Digital Image Processing: PIKS Inside, Third Edition. William K. Pratt

ISBNs: 0-471-37407-5 (Hardback); 0-471-22132-5 (Electronic)

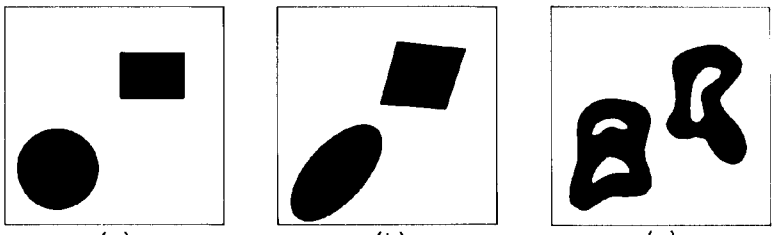
18

SHAPE ANALYSIS

Several qualitative and quantitative techniques have been developed for characterizing the shape of objects within an image. These techniques are useful for classifying objects in a pattern recognition system and for symbolically describing objects in an image understanding system. Some of the techniques apply only to binary-valued images; others can be extended to gray level images.

18.1. TOPOLOGICAL ATTRIBUTES

Topological shape attributes are properties of a shape that are invariant under *rubber-sheet* transformation (1–3). Such a transformation or mapping can be visualized as the stretching of a rubber sheet containing the image of an object of a given shape to produce some spatially distorted object. Mappings that require cutting of the rubber sheet or connection of one part to another are not permissible. Metric distance is clearly not a topological attribute because distance can be altered by rubber-sheet stretching. Also, the concepts of perpendicularity and parallelism between lines are not topological properties. Connectivity is a topological attribute. Figure 18.1-1*a* is a binary-valued image containing two connected object components. Figure 18.1-1*b* is a spatially stretched version of the same image. Clearly, there are no stretching operations that can either increase or decrease the connectivity of the objects in the stretched image. Connected components of an object may contain holes, as illustrated in Figure 18.1-1*c*. The number of holes is obviously unchanged by a topological mapping.



(*a*) Two objects

(*b*) Two objects after rubber sheet stretching

(*c*) Two objects with holes $C = 2, H = 3, E = 1$

FIGURE 18.1-1. Topological attributes.

There is a fundamental relationship between the number of connected object components C and the number of object holes H in an image called the *Euler number*, as defined by

$$E = C - H \tag{18.1-1}$$

The Euler number is also a topological property because C and H are topological attributes.

Irregularly shaped objects can be described by their topological constituents. Consider the tubular-shaped object letter R of Figure 18.1-2*a*, and imagine a rubber band stretched about the object. The region enclosed by the rubber band is called the *convex hull* of the object. The set of points within the convex hull, which are not in the object, form the *convex deficiency* of the object. There are two types of convex deficiencies: regions totally enclosed by the object, called *lakes*; and regions lying between the convex hull perimeter and the object, called *bays*. In some applications it is simpler to describe an object indirectly in terms of its convex hull and convex deficiency. For objects represented over rectilinear grids, the definition of the convex hull must be modified slightly to remain meaningful. Objects such as discretized circles and triangles clearly should be judged as being convex even though their

(*a*) Object (*b*) Convex hull (*c*) Bays and lake

FIGURE 18.1-2. Definitions of convex shape descriptors.

boundaries are jagged. This apparent difficulty can be handled by considering a rubber band to be stretched about the discretized object. A pixel lying totally within the rubber band, but not in the object, is a member of the convex deficiency. Sklansky et al. (4,5) have developed practical algorithms for computing the convex attributes of discretized objects.

18.2. DISTANCE, PERIMETER, AND AREA MEASUREMENTS

Distance is a real-valued function $d\{(j_1, k_1), (j_2, k_2)\}$ of two image points (j_1, k_1) and (j_2, k_2) satisfying the following properties (6):

$$d\{(j_1, k_1), (j_2, k_2)\} \geq 0 \tag{18.2-1a}$$

$$d\{(j_1, k_1), (j_2, k_2)\} = d\{(j_2, k_2), (j_1, k_1)\} \tag{18.2-1b}$$

$$d\{(j_1, k_1), (j_2, k_2)\} + d\{(j_2, k_2), (j_3, k_3)\} \geq d\{(j_1, k_1), (j_3, k_3)\} \tag{18.2-1c}$$

There are a number of distance functions that satisfy the defining properties. The most common measures encountered in image analysis are the *Euclidean distance,*

$$d_E = \left[(j_1 - j_2)^2 + (k_1 - k_2)^2\right]^{1/2} \tag{18.2-2a}$$

the *magnitude distance,*

$$d_M = |j_1 - j_2| + |k_1 - k_2| \tag{18.2-2b}$$

and the *maximum value distance,*

$$d_X = \text{MAX}\{|j_1 - j_2|, |k_1 - k_2|\} \tag{18.2-2c}$$

In discrete images, the coordinate differences $(j_1 - j_2)$ and $(k_1 - k_2)$ are integers, but the Euclidean distance is usually not an integer.

Perimeter and area measurements are meaningful only for binary images. Consider a discrete binary image containing one or more objects, where $F(j, k) = 1$ if a pixel is part of the object and $F(j, k) = 0$ for all nonobject or background pixels.

The perimeter of each object is the count of the number of pixel sides traversed around the boundary of the object starting at an arbitrary initial boundary pixel and returning to the initial pixel. The area of each object within the image is simply the count of the number of pixels in the object for which $F(j, k) = 1$. As an example, for

a 2×2 pixel square, the object area is $A_O = 4$ and the object perimeter is $P_O = 8$. An object formed of three diagonally connected pixels possesses $A_O = 3$ and $P_O = 12$.

