LECTURE NOTES ON MORPHOLOGICAL OPERATORS

Jean Serra
Centre de Morphologie Mathématique
École Nationale Supérieure des Mines de Paris
35, rue Saint-Honoré
77305 Fontainebleau Cedex
FRANCE
serra@cmm.ensmp.fr

First French-Nordic Summer Course in Mathematics
Uppsala University, SWEDEN
June 2001

Contents

Foreword

As often happens, it is in response to a particular request that the following text was written. Prof. Christer Kiselman, who organised the first French-Nordic Summer School in Mathematics, requested that I write these lecture notes in the framework of his course, to be given at the University of Uppsala in June 2001.

Exactly twelve years ago, also in the northern countries, but this time at the University of Tampere, I gave a similar summer course, in co-operation with L. Vincent. The comparison between the first lecture notes [[66]] and the current one is instructive: the previous material represents only twenty five per cent of the new one! In the field of image processing, matters evolve rapidly.

However, some methodological bases, some a-priori points of view to understand the physical world remain unchanged. And regarding Mathematical Morphology, this behaviour consists in privileging ordering relations between regions rather than their metric characteristics. In particular, it leads directly to set descriptors. Therefore, a certain epistemology is chosen which appears all along the text, as soon as one looks beyond the mathematical layer. For instance, the notion of a granulometry summarises the essentials of all size measurement techniques, and even of what the human language intends to mean by the word "size".

The above choice requires some comments. If one wanted to explain the difference between the natural sciences and physics, the main point of departure would probably be that for the natural sciences, morphological description precedes the determination of laws, whereas for physics it follows. The main interest of law comes classically from the elimination of the morphological characteristics (volume, etc.).

However, in the great majority of areas where mathematical morphology can be used, both the structural description and the determination of laws co-exist, and we must try to bring them together. It is not always an easy task, and if we take only one of these two points of view, our ship is in danger of sinking. For example, a pure physicist might consider that the flow in a porous medium is completely described as soon as the partial differential equation is written down, i.e. when he knows the Navier-Stokes law, plus the initial and boundary conditions. But in practice, the boundary conditions may be so complex and locally unpredictable, that finally the physicist changes his approach (and the observation scale) and uses the Darcy law (flow is proportional to pressure with the coefficient of proportionality being the permeability). At this scale the Navier-Stokes law is virtually meaningless.

Now, from a natural scientist's point of view, porous media can be classified according to their petrographic family : sandstone with or without clay, quartzites, etc Their permeabilities depend partially on this classification, but the relationships are too fuzzy to become operational tools in the hands of the hydro-geologist. Thus, two opposite points of view arise : the descriptive point of view of the naturalist, and the functional point of view of the physicist.

In order to try and bridge the gap between the two approaches, Mathematical Morphology has conceived a series of descriptors, and has developed its theory in three ways. Firstly, it proposes a gamut of operators, that express some characteristics of the medium of the image under study (granulometry, morphological filters, connections, etc). In parallel, Mathematical Morphology has elaborated a comprehensive range of random models, and thirdly has reached a synthesis between textures and physical properties at least in some fields of physics, such as hydrodynamics. The lecture notes that follow are exclusively devoted to the first of these three branches, i.e. the operators.

Although these operators were initially set oriented and designed for physics, they evolved by themselves, and also according to the questions that arose: a tool exists independently of its initial finality. Indeed, the boom of multi-media applications during the last decade has oriented some developments of the method, and notions such as connection, levelling or multiscale processing were suggested by multimedia problems.

Pedagogically speaking, the course comprises the text that follows, plus two other elements. There is on the one hand the series of transparencies for the verbal presentation, but also a collection of 80 case studies. They cover all the current fields of application, and all the modalities of concatenation of the various operators. They can run in real time by means of the "Micromorph" software. A last point. For size reasons I have removed sixty pages of the manuscript, in order to adapt the text to a course which is finally rather short. The complete version is available at the Ecole des Mines de Paris.

Jean SERRA
June 2001

Chapter 1

Complete lattices

1.1 Introduction

This chapter reminds several notions from the theory of complete lattices. Some of them are classical (e.g. boolean lattices), other are less (e.g. Matheron's characterization of $\mathcal{P}$(E) type lattices, ...). Here a rapid historical survey is probably useful. Lattice theory was born at the beginning of the twentieth century. To give an example, the cruxial notion of an algebraic closing and its characterization from its invariants sets goes back to E.H. Moore[50], in 1910. Number of famous names, such as that of J. von Neumann, participated in the development of the theory (in this case in particular, by trying to bend the very algebraic structure of lattice towards topological modalities). The "bible" book of G. Birkhoff [8], whose first edition dates from 1940, provides an excellent perspective view of this historical phase.

However, during this period, few attention was paid to *complete* lattices (i.e. the non finite case), which are the basic ones in Mathematical Morphology, and even in the other theories for image processing. In Birkhoff's book, (3rd edition, 1983), 22 pages, over 418, are devoted to the subject. Indeed, this re-orientation results from the apparition of new domains of application. Before the 2nd world war, and till the sixties, emphasis was put on logical descriptors (e.g. in electronics), that can afford to be finite, and for which the boolean lattices were invented. Emphasis was also put on taxonomy, or classification techniques, in domains such as statistics or semantics, which again are intrinsically finite.

At the beginning of the seventies, the development of lattice theory turned out to have a break. The classical applications had been exhaustively modelled, and the new one, namely image analysis, still in its infancy, used to borrow its concepts from vector spaces and convolution. It is at the beginning of the eighties only, that G. Matheron and J. Serra resorted to complete lattices as a common framework for the four situations met in practice : binary and grey tone images, graphs and partitions[62]. This gave the lattice theory a fresh boost.

Today, apart from the pioneer book [62], the three reference books about the recent developments are that of H.J.A.M. Heijmans on morphological operators[24], G. Matheron's monography on compact lattices ([41], in french) and G. Giers and Al. compendium [17], the latter being relatively independent of image analysis background. But the theory extends above all by means of papers, in particular in the two directions of partial derivative equations and of connectivity. This last, but not exhaused, topic for example, yielded the algrebraic notion of a connection on lattices (J. Serra [71], [72], C. Ronse [56]), connected operators (J. Crespo and Al [12], H. Heijmans [25]), including the important class of the levelings (F. Meyer [49]), and geodesical operators (Ronse and Serra [57]).

1.2 Symbols

In this course, the term "lattice" always means "complete lattice", excepted in some specific cases of the present chapter, where the contrary is explicitely mentionned. The generic symbol for a lattice is $\mathcal{L}$, with some curvilinear variants for other commonly used lattices : $\mathcal{P}(E)$ (for the family of the subsets of set E, $\mathcal{T}$ for a totally ordered set, $\mathcal{D}(E)$ for all partitions of space E, $\mathcal{F}$ for function lattices, for example. The curvilinear letters designate also families of elements, such as the class $\mathcal{B}$ of the invariant sets of an opening, or such as a connection $\mathcal{C}$ on $\mathcal{P}(E)$.

The elements of a lattice may be given small letter (e.g. a numerical function f) or capital letters, according to the context. The choice becomes imperative in the set oriented case, where the elements x of E should be distinguished from thoses, X say, of $\mathcal{P}(E)$. Also, in this case, when point $x \in E$ is considered as an element of $\mathcal{P}(E)$, one writes $\{x\}$.

The operators acting on a given lattice are often given small greek letters; in particular $\delta, \varepsilon, \gamma$ and φ stand for dilation, erosion, opening and closing respectively.

1.3 Partially ordered sets

Definition 1.1 *Given a nonempty set $\mathcal{L}$, a binary relation $\leq$ on $\mathcal{L}$ is called a partial ordering if the following properties hold :*

$x \leq x$; *(reflexitivity)*

$x \leq y$ *and* $y \leq x$ *implies* $x = y$ *(anti-symmetry)*

$x \leq y$ *and* $y \leq z$ *implies* $x \leq z$ *(transitivity)*

for every x, y, $z \in \mathcal{L}$. A set $\mathcal{L}$ which carries a partial ordering $\leq$ is called a partially ordered set, or briefly poset*, and is denoted by ($\mathcal{L}$,$\leq$). We say that the poset ($\mathcal{L}$,$\leq$) is* totally ordered *if*

$$x \leq y \text{ or } y \leq x, \text{ for every pair } x,y \in \mathcal{L}.$$

A totally ordered poset is also called a chain.

Examples

(a) Let X be the set of all numbers of the open interval]01[, and let x $\leq$ y have its usual meaning; then X is a chain. The set $\mathbb{R}^d$ with the relation "(x_1, x_2, ..., x_d) $\leq$ (y_1, y_2, ..., y_d) if $x_i \leq y_i$ for all i" is a poset. It is a chain if and only if d = 1.

(b) Given a set E, the power set $\mathcal{P}(E)$comprising all subsets of E becomes a poset under the inclusion relation, that is, "X $\leq$ Y if and only if X $\subseteq$ Y , X,Y $\in \mathcal{P}(E)$". The empty set is denoted by $\emptyset$.

Duality Principle. If ($\mathcal{L}$,$\leq$) is a poset, then ($\mathcal{L}$,$\leq'$) is a poset too, called the dual poset. To every definition, property, statement, etc., referring to ($\mathcal{L}$,$\leq$) there corresponds a dual one referring to ($\mathcal{L}$,$\leq'$), interchanging the role of $\leq$ and $\leq'$.
Note that the second dual partial ordering $(\leq')'$ coincides with $\leq$. The Duality Principle, seemingly a trivial and rather useless observation, plays a prominent role in this course. Its major implication is that every notion and statement concerning posets has a dual counterpart (e.g. opening versus closing).

Given a poset $\mathcal{L}$ and a subset of $\mathcal{K}$ of $\mathcal{L}$, and element $a \in \mathcal{K}$ is called *a least element* of $\mathcal{K}$ if $a \leq x$ for all $x \in \mathcal{K}$. An element $b \in \mathcal{K}$ is called *a greatest element* of $\mathcal{K}$ if $b \geq x$ for all $x \in \mathcal{K}$. An element $a \in \mathcal{L}$ is called *a lower bound* of $\mathcal{K}$ if $a \leq x$ for every $x \in \mathcal{K}$; note that a need not lie in $\mathcal{K}$. If the set of lower bounds of $\mathcal{K}$, which is a subset of $\mathcal{L}$, contains a greatest element a_0, then this is called *the greatest lower bound*, or *infimum*, of $\mathcal{K}$. Note that a_0 satisfies

(i) $a_0 \leq x$ for $x \in \mathcal{K}$(a_0 is a lower bound of $\mathcal{K}$);
(ii) $a \leq a_0$ for every other lower bound a of $\mathcal{K}$.

The notions *upper bound* and *least upper bound*, also called *supremum,* are defined analogously. In fact, infimum and supremum are dual notions in the sense of the Duality Principle. The infimum (resp. supremum) of a subset $\mathcal{K}$, if it exists, is unique, and is denoted by inf $\mathcal{K}$ or $\bigwedge \mathcal{K}$ (resp. sup $\mathcal{K}$. or $\bigvee \mathcal{K}$). If $x_i \in \mathcal{L}$ for all i in some possibly infinite family of indexes i, then we write $\bigwedge_{i \in I} x_i$ or $\bigwedge \{x_i \mid i \in I\}$, and $\bigvee_{i \in I} x_i$ or $\bigvee \{x_i \mid i \in I\}$.

1.4 Lattices

Definition 1.2 *A poset $\mathcal{L}$ is called a* lattice *if every finite subset of $\mathcal{L}$ has an infimum and a supremum. A lattice is called* complete *if every subset of $\mathcal{L}$ has an infimum and a supremum.*

Every totally ordered poset is a lattice, for every finite set of elements of a chain can be arranged in increasing order, but this lattice may not be complete (e.g. the open interval]01[).

By definition, every complete lattice $\mathcal{L}$ contains a least element o and a greatest element m, called the *extreme*, or *universal bounds* of $\mathcal{L}$.

A poset $\mathcal{L}$ on which every subset admits an infimum only is called a complete *inf semi-lattice.* According to proposition 1.5 below, a complete inf semi-lattice with a maximum element is a complete lattice.

A subset $\mathcal{M}$ of a complete lattice $\mathcal{L}$ is a *complete sub lattice* of $\mathcal{L}$ when finite or infinite infima and suprema of families in $\mathcal{M}$ lie in $\mathcal{M}$ and when $\mathcal{M}$ contains the two extreme elements of $\mathcal{L}$.

Definition 1.3 (Anamorphoses) *Let $\mathcal{L}, \mathcal{M}$ be complete lattices. The mapping $\alpha : \mathcal{L} \longrightarrow \mathcal{M}$ is called an* anamorphosis, *or a lattice isomorphism, if*

α is a bijection (one-one and onto), and if α as well as its inverse α^{-1} are order-preserving, that is,

$$x \leq y \qquad \text{if and only if} \qquad \alpha(x) \leq \alpha(y) \tag{1.1}$$

for x,y $\in \mathcal{L}$.

The definition is slightly redundant. Indeed, if mapping α is onto, then according to rel.1.1 $\alpha(x) = \alpha(y)$ implies $x < y$ and $y < x$ hence $x = y$, and α is one-one. An equivalent formulation of ananamorphosis is the following

Proposition 1.4 *The mapping $\alpha : \mathcal{L} \longrightarrow \mathcal{M}$ is an anamorphosis if and only if α is a bijection that preserves infima and suprema, that is,*

$$\begin{aligned} \alpha(\bigwedge\{x_i \mid i \in I\}) &= \bigwedge\{\alpha(x_i) \mid i \in I\}\,, \\ \alpha(\bigvee\{x_i \mid i \in I\}) &= \bigvee\{\alpha(x_i) \mid i \in I\}\,, \end{aligned}$$

for any family $\{x_i \mid i \in I\}$ in $\mathcal{L}$.

proof: We prove the first relation. Since α is order preserving, it follows from $\bigwedge x_i \leq x_i$ that $\alpha(\bigwedge x_i) \leq \bigwedge \alpha(x_i)$. Assume now that $z \leq \bigwedge \alpha(x_i)$. Since α^{-1} is order preserving, we have $\alpha^{-1}(z) \leq x_i$ for all $i \in I$, hence $\alpha^{-1}(z) \leq \bigwedge x_i$ i.e. $z \leq \alpha(\bigwedge x_i)$, which results in the first equality of the proposition.Conversely, rel.1.1 is obviously necessary, it suffices to take the family $\{x_1, x_2\}$ with $x_1 \leq x_2$. $\square$

The word anamorphosis was introduced in painting, during the Renaissance, to designate geometrical distortions of the space. Indeed such distortions on space E induce an *anamorphosis* (in the above sense) on $\mathcal{P}(E)$. But there are many other ones, suh as $x \to \log x$, for $x > 0$: it is an anamorphosis $\overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}$, that extends to the numerical functions on E. Incidently, the proposition introduces two types of operations, that preverve $\bigvee$ or $\bigwedge$. They are called *dilation* and *erosion* respectively, and play a central role in the theory of Mathematical Morphology.

Proposition 1.5 *Given a poset $\mathcal{L}$, the following three statements are equivalent :*

(i) *$\mathcal{L}$ is a complete lattice;*
(ii) *$\mathcal{L}$ has a least element o and every subset of $\mathcal{L}$ has a supremum;*
(iii) *$\mathcal{L}$ has a greatest element $m$and every subset of $\mathcal{L}$ has an infimum.*

proof: It is obvious that (i) implies (ii) and (iii). We show that (ii) implies (i). The other implication follows from the Duality Principle. Assume that (ii) holds, and let $\mathcal{K} \subseteq \mathcal{L}$. Denote the set of lower bounds of $\mathcal{K}$ by $\mathcal{T}$. Then$\mathcal{T} \neq \emptyset$ since $O \in \mathcal{T}$. Let a = sup $\mathcal{T}$; we show that a is the infimum of $\mathcal{K}$. Note first that every x $\in \mathcal{K}$ is an upper bound of $\mathcal{T}$; since a is the least upper bound, $a \leq x$. Let a' be a lower bound of $\mathcal{K}$; then $a' \in \mathcal{T}$, and so $a' \leq a$. This proves that $a = \inf \mathcal{K}$. □

In other words, a complete inf semi-lattice with a gratest element is a complete lattice.The next result lists some basic properties of the infimum and supremum.

Proposition 1.6 *Let $\mathcal{L}$ be a poset and $x, y, z \in \mathcal{L}$. We have*

(a)	$x \wedge y = x \vee x = x$	(*idempotence*)
(b)	$x \wedge y = y \wedge x, x \vee y = y \vee x;$	(*commutativity*)
(c)	$x \wedge (y \wedge z) = (x \wedge y) \wedge z;$	(*associativity*)
(d)	$x \wedge (x \vee y) = x \vee (x \wedge y) = x.$	(*absorption*)

1.5 Remarkable Elements

Definition 1.7 *(*atoms, co-primes) *A non zero element a of complete lattice $\mathcal{L}$ is an* atom *if $x \leq a$ implies $x = 0$ or $x = a$. For example, when $\mathcal{L}$ is of the type $\mathcal{P}(E)$, the points of E are atoms in $\mathcal{P}(E)$.*

An element $a \in \mathcal{L}$, $a \neq 0$ is said to be co-prime *when $a \leq x \bigvee y$ implies $a \leq x$ or $a \leq y$, in a non exclusive manner.*

We will complete these two classical definitions (see Heijmans [24], or Gierz et al. [17]) with a third one, from Matheron [41], according to which $a \in \mathcal{L}$ is *strongly* co-prime when for any family B in $\mathcal{L}$ (finite or not), $a \leq \bigvee \{b : b \in B\}$ implies the existence of a $b \in B$ with $a \leq b$.

Definition 1.8 *(*Sup-generators) *Let L be a lattice and $X \subseteq \mathcal{L}$ a family in $\mathcal{L}$. The class X is a* sup generator *when every element $a \in \mathcal{L}$ is the supremum of the elements of X that it majorates :*

$$a = \bigvee \{x \in X,\, x \leq a\} \quad .$$

Lattice $\mathcal{L}$ is said to be atomic *(resp.* co-prime, strongly co-prime*) when it is generated by a class of atoms (res. co-prime, strong co-primes).*

Clearly, every atom and every strong co-prime belong to every sup-generating family.

Definition 1.9 *(*Complements*) Let $\mathcal{L}$ be a lattice with extreme elements o and m. If $x, y \in \mathcal{L}$ are such that*

$$x \wedge y = 0 \qquad \text{and} \qquad x \vee y = m$$

then y is called a complement of x (and vice versa). The lattice $\mathcal{L}$ is said to be complemented when all its elements have a complement

1.5.1 Matheron characterization

Here, one result is worth mentioning. Due to G. Matheron ([41], p. 179), it combines the notions we have just introduced.

Theorem 1.10 *For a complete lattice $\mathcal{L}$, the three following statements are equivalent :*
a/ $\mathcal{L}$ is co-prime and complemented
b/ $\mathcal{L}$ is atomic and strongly co-prime
c/ If Q, Q_a and Q_f denote the classes of co-primes, atoms and strong co-primes respectively, then

$$Q = Q_a = Q_f$$

and $\mathcal{L}$ is isomorphic to lattice $\mathcal{P}(Q)$.

This theorem shows how demanding are the assumptions of atomicity and of strong co-primarity, which in fact restricts the approach to the set-oriented case.

In this reminder, and in the course which follows, the emphasis is put on the supremum. But it is clear that each of the above notions admits a dual form. It suffices to consider the dual lattice $\mathcal{L}^*$ of $\mathcal{L}$ (where inequalities, and sup and inf are inverted). Atoms, co-prime, strong co-prime and sup-generators on $\mathcal{L}^*$ define, on $\mathcal{L}$, dual atoms (also called anti-atoms), prime, strong prime and inf-generators respectively.

1.5.2 De Morgan's law

A major property of the complemented lattices is the classical De Morgan's law, which plays a central role in logic and whose statement is the following.

Proposition 1.11 *Consider a possibly non complete lattice $\mathcal{L}$ with a complement $x \longmapsto x^*$. For every finite family x_i in $\mathcal{L}$, we have*

$$\begin{aligned} (i \in I \bigwedge x_i)^* &= i \in I \bigvee x_i^* \\ (i \in I \bigvee x_i)^* &= i \in I \bigwedge x_i^*. \end{aligned}$$

If $\mathcal{L}$ is complete, then these laws are also valid for infinite families.

1.6 Distributivity

Several useful properties involve distributivity, or rather, distributivities. Remember that a lattice $\mathcal{L}$ is *distributive* if

$$\begin{aligned} x \bigwedge \left(y \bigvee z\right) &= \left(x \bigwedge y\right) \bigvee \left(x \bigwedge z\right) \\ x \bigvee \left(y \bigwedge z\right) &= \left(x \bigvee y\right) \bigwedge \left(x \bigvee z\right) \end{aligned}$$

for all x, y, $z \in \mathcal{L}$. The two equalities are equivalent. When the collection of elements between parentheses is allowed to extend to infinity, i.e. when

$$x \bigwedge \left(\bigvee y_i \;, i \in I\right) = \bigvee \left\{\left(x \bigwedge y_i\right), i \in I\right\} \qquad \left(\text{infinite } \bigvee\text{-distributivity}\right) \tag{1.2}$$

$$x \bigvee \left(\bigwedge y_i \;, i \in I\right) = \bigwedge \left\{\left(x \bigvee y_i\right), i \in I\right\} \qquad \left(\text{infinite } \bigwedge\text{-distributivity}\right) \tag{1.3}$$

for any collection $\{y_i\} \in \mathcal{L}$ and for $x \in \mathcal{L}$, then lattice L is *infinite distributive*. Unlike the finite case, the two equalities1.2 and1.3 are not equivalent. There exists a more severe distributivity, called *complete distributivity* by G. Birkhoff [8], and which has been discussed by G. Matheron [[41], p.77] under the name of *total distributivity*. If $\mathcal{B}$ stands for a family of subsets of an arbitrary set E, we denote $h(\mathcal{B})$ the class of those parts H of E which are obtained by taking one point in each $B \in \mathcal{B}$. Then a lattice L is *totally distributive* when

$$\bigwedge_{B\in\mathcal{B}} \bigvee B = \bigvee_{H\in h(\mathcal{B})} \bigwedge H \qquad (\mathcal{B} \in \mathcal{P}(\mathcal{L}))$$

or equivalently

$$\bigvee_{B\in\mathcal{B}} \bigwedge B = \bigwedge_{H\in h(\mathcal{B})} \bigvee H \qquad (\mathcal{B} \in \mathcal{P}(\mathcal{L}))$$

Total distributivity implies the two other ones, and can be identified with the very strong property of *mono-separation* ([[41], th. 8-11). In particular, every class of functions which is closed under numerical sup and inf forms a totally distributive lattice.

Coprimarity and distributivity in lattices are related to each other. Half of Matheron's monography [41] is devoted to this matter. More modestly, we shall restrict ourselves to the three following results.

Proposition 1.12 *a/ Any co-prime lattice is infinite $\bigwedge$-distributive ([41], th. 8-11)*

b/ In a distributive lattice, every atom is co-prime([41], p. 101) ([24], prop. 2-37).

c/ In an infinite $\bigvee$-distributive lattice, every atom is strong co-prime.

1.7 Boolean lattices

A very popular class of lattices are the so-called *Boolean lattices*, which are defined as follows.

Definition 1.13 *A (non necessarily complete) lattice $\mathcal{L}$ is called Boolean when it is* complemented *and* distributive.

Proposition 1.14 *If $\mathcal{L}$ is a distributive lattice, then every element of $\mathcal{L}$ has at most one complement.*

proof: Let $x \in \mathcal{L}$ and assume that y, z are both complements of x. Then

$$y = y \wedge m = y \wedge (z \vee x) = (y \wedge z) \vee (y \wedge x) = y \wedge z.$$

Analogously $z = y \wedge z$, and we conclude that $y = z$. $\square$

It derives directly from this proposition that in a Boolean lattice every element x possesses a unique complement, denoted by x^*. Clearly, we have

$$(x^*)^* = x.$$

In complete Boolean lattices, finite distributivity and infinite one are equivalent (but not the total one). Indeed, we have

Proposition 1.15 *In every complete Boolean lattice the infinite distributive laws 1.2 and 1.3 hold.*

Let $\mathcal{L}$ be a complete Boolean lattice. We show that 1.2 holds : then1.3 follows by duality. We must show that

$$a \wedge i \in I \bigvee x_i = i \in I \bigwedge (a \wedge x_i)$$

if $a, x_i \in \mathcal{L}$. Since $a \wedge \bigvee_{i \in I} x_i \geq a \wedge x_i$, the inequality $\geq$ follows immediately.To prove $\leq$ put $y = \bigvee_{i \in I}(a \wedge x_i)$. Then, $a \wedge x_i \leq y$ for every i; hence

$$x_i = (a \wedge x_i) \vee (a^* \wedge x_i) \leq y \vee a^*,$$

so that

$$a \wedge \bigvee (x_i, i \in I) \leq a \wedge (y \vee a^*) = a \wedge y \leq y$$

Here is now a stonger result, from [41], p.175, that provides a necessary condition for a complete lattice to be Boolean.

Proposition 1.16 *Let $\mathcal{L}$ be a complete lattice. If there exists a one-to-one* or *onto mapping C from into itself that satisfies the relation*

$$y \leq \mathsf{C}(x) \Leftrightarrow y \wedge x = 0, \tag{1.4}$$

then C is a complement over $\mathcal{L}$, and $\mathcal{L}$ is infinite distributive and complemented.

The major reason for introducing Boolean lattices comes from that, *in the finite case*, they are identical to the $\mathcal{P}(E)$ type lattices. Unfortunately, this cruxial property does not extend to all infinite situations, as shown by example 3 below, that we borrow from [41] and [24].

