COMPUTER GRAPHICS AND IMAGE PROCESSING 5, 52–67 (1976)

Texture Gradient as a Depth Cue

RUZENA BAJCSY AND LAWRENCE LIEBERMAN*

University of Pennsylvania, The Moore School of Electrical Engineering, Department of Computer and Information Science, Philadelphia, Pennsylvania 19174

Communicated by A. Rosenfeld

Received February 10, 1975

A texture operator for use in computer vision programs is described. The operator classifies texture according to characteristics of the Fourier transform of local image windows. Gradients of the texture are found by comparing and associating quantitative and qualitative values of adjacent windows. The gradients are then interpreted as a depth cue for longitudinal (receding) surfaces. Experimental results with natural, outdoor scenes are reported.

1. INTRODUCTION

In this paper we shall concentrate on a special aspect of computer vision programs: namely, on their capability to make three-dimensional interpretations of objects from individual pictures. Shortly, we shall discuss monocular depth cues, in general, and texture gradient in particular, and their possible computer implementation. It is a prevailing theory in psychology that *gradients* play a large role in human perception of three-dimensional forms from their two-dimensional representations. We have developed a method of measuring texture gradients. The domain in which we experiment is that of natural outdoor scenes, such as those shown in Figs. 1 and 2. Fourier descriptors of texture are implemented whose quantitative components vary in a manner which is consistent with three-dimensional surface interpretation. However, for a computer to make three-dimensional interpretation of surfaces in the scene, besides having the ability to measure gradients, it needs to have some knowledge of the world in an internal representation. The representation of this knowledge, called the model of the scene, includes the model of the world, the model of the observer, the illumination model, and the environmental model.

2. THE PSYCHOLOGICAL BACKGROUND OF GRADIENTS IN VISION

Nearly all so-called pictorial (monocular) cues to depth were discovered by the artists [8]. The contribution of psychological theory was Gibson's *The Perception of the Visual World* [5].

* Current adress: IBM, Tomas J. Watson Research Center, Yorktown Heights, N. Y. 10598.

FIG. 1. A grassy field and forest.

Gibson's theory is primarily concerned with the three-dimensional visual world. He makes an important distinction between the visual world of 3-D objects and the visual field which is a direct correspondence to the retinal image. The visual world is defined by surfaces and edges and is perceived as "an array of physical

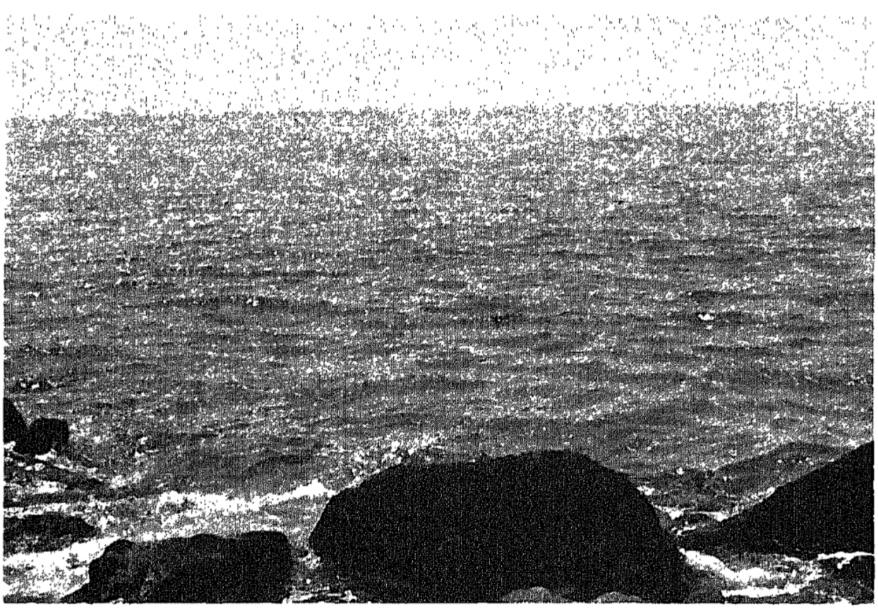

FIG. 2. An ocean view.

surfaces reflecting light and projected on the retina." Surfaces are primary percepts which fill up our world. When the edge of a surface is not clearly delineated we may see that surface as "background." There are basically two classes of surfaces in the visual world, longitudinal and frontal. These directions are defined with respect to the observer's line of sight; longitudinal is parallel to the line of sight and frontal is transverse to it. They in general correspond to ground (also floor, ceiling, walls) and to individual objects, respectively. Longitudinal surfaces are responsible for our perception of distance.

Unlike his predecessors who based visual perception on a spatial system of geometric coordinates and points of light projected on the retina, Gibson states that it is surfaces that are perceived and that the physical stimuli due to these surfaces are not points of light but *gradients of physical features*. Such gradients are forms of ordinal stimulation, i.e., "a simultaneous variation over the set of different receptors, or a differential excitation of different receptors." In this manner he concludes that "the general condition for the perception of an edge, and hence for the perception of a bounded surface in the visual field, is the type of ordinal stimulation consisting of an abrupt transition." This is a concept familiar to computer picture processing; it is the basis for most edge-finding operators. More important, though, is the statement that "the perception of depth, distance, or the so-called third dimension, is reducible to the problem of the perception of longitudinal surfaces ... the general condition for the perception of a surface is the type of ordinal stimulation which yields texture."