The enclosed area of an object is defined to be the total number of pixels for which $F(j,k) = 0$ or 1 within the outer perimeter boundary P_E of the object. The enclosed area can be computed during a boundary-following process while the perimeter is being computed (7,8). Assume that the initial pixel in the boundary-following process is the first black pixel encountered in a raster scan of the image. Then, proceeding in a clockwise direction around the boundary, a crack code $C(p)$, as defined in Section 17.6, is generated for each side p of the object perimeter such that $C(p) = 0, 1, 2, 3$ for directional angles 0, 90, 180, 270°, respectively. The enclosed area is

$$A_E = \sum_{p=1}^{P_E} j(p-1)\,\Delta k(p) \tag{18.2-3a}$$

where P_E is the perimeter of the enclosed object and

$$j(p) = \sum_{i=1}^{p} \Delta j(i) \tag{18.2-3b}$$

with $j(0) = 0$. The delta terms are defined by

$$\Delta j(p) = \begin{cases} 1 & \text{if } C(p) = 1 \quad (18.2\text{-}4a) \\ 0 & \text{if } C(p) = 0 \text{ or } 2 \quad (18.2\text{-}4b) \\ -1 & \text{if } C(p) = 3 \quad (18.2\text{-}4c) \end{cases}$$

$$\Delta k(p) = \begin{cases} 1 & \text{if } C(p) = 0 \quad (18.2\text{-}4d) \\ 0 & \text{if } C(p) = 1 \text{ or } 3 \quad (18.2\text{-}4e) \\ -1 & \text{if } C(p) = 2 \quad (18.2\text{-}4f) \end{cases}$$

Table 18.2-1 gives an example of computation of the enclosed area of the following four-pixel object:

TABLE 18.2-1. Example of Perimeter and Area Computation

p	$C(p)$	$\Delta\ j(p)$	$\Delta\ k(p)$	$j(p)$	$A(p)$
1	0	0	1	0	0
2	3	–1	0	–1	0
3	0	0	1	–1	–1
4	1	1	0	0	–1
5	0	0	1	0	–1
6	3	–1	0	–1	–1
7	2	0	–1	–1	0
8	3	–1	0	–2	0
9	2	0	–1	–2	2
10	2	0	–1	–2	4
11	1	1	0	–1	4
12	1	1	0	0	4

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

18.2.1. Bit Quads

Gray (9) has devised a systematic method of computing the area and perimeter of binary objects based on matching the logical state of regions of an image to binary patterns. Let $n\{Q\}$ represent the count of the number of matches between image pixels and the pattern Q within the curly brackets. By this definition, the object area is then

$$A_O = n\{1\} \tag{18.2-5}$$

If the object is enclosed completely by a border of white pixels, its perimeter is equal to

$$P_O = 2n\{0\ 1\} + 2n\left\{\begin{matrix}0\\1\end{matrix}\right\} \tag{18.2-6}$$

Now, consider the following set of 2×2 pixel patterns called *bit quads* defined in Figure 18.2-1. The object area and object perimeter of an image can be expressed in terms of the number of bit quad counts in the image as

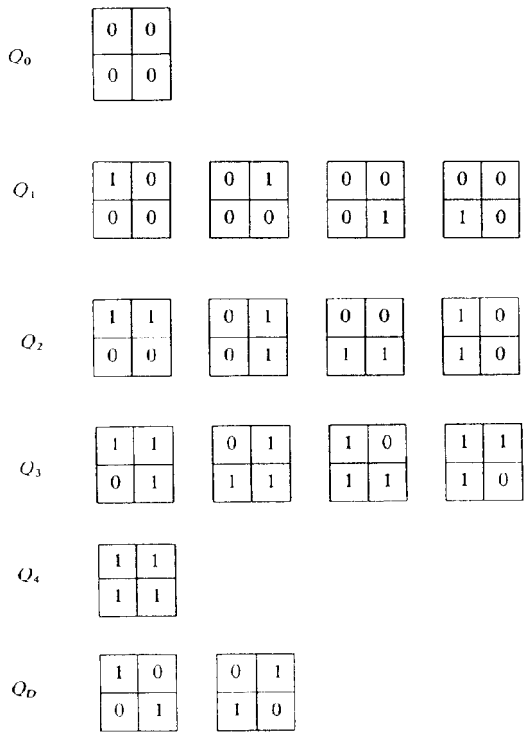


FIGURE 18.2-1. Bit quad patterns.

$$A_O = \frac{1}{4}[n\{Q_1\} + 2n\{Q_2\} + 3n\{Q_3\} + 4n\{Q_4\} + 2n\{Q_D\}] \quad (18.2\text{-}7a)$$

$$P_O = n\{Q_1\} + n\{Q_2\} + n\{Q_3\} + 2n\{Q_D\} \quad (18.2\text{-}7b)$$

These area and perimeter formulas may be in considerable error if they are utilized to represent the area of a continuous object that has been coarsely discretized. More accurate formulas for such applications have been derived by Duda (10):

$$A_O = \frac{1}{4}n\{Q_1\} + \frac{1}{2}n\{Q_2\} + \frac{7}{8}n\{Q_3\} + n\{Q_4\} + \frac{3}{4}n\{Q_D\} \quad (18.2\text{-}8a)$$

$$P_O = n\{Q_2\} + \frac{1}{\sqrt{2}}[n\{Q_1\} + n\{Q_3\} + 2n\{Q_D\}] \quad (18.2\text{-}8b)$$

Bit quad counting provides a very simple means of determining the Euler number of an image. Gray (9) has determined that under the definition of four-connectivity, the Euler number can be computed as

$$E = \tfrac{1}{4}[n\{Q_1\} - n\{Q_3\} + 2n\{Q_D\}] \tag{18.2-9a}$$

and for eight-connectivity

$$E = \tfrac{1}{4}[n\{Q_1\} - n\{Q_3\} - 2n\{Q_D\}] \tag{18.2-9b}$$

It should be noted that although it is possible to compute the Euler number E of an image by local neighborhood computation, neither the number of connected components C nor the number of holes H, for which $E = C - H$, can be separately computed by local neighborhood computation.

18.2.2. Geometric Attributes

With the establishment of distance, area, and perimeter measurements, various geometric attributes of objects can be developed. In the following, it is assumed that the number of holes with respect to the number of objects is small (i.e., E is approximately equal to C).

The *circularity* of an object is defined as

$$C_O = \frac{4\pi A_O}{(P_O)^2} \tag{18.2-10}$$

This attribute is also called the *thinness ratio*. A circle-shaped object has a circularity of unity; oblong-shaped objects possess a circularity of less than 1.

