Digital Image Processing

Fourth Edition

Chapter 12

Feature Extraction

Gonzalez 2018, pp. 1075

Contents

Background

Boundary Preprocessing

Boundary Feature Descriptors

Region Feature Descriptors

Principal Components as Feature Descriptors

Whole-Image Features

Scale-Invariant Feature Transform (SIFT)

Figure 12.1

Illustration of the first few steps in the boundary-following algorithm. The point to be processed next is labeled in bold, black; the points yet to be processed are gray; and the points found by the algorithm are shaded. Squares without labels are considered background (0) values.
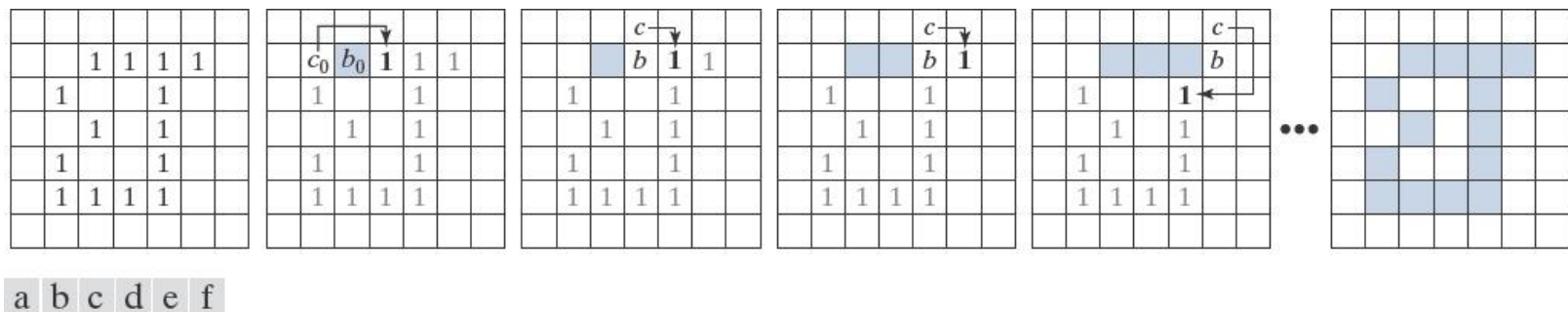

a b c d e f

Figure 12.2

Examples of boundaries that can be processed by the boundary-following algorithm. (a) Closed boundary with a branch. (b) Self-intersecting boundary. (c) Multiple boundaries (processed one at a time).
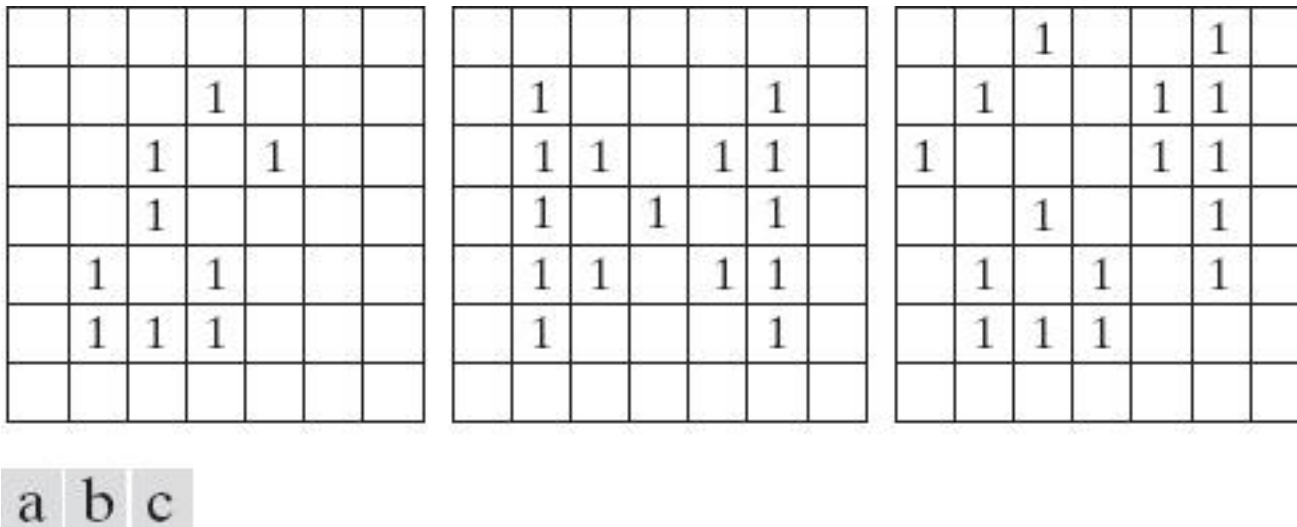

a b c

Pearson

Figure 12.3

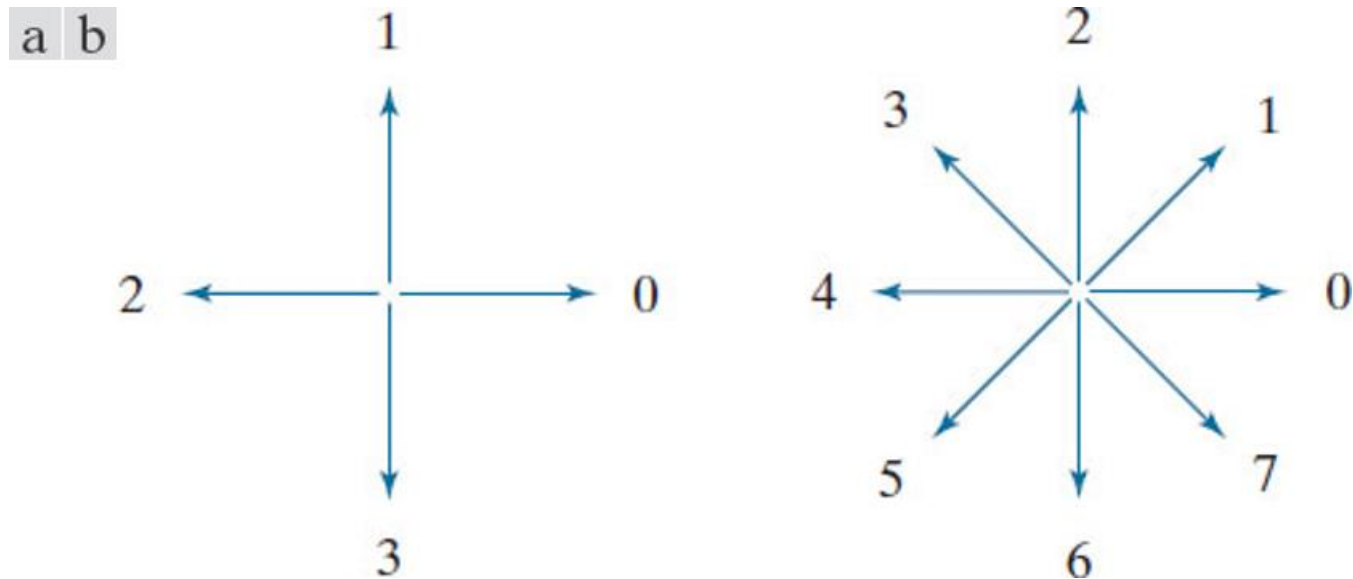Direction numbers for (a) 4-directional chain code, and (b) 8-directional chain code.

Figure 12.4

(a) Digital boundary with resampling grid superimposed. (b) Result of resampling. (c) 8-directional chain-coded boundary.
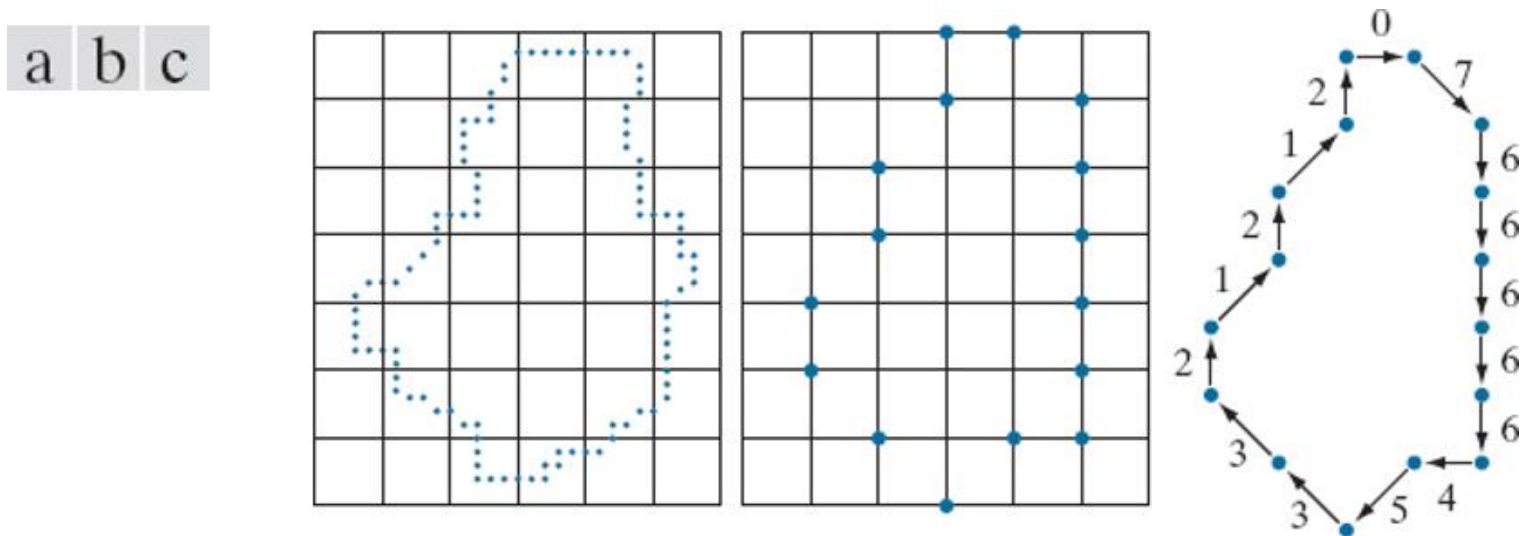
Figure 12.5

(a) Noisy image of size 570×570 pixels (b) Image smoothed with a 9×9 box kernel. (c) Smoothed image, thresholded using Otsu's method. (d) Longest outer boundary of (c). (e) Subsampled boundary (the points are shown enlarged for clarity). (f) Connected points from (e).
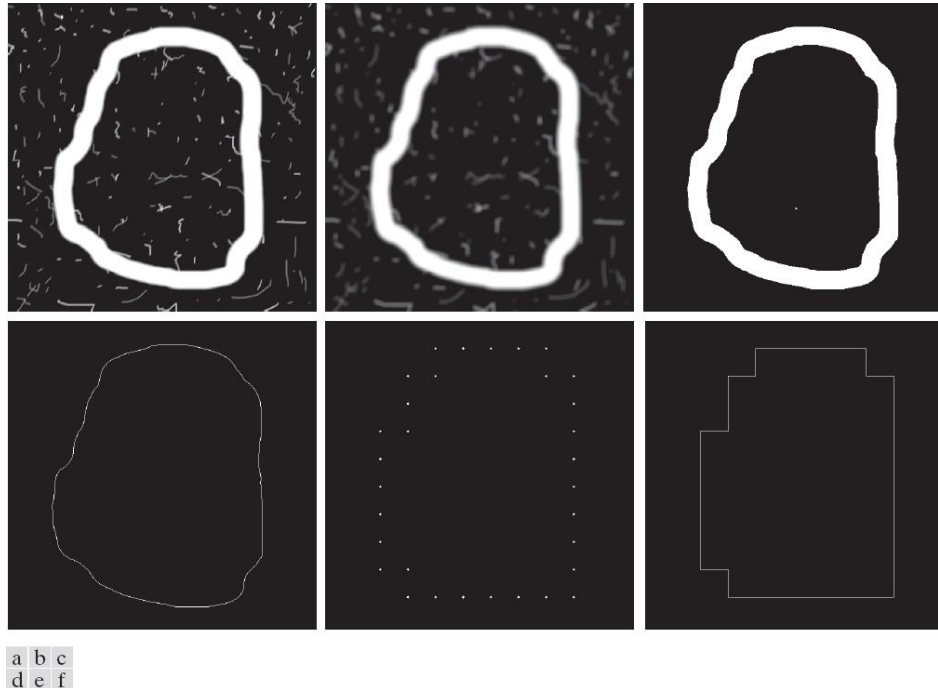

a b c
d e f

Figure 12.6

(a) An open curve. (b) A straight-line segment. (c) Traversing the curve using circumferences to determine slope changes; the dot is the origin (starting point). (d) Range of slope changes in the open interval (−1,1) (the arrow in the center of the chart indicates direction of travel). There can be ten subintervals between the slope numbers shown.(e) Resulting coded curve showing its corresponding numerical sequence of slope changes. (Courtesy of Professor Ernesto Bribiesca, IIMAS-UNAM, Mexico.)
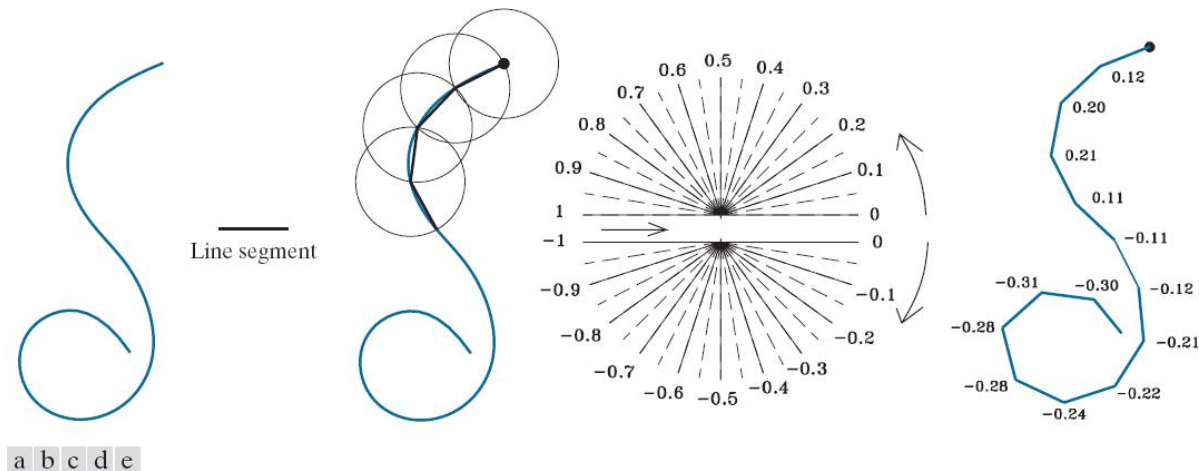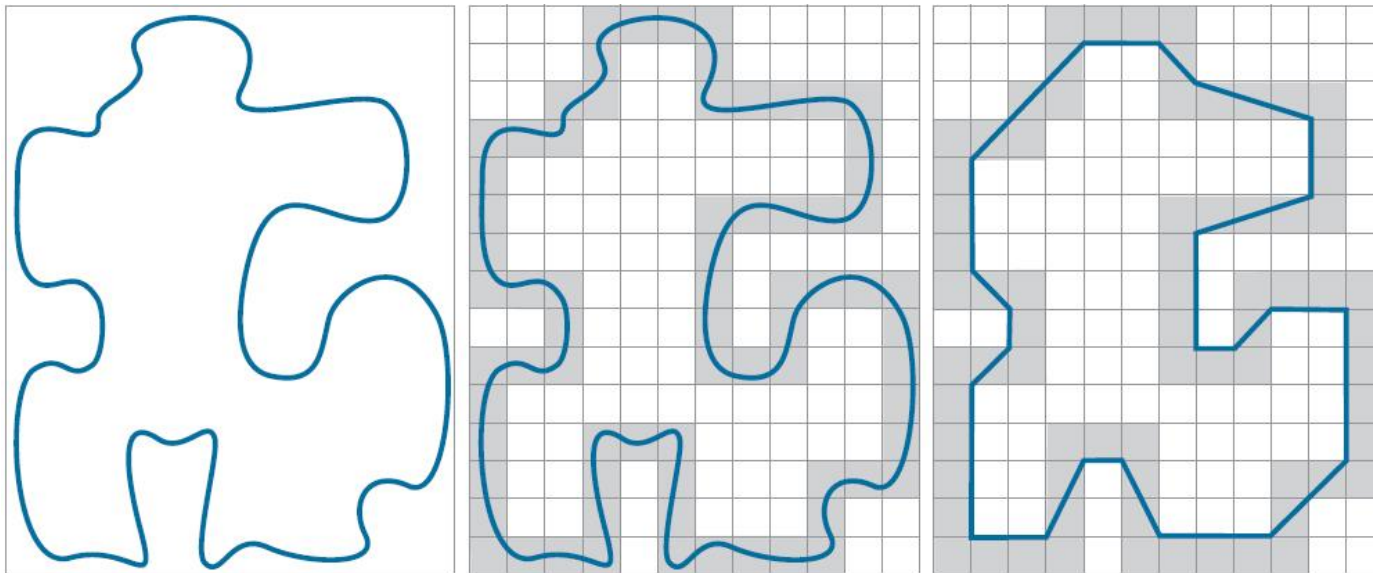
Figure 12.7

(a) An object boundary. (b) Boundary enclosed by cells (shaded). (c) Minimum-perimeter polygon obtained by allowing the boundary to shrink. The vertices of the polygon are created by the corners of the inner and outer walls of the gray region.