This counter example suggests that the characterization of $\mathcal{P}(E)$ type lattices requires a stronger notion than infinite distributivity. Here,we meet again the total distributivity, that yields the following representation theorem ([21], p 70).

Theorem 1.17 *In a Boolean lattice $\mathcal{L}$, the following three statements are equivalent :*
  (i) *$\mathcal{L}$ is complete and atomic ;*
  (ii) *$\mathcal{L}$ is complete and totally distributive*
  (iii) *$\mathcal{L}$ is isomorphic to the class $\mathcal{P}(E)$ for some set E.*

The comparison of the two theorems 1.10 and 1.17 is instructive. Matheron's result is more precise (the set E is identified by several sup-generators), and more adapted to complete lattices. But above all, his characterization which requires two axioms only (co-prime and complemented) is logically more economic that the three ones of the above theorem (distributive, complemented and atomic). Finally, Matheron's approach does not need the distinction between finite and various infinite cases.

1.8 Topology and CCO lattices

Topological spaces may also be complete lattices, such as the open sets, or the upper semi continous functions presented below. But conversely, given lattice $\mathcal{L}$, how can we equip it with nice topologies? The convenient notion here is that of a compact and Hausdorff complete lattice with closed ordering [41] (in brief : a CCO lattice). But before presenting it, we would like to recall the *purely algebraic* notions of monotonous convergence, and monotonous continuity.

Definition 1.18 *A family $\{x_i, i \in I\}$ in the complete lattice $\mathcal{L}$ is said to be increasingly (resp. decreasing) filtering if the set I of labels is a poset that satisfies the two properties*
  a) $i \geq j \Rightarrow x_i \geq x_j$ $(resp. x_i < x_j)$

b) for all $i, j \in I$, there exists k larger than i and j.
Then the element x of $\mathcal{L}$ defined *by*

$$x = \vee\{x_i, i \in I\} \quad (resp.\ x = \wedge\{x_i, i \in I\})$$

is called monotonous limit *of the filering family* $\{x_i\}$.

One usually writes $x_i \uparrow x$ (resp.$x_i \downarrow x$). This definition yields the *non topological* notion of monotonous continuity. Let $f : \mathcal{L} \to \mathcal{M}$ be an increasing mapping between lattices $\mathcal{L}$ and $\mathcal{M}$ .We say that f is $\uparrow$ *−continuous* (resp.$\downarrow$ *−continuous*) if

$$x_i \uparrow x in\ \mathcal{L}\ implies\ f(x_i) \uparrow f(x) in\ \mathcal{M}$$

$$(resp.\ x_i \downarrow x\ in\ \mathcal{L}\ \ implies\ f(x_i) \downarrow f(x) in\ \mathcal{M})$$

The monotonous continuity, as introduced here, is a stronger notion than the usual *sequential* monotonous continuity, where the set of labels $I = \mathbb{Z}_+$.We now continue with actual topological concepts.

Definition 1.19 *A* topological *complete lattice $\mathcal{L}$ is said to have a* closed ordering *when for all $x, y \in \mathcal{L}$,the set $\{(x, y) : x < y\}$ is closed in the product space $\mathcal{L} \otimes \mathcal{L}$.*

A simple criterion is the following: if two families x_i and y_i $i \in I$ are filtered by a same base $\mathcal{B}$ and satisfiy $x_i < y_i, x_i \to x, y_i \to y$ in $\mathcal{L}$, then $x < y$. Since the whole set of morphological operators rests on $\bigvee$and $\bigwedge$, we can wonder under which conditions both sup and inf mappings, from the closed sets $\mathcal{F}(\mathcal{L})$ into $\mathcal{L}$ are continuous. The following theorem [41], p.60, provides a criterion to answer this question

Theorem 1.20 *An $\mathcal{L}$ an algebraic complete lattice. There exists on $\mathcal{L}$ a separated topology for which monotonous and topological limits are identical if and only if, for all $(x, y) \in \mathcal{L}$, one can find two elements x' and y' in $\mathcal{L}$ such that*

$$x \notin M_{y'}\ ,\ \ y \notin M^{x'} \quad M^{x'} \cup M_{y'} = \mathcal{L} \tag{1.5}$$

where $M^{x'}$stands for the set of the elements smaller than x' and $M_{y'}$ for that of the elements larger than y'. This separated topology, necessarily unique, makes $\mathcal{L}$ a CCO lattice, where both sup and inf operations are continous.

Remarkably, the theorem does not involve any a priori topological status for lattice $\mathcal{L}$. Criterion 1.5 of the theorem provides at once existence and unicity for the topology, and continuity for both sup and inf.

1.9 Examples

There are numerous lattices associated with image processing. We list here the most common ones, plus a few other ones, instructive because they illustrate differently atoms, co-primes and distributivity. They will be completed, further on, by lattices modelling morphological *operators*,and no longer objects under study.

1.9.1 $\mathcal{P}(E)$ type lattice

Start from an arbitrary set E. Obviously, the set $\mathcal{P}(E)$ of the subsets of E, which is ordered for the inclusion relationship, is a complete lattice for the operations $\cup$ (union) and $\cap$ (intersection). Moreover, with each $X \in \mathcal{P}(E)$, there exists a unique $X^C \in \mathcal{P}(E)$, called the *complement* of X, such that:

$$X \cap X^C = \emptyset \qquad \text{and} \qquad X \cup X^C = E. \tag{1.6}$$

The points of E are atoms, strong co-primes and sup-generators of $\mathcal{P}(E)$; $\mathcal{P}(E)$ is also complemented, hence strongly co-prime, and totally distributive, i.e. for all $Y \in \mathcal{P}(E)$ and any family (X_i) of elements of $\mathcal{P}(E)$, we have:

$$\begin{aligned} (\bigcup X_i) \cap Y &= \bigcup (X_i \cap Y), \\ (\bigcap X_i) \cup Y &= \bigcap (X_i \cup Y); \end{aligned} \tag{1.7}$$

it accumulates all nice features.

1.9.2 Open sets

When E is a topological space, its open sets generate a complete lattice for the inclusion, where the sup coincides with the union and where $\inf(X_i)$ is the interior of $\bigcap X_i$. This lattice is not complemented; it does not admit atoms, but the complements of the points are dual atoms. They are not co-prime, but form a inf-generating family, and indeed this lattice is not infinite $\bigwedge$-distributive (contraposed form of property (a) of proposition 1.12). Similar structures are derived for the *closed* sets and the *compact* sets.

1.9.3 Open/closed sets

Let E be a topological space, where $\mathcal{F}$ and $\mathcal{G}$ stand for the classes of the closed sets and the open sets of E respectively. Denote by $\mathcal{A} \subseteq \mathcal{F}$ the class of the closed sets A such that

$$A = \overline{\overset{\circ}{A}}$$

Class $\mathcal{A}$ turns out to be a complete lattice for the inclusion ordering, and for the following supremum and infimum :

$$\vee A_i = \overline{\cup} A_i \qquad ; \qquad \wedge A_i = \overline{\overset{\circ}{\cap}} A_i$$

In addition, the operator $\complement$ defined by

$$\complement(A) = (\overset{\circ}{A})^c = \overline{(A^c)}$$

is a bijection on $\mathcal{A}$ such that

$$y \leq \complement(x) \Leftrightarrow x \wedge y = O$$

Then, according to proposition1.16, operator $\complement$ is a complement on $\mathcal{A}$, and $\mathcal{A}$ is infinite distributive and complemented ... but not of $\mathcal{P}$(E) type, since obviously, it has no atom.

1.9.4 Minkowski dilates

Lattice of the Minkowski dilates by a disc B, in $\mathbb{R}^2$, i.e.

$$\mathcal{L} = \left\{ X \oplus B, X \in \mathcal{P}(\mathbb{R}^2) \right\}$$

Here, the sup coincides with the usual union, but the inf is the opening by B of the intersection (Fig.1.1). The discs B_x, $x \in \mathbb{R}^2$, are sup-generating atoms, but not co-primes. Again, property (b) of proposition 1.12 appears (in a contraposed form) since this lattice does not satisfy any distributivity.

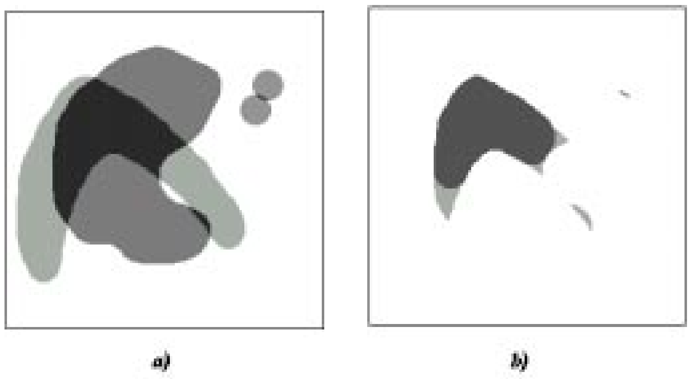


Figure 1.1: a) Two large particles and two atoms in the lattice of the dilates by a disc, b) In light grey, the corresponding intersections, and in dark grey, the inf in the dilates lattice sense. Since the two small discs have an empty inf, they are seen as disjoint particles in the dilate lattice, whereas they intersect with each other.

1.9.5 Convex sets

A set $X \subseteq \mathbb{R}^2$ is *convex* if for every two points $x, y \in X$ the entire straight line segment between x and y is contained in X. Note that this definition includes the points and the empty set. Denote the convex subsets of $\mathbb{R}^2$ by $\mathcal{C}(\mathbb{R}^2)$. With the inclusion as partial ordering this set becomes a complete lattice. The infimum is the ordinary set intersection, but this is no longer true for the union. Define the *convex hull* $\mathrm{co}(X)$ of a set $X \subseteq \mathbb{R}^2$ as the smallest convex set that contains X. It is evident that $\mathrm{co}(X)$ is the intersection of all convex sets which contain X. Clearly, the supremum of a family $X_i \in \mathcal{C}$ $(\mathbb{R}^2)$ is given by

$$\vee\{X_i, i \in I\} = co(\cup\ X_i, i \in I)$$

This lattice is atomic, but not sup-generated, neither complemented nor distributive.

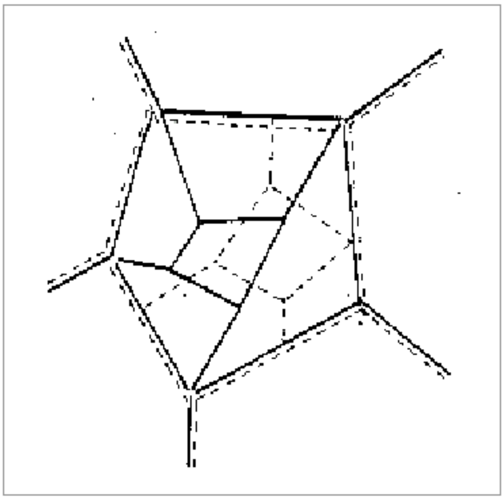


Figure 1.2: Supremum of two partitions

1.9.6 Partitions lattices

A number of attractive properties of the connected classes come from their ability to *partition* each element of L into its components. In order to describe them, we will first define the notion of a partition $\mathcal{D}$ ($\mathcal{D}$ as "division").

Definition 1.21 *(*Partition*) Let E be an arbitrary set. A partition $\mathcal{D}$ of E is a mapping $x \to D(x)$ from E into $\mathcal{P}(E)$ such that*
(i) for all $x \in E : x \in D(x)$
(ii) for all $x, y \in E : D(x) = D(y)$ or $D_a(x) \cap D_a(y) = \varnothing$
$D(x)$ is called the class of the partition of origin x.

In the set of the partitions of an arbitrary set E, we can introduce the following ordering: a partition A is smaller than a partition B when each class of A is included in a class of B. This leads to a lattice which is complete, the greatest element has one class only, namely set E itself, and the finest partition has all the points of E as classes.

Given a family $\{\mathcal{D}_i, i \in I\}$ of partitions, the mapping $\mathcal{D}$ of E into $\mathcal{P}(E)$ defined via its classes

$$D(x) = \cap\{D_i(x), i \in I\}$$

generates obviously a partition where for all $x \in E$, $D(x)$ is the largest element of $\mathcal{P}(E)$ that is contained in each $D_i(x)$. Therefore, $\mathcal{D}$ is the inf of the $\mathcal{D}_i$, in the sense of the partition lattice. The expression of the supremum is more complex, and $\mathcal{D} = \vee \mathcal{D}_i$ means that for all x and any i, class $D_i(x)$ is the smallest set that is a union of classes $D_i(y), y \in E$ (see fig.1.2). The lattice of the partitions of E is neither distributive, nor atomic, nor co-prime, nor complemented...but sup-generated.

1.9.7 Complete chains $\mathcal{T}$

The set of real numbers $\mathbb{R}$ is a lattice; even more, it is a chain. It is not complete for it does not contain a least and greatest element. To make it complete one has to add -∞ (as the least element), and $+\infty$ (as the greatest element). Henceforth, the set $\mathbb{R} \cup \{-\infty, +\infty\}$ will be denoted by $\overline{\mathbb{R}}$. Analogously, $\overline{\mathbb{Z}} = \mathbb{Z} \cup \{-\infty, +\infty\}$ is a complete chain.The sets $\overline{\mathbb{R}_+}$ and $\overline{\mathbb{Z}_+}$ comprising, respectively, the positive real numbers and positive integers including ∞, are complete chains, too. Clearly, the three totally ordered lattices $\overline{\mathbb{R}}$, $\overline{\mathbb{R}_+}$ and $[0, 1]$ are isomorphic, as well as the two lattices $\overline{\mathbb{Z}}$ and $\overline{\mathbb{Z}_+}$.

However, a chain complete lattice is not always isomorphic to $\overline{\mathbb{R}}$ or to $\overline{\mathbb{Z}}$. Structures such as a lexicographic ordering in $\mathbb{R}^2$ generate complete lattice which are totally ordererd but not reducible to the completed straight line.

One easily derives from theorem1.20 that for the order topology on $\overline{\mathbb{R}}$, both $\bigvee$and $\bigwedge$ are continuous.

1.9.8 Numerical functions $\mathcal{T}^E$

Let E be an arbitrary space. The class $\mathcal{T}^E$ of the real valued functions $f : E \longrightarrow \mathcal{T}$, where $\mathcal{T}$ is a complete chain, is obviously ordered by the relation: $f \leq g$ if for each $x \in E$, $f(x) \leq g(x)$ and constitutes a complete lattice. The sup and the inf are given by the relationships:

$$\begin{array}{lclr} f = \vee f_i & \iff & f(x) = \sup f_i(x), & \forall x \in E, \\ f' = \wedge f_i & \iff & f'(x) = \inf f_i(x), & \forall x \in E. \end{array} \tag{1.8}$$

Class $\mathcal{T}^E$ is a topological space for the product topology induced by the CCO topology on $\mathcal{T}$. One derives from theorem 1.20 that, just as in lattice $\mathcal{T}$, the two operations of sup and inf from $\mathcal{F}(\mathcal{T}^E)$ into $\mathcal{T}^E$ are continuous.

The pulse functions:

$$\begin{cases} i_{x,t}(y) = t & if \quad y = x \\ i_{x,t}(y) = 0 & if \quad x \neq y \end{cases} \tag{1.9}$$

associated with each $x \in E$ and $t \in \mathcal{T}$ are sup-generating co-primes but not atoms (except when $\mathcal{T} = \{0,1\}$) and generally not strong co-primes. However, when $\mathcal{T}$ is discrete ($\mathcal{T}$ finite, or $\mathcal{T} = \overline{\mathbb{Z}}$, etc.), then the pulses are strong co-primes. This lattice is totally distributive but not complemented. Rel. (1.8) implies that $f(x)$ may equal $+\infty$.

Note that, by anamorphosis, the previous comments apply equivalently to non negative functions, with $\mathcal{T} = [0, +\infty]$ or to bounded ones, with $\mathcal{T} = [0, 1]$.

1.9.9 φ-continous functions

We now suppose that $\mathcal{T}$ is either finite, or isomorphic to $\overline{\mathbb{R}}$ or to $\overline{\mathbb{Z}}$ (e.g. $[0,1]$, $[0,+\infty]$, etc.). Concerning the starting space E, we assume that it is metric, with distance d. We are often more interested in some sublattices of $\mathcal{T}^E$, rather than in $\mathcal{T}^E$ itself, which includes number of "exotic" functions and rather irrealistic to model physical phenomena. A sublattice $\mathcal{L}' \subseteq \mathcal{T}^E$ is a class of functions which is closed under $\bigvee$ and $\bigwedge$ of $\mathcal{T}^E$ and which admits the same extrema as $\mathcal{T}^E$ itself. For example, the Lipschitz functions $\mathcal{L}_k$ of module k are defined by

$$f \in \mathcal{L}_k \qquad \Leftrightarrow \qquad |f(x) - f(y)| < kd(x,y) \qquad \forall\, x, y \in E \quad .$$

form a complete sublattice of $\mathcal{T}^E$ ([10], p.136). More generally, if we replace $kd(x,y)$ by $\varphi(d(x,y))$ in the above inequality, were $\varphi : \mathbb{R}_+ \to \overline{\mathbb{R}_+}$ is continuous at the origin, we delineate the class of the so-called "φ-continuous functions" [10]. For each φ, the associated φ-continuous functions form a complete sub-lattice of $\mathcal{T}^E$, i.e. a lattice where $\bigvee$ and $\bigwedge$ are provided by the *numerical* supremum and infimum at each point[68]. All theses lattices are sup-generated. It has been proved that under broad conditions, usual operators, such as dilations, openings, morphological filters, etc. map every φ-continuous lattice into another one [68], and that $\bigvee$ and $\bigwedge$ are continuous operators([41], p.64). Moreover, every function lattice where $\bigvee$ and $\bigwedge$ are identical to the numerical ones is totally distributive [41].

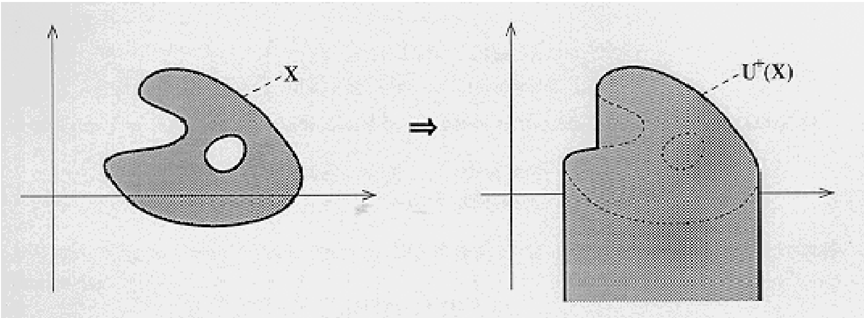


Figure 1.3: Umbra $U(X)$ of a set $X \subset \mathbb{R} \times \overline{\mathbb{R}}$.

In contrast, in the same context, the class $\mathcal{F}$ of the upper semi-continuous functions, that we introduce just below, forms a complete lattice, where ordering and infinum are the same as for $\mathcal{L}^E$, but where the supremum of $\{f_i\}$is the *closure* of $\bigvee f_i$. Therefore, $\mathcal{F}$ is not a sublattice of $\mathcal{L}^E$.

Finally, note that the situation of the continuous functions is worse, since they do not even form a complete lattice. For example, the infimum of all functions $\{x^\alpha,\ 0 \leq x \leq 1,\ \ \alpha \in \mathbb{Z}_+\}$ is not a continuous function.

1.9.10 Upper semi-continuous functions

Comments on functions and umbrae

Is it possible to identify the function lattice $\mathcal{F}$ with the *set* class of the associated subgraphs, or umbrae? Remember that to every function $f : E \longrightarrow \overline{\mathbb{R}}$ (and more generally to every set in $E\times\overline{\mathbb{R}}$, see Fig.1.3), corresponds the set $U(f)$ of $E \times \overline{\mathbb{R}}$ defined by the relations:

$$U(f) = \{(x, z) \in E \times \overline{\mathbb{R}}, f(x) \leq z\} \cup E_{-\infty}.$$

Although we have

$$f \leq g \iff U(f) \subseteq U(g) \ ,$$

the correspondence "function - umbra" is not an anamorphosis, for it is not a bijection (an umbra and its topological closure can yield the same function).

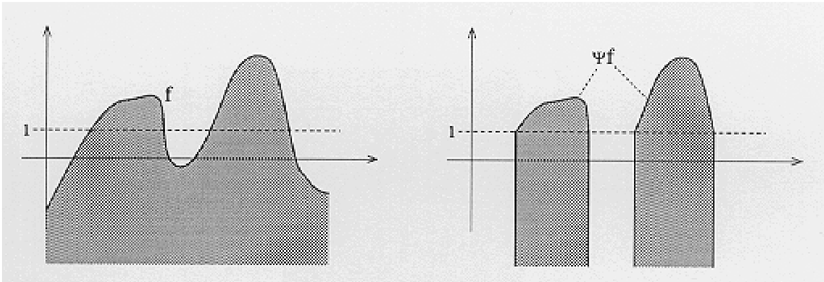


Figure 1.4: The threshold mapping ψ.

We will illustrate this point by considering the *threshold mapping* defined as follows:

$$[\psi(f)](x) = \begin{cases} f(x) & \text{when } f(x) \geq 1, \\ -\infty & \text{otherwise.} \end{cases} \tag{1.10}$$

This operation is shown on Fig. 1.4.

In set terms, the transformation ψ consists in intersecting the umbra $U(f)$ by the closed half space

$$E_1 = \{(x,y), x \in E, z \geq 1\},$$

and in taking the umbra of the result:

$$U(\psi(f)) = U[E_1 \cap U(f)] \cup E_{-\infty}. \tag{1.11}$$

If functions and umbrae were equivalent, then the two algorithms (1.10) and (1.11) should give the same result. Let's try and apply the two algorithms to the sup of the following family (see Fig. 1.5):

$$\begin{cases} f_i(x) &= 1 - 1/i & \text{when } |x| \leq 1, \\ f_i(x) &= -\infty & \text{otherwise.} \end{cases}$$

If the sup f of this family is understood in the sense of the function lattice, it is equal to:

$$\begin{cases} f(x) &= 1 & \text{when } |x| \leq 1, \\ f(x) &= -\infty & \text{otherwise,} \end{cases}$$

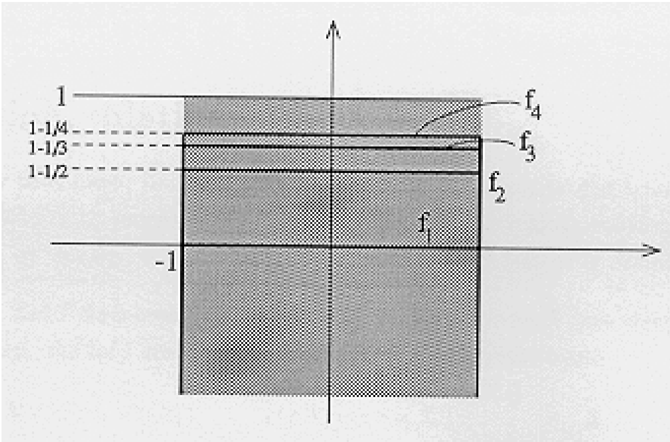


Figure 1.5: The family of functions (f_i).

and according to the rel. (1.10), $\psi f = f$. But if the sup is understood in the sense of the umbrae lattice, i.e.

$$U(f) = \bigcup_i U(f_i),$$

then from rel. (1.11), we derive $U[\psi(f)] = E_{-\infty}$, i.e. $\forall x \in E, \psi f(x) = -\infty$. In other words, in the Euclidean case, the function lattice and the set oriented lattice of umbrae are not equivalent at all. Nevertheless, in the discrete case of functions $f : \mathbb{Z}^n \longrightarrow \mathbb{Z}$, the two approaches coincide and one can transpose the way of reasoning from sets to functions.

Lattices of semi-continuous functions

Given a topological space E, a numerical function $f : E \to \overline{\mathbb{R}}$ is *upper semi-continuous* (u.s.c.) when its umbra $U(f)$ is a closed set in $E \times \overline{\mathbb{R}}$. The use of semi-continuity appears as soon as extrema are involved in the analysis under study, at least in continuous cases. For example, could we extract the maxima of the following function in $\mathbb{R}$ (see Fig. 1.6):

$$f(x) = \begin{cases} 1 - x^2 & \text{when } 0 < |x| < 1, \\ 0 & \text{when } |x| \geq 1 \text{ or } x = 0 \end{cases} \qquad ?$$

Actually, the maximum of such a function, although it is bounded, *does not exist.* Conversely, as soon as we refer to the "maximum" of a function over

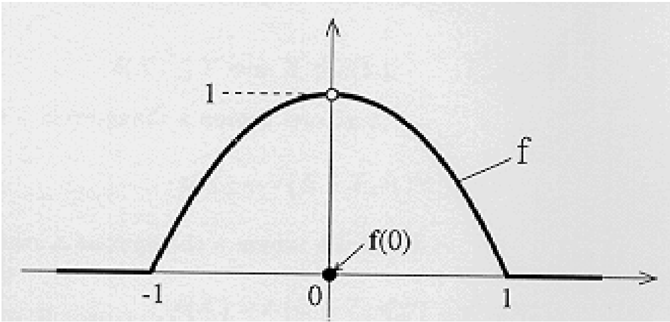


Figure 1.6: A function without a maximum (it is not u.s.c.).

a continuous space, we implicitely introduce the requirement that it is u.s.c. (or lower semi-continuous when looking for minima).

