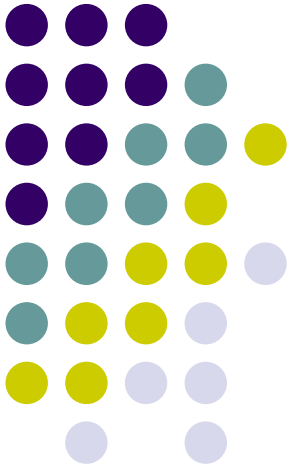
Chapter 10
Image Segmentation

Yinghua He



- **The whole is equal to the sum of its parts.**

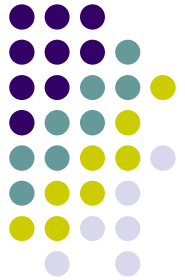
  -Euclid

- **The whole is greater than the sum of its parts.**

  -Max Wertheimer

- **The Whole is Not Equal to the Sum of Its Parts: An Approach to Teaching the Research Paper.**

  -by Mangum, Bryant

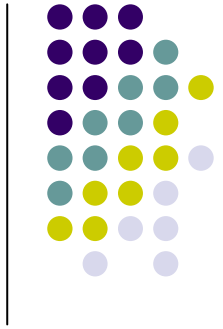


- **Detection of Discontinuities**
- **Edge Linking and Boundary Detection**
- **Thresholding**
- **Region-Based Segmentation**
- **Segmentation by Morphological Watersheds**
- **The Use of Motion in Segmentation**

FIGURE 10.1 A general 3×3 mask.

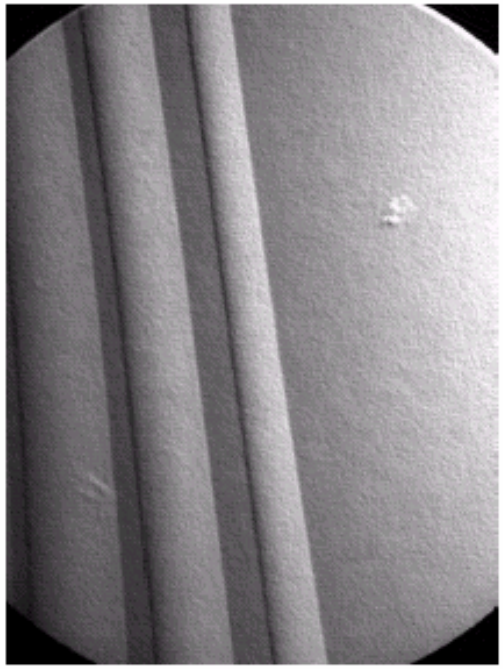
w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

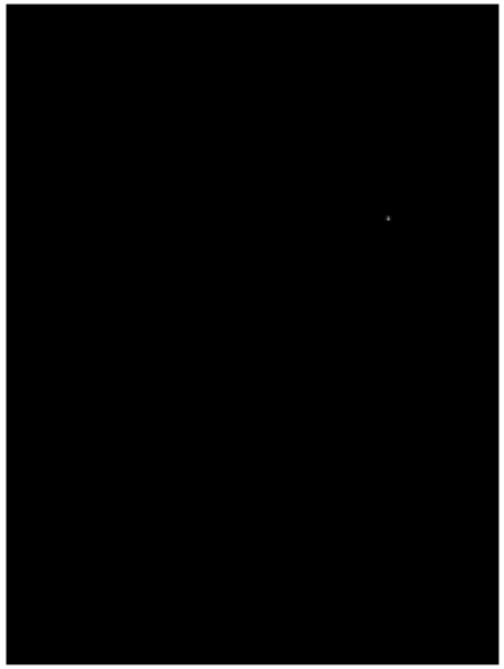
$$R = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$



- Point Detection
- Line Detection
- Edge Detection

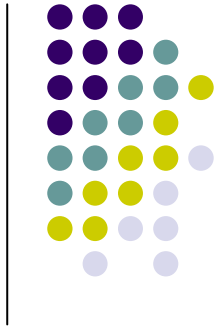
−1	−1	−1
−1	8	−1
−1	−1	−1





a
b c d

FIGURE 10.2
(a) Point detection mask. (b) X-ray image of a turbine blade with a porosity. (c) Result of point detection. (d) Result of using Eq. (10.1-2). (Original image courtesy of X-TEK Systems Ltd.)



- Point Detection
- Line Detection
- Edge Detection

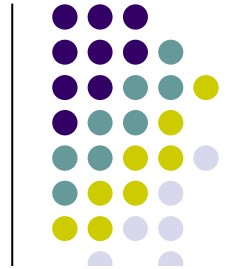


FIGURE 10.3 Line masks.

−1	−1	−1
2	2	2
−1	−1	−1

Horizontal

−1	−1	2
−1	2	−1
2	−1	−1

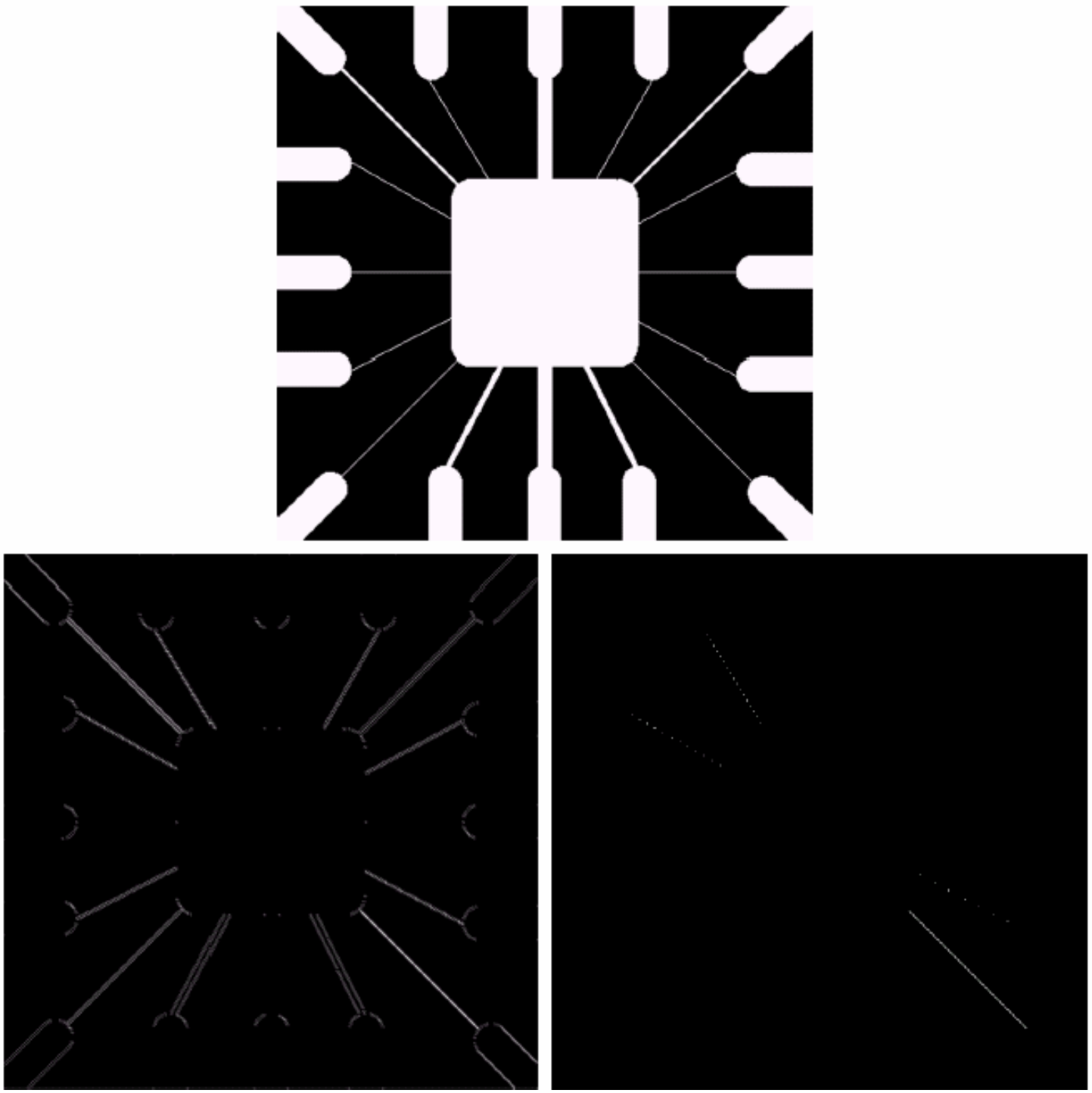
+45°

−1	2	−1
−1	2	−1
−1	2	−1

Vertical

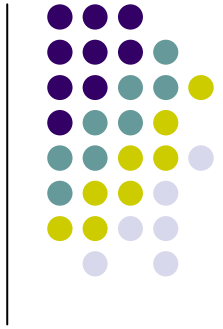
2	−1	−1
−1	2	−1
−1	−1	2

−45°

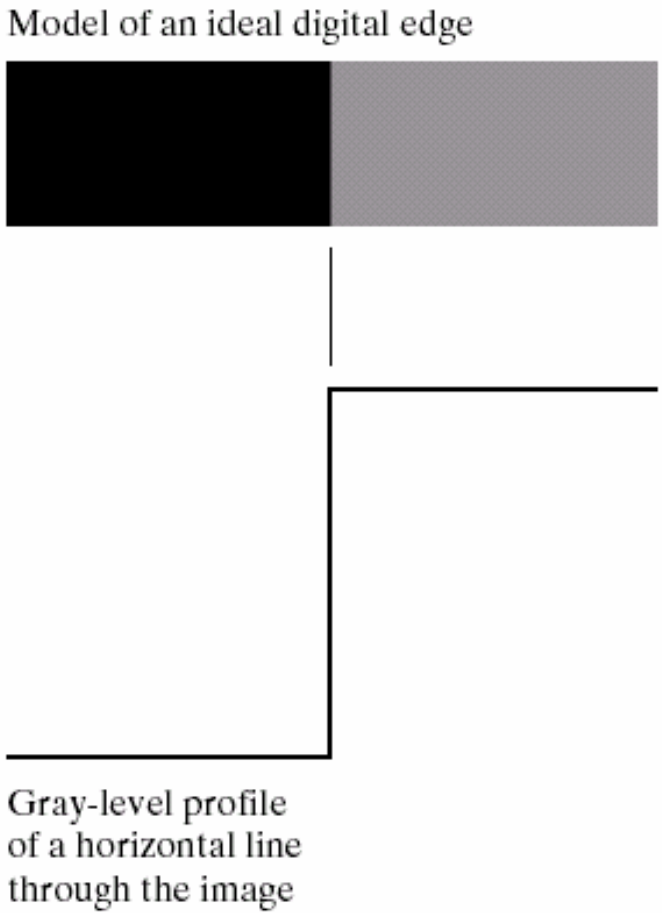


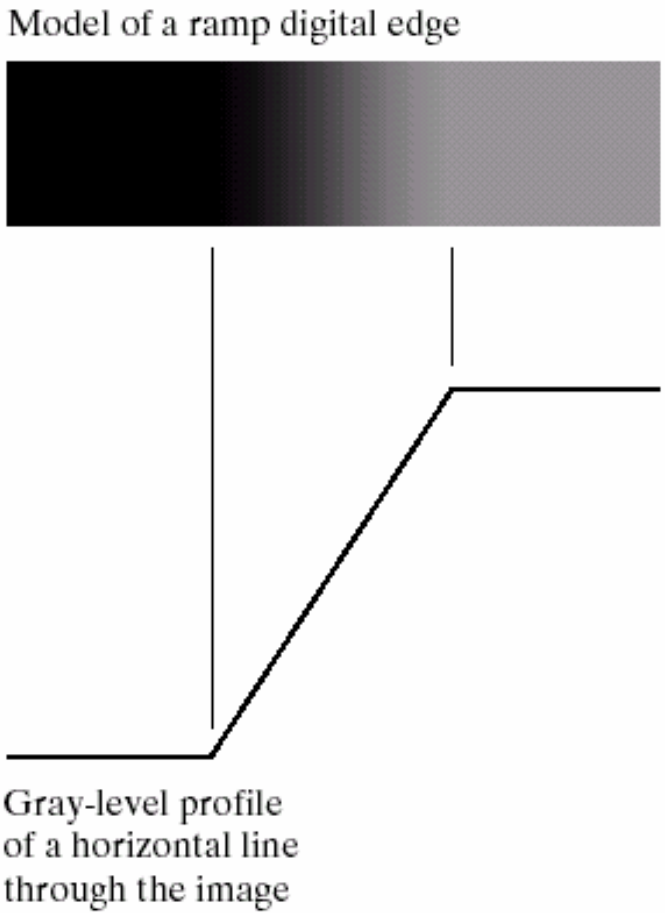
a
b c

FIGURE 10.4
Illustration of line detection.
(a) Binary wire-bond mask.
(b) Absolute value of result after processing with −45° line detector.
(c) Result of thresholding image (b).



- Point Detection
- Line Detection
- Edge Detection



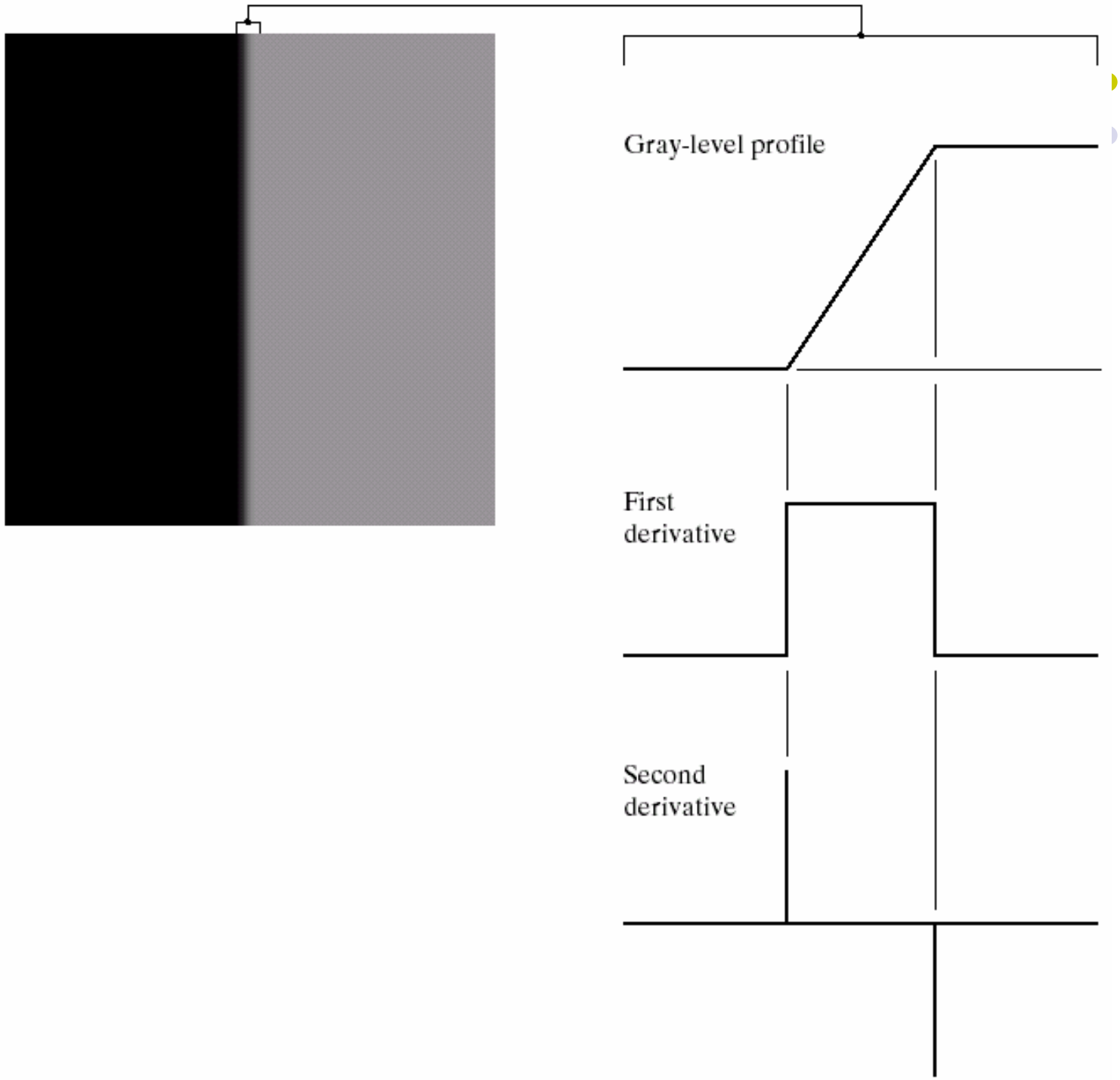


a b

FIGURE 10.5
(a) Model of an ideal digital edge. (b) Model of a ramp edge. The slope of the ramp is proportional to the degree of blurring in the edge.

a b

FIGURE 10.6
(a) Two regions separated by a vertical edge.
(b) Detail near the edge, showing a gray-level profile, and the first and second derivatives of the profile.