a b c

Figure 12.8

(a) Region (dark gray) resulting from enclosing the original boundary by cells (see Fig. 12.7). (b) Convex (white dots) and concave (blue dots) vertices obtained by following the boundary of the dark gray region in the counterclockwise direction. (c) Concave vertices (blue dots) displaced to their diagonal mirror locations in the outer wall of the bounding region; the convex vertices are not changed. The MPP (solid boundary) is superimposed for reference.
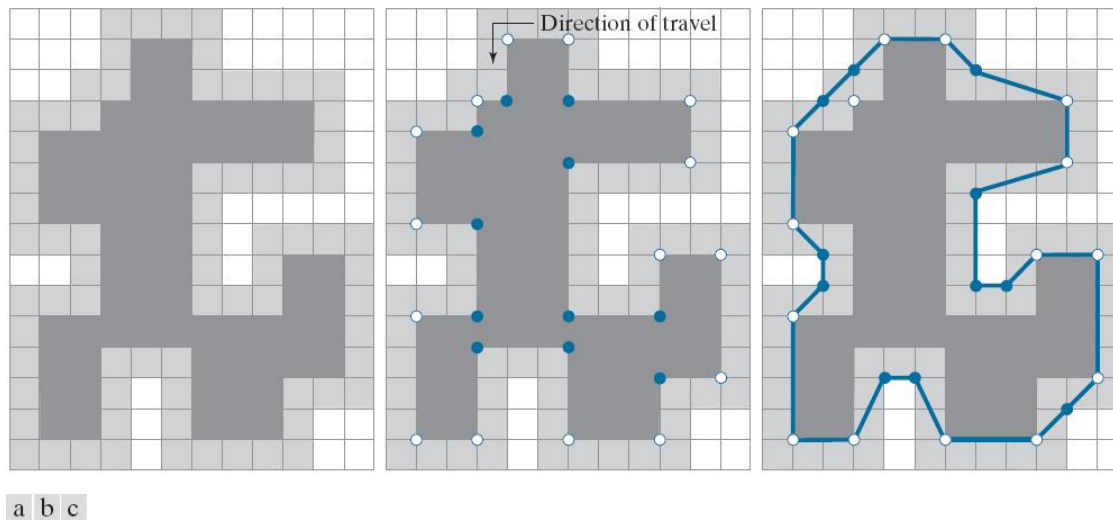

a b c

Figure 12.9

(a) 566×566 binary image. (b) 8-connected boundary. (c) through (h), MMPs obtained using square cells of sizes 2, 4, 6, 8, 16, and 32, respectively (the vertices were joined by straight-line segments for display). The number of boundary points in (b) is 1900. The numbers of vertices in (c) through (h) are 206, 127, 92, 66, 32, and 13, respectively. Images (b) through (h) are shown as negatives to make the boundaries easier to see.
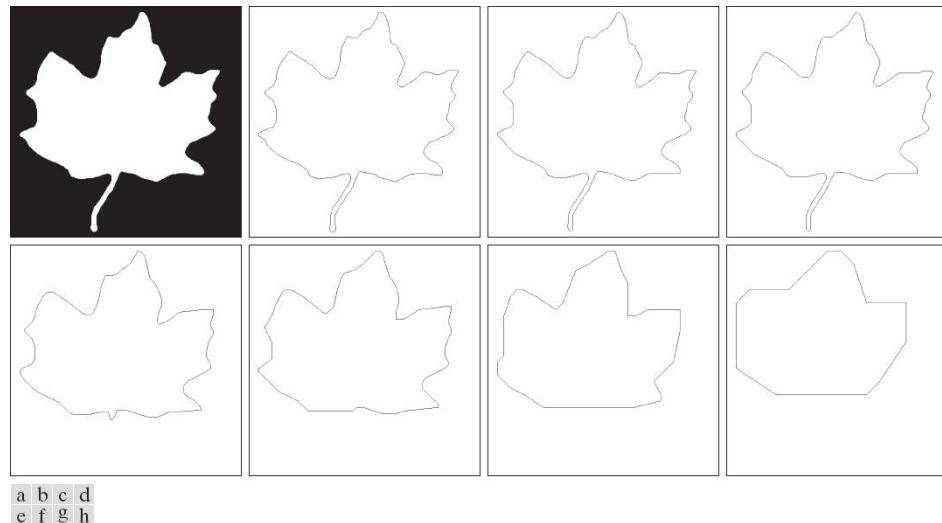

a b c d
e f g h

Figure 12.10

Distance-versus angle signatures. In (a), $r(\theta)$ is constant. In (b), the signature consists of repetitions of the pattern $r(\theta) = A\sec\theta$ for $0 \le \theta \le \dfrac{\pi}{4}$, and $r(\theta) = A\csc\theta$ for $\dfrac{\pi}{4} < \theta \le \dfrac{\pi}{2}$
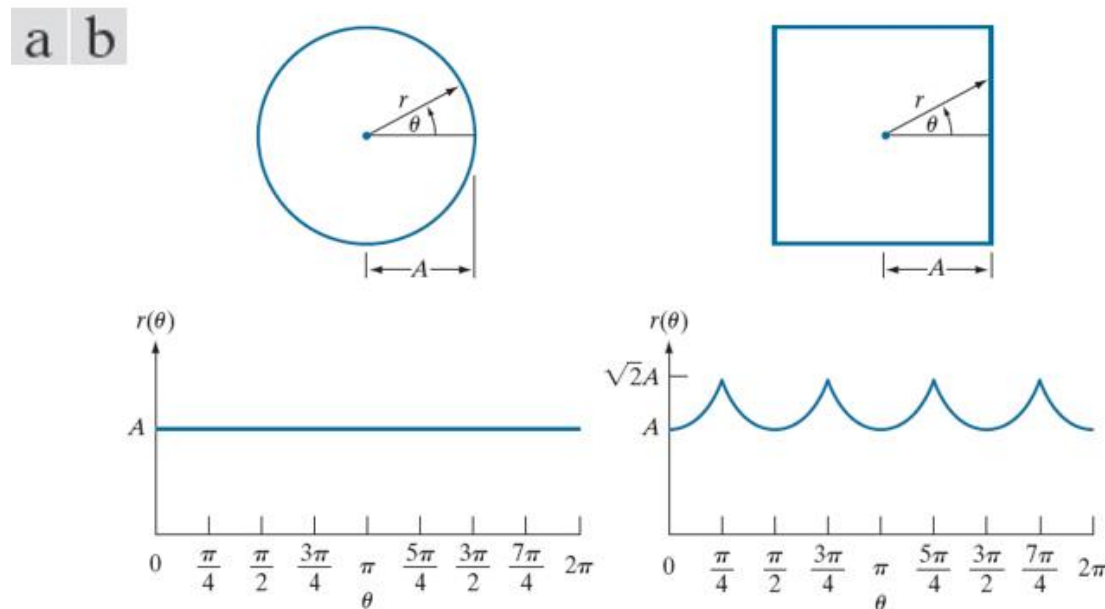
Figure 12.11

(a) and (d) Two binary regions, (b) and (e) their external boundaries, and (c) and (f) their corresponding $r(\theta)$ signature. The horizontal axes in (c) and (f) correspond to angles from 0 degree to 360 degree, in increments of 1 degree.
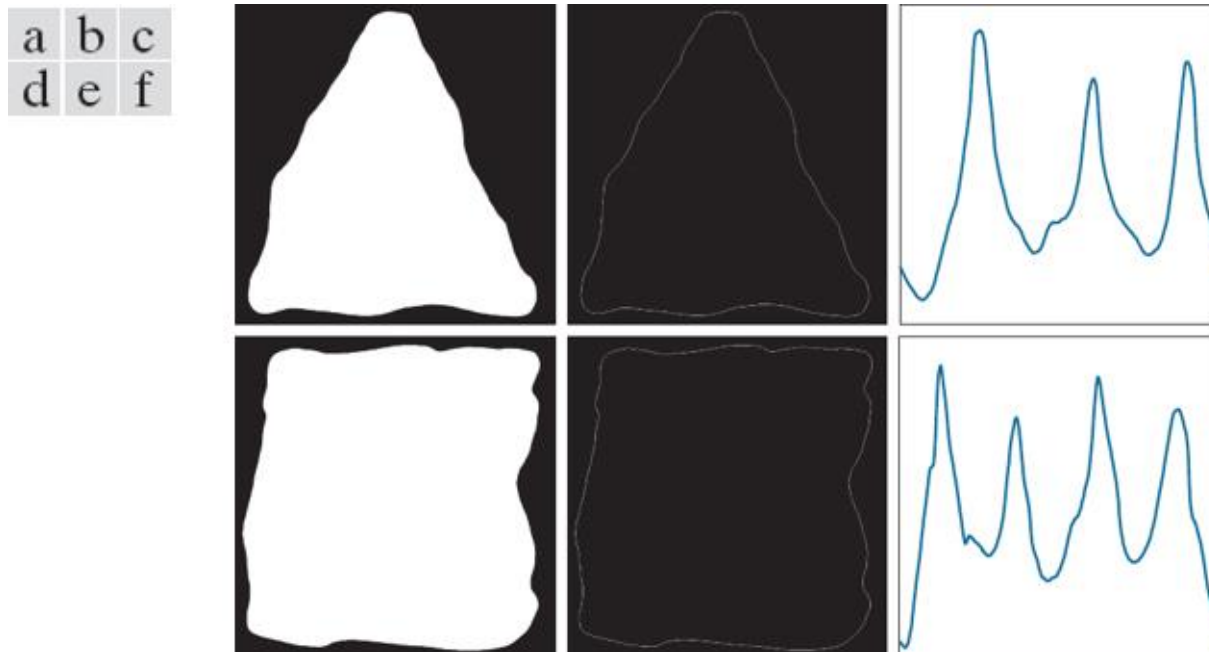
Figure 12.12

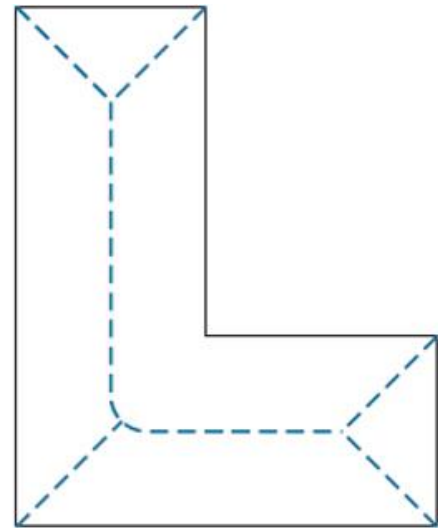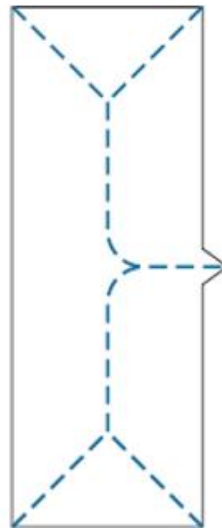Medial axes (dashed) of three simple regions.

Figure 12.13

(a) A small image and (b) its distance transform. Note that all 1-valued pixels in (a) have corresponding 0's in (b). (c) A small image, and (d) the distance transform of its complement. (e) A larger image, and (f) the distance transform of its complement. The Euclidian distance was used throughout.
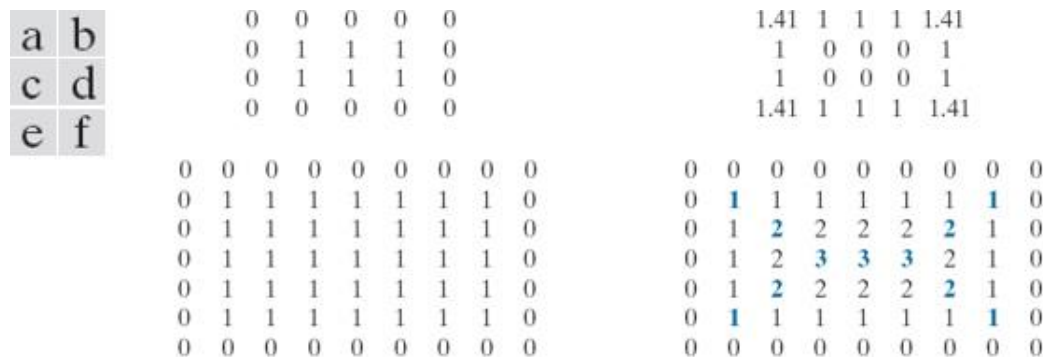
Figure 12.14

(a) Thresholded image of blood vessels. (b) Skeleton obtained by thinning, shown superimposed on the image (note the spurs). (c) Result of 40 passes of spur removal. (d) Skeleton obtained using the distance transform.
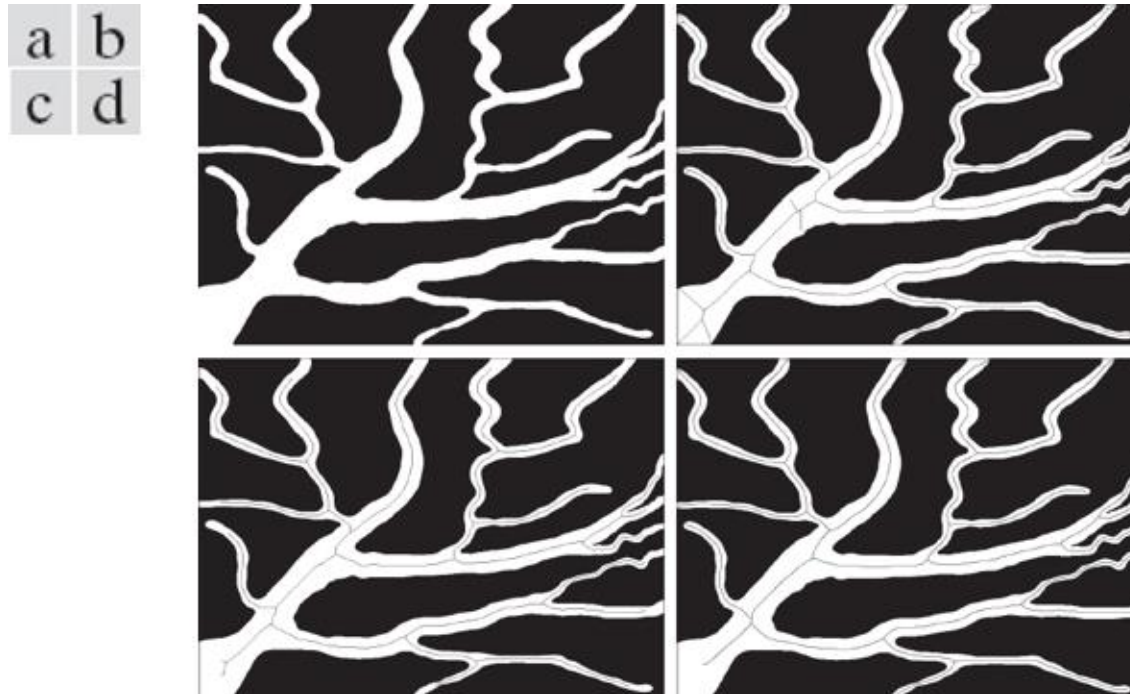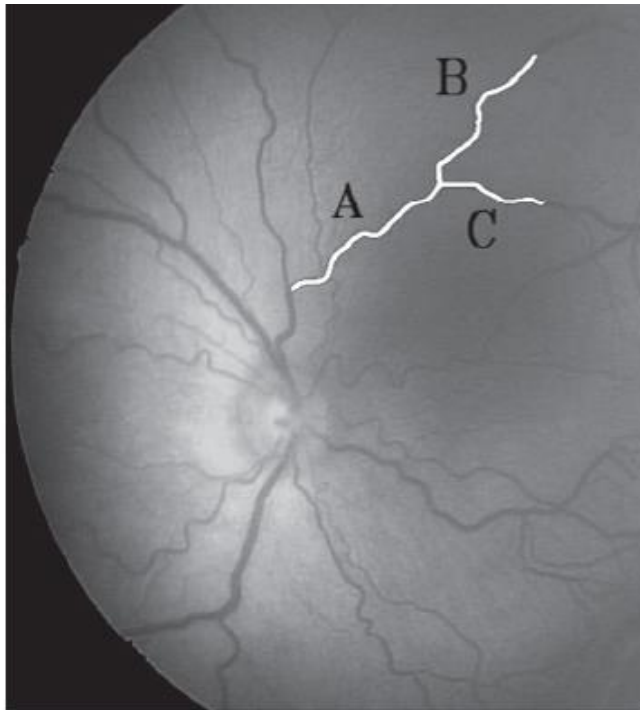
Figure 12.15

(a) Fundus image from a prematurely born baby with ROP. (b) Tortuosity of vessels A, B, and C. (Courtesy of Professor Ernesto Bribiesca, IIMASUNAM, Mexico.)

Curve	n	τ
A	50	2.3770
B	50	2.5132
C	50	1.6285

Figure 12.16

All shapes of order 4, 6, and 8. The directions are from Fig. 12.3(a), and the dot indicates the starting point.