The class$\mathcal{F}_u$ of the upper semi-continuous functions $f : E \to \overline{\mathbb{R}}$ forms a complete lattice. In this lattice, the inf of a family $\{f_i\}$ is the function which admits, at each point $x \in \mathbb{R}^2$, the numerical inf of the $f_i(x)'$s, but the sup is the function whose umbra is the topological closure of the union of the umbrae of the f_i's.We have

$$\begin{aligned} \inf_i f_i &= \{f \in \mathcal{F}_u, U^+(f) = \bigcap_i U(f_i)\}, \\ \sup_i f_i &= \{f \in \mathcal{F}_u, U^+(f) = \overline{\bigcup_i U(f_i)}\}, \end{aligned}$$

a notation which shows that the lattice $\mathcal{F}_u$ and that of the closed upper umbrae are *anamorphic* (in this case, the identification between sets and functions works).

The cross sections

$$X_t(f) = \{x : f(x) \geq t\} \qquad -\infty \leq t \leq +\infty$$

are closed and monotonically decreasing sets of E, i.e.

$$t' < t \Longrightarrow X_{t'} \supseteq X_t \quad \text{and} \quad t' \uparrow t \Longrightarrow X_{t'} \downarrow X_t \ .$$

Conversely, any stack $\{X_t\}$ of closed sets satisfying these two conditions generates a unique u.s.c. function f with

$$f(x) = \sup\{t : x \in X_t\} \qquad x \in E$$

(see [60], p.426). In other words, $\mathcal{F}_u$ lattice turns out to be the most direct transposition of the closed sets to numerical functions. The pulses, which belong to lattice $\mathcal{F}_u$, are sup-generating co-primes, but not strong ones. Lattice $\mathcal{F}_u$ is distributive, and infinite $\bigwedge$-distributive, but not infinite $\bigvee$-distributive. For the classical Choquet toplolgy on closed sets, lattice $\mathcal{F}_u$ is CCO, but the nice symmetry between $\bigvee$ and $\bigwedge$ of the φ-continuous functions is now lost, since the $\bigvee$ only is continuous. This means that the only models for situations where both maxima and minima are involved are the φ-continuous lattices.

1.9.11 Multi-spectral images

Given the complete lattices $(\mathcal{L}_1, \leq_1)$, $(\mathcal{L}_2, \leq_2)$, ..., $(\mathcal{L}_d, \leq_d)$, define $\mathcal{M} = \mathcal{L}_1\times \mathcal{L}_2\times$ ... $\times$ $\mathcal{L}_d$; that is, $\mathcal{M}$ contains all d-tuples $(x_1, x_2,$ ..., $x_d)$ where $x_k \in \mathcal{L}_k$, $k = 1, 2, ..., d$. Furthermore, define the relation $\leq$ on $\mathcal{M}$ by $(x_1, x_2,$ ..., $x_d) \leq$ $(y_1, y_2,$ ..., $y_d)$, if $x_k \leq_k y_k$ for every k = 1,2,...,d.

Hereafter, we refer to this ordering as the *product ordering.* $(\mathcal{M},\leq)$ is a complete lattice. Product lattices are involved in color spaces for example.Clearly, multi-spectral images generate a lattice where the ordering, the sup and the inf are taken separately channel by channel.

1.9.12 Comparisons

The comparison of the five models is instructive. Some of them admit co-prime but not atoms (n^{os} 8-10) or *vice-versa* (n^{o} 4). Sup-generating families exist, which do not consist of atoms or co-primes (n^{o} 2).Some of them are Boolean , but not of $\mathcal{P}(E)$ type (n^{o} 3). Distributivity is also a deep distinction between lattices.

.

Chapter 2

Erosion, dilation

In the same way that linear image processing puts the emphasis on the transformations that commute with addition, morphology naturally stresses the transformations that commute with the sup or, by duality, with the inf. This results in the following definition:

Definition 2.1 *Let $\mathcal{L}$,$\mathcal{M}$ be two complete lattices. The mappings from $\mathcal{L}$ into $\mathcal{M}$ which commute with the* sup *(resp. the* inf*) are called dilations δ (resp. erosions ε)*

$$\delta(\vee x_i) = \vee\delta(x_i), \qquad \varepsilon(\wedge x_i) = \wedge\varepsilon(x_i), \quad x_i \in \mathcal{L} \tag{2.1}$$

with in particular $\delta(0_{\mathcal{L}}) = 0_{\mathcal{M}}$ and $\delta(m_{\mathcal{L}}) = m_{\mathcal{M}}$.

The following theorem characterizes these two operations and their links:

Theorem 2.2 *Let $\mathcal{L}$,$\mathcal{M}$ be two complete lattices. The dilations $\delta : \mathcal{L} \to \mathcal{M}$ and of the erosions $\varepsilon : \mathcal{M} \to \mathcal{L}$ correspond to one another through the duality relation*

$$\delta(x) \leq y \Longleftrightarrow x \leq \varepsilon(y), \qquad x, y \in \mathcal{L} \ , \tag{2.2}$$

called "Gallois adjunction" *and which occurs if and only if δ is a dilation and ε an erosion. To each dilation δ corresponds then the unique erosion ε:*

$$\varepsilon(x) = \vee\{b \in \mathcal{L} \ , \delta(b) \leq x\} \tag{2.3}$$

and to each erosion ε corresponds the unique dilation δ:

$$\delta(x) = \wedge\{b \in \mathcal{L}\ , \varepsilon(b) \geq x\}. \tag{2.4}$$

proof: We first show that rel(2.2)⇒rel(2.3). Denote by $\varepsilon^*(x)$ the right member of rel(2.3). We see from rel(2.2) that every element $b \leq \varepsilon^*(x)$ is smaller than $\varepsilon(x)$, hence $\varepsilon^*(x) \leq \varepsilon(x)$; conversely, the same adjunction, when applied to inequality $\varepsilon(x) \leq \varepsilon(x)$ shows that $\delta\varepsilon(x) \leq x$,hence $\varepsilon(x) \leq \varepsilon^*(x)$. This results in rel(2.3), and proves also the uniqueness, since if ε exists, it must be of the form (2.3).

We now look at implication (2.2)⇒(2.1) (first equality). Let $\{x_i, i \in I\}$ be a family in lattice $\mathcal{L}$. Applying adjunction (2.2) by taking x_i for x and $\vee\delta(x_i)$ for y, we have for all $i \in I$

$$\delta(x_i) \leq \vee\delta(x_i) \Leftrightarrow x_i \leq \varepsilon[\vee\delta(x_i)]$$

hence

$$\vee x_i \leq \varepsilon[\vee\delta(x_i)] \Leftrightarrow \delta(\vee x_i) \leq \vee\delta(x_i).$$

But we have also the reverse inequality, by increasingness of δ, so that the first equality of rel.(2.1) follows.

It remains to prove implication (2.1)⇒(2.2). With any dilation one can always associate the operation $\varepsilon^*(x)$ defined by the right member of rel(2.3). The form of this relation shows that $\delta(b) \leq x \Rightarrow b \leq \varepsilon^*(x)$. Conversely, since δ commutes under $\vee$, we have

$$b \leq \varepsilon^*(x) \Rightarrow \delta(b) \leq \delta\varepsilon^*(x) \Rightarrow \vee\{\delta(b), b \in \mathcal{L}\ , \delta(b) \leq x\} \leq x.$$

Finally, the increasingness of ε, and the fact that it commutes under $\wedge$ derive easily from rel(2.3), and rel(2.4) is a consequence of the general duality principle. □

The above notion of an *adjunction*, which goes back to E.Galois ([8], p.124, see also[17]) is well known in algebra. The concept of dilation was introduced by H. Minkowski (1901) for the Eclidean translation invariant case, and developped in lenght, in association with erosion and adjunction opening and closing, by G.Matheron in his two books of 1968 and 1975. The extension of these four operators to complete lattices appears in T.S. Blyth and M.F. Janowitz's book [9]. The above theorem was proved by J. Serra in the framework of complete lattices, for $\mathcal{L} = \mathcal{M}$ (the proof is the same as for $\mathcal{L} \neq \mathcal{M}$)([62, page 24], p.17), and enlarged to $\mathcal{L} \neq \mathcal{M}$ by H.J.A.M. Heijmans and Ch. Ronse ([23] and [24],p.51), among other results.

Corollary 2.1 *The classes of dilations $\delta : \mathcal{L} \to \mathcal{M}$ and of the erosions $\varepsilon : \mathcal{M} \to \mathcal{L}$ are two complete anamorphic lattices.*

Not only dilations and erosions are themselves increasing mappings, but both of them generate the whole class of increasing mappings. Indeed, we have the following representation theorem of J.Serra ([62, page 20], p.20):

Theorem 2.3 *Any mapping $\psi : \mathcal{L} \longrightarrow \mathcal{L}$ such that $\psi(E) = E$ is increasing if and only if it can be written as*

$$\psi = \vee\{\varepsilon_b, b \in \mathcal{L}\},$$

with the erosions ε_B given by

$$\varepsilon_b(x) = \begin{cases} m & if\ x = m \\ \psi(b) & if\ x \geq b \text{ and } x \neq m \\ 0 & otherwise. \end{cases} \tag{2.5}$$

(dual result for the dilation.)

proof: It is easy to verify that Eq (2.5) defines an erosion on $\mathcal{L}$. Now, given element x and taking the supremum of all images $\varepsilon_b(x)$ as b spans lattice, we find, by increasingness of ψ

$$\vee\{\varepsilon_b(x), b \in \mathcal{L}\} = \vee\{\psi(b) : b \leq x\} = \psi(x)$$

Conversely, since erosion is increasing, the upper bound of an arbitrary family of erosions is also inreasing □

2.1 $\mathcal{P}(E)$ type lattice

In this section and in the following, we would like to compare the two lattices which model the binary and the grey-tone images.

The first one is the boolean lattice $\mathcal{P}(E)$, where E is $\mathbb{R}^n$ or $\mathbb{Z}^n$ for example. We can look at mappings from $\mathcal{P}(E)$ into itself as extensions of mappings from E into $\mathcal{P}(E)$. In the following, lower-case letters such as x, y, a, b denote elements of E, or points, and capital letters denote elements of $\mathcal{P}(E)$. A point $x \in E$, when considered as an element of $\mathcal{P}(E)$, is written as $\{x\}$. The letter δ denotes the mapping $E \longrightarrow \mathcal{P}(E)$, which generates a dilation, as well as the dilation from $\mathcal{P}(E)$ into itself. We define a **structuring function** on $\mathcal{P}(E)$ as any mapping $\delta : E \longrightarrow \mathcal{P}(E)$. Then, we have [62, page 41]:

Theorem 2.4 *Let E be an arbitrary set. The datum of any mapping $\delta : E \longrightarrow \mathcal{P}(E)$ is equivalent to that of a dilation from $\mathcal{P}(E)$ into itself, again symbolized by δ, and defined by the relation*

$$\delta(X) = \bigcup_{x \in X} \delta(x), \qquad X \in \mathcal{P}(E). \tag{2.6}$$

Conversely, any dilation of $\mathcal{P}(E)$ into itself determines a unique structuring function $\delta : E \longrightarrow \mathcal{P}(E)$.

2.1.1 The three dualities

In any boolean algebra $\mathcal{P}(E)$, the duality w.r. to the complementation associates with each mapping ψ the operation $\psi^* = \complement \psi \complement$, where $\complement$ designates the complement operator, as expressed by

$$\psi^*(X) = [\psi(X^C)]^C.$$

In the case of the dilation δ, we find

$$\delta^*(X) = [\bigcup_{x \in X^C} \delta(x)]^C = \bigcap_{x \in X^C} [\delta(x)]^C. \tag{2.7}$$

δ^*, which obviously commutes with the inf, is an erosion. $\delta^*(X)$ consists of the points that are not *descendant* from any point in the complement of X (that are not included in any $\delta(x)$ when $x \in X^C$), i.e:

1. those whose ancestors are all included in X,
2. those that do not have ancestors (a fixed part S, which remains the same for any set X).

We form another duality notion by operating on the structuring function with the **transposition** $\delta \longmapsto \check{\delta}$, i.e:

$$\check{\delta}(x) = \{y \in E, x \in \delta(y)\}.$$

The transpose $\check{\delta}(x)$ of $\delta(x)$ is made of the set of points from which x descends, hence $\check{\check{\delta}} = \delta$. The structuring function $\check{\delta}$ generates the dilation $\check{\delta}$:

$$\check{\delta}(X) = \{y \in E, \delta(y) \cap X \neq \emptyset\}. \tag{2.8}$$

From the two relations (2.7) and (2.8), we derive the links between these two dualities, and the basic one, namely $\delta \leftrightarrow \varepsilon$ (see rel. (2.2)):

$$\varepsilon = (\check{\delta})^* \qquad \check{\varepsilon} = \delta^* \qquad \varepsilon^* = \check{\delta}. \tag{2.9}$$

2.1.2 Translation invariance

We now assume that E is equipped with a translation (e.g. $E = \mathbb{Z}^n$ or $E = \mathbb{R}^n$), and that the dilation δ is translation invariant, i.e. is a t-dilation. In other words, the structuring function $\delta(\{h\})$ at point $\{h\}$ is deduced from that of the origin (denoted $\delta(\{o\} = B)$ by translation: $\delta(\{h\} = B_h = \{B+h, b \in B\}$. The set B is called structuring element. We see from relation (2.6) that

$$\begin{aligned} \delta(X) &= X \oplus B = B \oplus X \\ &= \bigcup_{x \in X} B_x \\ &= \{b + x, x \in X, b \in B\} \\ &= \bigcup_{b \in B} X_b. \end{aligned}$$

The t-dilation δ is classically known as **Minkowski addition** between sets X and B. By duality under complementation, it gives

$$\delta^*(X) = \check{\varepsilon}(X) = \bigcap_{b \in B} X_b = X \ominus B$$

and by lattice duality:

$$\varepsilon(X) = \bigcap_{b \in \check{B}} X_b, \qquad \text{with } \check{B} = \{-b, b \in B\}.$$

Both operations are Minkowski subtractions of X by B and $\check{B}$ respectively. According to a classical result due to G. Matheron, any increasing t-mapping is a union of t-erosions, and also an intersection of t-dilations [38, page 221]. More precisely,

Theorem 2.5 *Let ψ be a translation invariant increasing mapping. Then, for any $X \in \mathcal{P}(E)$,*

$$\psi(X) = \bigcup\{X \ominus \check{B}, B \in \mathcal{P}(E), o \in \psi(B)\} = \bigcap\{X \oplus \check{B}, B \in \mathcal{P}(E), o \in \psi^*(B)\}.$$

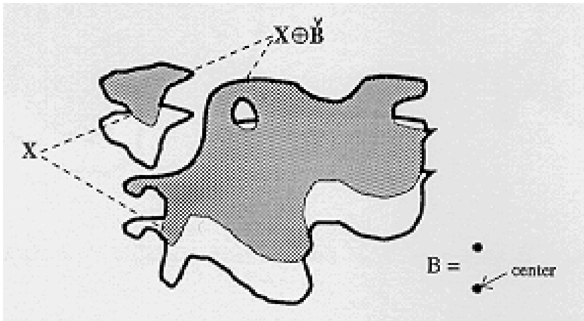


Figure 2.1: Dilation of a set X by a bipoint $\check{B}$

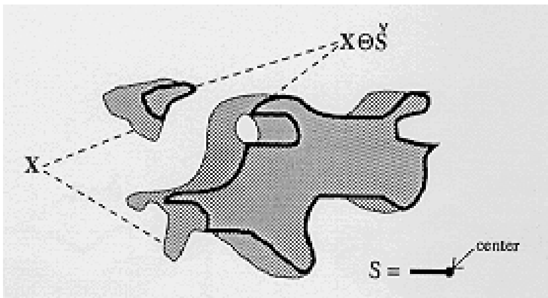


Figure 2.2: Erosion of a set X by a segment $\check{S}$.

2.1.3 Examples

We will now present some examples of dilations and erosions in the lattice of sets $\mathcal{P}(\mathbb{R}^2)$. Fig. 2.1 shows the dilation of a set X by a bipoint B, i.e. the set $X \oplus \check{B}$. Similarly, Fig. 2.2 illustrates the effect of an erosion of X by a segment S. On Fig. 2.3, the same set is dilated and eroded by a disc D (*Euclidean* dilation and erosion). The dilation by a disc is then compared, on Fig. 2.4, to that by an hexagon H of similar size. One can remark that many parts on the boundary of $X \oplus \check{H}$ are parallel to the vertices of H.

Lastly, Fig. 2.5 illustrates the algorithm which is used for performing a *geodesic* dilation of a set Y inside a set X (see Chapter 7).

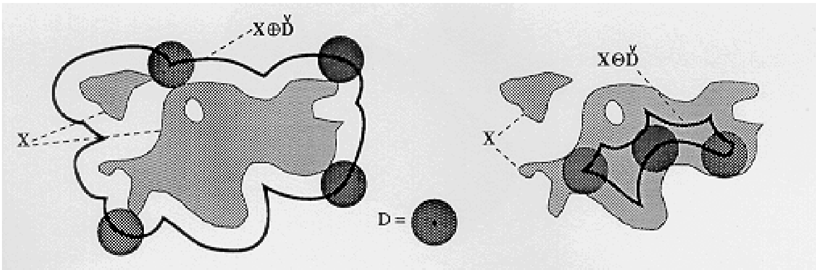


Figure 2.3: Dilation and erosion of X by a disc $D = \check{D}$.

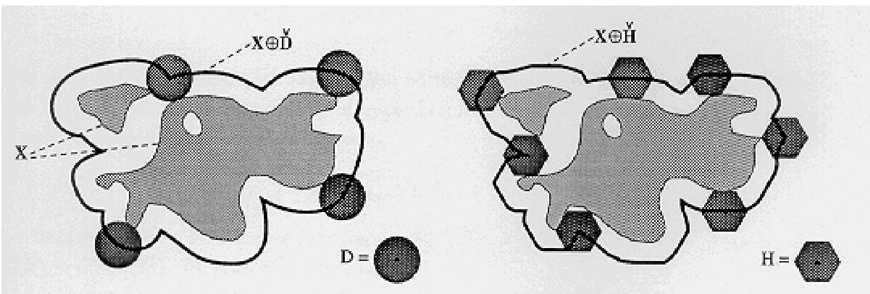


Figure 2.4: Comparison between the dilations of X by a disc and by an hexagon. Note that these structuring elements are symmetrical: $D = \check{D}$ and $H = \check{H}$.

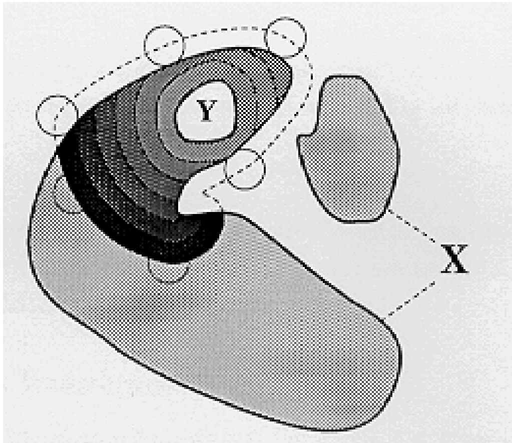


Figure 2.5: Successive geodesic dilations of set Y inside set X.

2.2 Lattices of functions

The lattice $\mathcal{F}(E, \overline{\mathbb{R}})$ of the functions $f : E \longrightarrow \overline{\mathbb{R}}$ shares several properties with the previous one, but it differs from $\mathcal{P}(E)$ by two major aspects:

1. it is not complemented,
2. when additions or subtractions are involved, they may lead to indetermination, of the type $+\infty - \infty$, since the range of variation is $\overline{\mathbb{R}}$.

We will now study $\mathcal{F}(E, \overline{\mathbb{R}})$ by following the same plan as for $\mathcal{P}(E)$.

2.2.1 Generation of dilations from structuring functions

Call *pulse* $u_{h,z}$ a function whose value is z at point $h \in E$, and $-\infty$ elsewhere [23]:

$$\forall x \in E, \quad u_{h,z}(x) = \begin{cases} z & \text{when } x = h, \\ -\infty & \text{otherwise.} \end{cases}$$

For numerical functions, the class $\mathcal{I}(E)$ of the pulses is a sup-generator equivalent to that of the points $(h, z) \in E \times \overline{\mathbb{R}}$. Clearly, any function $f \in \mathcal{F}(E, \overline{\mathbb{R}})$

is the sup of upper bounded pulses smaller than itself (just as a set is the union of the points it contains):

$$f = \sup\{u_{h,z}, h \in E, z < f(h)\}.$$

Introduce now a structuring function on $\mathcal{F}(E, \overline{\mathbb{R}})$ as any upper bounded mapping $\delta : \mathcal{I}(E) \longrightarrow \mathcal{F}(E, \overline{\mathbb{R}})$. We then have [62, page 185]:

Theorem 2.6 *any structuring function is equivalent to a dilation from $\mathcal{F}(E, \overline{\mathbb{R}})$ into itself, defined by the relation*

$$\delta(f) = \sup\{\delta(u_{h,z}), h \in E, z < f(h)\}. \tag{2.10}$$

Conversely, any dilation $\delta : \mathcal{F}(E, \overline{\mathbb{R}}) \longrightarrow \mathcal{F}(E, \overline{\mathbb{R}})$ induces a **unique** *structuring function obtained by restricting δ to $\mathcal{I}(E)$.*

2.2.2 Dualities

The transposition duality extends immediately to functions, by replacing points by pulses. The duality with respect to the complementation is replaced by all those given by the relation

$$\psi^*(f) = m - \psi(m - f), \tag{2.11}$$

as m spans the class of the real numbers. In practice, ψ often commutes with vertical shifts, i.e. $\psi(f + m) = \psi(f) + m$. Then, all the relations (2.11) are equal to $\psi^*(f) = -\psi(-f)$, and the three expressions (2.9) extend to functions.

2.2.3 Translation invariances

We can consider either a translation operation t'_h, by vector $h \in E$, or a translation operation $t_{h,z}$ by a vector $(h, z) \in E \times \overline{\mathbb{R}}$. The two corresponding formulas are:

$$\begin{aligned} (t_{h,z} f)(x) &= f(x - h) + z, \\ (t'_h f)(x) &= f(x - h). \end{aligned}$$

We shall focus on the t-invariant mappings, which are the most useful in practice. Saying that dilation δ is invariant with respect to translations is equivalent to saying that the *structuring function* δ is the same everywhere, i.e. if $g = \delta u_{0,0}$ is the transform of the origin-impulse, then $\forall x \in E, \delta u_{h,z}(x) = g(x-h)+z$. Then, the expression (2.10) of the dilation δ takes the following simpler form:

$$(\delta f)(x) = \sup\{g(x-h)+z, z < f(h), h \in E\}, \quad f \in \mathcal{F}(E, \overline{\mathbb{R}}).$$

Note that the operand $g(x-h)+z$ cannot take the undetermined form $+\infty - \infty$ since, for all h, x and z, each of the two numbers $g(x-h)$ and z is $< +\infty$. Hence, we have finally

$$(\delta f)(x) = \sup\{g(x-h)+f(h), h \in E\}, \quad f \in \mathcal{F}(E, \overline{\mathbb{R}}). \tag{2.12}$$

The two dual erosions ε and $\check{\varepsilon}$ of δ are given by the following formulae:

$$\begin{aligned} (\varepsilon f)(x) &= \inf\{f(h) - g(h-x), h \in E\}, \\ (\check{\varepsilon} f)(x) &= \inf\{f(h) - g(x-h), h \in E\}. \end{aligned}$$

Similarly to theorem 2.5, any increasing mapping $\psi : E \longrightarrow \overline{\mathbb{R}}$ which is t-invariant may be decomposed into a sup of erosions as well as into an inf of dilations (same proof as for theorem 2.5).

2.2.4 Planar increasing mappings

An increasing mapping $\psi : \mathcal{F}(E, \overline{\mathbb{R}}) \longrightarrow \mathcal{F}(E, \overline{\mathbb{R}})$ is said to be planar, or flat, or again to be a *stack mapping* [79] when for any $z \in \overline{\mathbb{R}}$, the class $\mathcal{C}_z$ of the half cylinders of top level z is closed under ψ:

$$\begin{aligned} G_z \in \mathcal{C}_z \quad &\Longleftrightarrow \quad \forall x \in E, G_z(x) = z \text{ or } G_z(x) = -\infty. \\ \psi \text{ is planar} \quad &\Longleftrightarrow \quad \forall z \in \overline{\mathbb{R}}, \psi(\mathcal{C}_z) = \mathcal{C}_z \Longleftrightarrow \forall z \in \overline{\mathbb{R}}, \forall G_z \in \mathcal{C}_z, \psi(G_z) \in \mathcal{C}_z. \end{aligned}$$

With any function $f \in \mathcal{F}(E, \overline{\mathbb{R}})$, we can associate its maximum cylinders $C_z(f)$:

$$C_z(f) = \sup\{G_z \in \mathcal{C}_z, G_z \leq f\}.$$

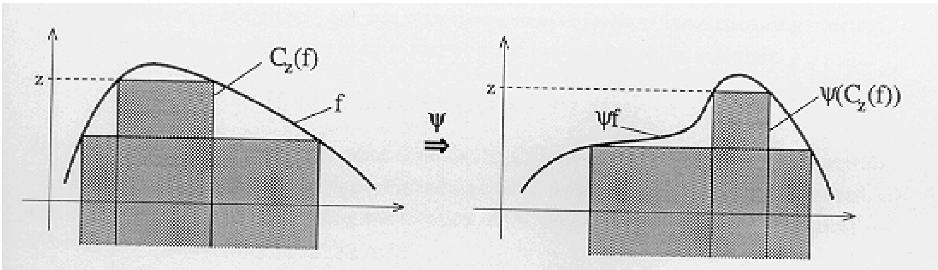


Figure 2.6: An example of a planar increasing mapping.