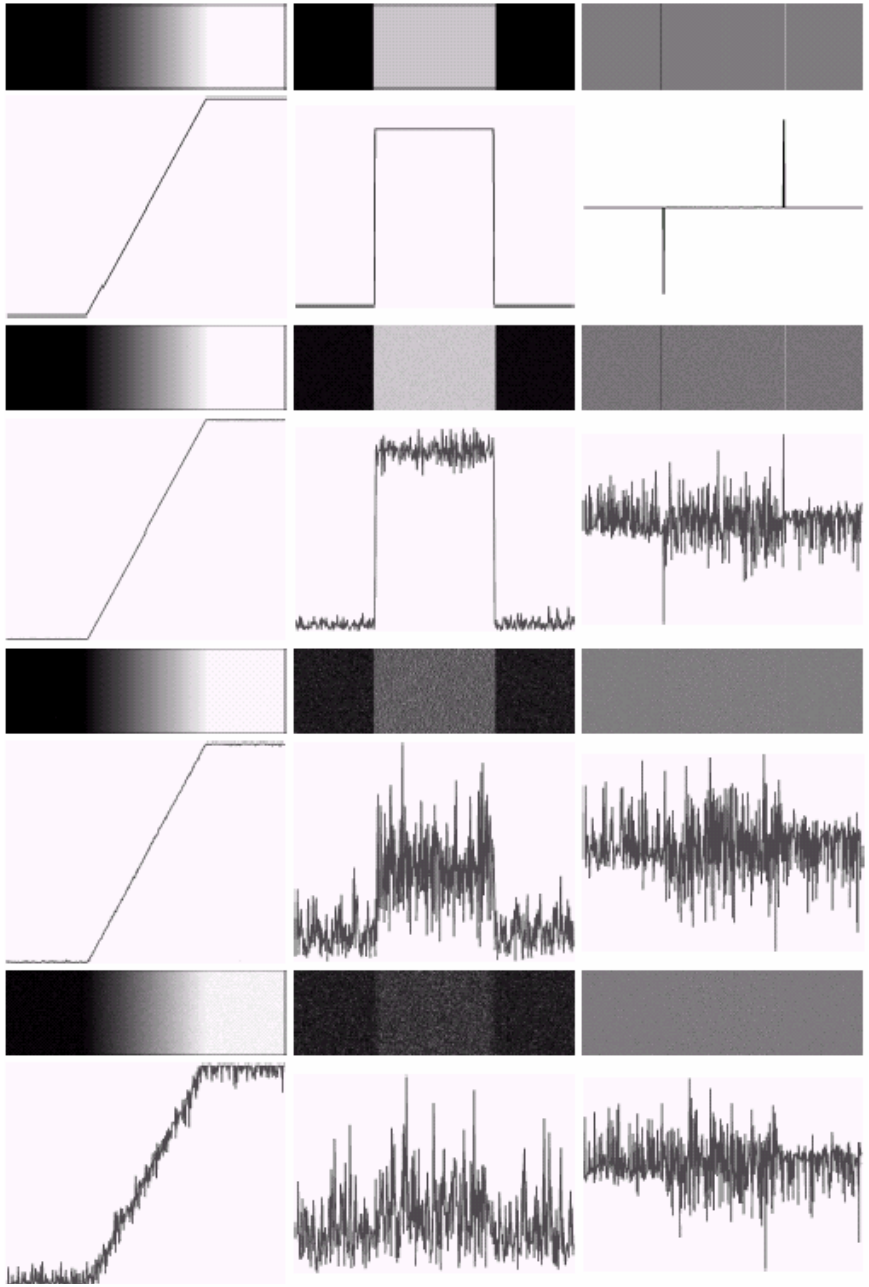


FIGURE 10.7 First column: images and gray-level profiles of a ramp edge corrupted by random Gaussian noise of mean 0 and $\sigma = 0.0, 0.1, 1.0,$ and 10.0, respectively. Second column: first-derivative images and gray-level profiles. Third column: second-derivative images and gray-level profiles.

a
b
c
d

Gradient operators

- The gradient of an image f(x,y) at location (x,y) is defined as the vector

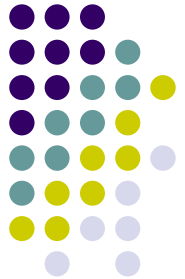
$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- An important quantity in edge detection is the magnitude of this vector, denoted ∇f ,where

$$\nabla f = mag(\nabla f) = \left[G_x^2 + G_y^2 \right]^{1/2}$$

- The direction of the gradient vector also is an important quantity. Let $\alpha(x, y)$ represent the direction angle of the vector ∇f at (x,y). Then, from vector analysis,

$$\alpha(x, y) = tag^{-1}\left(\frac{G_y}{G_x}\right)$$



- Roberts cross-gradient operators:

$$G_x = (z_9 - z_5)$$

and

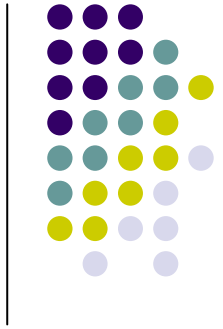
$$G_y = (z_8 - z_6)$$

- An approach using masks of size 3*3 is given by

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

- and

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

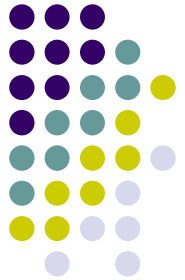


- A slight variation of these two equations uses a weight of 2 in the center coefficient:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

and

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$



- An approach used frequently is to approximate the gradient by absolute values:

$$\nabla f \approx |G_x| + |G_y|$$

The Laplacian

- The Laplacian of a 2-D function f(x,y) is a second-order derivative defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- For a 3*3 region, one of the two forms encountered most frequently in practice is

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

- A digital approximation including the diagonal neighbors is given by

$$\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$$

a
b c
d e
f g

FIGURE 10.8
A 3 × 3 region of an image (the z's are gray-level values) and various masks used to compute the gradient at point labeled z_5.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

−1	0
0	1

0	−1
1	0

Roberts

−1	−1	−1
0	0	0
1	1	1

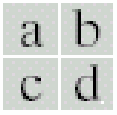
−1	0	1
−1	0	1
−1	0	1

Prewitt

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

Sobel



0	1	1	−1	−1	0
−1	0	1	−1	0	1
−1	−1	0	0	1	1

Prewitt

0	1	2	−2	−1	0
−1	0	1	−1	0	1
−2	−1	0	0	1	2

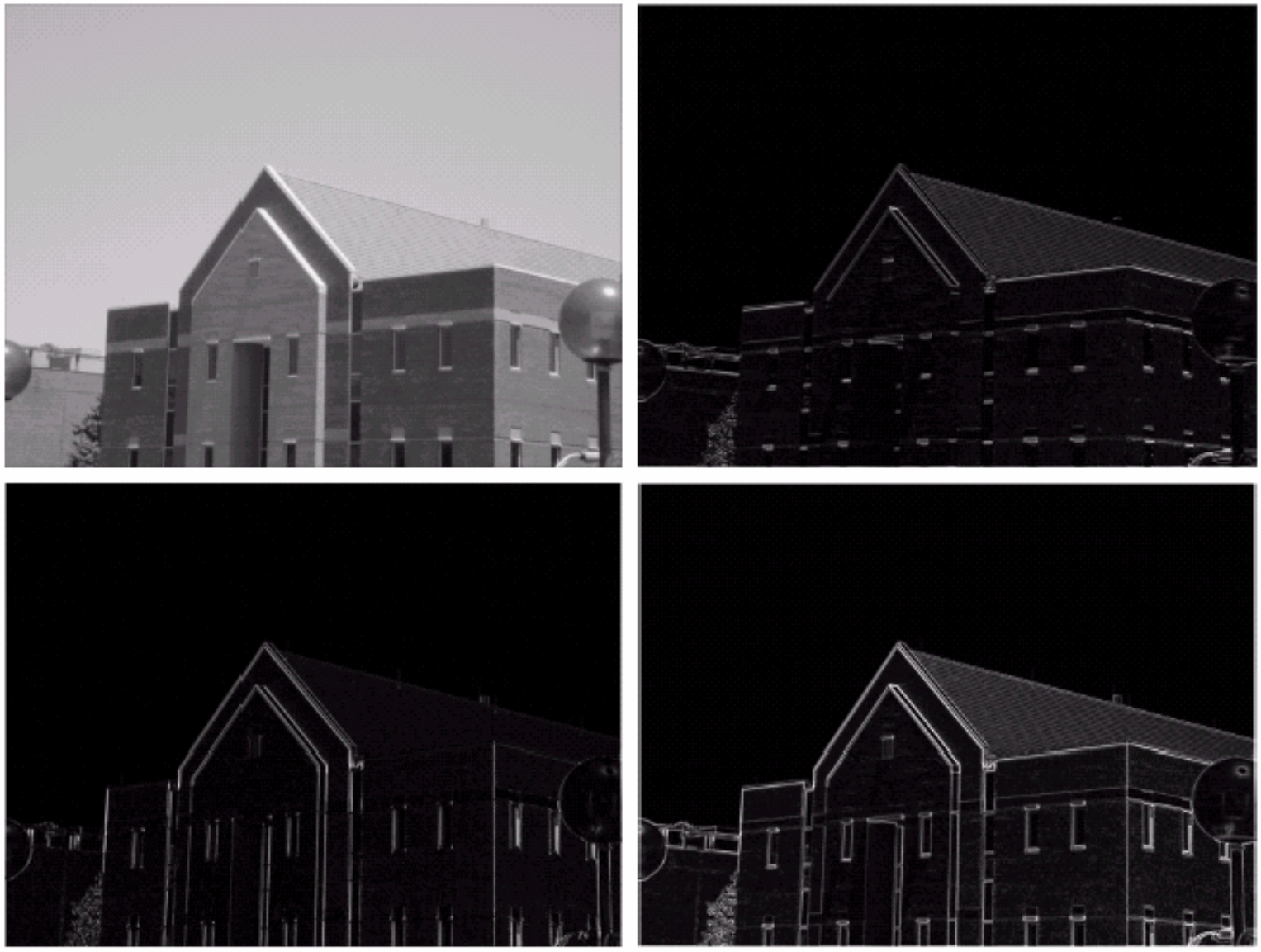
Sobel

a b
c d

FIGURE 10.9 Prewitt and Sobel masks for detecting diagonal edges.

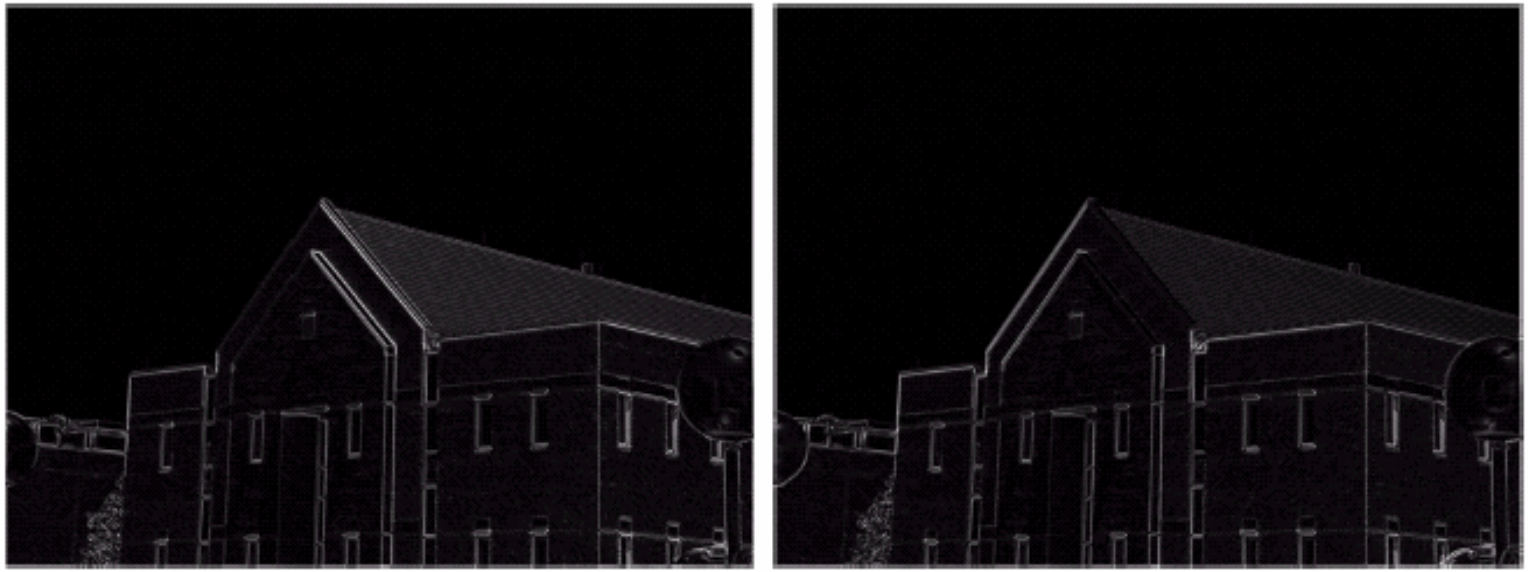
a b
c d

FIGURE 10.10
(a) Original image. (b) $|G_x|$, component of the gradient in the x-direction. (c) $|G_y|$, component in the y-direction. (d) Gradient image, $|G_x| + |G_y|$.



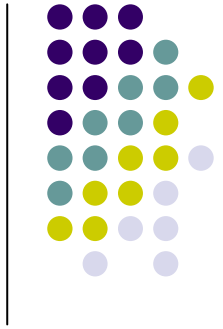
a b
c d

FIGURE 10.11
Same sequence as in Fig. 10.10, but with the original image smoothed with a 5×5 averaging filter.



a b

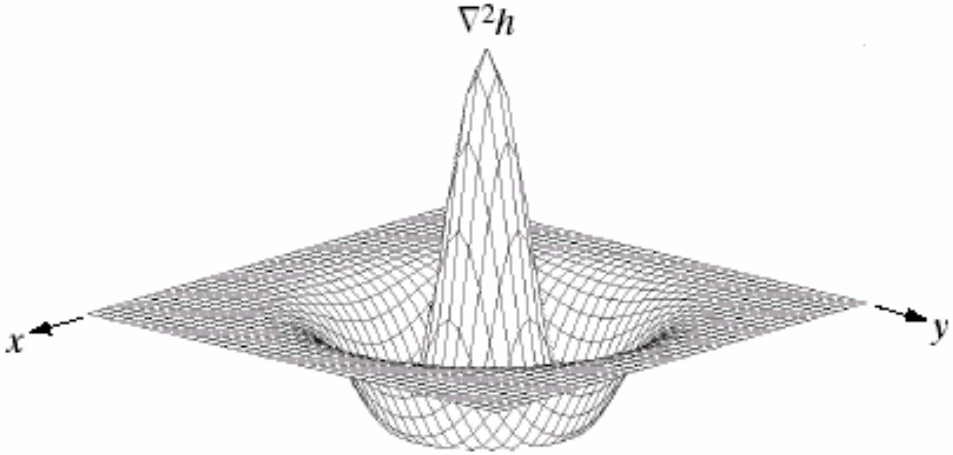
FIGURE 10.12
Diagonal edge detection.
(a) Result of using the mask in Fig. 10.9(c).
(b) Result of using the mask in Fig. 10.9(d). The input in both cases was Fig. 10.11(a).

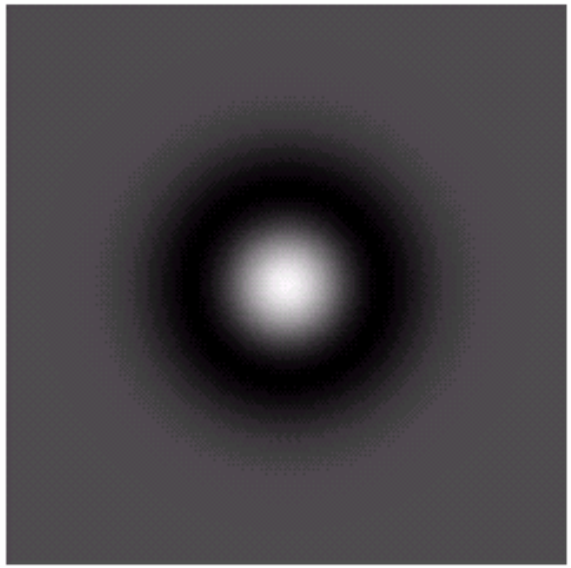


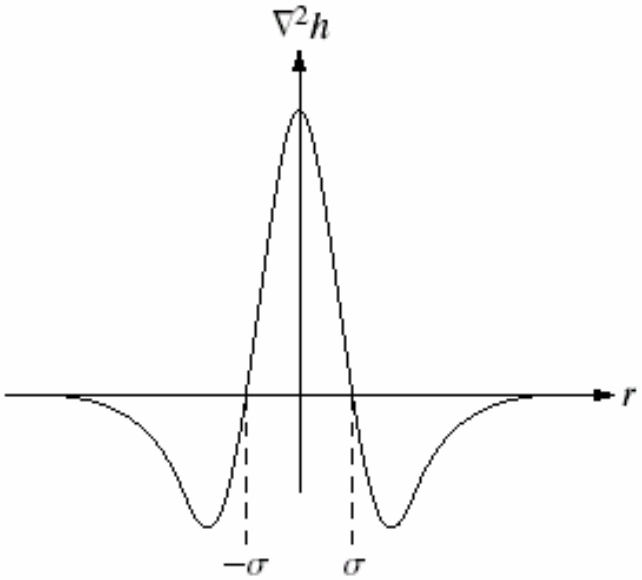
0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

FIGURE 10.13
Laplacian masks used to implement Eqs. (10.1-14) and (10.1-15), respectively.



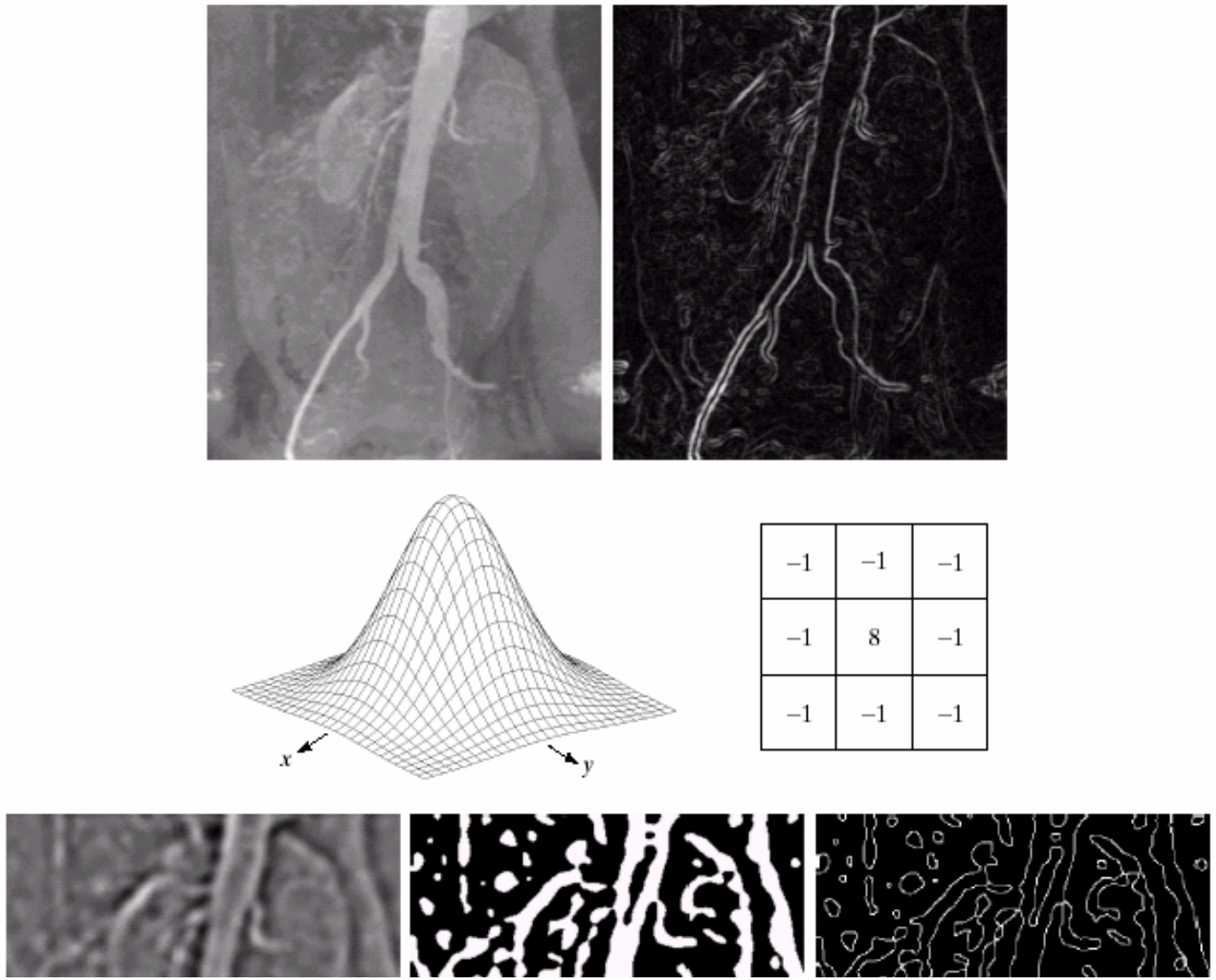




0	0	−1	0	0
0	−1	−2	−1	0
−1	−2	16	−2	−1
0	−1	−2	−1	0
0	0	−1	0	0

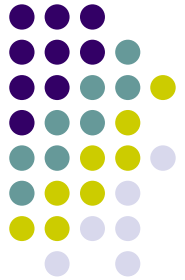
a b
c d

FIGURE 10.14 Laplacian of a Gaussian (LoG). (a) 3-D plot. (b) Image (black is negative, gray is the zero plane, and white is positive). (c) Cross section showing zero crossings. (d) 5 × 5 mask approximation to the shape of (a).