The *gradient of the texture* will be due to the (perceived) decreasing size of texture elements or shrinking of the element spacing, etc. This is the primary depth cue. There are other important cues which we chose to ignore because of the inapplicability to our work with momentary, monocular views. They are binocular disparity (stereoscopic vision), degree of convergence of the eyes, and motion parallax. There are other monocular depth cues, however, which exist and are used by humans when the primary cues are lacking. There is the convergence of edges of a longitudinal surface, the common "perspective" effect of parallel lines appearing closed together as they recede from the viewer. This is a special case of a *gradient of the distance* between the lines.

Another commonly used cue which comes into play when a surface gradient is not present is the *gradient of size* of known, similar objects. This cue is obviously experience dependent; it is learned. Without getting into the controversy of what perceptual skills are learned, we will simply state that Gibson indicates certain 3-D perceptual processes, particularly those involving the primary cues, are innate.

Gradients of surface texture are dependent only on the physical characteristics of the surface and the observer's viewpoint. There are other gradients, namely, shading gradients, which are also functions of the illumination. Shading gradients give us cues to interpret the "depth-shape" of object surfaces. Models for this depth-shape cue have been derived and employed in computer vision studies [7,9].

A very weak depth cue, but one which is easily observable, is a hue gradient known as *aerial perspective*. This is the apparent blue haziness seen at a distance

in outdoor views. It is due in part to the atmosphere, and is a kind of "last resort" cue for distance perception.

Gibson classifies forms into three major groups [4]. The first group is the 3-D world of objects consisting of solids and surfaces. The second is the set of representation forms, i.e., 2-D representations of objects and includes outline, pictorial, plan, perspective, and nonsense forms. The third class is that of abstract geometrical forms. It is Gibson's contention that we are inherently able to perceive the surface and solid forms of his first grouping, but that the perception of other forms is learned. We naturally perceive three dimensions; we must interpret 2-D representations.

Pictorial forms (in our case, photographs) are two-dimensional representations of the real world. But, as Gibson points out, if the representation (film) is good enough and if one limits the observer to a "peephole" view (not showing any of the frame), then the observer using one eye may perceive a 3-D scene. Thus, if we can assure that sufficient gradient information exists in our (digitized) data, then much of the theory of 3-D cues can be applied to our scenes. This is without using any explicit 3-D data such as laser ranging information, or third coordinate digitization.

3. A TEXTURE OPERATOR

It is a straightforward statement that in order to measure or detect texture gradient one first has to have some means of describing the texture. In this work we have adopted a technique developed by Bajcsy [1], which generates texture descriptors on a local windowed basis from features of the power spectrum of the Fourier transform of the image.

A digitized picture can be considered a real function of two variables, $g(x, y)$, where x and y vary over the natural numbers up to the size of the image (x and y are the Cartesian coordinates of a picture point). The two-dimensional discrete Fourier transform of $g(x, y)$ is

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

where p is the dimension of the image matrix. The power spectrum of $F(n, m)$ is the magnitude of that complex function and is given by

$$P(n, m) = [F_{\mathrm{Re}}^2 (n, m) + F_{\mathrm{Im}}^2 (n, m)]^{\frac{1}{2}}.$$

The phase spectrum is

$$\psi(n, m) = \arctan [F_{\mathrm{Im}} (n, m)/F_{\mathrm{Re}} (n, m)].$$

The Fourier operator has properties which can be utilized for texture analysis. The power spectrum is invariant to translation in the spatial domain but not invariant to rotation. This implies that the directionality of a pattern in the image is preserved in $P(n, m)$. The Fourier transform compresses data if it is periodic or almost periodic, which is one characteristic of texture. The phase function contains information on the position in the image window and is of little

use in texture processing. The entire Fourier transform is symmetric with respect to the origin in the complex plane if the original function is real. To detect the directionality information contained in the power spectrum, it is transformed from the Cartesian system (n, m) into a polar system (r, phi). Gross shape properties of the texture elements may be inferred from $P(r, \text{phi})$. The function $P(r, \text{phi})$ is decomposed into $PPHI(r)$ for each phi, and into $PR(\text{phi})$ for each r yielding a pair of functions $[PPHI(r), PR(\text{phi})]$ to be used by the texture operator.