If an image contains many components but few holes, the Euler number can be taken as an approximation of the number of components. Hence, the average area and perimeter of connected components, for $E > 0$, may be expressed as (9)

$$A_A = \frac{A_O}{E} \tag{18.2-11}$$

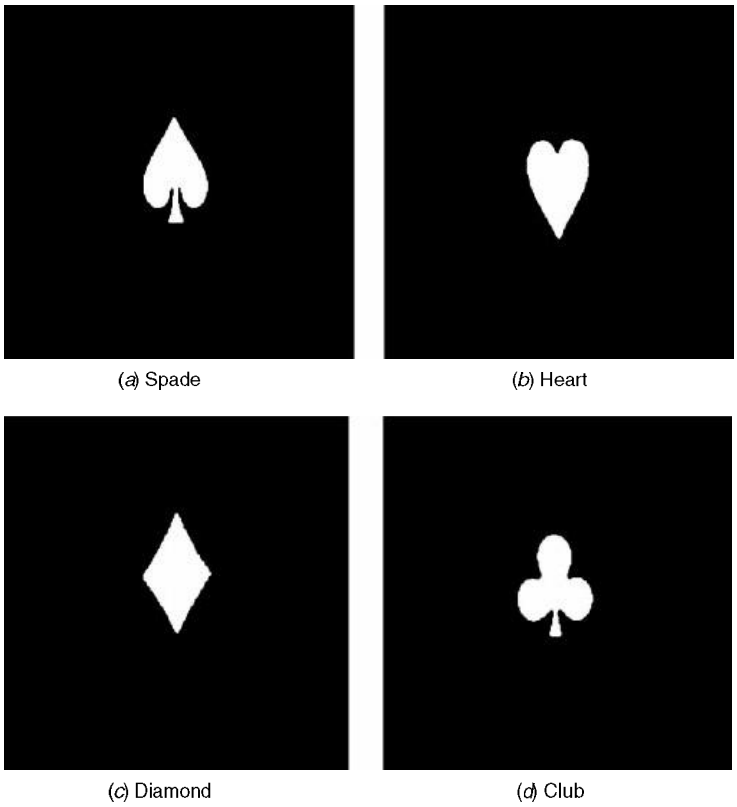
$$P_A = \frac{P_O}{E} \tag{18.2-12}$$

For images containing thin objects, such as typewritten or script characters, the average object length and width can be approximated by

$$L_A = \frac{P_A}{2} \tag{18.2-13}$$

$$W_A = \frac{2A_A}{P_A} \tag{18.2-14}$$

These simple measures are useful for distinguishing gross characteristics of an image. For example, does it contain a multitude of small pointlike objects, or fewer bloblike objects of larger size; are the objects fat or thin? Figure 18.2-2 contains images of playing card symbols. Table 18.2-2 lists the geometric attributes of these objects.



(*a*) Spade (*b*) Heart

(*c*) Diamond (*d*) Club

FIGURE 18.2-2. Playing card symbol images.

TABLE 18.2-2 Geometric Attributes of Playing Card Symbols

Attribute	Spade	Heart	Diamond	Club
Outer perimeter	652	512	548	668
Enclosed area	8,421	8,681	8.562	8.820
Average area	8,421	8,681	8,562	8,820
Average perimeter	652	512	548	668
Average length	326	256	274	334
Average width	25.8	33.9	31.3	26.4
Circularity	0.25	0.42	0.36	0.25

18.3. SPATIAL MOMENTS

From probability theory, the (m, n)th moment of the joint probability density $p(x, y)$ is defined as

$$M(m, n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^m y^n p(x, y)\, dx\, dy \tag{18.3-1}$$

The central moment is given by

$$U(m, n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \eta_x)^m (y - \eta_y)^n p(x, y)\, dx\, dy \tag{18.3-2}$$

where η_x and η_y are the marginal means of $p(x, y)$. These classical relationships of probability theory have been applied to shape analysis by Hu (11) and Alt (12). The concept is quite simple. The joint probability density $p(x, y)$ of Eqs. 18.3-1 and 18.3-2 is replaced by the continuous image function $F(x, y)$. Object shape is characterized by a few of the low-order moments. Abu-Mostafa and Psaltis (13,14) have investigated the performance of spatial moments as features for shape analysis.

18.3.1. Discrete Image Spatial Moments

The spatial moment concept can be extended to discrete images by forming spatial summations over a discrete image function $F(j, k)$. The literature (15–17) is notationally inconsistent on the discrete extension because of the differing relationships defined between the continuous and discrete domains. Following the notation established in Chapter 13, the (m, n)th spatial moment is defined as

$$M_U(m, n) = \sum_{j=1}^{J}\sum_{k=1}^{K} (x_k)^m (y_j)^n F(j, k) \tag{18.3-3}$$

where, with reference to Figure 13.1-1, the scaled coordinates are

$$x_k = k - \frac{1}{2} \tag{18.3-4a}$$

$$y_j = J + \frac{1}{2} - j \tag{18.3-4b}$$

The origin of the coordinate system is the lower left corner of the image. This formulation results in moments that are extremely scale dependent; the ratio of second-order ($m + n = 2$) to zero-order ($m = n = 0$) moments can vary by several orders of magnitude (18). The spatial moments can be restricted in range by spatially scaling the image array over a unit range in each dimension. The (m, n)th scaled spatial moment is then defined as

$$M(m, n) = \frac{1}{J^n K^m} \sum_{j=1}^{J} \sum_{k=1}^{K} (x_k)^m (y_j)^n F(j, k) \tag{18.3-5}$$

