Distance, granulometry, skeleton

Michel Couprie, Hugues Talbot

Chapter 10

Distance, granulometry, skeleton

In this chapter, we present a series of concepts and operators based on the notion of distance. As often with mathematical morphology, there exists more than one way to present ideas, that are simultaneously equivalent and complementary. Here, our problem is to present methods to characterize sets of points based on metric, geometry and topology considerations. An important concept is that of the skeleton, which is of fundamental importance in pattern recognition, and has many practical applications.

10.1. Skeletons

The notion of skeleton, which we will define shortly in more than one way, is very useful in many applications. Intuitively, the skeleton of an object (of a set of points) X is a set of "median" lines, constituted of points that are located at an equal distance from distinct areas of the border of X. Historically, a definition of the general concept was proposed by Listing in 1861 [LIS 61], under the name of "cyclomatic diagram". This diagram results from the contraction of a closed curve in the Euclidean plane, under topology preservation constraints. This notion is essentially a homology concept.

A more precise and geometric definition, that of the *median axis* was proposed by Blum almost exactly one hundred years later [BLU 61, BLU 67]. This definition requires the notion of *maximal ball*.

Chapter written by Michel COUPRIE and Hugues TALBOT.

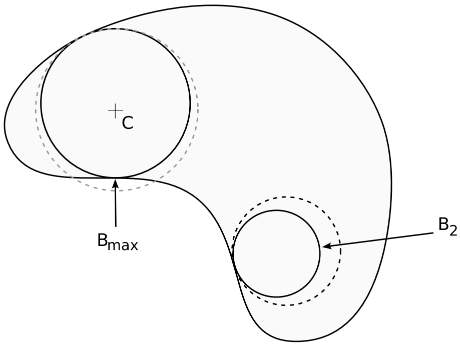


Figure 10.1. *Illustration of a maximal ball in the Euclidean case. The Ball B_{max} is maximal because no other ball can contain it and simultaneously be included in the set. In contrast, the ball B_2 is not maximal.*

10.1.1. *Maximal balls*

Maximal balls are a simple geometric concept that will allow us to introduce the notion of skeleton of a set of points.

DEFINITION 10.1 *(Ball).– The ball $B_r(x)$ of radius r and center x is the set of points y of $\mathbb{R}^n$ such that $d(x, y) < r$, where d is a distance (e.g. the Euclidean distance or the city-block distance)*

DEFINITION 10.2 *(Maximal ball).– A maximal ball B_r^{max} of radius r for set X is a ball included in X such that $\forall s > r$, if $B_r^{max} \subset B_s$ then $B_s \not\subset X$.*

This notion is illustrated on Fig 10.1. We remark that on this figure, where the set under study exhibit a regular border, the maximal ball B_{max} is tangent to the border of the set in two distinct points. This is common but not the general case: sometime there are more than two points, and even a non-countable number. The point C on this figure is the center of B_{max}. It is clearly located at the same distance of the two points of B_{max} that are tangent to the border of the set.

10.1.2. *Firefronts*

Almost simultaneously to the definition of Blum appear those of Calabi and Hartnett [CAL 68], as well as that of Montanari [MON 68], that use the notion of *firefront*.

DEFINITION 10.3 *(Firefront and quench function).– Suppose a fire is lit on the border of X and propagates toward the interior of X at a constant rate, in an isotropic manner. Assuming that a burnt point does not lit itself again, then the locus of points*

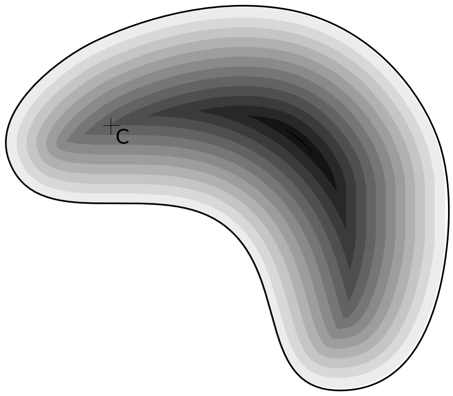


Figure 10.2. *Illustration of the concept of firefront. Fire propagates from the border of the object toward its interior, in an isotropic manner. The point C is a quench point and is a center of maximal ball.*

of X where several firefronts meet are called the quench points*. The function that at each quench point associates the time at which the point stopped burning is equivalent to the distance to the border of the set. This function is called the* quench function*.*

This definition is useful conceptually, because it forms the basis of many algorithmic ideas, in particular regarding skeletonization methods that use the idea of "eating up" at the border of the object.

PROPERTY 10.4.– *For a sufficiently regular compact set X [1] in the Euclidean space, points p are centers of maximal balls of X if and only if they are quench points.*

It is not immediately obvious that the two notions are equivalent, but it is not very difficult to prove. Intuitively, the meeting point C of several firefronts is located at the same distance of n distinct points $A_1, A_2, \ldots, A_n$ of the border of X, from where the firefronts originate, due to the constant propagation rate. Because of the isotropic propagation, the segments $[CA_i]$ all have the same length, are normal to the border of X and are radii of a ball B_{C,A_i}, centered in C. In addition, this ball is necessarily included in X, otherwise another distinct point from the border or X would have extinguished C before the firefronts coming from the points A_I. Finally, any ball that would include B_{C,A_I} and included in X must also be normal to ∂X on all points A_i. It is therefore necessarily centered in C and of radius $[CA_i]$, therefore the ball B_{C,A_i} is maximal.

1. Here we admit we need to define the notion of tangent disk to the border ∂X of X. As a regularity condition, we can for instance take that ∂X be everywhere twice differentiable.

Conversely, let B_C a maximal ball of X, centered in C. Since it is maximal, it intersects ∂X on at least two points A and B. Indeed, if we suppose that B_C does not intersect ∂X at any point, then by definition, we can include a circle strictly in X B_C, and so B_C is not maximal. This is still true if B_C intersects ∂X on a single point A, assuming all the following curves are tangent in A: the border of X, that of B_C and that of the circle. Finally, the two firefronts issuing from A and B meet in C, and so C is a quench point.

The definition of firefronts is useful as an illustration but not very rigorous, and linked to the continuous domain. It may be useful to consider the properties of the skeleton in this domain.

10.1.3. *Properties of the skeleton in the continuum*

With standard disclaimers, the preceding definitions apply to the continuous domain, in this case arbitrary dimension manifolds, equipped with a Riemannian metric. Here we shall only describe the $\mathbb{R}^n$ case with the Euclidean metric. We will consider as before the set X to be skeletonized as a compact set with a border everywhere twice differentiable. To avoid some limit cases, we shall consider that the interior of X, denoted $\mathring{X}$ is connected.

In this context, the two definitions of the skeleton given earlier are equivalent, and exhibit the following properties:

PROPERTY 10.5.– *Properties of the skeleton in the continuous domain*

– Non-continuity: a continuous deformation of the set can induce non-continuous changes in the associated skeleton. For instance, the euclidean skeleton of a disk is a single point : its center. However, any infinitesimal change on the border of the disk will result in at least one branch in the skeleton.

– Homotopy: the skeleton is homotopic to the original set.

– Negligibility: the Lebesgue measure of the skeleton is zero. This translates into the property that the skeleton does not contain any neighborhood.

