

A Multiscale Region Detector*

DOROTHEA BLOSTEIN[†] AND NARENDRA AHUJA

*Coordinated Science Laboratory, University of Illinois, 1101 West Springfield Avenue,
Urbana, Illinois, 61801*

Received June 8, 1987; accepted June 1, 1988

A multiscale region detector is derived from properties of a $\nabla^2 G$ (Laplacian-of-Gaussian) scalespace. The detector finds all image regions that have small gray-level variation relative to a neighborhood of their size. These regions may be nested. Since the region detector is designed as a coarse estimator of region size and shape, it is better suited for texture element extraction than for general image segmentation. Uniform gray level regions are represented by a configuration of best-fitting disks of appropriate sizes and locations contained within the regions. The sizes and locations of best disks are estimated from the observed response of the $\nabla^2 G$ operator across the image and across scales, using an analytic expression for the response of $\nabla^2 G$ to disks. Both positive-contrast and negative-contrast regions are detected, and an estimate of region contrast is computed. Results are shown for a variety of natural textures.

* 1989 Academic Press, Inc.

1. INTRODUCTION

A low-level operation often required for image analysis is the extraction of regions of relatively uniform gray level. The desirable properties of a region detector depend on the application. In this paper we focus on region detection for texture analysis. Region detectors used in other contexts are described by Haralick and Shapiro [8] (image segmentation) and Schachter [18], Hartley *et al.* [9] (target detection). For texture analysis, a region detector must respond to densely placed texture elements, which may vary greatly in size (especially when the texture is seen under perspective). Regions used to represent texture elements may be nested, as when subtexture is present.

Several approaches to texture-element extraction have been described in the literature. Wang *et al.* [21] describe three methods based on gray-level thresholding: (1) thresholding with a fixed percentage of pixels above threshold, (2) histogram peak sharpening, and (3) the "superslice" algorithm, which evaluates thresholds based on agreement with an edge map. Yet another threshold-selection method is used by Matsuyama *et al.* [15]: their criterion for threshold selection is that many regions with similar properties be produced. Voorhees and Poggio [20] extract texture elements by convolving the image with a $\nabla^2 G$ filter and then selecting components of above-threshold pixels that have suitable geometric properties, such as compactness. Vilnrotter *et al.* [19] construct edge repetition arrays to estimate the size and spacing of texture elements; these estimates are then used to extract the texture elements. Maleson *et al.* [12] construct elliptical approximations to regions of uniform gray level. Davis [7] detects texture elements of known shape using a Hough transform of the edge map. Texture coarseness has been estimated by

*The support of the Air Force Office of Scientific Research under Grant AFOSR 86-0009 is gratefully acknowledged. Part of the work was supported by a graduate fellowship from IBM.

[†]Currently at Department of Computer and Information Science, Queen's University, Kingston, Ontario K7L 3N6 Canada.

comparing the outputs of coarse and fine spot detectors [10, 23]. Hong *et al.* [11] extract texels by grouping facing pairs of edge points into region boundaries.

Most of the texture-element extraction methods reviewed above were applied to texture discrimination problems. Many of these methods rely on a frontal view of the texture, where the average texture-element size does not vary throughout the image. Such methods are not suitable for the more difficult task of extracting texture elements when texture is seen under perspective, as must be done for obtaining surface orientation information from a perspective view of the texture. Ohta *et al.* [17] discuss one approach to this problem, but do not address the problem of finding texture elements in real images. The region detector we describe here has been used to extract texture gradients in images of natural textures, such as the ones shown in parts (a) of Figs. 1 to 6 (Blostein and Ahuja [3]). The goal is to obtain estimates of the location, size, shape, and contrast of the texture elements. The detector does not precisely localize region boundaries and hence may not be suitable for image segmentation. However, since texture elements are inherently variable in size and shape, an approximate detection of texture-element shape is adequate for many applications. Our detector does not rely on threshold selection. It is tolerant of gray-level noise and insensitive to lighting changes. Uniform image regions of all sizes are detected; these regions may be nested.

2. OVERVIEW OF THE REGION DETECTOR

Uniform image regions vary in size, shape, and contrast. The uniformity and contrast of small regions have significance when measured relative to a small surrounding neighborhood in the image, whereas the uniformity and contrast of large regions have significance relative to a proportionally larger image neighborhood. A uniform image region is composed of gray levels which are uniform relative to the gray-level changes occurring at its own scale; such a region may nevertheless contain internal variations of gray level that are significant at a smaller scale. In applications where the size of interesting regions is unknown in advance, the region detector must be capable of identifying nested regions. The design of a region detector is complicated by the fact that uniform image regions may have any shape.

The Laplacian-of-Gaussian ($\nabla^2 G$) filter has been useful in a variety of multiscale image analysis tasks; it has desirable noise suppression properties and is suitable for local image analysis as in edge detection and spot detection. As described in Section 3, a scale space of $\nabla^2 G$ -convolved images can be used to robustly extract all circular image regions of relatively uniform gray level. This disk detection is performed by comparing the local image response to the response expected for an image of an ideal circular region of constant gray level. Unions of overlapping and adjacent disks are used to represent arbitrary regions (Section 4): large disks define the rough shape of a region, with overlapping smaller disks capturing finer shape details such as protrusions and concavities. Section 5 describes an implementation of the region detector; the performance of the region detector is analyzed in Section 6.

3. FINDING CIRCULAR DISKS WHICH BEST FIT UNIFORM IMAGE REGIONS

The region-extraction algorithm operates by finding best-fitting disks and then constructing unions of overlapping disks to describe uniform image regions. The

disk-fitting process is based on an analysis of the *scale-space* behavior of uniform image regions. The term *scale space*, introduced by Witkin [22], denotes a representation of the image response to convolution with $\nabla^2 G$ filters over a range of scales. A scale-space representation of a signal has many features that could be analyzed. Witkin [22] concentrates on the behavior of zero crossings over a range of scales, whereas Crowley and Parker [6] represent peaks and ridges extracted over a range of scales.¹ We summarize these papers before turning to our own scale-space analysis.

3.1. Scale Space

Witkin [22] builds on the theory of edge detection developed by Marr and Hildreth [13] (see also Marr [14]) in which edges are located as the zero crossings in the Laplacian of a Gaussian-smoothed image. Marr and Hildreth suggest using a selection of filter in order to capture edges at different scales, but do not adequately address the problem of combining the edge images obtained from various filter sizes. Witkin introduces a scale-space representation of $\nabla^2 G$ zero crossings over a continuous range of scales. A scale-space representation is constructed by convolving the original signal with $\nabla^2 G$ filters for all possible choices of the filter size σ . The scale-space representation of 1-dimensional signal occupies an x - σ plane, whereas the scale-space representation of a 2-dimensional signal (such as an image) occupies an x - y - σ volume. Gaussian smoothing has two effects: simplification through removal of fine-scale features and distortion through dislocation, broadening, and flattening of the surviving features. Salient zero-crossing contours may be identified at coarse scales and then traced to fine scales for accurate localization. Witkin [22] describes an efficient representation of the zero crossings of a 1-dimensional signal (in the x - σ plane). Extending this representation to encode the zero crossings of 2-dimensional signals is difficult.

Crowley and Parker [6] represent scale-space peaks and ridges. The pattern of scale-space peaks and connecting ridges characterizes object shape in a form that is suited to object recognition or matching: the coarse shape information captured by the large filter sizes is used to bring the objects into approximate registration, and then the more detailed shape information captured by the small filter sizes is used to refine the matching.