Then, the sup of the family $(C_z(f))_{z\in\overline{\mathbb{R}}}$ generates f:

$$f = \sup_{z\in\overline{\mathbb{R}}} C_z(f).$$

If ψ is increasing, we have, by growth

$$\psi C_z(f) \geq \sup\{\psi G_z, G_z \in \mathcal{C}_z, G_z \leq f\}.$$

Furthermore, if ψ is planar, then $\psi G_z \in \mathcal{C}_z$ and the inequality becomes

$$\psi C_z(f) \geq \sup\{G'_z, G'_z \in \mathcal{C}_z, G'_z \leq \psi f\}.$$

Now, by construction, $\psi C_z(f)$ is itself one of the G'_z, so that the above relation turns out to be an equality. In this equality, the right member is nothing but $C_z\psi(f)$. Finally:

$$\psi C_z(f) = C_z\psi(f),$$

hence:

$$\psi(f) = \sup_{z\in\overline{\mathbb{R}}} C_z\psi(f) = \sup_{z\in\overline{\mathbb{R}}} \psi C_z(f). \tag{2.13}$$

This relation is illustrated by Fig. 2.6.

In other words, the planarity of the mapping ψ allows us to process f threshold by threshold. A series of results derive directly from the key relation (2.13), namely:

Theorem 2.7 *Let ψ be a planar increasing mapping from $\mathcal{F}(E, \overline{\mathbb{R}})$ into itself. Then:*

1. *the class of step functions with k levels (k arbitrary) is closed under ψ,*
2. *ψ commutes with anamorphosis.*

An anamorphosis is a strictly increasing and continuous point mapping $s : \overline{\mathbb{R}} \longrightarrow \overline{\mathbb{R}}$. For example, if $s = \exp$, then for any planar increasing mapping ψ, we have:

$$\psi \exp(f) = \exp \psi(f), \qquad f \in \mathcal{F}(E, \overline{\mathbb{R}}),$$

i.e. "vertical" and "horizontal" dimensions are treated **independently.**

A basic example of a planar increasing mapping consists in the dilation of f by a function b which is equal to 0 on its support B and to $-\infty$ everywhere else. Then, the relation (2.12) becomes:

$$(\delta f)(x) = \sup\{f(x-h), h \in B\} = (f \oplus B)(x). \tag{2.14}$$

By duality w.r. to lattice, relation (2.13) yields:

$$(\varepsilon f)(x) = \inf\{f(x+h), h \in B\} = (f \ominus \check{B})(x). \tag{2.15}$$

Fig. 2.7 shows an example of a planar dilation and of a planar erosion of a function f.

2.3 Digital implementations

In this section, we concentrate upon implementations of t-dilations (and t-erosions), which are the basic stones for building up more sophisticated algorithms.

When the dilation is planar, it is produced for functions in the same way as for sets. One has merely to replace union by sup and intersection by inf (e.g. refer to relation (2.14)). When the dilation is not planar, one can scan the successive levels of the structuring function, or use Steiner decomposition. In both cases, we shall use the following notation:

$$[f \oplus (1\ \mathbf{0})](x) = \sup\{f(x+1)+1, f(x)\}.$$

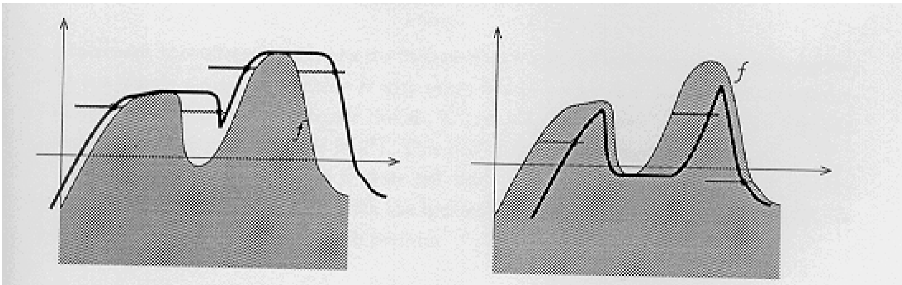


Figure 2.7: Dilation (a) and erosion (b) of a function f by a planar structuring element B.

The number associated with each point denotes the altitude of the corresponding structuring function (here a function whose support is reduced to an horizontal doublet). When needed, a bold character is used for indicating the location of the origin. The elementary "spherical"—and centered—structuring functions are:

- the cube: 9 pixels, on two successive levels
- the octahedron: 5 pixels, on two successive levels
- the rhombododecahedron: 9 pixels, on two successive levels
- the cuboctahedron: 9 pixels, on two successive levels.

They are represented on Fig. 2.8.

The elementary *rhombododecahedron* R can be represented (as in Fig. 2.9) by taking the spacing of the horizontal square grid to be $\sqrt{2}$.

The Steiner rhomb kR of size k is obtained by taking k dilations of R:

$$kR = \underbrace{R \oplus R \oplus \ldots \oplus R}_{k \text{ times}} = R^{\oplus k}.$$

As k increases, the difference between the Steiner rhomb and the ball becomes more apparent, but it is a simple matter to combine R with other Steiner polyhedra, such as the cuboctahedron, or simply with another Steiner rhomb,

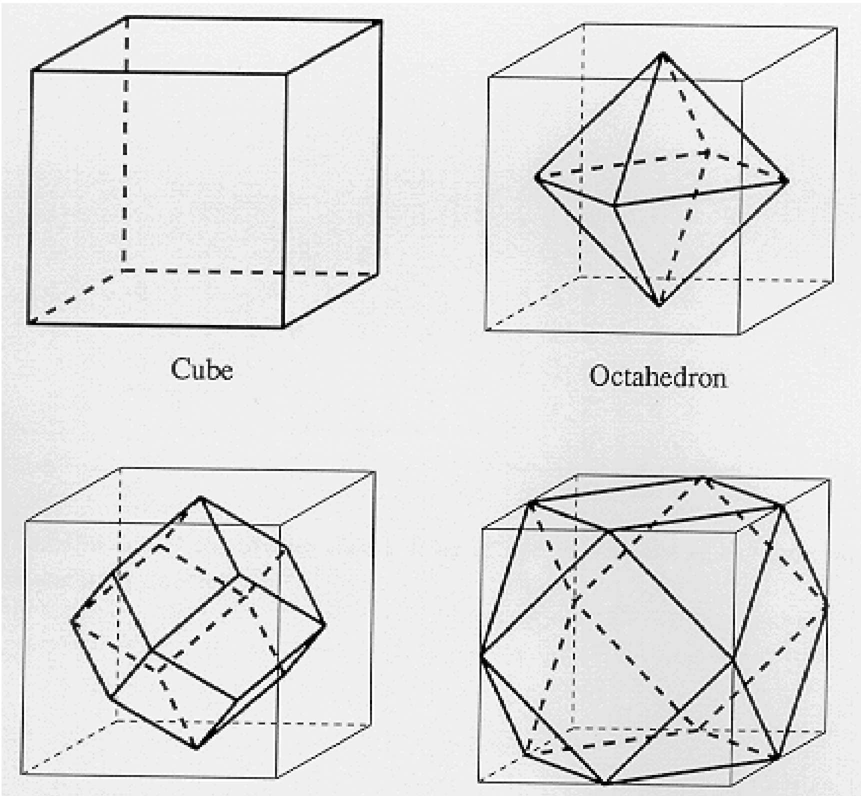


Figure 2.8: Basic spherical shapes in $\mathbb{Z}^3$. The plane $z = 0$ corresponds to the median horizontal section of the cube. The structuring functions derive from these sets by taking their umbrae.

$$\begin{array}{ccccc} . & . & . & . & . \\ . & 0 & 1 & 0 & . \\ . & 1 & 2 & 1 & . \\ . & 0 & 1 & 0 & . \\ . & . & . & . & . \end{array} = (1\ 0) \oplus (0\ 1) \oplus \begin{pmatrix} -1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Figure 2.9: The elementary rhombododecahedron R (i.e. Steiner rhomb of $\mathbb{Z}^3$) and its decomposition in four dilations by segments. The complex shape of the polyhedron has been decomposed into four simple structuring functions, whose implementation is very simple and extremely efficient.

$$\underbrace{\begin{array}{ccc} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{array}}_{R} \oplus \underbrace{\begin{array}{ccccc} & & 0 & & \\ & 1 & . & 1 & \\ 0 & . & 2 & . & 0 \\ & 1 & . & 1 & \\ & & 0 & & \end{array}}_{R^*} = \underbrace{\begin{array}{ccccccc} & & 0 & 1 & 0 & & \\ & 1 & 2 & 2 & 2 & 1 & \\ 0 & 2 & 3 & 3 & 3 & 2 & 0 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 0 & 2 & 3 & 3 & 3 & 2 & 0 \\ & 1 & 2 & 2 & 2 & 1 & \\ & & 0 & 1 & 0 & & \end{array}}_{R \oplus R^*}$$

Figure 2.10: Dilation of the Steiner rhomb R by R^* (R^* is obtained in the same way as R, by four dilations by doublets).

$$\underbrace{\begin{matrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{matrix}}_{C} = \underbrace{\begin{matrix} 0 & . & 0 \\ . & . & . \\ 0 & . & 0 \end{matrix}}_{C_0} \cup \underbrace{\begin{matrix} . & 1 & . \\ 1 & 1 & 1 \\ . & 1 & . \end{matrix}}_{C_1}$$

Figure 2.11: The elementary cuboctahedron C and its decomposition into two successive horizontal planes.

R^*, that is constructed at 45° to the first one (exactly as we construct octagons in $\mathbb{Z}^2$). This possibility is illustrated in Fig. 2.10.

The elementary *cuboctahedron* C does not lead to a sequence of segments. It has the decomposition shown in Fig. 2.11 (with the horizontal spacing being again equal to $\sqrt{2}$).

To dilate a function f by C, it suffices to perform

$$f_1 = f \oplus \begin{pmatrix} 0 & . & 0 \\ . & . & . \\ 0 & . & 0 \end{pmatrix}, \qquad f_2 = f \oplus \begin{pmatrix} . & 0 & . \\ 0 & 0 & 0 \\ . & 0 & . \end{pmatrix}$$

and then to compute the sup between f_1 and $f_2 + 1$:

$$f \oplus C = \sup\{f_1, f_2 + 1\}.$$

Chapter 3

Openings and closings

This chapter is a go-between from dilations to morphological filtering. Here, the two basic references are [62, chapters 7 & 17] and [38, chapters 1 & 5]. We shall see how, by looking for an *inverse* to the dilation—i.e. for an impossibility—we find a new operation, the *morphological closing,* whose three basic properties are extremely useful. We shall then try and keep these properties as axioms for the general concept of an (algebraic) closing. The notion of an *opening* is introduced by duality. It satisfies two of the three basic properties of the closing, that will become the two axioms of the morphological filtering in the next chapter.

3.1 Adjunction opening and closing

Generally, in a complete lattice $\mathcal{L}$, the dilation $X \longrightarrow \delta(X)$ and the erosion $X \longrightarrow \varepsilon(X)$ do not admit inverses and there is no way for determining **one** element X from the images $\delta(X)$ or $\varepsilon(X)$. However, starting from a dilation and then performing the dual erosion (or the contrary), we always have either an upper, or a lower bound according to the situation at hand.

Indeed, if we take $\delta(X)$ for the set Y in Galois adjunction (2.2), the left inclusion is satisfied, so $X \leq \varepsilon\delta(X)$, and by duality:

$$\delta \circ \varepsilon(X) \leq X \leq \varepsilon \circ \delta(X),$$

or in terms of operators:

$$\delta\varepsilon \leq I \leq \varepsilon\delta.$$

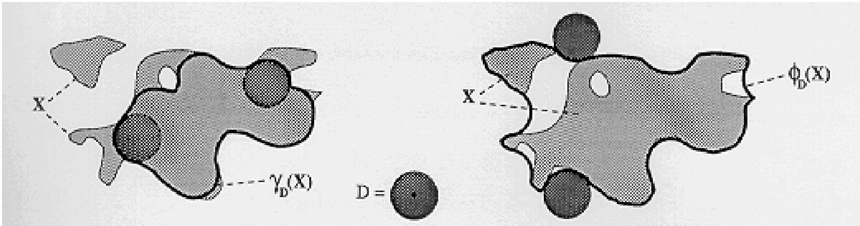


Figure 3.1: Examples of a morphological opening and of a morphological closing of a set X by a disc D.

We say that $\varepsilon\delta$ is **extensive** (larger than the identity mapping) and that $\delta\varepsilon$ is **anti-extensive**. Both operations are also **increasing** as the product of increasing mappings. Now, $\varepsilon\delta \geq I$ implies, by growth, that $\delta\varepsilon\delta\varepsilon \geq \delta\varepsilon$, whereas $\delta\varepsilon \leq I$ implies the inverse inequality. Hence $\delta\varepsilon = \delta\varepsilon\delta\varepsilon$, i.e. is **idempotent** (as well as $\varepsilon\delta$, by duality). The three properties of $\varepsilon\delta$ characterize what is called a *closing,* in algebra, and those of $\delta\varepsilon$ an *opening.* We shall call these two operators *adjunction*, or sometimes *morphological,* to indicate that they are generated from a dilation and its dual erosion, and we denote:

$$\gamma_m = \delta\varepsilon \qquad\qquad \varphi_m = \varepsilon\delta \tag{3.1}$$

Fig. 3.1 shows an example of a adjunction opening and of a adjunction closing of a set X in the plane. In this 2-D case, a adjunction opening may remove three types of features: capes, isthma and islands. By duality, a adjunction closing may fill gulfs, channels and lakes.

Let $Z = \delta(X)$ be the dilation image of an arbitrary element $X \in \mathcal{L}$. We have:

$$\gamma_m(Z) = \delta\varepsilon\delta(X) \quad \begin{array}{ll} \geq \delta(X) & \text{by extensivity of } \varepsilon\delta \\ \leq \delta(X) & \text{by anti-extensivity of } \delta\varepsilon \end{array}$$

Hence $\gamma_m(Z) = Z$, i.e. Z belongs to the class $\mathcal{B}$ of the *invariant elements* of $\mathcal{L}$ under γ_m. Conversely if $Z \in \mathcal{B}$, then $Z = \delta(\varepsilon(Z))$ i.e. is the dilate of an element of $\mathcal{L}$. To summarize, we have the following theorem:

Theorem 3.1 *Given a dilation δ on lattice $\mathcal{L}$ and its dual erosion ε, the composition products $\gamma_m = \delta\varepsilon$ and $\varphi_m = \varepsilon\delta$ are respectively an opening and a*

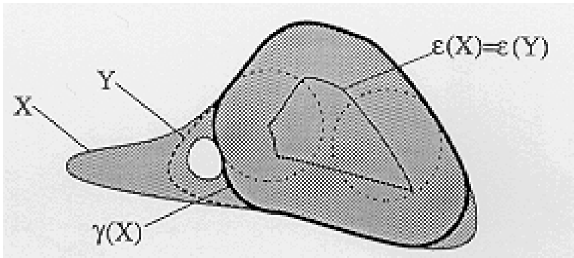


Figure 3.2: $\gamma_m(X)$ is the smallest element $Y \subseteq \mathcal{T}$ such that $\varepsilon(Y) = \varepsilon(X)$.

closing on $\mathcal{L}$, called adjunction opening and closing. The invariance domain of the former is the image of $\mathcal{L}$ under δ and that of the latter forms the image of $\mathcal{L}$ under ε.

Corollary 3.1 *Given $X \in \mathcal{L}$, $\gamma_m(X)$ is the smallest inverse image of X under ε, and $\varphi_m(X)$ is the largest one under δ.*

This corollary is illustrated by Fig. 3.2.

proof: Suppose that $Y \in \mathcal{L}$ is such that $\varepsilon(Y) = \varepsilon(X)$. Then, a fortiori, $Z = \varepsilon(X) \leq \varepsilon(Y)$ and thus, applying Galois adjunction (2.2), $\delta(Z) \leq Y$, or:

$$\gamma_m(X) \leq Y.$$

By duality, we have also

$$\forall Y \in \mathcal{L}, \delta(Y) = \delta(X) \Longrightarrow Y \leq \psi_m(X),$$

which completes the proof. $\square$

Corollary 3.2 *If $\mathcal{B}$ and $\mathcal{B}'$ stand for the invariance domains of γ_m and φ_m respectively, then*

$$\begin{aligned} \gamma_m(X) &= \vee\{B, B \in \mathcal{B}, B \leq X\} \\ \varphi_m X) &= \wedge\{B, B \in \mathcal{B}', B \geq X\} \end{aligned} \tag{3.2}$$

proof: From relation (2.3), we have

$$\begin{aligned}\gamma_m(X) &= \delta(\varepsilon(X)) = \delta(\vee\{B \in \mathcal{L}\ , \delta(B) \leq X\}) \\ &= \vee\{\delta(B), B \in \mathcal{L}\ , \delta(B) \leq X\},\end{aligned}$$

but according to the theorem, $\mathcal{B} = \{\delta(B), B \in \mathcal{L}\ \}$. Hence, we get relation (3.2). As concerns relation (3.2), it has a dual proof. □

Example:

We have seen in § 2.2.4 that **planar** increasing mappings preserve vertical walls. Fig. 3.3 typically illustrates this point by showing the morphological opening of a 1-D function by an horizontal segment. Unlike this kind of opening, **circular** openings (i.e. openings with discs) do not preserve the vertical parts of the 1-D functions on which they act. In this case, Fig. 3.3 clearly indicates changes of slope. The same remarks apply in the 2-D case and the experimenter must choose between one approach or the other acording to his purpose. It should be noticed that "planar" structuring elements are most of the time preferred, since the computation the corresponding openings and closings can be done more efficiently than with 3-D structuring elements.

3.2 Algebraic openings and closings

The important corollary 3.2 directly associates an opening γ_m with its invariant elements, without referring to the intermediary erosion and dilation. Should it be also true for any algebraic opening γ, i.e. for any operation on $\mathcal{L}$ which is increasing, anti-extensive and idempotent? Let $\mathcal{B}$ be the invariance domain of such a γ, and B be an invariant element, $B \leq X$. Then (by increasingness) $B = \gamma(B) \leq \gamma(X)$, hence $\gamma(X) \geq \vee\{B, B \in \mathcal{B}, B \leq X\}$. But $\gamma(X) \in \mathcal{B}$ (by idempotence) and $\gamma(X) \leq X$ (by anti-extensivity), therefore $\gamma(X)$ is one of the B of the right member. Thus, relation (3.2) is valid for any opening.

Conversely, start from an arbitrary part $\mathcal{B}_0$ of lattice $\mathcal{L}$ and let $\mathcal{B}$ be the class closed under union generated by $\mathcal{B}_0$. The operation defined by

$$\gamma(X) = \vee\{B, B \in \mathcal{B}_0, B \leq X\}$$

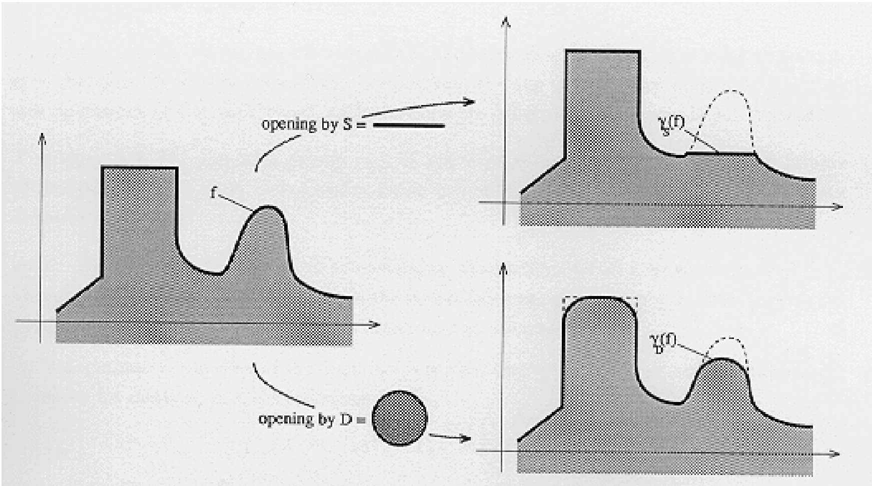


Figure 3.3: Comparison between the openings of a 1-D function by an horizontal segment S and by a disc D.

is increasing and anti-extensive. Moreover, $\gamma(X) = X$ iff $X \in \mathcal{B}$. The product $\gamma \circ \gamma$ is smaller than γ (growth and anti-extensivity), but also:

$$\begin{aligned}\gamma\gamma(X) &\geq \vee\{\gamma(B), B \in \mathcal{B}_0, B \leq X\} \\ &= \vee\{B, B \in \mathcal{B}_0, B \leq X\} \\ &= \gamma(X).\end{aligned}$$

We may therefore state the following:

Theorem 3.2 *An operation γ (resp. φ) on $\mathcal{L}$ is an opening (resp. a closing) if and only if there exists a class $\mathcal{B} \subseteq \mathcal{L}$, closed under union (resp. intersection) such that*

$$\begin{aligned}\gamma(X) &= \vee\{B, B \in \mathcal{B}, B \leq X\} \\ \varphi(X) &= \wedge\{B, B \in \mathcal{B}, B \geq X\}.\end{aligned}$$

$\mathcal{B}$ *is the* invariance domain *of γ (resp. φ).*

In other words, we can approach openings and closings either directly or via their invariance domains. Now, what about the *composition,* the *sup* or the *inf* of openings. Are they still operations of the same type? As far as sups are concerned, the anwer is yes. Indeed:

Theorem 3.3 *The sup of a family* (γ_i) *of openings is again an opening, whose domain of invariance is the class closed under union generated by the union of the* $\mathcal{B}_i$ *(invariance domains of the* γ_i*'s).*

proof: Clearly, $\vee\gamma_i$ is increasing and anti-extensive. Furthermore, for all i, we have $\gamma_i \circ (\vee\gamma_i) \geq \gamma_i$. Therefore, $(\vee\gamma_i) \circ (\vee\gamma_i) \geq (\vee\gamma_i)$, and also the inverse inclusion, since $(\vee\gamma_i) \leq I$. This gives us the idempotence. The domain of invariance is determined as was done before. $\square$

Unfortunately, the class of the openings is neither closed under $\wedge$, nor under composition. Consider for instance in $\mathbb{Z}$ the following set:

$$X = \texttt{..1111..111111..1111...}$$

and the two structuring elements

$$A = \texttt{.1.....1.} \quad \text{and} \quad B = \texttt{.11111.}$$

Denoting γ_A and γ_B, the associated morphological openings, we have:

$$\gamma_A(X) = X \quad \text{and} \quad \gamma_B(X) = \quad \texttt{.111111.}$$

and

$$\gamma_B \circ \gamma_A(X) = \gamma_B(X) \neq \gamma_A \circ \gamma_B(X) = \emptyset.$$

Hence:

$$(\gamma_B\gamma_A)(\gamma_B\gamma_A) \neq (\gamma_B\gamma_A),$$

and

$$(\gamma_B \wedge \gamma_A)(\gamma_B \wedge \gamma_A)(X) = \emptyset \neq (\gamma_B \wedge \gamma_A)(X) = \gamma_B(X).$$

3.3 Algebraic openings and adjunction openings

This last theorem shows in particular that any upper bound of adjunction openings γ_i, with corresponding domains of invariance $\mathcal{B}_i$, is still an opening, but which no longer results from an adjunction. It has as for domain of invariance the class closed under the sup (in $\mathcal{L}$) generated by the union of the $\mathcal{B}_i$ (in the space $\mathcal{P}(\mathcal{L})$ of all subsets of $\mathcal{L}$).

Consider now the inverse problem. Starting with an arbitrary opening $\gamma : \mathcal{L} \to \mathcal{L}$, we can consider it as a supremum of adjunction openings? Since any algebraic opening γ on $\mathcal{L}$, with invariant domain $\mathcal{B}$, is the smallest extension to $\mathcal{L}$ of the identity on $\mathcal{B}$, we can write

$$\gamma(X) = \vee\{B : B \in \mathcal{B}, B \leq X\} \qquad \forall X \in \mathcal{L}. \tag{3.3}$$

Therefore class $\mathcal{B}$ is closed under supremum. Conversely, the class closed under supremum generated by an arbitrary class, $\mathcal{B}_0 \in \mathcal{P}(\mathcal{L})$, defines, with the aid of rel.3.3, a mapping that is opening. Now, associate with each $B \in \mathcal{B}$ the dilation

$$\delta_B(A) = B \qquad \text{if } A \not\leq B; \qquad \delta_B(A) = \varnothing \qquad \text{if } A \leq B.$$

Its corresponding adjunctionl opening is

$$\gamma_B(X) = \vee\{\delta_B(A) : \delta_B(A) \leq X\} = \begin{cases} B & \text{if } B \leq X \\ \varnothing & \text{if } B \not\leq X, \end{cases}$$

so that rel. 3.3 becomes

$$\gamma = \vee\{\gamma_B, B \in \mathcal{B}\}.$$

In other words, we have the following theorem ([62], p.22).

Theorem 3.4 *If a mapping $\gamma : \mathcal{L} \to \mathcal{L}$ is an opening then it has a representation of the form*

$$\gamma = \vee\{\gamma_B, B \in \mathcal{B}\},$$

where γ_B is the adjunction opening associated with the dilation $\delta_B = \varnothing$ if $A \leq B$ and $\delta_B(A) = B$ otherwise, and where $\mathcal{B}$ is the domain of invariance *of γ. Conversely, if $\mathcal{B}$ denotes the class closed under union generated by $\mathcal{B}_0$, an arbitrary class of elements of $\mathcal{L}$, then the mapping defined by γ is an opening.*

In the Euclidean case, if X_B designates the adjunction opening $(X \ominus B) \oplus B$,we obtain the following more precise result ([38], p.190).