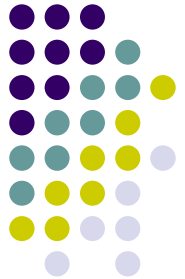
a b
c d
e f g

FIGURE 10.15 (a) Original image. (b) Sobel gradient (shown for comparison). (c) Spatial Gaussian smoothing function. (d) Laplacian mask. (e) LoG. (f) Thresholded LoG. (g) Zero crossings. (Original image courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)



- Detection of Discontinuities
- Edge Linking and Boundary Detection
- Thresholding
- Region-Based Segmentation
- Segmentation by Morphological Watersheds
- The Use of Motion in Segmentation

- Local Processing
- Global Processing via the Hough Transform
- Global Processing via Graph-Theoretic Techniques



- An edge pixel with coordinates (x_0,y_0) in a predefined neighborhood of (x,y), is similar in magnitude to the pixel at (x,y) if

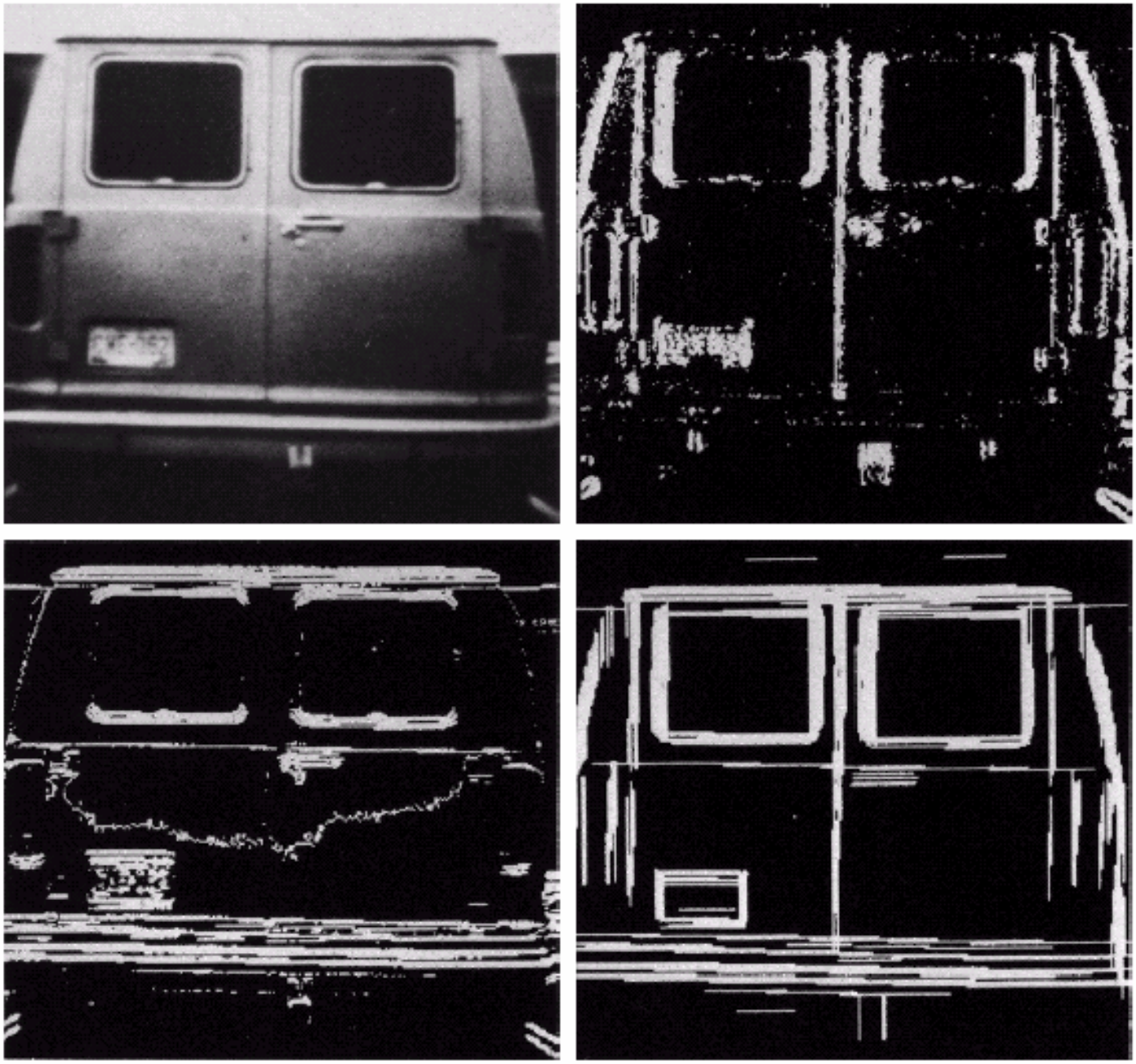
$$\left|\nabla f(x,y) - \nabla f(x_0,y_0)\right| \leq E$$

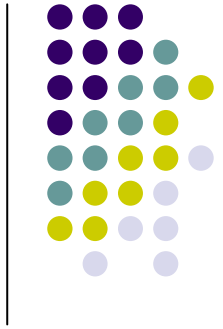
- An edge pixel at (x_0,y_0) in the predefined neighborhood of (x,y) has an angle similar to the pixel at (x,y) if

$$\left|\alpha(x,y) - \alpha(x_0,y_0)\right| < A$$

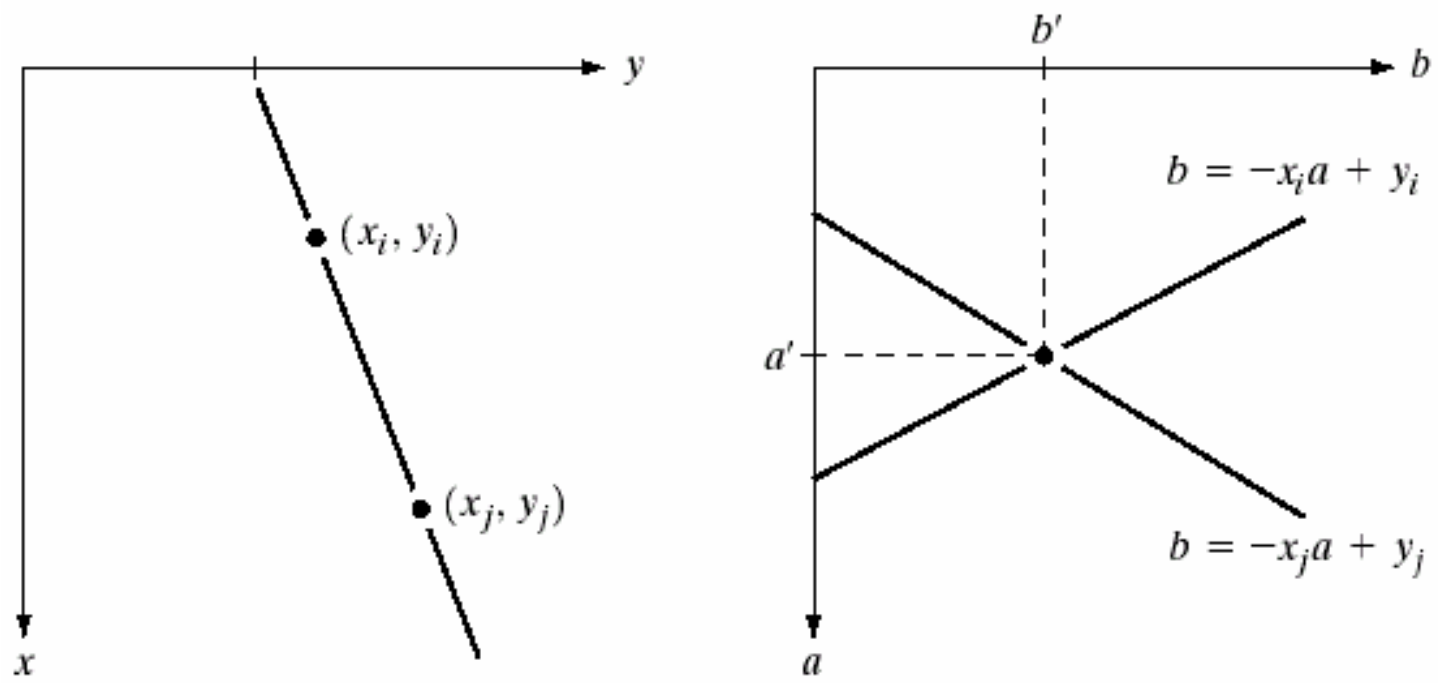
a b
c d

FIGURE 10.16
(a) Input image.
(b) G_y component of the gradient.
(c) G_x component of the gradient.
(d) Result of edge linking. (Courtesy of Perceptics Corporation.)





- Local Processing
- Global Processing via the Hough Transform
- Global Processing via Graph-Theoretic Techniques



a b

FIGURE 10.17
(a) xy-plane. (b) Parameter space.

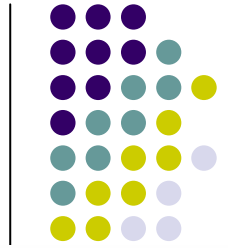
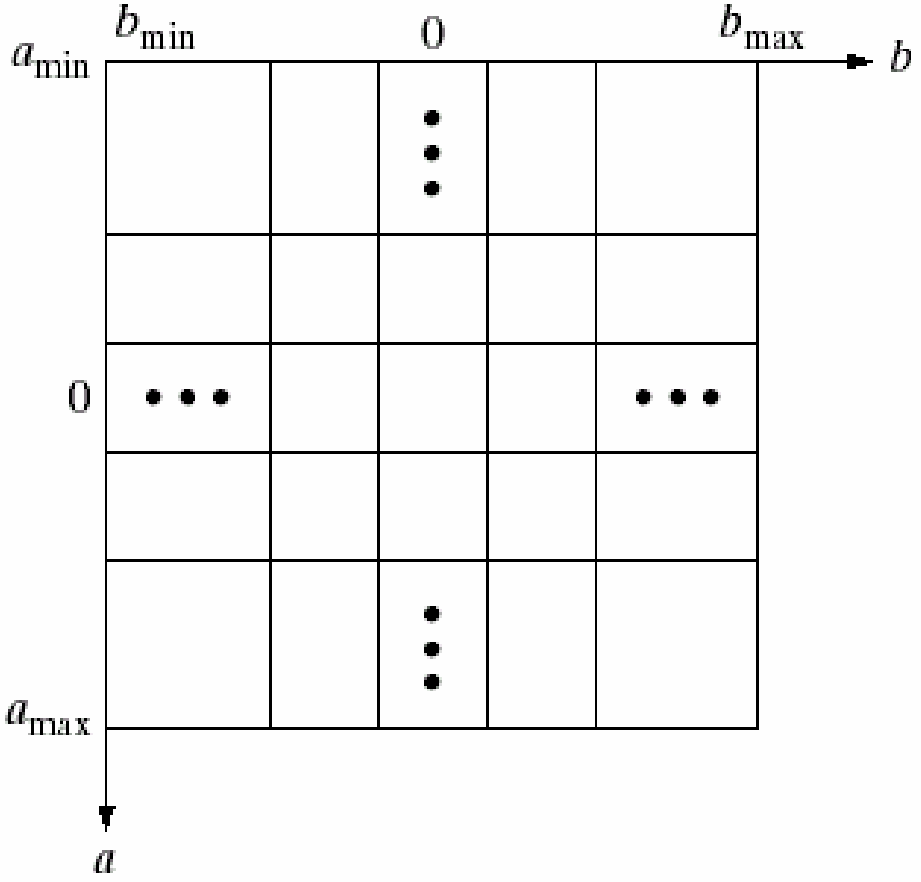
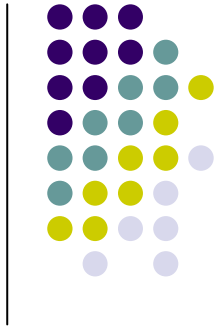
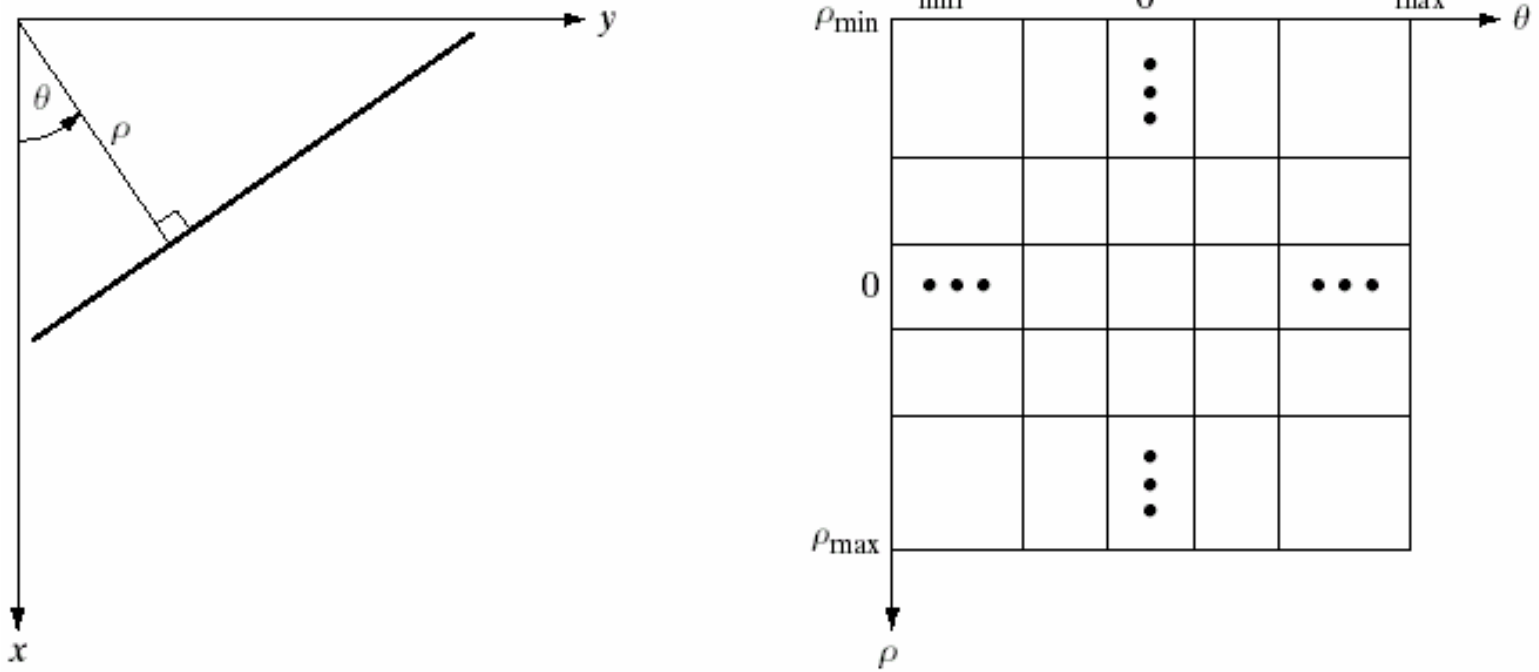


FIGURE 10.18
Subdivision of the parameter plane for use in the Hough transform.



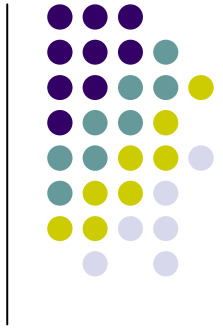


$$x\cos\theta + y\sin\theta = \rho$$



a b

FIGURE 10.19
(a) Normal representation of a line.
(b) Subdivision of the $\rho\theta$-plane into cells.