As described in Section 3.2., we use the $(\partial/\partial\sigma)\nabla^2 G$ filter as part of the region-detection algorithm, for analyzing the behavior of images across scales.² The other partial derivatives of $\nabla^2 G$, $(\partial/\partial x)\nabla^2 G$ and $(\partial/\partial y)\nabla^2 G$, are commonly used to calculate edge strength at zero crossings [14]. Normally x and y partial derivatives are estimated by subtracting neighboring pixels in the $\nabla^2 G * I$ image. The equivalent method of estimating the σ partial is to subtract the $\nabla^2 G$ convolution outputs at two different filter sizes. However, such discrete derivative approximations are very error prone, especially if the neighboring σ values are not very close. Direct convolution with a $(\partial/\partial\sigma)\nabla^2 G$ filter gives accurate results and allows a rather sparse set of σ values to be used to sample the scale space.

¹Crowley and Parker use a difference-of-Gaussian operator, which is a discrete approximation to $(\partial/\partial\sigma)G$ and hence to $\nabla^2 G$. By the diffusion equation, $\nabla^2 G = (1/\sigma)(\partial/\partial\sigma)G$.

²We originally started investigating this filter as a tool for analyzing the movement of $\nabla^2 G$ zero crossings across scales.

3.2. A Closed-Form Expression for the $\nabla^2 G$ Response of a Disk Image

We now describe a method of estimating the size of uniform disks from the variation of $\nabla^2 G$ response with scale. As discussed in Section 4, unions of overlapping disks constitute the final output of the region detector. In order to judge how to fit disks to an image, we first analyze the $\nabla^2 G$ response expected from an idealized image of a circular region on a uniform background. We use this ideal response as a reference for fitting disks to an arbitrary image. Two parameters are needed to describe a disk at a given image location: disk radius and disk contrast. At least two image measurements must be made to recover these two parameters. The two measurements we use are (1) the $\nabla^2 G$ response and (2) the $(\partial/\partial\sigma)\nabla^2 G$ response, both measured at the center of the disk. We now derive the expected values of these two measurements for an ideal disk.

Given a function $I(x, y)$ which specifies the intensity of an image, the $\nabla^2 G$ response³ of this image at (x, y) is given by the convolution

$$\nabla^2 G(x, y) * I(x, y) = \iint_{-\infty}^{+\infty} \frac{2\sigma^2 - (u^2 + v^2)}{\sigma^4} e^{-(u^2 + v^2)/2\sigma^2} I(x - u, y - v) du dv. \quad (1)$$

Mathematical analysis of the response of the $\nabla^2 G$ filter to most images is difficult because the convolution integrals of Eq. (1) do not have closed-form solutions. However, a closed-form solution can be derived for the center point of a circular disk of constant intensity. We analyze the $\nabla^2 G$ response at the center of an ideal circular disk in the continuous domain; to generate the $\nabla^2 G$ convolution of digitized images, we sample the $\nabla^2 G$ filter values and perform a discrete convolution. The image of a disk of diameter D and contrast C is defined by

$$\text{disk image: } I(x, y) = \begin{cases} C & \text{if } x^2 + y^2 \leq D^2/4 \\ 0 & \text{elsewhere} \end{cases} \quad (2)$$

We set x and y to zero in Eq. (1) to derive the $\nabla^2 G$ response at the center of the disk (see [2] for a more detailed derivation):

$$\nabla^2 G(x, y) * I(x, y) = \iint_{-\infty}^{+\infty} \frac{2\sigma^2 - (u^2 + v^2)}{\sigma^4} e^{-(u^2 + v^2)/2\sigma^2} I(-u, -v) du dv.$$

Changing to polar coordinates and using $I(x, y)$ from Eq. (2),

$$\begin{aligned} \nabla^2 G(x, y) * I(x, y) &= \frac{C}{\sigma^4} \int_0^{D/2} \int_{-\pi}^{\pi} (2\sigma^2 - \rho^2) e^{-\rho^2/2\sigma^2} \rho d\theta d\rho \\ &= \frac{2\pi C}{\sigma^4} \left(\int_0^{D/2} 2\rho\sigma^2 e^{-\rho^2/2\sigma^2} d\rho - \int_0^{D/2} \rho^3 e^{-\rho^2/2\sigma^2} d\rho \right). \end{aligned}$$

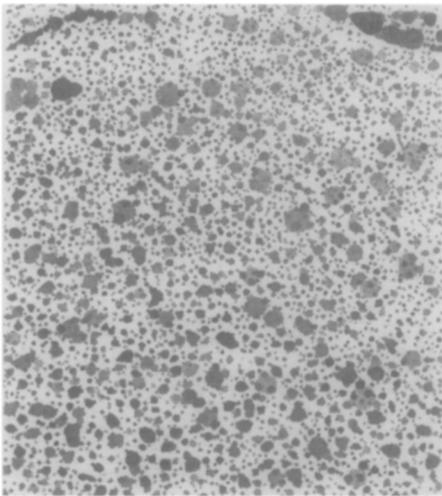
³The expression we use for $\nabla^2 G$ is the Laplacian of an unnormalized Gaussian ($e^{-r^2/2\sigma^2}$). In keeping with tradition in the literature, we negate the $\nabla^2 G$ equations, so that filters with a positive center lobe result.

Using integration by parts on the second integral, we cancel the first integral to obtain

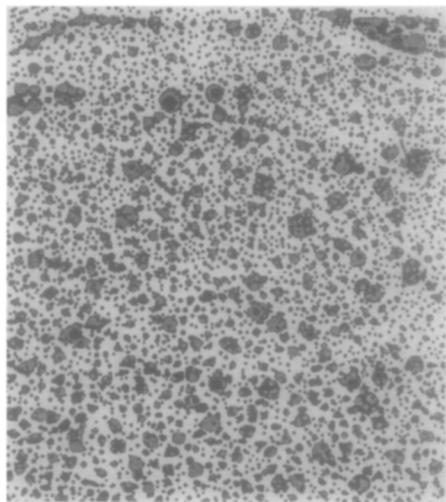
$$\nabla^2 G \text{ response at the center of the disk} = \frac{\pi CD^2}{2\sigma^2} e^{-D^2/8\sigma^2}. \quad (3)$$



(a)



(b)



(c)

FIGS. 1a-c. (a) The rock pile. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. This set of disks includes all of the disks from (g), (i), (k), and (m). At pixel locations covered by several disks, the disk of higher contrast is displayed. (c) Positive-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray and doubly-nested regions are again black. (Regions nested more than two levels deep have been omitted.)

Taking the derivative with respect σ , we obtain

$$\frac{\partial}{\partial \sigma} \nabla^2 G \quad \text{response at the center of the disk} = \frac{\pi C D^2}{2} \left(\frac{D^2}{4\sigma^5} - \frac{2}{\sigma^3} \right) e^{-D^2/8\sigma^2}. \quad (4)$$

These equations have been verified experimentally using discrete convolutions of $\nabla^2 G$ and $(\partial/\partial\sigma)\nabla^2 G$ masks with synthesized images of isolated disks. The $\nabla^2 G$ and $(\partial/\partial\sigma)\nabla^2 G$ values at the centers of the disks match the values predicted by the equations to within roundoff and discretization errors.