The operator looks for significant peaks in $PR(\text{phi})$ which indicate directionality of the texture: flat $[PR(\text{phi})]$ implies nondirectional, a few peaks mean directional. An example of a highly directional texture is a grassy field. Two significant peaks may indicate a bidirectional texture, given certain constraints. A peak is considered significant if it lies above a threshold equal to the mean of the function $[PR(\text{phi})]$ plus 1.5 times the standard derivation. The rationale for choosing a factor of 1.5 times the standard deviation is that if the function were sinusoidal then none of its peak should be considered significant. The standard deviation of a sine wave is the square root of 2 (1.414). Thus, a factor of 1.5 would threshold out all the maxima of a sine curve. In the case of nondirectional texture, $PPHI(r)$ indicates a noisy texture if it is flat, and bloblike if it has some peaks. A homogeneous texture, i.e., a smooth area, has an almost constant $PPHI(r)$ and a large value for $PPHI(0)$ (the D.C. constant).

Thus, the texture operator provides a quantitative as well as a qualitative description of the texture. The quantitative components of the texture descriptor consist of the average gray value (the D.C. value), the prime textural directionalities, the maximum power corresponding to prime directionalities $\{\text{MAX}[PR(phi)]\}$, the maximum power in each prime direction $\{\text{MAX}[PPHI(r_i)]\}$ and their corresponding spatial frequencies (RMAX) and wavelength.

Qualitative descriptors include the name of the texture class (bloblike, monodirectional, noisy, homogeneous, etc.) grades of contrast (sharp, medium, weak), grades of brightness (bright, dark), and grades of granulation (large, medium, small). These qualitative descriptors, naturally, are derived from the quantitative descriptors. They represent gross classification on a local basis that should help in a region-growing process, which consists of finding similarities and dissimilarities between adjacent texture windows.

Thus, the texture classification is based on the shape of the power function in different directionalities $[PR(phi)]$ while the contrast is derived from the maximum power of $PPHI(r)$. The brightness is measured by the average gray value, provided that the texture is homogeneous (smooth). Finally, the granulation is dependent on the wavelength.

An example of a texture descriptor record for a "grassy" window is shown in Table 1.

For a detailed description of the texture operator, see [10].

4. USING TEXTURE GRADIENTS

As was pointed out, texture gradient is, essentially, a gradient in the size of texture elements and their spacing. Gradients of these components as a function

TABLE 1

Texture Descriptor Record for a Grassy Window

Texture class	3 (bloblike)
Number of directionalities	2
D.C. value	8.9 $E + 0.05$
Max(PR(phi))	2.3 $E + 0.05$, 1.8 $E + 0.05$
Directionalities	90°, 0°
Max $(PPHI(r))$	1.0 $E + 0.05$, 1.4 $E + 0.05$
$r$max	5, 7
Wavelengths	25, 18
Contrast	weak
Size	medium

of distance in a particular image direction can be measured given that certain conditions on the image data are met. The interpretation of such a gradient is a higher-level process. Furthermore, it is a general process, global to all 3-D interpretation of surfaces.

The texture gradient plays a particularly strong role in outdoor scenes since the primary longitudinal surface, the "ground," is the main cue for distance and depth in the entire scene and may be identified or verified through the gradient. In addition, if the model of the observer is detailed enough, i.e., if we have the camera focal length, field of view angle, the height above the ground, etc., then quantitative distance inferences can be made. The depths thus obtained are only relative; the absolute scale cannot be determined unless we also know the actual size of the object space texture elements.

A geometric model for this kind of interpretation is given in Fig. 3.

l_i is the observed texture element size in the image plane. Y_i is the center of the window in which l_i is found. The texture elements on the ground are all assumed to be of the same size ($=t$). Y_A, Y_B, and Y_C are the Y values of the projections of points A, B, and C in the image plane. The problem is then to determine the ratio

$$\overline{AC}/\overline{AB}.$$

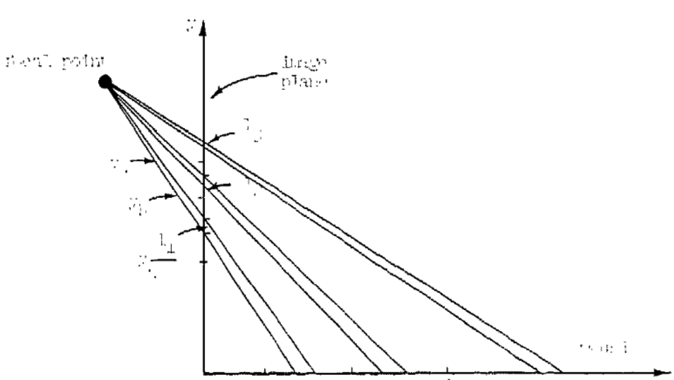

FIG. 3. A geometric model for interpreting the texture gradient.

Consider a projection function $P(Y)$ which indicates how distance on the ground changes with distance in the image plane. P is of the form

$$P = (1/k_1)(ds/dy),$$

where k_1 is a proportionality constant dependent on geometric parameters of the system.