$$(x - c_1)^2 + (y - c_2)^2 = c_3^2$$

a b
c d

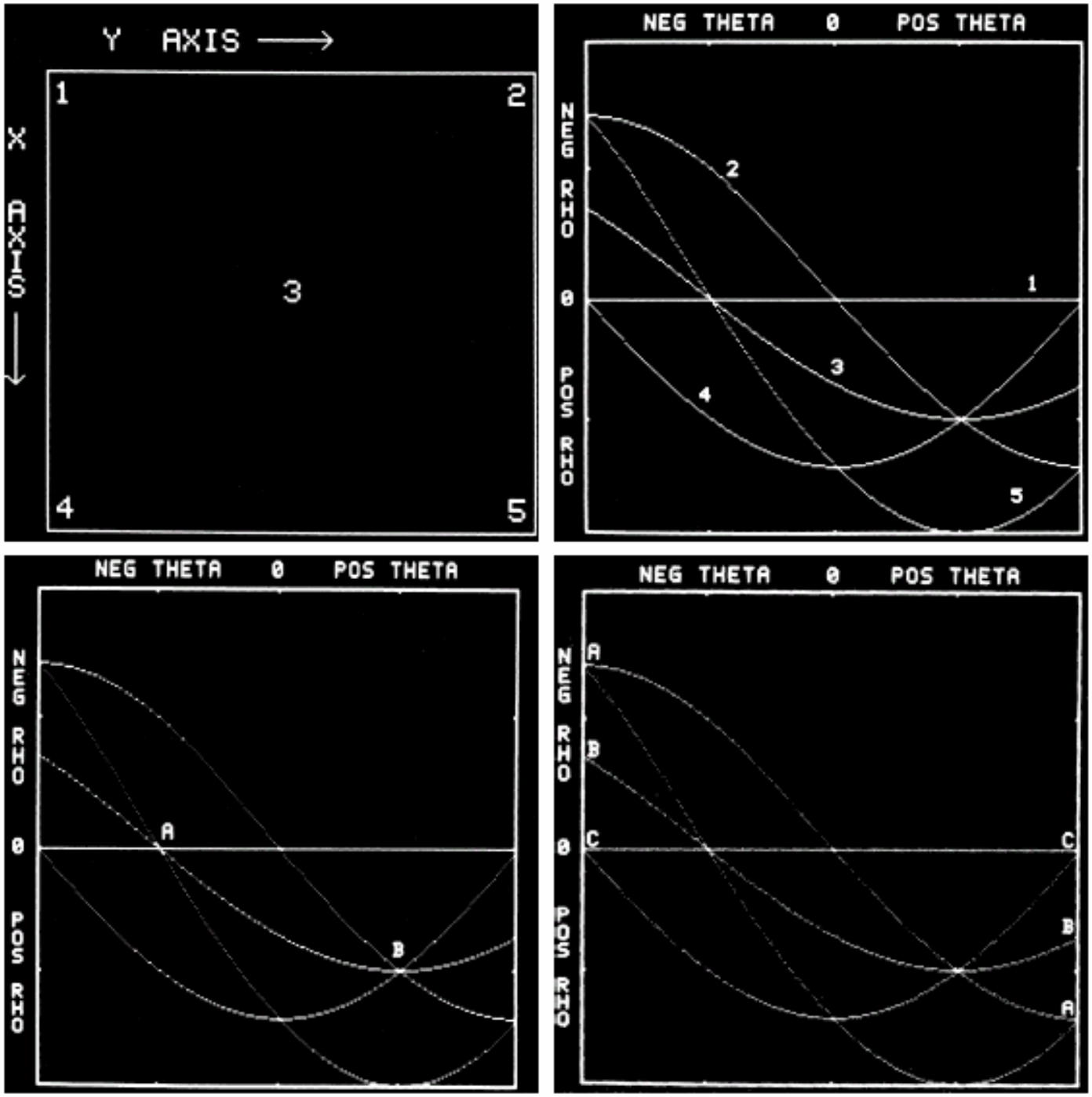
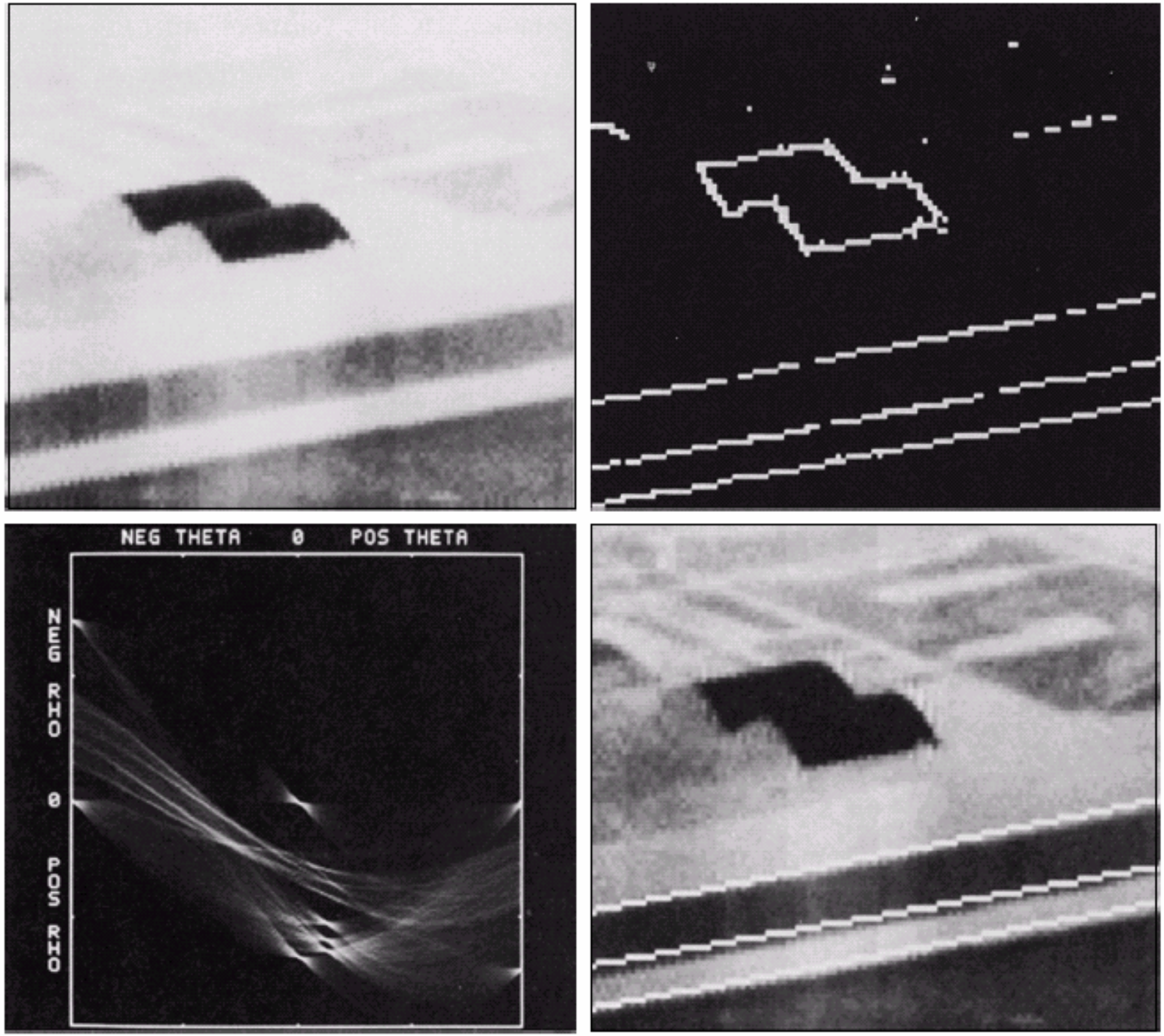


FIGURE 10.20
Illustration of the Hough transform. (Courtesy of Mr. D. R. Cate, Texas Instruments, Inc.)



a b
c d

FIGURE 10.21
(a) Infrared image.
(b) Thresholded gradient image.
(c) Hough transform.
(d) Linked pixels.
(Courtesy of Mr. D. R. Cate, Texas Instruments, Inc.)

- Local Processing
- Global Processing via the Hough Transform
- Global Processing via Graph-Theoretic Techniques

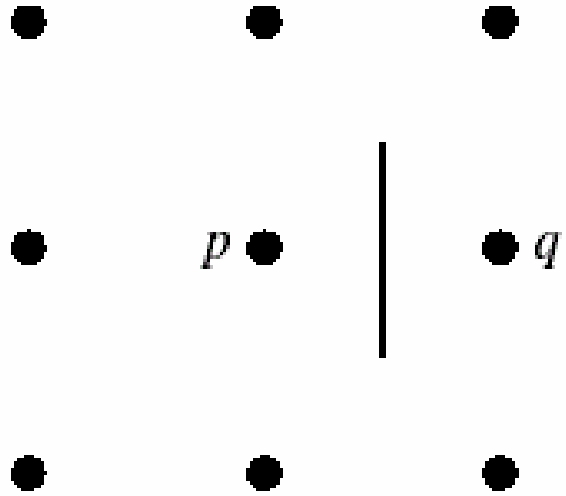
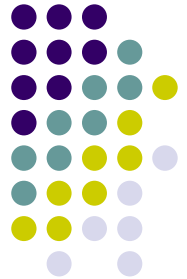
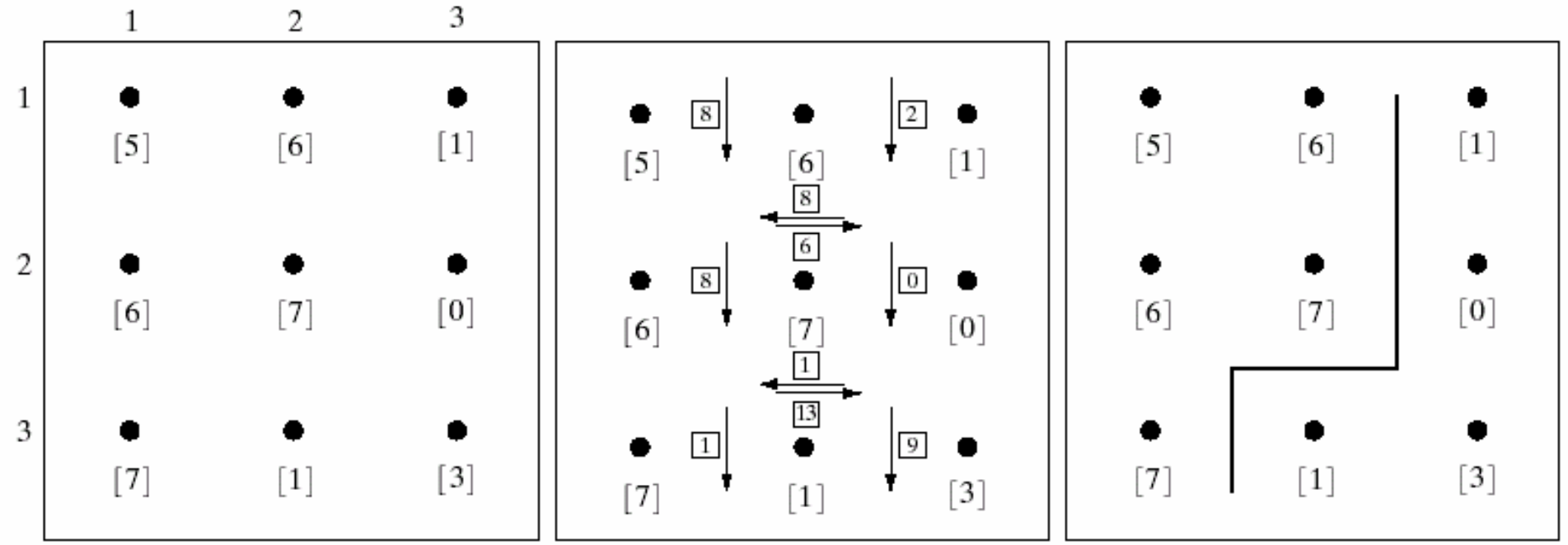


FIGURE 10.22
Edge element between pixels p and q.



- A sequence of nodes $n_1, n_2, \ldots, n_k$, with each node n_i being a successor of node n_{i-1}, is called a path from n_1 to n_k. The cost of the entire path is:

$$c = \sum_{i=2}^{k} c(n_{i-1}, n_i)$$



a b c

FIGURE 10.23 (a) A 3 × 3 image region. (b) Edge segments and their costs. (c) Edge corresponding to the lowest-cost path in the graph shown in Fig. 10.24.

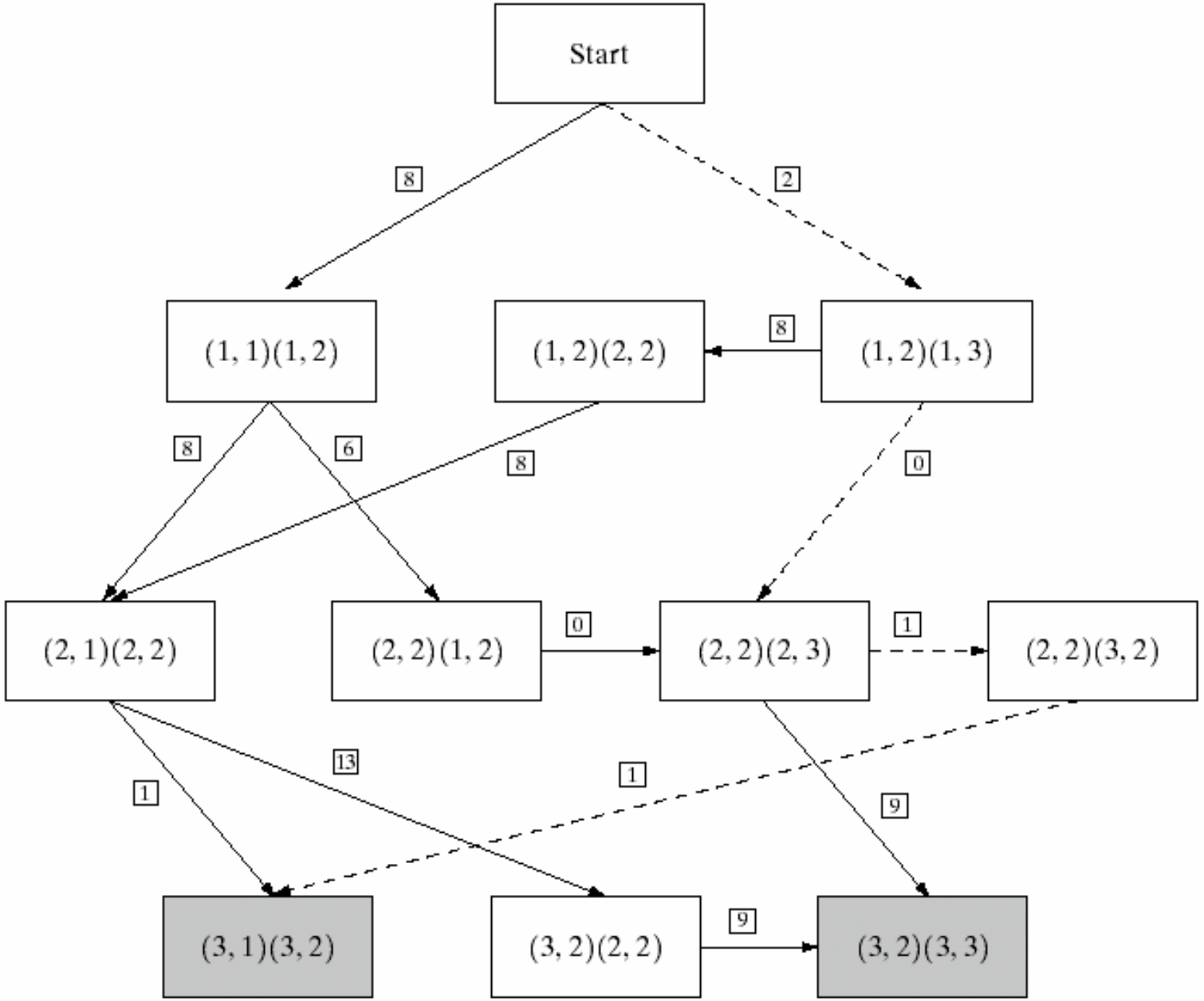
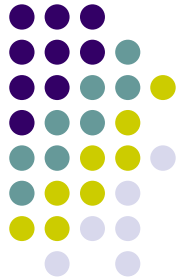
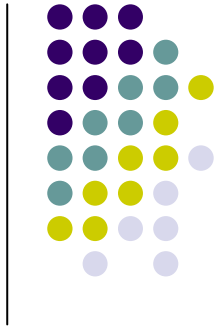


FIGURE 10.24
Graph for the image in Fig. 10.23(a). The lowest-cost path is shown dashed.

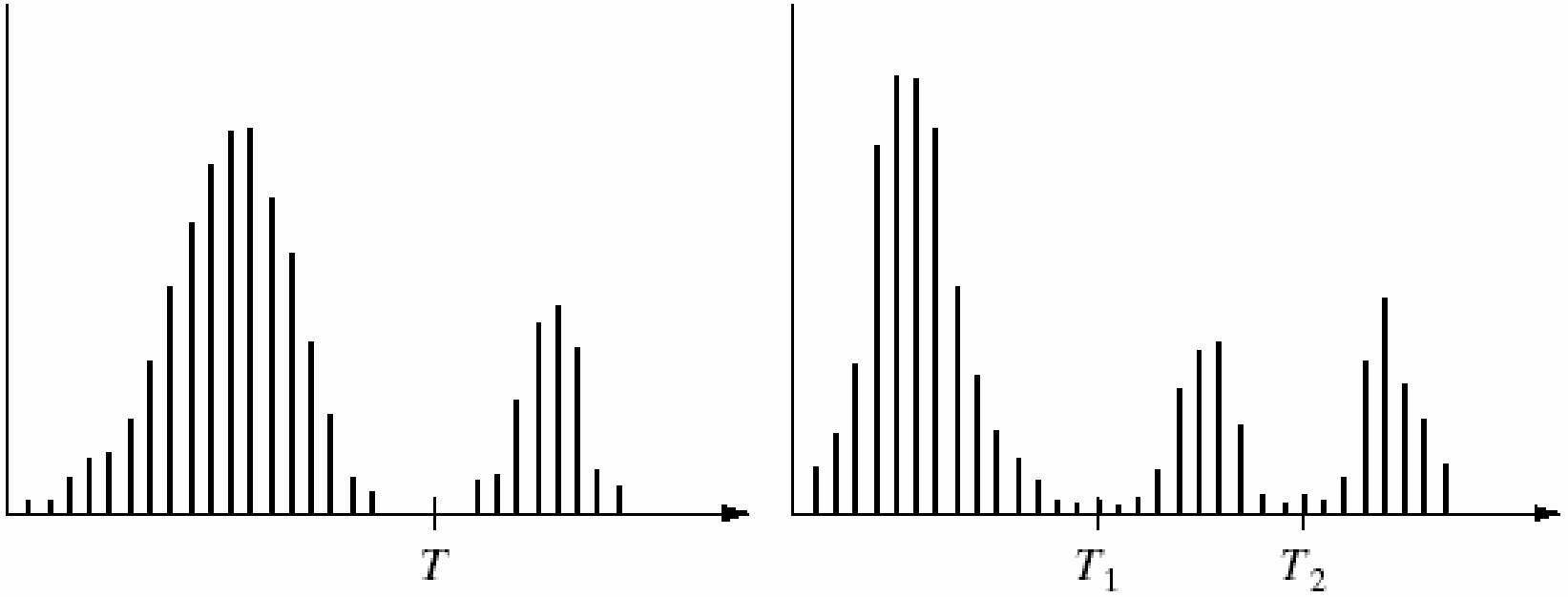
FIGURE 10.25
Image of noisy chromosome silhouette and edge boundary (in white) determined by graph search.



- Detection of Discontinuities
- Edge Linking and Boundary Detection
- Thresholding
- Region-Based Segmentation
- Segmentation by Morphological Watersheds
- The Use of Motion in Segmentation



- Foundation
- The Role of Illumination
- Basic Global Thresholding
- Basic Adaptive Thresholding



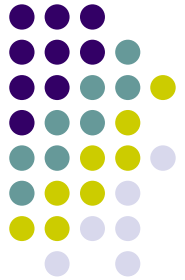
a b

FIGURE 10.26 (a) Gray-level histograms that can be partitioned by (a) a single threshold, and (b) multiple thresholds.

