

- [Uray 97] P Uray. From 3D point clouds to surface and volumes. PhD thesis, Technische Universität Graz, Austria, October 1997.
- [Waltz 75] D L Waltz. Understanding line drawings of scenes with shadows. In P H Winston, editor, *The Psychology of Computer Vision*, pages 19–91. McGraw-Hill, New York, 1975.
- [Weiss 88] I Weiss. Projective invariants of shapes. In *Proceedings of the DARPA Image Understanding Workshop*, Cambridge, MA, volume 2, pages 1125–1134. DARPA, 1988.
- [Weiss and Rosenfeld 96] I Weiss and A Rosenfeld. A simple method for ellipse detection. Technical Report CS-TR-3717, University of Maryland, Center for Automation Research, College Park, MD, 1996.
- [Werner et al. 95] T Werner, R D Hersch, and V Hlaváč. Rendering real-world objects using view interpolation. In *5th International Conference on Computer Vision*, MIT, pages 957–962, IEEE, Piscataway, NJ, 1995.

- [Werner et al. 96] T Werner, V Hlaváč, A Leonardis, and T Pajdla. Selection of reference views for image-based representation. In *Proceedings of the 13th International Conference on Pattern Recognition*, volume I—Track A: Computer Vision, pages 73–77, IEEE Computer Society Press, Los Alamitos, CA, Vienna, Austria, August 1996.
- [Werner et al. 97] T Werner, T Pajdla, and V Hlaváč. Visualizing 3-D real-world scenes using view extrapolation. Technical Report K335-CMP-1997-137, Department of Control, Faculty of Electrical Engineering, Czech Technical University, Prague, June 1997.
- [Witkin 81] A P Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.
- [Worring and Smeulders 93] M Worring and A W M Smeulders. Digital curvature estimation. *CVCP: Image Understanding*, 58(3):366–382, November 1993.

Chapter 11

Mathematical morphology

11.1 Basic morphological concepts

Mathematical morphology, which started to develop in the late 1960s, stands as a relatively separate part of image analysis. It is based on the algebra of non-linear operators operating on object shape and in many respects supersedes the linear algebraic system of convolution. It performs in many tasks—pre-processing, segmentation using object shape, and object quantification—better and more quickly than the standard approaches. The main obstacle for the novice user of mathematical morphology tools is the slightly different algebra than is usual in standard algebra and calculus courses.

The main protagonists of mathematical morphology were Matheron [Matheron 67] and Serra [Serra 82], whose monographs are highly mathematical books, and more recent books are typically written in a similar spirit, e.g., [Giardina and Dougherty 88, Dougherty 92, Heijmans 94]. Other significant references are [Maragos and Schafer 87a, Maragos and Schafer 87b, Serra 87, Roerdink and Heijmans 88].

Our aim is to present morphology in a manner that is relatively easy to follow [Haralick and Shapiro 92, Vincent 95]. Morphological tools are implemented in most advanced image analysis packages, and we hope the reader will learn enough to apply them in a qualified way. Mathematical morphology is very often used in applications where shape of objects and speed is an issue—for example, analysis of microscopic images (in biology, material science, geology, and criminology), industrial inspection, optical character recognition, and document analysis.

The non-morphological approach to image processing is close to calculus, being based on the point-spread function concept and linear transformations such as convolution, and we have discussed image modeling and processing from this point of view in earlier chapters. Mathematical morphology uses tools of non-linear algebra and operates with point sets, their connectivity and shape. Morphological operations simplify images, and quantify and preserve the main shape characteristics of objects.

Morphological operations are used predominantly for the following purposes:

- Image pre-processing (noise filtering, shape simplification)
- Enhancing object structure (skeletonizing, thinning, thickening, convex hull, object marking)

- Segmenting objects from the background
- Quantitative description of objects (area, perimeter, projections, Euler-Poincaré characteristic)

Mathematical morphology exploits point set properties, results of integral geometry, and topology. The initial assumption states that real images can be modeled using point sets of any dimension (e.g., N -dimensional Euclidean space); the Euclidean 2D space \mathcal{E}^2 and its system of subsets is a natural domain for planar shape description. Understanding of inclusion (\subset or \supset), intersection (\cap), union (\cup), the empty set \emptyset , and set complement (\complement) is assumed. Set difference is defined by

$$X \setminus Y = X \cap Y^c \quad (11.1)$$

Computer vision uses the digital counterpart of Euclidean space—sets of integer pairs ($\in \mathbb{Z}^2$) for binary image morphology or sets of integer triples ($\in \mathbb{Z}^3$) for gray-scale morphology or binary 3D morphology.

We begin by considering binary images that can be viewed as subsets of the 2D space of all integers, \mathbb{Z}^2 . A point is represented by a pair of integers that give co-ordinates with respect to the two co-ordinate axes of the digital raster; the unit length of the raster equals the sampling period in each direction. We talk about a **discrete grid** if the neighborhood relation between points is well defined. This representation is suitable for both rectangular and hexagonal grids, but a rectangular grid is assumed hereafter.

A binary image can be treated as a 2D point set. Points belonging to objects in the image represent a set X —these points are pixels with value equal to one. Points of the complement set X^c correspond to the background with pixel values equal to zero. The origin (marked as a diagonal cross in our examples) has co-ordinates $(0, 0)$, and co-ordinates of any point are interpreted as (x, y) in the common way used in mathematics. Figure 11.1 shows an example of such a set—points belonging to the object are denoted by small black squares. Any point x from a discrete image $X = \{(1, 0), (1, 1), (1, 2), (2, 2), (0, 3), (0, 4)\}$ can be treated as a vector with respect to the origin $(0, 0)$.

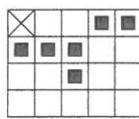


Figure 11.1: A point set example.

A morphological transformation Ψ is given by the relation of the image (point set X) with another small point set B called a **structuring element**. B is expressed with respect to a local origin O (called the representative point). Some typical structuring elements are shown in Figure 11.2. Figure 11.2c illustrates the possibility of the point O not being a member of the structuring element B .

To apply the morphological transformation $\Psi(X)$ to the image X means that the structuring element B is moved systematically across the entire image. Assume that B is positioned

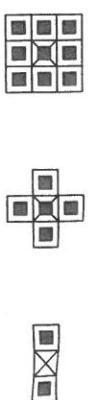


Figure 11.2: Typical structuring elements.

at some point in the image; the pixel in the image corresponding to the representative point O of the structuring element is called the **current pixel**. The result of the relation (which can be either zero or one) between the image X and the structuring element B in the current position is stored in the output image in the current image pixel position.

The duality of morphological operations is deduced from the existence of the set complement; for each morphological transformation $\Psi(X)$ there exists a dual transformation $\Psi^*(X)$

$$\Psi(X) = [\Psi^*(X^c)]^c \quad (11.2)$$

The translation of the point set X by the vector h is denoted by X_h ; it is defined by

$$X_h = \{p \in \mathcal{E}^2, p = x + h \text{ for some } x \in X\} \quad (11.3)$$

This is illustrated in Figure 11.3.

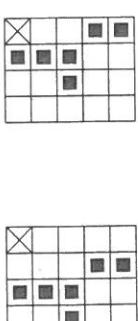


Figure 11.3: Translation by a vector.

11.2 Four morphological principles

It is appropriate to restrict the set of possible morphological transformations in image analysis by imposing several constraints on it; we shall briefly present here four morphological principles that express such constraints. These concepts may be difficult to understand, but an understanding of them is not essential to a comprehension of what follows, and they may be taken for granted. A detailed explanation of these matters may be found in [Serra 82].

Humans have an intuitive understanding of spatial structure. The structure of the Alps versus an oak tree crown is perceived as different. Besides the need for objective descriptions of such objects, the scientist requires a quantitative description. Generalization is expected as well; the interest is not in a specific oak tree, but in the class of oaks.

The morphological approach with quantified results consists of two main steps: (a) geometrical transformation and (b) the actual measurement. [Serra 82] gives two examples. The

first is from chemistry, where the task is to measure the surface area of some object. First, the initial body is reduced to its surface, e.g., by marking by some chemical matter. Second, the quantity of the marker needed to cover the surface is measured. Another example is from sieving analysis, often used in geology, when the distribution of sizes of milled rocks is of interest. The milled product is passed through sieves with different sizes of holes from coarse to fine. The result is a sequence of subsets of milled product. For each sieve size some oversize particles remain on it after sieving, and these are measured.

A morphological operator is (by definition) composition of a mapping Ψ (or geometrical transformation) followed by a measure μ which is a mapping $Z \times \dots \times Z \rightarrow R$. The geometrically transformed set $\Psi(X)$ can be the boundary, oversized particles in sieving analysis, etc., and the measure $\mu[\Psi(X)]$ yields a number (weight, surface area, volume, etc.). The discussion here is simplified just to the transformation Ψ , but the axiomatics can be transposed to measures as well.

A morphological transformation is called quantitative if and only if it satisfies four basic principles [Serra 82].

- Compatibility with translation: Let the transformation Ψ depend on the position of the origin O of the co-ordinate system, and denote such a transformation by Ψ_O . If all points are translated by the vector $-h$, it is expressed as Ψ_{-h} . The *compatibility with translation* principle is given by

$$\Psi_O(X_h) = [\Psi_{-h}(X)]_h \quad (11.4)$$

If Ψ does not depend on the position of the origin O , then the compatibility with translation principle reduces to invariance under translation:

$$\Psi(X_h) = [\Psi(X)]_h \quad (11.5)$$

- Compatibility with change of scale: Let λX represent the homothetic scaling of a point set X (i.e., the co-ordinates of each point of the set are multiplied by some positive constant λ). This is equivalent to change of scale with respect to some origin. Let Ψ_λ denote a transformation that depends on the positive parameter λ (change of scale). *Compatibility with change of scale* is given by

$$\Psi_\lambda(X) = \lambda \Psi\left(\frac{1}{\lambda} X\right) \quad (11.6)$$

If Ψ does not depend on the scale λ , then compatibility with change of scale reduces to invariance to change of scale:

$$\Psi(\lambda X) = \lambda \Psi(X) \quad (11.7)$$

- Local knowledge: The local knowledge principle considers the situation in which only a part of a larger structure can be examined—this is always the case in reality, due to the restricted size of the digital grid. The morphological transformation Ψ satisfies the *local knowledge principle* if for any bounded point set Z' in the transformation $\Psi(X)$ there exists a bounded set Z knowledge of which is sufficient to provide Ψ . The local knowledge principle may be written symbolically as

$$[\Psi(X \cap Z)] \cap Z' = \Psi(X) \cap Z'$$

$$(11.8)$$

- Upper semi-continuity: The upper semi-continuity principle says that the morphological transformation does not exhibit any abrupt changes. A precise explanation needs many concepts from topology and is given in [Serra 82].