– Invariance by isometric transform: applying any isometric transform (i.e: a translation or rotation) before or after the computation of the skeleton leads to the same result.

– Links with the distance transform:

- The function that associates any point of the skeleton with its distance to the border of the set, also known as the quench function *seen earlier can be used to reconstruct the original set perfectly. This is achieved simply by dilating each point of the skeleton with the ball of radius equal to the value of the quench function at that point.*

- The skeleton is centered in the object, in the sense that each point of the skeleton is located at equal distance from at least two distinct points from the border of the object.

- For each point belonging to the skeleton, the line of greatest ascent for the distance transform issuing from that point is in the direction of the skeleton. We say the upstream *of every skeleton point also belongs to the skeleton.*

- For every point of the skeleton, the slope of the upstream is strictly less than 1.

- The centers of maximal balls are special points of the distance transform: they form the ridge *point of this transform.*

These properties are explored in detail in [MAT 88a, MAT 88c]. Many are demonstrated in [RIV 87]. The non-continuity of the skeleton translates in practice into a high sensitivity to noise. Just one point added to the border of an object can induce the formation of a new branch of the skeleton. A small arbitrary change of X may involve a change of topology and radical changes in the appearance of the skeleton. In general we try to regulate these effects by filtering either on the shape of X, or on the skeleton itself [VIN 91a, ATT 95].

The homotopy of the skeleton is not guaranteed by all discrete skeletonization algorithms. Most of them however, try to ensure this important property, which is discussed in Section 10.6. Preservation of the homotopy of the skeleton is a property used in pattern recognition.

The fact that the skeleton is negligible in the continuum is sometimes reflected in need to be thin, *i.e.*, it can be erased by an erosion with the elementary ball for the connectivity of the grid used for computing it. Some algorithms attempt to produce a skeleton of thickness of one a pixel, which is not necessarily compatible with other properties of the skeleton, such as centering.

10.2. Skeletons in discrete spaces

We can not calculate, much less use in a practical way, the continuous skeleton, but we will seek, in the discrete domain, to reproduce all or part of its characteristics.

First, note that there exist skeletonization algorithms that start from the definition of the continuum, and formulate the problem in this area. For example, we can model a forest fire through an evolution of the contour by Partial Differential Equation (PDE), and find the skeleton points by detecting points of shocks, which are points in space where only a weak solution of the PDE exists, or maintaining the Center of Maximal Balls (CMB) as anchors for the evolution of the contour [LEY 91, SID 99, TOR 03]. Ultimately, despite its theoretical elegance, these methods are still a discrete skeleton due do the fact that the PDE used for the calculation are discretized. It is difficult to prove anything about the results because the solutions are only approximate. Finally things can become even more complicated in 3D and more.

We can not reproduce in a discrete setting all the properties of continuous skeletons, but we may instead choose a particular compromise as needed.

We will try to calculate, for a discrete object X, an object $S(X)$

– homotopic to the set X;

– thin: that is to say, a minimum number of pixel or voxel large, or an object erasable the elementary erosion;

– which contains the centers of discrete maximal balls;

– centered in the object X, that is to say that every point of $S(X)$ should be at equal distance, for the underlying grid, of two distinct points of the boundary of X.

The latter property is usually impossible to satisfy strictly. For example, for the 8-connectivity in the plane $\mathbb{Z}^2$, for a set X that consists of a 10×10 square, the centers of maximal balls, which are within this grid squares of size e $n \times n$ with n odd, are not located at equal distance from all points of the border of X.

In the following sections we present two approaches that allows to define the skeleton of a shape in a discrete space. The first approach, described in Section 10.3, relies on the granulometry concept, familiar in mathematical morphology. It has the advantage of inducing interesting generalizations based on such family of structuring elements non-symmetrical or not connected. The second approach involves using discrete distances to define the notions of ball and maximal ball relatively to a given object (Section 10.4). This latter approach allows the use of the exact Euclidean distance, which provides for the skeleton less sensitivity to the orientation of the object in space, a full rotation invariance being impossible to guarantee.

10.3. Granulometric families and skeletons

10.3.1. *Granulometric family*

We first start by defining a way to characterize the size of a set of points. We shall see later that this characterization is rich in consequences.

DEFINITION 10.6.– *We first define the concept of a* granulometric family *using the following steps:*

1) we define a family of erosions $\{\alpha_i\}$ *called* elementary *by convex structuring elements indexed by an integer* i*;*

2) we construct the family of structuring elements $\kappa_i = \alpha_i \dots \alpha_2\alpha_1$*;*

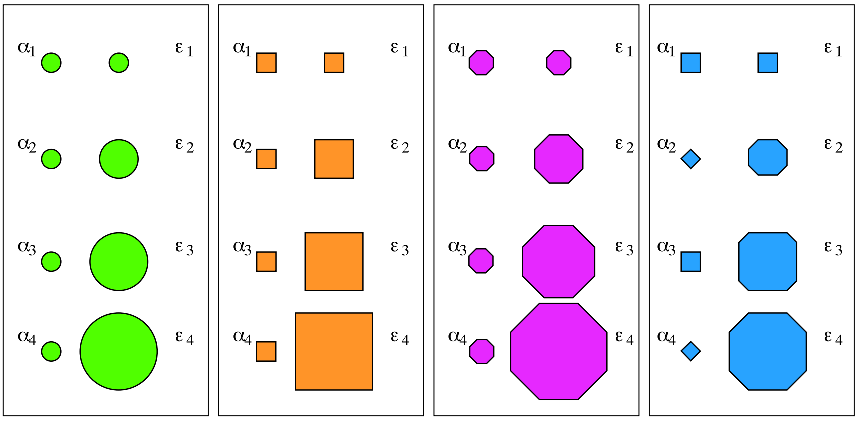


Figure 10.3. *Granulometric families that are homogeneous (a, b, c) and heterogeneous (d)*

3) these structuring elements define a granulometric *family* $\{\kappa_i\}$*;*

4) in most cases we use a homogeneous family, that is to say when all elementary erosions are identical: $\forall i$*,* $\alpha_i = \alpha_1$*, implying* $\kappa_i = (\alpha_1)^i$.

Heterogeneous families can be used to obtain a finer control over the granulometry, in particular, they allow the use of more isotropic structuring elements. These families are illustrated in Figure 10.3.

DEFINITION 10.7.– *We have the following property, by construction:* $\forall j \leq i, \gamma_{\kappa_i} \circ \gamma_{\kappa_j} = \gamma_{\kappa_i}$*. We say that each* κ_i *is open by all* κ_j *for* $j \leq i$*. This property is called "absorption property".*

10.3.2. *Applications of granulometries*

Granulometric families allows to define incremental operations with desirable properties.

– We can trivially extend the definition to grayscale.

– For example, binary or greyscale granulometries can be used to characterize texture or to obtain information about the image without segmentation.

– We give some examples of applications of granulometries in Chapter 1 and in Chapters 13 and 19.

