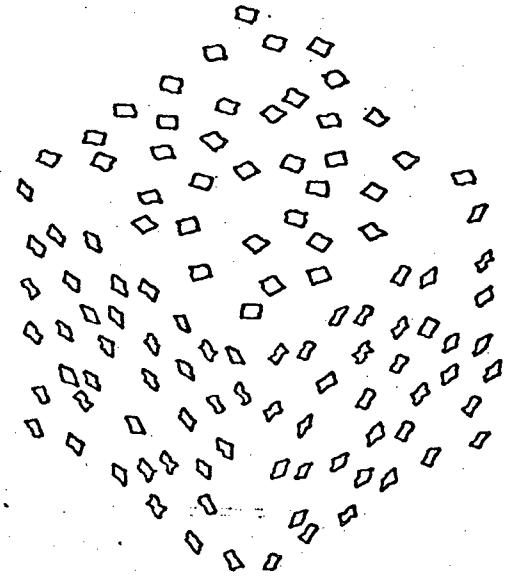
Figure f3.3.12 shows a plot of I_{yy} versus I_{xy} taken of Figure f3.3.1 showing the good separation obtained even for this two-dimensional case. An example of the algorithm at work is shown in Figure f3.3.13.

Now that the image is divided up into regions, it is necessary only to determine the slant of each surface. This can be done one surface at a time since the techniques for handling the single surface case have already been described. An example of the entire process at work is shown on pages 86 to 89.



I_{yy} v.s. I_{xy} taken of Figure f3.3.1

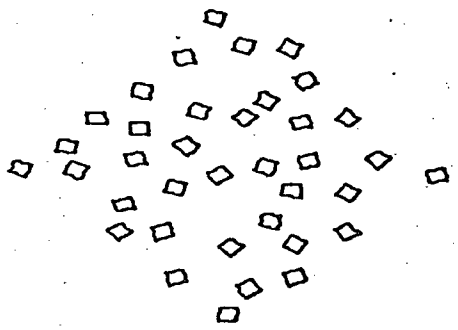
Figure f3.3.12



THIS IS THE INPUT IMAGE

Application of the clustering algorithm

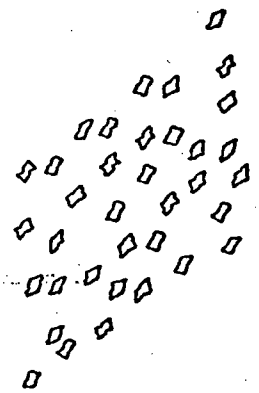
Figure f3.3.13



DETECTED SURFACE

Application of the clustering algorithm

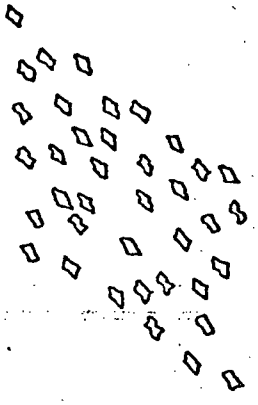
Figure f3.3.13



DETECTED SURFACE

Application of the clustering algorithm

Figure f3.3.13



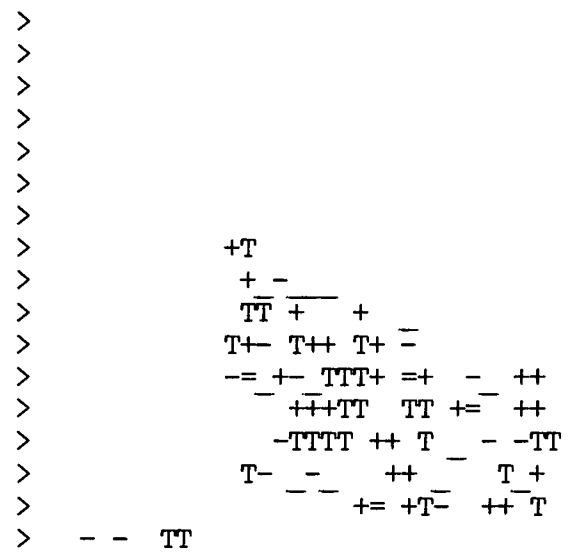
DETECTED SURFACE

Application of the clustering algorithm

Figure f3.3.13

```
* (grad i)
*
*
*
*                              +-
*                           T_+    +   +
*                        ++    +   _=- T-
*                   + +++  ++  __T  -    -+  -
*                 =- TTT-  +-+      -   =+ T T   ---  +-=
*               +T        T=_   T       +--+  -        T ++
*                + -       -  -- T    _-_ T+T__  - -      TT_
*                TT +    +         ++           +=_       _T+ T   T
*               T+- T++ T+ -          +-+       ++ +  +T++   T
*               -=_+-_TTT+ =+   -_  ++ T + + TT+= ++++ T+=-T
*                   +++TT   TT +=   ++   T T += T+ ++ + +T ++
*                  -TTTT ++ T _  - -TT_=- +T  ++  +TT  --
*                T-  _-     ++      T +     +T ++ T   |  ++ -
*                        += +T-   ++ T   ++-   +T++ +T TT
*                           - -  TT        +T+ T+
*                                T+ __    + T+T
*                                         T
*                  This is the input image
*
>   This is the input image ok?
* T
```

```
* FREE SPACE EXPAND
*
*
*
*                                 +-  _  _
*                             T_+   -+  -+_
*                          ++   _+  __=- T-  _
*                   + +++  ++ __T  -    -+  -
*               _=- TTT-  +--+ _  _-  _=+ T T   ---   +-=
*                        T=    T        +--+  -        T
*                       -_ --_T   _-  T+T__ - -
*                              ++        +=_
*                                 +-+-.
*                                 T T
*
*
*
*
*
*
*
*                          Detected surface
*
>    Detected surface ok?
* T
>    This surface is tilted by an angle phi= ±0.767009
>    about a line an angle sigma_n = -3.964055E-3 from the X-axis
>    Therefore, its gradient vector is given by
>    p = 3.820867E-3      or     p = -3.820867E-3
>    q = 0.963873                q = -0.963873
>
>
>    actual value of phi = ±π/4 ≈ ±0.785398
>
>    actual value of sigma_n  = 0
```



Detected surface

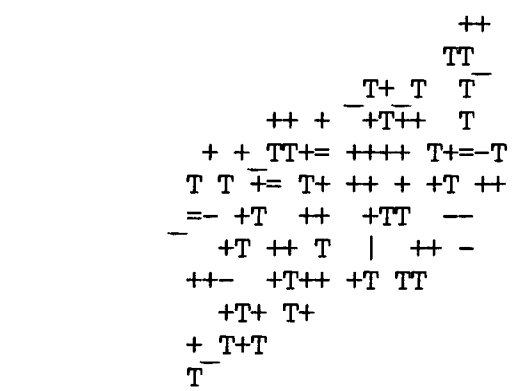
Detected surface ok?

* T

This surface is tilted by an angle phi= ±1.021549
about a line an angle theta0 = -0.94459 from the X-axis
Therefore, its gradient vector is given by
p = -1.323798 or p = 1.323798
q = 0.95753 q = -0.95753

actual value of phi = ±π/3 ≈ ±1.047197

actual value of sigma_n = $-\tan^{-1}(\mathrm{SQRT}(2))$ ≈ -0.955316



Detected surface

Detected surface ok?

* T

This surface is tilted by an angle phi= ±0.923183
about a line an angle sigma_n = 0.868427 from the X-axis
Therefore, its gradient vector is given by
p = 1.00908 or p = -1.00908
q = 0.854031 q = -0.854031

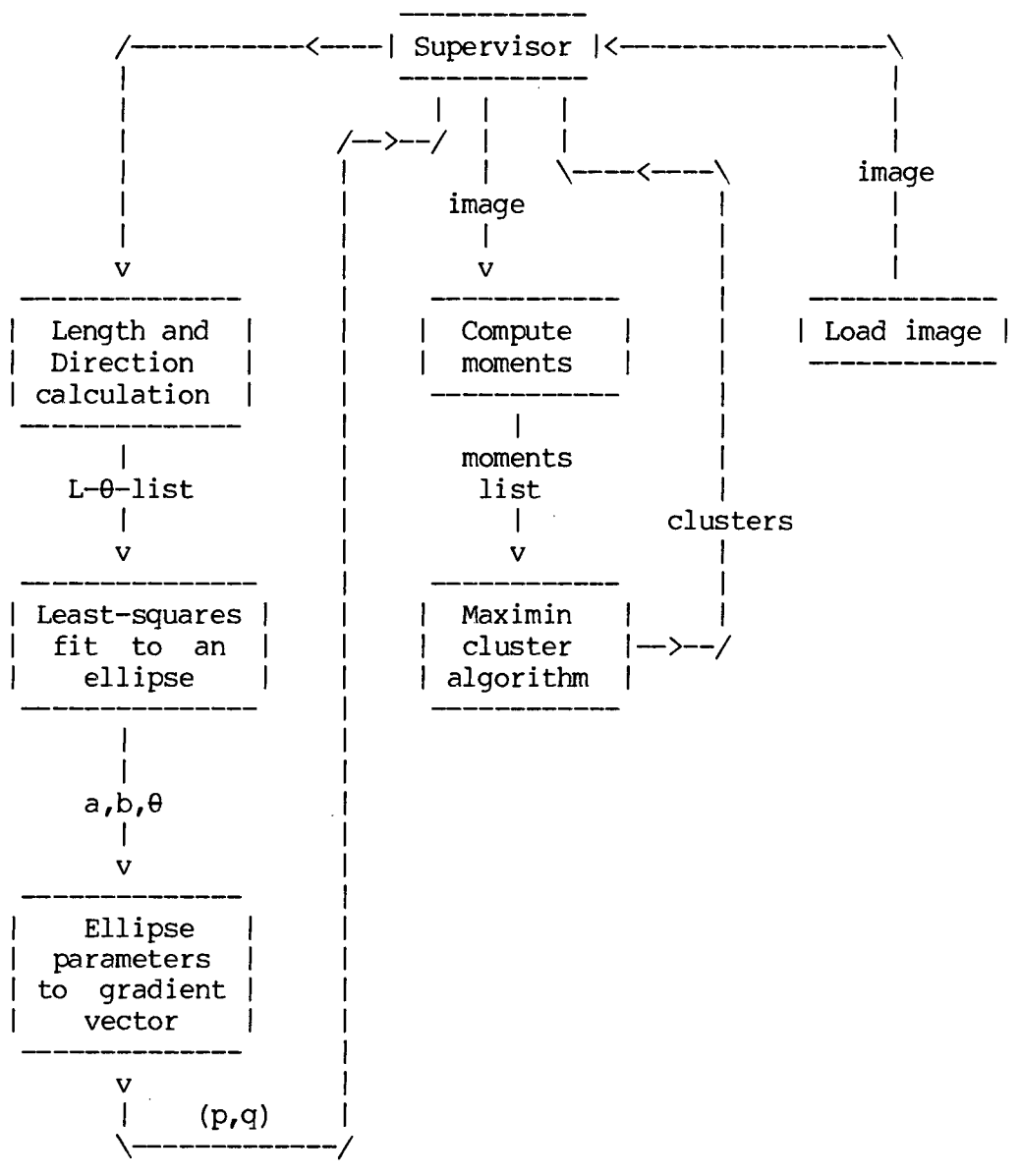
actual value of phi = ±π/3 ≈ ±1.047197
actual value of sigma_n = $\tan^{-1}$(SQRT(2)) ≈ 0.94459

Details of the implementation

The process of detecting surface slant from texture is composed of three major steps in this program: load the image, segment the image into regions corresponding to surfaces, and finally compute the slant information for each surface. The "load image" module is responsible for the pre-processing of the input image. The entire process is illustrated in Figure f3.3.14. The purpose of the supervisor is two-fold: it must maintain the flow of control and it must transform data structures to allow communication between modules.