11.3 Binary dilation and erosion

The sets of black and white pixels constitute a description of a binary image. Assume that only black pixels are considered, and the others are treated as a background. The primary morphological operations are dilation and erosion, and from these two, more complex morphological operations such as opening, closing, and shape decomposition can be constituted. We present them here using Minkowski's formalism [Haralick and Shapiro 92]. The Minkowski algebra is closer to the notions taught in standard mathematics courses (an alternative is Serra's formalism based on stereological concepts [Serra 82]).

11.3.1 Dilation

The morphological transformation dilation \oplus combines two sets using vector addition (or Minkowski set addition, e.g., $(a, b) + (c, d) = (a+c, b+d)$). The dilation $X \oplus B$ is the point set of all possible vector additions of pairs of elements, one from each of the sets X and B .

$$X \oplus B = \{p \in \mathcal{E}^2 : p = x + b, x \in X \text{ and } b \in B\} \quad (11.9)$$

Figure 11.4 illustrates an example of dilation.

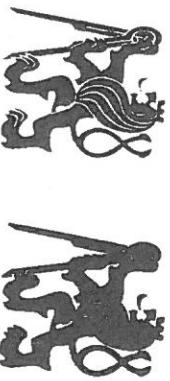
$$\begin{aligned} X &= \{(1, 0), (1, 1), (1, 2), (2, 2), (0, 3), (0, 4)\} \\ B &= \{(0, 0), (1, 0)\} \\ X \oplus B &= \{(1, 0), (1, 1), (1, 2), (2, 2), (0, 3), (0, 4), \\ &\quad (2, 0), (2, 1), (2, 2), (3, 2), (1, 3), (1, 4)\} \end{aligned}$$



Figure 11.4: Dilation.

Figure 11.5 shows a 256×256 original image (the emblem of the Czech Technical University) on the left. A structuring element of size 3×3 , see Figure 11.2a, is used. The result of dilation is shown on the right side of Figure 11.5. In this case the dilation is an isotropic expansion (it behaves the same way in all directions). This operation is also sometimes called *fill or grow*.

Dilation with an isotropic 3×3 structuring element might be described as a transformation which changes all background pixels neighboring the object to object pixels.

Figure 11.5: *Dilation as isotropic expansion.*

Dilation has several interesting properties that may ease its hardware or software implementation; we present some here without proof. The interested reader may consult [Serra 82] or the tutorial paper [Haralick et al. 87].

The dilation operation is commutative,

$$X \oplus B = B \oplus X \quad (11.10)$$

and is also associative,

$$X \oplus (B \oplus D) = (X \oplus B) \oplus D \quad (11.11)$$

Dilation may also be expressed as a union of shifted point sets,

$$X \oplus B = \bigcup_{b \in B} X_b \quad (11.12)$$

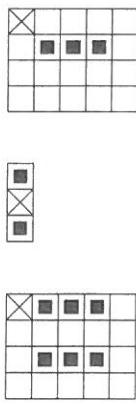
and is invariant to translation,

$$X_h \oplus B = (X \oplus B)_h \quad (11.13)$$

Equations (11.12) and (11.13) show the importance of shifts in speeding up implementations of dilation, and this holds for implementations of binary morphology on serial computers in general. One processor word represents several pixels (e.g., 32 for a 32-bit processor), and shift or addition corresponds to a single instruction. Shifts may also be easily implemented as delays in a pipeline parallel processor.

Dilation is an increasing transformation;

$$\text{If } X \subseteq Y \text{ then } X \oplus B \subseteq Y \oplus B \quad (11.14)$$

Figure 11.6: *Dilation where the representative point is not a member of the structuring element.*

Dilation is used to fill small holes and narrow gulls in objects. It increases the object size—if the original size needs to be preserved, then dilation is combined with erosion, described in the next section.

Figure 11.6 illustrates the result of dilation if the representative point is not a member of the structuring element B ; if this structuring element is used, the dilation result is substantially different from the input set. Notice that the connectivity of the original set has been lost.

11.3.2 Erosion

Erosion \ominus combines two sets using vector subtraction of set elements and is the dual operator of dilation. Neither erosion nor dilation is an invertible transformation.

$$X \ominus B = \{p \in \mathcal{E}^2 : p + b \in X \text{ for every } b \in B\} \quad (11.15)$$

This formula says that every point p from the image is tested; the result of the erosion is given by those points p for which all possible $p + b$ are in X . Figure 11.7 shows an example of the point set X eroded by the structuring element B .

$$\begin{aligned} X &= \{(1, 0), (1, 1), (1, 2), (0, 3), (1, 3), (2, 3), (3, 3), (1, 4)\} \\ B &= \{(0, 0), (1, 0)\} \\ X \ominus B &= \{(0, 3), (1, 3), (2, 3)\} \end{aligned}$$

Figure 11.7: *Erosion.*

Figure 11.8 shows the erosion by a 3×3 element (see Figure 11.2a) of the same original as in Figure 11.5. Notice that single-pixel-wide lines disappear. Erosion (such as Figure 11.8) with an isotropic structuring element is called *shrink* or *reduce* by some authors.

Basic morphological transformations can be used to find the contours of objects in an image very quickly. This can be achieved, for instance, by subtraction from the original picture of its eroded version—see Figure 11.9.

Erosion is used to simplify the structure of an object—objects or their parts with width equal to one will disappear. It might thus decompose complicated objects into several simpler ones.

There is an equivalent definition of erosion [Matheron 75]. Recall that B_p denotes B translated by p :

$$X \ominus B = \{p \in \mathcal{E}^2 : B_p \subseteq X\} \quad (11.16)$$

Denote by \check{B} the symmetrical set to B (called the transpose [Serra 82] or rational set [Haralick et al. 87] by some authors) with respect to the representative point O :

$$\check{B} = \{-b : b \in B\}$$

For example,

$$\begin{aligned} B &= \{(1, 2), (2, 3)\} \\ \check{B} &= \{(-1, -2), (-2, -3)\} \end{aligned} \quad (11.21)$$

Figure 11.8: Erosion as isotropic shrink.

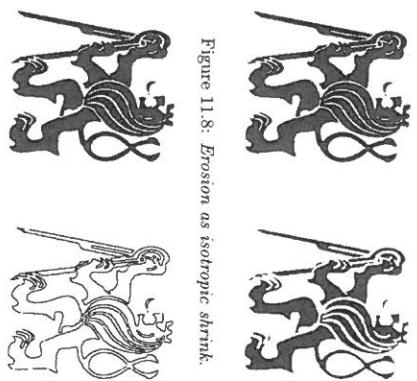


Figure 11.9: Contours obtained by subtraction of an eroded image from an original (left).

The erosion might be interpreted by structuring element B sliding across the image X ; then, if B translated by the vector p is contained in the image X , the point corresponding to the representative point of B belongs to the erosion $X \ominus B$.

An implementation of erosion might be simplified by noting that an image X eroded by the structuring element B can be expressed as an intersection of all translations of the image X by the vector¹ $-b \in B$:

$$X \ominus B = \bigcap_{b \in B} X_{-b} \quad (11.17)$$

If the representative point is a member of the structuring element, then erosion is an anti-extensive transformation; that is, if $(0, 0) \in B$, then $X \ominus B \subseteq X$. Erosion is also translation invariant,

$$X_h \ominus B = (X \ominus B)_h \quad (11.18)$$

and, like dilation, is an increasing transformation.

$$\text{If } X \subseteq Y \text{ then } X \ominus B \subseteq Y \ominus B \quad (11.20)$$

If B, D are structuring elements, and D is contained in B , then erosion by B is more aggressive than by D ; that is, if $D \subseteq B$, then $X \ominus B \subseteq X \ominus D$. This property enables the ordering of erosions according to structuring elements of similar shape but different sizes.

¹This definition of erosion, \ominus , differs from that used in [Serra 82]. There \ominus denotes Minkowski subtraction, which is an intersection of all translations of the image by the vector $b \in B$. In our case the minus sign has been added. In our notation, if convex sets are used, the dilation of erosion (or the other way around) is identity.

We have already mentioned that erosion and dilation are dual transformations. Formally, $(X \ominus Y)^C = X^C \oplus Y$

$$\begin{aligned} X \ominus B &\neq B \ominus X \\ \text{The properties of erosion and intersection combined together are} \\ (X \cap Y) \ominus B &= (X \ominus B) \cap (Y \ominus B) \\ B \ominus (X \cap Y) &\supseteq (B \ominus X) \cup (B \ominus Y) \end{aligned} \quad (11.25)$$

On the other hand, image intersection and dilation cannot be interchanged; the dilation of the intersection of two images is contained in the intersection of their dilations:

$$(X \cap Y) \oplus B = B \oplus (X \cap Y) \subseteq (X \oplus B) \cap (Y \oplus B) \quad (11.26)$$

The order of erosion may be interchanged with set union. This fact enables the structuring element to be decomposed into a union of simpler structuring elements:

$$B \oplus (X \cup Y) = (X \cup Y) \oplus B = (X \oplus B) \cup (Y \oplus B)$$

$$\begin{aligned} (X \cup Y) \ominus B &\supseteq (X \ominus B) \cup (Y \ominus B) \\ B \ominus (X \cup Y) &= (B \ominus X) \cap (B \ominus Y) \end{aligned} \quad (11.27)$$

Successive dilation (respectively, erosion) of the image X first by the structuring element B and then by the structuring element D is equivalent to the dilation (erosion) of the image X by $B \oplus D$:

$$\begin{aligned} (X \oplus B) \oplus D &= X \oplus (B \oplus D) \\ (X \ominus B) \ominus D &= X \ominus (B \ominus D) \end{aligned} \quad (11.28)$$

11.3.3 Hit-or-miss transformation

The hit-or-miss transformation is the morphological operator for finding local patterns of pixels, where *local* means the size of the structuring element. It is a variant of template matching that finds collections of pixels with certain shape properties (such as corners, or border points). We shall see later that it may be used for thinning and thickening of objects (Section 11.5.3).

Operations described hitherto used a structuring element B , and we have tested points for their membership of X ; we can also test whether some points do not belong to X . An operation may be denoted by a pair of disjoint sets $B = (B_1, B_2)$, called a composite structuring element. The hit-or-miss transformation \otimes is defined as

$$X \otimes B = \{x : B_1 \subset X \text{ and } B_2 \subset X^c\} \quad (11.29)$$

This means that for a point x to be in the resulting set, two conditions must be fulfilled simultaneously: First the part B_1 of the composite structuring element that has its representative point at x must be contained in X , and second, the part B_2 of the composite structuring element must be contained in X^c .