10.3.3. *Ultimate eroded formula*

The *ultimate eroded* consists of specific points of the skeleton, and is defined as follows:

DEFINITION 10.8 *(Ultimate eroded).–*

$$U(X) = \bigcup_{i \in \mathbb{N}} U_i(X), \text{ where } U_i(X) = \varepsilon_{\kappa_i}(X) \setminus \delta^{\infty}_{\varepsilon_{\kappa_i}(X)}[\varepsilon_{\kappa_{i+1}}(X)] \tag{10.1}$$

Here, denotes $\delta^{\infty}_{A}(B)$ the reconstruction of the set B in the mask A, as introduced in Chapter 1. It uses the basic structuring element of the grid to make the operation (see below).

The ultimate eroded has the following properties:

– Increasingness

– Anti-extensivity

– Idempotence (it is thus an opening)

– All the U_i are disjoint

– If we dilate a point of U_i by the structuring element κ_i, the result is called a *maximal ball*.

The definition of a discrete ball is close to the one of the continuum: in the continuum we have a continuous family of balls.

The process of obtaining the ultimate eroded is illustrated in Figure 10.4. Each point of an ultimate eroded corresponds to a center of maximal ball, so convex as defined by the family of structuring elements α_i. The connected components of the ultimate erodeds identify and possibly separate approximately convex parts of a particle [SER 82]. The ultimate eroded of each component is the last step before his disappearance by successive erosions, hence its name.

10.3.4. *Lantuéjoul formula*

From its definition, it follows that the ultimate eroded is a subset of the skeleton. We can adapt the formula of ultimate eroded to produce a different definition of the skeleton. This formula is due to Christian Lantuéjoul [LAN 78, LAN 80b]:

DEFINITION 10.9 *(Lantuéjoul formula).– The skeleton $S(X)$ of a set X is the set of points defined with the following formula:*

$$S(X) = \bigcup_{i \in \mathbb{N}} S_i(X), \text{ where } S_i(X) = \varepsilon_{\kappa_i}(X) \setminus \gamma_1[\varepsilon_{\kappa_i}(X)] \tag{10.2}$$

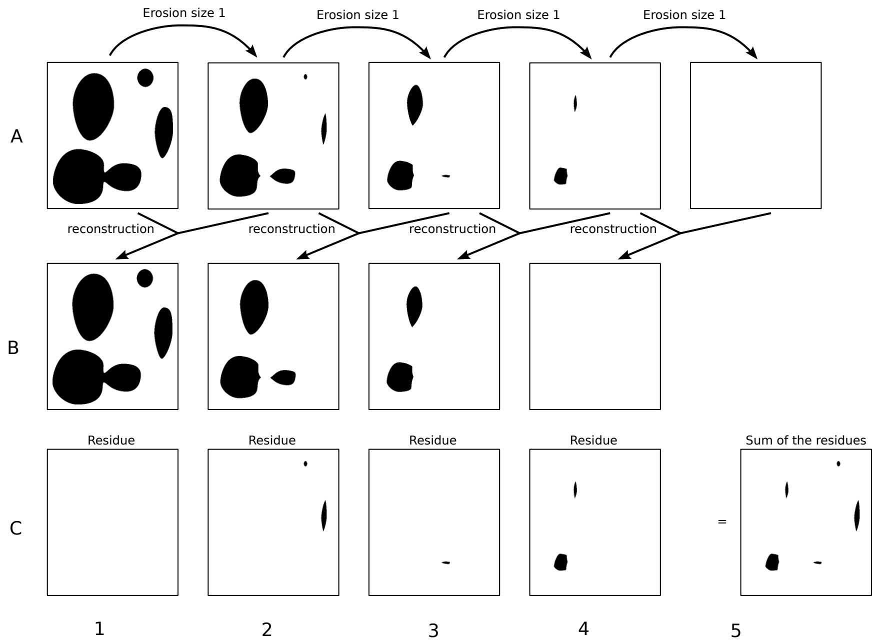


Figure 10.4. *Process for obtaining the ultimate eroded. In the first row (line A), we have the successive erosions A_1, A_2, etc., with a unit ball as structuring element. In the second row (line B), the successive binary reconstructions of A_{i+1} in A_i. So B_2 is the reconstruction of A_3 in A_2. Third row (line C) we have the residues $C_i = A_i - B_i$. Finally, the ultimate eroded is the union $\cup_i C_i$.*

Here, γ_1 is the opening by the unit ball, that is achieved with the smallest structuring element of the chosen granulometric family. In the continuum it may be infinitesimal [MAT 92], but this is not a necessity. In the discrete framework, it is done with α_1, which is often the basic structuring element of the grid (*e.g.* a square 3×3 in 8-connectivity). This definition is inspired from the ultimate eroded, by replacing the opening by reconstruction by the unit opening. Residuals of the unit opening of each successive erosion are the elements of Lantuéjoul skeleton. This procedure is illustrated in two parts, first Figure 10.5 is the illustration of the part $A \backslash \gamma_1(A)$ of the equation 10.2, which corresponds to residues of the opening. Then in Figure 10.6, we show the union of residues for the family erosions $A = \varepsilon_{\kappa_i}(X)$ with X being a rectangle. In both cases we take a disk as structuring element.

In a purely discrete setting, the Lantuéjoul formula still applies. In Figure 10.7 we have the example of the skeleton of Lantuéjoul of a simple discrete figure, taking as unit ball the four-connected neighborhood (the center point plus its four relative nearest neighbors). The key here is to choose a granulometric family ε_{κ_i}, for example with a constant α_i.

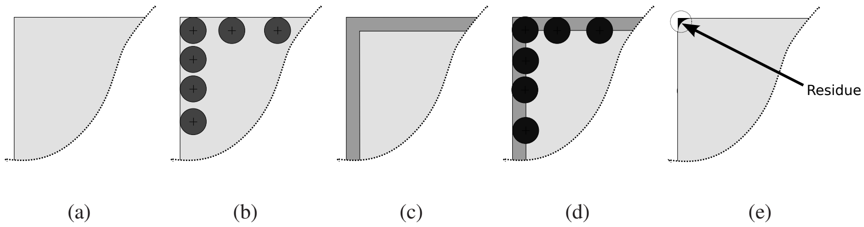


Figure 10.5. *A residue using the Lantuéjoul formula,* i.e. *the result of $A \backslash \gamma_1(A)$: (a) part of an object A, in light grey, (b) erosion by a unit ball, (c) result of the erosion, in light gray on a dark background, (d) dilation process by the same ball, (e) result of the opening, and its residue in black*

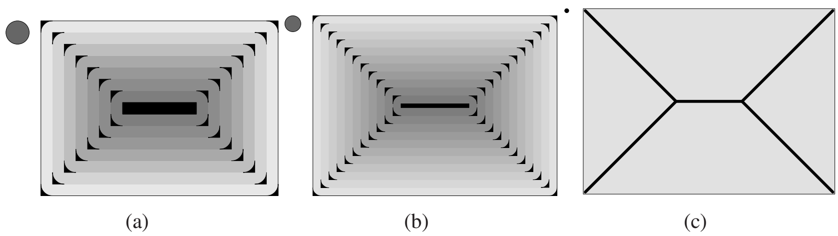


Figure 10.6. *The family of residues of the opening for a simple figure: a rectangle. The successive erosions form nested rectangles, from the lightest to the darkest. The used unit ball is shown in the top left of each rectangle. Residues are in black. In (a) we show the result with a large unit ball for the opening. If the diameter of the ball tends to zero (b),* i.e. *the case of the infinitesimal unit ball, we finally obtain the limiting case of continuous skeleton, in (c) .*