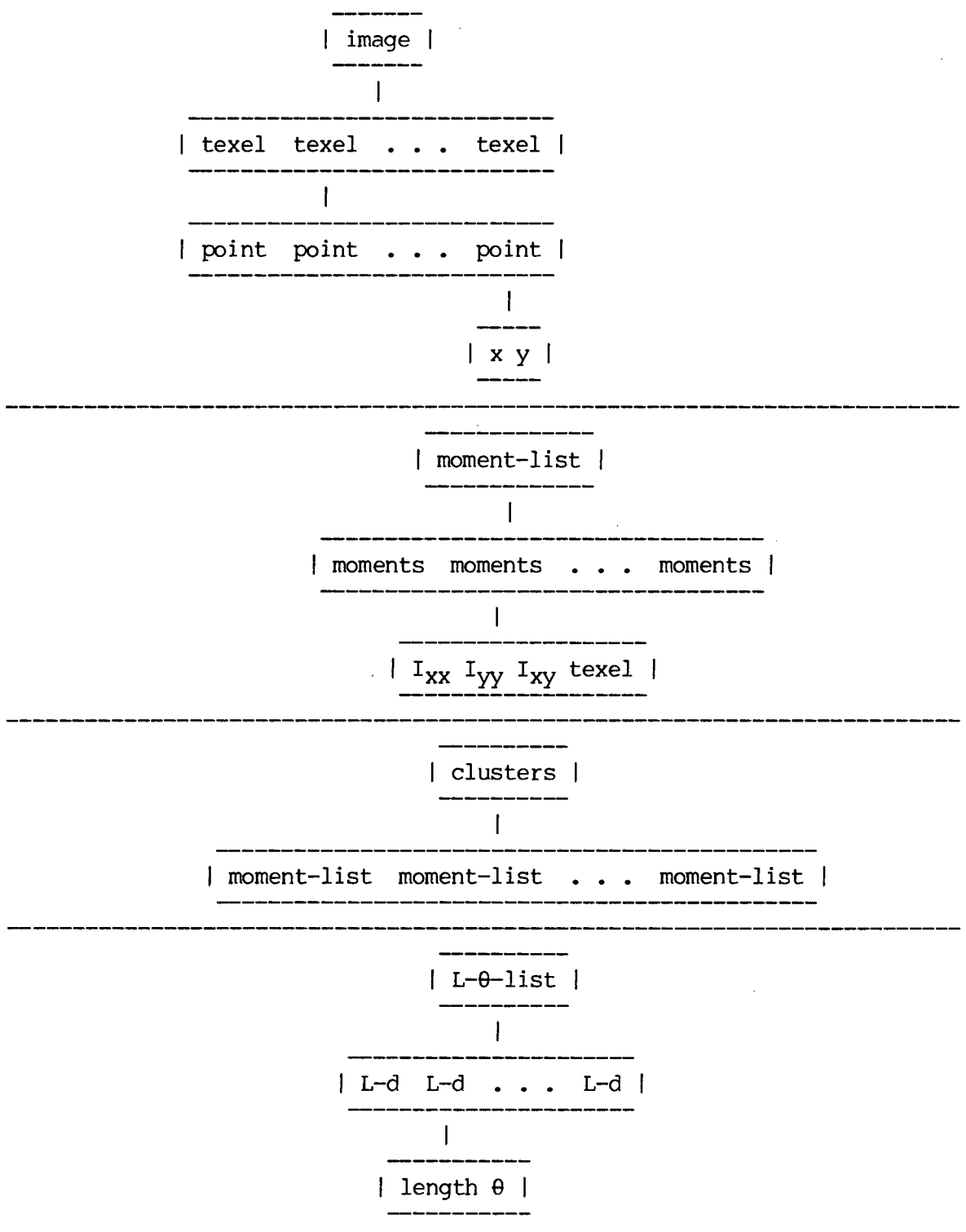
Flow of control is simple. The image is first loaded. This information is then passed to the region detector. The supervisor then passes one region at a time to the gradient vector calculation routines.

The real need for the supervisor is ensuring that each module receives the necessary information. The data structures used are shown in Figure f3.3.15. From the flow of control diagram, we see that the maximin cluster algorithm, for example, requires a list of moments as input data. Output would consist of a list of clusters, each cluster being a list of moments. However, a list of clustered moments is of no help in computing the gradient vectors since they are not used nor even needed for this task. Therefore, the fourth element of the "moments" data structure is a "texel" data structure. This element is simply carried along (but otherwise ignored) by the maximin cluster algorithm. The supervisor then "extracts" the texels from each cluster to form a region.



Flow of Control of the Micro-structural Analysis Algorithm

Figure f3.3.14



Data Structures for the Micro-structural Analysis Algorithm

Figure f3.3.15

Conclusions

We have seen that surface slant can be computed with a structural analysis of the texels in an image of the surface. This is found to be possible under the assumptions that the surfaces in the scene are planar, that exactly one micro-structure is used, that the micro-structure is randomly rotated before being placed on the surface, and that the (orthographic) image is pre-processable to yield a list of texels. The title of the section is "Structural Approaches", but perhaps the algorithm should be examined a little more closely to see just why it should be considered structural.

While it would be incorrect to say that the technique is statistical (probabilities or other statistical values are not calculated at any time), it would be correct to point out that it relies on statistical assumptions. In particular, the findings of this section agree with Schatz' (Schatz,1977) restricted version of Julesz' (Julesz, 1973) hypothesis concerning texture discrimination. That is, the data, L-θ-list, that is used to calculate the surface's gradient vector corresponds to the second order statistics used by Julesz and Schatz. We also restricted ourselves to those points connected by edges or diagonals (lines and virtual lines (Schatz,1977)). For details of the work by Julesz and by Schatz see Chapter 2.

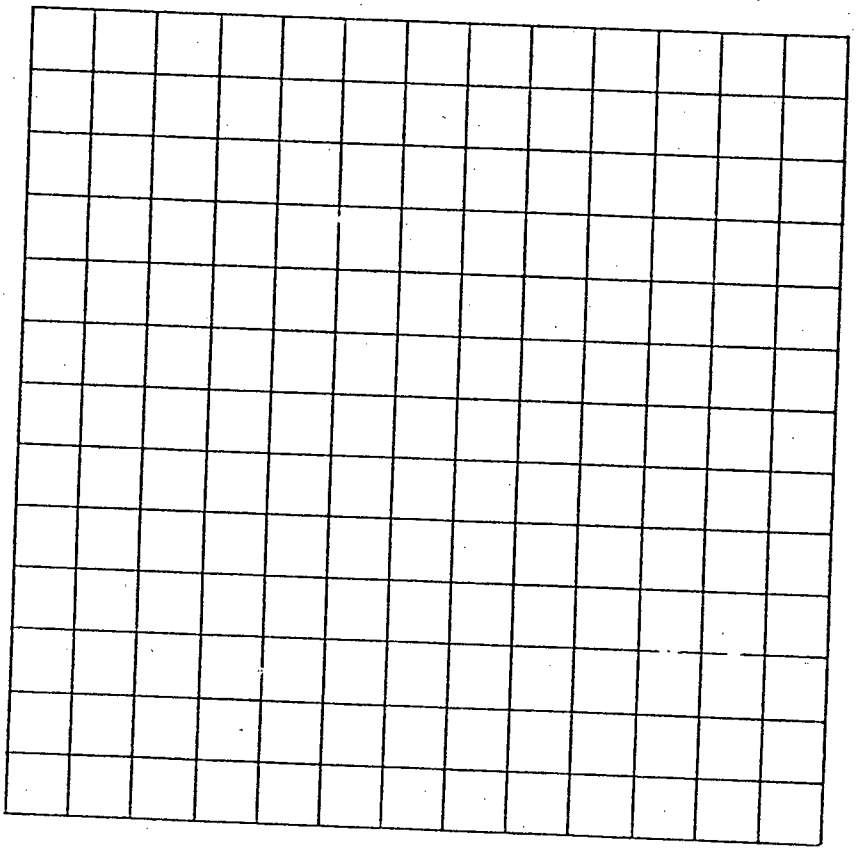
The difference between our approach and their approach is our treatment of the L-θ-list data as a function, $L = F(\theta)$. We are not interested in how often a given point, (L,θ), occurs, as statistical approaches are, rather we are interested in the nature of the function, F. It is this interest that makes our approach a structural one.

3.4 STRUCTURAL APPROACHES: THE MACRO-STRUCTURE

In the last section an algorithm was developed for detecting surface slant from an image which has been pre-processed to yield a list of texels, where each texel had a structure containing surface slant information. This pre-processing stage is a valid assumption for the textures examined in the last section, except, perhaps, for the unconnected micro-structures, but it is not possible for all images. For example, the square texels in Figure f3.4.1 are not easily discernible (it would be necessary to compute intersections of lines, etc.), and the line texels of figure f3.4.2 do not appear to be very useful. Indeed, the useful information in both cases seems to lie in the macro-structure of the texture. In this case, the texels can be viewed as being line segments only. Many successful line finding algorithms have been proposed (Shirai, 1973) (Horn, 1971) (O'Gorman and Clowes, 1973) so the texel detection assumption now sits on a firmer foundation than it did before. In the section, one technique, a modified Hough transform, for analyzing the macro-structure will be examined.

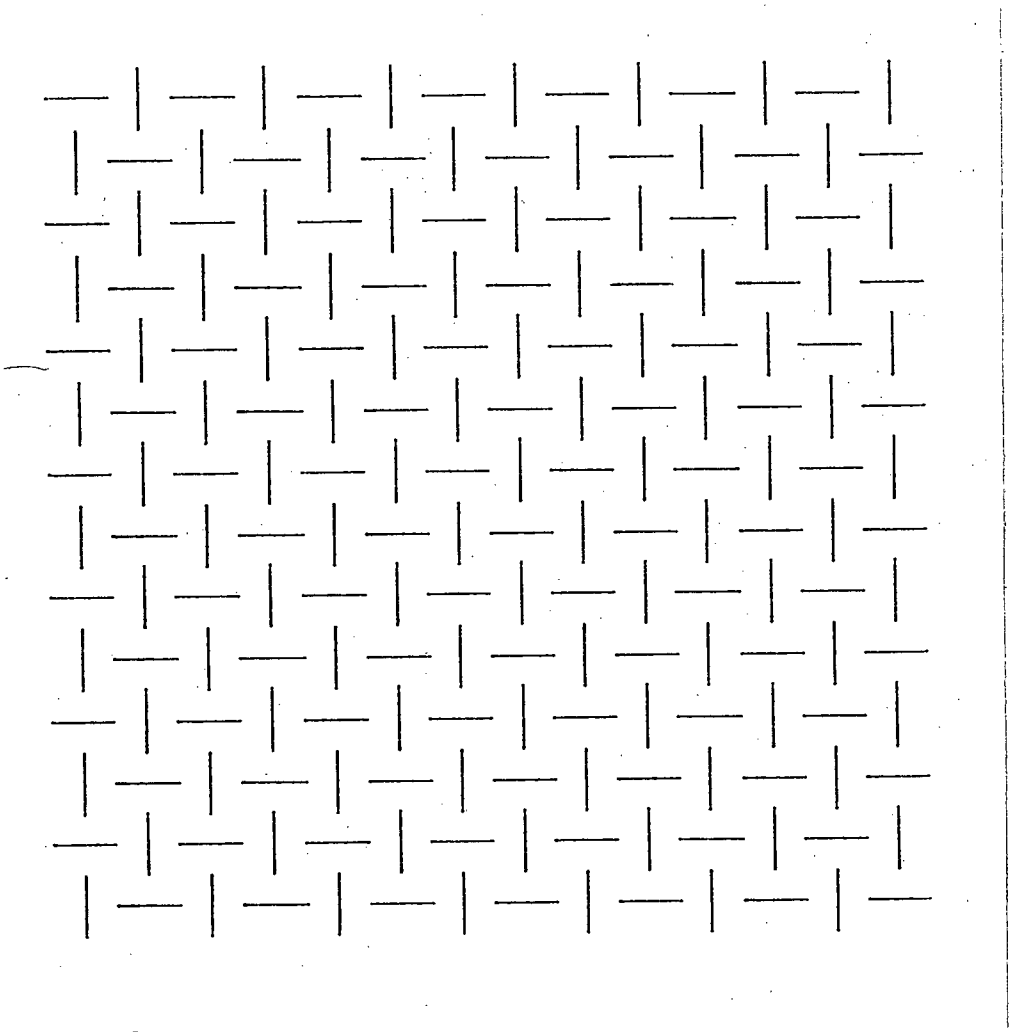
"Grid-like" Textures

It will be assumed that the image is of a single surface, relying on the texture discrimination work described in Chapter 2 to do so. The micro-structures are restricted to line segments of arbitrary length. There may be several such micro-structures for any given image. These micro-structures are placed on the lines of an invisible square grid of fixed spatial frequency. This frequency is the same in both directions of the grid. Both Figures f3.4.1 and f3.4.2 are examples of this kind



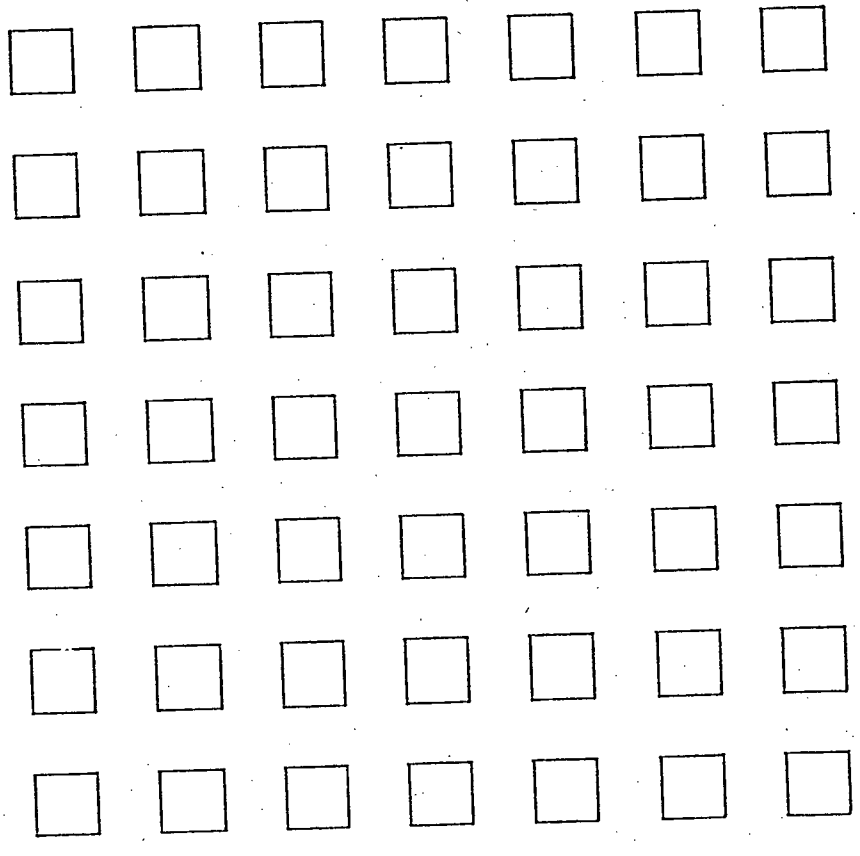
"complete" grid texture

Figure f3.4.1



"line-segments" grid texture

Figure f3.4.2



"boxes" grid texture

Figure f3.4.3

of texture. Another example is shown in Figure f3.4.3. Note that this class of textures is very similar to that of (Kender, 1979) discussed in Chapter 2.2. The texture examined in this thesis is different in that a square grid and a fixed spatial frequency are reguired, which is not the case for Kender's textures. Kender uses Hough transforms to exploit the phenomenon of converging parallel lines in persepective images. Since this thesis is restricted to orthographic images this feature cannot be relied on. Nevertheless, Hough transforms are so natural for this domain of textures that other features can easily be exploited, as was hinted in Chapter 2.2.