Order 4

Chain code: 0 3 2 1
Difference: 3 3 3 3
Shape no.: 3 3 3 3

Order 6

0 0 3 2 2 1
3 0 3 3 0 3
0 3 3 0 3 3

Order 8

Chain code: 0 0 3 3 2 2 1 1   0 3 0 3 2 2 1 1   0 0 0 3 2 2 2 1
Difference: 3 0 3 0 3 0 3 0   3 3 1 3 3 0 3 0   3 0 0 3 3 0 0 3
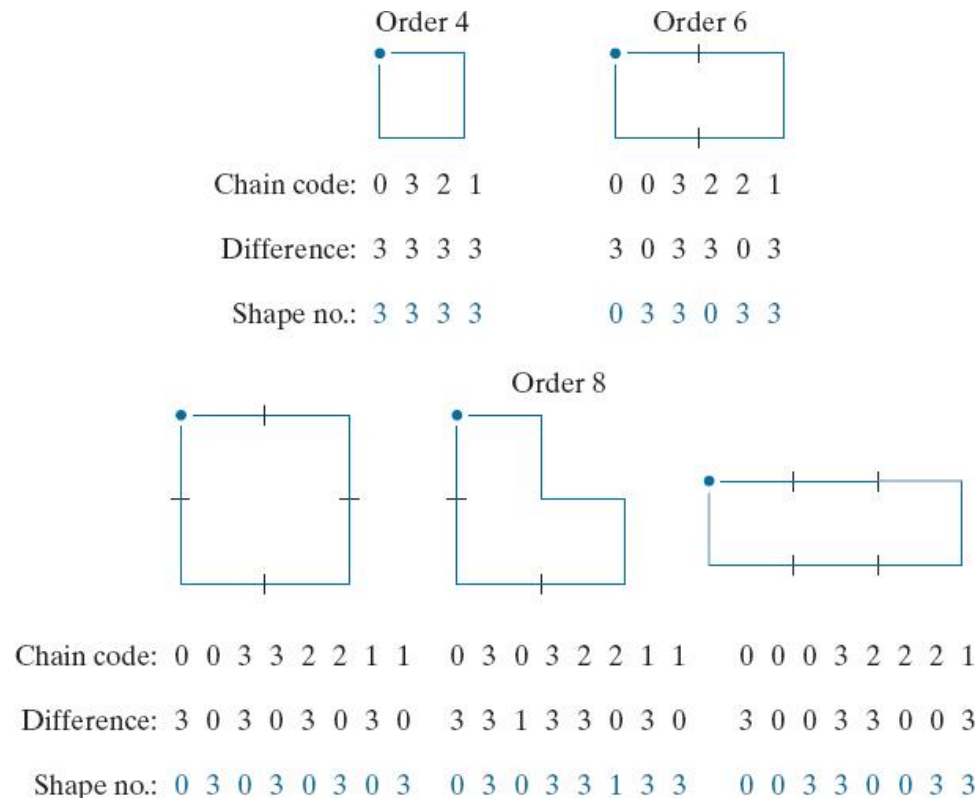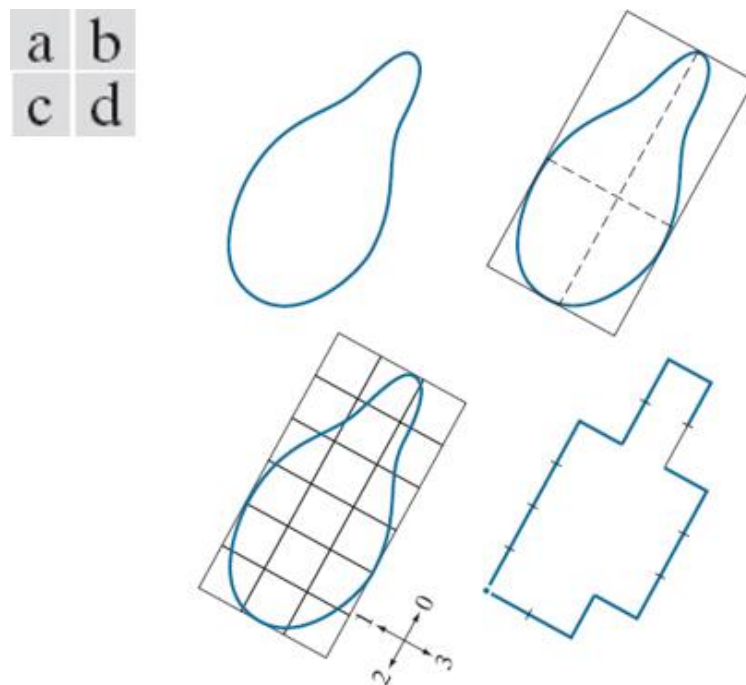Shape no.: 0 3 0 3 0 3 0 3   0 3 0 3 3 1 3 3   0 0 3 3 0 0 3 3

Figure 12.17

Steps in the generation of a shape number.

a b
c d

Chain code: 0 0 0 0 3 0 0 3 2 2 3 2 2 2 1 2 1 1

Difference: 3 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0

Shape no.: 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0 3

P Pearson

Figure 12.18

A digital boundary and its representation as sequence of complex numbers. The points (x_0, y_0) and (x_1, y_1) are (arbitrarily) the first two points in the sequence.
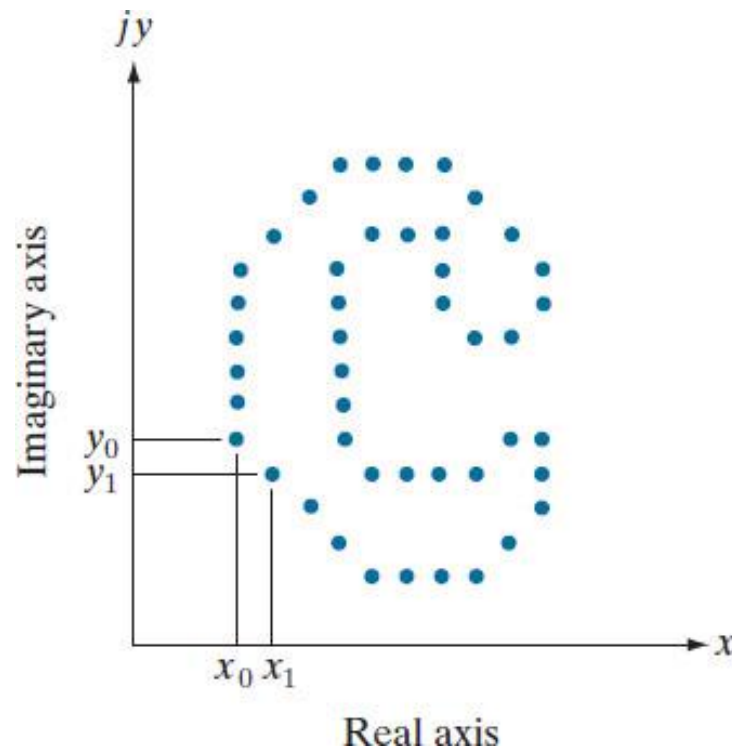
Figure 12.19

(a) Boundary of a human chromosome (2868 points). (b)–(h) Boundaries reconstructed using 1434, 286, 144, 72, 36, 18, and 8 Fourier descriptors, respectively. These numbers are approximately 50%, 10%, 5%, 2.5%, 1.25%, 0.63%, and 0.28% of 2868, respectively. Images (b)–(h) are shown as negatives to make the boundaries easier to see.
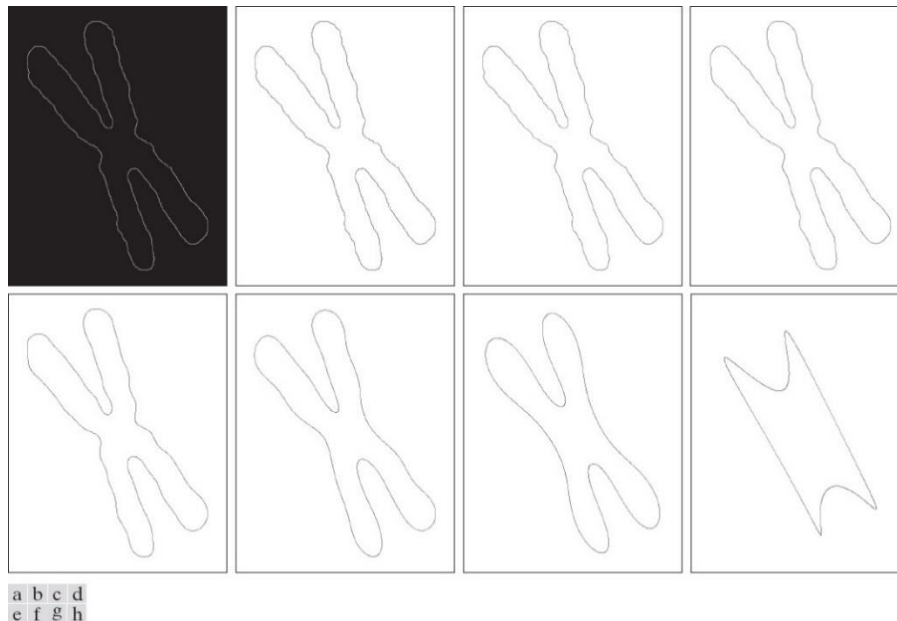

a b c d
e f g h

Table 12.1

Some basic properties of Fourier descriptors.

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

Figure 12.20

Sampled signature from Fig. 12.10(b) treated as an ordinary, discrete function of one variable.
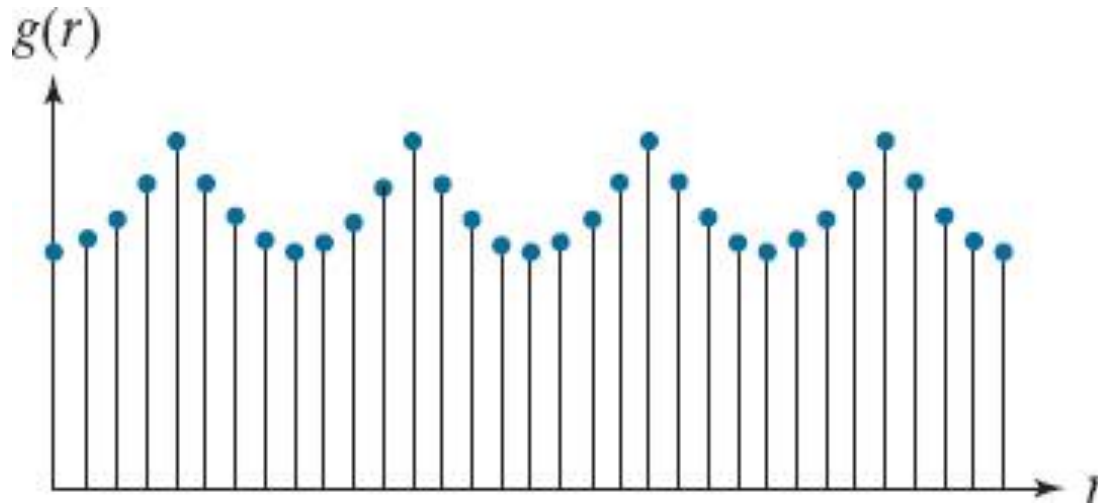
Figure 12.21

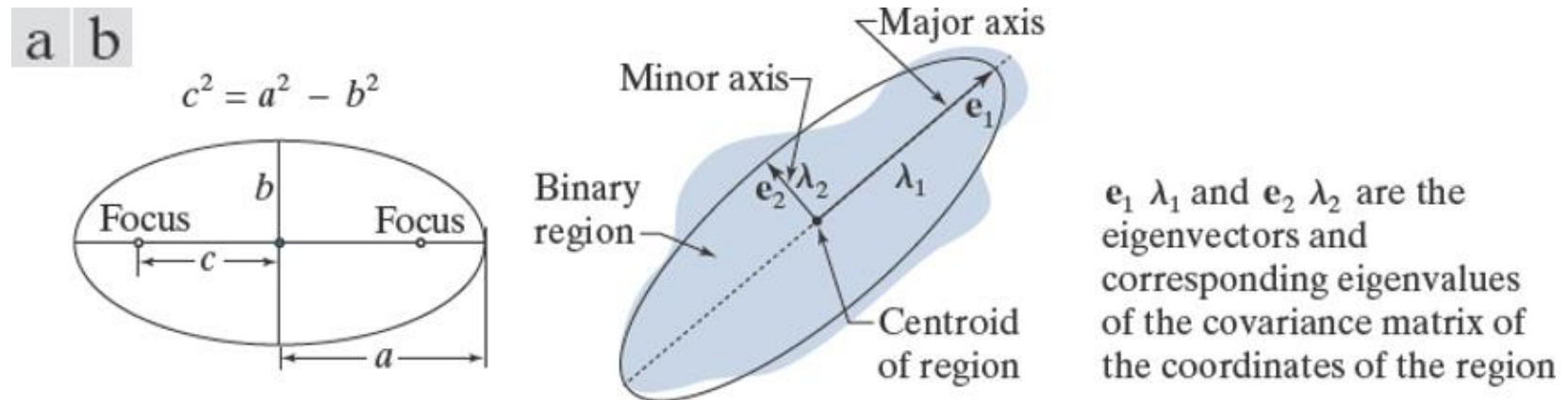(a) An ellipse in standard form. (b) An ellipse approximating a region in arbitrary orientation.

Figure 12.22

Compactness, circularity, and eccentricity of some simple binary regions.

a b c d				
Descriptor				
Compactness	10.1701	42.2442	15.9836	13.2308
Circularity	1.2356	0.2975	0.7862	0.9478
Eccentricity	0.0411	0.0636	0	0.8117

Figure 12.23

The descriptors from Fig.12.22 in 3-D feature space. Each dot shown corresponds to a feature vector whose components are the three corresponding descriptors in Fig.12.22.
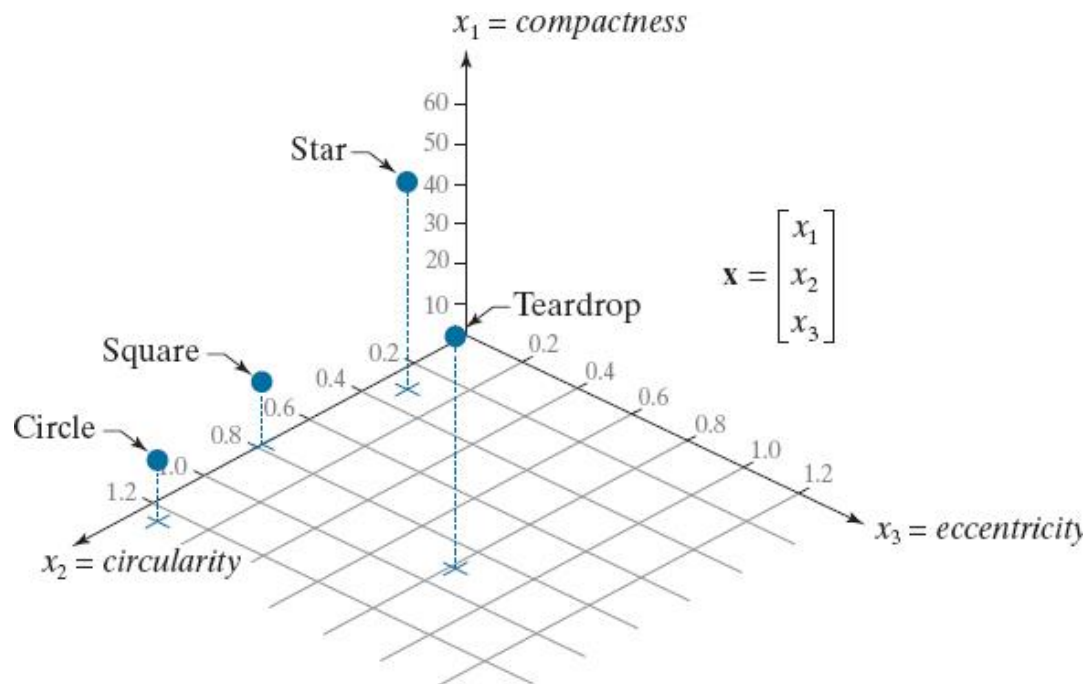
Figure 12.24

Infrared images of the Americas at night. (Courtesy of NOAA.)

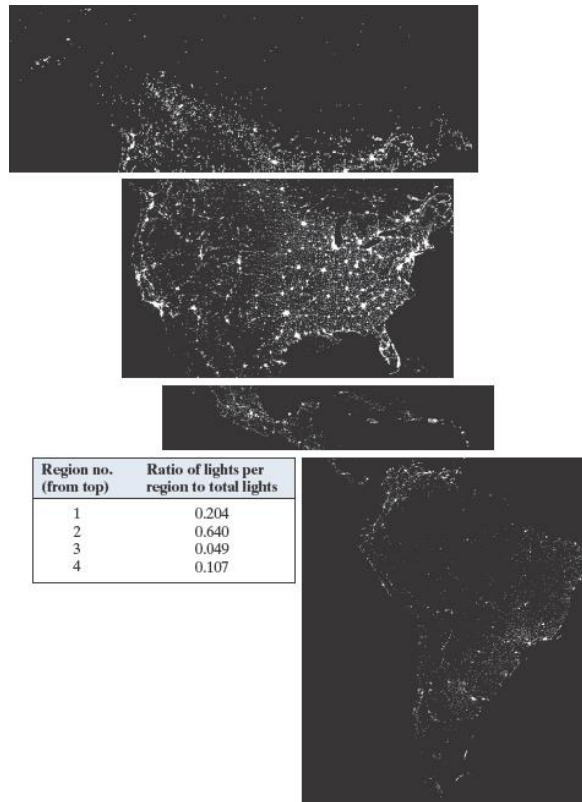| Region no. (from top)	Ratio of lights per region to total lights
1	0.204
2	0.640
3	0.049
4	0.107

Figure 12.25

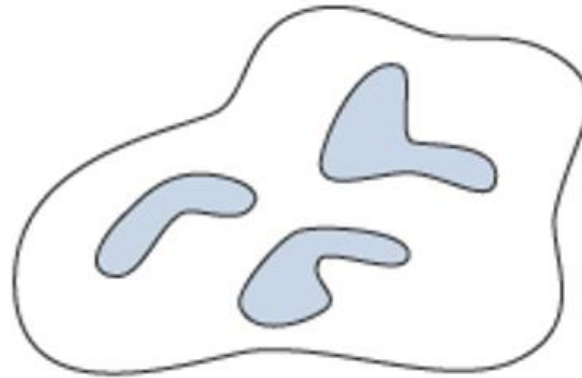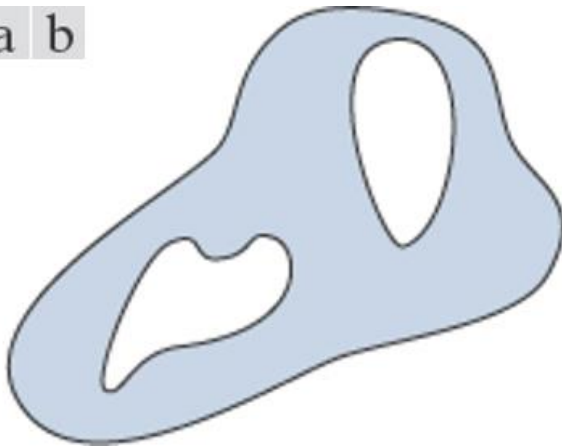(a) A region with two holes. (b) A region with three connected components.