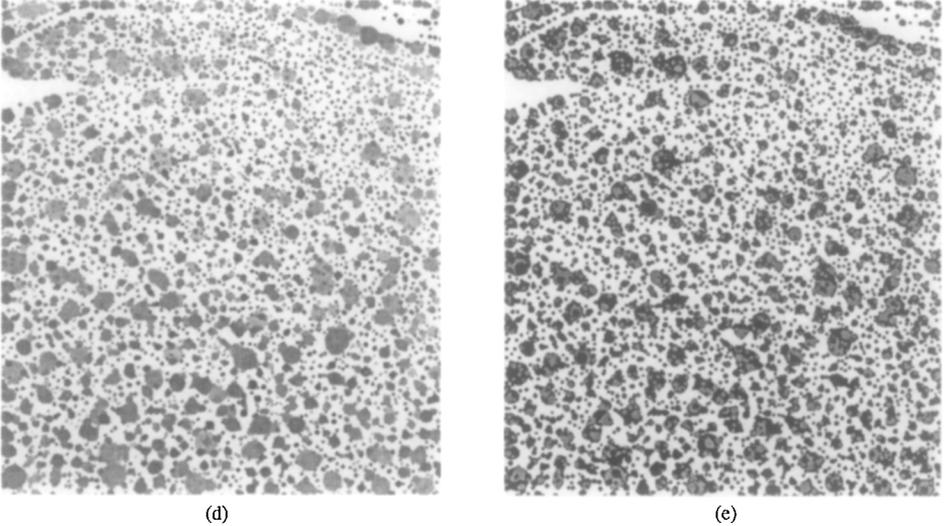
If we have measured values for $\nabla^2 G * I$ and $(\partial/\partial\sigma)\nabla^2 G * I$ at the center of the disk, the disk diameter and contrast can be recovered using Eqs. (3) and (4). First the disk diameter D is calculated as

$$D = 2\sigma \sqrt{\left\{ \sigma \left(\frac{\partial}{\partial \sigma} \nabla^2 G * I \right) / (\nabla^2 G * I) \right\} + 2}. \quad (5)$$

Substituting D into Eq. (3), we obtain the contrast C as

$$C = \frac{2\sigma^2}{\pi D^2} e^{D^2/8\sigma^2} (\nabla^2 G * I). \quad (6)$$

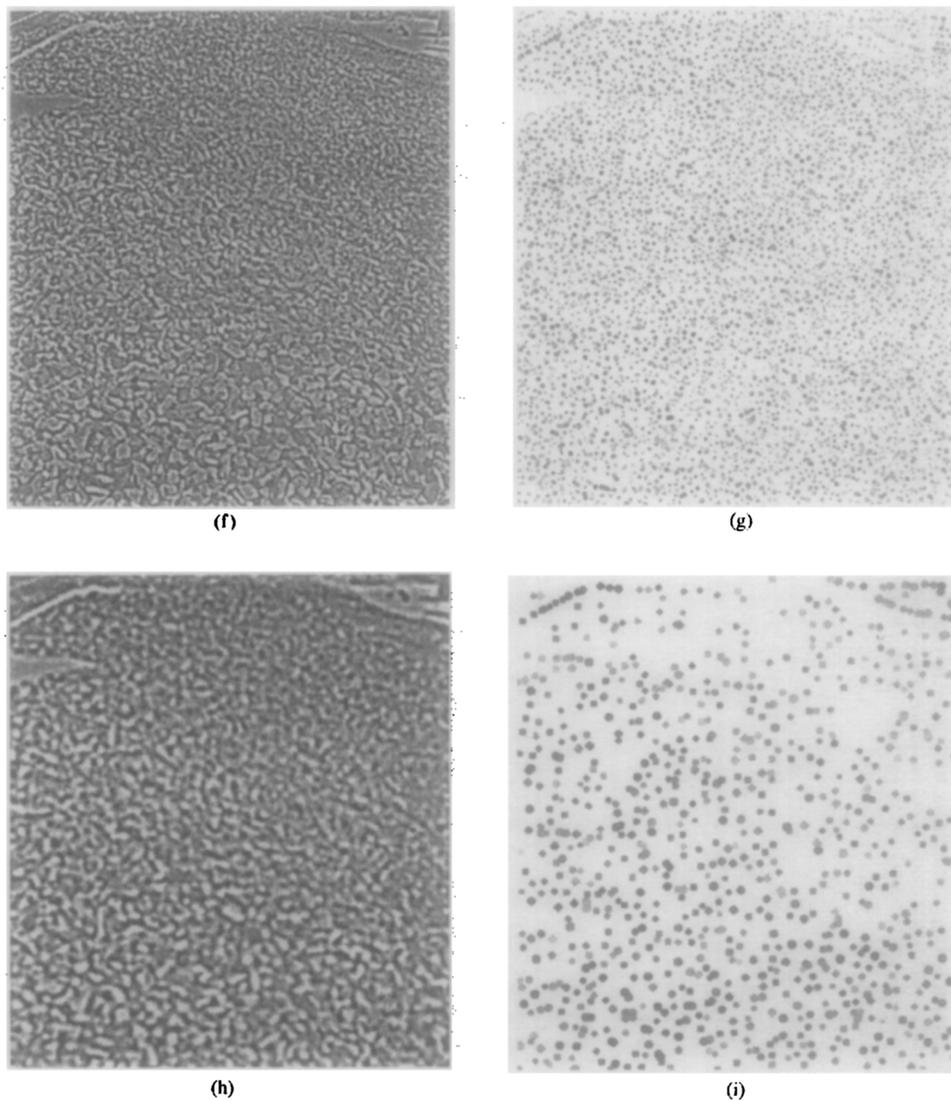
Application of these equations to a real image, where the disk center locations are not known a priori, is discussed in the next section.



FIGS. 1d, e. (d) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. At pixel locations covered by several disks, the disk of higher contrast is displayed. (e) Negative-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

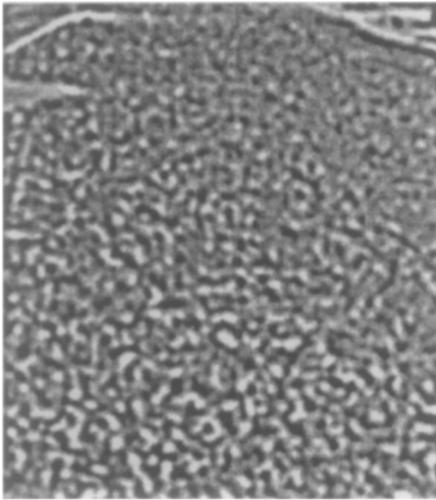
4. DETECTION OF IMAGE REGIONS OF RELATIVELY UNIFORM GRAY LEVEL

An arbitrary region of relatively uniform gray level can be represented by a configuration of uniform disks, each of which is locally maximal, i.e., contained within the region but not within any other uniform disk [4; 16, Chap. 11]. Locally maximal disks, in general, have significant overlap. Thus, each disk is surrounded not by background alone, but both by background and other locally maximal disks:

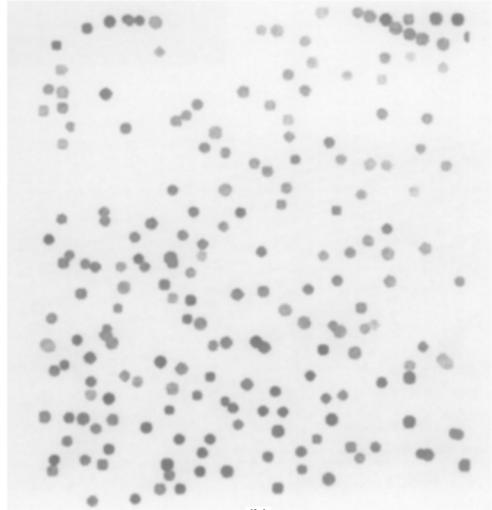


FIGS. 1f-i. Details of the disk-fitting process for positive-contrast regions. (f) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma = \sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 4 pixels. (g) Disks detected at this filter size. The disk diameters range from 2 to 6 pixels. (h) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma = 2\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 8 pixels. (i) Disks detected at this filter size. The disk diameters range from 6 to 10 pixels.