Then the distance in the object space can be expressed as

$$S = k_1 \int P \, dy.$$

We have a form of P in our measured data. The l_i's are the texture wavelengths or element sizes measured in adjacent windows along the vertical direction in the image.

$$P_i\ (Y_i) = k_2 t / l_i(Y_i),$$

where k_2 is another constant. We can fit a curve to P (using, for instance, a parabolic curve); call this curve P^*. Then the ratio is

$$\frac{\overline{AC}}{\overline{AB}} = \frac{k_1 k_2 t \displaystyle\int_{Y_A}^{Y_C} P^* dy}{k_1 k_2 t \displaystyle\int_{Y_A}^{Y_A} P^* dy} = \frac{\displaystyle\int_{Y_A}^{Y_C} P^* dy}{\displaystyle\int_{Y_A}^{Y_B} P^* dy}.$$

Figure 4 shows this graphically. In other words, if our goal is the relative distance comparison, the above method eliminates the need for knowing the constants of the system (focal length, height above the ground, field of view angle). The necessary information is acquired through the gradient function.

The conditions we referred to previously are minimal restrictions on the image characteristics and the texture operator itself. The desired texture operator must

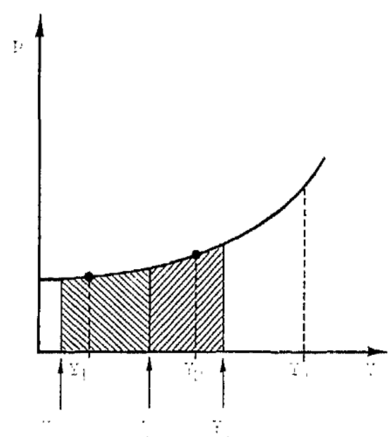

FIG. 4. The projection function.

be sensitive enough to register the repetition of those features of texture elements which perceptually group them to form a textured surface. It is our contention that rough shape class and size of elements and spatial organization (element spacing and placement) which are measured by the Fourier operator capture those features. (See [2, 5, 6, 11, 12], for a discussion of the analysis of texture and for a survey of the appropriate psychological studies.) The texture operator must not be sensitive to variations of outdoor scene texture elements or variations caused by the projection of the 3-D objects into the 2-D image plane. This last problem is the one of partial occlusion of one physical texture element by another. The result is a reflected "blob" in the image plane or "blobs" whose shapes vary widely from the expected shape of the object space texture element. The Fourier operator is an averaging technique which we feel is especially suited to natural, outdoor textures, since its qualitative descriptions are very rough (see Section 3). It can handle the above problems because it does not depend upon exact similarity of shape or upon perfectly uniform organization of elements as some current spatial texture analyzers do.

The conditions on the picture data are related to the spatial resolution and the overall image size. Basically, there must be a sufficient number of digitized points to be able to "see" the texture elements, and a large enough image so that contiguous (texture) windows have at least the minimum information for a gradient.

It is clear that these conditions are dependent on the resolution of the digitizer and on the information contained in the world model about the observer and the textural characteristics of objects. The Fourier operator measures major variations across a window (in various directions) which are due to reflections from the texture elements. Given the model of the observer, including the information shown in Fig. 3, and the attribute values for the object space texture elements taken from the world model, the program can then determine the correct window size so that texture element variations can be measured. Since for longitudinal surfaces we expect the elements to appear larger in the foreground the window size is selected with respect to the nearest distance (again see Fig. 3).

Now, determination of window size must be based on the largest object space texture element size the program can anticipate for longitudinal surfaces. This is accomplished by the control of the higher-level program which uses the information about the observer's position. For example, given an observer on the ground looking straight ahead, the higher-level interpretive program expects to find a ground object, and therefore a possible longitudinal surface, with a certain element size texture. Thus, the expected sizes of texture blobs in the image plane can then be calculated.

A window size is then selected which will allow at least two repetitions of the texture element so that its periodicity (spacing and wavelength) can be computed. The selection of appropriate window size with respect to foreground texture element sizes can thus be automated. We have assumed so far a fixed spatial (digitizer) resolution which has been verified a priori to be fine enough to register the variations. This assumption is made for our study because the scanner is passive. (The resolution for our experiments is discussed later.) A system with an active (controllable) scanner, however, could direct the depth of focus and field of view

of the device from the higher-level program to obtain acceptable resolution; i.e., it could accommodate to the situation (see [13]).