- Based on the preceding discussion, thresholding may be viewed as an operation that involves tests against a function T of the form

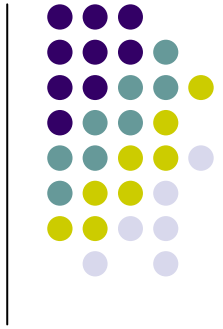
$$T = T\left[x, y, p(x, y), f(x, y)\right]$$

- A thresholded image g(x,y) is defined as

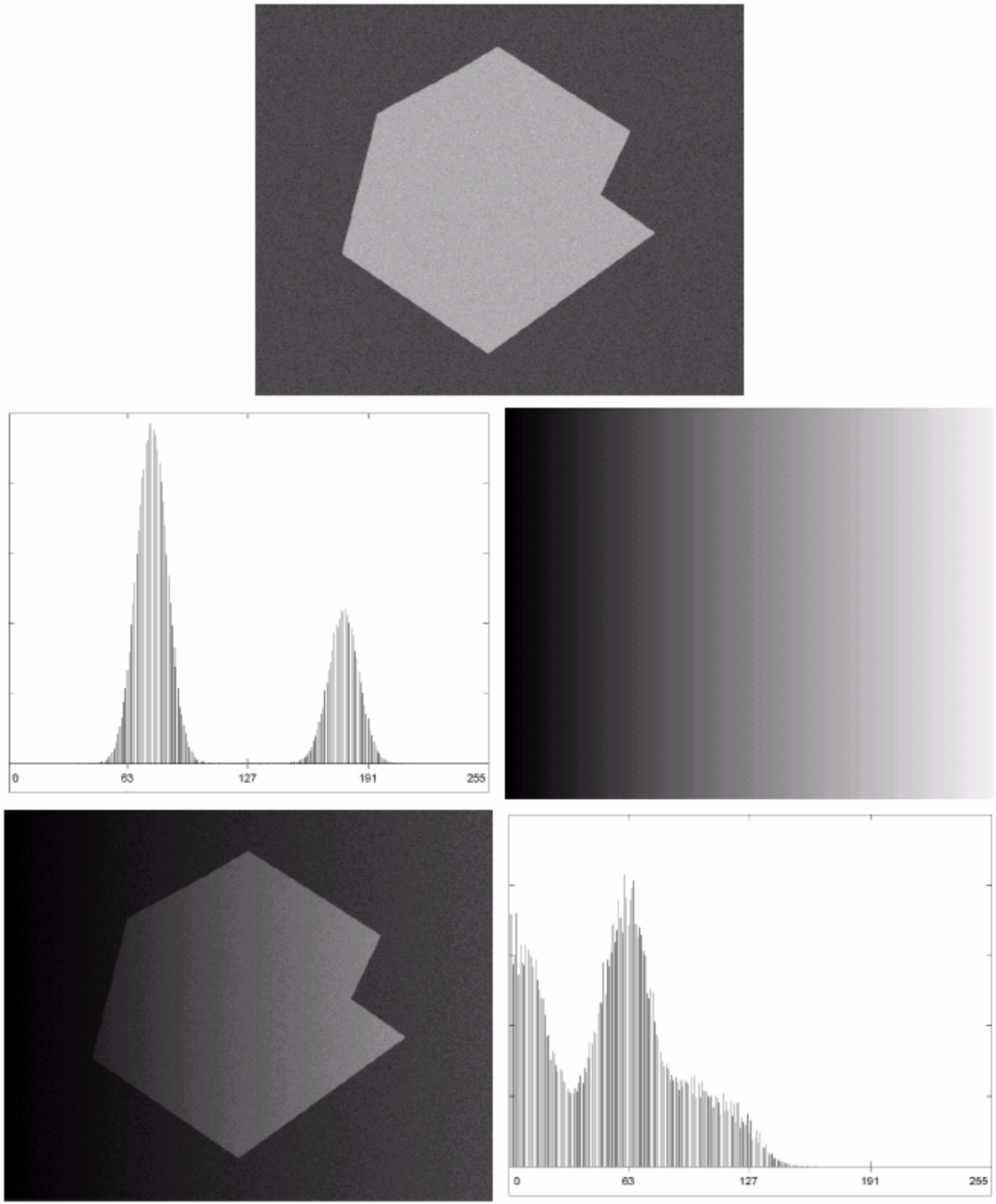


- A thresholded image g(x,y) is defined as

$$g(x,y) = \begin{cases} 1 & if \quad f(x,y) > T \\ 0 & if \quad f(x,y) \leq T \end{cases}$$

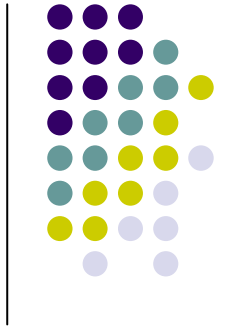


- Foundation
- The Role of Illumination
- Basic Global Thresholding
- Basic Adaptive Thresholding



a
b c
d e

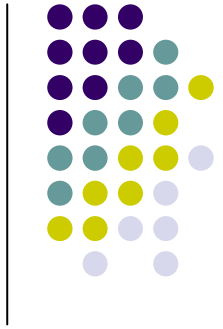
FIGURE 10.27
(a) Computer generated reflectance function.
(b) Histogram of reflectance function.
(c) Computer generated illumination function.
(d) Product of (a) and (c).
(e) Histogram of product image.



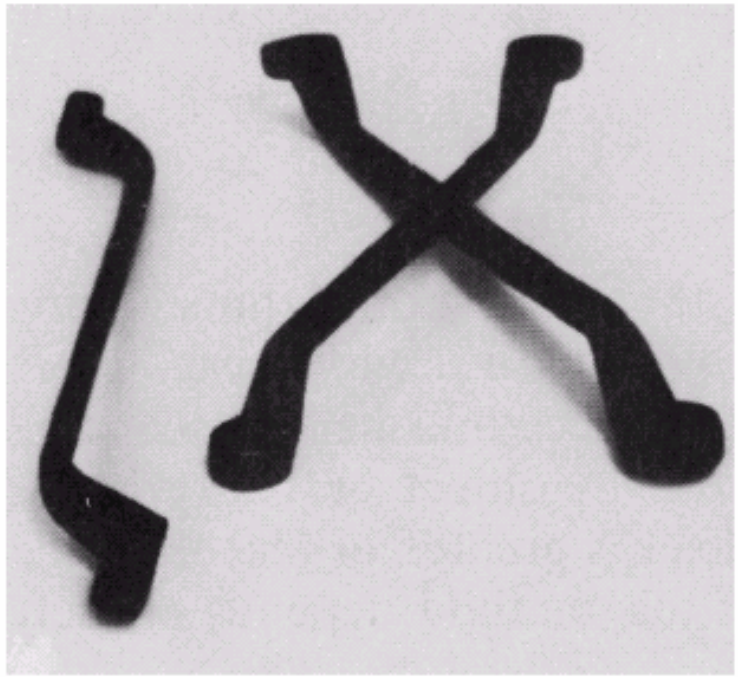
$$f(x,y) = i(x,y)r(x,y)$$

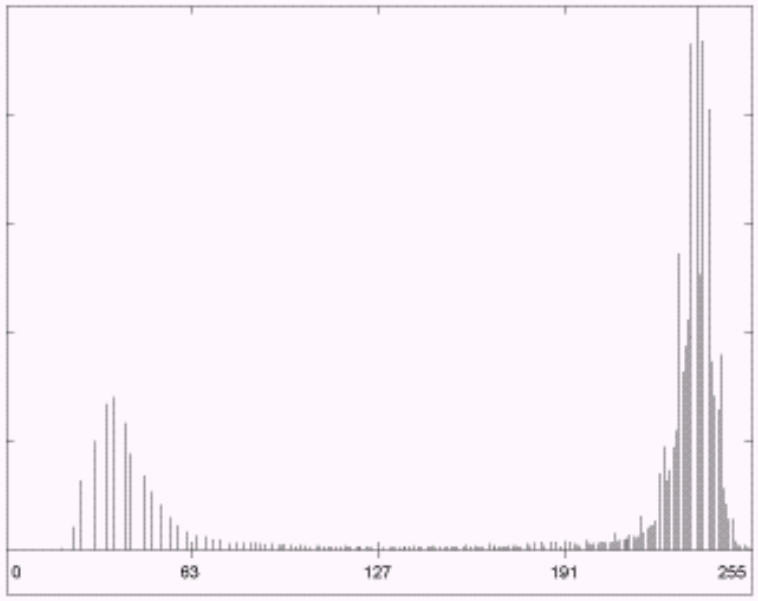
- Taking the natural logarithm of this equation yields a sum:

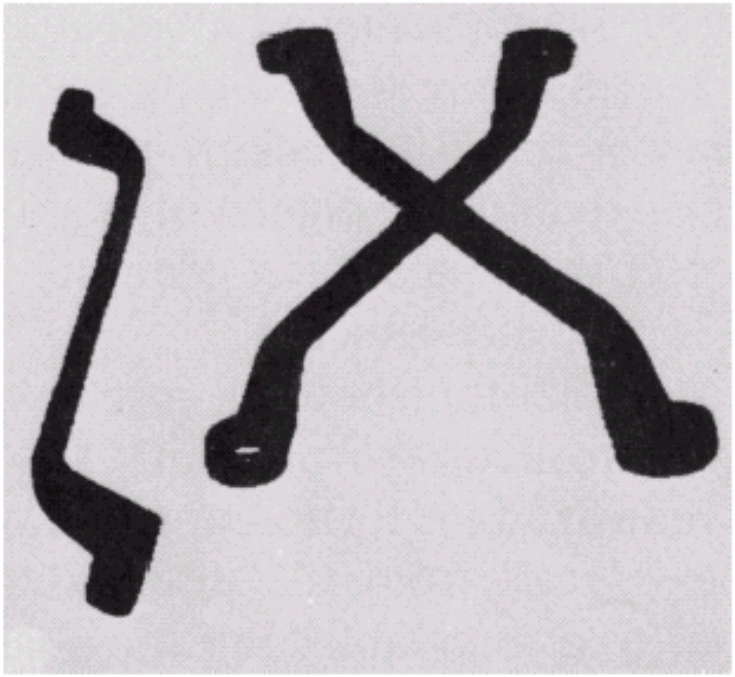
$$\begin{aligned} z(x,y) &= \ln f(x,y) \\ &= \ln i(x,y) + \ln r(x,y) \\ &= i'(x,y) + r'(x,y) \end{aligned}$$



- Foundation
- The Role of Illumination
- Basic Global Thresholding
- Basic Adaptive Thresholding

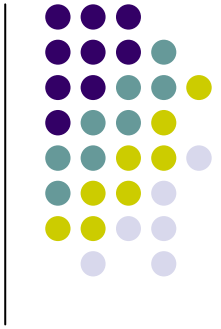






a
b c

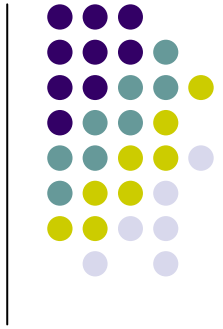
FIGURE 10.28
(a) Original image. (b) Image histogram. (c) Result of global thresholding with T midway between the maximum and minimum gray levels.



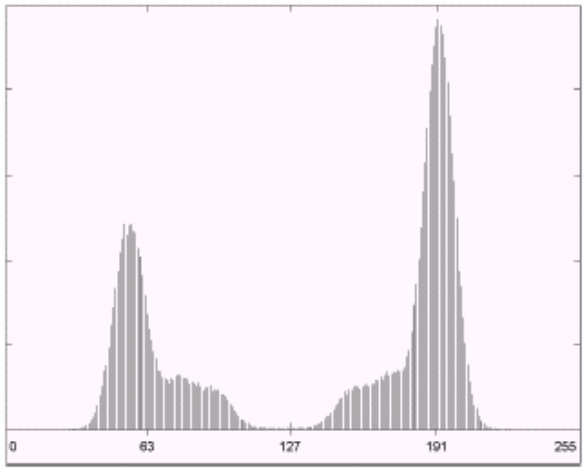
- The following algorithm can be used to obtain T automatically:
  - Select an initial estimate for T;
  - Segment the image using T. This will produce two groups of pixels: G_1 consisting of all pixels with grey level values>T and G_2 consisting of pixels with values <=T.
  - Compute the average gray level values μ_1 and μ_2 for the pixels in regions G_1 and G_2.
  - Compute a new threshold value:

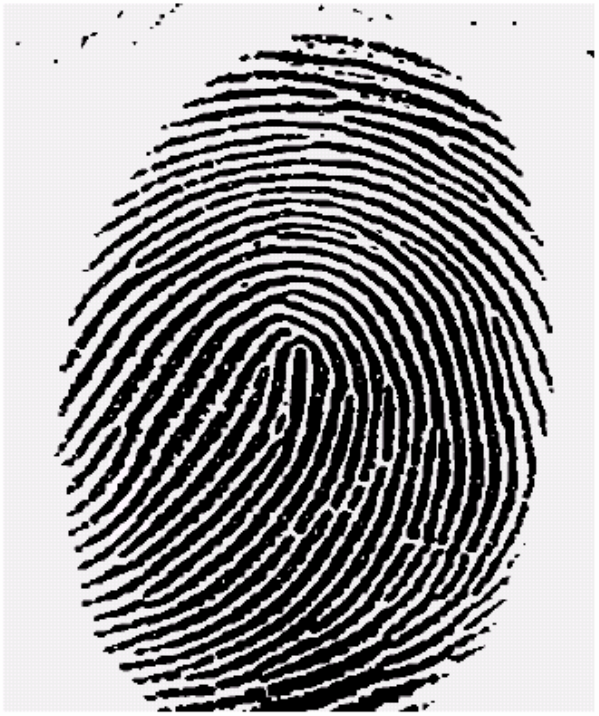
$$T = \frac{1}{2}(\mu_1 + \mu_2)$$

  - Repeat steps 2 through 4 until the difference in T in successive iterations is smaller than a predefined parameter T_o.



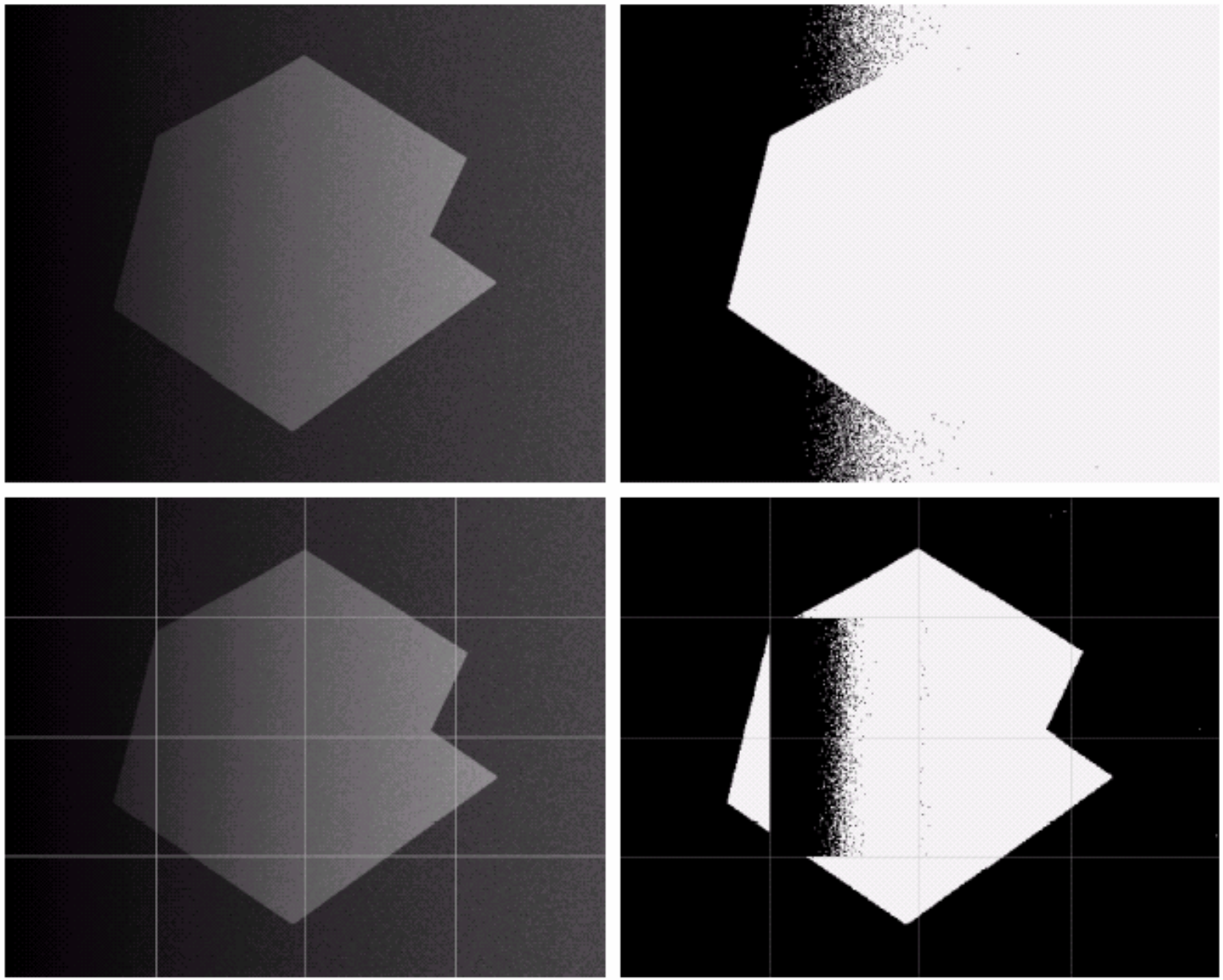
- Foundation
- The Role of Illumination
- Basic Global Thresholding
- Basic Adaptive Thresholding





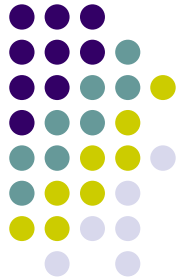
a b
c

FIGURE 10.29
(a) Original image. (b) Image histogram. (c) Result of segmentation with the threshold estimated by iteration. (Original courtesy of the National Institute of Standards and Technology.)

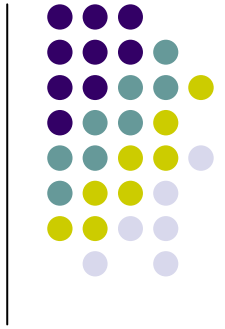


a b
c d

FIGURE 10.30
(a) Original image. (b) Result of global thresholding. (c) Image subdivided into individual subimages. (d) Result of adaptive thresholding.