Figure 12.26

Regions with Euler numbers equal to 0 and −1, respectively.
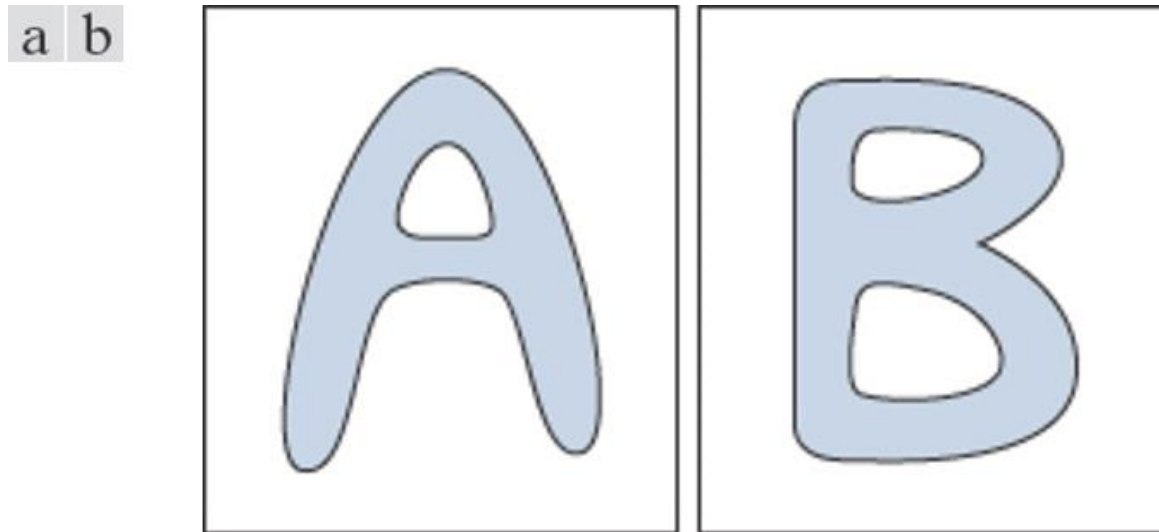
Figure 12.27

A region containing a polygonal network.
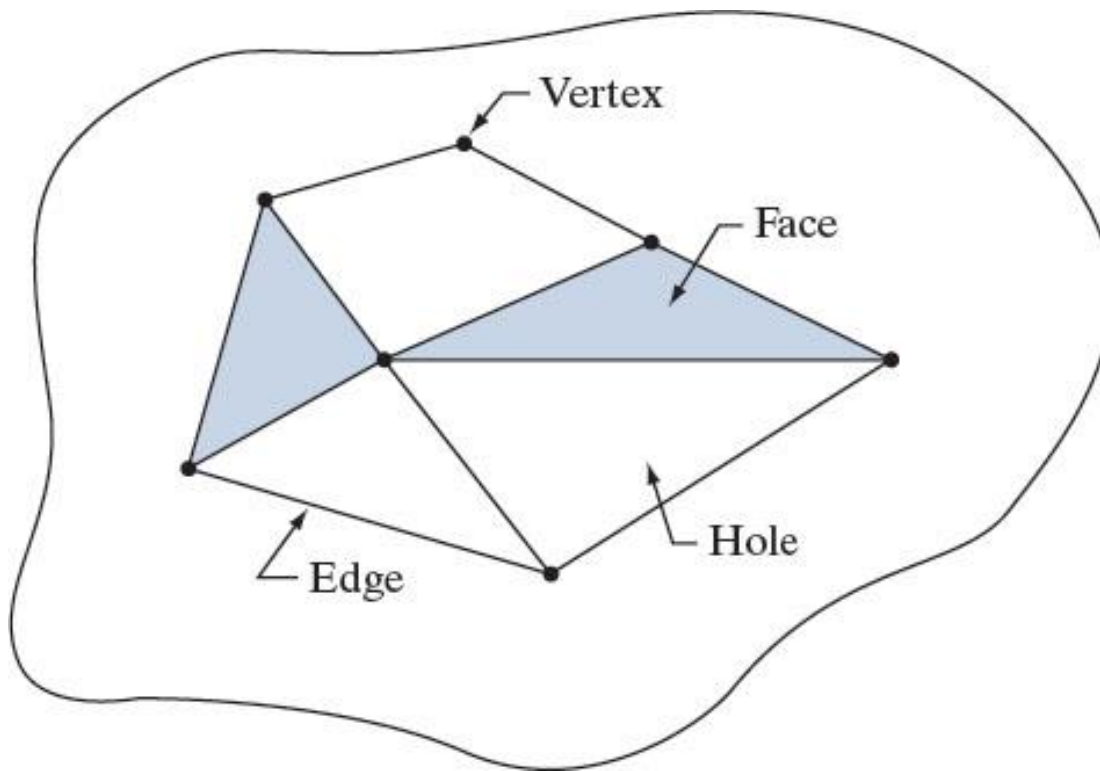
Figure 12.28

(a) Infrared image of the Washington, D.C. area. (b) Thresholded image. (c) The largest connected component of (b). (d) Skeleton of (c). (Original image courtesy of NASA.)
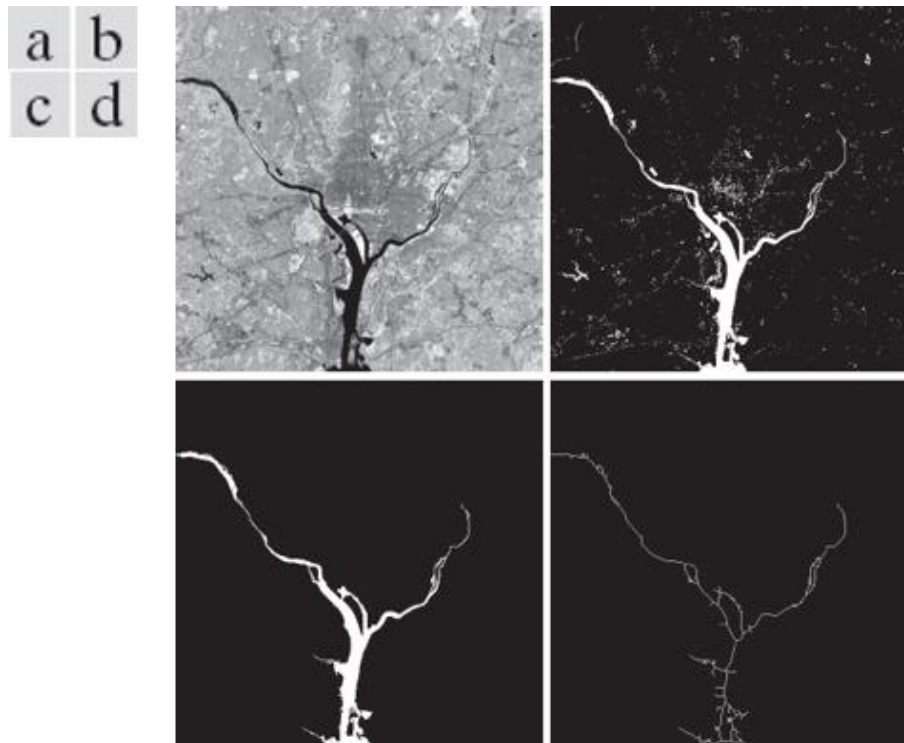
Figure 12.29

The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)
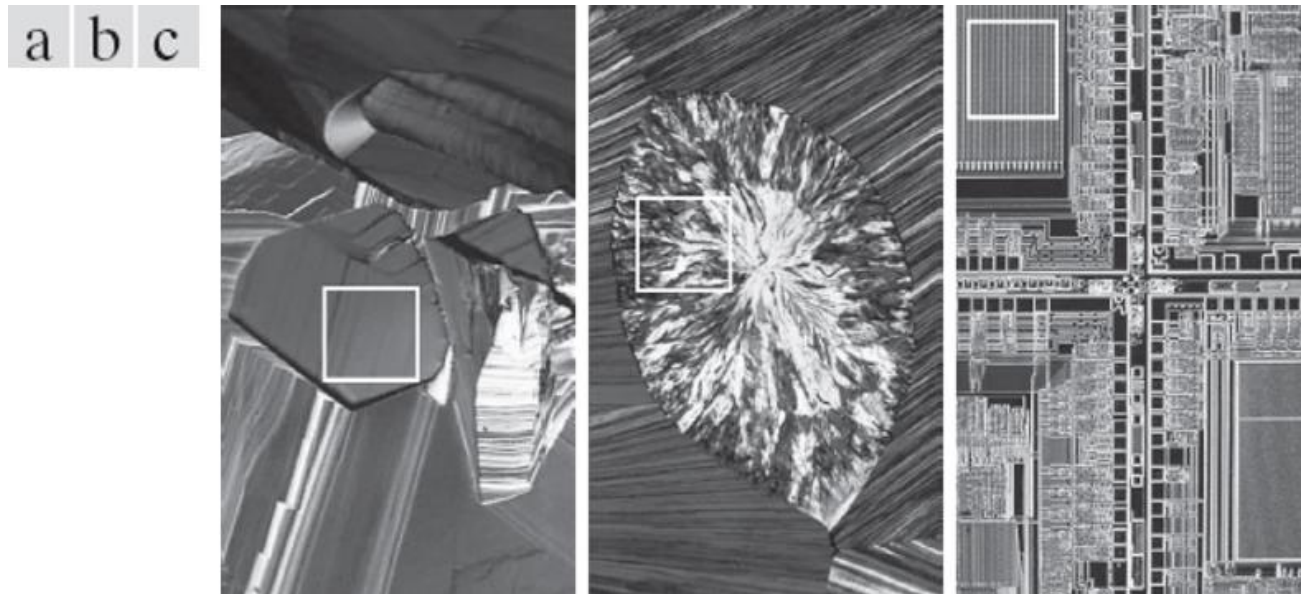
Table 12.2

Statistical texture measures for the subimages in Fig. 12.29.

Texture	Mean	Standard deviation	R (normalized)	3rd moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	-0.105	0.026	5.434
Coarse	143.56	74.63	0.079	-0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

Figure 12.30

How to construct a co-occurrence matrix.

Image f

Co-occurrence matrix **G**

Table 12.3

Descriptors used for characterizing co-occurrence matrices of size **K×K**. The term **p$_{ij}$** is the **ij-th** term of **G** divided by the sum of the elements of **G**.

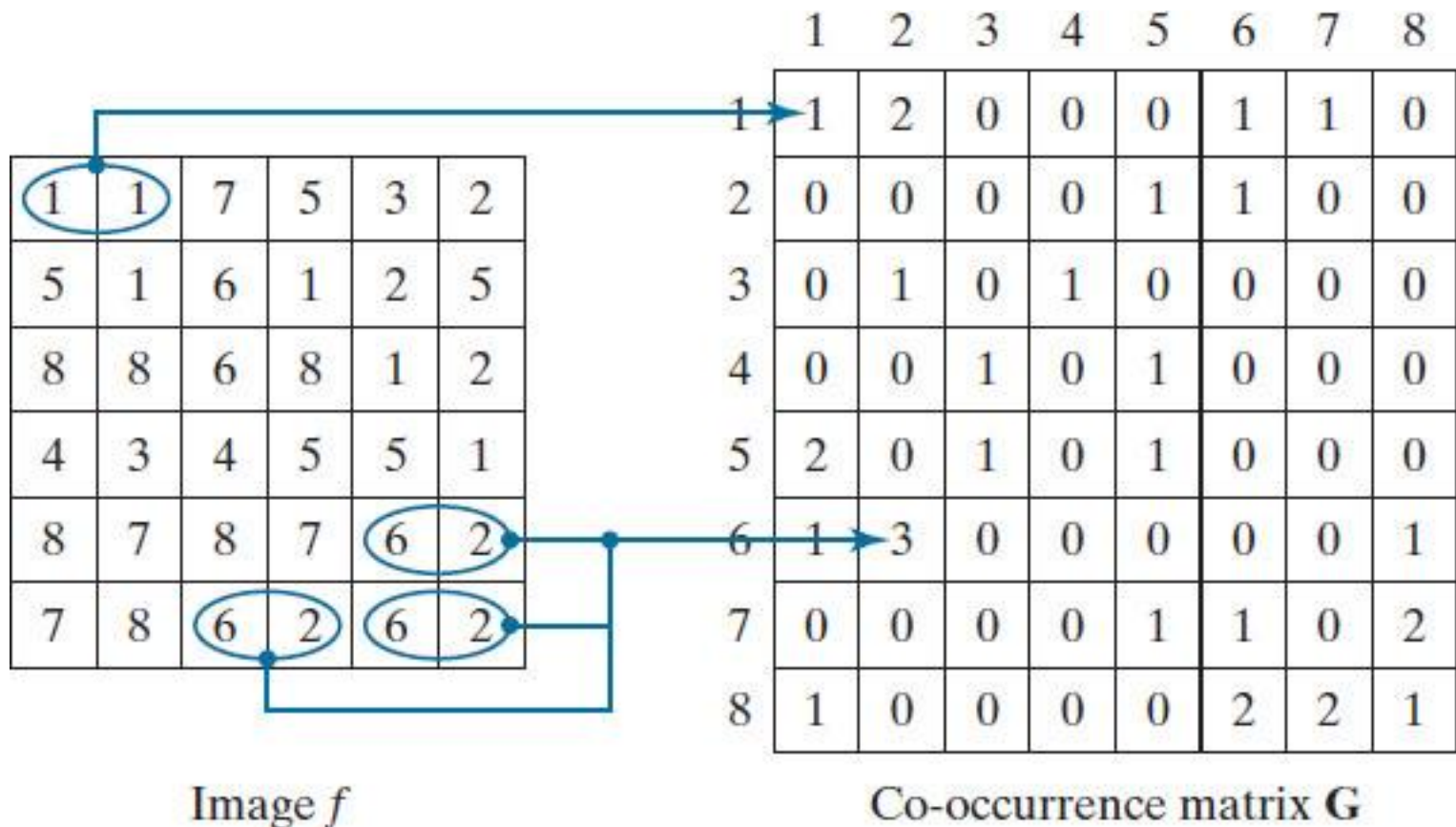| Descriptor	Explanation	Formula
Maximum probability	Measures the strongest response of **G**. The range of values is $[0, 1]$.	$\max\limits_{i,j}(p_{ij})$
Correlation	A measure of how correlated a pixel is to its neighbor over the entire image. The range of values is 1 to -1 corresponding to perfect positive and perfect negative correlations. This measure is not defined if either standard deviation is zero.	$\displaystyle\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r\,\sigma_c}$ $\sigma_r \neq 0; \sigma_c \neq 0$
Contrast	A measure of intensity contrast between a pixel and its neighbor over the entire image. The range of values is 0 (when **G** is constant) to $(K-1)^2$.	$\displaystyle\sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2\, p_{ij}$
Uniformity (also called Energy)	A measure of uniformity in the range $[0, 1]$. Uniformity is 1 for a constant image.	$\displaystyle\sum_{i=1}^{K}\sum_{j=1}^{K} p_{ij}^2$
Homogeneity	Measures the spatial closeness to the diagonal of the distribution of elements in **G**. The range of values is $[0, 1]$, with the maximum being achieved when **G** is a diagonal matrix.	$\displaystyle\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{p_{ij}}{1+|i-j|}$
Entropy	Measures the randomness of the elements of **G**. The entropy is 0 when all p_{ij}'s are 0, and is maximum when the p_{ij}'s are uniformly distributed. The maximum value is thus $2\log_2 K$.	$-\displaystyle\sum_{t=1}^{K}\sum_{j=1}^{K} p_{ij}\log_2 p_{ij}$

Figure 12.31

Images whose pixels have (a) random, (b) periodic, and (c) mixed texture patterns. Each image is of size 263×800 pixels.
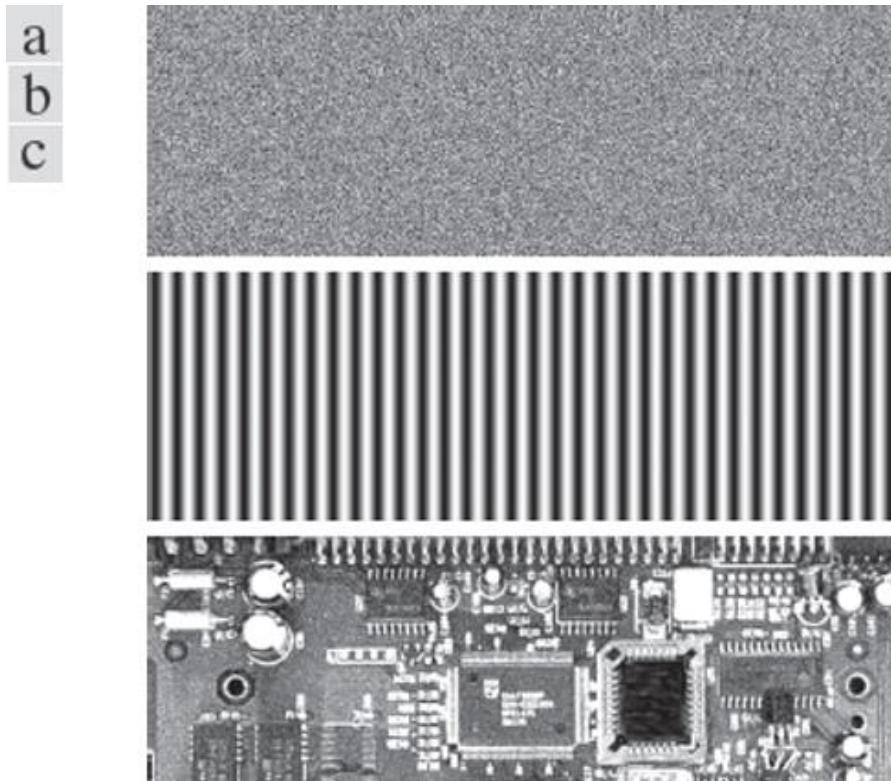
Figure 12.32