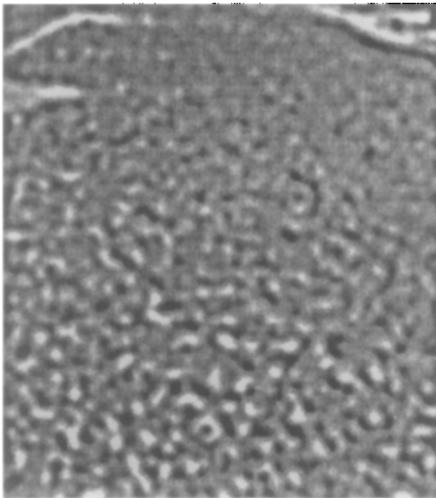
a disk usually touches the region boundary at two points or along a small fraction of its circumference. Since these disks do not fit the model of Eq. (2), the analysis for estimation of disk diameter and contrast of Section 3 cannot be applied without error. Often, Eq. (5) leads to an overestimation of the disk diameter when applied at the center point of a maximal disk. In Section 6 we discuss the extent of the overestimation error. In the rest of the paper we will use Eqs. (5) and (6) to estimate the diameter and contrast of the locally maximal disks, ignoring the estimation error.



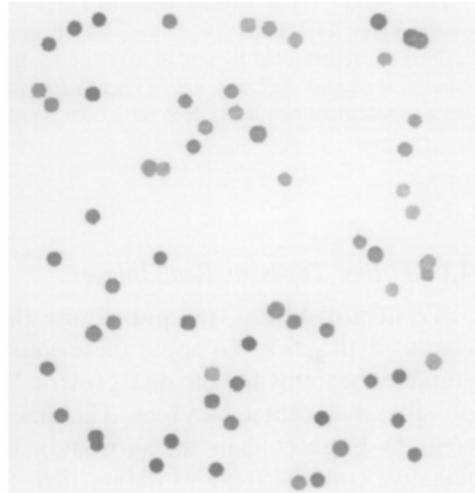
(j)



(k)

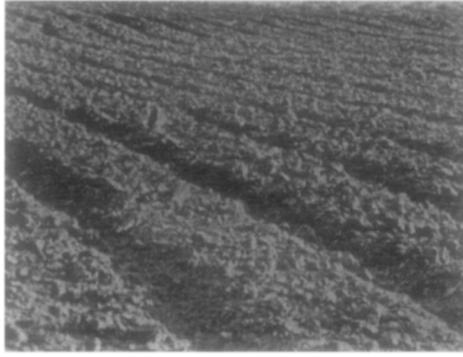


(l)

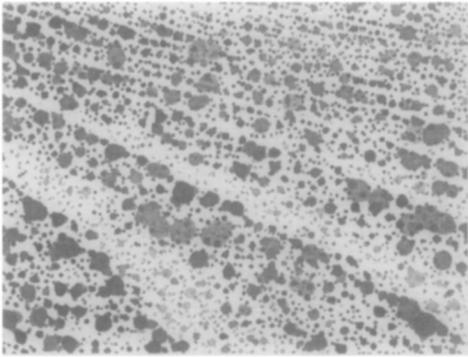


(m)

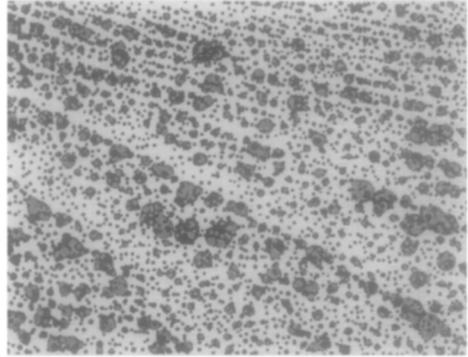
FIGS. 1j–m. (j) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma = 3\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 12 pixels. (k) Disks detected at this filter size. The disk diameters range from 10 to 14 pixels. (l) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma = 4\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 16 pixels. (m) Disks detected at this filter size. The disk diameters range from 14 to 18 pixels.



(a)



(b)



(c)

FIGS. 2a–c. (a) A plowed field. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (c) Positive-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

4.1. Fitting Disks to Real Images

To fit a disk, Eqs. (5) and (6) use the $\nabla^2 G * I$ and $(\partial/\partial\sigma)\nabla^2 G * I$ values at the center of the disk. To apply these equations to real images, we must first identify suitable locations for the disk centers. We choose extrema of the $\nabla^2 G * I$ images as possible disk-center locations. The disks fit to local maxima have positive contrast (regions brighter than the surround), whereas the disks fit to local minima have negative contrast (regions darker than the surround). Using extrema of $\nabla^2 G * I$ as disk centers is justified by the following considerations. For a given region, local $\nabla^2 G$ extrema occur at the centers of locally maximal disks for filters whose size matches the diameter of the disk ($w \approx D$, i.e., $D \approx 2\sqrt{2}\sigma$).⁴ If σ is too small for the

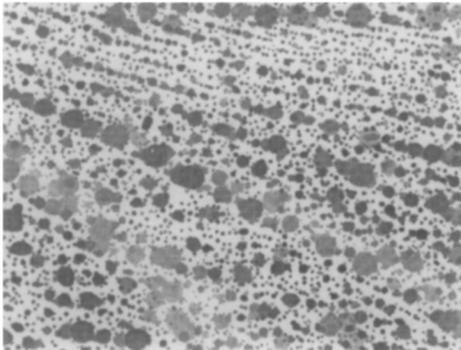
⁴The size of a $\nabla^2 G$ filter is characterized by σ , the standard deviation of the Gaussian distribution, or by w , the width of the positive center lobe of the $\nabla^2 G$ filter; $w = 2\sqrt{2}\sigma$.

given region, then the extrema occur off center (and application of Eq. (5) underestimates the region diameter). If σ is too large for the given region, then Gaussian smoothing merges neighboring regions, making Eq. (5) inappropriate for representing regions at the scale of the given region. Applying Eqs. (5) and (6) at $\nabla^2 G$ extrema for a selection of filter sizes assures that some σ value falls into the range at which it is appropriate to analyze the local shape of each region. To fit disks as accurately as possible, we accept a disk only if the computed diameter D is close to the width of the $\nabla^2 G$ center lobe: $D \approx 2\sqrt{2}\sigma$. If the latter condition does not hold for a given filter size σ and a given $\nabla^2 G$ extremum, the region size is more accurately analyzed using a different σ value.