The plan here is to examine the image and determine the nature of the parallelogram texels as if they existed. That is, the nature of the images of the square micro-structures of Figure f3.4.3. Of course, only an image of this particular figure will actually contain such texels; the proposed program will be required to "hallucinate" texels for the other images considered. This will be done by using information obtained from a rho-theta Hough transform of the image. Then, assuming that the micro-structure is a square, the gradient of the surface can be determined by analysing the transformation from micro-structure to texel. In preparation, the next few paragraphs will develop the relations necessary to do so.

Note that the lengths of the sides of such texels are related in some way to the normal distances between the line segments. The details of this relation are required because the normal distances mentioned are readily measurable in the Hough transform of the image. The lengths, $|L_1|$ and $|L_2|$, of the sides of a parallelogram are given by:

$|L_1| = |d_1|/\cos\sigma_1$ and

$|L_2| = |d_2|/\cos\sigma_2$

where

$|d_1|$ is the normal distance between sides of length $|L_2|$

$|d_2|$ is the normal distance between sides of length $|L_1|$

σ_1 is the angle between L_1 and the normal to L_2 (i.e. between L_1 and d_1)

σ_2 is the angle between L_2 and the normal to L_1 (i.e. between L_2 and d_2)

This is demonstrated in Figure f3.4.4.

It will be seen that $|d_i|$ and its direction, θ_i, can be computed from the Hough transform of the image. The goal is to use these values to determine the shape of the texels and, from this and the knowledge of the micro-structure shape, to determine the gradient of the surface.

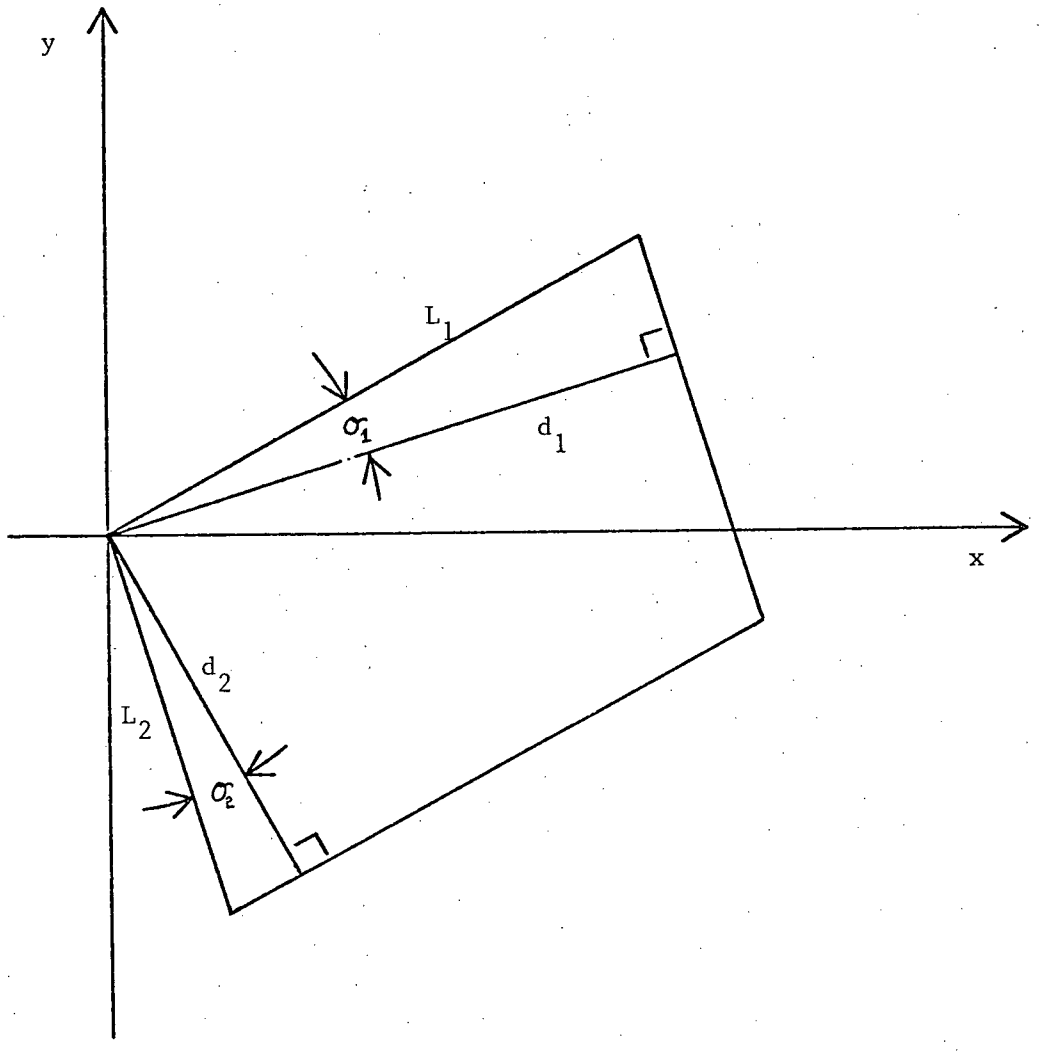
L_1 is normal to d_2 and L_2 is normal to d_1. So, referring to Figure f3.4.5, the directions of L_1 and L_2, μ_1 and μ_2 respectively, are given by

$$\mu_1 = \begin{cases} \theta_2 + \pi/2 & \text{if } \theta_2 \leq 0 \\ \theta_2 - \pi/2 & \text{if } \theta_2 > 0 \end{cases}$$

and

$$\mu_2 = \begin{cases} \theta_1 + \pi/2 & \text{if } \theta_1 \leq 0 \\ \theta_1 - \pi/2 & \text{if } \theta_1 > 0 \end{cases}$$

where θ_i is the direction of d_i. It will be seen shortly that d_i and θ_i are easily obtained from the Hough transform. It is, therefore, simple to calculate the cartesian co-ordinates of an "idealized" texel, one



The relationship between L_i and d_i

Figure f3.4.4

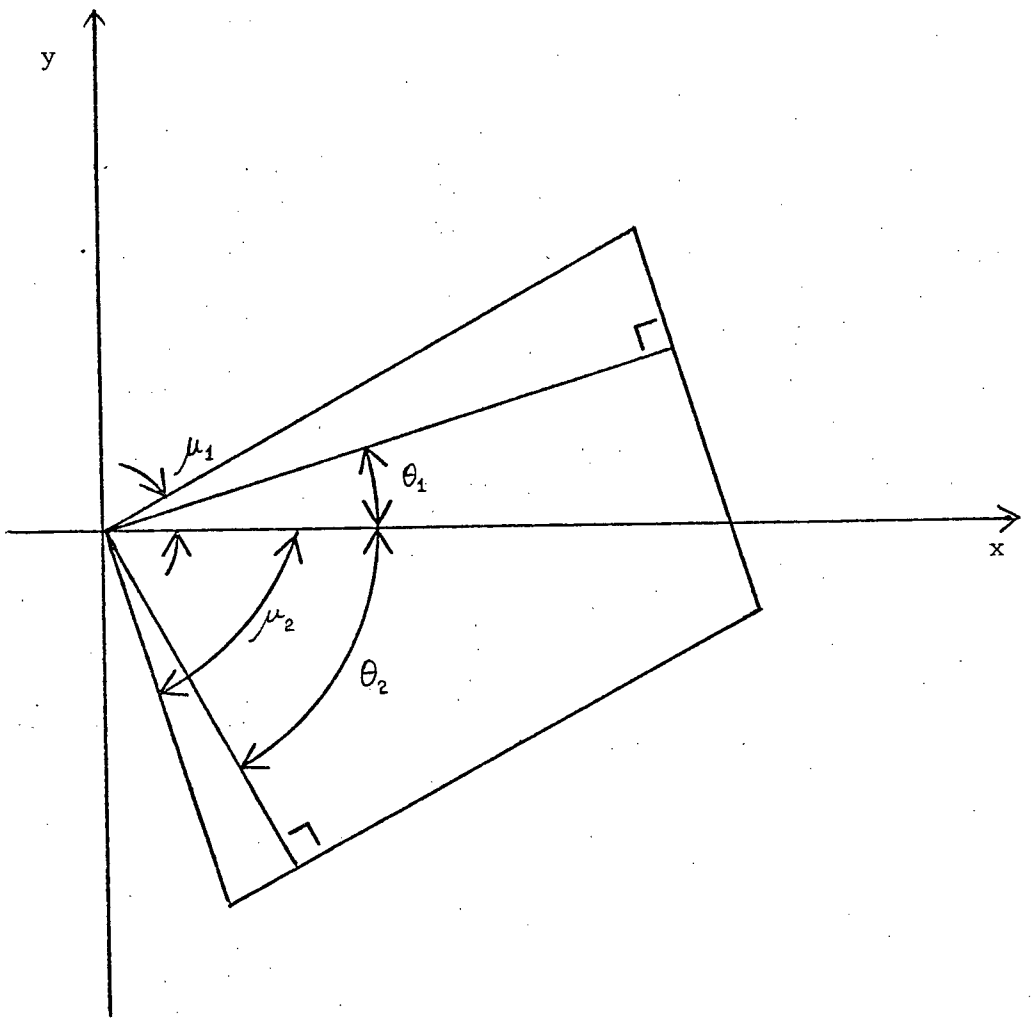


Illustration of μ_1 and μ_2

Figure f3.4.5

corner of which is at the origin, from the polar co-ordinates, (L_1,μ_1) and (L_2,μ_2). If $\mu_2 < \mu_1$, we can represent the texel by the matrix

$$P = \begin{pmatrix} L_2\cos\mu_2 & L_1\cos\mu_1 \\ L_2\sin\mu_2 & L_1\sin\mu_1 \end{pmatrix}$$

The columns of P are interpreted to be cartesian vectors representing the sides of the texel. So P will be viewed as the picture or image of the square micro-structure. Now, (Mackworth, 1974) has shown that, knowing the "true shape" of a "surface" (read "micro-structure"), one can deduce the gradient vector of the surface from its image (read "texel"). In this case the true shape is given by:

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and the apparent shape is given by P (S for square, P for picture).

So the task of determining surface slant is accomplished by applying Mackworth's algorithm to S and P. The details involved in extracting the measures mentioned will now be discussed.

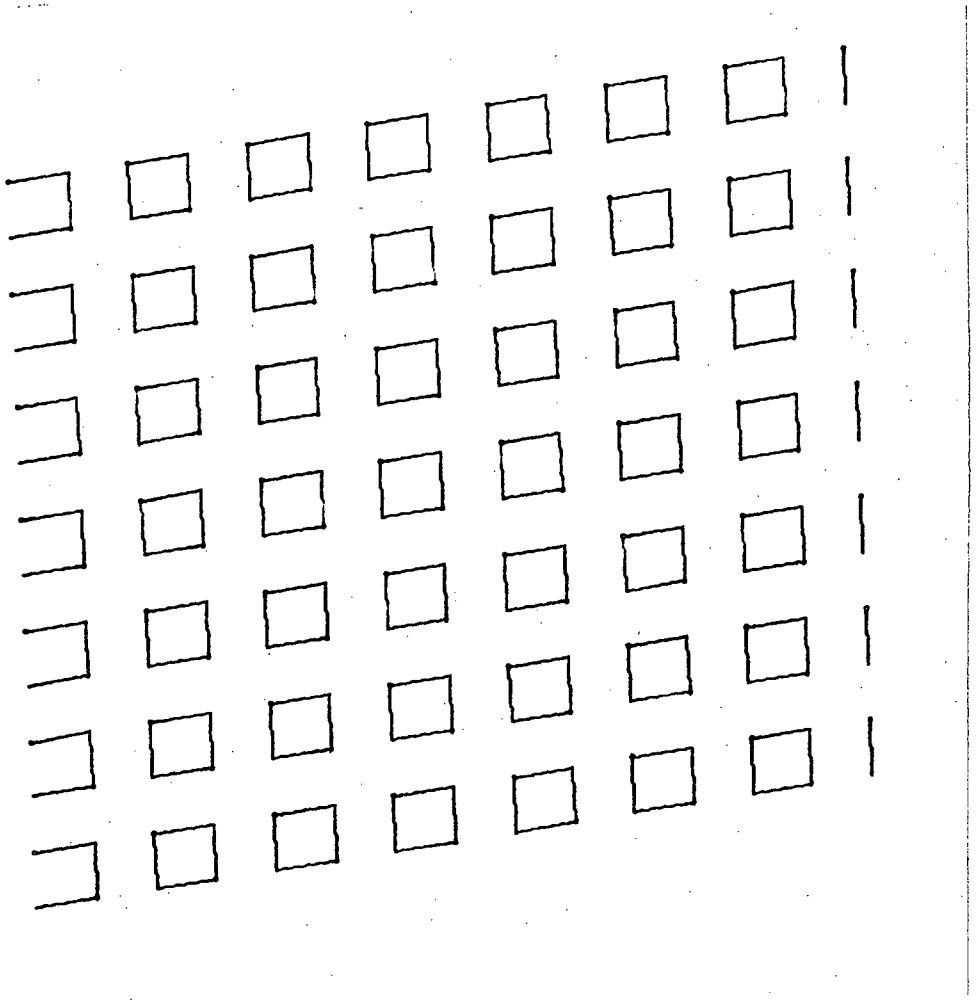
The application of theorem 3.4.1 is required to get the values of $|L_1|$, $|L_2|$, μ_1, and μ_2 since a texel may not exist. However, it is still necessary to compute d_1 and d_2, the normal-distance-between-parallel-lines vectors.