- Detection of Discontinuities
- Edge Linking and Boundary Detection
- Thresholding
- Region-Based Segmentation
- Segmentation by Morphological Watersheds
- The Use of Motion in Segmentation



- Region Growing
- Region Splitting and Merging

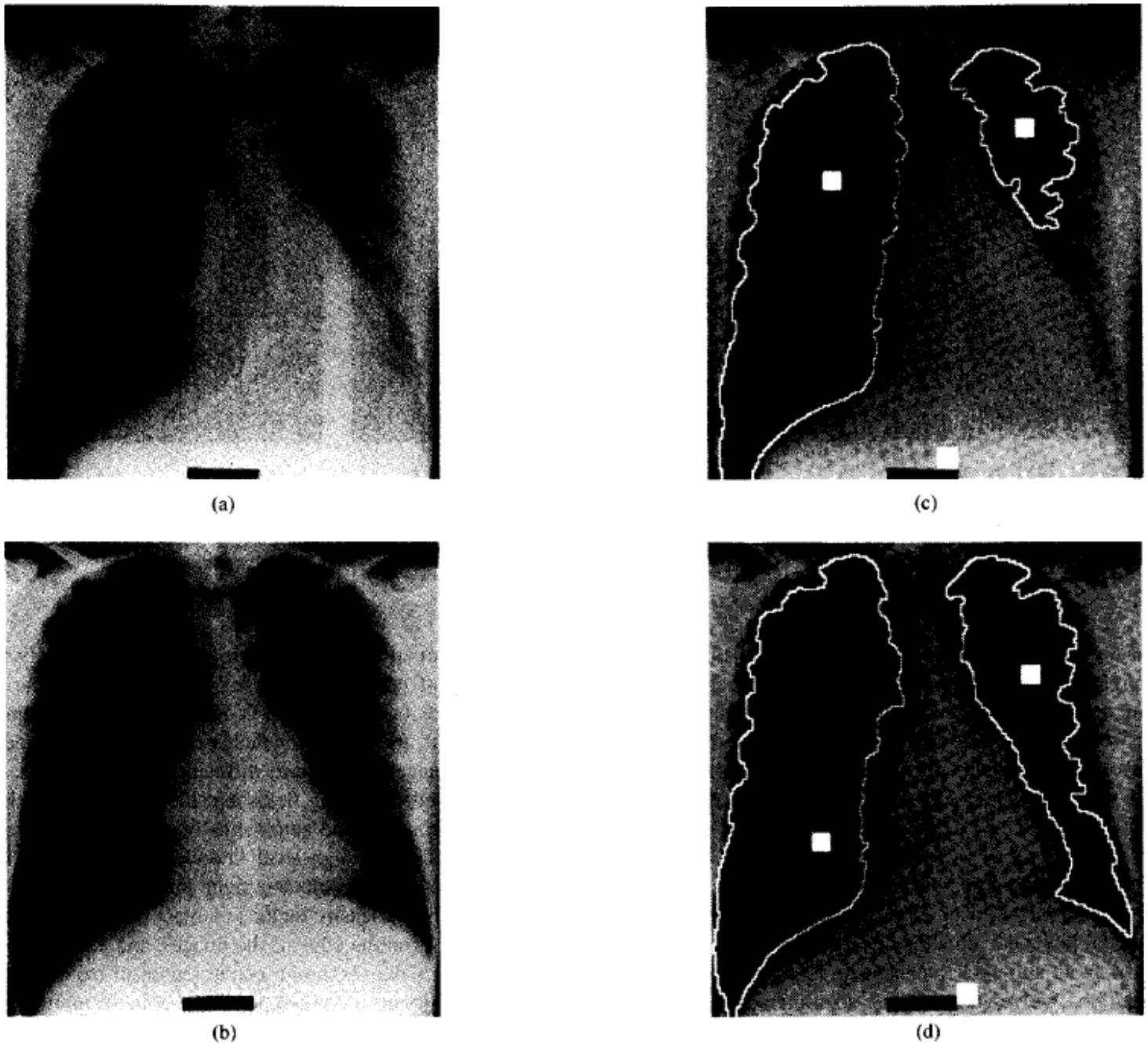
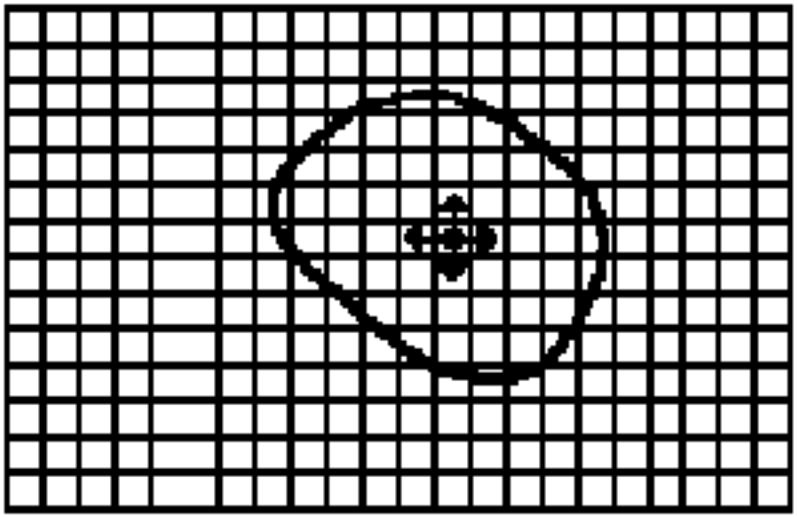


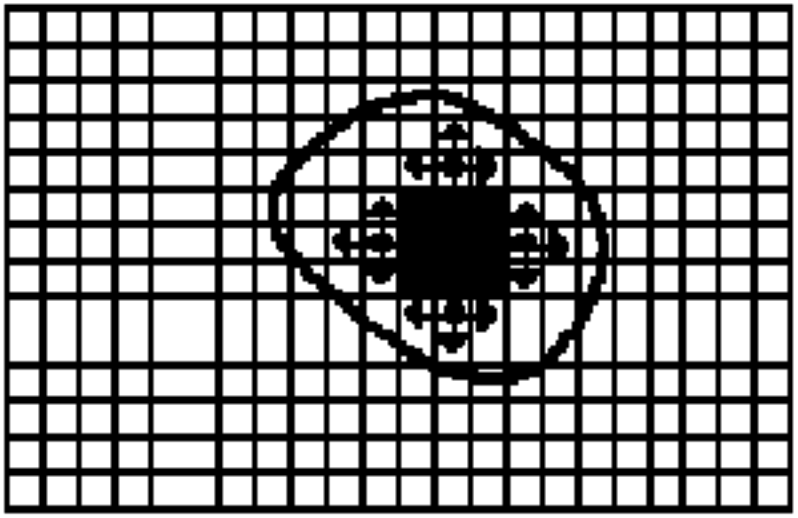
Fig. 5. (a) X-ray image of human chest in which the left and right lung fields are to be segmented. (b) A second X-ray image of human chest with very different left lung field. (c) Result of segmentation of (a) by SRG using automatically derived seed areas (10×10 boxes). These seeds, for the left and right lung fields and the region in between them, are found using a converging squares algorithm. (d) Result of segmentation of (b).



Seed Pixel

Direction of Growth

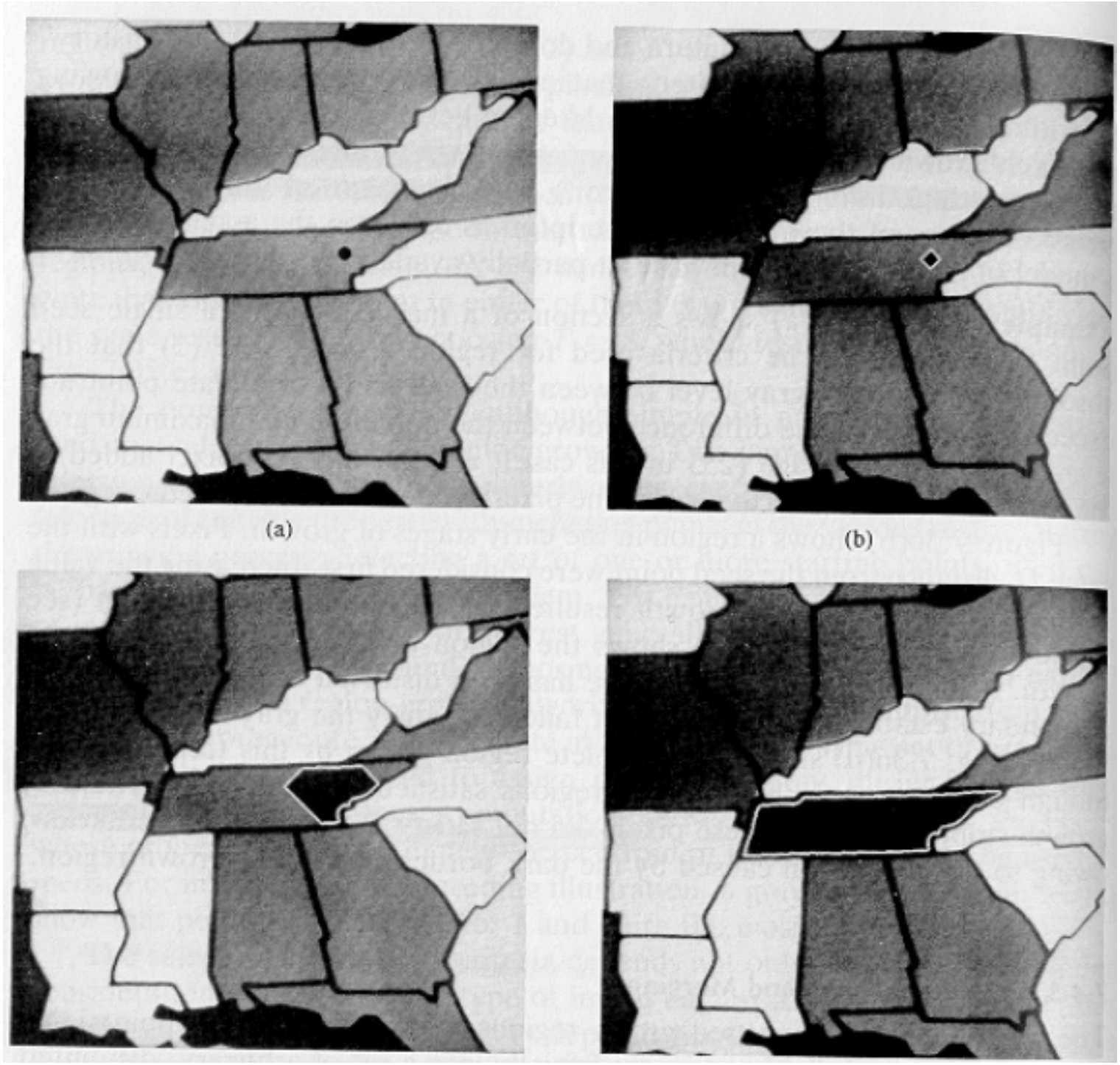
(a) Start of Growing a Region

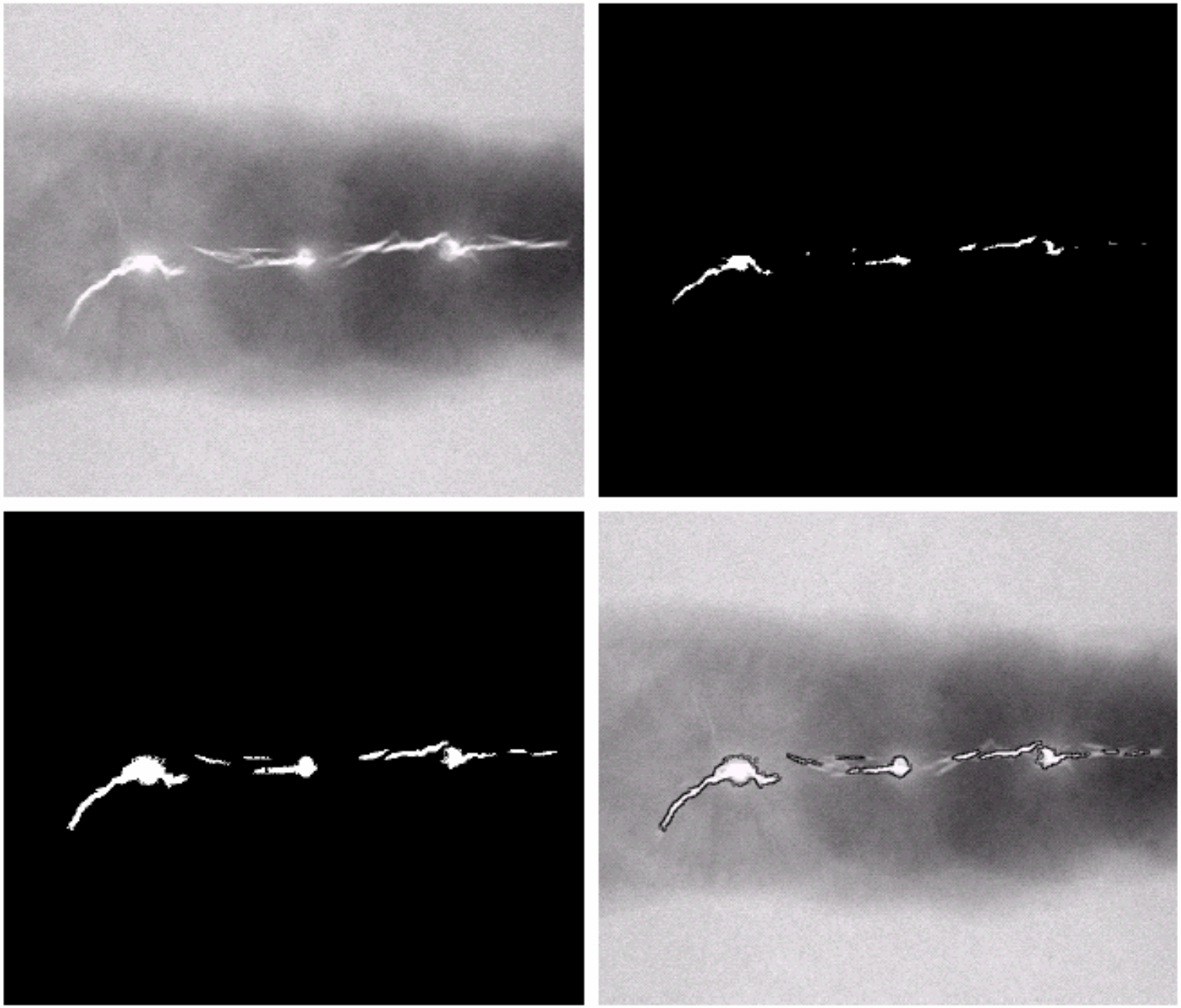


Grown Pixels

Pixels Being Considered

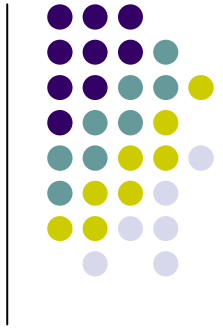
(b) Growing Process After a Few Iterations


(a)
(b)



a b
c d

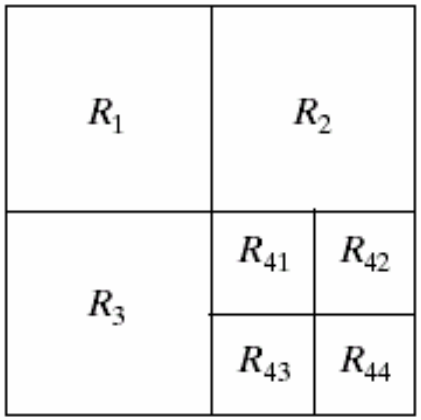
FIGURE 10.40
(a) Image showing defective welds. (b) Seed points. (c) Result of region growing. (d) Boundaries of segmented defective welds (in black). (Original image courtesy of X-TEK Systems, Ltd.).

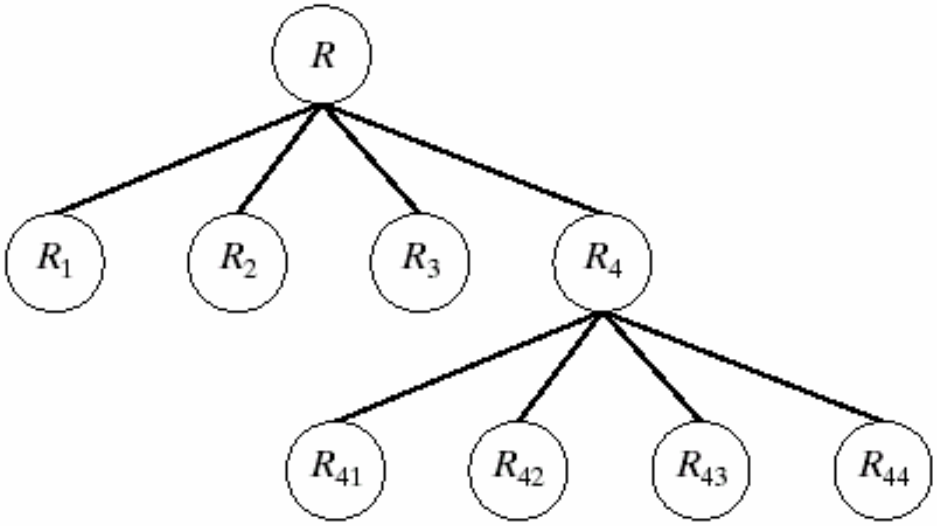


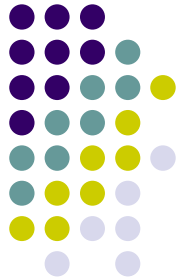
- Region Growing
- Region Splitting and Merging

a b

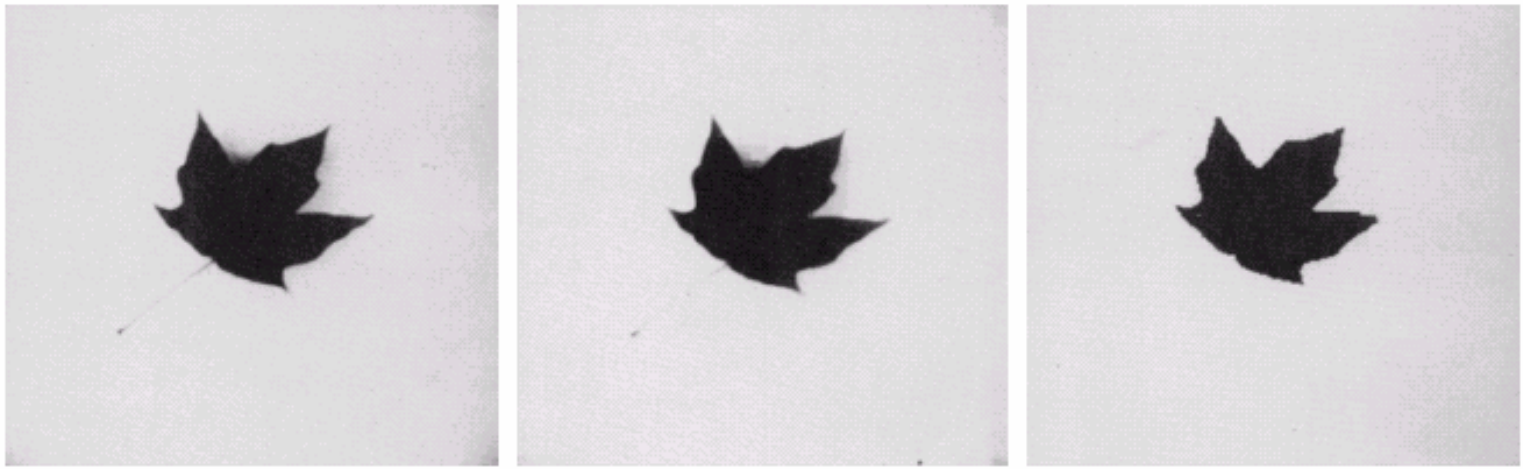
FIGURE 10.42
(a) Partitioned image.
(b) Corresponding quadtree.





