March 1989

Estimation of Textured Surface Inclination by Parallel Local Spectral Analysis

M. R. Turner
University of Pennsylvania

Ruzena Bajcsy
University of Pennsylvania

G. L. Gerstein
University of Pennsylvania

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-89-47.

Estimation of Textured Surface Inclination by Parallel Local Spectral Analysis

Abstract

When an inclined, uniformly textured surface is viewed by an observer or imaged by a camera, the systematic distortions of the perspective transformation will induce a predictable distribution of shifts in the projected spatial frequencies which compose the texture. By measuring these shifts using a set of filters having suitable spatial, frequency, and orientation resolution, the inclination angles of the original textured surface may be estimated. An algorithm is presented which uses the amplitude distributions of 2D Gabor filters to perform such a calculation. Central to the algorithm is a pair of iteratively executed routines. The fist adjusts local sets of parameters to reduce the error between predicted and measured filter amplitudes. The second propagates the local parameters to neighboring regions to consolidate the estimates of inclination. The algorithm is capable of operating in parallel on any number of regions in the image and with a diverse set of filter inputs.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-89-47.

Estimation Of Textured Surface Inclination By Parallel Local Spectral Analysis

MS-CIS-89-47
GRASP LAB 186

M. R. Turner
R Bajcsy
G. L. Gerstein

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389

July 1989

Acknowledgements:
This research was supported in part by Air Force AFOSR-F49620-85-K-0018, US Army grants DAA29-84-K-0061, DAA29-84-9-0027, N00014-85-K-0807, NSF grants MCS-8219196-CER, IRI84-10413-A02, INT85-14199, DMC85-17315, NIH NS-10930-11 as part of the Cerebro Vascular Research Center, NIH 1-R01-NS-236-36-01, NATO grant 0024/85, NASA NAG5-1045, ONR SB-35923-0, DARPA grant N00014-85-K-0018 and by DEC Corp, IBM Corp. and LORD Corp.

Estimation of Textured Surface Inclination by Parallel Local Spectral Analysis

M. R. Turner*† R. Bajcsy*
G. L. Gerstein†

*Department of Computer and Information Science
†Department of Physiology
University of Pennsylvania
Philadelphia, Pennsylvania 19104

March 28, 1989

Abstract

When an inclined, uniformly textured surface is viewed by an observer or imaged by a camera, the systematic distortions of the perspective transformation will induce a predictable distribution of shifts in the projected spatial frequencies which compose the texture. By measuring these shifts using a set of filters having suitable spatial, frequency, and orientation resolution, the inclination angles of the original textured surface may be estimated. An algorithm is presented which uses the amplitude distributions of 2D Gabor filters to perform such a calculation. Central to the algorithm is a pair of iteratively executed routines. The first adjusts local sets of parameters to reduce the error between predicted and measured filter amplitudes. The second propagates the local parameters to neighboring regions to consolidate the estimates of inclination. The algorithm is capable of operating in parallel on any number of regions in the image and with a diverse set of filter inputs.

1 Introduction

Gradients and transitions in visual textures can provide a number of cues for the analysis of objects and surfaces in the visual world. The image of an inclined, uniformly textured surface will exhibit systematic variations in the projected size, shape and density of elements which compose the texture. Unlike segmentation, in which boundaries between objects are denoted by abrupt textural transitions, the distortions associated with surface inclination are more gradual and smoothly varying across the image. With appropriate means of measurement, they may be used to estimate several three dimensional shape properties of the object, one of which is the angle of inclination of the surface relative to the viewer.

Although a precise definition of texture remains elusive, texture has been described as a two dimensional distribution of elements having a kind of spatial shift invariance (Levine, 1985) in that any patch of the texture perceptually resembles any other similarly sized patch. The elements making up the texture (or texels) may be highly structured (as in a brick wall) or irregular (as in patterns of foam on the sea). In either case, the perspective imaging process will affect the projection of textural patterns.

Two types of systematic distortion are introduced into the image of an inclined textured surface by perspective. The first occurs because points at different locations on the textured surface are different distances from the viewer. This is simply a distance effect: those areas farther away will project to smaller areas in the image plane resulting in a decrease in projected texture element size and an increase in projected element density. The second effect results from differences in angle between the textured surface and the line of sight to the viewer. As the inclined textured surface recedes into the distance, the angle between the line of sight and the surface tangent becomes more acute. This causes an anisotropic compression of apparent distance along the direction of maximum inclination (i.e. along the tilt) which distorts or "foreshortens" the shape of texture elements and increases their density along this dimension. Perspective distortions result from a combination of these two separate effects.

Accordingly, a uniformly textured inclined planar surface will exhibit textural compression and density changes which vary when measured at different orientations and positions in the image plane. If we take measurements at a number of locations with an orientation across the tilt of the surface in

the image (perpendicular to the angle along which the surface recedes most rapidly from the viewer), we will detect linear compressions of texture as a function of location caused by the distance effect. If we measure at a number of locations with an orientation along the tilt, then the distance and foreshortening effects combine and we will detect quadratic shape and size distortions of the texture.

Gibson (1950a) used the term "texture gradient" to describe these systematic variations and proposed that perception of surfaces is due to gradients of textural or other physical features. He suggested that such gradients are perceived by simultaneous variation over a set of different receptors, or differential excitation of different receptors. The work presented in this paper is in many ways faithful to the spirit of Gibson's conjecture: we make estimates of surface inclination by analysis of the output of a number of filters which have the same mathematical form as those found to be an accurate model for simple cells in the visual cortex.

One of the strongest controversies in studies of the visual system has been over the preferred elemental stimulus and "tuning curves" of cortical neurons. At one extreme, the elemental stimulus is a long bar, and tuning is to its orientation (Hubel and Wiesel, 1959, 1962, 1974). At the other extreme, the elemental stimulus is a grating, and tuning is with respect to its spatial frequency (Campbell and Robson, 1968; Albrecht et al., 1979; Tootell et al., 1981; DeValois et al., 1979). In the recent past a possible compromise has appeared with the revival of interest in the functions first proposed by Dennis Gabor (Gabor, 1946; Daugman, 1980, 1983, 1985a; Marcelja, 1981; Mackay, 1981). These functions have the attribute of maximizing the simultaneous information about two different stimulus properties, an idea analogous to the uncertainty principle from physics, but applied to information. Daugman (1985a) has explicitly delineated the necessary formalism in extending the idea to the two dimensional receptive field of a visual neuron. At the same time, Jones and Palmer (1987a, 1987b, Jones et al., 1987) have shown that the two dimensional receptive field of many simple cells in area 17 of the cat may be accurately and economically represented by the same 2D Gabor functions. Thus it would seem that 2D Gabor functions are highly appropriate candidates for the filtering operators that do early (in the cognitive sense) analysis of the visual scene.

A 2D Gabor function may be visualized as the product of a sinusoidal plane wave (grating) of some frequency and orientation with a two dimen-

sional elliptical (circular in the filters used in this paper) Gaussian (see figure 1). They afford a means of making visual measurements which are spatially as well as frequency and orientation selective and, under certain norms, optimize the tradeoff in resolution which inevitably results from the uncertainty relation using a linear filter in these two domains (Daugman, 1985a). The kind of local spectral analysis afforded by filters of this type have been shown to be applicable to a number of visual processing tasks. Several authors have shown that such local spectral measures are conducive to the segmentation of images by textural differences, particularly those textures which have been found to be preattentively (rapidly) discriminable by humans (Turner, 1985, 1986; Gerstein and Turner, 1987; Daugman, 1985b; Clark et al., 1987). It has also been demonstrated that a number of interesting information compaction properties accrue from a Gabor representation (Daugman, 1988; Porat and Zeevi, 1988). Other authors (Adelson and Bergen, 1985; Heeger, 1987, 1988) have shown a temporal extension of Gabor functions to be good primitive operators for early motion analysis. Orientation selective receptive fields much like Gabor functions can play a role in detecting shape from shading (Lehky and Sejnowski, 1988). It is likely, then, that such local spectral receptive fields may have evolved because this representation of visual information may serve a number of different computational functions.

The systematic spatial distortions of size and density under perspective must necessarily affect the projection of the spatial frequencies which comprise the textured surface. Therefore, the inclination of a textured surface may be recovered from the unique and predictable frequency shifts which occur at different locations in its perspective image. However, there are two conditions necessary for any vision system to utilize these shifts in estimating surface inclination. First, the system must have adequate means of measuring the frequency components of the projected texture in local image regions. The manner in which the frequency components vary from region to region depends upon the direction of tilt of the surface and, because of the anisotropic nature of the distortions, the orientation of the spectral measurement. Consequently, the early representation of image information must possess a reasonable degree of spatial, frequency, and orientation selectivity. There are a number of filter forms which could satisfy these simultaneous requirements. The 2D Gabor function is one member of this family, selected for this research because of its physiological relevance.

The second requirement pertains to the textured surface itself. The texture must be appropriate in several ways for the accurate determination of the frequency shifts. In part, this requires that the spectral signature of the texture be sufficiently uniform that variations measured from region to region derive largely from perspective distortions and not from systematic irregularities in the texture or other influences (such as occlusion or shadowing of texels which project out of the texture surface). As discussed later in this paper and in greater depth in another paper (Turner et al., 1989a), these are necessary but not sufficient conditions for the accurate estimation of surface inclination.

Local spectral analysis of texture gradients was first developed by Bajcsy and Lieberman (1976) using windowed Fourier transforms. Utilizing algorithms formulated by Bajcsy (1972, 1973), they identified frequency shifts in spectral components occurring as a result of perspective projection of textured surfaces in natural scenes. The work described in this paper is in some ways similar. However, the goals of this work go beyond development of computational techniques appropriate only to machine vision. By using a low level filter similar to the receptive field of cortical simple cells and by constraining the higher level processes to local and parallel computational steps, the system developed in this paper may provide insights into problems and solutions common to both biological and machine vision. As an example, it has been found that the program described later in this paper replicates errors of slant angle underestimation with "irregular" textures that humans make in psychophysical experiments (Gibson, 1950b; Gruber and Clark, 1956; Flock and Moscatelli, 1964). Moreover, efforts to realize this system in a simulated neural network architecture have led to the postulation of a new, higher level receptive field organization for biological vision systems which has not yet been described in the physiological literature.

This paper specifically addresses the computational aspects of the problem and develops a machine vision algorithm by which textured surface inclination may be estimated from local spectral measurements. It differs from most other computer based techniques in that it does not require the prior identification of texels (Blostein and Ahuja, 1987; Aloimonos, 1988), a difficult task in natural environments, nor does it make strong assumptions about the isotropy of edge orientations (Witkin, 1981; Kanatani, 1984) or the regularity of known texture elements (Ikeuchi, 1984) on the planar surface. The paper is organized as follows. After introducing the the perspective imaging

model and filter form, section 4 describes the analytic relationship between the perspective projection of texture frequencies and the amplitude distributions of Gabor functions. Section 5 discusses the complications arising from the application of this model to naturally occurring textured surfaces and presents a parallel algorithm which recovers surface inclination using estimates obtained from local image patches. Section 6 describes a computer implementation of the method and presents results from a variety of different textured surfaces. The final section discusses certain limitations of the approach, areas for future investigation, and relevance of the model for biological vision systems.

2 Perspective Imaging Model

The imaging model, shown in figure 2, consists of an inclined textured plane a distance d from a pinhole camera with a flat image plane and a focal length f. Although this model is an oversimplification of that occurring in a lensed camera or an eye (with a curved retina), these simplifications make the situation considerably more tractable by eliminating such complications as depth of field, yet the general applicability of the results to other systems is not significantly reduced.

Imagine that the textured plane is a uniform sine wave grating of some frequency and orientation and that the plane has an oblique inclination relative to the viewer. As seen in figure 3, the perspective distortions vary the projected frequency from region to region in the image.

3 Gabor Filters

The filters used in this work are discrete realizations of the following function:

$$G_\phi(x,y) = \exp\left(-\frac{(x-x_c)^2+(y-y_c)^2}{2\sigma^2}\right)\sin\omega(x\cos\alpha - y\sin\alpha + \phi) \quad (1)$$

where

- x_c and y_c are the center locations of the Gabor filter Gaussian envelope

- σ is the standard deviation of the Gaussian envelope.
- ω is the frequency of the sinusoidal plane wave.
- α is the orientation of the plane wave.
- ϕ is its phase.

The characteristics and selectivities of a filter are determined by these variables with ω and α the frequency – orientation peaks, and σ the spatial size of the Gaussian (and thereby also the frequency bandwidth). For simplicity, only circularly symmetric Gaussians have been used in this work. The value of ϕ determines the phase of the filter. For each frequency – orientation, a pair of filters is generated differing in spatial phase by 90 degrees (approximately quadrature). With a number of different frequency, orientation and spatial envelope combinations a set of filters as shown in figure 1 is generated.

The 2D Gabor functions are applied to an image in a set of overlapping regions (much like receptive fields). A filter is applied to a region (centered at (x_c, y_c)) by taking the inner product of the filter with the image material falling within the region.

$$V_\phi(I, G_\phi) = \sum_x \sum_y I(x, y)\, G_\phi(x, y) \tag{2}$$

From each phase pair of filters G at each location (x_c, y_c) we get 2 values. An amplitude A is computed as:

$$A(I, G_\phi) = \sqrt{V^2_{\phi_1}(I, G_\phi) + V^2_{\phi_2}(I, G_\phi)} \tag{3}$$

where V_{ϕ_1} and V_{ϕ_2} are the values computed from the same image region using the members of G which differ in phase. (Further details are available in Turner, 1986.)

A set of Gabor functions applied to the perspective image of the inclined grating produce outputs which vary from location to location depending upon the local match between the projected spectral value and the frequency – orientation of the Gabor filter pair. Figure 4 shows the output of one particular filter pair at a number of different regions in the projected image.

The amplitudes measured in the upper band of the image (where the projected frequencies and orientations more nearly match those of this filter) are considerably higher than those of the bottom and top right of the image.

Application of a number of Gabor functions to a grating image or to the projected image of a textured surface produces a two dimensional amplitude distribution for each filter specification in the set. From these, the algorithm attempts to estimate the slant – tilt values consistent with the distributions. In order to do so, however, the relationship between amplitude distributions and surface inclination must be described.