The hit-or-miss transformation operates as a binary matching between an image X and the structuring element (B_1, B_2) . It may be expressed using erosions and dilations as well

$$X \otimes B = (X \ominus B_1) \cap (X^c \ominus B_2) = (X \ominus B_1) \setminus (X \oplus B_2) \quad (11.30)$$

11.3.4 Opening and closing

Erosion and dilation are not inverse transformations—if an image is eroded and then dilated, the original image is not re-obtained. Instead, the result is a simplified and less detailed version of the original image.

Erosion followed by dilation creates an important morphological transformation called **opening**. The opening of an image X by the structuring element B is denoted by $X \circ B$ and is defined as

$$X \circ B = (X \ominus B) \oplus B \quad (11.31)$$

Dilation followed by erosion is called **closing**. The closing of an image X by the structuring element B is denoted by $X \bullet B$ and is defined as

$$X \bullet B = (X \oplus B) \ominus B \quad (11.32)$$

If an image X is unchanged by opening with the structuring element B , it is called *open with respect to B* . Similarly, if an image X is unchanged by closing with B , it is called *closed with respect to B* .

Opening and closing with an isotropic structuring element is used to eliminate specific image details smaller than the structuring element—the global shape of the objects is not distorted. Closing connects objects that are close to each other, fills up small holes, and smooths the object outline by filling up narrow gulls. Meanings of ‘near’, ‘small’, and ‘narrow’ are related to the size and the shape of the structuring element. Opening is illustrated in Figure 11.10, and closing in Figure 11.11.



Figure 11.10: Opening (original on the left).

Unlike dilation and erosion, opening and closing are invariant to translation of the structuring element. Equations (11.14) and (11.20) imply that both opening and closing are increasing transformations. Opening is anti-extensive ($X \circ B \subseteq X$) and closing is extensive ($X \subseteq X \bullet B$).

Opening and closing, like dilation and erosion, are dual transformations:

$$(X \bullet B)^C = X^C \circ \bar{B} \quad (11.33)$$

Another significant fact is that iteratively used openings and closings are **idempotent**, meaning that reapplication of these transformations does not change the previous result. Formally,

$$X \circ B = (X \circ B) \circ B \quad (11.34)$$

$$X \bullet B = (X \bullet B) \bullet B \quad (11.35)$$

11.4 Gray-scale dilation and erosion

Binary morphological operations acting on binary images are easily extendible to gray-scale images using the ‘min’ and ‘max’ operations. Erosion (respectively, dilation) of an image is the operation of assigning to each pixel the minimum (maximum) value found over a neighborhood of the corresponding pixel in the input image. The structuring element is more rich than in the binary case, where it gave only the neighborhood. In the gray-scale case, the structuring element is a function of two variables that specifies the desired local gray-level

property. The value of the structuring element is added (subtracted) when the maximum (or minimum) is calculated in the neighborhood.

This extension permits a **topographic view** of gray-scale images—the gray-level is interpreted as the height of a particular location of a hypothetical landscape. Light and dark spots in the image correspond to hills and hollows in the landscape. Such a morphological approach permits the location of global properties of the image, i.e., to identify characteristic topographic features on images as valleys, mountain ridges (crests), and watersheds.

We follow the explanation used in [Haralick and Shapiro 92] for gray-scale dilation and erosion, where the concept of **umbra** and **top of the point set** is introduced. Gray-scale dilation is expressed as the dilation of umbras.

11.4.1 Top surface, umbra, and gray-scale dilation and erosion

Consider a point set A in n -dimensional Euclidean space, $A \subseteq \mathcal{E}^n$, and assume that the first $(n - 1)$ co-ordinates of the set constitute a spatial domain and the n^{th} co-ordinate corresponds to the value of a function or functions at a point ($n = 3$ for gray-scale images). This interpretation matches the topographic view for a 2D Euclidean space, where points are given by triples of co-ordinates; the first two co-ordinates locate the position in the 2D support set and the third co-ordinate gives the height.

The **top surface** of a set A is a function defined on the $(n - 1)$ -dimensional support. For each $(n - 1)$ -tuple, the top surface is the highest value of the last co-ordinate of A , as illustrated in Figure 11.12. If the space is Euclidean the highest value means supremum.

$$y = f(x_1, x_2)$$

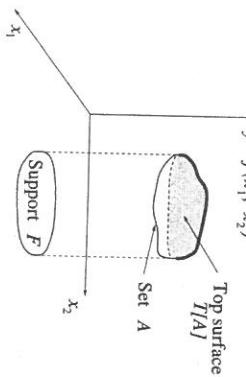


Figure 11.12: *Top surface of the set A corresponds to maximal values of the function $f(x_1, x_2)$.*

Let $A \subseteq \mathcal{E}^n$ and the support $F = \{x \in \mathcal{E}^{n-1} \text{ for some } y \in \mathcal{E}, (x, y) \in A\}$. The **top surface** of A , denoted by $T[A]$, is a mapping $F \rightarrow \mathcal{E}$ defined as

$$T[A](x) = \max\{y, (x, y) \in A\} \quad (11.36)$$

The next concept is the **umbra** of a function f defined on some subset F (support) of $(n - 1)$ -dimensional space. The usual definition of umbra is a region of complete shadow resulting from obstructing the light by a non-transparent object. In mathematical morphology, the umbra of f is a set that consists of the top surface of f and everything below it; see Figure 11.13.

Figure 11.13: *Umbra of the top surface of a set is the whole subspace below it.*

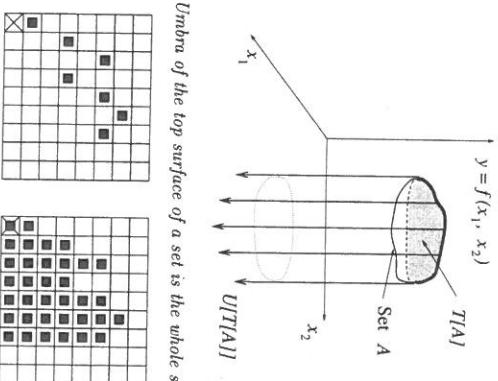


Figure 11.14: *Example of a 1D function (left) and its umbra (right).*

Formally, let $F \subseteq \mathcal{E}^{n-1}$ and $f : F \rightarrow \mathcal{E}$. The **umbra** of f , denoted by $U[f]$, $U[f] \subseteq F \times \mathcal{E}$, is defined by

$$U[f] = \{(x, y) \in F \times \mathcal{E}, y \leq f(x)\} \quad (11.37)$$

We see that the umbra of an umbra of f is an umbra.

We can illustrate the top surface and umbra in the case of a simple 1D gray-scale image. Figure 11.14 illustrates a function f (which might be a top surface) and its umbra.

We can now define the gray-scale dilation of two functions as the top surface of the dilation of their umbras. Let $F, K \subseteq \mathcal{E}^{n-1}$ and $f : F \rightarrow \mathcal{E}$ and $k : K \rightarrow \mathcal{E}$. The dilation \oplus of f by k , $f \oplus k : F \oplus K \rightarrow \mathcal{E}$ is defined by

$$f \oplus k = T\{U[f] \oplus U[k]\} \quad (11.38)$$

Notice here that \oplus on the left-hand side is dilation in the gray-scale image domain, and \oplus on the right-hand side is dilation in the binary image. A new symbol was not introduced since no confusion is expected here; the same applies to erosion \ominus in due course.

Similarly to binary dilation, one function, say f , represents an image, and the second, k , a small structuring element. Figure 11.15 shows a discretized function k that will play the role of the structuring element. Figure 11.16 shows the dilation of the umbra of f (from the example given in Figure 11.14) by the umbra of k .



Figure 11.15: A structuring element: 1D function (left) and its umbra (right).

$$\begin{array}{c} Uff \oplus U[k] \\ \hline \text{---} \\ T[Uff \oplus U[k]] = f \oplus k \end{array}$$

Figure 11.16: 1D example of gray-scale dilation. The umbras of the 1D function f and structuring element k are dilated first, $U[f] \oplus U[k]$. The top surface of this dilated set gives the result, $f \oplus k = T[Uf] \oplus U[k]$.

This definition explains what gray-scale dilation means, but does not give a reasonable algorithm for actual computations in hardware. We shall see that a computationally plausible way to calculate dilation can be obtained by taking the maximum of a set of sums:

$$(f \oplus k)(x) = \max\{f(x - z) + k(z), z \in K, x - z \in F\} \quad (11.39)$$

The computational complexity is the same as for convolution in linear filtering, where a summation of products is performed.

The definition of gray-scale erosion is analogous to gray-scale dilation. The gray-scale erosion of two functions (point sets)

1. Takes their umbras
2. Erodes them using binary erosion
3. Gives the result as the top surface

Let $F, K \subseteq \mathcal{E}^{n-1}$ and $f : F \rightarrow \mathcal{E}$ and $k : K \rightarrow \mathcal{E}$. The erosion \ominus of f by k , $f \ominus k$: $F \ominus K \rightarrow \mathcal{E}$ is defined by

$$f \ominus k = T\{Uf\} \ominus U[k] \quad (11.40)$$

Erosion is illustrated in Figure 11.17. To decrease computational complexity, the actual computations are performed in another way as the minimum of a set of differences (notice the similarity to correlation):

$$(f \ominus k)(x) = \min_{z \in K} \{f(x + z) - k(z)\} \quad (11.41)$$



Figure 11.18: Morphological pre-processing: (a) cells in a microscopic image corrupted by noise; (b) eroded image; (c) dilation of (b), the noise has disappeared; (d) reconstructed cells. Courtesy P. Kodl, Rockwell Automation Research Center, Prague, Czech Republic.

We illustrate morphological pre-processing on a microscopic image of cells corrupted by noise in Figure 11.18a; the aim is to reduce noise and locate individual cells. Figure 11.18b shows erosion of the original image, and Figure 11.18c illustrates dilation of the original image. A 3×3 structuring element was used in both cases—notice that the noise has been considerably reduced. The individual cells can be located by the reconstruction operation (to be explained in Section 11.5.4). The original image is used as a mask and the dilated image in Figure 11.18c is an input for reconstruction. The result is shown in image 11.18d, in which the black spots depict the cells.

11.4.2 Umbral homeomorphism theorem, properties of erosion and dilation, opening and closing

The top surface always inverts the umbral operation; i.e., the top surface is a left inverse of the umbral, $T[Uf] = f$. However, the umbral is not an inverse of the top surface. The strongest conclusion that can be deduced is that the umbral of the top surface of a point set A contains A (recall Figure 11.13).