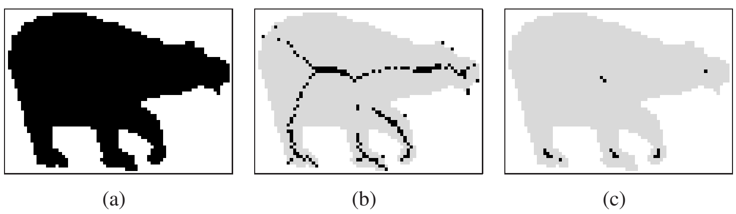


Figure 10.7. *Lantuéjoul skeleton in the discrete case: in (a) the silhouette of a bear, in (b) its Lantuéjoul skeleton, with as unit ball the four-connected neighborhood, in (c) the ultimate eroded with the same unit ball*

We have the important following property:

PROPERTY 10.10.– *Lantuéjoul skeleton is composed of the center of the maximal balls.*

Without bringing a formal proof, we can see why this is so. The LANTUÉJOUL residues at iteration i of the set X are the points kept by an erosion of size i and removed through an opening of size $i + 1$. They are thus the center of a ball of size i contained in X, but no ball of size $i + 1$ or higher can be centered at this point.

Let us observe that any set of points not reconstructed by the opening by reconstruction in the formula of ultimate erodeds would also not be reconstructed with the Lantuéjoul formula, thus the following property holds.

PROPERTY 10.11.– *The ultimate eroded is a subset of the Lantuéjoul skeleton.*

10.4. Discrete distances

This Section provides a brief introduction to the main discrete distances used in image processing. For a more detailed treatment, the reader can refer to [THI 07]. From one of these distances, we deduce immediately the notions of ball and maximal ball for a given object X, and we call *medial axis* of X all the centers of the maximal balls for X.

Let us consider a set E (in the sequel we generally take $E = \mathbb{Z}^n$, with, in most of the examples, $n = 2$), a *distance on* E is an application d from $E \times E$ to $\mathbb{R}^+$ that satisfies:

$$\forall x, y \in E, \quad d(x, y) = d(y, x) \quad \text{(symmetry)} \tag{10.3}$$

$$\forall x, y \in E, \quad d(x, y) = 0 \Leftrightarrow x = y \quad \text{(separation)} \tag{10.4}$$

$$\forall x, y, z \in E, \quad d(x, z) \leq d(x, y) + d(y, z) \quad \text{(triangular inegality)} \tag{10.5}$$

If we consider a subset Y of E and a point x in E, the *distance from x to Y* is defined by $d(x, Y) = \min\{d(x, y) \mid y \in Y\}$.

Let X be a strict subset of E, called “object”, the *distance map of X* is the application D_X from E to $\mathbb{R}^+$ defined by:

$$\forall x \in E,\ D_X(x) = d(x, \overline{X}) \tag{10.6}$$

where $\overline{X}$ denotes the complement of X in E. By abuse of language, we will retain the name "distance map" even if d does not verify the triangular inequality.

In digital image processing, we have extensively used the distances known under the names of "city-block distance" (or "Manhattan distance") and "chessboard distance" (in 2D), because these distances are the easiest to calculate [ROS 68]. We denote them respectively d_4 and d_8, referring to the number of points which are distance 1 of a given point. They are defined by:

$$d_4(x,y) = \sum_{i=1}^{n} |y_i - x_i| \tag{10.7}$$

$$d_8(x,y) = \sup_{i=1}^{n} |y_i - x_i| \tag{10.8}$$

The major problem of these distances is their non-rotational invariance. In practice this means that if one performs a distance measurement on an object taken from a digital image, we can obtain significantly different results depending on the orientation of the object during the shooting. Note that a full rotation invariance can not be achieved when one uses discrete images, however we would like that the effects of a rotation on the measurement of distances do not exceed the size of the discretization step.

In an attempt to overcome this defect of distances d_4 and d_8, the so-called *chamfer distances* have been introduced and studied [MON 68, BOR 84]. To define these distances we must assume that the set E is equipped with a module structure (see [THI 07]). We call *chamfer mask* a finite set of pairs consisting of a displacement v_k and a weight p_k:

$$M = \{(v_k, p_k) \mid v_k \in E, p_k \in \mathbb{R}^+, k \in \{1, \ldots, m\}\} \tag{10.9}$$

satisfying the following conditions:

i) each of the v_k and of the p_k is not null,

ii) M has a central symmetry, that is to say that if (v,p) is in M then $(-v,p)$ is also in M,

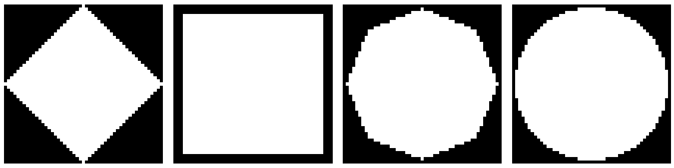
iii) M contains a base of E, that is to say that for all x in E, there exists a m-uplet $(a_1 \ldots a_m)$ of positive or null integers such that $\sum_{1 \leq k \leq m} a_k v_k = x$.

Let x and y two points of E, we set

$$d_M(x,y) = \min\{\sum_{1 \leq k \leq m} a_k p_k \mid y = x + \sum_{1 \leq k \leq m} a_k v_k, a_k \in \mathbb{N}\} \tag{10.10}$$

For example, the following mask allows to retrieve the d_4 distance:

$$M_4 = \{((1,0),1),((0,1),1),((-1,0),1),((0,-1),1)\}.$$



d_4 d_8 $d_{M_{5,7}}$ d_E

Figure 10.8. *Discrete balls obtained from the d_4, d_8 and $d_{M_{5,7}}$ distances and from the Euclidean distance d_E*

The mask

$$M_{5,7} = \{((1,0),5),((1,1),7),((0,1),5),((-1,1),7),\\((-1,0),5),((-1,-1),7),((0,-1),5),((1,-1),7)\}$$

is often used in 2D. By varying the number of pairs in the mask and the weights used, it is possible to obtain more or less accurate approximations of the Euclidean distance, which we discuss below. We can compare on Figure 10.8 the forms of the balls obtained from different distances.

However, the lowest sensitivity to the effects of rotations can only be achieved by using the *Euclidean distance* d_E:

$$d_E^2(x,y) = (y-x)^2 \quad = \quad \sum_{i=1}^{n}(y_i - x_i)^2 \tag{10.11}$$

$$d_E(x,y) \quad = \quad \sqrt{d_E^2(x,y)} \tag{10.12}$$

The *Euclidean squared distance* d_E^2, which is not a distance because it does not verify the triangukar inequality is however sufficient in many applications, and easier to handle computationally than the Euclidean distance: indeed, it only involves integers when the points are integer coordinates. Moreover, if we can calculate a distance map relative to d_E^2, it is easy to deduce a map of Euclidean distance.