4 Gabor Filter Amplitudes and Texture Frequency Gradients

In this section an equation is derived for estimating the output or amplitude of a 2D Gabor filter pair applied to the perspective projection of the frequency component of a textured planar surface. The equation takes two sets of parameters as arguments. The first contains those values which are part of the imaging and measuring system and are, therefore, known quantities. This set includes the specifications of the Gabor filter pair and the location within the image plane to which the filters are applied. It also contains the relevant properties of the imaging system, in this case the focal length of the pinhole camera. The second set contains values associated with the textured surface in the real world. In most situations, these are the variables which must be determined by the algorithm. The two most important of these are, of course, the representation of slant and tilt of the surface. In order to evaluate these, however, the algorithm must also determine the texture frequency being projected (and measured by a particular Gabor filter pair). The algorithm makes these determinations by adjusting the second set of parameters to reduce the error between a set of amplitudes measured from the image and the estimates of the equation derived here.

4.1 Encoding of Slant – Tilt

Stevens (1983a, 1983b) has presented a method for encoding slant and tilt (ς, τ) which is in common use for texture gradient studies. In his scheme the slant of a textured plane is the maximum angle between the texture plane

and the image plane. The tilt is the angle between the projection of the texture plane surface normal and the x axis on the image plane. In other words, the tilt is the angle in the image plane along which the texture plane recedes most rapidly from the viewer while the slant is related to the rate at which the plane recedes. The slant angle may vary between 0 and 90 degrees while the tilt varies between 0 and 360 degrees (with 90 degrees representing a texture surface farthest away from the viewer at the top of the image).

While this slant – tilt representation is both intuitively simple and capable of describing any surface inclination, a slightly different method is used within the algorithm. Instead of slant – tilt the algorithm uses a measure of vertical slant and horizontal slant. Both the image plane and the texture plane are described by Cartesian coordinate systems with (x, y) representing points in the image plane and (u, v) representing points in the texture plane. The vertical slant (θ) is the angle between the positive vertical axis on texture plane and the positive vertical axis in the image plane. The horizontal slant (ψ) is similarly defined.

The inclination of a textured surface is constructively defined by the sequence of coordinate transformations necessary to orient an imaginary textured surface parallel to the image plane (perpendicular to the line of sight) with that of the actual textured surface being imaged. The vertical slant transformation is first applied by rotating the plane around its horizontal axis by the vertical slant angle (θ). This is followed by a rotation around a vertical axis (through the origin of the texture plane) by the angle of its horizontal slant (ψ). (All rotations are clockwise when looking toward the origin from the positive side of the axis.) This has the advantage of favoring the horizontal axis in that points on the u axis ($v = 0$) of the texture plane always project to points on the x axis in the image plane. Using this system internally, an algorithmic adjustment to the vertical slant will have limited effect on horizontal frequency estimates. For low values of vertical slant, this same advantage applies to vertical frequency estimates with changes in horizontal slant. Using Stevens' method internally would involve rotating the coordinate systems relative to each other (necessitating a corresponding change in both horizontal and vertical frequency estimates) with any change in tilt estimate.

The horizontal and vertical slant encoding may be transformed into Stevens' slant – tilt form using

$$\varsigma = \sin^{-1}\sqrt{\sin^2\psi + \sin^2\theta\cos^2\psi} \quad (4)$$

$$\tau = \tan^{-1}\frac{-\sin\theta\cos\psi}{\sin\psi} \quad (5)$$

4.2 Estimation of Filter Amplitude

The derivation proceeds in three parts. First a projection function is presented describing the mapping from image coordinates to those on the textured plane with the perspective transformation model. Second, the gradient of the projection function is used to determine the perspective distortions of spectral wavelengths which result from the imaging process. Third, the Fourier transform of a Gabor filter is used to estimate the output which results from using a filter pair to measure the projected spectra in the image plane.

4.2.1 Projection Function

As described above, the projection of points on the texture plane to the image plane is obtained by a sequence of coordinate transformations involving rotations and perspective, giving us

$$x = \frac{fu\cos\psi + fv\sin\theta\sin\psi}{d - u\sin\psi + v\sin\theta\cos\psi} \quad (6)$$

$$y = \frac{fv\cos\theta}{d - u\sin\psi + v\sin\theta\cos\psi} \quad (7)$$

Inverting these for (u, v) gives

$$u = \frac{xd\cos\theta - yd\sin\theta\sin\psi}{f\cos\theta\cos\psi - y\sin\theta + x\cos\theta\sin\psi} \quad (8)$$

$$v = \frac{yd\cos\psi}{f\cos\theta\cos\psi - y\sin\theta + x\cos\theta\sin\psi} \quad (9)$$

4.2.2 Projection of Texture Spectra

The frequency projected to the image plane at the point (x, y) is the product of the horizontal and vertical components of the texture frequency (F_u, F_v) and the Jacobian of the projection function

$$\begin{bmatrix} F_x & F_y \end{bmatrix} = \begin{bmatrix} F_u & F_v \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} \tag{10}$$

Where

$$\frac{\partial u}{\partial x} = \frac{fd\cos^2\theta\cos\psi - yd\sin\theta\cos\theta\cos^2\psi}{(f\cos\theta\cos\psi - y\sin\theta + x\cos\theta\sin\psi)^2} \tag{11}$$

$$\frac{\partial u}{\partial y} = \frac{-fd\sin\theta\cos\theta\sin\psi\cos\psi + xd\sin\theta\cos\theta\cos^2\psi}{(f\cos\theta\cos\psi - y\sin\theta + x\cos\theta\sin\psi)^2} \tag{12}$$

$$\frac{\partial v}{\partial x} = \frac{-yd\cos\theta\sin\psi\cos\psi}{(f\cos\theta\cos\psi - y\sin\theta + x\cos\theta\sin\psi)^2} \tag{13}$$

$$\frac{\partial v}{\partial y} = \frac{fd\cos\theta\cos^2\psi + xd\cos\theta\sin\psi\cos\psi}{(f\cos\theta\cos\psi - y\sin\theta + x\cos\theta\sin\psi)^2} \tag{14}$$

4.2.3 Gabor Filter Amplitude

With the projected texture frequency (F_x, F_y), the amplitude of a filter pair may be calculated using the Fourier transform of the Gabor function resulting in

$$A' = C2\pi\sigma^2 \exp\left\{-2\pi^2\sigma^2\left[(g_x - F_x)^2 + (g_y - F_y)^2\right]\right\} \tag{15}$$

where

- C is the contrast of the texture frequency.
- σ is the standard deviation of the Gaussian envelope of the 2D Gabor filter pair.
- g_x and g_y are the horizontal and vertical components of the Gabor filter modulating frequency.
- F_x and F_y are the horizontal and vertical frequency components from equation (10) projected to the image plane.

Whereas the projected texture frequency described in equation (10) is a point function, the 2D Gabor filter effectively occupies a finite spatial area. The amplitude of the filter pair is described by equation (15) as if derived from a uniform frequency at all image points. However, for non-perpendicular textured surface inclinations, the projected frequency varies smoothly across the image. Thus, for function (15) to accurately approximate those amplitudes actually measured from an image, the spatial size of the filters must be kept reasonably small to minimize the frequency variation over the effective area of the Gabor filter. In its present form, then, equation (15) is an accurate approximation for filters of moderate bandwidth. For the examples in this paper, the filter bandwidth has been kept constant at 1.5 octaves. This is well within the range of physiological estimates (Kulikowski et al., 1982). To maintain constant bandwidth, σ varies with the frequency, making all filters scalings and rotations of each other (Mallat, 1988; Daugman, 1988).

With equation (15) and information about the texture being imaged, we could predict the amplitude distributions obtainable in the image plane using a particular filter pair. However, the problem we are trying to solve is the inverse one. The algorithm *extracts* sets of amplitude distributions from an image using a known set of 2D Gabor functions and from these values must make surface inclination estimates. This process is complicated by the fact that many natural textures do not have perfectly uniform or simple spectral signatures. Such textures are often composed of a number of frequency components which, due to textural irregularities, vary the magnitude of local spectral measurements from region to region. In addition, there are often strong contrast gradients which can dominate the distributions of Gabor filter amplitudes, thereby masking the frequency induced gradients critical for the estimation process. Thus, the algorithm must attempt to utilize a potentially variable and misleading source of information for the recovery of surface inclination.

5 Description of Algorithm

An algorithm which utilizes the preceding model is described in the following sections. Its major components are shown in the diagram of figure 5. The image is initially transformed by a point logarithmic function (described below) to compress its luminance dynamic range, thereby reducing the effect of contrast gradients. To this transformed image a set of Gabor filters are applied to extract the amplitudes from which the inclination estimates are made. The heart of the algorithm is the parallel application of a pair of operations which are iteratively executed to form and consolidate the slant – tilt estimates at a number of locations in the image.

In the first of this pair the algorithm attempts to fit parameter sets (including horizontal and vertical slant) to each of a number of local patches of the image by an error reduction procedure. In the second the local parameter estimates are propagated to neighboring regions for consolidation. This phase ensures that variations in the local spectra of the texture will not result in unrealistically extreme local estimates. After completing the second phase the algorithm repeats the sequence of phases until termination conditions are met.

This portion of the algorithm may be viewed as operating in parallel on a number of processors, each associated with one of the 2D Gabor functions applied to the image. Therefore, each processor is associated with a specific spatial location in the image and with a specific spectral value (those of the Gabor function). In the extreme case one processor could be created for each Gabor filter frequency and orientation specification at each image location to which the filter is applied. (However, such a calculation would run very slowly on a sequential architecture.) In making its slant – tilt estimates, each processor uses the outputs of the filter pair with which it is directly associated as well as those neighboring filters in the image with similar frequencies and orientations. Each processor communicates with other processors in a surrounding spatial neighborhood.

Each of the sections of the algorithm is now described in further detail.

5.1 Reducing the Influence of Contrast Gradients

Changes in contrast across an image source will induce gradients in the Gabor filter amplitudes which may mask or distort those induced by the perspective

projections of texture frequencies. These combined effects, common in outdoors or natural scenes with nonuniform levels of illumination or reflectance, would result in inaccurate estimates of surface slant and tilt if not corrected by other means. Two general strategies are appropriate to counteract their effects. In the first, the filter values are "normalized" by using ratios of single filter outputs to local ensemble sums (eg. take the ratio of a single filter output to the sum of outputs of filters of all frequencies and/or orientations at the same spatial location in the image). This technique is similar to that effectively used in spatiotemporal filter studies for motion detection (Adelson and Bergen, 1985). In the second technique compression of the luminance dynamic range is achieved by application of a point nonlinear function (such as log) to the image before the filters are applied (Levine, 1985). This kind of compressive nonlinear transformation is also consistent with both physiological and psychophysical data (Fuortes, 1958, 1959; Rushton, 1961; Hurvich and Jameson, 1966; Stockham, 1972).

Either technique would be equally effective for this application; for the examples shown in this paper the second has been selected. A new image I' is created by transforming the gray level $I(x, y)$ of each location x, y in source image I by

$$I'(x, y) = \ln (I(x, y) + 1) \tag{16}$$

This operation serves to compress contrast gradients in I' while leaving the frequency gradients relatively unaffected. It is to this transformed image, then, that the Gabor filters are applied.

5.2 Generation and Application of Gabor Filters

The filters used by the algorithm are discrete realizations of equation (1). While the frequency – orientation selections for the filter set must ensure that the projected gradients of strong texture frequencies are adequately measured for the iterative sections, the algorithm requires neither completeness nor orthogonality for its operation. A wider frequency span of filters (as might exist in biological vision systems) would allow the same system to perform with a larger variety of stimuli by ensuring that some subset of the filters would be tuned to projected spectral peaks. However, it has been found that even fairly small filter sets can perform correctly with a diverse variety of

texture types. (For example, the same set of 48 filters was used on all the image examples shown in this paper.) For reasons which will be described further in the next section, incorrect minima in the error surface used in the iterative sections are eliminated by overconstraining estimates to satisfy several filters which are closely tuned in frequency. Consequently, the filter set will contain a closer spacing of frequencies than might be required for other visual processing tasks.

The algorithm is capable of compensating for any arrangement of spacing positions (even random) and sampling density has not been found in practice to be a critical variable for this application. Filters used in the examples in this paper have been spaced horizontally and vertically at a distance which results in 90 overlapping regions per dimension.

5.3 Local Parameter Adjustment

Each processor maintains an internal set of parameters for texture frequency, horizontal and vertical slant. By evaluating equation (15) for the specifications of each Gabor filter pair within its spatial and spectral neighborhood (defined below), an expected distribution of outputs is generated from the current set of parameters. These are then compared with the actual values measured from the image by the Gabor filter pairs and an error is computed. Test sets of parameters are generated by taking combinations of positive and negative steps from the current parameter set. The algorithm computes an error for each of the test sets and adopts the set with the lowest error as the new estimate for this processor. Each processor, then, performs a hill climb, evaluating sets of parameters around the current set and always choosing the new set in the direction of lowest error. This operation is performed for all processors in the system.

Equation (15) has five unknown variables: two slant terms, two texture frequency terms and a contrast term. It has been found that the system attains a solution more quickly if the contrast term is eliminated by taking a ratio of sums and differences of amplitudes. For the Gabor amplitudes taken from the image:

$$R(c,r) = \frac{A_c - A_r}{A_c + A_r} \tag{17}$$

where

- A_c is the Gabor amplitude of the filter directly associated with the processor.
- A_r is any other Gabor amplitude within the processor's spatial and spectral neighborhood.

A similar ratio is computed from a parameter estimate using (15):

$$R'(c,r) = \frac{A'_c - A'_r}{A'_c + A'_r} \tag{18}$$

From (17) and (18) a squared error is computed

$$E(c) = \frac{\sum_r W(c,r)\left(R(c,r) - R'(c,r)\right)^2}{\sum_r W(c,r)} \tag{19}$$

The weight $W(c,r)$ defines the spatial and spectral neighborhood of the processor. It determines which filter amplitudes are used in computing the error and with what importance. As will be shown later, the choice of weighting function has a considerable bearing on the shape of the error surface used in the hill climb operation.

The weighting function used in the current implementations of the algorithm is described below. It gives greater influence in the error evaluation to outputs from filter pairs which are both spatially near each other in the image and similar in frequency and orientation. It is formed as the product of two terms.

$$W(c,r) = D\left[(x_c, y_c), (x_r, y_r)\right] S(G_c, G_r) \tag{20}$$

D is a spatial weighting function (a Gaussian in this version of the algorithm) which gives greater influence to values from filters which are spatially near to each other.

$$D\left[(x_c, y_c), (x_r, y_r)\right] = \exp\left(-\frac{\sqrt{(x_c - x_r)^2 + (y_c - y_r)^2}}{2\sigma_W^2}\right) \tag{21}$$

- (x_c, y_c) is the image location of the Gabor filter associated with the processor.