Proposition 3.5 *A mapping $\gamma : \mathcal{P}(\mathbb{R}^n) \to \mathcal{P}(\mathbb{R}^n)$ is an opening invariant under translation if and only if it admits the representation $\gamma(X) = \cup \{X_B, B \in \mathcal{B}_0\}$ for a an arbitrary class $\mathcal{B}_0 \subseteq \mathcal{P}(\mathbb{R}^n)$. The domain of invariance of γ is the class closed under translation and union generated by $\mathcal{B}_0$.*

3.3.1 Closings and duality

The criteria of extensivity and anti-extensivity are duals of one another in the sense of Chapter 1. Consequently, to each property of openings there corresponds a symmetrical property of closings, which is obtained by changing the direction of the inequalities, or by replacing union by intersection, etc. Thus, Propositions 13.1, 13.2, etc. may be transposed, and the first, for example, becomes as follows.

Proposition 3.6 *The closing ϕ, with domain of invariance $\mathcal{B}$, is the largest extension to $\mathcal{L}$ of the identity on $\mathcal{B}$, i.e.*

$$\phi(X) = \wedge \{B : B \in \mathcal{B}, B \geq X\}.$$

3.4 (Non exhaustive) catalog of openings and closings

Although theorem 3.4 is heuristically deep, we may have difficulties in applying it directly, as the number of terms γ_i necessary for generating a given γ becomes prohibitive. Actually, there are various starting points for creating openings, namely:

- *the adjunction openings,*

- *the trivial openings,*
- *the envelope openings.*
- *the sandwich openings*
- *the annular openings*

lus any derivation obtained by cross-union of these types, plus all the openings associated with connections (chapter 6). The first mode has already been developed. We will now present the other four.

3.4.1 Trivial openings

A criterion T is said to be *increasing* when, for all $X \in \mathcal{L}$:

$$\begin{cases} X \text{ satisfies } T \text{ and } Y \geq X & \Longrightarrow \quad Y \text{ satisfies } T, \\ X \text{ does not satisfy } T \text{ and } Y \leq X & \Longrightarrow \quad Y \text{ does not satisfy } T. \end{cases}$$

For example, in $\mathbb{R}^n$, for X to hit a given set A_0, as well as to have a Lebesgue measure larger than a given value λ_0 are both increasing criteria.

Proposition 3.7 *Given an increasing criterion T over lattice $\mathcal{L}$, the operation*

$$\gamma_1(X) = \begin{cases} X & \textit{when } X \textit{ satisfies } T, \\ \emptyset & \textit{otherwise.} \end{cases}$$

(with $\gamma_1(\emptyset) = \emptyset$) is an opening called the trivial opening associated with criterion T.

3.4.2 Envelope openings

Consider a **finite** lattice $\mathcal{L}$ and an increasing mapping $\psi : \mathcal{L} \longrightarrow \mathcal{L}$. Then, for any $X \in \mathcal{L}$, the sequence $[X \cap \psi(X)]^n$ decreases with n, and finally stops for a certain n_0, since $\mathcal{L}$ is finite. The operator

$$\check{\psi} = (I \wedge \psi)^{n_0} \tag{3.4}$$

is therefore an opening. Moreover, if h is an opening smaller than ψ, then $h \leq I \wedge \psi$. Hence $h = h^n \leq (I \wedge \psi)^n$ for every n and thus $h \leq \check{\psi}$. In other words:

Theorem 3.8 *Let $\mathcal{L}$ be a finite lattice. Then, for every increasing mapping $\psi : \mathcal{L} \longrightarrow \mathcal{L}$, there exists an upper envelope $\check{\psi}$ of the openings which minorate ψ. It is itself an opening and is given by the relation*

$$\check{\psi} = (I \wedge \psi)^{n_0} \qquad \textit{for a finite } n_0.$$

(dual result with $\hat{\psi} = (I \vee \psi)^{n_0}$.)

N.B: *(i)* The iterations may well stop at the first step. In § **??**, the example of the rank-operators illustrates this point.

(ii) Under conditions which are always fulfilled in practice, theorem 3.8 extends to the lattice of the functions $f : \mathbb{Z}^n \longrightarrow \mathbb{Z}$ [64].

3.4.3 Sandwich openings

Given arbitrary opening γ and dilation δ, of adjoint erosion ε, the composition product

$$\psi = \delta\gamma\varepsilon$$

is still an opening. Idempotence results from that we have $\delta\gamma\,(\varepsilon\delta)\,\gamma\varepsilon \geq \delta\gamma\varepsilon$, but also the inverse inequality, by extensivity of γ and of $\delta\varepsilon$. An example is provided by taking for γ the removal of all particles smaller than a certain size during te processing of a adjunction opening.

3.4.4 Annular opening

Consider the pair of points $B = \{o, b\}$, made of the origin o and a point b in direction α in $\mathbb{R}^2$ or in $\mathbb{Z}^2$. Clearly, the morphological opening γ_b with respect to B is equivalent to

$$\gamma_b = I \wedge \delta_{B'}$$

where $\delta_{B'}$ is the t-dilation by the bi-point $B' = \{-b; +b\}$. Now, make vary b in a certain domain D which does not contain the origin (e.g. three consecutive vertices of an hexagon centered on o, half a circle, ...) and take the sup γ:

$$\gamma = \vee\{\gamma_b, b \in B\} = I \wedge \{\vee\delta_{B'}, b \in D\}$$

i.e. since the dilation commutes with $\vee$:

$$\gamma = I \wedge \delta_{D\cup\check{D}} \tag{3.5}$$

where $\delta_{D\cup\check{D}}$ is the dilation by $D \cup \check{D} = \bigcup_{b\in D}\{-b;+b\}$. The effect of this *annular opening* γ is shown on Fig. 3.4. γ eliminates the components of a given set X as a function of their environment more than of their size or shape. On the example presented here, $D \cup \check{D}$ is taken to be a circle and γ eliminates the central particle without touching the others.

To illustrate the specific action of γ, we can compare it with the morphological opening γ' by a disc and with the union γ'' of the morphological openings by segments in various directions (see Fig. 3.4).

3.5 Granulometries

3.5.1 Matheron's axiomatics

The *Granulometries,* which are size distributions based on an appropriate axiomatics, deal with families of openings or closings that are parametrized by a positive number (the size) [37]. More precisely, we have the following

Definition 3.9 *A family $\{\gamma_\lambda\}$ of mappings $\mathcal{L} \to \mathcal{L}$, depending on a positive parameter λ is a granulometry when*

$$\begin{array}{c} \textit{(i) } \gamma_\lambda \textit{ is an opening } \forall\lambda \geq 0, \\ \textit{(ii) } \lambda,\mu \geq 0 \Rightarrow \gamma_\lambda\gamma_\mu = \gamma_\mu\gamma_\lambda = \gamma_{\sup(\lambda,\mu)}. \end{array} \tag{3.6}$$

These conditions are called Matheron's axioms for granulometry ([38], p.192). It is easy to verify that these conditions are satisfied by every process that common sense would qualify as a size distribution (see for example [60] ch.X). System (3.6) is equivalent to

$$\begin{array}{c} \textit{(i) } \gamma_\lambda \textit{is an opening } \forall\gamma, \\ \textit{(iii) } \lambda \geq \mu \geq 0 \Rightarrow \gamma_\lambda \geq \gamma_\mu. \end{array} \tag{3.7}$$

This second version of Matheron's axioms emphasizes the two monotonicities with respect to X (opening axiom), and with respect to λ. The first presentation, by system (3.6), shows how openings are composed and highlights their *semi-group* structure. A third representation involves the invariant sets as follows. Let $\mathcal{B}_\lambda = \gamma_\lambda(\mathcal{L})$ denote the domain of invariance of γ_λ, i.e. the family of B's such that $B = \gamma_\lambda(B)$ then Matheron's system is equivalent to

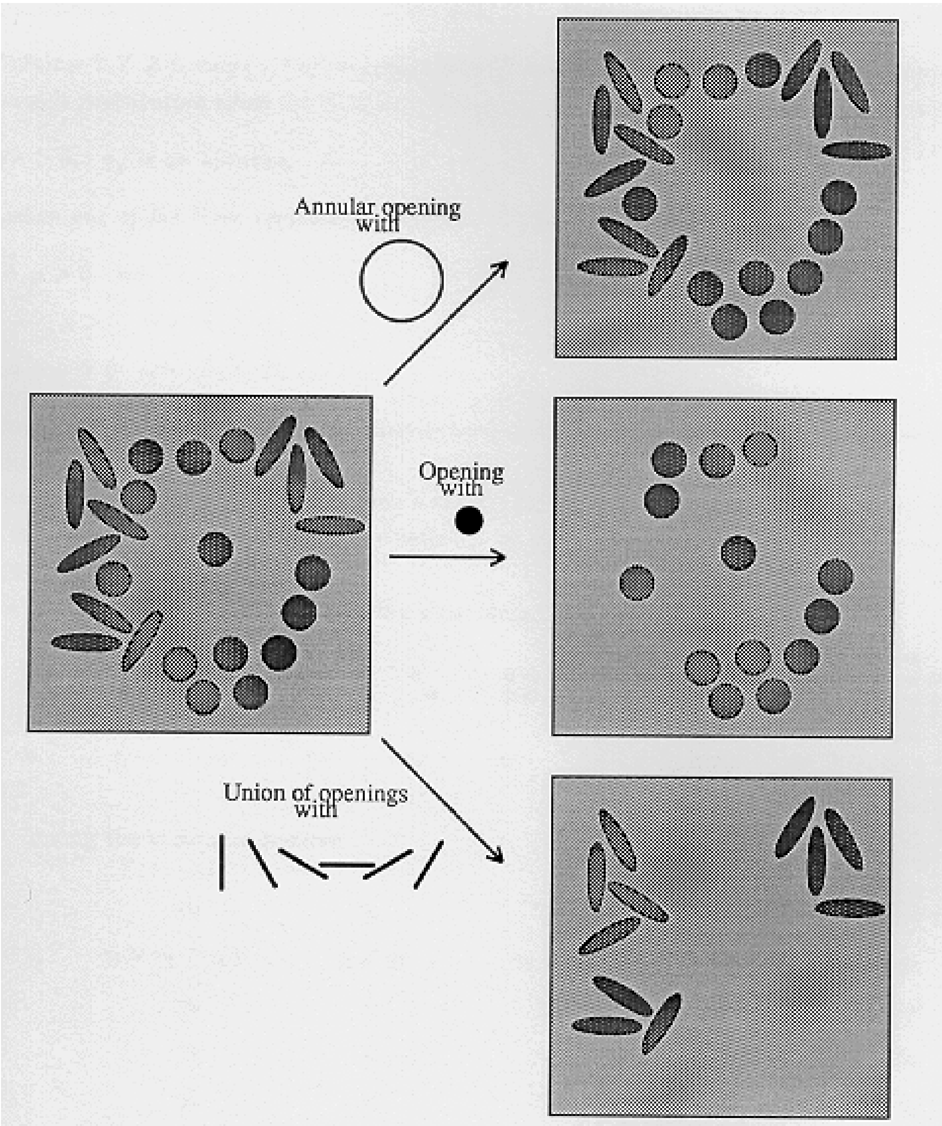


Figure 3.4: Annular opening γ versus a classical opening by a disc and a union of openings by segments.

$$\begin{cases} \text{(i) } \gamma_\lambda \text{is an opening } \forall \lambda \geq 0, \\ \text{(iv) } \lambda \geq \mu \geq 0 \Rightarrow \mathcal{B}_\lambda \subseteq \mathcal{B}_\mu. \end{cases}$$

By duality, we introduce *anti-granulometry* as the families of closings ϕ_λ, $\lambda \geq 0$, such that one of the following three equivalent properties is satisfied:

$$\begin{aligned} &\text{(ii}'\text{) } \lambda, \mu \geq 0 \Rightarrow \phi_\lambda \phi_\mu = \phi_\mu \phi_\lambda = \phi_{\sup(\lambda,\mu)}; \\ &\text{(iii}'\text{) } \lambda \geq \mu \geq 0 \Rightarrow \phi_\lambda \geq \phi_\mu; \\ &\text{(iv}'\text{) } \lambda \geq \mu \geq 0 \Rightarrow \mathcal{B}_\lambda \subseteq \mathcal{B}_\mu. \end{aligned}$$

3.5.2 Euclidean granulometries

When dealing with Euclidean spaces, we are particularly interested in granulometries that are translationally invariant and compatible with homothetics. These are *Euclidean* granulometries. A family of mappings γ_λ from a complete sublattice of $\mathcal{P}(\mathbb{R}^n)$ into itself is a Euclidean granulometry if and only if there exists a class $\mathcal{B}_0 \subseteq \mathcal{P}(\mathbb{R}^n)$ such that

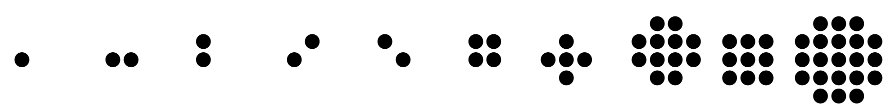
$$\gamma_\lambda(X) = \cup\{X_{\mu B},\ B \in \mathcal{B}_0,\ \mu \geq \lambda\} \qquad X \subseteq \mathbb{R}^n, \tag{3.8}$$

where $X_{\mu B}$ denotes the morphological opening of X by μB. Then the domains of invariance $\mathcal{B}_\lambda$ are equal to $\lambda\mathcal{B}$, where $\mathcal{B}$ is the class closed under union, translation and homothetics ≥ 1, which is generated by $\mathcal{B}_0$. The relation (3.8), which involves a double union, is complicated; even if we reduce the class $\mathcal{B}$ to a single element B, the associated size distribution becomes

$$\gamma_\lambda(X) = \cup\{X_{\mu B},\ \mu \geq \lambda\} \qquad X \subseteq \mathbb{R}^n, \tag{3.9}$$

which brings into play an infinite union.

As examples, consider the following structuring elements:



e_0 e_1 e_2 e_3 e_4 e_5 e_6 e_7 e_8 e_9

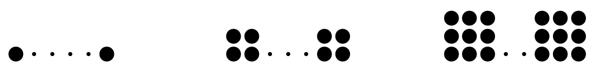
Then, among the various sequences

$$\begin{array}{cccccc}
\gamma_0 & \gamma_0 & \gamma_0 & \gamma_0 & \gamma_0 & \gamma_0 \\
\gamma_1 \vee \gamma_2 & \gamma_1 \vee \gamma_2 \vee \gamma_3 \vee \gamma_4 & \gamma_1 \vee \gamma_2 & \gamma_3 \vee \gamma_4 & \gamma_1 \vee \gamma_2 & \gamma_5 \\
\gamma_5 & \gamma_6 & \gamma_5 & \gamma_5 & \gamma_5 & \gamma_7 \\
\gamma_7 & \gamma_7 & \gamma_7 & \gamma_8 & \gamma_6 & \gamma_8 \\
\gamma_9 & \gamma_9 & \gamma_9 & \gamma_9 & \gamma_8 & \gamma_9
\end{array}$$

the four former ones lead to size distributions, but not the last two ones.

3.5.3 Granulometry and convexity

Remark that the convexity of the structuring elements is not a necessity. For instance, the sequence



e_1 e_2 e_3 etc ...

induces a size distribution where the union involved in rel.(3.9) is reduced to one element for each λ. However, in the Euclidean space $\mathbb{R}^n$, a family $(B_\lambda)_{\lambda \geq 0}$ of structuring elements generates a size distribution $(\gamma_\lambda)_{\lambda \geq 0}$ which is compatible with magnification, i.e.

$$\forall \lambda \geq 0, \forall X \subset \mathbb{R}^n, \quad \gamma_\lambda(X) = \lambda \gamma_1(X/\lambda), \tag{3.10}$$

if and only of the B_λ's are the homotethics of a compact **convex** set B. Moreover, if set X is closed ($X \in \mathcal{F}$) or compact ($X \in \mathcal{K}$) then the mapping $(\lambda, X) \rightarrow \gamma_\lambda(X)$ from $\mathbb{R}^+ \times \mathcal{F}$ (resp. $\mathbb{R}^+ \times \mathcal{K}$) into $\mathcal{F}$ (resp. $\mathcal{K}$) is upper-semicontinous. The signification of rel.(3.10) is clear: it just means that γ_λ acts on λX just as γ_1 does on X. Such a property, which is always satisfied for convolution products, may not exist for morphological filters. However, in the two important cases of the size distributions and of the alternating sequential filters, we easily obtain it. From an epistomological point of view, it is instructive to remark that mathematical morphology exchanges the property of monotonicity of $\lambda \rightarrow \gamma_\lambda(X)$ for a given set X (which depends on X) with the convexity of the structuring element, which is independent of the set under study.

Chapter 4

Morphological filters

4.1 12-1 - Introduction

When dealing with a signal in one- or multidimensional space, the filter is commonly defined as ant operator that is *linear*, *continuous* and *invariant under translation*. According to a classical result, any filter, in the above sense, can be expressed as the convolution product $f * \phi$ of the signal f by a convolution distribution ϕ. In addition to these three intrinsic properties, it is common practice to consider these filters as band-pass devices, even if this is not exactly true. We say that a hi-fi amplifier "covers up to 30 000 Hz", or that coated glass is "monochromatic", etc. Implicitly, this confers on the filtering operation the property of not acting by iteration: a signal that has lost the part of the spectrum above 30 000 Hz will not be modified if processed by a second identical filter. Algebraically, this is known as *idempotence*.

Let us examine more closely the physical significance of linearity, which states that the transformation $f + f'$ is the sum of the transforms f and f'. When we listen to a piano and violin duo on the radio, for example, it is clear that the radio amplifies the individual sounds produced by the piano and the violin in proportion to their intensities. Likewise, the human ear *sums* the intensities, or their logarithms, that are produced by different sources. In image analysis, when we should like to correct for lack of focusing, or camera movement in a photograph or in a satellite transmission, we find that the associated physical phenomena are also linear (camera movement can be considered as the sum of several photos). This situation justifies the use of convolutions for the improvement of these types of image defects.

Although acoustic signals are summed, visual signals are not compounded in this manner. The world around us is not translucent; on the contrary, it is composed of opaque objects that hide one another. The notion of *inclusion* is used to express this fundamental law. For example, "if the contour of the nearest object contains the contour of any object further away then the latter will be completely hidden from sight; if, on the other hand, the nearer contour is contained by the more distant one then it cannot hide any more distant objects that are not already hidden by further object", and so on .

This inclusion relation in the visual universe is as basic as additivity for acoustic perception. Thus the first prerequisite for any morphological filter ψ should be that it preserves *ordering relations,* i.e. that ψ be increasing.

Can we conceive of a morphological filter which is also linear? Yes, but only for those convolutions whose kernel is non negative (such a restriction rejects almost all deconvolutions met in applied signal, or image, processing). Indeed, the sup $f \vee g$ is different from the sum $f + g$ in that it has no inverse, and $(f \vee g) \wedge g$ cannot regenerate f as could $(f + g) - g$. Not only are both properties exclusive, in practice at least, but there is a conflict in the very philosophy of the image analysis involved. Linearity often gives a group structure to convolutions, and thus permits us to deconvolve, i.e. to produce a clear image from a blurry one, or to find the point of a function after one-dimensional integration, as in tomography. This implies that there was no information loss in the convolution that produced the blur or in the integration. On the contrary, an increasing transformation generally produces a *loss of information.* It is for this reason that one cannot find an equivalent for Fourier space that would replace increasing mappings by multiplications, or by any other reversible operation. More generally, we are not seeking to replace reversible operations by others, more or less adapted to a particular problem, but rather to accept this information loss as inevitable, and to try to control it.

To attain this, we shall add a second (and final) axiom to the definition of a morphological filter, namely *idempotence(*[62], Ch.5and 6) . This last condition stops the simplifying action of increasingness at the first stage, and thus makes formal the intuition we observed concerning band-pass and monochromatic filters[1].

4.2 The lattice of the increasing mappings

This chapter constitutes an overview of the theory of morphological filtering, due to G. Matheron [39, chapter 6]. The lattice examples introduced in chapter 1 concerned the scenes under study. We will now consider classes of *operations* working on these objects. Let ψ be such an operator, i.e. be a mapping from a complete lattice $\mathcal{L}$ into itself. We assume that ψ is increasing, i.e. that it preserves the ordering relation of $\mathcal{L}$:

$$\forall A, A' \in \mathcal{L}, \qquad A \geq A' \Longrightarrow \psi(A) \geq \psi(A'). \tag{4.1}$$

The set $\mathcal{L}'$ of the increasing mappings on the complete lattice $\mathcal{L}$ satisfies the following properties:

1. $\mathcal{L}'$ is a semi-group for the composition product $\circ$, with a unit element, namely the identity mapping I $(\forall A \in \mathcal{L}, I(A) = A)$.
2. $\mathcal{L}'$ is a complete lattice for the ordering relation:

   $$f \geq g \iff \forall A \in \mathcal{L},\ f(A) \geq g(A),$$

   since the following identities

   $$(\vee_{\mathcal{L}'} f_i)(A) = (\vee_{\mathcal{L}} f_i) \qquad \text{and} \qquad (\wedge_{\mathcal{L}\ '} f_i)(A) = (\wedge_{\mathcal{L}}\ f_i)$$

   generate a supremum and an infimum in the set $\mathcal{L}'$.

The two basic structures of the semi-group and of the lattice interact with each other, and we have, for all f, g, h and (f_i) in $\mathcal{L}'$:

$$\begin{array}{lcl} (\vee f_i) \circ g = \vee(f_i \circ g) & ; & g \circ (\vee f_i) \geq \vee(g \circ f_i) \\ (\wedge f_i) \circ g = \wedge(f_i \circ g) & ; & g \circ (\wedge f_i) \leq \wedge(g \circ f_i) \end{array} \tag{4.2}$$

and

$$f \geq g \Longrightarrow \begin{cases} f \circ h & \geq & g \circ h \\ h \circ f & \geq & h \circ g \end{cases}$$

In the following, the two classes of the *overfilters* (i.e. the mappings $f \in \mathcal{L}'$ such that $f \circ f \geq f$) and of the *underfilters* play a major role. Indeed:

[1] In literature, the term “filter” may also be associated with growth only [34, 35], and can even be a synonymous with mapping [80]

Theorem 4.1 *the class of the underfilters (resp. overfilters) is closed under $\wedge$ (resp. $\vee$) and under self-composition.*

proof: For example, let $(f_j)_{j \in J}$ be a family of underfilters. From (4.1) and (4.2), we get:

$$(\vee_{j\in J} f_j) \circ (\vee_{j\in J} f_j) = \vee_{i\in J}(f_i \circ \vee_{j\in J} f_j) \leq \vee_{i\in J}(f_i \circ f_i) \leq \vee_{i\in J} f_i,$$

so that $\vee_{j\in J} f_j$ is an underfilter. Moreover, given an underfilter f, $ff \leq f$ implies, by growth, that $ff \circ ff \leq ff$, so that the self-composition ff is an underfilter. $\Box$

4.3 Morphological filters

We now define the notion of a **morphological filter** as follows:

Definition 4.2 *The elements of $\mathcal{L}'$ which are both underfilters and overfilters are called (morphological)* **filters.**

In other words, the morphological filters are the transformations acting on the scenes under study (i.e. the lattice $\mathcal{L}$) and which are **increasing** and **idempotent.** We shall denote by $\mathcal{V}$ the class of the filters, with $\mathcal{V} \subseteq \mathcal{L}'$. Remark that the class $\mathcal{V}$ is not closed either under $\vee$, or $\wedge$, or under composition (a counter-example, based on openings, has been exhibited in § 3.2).

This apparent drawback suggests us to investigate more accurately the possible connections of class $\mathcal{V}$ with the composition product and with extrema. Can we find, for example, pairs (f, g) of filters such that $f \circ g$, $g \circ f$, $f \circ g \circ f$, etcre surely filters (composition problem)? Can we keep the usual ordering relation in $\mathcal{V}$ and equip $\mathcal{V}$ with **new** sup and inf, such that it turns out to become a complete lattice (extrema problem)? These two sorts of questions will build the subject of the next two sections.

However, we can already notice that filters allow to generate the under and overfilters in the following way.