Clearly,

$$M(m, n) = \frac{M_U(m, n)}{J^n K^m} \tag{18.3-6}$$

It is instructive to explicitly identify the lower-order spatial moments. The zero-order moment

$$M(0, 0) = \sum_{j=1}^{J} \sum_{k=1}^{K} F(j, k) \tag{18.3-7}$$

is the sum of the pixel values of an image. It is called the *image surface*. If $F(j, k)$ is a binary image, its surface is equal to its area. The *first-order row moment* is

$$M(1, 0) = \frac{1}{K} \sum_{j=1}^{J} \sum_{k=1}^{K} x_k F(j, k) \tag{18.3-8}$$

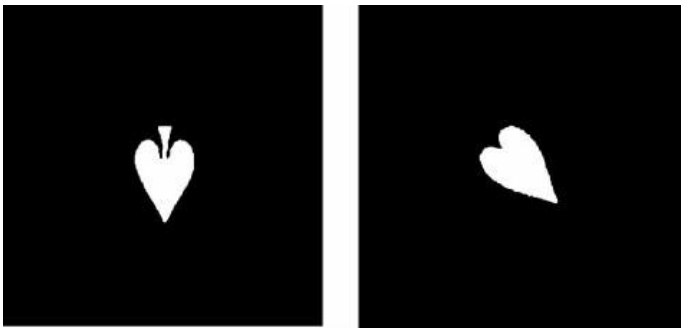
and the *first-order column moment* is

$$M(0, 1) = \frac{1}{J} \sum_{j=1}^{J} \sum_{k=1}^{K} y_j F(j, k) \tag{18.3-9}$$

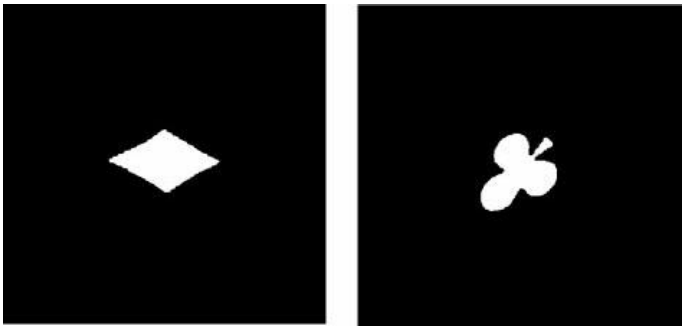
Table 18.3-1 lists the scaled spatial moments of several test images. These images include unit-amplitude gray scale versions of the playing card symbols of Figure 18.2-2, several rotated, minified and magnified versions of these symbols, as shown in Figure 18.3-1, as well as an elliptically shaped gray scale object shown in Figure 18.3-2. The ratios

TABLE 18.3-1. Scaled Spatial Moments of Test Images

Image	$M(0,0)$	$M(1,0)$	$M(0,1)$	$M(2,0)$	$M(1,1)$	$M(0,2)$	$M(3,0)$	$M(2,1)$	$M(1,2)$	$M(0,3)$
Spade	8,219.98	4,013.75	4,281.28	1,976.12	2,089.86	2,263.11	980.81	1,028.31	1,104.36	1,213.73
Rotated spade	8,215.99	4,186.39	3,968.30	2,149.35	2,021.65	1,949.89	1,111.69	1,038.04	993.20	973.53
Heart	8,616.79	4,283.65	4,341.36	2,145.90	2,158.40	2,223.79	1,083.06	1,081.72	1,105.73	1,156.35
Rotated Heart	8,613.79	4,276.28	4,337.90	2,149.18	2,143.52	2,211.15	1,092.92	1,071.95	1,008.05	1,140.43
Magnified heart	34,523.13	17,130.64	17,442.91	8,762.68	8,658.34	9,402.25	4,608.05	4,442.37	4,669.42	5,318.58
Minified heart	2,104.97	1,047.38	1,059.44	522.14	527.16	535.38	260.78	262.82	266.41	271.61
Diamond	8,561.82	4,349.00	4,704.71	2,222.43	2,390.10	2,627.42	1,142.44	1,221.53	1,334.97	1,490.26
Rotated diamond	8,562.82	4,294.89	4,324.09	2,196.40	2,168.00	2,196.97	1,143.83	1,108.30	1,101.11	1,122.93
Club	8,781.71	4,323.54	4,500.10	2,150.47	2,215.32	2,344.02	1,080.29	1,101.21	1,153.76	1,241.04
Rotated club	8,787.71	4,363.23	4,220.96	2,196.08	2,103.88	2,057.66	1,120.12	1,062.39	1,028.90	1,017.60
Ellipse	8,721.74	4,326.93	4,377.78	2,175.86	2,189.76	2,226.61	1,108.47	1,109.92	1,122.62	1,146.97



(*a*) Rotated spade (*b*) Rotated heart



(*c*) Rotated diamond (*d*) Rotated club

(*e*) Minified heart (*f*) Magnified heart

FIGURE 18.3-1 Rotated, magnified, and minified playing card symbol images.

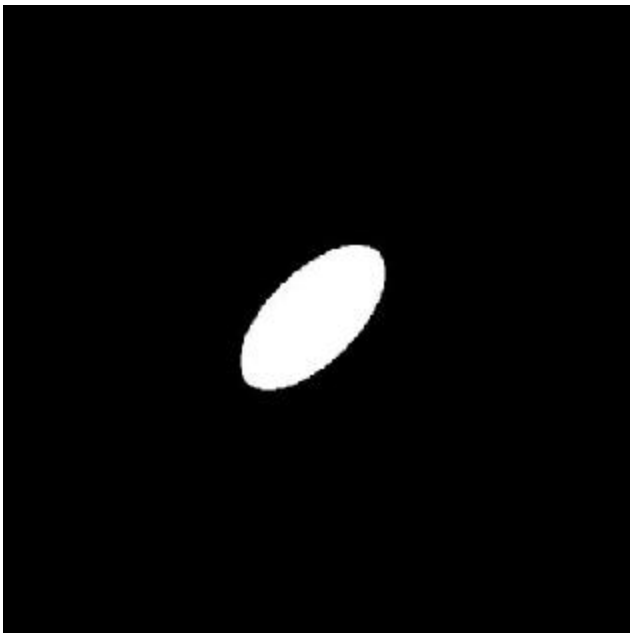


FIGURE 18.3-2 Eliptically shaped object image.

$$\bar{x}_k = \frac{M(1,0)}{M(0,0)} \tag{18.3-10a}$$

$$\bar{y}_j = \frac{M(0,1)}{M(0,0)} \tag{18.3-10b}$$

of first- to zero-order spatial moments define the *image centroid.* The centroid, called the *center of gravity*, is the balance point of the image function $F(j,k)$ such that the mass of $F(j,k)$ left and right of $\bar{x}_k$ and above and below $\bar{y}_j$ is equal.