To detect a gradient the program requires at least two points. By this we mean

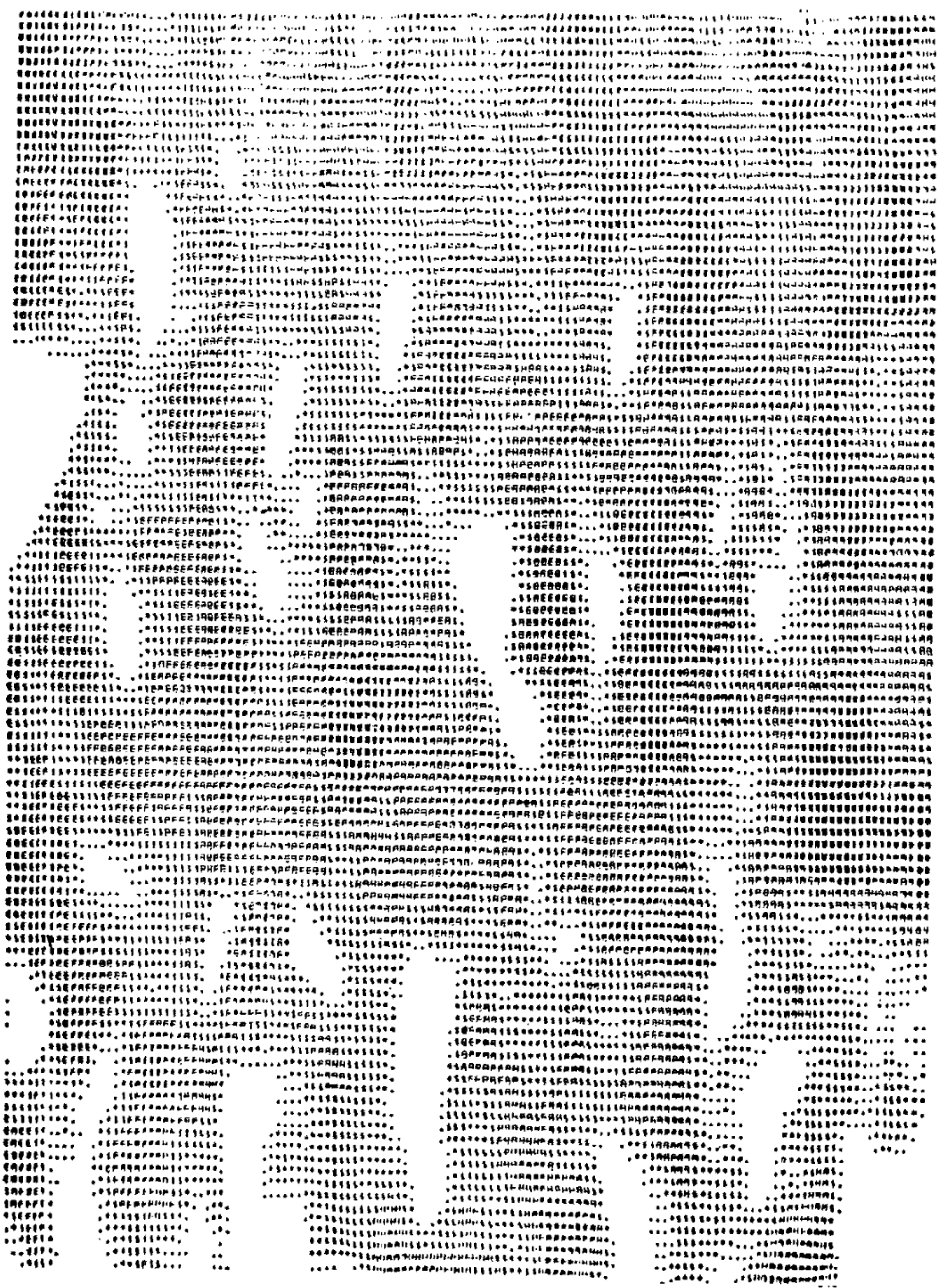

FIG. 5. A window of a grass texture.

FIG. 6. A window of ocean waves.

that the texture size it to be measured over at least two windows at positions in the image corresponding to the longitudinal (line of sight) direction. For our observer on the ground that direction is always from bottom to top in the image.

The two-window "height" can be achieved either by choosing (vertically) adjacent windows or by overlapping them. In either case it can be seen that with a chosen window size a minimum area is needed to observe any gradient. Thus, the second restriction that the image (actually the portion of the image which is the longitudinal surface) have at least that minimal, two-window extent is obtained.

The psychological model of gradients as depth cues also tells us that no gradient is perceived from texture elements on a longitudinal surface along the direction transverse to the line of sight. This implies that there is a continuity of texture features horizontally in the image.

5. TEXTURE GRADIENT RESULTS

The "ground" regions of Figs. 1 and 2 were hypothesized in the higher-level program by color, size, and position. The texture verification only dealt with the class of texture expected for a grassy field or ocean waves. Each of these objects is a class of "ground" and therefore they constitute longitudinal surfaces (with respect to an observer on the ground looking straight ahead). Additional verification is accomplished by searching for a gradient and thereby making some general 3-D inferences.

For both of these scenes the window size was chosen as follows. The exact parameters of the observer and the actual size of texture elements are not known in the current experiment. Line printer pictures of the total brightness function of various foreground sections of the image were examined manually. Figures 5 and 6 are typical segments. It was visually verified that at least two repetitions of the texture elements appeared both vertically and horizontally for a specific window size. The size chosen was a power of 2 (for efficient FFT calculation) with