- (x_r, y_r) is the image location of another filter pair used in the error evaluation.

- σ_W is the standard deviation of the Gaussian spatial weighting function.

S is a measure of filter similarity in frequency and orientation. In the current algorithm it is formed as an inner product amplitude between the different filter pairs. Equation (3) computes just such an inner product amplitude for regions of the image and, therefore, may be used in the description of S. Substituting the phase pair members of G_r in place of the image and centering them at the same spatial location (x_c, y_c) as G_c:

$$S(G_c, G_r) = \frac{A(G_c, G_{r\phi_1}) + A(G_c, G_{r\phi_2})}{2} \tag{22}$$

in which $G_{r\phi_1}$ and $G_{r\phi_2}$ are the members of G_r differing in phase by 90 degrees.

Equation (20) may be thought of as a rough measure of distance or overlap between filters in a four-dimensional space (with two spatial and two spectral dimensions) which gives the algorithm an indication of the likelihood that amplitudes in two filters derive from the same frequencies in the texture. It also produces an interconnection "granularity" in this four-dimensional space in that it defines and limits the effective or required connectivity between processors. While it is probable that a number of other functions could serve this purpose, certain tradeoffs constraining the choice of weighting function have been identified and these are described below.

5.4 Tradeoffs in the Choice of Weighting Function

5.4.1 Wide Spectral and Spatial Support

Since equation (15) has four variables, the error space is four dimensional. In order to provide a visual representation of the surface a two dimensional slice may be taken by adjusting the frequency parameters with the horizontal and vertical slant values to show the lowest possible error for each slant pair. Figure 6 shows such a slice using a sine wave grating slanted 30 degrees. For this example a weighting function has been selected which uses the amplitudes from only *one* filter pair frequency – orientation specification on a

small patch in the center of the image. The horizontal and vertical slant parameters vary along the x and y axes respectively. Each point on the surface represents the lowest possible error for this combination of horizontal and vertical slant. This error value is represented as the height of the peak. (The higher the peak the lower the error.) For one filter pair there are two peaks of very similar size. The one in back represents the correct inclination of 30 degrees slant. The second peak is the reverse of this inclination (inclined in the reverse direction from perpendicular by an angle similar to the correct inclination).

A grating which recedes from the viewer at the top projects frequencies in the image plane which increase from bottom to top. A grating inclined toward the viewer at the top has these projected frequencies reversed. With its symmetrical (Gaussian) frequency response (shown in one dimensional form in figure 7), a rising gradient of amplitudes may represent either an increasing frequency gradient (inclined away from viewer at the top) on the low frequency side of the Gabor filter's response or a decreasing frequency gradient (inclined toward viewer) on the high frequency side of the filter's response. As evident in the error surface, both situations can produce an accurate fit to the amplitude distributions. If the algorithm makes the wrong choice of texture frequency in its initial hill climb steps then it will likely select the error minimum representing slant values which are the reverse of the correct inclination. It becomes trapped in the reversed inclination minimum. (It is interesting that one of the subjects in Gibson's 1950 psychophysical study repeatedly perceived the slant of one textured surface as reversed in exactly this manner. p. 379)

This situation occurs in other sensory modalities of biological organisms (Erickson, 1963) in which one sensor provides ambiguous information for the solution of a problem. Adding the constraint that several sensors of slightly different response must be simultaneously satisfied frequently disambiguates the problem. In this case, local estimates may be constrained to satisfy several filters of similar but not identical frequency and orientation. The addition of another filter having spectral overlap with the first (figure 7d) provides a clear means to distinguish the two cases. Each frequency value is now associated with a *unique* set of amplitude values. Moreover, the amplitude distributions of the second filter in both forward and backward inclination cases are sufficiently different to ensure that the reversed inclination estimate is associated with a large error. By using a weighting function

which includes different filter distributions, the error surface is considerably changed, affording the system a more rapid identification of the projected texture frequency and surface inclination. In figure 8 an error surface like that of figure 6 is shown. In this case, however, the weighting function has been altered to include a set of three filters with considerable overlap in frequency – orientation. The peak representing the correct inclination values is considerably sharper than it was in the single filter surface and the reversed inclination peak is eliminated.

Irregularities in small patches of the textured surface may induce amplitude distributions which are consistent with extreme inclination values. By including amplitudes from a larger portion of the image the influence of these local irregularities are reduced and the estimates of each processor are more likely to be consistent over the entire planar surface. For these reasons it would seem that the weighting function should be selected to provide a wide spatial and spectral support to both improve the accuracy and speed of the hill climb (by changing the error surface) and to reduce the effect of textural irregularities on local inclination estimates.

5.4.2 Narrow Spectral and Spatial Support

The problem with a weighting function that admits wide spatial and spectral support is that, while it works quite well with synthetic or natural textures having one dominant frequency (like a sine wave grating), the world is filled with textures of considerably greater spectral complexity. Naturally occurring textures are frequently composed of a number of spectral components of similar amplitude which can be closely spaced in frequency and orientation. Using a weighting function with wide spectral support broadens the range of frequencies over which the error term is computed. The wider the frequency – orientation range of the Gabor filter pairs used in computing an error, the more likely the amplitude distributions will derive from different spectral components of the texture. This makes the algorithm more sensitive to the presence of multiple spectral peaks in the texture, even when they are spaced sufficiently widely to have minimal effect on local amplitude distributions of single filter types taken alone. The resulting effect is a distortion of the error surface which introduces inaccuracies (particularly those of surface angle underestimation) into the parameter evaluation.

Similarly, use of amplitudes from spatially distant filter pairs is more likely

to be inaccurate than from local spatial neighborhoods. Different texture frequencies may project to the same image frequency in spatially separated locations of the image. This can distort the amplitude distributions and gradients in ways that introduce further inaccuracies into the processor's parameter estimates.

The tradeoff in spatial and spectral support depends, in large measure, upon the characteristics of the textured surface being imaged. Very regular, periodic textures (many manmade surfaces, for example) allow rapid evaluation of surface inclination by the inclusion of a wide variety of spatial and spectral inputs to the error evaluation. On the other hand, complex natural textures necessitate the narrowing of inputs which has the inevitable effect of making the local estimates more sensitive to the irregularities accompanying these textures and, therefore, more likely to be in error. In this situation the identification of structure in the distribution of local estimates must be performed in part at the higher parameter level rather than the level of the Gabor amplitudes. The weighting function described in this paper represents a compromise which has proven satisfactory on a wide variety of textured materials. Nevertheless, the diversity of texture characteristics argues for an adaptable system which both adjusts the weighting span of input terms for different textures by analysis of filter amplitude statistics, and determines the optimal level upon which to evaluate surface structure. Investigations into such adaptive systems are appropriate for the future.

5.5 Lateral Parameter Propagation

Texture variations frequently produce local output distributions consistent with extreme inclination values. As described in the previous section, however, errors may result from the obvious solution of increasing the spatial span of amplitudes from the image for parameter adjustment. Therefore, the parameter adjustment phase is followed by one in which distribution of parameters between neighboring processors is used to adjust and consolidate the parameter values. Each processor is connected to other processors in a spatial neighborhood around it. At each iteration the values are received from its neighbors and a weighted average for each parameter is computed. This weighted average becomes the new parameter value for the next execution of the local parameter adjustment step.

This weighting function has two parts. First, greater importance is given

to those estimates with lower error residuals (eq. 19). Second, for slant estimates, greater importance is given to parameters from nearby processors (eq. 21). For frequency estimates, spectral similarity weighting between filters is also included (eq. 22).

$$P_{n+1}(t,c) = \begin{cases} \frac{\sum_r [D(c,r)P_n(t,r)]/[E(r)]}{\sum_r [D(c,r)]/[E(r)]} & \text{if } t \epsilon \{\theta, \psi\} \\ \\ \frac{\sum_r [D(c,r)S(c,r)P_n(t,r)]/[E(r)]}{\sum_r [D(c,r)S(c,r)]/[E(r)]} & \text{if } t \epsilon \{F_u, F_v\} \end{cases} \tag{23}$$

where t is the parameter type, P_n is the current parameter value, P_{n+1} is the new parameter average. In some ways, this is like a relaxation labeling process (Rosenfeld et al., 1976; Hummel and Zucker, 1983; Landy, 1987). Instead of using a wide spatial and spectral span of inputs for each evaluation, the accuracy and reliability of inclination estimates are increased by lateral propagation of parameter values, essentially increasing the span of influence by label consolidation on the higher parameter level rather than at the low level of Gabor amplitudes.

6 Implementation Details and Results

This algorithm has been implemented in a set of programs which run on a Sun 4/260. The images of textured surfaces are 512 x 512 pixels with pixel grayscales varying between 0 and 255. The set of 48 2D Gabor filters (figure 1) used with these images contain 6 frequencies (8, 10, 12, 14, 16, 18 pixel wavelengths) at 4 orientations (0, 45, 90, 135 degrees) with 2 phase pairs (0 and 90 degrees) for each frequency – orientation combination. The spatial size for each frequency choice is adjusted to give each filter a frequency bandwidth of 1.5 octaves. With the exception of pure sine phase, Gabor filters have a slight DC component which makes them sensitive to the phase of low texture frequencies. To prevent this from interfering with frequency gradients each filter has been adjusted by subtracting its mean pixel value from each pixel in its filter array. (This is discussed in Turner, 1986.) The filters are applied to the images in a set of overlapping regions separated by 5 pixels along each dimension.

6.1 Reduction of Gabor filter outputs

In order to improve the efficiency of the simulation on this sequential architecture, the number of Gabor filter outputs used by the algorithm is reduced by averaging the amplitude of Gabor filters in a number of regions of the image. This averaging or smoothing step also serves to reduce the effect of local texture variations on the filter output (and, thereby, inclination) estimates. An image is divided into 36 regions (6 x 6). Around each region a Gaussian weighted average (of 25 pixels std. dev.) is made of the output of each filter type. For each filter pair frequency and orientation value, this step produces 36 averaged output values for the image. These averaged output values are then used in place of the raw filter pair amplitudes. While this averaging process induces some error into the filter output gradients, in practice this has not been found to significantly alter the accuracy of the results.

6.2 Thresholding and Processor Creation

The number of amplitude values used in fitting the parameters is further reduced by a thresholding operation. Particular filters may not be sensitive to the spectral components of the imaged texture throughout the image plane. Their outputs will tend to be of low and relatively uniform amplitude, supplying little information about the gradient. (For example, in figure 4, the bright (high amplitude) region for this filter only occupies a band in the upper part of the image. The other sections of this image induce low amplitudes in this filter pair.) Moreover, fluctuations in these low amplitude values caused by local variations in the texture, if treated with the same importance as gradients in the higher amplitude values, would tend to introduce inaccuracies into the local slant estimates.

The selection process has been accomplished by a thresholding operation which identifies the highest percentage (20 percent for the examples shown in this paper) of averaged Gabor amplitudes. A processor is assigned to the spatial location, frequency and orientation of each such high amplitude Gabor filter pair. This serves to conserve computational resources for those filters and image regions which contain the most information about the texture gradient. (This is in many ways similar to the information compaction procedure used by Daugman (1988) in which a recognizable image is constructed from only the Gabor filters with the highest output amplitudes.)

The choice of percentage has not been found to be a critical parameter and a variety of other techniques would work as well here. In a truly parallel architecture a more appropriate solution would be to modify the lateral parameter propagation sections described previously to place greater emphasis upon the higher amplitude values. The combined averaging and thresholding techniques described in these sections reduce the number of amplitudes used in the iterative portions of the algorithm from the original 194,400 extracted from the image to 173.

6.3 Initial and Termination Conditions

Parameters for all processors are set to the same initial values. Horizontal and vertical slant are both set to 0 degrees (perpendicular to the viewer) and frequency estimates are set to 0 frequency (DC). After some number of iterations during which the parameters change rapidly, the rate of change of average slant estimates for all processors in the system will slow. For the examples shown here, inclination values are shown at the time when average slant estimates do not vary more than 4 degrees over 20 iterations. For very simple regular textures this occurs quickly (50 iterations or less). For very irregular natural textures with complicated error surfaces the system takes longer, usually between 100 and 200 iterations.

6.4 Image Examples

Each processor computes its own estimate of surface inclination from the subset (by frequency and spatial location) of Gabor amplitudes used in computing its error value. While the lateral propagation operation serves to consolidate local estimates by pulling processor slant values toward the average of its neighbors (and, perhaps, out of an incorrect minimum), there is still some variability from processor to processor. For non-planar textures, such systematic differences could characterize the local planar fit to the surface curvature. For the examples which follow, however, only a single inclination estimate is shown which is the average of all processor estimates.

Figure 9a contains the image of a regular synthesized checkerboard texture which has been given a perspective projection of 45 degrees slant, 60 degrees tilt. After 57 iterations the program estimate of its inclination was 44.2 degrees slant, 60.1 degrees tilt. This level of accuracy is typical of the

system's performance for regular textures with strong spectral peaks. In order to provide a subjective means for comparison, the program also produces the perspective projection of a planar surface of grid lines inclined by the same horizontal and vertical slant transformation used to represent inclination within the program (see figure 9b).

Results are shown in similar form for the natural images of figures 10 through 15. Although the precise inclination values for these photographs were unavailable, the grid planes are in subjective agreement with those of the textured surfaces. Note that in figure 14, the surface is, in fact, slightly curved. In its present form the algorithm gives the planar approximation of this surface.

7 Discussion

There is an uncertainty relation which limits the resolution simultaneously attainable in space and frequency with linear filters (Gabor, 1946; Daugman, 1985a). Optimally, the measurement of frequency gradients would be performed with filters of perfect resolution in both domains. However, this is not possible because of the upper limit on simultaneous resolution; the local spectral measurements necessarily admit a band of frequencies over a spatial region of the image. Reducing the bandwidth of the filters has the inevitable effect of increasing their spatial size, thereby increasing the image region (over which the projected frequencies are systematically varying) from which the amplitude is obtained. Increasing the spatial resolution by decreasing the spatial size has the effect of admitting a larger band of frequencies into each measurement. When a texture is composed of one predominant and uniform frequency, this tradeoff is not a significant problem. Many naturally occurring textured surfaces, however, have a different character.