With the centroid established, it is possible to define the scaled spatial central moments of a discrete image, in correspondence with Eq. 18.3-2, as

$$U(m,n) = \frac{1}{J^n K^m} \sum_{j=1}^{J} \sum_{k=1}^{K} (x_k - \bar{x}_k)^m (y_j - \bar{y}_j)^n F(j,k) \tag{18.3-11}$$

For future reference, the (m,n)th unscaled spatial central moment is defined as

$$U_U(m,n) = \sum_{j=1}^{J} \sum_{k=1}^{K} (x_k - \bar{x}_k)^m (y_j - \bar{y}_j)^n F(j,k) \tag{18.3-12}$$

where

$$\tilde{x}_k = \frac{M_U(1,0)}{M_U(0,0)} \tag{18.3-13a}$$

$$\tilde{y}_j = \frac{M_U(0,1)}{M_U(0,0)} \tag{18.3-13b}$$

It is easily shown that

$$U(m,n) = \frac{U_U(m,n)}{J^n K^m} \tag{18.3-14}$$

The three second-order scaled central moments are the *row moment of inertia,*

$$U(2,0) = \frac{1}{K^2} \sum_{j=1}^{J} \sum_{k=1}^{K} (x_k - \bar{x}_k)^2 F(j,k) \tag{18.3-15}$$

the *column moment of inertia,*

$$U(0,2) = \frac{1}{J^n} \sum_{j=1}^{J} \sum_{k=1}^{K} (y_j - \bar{y}_j)^2 F(j,k) \tag{18.3-16}$$

and the *row–column cross moment of inertia,*

$$U(1,1) = \frac{1}{JK} \sum_{j=1}^{J} \sum_{k=1}^{K} (x_k - \bar{x}_k)(y_j - \bar{y}_j) F(j,k) \tag{18.3-17}$$

The central moments of order 3 can be computed directly from Eq. 18.3-11 for $m + n = 3$, or indirectly according to the following relations:

$$U(3,0) = M(3,0) - 3\bar{y}_j M(2,0) + 2(\bar{y}_j)^2 M(1,0) \tag{18.3-18a}$$

$$U(2,1) = M(2,1) - 2\bar{y}_j M(1,1) - \bar{x}_k M(2,0) + 2(\bar{y}_j)^2 M(0,1) \tag{18.3-18b}$$

$$U(1,2) = M(1,2) - 2\bar{x}_k M(1,1) - \bar{y}_j M(0,2) + 2(\bar{x}_k)^2 M(1,0) \quad (18.3\text{-}18c)$$

$$U(0,3) = M(0,3) - 3\bar{x}_k M(0,2) + 2(\bar{x}_k)^2 M(0,1) \quad (18.3\text{-}18d)$$

Table 18.3-2 presents the horizontal and vertical centers of gravity and the scaled central spatial moments of the test images.

The three second-order moments of inertia defined by Eqs. 18.3-15, 18.3-16, and 18.3-17 can be used to create the moment of inertia covariance matrix,

$$\mathbf{U} = \begin{bmatrix} U(2,0) & U(1,1) \\ U(1,1) & U(0,2) \end{bmatrix} \quad (18.3\text{-}19)$$

Performing a singular-value decomposition of the covariance matrix results in the diagonal matrix

$$\mathbf{E}^T\mathbf{U}\mathbf{E} = \mathbf{\Lambda} \quad (18.3\text{-}20)$$

where the columns of

$$\mathbf{E} = \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix} \quad (18.3\text{-}21)$$

are the eigenvectors of $\mathbf{U}$ and

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (18.3\text{-}22)$$

contains the eigenvalues of $\mathbf{U}$. Expressions for the eigenvalues can be derived explicitly. They are

$$\lambda_1 = \tfrac{1}{2}[U(2,0) + U(0,2)] + \tfrac{1}{2}[U(2,0)^2 + U(0,2)^2 - 2U(2,0)U(0,2) + 4U(1,1)^2]^{1/2} \quad (18.3\text{-}23a)$$

$$\lambda_2 = \tfrac{1}{2}[U(2,0) + U(0,2)] - \tfrac{1}{2}[U(2,0)^2 + U(0,2)^2 - 2U(2,0)U(0,2) + 4U(1,1)^2]^{1/2} \quad (18.3\text{-}23b)$$

TABLE 18.3-2 Centers of Gravity and Scaled Spatial Central Moments of Test Images

Image	Horizontal COG	Vertical COG	$U(2,0)$	$U(1,1)$	$U(0,2)$	$U(3,0)$	$U(2,1)$	$U(1,2)$	$U(0,3)$
Spade	0.488	0.521	16.240	–0.653	33.261	0.026	–0.285	–0.017	0.363
Rotated spade	0.510	0.483	16.207	–0.366	33.215	–0.013	0.284	–0.002	–0.357
Heart	0.497	0.504	16.380	0.194	36.506	–0.012	0.371	0.027	–0.831
Rotated heart	0.496	0.504	26.237	–10.009	26.584	–0.077	–0.438	0.411	0.122
Magnified heart	0.496	0.505	262.321	3.037	589.162	0.383	11.991	0.886	–27.284
Minified heart	0.498	0.503	0.984	0.013	2.165	0.000	0.011	0.000	–0.025
Diamond	0.508	0.549	13.337	0.324	42.186	–0.002	–0.026	0.005	0.136
Rotated diamond	0.502	0.505	42.198	–0.853	13.366	–0.158	0.009	0.029	–0.005
Club	0.492	0.512	21.834	–0.239	37.979	0.037	–0.545	–0.039	0.950
Rotated club	0.497	0.480	29.675	8.116	30.228	0.268	–0.505	–0.557	0.216
Ellipse	0.496	0.502	29.236	17.913	29.236	0.000	0.000	0.000	0.000

Let $\lambda_M = \text{MAX}\{\lambda_1, \lambda_2\}$ and $\lambda_N = \text{MIN}\{\lambda_1, \lambda_2\}$, and let the orientation angle θ be defined as

$$\theta = \begin{cases} \arctan\left\{\dfrac{e_{21}}{e_{11}}\right\} & \text{if } \lambda_M = \lambda_1 \quad (18.3\text{-}24a) \\ \arctan\left\{\dfrac{e_{22}}{e_{12}}\right\} & \text{if } \lambda_M = \lambda_2 \quad (18.3\text{-}24b) \end{cases}$$

The orientation angle can be expressed explicitly as

$$\theta = \arctan\left\{\frac{\lambda_M - U(0,2)}{U(1,1)}\right\} \quad (18.3\text{-}24c)$$

The eigenvalues λ_M and λ_N and the orientation angle θ define an ellipse, as shown in Figure 18.3-2, whose major axis is λ_M and whose minor axis is λ_N. The major axis of the ellipse is rotated by the angle θ with respect to the horizontal axis. This elliptically shaped object has the same moments of inertia along the horizontal and vertical axes and the same moments of inertia along the principal axes as does an actual object in an image. The ratio

$$R_A = \frac{\lambda_N}{\lambda_M} \quad (18.3\text{-}25)$$

of the minor-to-major axes is a useful shape feature.