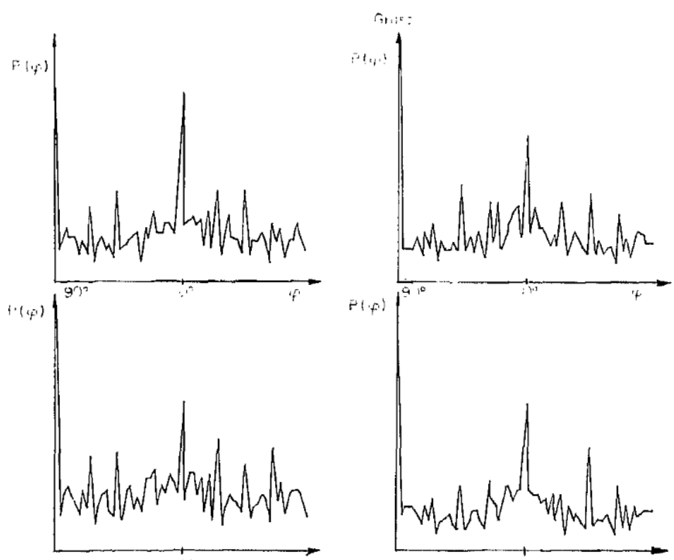

FIG. 7. Directionality for a grass texture.

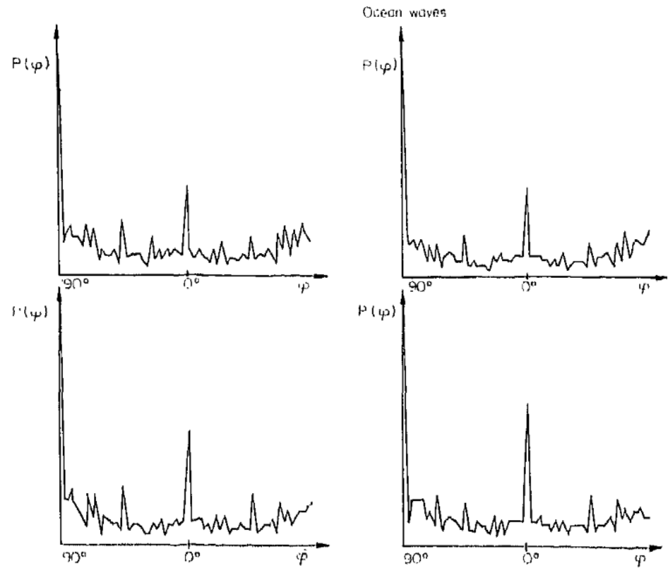

FIG. 8. Directionality for ocean waves.

sufficient area. For both Figs. 1 and 2 the size was 128 × 128. We chose to use adjacent (nonoverlapping) windows and found that the "height" of the grassy region was 3 windows while that of the ocean was 5 windows, thus satisfying the second image condition. There are 17 windows across each row; in the grassy field the rows extend from the bottom edge of the image to the bottom of the

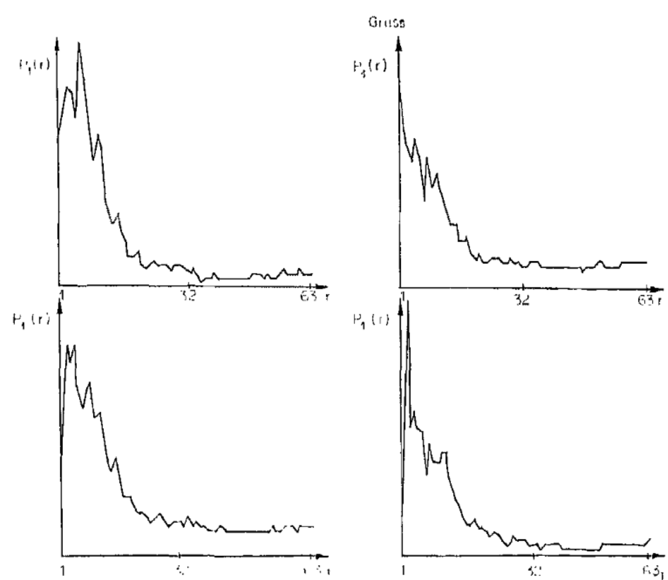

FIG. 9. The frequency spectrum in the vertical direction for a grass texture.

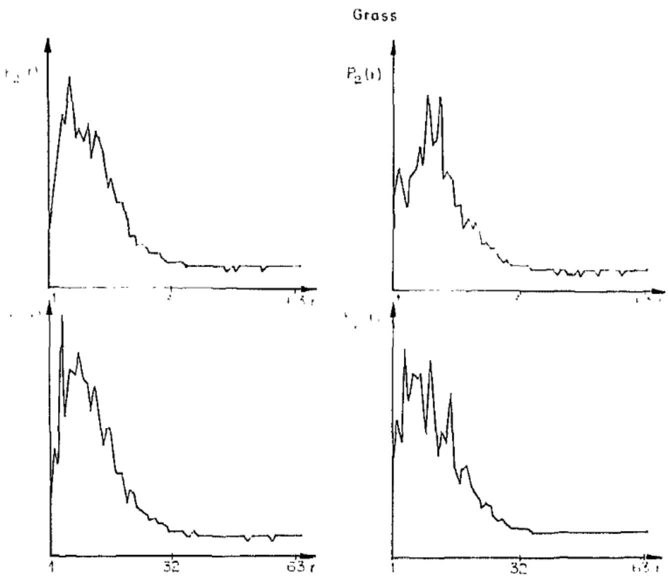

FIG. 10. The frequency spectrum in the horizontal direction for a grass texture.

tree line, and in the ocean scene the rows begin directly above the rocks and end at the skyline.