In the description of weighting function it was suggested that the wider the frequency span of the filters over which the error term is computed, the more likely it is that the different amplitude distributions derive from different frequencies in the texture. However, when a texture is composed of *closely* spaced frequency peaks, then the amplitude distributions of *individual* filter types can be altered in ways which make accurate inclination estimation difficult if not impossible. In this situation, each filter amplitude represents a sum of discernible responses to each individual texture frequency. The am-

plitude gradients which result tend to be flatter, resembling those of single frequencies at shallower angles of inclination. (As an extreme example, the perspective image of a surface with a flat power spectrum would contain no frequency gradients.) The algorithm in its present form will tend to underestimate the inclination of such surfaces. It has been found in psychophysical experiments that humans also tend to underestimate the angle of textured surfaces with such irregular appearance (Gibson, 1950b; Gruber and Clark, 1956; Flock and Moscatelli, 1964) suggesting the possibility that a similar process may be occurring.

This is, then, a property of certain textures which makes the measurement of frequency gradients difficult or impossible with filters of finite spatial and frequency resolution. In addition to the texture attribute of reasonable spectral uniformity, the kind of local spectral analysis of texture gradients developed in this paper depends upon the texture frequency peaks being sufficiently separated from each other and from surrounding spectral values so as not to interfere significantly with the local amplitude distributions they independently induce in individual filters. To the extent that they deviate from this condition, the inclination estimates will deviate from the correct values, with a particular tendency toward underestimation of inclination angles. (This is treated in greater detail in Turner et al., 1989a. In this paper slant estimates are compared with the known inclination angles of a set of textured surfaces of increasing spectral complexity.)

However, it is a highly artificial circumstance in which an animal or robotic system encounters only a monocular, stationary view of a textured surface (such as provided to this algorithm). In these cases, the visual system has been deprived of most other natural environmental cues to surface inclination such as stereopsis, convergence of surface borders, relative movement of the texture elements with the motion of the observer, hue changes with increasing distance, depth from focus, etc. Under normal conditions, deficiencies in the visual information available to one system can be overcome by other forms of analysis. It is certainly likely that all of the analysis mechanisms available to biological vision systems operate together to improve the accuracy and reliability of surface perceptions. Even limited to stationary monocular views, a variety of edge and texel based algorithms which differ from the local spectral technique described here are available to machine vision systems, each with particular strengths and weaknesses. This suggests that a robust system for analysis of surface properties such as inclination

should employ and integrate a variety of mechanisms.

The concerns of robustness and reliability also extend to internal aspects of these models. It might be unrealistic to demand a very high degree of precision for a biological information processing system (Huggins and Licklider, 1951). Within a neural assembly, any misarrangement of interconnections, noise, or the death of cells could result in firing rates not perfectly consistent with an optimal or precise representation of information. Although evolution may select for elegant mathematical or engineering solutions, deficiencies or inaccuracies in their implementation may result in a less than perfect execution of those principles. Consequently, an algorithm intended to serve in any sense as a biological model must be robust in that it is capable of adjusting or compensating for the variability of the information provided to it by the lowest level filtering operations and should degrade gracefully with selective deficits in that information.

The set of filters applied to these images is neither complete nor orthogonal. The coefficients derived from their application are inappropriate for such tasks as image representation. Certain spectral values are overrepresented while others are significantly underrepresented. In fact, as described in section 6.2, the current implementation disposes of the low amplitude filter pair values which contain little information about the texture gradient. Because the iterative algorithm adapts to the spatial and spectral relationships between the Gabor filters, it is capable of operating with an almost randomly selected set of filters as input. While the accuracy of the results would vary depending upon the match between the spectral components of the image and the filter set, the absence of particular coefficients would not prohibit the operation of the algorithm. Systematic studies of the relationship between texture spectra and the choice of filter sets, as well as the effects of amplitude variability and noise on the accuracy of results are planned for the future.

The weighting function used in the iterative section is a measure of spectral similarity and spatial distance between filters. In some sense these represent the likelihood that two filters are measuring the same spectral component in the image. In a self-organizing or adaptive system, the weighting function could be implemented as a modifiable measure of simultaneous activity within populations of filter elements. Two elements which are similar in frequency and orientation preference are more likely to be simultaneously stimulated by a wide variety of image material than would those of very

different filters. Furthermore, because image information tends to be locally similar, filter elements with similar receptive fields which are also spatially close are more likely to be simultaneously stimulated. This provides a mechanism by which assemblies of elements (or neurons) could form for the performance of a particular computational function. Filters which are comparable and close will have greater tendency to be simultaneously active and will, therefore, tend to have the most influence on each other, exactly the criterion found useful for the weighting function in this application. In real nervous systems such dynamic formation of assemblies, although long postulated, has only recently been observed (Gerstein et al., 1989).

If this model is to provide insight into mechanisms for biological vision then an additional issue must be addressed. Details of this portion of the research are fully reported elsewhere (Turner et al., 1989b) so they are only briefly described here. Electrophysiological experiments indicate that visual cells frequently operate by application of receptive fields to image material. These cells may be organized in a hierarchical manner to extract increasingly complex and subtle information from the the image. If a mechanism for analysis of surface inclination by local spectral gradients exists in cortex, then it is unlikely to operate by explicit evaluation of nonlinear equations as does the computer based algorithm. However, complicated algorithmic transformations may be realized by sets of receptive fields which represent specific instances of the transformation.

For any value of surface inclination and texture frequency we may systematically vary the Gabor function frequency and spatial location parameters in equation (15) to describe a distribution of amplitudes which would occur over an array of Gabor functions. With the array represented in a four dimensional space (x, y, g_x, g_y) the resulting distribution of amplitudes may be considered to be model for a higher level receptive field which takes Gabor filters amplitudes as input and signals by its output level the presence of a particular inclination angle and texture frequency in the input array. The higher amplitude values in this array represent excitatory regions and lower amplitudes inhibitory ones. The receptive field which results can be described as a central, elongated excitatory band flanked by inhibitory regions. This would be a familiar description of cells in visual cortex if the dimensions of the coordinate system were only spatial, rather than spatial and spectral as in this case. Neurons with receptive fields such as these would have systematically varying preferred spatial frequencies across the spatial dimensions of

their receptive fields. In fact, elements with this receptive field organization have been formed in the hidden layer of a neural network simulation which, using a back-propagation algorithm (Rummelhart et al., 1986), learned to solve a simplified version of the texture gradient problem.

At higher visual levels in cortex, receptive fields are often sufficiently complicated to make an understanding of their purpose difficult if not impossible without a model of the computation being performed. Neurons with receptive field organizations like those described here have not been described in the physiological literature. However, it is also unlikely that an experiment to detect such cells would have been conducted without a predictive theoretical model. It is hoped that computational studies such as these will provide a basis upon which to design the relevant physiological investigations.

By computing parallel inclination estimates on a number of local image regions, the algorithm is well suited to future extension into curved surface description by local planar approximation. In its current form, the lateral parameter propagation phase computes weighted averages which implicitly assume a similarity of inclination across regions characteristic of planar or very slightly curved surfaces. For surfaces with constant rates of curvature, the inclination estimates should vary smoothly from region to region across the image. Because of textural irregularities, however, the actual local estimates will have considerable variability. The current lateral propagation step is, in some ways, an attempt to identify and reinforce the presence of planar structure within the set of inclinations from the parameter adjustment step. If, instead, the algorithm identified higher order regularities within the distribution of inclination values, it should be possible to converge upon a description of certain types of curvature in the textured surface. Various techniques for identifying these regularities are currently under investigation. Whether the low level Gabor amplitudes obtained from natural surfaces have sufficient structure to allow identification of any but the simplest curvatures remains to be determined.

Acknowledgments

We thank Dorothea Blostein for generously providing the textured surface images of figures 10 through 15. We would also like to thank Marcos Salganicoff for his critical reading of the manuscript and many useful discussions. Support for this research provided by ONR N00014-87-K-0766 and AFOSR 88-0296.

References

[1] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2:284–299, 1985.

[2] D. G. Albrecht, R. L. DeValois, and L. G. Thorell. Visual cortical neurons: are bars or gratings the optimal stimuli? *Science*, 207:88–90, 1980.

[3] J. Aloimonos. Shape from texture. *Biol. Cybern.*, 58:345–360, 1988.

[4] R. Bajcsy. *Computer identification of textured visual scenes.* PhD thesis, Stanford University, Stanford, Calif., 1972.

[5] R. Bajcsy. Computer description of textured surfaces. In *IJCAI*, Stanford University, Stanford, Calif, August 1973.

[6] R. Bajcsy and L. Lieberman. Texture gradient as a depth cue. *Computer Graphics and Image Processing*, 5:52–67, 1976.

[7] D. Blostein and N. Ahuja. Representation and three-dimensional interpretation of image texture: An integrated approach. Technical Report UILU-ENG-87-2226, University of Illinois at Urbana-Champaign, Coordinated Science Laboratory, April 1987.

[8] F. W. Campbell and J. G. Robson. Application of fourier analysis to the visibility of gratings. *J. Physiol. (London)*, 197:551–566, 1968.

[9] M. Clark, A. Bovik, and W. Geisler. Texture segmentation using a class of narrowband filters. *Proc. Int'l. Conf. Accoustics, Speech, and Signal Processing - 1987*, pages 571–574, 1987.

[10] J. G. Daugman. Six formal properties of two-dimensional anisotropic visual filters: Structural principles and frequency/orientation selectivity. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:882–887, 1983.

[11] J. G. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Res.*, 20:847–856, 1984.

[12] J. G. Daugman. Image analysis by local 2-d spectral signatures. *J. Opt. Soc. Am. A*, 2:P74, 1985.

[13] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2:1160–1169, 1985.

[14] J. G. Daugman. Complete discrete 2-d gabor transforms by neural networks for image analysis and compression. *IEEE Transactions Acoustics, Speech, and Signal Processing*, 1988.

[15] K. K. DeValois, R. L. DeValois, and E. W. Yund. Responses of striate cortex cells to grating and checkerboard patterns. *J. Physiol. (London)*, 291:483–505, 1979.

[16] R. P. Erickson. Sensory neural patterns and gustation. In Y. Zotterman, editor, *Olfaction and Taste*, pages 205–213. MacMillan Company, New York, 1963.

[17] H. Flock and A. Moscatelli. Variables of surface texture and accuracy of space perceptions. *Perceptual and Motor Skills*, 19:327–334, 1964.

[18] M. G. F. Fuortes. Electrical activity of cells in the eye of the limulus. *American Journal of Ophthalmology*, 46:210–223, 1958.

[19] M. G. F. Fuortes. Initiation of impulses in the visual cells of limulus. *J. Physiol. (London)*, 148:14–28, 1959.

[20] D. Gabor. Theory of communication. *J.IEE. (London)*, 93:429–457, 1946.

[21] G. L. Gerstein, P. Bedenbaugh, and A. M. H. J. Aertsen. Neuronal assemblies. *IEEE Trans. on Biomedical Engineering*, 36:4–14, 1989.

[22] G. L. Gerstein and M. R. Turner. Neural assemblies as building blocks of cortical computation. In Eric Schwartz, editor, *Computational Neuroscience*. MIT Press, Cambridge, MA, 1987.

[23] J. J. Gibson. *The perception of the visual world.* Houghton Mifflin Company, Boston, 1950.

[24] J. J. Gibson. The perception of visual surfaces. *Am. J. Psychol.*, 63:367–384, 1950.

[25] H. E. Gruber and W. C. Clark. Perception of slanted surfaces. *Perceptual and Motor Skills*, 6:97–106, 1956.

[26] D. J. Heeger. Model for the extraction of image flow. *J. Opt. Soc. Am. A*, 4(8):1455–1471, 1987.

[27] D. J. Heeger. Optical flow using spatiotemporal filters. *Int. J. of Comp. Vis.*, 1(4):279–302, 1988.

[28] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurons in the cat's striate cortex. *J. Physiol. (London)*, 148:574–591, 1959.

[29] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (London)*, 160:106–154, 1962.

[30] D. H. Hubel and T. N. Wiesel. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J. Comp. Neurol.*, 158:267–294, 1974.

[31] R. A. Hummel and S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE Transactions Pattern Anal. Machine Intell.*, PAMI-5:267–287, 1983.

[32] L. M. Hurvich and D. Jameson. *The perception of brightness and darkness.* Allyn and Bacon, Inc., Boston, 1966.

[33] K. Ikeuchi. Shape from regular patterns. *Artif. Intell.*, 22:49–75, 1984.

[34] J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1233–1258, 1987.

[35] J. P. Jones and L. A. Palmer. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1187–1211, 1987.

[36] J. P. Jones, A. Stepnoski, and L. A. Palmer. The two-dimensional spectral structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1212–1232, 1987.

[37] K. Kanatani. Detection of surface orientation and motion from texture by a stereological technique. *Artif. Intell.*, 23:213–237, 1984.

[38] J. J. Kulikowski, S. Marcelja, and P. O. Bishop. Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biol. Cybern.*, 43:187–198, 1982.

[39] M. S. Landy. Parallel model of kinetic depth effect using local computations. *J. Opt. Soc. Am. A*, 4:864–877, 1987.

[40] S. R. Lehky and T. J. Sejnowski. Network model of shape-from-shading: neural function arises from both receptive and projective fields. *Nature*, 333:452–454, 1987.

[41] M. D. Levine. *Vision in man and machine.* McGraw-Hill, New York, 1985.

[42] D. M. MacKay. Strife over visual cortical function. *Nature (London)*, 289:117–118, 1981.

[43] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10, 1988.

[44] S. Marcelja. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, 70:1297–1300, 1980.

[45] M. Porat and Y. Y. Zeevi. The generalized gabor scheme of image representation in biological and machine vision. *IEEE Transactions Pattern Anal. Machine Intell.*, 1988.

[46] A. Rosenfeld, R. A. Hummel, and S. W. Zucker. Scene labelling by relaxation operations. *IEEE Trans. Syst. Man Cybern.*, SMC-6:420–433, 1976.

[47] D. E. Rummelhart, G. E. Hinton, and R. J. WIlliams. Learning internal representations by error propagation. In D. E. Rummelhart and J. L. McClelland, editors, *Parallel Distributed Processing: Exploring the Microstructures of Cognition*, pages 318–364. MIT Press, Cambridge, MA, 1986.

[48] W. A. H. Rushton. Peripheral coding in the nervous system. In W. A. Rosenblith, editor, *Sensory Communication*, pages 169–181. MIT Press, Cambridge, MA, 1961.

[49] K. Stevens. Slant – tilt: The visual encoding of surface orientation. *Biol. Cybern.*, 46:183–195, 1983.

[50] K. Stevens. Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception and Psychophysics*, 33(3):241–250, 1983.

[51] T. G. Stockham Jr. Image processing in the context of a visual model. *Proc. IEEE*, 60:828–842, 1972.

[52] R. B. Tootell, M. S. Silverman, and R. L. DeValois. Spatial frequency columns in primary visual cortex. *Science*, 214:813–815, 1981.