Theorem 4.3 *Any underfilter $f \in \mathcal{L}'$ is the infimum of the filters $\psi \geq f$. Any overfilter $g \in \mathcal{L}'$ is the supremum of the filters $\psi \leq g$.*

proof: Let, for instance, g be an overfilter. For any $A \in \mathcal{L}$, we consider the mapping $\psi_A \in \mathcal{L}'$ defined as follows:

$$\psi_A(A') = \begin{cases} g(A) & \text{if} \quad A' \geq A \quad \text{or} \quad A' \geq g(A), \\ \varnothing & \text{otherwise.} \end{cases}$$

Clearly, ψ_A is a filter. If $A' \geq A$, we have $\psi_A(A') = g(A) \leq g(A')$. If $A' \geq g(A)$ then we find $\psi_A(A') = g(A) \leq g(g(A))$, because g is an overfilter. But $g(A) \leq A'$ implies $g(g(A)) \leq g(A')$, and then $\psi_A(A') \leq g(A')$. We conclude that $\psi_A \leq g$ for any $A \in \mathcal{L}$. But $\psi_A(A) = g(A)$, for each $A \in \mathcal{L}$, and thus g is the supremum of the filters $\psi_A, A \in \mathcal{L}$. □

4.4 Composition of morphological filters

With any increasing mapping $\psi : \mathcal{L} \longrightarrow \mathcal{L}$, associate:

1. the **image domain** $\psi(\mathcal{L})$, i.e. the set of the transforms by ψ:

$$\psi(\mathcal{L}) = \{\psi(A), A \in \mathcal{L}\},$$

2. the **invariance domain** $\mathcal{B}_\psi$, i.e. the class of those $B \in \mathcal{L}$ which are left unchanged under

$$\psi : \mathcal{B}_\psi = \{B \in \mathcal{L}, \psi(B) = B\}.$$

When ψ is a filter, $\mathcal{B}_\psi$ is sometimes called the *root* of ψ in literature.We always have $\mathcal{B}_\psi \subseteq \psi(\mathcal{L})$, an inclusion which becomes an equality

$$\mathcal{B}_\psi = \psi(\mathcal{L})$$

if and only if ψ is **idempotent.** This preliminary remark leads to the following three criteria:

Criterion 4.4 *For any mappings f, g from $\mathcal{L}$ into itself,*

$$fg = g \Longleftrightarrow g(\mathcal{L}\ ) \subseteq \mathcal{B}_f.$$

In particular, when g is idempotent:

$$fg = g \Longleftrightarrow \mathcal{B}_g \subseteq \mathcal{B}_f. \tag{4.3}$$

Criterion 4.5 *Two mappings f and g from $\mathcal{L}$ into itself are idempotent and admit the same invariance domain $\mathcal{B}_f = \mathcal{B}_g$ if and only if:*

$$fg = g \qquad \text{and} \qquad gf = f. \tag{4.4}$$

proof: Criterion 4.4 is obvious. Now, if rel. (4.4) is satisfied, then $ff = f \circ gf = gf = f$, i.e. f, and similarly g, are idempotent. Hence, from (4.3), $B_f \subseteq B_g$ and $B_g \subseteq B_f$. Conversely, when f and g, idempotent, have the same invariance domain, rel.(4.3) implies rel.(4.4).Criterion4.5 thus follows.

Criterion 4.6 *Let ψ be idempotent on $\mathcal{L}$. Then, for any mapping f from $\mathcal{L}$ into itself such that*

$$f\psi = \psi$$

ψf is idempotent and $\mathcal{B}_{\psi f} = \mathcal{B}_{\psi}$.

proof: If $f\psi = \psi$ then we have $\psi f.\psi = \psi\ .\psi = \psi$ and $\psi.\psi f = \psi f$ Thus ψf is idempotent, and, from Criterion4.5 $\mathcal{B}_{\psi f} = \mathcal{B}_{\psi}$. □

□

In these three criteria, the ordering $\leq$ does not intervene. From now on, we shall only consider the **increasing mappings** ψ, i.e. $\psi \in \mathcal{L}'$. For any filter ψ, the class of the filters ψ' that have the same invariance domain $\mathcal{B}_\psi$ as ψ will be denoted $\mathcal{I}d(\psi)$. The following theorem is the key result concerning the composition of filters:

Theorem 4.7 *Let f and g be two filters on $\mathcal{L}$ such that $f \geq g$. Then:*

(i) $f \geq fgf \geq gf \vee fg \geq gf \wedge fg \geq gfg \geq g$,

(ii) gf, fg, fgf and gfg are filters, and $fgf \in \mathcal{I}d(fg)$, $gfg \in \mathcal{I}d(gf)$,

(iii) fgf is the smallest filter greater than $gf \vee fg$ and gfg is the greatest filter smaller than $gf \wedge fg$,

(iv) the following equivalences hold:

$$\begin{aligned} \mathcal{B}_{fg} = \mathcal{B}_{gf} \iff\ & \mathcal{B}_{fg} = \mathcal{B}_f \cap \mathcal{B}_g \iff \mathcal{B}_{gf} = \mathcal{B}_f \cap \mathcal{B}_g \\ \iff\ & fgf = gf \qquad\quad\ \iff gfg = fg \\ \iff\ & gf \geq fg. \end{aligned}$$

proof: The inequalities *(i)* are obvious. From the relationships

$$fg = fffg \geq fgfg \geq fggg = fg,$$

we conclude that fg is a filter. By the dual inequalities, gf is also a filter. Now, we have

$$\begin{aligned} fgf \circ fg &= fg(ff)g = fgfg = fg, \\ fg \circ fgf &= fgfg \circ f = (fg \circ fg)f = fgf, \end{aligned}$$

and thus, $fgf \in \mathcal{I}d(fg)$, by criterion 4.5. In the same way, we find that $fgf \in \mathcal{I}d(gf)$, so that *(ii)* is proved.

Now, fgf is a filter (by *(ii))* and $fgf \geq gf \vee fg$ (by *(i))*. Let ψ be a filter such that $\psi \geq fg$ and $\psi \geq gf$. It follows that $\psi = \psi\psi \geq fggf = fgf$. Thus, fgf is the smallest filtering upper bound of fg and gf. Hence *(iii)* is proved.

By criterion 4.5, we have $\mathcal{B}_{fg} = \mathcal{B}_{gf}$ if and only if

$$fg \circ gf = fgf = gf \quad \text{and} \quad gf \circ fg = gfg = fg.$$

These relations actually imply one another. For instance, $fgf = gf$ implies $fgf \circ g = gfgf$, i.e. $fg = gfg$. By *(iii)*, these relations are equivalent to $gf \geq fg$.

The inclusions

$$\mathcal{B}_f \cap \mathcal{B}_g \subseteq \mathcal{B}_{fg} \subseteq \mathcal{B}_f$$

always hold, so that $\mathcal{B}_{fg} = \mathcal{B}_f \cap \mathcal{B}_g$ if and only if $\mathcal{B}_{fg} \subseteq \mathcal{B}_g$, i.e, by criterion 4.4, if and only if $gfg = fg$. This completes the proof. □

Examples:

1. Start from an arbitrary opening γ and an arbitrary closing φ. Since

   $$\gamma \leq I \leq \varphi,$$

   by theorem 4.7, $\gamma\varphi$, $\varphi\gamma$, $\gamma\varphi\gamma$ and $\varphi\gamma\varphi$ are filters. The composition products of φ by γ, then by φ, etcgenerates the oscillating sequence

   $$\varphi \longrightarrow \gamma\varphi \longrightarrow \varphi\gamma\varphi \longrightarrow \gamma\varphi \longrightarrow \dots$$

   Remark that when $\gamma\varphi \geq \varphi\gamma$ (which is generally not the case), we have $\gamma\varphi = \varphi\gamma\varphi$ and the oscillations are stopped after the first step.

2. There is a more particular example, which illustrates point *(iv)* of the theorem. In $\mathcal{P}(\mathbb{R}^n)$ (or $\mathcal{P}(\mathbb{Z}^n)$, or $\mathcal{F}(\mathbb{R}^n, \mathbb{R})$, or $\mathcal{F}(\mathbb{Z}^n, \mathbb{Z})$), consider the morphological opening γ_l by a segment of length l in the horizontal direction. For a given X, $\gamma_l(X)$ is made of horizontal segments of length $\geq l$. Moreover, closing this set by the dual closing φ_l, i.e. determining $\varphi_l\gamma_l(X)$ may only suppress intervals between two such segments, hence increase the length of the horizontal intercepts; Therefore, $\gamma_l\varphi_l\gamma_l(X) = \varphi_l\gamma_l(X)$, and by the theorem, $\varphi_l\gamma_l \leq \gamma_l\varphi_l$.

4.5 Structure of the invariance domain B_ψ.

In general, if $\mathcal{B}$ is an arbitrary subset of $\mathcal{L}$ then there exist no filter ψ having $\mathcal{B}$ as its invariance domain, and $\mathcal{J}\mathrm{d}(\mathcal{B}) = \varnothing$. Under what condition is $\mathcal{J}\mathrm{d}(\mathcal{B})$ not empty? A sufficient condition is that $\mathcal{B}$ be closed under $\wedge$ or under $\vee$, because in this case the closing $\widetilde{\varphi}_\mathcal{B}$ or the opening $\widetilde{\gamma}_\mathcal{B}$ generated from the invariant elements $\mathcal{B}$ belongs to $\mathcal{J}\mathrm{d}(\mathcal{B})$. We shall see that the necessary and sufficient condition is in a certain sense a generalization of these two particular cases.

Let ψ be a filter, and $\mathcal{B} = \mathcal{B}_\psi$ its invariance domain. Denote by $\widetilde{\gamma} = \widetilde{\gamma}_\mathcal{B}$ (resp.$\widetilde{\varphi} = \widetilde{\varphi}_\mathcal{B}$) the smallest (resp. the greatest) increasing extension on $\mathcal{L}$ of the identity mapping on $\mathcal{B}$. Explicitely :

$$\widetilde{\gamma}(A) = \vee\{B : B \in \mathcal{B}, B \leq A\}$$
$$\widetilde{\varphi}(A) = \wedge\{B : B \in \mathcal{B}, B \geq A\}$$

The opening $\widetilde{\gamma}$ and the closing $\widetilde{\varphi}$ satisfy the inequalities

$$\widetilde{\gamma} \leq \psi \leq \widetilde{\varphi} \quad . \tag{4.5}$$

Moreover, by Criterion 4.4, the inclusions $\mathcal{B} \subseteq \mathcal{B}_{\widetilde{\gamma}}$ and $\mathcal{B} \subseteq \mathcal{B}_{\widetilde{\varphi}}$ imply that

$$\widetilde{\gamma}\psi = \widetilde{\varphi}\psi = \psi \tag{4.6}$$

so that, by Criterion 4.6, we also have

$$\psi\widetilde{\gamma} \in \mathcal{J}\mathrm{d}(\mathcal{B}) \qquad , \qquad \psi\widetilde{\varphi} \in \mathcal{J}\mathrm{d}(\mathcal{B}).$$

More precisely $\psi_M = \psi\widetilde{\varphi}$ is the *greatest element* of $\mathcal{J}\mathrm{d}(\mathcal{B})$, and $\psi_m = \psi\widetilde{\gamma}$ is its *smallest element*. In fact, by rel.(4.6), we have for instance

$$\psi_M = \psi\widetilde{\varphi} \quad = \widetilde{\gamma}\ \psi\widetilde{\varphi} \quad ,$$

and the inequalities 4.5 imply

$$\widetilde{\gamma}\ \widetilde{\varphi} = \widetilde{\gamma}\ \widetilde{\gamma}\widetilde{\varphi}\ \subseteq \widetilde{\gamma}\psi\widetilde{\varphi} = \psi_M \subseteq \widetilde{\gamma}\widetilde{\varphi}\widetilde{\varphi} = \widetilde{\gamma}\widetilde{\varphi} \quad .$$

Hence $\psi_M = \widetilde{\gamma}\ \widetilde{\varphi}$, and in the same way $\psi_m = \widetilde{\varphi}\ \widetilde{\gamma}$. But the filter $\widetilde{\gamma}\widetilde{\varphi}$ depends only on $\mathcal{B}$, and not on the choice of the particular element $\psi \in \mathcal{J}\mathrm{d}(\mathcal{B})$. Hence, we have $\psi_M = \psi\widetilde{\varphi} \supseteq \psi$ for any $\psi \in \mathcal{J}\mathrm{d}(\mathcal{B})$, and ψ_M is the greatest element of $\mathcal{J}\mathrm{d}(\mathcal{B})$. In the same way, $\psi_m = \psi\widetilde{\gamma} = \widetilde{\varphi}\ \widetilde{\gamma}$ is the smallest element of $\mathcal{J}\mathrm{d}(\mathcal{B})$. Now, by applying theorem 4.7(*iv*), with $f = \widetilde{\varphi}$ and $g = \widetilde{\gamma}$, we find

$$\mathcal{B} = \mathcal{B}_{\widetilde{\gamma}} \cap \mathcal{B}_{\widetilde{\varphi}} \qquad ; \qquad \widetilde{\gamma}\ \widetilde{\varphi} \quad = \widetilde{\varphi}\ \widetilde{\gamma}\ \widetilde{\varphi} \qquad ; \qquad \widetilde{\varphi}\ \widetilde{\gamma}\ = \widetilde{\gamma}\ \widetilde{\varphi}\ \widetilde{\gamma} \quad .$$

By the same theorem 4.7, these necessary conditions are also sufficient. More precisely, we may summarize our results in the following **structural theorem**

Theorem 4.8 *Let $\mathcal{B}$ be a subset of $\mathcal{L}$. Then $\mathcal{J}d(\mathcal{B})$ is not empty if and only if the condition*

$$\mathcal{B} = \mathcal{B}_{\widetilde{\gamma}} \cap \mathcal{B}$$

and one of the following three equivalent conditions are satisfied:

Theorem 4.9 $i)\widetilde{\gamma}\widetilde{\varphi} \geq \widetilde{\varphi}\ \widetilde{\gamma} \quad ;$

$ii)\widetilde{\gamma}\widetilde{\varphi} = \widetilde{\varphi}\ \widetilde{\gamma}\ \widetilde{\varphi} \quad ;$

$iii)\ \widetilde{\varphi}\ \widetilde{\gamma} = \widetilde{\gamma}\ \widetilde{\varphi}\ \widetilde{\gamma} \quad .$

If so, then $\mathcal{J}\mathrm{d}(\mathcal{B})$ has a greatest element ψ_M which is a $\vee-$ filter, and a smallest element ψ_m which is a $\wedge-$ filter. Moreover, we have for any other filter $\psi \in \mathcal{J}\mathrm{d}(\mathcal{B})$:

Theorem 4.10

$$\begin{cases} \psi_m = \widetilde{\varphi}\,\widetilde{\gamma} = \widetilde{\gamma}\,\widetilde{\varphi}\,\widetilde{\gamma} = \psi\widetilde{\gamma} & , \\ \psi_M = \widetilde{\gamma}\widetilde{\varphi} = \widetilde{\varphi}\,\widetilde{\gamma}\,\widetilde{\varphi} = \psi\widetilde{\varphi} & , \\ \psi = \widetilde{\gamma}\,\psi = \widetilde{\varphi}\,\psi, & \\ \widetilde{\gamma} \le \psi_m \le \psi \le \psi_M \le \widetilde{\varphi} & . \end{cases}$$

The same theorem may be restated in another way. If $\mathcal{J}\mathrm{d}(\mathcal{B}) \neq \varnothing$ then let B_i be a family of elements of $\mathcal{B}$. We have $\vee\, B_i \in {\sim}B$, and thus $\widetilde{\gamma}\,(\vee\, B_i) = \vee B_i$. From the first relation above, it follows for any $\psi \in \mathcal{J}\mathrm{d}(\mathcal{B})$, that

$$\psi\,(\vee\, B_i) = \psi\widetilde{\gamma}\,(\vee\, B_i) = \widetilde{\varphi}\,\widetilde{\gamma}\,(\vee\, B_i)\,.$$

But $\widetilde{\gamma}\,(\vee\, B_i) = \vee\, B_i$, so that

$$\widetilde{\varphi}\,(\vee B_i) = \psi\,(\vee B_i) \in \mathcal{B}.$$

In the same way, we also obtain

$$\widetilde{\gamma}\widetilde{\varphi}\,(\wedge\, B_i) = \widetilde{\gamma}\,(\wedge\, B_i) = \psi\,(\wedge\, B_i) \in \mathcal{B}.$$

In other words, $\mathcal{B}$ is a *complete lattice* with respect to the ordering on $\mathcal{B}$ induced by $\le$, i.e. any family B_i in $\mathcal{B}$ has a smallest upper bound $\widetilde{\varphi}\,(\vee\, B_i)\,\mathcal{B}$ and a greatest lower bound $\widetilde{\gamma}\,(\wedge\, B_i) \in \mathcal{B}$.

Conversely, let us assume that $\mathcal{B}$ is a complete lattice. Thus, for any $A \in \mathcal{L}$, the family $\{B : B \in \mathcal{B}, B \ge A\}$ has in $\mathcal{B}$ a greatest lower bound, which is

$$\widetilde{\gamma}\,(\wedge\,\{B : B \in \mathcal{B}, B \ge A\}) = \widetilde{\gamma}\,\widetilde{\varphi}\,(A) \in \mathcal{B}.$$

But this implies $\mathcal{B}_{\psi_M} \subseteq \mathcal{B}$ for the filter $\psi_M = \widetilde{\gamma}\,\widetilde{\varphi}$. Conversely, for any $B \in \mathcal{B}$ we have $\widetilde{\gamma}\,(B) = \widetilde{\varphi}\,(B) = B$, and thus $\psi_M\,(B) = B$, i.e. $\mathcal{B} \subseteq \mathcal{B}_{\psi_M}$.We conclude that $\mathcal{B}_{\psi_M} = \mathcal{B}$, and $\mathcal{J}\mathrm{d}(\mathcal{B})$ is not empty. In other words, we have the following.

Theorem 4.11 *Let $\mathcal{B}$ be a subset of $\mathcal{L}$. Then $\mathcal{J}d(\mathcal{B})$ is not empty if and only if $\mathcal{B}$ is a complete lattice with respect to the ordering on $\mathcal{B}$ induced by $\le$, i.e.*

$\widetilde{\varphi}\,(\vee\, B_i) \in \mathcal{B}$, $\qquad \widetilde{\gamma}\,(\wedge\, B_i) \in \mathcal{B}$ *for any family B_i in $\mathcal{B}$.*

If so, then we have

$\widetilde{\varphi}\,(\vee\, B_i) = \psi\,(\vee\, B_i)$, $\qquad \widetilde{\gamma}\,(\wedge\, B_i) = \psi\,(\wedge\, B_i)$ *for any $\psi \in Jd(B)$.*

4.6 Complete lattice of the filters on $\mathcal{L}$

We now go back to the complete lattice $\mathcal{L}'$ of the increasing mappings on $\mathcal{L}$, and we purpose to prove that the class $\mathcal{V} \subseteq \mathcal{L}\,'$ of the filters on $\mathcal{L}$ forms a complete lattice. We shall use the theorem of structure 4.8. Clearly, if (ψ_i) is a family of elements of $\mathcal{V}$, then $\vee\psi_i$ is an overfilter,

$$(\vee_i\psi_i)(\vee_i\psi_i) = \vee_i(\psi_i(\vee_j\psi_j)) \geq \vee_i(\psi_i \circ \psi_i) = \vee_i\psi_i,$$

and similarly, $\wedge\psi_i$ is an underfilter. Therefore, we have to proe that every overfilter admits a smaller filter that majorates it.

Given $\psi \in \mathcal{L}'$, consider the overfilter $g = G\psi$. The class $\mathcal{C}$ of the underpotent of upper bounds g, which is closed under $\wedge$, admits by definition $f = Fg$ as its smaller element. Now, $\mathcal{C}$ is also closed under self-composition, since $h \geq g$ implies $hh \geq gg \geq g$ (g is an overfilter). Therefore $ff \in \mathcal{C}$, hence $ff \geq f$. But we have also the inverse inequality, for f is an underfilter, so that f is idempotent. Then, for all $\psi \in \mathcal{L}'$, operator $f = FG\psi$ being idempotent is also overfilter, which implies, by definition of G, that

$$GFG\psi = FG\psi \qquad \forall\psi \in \mathcal{L}'$$

i.e. from theorem 4.7, $GF \geq FG$. The condition of the theorem of structure are satisfied, which allows to state the following.

Theorem 4.12 *The class $\mathcal{V}$ of the filters on a complete lattice $\mathcal{L}$ is itself a complete lattice for the ordering induced by $\mathcal{L}$, and where for any family $\{\psi_i\} \leq \mathcal{V}$*
$F(\vee\psi_i)$ is the smaller filter that majorates the ψ_i's
$G(\wedge\psi_i)$ is the larger filter that minorates the ψ_i's

Note that this theorem of existences does not provides by itself a mean to calculate $F(\vee\psi_i)$, or Fg, for a given overfilter g. Here is a second, and more practical characterization.

Corollary 4.1 *Let $g \in \mathcal{L}'$ be an overfilter. Then Fg is nothing but the largest element of the class closed under $\vee$ and selfcomposition which is generated by g.*

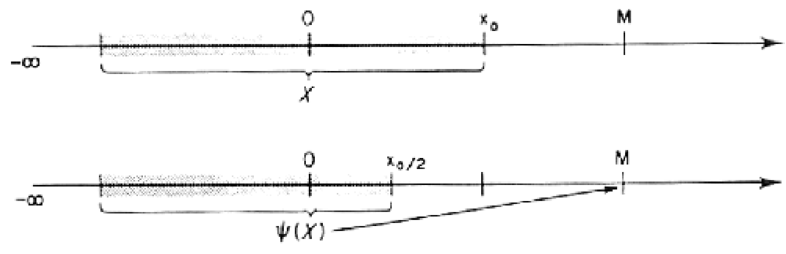


Figure 4.1: A counterexample showing that when an iterative appraoch works in Euclidean space, it may fail completely.

proof: Associate with $f = Fg$ the class C_1 of all overfilters $\leq f$. This class contains g and is closed under $\vee$ and selfcomposition. Hence it also contains the class C_0 closed under $\vee$ and selfcomposition which is generated by g. Now C_0 admits a largest element f_0, hence $f_0 \leq f$. but f_0 is idempotent ($f_0 f_0 \in C_0$, by closure under selfcomposition, i.e. $f_0 f_0 \leq f_0$, and $f_0 f_0 \geq f_0$ as an overfilter). Therefore f_0 is an underfilter larger than g, hence larger than $f = Fg$, which is by definition the smaller underfilter larger than g. □

In particular , when lattice $\mathcal{L}$ is finite, we always have, for n large enough

$$F = Fg = g^n$$

In such a case, filter f (i.e. the filter supremum of the family $\{\psi_i\}$ of filters when $g = \vee\psi_i$) is just obtained as the limit iteration of g.However, such a simple technique fails as soon as lattice $\mathcal{L}$ is no longer finite. Here is an example

Consider, in $\mathbb{R}^1$, a point M that has a positive abscissa, and the set $E =]-\infty, M]$. Define the mapping ψ on the lattice $\mathcal{P}(E)$ as follows:

$$\psi(X) = \begin{cases} \varnothing & \text{when } x_0 = \sup\{x \in X\} \leq 0, \\ M \cup \left(X \cap \left]-\infty, \tfrac{1}{2}x_0\right[\right) & \text{when } x_0 \geq 0, \end{cases}$$

where $X \in \mathcal{P}(E)$. It is easy to verify that $\psi(X)$ is increasing and idempotent. After n iterations of $I \wedge \psi$, we obtain

$$(I \wedge \psi)^n (X) = X \cap]-\infty, x_0/2^n[$$

and

$$(I \wedge \psi)^\infty (X) = X \cap]-\infty, 0]\,;$$

hence

$$(I \wedge \psi)^{\infty+1} (X) = (I \wedge \psi)(X \cap]-\infty, 0]) = \varnothing,$$

i.e.

$$(I \wedge \psi)^\infty (X) \neq (I \wedge \psi)^{\infty+1} (X) \ldots!!$$

Examples:

1. Lattice of the openings: take the class $\mathcal{V}' \subseteq \mathcal{V}$ of the openings on $\mathcal{L}$. We have seen that for every family (γ_i) in $\mathcal{V}'$, $\vee\gamma_i$ is still an opening (theorem 3.3). Moreover, from theorem 4.12, there exists a largest filter g which is smaller than all the γ_i's. g being obviously anti-extensive, $\mathcal{V}'$ is a complete lattice.

2. Start from an arbitrary increasing mapping ψ. Then, the extensive mapping $I \vee \psi$ is an overfilter and the proof of the theorem shows that there exists a smaller filter $\hat{\psi}$—hence a smaller **closing**—that majorates ψ. In the finite case, we find again theorem 3.8.

4.7 $\vee-$ and $\wedge-$ Filters

4.7.1 Introduction

We say that a mapping $f : \mathcal{L} \longrightarrow \mathcal{L}$ is a $\vee$-mapping when

$$f = f \circ (I \vee f) \tag{4.7}$$

and a $\wedge$-mapping when

$$f = f \circ (I \wedge f). \tag{4.8}$$

Basically, this property is something new and independent from the two axioms which build the definition of the morphological filters. If now f is increasing and satisfies rel. (4.7), we shall call it a $\vee$-underfilter. Indeed, any

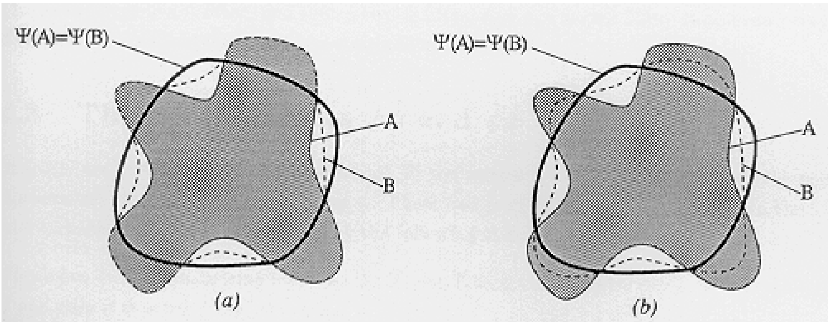


Figure 4.2: An example of a $\vee$-filter (a) and of a strong filter (b).

$\vee$-underfilter is an underfilter and similarly, any $\wedge$-overfilter is an overfilter. If f is, for instance, a $\vee$-underfilter, then we have:

$$f = f \circ (I \vee f) \geq f \vee ff \geq f.$$

Thus, $f = f \vee ff$ is an underfilter.