Note first that it is a property of the orthographic imaging process that line segments map into line segments and parallel line segments map into parallel line segments. This assures that the domain of discourse is not vacuous.

The intention is to extract d_1 and d_2 from a rho-theta Hough transform (see Chapter 2 for details of the transform). So each line

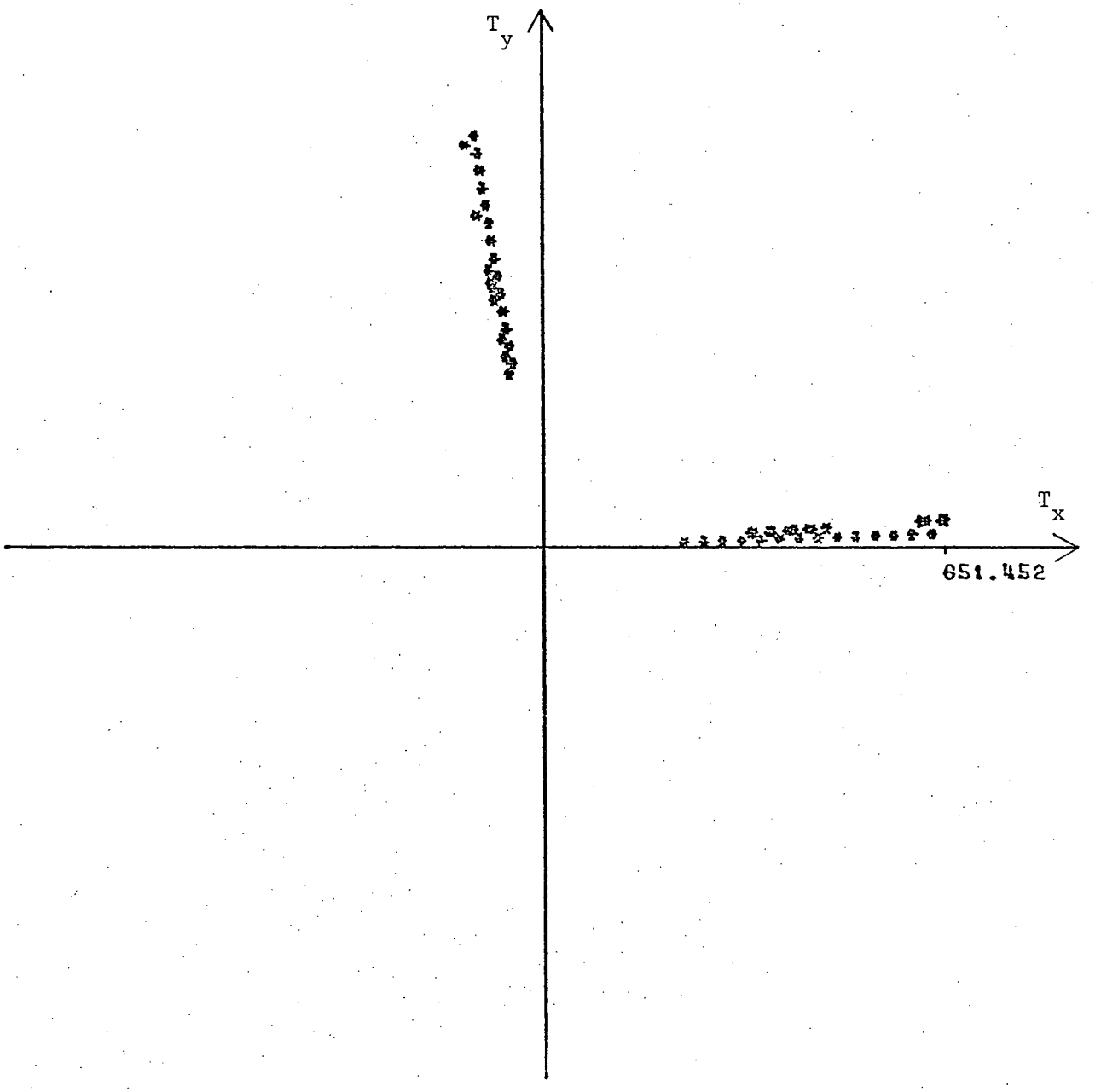
segment is first parameterized in terms of rho and θ. In our implementation this step must actually be carried out since the input is a list of line-segments represented as a list of end points (the data structures used are described later in this section). However, if the input image is in the form of an array of pixels, then it is conceivable that the line detection stage, assumed in this thesis, will be performed with the aid of a rho-theta Hough transform. See Chapter 2 for details of using the Hough transform as a line finder. In this case it would not be necessary to reconvert the data into the endpoint list representation used in this thesis. Input may simply be the accumulation points in rho-theta space, as this is precisely the information we want. So, having done this, the broken lines vanish and only solid lines result in the parameterization For example, a polar-plot of the parameterization of Figure f3.4.6 is shown in Figure f3.4.7. Notice that this transform is exactly that of the complete grid "underlying" Figure f3.4.6. Each set of parallel lines in the image contains identical, unique to the set, values of theta (θ). It is now pointed out that the normal distance between lines is simply the difference, delta-rho, in the value, rho, between adjacent points in the transform.

This task, then, is simple; group the parameterized points by θ and, for each group, measure delta-rho. Then d_1 and d_2 are set to the the two delta-rhos and θ_1 and θ_2 are set to the corresponding θs from the transform. The angles μ_1 and μ_2 are then set to the complements of θ_2 and θ_1 respectively. Then



Slanted surface textured with "boxes" grid

Figure f3.4.6



Polar plot of Hough transform of Figure f3.4.6

Figure f3.4.7

$\sigma_1 = \mu_1 - \theta_1$

and

$\sigma_2 = \mu_2 - \theta_2$

(see theorem 3.4.1 and Figures f3.4.4 and f3.4.5 again)

Finally, $L_1 = d_1 / \cos \sigma_1$

$L_2 = d_2 / \cos \sigma_2$

as dictated by theorem 3.4.1. The matrix P is then constructed as described previously. Then using a specialized version of Mackworth's algorithm (see the appendix to this section), the amount of slant, ϕ, and the direction of the gradient vector, σ are calculated.

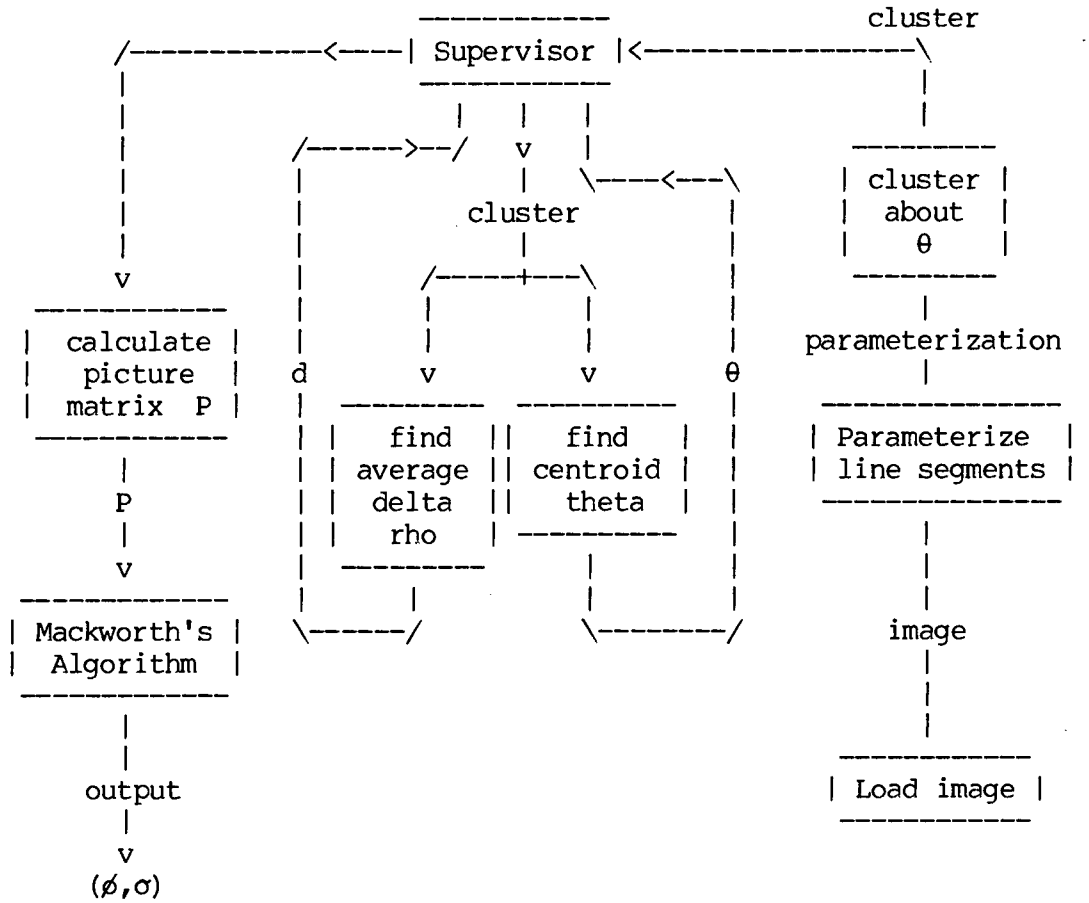
Results

This system has been implemented and tested on several images. We found the results to be somewhat more accurate than the results of the section on micro-structural analysis. In that section, the angles ϕ and σ_n were obtained to within ±0.09 radians. The macro-structural analysis implementation obtained the angles ϕ and σ to within ±0.01 radians. This difference in accuracy may be attributed to the statistical nature of the algorithm of that section. No such statistical assumptions are made here (the assumption of "enough" texels is not made nor are values such as σ_n guessed at.) Some sample runs are given on the following pages.

The entire algorithm is illustrated in figure f3.4.8. The supervisor serves the same purpose here as in the last section. For a noiseless image, the values of the θs and delta-rhos for each group would be identical. Due to noise, however, it was found necessary to

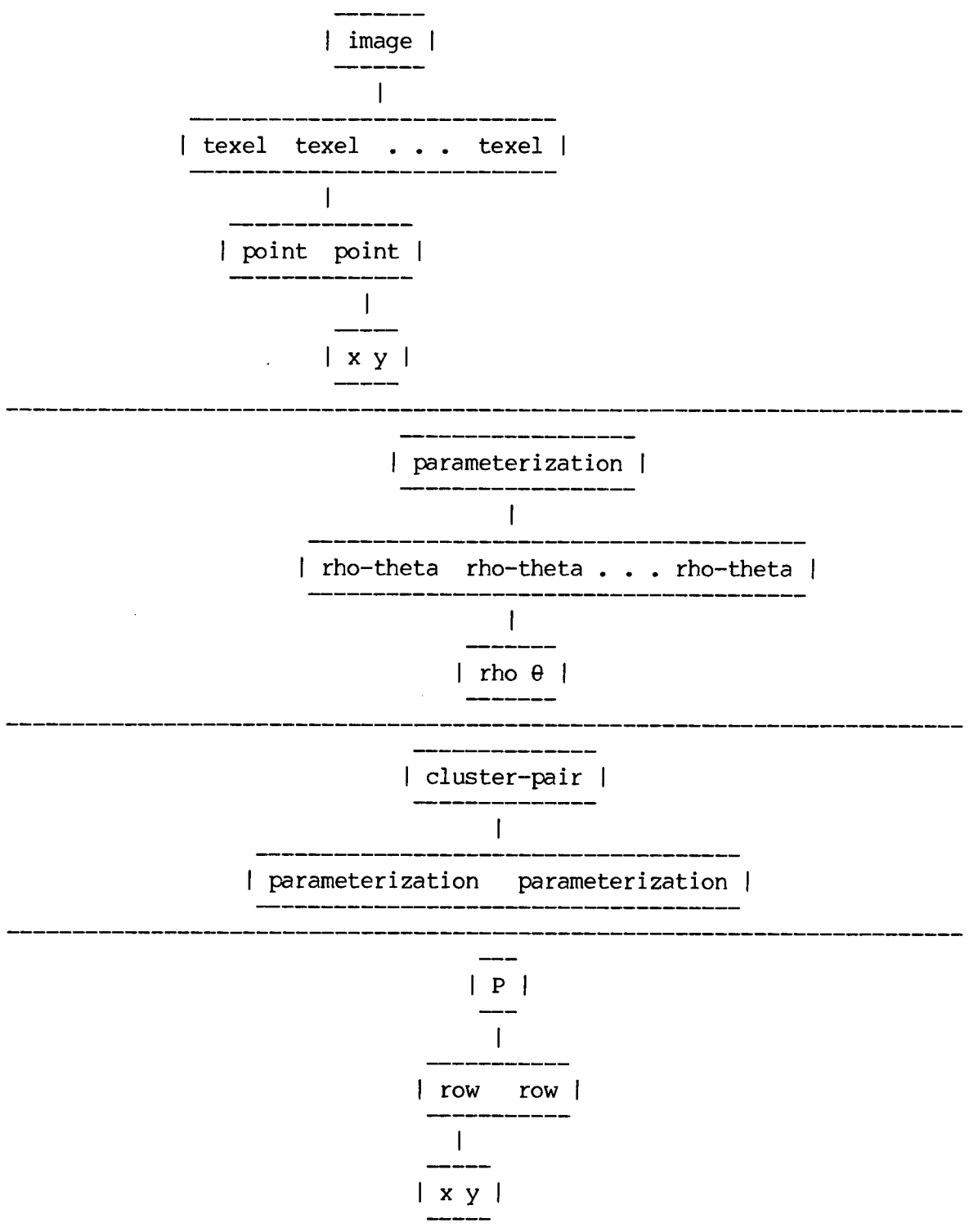
calculate the value of theta for each group's centroid and the average value of delta-rho.