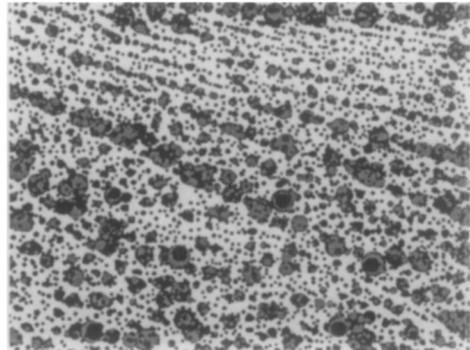
4.2. Representing Image Regions as a Union of Disks

After all disks have been detected for a particular image, sets of overlapping disks are created to describe uniform image regions. Typically a few large disks model the basic shape of the region, with smaller disks near the region border corresponding to protrusions in the region shape. A large detected region has gray levels that are uniform relative to its own scale; it may nevertheless contain internal variations of gray levels that are uniform relative to a smaller scale. Nested regions are constructed whenever small disks are spatially contained within larger regions. The contrast of a region as a whole is computed as a weighted average of the contrasts of the disks constituting the region.

It may be noted that the detected regions often serve as the starting point for further analysis, where the goal is to extract image structure defined not only by gray-level criteria, but also by other constraints or world knowledge. For example, identification of texture elements in perspective views involves enforcing consistency in the geometric properties of the texture elements [3]. Here, however, we restrict our attention to detecting regions having relatively uniform gray levels.



(d)



(e)

FIGS. 2d, e. (d) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (e) Negative-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

5. IMPLEMENTATION

The first step in processing an image I is to compute $\nabla^2 G^* I$ and $(\partial/\partial\sigma)\nabla^2 G^* I$ for a selection of filter sizes. To compute $\nabla^2 G^* I$ for a particular σ value, the image is convolved with a mask whose coefficients are taken from

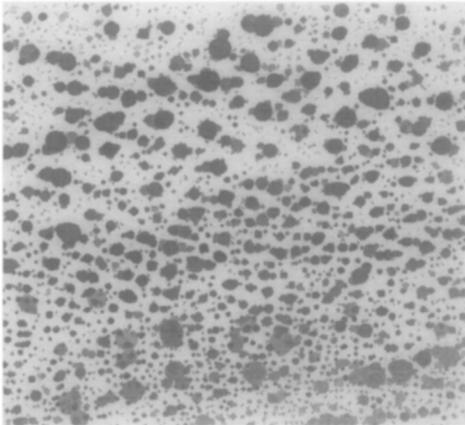
$$\frac{2\sigma^2 - r^2}{\sigma^4} e^{-r^2/2\sigma^2}.$$

To compute $(\partial/\partial\sigma)\nabla^2 G^* I$ for a particular σ value, the image is convolved with a mask whose coefficients are taken from

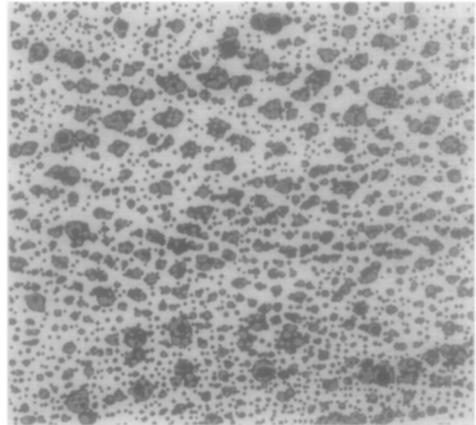
$$\frac{6r^2\sigma^4 - r^4 - 4\sigma^4}{\sigma^7} e^{-r^2/2\sigma^2}.$$



(a)



(b)



(c)

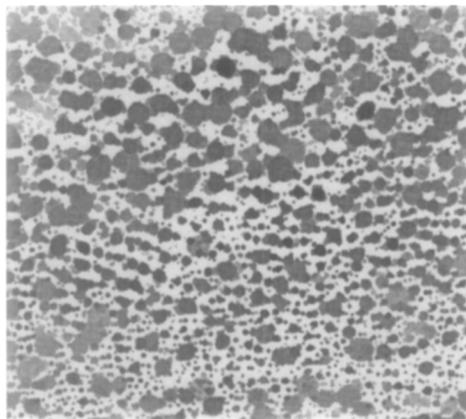
FIGS. 3a-c. (a) Fleecy clouds. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (c) Positive-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

In our current implementation, six different $\nabla^2 G$ and $(\partial/\partial\sigma)\nabla^2 G$ convolutions are evaluated, using σ values of $\sqrt{2}$, $2\sqrt{2}$, $3\sqrt{2}$, $4\sqrt{2}$, $5\sqrt{2}$, and $6\sqrt{2}$. The center lobes of the six $\nabla^2 G$ filters have diameters of 4, 8, 12, 16, 20, and 24 pixels, respectively. We perform the convolutions via multiplication in the Fourier domain. This is the only computation-intensive step of the region detection. Computation time can be reduced by using algorithms designed for simultaneous computation of $\nabla^2 G$ convolutions for a range of σ values. Burt [5] and Adelson and Burt [1] present efficient recursive algorithms for computing convolutions with Gaussians of many different sizes. Differences of Gaussians can be used to construct approximations to $\nabla^2 G$ convolutions. Note that the $(\partial/\partial\sigma)\nabla^2 G * I$ values are needed only at sparse image locations (at extrema in the $\nabla^2 G * I$ images) and could therefore be efficiently computed in the spatial domain.

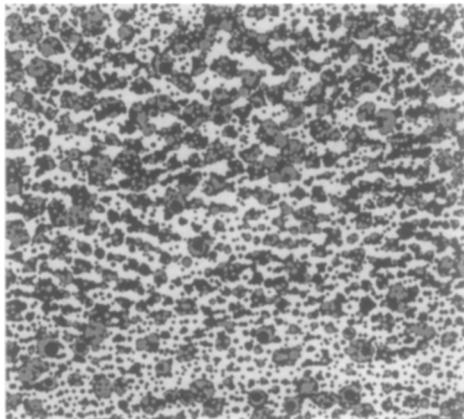
The second step in processing the image I is to mark the locations where disks will be fit. In order to find disks that model positive-contrast image regions, each $\nabla^2 G * I$ image is scanned to find local maxima: any pixel larger than all eight of its neighbors is marked as a disk-center location. Similarly, in order to find disks that model negative-contrast image regions, each $\nabla^2 G * I$ image is scanned to find local minima: any pixel smaller than all eight of its neighbors is marked as a disk-center location.

Next, disk-diameter D and disk-contrast C are computed at each of the marked locations, using the $\nabla^2 G * I$ and $(\partial/\partial\sigma)\nabla^2 G * I$ values observed at that location (Eqs. (5) and (6)):

$$D = 2\sigma \sqrt{\left\{ \sigma \left(\frac{\partial}{\partial\sigma} \nabla^2 G * I \right) \right\} / (\nabla^2 G * I)} + 2, \quad C = \frac{2\sigma^2}{\pi D^2} e^{D^2/8\sigma^2} (\nabla^2 G * I).$$



(d)



(e)

FIGS. 3d, e. (d) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (e) Negative-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

Disks are detected most accurately at a filter size close to their diameter ($D \approx w = 2\sqrt{2}\sigma$); therefore only a restricted range of disks diameters is accepted from each filter size. In the current implementation, the detected disk diameter must be within two pixels of the filter size. (Adjacent filter sizes differ by four pixels.) For example, of the disks detected by the filter of width 12 pixels ($\sigma = 3\sqrt{2}$), we keep only those with diameters in the range 10 to 14 pixels.