To calculate a distance map relative to d_E^2, the naive algorithm (direct application of the definition) has a quadratic complexity in the number of points in the image, which is very inefficient in practice. In 1980, Danielsson [DAN 80] proposes a linear time algorithm for computing an very correct approximation for this kind of distance map for two-dimensional images.

However, it is only relatively recently that this problem has received a truly satisfactory solution: an algorithm giving an exact result, linear in computation time, and that can be generalized whatever the dimension n of the space. This algorithm, published in 1996 by Hirata [HIR 96], based on earlier work [SAI 94] of Saito and Toriwaki, and was found independently by Meijster *et al.* [MEI 00]. Another approach [MAU 03], based on the notion of Voronoï diagram, also provides an efficient and accurate algorithm. In both cases, these algorithms called "separable", compute recursively a distance map of dimension n from maps of dimension $n-1$ computed independently of each other (line by line, plan by plan, etc.) We describe in detail the first of these algorithms in the case $n = 2$.

Let us set $E = [0 \ldots \ell - 1] \times [0 \ldots h - 1]$. To compute the distance map D_X relative to a subset X of E, we have to compute, for any point x of E, the quantity

$$d_X(x) = \min\{(y_1 - x_1)^2 + (y_2 - x_2)^2 \mid y \in \overline{X}\} \tag{10.13}$$

Let us set, for any $j \in [0 \ldots h-1]$

$$d_{X1}(x_1, j) = \min\{(y_1 - x_1)^2 \mid (y_1, j) \in \overline{X}\} \tag{10.14}$$

We can write Eq. (10.13) under the form

$$d_X(x) = \min\{d_{X1}(x_1, j) + (j - x_2)^2 \mid 0 \leq j < h\} \tag{10.15}$$

The calculation of the values of d_{X1} can be done independently, thanks to Eq. (10.14), for each image line, and the d_{X1} are stored in a table. Then, we calculate the values d_X by Eq.(10.15) independently column by column.

To interpret Eq. (10.15) geometrically, note that when x_1 is fixed, as is the case when performing the calculations for a given column, and for a given j, the expression $d_{X1}(x_1, j) + (j - x_2)^2$ defines a parabola whose minimum is reached for $x_2 = j$ and is $d_{X1}(x_1, j)$. The calculations relative to Eq. (10.15) are therefore tantamount to finding the lower envelope of a family of parabolas (see Figure 10.9).

The algorithm described in [HIR 96, MEI 00] consists in calculating the points of integer abscissa of the envelope in two passes, one considering the ascending half-parabolas and the other, the descending half-parabolas. A stack structure allows to act on each parabola with only two operations: one of stacking, the other of unstacking, which ensures the linearity of the algorithm. Moreover, its implementation requires only twenty lines of code.

Following the same approach, D. Coeurjolly proposed algorithms to calculate with an optimal complexity a subset of the exact Euclidean medial axis, sufficient to reconstruct the original object, to make this reconstruction, and to calculate the function that for each point x of an object X associates the set of points of $\overline{X}$ which are minimum distance of xd [COE 03, COU 07a]. This fonction, called *projection on* $\overline{X}$, plays an essential role in defining and calculating the bisector function, which is the subject of Section 10.5 of this chapter.

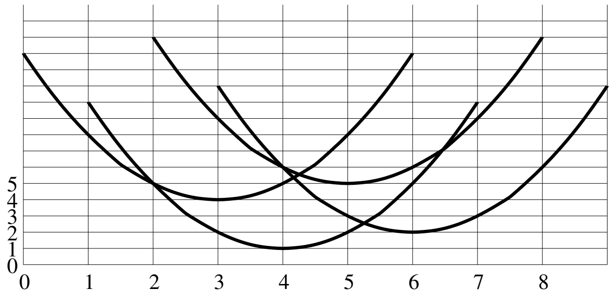


Figure 10.9. *Parabolas obtained for:* $j = 3, d_{X1}(x_1, j) = 4$*;* $j = 4, d_{X1}(x_1, j) = 1$*;* $j = 5, d_{X1}(x_1, j) = 2$*;* $j = 6, d_{X1}(x_1, j) = 5$.

10.5. Bisector function

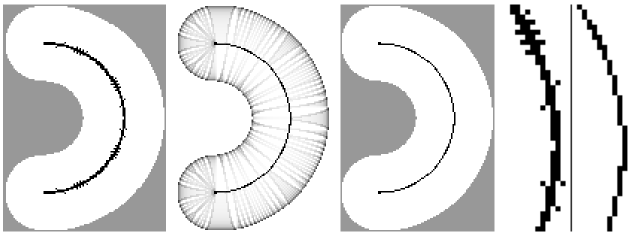
We denote by E the set $\mathbb{R}^n$ or $\mathbb{Z}^n$. Let S be a non-empty subset pf E, and let $x \in E$. The *projection of* x *on* S, denoted by $\Pi_S(x)$, is the set of all points y in S that are closest to x; more precisely, $\Pi_S(x) = \{y \in S, \forall z \in S, d(y, x) \leq d(z, x)\}$.

The bisector angle of a point x in X can be defined, in the continuous framework, as the maximal unsigned angle formed by x (as the vertex) and any two points in the projection of x on $\overline{X}$ [MEY 79, TAL 92]. In particular, if $\#\Pi_{\overline{X}}(x) = 1$, then the bisector angle of x is zero. The *bisector function of* X is the function which associates to each point x of X, its bisector angle in X.

This very definition of the bisector function was used in [ATT 96] in order to provide a filtering criterion for skeletons based on Voronoi diagrams in the continuous plane. It has been also adapted to the discrete case in [TAL 92, MAL 98, COU 07a]. We give here the definition proposed in [COU 07a].

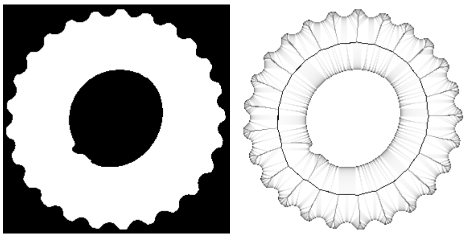
DEFINITION 10.12.– *Let* $X \subset E$*, and let* $x \in X$*. The* extended projection of x on $\overline{X}$*, denoted by* $\Pi^e_{\overline{X}}(x)$*, is the union of the sets* $\Pi_{\overline{X}}(y)$*, for all* y *in* $\Gamma_4(x)$ *such that* $d^2(y, \overline{X}) \leq d^2(x, \overline{X})$*.*
The (discrete) bisector angle of x in X*, denoted by* $\theta_X(x)$*, is the maximal unsigned angle between the vectors* $\overrightarrow{xy}, \overrightarrow{xz}$*, for all* y, z *in* $\Pi^e_{\overline{X}}(x)$*. In particular, if* $\#\Pi^e_{\overline{X}}(x) = 1$*, then* $\theta_X(x) = 0$*. The* (discrete) bisector function of X*, denoted by* θ_X*, is the function which associates to each point* x *of* X*, its discrete bisector angle in* X*.*

Thanks to the algorithm introduced in [COU 07a], the Voronoi mapping $\Pi_{\overline{X}}$ can be computed in optimal time. For each object point x, we must then compute $\Pi^e_{\overline{X}}(x)$ using the adjacency relation Γ and the distance map D^2_X. The last step to obtain the bisector angle consists in the computation of the maximum unsigned angle between all the pairs of vectors $\{\overrightarrow{xy}, \overrightarrow{xz}\}$ for all y, z in $\Pi^e_{\overline{X}}(x)$. In practice, the mean cardinal