256×256 co-occurrence matrices G_1, G_2 and G_3, corresponding from left to right to the images in Fig. 12.31.
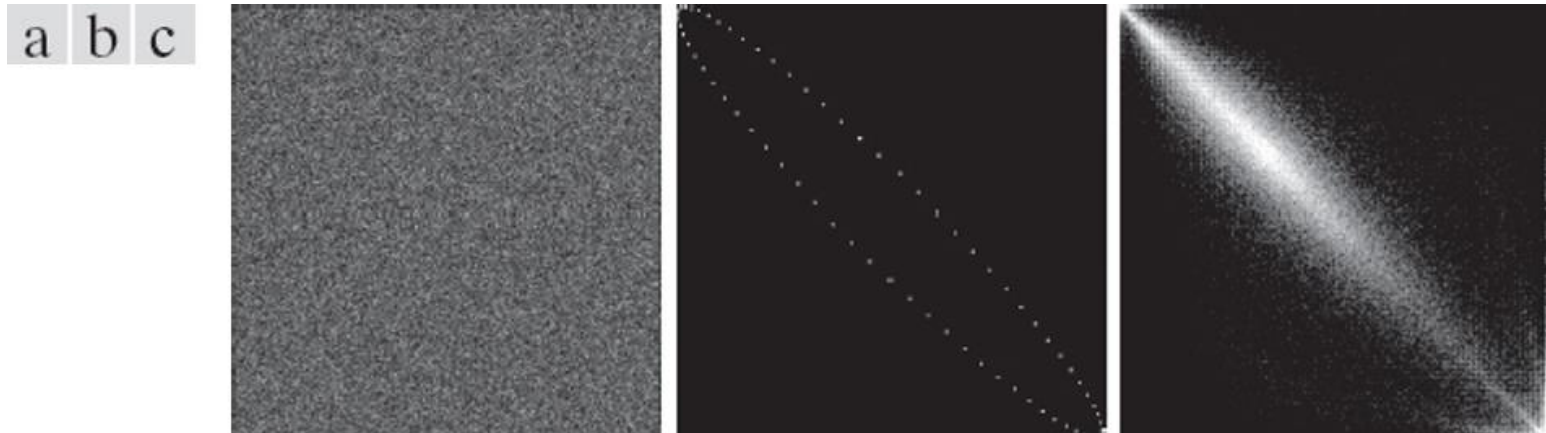

a b c

Table 12.4

Descriptors evaluated using the co-occurrence matrices displayed as images in Fig. 12.32.

Normalized Co-occurrence Matrix	Maximum Probability	Correlation	Contrast	Uniformity	Homogeneity	Entropy
G_1/n_1	0.00006	−0.0005	10838	0.00002	0.0366	15.75
G_2/n_2	0.01500	0.9650	00570	0.01230	0.0824	06.43
G_3/n_3	0.06860	0.8798	01356	0.00480	0.2048	13.58

Figure 12.33

Values of the correlation descriptor as a function of offset (distance between "adjacent" pixels) corresponding to the (a) noisy, (b) sinusoidal, and (c) circuit board images in Fig. 12.31.
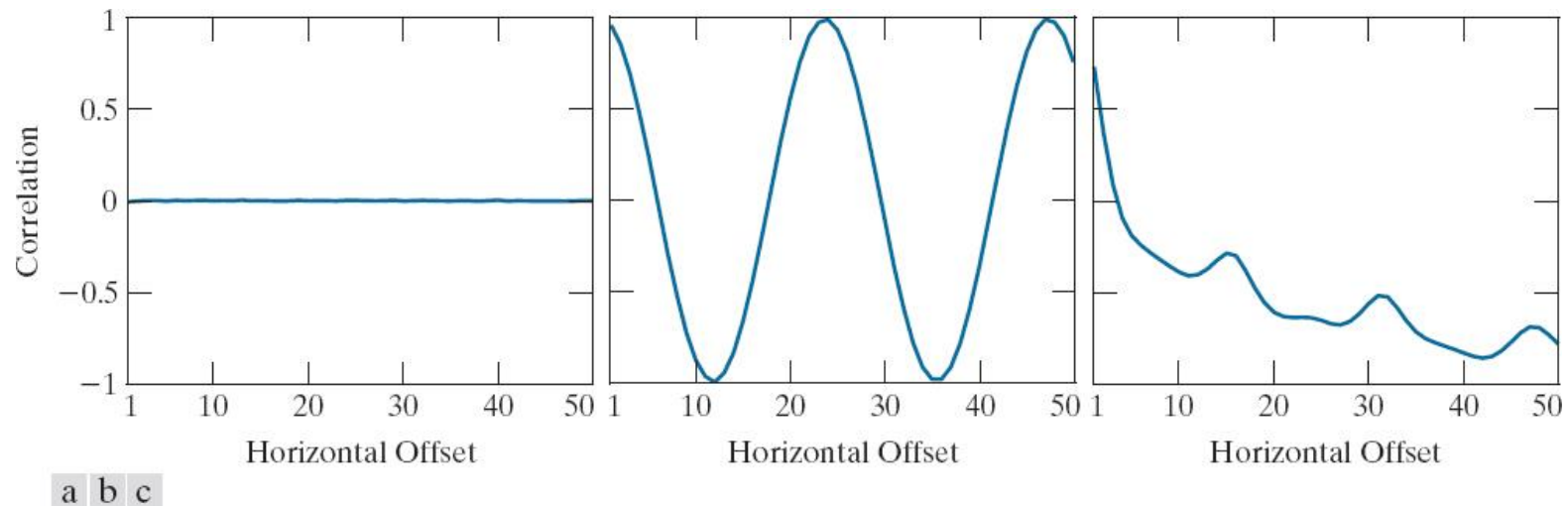
Figure 12.34

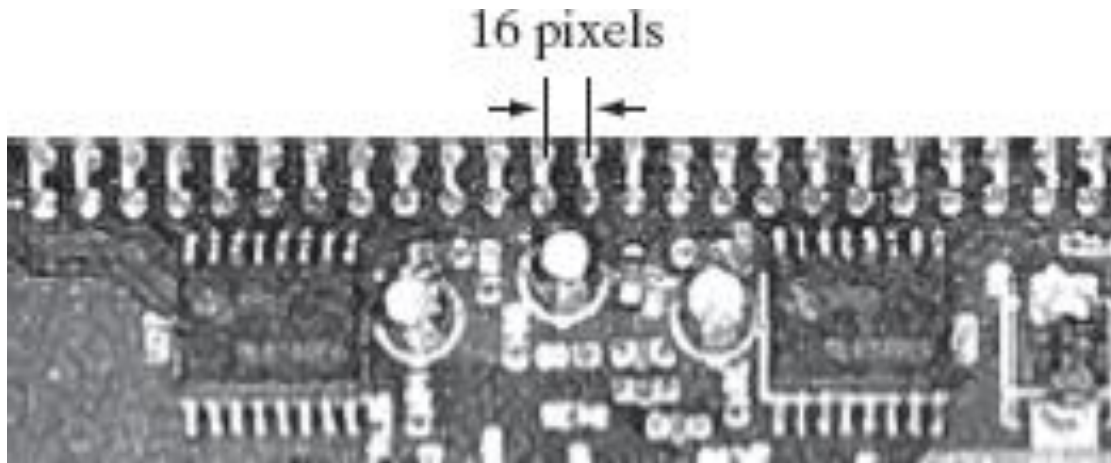A zoomed section of the circuit board image showing periodicity of components.

16 pixels

Figure 12.35

(a) and (b) Images of random and ordered objects. (c) and (d) Corresponding Fourier spectra. All images are of size 600×600 pixels.
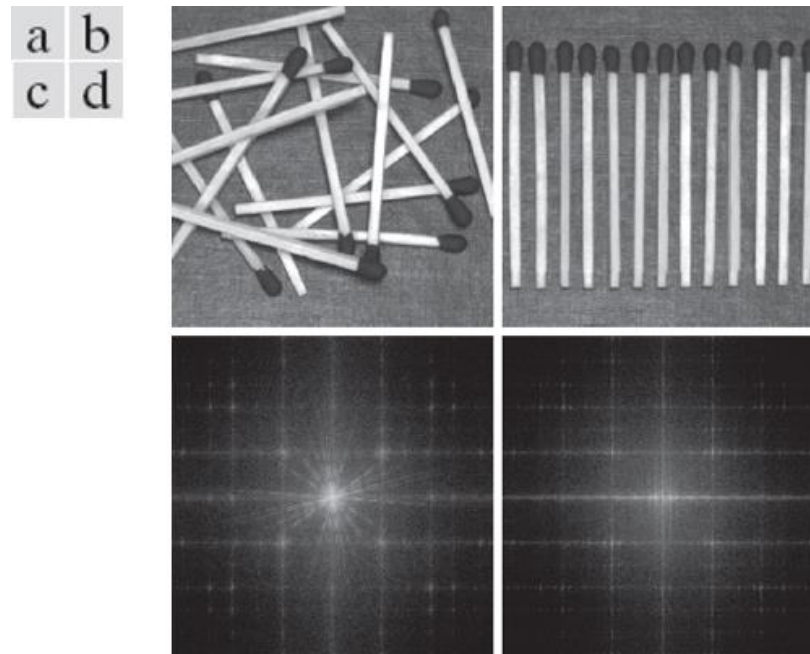
Figure 12.36

(a) and (b) Plots of $S(r)$ and $S(\theta)$ for Fig. 12.35(a).

(c) and (d) Plots of $S(r)$ and $S(\theta)$ for Fig. 12.35(b).
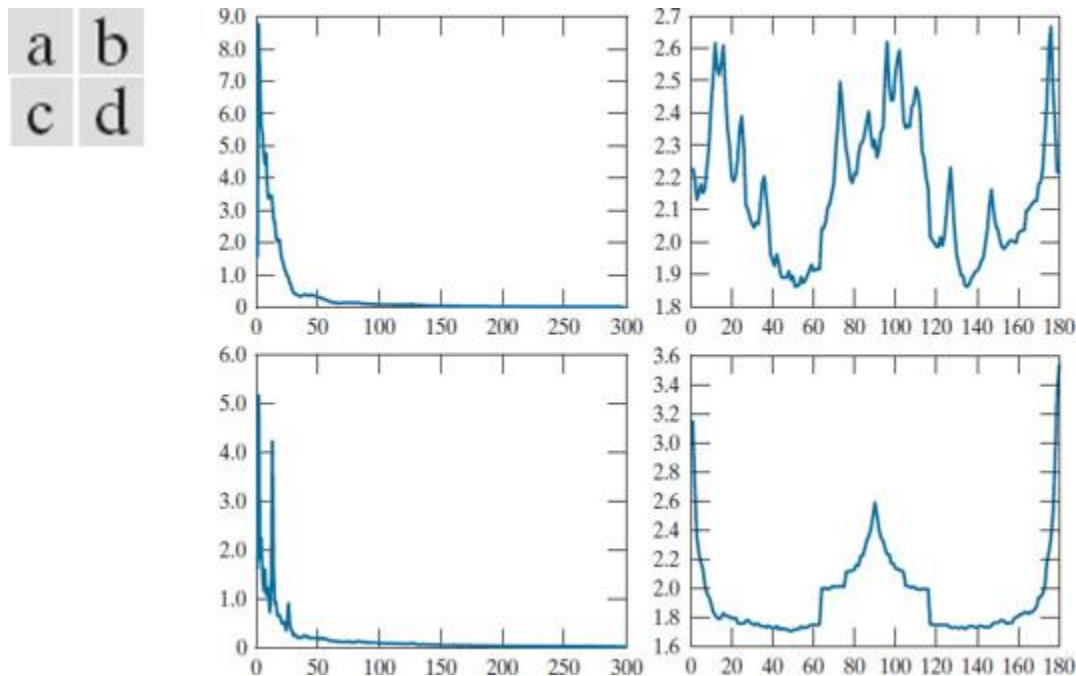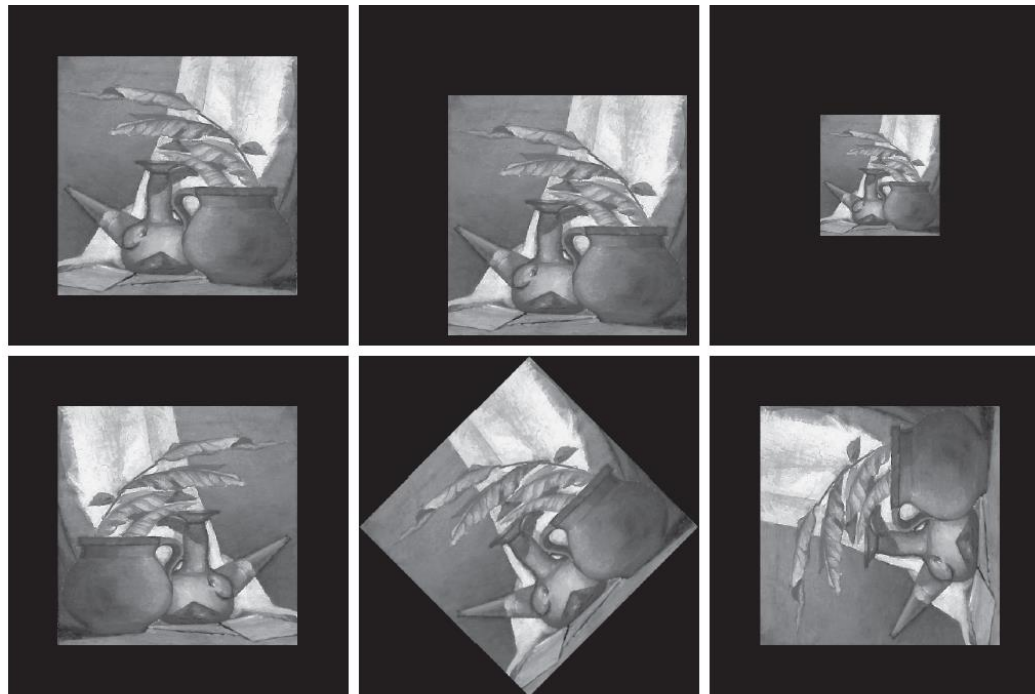
All vertical axes are $\times 10^5$.

Figure 12.37

(a) Original image. (b)–(f) Images translated, scaled by one-half, mirrored, rotated by 45 degree, and rotated by 90 degree , respectively.

a b c
d e f

Table 12.5

Moment invariants for the images in Fig. 12.37.