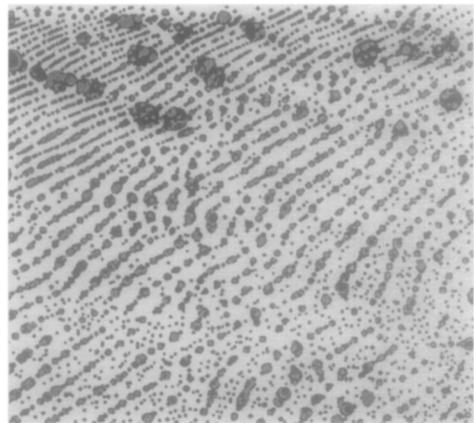
The details of disk detection (the first step in region detection) are illustrated in Fig. 1 for an image of a rock pile. Parts (b) and (f) to (m) of Fig. 1 show the positive-contrast disks in detail; the disks detected at various filter sizes are shown separately. The image area occupied by each disk is given an intensity proportional to the disk contrast. Note that more disks are detected at small filter sizes than at



(a)



(b)



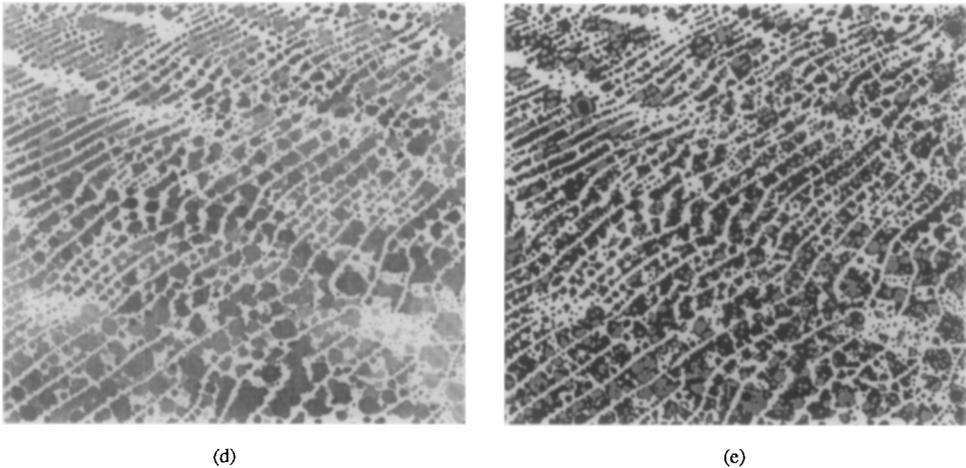
(c)

FIGS. 4a-c. (a) Ripple marks in a shallow area. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (c) Positive-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

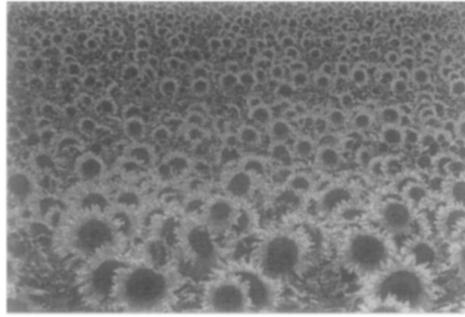
large filter sizes: the expected distance between $\nabla^2 G * I$ zero crossings is proportional to σ [14, p. 136], and hence the density of local maxima (or minima) is proportional to $1/\sigma^2$. Part (b) of Fig. 1 shows the complete set of disks detected for the rockpile image. Only one disk can be displayed at pixel locations covered by several disks; since we display the disk of higher contrast, low-contrast disks that are spatially contained within high-contrast disks are not visible. Sets of overlapping disks are used to represent complete image regions; the rockpile regions are illustrated in part (c) of Fig. 1.

6. PERFORMANCE ANALYSIS

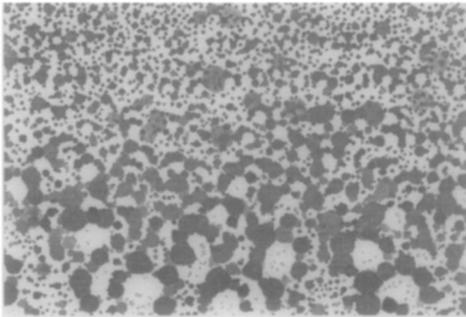
Figures 1 to 6 illustrate the performance of the region extraction on a variety of images. (These are a subset of the seventeen images presented in Blostein [2].) The images contain positive-contrast regions, which are brighter than the surround, as well as negative-contrast regions, which are darker than the surround. The positive-contrast and negative-contrast regions detected in each image are illustrated separately. Part (a) of each figure shows the original image. The images are rephotographed illustrations from books; they were digitized off the photographic negatives using a drum scanner. The images are 512 by 512 pixels; the image sizes in the figures vary because image borders have been trimmed. Parts (b) and (c) illustrate the positive-contrast regions: the maximal disks are shown in (b) and the complete regions are shown in (c). Parts (d) and (e) similarly illustrate the negative-contrast regions. The disks in parts (b) and (d) are shown with an intensity proportional to disk-contrast. In these displays, preference is given to high-contrast disks wherever several disks overlap. The regions in parts (c) and (e) are shown with intensities that were chosen to display nested regions well: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. To reduce



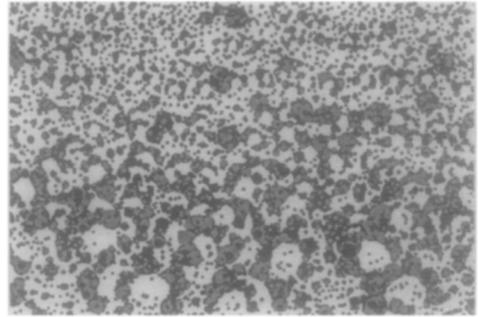
FIGS. 4d,e. (d) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (e) Negative-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)



(a)



(b)



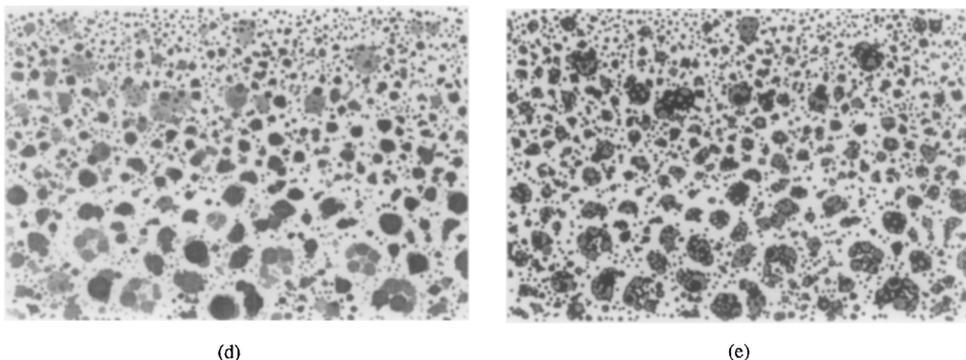
(c)

FIGS. 5a-c. (a) Sunflowers. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (c) Positive-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

the complexity of the displays, regions nested more than two levels deep have been omitted. There are typically less than 30 such regions per image.

As can be seen from the figures, regions are detected quite reliably over a large range of sizes and in the face of significant noise. (Very large regions, such as the snow-tongue in Fig. 1, are fragmented because our implementation does not generate disks with a diameter greater than 26 pixels. A larger $\nabla^2 G$ filter size could easily be added if larger regions are desired.) The region detector is insensitive to lighting changes; this can be seen well in Fig. 2, where clumps of dirt are detected both in the sunlit as well as the shadowed parts of the field.