The data-structures used are illustrated in figure f3.4.9. Note that the "image" data structure is the same as in the last section, except that each texel contains only two points, the end-points of each line segment. Two sample runs are shown on pages 110-111.



Flow of Control for the Macro-structural Analysis Algorithm

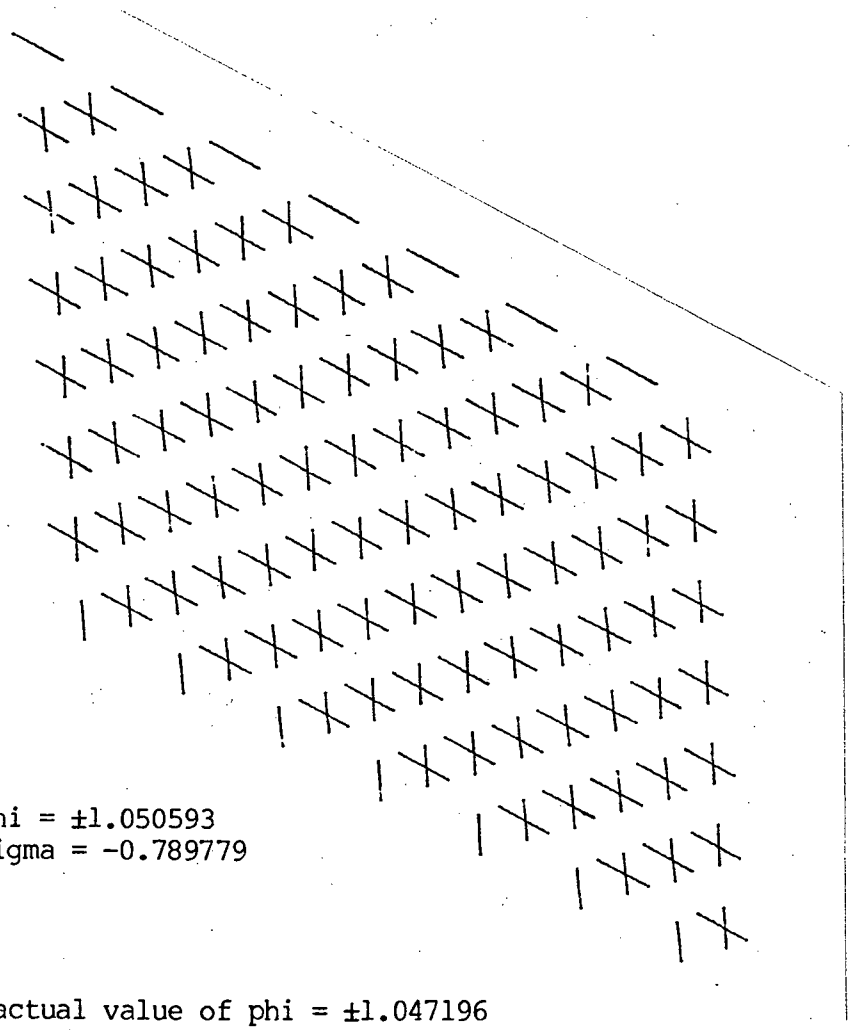
figure f3.4.8



Data Structures for the Macro-structural Analysis Algorithm

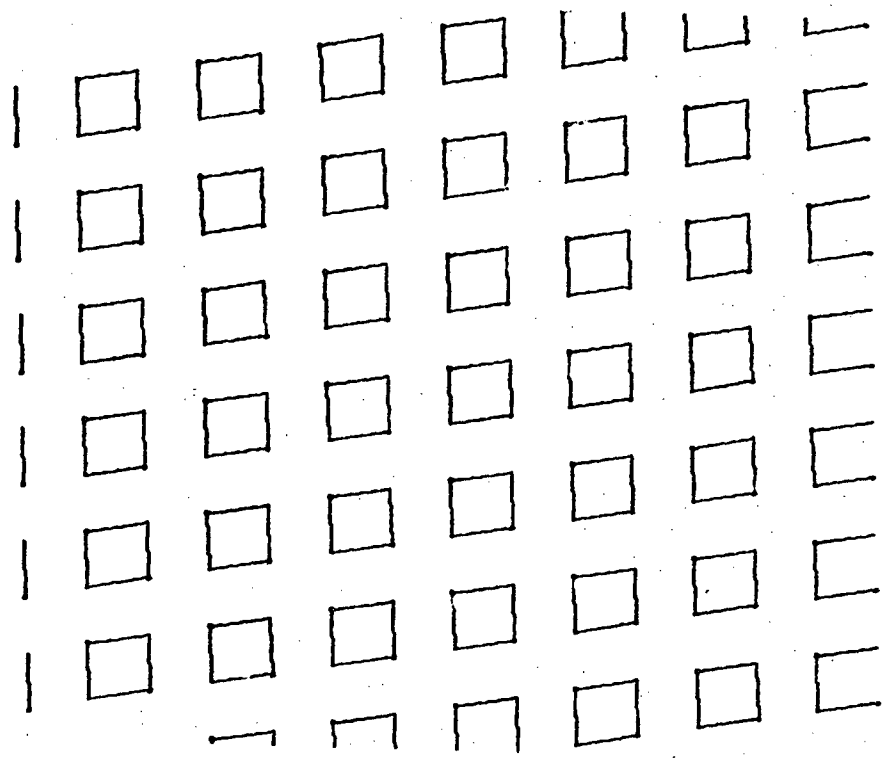
Figure f3.4.9

```
* (supervisor (clusters (parametrize (image>
* THERE ARE  1 PLOTS IN THE FILE.
*  ENTER NUMBER OF PLOT TO BE DISPLAYED. * 1
* ------------------------------------------
* VECR CONTAINS EIGENVECTORS
*   0.710198  0.704001
*  -0.704002  0.710199
* ------------------------------------------
* EVR CONTAINS EIGENVALUES
*   4.417268  2.195629
* ------------------------------------------
*
```



```
>    phi = ±1.050593
>    sigma = -0.789779
*
*
*
*
> The actual value of phi = ±1.047196
*
> The actual value of sigma = -0.785398
*
```

```
* (supervisor (clusters (parametrize (image>
* THERE ARE  1 PLOTS IN THE FILE.
*  ENTER NUMBER OF PLOT TO BE DISPLAYED. * 1
* ---------------------------------------------
* VECR CONTAINS EIGENVECTORS
*   0.881348 -0.472472
*   0.472467  0.881346
* ---------------------------------------------
* EVR CONTAINS EIGENVALUES
*  17.906174 14.963252
* ---------------------------------------------
*
```



```
>     phi = ±0.581486
*
>     sigma = 1.078706
*
*
*
> The actual value of phi =  ±0.523598
*
> The actual value of sigma =  0.942477
*
```

4. AN INTERPRETATION OF THE RESULTS

4.1 SUMMARY

The previous chapter examined texture from three different points of view:

- statistical analysis
- micro-structural analysis
- macro-structural analysis

Section 3.1, the introduction to the chapter, defined "orthographic projection" and proved a theorem illustrating the foreshortening property of this projection. It was also noted that this property is not sufficient to distinguish between Necker reversals of a scene.

Statistical approaches to detecting surface slant from texture measures were discussed in Section 3.2. These measures were in the form of 2x2 co-occurence matrices and were applied to images of regular dots and of circles. It was concluded that such a blind application of statistical techniques is not very useful to the problem of detecting surface slant, due to data from complex images that proved difficult to interpret. A theorem was then proved stating that there is no directionality in the co-occurence matrices computed for images of truly random textures. It was concluded from this theorem that co-occurrence matrices can not be used to determine surface slant from such images.

Section 3.3 examined the analysis of the texels' structures. It was assumed that each texel contained a line (or virtual line) which is identifiable as the image of a given line (or virtual line) in the micro-structure. In the algorithm developed within this section, the

presence of a sufficiently long diagonal was required to assure an identifiable line in the texel.

The use of the macro-structure of a texture was examined in Section 3.4. Grid-like textures were assumed; they were analysed by determining the shape of the texels as if they existed. This was accomplished by applying a Hough transformation to the image and retrieving the needed information from the resulting transformed image. Then, assuming the micro-structure to be a square, the gradient of the surface was computed with an algorithm due to Mackworth.

4.2 FALLACIES, OR WHY 3.2, 3.3, AND 3.4 ALL START WITH 3

At first glance, these 3 approaches are quite independent of one another. A closer look reveals marked similarities in the underlying assumptions.

The Fallacy of the Statistical-Structural Dichotomy

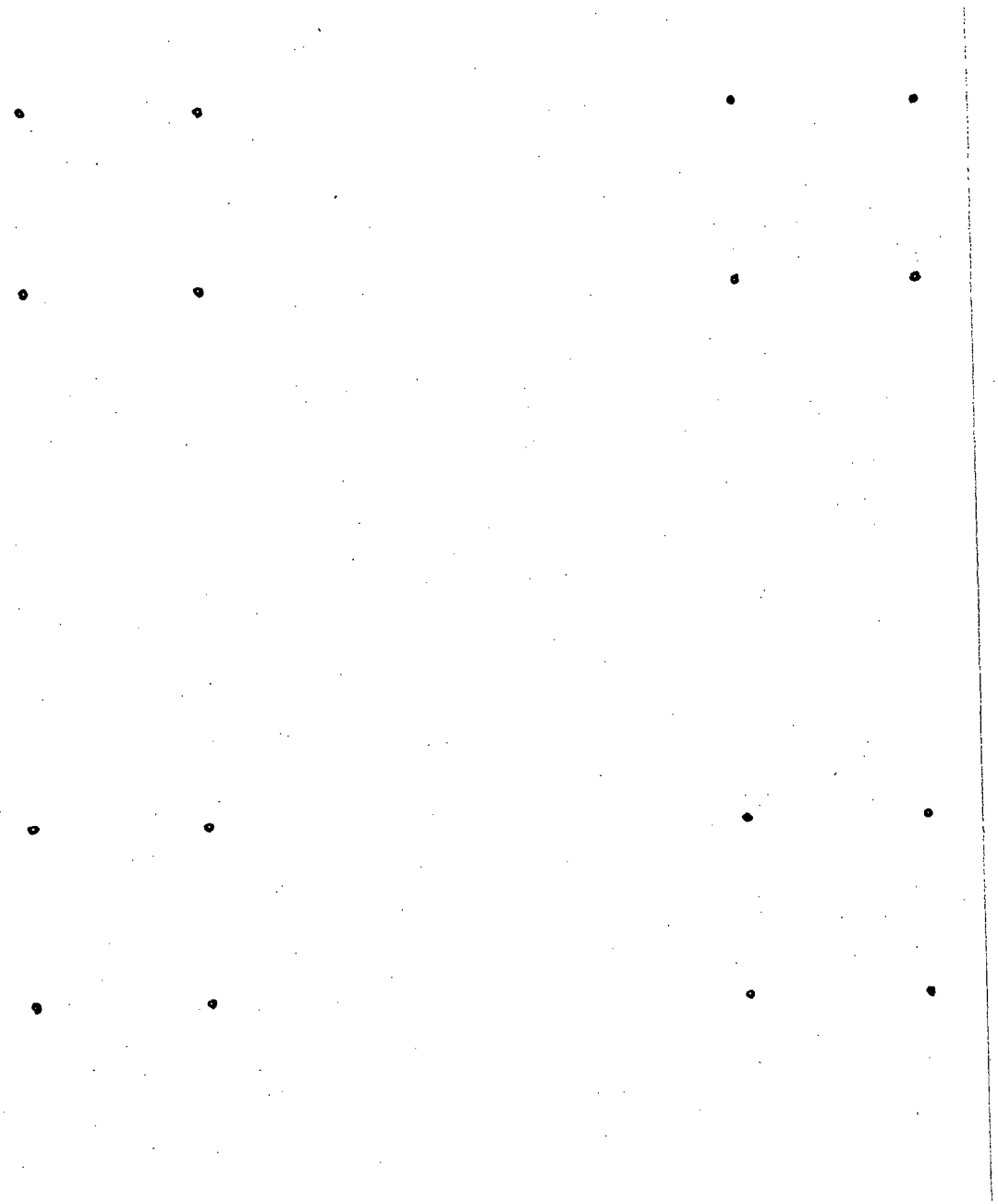
The texture of randomly, uniformly distributed dots considered in theorem 3.2.1 (the non-directionality of random textures) can be described as one having a micro-structure consisting of a dot of negligible size and a macro-structure defined as randomly distributed micro-structures on a surface. In light of Sections 3.3 and 3.4, neither of these structures is useful. Hence, it is not surprising that statistical techniques fail on this image as well. Throughout Chapter 3.2, the use of second order statistics was "blind", that is, they were applied over the whole image just to "see what could be seen".