The notion of top surface and umbral provides an intuitive relation between gray-scale and binary morphology. The umbral homeomorphism theorem states that the umbral operation is a homeomorphism from gray-scale morphology to binary morphology. Let $F, K \subseteq$



Figure 11.17: 1D example of gray-scale erosion. The umbras of 1D function f and structuring element k are eroded first, $Uf \ominus U[k]$. The top surface of this eroded set gives the result, $f \ominus k = T[Uf] \ominus U[k]$.

$$\begin{array}{c} Uff \ominus U[k] \\ \hline \text{---} \\ T[Uff \ominus U[k]] = f \ominus k \end{array}$$

$\mathcal{E}^{n-1}, f : F \rightarrow \mathcal{E}$, and $k : K \rightarrow \mathcal{E}$. Then

$$\begin{aligned} (a) \quad U[f \oplus k] &= U[f] \oplus U[k] \\ (b) \quad U[f \ominus k] &= U[f] \ominus U[k] \end{aligned} \quad (11.42)$$

(the proof may be found elsewhere [Haralick and Shapiro 92]). The umbra homeomorphism is used for deriving properties of gray-scale operations. The operation is expressed in terms of umbra and top surface, then transformed to binary sets using the umbra homeomorphism property, and finally transformed back using the definitions of gray-scale dilation and erosion. Using this idea, properties already known from binary morphology can be derived, e.g., commutativity of dilation, the chain rule that permits decomposition of large structural elements into successive operations with smaller ones, duality between erosion and dilation.

Gray-scale opening and closing is defined in the same way as in the binary morphology. **Gray-scale opening** is defined as $f \circ k = (f \ominus k) \oplus k$. Similarly, **gray-scale closing** $f \bullet k = (f \oplus k) \ominus k$. The duality between opening and closing is expressed as (recall that mean means the transpose, i.e., symmetric set with regards to origin of co-ordinates)

$$-(f \circ k)(x) = [(-f) \bullet \check{k}](x) \quad (11.43)$$

There is a simple geometric interpretation of gray-scale opening; see [Haralick and Shapiro 92] for derivation and details. The opening of f by structuring element k can be interpreted as sliding k on the landscape f . The position of all highest points reached by some part of k during the slide gives the opening, and a similar interpretation exists for erosion.

Gray-scale opening and closing is often used in applications to extract parts of a gray-scale image with given shape and gray-scale structure.

11.4.3 Top hat transformation

The top hat transformation is used as a simple tool for segmenting objects in gray-scale images that differ in brightness from background, even when the background is of uneven gray-scale. The top hat transform is superseded by the watershed segmentation (to be described in Section 11.7.3) for more complicated backgrounds.

Assume a gray-level image X and a structuring element K . The residue of opening as compared to original image $X \setminus (X \circ K)$ constitutes a new useful operation called a top hat transformation [Meyer 78].

The top hat transformation is a good tool for extracting light objects (or, conversely, dark ones, of course) on a dark (or light) but slowly changing background. Those parts of the image that cannot fit into structuring element K are removed by opening. Subtracting the opened image from the original provides an image where removed objects stand out clearly. The actual segmentation can be performed by simple thresholding. The concept is illustrated for the 1D case in Figure 11.19, where we can see the origin of the transformation name. If an image were a hat, the transformation would extract only the top of it, provided that the structuring element is larger than the hole in the hat.

An example from visual industrial inspection provides a practical application of gray-level morphology and the top hat transformation. A factory producing glass capillaries for mercury maximal thermometers had the following problem: The thin glass tube should be narrowed

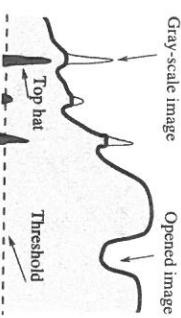


Figure 11.19: The top hat transform permits the extraction of light objects from an uneven background.

in one particular place to prevent mercury falling back when the temperature decreases from the maximal value. This is done by using a narrow gas flame and low pressure in the capillary. The capillary is illuminated by a collimated light beam—when the capillary wall collapses due to heat and low pressure, an instant specular reflection is observed and serves as a trigger to cover the gas flame. Originally the machine was controlled by a human operator who looked at the tube image projected optically on the screen: the gas flame was covered when the specular reflection was observed. This task had to be automated and the trigger signal learned from a digitized image. The specular reflection is detected by a morphological procedure—see Figure 11.20.

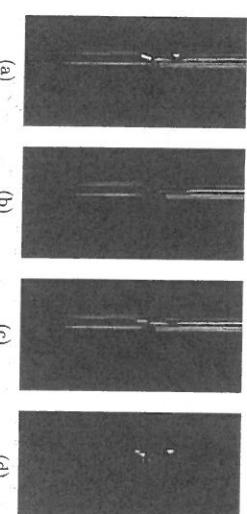


Figure 11.20: An industrial example of gray-scale opening and top hat segmentation, i.e., image-based control of glass tube narrowing by gas flame: (a) original image of the glass tube, 512×256 pixels; (b) erosion by a one-pixel-wide vertical structuring element 20 pixels long; (c) opening with the same element; (d) final specular reflection segmentation by the top hat transformation. Courtesy V. Smutny, R. Šára, CTU Prague, P. Kodl, Rockwell Automation Research Center, Prague, Czech Republic.

11.5 Skeletons and object marking

11.5.1 Homotopic transformations

Topological properties are associated with continuity (Section 2.3.1), and mathematical morphology can be used to study such properties of objects in images. There is an interesting group among morphological transformations called homotopic transformations [Serra 82].

A transformation is homotopic if it does not change the continuity relation between regions and holes in the image. This relation is expressed by the homotopic tree; its root corresponds to the background of the image; first-level branches correspond to the objects (regions), second-level branches match holes within objects, etc.

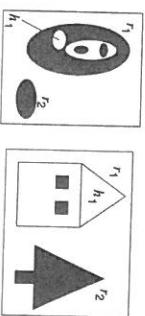


Figure 11.21: The same homotopic tree for two different images.

Figure 11.21 shows an example of a homotopic tree in which there are two different images with the homotopic tree below them. On the left side are some biological cells and on the right side a house and a spruce tree; both images have the same homotopic tree. Its root b corresponds to the background, node r_1 matches the larger cell (the outline of the house), and node r_2 matches the smaller cell (the spruce tree). Node h_1 corresponds to the empty hole in the cell r_1 (the hole inside the roof of the house)—the other correspondences to nodes should now be clear. A transformation is homotopic if it does not change the homotopic tree.

11.5.2 Skeleton, maximal ball

It is sometimes advantageous to convert an object to an archetypical stick figure called a skeleton (also considered in Section 6.3.4). We shall explain this in the context of 2D Euclidean space first, which is more illustrative than on the digital grid that we shall consider later.

The idea of skeleton was introduced by Blum under the name medial axis transform [Blum 67] and illustrated on the following ‘grassfire’ scenario. Assume a region (point set) $X \subset \mathcal{R}^2$: A grassfire is lit on the entire region boundary at the same instant, and propagates towards the region interior with constant speed. The skeleton $S(X)$ is the set of points where two or more firefronts meet; see Figure 11.22.

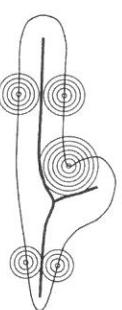


Figure 11.22: Skeleton as points where two or more wavefronts of grassfire meet.

A more formal definition of skeleton is based on the concept of maximal ball. A ball $B(p, r)$ with center p and radius r , $r \geq 0$, is the set of points with distances d from the center less than or equal to r .

The ball B included in a set X is said to be maximal if and only if there is no larger ball included in X that contains B , i.e., each ball B' , $B \subseteq B' \subseteq X \Rightarrow B' = B$. Balls and maximal balls are illustrated in Figure 11.23.

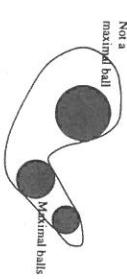


Figure 11.23: Ball and maximal balls in Euclidean plane.

The distance metric d that is used depends on the grid and definition of connectivity. Unit balls in a plane (i.e., unit disks) are shown in Figure 11.24.

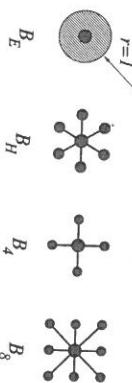


Figure 11.24: Unit-size disk for different distances, from left side: Euclidean distance, 6-, 4-, and 8-connectivity, respectively.

The plane \mathcal{R}^2 with the usual Euclidean distance gives the ball B_E . Three distances and balls are often defined in the discrete plane \mathbb{Z}^2 . If a hexagonal grid and 6-connectivity is used, the hexagonal ball B_H is obtained. If the support is a square grid, two unit balls are possible: B_4 for 4-connectivity and B_8 for 8-connectivity.

The skeleton by maximal balls $S(X)$ of a set $X \subset \mathbb{Z}^2$ is the set of centers p of maximal balls:

$$S(X) = \{p \in X : \exists r \geq 0, B(p, r) \text{ is a maximal ball of } X\}$$

This definition of skeleton has an intuitive meaning in the Euclidean plane. The skeleton of a disk reduces to its center, the skeleton of a stripe with rounded endings is a unit thickness

line at its center; etc.

Figure 11.25 shows several objects together with their skeletons—a rectangle, two touching balls, and a ring. The properties of the (Euclidean) skeleton can be seen here—in particular, the skeleton of two adjacent circles consists of two distinct points instead of a straight line joining these two points, as might be intuitively expected.

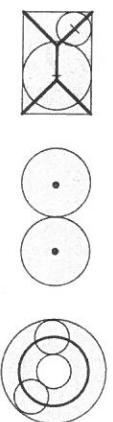


Figure 11.25. *Skeletons of rectangle, two touching balls, and a ring.*

The skeleton by maximal balls has two unfortunate properties in practical applications. First, it does not necessarily preserve the homotopy (connectivity) of the original set; and second, some of the skeleton lines may be wider than one pixel in the discrete plane. We shall see later that the skeleton is often substituted by sequential homotopic thinning that does not have these two properties.

Dilation can be used in any of the three discrete connectivities to create balls of varying radii. Let nB be the ball of radius n :

$$nB = B \oplus B \oplus \dots \oplus B \quad (11.44)$$

The skeleton by maximal balls can be obtained as the union of the residues of opening of the set X at all scales [Serra 82]:

$$S(X) = \bigcup_{n=0}^{\infty} [(X \ominus nB) \setminus (X \ominus nB) \circ B] \quad (11.45)$$

The trouble with this is that the resulting skeleton is completely disconnected and this property is not useful in many applications. Thus homotopic skeletons that preserve connectivity are often preferred. We present an intuitive homotopic skeletonization algorithm based on consecutive erosions (thinning) in Section 11.5.3.