[53] M. R. Turner. Gabor functions and textural segmentation. *Opt. Soc. Am. Tech. Digest*, 48, 1985.

[54] M. R. Turner. Texture discrimination by gabor functions. *Biol. Cybern.*, 55:71–82, 1986.

[55] M. R. Turner. A model for the underestimation of texture slant by human observers. *In preparation*, 1989.

[56] M. R. Turner. A receptive field model for the perception of textured surface slant in visual cortex. *In preparation*, 1989.

[57] A. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.

Figure Captions

Figure 1: A set of 2D Gabor filters. The set of 48 filters in the top half of the figure are the ones used in calculating inclination of all the image examples which follow in the paper. The set (shown in actual size relative to the textured image examples) contains 6 frequencies (8, 10, 12, 14, 16, 18 pixel wavelengths), 4 orientations (0, 45, 90, 135 degrees) with 2 phases (0 and 90 degrees) for each frequency and orientation specification. All filters have a bandwidth of 1.5 octaves, thereby making them scalings and rotations of the two larger filters (shown to provide a more detailed view of the smaller filters actually used) in the bottom half of the figure.

Figure 2: Perspective imaging model. A textured planar surface with coordinate system (u, v) is distance d from a pinhole camera with a flat image plane, coordinate system (x, y) and focal length f. The inclination of the textured surface to the line of sight is indicated by the two angles θ and ψ. Note that the origins of the (x, y) and (u, v) coordinate systems are on the line of sight at the centers of their respective planes.

Figure 3: The perspective projection of a sine wave grating produces an image in which the projected frequencies vary smoothly across the image.

Figure 4: Application of a particular 2D Gabor filter pair to the perspective sine wave image of the previous figure. The amplitude of the filter pair (indicated by the gray level intensity with the highest amplitude indicated by white) varies from location to location in the image depending upon the local match between the projected sine wave frequency in the image plane and the frequency – orientation of the Gabor filter.

Figure 5: Diagram of the algorithm for estimating textured surface inclination. The algorithm may be viewed as operating in parallel on a number of processors, each with access to a local section of the image and with interconnections to a subset of its neighboring processors. The surface inclination estimates are made in a pair of iteratively executed operations labeled "Local Parameter Adjustment" and "Lateral Parameter Propagation". The first of this pair attempts to fit inclination and other parameters to the Gabor filter amplitudes from its local patch of image. In the second phase, the parameters are propagated laterally to neighboring processors for consolidation by calculation of weighted parameter averages.

Figure 6: Error surface for one Gabor filter frequency – orientation specification applied to the image of an inclined sine wave grating. Each location on the surface denotes one surface inclination value (θ, ψ). The height at each point indicates the degree to which the Gabor amplitudes measured from the image fit this inclination value. (The greater the height, the better the fit.) When only one frequency – orientation filter specification is used, there are two surface inclinations which have minima. The one in back is the correct inclination. The one in front is the reverse of this inclination (similar slant angle in the reverse direction from perpendicular). If the algorithm makes the wrong initial parameter choices it can easily select a surface inclination which is the reverse of the correct one.

Figure 7: a. Frequency tuning curve for a single Gabor function. With its symmetric Gaussian frequency response, a rising spatial gradient of amplitudes from the bottom to the top of an image may represent either b. a rising frequency gradient (with the surface inclined away from the viewer at the top) on the low frequency side of the filter's tuning curve or c. a decreasing gradient (inclined toward the viewer) on the high frequency side of the tuning curve. d. Adding a second filter having considerable frequency overlap with the first affords an unambiguous means of determining which side of the Gaussian tuning curve an amplitude gradient represents. Although the amplitude gradients of the first filter are similar, the amplitudes of the second are significantly different in e. and f.

Figure 8: Error surface for three Gabor filter frequency – orientation specifications. Constraining the solution to satisfy three filters with similar but not identical frequencies disambiguates the previous situation. Only one minimum appears (the correct inclination angles) and the peak is more sharply defined, allowing the algorithm to arrive more rapidly to the solution.

Figure 9: a. Synthesized Checkerboard with a 45 degree slant, 60 degree tilt. b. After 57 iterations the average inclination estimate for all processors is 44.2 degrees slant and 60.1 degrees tilt. For subjective comparison, a planar surface of grid lines is shown with the same surface inclination as that computed by the program.

Figure 10: Image of birds flying over water. The average inclination estimate of all processors for this image is shown in the same planar grid format as for the checkerboard.

Figure 11: Overhead clouds.

Figure 12: Field of clover.

Figure 13: Field of mud.

Figure 14: Electron micrograph of a pen.

Figure 15: Leaves.