Admittedly, not much is seen. But suppose that more judicious use is made of them, as it is in (Schatz, 1978). Schatz (see Chapter 2 for details) considers only the second order statistics taken for end points of lines or "virtual" lines. This is, effectively, the same information that is calculated and used in Chapter 3.3. In fact, this thesis maintains that second order statistics are best thought of as dipole structures in the image. It is true that it would be hard to extract just what is needed from Schatz' data since the concept of "longest diagonal" will get buried in all the other length measurements and that Schatz is interested in the occurrence statistics, whereas our interest is in the length as a function of direction, but these are largely matters of proposed application.

So Schatz' statistical work is, in some sense, micro-structural analysis since statistics are confined to pairs of points connected by structures (lines) and the micro-structural analysis (Chapter 3.3) is, in some sense, statistical since the L-θ-list can be reegarded as statistical measures taken of the texels.

The Fallacy of the Micro/Macro-Structure Dichotomy

Similarily, the distinction between micro- and macro-structure becomes less well defined as it is examined. There was little effort involved in distinguishing micro- from macro-structure in Chapter 3.3. The texels were simply considered to be the "connected" line segments (although we did wave our hands and allow unconnected texels). The problem arises when the connections are not so convenient. Figure f4.2.1, for example, may be considered to be dots arranged in a complex



Hierarchical texture I

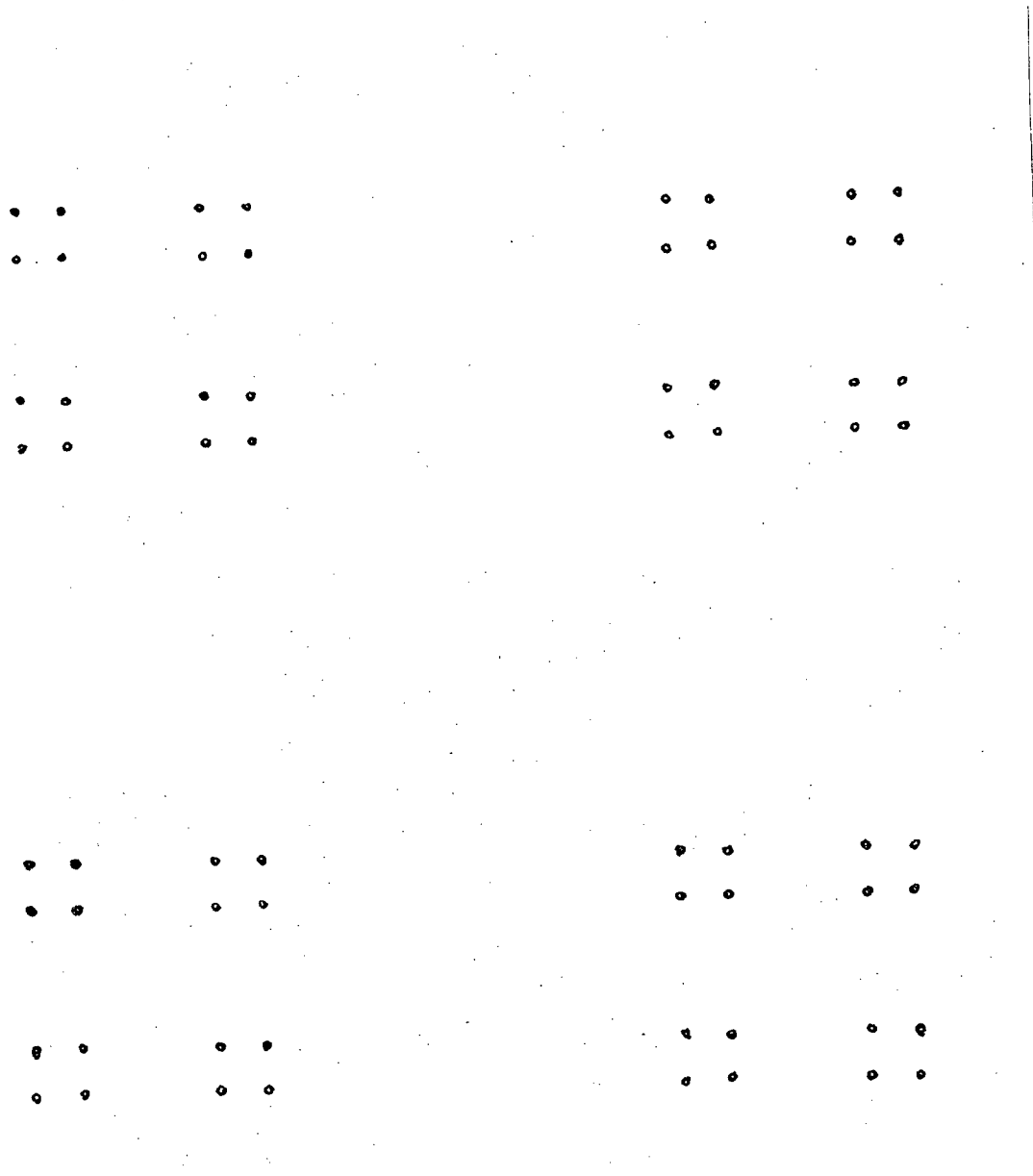
Figure f4.2.1

macro-structure. However, it may be more profitable to consider it as square texels, made up of dots, arranged in a simple macro-structure, namely a square. Note that the labelling of macro- and micro- structure depends on the scale of the image. For example, Figure f4.2.1 may just be a micro-structure of Figure f4.2.2.

In fact, this thesis proposes that a macro-structure is best expressed as a micro-structure, and this is precisely what is done in Chapter 3.4 when the dimensions of a parallelogram texel are found, even though it may not exist.

This thesis also maintains that the labels "micro-structure" and "macro-structure" are misleading and may be forcing the acceptance of an unnatural view of texture. Certainly, all of the views discussed in Chapter 2.1 have fallen into this trap. Since there is no real distinction between micro- and macro- structure, it is suggested that it would be best to express a texture as nested structures, all of which are expressed as micro-structures. For example, Figure f4.2.2 should be described as "a (square of (squares of (squares of dots)))". It would then be possible for a program (or user) to decide on what level a "texel" should exist.

Under this interpretation, then, all of the techniques of the last chapter are structural. They appear different at first because each constructed an artificial boundary to the level of structural analysis it would undertake. Now, the claim is not made that it is impossible to do any sort of non-structural statistical analysis of texture, but it is claimed that such an analysis has no place in the domains this thesis is restricted to. This appears to be due to the nature of the orthographic imaging process and is discussed in the next chapter.



Hierarchical texture II

Figure f4.2.2

5. CONCLUSION

Texture has been shown to be a useful cue for the detection of surface slant in orthographic images. Texture is used by applying structural analysis techniques to it, determining surface slant by relying on the "foreshortening" property of the orthographic imaging process (see theorem 3.1.1). In all cases this has been done without considering texture gradient at all. As mentioned in Chapter 2, all previous works described have used texture gradient since they assumed a perspective imaging process.

Similarily, purely statistical techniques were found not to be useful in the domain of apparent textures under orthographic projection. Let us point out, however, that if the statistics (notably first and second order statistics) changed in some determinable way with the distance of the textured object from the viewer, then surface slant could be determined by a purely statistical analysis. This is the case for perspective images and Bajcsy uses the change in first-order statistics (texture gradient) in her work, as does Gibson (see Chapter 2). This is not the case for orthographic projections since the location of the viewer is indeterminate on the Z-axis (see Chapter 3.1).

We have put forth several opinions suggesting a "nested" structural description of texture. It was found that this notion clashes with previous definitions (see Chapter 2) which defined texture in terms of a micro-structure and a macro-structure. It was also found that the new definition makes difficulties such as texel determination vanish by making them non-problems. One reason for the lack of greater progress in texture and the lack of acceptance of texture by related concerns in scene analysis is the unavailability of such a definition. It is in the

spirit of providing a better basis for texture research, then, that we propose it.

5.1 DIRECTIONS FOR FURTHER RESEARCH

Extensions

Immediate extensions of this work would be to relax the various assumptions made throughout the paper. Such extensions might include:

Incorporation of "Real-World" Textures

This would basically take the form of introducing variation into the strictly structural textures mentioned. For example, random, many-sided polygons distributed uniformly over a surface could simulate cork-board or plastered-wall texture. Pebble walls and concrete walks are of the type of texture studied in Chapter 3.3 but with uneven "texels". However, many man-made surfaces, such as tiled floors, woven cloth, and windowed building walls, are textured in the way studied in Chapter 3.4 with little or no changes. These, too, should be studied more closely.

Non-planar Surfaces

This would probably present a very difficult problem for the class of texture examined in Chapter 3.3 since so little information is available "locally" in the image. In Chapter 3.3, we computed the surface slant from texture measures made over the entire surface. We relied on many such measures to accurately fit an ellipse to the data. Hence, the technique of that chapter would not be applicable to non-

planar surfaces. As (Kender, 1979) also points out, such processing would be in the manner of (Woodham, 1977).

However, non-planar surfaces should present little difficulty if the texture is "continuous", as the grid-like textures of Chapter 3.4 are. This class of textures then, would be a logical starting point for such work. Such textures are often used in line drawings by artists and workers in computer graphics precisely because the surface gradient and curvature are so easily recovered.

Perspective Projections

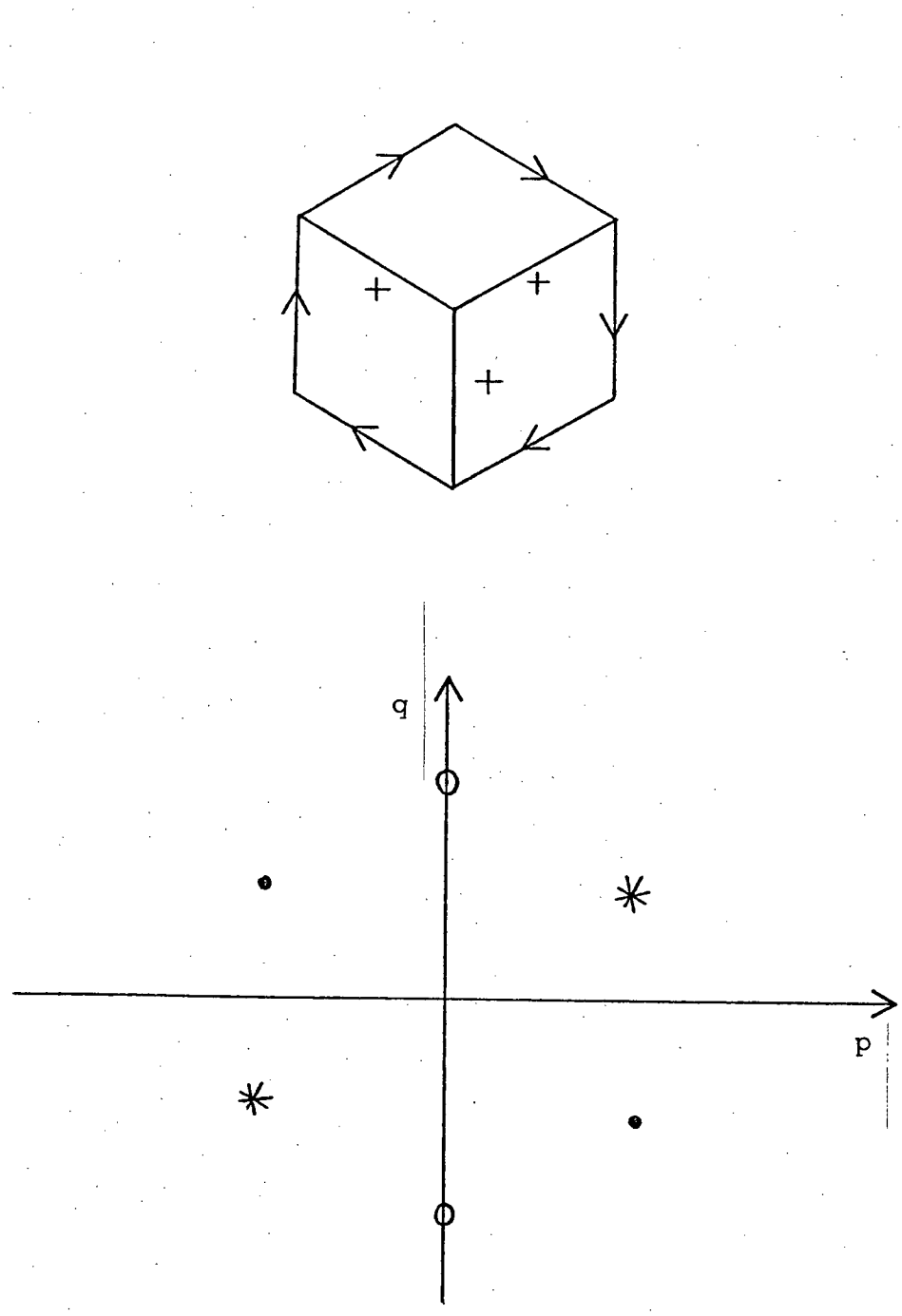
This paper has dealt solely with orthographic projections and, while this is a good approximation to many "real" images that one may want to deal with, it is not a valid assumption for all cases. In fact, all single-camera, fixed-viewpoint images are perspective images and we should be able to deal with these. It is not clear to us how the algorithms in this paper could be generalized but we do not doubt that they could be since the foreshortening aspect of orthographic projections is also a part of perspective projections. Generalization to perspective images is a promising extension, for, as Bajcsy and Kender have discovered, there is much more information present due to the fixed position of the viewer.

Formalize the Definition

Our definition of texture as "nested" structures is very informal. We feel that this definition is important enough to require formalization. It would then be necessary to explore the power and restrictions of the definition and to apply it to such tasks as texture description and generation.