Moment Invariant	Original Image	Translated	Half Size	Mirrored	Rotated 45°	Rotated 90°
ϕ_1	2.8662	2.8662	2.8664	2.8662	2.8661	2.8662
ϕ_2	7.1265	7.1265	7.1257	7.1265	7.1266	7.1265
ϕ_3	10.4109	10.4109	10.4047	10.4109	10.4115	10.4109
ϕ_4	10.3742	10.3742	10.3719	10.3742	10.3742	10.3742
ϕ_5	21.3674	21.3674	21.3924	21.3674	21.3663	21.3674
ϕ_6	13.9417	13.9417	13.9383	13.9417	13.9417	13.9417
ϕ_7	−20.7809	−20.7809	−20.7724	20.7809	−20.7813	−20.7809

Figure 12.38

Multispectral images in the (a) visible blue, (b) visible green, (c) visible red, (d) near infrared, (e) middle infrared, and (f) thermal infrared bands. (Images courtesy of NASA.)
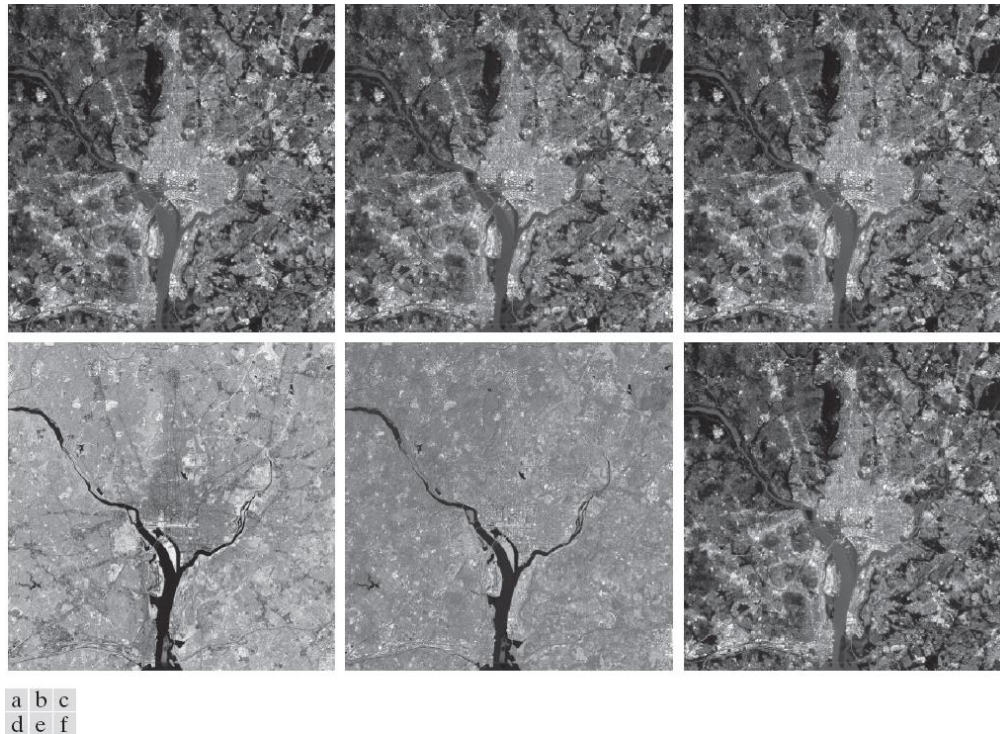

a b c
d e f

Figure 12.39

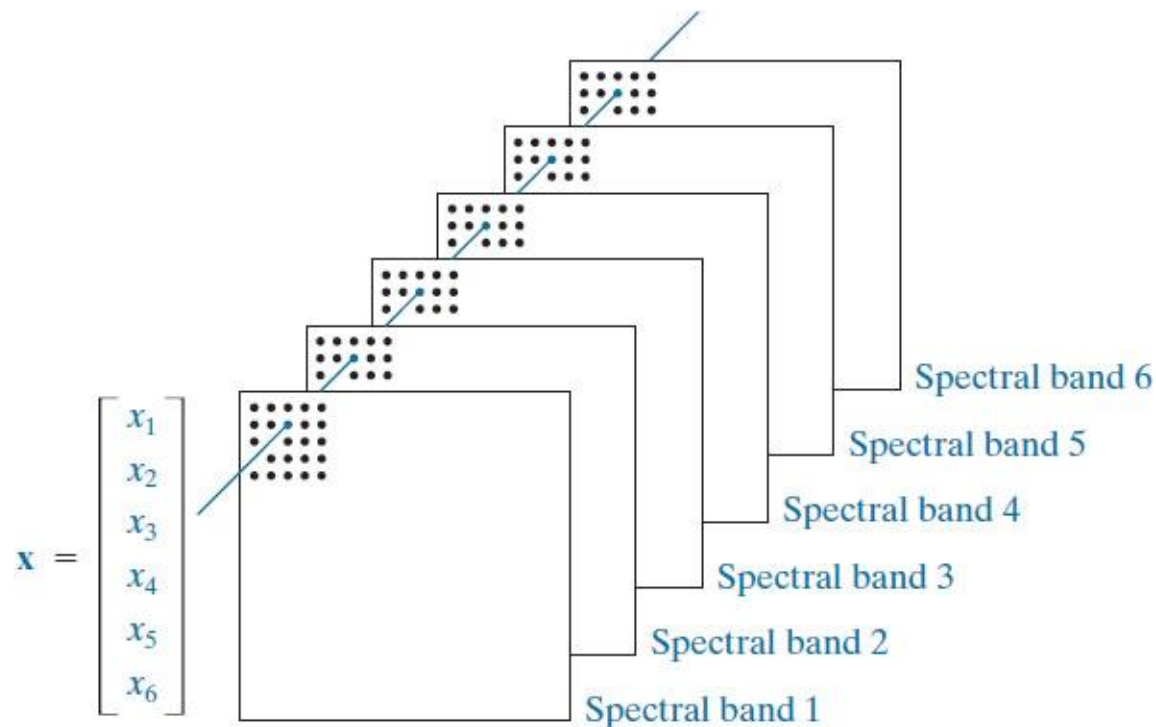Forming of a feature vector from corresponding pixels in six images.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

Spectral band 6
Spectral band 5
Spectral band 4
Spectral band 3
Spectral band 2
Spectral band 1

Table 12.6

Eigenvalues of C_x obtained from the images in Fig. 12.38.

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
10344	2966	1401	203	94	31

Figure 12.40

The six principal component images obtained from vectors computed using Eq. (12-49). Vectors are converted to images by applying Fig. 12.39 in reverse.
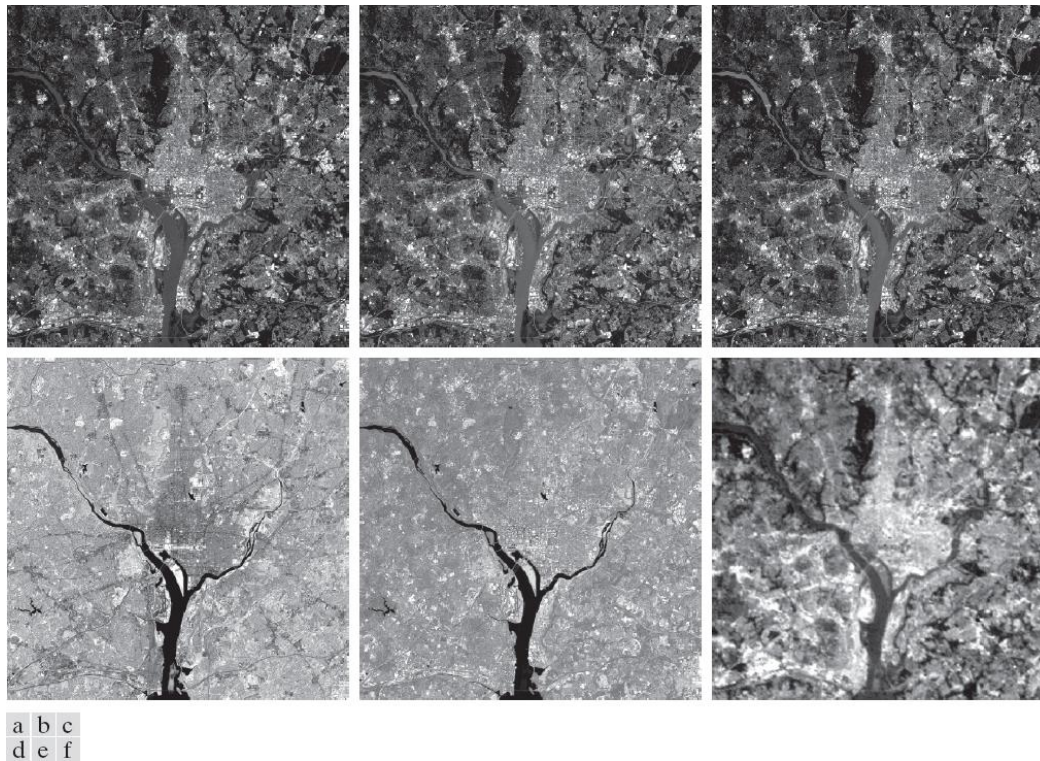

a b c
d e f

Figure 12.41

Multispectral images reconstructed using only the two principal component images corresponding to the two principal component vectors with the largest eigenvalues. Compare these images with the originals in Fig. 12.38.
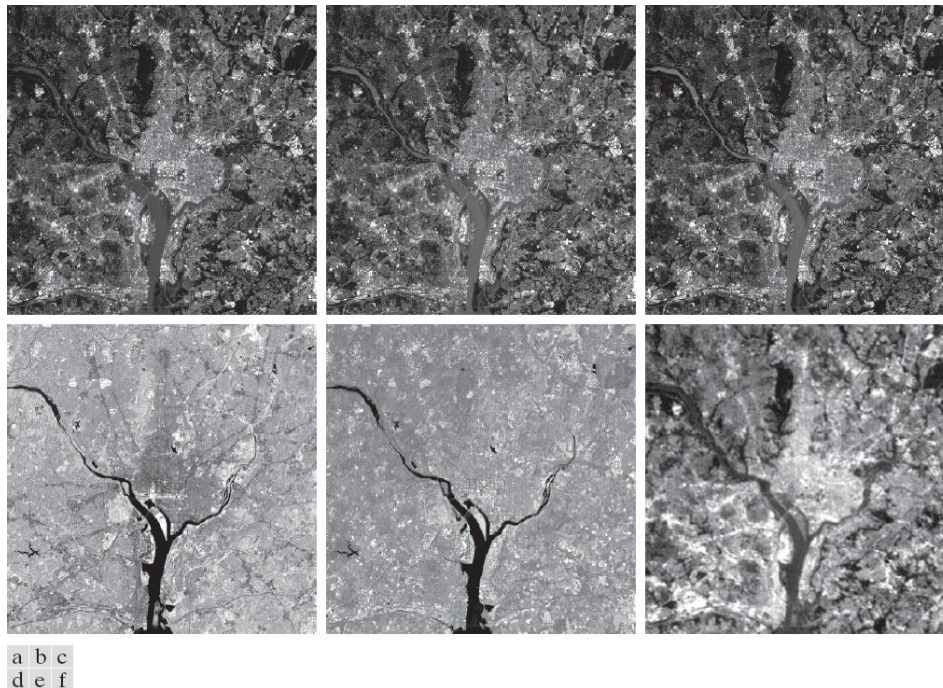

a b c
d e f

Figure 12.42

Differences between the original and reconstructed images. All images were enhanced by scaling them to the full [0, 255] range to facilitate visual analysis.
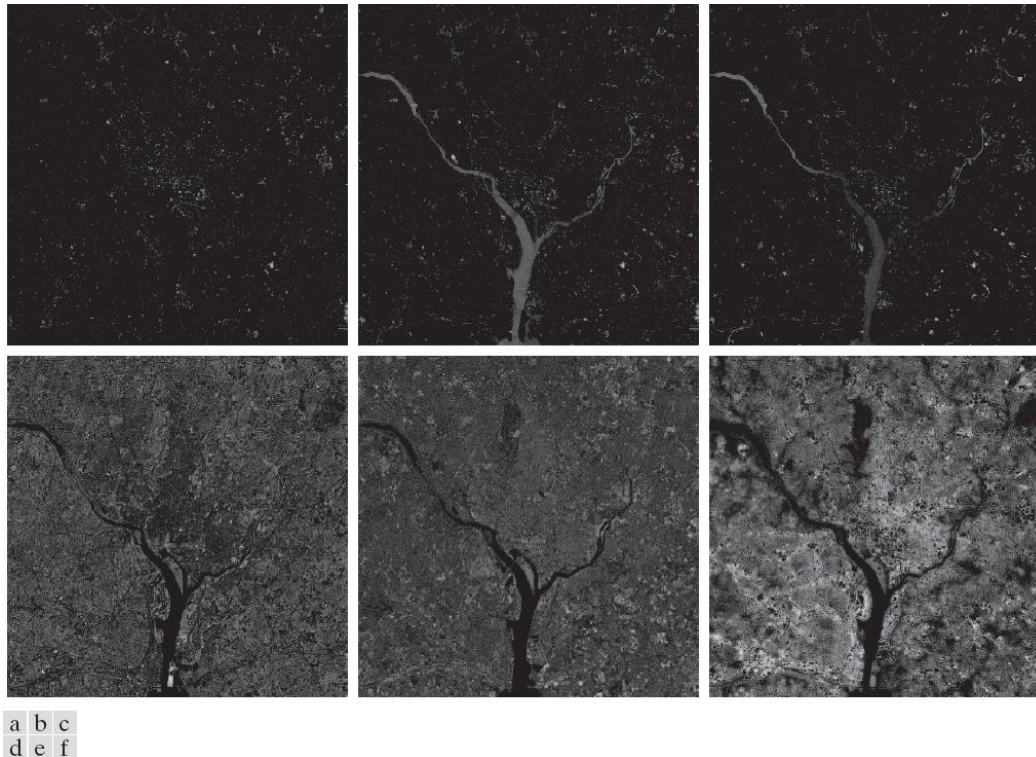

a b c
d e f

Figure 12.43

(a) An object. (b) Object showing eigenvectors of its covariance matrix. (c) Transformed object, obtained using Eq. (12-49). (d) Object translated so that all its coordinate values are greater than 0.
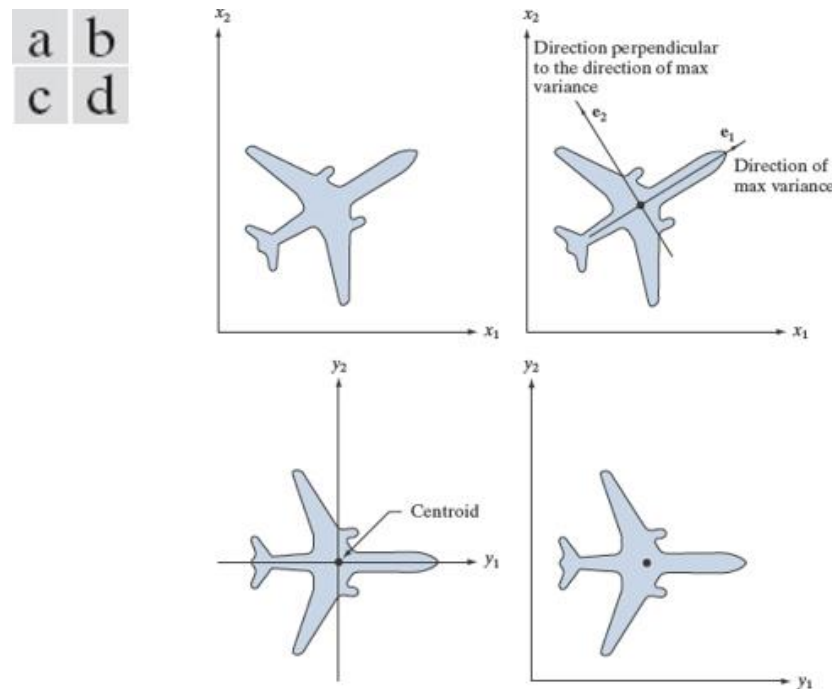
Figure 12.44

A manual example. (a) Original points. (b) Eigenvectors of the covariance matrix of the points in (a). (c) Transformed points obtained using Eq. (12-49). (d) Points from (c), rounded and translated so that all coordinate values are integers greater than 0. The dashed lines are included to facilitate viewing. They are not part of the data.
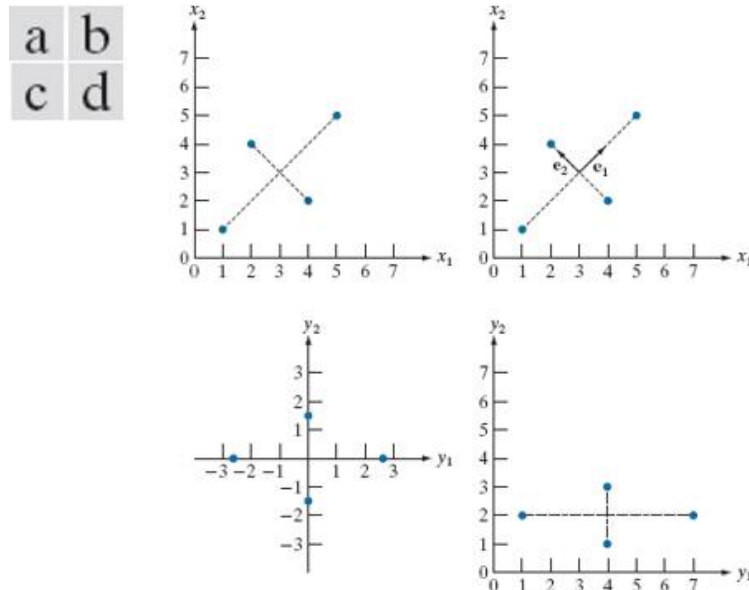
Figure 12.45

Illustration of how the Harris-Stephens corner detector operates in the three types of subregions indicated by A (flat), B (edge), and C (corner). The wiggly arrows indicate graphically a directional response in the detector as it moves in the three areas shown.
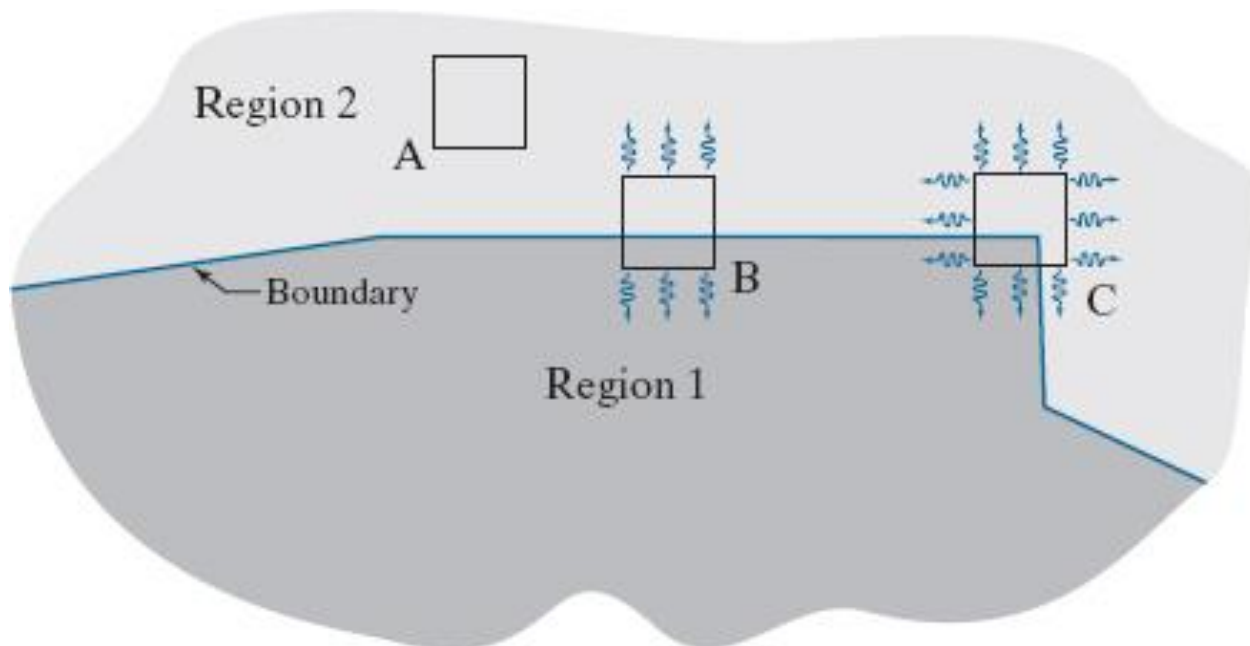
Figure 12.46