The texture of the grass is "elongated bloblike" with the major axis of the texture elements oriented vertically. The ocean waves also have an elongated bloblike texture, but with a horizontal prime axis. The texture of the waves appears as reflected blobs rather than long reflected lines because of the choppiness of the ocean at the time. The power spectrum for each window shows that the prime directionalities are 90° and 0° (from horizontal) which will be referred to as directions 1 and 2, respectively. $P(phi)$ for a grassy segment is shown in Fig. 7. A typical ocean segment $P(phi)$ is given in Fig. 8.

Figures 9 and 10 are typical $P(r)$ for grass in directions 1 and 2, respectively. The threshold for significant peaks is equal to the mean energy $+1.5 \times$ (standard deviation), where the statistics are calculated from the $P(r)$ in the window. Figure 11 is a chart of the texture wavelengths in direction 1 in each window over the grassy area. Each value in the chart (except those in the mean wavelength

Top of image (background)																	Mean wavelength
26	26	26	26	26	26	26	26	26	26	21	26	26	26	26	26	26	26
32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
32	32	32	32	32	32	32	32	32	32	32	32	32	32	26	32	32	32
Bottom of image (foreground)																	

FIG. 11. Texture wavelength chart. Grass texture in vertical direction. 128^2 windows. (*Note:* Wavelengths in the chart correspond to "preferred" frequencies in that region.)

Top of image (background)																	Mean
26	18	18	18	18	18	18	18	18	18	15	18	18	18	18	18	18	18
32	21	21	21	21	21	21	18	21	18	21	21	18	21	21	21	21	21
26	26	26	26	26	26	32	26	26	26	26	26	21	26	21	26	26	26
Botton of image (foreground)																	

FIG. 12. Texture wavelength chart. Grass texture in horizontal direction. 128^2 windows. (*Note:* Wavelengths in the chart correspond to "preferred" frequencies in that region.)

column) is computed as $W/RMAX$. The $RMAX$ for a window is determined as follows. The absolute maximum power is found in each window of a row. The frequencies for these peaks are averaged for the row. That frequency is declared the "preferred" frequency for the texture at that vertical position (row) in the image. Each window is reexamined to find all the significant (above-threshold) energy values. The significant energy peak whose frequency is closest to the preferred frequency is chosen as $RMAX$ for that window. This is justified, and follows directly from, the principle of continuity of texture features in the direction transverse to the line of sight. The continuity of the texture descriptor from window to window across the image row is preserved by averaging the window operator results.

Thus, each of the wavelength entries is related to the $RMAX$ calculated with this principle of continuity of texture. Figure 12 is a chart of the wavelength in direction 2. A weak gradient exists in the direction 1 wavelength, but a much more significant one can be seen in the direction 2 chart. The interpretation of this result is that the texture appears to shrink in terms of horizontal spacing of the

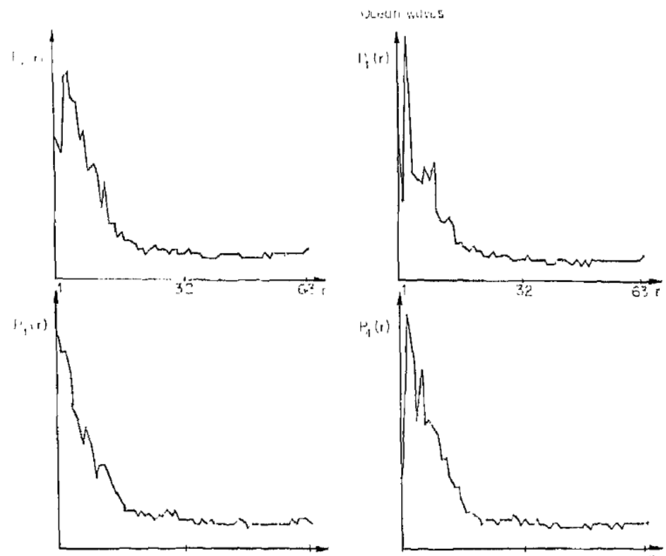

FIG. 13. The frequency spectrum in the vertical direction for ocean waves texture.

Top of image (background)																	Mean wavelength
12	21	21	18	18	21	18	21	21	21	21	21	21	21	21	21	21	20
26	26	26	26	26	26	26	26	21	26	26	26	26	26	26	26	26	26
32	32	32	32	32	32	32	32	32	26	32	32	32	32	32	32	32	32
32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
43	43	43	43	43	32	43	43	43	43	43	43	43	43	43	43	64	44
Bottom of image (foreground)																	

FIG. 14. Texture wavelength chart. Ocean waves texture in the vertical direction. 128^2 windows. (*Note:* Wavelengths in the chart correspond to "preferred" frequencies in that region.)

grass blades proceeding along the line of sight. The "height" of the blades does not shrink in the same way. We think this can be explained by considering that what the operator is measuring is the brightness of reflected blobs. These blobs in the image plane are caused by the many overlapping grass blades. In the far distance the occlusion of the blades by each other causes a reflected projection in the image which is merged vertically but not as much horizontally.