11.5.3 Thinning, thickening, and homotopic skeleton

One application of the hit-or-miss transformation (Section 11.3.3) is thinning and thickening of point sets. For an image X and a composite structuring element $B = (B_1, B_2)$ (notice that B here is not a ball), *thinning* is defined as

$$X \oslash B = X \setminus (X \otimes B) \quad (11.46)$$

and *thickening* is defined as

$$X \odot B = X \cup (X \otimes B) \quad (11.47)$$

When thinning, a part of the boundary of the object is subtracted from it by the set difference operation. When thickening, a part of the boundary of the background is added to the object. Thinning and thickening are dual transformations:

$$(X \odot B)^c = X^c \oslash B \quad B = (B_2, B_1) \quad (11.48)$$

Thinning and thickening transformations are very often used sequentially. Let $\{B^{(1)}, B^{(2)}, B^{(3)}, \dots, B^{(n)}\}$ denote a sequence of composite structuring elements $B^{(i)} = (B_{i1}, B_{i2})$. Sequential thinning can then be expressed as a sequence of n structuring elements for square rasters,

$$X \oslash (B^{(i)}) = (((X \odot B^{(1)}) \oslash B^{(2)}) \dots \oslash B^{(n)}) \quad (11.49)$$

and sequential thickening as

$$X \odot \{B^{(i)}\} = (((X \odot B^{(1)}) \odot B^{(2)}) \dots \odot B^{(n)}) \quad (11.50)$$

There are several sequences of structuring elements $\{B^{(i)}\}$ that are useful in practice. Many of them are given by a permissible rotation of a structuring element in the appropriate digital raster (e.g., hexagonal, square, or octagonal). These sequences, sometimes called the Golay alphabet [Golay 69], are summarized for the hexagonal raster in [Serra 82]. We shall present structuring elements from which the other rotations can easily be derived.

A composite structuring element can be expressed by a single matrix only. A value of one in it means that this element belongs to B_1 (it is a subset of objects in the hit-or-miss transformation), and a value zero belongs to B_2 and is a subset of the background. An asterisk * in the matrix denotes an element that is not used in the matching process, so its value is not significant.

Thinning and thickening sequential transformations converge to some image—the number of iterations needed depends on the objects in the image and the structuring element used. If two successive images in the sequence are identical, the thinning (or thickening) is stopped.

Sequential thinning by structuring element L

This sequential thinning is quite important, as it serves as the homotopic substitute of the skeleton, the final thinned image consists only of lines of width one and isolated points. The structuring element L from the Golay alphabet is given by

$$L_1 = \begin{bmatrix} 0 & 0 & 0 \\ * & 1 & * \\ 1 & 1 & 1 \end{bmatrix} \quad L_2 = \begin{bmatrix} * & 0 & 0 \\ 1 & 1 & 0 \\ * & 1 & * \end{bmatrix} \dots \quad (11.51)$$

(The other six elements are given by rotation). Figure 11.26 shows the result of thinning with the structuring element L , after five iterations to illustrate an intermediate result, and Figure 11.27 shows the homotopic substitute of the skeleton when the idempotency was reached (in both cases, the original is shown on the left).

Sequential thinning by structuring element E

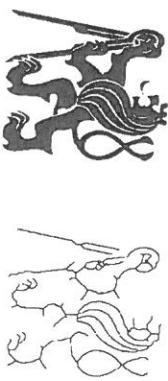
Assume that the homotopic substitute of the skeleton by element L has been found. The skeleton is usually jagged, because of sharp points on the outline of the object, but it is possible to ‘smooth’ the skeleton by sequential thinning by structuring element E . Using n iterations, several points (whose number depends on n) from the lines of width one (and isolated points as well) are removed from free ends. If thinning by element E is performed until the image does not change, then only closed contours remain.

Figure 11.26: Sequential thinning using element L after five iterations.Figure 11.27: Homotopic substitute of the skeleton (element L).

The structuring element E from the Golay alphabet is given again by eight rotated masks,

$$E_1 = \begin{bmatrix} * & 1 & * \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_2 = \begin{bmatrix} 0 & * & * \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \dots \quad (11.52)$$

Figure 11.28 shows sequential thinning (five iterations) by the element E of the skeleton from Figure 11.27. Notice that lines have been shortened from their free ends.

Figure 11.28: Five iterations of sequential thinning by element E .

There are three other elements M , D , C in the Golay alphabet [Golay 69]. These are not much used in practice at present, and other morphological algorithms are used instead to find skeletons, convex hulls, and homotopic markers.

The computationally most efficient algorithm of which we are aware creates the connected skeleton as the minimal superset of the skeleton by maximal balls [Vincent 91]. Its performance is shown in Figure 11.29. The homotopy is preserved.

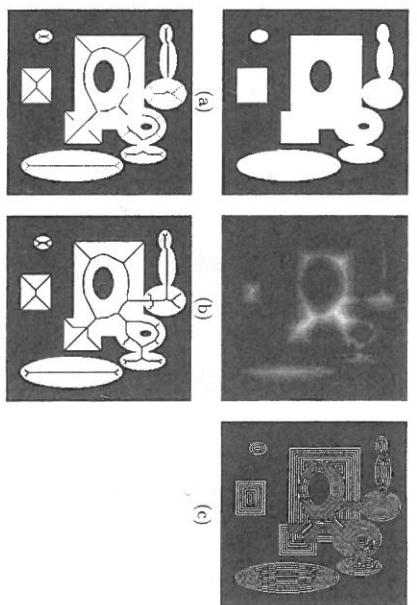


Figure 11.29: Performance of Vincent's quick skeleton by maximal balls algorithm: (a) original binary image; (b) distance function (to be explained later); (c) distance function visualized by contouring; (d) non-continuous skeleton by maximal balls; (e) final skeleton.

The skeleton can be applied to native 3D images as well, e.g., in the analysis of computer tomography images. Figure 11.30 illustrates examples of thinning of 3D point sets; parallel algorithms are available [Ma and Sonka 96].

11.5.4 Quench function, ultimate erosion

The binary point set X can be described using maximal balls B . Every point p of the skeleton $S(X)$ by maximal balls has an associated ball of radius $q_X(p)$; the term quench function is used for this association. An important property of $q_X(p)$ is that the quench function permits the reconstruction of the original set X completely as a union of its maximal balls B :

$$X = \bigcup_{p \in S(X)} [p + q_X(p)B] \quad (11.53)$$

This formula allows lossless compression of a binary image. Similar ideas are used for encoding documents in CCITT group 4 compression algorithms.

It is useful to distinguish several types of extrema, and use the analogy of a topographic view of images to make the explanation more intuitive. The **global maximum** is the pixel with highest value (lightest pixel, highest summit in the countryside); similarly, the **global minimum** corresponds to the deepest chasm in the countryside.

A pixel p of a gray-scale image is a local maximum if and only if for every neighboring pixel q of a pixel p , $I(p) \geq I(q)$. For example, the local maximum may mean that

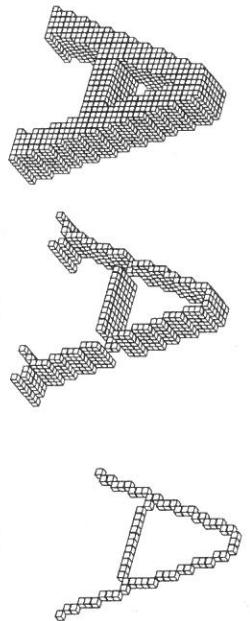


Figure 11.30: Morphological thinning in 3D: (a) original 3D data set, a character A ; (b) thinning performed in one direction; (c) resulting of one pixel thick skeleton obtained by thinning image (b) in second direction. Courtesy K. Palágyi, University of Szeged, Hungary.

the landscape around is studied in a small neighborhood of the current position (neighborhood in morphology is defined by the structuring element). If no ascent is seen within the neighborhood, the pixel is at a local maximum.

The regional maximum M of a digital gray-scale image (function) I is a connected set of pixels with an associated value h (plateau at altitude h), such that every neighboring pixel of M has strictly lower value than h . There is no connected path leading upwards from a regional maximum. Topographically, regional extrema correspond to geographic summits and hollows. If M is a regional maximum of I and $p \in M$, then p is a local maximum. The converse does not hold. Global, local, and regional maxima are illustrated for the 1D case in Figure 11.31.

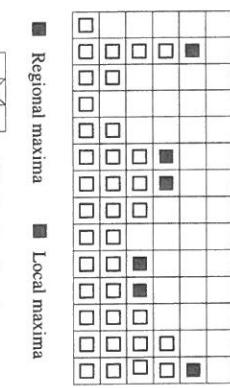


Figure 11.31: 1D illustration of global, regional, and local maxima.

The definition of various maxima allows us to analyze the quench function. The quench function is also useful to define ultimate erosion, which is often used as a marker of convex

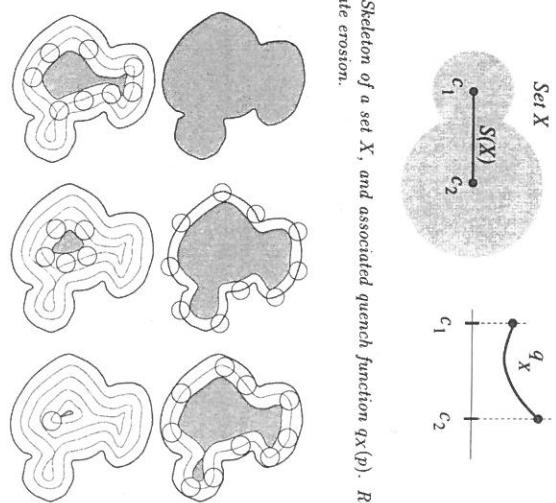
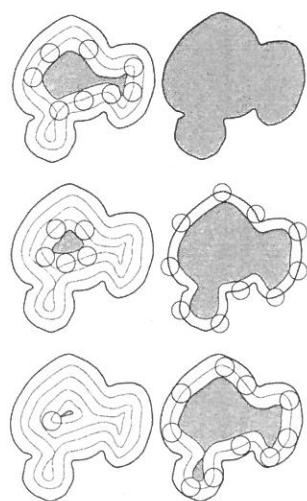


Figure 11.32: Skeleton of a set X , and associated quench function $q_X(p)$. Regional maxima give the ultimate erosion.