(a)–(c) Noisy images and image patches (small squares) encompassing image regions similar in content to those in Fig. 12.45. (d)–(f) Plots of value pairs (f_x, f_y) showing the characteristics of the eigenvalues of **M** that are useful for detecting the presence of a corner in an image patch.
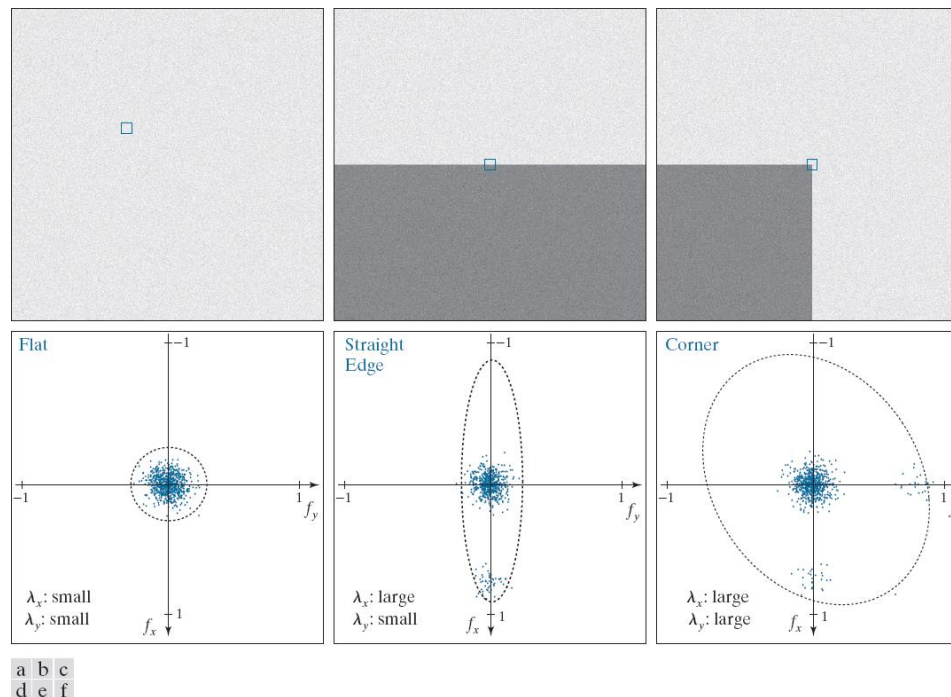
Figure 12.47

(a) A 600×600 image with values in the range [0,1], corrupted by additive Gaussian noise with 0 mean and variance of 0.006. (b) Result of applying the HS corner detector with **k** = 0.04 and **T** = 0.01 (the defaults). Several errors are visible. (c) Result using **k** = 0.1 and **T** = 0.01. (d) Result using **k** = 0.1 and **T** = 0.1. (e) Result using **k** = 0.04 and **T** = 0.1. (f) Result using **k** = 0.04 and **T** = 0.3 (only the strongest corners on the left were detected).

a b c
d e f

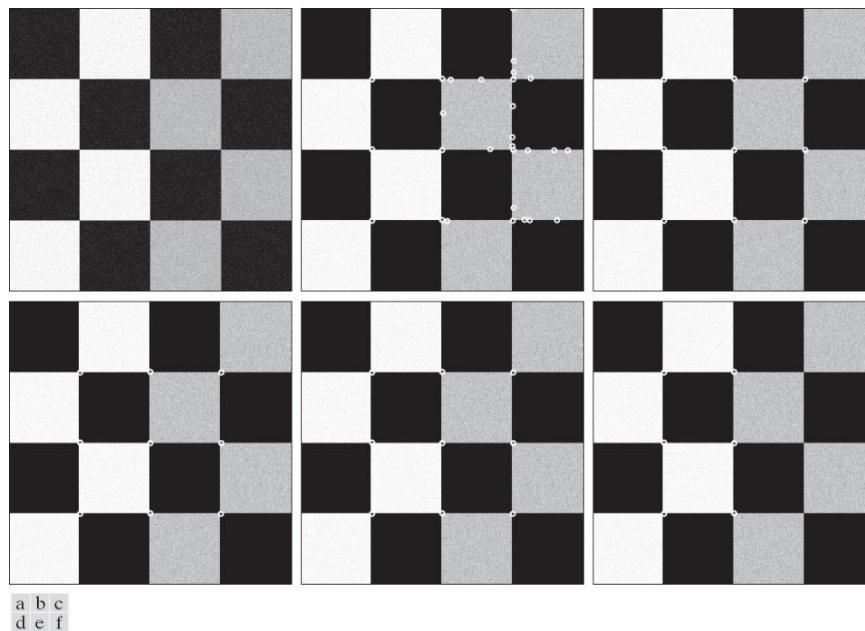# Figure 12.48

(a) Same as Fig. 12.47(a), but corrupted with Gaussian noise of mean 0 and variance 0.01. (b) Result of using the HS detector with **k** = 0.04 and **T** = 0.01 [compare with Fig. 12.47(b)]. (c) Result with **k** = 0.249, (near the highest value in our implementation), and **T** = 0.01. (d) Result of using **k** = 0.04 and **T** = 0.15.
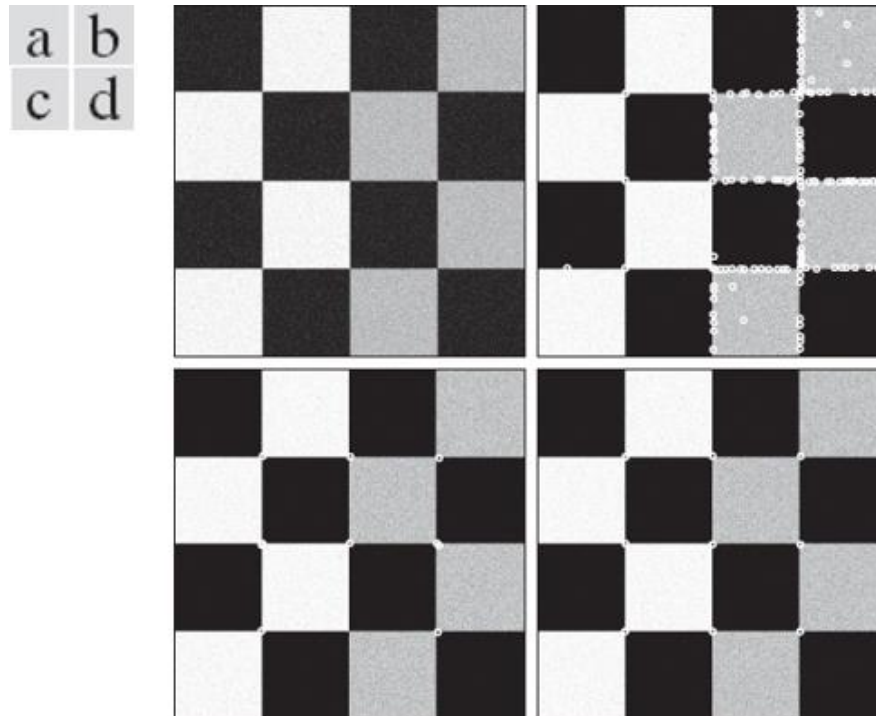
Figure 12.49

600×600 image of a building. (b) Result of applying the HS corner detector with **k** = 0.04 and **T** = 0.01 (the default values in our implementation). Numerous irrelevant corners were detected. (c) Result using **k** = 0.249 and the default value for **T**. (d) Result using **k** = 0.17 and **T** = 0.05. (e) Result using the default value for **k** and **T** = 0.05. (f) Result using the default value of **k** and **T** = 0.07.

a b c
d e f

Figure 12.50

(a) Image rotated 5 degree. (b) Corners detected using the parameters used to obtain Fig. 12.49(f).

a b

Pearson

Figure 12.51

Detecting MSERs. Top: Grayscale image. Left: Thresholded images using **T** = 10 and $\Delta T = 50$. Right: Component tree, showing the individual regions. Only one MSER was detected (see dashed tree node on the rightmost branch of the tree). Each level of the tree is formed from the thresholded image on the left, at that same level. Each node of the tree contains one extremal region (connected component) shown in **white**, and denoted by a subscripted **R**
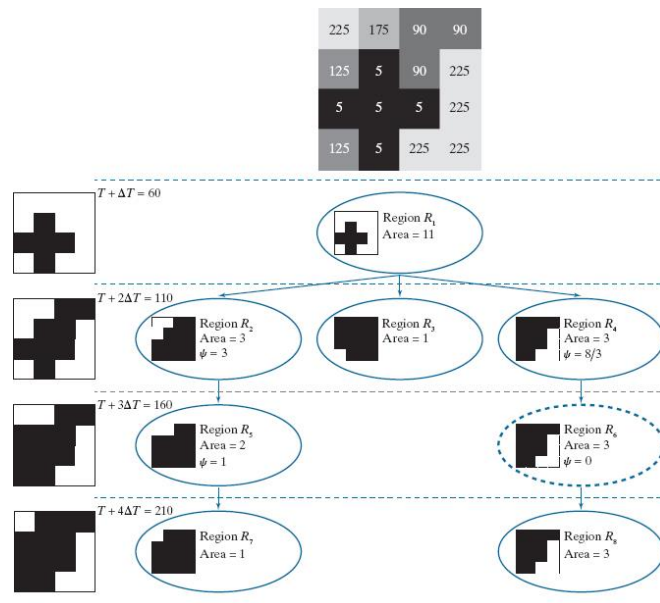
Figure 12.52

(a) 600×570 CT slice of a human head. (b) Image smoothed with a box kernel of
of size 15×15 elements. (c) A extremal region along the path of the tree
containing one MSER. (d) The MSER. (All MSER regions were limited to the
range 10,260 – 34,200 pixels, corresponding to a range between 3% and 10% of
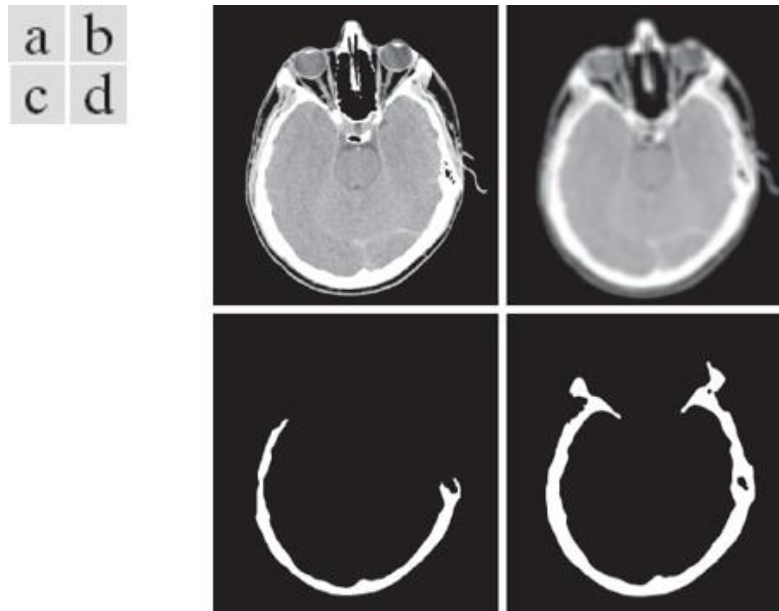image size.) (Original image courtesy of Dr. David R. Pickens, Vanderbilt
University.)

Figure 12.53

(a) Building image of size 600×600 pixels. (b) Image smoothed using a 5×5 box kernel MSERs detected using $T = 0$, $\Delta T = 10$, and MSER size range between 10,000 and 30,000 pixels, corresponding approximately to 3% and 8% of the area of the image. (e) Composite image.

a b
c d e

Figure 12.54

(a) Building image rotated 5 degree counterclockwise. (b) Smoothed image using the same kernel as in Fig. 12.53(b). (c) Composite MSER detected using the same parameters we used to obtain Fig. 12.53(e). The MSERs of the original and rotated images are almost identical.

a b c

Figure 12.55

(a) Building image reduced to half-size. (b) Image smoothed with a 3×3 box kernel. (c) Composite MSER obtained with the same parameters as Fig. 12.53(e), but using a valid MSER region size range of 2,500 - –7,500 pixels.

a b c

Figure 12.56

Scale space, showing three octaves. Because $s = 2$ in this case, each octave has five smoothed images. A Gaussian kernel was used for smoothing, so the space parameter is σ.
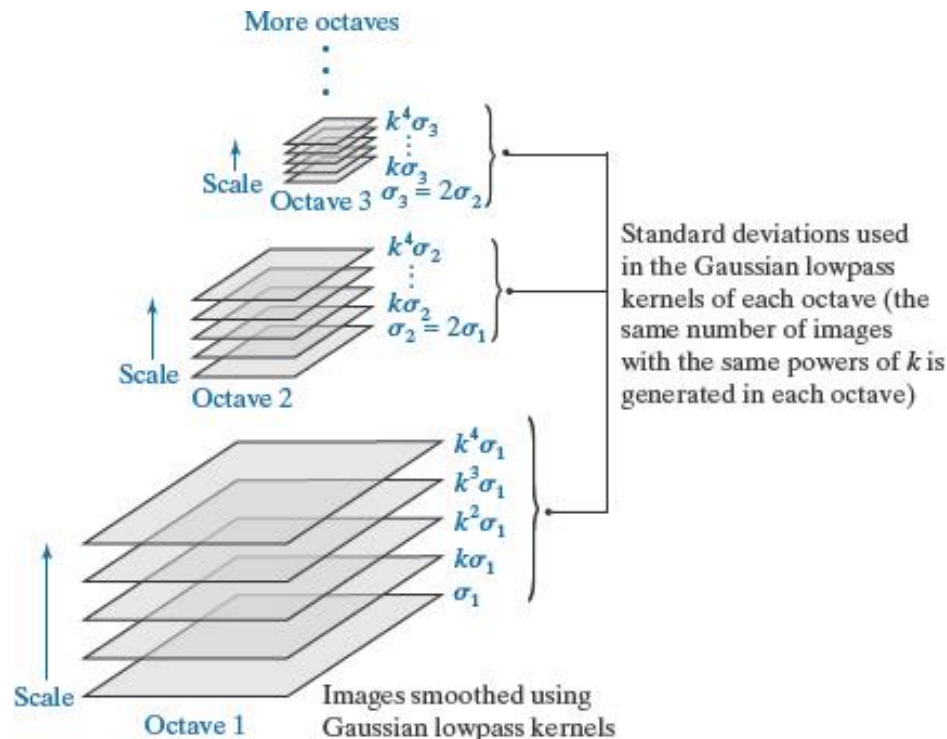
Figure 12.57

Illustration using images of the first three octaves of scale space in SIFT. The entries in the table are values of standard deviation used at each scale of each octave. For example the standard deviation used in scale 2 of octave 1 is $K\sigma_1$, which is equal 1.0. (The images of octave 1 are shown slightly overlapped to fit in the figure space.)
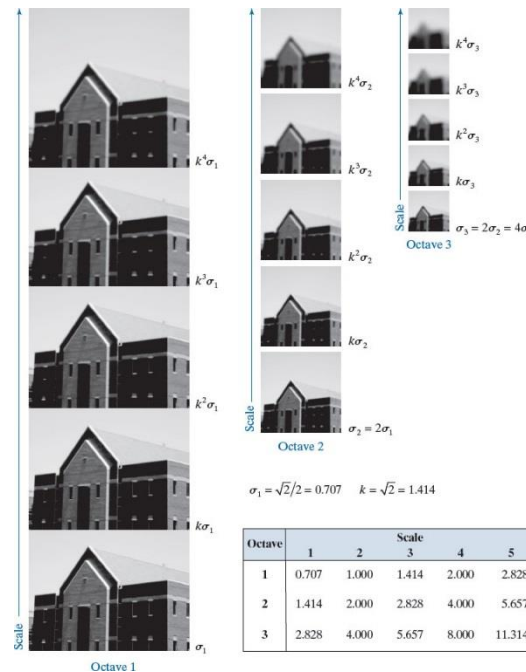
Figure 12.58

How Eq. (12-69) is implemented in scale space. There are $s+3$ $L(x, y, \sigma)$ images and $s + 2$ corresponding $D(x, y, \sigma)$ images in each octave.