A filter which satisfies rel. (4.7) (resp rel. (4.8)) will be called a $\vee$-filter (resp. a $\wedge$-filter). When it satisfies both rel. (4.7) and (4.8) it will be said to be a **strong filter.** The geometrical interpretation of $\vee$-filtering and of $\wedge$-filtering are very easy. Indeed, ψ is a $\vee$-filter if and only if, for any $A \in \mathcal{L}$, every B between A and $A \vee \psi(A)$ has the same transform as A itself (see Fig.4.2), i.e.

$$\psi \ \vee -\text{filter} \quad \forall A \in \mathcal{L}, (A \leq B \leq A \vee \psi(A) \Longrightarrow \psi(B) = \psi(A)).$$

Similarly, we have

$$\psi \ \wedge -\text{filter} \quad \forall A \in \mathcal{L}, (A \wedge \psi(A) \leq B \leq A \Longrightarrow \psi(B) = \psi(A)).$$

Examples:

If γ is an *opening* and φ a *closing,* we have $\gamma \leq I \leq \varphi$, and then

- γ and φ are strong filters,

- $\gamma\varphi$ and $\varphi\gamma\varphi$ are $\vee$-filters,
- $\varphi\gamma$ and $\gamma\varphi\gamma$ are $\wedge$-filters.

Moreover, if $\gamma\varphi$ is a $\wedge$-filter, and thus a strong filter, $\varphi\gamma\varphi$ is a strong filter. In the same way, if $\varphi\gamma$ is a strong filter, then $\gamma\varphi\gamma$ is a strong filter.

4.7.2 Characterization of the $\gamma\varphi$ type filters

We have associated with each $\psi \in \mathcal{L}'$ the largest opening $\check{\psi}$ that minorates ψ and the smallest closing $\hat{\psi}$ which majorates ψ. These two primitives play a central role in the $\vee$- and $\wedge$-characterizations, as is shown by the following theorem:

Theorem 4.13 *An increasing mapping $\psi : \mathcal{L} \longrightarrow \mathcal{L}$ is a $\vee$-underfilter (resp. a $\wedge$-overfilter) if and only if $\psi = \psi\hat{\psi}$ (resp. $\psi = \psi\check{\psi}$).*

proof: If $\psi = \psi(I \vee \psi)$, then

$$(I \vee \psi)(I \vee \psi) = I \vee \psi \vee \psi(I \vee \psi) = I \vee \psi \vee \psi = I \vee \psi.$$

The mapping $I \vee \psi$, which is idempotent and which majorates I, is nothing but $\hat{\psi}$, and $\psi(I \vee \psi) = \psi\hat{\psi}$.

Conversely, start from

$$\hat{\psi} \leq \hat{\psi}(I \vee \psi) \leq \hat{\psi}(I \vee \hat{\psi}) = \hat{\psi},$$

which implies $\hat{\psi} = \hat{\psi}(I \vee \psi)$. Now, if $\psi = \psi\hat{\psi}$, then

$$\psi = \psi\hat{\psi} = \psi\hat{\psi}(I \vee \psi) = \psi(I \vee \psi).$$

$$\check{\psi} \leq \psi\check{\psi} \leq \psi \leq \psi\hat{\psi} \leq \hat{\psi}. \tag{4.9}$$

□

We have seen that the product $\gamma\varphi$ of any closing followed by any opening was a $\vee$-filter. We will prove now that the converse is true, so that the $\vee$-property **characterizes** the class of the filters of the type $\gamma\varphi$.

Theorem 4.14 *A mapping $\psi \in \mathcal{L}'$ is a $\vee$-filter (resp. a $\wedge$-filter) if and only if there exist an opening γ and a closing φ such that $\psi = \gamma\varphi$ (resp. $\psi = \varphi\gamma$).*

proof: Assume that ψ is a $\vee$-filter and consider its invariance domain $\mathcal{B}$. Denote by $\underset{\sim}{\mathcal{B}}$ the class closed for the sup which is generated by $\mathcal{B}$, and by $\widetilde{\gamma}$ the associated opening. Clearly, we have $\psi \geq \widetilde{\gamma}$. Moreover, according to criterion 4.5, $\mathcal{B} \subseteq \underset{\sim}{\mathcal{B}}$ implies $\psi = \widetilde{\gamma}\psi$ and (theorem 4.13) $\psi = \widetilde{\gamma}\psi\hat{\psi}$. Thus, we may write:

$$\psi = \widetilde{\gamma}\psi = \widetilde{\gamma}\psi\hat{\psi} \begin{cases} \geq \widetilde{\gamma}\hat{\psi} & \text{for} \quad \psi \geq \widetilde{\gamma}, \\ \leq \widetilde{\gamma}\hat{\psi} & \text{for} \quad \psi \leq \hat{\psi}. \end{cases}$$

Hence, $\psi = \widetilde{\gamma}\hat{\psi}$, i.e. the composition product of a closing by an opening. □

Remark: the above decomposition is not unique. We also have, for a $\vee$-filter ψ, $\psi = \check{\psi}\hat{\psi}$.

4.8 Conclusion

In the framework of morphological filtering, the $\vee$- and $\wedge$-properties are weaker substitutes for extensivity and anti-extensivity (closings and openings are strong filters, but the converse is false). Such a weaker version allows us to combine both properties in filters that are not trivial, whereas the only strong filter to be extensive and anti-extensive at the same time is the identity mapping I.

Chapter 5

Alternating Sequential Filters

5.1 Introduction

The appearance of the alternating sequential filters in the world of mathematical morphology is due to an experimental work of S.R.Sternberg [75]. His study consisted in taking a polyhedric form (namely a cuboctahedron, see § 2.6), altering it by the addition of a largely varying white noise, and then trying to clean the resulting image X . To attain this goal, X was first filtered by a small closing φ_1, followed by a small opening γ_1, then by a slightly larger closing φ_2 followed by a slightly larger opening γ_2, etcThe final operator produced by this succession of openings and closings was

$$M = (\gamma_i \varphi_i) \circ \ldots \circ (\gamma_2 \varphi_2) \circ (\gamma_1 \varphi_1).$$

The family (φ_i) that was used in this example consisted in morphological closings by homothetic structuring elements, whereas (γ_i) was the dual family of openings. After this experiment, a certain number of questions arose: is the operator M a filter? To what extend does it depend on the totality of the sequence of parameters 1, 2, i? Must these parameters be integers? Is it essential to use a size distribution (γ_i) and its dual (φ_i)? Is the product of these operators an operator of the same type?

The theory which answers these questions is due to J.Serra ([62]Ch. X); it covers both continuous and discrete cases. Here, for the sake of simplicity, we firtsly present the basic results of [62] in the general framework of the complete lattices, and then we restrict ourselves to the discrete case (i.e. $\mathcal{P}(\mathbb{Z}^n)$,

$\mathcal{F}(\mathbb{Z}^n, \mathbb{Z})$ or planar graphs), for establishing the various derived properties of these operators (e.g. dual form, symmetrical form, laws of composition, etc.). The end of the chapter is concerned with more practical problems, such as computation time and the use of these filters for concrete applications.

5.2 Definition of an A.S.F.

Given a complete lattice $\mathcal{L}$, define *a pair of primitives* as a granulometry $(\lambda, X) \to \gamma_\lambda(X)$ and an anti-granulometry$(\lambda,X) \to \varphi_\lambda(X)$ from $\mathbb{R}^+ \times \mathcal{L}$ into $\mathcal{L}$. In other words, for all $\lambda > 0$, γ_λ is an opening and φ_λ is a closing and such that

$$\lambda \geqslant \mu \Longrightarrow \gamma_\lambda \leq \gamma_\mu \qquad \text{and} \qquad \varphi_\lambda \geq \varphi_\mu.$$

These two mappings are chosen independently of one another (which does not necessarily bring us to self-duality). Moreover, we assume that, for all λ, the two mappings γ_λ and φ_λ are $\downarrow$-continuous. Now set

$$m_\lambda = \gamma_\lambda \varphi_\lambda$$

(which is obviously a filter), and for all pairs $\lambda, \lambda' \in \mathbb{R}^+$, with $\lambda' \geq \lambda$, construct the sequence of products

$$\begin{aligned} M_\lambda^0 &= m_\lambda \\ M_\lambda^1 &= m_{\lambda/2} m_\lambda \end{aligned}$$

$$M_\lambda^k = m_{\lambda.2^{-k}} ... m_{i\lambda.2^{-k}} ... m_\lambda$$

which are obtained by dividing the segment $[0, \lambda]$ into 2^k equal parts and then taking the extremities of these parts as increasing indices for m. It is easy to see that

$$\lambda \geq \mu \qquad \Longrightarrow \qquad m_\lambda m_\mu \leq m_\lambda \quad \text{and} \quad m_\mu m_\lambda \geq m_\lambda$$

(For example, from $\varphi_\lambda \varphi_\mu = \varphi_\lambda$we find $\varphi_\lambda \gamma_\mu \varphi_\mu \leq \varphi_\lambda$ and $m_\lambda m_\mu \leq m_\lambda$). It then follows that for all k, $M_\lambda^k M_\lambda^k = M_\lambda^k$ and $k' \geq k \Longrightarrow M_\lambda^{k'} \leq M_\lambda^k$, and the

M_λ^k are decreasing filters w.r.t. $k \in \mathbb{Z}_+$. Therefore we naturally introduce the infimum w.r.t. k of the M_λ^k :

$$M_\lambda = \wedge\{M_\lambda^k,\ k \in \mathbb{Z}_+\}$$

Mapping M_λ is increasing as an inf of increasing mappings. It is also idempotent: since M_λ^k decreases as a function of k, we can write

$$M_\lambda^{k\vee k'} \leq M_\lambda^k M_\lambda^{k'} \leq M_\lambda^{k\wedge k'}$$

This gives us

$$M_\lambda \leq \wedge_{k'} M_\lambda^k M_\lambda^{k'} \leq M_\lambda^k$$

and because of the $\downarrow$continuity of the primitives γ_λ and φ_λ,

$$M_\lambda \leq \wedge_k \wedge_{k'} M_\lambda^k M_\lambda^{k'} = \wedge_k M_\lambda^k[\wedge_{k'} M_\lambda^{k'}] = M_\lambda M_\lambda \leq \wedge_k M_\lambda^k = M_\lambda$$

In summary, we have the following.

Theorem 5.1 *Let* $\{\gamma_\lambda\}$ *be a granulometry and* $\{\varphi_\lambda\}$*be an anti-granulometry, both defined on a complete lattice* $\mathcal{L}$. *The indexed sequence of filters*

$$M_\lambda^k = m_{\lambda.2^{-k}}...m_{i\lambda.2^{-k}}...m_\lambda$$

with $\lambda \geq 0$, $m_\lambda = \gamma_\lambda\ \varphi_\lambda$ *and* $0 \leq i \leq 2^k$ *allows us to define an operation*

$$M_\lambda = \wedge\{M_\lambda^k, k \in \mathbb{Z}_+\}$$

which is a morphological filter called an alternating sequential filter *of primitives* $\{\gamma_\lambda\}$ *and* $\{\varphi_\lambda\}$ *and of size* λ .

Corollary 5.1 *The ASF M is* $\downarrow$ *continous.*

proof: For any sequence $\{X_p\}$ of elements of $\mathcal{L}$ such that $X_p \downarrow X, X \in \mathcal{L}$, the $\downarrow$ continuity of γ_λ and φ_λ implies that $M_\lambda^k(X_p) \downarrow M_\lambda^k(X)$. Therefore

$$M_\lambda(X) = \wedge_k M_\lambda^k(X) = \wedge_k \wedge_p M_\lambda^k(X_p) = \wedge_p \wedge_k M_\lambda^k(X_p) = \wedge_p M_\lambda(X_p).$$

□

To obtain stronger properties, we must provide $\mathcal{L}$ with a topological status in which the infimum operation is upper semi-continuous (e.g. in Euclidean spaces : the lattices of the closed sets, the u.c.s. numerical functions, the Lipschitz functions). In this framework, we assume both primitives γ_λ and φ_λ to be upper-semicontinuous.Then we have the following.

Theorem 5.2 *Let $\mathcal{L}$ be a topological lattice in which the infimum operator is u.s.c.. If the two primitives $(\lambda, X) \to \gamma_\lambda(X)$ and $(\lambda, X) \to \varphi_\lambda(X)$ are u.s.c., then*

1/ the ASF $M_\lambda : (\lambda, X) \to M_\lambda(X)$ is also an u.s.c. mapping from $\mathbb{R}^+ \times \mathcal{L}$ into $\mathcal{L}$;

2/ as varies, the ASF M_λ satisfy the following law of absoption:

$$\lambda \geq \mu > 0 \Longrightarrow M_\lambda M_\mu = M_\lambda \tag{5.1}$$

(Proof in[62], p.208). Absorption law 5.1 is not commutative in general; we have only

$$\lambda \geq \mu > 0 \Longrightarrow M_\mu M_\lambda \geq M_\lambda,$$

equality being reached in the important case of the *connected filters* (see chapter 6 below).

5.3 Discrete Sequential Alternating Filters

In the following, $\mathcal{L}$ denotes a discrete lattice such as $\mathcal{P}(\mathbb{Z}^n)$ or $\mathcal{F}(\mathbb{Z}^n, \mathbb{Z})$. We also consider two families $(\gamma_i)_{i\geq 1}$ and $(\varphi_i)_{i\geq 1}$—indexed on $\mathbb{Z}^+/\{0\}$—of operators on $\mathcal{L}$, which are a size distribution and an anti-size distribution respectively, i.e:

$$\begin{aligned} &\forall i \in \mathbb{Z}, i \geq 1, \quad \gamma_i \text{ is an opening and } \phi_i \text{ is a closing.} \\ &\forall i, j \in \mathbb{Z}, 1 \leq i \leq j, \quad \gamma_j \leq \gamma_i \leq I \leq \varphi_i \leq \varphi_j. \end{aligned} \tag{5.2}$$

Moreover, we have seen in definition 2.7 that the inequalities (5.2) are equivalent to the following property:

$$\forall i, j \in \mathbb{Z}, \qquad \gamma_i\gamma_j = \gamma_j\gamma_i = \gamma_{\max(i,j)} \text{ and } \varphi_i\varphi_j = \varphi_j\varphi_i = \varphi_{\max(i,j)} \tag{5.3}$$

Remark that these two families are chosen **independently** from one another (although they are often taken as dual of each other in practice).

Now, we know from theorem 3.5 that by composing two filters f and g such that $f \leq g$ we get new filters. In particular for all $i \in \mathbb{Z}, i \geq 1$the two composition products

$$m_i = \gamma_i \varphi_i \qquad n_i = \varphi_i \gamma_i$$

defined the so-called *alternated filters*

Proposition 5.3 *For $i, j \in \mathbb{Z}$ such that $1 \leq i \leq j$, we have:*

$$m_j m_i \leq m_j \leq m_i m_j \tag{5.4}$$

$$n_i n_j \leq n_j \leq n_j n_i. \tag{5.5}$$

proof: Let us show first that $m_j m_i \leq m_j \leq m_i m_j$. We know that $\gamma_i \leq I$, hence, $\gamma_i \varphi_i \leq \varphi_i$. Thus, since $\varphi_i \leq \varphi_j$, the preceding inequality implies

$$\gamma_i \varphi_i \leq \varphi_j.$$

Now, φ_j being increasing, $\varphi_j \gamma_i \varphi_i \leq \varphi_j \varphi_j = \varphi_j$ (by idempotence of φ_j). γ_j is also increasing, hence we finally obtain $\gamma_j \varphi_j \gamma_i \varphi_i \leq \gamma_j \varphi_j$, i.e:

$$m_j m_i \leq m_j,$$

which is the first inequality. Similarly, $I \leq \varphi_i$ yields (γ_i increasing) $\gamma_i \leq \gamma_i \varphi_i$. Thus, the family (γ_i) being a size distribution, $\gamma_j \leq \gamma_i \varphi_i$, which in turn implies $\gamma_j \gamma_j = \gamma_j \leq \gamma_i \varphi_i \gamma_j$ (γ_j idempotent). Therefore, $\gamma_j \varphi_j \leq \gamma_i \varphi_i \gamma_j \varphi_j$, i.e:

$$m_j \leq m_i m_j.$$

A similar proof is used for the first part of relation (5.5). □

Proposition 5.4 *Let $(i_k)_{1 \leq k \leq p}$ be p numbers such that*

$$\forall k, i_k \in \mathbb{Z}, 1 \leq i_k \leq i_1 = i_p.$$

then:

$$\begin{cases} m_{i_p} m_{i_{p-1}} \dots m_{i_2} m_{i_1} & = & m_{i_p} & = & m_{i_1} \\ n_{i_p} n_{i_{p-1}} \dots n_{i_2} n_{i_1} & = & n_{i_p} & = & n_{i_1} \end{cases}$$

proof: Let us prove for instance the first relation. $i_2 \leq i_1$, hence, by rel. (5.4), $m_{i_2} m_{i_1} \geq m_{i_1}$. Then, m_{i_3} being increasing, we have $m_{i_3}(m_{i_2} m_{i_1}) \geq m_{i_3} m_{i_1}$, and applying again rel. (5.4) gives, since $i_3 \leq i_1$: $m_{i_3}(m_{i_2} m_{i_1}) \geq m_{i_1}$. If we iterate this process, we finally get

$$m_{i_p} m_{i_{p-1}} \ldots m_{i_2} m_{i_1} \geq m_{i_1}.$$

Conversely, $i_{p-1} \leq i_p$ yields, by rel.(5.4), $m_{i_{p-1}} m_{i_p} \leq m_{i_p}$. Then, $m_{i_{p-2}}$ being increasing, we have $m_{i_{p-2}}(m_{i_{p_1}} m_{i_p}) \leq m_{i_{p-2}} m_{i_p}$, and applying again rel.(5.4), we obtain ($i_3 \leq i_1$): $m_{i_{p-2}}(m_{i_{p-1}} m_{i_p}) \leq m_{i_p}$. After iterating this process, we finally get

$$m_{i_p} m_{i_{p-1}} \ldots m_{i_2} m_{i_1} \leq m_{i_p},$$

which completes the proof of the first relation. Similar proof for the second one. □

We can now give the definition of the alternating sequential filters and prove that they are effectively filters:

Definition 5.5 *For all $i \in \mathbb{Z}, i \geq 1$, the following operators:*

$$M_i = m_i m_{i-1} \ldots m_2 m_1 \qquad N_i = n_i n_{i-1} \ldots n_2 n_1$$

are called alternating sequential filters of order i.

Proposition 5.6 $\forall i \in \mathbb{Z}, i \geq 1$, *the operators M_i and N_i are filters.*

proof: These operators are increasing as compositions of increasing mappings. Moreover, we have:

$$\begin{aligned} M_i M_i &= (m_i m_{i-1} \ldots m_2 m_1)(m_i m_{i-1} \ldots m_2 m_1) \\ &= (m_i m_{i-1} \ldots m_2 m_1 m_i)(m_{i-1} \ldots m_2 m_1) \\ &= m_i(m_{i-1} \ldots m_2 m_1) \\ &= M_i. \end{aligned}$$

The idempotence of M_i is thus proved. That of N_i has a similar proof. □

5.4 Discrete ASF Properties

We have already said that there is no need for duality between the γ_i's and the ψ_i's. Anyway, an ASF *cannot be self dual,* since $N_i \neq M_i$. In the present section, some properties of the ASF's with respect to the composition product and the order relationship between operators are presented. We also deal with new filters, which are derived from the previous ones and which are called *transposed ASF's.* Lastly, *symmetrical alternating filters* are introduced, and some of their properties reviewed.

Proposition 5.7 (Absorption laws) *For $i, j \in \mathbb{Z}$ such that $1 \leq i \leq j$, we have the following relations:*

$$M_j M_i = M_j \leq M_i M_j \tag{5.6}$$

$$N_i N_j \leq N_j = N_j N_i. \tag{5.7}$$

proof: We shall only proof (5.6), since the two relations (5.7) derive by duality. The following egality holds: $M_j M_i = m_j m_{j-1} \dots m_{i+1} M_i M_i$. Therefore, M_i being idempotent, $M_j M_i = m_j m_{j-1} \dots m_{i+1} M_i$, i.e. $M_j M_i = M_j$.

Let us show now the second part of this equation, i.e. $M_j \leq M_i M_j$: we can write $M_j = \gamma_j M_j = (\gamma_i \gamma_{i-1} \dots \gamma_1)\gamma_j M_j$. Now, $I \leq \varphi_1$ (extensivity), which implies $\gamma_1 \leq \gamma_1 \varphi_1$. Thus, $\gamma_1 \leq \varphi_2 \gamma_1 \varphi_1$, which in turn implies the inequality $\gamma_2 \gamma_1 \leq \gamma_2 \varphi_2 \gamma_1 \varphi_1$. By iterating theis process, we finally obtain

$$\gamma_i \gamma_{i-1} \dots \gamma_1 \leq \gamma_i \varphi_i \gamma_{i-1} \varphi_{i-1} \dots \gamma_1 \varphi_1 = M_i.$$

Therefore, $M_j \leq M_i \gamma_j M_j$, i.e.

$$M_j \leq M_i M_j.$$

This completes the proof of the first relation. There is a similar proof for the second part of rel. (5.7). □

Compatibility under magnification

When the two primitive families (γ_i) and (φ_i) are compatible under magnification, i.e.

$$\forall X, \forall k > 0, \quad \left\{ \begin{array}{lcl} \varphi_k(kX) & = & k\varphi_1(X), \\ \gamma_k(kX) & = & k\gamma_1(X), \end{array} \right. \tag{5.8}$$

this property is transmitted to the corresponding ASF's. The physical interpretation of this is clear: it means that the ASF of order k works on the k times magnified image exactly as does the ASF of order 1 on the initial image.

5.5 Applications, computation time

These alternating filters are among the most useful filters in mathematical morphology. They actually provide efficient filtering in many image cleaning problems, and can be finely adjusted to each case. Indeed, we can operate on:

- the families (γ_i) and (φ_i) (most of the time, this comes down to choosing families of structuring elements),
- the type of filter (alternating sequential filter, alternating transposed filter, alternating symmetrical filter,
- the "size" of the filter,
- etc

The only problem with such filters (and with most filters used for image cleaning) is that of the computation time. Although the formulas for computing these filters can be "compacted", a certain number of elementary operations has yet to be performed in each case. Thus, the computation time of an ASF may well be very long on a non specialized equipment.

As an example, suppose that the computation times of φ_i and γ_i are equal to $i \times \Delta t$, for a fixed image size. Then, the computation time of M_i, N_i, etcis proportional to i^2, as shown by table 5.1 (we suppose that the most efficient fomulas are used):

Filter	Computation time	case of $i = 5$
M_i, N_i, M_i^t, N_i^t	$i(i+1)\Delta t$	$30 \times \Delta t$
R_i, S_i, R_i^t, S_i^t	$i(i+2)\Delta t$	$35 \times \Delta t$
$\tilde{M}_i, \tilde{N}_i$	$2i(i+1)\Delta t$	$60 \times \Delta t$
$\tilde{R}_i, \tilde{S}_i$	$(2i^2+2i+1)\Delta t$	$65 \times \Delta t$

Table 5.1: Computation time of some sequential filters, provided that the time required for computing γ_i or ϕ_i equals $i \times \Delta t$ (for a fixed image size).

In order to reduce the above computation times, it is sometimes useful to introduce the following type of alternating filters [62, page 213]:

$$\begin{aligned} M_i(2) &= m_i m_{i-2} \dots m_3 m_1 \\ M_i'(2) &= \varphi_i \gamma_{i-1} \varphi_{i_2} \dots \gamma_2 \varphi_1, \end{aligned}$$

i being here an odd number. These filters obey to similar absorption laws as those which were presented above. Moreover, their filtering capabilities are, in most concrete problems, practically as good as the capabilities of the alternating filters described in this chapter. However, they can be computed two times faster than the "regular" alternating sequential filters. This is extremely interesting in practice.

.

Chapter 6

Connections and Connected Filters

6.1 Classical connectivity and image analysis

In mathematics, the concept of connectivity is formalized in the framework of topological spaces and is introduced in two different ways. First, a set is said to be connected when one cannot partition it into two non empty closed (or open) sets. This definition makes precise the intuitive idea that $[0,1] \cup [2,3]$ consists of two pieces, while $[0,1]$ consists of only one. But this first approach, extremely general, does not derive any advantage from the possible regularity of some spaces, such as the Euclidean ones. In such cases, the notion of *arcwise connectivity* turns out to be more convenient. According to it, a set A is connected when, for every $a, b \in A$, there exists a continuous mapping ψ from $[0,1]$ into A such that $\psi(0) = a$ and $\psi(1) = b$. Arcwise connectivity is more restrictive than the general one ; however, in $\mathbb{R}^d$, any open set which is connected in the general sense is also arcwise connected.

A basic result governs the meaning of connectivity ; namely, the union of connected sets whose intersection is not empty is still connected :

$$\{A_i \text{ connected}\} \text{ and } \{\cap A_i \neq \emptyset\} \Rightarrow \{\cup A_i \text{ connected}\} \tag{6.1}$$

In discrete geometry, the digital connectivities transpose the arcwise corresponding notion of the Euclidean case, by introducing some elementary arcs between neighboring pixels. This results in the classical 4- and 8-square connectivities, as well as the hexagonal one, or the cuboctahedric one in 3-D space. Is such a metric approach to connectivity adapted to image analysis? We can argue that

a/ Certain arcwise connections seem smowhat shaky, *e.g.* when they do not treat equally a set and its complement;

b/ In discrete motion analysis, the trajectories of fast moving objects often appear as dotted tubes, and arcwise connections are unable to handle such situations;

c/ more deeply, one can wonder what is actually needed in image processing. As a matter of fact, when we examine the requirements for connectivity, we observe that the basic operation they involve consists, given a set A and a point $x \in A$, in extracting the particle of A at point x. For such a goal, an arcwise approach is obviously sufficent. But is it necessary?

6.2 The notion of a connection

These criticisms led G. Matheron and J. Serra to propose a new approach, in 1988 ([62], Chap. 2 and 7) where they take not rel.(6.1) as a consequence, but as a starting point. However, their definition is rather general and stated as follows.