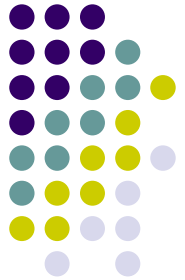


- Split into four disjoint quadrants any region R_i for which $P(R_i)$=FALSE;
- Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k) = TRUE$.
- Stop when no further merging or splitting is possible.



a b c

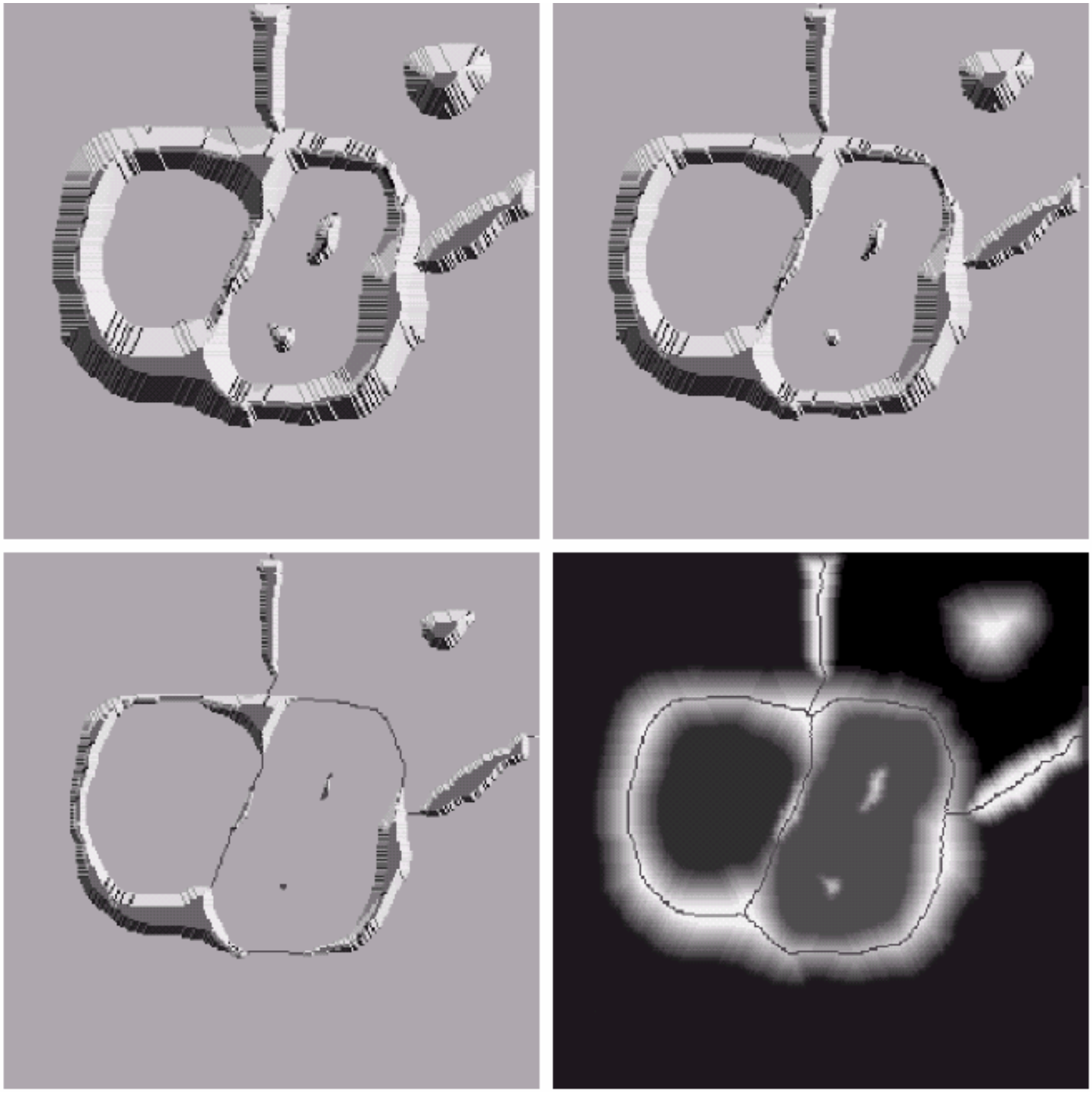
FIGURE 10.43
(a) Original image. (b) Result of split and merge procedure. (c) Result of thresholding (a).



- Detection of Discontinuities
- Edge Linking and Boundary Detection
- Thresholding
- Region-Based Segmentation
- Segmentation by Morphological Watersheds
- The Use of Motion in Segmentation

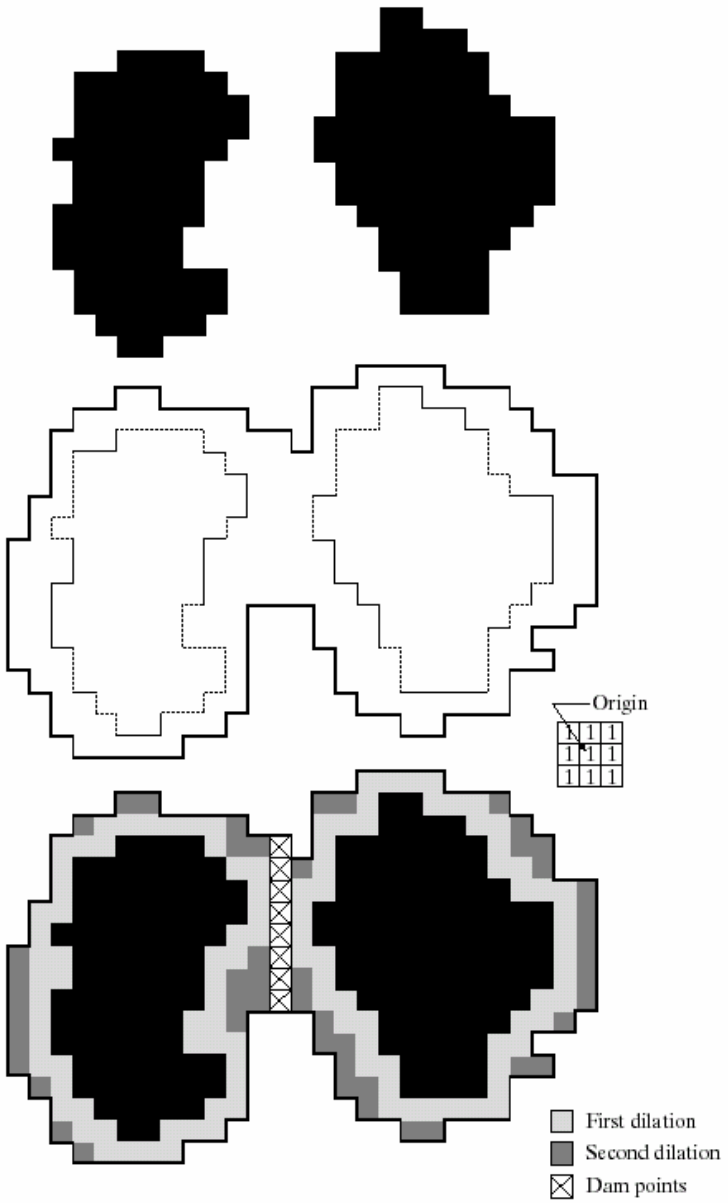
a b
c d

FIGURE 10.44
(a) Original image.
(b) Topographic view. (c)–(d) Two stages of flooding.



e f
g h

FIGURE 10.44
(Continued)
(e) Result of further flooding. (f) Beginning of merging of water from two catchment basins (a short dam was built between them). (g) Longer dams. (h) Final watershed (segmentation) lines. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)



a
b
d c

FIGURE 10.45 (a) Two partially flooded catchment basins at stage $n - 1$ of flooding. (b) Flooding at stage n, showing that water has spilled between basins (for clarity, water is shown in white rather than black). (c) Structuring element used for dilation. (d) Result of dilation and dam construction.

a b
c d

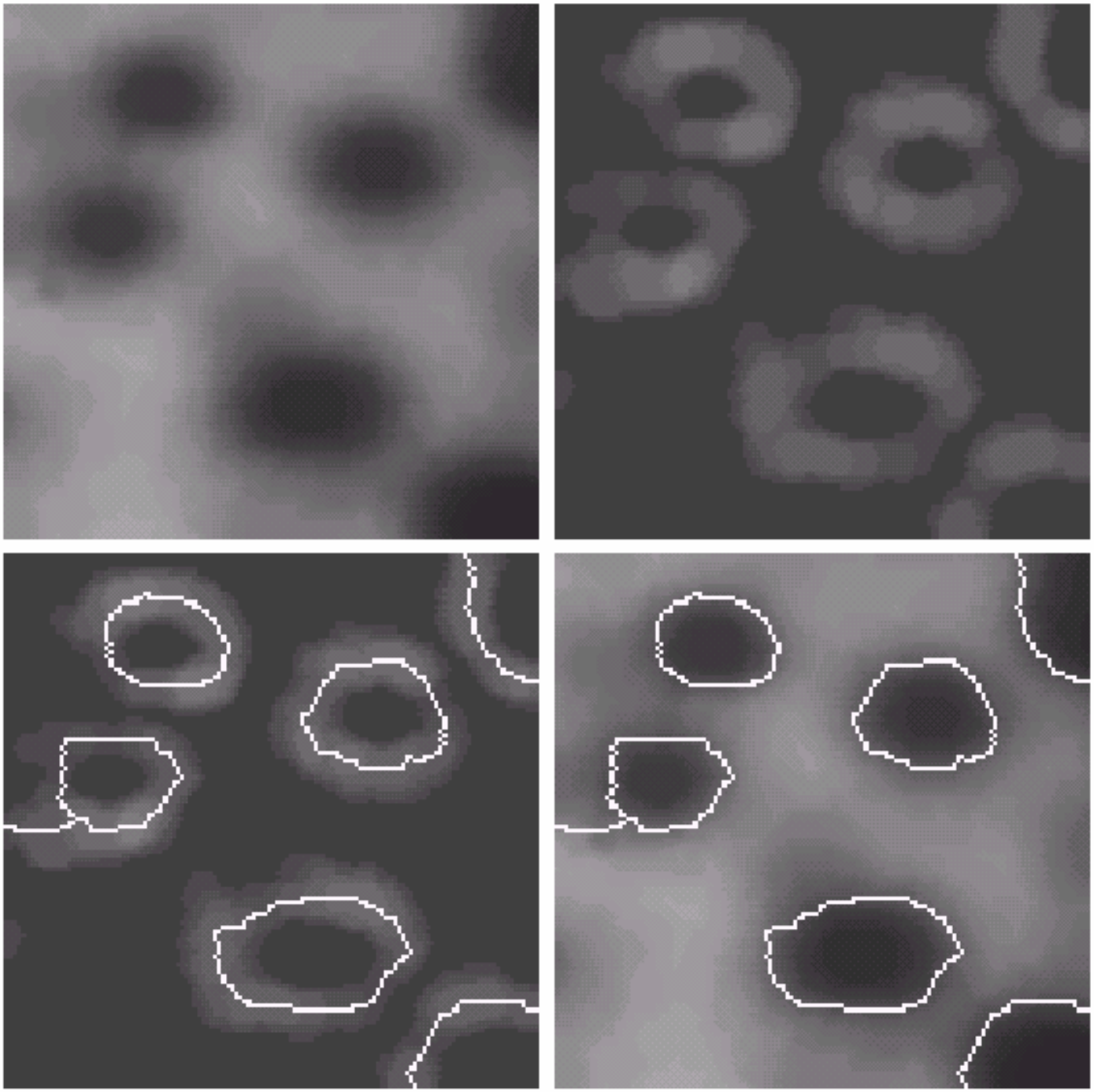
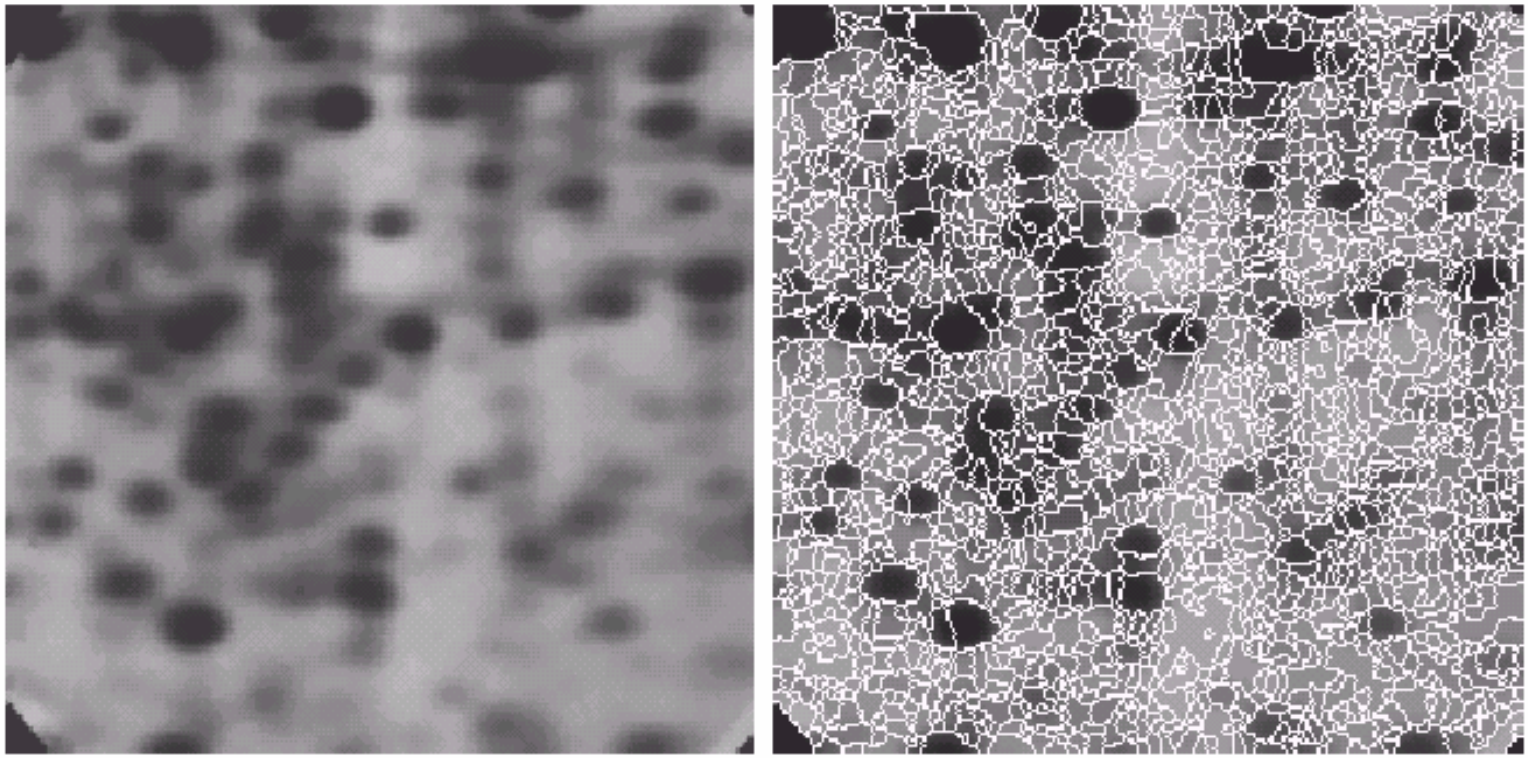
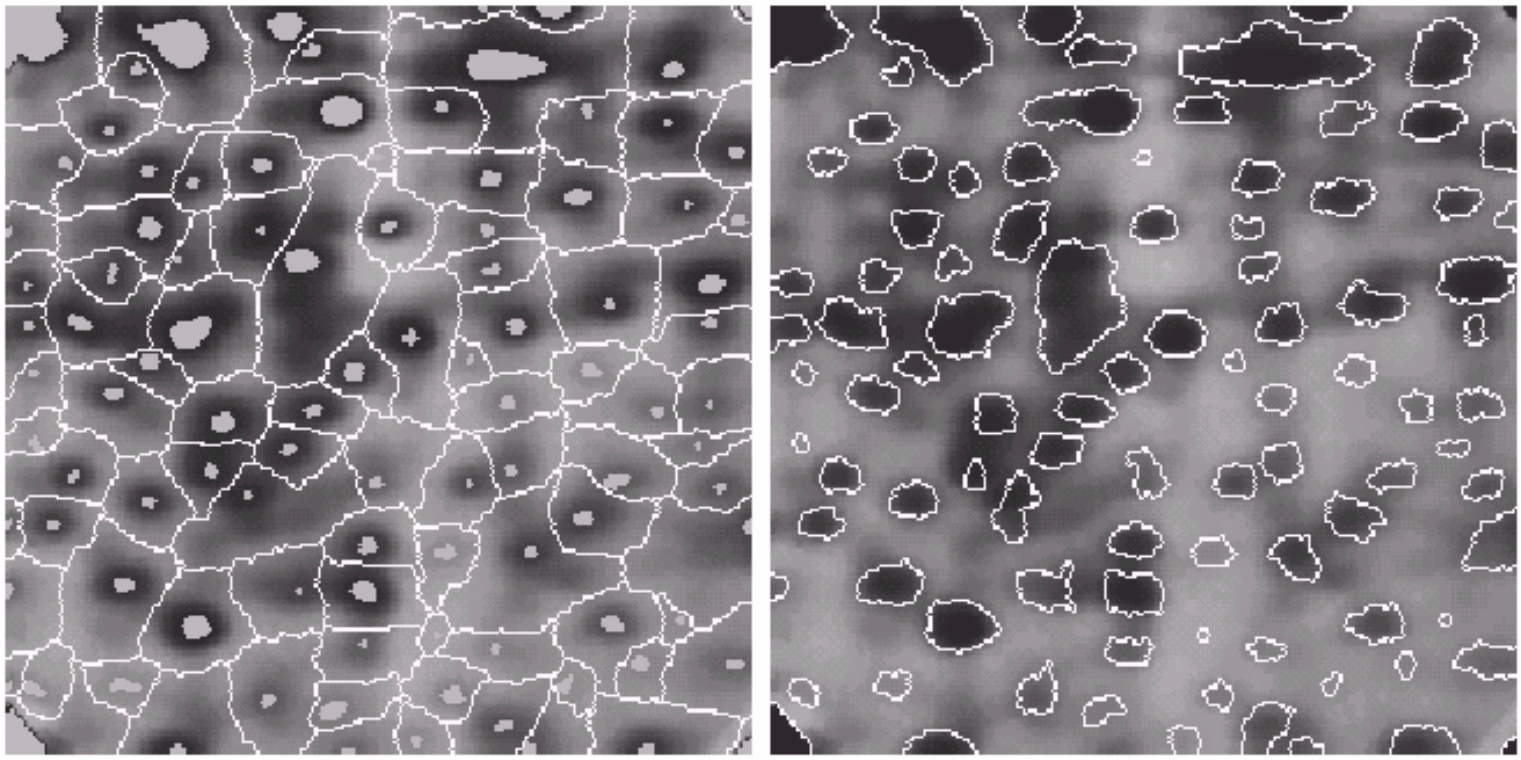


FIGURE 10.46
(a) Image of blobs. (b) Image gradient. (c) Watershed lines. (d) Watershed lines superimposed on original image. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)



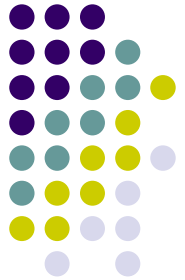
a b

FIGURE 10.47
(a) Electrophoresis image. (b) Result of applying the watershed segmentation algorithm to the gradient image. Oversegmentation is evident. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)