Octave 3

Octave 2

Octave 1

Scale

Gaussian-filtered images, $L(x, y, \sigma)$
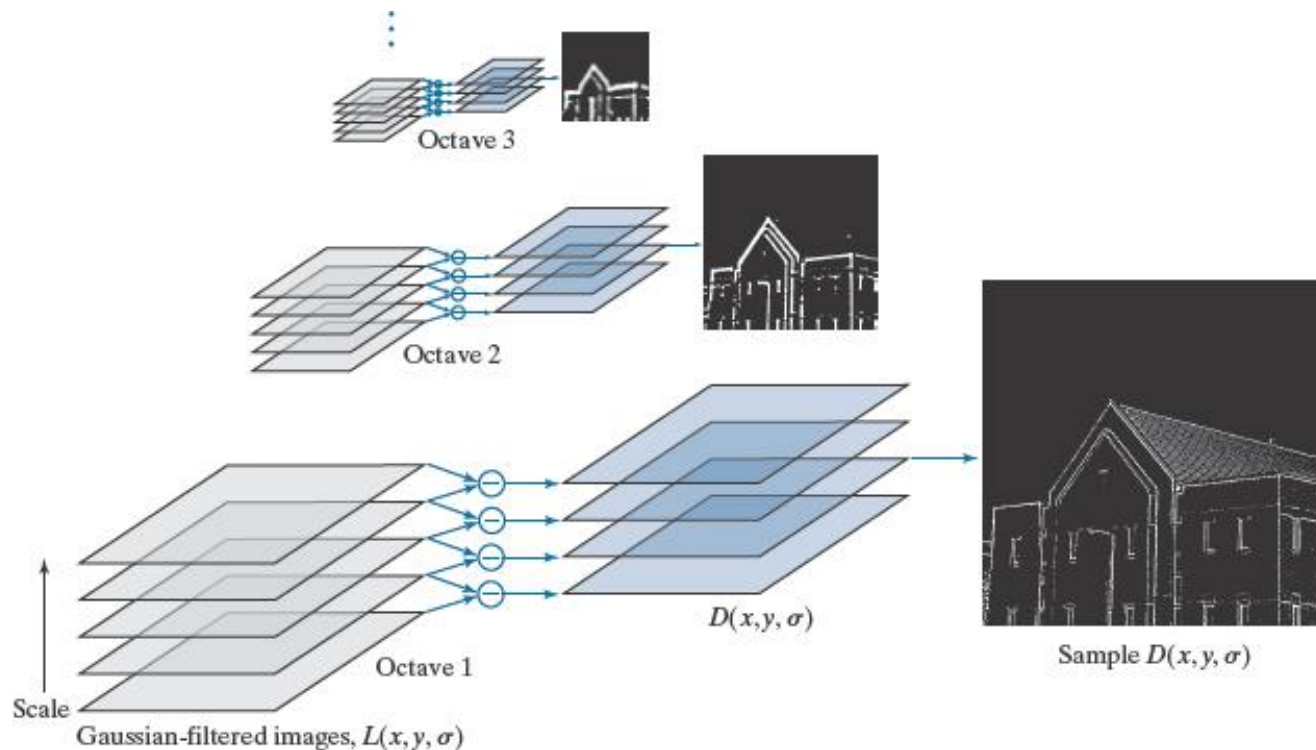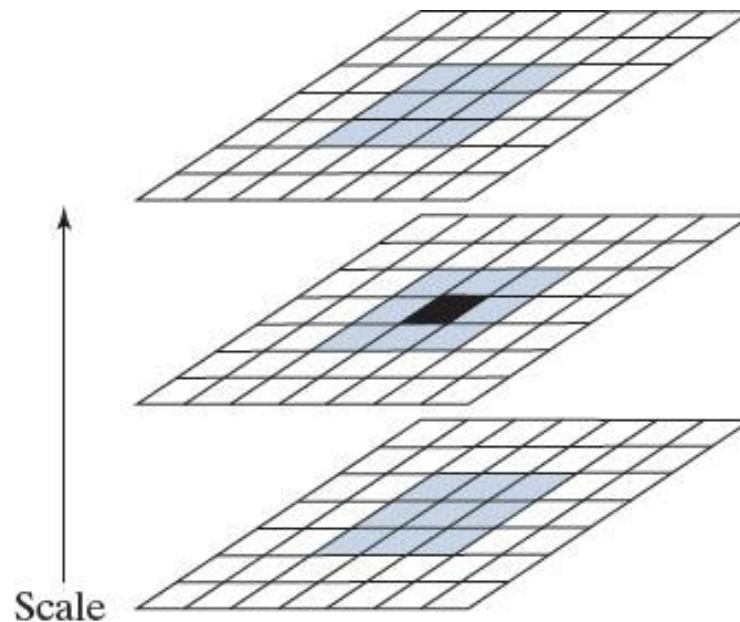
$D(x, y, \sigma)$

Sample $D(x, y, \sigma)$

Figure 12.59

Extrema (maxima or minima) of the $D(x, y, \sigma)$ images in an octave are detected by comparing a pixel (shown in black) to its 26 neighbors (shown shaded) in 3×3 regions at the current and adjacent scale images.

Scale

Corresponding sections of three contiguous $D(x, y, \sigma)$ images

Figure 12.60

SIFT keypoints detected in the building image. The points were enlarged slightly to make them easier to see.

Figure 12.61

The keypoints from Fig. 12.60 superimposed on the original image. The arrows indicate keypoint orientations.

Figure 12.62

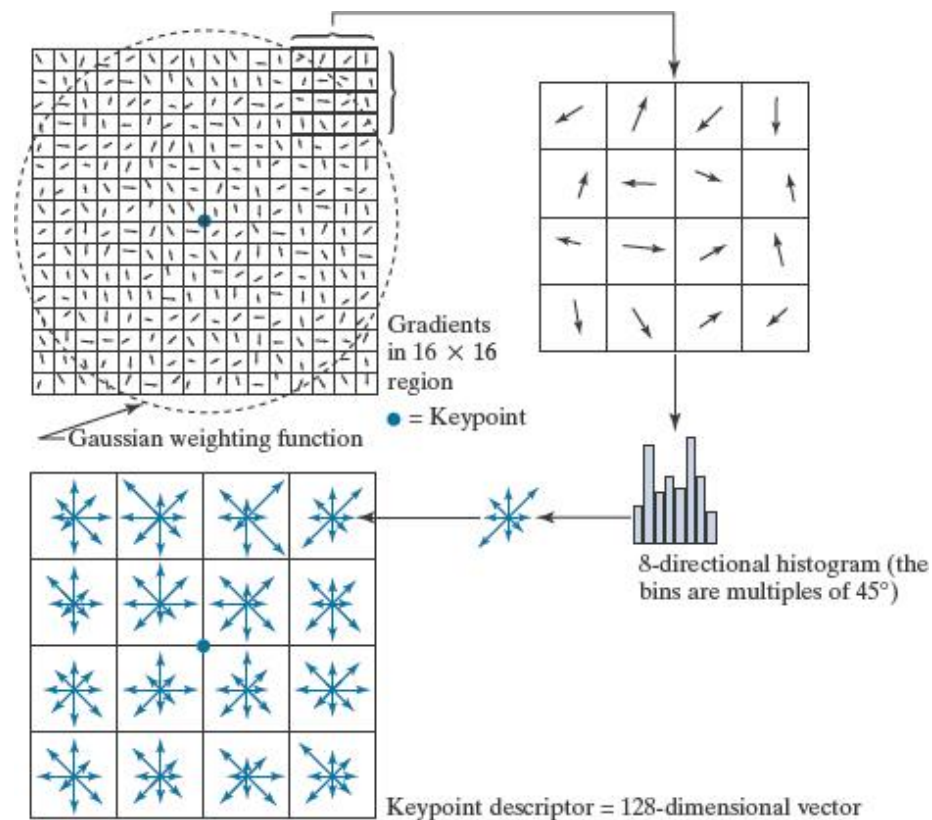Approach used to compute a keypoint descriptor.

Gradients in 16 × 16 region
● = Keypoint
Gaussian weighting function

8-directional histogram (the bins are multiples of 45°)

Keypoint descriptor = 128-dimensional vector

P Pearson

Figure 12.63

(a) Keypoints and their directions (shown as gray arrows) for the building image and for a section of the right corner of the building. The subimage is a separate image and was processed as such. (b) Corresponding key points between the building and the subimage (the straight lines shown connect pairs of matching points). Only three of the 36 matches found are incorrect.


a b

Pearson
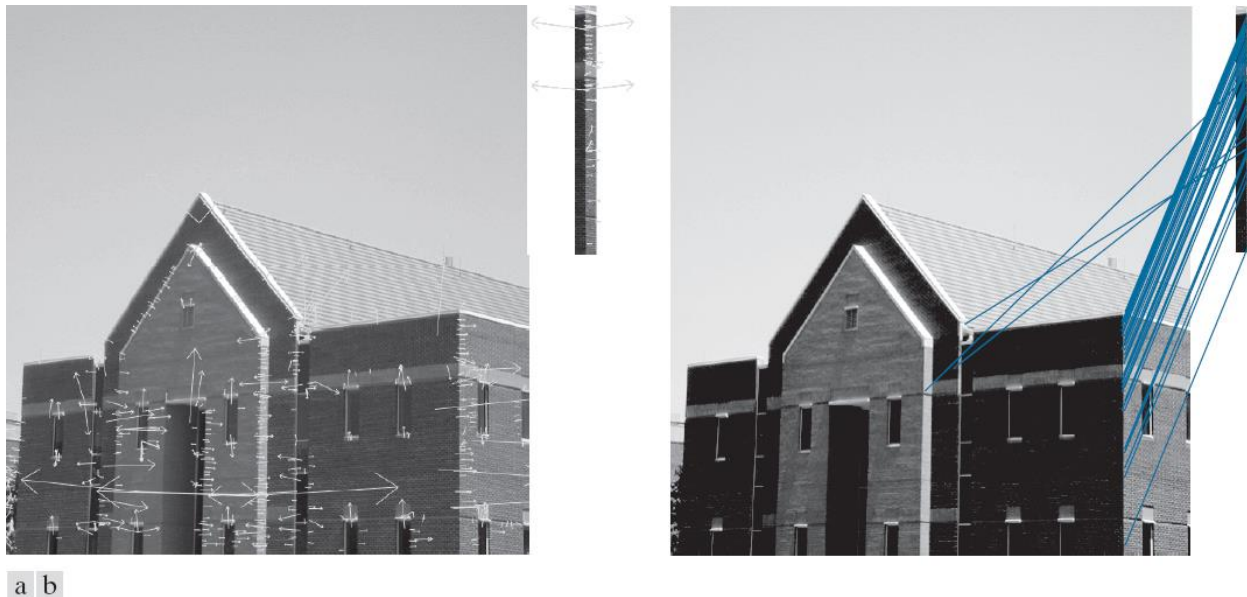
Figure 12.64

(a) Keypoints for the rotated (by 5 degree) building image and for a section of the right corner of the building. The subimage is a separate image and was processed as such. (b) Corresponding keypoints between the corner and the building. Of the 26 matches found, only two are in error.


a b

Figure 12.65

(a) Keypoints for the half-sized building and a section of the right corner. (b) Corresponding keypoints between the corner and the building. Of the seven matches found, only one is in error.
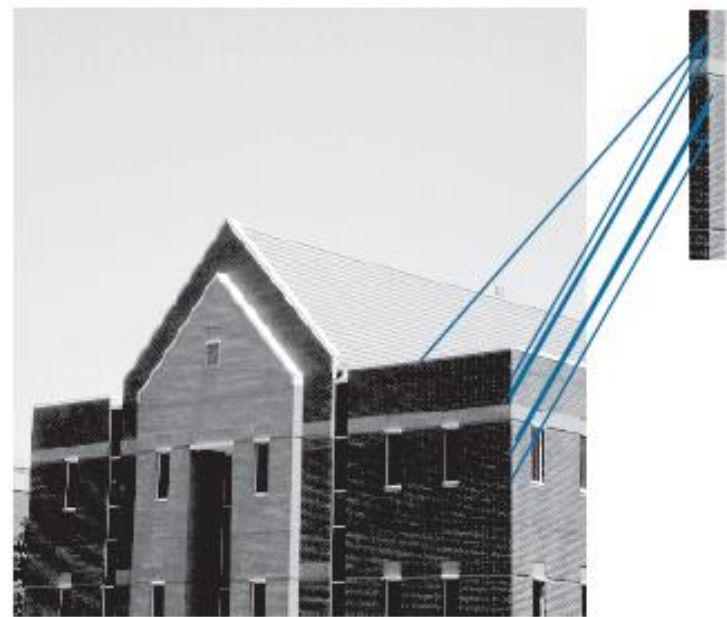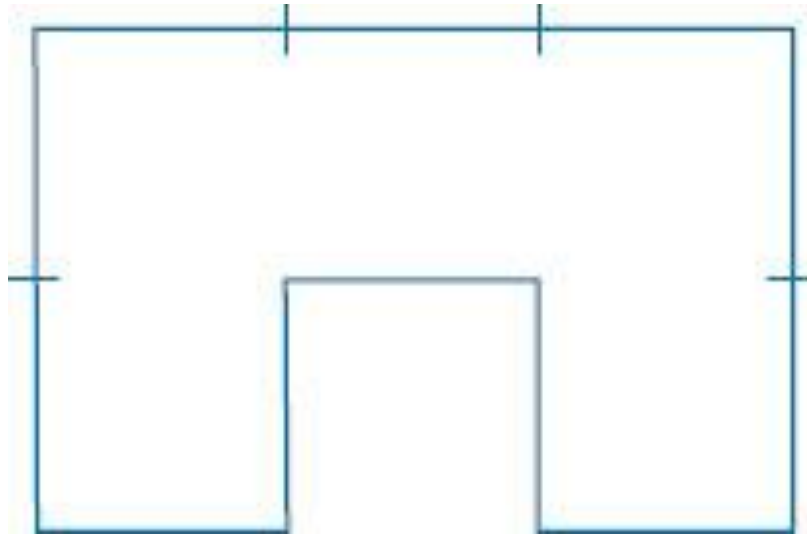

a b

Figure 12.66

(a) Matches between the original building image and a rotated version of a segment of its right corner. Ten matches were found, of which two are incorrect. (b) Matches between the original image and a half-scaled version of a segment of its right corner. Here, 11 matches were found, of which four were incorrect.
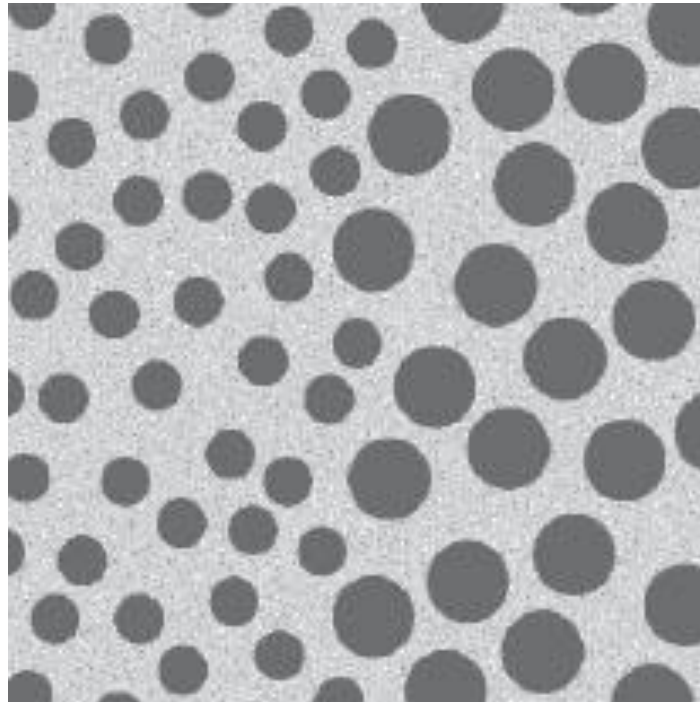

a b
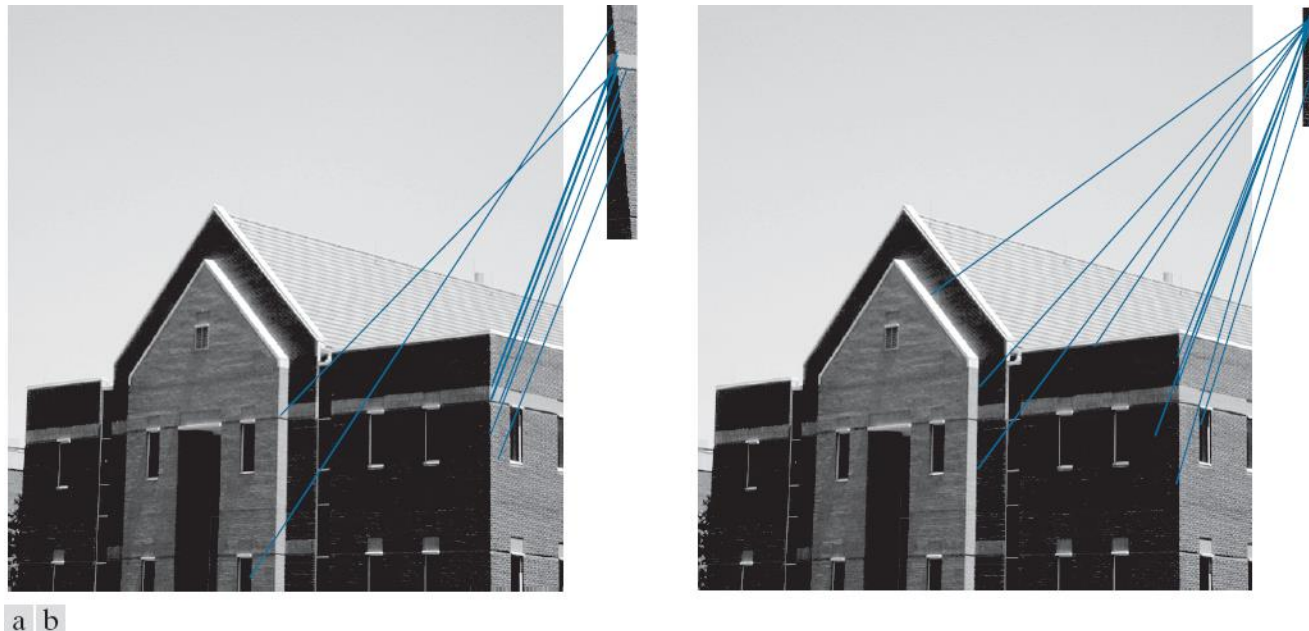
Unnumbered Figure 1

Unnumbered Figure 2

Unnumbered Figure 3

Unnumbered Figure 4

Copyright