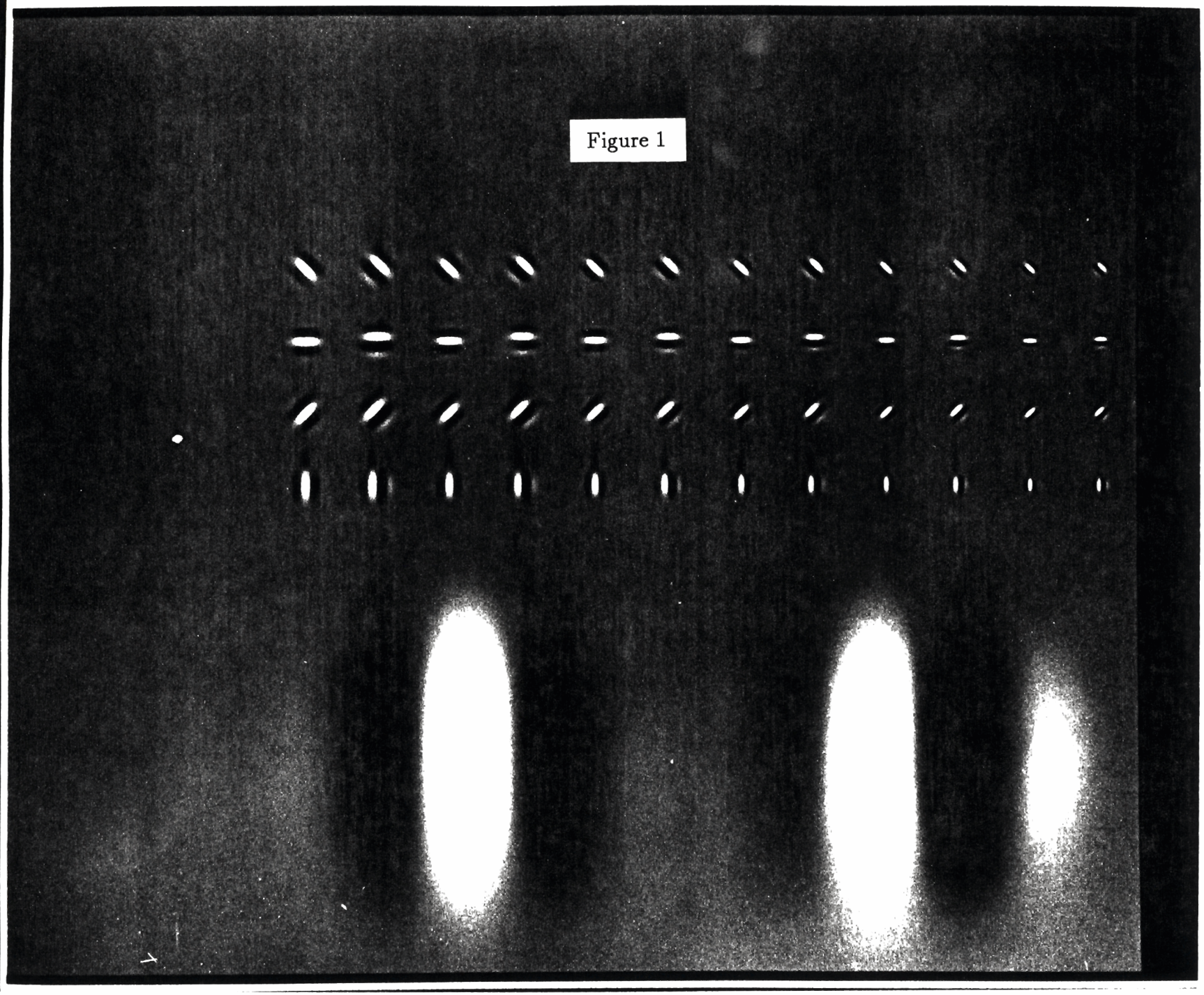


Figure 1

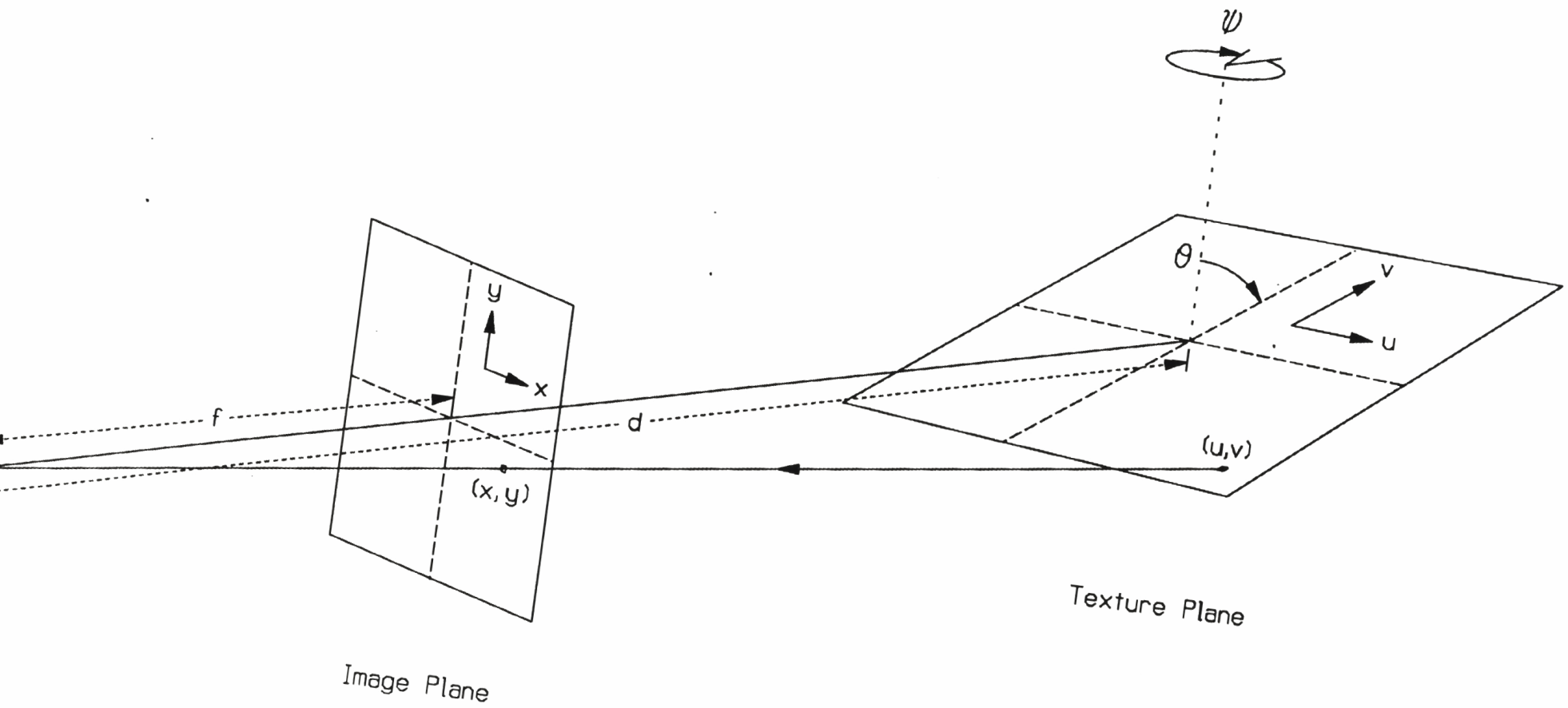


Figure 2

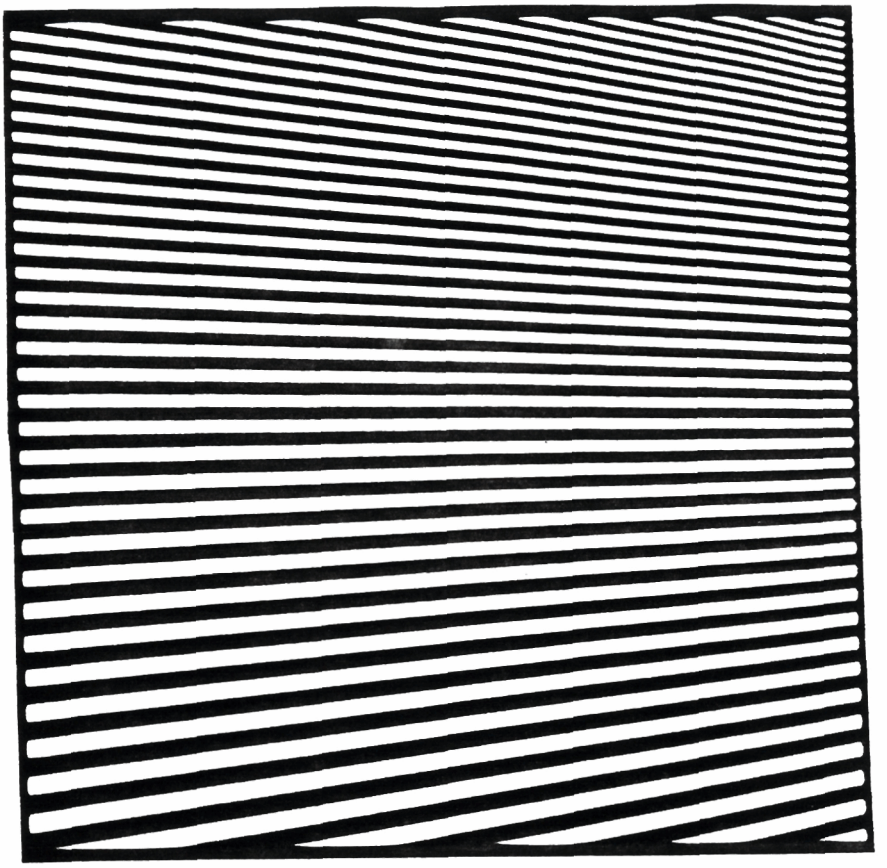


Figure 3

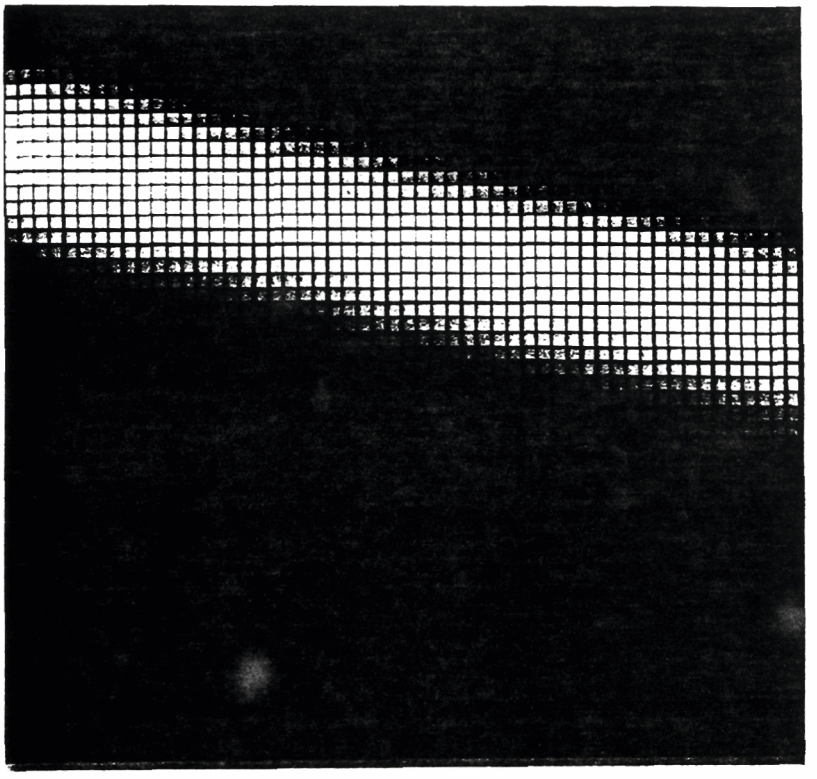


Figure 4

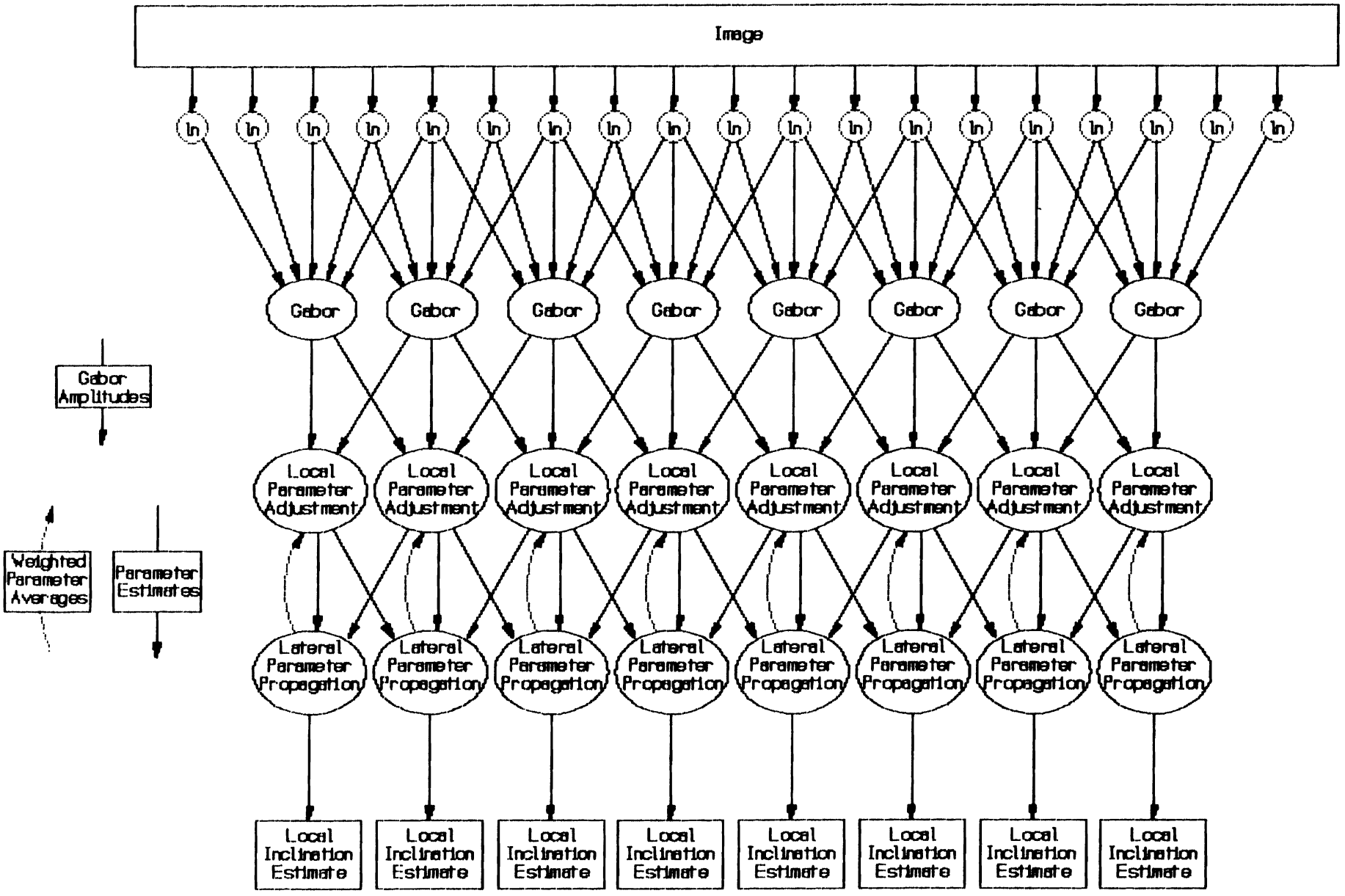


Image
In
Gabor
Gabor Amplitudes
Local Parameter Adjustment
Weighted Parameter Averages
Parameter Estimates
Lateral Parameter Propagation
Local Inclination Estimate

Figure 5

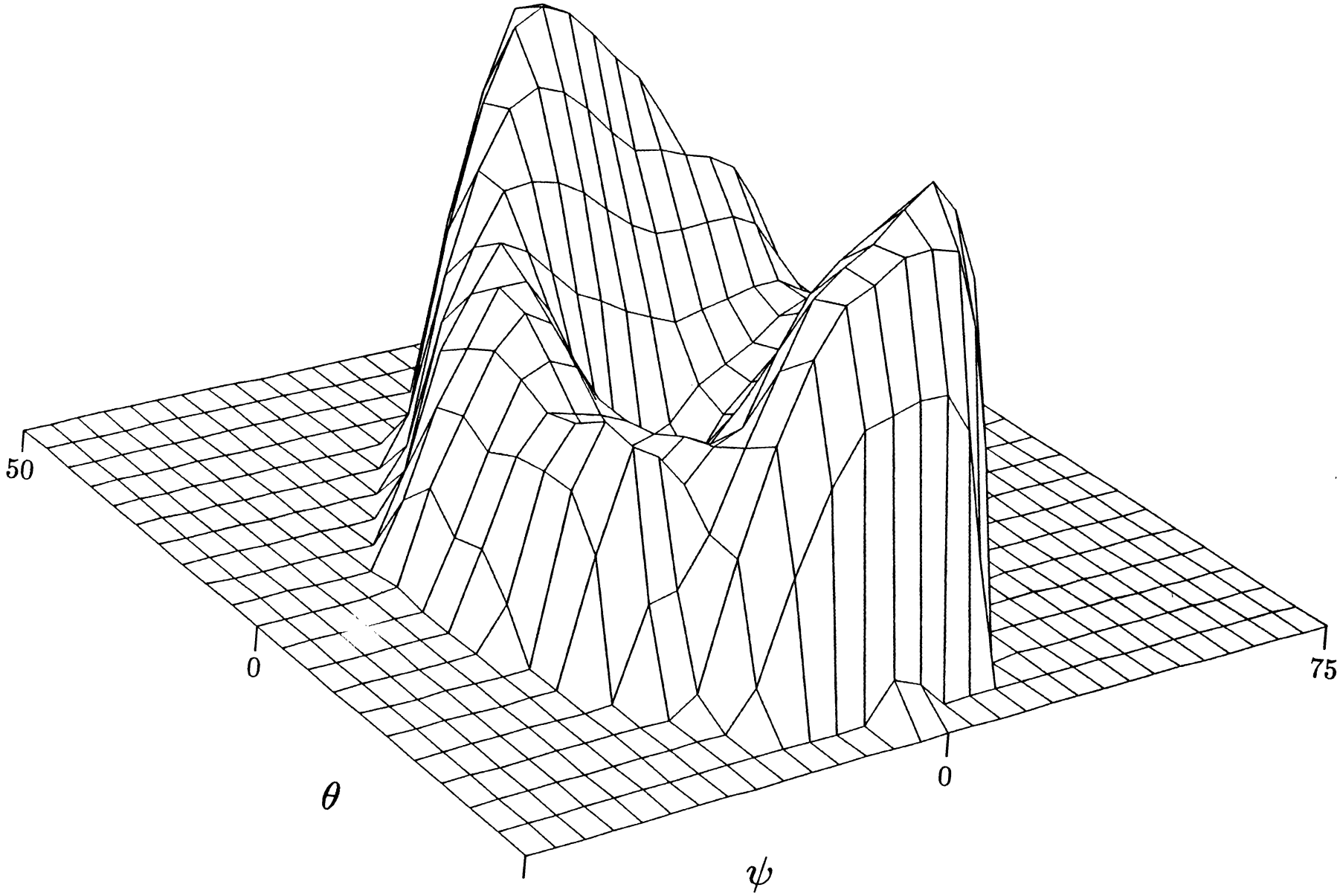


Figure 6

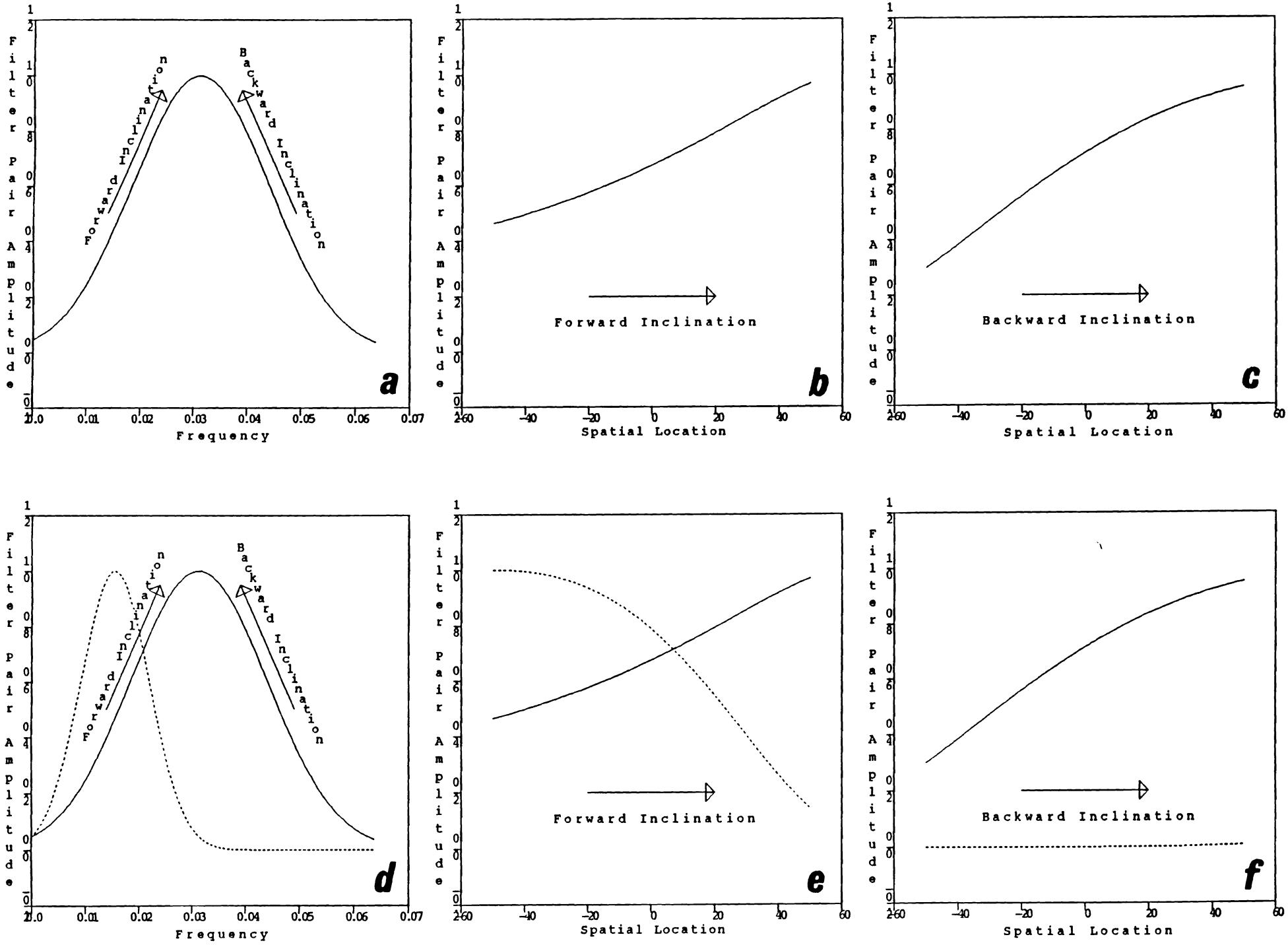

Filter pair Amplitude
Frequency
Forward Inclination
Backward Inclination
a
Spatial Location
Forward Inclination
b
Spatial Location
Backward Inclination
c
Filter pair Amplitude
Frequency
Forward Inclination
Backward Inclination
d
Spatial Location
Forward Inclination
e
Spatial Location
Backward Inclination
f