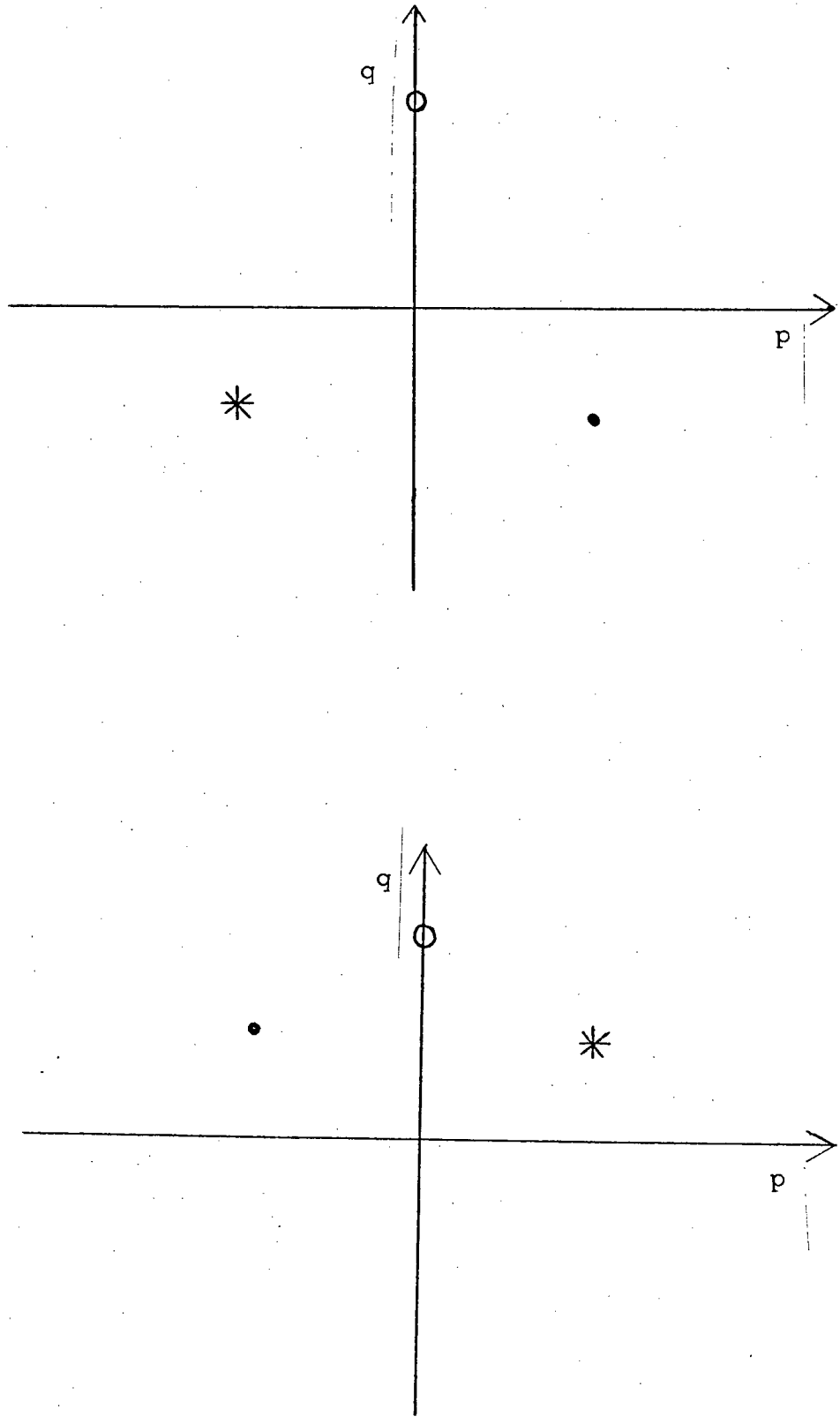
Application to the Origami World

The origami world was introduced to illustrate our motivations in Chapter 1. However, the application of our work to this domain is not as straight forward as we may have led the reader to believe. The difficulty is due mainly to the inability of texture analysis to distinguish between Necker reversals. We are not saying that it should, though. Nevertheless, this does introduce some complications. Let us look again at our "cube" example. Figure f5.1.1 illustrates the six gradient values that could be obtained from such a cube. But how are we to determine that the gradient configurations for the labelling of Figure f5.1.1 are given by Figure f5.1.2 a) and not by Figure f5.1.2 b)? The answer is that we must look again at the labelling and see which of the possible gradient configurations are "consistent" (Mackworth, 1973). Note that none of the configurations may be consistent. In this case, we have determined that the labelling is not correct for the given image.



"cube" labelling and gradients

Figure f5.1.1



Two "possible" gradient configurations of Figure f5.1.1

Figure f5.1.2

5.2 SYNOPSIS

Summing up, then, several algorithms operating on several classes of textures have been presented. It has been decided that these classes are not actually distinct, but that the distinction is forced by the currently held definition of texture. In light of these results, we have proposed a new definition of texture which avoids this class distinction.

Several directions for further study, including an application to a problem which served as a motivation for this thesis, have been suggested. It is hoped that the reader will think seriously about at least one of these problems, for only by confronting such problems will (s)he gain a real feeling for the problems inherent in determining the nature of a scene from its image.

6. APPENDICES

6.1 APPENDIX TO 2.2

Derivation of the rho-theta Equation of a Line

Referring to Figure f6.6.1, we note that a point X = (x,y) is on the line precisely when the vector, X-N is perpendicular to the vector, N. That is, when,

$$(X-N)\cdot N = 0$$

$$X\cdot N - N\cdot N = 0$$

$$X\cdot N = N\cdot N$$

$$X\cdot N = |N|^2$$

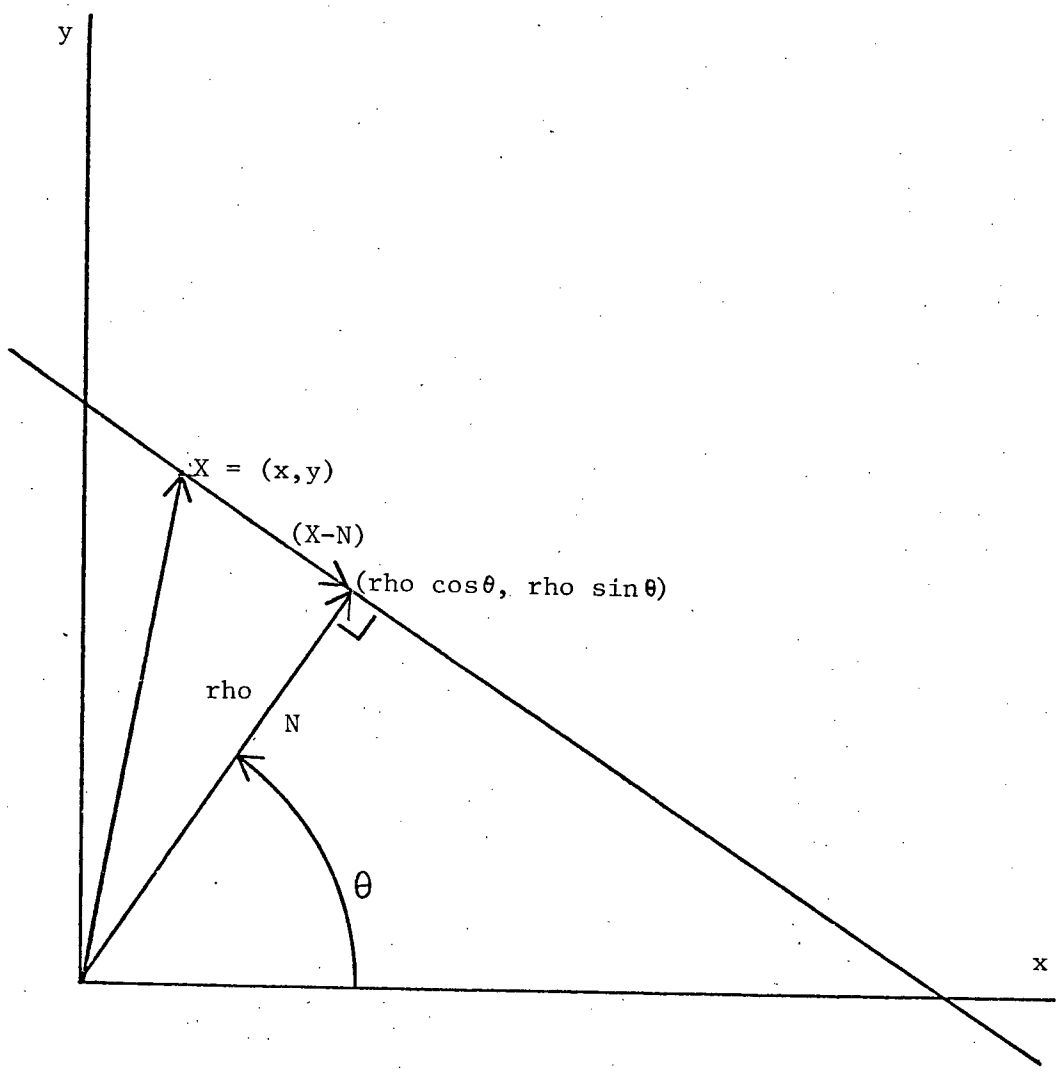
$$X\cdot N = rho^2$$

$$(x,y)\cdot(rho*\cos\theta, rho*\sin\theta) = rho^2$$

$$x*rho*\cos\theta + y*rho*\sin\theta = rho^2$$

$$x\cos\theta + y\sin\theta = rho$$

Which is therefore the rho-theta equation of the line.



Rho-theta equation of a line

Figure f6.1.1

6.2 APPENDIX TO 3.3

Least Squares Fit of Data to an Ellipse

Theorem 3.1.1 gives us good reason to believe that our data, L-θ-list, lies on ellipse with the equation:

$$F(\theta;a,b) = r^2$$

$$= (a^2b^2) \,/\, [a^2\sin^2(\theta-\sigma_n) + b^2\cos^2(\theta-\sigma_n)] \qquad (1)$$

where σ_n is the angle that the major axis of the ellipse makes with the line, $\theta = 0$.

Our task here is to determine a and b. In order to do so, it will be convenient to express $F(\theta)$ in a form that depends linearly on its parameters. Thus, let us not use $F(\theta) = r^2$, but rather $F(\theta) = 1/r^2$. Then

$$F(\theta;a,b) = 1/r^2 = [\sin^2(\theta-\sigma_n)] \,/\, b^2 + [\cos^2(\theta-\sigma_n)] \,/\, a^2$$

$$F(\theta;c_1,c_2) = c_1\phi_1(\theta) + c_2\phi_2(\theta) \qquad (2)$$

where $c_1 = 1/b^2$ $\quad c_2 = 1/a^2$

$$\phi_1(\theta) = \sin^2(\theta-\sigma_n)$$

$$\phi_2(\theta) = \cos^2(\theta-\sigma_n)$$

So we will be content to determine c_1 and c_2 and, from them, a and b. To get a least-squares approximation, we wish to minimize the sum of the squares of the errors. That is, we wish to minimize

$$E(c_1,c_2) = \sum_{i=1}^{n} [f_i - F(\theta_i;c_1,c_2)]^2$$

(Conte and de Boor,1965; pg.241-246)

where $f_i = 1/L_i^2$

n = number of elements in L-d-list

(L_i,θ_i) is the i^{th} element of L-d-list

It should be pointed out that we have made a major compromise by insisting on the linearly dependent form. Ideally, we would like to minimize

$$\sum_{I=1}^{n} [L_i - r]^2.$$

With $F(\theta) = 1/r^2$ we are actually minimizing

$$\sum_{I=1}^{n} [1/L_i^2 - 1/r^2]^2.$$

This "weights" the small values of L_i and r more than large values, hence the data will fit the minor-axis part of the ellipse found by this method better than the rest of the ellipse. Nevertheless, we will continue using this method because of its ease of applicability.

Since $E(c_1,c_2)$ is continuously differentiable with respect to c_1 and c_2, we can detect the minimum at the point where the 1^{st} partial derivatives vanish. That is

$$\partial E(c_1,c_2)/\partial c_j = \sum_{i=1}^{n} \partial[f_i - F(\theta_i c_1,c_2)]^2/\partial c_j$$

$$= -2 \sum_{i=1}^{n} [f_i - F(\theta_i c_1,c_2)]\ \partial F(\theta_i;c_1,c_2)/\partial c_j \qquad (3)$$

Now

$$\partial F(\theta_i;c_1,c_2)/\partial c_j = \partial[(c_1\phi_1(\theta_i) + c_2\phi_2(\theta_i)]/\partial c_j$$

$$= \phi_j(\theta_i) \qquad (4)$$

So substituting (4) into (3) we get

$$-2\sum_{i=1}^{n}[f_i - F(\theta_i c_1,c_2)]\phi_j(\theta_i) = 0$$

And substituting (2) into this we get

$$c_1\sum_{i=1}^{n}\phi_j(\theta_i)\ \phi_1(\theta_i) + c_2\sum_{i=1}^{n}\phi_j(\theta_i)\ \phi_2(\theta_i) = \sum_{i=1}^{n} f_i\phi_j(\theta_i)$$

Hence we have a system of two linear equations (j=1,2) in two unknowns (c_1,c_2). This system is then solved using Gaussian elimination (Conte and de Boor,1965;pg 110-127).

6.3 APPENDIX TO 3.4

Orientation of a Surface from True Shape and Projected Shape

This is taken nearly verbatim from (Mackworth, 1974).

The orthographic imaging process may be given as a matrix transformation by:

$$P = (\text{tilt})(\text{rotation})(\text{scale})F$$

where

P is of dimensions 2*(n-1) and contains the co-ordinates of all n vertices except the pair at the origin.