(a) (b) (c) (d)

Figure 10.10. *(a): a set X and its medial axis [RÉM 05] (in black); (b): the bisector function θ_X (dark colors correspond to wide angles); (c): filtered medial axis, based on the values of θ_X; (d): detail of the non-filtered and filtered medial axis.*



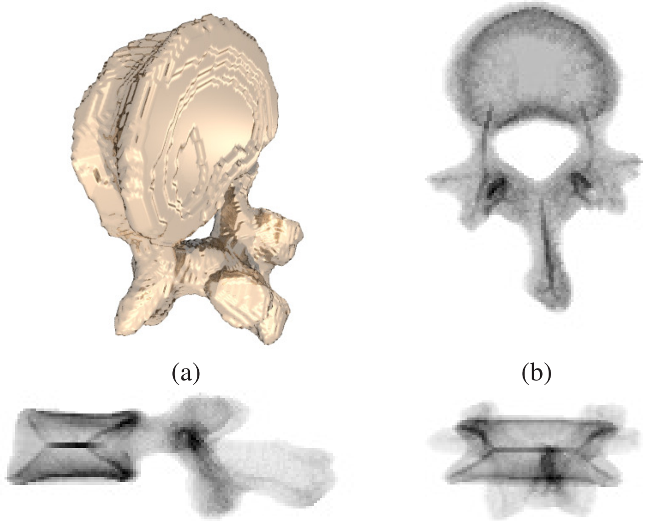
(a) (b)

Figure 10.11. *(a): a set X (in white); (b): the bisector function of X.*

of the extended projections for a given shape is usually quite small; thus considering all possible pairs constitutes the best choice. However, sub-quadratic algorithms exist for this task.

In Figure 10.10, we show a set X together with its medial axis (a) and the discrete bisector function θ_X (b). We illustrate the use of this function to eliminate spurious points of the medial axis: in (c), we show the points of the medial axis (in black) which have a bisector angle greater than 40 degrees. A zoomed detail of both axes is shown in (d). Notice that only the bisector angles of the medial axis points need to be computed for this application. Figure 10.11 shows the bisector function of a more complex 2D shape.

The definition and computation of this discrete bisector function can be straightforwardly extended to $\mathbb{Z}^3$. To conclude this section, we present in Figure 10.12 an illustration of the bisector function of a three-dimensional object (a vertebra).



(a) (b)

(c) (d)

Figure 10.12. *(a): a view of a subset X of $\mathbb{Z}^3$ (vertebra). (b,c,d): the bisector function, illustrated in an "X-ray" manner: the gray level of a point corresponds to the average of the bisector angles on a straight line parallel to one of the three axes.*

10.6. Homotopic transformations

As we saw before, the discrete medial axis is not topologically equivalent, in general, to the original object. Some algorithms used to compute the medial axis proceed by iterated thinning, that is, iterative removal of points from the object. The notion of simple point permits to garantee that such a transformation preserves topology.

Intuitively, a point of an object (a subset of $\mathbb{Z}^n$) is called simple if it can be deleted from this object without altering topology. This notion, pionneered by Duda, Hart, Munson [DUD 67], Golay [GOL 69] and Rosenfeld [ROS 70], has since been the subject of an abundant literature (see *e.g.*, [KON 89]). In particular, local characterizations of simple points have been proposed in 2D, 3D and even 4D, on which efficient implementation of thinning procedures are based [COU 09b].

In Figure 10.13, the points (or pixels) x, y, z, t are not simple: the removal of x from the set X of pixels would create a new connected component of the complement $\overline{X}$ of X ; the removal of y would merge two connected components of $\overline{X}$; the removal of z would split a connected component of X ; and the removal of t would delete a connected component of X. On the other hand, the pixels a, b and c are simple

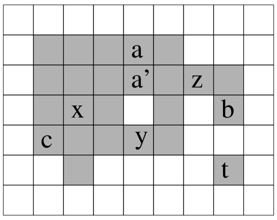


Figure 10.13. *Illustration of 2D simple pixels. The set X is made of the pixels in gray, a, b, c are simple whereas x, y, z, t are not simple.*

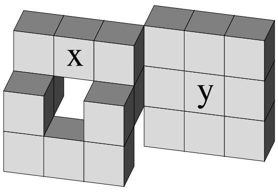


Figure 10.14. *A set X of voxels. The voxels x and y are not simple.*

pixels. We see that, in 2D, the notion of connectedness (for both X and $\overline{X}$) suffices to characterize simple pixels.

Things are more difficult in 3D. Consider the example of the set X depicted in Figure 10.14, removing the voxel x or the voxel y from X would not split, merge, create or suppress any component of X nor any component of $\overline{X}$. However neither x nor y is simple, for the deletion of x (resp. y) causes the suppression (resp. creation) of a tunnel. The notion of tunnel can be formalized thanks to the fundamental group introduced by Poincaré. This group is a topological invariant that is preserved by any continuous deformation.

However, it is still possible to characterize 3D simple points by local conditions that are only based on connectedness (see [BER 94a, BER 94b]), we give one such characterization in the sequel. The fact that a global notion, the one of simple point, can be characterized by a local test, is remarkable. It is even more surprising that connectedness alone suffices to provide such a local characterization in 3D. However, this is no longer true in 4D and higher dimensions. In [COU 09b], a definition of simple point based on the collapse operation is presented, and new local characterizations in 2D, 3D and 4D spaces are introduced. This article also analyzes the difficulties that make it impossible to extend this kind of characterizations to dimension 5 and higher.

10.6.1. *Neighborhoods, connectedness*

We recall here basic definitions of digital topology for binary images [KON 89].

A point $x \in \mathbb{Z}^d$ ($d = 2, 3$) is defined by $(x_1, \ldots, x_d)$ with $x_i \in \mathbb{Z}$.

We consider the neighborhood relations Γ_4 and Γ_8 defined, for any point $x \in \mathbb{Z}^2$, by:
$\Gamma_4(x) = \{y \in \mathbb{Z}^2 \mid |y_1 - x_1| + |y_2 - x_2| \leq 1\}$,
$\Gamma_8(x) = \{y \in \mathbb{Z}^2 \mid \max(|y_1 - x_1|, |y_2 - x_2|) \leq 1\}$,
and the neighborhood relations Γ_6, and Γ_{26} and Γ_{18} defined, for any point $x \in \mathbb{Z}^3$, by:
$\Gamma_6(x) = \{y \in \mathbb{Z}^3 \mid |y_1 - x_1| + |y_2 - x_2| + |y_3 - x_3| \leq 1\}$,
$\Gamma_{26}(x) = \{y \in \mathbb{Z}^3 \mid \max(|y_1 - x_1|, |y_2 - x_2|, |y_3 - x_3|) \leq 1\}$,
$\Gamma_{18}(x) = \{y \in \Gamma_{26}(x) \mid |y_1 - x_1| + |y_2 - x_2| + |y_3 - x_3| \leq 2\}$.