Figure 7

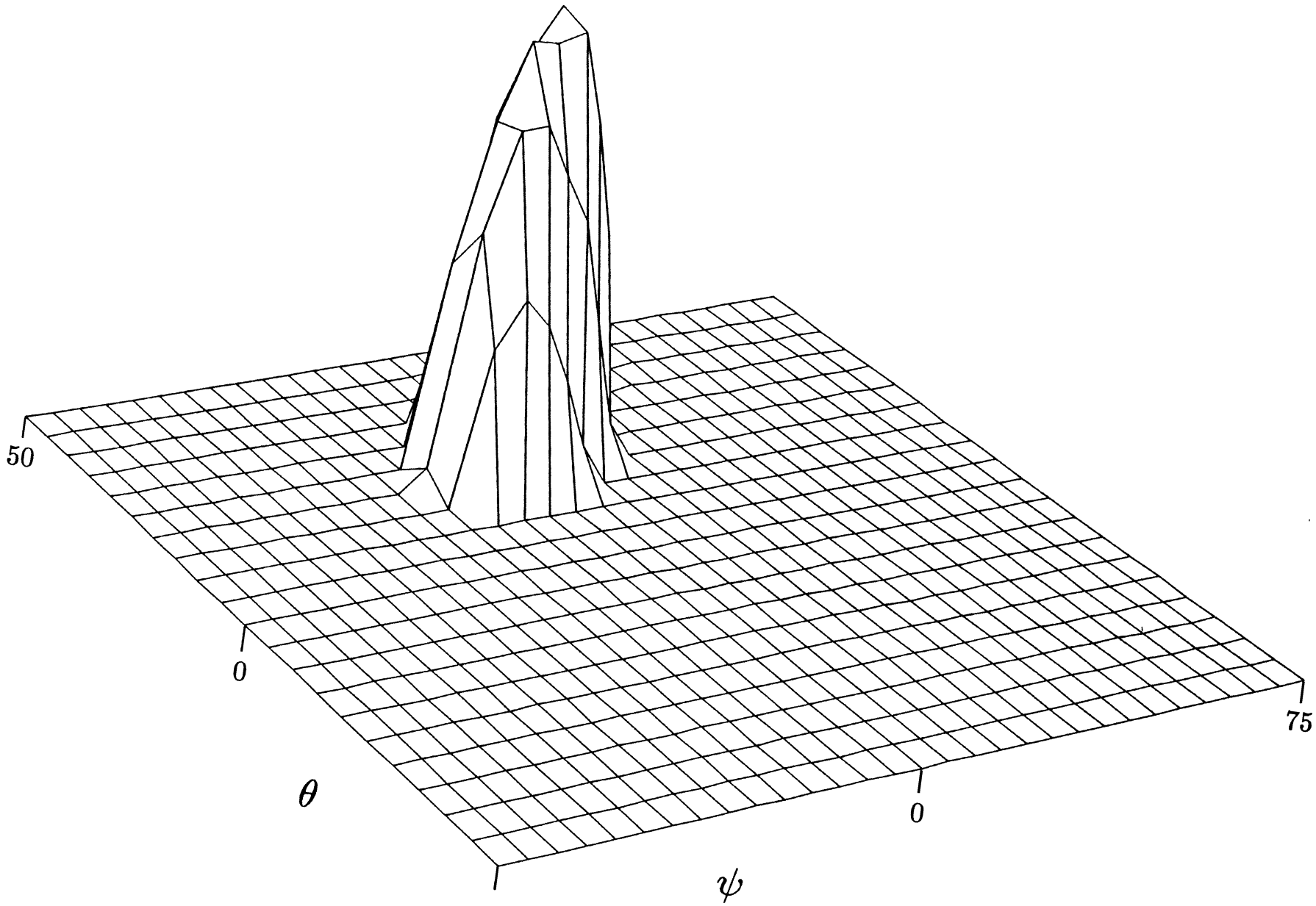


Figure 8

Figure 9

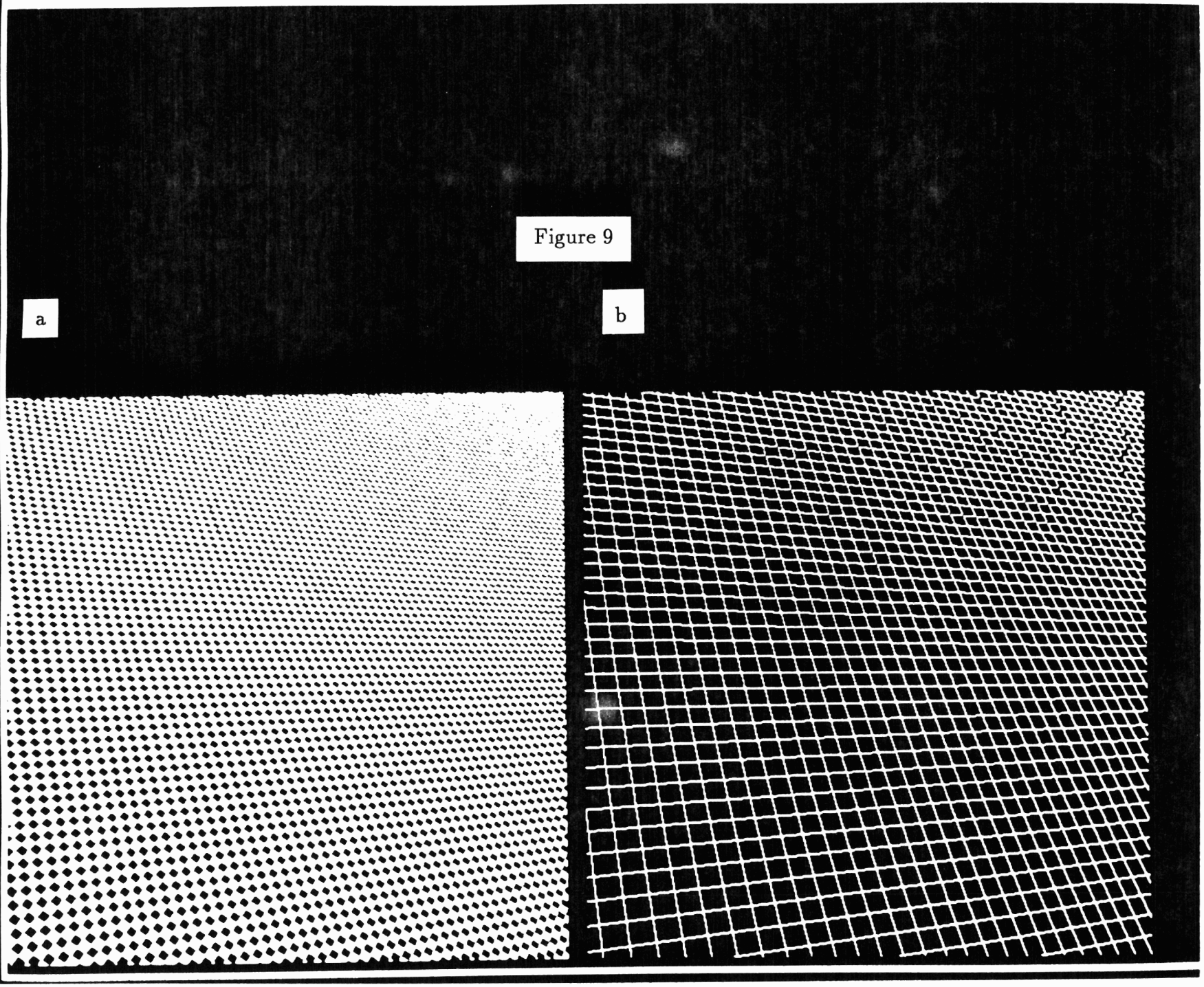
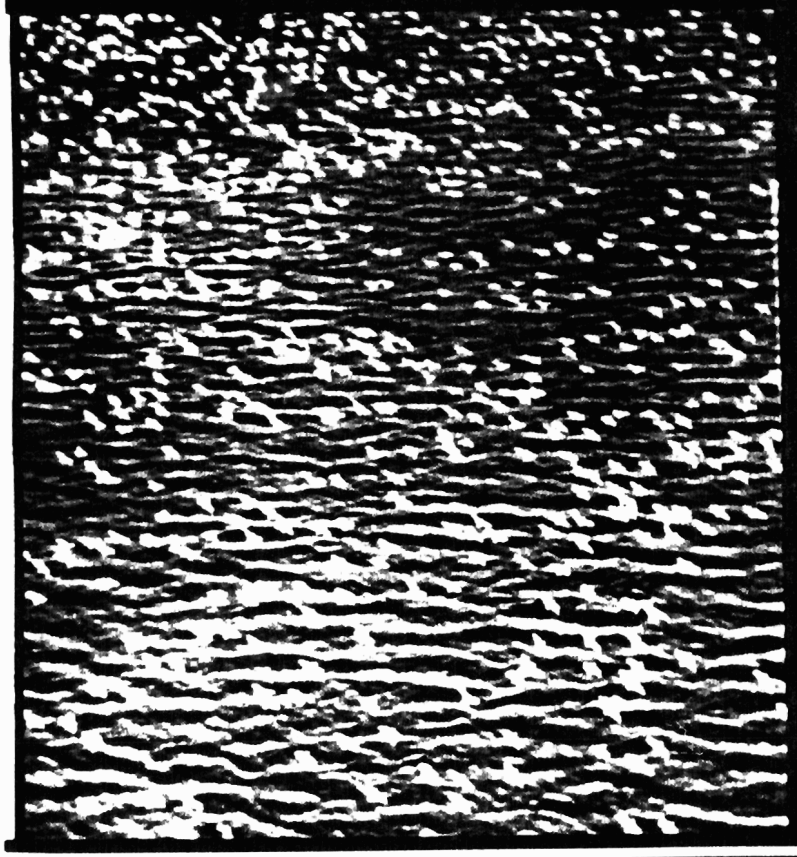