F is like P but contains the vertices of the true face.

$$(\text{scale}) = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

$$(\text{rotation}) = \begin{pmatrix} \cos(r) & \sin(r) \\ -\sin(r) & \cos(r) \end{pmatrix}$$

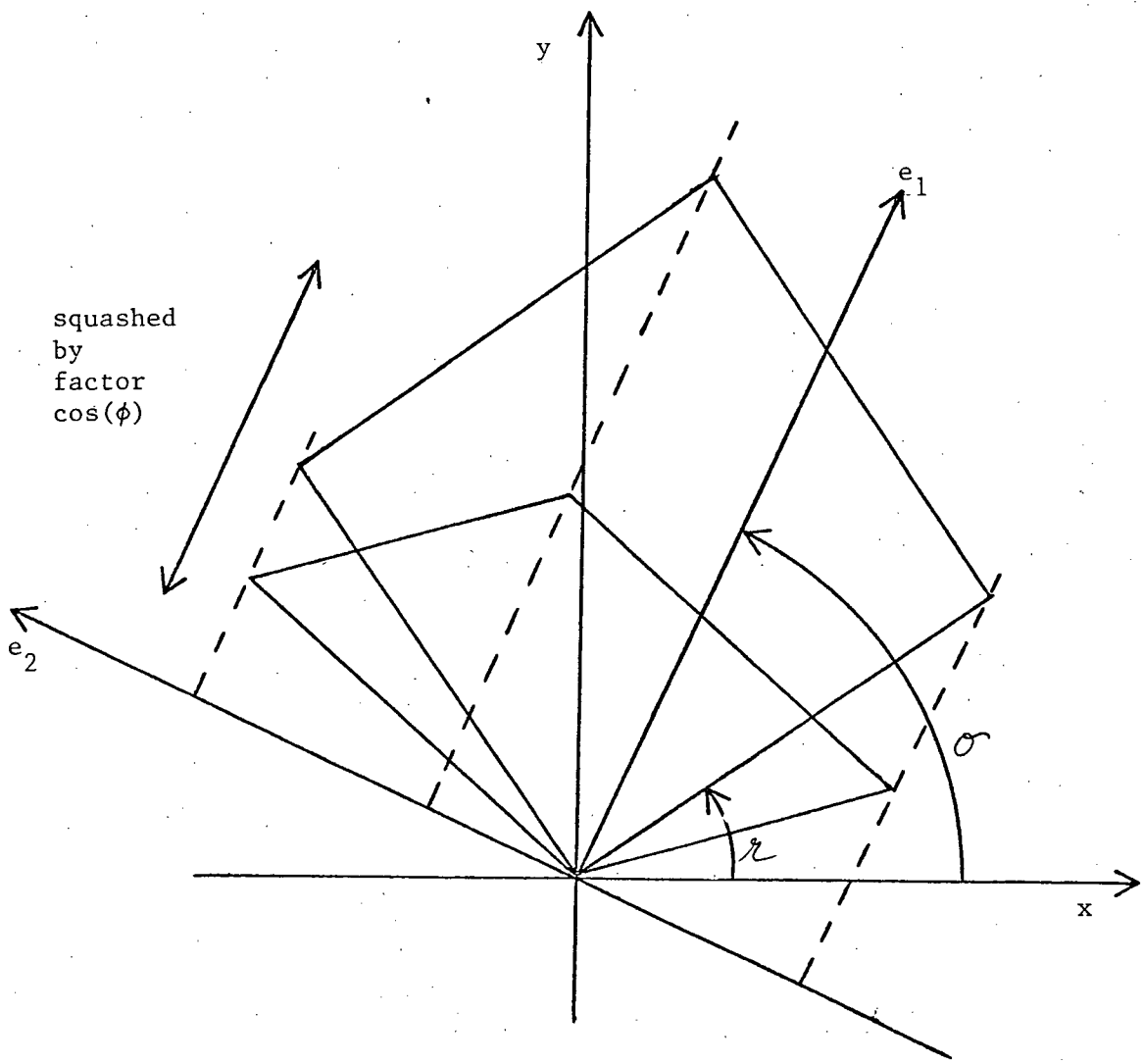
$$(\text{tilt}) = \begin{pmatrix} \cos(\sigma) & -\sin(\sigma) \\ \sin(\sigma) & \cos(\sigma) \end{pmatrix} \begin{pmatrix} \cos(\phi) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\sigma) & \sin(\sigma) \\ -\sin(\sigma) & \cos(\sigma) \end{pmatrix}$$

See Figure f6.3.1 (also from (Mackworth, 1974)).

In our case, F represents a square and is given by:

$$F = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ which is the identity matrix

So P = (tilt)(rotation)(scale). Now from P, we would like to obtain the direction and magnitude of the gradient vector, σ and $\tan(\phi)$



"squashing" effect of the orthographic projection

Figure f6.3.1

respectively.

We first compute r as follows:

(tilt)(rotation)(scale) = P

(tilt)(scale)(rotation) = P

(tilt)(scale) = P(rotation)$^{-1}$

$$= \begin{pmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{pmatrix} = J$$

Since (tilt) and (scale) commute and are both symmetric then the right hand side must also be symmetric:

$$\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} \cos(r) & -\sin(r) \\ \sin(r) & \cos(r) \end{pmatrix}$$

That is, $-p_{11}\sin(r) + p_{12}\cos(r) = p_{21}\cos(r) + p_{22}\sin(r)$

so $\tan(r) = (p_{12} - p_{21}) / (p_{11} + p_{22})$

Now, from J we compute the eigenvectors and the eigenvalues which must be

$E_1 = (\cos(\sigma), \sin(\sigma))$ and $\lambda_1 = k\cos(\phi)$

$E_2 = (-\sin(\sigma), \cos(\sigma))$ and $\lambda_2 = k$

(see Figure f6.3.1 again)

Hence $\phi = \cos^{-1}(\lambda_1/\lambda_2)$ and $\sigma = (E_1(1)/E_2(2))$.

Bibliography

Bajcsy, R. Computer identification of textured visual scenes. Technical-report AIM-180, Stanford Artificial Intelligence Laboratory, Stanford University, Oct. 1972.

Bajcsy, R. Computer description of textured surfaces. IJCAI3(1973),572-579.

Bajcsy, R. and Lieberman, L. Texture gradient as a depth cue. Computer Graphics and Image Processing 5,1 (March 1976),52-67.

Caelli, T. and Julesz, B. Perceptual analyzers underlying visual texture discrimination: part I. Biological Cybernetics 28 (1978),167-175.

Carton, E. J.,Weszka, J. S., and Rosenfeld, A. Some basic texture analysis techniques. Technical-report tr-288, Computer Science, University of Maryland, Jan. 1974.

Clowes, M.B. On seeing things. Artificial Intelligence 2,1 (1971),79-116.

Conte, S.D. and deBoor, C. Elementary Numerical Analysis: an Algorithmic Approach. McGraw-Hill, New York, .

Davis, L.S., Johns, S.A., and Aggarwal, J.K. Texture analysis using generalized co-occurrence matrices. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-1,3 (July 1979), 251-259.

Deutsch, E.S. and Belknap, N.J. Texture descriptors using neighborhood information. Computer Graphics and Image Processing 1,2 (Aug. 1972),145-168.

Duda, R.O. and Hart, P.E. Use of the Hough transform to detect lines and curves in pictures. CACM 15,1 (Jan. 1972),11-15.

Dyer, C. R. and Rosenfeld, A. Fourier texture features- suppression of aperture effects. Technical-report tr-391, Computer Science, University of Maryland, July 1975.

Dyer, C.R., Weszka, J.S. and Rosenfeld, A. Experiments in terrain classification on Landsat imagery by texture analysis. Technical-report tr-383, Computer Science, University of Maryland, June 1975a.

Dyer, C.R., Weszka, J.S., and Rosenfeld, A. Further experiments in terrain classification by texture analysis. Technical-report tr-417, Computer Science, University of Maryland, Sept. 1975b.

Dyer, C.R., Weszka, J.S., and Rosenfeld, A. Detection of 'hazy anomalies' in Landsat imagery by texture analysis. Technical-report tr-429, Computer Science, University of Maryland, Dec. 1975c.

Gibson, J.J. The Perception of the Visual World. Houghton Mifflin Co., Boston,Mass., 1950.

Haralick, R.M., Shanmugan, K., and Dinstein, I. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics SMC-3,6 (Nov. 1973),610-621.

Hawkins, J.K. Textural properties for pattern recognition. In (Lipkin,B. and Rosenfeld,A.,Eds.),347-370.

Horn, B.K.P. The Binford-Horn line finder. Memo 285, AI Lab, MIT, July 1971.

Hough, P.V.C. Method and means for recognizing complex patterns. U.S. patent 3,069,654, Dec. 18, 1962.

Huffman, D.A. Impossible objects as nonsense sentences. In Machine Intelligence 6(1971), Meltzer, B., Michie, D. (Eds),295-323.

Julesz, B. Foundations of Cyclopean Perception. University of Chicago Press, Chicago, 1971.

Julesz, B., Gilbert, E.N.,Shepp, L.A. and Frisch, H.L. Inability of humans to discriminate between visual textures that agree in second-order statistics-revisited. Perception 2 (1973),391-405.

Julesz, B. Experiments in the visual perception of textures. Scientific American ,232 (1975),34-43.

Kanade, T. A theory of origami world. Technical-report cmu-cs-78-144, Computer Science, Carnegie-Mellon University, Sept. 20, 1978.

Kelly, M.D. Edge detection by computer using planning. In Machine Intelligence 6(1971), Meltzer, B., Michie, D. (Eds),397-410.

Kender, J.R. Shape from texture: an aggregation transform that maps a class of textures into surface orientation. IJCAI-79,475-480.

Lipkin, B. and Rosenfeld, A. (Eds.) Picture Processing and Psychopictorics. Academic Press, New York, 1970.

Mackworth, A. K. Interpreting pictures of polyhedral scenes. Artificial Intelligence 4,2 (1973),121-137.

Mackworth, A.K. On the interpretation of drawings as three-dimensional scenes. D.Phil. thesis , University of Sussex, 1974.

Mackworth, A. K. Model-driven interpretation in intelligent vision systems. Perception 5 (1976),349-370.

Nevatia, R., Price, K., and Vilnrotter, F. Describing natural textures. IJCAI 79,642-644.

O'Gorman, F. and Clowes, M.B. Finding picture edges through collinearity of feature points. IJCAI3(1973),543-555.

Pickett, R.M. Visual analysis of texture in the detection and recognition of objects. In (Lipkin,B. and Rosenfeld,A.,Eds.),289-308.

Rosenfeld, A., Lee, Y.H., and Thomas, R.B. Edge and curve detection for texture discrimination. In (Lipkin,B. and Rosenfeld,A.,Eds.),381-393.

Rosenfeld, A. Visual texture analysis- an overview. Technical-report tr-406, Computer Science, University of Maryland, Aug. 1975.

Rosenfeld, A. and Lipkin, B.S. Texture synthesis. In (Lipkin,B. and Rosenfeld,A.,Eds.),309-345.

Schatz, B.R. The computation of immediate texture. PhD thesis AIM-426, Department of Artificial Intelligence, MIT, Aug. 1977.

Schatz, B.R. The computation of immediate texture discrimination. Technical-report cmu-cs-78-152, Computer Science Dept., Carnegie-Mellon University, Dec. 1978.

Shirai, Y. A context sensitive line finder for recognition of polyhedra. Artificial Intelligence 4,2 (1973),95-119.

Tamura, H., Mori, S. and Yamawaki, T. Textural features corresponding to visual perception. IEEE Transactions on Systems, Man, and Cybernetics SMC-8,6 (June 1978),460-473.

Tanimoto, S. and Pavlidis, T. A hierarchical data structure for picture processing. Computer Graphics and Image Processing 4,2 (June 1975),104-119.

Tomita, F., Shirai, Y. and Tsuji, S. Descriptions of textures by a structural analysis. IJCAI 79,884-889.

Tomita, F., Yachida, M. and Tsuji, S. Detection of homogeneous regions by structural analysis. IJCAI3(1973),364-371.

Tou, J.T. and Gonzalesz, R.C. Pattern Recognition Principles. Addison-Wesley, Reading,Mass., 1974.

Tsuji, S. and Tomita, F. A structural analyzer for a class of textures. Computer Graphics and Image Processing 2,3/4 (Dec. 1973),216-231.

Waltz, D.L. Understanding line drawings of scenes with shadows. In the Psychology of Computer Vision, Winston, P. (Ed),19-91.

Wang, S.,Velasco, F.R.D., and Rosenfeld, A. A comparison of some simple methods for extracting texture primitives and their effectiveness in texture discrimination. Technical-note tr-759, Computer Science Centre, Univ. of Md., April 1979.

Winston, P.H.(Ed.) The Psychology of Computer Vision. McGraw-Hill, New York, 1975.

Woodham, R.J. A cooperative algorithm for determining surface orientation from a single view. IJCAI5(1977),635-641.

Yachida, M., Ikeda, M. and Tsuji, S. Boundary detection of textured regions. IJCAI 79,992-994.

Zahn, C.T. Graph-theoretic methods for detecting and describing gestalt clusters. IEEE Trans. Computers vol. c-20,no. 1 (1971),68-86.

Zucker, S.W., Rosenfeld, A. and Davis, L.S. Picture segmentation by texture discrimination. IEEE Transactions on Computers C24 (Dec. 1975),1228-1233.

Zucker, Steven W. Toward a model of texture. Computer Graphics and Image Processing 5,2 (June 1976),190-202.