The most notable weakness of the region extraction is the inaccurate representation of elongated regions. This is not surprising, since the only shape primitive used is a circular disk. (The use of primitives with elongated shapes could lead to better detection of elongated regions.) Here we analyze the result produced by our region detector when it is applied to an elongated image region. Two sources of error are apparent: (1) the calculated disk diameters overestimate the widths of elongated regions and (2) long thin regions tend to appear as a string of disconnected disks. We discuss these two types of errors in turn.



FIGS. 5d,e. (d) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (e) Negative-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

We estimate the error of fitting disks to elongated regions by comparing the $\nabla^2 G$ response of a disk to the $\nabla^2 G$ response for an infinitely long bar. Consider an image of an infinitely long bar of width B , where the bar is brighter than the background. Our current implementation models this bar by fitting disks along the centerline of the bar, since this is where local maxima in the $\nabla^2 G$ image occur. (If the bar has perfectly uniform intensity, a ridge of uniformly high values occurs along the centerline of the bar. However, any small fluctuations in bar intensity make some ridge points higher than others, resulting in a series of local maxima along the centerline of the bar.) The disk diameters will overestimate the bar width since the formula

$$D = 2\sigma \sqrt{\left\{ \sigma \left(\frac{\partial}{\partial \sigma} \nabla^2 G * I \right) / (\nabla^2 G * I) \right\} + 2}$$

is used to obtain the diameters D that model the bar width B , whereas the correct formula for the bar (derived in the Appendix) is

$$B = 2\sigma \sqrt{\left\{ \sigma \left(\frac{\partial}{\partial \sigma} \nabla^2 G * I \right) / (\nabla^2 G * I) \right\} + 1}.$$

The first term under the square root symbols is small, since we accept a disk only if the diameter is close to the filter size: $D \approx w = 2\sqrt{2}\sigma$. When $D = 2\sqrt{2}\sigma$ we have $(\partial/\partial\sigma)\nabla^2 G * I / (\nabla^2 G * I) = 0$, so the calculated disk diameter overestimates the bar width by a factor of $\sqrt{2}$. Thus, in an image of an infinitely long region, the region width is overestimated by a factor of approximately $\sqrt{2}$. For regions that are more moderately elongated the overestimation is less serious.

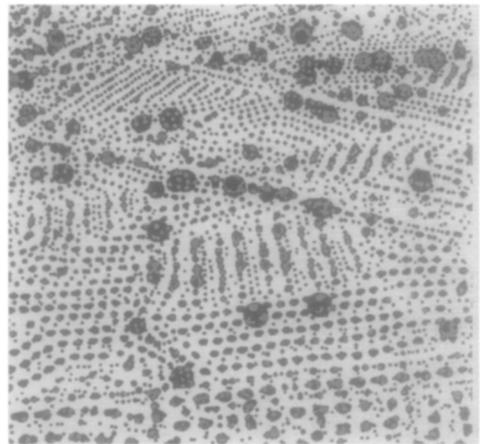
Using a circular disk as a shape primitive, we hope to model elongated regions by a string of overlapping disks. However, in the current implementation the disks that model an elongated region are often placed too sparsely, so that a disconnected



(a)



(b)

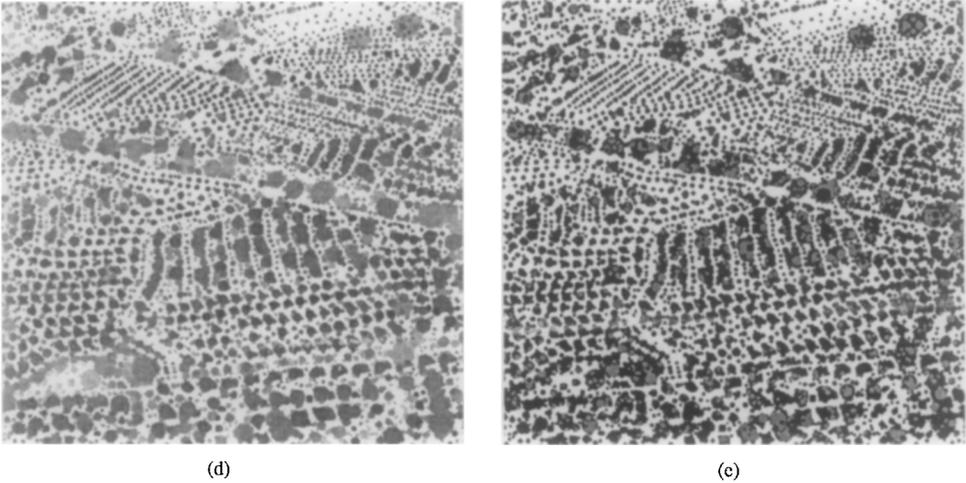


(c)

FIGS. 6a-c. (a) Aerial view of Levittown, Pennsylvania. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (c) Positive-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

chain of disks results. One possible remedy is to fit disks more closely. Currently we fit disks at local maxima (or minima) of the $\nabla^2 G * I$ images. An elongated region gives rise to a ridge of large values in the $\nabla^2 G * I$ image. Such a region could be better modeled by fitting a disk at each ridge point rather than just at each local maximum. However, it is difficult to formulate simple and accurate criteria for judging when a ridge point is significant enough to be used as a disk center.

Despite its shortcomings in representing elongated regions, the region detector permits fairly accurate estimation of the gradient of texture-element area, even for textures composed of elongated texture elements [3].



FIGS. 6d,e. (d) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to their contrast. (e) Negative-contrast regions are represented by sets of overlapping disks. Region intensity corresponds to nesting level: the outermost regions are black, nested regions are gray, and doubly nested regions are again black. (Regions nested more than two levels deep have been omitted.)

7. SUMMARY

A multiscale region detector is developed from the response of an ideal disk to convolution with a Laplacian-of-Gaussian ($\nabla^2 G$) over a range of scales. A region of arbitrary shape is detected as a union of disks contained in the region. Uniform regions are detected over a large range of sizes and in the face of significant noise. The region detector has a straightforward implementation which does not involve parameters that must be tuned to particular images. The region detector requires little computation beyond the generation of $\nabla^2 G$ convolutions for a range of filter sizes. The detector is designed as a coarse estimator of the location, size, shape, and contrast of texture elements. It is not proposed for general image segmentation.

APPENDIX

In this Appendix we analyze the $\nabla^2 G$ convolution output for a bar image. A bar of width B and contrast C is defined as

$$\text{bar image: } I(x, y) = \begin{cases} C & \text{if } 0 \leq x \leq B \\ 0 & \text{elsewhere.} \end{cases} \quad (7)$$

We show that at the center of the bar,

$$B = 2\sigma \sqrt{\left\{ \sigma \left(\frac{\partial}{\partial \sigma} \nabla^2 G * I \right) \right\} / (\nabla^2 G * I)} + 1. \quad (8)$$

This parallels the result of Eq. (5) obtained for a disk image. We begin with an analysis of the $\nabla^2 G$ response to a step-edge image; the bar response is computed as the sum of two step-edge responses.