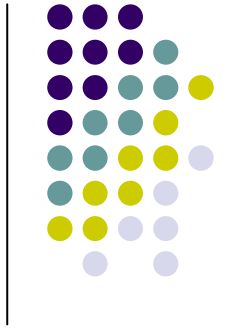


a b

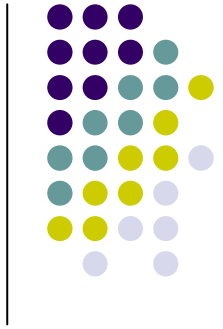
FIGURE 10.48
(a) Image showing internal markers (light gray regions) and external markers (watershed lines). (b) Result of segmentation. Note the improvement over Fig. 10.47(b). (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)



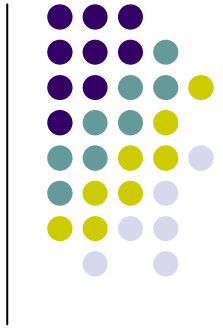
- Detection of Discontinuities
- Edge Linking and Boundary Detection
- Thresholding
- Region-Based Segmentation
- Segmentation by Morphological Watersheds
- The Use of Motion in Segmentation



- Spatial Techniques
- Frequency Domain Techniques



$$d_{ij}(x,y) = \begin{cases} 1 & \text{if} \quad \left| f(x,y,t_i) - f(x,y,t_j) \right| > T \\ 0 & \text{otherwise} \end{cases}$$

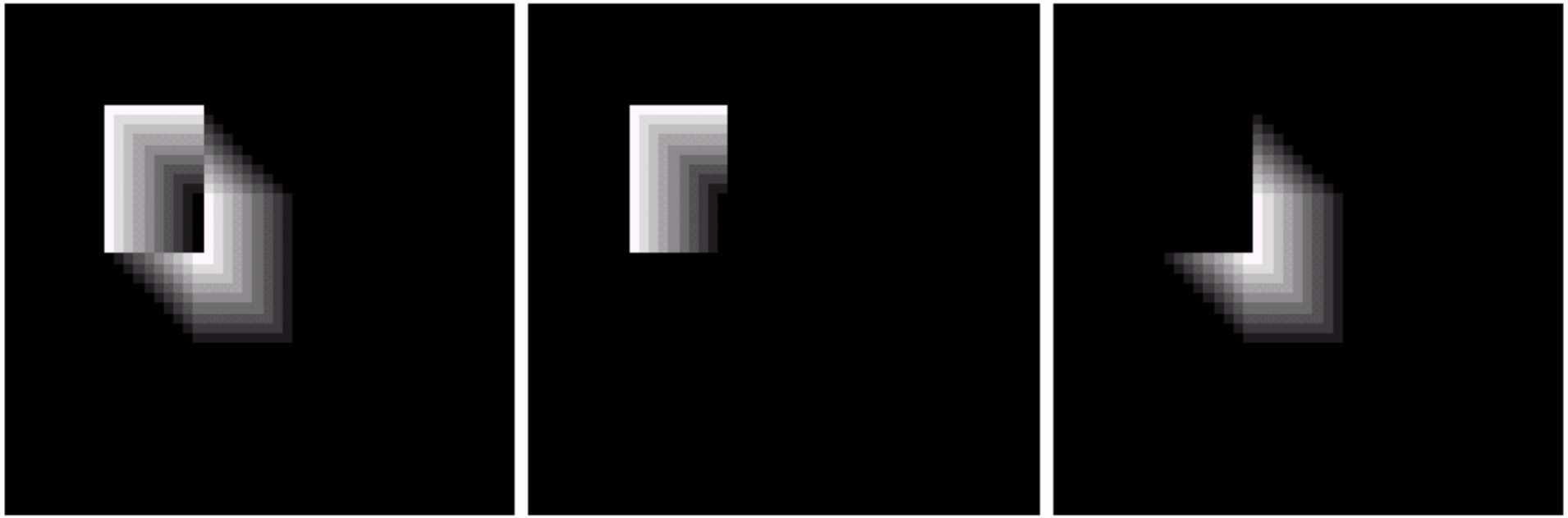


- $$A_k(x,y) = \begin{cases} A_{k-1}(x,y)+1 & \text{if } \left|R(x,y) - f(x,y,k)\right| > T \\ A_{k-1}(x,y) & \text{otherwise} \end{cases}$$

$$P_k(x,y) = \begin{cases} P_{k-1}(x,y)+1 & \text{if } \left|R(x,y) - f(x,y,k)\right| > T \\ P_{k-1}(x,y) & \text{otherwise} \end{cases}$$

- and

$$N_k(x,y) = \begin{cases} N_{k-1}(x,y)+1 & \text{if } \left|R(x,y) - f(x,y,k)\right| < -T \\ N_{k-1}(x,y) & \text{otherwise} \end{cases}$$

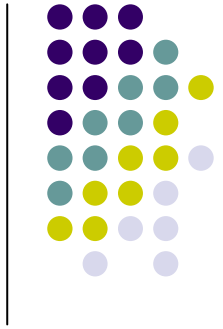


a b c

FIGURE 10.49 ADIs of a rectangular object moving in a southeasterly direction. (a) Absolute ADI. (b) Positive ADI. (c) Negative ADI.

a b c

FIGURE 10.50 Building a static reference image. (a) and (b) Two frames in a sequence. (c) Eastbound automobile subtracted from (a) and the background restored from the corresponding area in (b). (Jain and Jain.)



- Spatial Techniques
- Frequency Domain Techniques

$$e^{j2\pi a_1(x'+t)\Delta t} = \cos\left[2\pi a_1(x'+t)\Delta t\right] + j\sin\left[2\pi a_1(x'+t)\Delta t\right]$$

$$g_x(t, a_1) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y, t)e^{j2\pi a_1 x\Delta t} \qquad t = 0, 1, \ldots, K-1. \qquad (10.6\text{-}6)$$

Similarly, the sum of the projections onto the y-axis is

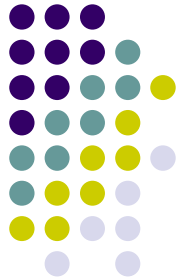
$$g_y(t, a_2) = \sum_{y=0}^{N-1}\sum_{x=0}^{M-1} f(x, y, t)e^{j2\pi a_2 y\Delta t} \qquad t = 0, 1, \ldots, K-1 \qquad (10.6\text{-}7)$$

The 1-D Fourier transforms of Eqs. (10.6-6) and (10.6-7), respectively, are

$$G_x(u_1, a_1) = \frac{1}{K}\sum_{t=0}^{K-1} g_x(t, a_1)e^{-j2\pi u_1 t/K} \qquad u_1 = 0, 1, \ldots, K-1 \qquad (10.6\text{-}8)$$

and

$$G_y(u_2, a_2) = \frac{1}{K}\sum_{t=0}^{K-1} g_y(t, a_2)e^{-j2\pi u_2 t/K} \qquad u_2 = 0, 1, \ldots, K-1. \qquad (10.6\text{-}9)$$



- The frequency-velocity relationship is

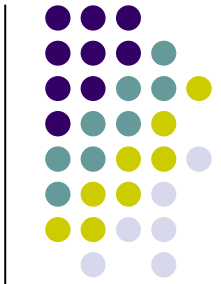
$$u_1 = a_1 v_1$$

and

$$u_2 = a_2 v_2.$$

- The actural physical speed in the x-direxion is

$$\begin{aligned} v_1 &= (10 \text{ pixels})(0.5 \text{ m/pixel})(2 \text{ frames/s})/(30 \text{ frames}) \\ &= 1/3 \text{ m/s}. \end{aligned}$$



The sign of the x-component of the velocity is obtained by computing

$$S_{1x} = \left. \frac{d^2 \mathrm{Re}[g_x(t, a_1)]}{dt^2} \right|_{t=n} \tag{10.6-12}$$

and

$$S_{2x} = \left. \frac{d^2 \mathrm{Im}[g_x(t, a_1)]}{dt^2} \right|_{t=n} \tag{10.6-13}$$

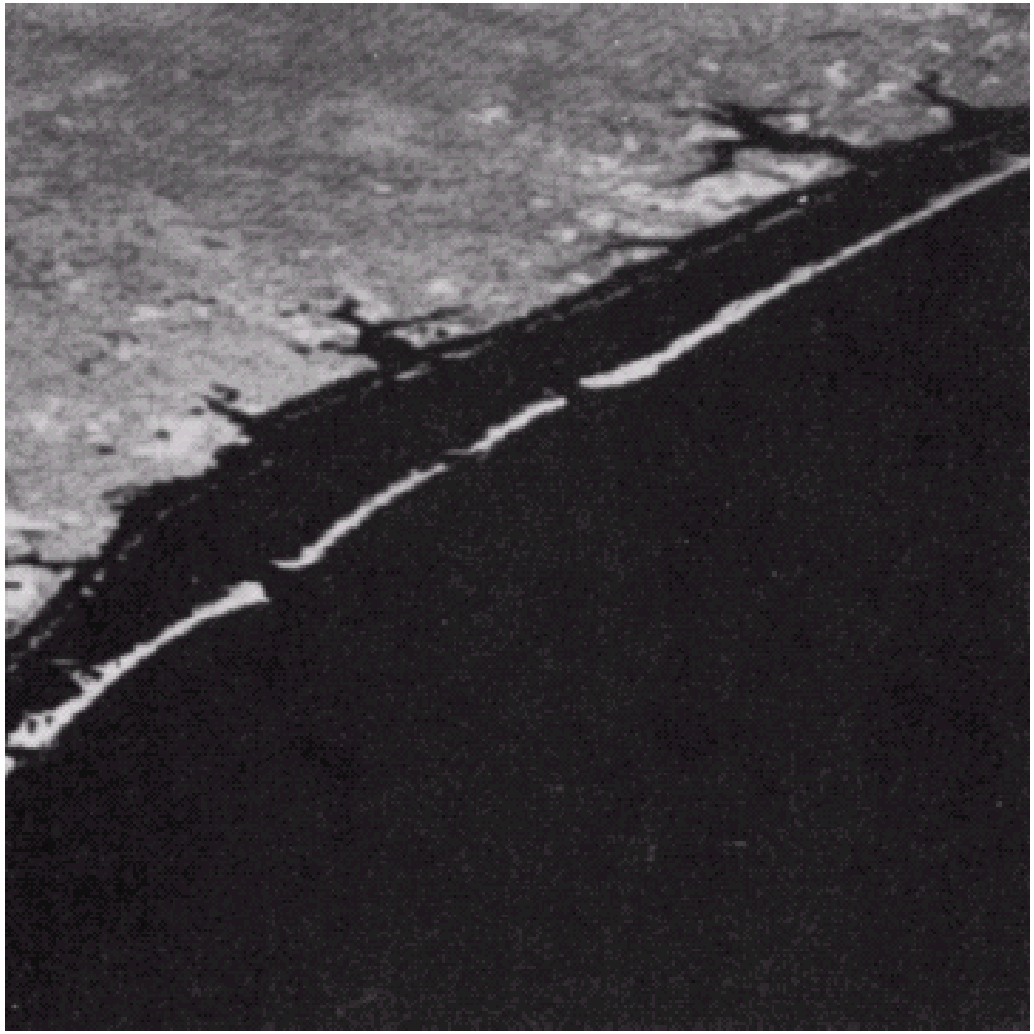


FIGURE 10.51
LANDSAT frame. (Cowart, Snyder, and Ruedger.)

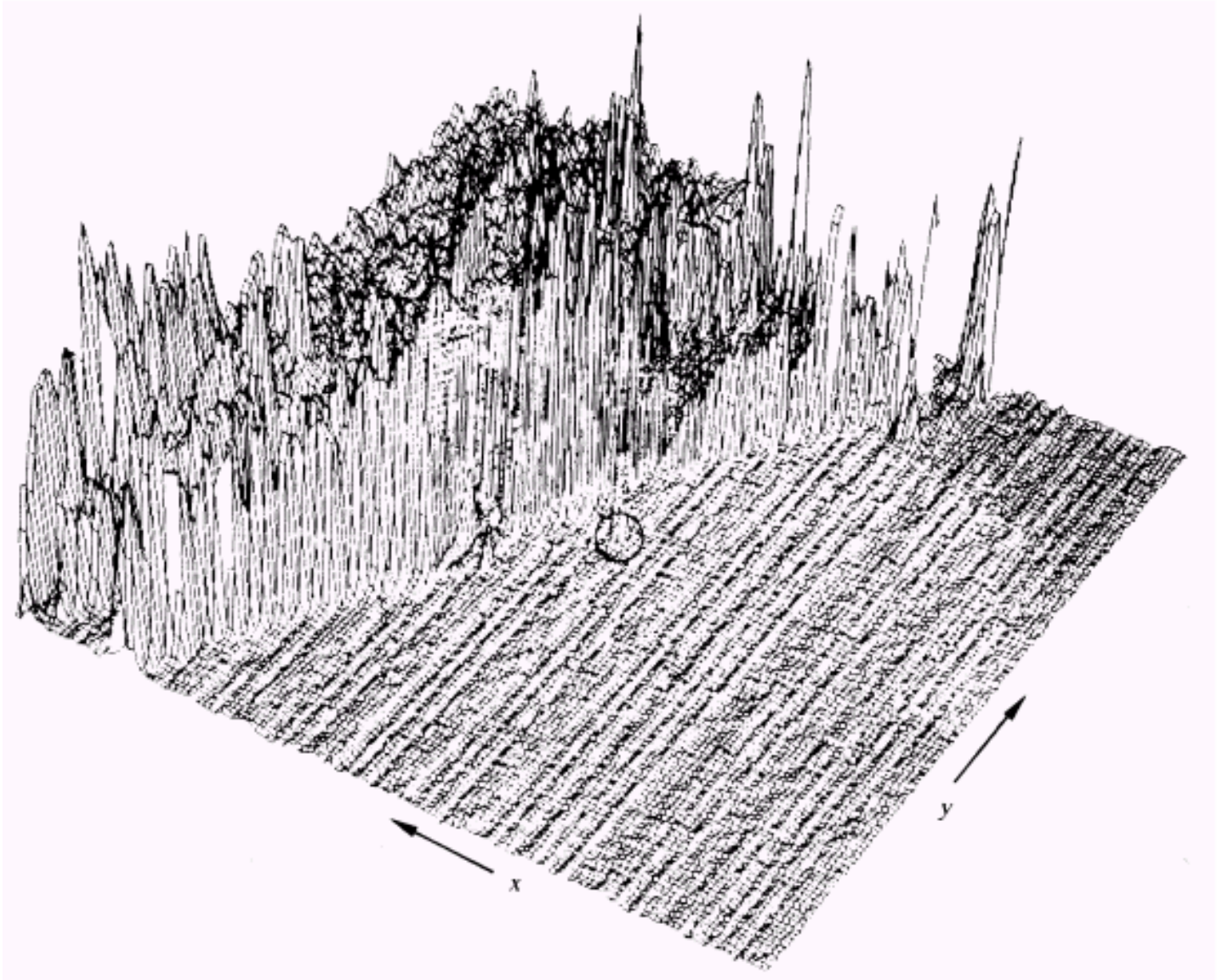


FIGURE 10.52
Intensity plot of the image in Fig. 10.51, with the target circled. (Rajala, Riddle, and Snyder.)

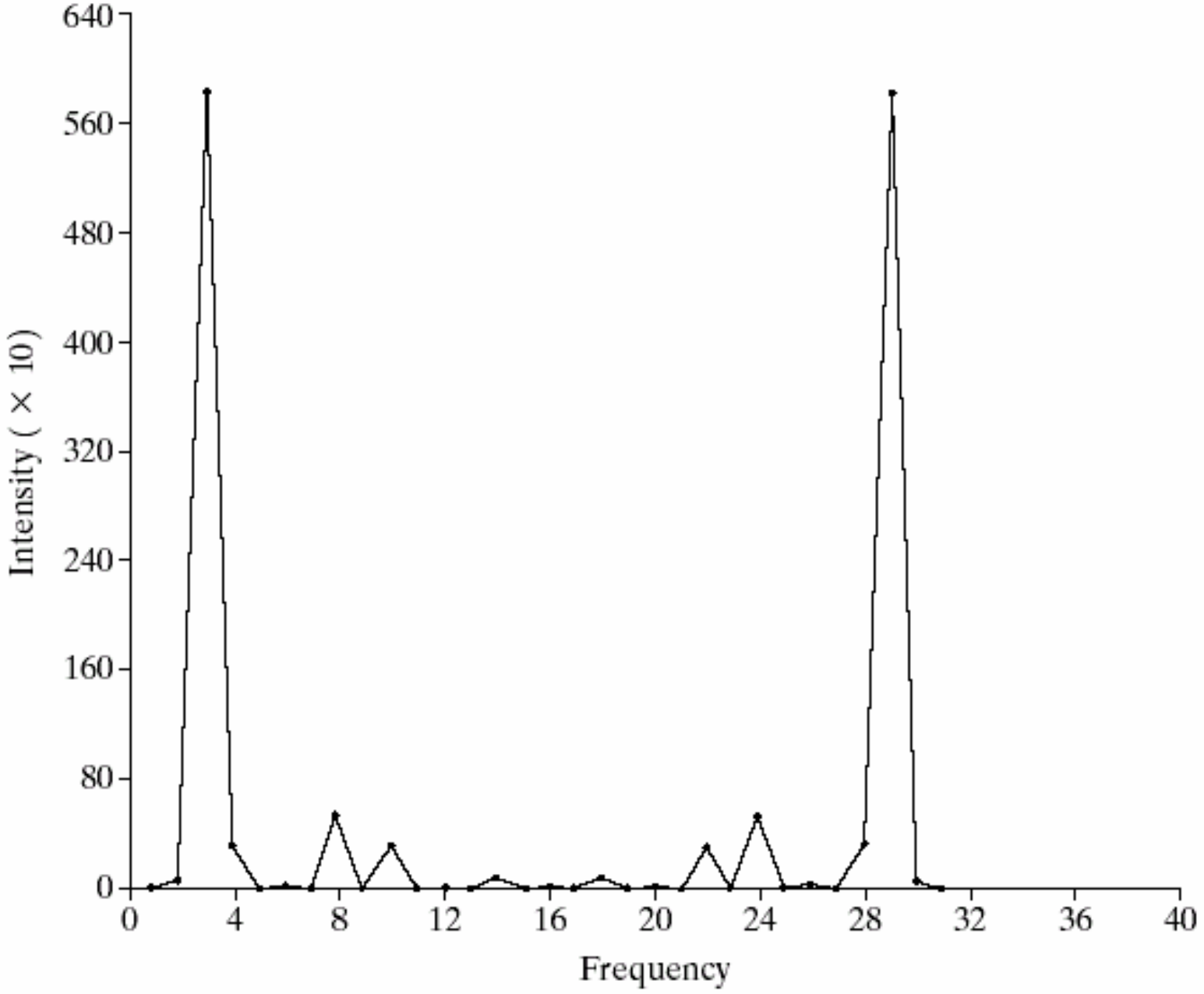


FIGURE 10.53 Spectrum of Eq. (10.6-8) showing a peak at $u_1 = 3$. (Rajala, Riddle, and Snyder.)

FIGURE 10.54
Spectrum of Eq. (10.6-9) showing a peak at $u_2 = 4$. (Rajala, Riddle, and Snyder.)