Figure 10



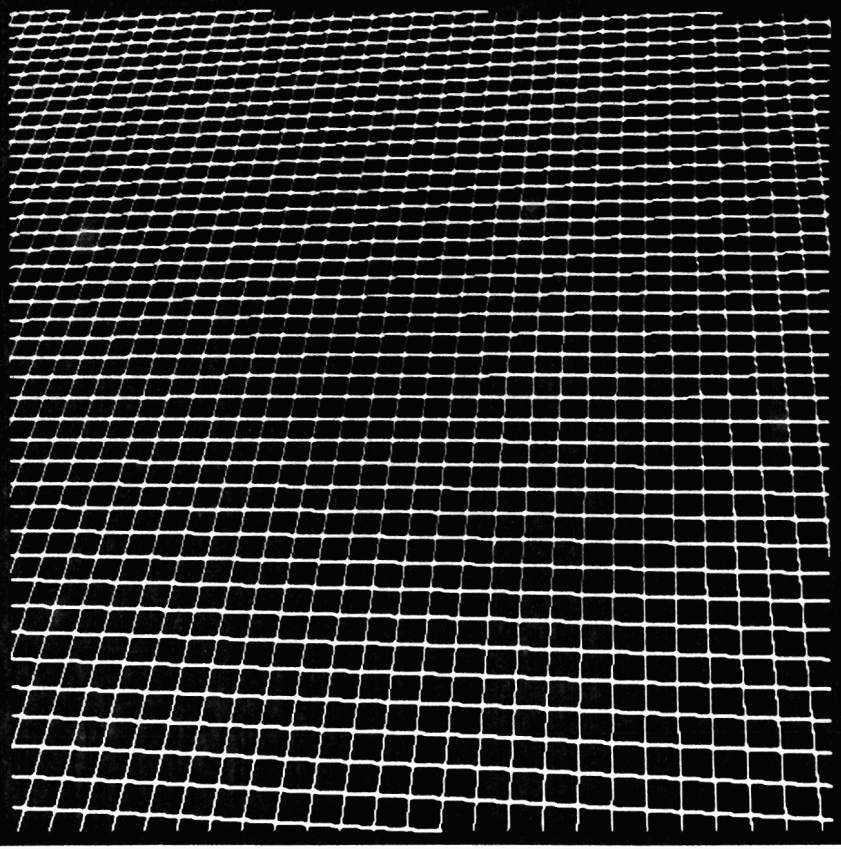


Figure 11

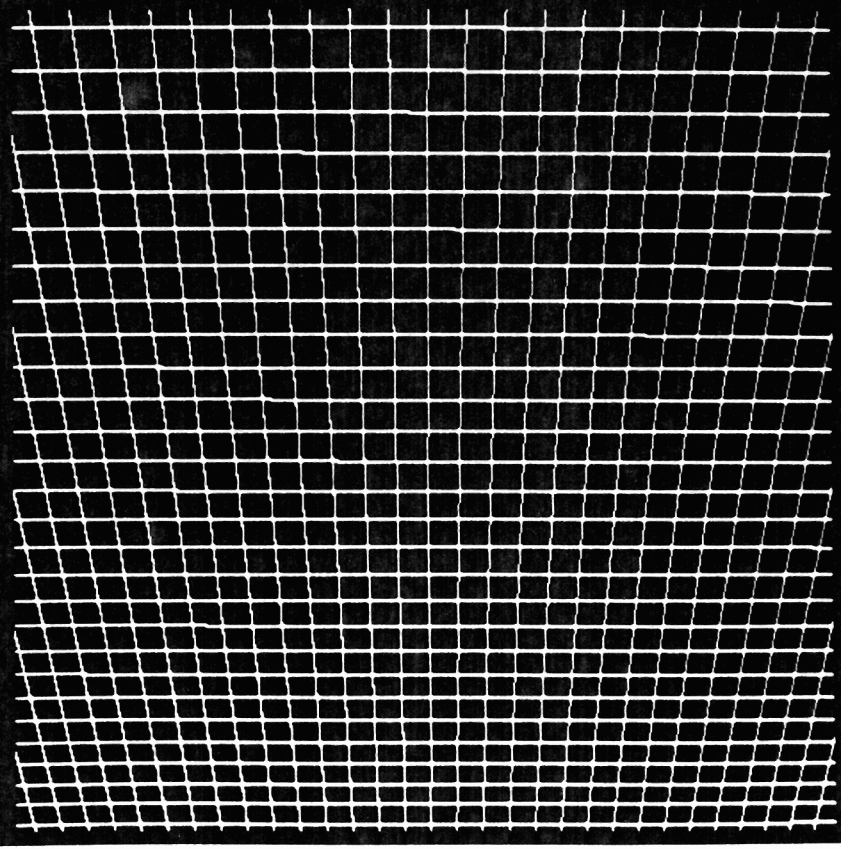


Figure 12

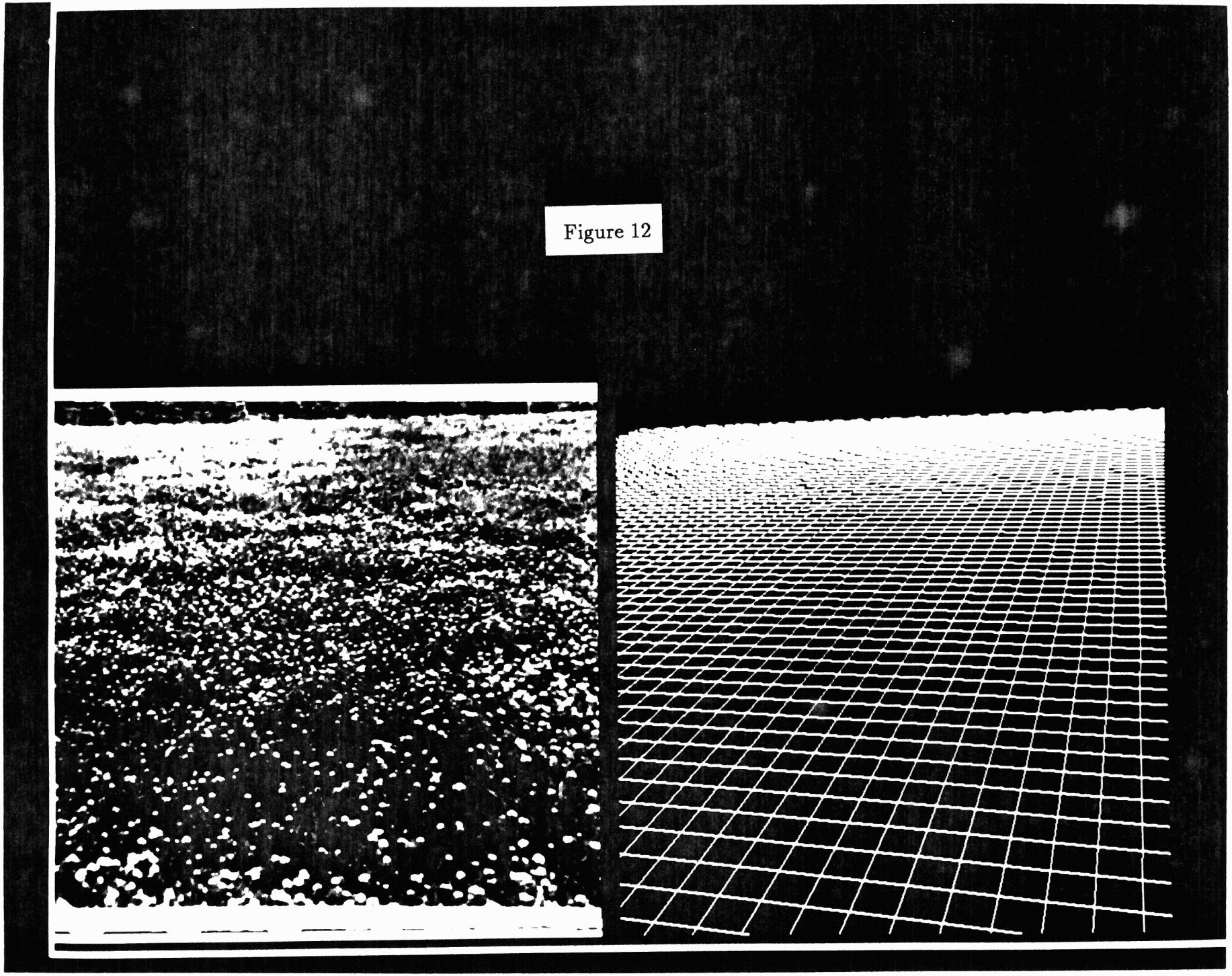
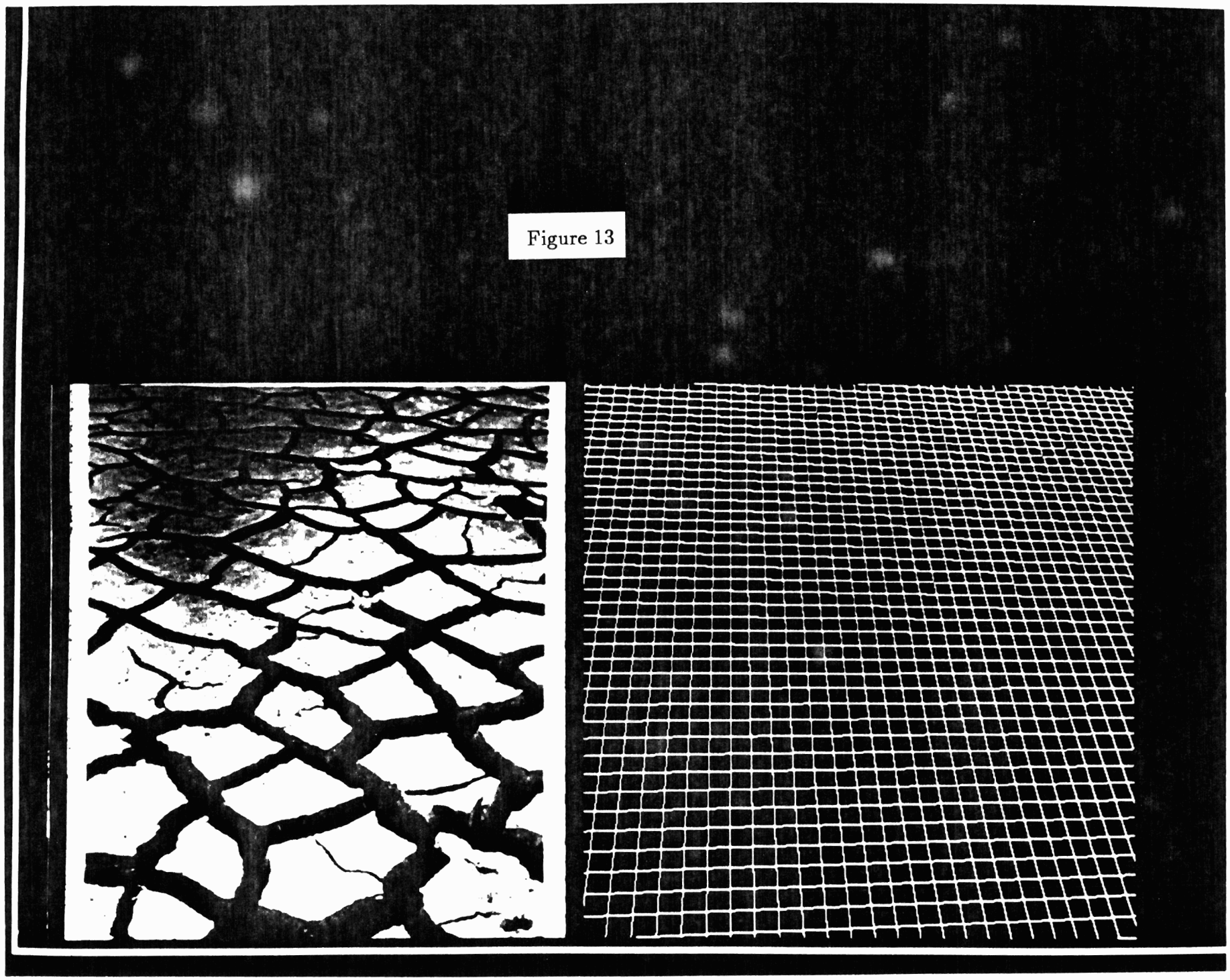


Figure 13



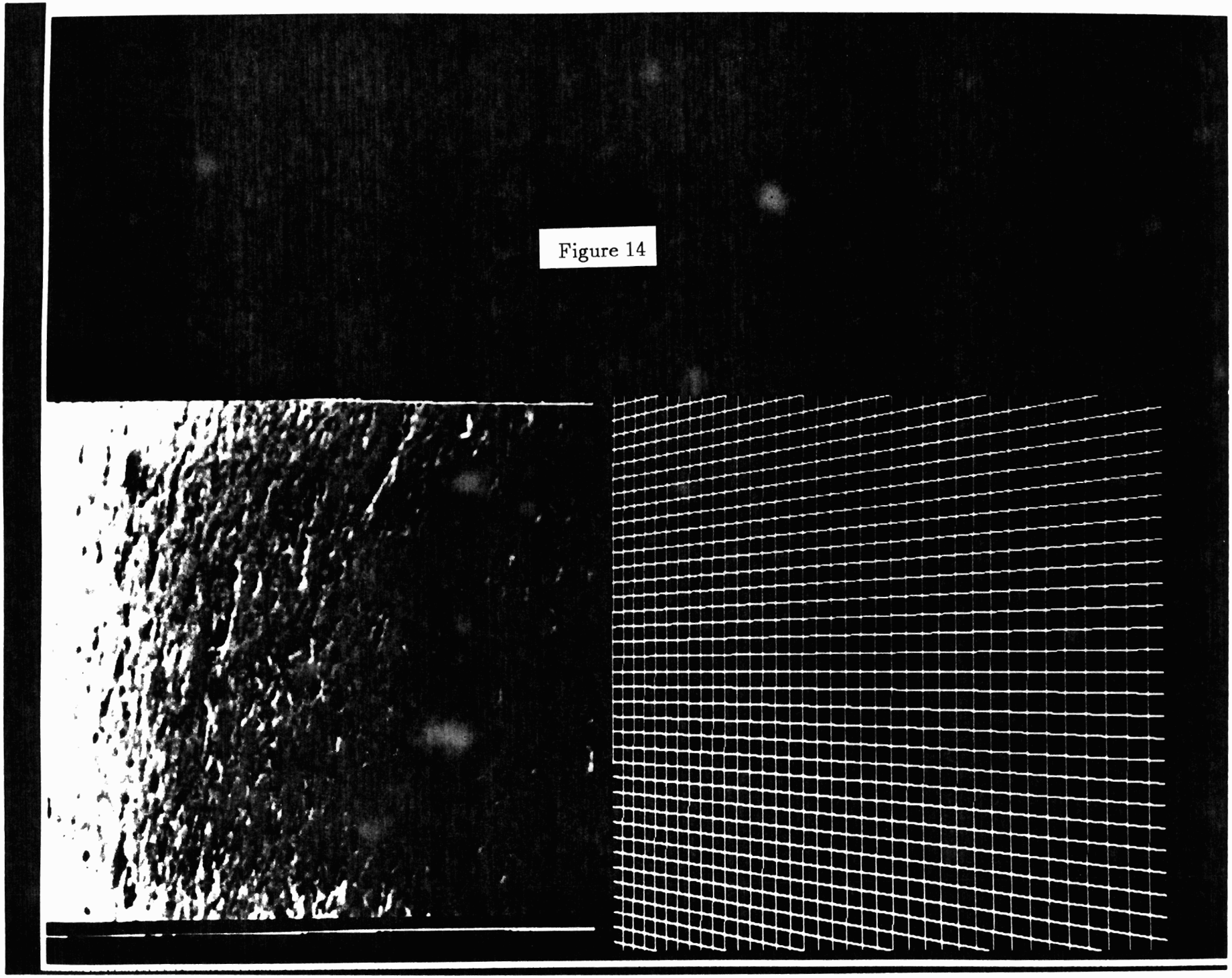


Figure 14

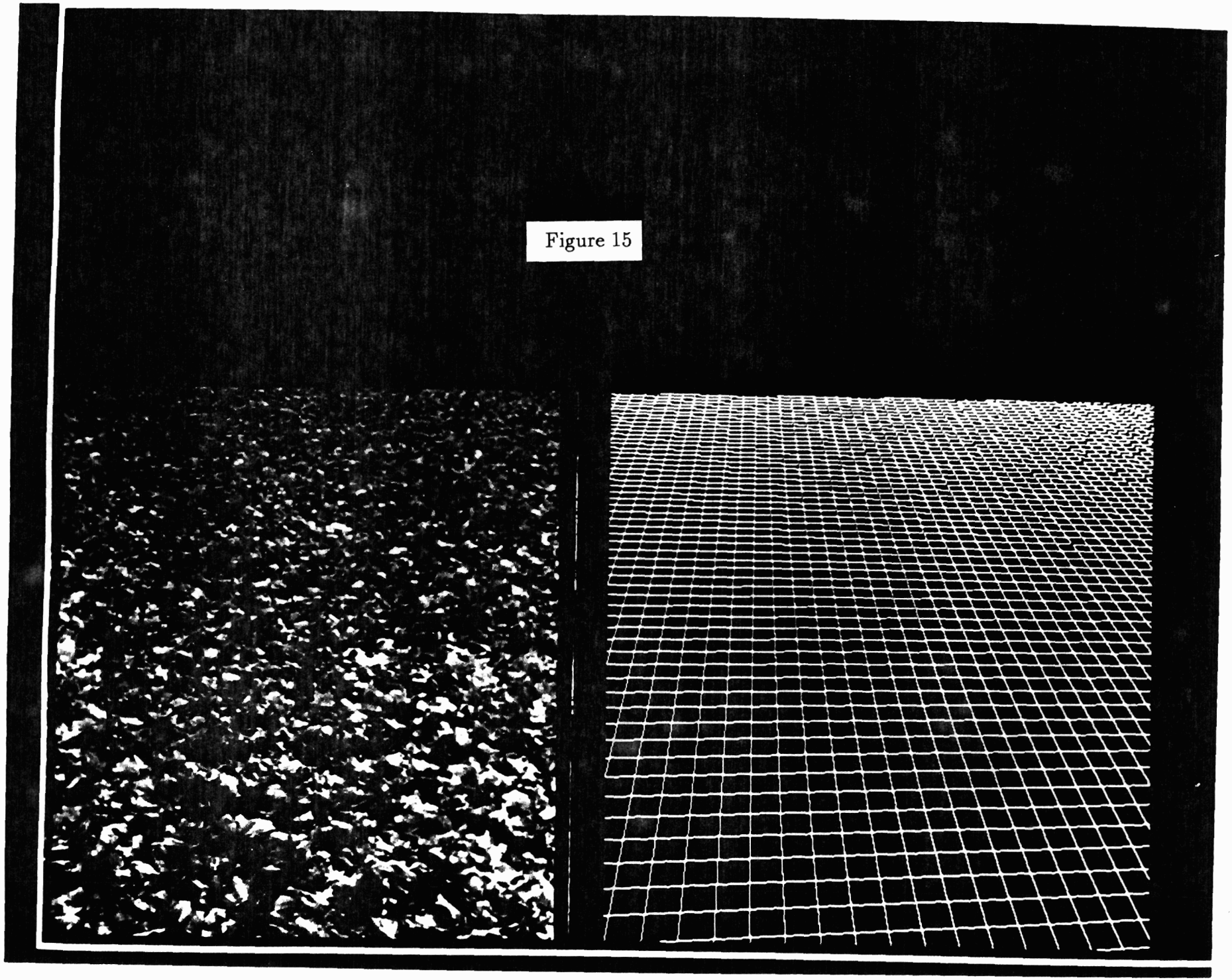


Figure 15