Table 18.3-3 provides moment of inertia data for the test images. It should be noted that the orientation angle can only be determined to within plus or minus $\pi/2$ radians.

TABLE 18.3-3 Moment of Intertia Data of Test Images

Image	Largest Eigenvalue	Smallest Eigenvalue	Orientation (radians)	Eigenvalue Ratio
Spade	33.286	16.215	–0.153	0.487
Rotated spade	33.223	16.200	–1.549	0.488
Heart	36.508	16.376	1.561	0.449
Rotated heart	36.421	16.400	–0.794	0.450
Magnified heart	589.190	262.290	1.562	0.445
Minified heart	2.165	0.984	1.560	0.454
Diamond	42.189	13.334	1.560	0.316
Rotated diamond	42.223	13.341	–0.030	0.316
Club	37.982	21.831	–1.556	0.575
Rotated club	38.073	21.831	0.802	0.573
Ellipse	47.149	11.324	0.785	0.240

Hu (11) has proposed a normalization of the unscaled central moments, defined by Eq. 18.3-12, according to the relation

$$V(m,n) = \frac{U_U(m,n)}{[M(0,0)]^{\alpha}} \tag{18.3-26a}$$

where

$$\alpha = \frac{m+n}{2} + 1 \tag{18.3-26b}$$

for $m + n = 2, 3,...$ These normalized central moments have been used by Hu to develop a set of seven compound spatial moments that are invariant in the continuous image domain to translation, rotation, and scale change. The *Hu invariant moments* are defined below.

$$h_1 = V(2,0) + V(0,2) \tag{18.3-27a}$$

$$h_2 = [V(2,0) - V(0,2)]^2 + 4[V(1,1)]^2 \tag{18.3-27b}$$

$$h_3 = [V(3,0) - 3V(1,2)]^2 + [V(0,3) - 3V(2,1)]^2 \tag{18.3-27c}$$

$$h_4 = [V(3,0) + V(1,2)]^2 + [V(0,3) - V(2,1)]^2 \tag{18.3-27d}$$

$$\begin{aligned} h_5 = {} & [V(3,0) - 3V(1,2)][V(3,0) + V(1,2)][[V(3,0) + V(1,2)]^2 - 3[V(0,3) + V(2,1)]^2] \\ & + [3V(2,1) - V(0,3)][V(0,3) + V(2,1)][3[V(3,0) + V(1,2)]^2 \\ & - [V(0,3) + V(2,1)]^2] \end{aligned} \tag{18.3-27e}$$

$$\begin{aligned} h_6 = {} & [V(2,0) - V(0,2)][[V(3,0) + V(1,2)]^2 - [V(0,3) + V(2,1)]^2] \\ & + 4V(1,1)[V(3,0) + V(1,2)][V(0,3) + V(2,1)] \end{aligned} \tag{18.3-27f}$$

$$\begin{aligned} h_7 = {} & [3V(2,1) - V(0,3)][V(3,0) + V(1,2)][[V(3,0) + V(1,2)]^2 - 3[V(0,3) + V(2,1)]^2] \\ & + [3V(1,2) - V(3,0)][V(0,3) + V(2,1)][3[V(3,0) + V(1,2)]^2 \\ & - [V(0,3) + V(2,1)]^2] \end{aligned} \tag{18.3-27g}$$

Table 18.3-4 lists the moment invariants of the test images. As desired, these moment invariants are in reasonably close agreement for the geometrically modified versions of the same object, but differ between objects. The relatively small degree of variability of the moment invariants for the same object is due to the spatial discretization of the objects.

TABLE 18.3-4 Invariant Moments of Test Images

Image	$h_1 \times 10^1$	$h_2 \times 10^3$	$h_3 \times 10^3$	$h_4 \times 10^5$	$h_5 \times 10^9$	$h_6 \times 10^6$	$h_7 \times 10^1$
Spade	1.920	4.387	0.715	0.295	0.123	0.185	–14.159
Rotated spade	1.919	4.371	0.704	0.270	0.097	0.162	–11.102
Heart	1.867	5.052	1.435	8.052	27.340	5.702	–15.483
Rotated heart	1.866	5.004	1.434	8.010	27.126	5.650	–14.788
Magnified heart	1.873	5.710	1.473	8.600	30.575	6.162	0.559
Minified heart	1.863	4.887	1.443	8.019	27.241	5.583	0.658
Diamond	1.986	10.648	0.018	0.475	0.004	0.490	0.004
Rotated diamond	1.987	10.663	0.024	0.656	0.082	0.678	–0.020
Club	2.033	3.014	2.313	5.641	20.353	3.096	10.226
Rotated club	2.033	3.040	2.323	5.749	20.968	3.167	13.487
Ellipse	2.015	15.242	0.000	0.000	0.000	0.000	0.000

The terms of Eq. 18.3-27 contain differences of relatively large quantities, and therefore, are sometimes subject to significant roundoff error. Liao and Pawlak (19) have investigated the numerical accuracy of moment measures.

18.4. SHAPE ORIENTATION DESCRIPTORS

The spatial orientation of an object with respect to a horizontal reference axis is the basis of a set of orientation descriptors developed at the Stanford Research Institute (20). These descriptors, defined below, are described in Figure 18.4-1.