Figure 11.33: When successively eroded, the components are first separated from the rest and finally disappear from the image. The union of residues just before disappearance gives ultimate erosion.



objects in binary images. The ultimate erosion of a set X , denoted $\text{Ult}(X)$, is the set of regional maxima of the quench function. The natural markers are centers of the largest maximal balls. The trouble arises if the objects are overlapping—here ultimate erosion comes into play. Consider first the simplest case, in which the set X consists of two overlapping disks (see Figure 11.32). The skeleton is a line segment between the centers. The associated quench function has regional maxima that are located at the disk centers in this particular example. These maxima are called ultimate erosion and can be used as markers of overlapping objects. The ultimate erosion provides a tool that extracts one marker per object of a given shape, even if objects overlap. The remaining trouble is that some objects are still multiply marked.

Consider a binary image, a set X , consisting of three rounded overlapping components of different size. When iteratively eroding the set by a unit-size ball, the set is shrunk, then separates, and finally disappears, as illustrated in Figure 11.33. During the successive erosions, the residuals of connected components (just before they disappear) are stored. Their union is the ultimate erosion of the original set X (Figure 11.34).

- The skeleton by maximal balls of a set X corresponds to the set of local maxima of the distance function of X .

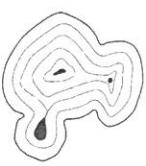


Figure 11.34: Ultimate erosion is the union of residual connected components before they disappear during erosions.

11.5.5 Ultimate erosion and distance functions

We seek to present this ultimate erosion procedure formally, and introduce the morphological reconstruction operator for this purpose. Assume two sets $A, B, B \subseteq A$. The reconstruction $\rho_A(B)$ of the set A from set B is the union of connected components of A with non-empty intersection with B (see Figure 11.35—notice that set A consists of two components). Notice that B may typically consist of markers that permit the reconstruction of the required part of the set A . Markers point to the pixel or small region that belongs to the object. Morphological reconstruction will be discussed in detail in Section 11.5.7.



Figure 11.35: Reconstruction $\rho_A(B)$ (in gray) of the set A from the set B . Notice that set A may consist of more than one connected component.

Let N be the set of integers. Ultimate erosion can be expressed by the formula

$$\text{Ult}(X) = \bigcup_{n \in N} ((X \ominus nB) \setminus \rho_{X \ominus nB}[X \ominus (n+1)B]) \quad (11.54)$$

There is a computationally effective ultimate erosion algorithm that uses the distance function (which is the core of several other quick morphological algorithms, as we shall see). The distance function $\text{dist}_X(p)$ associated with each pixel p of the set X is the size of the first erosion of X that does not contain p , i.e.,

$$\forall p \in X \quad \text{dist}_X(p) = \min\{n \in N, p \text{ not in } (X \ominus nB)\} \quad (11.55)$$

This behaves as one would expect: $\text{dist}_X(p)$ is the shortest distance between the pixel p and background X^C .

There are two direct applications of the distance function.

- The ultimate erosion of a set X corresponds to the union of the regional maxima of the distance function of X .

The last concept that will be introduced here is skeleton by influence zones, often abbreviated **SKIZ**. Let X be composed of n connected components $X_i, i = 1, \dots, n$. The influence zone $Z(X_i)$ consists of points which are closer to set X_i than to any other connected component of X .

$$Z(X_i) = \{p \in \mathcal{Z}^2, \forall i \neq j, d(p, X_i) \leq d(p, X_j)\} \quad (11.56)$$

The skeleton by influence zones denoted **SKIZ**(X) is the set of boundaries of influence zones $\{Z(X_i)\}$.

11.5.6 Geodesic transformations

Geodesic methods [Vincent 95] modify morphological transformations to operate only on some part of an image. For instance, if an object is to be reconstructed from a marker, say a nucleus of a cell, it is desirable to avoid growing from a marker outside the cell. Another important advantage of geodesic transformations is that the structuring element can vary at each pixel, according to the image.

The basic concept of geodesic methods in morphology is geodesic distance. The path between two points is constrained within some set. The term has its roots in an old discipline—geodesy—that measures distances on the Earth’s surface. Suppose that a traveler seeks the distance between London and Tokyo—the shortest distance passes *through* the Earth, but obviously the geodesic distance that is of interest to the traveler is constrained to the Earth’s surface.

The geodesic distance $d_X(x, y)$ is the shortest path between two points x, y while this path remains entirely contained in the set X . If there is no path connecting points x, y , we set the geodesic distance $d_X(x, y) = +\infty$. Geodesic distance is illustrated in Figure 11.36.

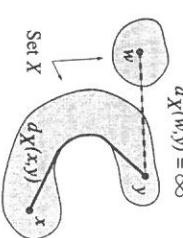


Figure 11.36: Geodesic distance $d_X(x, y)$.

The geodesic ball is the ball constrained by some set X . The geodesic ball $B_X(p, n)$ of center $p \in X$ and radius n is defined as

$$B_X(p, n) = \{p' \in X, d_X(p, p') \leq n\} \quad (11.57)$$

The existence of a geodesic ball permits dilation or erosion only within some subset of the image; this leads to definitions of geodetic dilations and erosions of a subset Y of X .

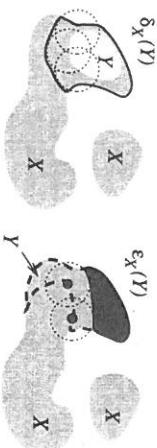


Figure 11.37: Illustration of geodesic dilation (left) and erosion (right) of the set Y inside the set X .

The geodesic dilation $\delta_X^{(n)}$ of size n of a set Y inside the set X is defined as

$$\delta_X^{(n)}(Y) = \bigcup_{p \in Y} B_X(p, n) = \{p' \in X : \exists p \in Y, d_X(p, p') \leq n\} \quad (11.58)$$

Similarly the dual operation of geodesic erosion $\epsilon_X^{(n)}(Y)$ of size n of a set Y inside the set X can be written as

$$\epsilon_X^{(n)}(Y) = \{p \in Y, B_X(p, n) \subseteq Y\} = \{p \in Y, \forall p' \in X \setminus Y, d_X(p, p') > n\} \quad (11.59)$$

Geodesic dilation and erosion are illustrated in Figure 11.37.

The outcome of a geodesic operation on a set $Y \subseteq X$ is always included within the set X . Regarding implementation, the simplest geodesic dilation of size 1 ($\delta_X^{(1)}$) of a set Y inside X is obtained as the intersection of the unit-size dilation of Y (with respect to the unit ball B) with the set X .

$$\delta_X^{(1)} = (Y \oplus B) \cap X \quad (11.60)$$

Larger geodesic dilations are obtained by iteratively composing unit dilations n times

$$\delta_X^{(n)} = \underbrace{\delta_X^{(1)}(\delta_X^{(1)}[\delta_X^{(1)}[\dots(\delta_X^{(1)})]]]}_{n \text{ times}} \quad (11.61)$$

The fast iterative way to calculate geodesic erosion is similar.

11.5.7 Morphological reconstruction

Assume that we want to reconstruct objects of a given shape from a binary image that was originally obtained by thresholding. All connected components in the input image constitute the set X . However, only some of the connected components were marked by markers that represent the set Y . This task and its desired result are shown in Figure 11.38.

Successive geodesic dilations of the set Y inside the set X enable the reconstruction of the connected components of X that were initially marked by Y . When dilating from the marker, it is impossible to intersect a connected component of X which did not initially contain a marker Y ; such components disappear.

Geodesic dilations terminate when all connected components set X previously marked by Y are reconstructed, i.e., idempotency is reached:

$$\forall n > n_0, \delta_X^{(n)}(Y) = \delta_X^{(n_0)}(Y) \quad (11.62)$$

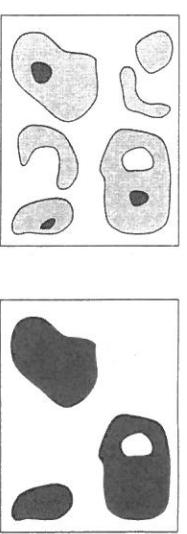


Figure 11.38: Reconstruction of X (shown in light gray) from markers Y (black). The reconstructed result is shown in black on the right side.

This operation is called reconstruction and denoted by $\rho_X(Y)$. Formally,

$$\rho_X(Y) = \lim_{n \rightarrow \infty} \delta_X^{(n)}(Y) \quad (11.63)$$

In some applications it is desirable that one connected component of X is marked by several markers Y . If it is not acceptable for the sets grown from various markers to become connected, the notion of influence zones can be generalized to geodesic influence zones of the connected components of set Y inside X . The idea is illustrated in Figure 11.39.

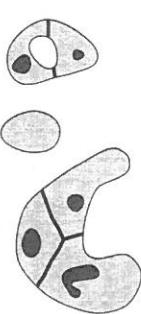


Figure 11.39: Geodesic SKIZ.

We are now ready to generalize the reconstruction to gray-scale images; this requires the extension of geodesy to gray-scale images. The core of the extension is the statement (which is valid for discrete images) that any increasing transformation defined for binary images can be extended to gray-level images [Serra 82]. By this transformation we mean a transformation Ψ such that

$$\forall X, Y \subset \mathcal{Z}^2, Y \subseteq X \Rightarrow \Psi(Y) \subseteq \Psi(X) \quad (11.64)$$

The generalization of transformation Ψ is achieved by viewing a gray-level image I as a stack of binary images obtained by successive thresholding—this is called the threshold decomposition of image I [Maragos and Ziff 90]. Let D_I be the domain of the image I , and the gray values of image I be in $\{0, 1, \dots, N\}$. The thresholded images $T_k(I)$ are

$$T_k(I) = \{p \in D_I, I(p) \geq k\} \quad k = 0, \dots, N \quad (11.65)$$

The idea of threshold decomposition is illustrated in Figure 11.40.

Threshold-decomposed images $T_k(I)$ obey the inclusion relation

$$\forall k \in [1, N], T_k(I) \subseteq T_{k-1}(I) \quad (11.66)$$

11.6 Granulometry

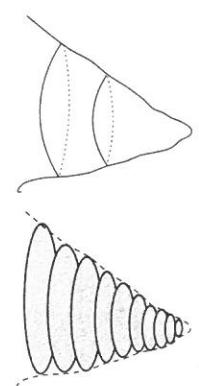


Figure 11.40: Threshold decomposition of a gray-scale image.