In the sequel, we denote by n a number such that $n \in \{4, 8, 6, 18, 26\}$. We define $\Gamma_n^*(x) = \Gamma_n(x) \setminus \{x\}$. The point $y \in E$ said to be *n-adjacent* to the point $x \in E$ if $y \in \Gamma_n^*(x)$. An *n-path* is a sequence of points $x_0 \ldots x_k$ such that x_i is n-adjacent to x_{i-1} for $i = 1 \ldots k$.

Let $X \subseteq E$, we say that two points x, y of X are *n-connected in X* if there exists a n-path in X between these two points. This defines an equivalence relation. The equivalence classes of this relation are the *n-connected components of X*. A subset X of E is said to be *n-connected* if it is composed of exactly one n-connected component.

The set of all n-connected components of X is denoted by $C_n(X)$. A subset Y of E is said to be *n-adjacent* to a point $x \in E$ if there exists a point $y \in Y$ that is n-adjacent to x. The set of n-connected components of X that are n-adjacent to x is denoted by $C_n^x(X)$. Remark that $C_n(X)$ and $C_n^x(X)$ are sets of subsets of X, not sets of points. Furthermore, if S is a set, we denote by $\#S$ the number of its elements.

10.6.2. *Connectivity numbers and simple points*

Let us first define connectivity numbers in the 2D case. We see later that the 3D case is more complex.
Let $X \subseteq \mathbb{Z}^2$ and $x \in \mathbb{Z}^2$. The *connectivity number* $T_n(x, X)$ is defined by: $T_n(x, X) = \#C_n^x[\Gamma_8^*(x) \cap X]$

If we use the n-connectivity for X, then we must use the $\overline{n}$-connectivity for $\overline{X}$, for example in 2D the 4-connectivity for X is associated with the 8-connectivity for

$\overline{X}$ (and vice-versa), and in 3D the 6-connectivity for X is associated with the 18- or the 26-connectivity for $\overline{X}$ (and vice-versa). This is necessary to have a correspondence between the topological characteristics of X and those of $\overline{X}$ (see [KON 89]). Furthermore, it is sometimes necessary, in 3D, to make a distinction between the 6-connectivity that is associated with the 18-connectivity, and the 6-connectivity that is associated with the 26-connectivity. In order to make this distinction explicit, we use the symbol 6^+ to denote the 6-connectivity that is associated with the 18-connectivity (see [BER 94a]). To summarize, we have the following possibilities in 2D: $(n, \overline{n}) = (4, 8)$ or $(8, 4)$; and in 3D: $(n, \overline{n}) = (6, 26)$, $(26, 6)$, $(6^+, 18)$ or $(18, 6^+)$.

In the 3D case, the definition of connectivity numbers is based on the notion of geodesic neighborhood. Let $X \subseteq \mathbb{Z}^3$ and $x \in \mathbb{Z}^3$, the *geodesic n-neighborhood of x in X of order k* is the set $\Gamma_n^k(x, X)$ recursively defined by:
$\Gamma_n^1(x, X) = \Gamma_n^*(x) \cap X$, and
$\Gamma_n^k(x, X) = \cup\{\Gamma_n(y) \cap \Gamma_{26}^*(x) \cap X, y \in \Gamma_n^{k-1}(x, X)\}$.
In other terms, $\Gamma_n^k(x, X)$ is the set composed of points y of $\Gamma_{26}^*(x) \cap X$ such that there exists a n-path π from x to y, the length of which is less than or equal to k, under the condition that all points of π but x must belong to $\Gamma_{26}^*(x) \cap X$. The *geodesic neighborhoods* $G_n(x, X)$ are defined by: $G_6(x, X) = \Gamma_6^2(x, X)$, $G_{6^+}(x, X) = \Gamma_6^3(x, X)$, $G_{18}(x, X) = \Gamma_{18}^2(x, X)$, et $G_{26}(x, X) = \Gamma_{26}^1(x, X)$.
We can now give a definition of connectivity numbers in 3D [MAL 93, BER 94a, BER 94b].

DEFINITION 10.13.– *Let $X \subseteq \mathbb{Z}^3$ and $x \in \mathbb{Z}^3$. The* connectivity number $T_n(x, X)$ *is defined by: $T_n(x, X) = \#C_n[G_n(x, X)]$*

Notice that a formulation in terms of geodesic neighborhoods also permits to retrieve the definition of connectivity numbers in 2D.

If we use the n-connectivity for X and the $\overline{n}$-connectivity for $\overline{X}$, the connectivity numbers $T_n(x, X)$ and $T_{\overline{n}}(x, \overline{X})$ describe the topological characteristics of point x in the object X. In particular, connectivity numbers can be used to detect whether a point is simple or not [BER 94a, BER 94b], in 2D and in 3D:

PROPERTY 10.14.– *Let $X \subseteq E$ and $x \in X$. The point x is n-simple if and only if $T_n(x, X) = 1$ and $T_{\overline{n}}(x, \overline{X}) = 1$.*

To give an intuitive interpretation of this caracterization, a point is simple if and only if there is in its neighborhood exactly one “object”component and one “background”component.

10.6.3. *Homotopic thinning*

It remains to decide in which order simple points will be deleted. We present, with Algorithm 1, a thinning strategy which consists of controlling this order based on a priority function. This function associates, to each point x of X, an integer or real number $P(x)$, which represents the priority of point x. The points of X having the lowest values of P will be considered first. Some points x can be given the priority $P(x) = +\infty$, meaning that such points cannot be deleted; in other words, these points with infinite priority are anchor points (see [DAV 81, VIN 91a, PUD 98]) for the thinnning.

Algorithm 1: Guided thinning

Data: $X \subseteq E, P$ a function from X to $\mathbb{Z}$ or $\mathbb{R}$
Result: X
repeat
    Let x be an element of X such that x is simple for X and $P(x)$ is minimal ;
    $X = X \setminus \{x\}$;
until *stability*;

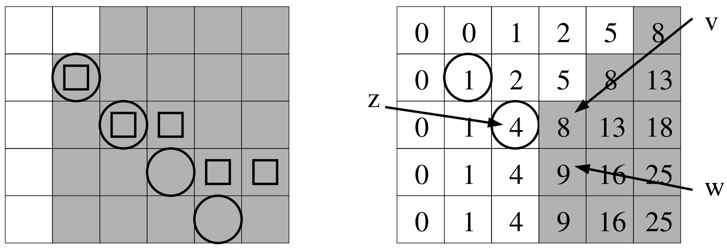
We will discuss later on (sec. 10.6.5) some specific problems that occur whenever the priority function is an Euclidean distance map.

10.6.4. *Sequential and parallel thinning algorithms*

Homotopic transforms discussed in this chapter are sequential by nature, in the sense that after each simple point deletion, this modification must be taken into account when testing the simplicity of other points. In other words, preserving topology is not guaranteed if one deletes several simple points simultaneously: for example in Figure 10.13, deleting both simple points a and a' would merge two connected components of the background.

Consequently, some arbitrary choices must sometimes be done with respect to the order in which simple points are treated. This may lead to different results for a same object. On the other hand, parallel thinning strategies, which are not covered by this chapter, produce skeletons that are uniquely defined.