A striking example of the texture gradient is found in the ocean scene. Figure 13 is $P(r)$ in direction 1 for a typical ocean window. The ocean waves' texture wavelengths for direction 1 are given in Fig. 14. These numbers physically represent the space between waves in windows in the image. The mean wavelength varies from 51 to 21, going from near to far (line of sight).

An example of the type of 3-D inference that can be deduced from the above longitudinal surface verification is that objects touching, i.e., adjacent to, such a surface region are really "on" the surface. In Fig. 14 this means that the bushes in the right half of the scene are on the ground at a point higher in the image and are therefore further back in 3-space than the trees found at the left. Using that information and the technique outlined in Section 4, a projection curve, P^*, is calculated. We then let point A be the bottom of the image, point B the bottom of the large tree's trunk, and point C the bottom edge of the bushes in back of and to the right of the large tree. The problem is to find the ratio of the distances AC to AB; i.e., how far is it to the bushes compared to the distance to the first tree? Evaluation of integrals yields a result of 1.06 for the ratio. This seems slightly smaller than the perceived distance in Fig. 1. A possible explanation is that the surface is not flat as in our model. A rising hill in the background has been the interpretation of some people who have viewed the photograph.

6. CONCLUSION

It has been our contention throughout this work that texture is the most significant feature of outdoor scenes.

The desired texture operator for a computer must be sensitive enough to register the repetition of those features of texture elements which perceptually group them to form a textured surface. The rough shape class and size of elements and spatial organization which are measured by the Fourier operator capture these features. The texture operator must not be sensitive to fine variations in shape and size

caused either by variation or by the projection of the 3-D objects into the 2-D image plane. The Fourier operator is an averaging technique which, if applied on a sufficiently large window, is especially suited to natural, outdoor textures.

Our experiments have shown that the Fourier texture operator is well suited to natural scene description. Three-dimensional surface interpretation is made possible through the use of texture gradient information obtained by higher-level correlation of the texture descriptors. It must be clear, however, that the texture gradient information of the surface can be used as a depth cue only if there is an additional knowledge about the world which can lead to a hypothesis about the surface. In other words, we need the world model, in particular the model of the observer and the scene, in order to make three-dimensional interpretations.

The associations of texture gradients with longitudinal surfaces is, of course, not limited to outdoor scenes. The principles of the mapping of the visual world into the visual field (image space) are universal.

ACKNOWLEDGMENTS

The authors are indebted to Dr. A. Rosenfeld and Mr. Andrew Pilipchuk of the University of Maryland Computer Science Center, who provided digitization of the photographs used in this study. They also thank Ms. M. Brown and Ms. R. Schultz for help in drawing the figures.

The work described in this paper has been supported in part by the National Science Foundation under Grant GJ-43440.

REFERENCES

1. R. Bajcsy, Computer identification of textured visual scenes, Memo AIM-180, Stanford University, October 1972.
2. E. S. Deutsch and N. J. Belknap, Texture descriptors using neighborhood information, *Computer Graphics Image Processing* 1, 1972, 145–168.
3. J. J. Gibson, *The Perception of the Visual World*, Houghton Mifflin, Boston, 1950.
4. J. J. Gibson, What is form?, *Psychol. Rev.* **58**, 1951, 430–412.
5. K. C. Hayes, A. N. Shah, and A. Rosenfeld, Texture coarseness: Further experiments, *IEEE Trans. Systems, Man, Cybernetics* SMC-September 1974, 467–472.
6. R. M. Haralick, K. Shanmugam, and I. Dinstein, Textural Features for Image Classification, *IEEE Transactions Systems, Man, Cybernetics* SMC-**3**, 1973, 610–621.
7. B. K. P. Horn, Shape From Shading: A method for obtaining the shape of a smooth opaque object from one view, MAC-TR-79, MIT, November 1970.
8. L. Kaufman, *Sight and Mind: An Introduction to Visual Perception*, Oxford University Press, New York, 1974.
9. L. J. Krakauer, Computer analysis of visual properties of curved objects, MAC-TR-82, MIT, May 1971.
10. L. I. Lieberman, Computer recognition and description of natural scenes, Moore School-TR-7408, University of Pennsylvania, June 1974.
11. R. M. Pickett, Visual analysis of texture in the detection and recognition of objects, in *Picture Processing and Psychopictorics* (B. S. Lipkin and A. Rosenfeld, Eds.), Academic Press, New York, 1970.
12. A. Rosenfeld, Picture Processing: 1973, TR-284, University of Maryland, Computer Science Center, January 1974.
13. J. M. Tenenbaum, Accommodation in computer vision, AIM-134, Artificial Intelligence Project, Stanford University, September 1970.