Consider the increasing transformation Ψ applied to each threshold-decomposed image; their inclusion relationship is kept. The transformation Ψ can be extended to gray-scale images using the following **threshold decomposition principle**:

$$\forall p \in D_1, \dots, \Psi(I)(p) = \max\{k \in [0, \dots, N], p \in \Psi(T_k(I))\} \quad (11.67)$$

Returning to the reconstruction transformation, the binary geodesic reconstruction ρ is an increasing transformation, as it satisfies

$$Y_1 \subseteq Y_2, X_1 \subseteq X_2, Y_1 \subseteq X_1, Y_2 \subseteq X_2, \implies \rho_{X_1}(Y_1) \subseteq \rho_{X_2}(Y_2) \quad (11.68)$$

We are ready to generalize binary reconstruction to **gray-level reconstruction** applying the threshold decomposition principle [equation (11.67)]. Let J, I be two gray-scale images defined on the same domain D , with gray-level values from the discrete interval $[0, 1, \dots, N]$. If, for each pixel $p \in D$, $J(p) \leq I(p)$, the gray-scale reconstruction $\rho_I(J)$ of image I from image J is given by

$$\forall p \in D, \rho_I(J)(p) = \max\{k \in [0, N], p \in \rho_{T_k}(J)\} \quad (11.69)$$

Recall that binary reconstruction grows those connected components of the mask which are marked. The gray-scale reconstruction extracts peaks of the mask I that are marked by J (see Figure 11.41).

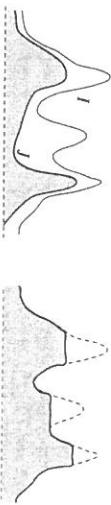


Figure 11.41: Gray-scale reconstruction of mask I from marker J .

The duality between dilation and erosion permits the expression of gray-scale reconstruction using erosion.

Granulometry was introduced by stereologists (mathematicians attempting to understand 3D shape from cross sections)—the name comes from the Latin *granulum*, meaning grain. Mathieu [Mathieu 67] used it as a tool for studying porous materials, where the distribution of pore sizes was quantified by a sequence of openings of increasing size. Currently, granulometry is an important tool of mathematical morphology, particularly in material science and biology applications. The main advantage is that granulometry permits the extraction of shape information without a priori segmentation.

Consider first a **sieving analysis** analogy; assume that the input is a heap of stones (or granules) of different sizes. The task is to analyze how many stones in the heap fit into several size classes. Such a task is solved by sieving using several sieves with increasing sizes of holes in the mesh. The result of analysis is a discrete function; on its horizontal axis are increasing sizes of stones and on its vertical axis the numbers of stones of that size. In morphological granulometry, this function is called a **granulometric spectrum** or **granulometric curve**.

In binary morphology, the task is to calculate a granulometric curve where the independent variable is the size of objects in the image. The value of the granulometric curve is the number of objects of given size in the image. The most common approach is that sieves with increasing hole sizes (as in the example) are replaced by a sequence of openings with structural elements of increasing size.

Granulometry plays a very significant role in mathematical morphology that is analogous to the role of frequency analysis in image processing or signal analysis. Recall that frequency analysis expands the signal into a linear combination of harmonic signals of growing frequency. The frequency spectrum provides the contribution of individual harmonic signals—it is clear that the granulometric curve (spectrum) is analogous to a frequency spectrum.

Let $\Psi = (\psi_\lambda)$, $\lambda \geq 0$, be a family of transformations depending on a parameter λ . This family constitutes a **granulometry** if and only if the following properties of the transformation ψ hold:

$$\forall \lambda \geq 0 \quad \psi_\lambda \text{ is increasing}$$

$$\forall \lambda \geq 0, \mu \geq 0 \quad \psi_\lambda \psi_\mu = \psi_{\mu \lambda} \quad \psi_{\max(\lambda, \mu)} \quad (11.70)$$

The consequence of property (11.70) is that for every $\lambda \geq 0$ the transformation ϕ_λ is idempotent, $(\phi_\lambda)_*, \lambda \geq 0$ is a decreasing family of openings (more precisely, algebraic openings [Serra 82] that generalize the notion of opening presented earlier). It can be shown that for any convex structuring element B , the family of openings with respect to $\lambda B = \{\lambda b, b \in B\}, \lambda \geq 0$, constitutes a granulometry [Mathieu 75].

Consider more intuitive granulometry acting on discrete binary images (i.e., sets). Here the granulometry is a sequence of openings ϕ_n indexed by an integer $n \geq 0$ —each opening result is smaller than the previous one. Recall the analogy with sieving analysis; each opening, which corresponds to one sieve mesh size, removes from the image more than the previous one. Finally, the empty set is reached. Each sieving step is characterized by some measure $m(X)$ of the set (image) X (e.g., number of pixels in a 2D image, or volume in 3D). The rate at which the set is sieved characterizes the set. The pattern spectrum provides such a characteristic.

The pattern spectrum, also called **granulometric curve**, of a set X with respect to the granulometry $\Psi = \psi_n$, $n \geq 0$ is the mapping

$$PS_\Psi(X)(n) = m[\psi_n(X)] - m[\psi_{n-1}(X)] \quad \forall n > 0 \quad (11.71)$$

The sequence of openings $\Psi(X)$, $n \geq 0$ is a decreasing sequence of sets, i.e., $[\psi_0(X) \supseteq \psi_1(X) \supseteq \psi_2(X) \supseteq \dots]$. The granulometry and granulometric curve can be used.

Suppose that the granulometric analysis with family of openings needs to be computed for a binary input image. The binary input image is converted into a gray-level image using a granulometry function $G_\Psi(X)$, and the pattern spectrum PS_Ψ is calculated as a histogram of the granulometry function.

The granulometry function $G_\Psi(X)$ of a binary image X from granulometry $\Psi = (\psi_n)$, $n \geq 0$, maps each pixel $x \in X$ to the size of the first n such that $x \notin \psi_n(X)$:

$$x \in X, G_\Psi(X)(x) = \min\{n > 0, x \notin \psi_n(X)\} \quad (11.72)$$

The pattern spectrum PS_Ψ of a binary image X for granulometry $\Psi = (\psi_n)$, $n \geq 0$, can be computed from the granulometry function $G_\Psi(X)$ as its histogram:

$$\forall n > 0, PS_\Psi(X)(n) = \text{card}\{p, G_\Psi(X)(p) = n\} \quad (11.73)$$

(where ‘card’ denotes cardinality). An example of granulometry is given in Figure 11.42. The input binary image with circles of different radii is shown in Figure 11.42a; Figure 11.42b shows one of the openings with a square structuring element. Figure 11.42c illustrates the granulometric power spectrum. At a coarse scale, three most significant signals in the power spectrum indicate three prevalent sizes of object. The less significant signals on the left side are caused by the artifacts that occur due to discretization. The Euclidean circles have to be replaced by digital entities (squares).

We see in this example that granulometries extract size information without the need to identify (segment) each object *a priori*. In applications, this is used for shape description, feature extraction, texture classification, and removal of noise introduced by image borders.

Until recently, granulometric analysis using a family of openings was too slow to be practically useful, but recent developments have made granulometries quick and useful: the reader interested in implementation may consult [Haralick et al. 95, Vincent 95]. For binary images, the basic idea towards speed-up is to use linear structuring elements for openings and more complex 2D ones derived from it, such as cross, square, or diamond (see Figure 11.43). The next source of computational saving is the fact that some 2D structuring elements can be decomposed as Minkowski addition of two 1D structuring elements. For example, the square structuring element can be expressed as Minkowski addition of horizontal and vertical lines.

Gray-scale granulometric analysis is another recent development that permits the extraction of size information directly from gray-level images. The interested reader should consult [Vincent 94].

11.7 Morphological segmentation and watersheds

11.7.1 Particles segmentation, marking, and watersheds

The concept “segmentation” commonly means finding objects of interest in the image. Mathematical morphology helps mainly to segment images of texture or images of particles—here

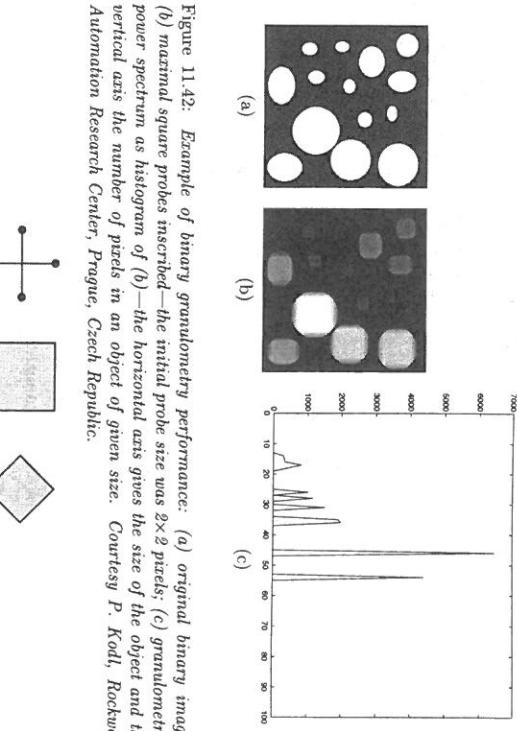


Figure 11.43: Structural elements used for fast binary granulometry are derived from line structuring elements, e.g., cross, square, and diamond.

we consider particle segmentation in which the input image can be either binary or gray-scale. In the binary case, the task is to segment overlapping particles; in the gray-scale case, the segmentation is the same as object contour extraction. The explanation here is inspired by Vincent’s view [Vincent 95], which is intuitive and easy to understand.

Morphological particle segmentation is performed in two basic steps: (1) location of particle markers, and (2) watersheds used for particle reconstruction. The latter is explained later in this section.

Marker extraction resembles human behavior when one is asked to indicate objects; the person just points to objects and does not outline boundaries. The marker of an object or set X is a set M that is included in X . Markers have the same homotopy as the set X , and are typically located in the central part of the object (particle).

A robust marker-finding technique will need to know the nature of the objects sought, and thus application-specific knowledge should be used. Often combinations of non-morphological and morphological approaches are used. Moreover, object marking is in many cases left to the user, who marks objects manually on the screen. Typically, software for analysis of microscopic images has user-friendly tools for manual or semi-automatic marking. When the objects are marked, they can be grown from the markers, e.g., using the water-

shed transformation (Section 11.7.3), which is motivated by the topographic view of images. Consider the analogy of a landscape and rain; water will find the steepest descent path until it reaches some lake or sea. We already know that lakes and seas correspond to regional minima. The landscape can be entirely partitioned into regions which attract water to a particular sea or lake—these will be called **catchment basins**. These regions are **influence zones** of the regional minima in the image. **Watersheds**, also called **watershed lines**, separate catchment basins. Watersheds and catchment basins are illustrated in Figure 11.44.