1. *Image-oriented bounding box:* the smallest rectangle oriented along the rows of the image that encompasses the object

2. *Image-oriented box height:* dimension of box height for image-oriented box

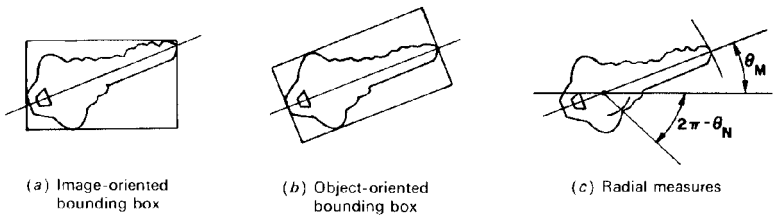


FIGURE 18.4-1. Shape orientation descriptors.

3. *Image-oriented box width:* dimension of box width for image-oriented box
4. *Image-oriented box area:* area of image-oriented bounding box
5. *Image oriented box ratio:* ratio of box area to enclosed area of an object for an image-oriented box
6. *Object-oriented bounding box:* the smallest rectangle oriented along the major axis of the object that encompasses the object
7. *Object-oriented box height:* dimension of box height for object-oriented box
8. *Object-oriented box width:* dimension of box width for object-oriented box
9. *Object-oriented box area:* area of object-oriented bounding box
10. *Object-oriented box ratio:* ratio of box area to enclosed area of an object for an object-oriented box
11. *Minimum radius:* the minimum distance between the centroid and a perimeter pixel
12. *Maximum radius:* the maximum distance between the centroid and a perimeter pixel
13. *Minimum radius angle:* the angle of the minimum radius vector with respect to the horizontal axis
14. *Maximum radius angle:* the angle of the maximum radius vector with respect to the horizontal axis
15. *Radius ratio:* ratio of minimum radius angle to maximum radius angle

Table 18.4-1 lists the orientation descriptors of some of the playing card symbols.

TABLE 18.4-1 Shape Orientation Descriptors of the Playing Card Symbols

Descriptor	Spade	Rotated Heart	Rotated Diamond	Rotated Club
Row-bounding box height	155	122	99	123
Row-bounding box width	95	125	175	121
Row-bounding box area	14,725	15,250	17,325	14,883
Row-bounding box ratio	1.75	1.76	2.02	1.69
Object-bounding box height	94	147	99	148
Object-bounding box width	154	93	175	112
Object-bounding box area	14,476	13,671	17,325	16,576
Object-bounding box ratio	1.72	1.57	2.02	1.88
Minimum radius	11.18	38.28	38.95	26.00
Maximum radius	92.05	84.17	88.02	82.22
Minimum radius angle	–1.11	0.35	1.06	0.00
Maximum radius angle	–1.54	–0.76	0.02	0.85

18.5. FOURIER DESCRIPTORS

The perimeter of an arbitrary closed curve can be represented by its instantaneous curvature at each perimeter point. Consider the continuous closed curve drawn on the complex plane of Figure 18.5-1, in which a point on the perimeter is measured by its polar position $z(s)$ as a function of arc length s. The complex function $z(s)$ may be expressed in terms of its real part $x(s)$ and imaginary part $y(s)$ as

$$z(s) = x(s) + iy(s) \tag{18.5-1}$$

The tangent angle defined in Figure 18.5-1 is given by

$$\Phi(s) = \arctan\left\{\frac{dy(s)/ds}{dx(s)/ds}\right\} \tag{18.5-2}$$

and the curvature is the real function

$$k(s) = \frac{d\Phi(s)}{ds} \tag{18.5-3}$$

The coordinate points $x(s)$, $y(s)$ can be obtained from the curvature function by the reconstruction formulas

$$x(s) = x(0) + \int_0^s k(\alpha) \cos\{\Phi(\alpha)\}\, d\alpha \tag{18.5-4a}$$

$$y(s) = y(0) + \int_0^s k(\alpha) \sin\{\Phi(\alpha)\}\, d\alpha \tag{18.5-4b}$$

where $x(0)$ and $y(0)$ are the starting point coordinates.

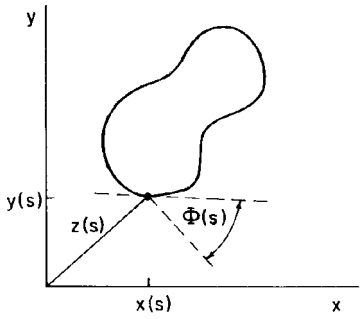


FIGURE 18.5-1. Geometry for curvature definition.

Because the curvature function is periodic over the perimeter length P, it can be expanded in a Fourier series as

$$k(s) = \sum_{n=-\infty}^{\infty} c_n \exp\left\{\frac{2\pi i n s}{P}\right\} \tag{18.5-5a}$$

where the coefficients c_n are obtained from

$$c_n = \frac{1}{P}\int_0^P k(s)\exp\left\{-\frac{2\pi i n}{P}\right\} ds \tag{18.5-5b}$$

This result is the basis of an analysis technique developed by Cosgriff (21) and Brill (22) in which the Fourier expansion of a shape is truncated to a few terms to produce a set of Fourier descriptors. These Fourier descriptors are then utilized as a symbolic representation of shape for subsequent recognition.

If an object has sharp discontinuities (e.g., a rectangle), the curvature function is undefined at these points. This analytic difficulty can be overcome by the utilization of a cumulative shape function

$$\theta(s) = \int_0^s k(\alpha)\, d\alpha - \frac{2\pi s}{P} \tag{18.5-6}$$

proposed by Zahn and Roskies (23). This function is also periodic over P and can therefore be expanded in a Fourier series for a shape description.

Bennett and MacDonald (24) have analyzed the discretization error associated with the curvature function defined on discrete image arrays for a variety of connectivity algorithms. The discrete definition of curvature is given by

$$z(s_j) = x(s_j) + iy(s_j) \tag{18.5-7a}$$

$$\Phi(s_j) = \arctan\left\{\frac{y(s_j) - y(s_{j-1})}{x(s_j) - x(s_{j-1})}\right\} \tag{18.5-7b}$$

$$k(s_j) = \Phi(s_j) - \Phi(s_{j-1}) \tag{18.5-7c}$$

where s_j represents the jth step of arc position. Figure 18.5-2 contains results of the Fourier expansion of the discrete curvature function.