Nevertheless, general conditions that permit to guarantee topology preservation while simultaneously deleting several simple points are much more difficult to establish than single simple point characterizations. Numerous attempts have been done during the last 40 years to solve this problem [COU 06a]. The notion of minimal non-simple set, introduced by C. Ronse [RON 88], allows for testing whether a set of



(a) (b)

Figure 10.15. *Illustration of the geometric distorsion of a skeleton. (a): A part of an object (in gray), skeleton points found by Algorithm 1 with the Euclidean distance map as priority function (circles), steepest descent path with regard to the same map (squares). (b): One step of thinning (see text). Numbers indicate the distance map values. .*

simple points can be removed while preserving topology. The notion of P-simple point proposed in [BER 95] has the same goal and constitutes fulthermore a general algorithmic scheme for designing 3D thinning algorithms, such that topology preservation is guaranteed by construction. Recently, a general framework for the study of parallel thinning in arbitrary dimension has been developped by G. Bertrand [BER 07a]. This framework, centered on the notion of critical kernel, generalizes both the one of minimal non-simple sets and the one pf P-simple points. The interested reader can find a complete review on critical kernels on the site `http://www.esiee.fr/~info/ck` .

10.6.5. *Skeleton based on the Euclidean distance*

The skeletonization methods which are based on homotopic thinnings, in the sense of section 10.6.3, provide a formal guarantee that the skeleton and the original object have the same topology. The simplest such method consists in computing an ultimate homotopic skeleton of the object X constrained by the medial axis of X, that is, removing iteratively simple points from X which do not belong to $\mathrm{MA}(X)$, taking the distance map as a priority function in order to select first the points which are closest to the background. This can be done using the guided thinning procedure described previously (Algorithm 1), with $P = D_X$ and $Y = \mathrm{MA}(X)$.

The drawback of this method has been well analyzed in [TAL 92]. Roughly speaking the method does not guarantee that points of the homotopic skeleton outside the medial axis are "well centered" in the object. In Figure 10.15, we give an example that illustrates the kind of problem that may occur.

In Figure 10.15b, numbers correspond to squared Euclidean distances from each object point to the nearest background point. The circled point with value 1, is one of the points that belong to the constraint set Y (see Algorithm 1), that is, a point of the medial axis. Assume that all points with value less to or equal to 8 have been treated by the algorithm. At this step, points in gray are still in the object X, as well as the two circled points (the point valued 1 for it belongs to Y, and the one valued 4 for it has been detected as non-simple when it has been examined). All other points are outside of X. Clearly, the point v valued 8, adjacent to z valued 4, will be selected before its neighbor w valued 9, and since it will be simple at this stage, it will be removed from X. This behaviour will be reproduced during subsequent steps, creating a diagonal skeleton branch. This contradicts a property of skeleton in the continuous space, that asserts that such a branch should follow a steepest descent path on the distance map. To check this, let us compute the slopes of segments zv and zw in our example configuration: $(\sqrt{8}-\sqrt{4})/1 \approx 0.83$, and $(\sqrt{9}-\sqrt{4})/\sqrt{2} \approx 0.71$. Thus, point v should be kept in the skeleton rather than point w, following this criterion .

To solve this problem, a strategy proposed in [COU 07a] consists of defining a priority function that takes into account both the distance map and an auxiliary function defined in the neighborhood of each dynamically detected skeleton point. Let x be such a point, to any neighbor y of x that is still in X but not in Y, we associate the value $p_y = D_X(x) + (D_X(y) - D_X(x))/d(x,y)$, with $D_X(x) = \sqrt{D_X^2(x)}$. The new priority function, for any point y, is defined by $\min(p_y, D_X(y))$. We see that $(D_X(y) - D_X(x))/d(x,y)$ is the slope of xy, thus the neighbors of x will be treated in increasing order of slope, since the value p_y is always less than or equal to the value $D_X(y)$ (for all x, y in $\mathbb{Z}^2$ or $\mathbb{Z}^3$ with $x \neq y$, we have $d(x,y) \geq 1$).

For example, in the previous case, we have $D_X(v) = \sqrt{8} \approx 2.83$, $D_X(w) = 3$, $p_v = \sqrt{4} + (\sqrt{8} - \sqrt{4})/1 = \sqrt{8}$ and $p_w = \sqrt{4} + (\sqrt{9} - \sqrt{4})/\sqrt{2} \approx 2.71$; thus the point w will be selected before v with this strategy. The algorithm is described below.

The time complexity of this algorithm depends on the data structure used to represent the sets Q and R. Specifically, this data structure must allow for efficiently choosing (p,x) at the beginning of the "while"loop, and also for fast insertion of new couples. If one uses *e.g.*, a balanced binary tree [COR 09], the overall complexity of the algorithm is $O(n \log n)$, where n is the number of points in the image.

10.7. Conclusion

In this chapter we introduced the concepts of skeleton, medial axis, distance, granulometry and thinning, and some links between these concepts.

The definition of a series of dilations of a point by a family of structuring elements defines a granulometric family, by analogy with screening in geology. This family produces a notion of distance, and an operator: the distance function. The latter, in addition to its intrisic interest, can be used to define some characteristic points of a

Algorithm 2: Euclidean skeleton

Data: $X \subseteq E, D_X$ the Euclidean distance map of $X, Y \subseteq X$
Result: Z
Z =X;
Q =$\{(D_X(x), x)$; where x is an arbitrary point of $X \setminus Y\}$;
R =$\{(p_x, x)$; where x is an arbitrary point of $X \setminus Y$ adjacent to Y, and where $p_x = \min\{D_X(z) + (D_X(x) - D_X(z))/d(x,z), z \in Y\}$ $\}$;
while $Q \neq \emptyset$ *or* $R \neq \emptyset$ **do**
    choose (p, x) in $Q \cup R$ such that p be minimal;
    Q =$Q \setminus \{(p, x)\}$; R =$R \setminus \{(p, x)\}$;
    if $x \in Z \setminus Y$ **then**
        if x *is simple for* Z **then**
            Z =$Z \setminus \{x\}$;
        else
            Y =$Y \cup \{x\}$;
            R =$R \cup \{(p_y, y)$; with $y \in \Gamma(x) \cap (Z \setminus Y)$ and with
            $p_y = D_X(x) + (D_X(y) - D_X(x))/d(x,y)\}$;

binary form, the centers of maximal balls. We showed how to compute these centers in continuous and discrete spaces, by the means of a granulometric family.

To retrieve certain desirable properties of the continuous skeleton in the discrete space, in particular the conservation of topological characteristics, we have been led to introduce the concepts of simple point and homotopic thinning. These concepts apply in dimensions greater than 2, in grayscale images and also in association to discrete Euclidean metric.

With certain precautions, it is now possible do define a skeleton with good robustness properties and rotation invariance. Studying the stability of skeletons is still a subject of active research, with contributions from different disciplines, particular from computational geometry. Attali, Boissonnat and Edelsbrunner identify in [ATT 09] prominent contributions in this area. These include in particular the λ-*medial axis* introduced by Chazal and Lieutier in [CHA 05], having a continuity property limited to certain values of the parameter λ, which has recently been adapted to the discrete framework in [CHA 09].