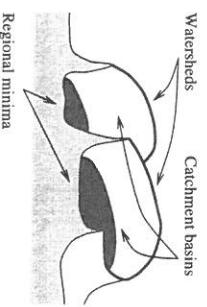


Figure 11.44: Illustration of catchment basins and watersheds in a 3D landscape view.

11.7.2 Binary morphological segmentation

If the task is to find objects that differ in brightness from an uneven background, the top hat transformation (Section 11.5.4) is a simple solution. The top hat approach just finds peaks in the image function that differ from the local background. The gray-level shape of the peaks does not play any role, but the shape of the structuring element does. Watershed segmentation takes into account both sources of information and supersedes the top hat method.

Morphological segmentation in binary images aims to find regions corresponding to individual overlapping objects (typically particles), and most of the tools for performing this task have already been explained. Each particle is marked first—ultimate erosion may be used for this purpose (Section 11.5.4), or markers may be placed manually. The next task is to grow objects from the markers provided they are kept within the limits of the original set and parts of objects are not joined when they come close to each other.

The oldest technique for this purpose is called **conditional dilation**. Ordinary dilation is used for growing, and the result is constrained by the two conditions (remain in the original set, and do not join particles).

Geodesic reconstruction (Section 11.5.7) is more sophisticated and performs much faster than conditional dilation. The structuring element adapts according to the neighborhood of the processed pixel.

Geodesic influence zones (Section 11.5.7) are sometimes used for segmenting particles. Figure 11.45 shows that the result can differ from our intuitive expectation.

The best solution is the **watershed transformation**. Only the basic idea will be described here—the reader interested in theory and fast implementation is referred to Section 11.7.3 and [Bleau et al. 92, Vincent 93, Vincent 95]. The original binary image is

Figure 11.45: Segmentation by geodesic influence zones (SKIZ) need not lead to correct results.

converted into gray-scale using the negative distance transform— dist [equation (11.55)]. If a drop of water falls onto a topographic surface of the $-\text{dist}$ image, it will follow the steepest slope towards a regional minimum. This idea is illustrated in Figure 11.46.

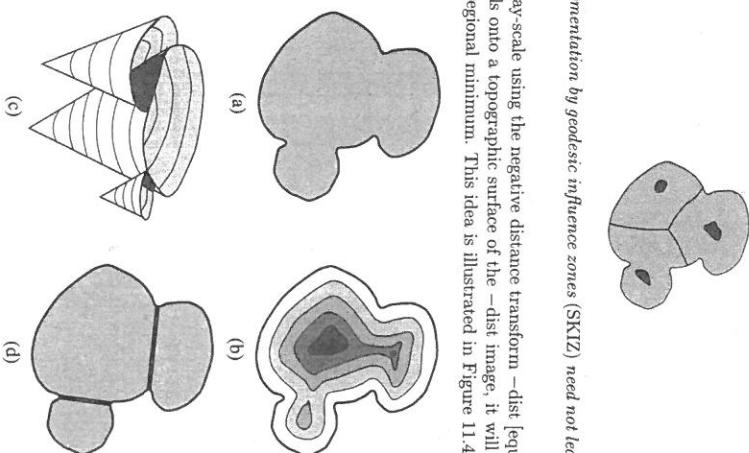


Figure 11.46: Segmentation of binary particles: (a) input binary image, (b) gray-scale image created from (a) using the $-\text{dist}$ function, (c) topographic notion of the catchment basin, (d) correctly segmented particles using watersheds of image (b).

Application of watershed particle segmentation is shown in Figure 11.47. We selected an image of a few touching particles as an input Figure 11.47a. The distance function calculated from the background is visualized using contours in Figure 11.47b for better understanding. The regional maxima of the distance function serve as markers of the individual particles, see Figure 11.47c. The markers are dilated in Figure 11.47d. In preparation for watershed segmentation, the distance function is negated, and is shown together with the dilated markers in Figure 11.47e. The final result of particle separation is illustrated in Figure 11.47f, where particle contours are shown.

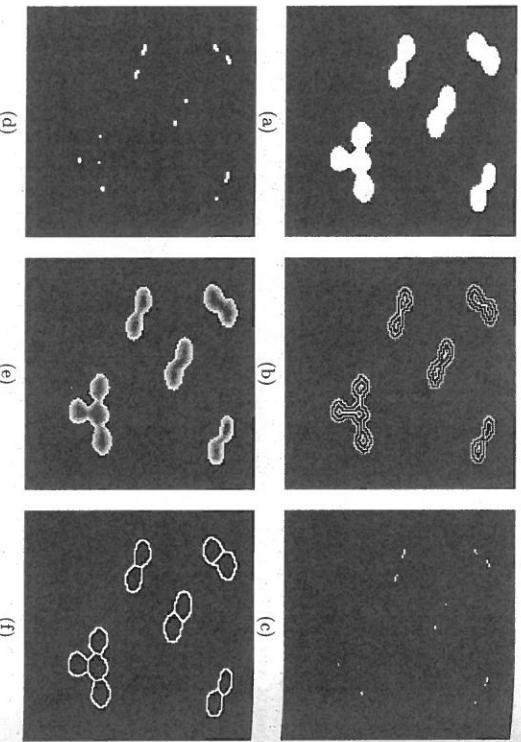


Figure 11.47: Particle segmentation by watershed: (a) original binary image; (b) distance function visualized using contours; (c) regional maxima of the distance function used as particle markers; (d) dilated markers; (e) inverse of the distance function with the markers superimposed; (f) resulting contours of particles obtained by watershed segmentation. Courtesy P. Kold, Rockwell Automation Research Center, Prague, Czech Republic.

11.7.3 Gray-scale segmentation, watersheds

The markers and watersheds method can also be applied to gray-scale segmentation. Watersheds are also used as crest-line extractors in gray-scale images. The contour of a region in a gray-level image corresponds to points in the image where gray-levels change most quickly—this is analogous to edge-based segmentation considered in Chapter 5. The watershed transformation is applied to the gradient magnitude image in this case (see Figure 11.48). There is a simple approximation to the gradient image used in mathematical morphology called Beucher's gradient [Serra 82], calculated as the algebraic difference of unit-size dilation and unit-size erosion of the input image X .

$$\text{grad}(X) = (X \oplus B) - (X \ominus B) \quad (11.74)$$

The main problem with segmentation via gradient images without markers is **oversegmentation**, meaning that the image is partitioned into too many regions (Figure 11.47c). Some techniques to limit oversegmentation in watershed segmentation are given in [Vincent

Figure 11.48: Segmentation in gray-scale images using gradient magnitude.

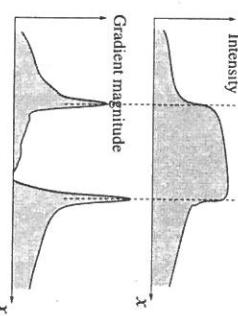
[93]. The watershed segmentation methods with markers do not suffer from oversegmentation, of course.

An example from ophthalmology will illustrate the application of watershed segmentation. The input image shows a microscopic picture of part of a human retina, Figure 11.49a—the task is to segment individual cells on the retina. The markers watershed paradigm was followed, with markers being found using a carefully tuned Gaussian filter (see Figure 11.49b). The final result with the outlined contours of the cells is in Figure 11.49c.

11.8 Summary

- Mathematical morphology

- Mathematical morphology stresses the role of **shape** in image pre-processing, segmentation, and object description. It constitutes a set of tools that have a solid mathematical background and lead to fast algorithms. The basic entity is a **point set**. Morphology operates using transformations that are described using operators in a relatively simple non-linear algebra. Mathematical morphology constitutes a counterpart to traditional signal processing based on linear operators (such as convolution).
- Mathematical morphology is usually divided into **binary mathematical morphology** which operates on binary images (2D point sets), and **gray-level mathematical morphology** which acts on gray-level images (3D point sets).
- Morphological operations
 - In images, morphological operations are relations of two sets. One is an image and the second a small probe, called a structuring element, that systematically traverses the image; its relation to the image in each position is stored in the output image.
 - Fundamental operations of mathematical morphology are **dilation** and **erosion**. Dilation expands an object to the closest pixels of the neighborhood. Erosion



11.9 Exercises

Short-answer questions

1. What is mathematical morphology?
2. What is a structuring element? What is its role in mathematical morphology?
3. Give the definition of erosion and dilation for binary images.
4. Is erosion a commutative operation?
5. What is the difference between the result of opening performed once and twice? What is idempotency?
6. Dilation and erosion for gray-level images is a generalization of the same notions in binary images. Describe how they are done.
7. What is the top hat transformation and when is it used?
8. What is the skeleton by maximal balls? What does the skeleton of two touching filled circles look like?
9. What does the homotopic skeleton of two touching balls look like?
10. Explain the role of ultimate erosion for marking particles.
11. What is a regional maximum?
12. What is geodesic distance? How is it used in mathematical morphology?
13. What is granulometry?
14. What is the relation between a granulometric curve (spectrum) and the Fourier spectrum?
15. What is a watershed? How is it used for morphological segmentation?

Problems

1. Prove that dilation is commutative and associative.
2. Consider a one-dimensional gray-scale image (signal). Draw a picture that demonstrates that gray-scale dilation fills narrow bays in the image.
3. Explain how the top hat transformation permits the segmentation of dark characters on a light background with varying intensity. Draw a picture of a one-dimensional cross section through the image.
4. Explain the calculation steps of the skeleton by maximal balls in Figure 11.29.
5. Draw a one-dimensional gray-level continuous function with local, regional, and global maxima (minima). Explain the role of regional extremes in mathematical morphology.
6. Name three application areas in which granulometry analysis may be used.
7. Explain the role of markers in morphological segmentation. Why would an attempt to perform watershed segmentation without markers lead to oversegmentation?

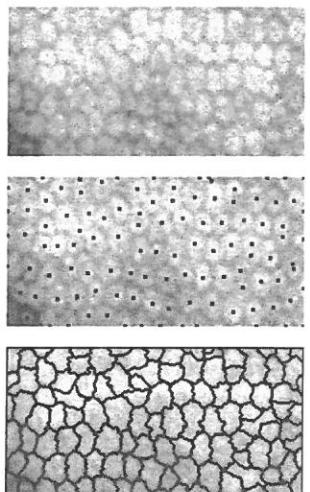


Figure 11.49: Watershed segmentation on the image of a human retina: (a) original gray-scale image; (b) dots are superimposed markers found by nonmorphological methods; (c) boundaries of retina cells found by watersheds from markers (b). Data and markers courtesy R. Šára, Czech Technical University, Prague, segmentation courtesy P. Kodl, Rockwell Automation Research Center Prague, Czech Republic.