Consider an image of a vertical step edge at $x = B$ defined by

$$\text{step-edge image: } I(x, y) = \begin{cases} C & \text{if } x \geq B \\ 0 & \text{elsewhere.} \end{cases} \quad (9)$$

Using this definition of $I(x, y)$ in Eq. (1) gives

$$\begin{aligned} & \int_{-\infty}^{x-b} \int_{-\infty}^{+\infty} \frac{2\sigma^2 - (u^2 + v^2)}{\sigma^4} e^{-(u^2+v^2)/2\sigma^2} C \, dv \, du \\ &= \frac{2C}{\sigma^2} \int_{-\infty}^{x-B} e^{-u^2/2\sigma^2} \int_{-\infty}^{+\infty} \left\{ \left(1 - \frac{u^2}{2\sigma^2} \right) e^{-v^2/2\sigma^2} - \frac{v^2}{2\sigma^2} e^{-v^2/2\sigma^2} \right\} \, dv \, du \\ &= \frac{2C}{\sigma^2} \int_{-\infty}^{x-B} e^{-u^2/2\sigma^2} \left\{ \left(1 - \frac{u^2}{2\sigma^2} \right) \sqrt{2\pi} \sigma - \frac{1}{2\sigma^2} \sqrt{2\pi} \sigma^3 \right\} \, du \\ &= \frac{\sqrt{2\pi} C}{\sigma} \int_{-\infty}^{x-B} \left\{ 1 - \frac{u^2}{\sigma^2} \right\} e^{-u^2/2\sigma^2} \, du \\ &= \sqrt{2\pi} C \int_{-\infty}^{(x-B)/\sigma} (1 - t^2) e^{-t^2/2} \, dt \\ &= \sqrt{2\pi} C \int_{-\infty}^{(x-B)/\sigma} e^{-t^2/2} \, dt \\ &\quad - \sqrt{2\pi} C \left\{ -\frac{x-B}{\sigma} e^{-(x-B)^2/2\sigma^2} + \int_{-\infty}^{(x-B)/\sigma} e^{-t^2/2} \, dt \right\}. \end{aligned}$$

Thus

$$\nabla^2 G \text{ response of step edge: } \frac{\sqrt{2\pi} C (x - B)}{\sigma} e^{-(x-B)^2/2\sigma^2}. \quad (10)$$

This expression is the closed-form solution for the $\nabla^2 G$ response of a vertical step edge of contrast C located at $x = B$. The $\nabla^2 G$ response of the vertical bar of width B and contrast C , as defined in Eq. (7), is the sum of two step-edge responses:

$$\begin{aligned} \nabla^2 G \text{ response of a bar} &= (\text{step up at } x = 0) + (\text{step down at } x = B) \\ &= \sqrt{2\pi} \frac{C}{\sigma} \left\{ x e^{-x^2/2\sigma^2} - (x - B) e^{-(x-B)^2/2\sigma^2} \right\}. \end{aligned} \quad (11)$$

Substituting $x = B/2$ into Eq. (11),

$$\nabla^2 G \text{ response at the center of a bar: } \frac{\sqrt{2\pi} CB}{\sigma} e^{-B^2/8\sigma^2}. \quad (12)$$

Taking the derivative with respect to sigma,

$$\frac{\partial}{\partial \sigma} \nabla^2 G \text{ response at the center of a bar: } \sqrt{2\pi} CB \left\{ \frac{B^2}{4\sigma^4} - \frac{1}{\sigma^2} \right\} e^{-B^2/8\sigma^2}. \quad (13)$$

Equations (12) and (13) can be combined to yield the desired expression

$$B = 2\sigma \sqrt{\left\{ \sigma \left(\frac{\partial}{\partial \sigma} \nabla^2 G * I \right) \right\} / (\nabla^2 G * I)} + 1 .$$

REFERENCES

1. E. Adelson and P. Burt, Image data compression with the Laplacian pyramid, *IEEE Pattern Recognition and Image Processing, Dallas, Texas, August 1981*, pp. 218–223.
2. D. Blostein, *Recovering the Orientation of Textured Surfaces in Natural Scenes*, Ph.D. thesis, Coordinated Science Laboratory Report UILU-ENG-87-2219, University of Illinois, April 1987.
3. D. Blostein and N. Ahuja, Representation and three-dimensional interpretation of image texture: An integrated approach, in *Proc. IEEE First International Conference on Computer Vision, London, June 1987*, pp. 444–449.
4. H. Blum, A transformation for extracting new descriptors of shape, in *Models for the Perception of Speech and Visual Form* (W. Wathen-Dunn, Ed.), pp. 362–380, MIT Press, Cambridge, MA, 1967.
5. P. Burt, Fast filter transforms for image processing, *Comput. Graphics Image Process.*, **16**, 1981, 20–51.
6. J. Crowley and A. Parker, A representation for shape based on peaks and ridges in the difference of low pass transform, *IEEE Pattern Anal. Mach. Intell.* **6**, No. 2, 1984, 156–170.
7. L. Davis, Computing the spatial structure of cellular textures, *Comput. Graphics Image Process.* **11**, 1979, 111–122.
8. R. Haralick and L. Shapiro, SURVEY: Image segmentation techniques, *Comput. Vision Graphics Image Process.* **29**, 1985, 100–132.
9. R. Hartley, L. Kitchen, C. Wang, and A. Rosenfeld, Segmentation of FLIR images: A comparative study, *DARPA Image Understanding Workshop, September 1982*, pp. 323–326.
10. K. Hayes, A. Shah, and A. Rosenfeld, Texture coarseness: Further experiments, *IEEE Sys. Man Cybern.* **4**, 1974, 467–472.
11. T. Hong, C. Dyer, and A. Rosenfeld, Texture primitive extraction using an edge-based approach, *IEEE Sys. Man Cybern.* **10**, 1980, 659–675.
12. J. Maleson, C. Brown, and J. Feldman, Understanding natural texture, *DARPA Image Understanding Workshop, October, 1977*, pp. 19–27.
13. D. Marr and E. Hildreth, Theory of edge detection, *Proc. Roy. Soc. London B* **207**, 1980, 187–217.
14. D. Marr, *Vision*, Freeman, San Francisco, 1987.
15. T. Matsuyama, K. Saburi, and M. Nagao, A structural analyzer for regularly arranged textures, *Comput. Graphics Image Process.* **18**, 1982, 259–278.
16. A. Rosenfeld and A. Kak, *Digital Picture Processing*, 2nd ed., Academic Press, New York, 1982.
17. Y. Ohta, K. Maenobu, and T. Sakai, *Proceedings International Joint Conference on Artificial Intelligence, 1981*, 746–751.
18. B. Schachter, A survey and evaluation of FLIR target detection/segmentation algorithms, *DARPA Image Understanding Workshop, September 1982*, pp. 49–57.
19. F. Vilnrotter, R. Nevatia, and K. Price, Structural analysis of natural textures, *IEEE Pattern Anal. Mach. Intell.* **8**, No. 1, 1986, 76–89.
20. H. Voorhees and T. Poggio, Detecting blobs as textons in natural textures, *DARPA Image Understanding Workshop, February 1987*, pp. 892–899.
21. S. Wang, R. Velasco, A. Wu, and A. Rosenfeld, Relative effectiveness of selected texture primitive statistics for texture discrimination, *IEEE Pattern Anal. Mach. Intell.* **11**, No. 5, 1981, 360–370.
22. A. P. Witkin, Scale space filtering, *Eighth International Joint Conference on Artificial Intelligence, Karlsruhe, West Germany, August 1983*, pp. 1019–1022.
23. S. Zucker, A. Rosenfeld, and L. Davis, Picture segmentation by texture discrimination, *IEEE Trans. Comput.* **24**, No. 12, 1975, 1228–1233.