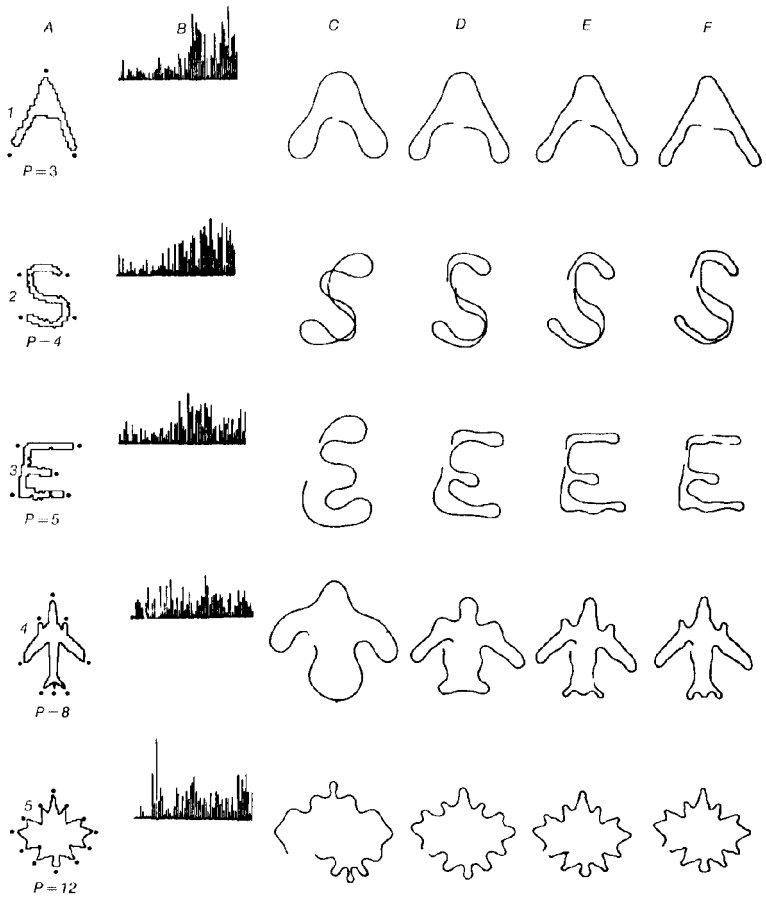


FIGURE 18.5-2. Fourier expansions of curvature function.

REFERENCES

1. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley-Interscience, New York, 1973.
2. E. C. Greanis et al., "The Recognition of Handwritten Numerals by Contour Analysis," *IBM J. Research and Development*, **7**, 1, January 1963, 14–21.

3. M. A. Fischler, "Machine Perception and Description of Pictorial Data," *Proc. International Joint Conference on Artificial Intelligence*, D. E. Walker and L. M. Norton, Eds., May 1969, 629–639.
4. J. Sklansky, "Recognizing Convex Blobs," *Proc. International Joint Conference on Artificial Intelligence*, D. E. Walker and L. M. Norton, Eds., May 1969, 107–116.
5. J. Sklansky, L. P. Cordella, and S. Levialdi, "Parallel Detection of Concavities in Cellular Blobs," *IEEE Trans. Computers*, **C-25**, 2, February 1976, 187–196.
6. A. Rosenfeld and J. L. Pflatz, "Distance Functions on Digital Pictures," *Pattern Recognition*, **1**, July 1968, 33–62.
7. Z. Kulpa, "Area and Perimeter Measurements of Blobs in Discrete Binary Pictures," *Computer Graphics and Image Processing*, **6**, 5, October 1977, 434–451.
8. G. Y. Tang, "A Discrete Version of Green's Theorem," *IEEE Trans. Pattern Analysis and Machine Intelligence*, **PAMI-7**, 3, May 1985, 338–344.
9. S. B. Gray, "Local Properties of Binary Images in Two Dimensions," *IEEE Trans. Computers*, **C-20**, 5, May 1971, 551–561.
10. R. O. Duda, "Image Segmentation and Description," unpublished notes, 1975.
11. M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Information Theory*, **IT-8**, 2, February 1962, 179–187.
12. F. L. Alt, "Digital Pattern Recognition by Moments," *J. Association for Computing Machinery*, **9**, 2, April 1962, 240–258.
13. Y. S. Abu-Mostafa and D. Psaltis, "Recognition Aspects of Moment Invariants," *IEEE Trans. Pattern Analysis and Machine Intelligence*, **PAMI-6**, 6, November 1984, 698–706.
14. Y. S. Abu-Mostafa and D. Psaltis, "Image Normalization by Complex Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence*, **PAMI-7**, 6, January 1985, 46–55.
15. S. A. Dudani et al., "Aircraft Identification by Moment Invariants," *IEEE Trans. Computers*, **C-26**, February 1962, 179–187.
16. F. W. Smith and M. H. Wright, "Automatic Ship Interpretation by the Method of Moments," *IEEE Trans. Computers*, **C-20**, 1971, 1089–1094.
17. R. Wong and E. Hall, "Scene Matching with Moment Invariants," *Computer Graphics and Image Processing*, **8**, 1, August 1978, 16–24.
18. A. Goshtasby, "Template Matching in Rotated Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, **PAMI-7**, 3, May 1985, 338–344.
19. S. X. Liao and M. Pawlak, "On Image Analysis by Moments,"*IEEE Trans. Pattern Analysis and Machine Intelligence*, **PAMI-18**, 3, March 1996, 254–266.
20. Stanford Research Institute, unpublished notes.
21. R. L. Cosgriff, "Identification of Shape," Report 820-11, ASTIA AD 254 792, Ohio State University Research Foundation, Columbus, OH, December 1960.
22. E. L. Brill, "Character Recognition via Fourier Descriptors," *WESCON Convention Record*, Paper 25/3, Los Angeles, 1968.
23. C. T. Zahn and R. Z. Roskies, "Fourier Descriptors for Plane Closed Curves," *IEEE Trans. Computers*, **C-21**, 3, March 1972, 269–281.
24. J. R. Bennett and J. S. MacDonald, "On the Measurement of Curvature in a Quantized Environment," *IEEE Trans. Computers*, **C-25**, 8, August 1975, 803–820.