Definition 6.1 *Let E be an arbitrary space. We call connected class or* connection $\mathcal{C}$ *any family in $\mathcal{P}(E)$ such that*

(i) $\varnothing \in \mathcal{C}$ *and for all $x \in E$,* $\{x\} \in \mathcal{C}$

(ii) for each family $\{C_i\}$ in $\mathcal{C}$, $\cap\, C_i \neq \varnothing$ *implies* $\cup C_i \in \mathcal{C}$.

As we can see, the topological background has been deliberately thrown out. The classical notions (e.g. connectivity based on digital or Euclidean arcs) are indeed particular cases, but the emphasis is put on another aspect, that answers the above criticism c/ in the following manner ([62], Chap. 2) :

Theorem 6.2 *The datum of a connection $\mathcal{C}$ on $\mathcal{P}(E)$ is equivalent to the family $\{\gamma_x\,, x \in E\}$ of openings such that*

(iii) for all $x \in E$, we have $\gamma_x(x) = \{x\}$

(iv) for all $A \subseteq E$, $x, y \in E$, $\gamma_x(A)$ and $\gamma_y(A)$ are equal or disjoint

(v) for all $A \subseteq E$, and all $x \in E$, we have $x \notin A \Rightarrow \gamma_x(A) = \varnothing$.

proof: First we show that the datum of C brings us to the openings γ_x. Axiom *(iii)* results from $\{x\} \in C$. To prove *(iv)*, note that $\gamma_x(A) \cap \gamma_y(A) \neq \varnothing$ implies

$$C = \gamma_x(A) \cup \gamma_y(A) \in \mathcal{C}, \qquad \text{with } C \subseteq A.$$

On the other hand, $\gamma_x(A)$ being non-empty gives

$$x \in \gamma_x(A) \Rightarrow x \in C \Rightarrow C \in \mathcal{C}_x \Rightarrow C \subseteq \gamma_x(A) \Rightarrow \gamma_y(A) \subseteq \gamma_x(A).$$

We show the reverse inclusion, thus equality, in the same way.

Conversely, suppose that we define the class $\mathcal{C}$ as the family of invariant sets of the γ_x, i.e.

$$\mathcal{C} = \{\gamma_x(A), x \in A, A \subseteq E\}.$$

For $A = \varnothing$ we find $\gamma_x(\varnothing) = \varnothing \in \mathcal{C}$. For $A = \{x\}$axiom *(iii)* implies that $\gamma_x(x) = \{x\} \in \mathcal{C}$, and axiom *(i)* is satisfied. Now let $\{C_i\}$ be a family with a non-empty intersection in $\mathcal{C}$ and $x \in \cap\ C_i$. As $C_i \in \mathcal{C}$, we can find a point y_i for each i such that $C_i = \gamma_{y_i}(C_i)$. But $x \in C_i$ therefore from *(iii)*, $\{x\} = \gamma_x(\{x\}) \subseteq \gamma_x(C_i)$. Thus $\gamma_{y_i}(C_i)$ and $\gamma_x(C_i)$ contain point x, and from *(iv)* we have $C_i = \gamma_{y_i}(C_i) = \gamma_x(C_i)$. So $\cup C_i = \cup\gamma_x(C_i)$ is invariant under $\gamma_x$and belongs to the class $\mathcal{C}$. Thus we have axiom *(ii)*

We still have to prove that the connected openings associated with this class $\mathcal{C}$ coincide with the γ_x themselves; i.e. to identify the following two classes

$$\mathcal{C}'_x = \{\gamma_x(A) : \gamma_x(A) \neq \varnothing, A \subseteq E\},$$

$$\mathcal{C}_x = \{\gamma_y(A) : y \in E, A \subseteq E, \gamma_y(A) \supseteq \{x\}\}.$$

From axiom*(iii)* we have $\{x\} \in \mathcal{C}_x$. Let $\gamma_x(A)$ be an element of $\mathcal{C}_x$; then $x \in \gamma_y(A) \subseteq A$ implies that $x \in \gamma_x(A)$, i.e. from axiom *(iv)*, that $\gamma_y(A) = \gamma_x(A)$. Hence $\mathcal{C}_x \subseteq \mathcal{C}'_x$. Conversely, set $\gamma_x(A) \in \mathcal{C}'_x$. Since $\gamma_x(A) \neq \varnothing$ axiom *(v)* implies that, $x \in A$; thus $\{x\} \subseteq \gamma_x(A)$and , i.e. $\mathcal{C}'_x \subseteq \mathcal{C}_x$. □

An alternative, but equivalent, axiomatics has been proposed by Ch. Ronse [56] ; it contains, as a particular case, another one by R.M. Haralick and L.G. Shapiro [22] ; however, both approaches are still set-oriented. The extension from sets to the general framework of complete lattices and in particular to numerical functions has been performed by J. Serra [71]

At first sight, this theorem just indicates that the operation shown in Fig.6.1 is an opening called *the connected component of A that contains point x,* which is somewhat obvious. But some other connectionx, below, are less obvious.

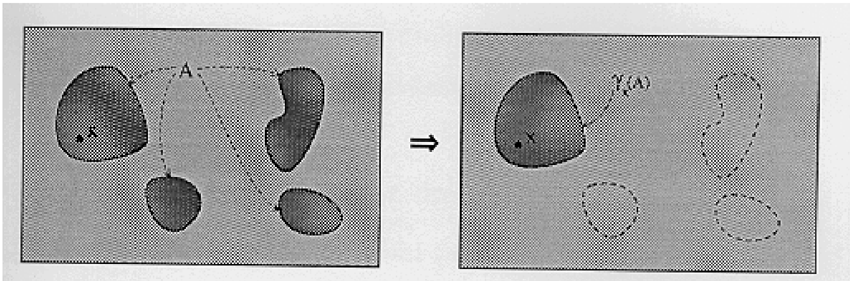


Figure 6.1: The opening called the connected component of A that contains point x.

Historically speaking, the number of applications or of theoretical developments which was suggested (and permitted) by this theorem is considerable (see, among others[36][48][58]). It has opened the way to an object-oriented approach for segmentation, compression and understanding of still and moving images.

6.2.1 Examples of connections on $\mathcal{P}(E)$

Several instructive examples of connections on $\mathcal{P}(E)$ can be found in [25], in [56] and in [71]. Here we just recall a few of them.

i/ All arcwise connectivities on digital spaces are connections in the sense of definition 6.1 (see6.2);

ii/ In [62]p.55, J. Serra provides E with a first connection $\mathcal{C}$ and considers an extensive dilation $\delta : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ that preserves $\mathcal{C}$ (*i.e.* $\delta(\mathcal{C}) \subseteq \mathcal{C}$). Then the inverse image $\mathcal{C}' = \delta^{-1}(\mathcal{C})$ of $\mathcal{C}$ under δ defines a new connection on $\mathcal{P}(E)$, which is richer. The $\mathcal{C}$ -components of $\delta(A)$, $A \in \mathcal{P}(E)$, are exactly the images $\delta(Y_i')$ of the $\mathcal{C}'$-components of A. If γ_x stands for the connected opening associated with connection $\mathcal{C}$ and ν_x for that associated with $\mathcal{C}'$, we have

$$\begin{array}{lll} \nu_x(A) = \gamma_x \circ \delta(A) \cap A & \textit{when} \quad x \in A & ; \\ \nu_x(A) = \varnothing & \textit{when} \quad x \notin A & \end{array}$$

(where δ is a dilation by a disc) is again a connected opening associated with the connectivity shown in Fig. 6.2, which is not so obvious.

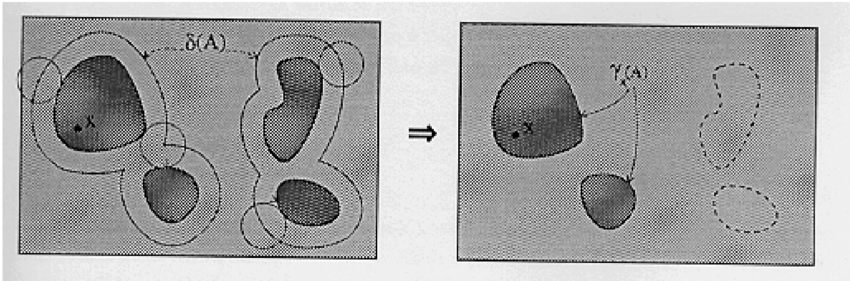


Figure 6.2: A less obvious connectivity notion, associated with the opening ν_x.

In practice, the openings ν_x characterize the *clusters* of objects from a given distance d apart. Fig. 6.3 illustrates this point by "reconnecting" dotted lines trajectories. But *a contrario*, such an approach can also provide a means to extract the objects which are isolated.They will be defined by the fact that for them $\nu_x(A) = \gamma_x(A)$, an equality which yields easy implementation [72].

iii/ In [56], Ch. Ronse starts also from a first connection on $\mathcal{P}(E)$, and proposes, as a new connection, the class generated by the points and the connected sets opened by a given structuring element B. If $x \in X \circ B$, then $\gamma_x(X)$ is the initial connected component of $X \circ B$ containing x, and when point $x \in X \backslash X \circ B$, then $\gamma_x(X) = \{x\}$. An example is shown in Fig.6.9 for such an "open"connection by a 3x3x square, the set of 6.9a has 16 particles: the two surrounded squares, plus 14 isolated points. Also, the six points of the vertical gulf are isolated pores.

6.3 Application to Function Segmenting

In this section, we investigate segmentation of numerical functions $f : E \to \mathcal{L}$ when a connection $\mathcal{C}$ is defined on $\mathcal{P}(E)$. We shall formulate the question as follows : "is there a largest partition of space E into connected classes such that function f satisfies a given criterion σ inside each class? Clearly, criterion σ must be consistent with the axioms of a connection.Therefore, we assume that, given any arbitrary function f

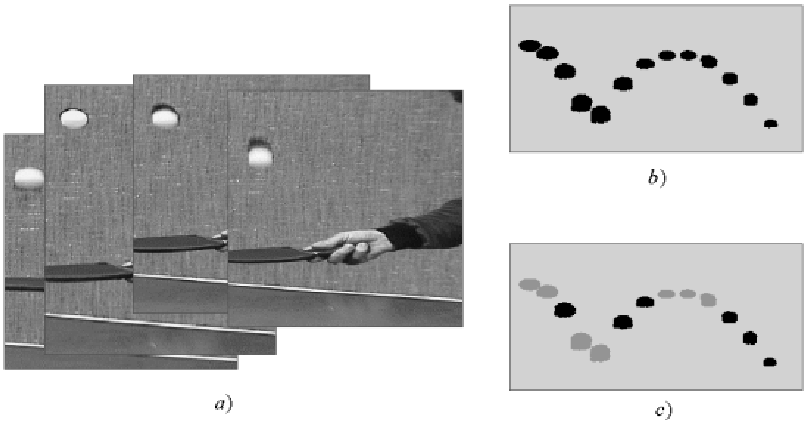


Figure 6.3: a) Sequence of images, b) Space-time display of the ball, c) In a dilation based connection, the three clusters in grey are considered as particules (they correspond to slow motions).

i/ for all $x \in E, f(x)$ fulfils criterion σ;

ii/ for all $A, B, \in \mathcal{C}$ with $A \cap B \neq \varnothing$, if f fulfils σ on A and on B, then f fulfils σ on $A \cup B$.

Hence, criterion σ generates a subclass $\mathcal{C}_\sigma$ of $\mathcal{C}$ which turns out to be a second connection on $\mathcal{P}(E)$. Therefore class $\mathcal{C}_\sigma$ *partitions* set E into the maximal classes satisfying criterion σ. Here are three examples of such segmentations

6.3.1 Flat Zones Connection

The segmentation of f by flat zones is the concern of this first approach. Given point $x \in E$, denote by Z_x the subclass of those elements of $\mathcal{C}$ that contains point $\{x\}$, and over every element of which function f is constant, i.e.

$$Z_x = \{x \in Z \subseteq \mathcal{C}, y \in Z \Rightarrow f(x) = f(y)\}\,.$$

Class Z_x is closed under union, so that its supremum

$$Z_x = \cup \{Z \in Z_x\}$$

is the largest connected component containing point x and on which function f is constant. Consequently, the mapping $x \rightarrow Z(x)$ is the largest partition of space E into flat zones of f. Furthermore, the family $\{Z_x \cup \emptyset\}$ defines the invariant sets of a connected opening at point x. This means that any set $Y \subseteq E$ is partitioned into the classes $Z(x) \cap Y$, as x spans E. By doing so, we have generated a new connection $\mathcal{C}_\sigma$ on E, by combining the initial one with some features of function f.

6.3.2 Smooth Path Connection

Provide $\mathcal{P}(\mathbb{R}^n)$ with the arcwise connection. Consider an arbitrry, but fixed, function $f : \mathbb{R}^n \to L$ and the class $\mathcal{C}_\sigma$ made of (a) all singletons plus the empty set, and (b) all connected sets Y of $\mathcal{P}(\mathbb{R}^n)$ such that function f is k-Lipschitz on set Y for the arcwise metric induced on this set (i.e. the so-called "geodesic metric" in practice).

The class of such sets Y is obviously closed under union, therefore it admits a largest element $Y(x)$. As x spans E, the sets $Y(x)$ partition space E, just as the flat zones $Z(x)$ did previously. And just as before, class $\mathcal{C}_\sigma$defines a second connection on $\mathcal{P}(\mathbb{R}^n)$, where the variation of f, over each set of class (b) is *"smooth"*in the k-Lipschitz sense.

In $\mathbb{Z}^2$, the implementation of this "*smooth path*" criterion is particularly easy. If $D(x)$ stands for the unit disc at a point x (with five, seven or nine points), then the partition has, for non-point classes, the arcwises connected components of all sets X such that

$$X = \cup \left\{ x \in \mathbb{Z}^2 : \vee \left| f(x) - f(y) \right|, y \in D(x) \leq k \right\}$$

The points of the complement set X^c correspond to connected components reduced to individual pixels.

An example of smooth path connection is given in Fig.6.4a, which represents an electron micrograph of concrete made of three phases : a white one and two grey ones. The histograms of the two grey phases are almost identical, but one is more continuous than the other. By segmenting Fig.6.4a according to the smooth path connection, with a slope $k = 6$, we obtain a correct pre-extraction, which has to be amended by some filtering. For the same image, the best jump connection is obtained by taking a range $h = 15$, and yields the rather poor result depicted in Fig.6.4c.

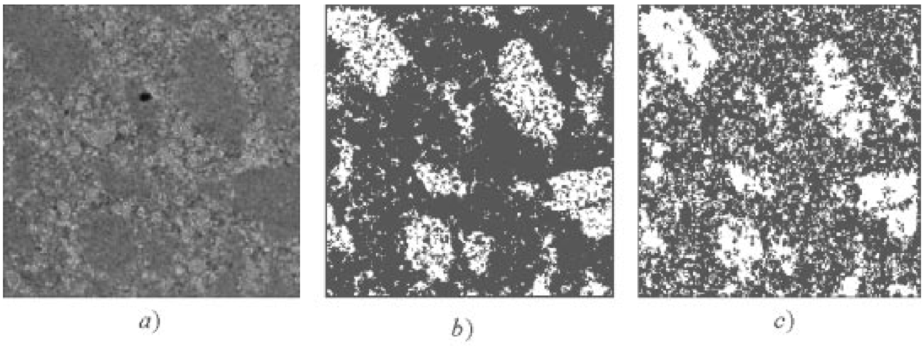


Figure 6.4: a) Original micrograph of beton, b) smooth path connection for k=6 (in dark: the connected component reduced to points), c) jump connection for h = 12

6.3.3 Jump Connection

Provide $\mathcal{P}(\mathbb{R}^n)$ with connection $\mathcal{C}$. Consider a given function $f : \mathbb{R}^n \to L$ and the class $\mathcal{C}_\sigma$ made of (a) the singletons plus the empty set, and (b) all connected sets around each minimum, and where the value of f is less than h above the minimum.

Clearly class $\mathcal{C}_\sigma$ forms a second connection on $\mathcal{P}(\mathbb{R}^n)$, called *jump connection from minima*, and induces a maximum partition on E. Let X_1 stand for the union of the non-point classes of this partition. By iterating the process on the restriction of f on X_1^c, we obtain a second family of non-point classes, of union X_2, etc... Finally, if function f is regular enough (or in the digital cases) the series of iterations results in a partition of the space into connected components (in the sense of $\mathcal{C}$) where the variation of f is smaller than h. In a variant of the procedure, one can alternate jump connection from minima and from the maxima in order to obtain a more self-dual result. The following example illustrates the use of such a transformation. Fig.6.5a depicts an optical micrograph of a polished section of alumina grains. The partition of the space under jump connection is depicted in , whereas Fig.6.5c depicts the superposition of the skeleton by influence zones of the set in (b) on the original image(a).

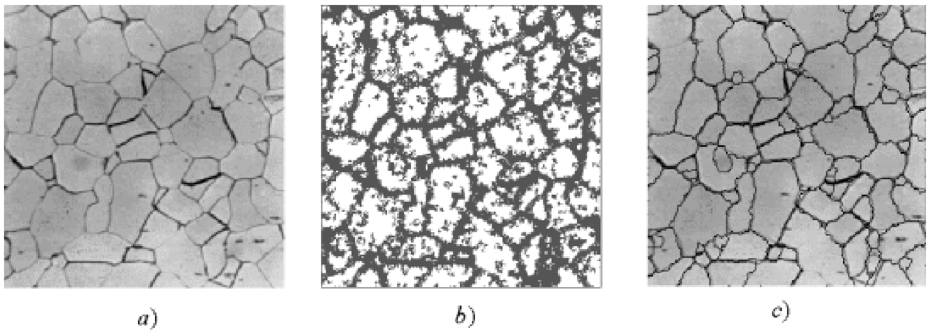


Figure 6.5: (a) original micrograph of alumina, (b) jump connection from the maxima , with h=15, (c) derived SKIZ.

6.4 Connected Filters

For now on E is an arbitrary set, and $\mathcal{P}(E)$ is supposed to be equipped with connection $\mathcal{C}$. For every set $A \in \mathcal{P}(E)$, the two families of the connected components of A (the "grains") and of A^c (the "pores") partition space E. Then, an operation $\psi : \mathcal{P}(E) \to \mathcal{P}(E)$ is said to be *connected* when the partition associated with $\psi(A)$ is coarser that that of A [69]. Clearly, taking the complement of a set, or removing some grains, or filling pores generate connected operators. The major class of mappings we have in view is that of the connected morphological filters.

6.4.1 Set opening by reconstruction and some derivatives

A comprehensive class of connected filters derives from the classical *opening by reconstruction*. Its definition appears in [62], chap.7.8. Significant studies which use this notion may be found in literature, such as [69] (connected operators),[13],(stable operators) [48],(spanning trees), [25],(grain oprators).

An opening by reconstruction is obtained by starting from an increasing binary criterion τ (e.g. "*the area of A is $\geq$ 10*"), to which one associates the trivial opening, in the sense of section 3.4.1.above, i. e.

$$\begin{array}{ll} \gamma^{\tau}(A) = A & \text{when } A \text{ satisfies the criterion} \\ \gamma^{\tau}(A) = \emptyset & \text{when not} \end{array}$$

The corresponding opening by reconstruction γ is then generated by applying the criterion to all grains of A, independently of one another, and by taking the union of the results :

$$\gamma(A) = \cup \{\gamma^{\tau}\gamma_x(A), \quad x \in E\}$$

A series of algorithms are based on this approach. For example, for 2-D binary images: *keep the connected components of X whose circular opening of size k is not empty and filter out the others.*

Reconstruction opening extends to numerical functions via their horizontal sections, but can be directly implemented in terms of numerical operators. If ε_k and δ_k stand respectively for the isotropic erosion and dilation of size k (square, hexagonal, octogonal in $\mathbb{Z}^2$ (sets), as well as in $\mathcal{F}(\mathbb{Z}^2, \mathbb{Z}^+)$ (functions from $\mathbb{Z}^2$ into $\mathbb{Z}^+$), we can write:

$$Y = \varepsilon_k(X) \ \text{(set case)} \qquad g = \varepsilon_k(f) \ \text{(function case)},$$

and then:

$$\mu^{(1)}(Y) = \delta_1(Y) \cap X \qquad \mu^{(1)}(g) = \delta_1(g) \wedge f. \tag{6.2}$$

By iterations, we compute $\mu^{(2)}(Y) = \mu \circ \mu(Y)$, $\mu^{(3)}(Y) = \mu \circ \mu^{(2)}(Y)$,etcThe sequence of the $\mu^{(i)}$s increases till an idempotent limit, which provides the desired opening (see Fig.6.6).The underlying methods, which are called *geodesic* ones, are presented in more details in chapter 11 below.

The closing by reconstruction φ (for the same criterion) is the dual of γ for the complement, *i.e.* if $\complement$ stands for the complement operator, then

$$\varphi = \complement\gamma\complement\,.$$

For example, in R^2, if we take for criterion τ, " *have an area* ≥ 10", then $\gamma(A)$ is given by the union of grains of A whose areas are ≥ 10, and $\varphi(A)$ is the union of A and all its pores whose areas are ≤ 10. Similarly, if criterion τ is expressed by "hit a fixed marker M", then $\gamma(A)$ is the union of the grains that hit A, wheras $\varphi(A)$ is composed of A and of all pores that miss M .

The operators by reconstruction satisfy number of nice properties. The three following ones are typical examples of them.

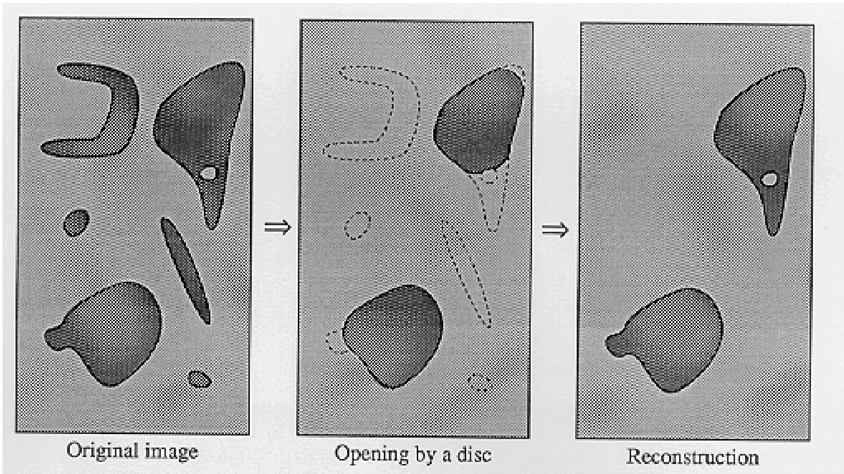


Figure 6.6: Algebraic opening as a morphological opening by a disc followed by a reconstruction.

Proposition 6.3 *[69]Let γ be an opening by reconstruction, and φ be a closing that does not create connected components, i.e. such that*

$$x \in \varphi(A) \Longrightarrow A \cap \gamma_x \varphi(A) \neq \varnothing \tag{6.3}$$

Then the associted alternated filters are ordered, and we have $\gamma\varphi \geq \varphi\gamma$

proof: Consider $\varphi\gamma(A)$, for $A \subseteq E$. Since the (extensive) closing φ does not create new connected components, it can only enlarge those of $\gamma(A)$; now γ acts grain by grain, hence $\gamma\varphi\gamma = \varphi\gamma$. According to theorem4.7 this equality is equivalent to $\gamma\varphi \geq \varphi\gamma$. □

The most common closings may not satisfy condition (6.3). It is the case for intersections of closings by segments, for example. However, if starting from an arbitrary closing φ, we restrict $\varphi(A)$ to its grains that contain at least one point $x \in A$, the resulting operation is still a closing. It is the reason for which condition (6.3) is always assumed implicity in practice.

Corollary 6.1 *Let $\{\gamma_i\}$ be a granulometry by reconstruction, and $\{\varphi_i\}$ be a anti-granulometry that does not create connected components, then the A.S.F.*

Figure 6.7: An example of a pyramid of connected alternated sequential filters. Each contour is preserved or suppressed, but never deformed : the initial partition increases under the successive filters, which are strong and form a semi-group.

$M_n = m_n...m_i...m_1$, *with* $m_i = \gamma_i\varphi_i$ *satisfy the semi-group relation*

$$M_nM_p = M_pM_n = M_{\sup(n,p)}$$

proof: We already know that $p \geq n \Rightarrow m_nm_p \geq m_p$ and $M_pM_n = M_p$. But we draw from the proposition that $m_nm_p = \gamma_n\varphi_n\gamma_p\varphi_p \leq \gamma_n\gamma_p\varphi_n\varphi_p = m_p$. Hence $p \geq i \Rightarrow m_im_p = m_p$, and $M_nM_p = m_n...m_i...m_1m_p...m_i...m_1 = M_p$. □

This corollary explains, partly at least, why the A.S.F. by reconstruction are so often involved in pyramids, for coding, segmentation, or indexation purposes. In such pyramids, the additional information to get finer levels is concentrated in subdivisions the flat zones [48]. An example of such a behaviour is presented in Fig.6.7. Each cross section of the gray tone image has processed by an alternating sequential filter by reconstruction. The underlying binary criterion was here associated with the size of the disc inscribable in each grain.

Another point of interest is the following. The infimum of openings is generally not idempotent. But consider a family $\{\gamma_i, i \in I\}$ of openings by reconstruction associated with criteria $\{\tau_i\}$. Clearly, their infimum $\gamma = \cap\gamma_i$ is still an opening, where each grain of A must fullfill all criteria τ_i to be retained. On the other hand, $\cup\gamma_i$ is the opening by reconstruction where each grain must satisfy one criterion τ_i at least. However, the largest opening is here the identity mapping, and not the largest increasing operator (*i.e.* $A \to E, \forall A \in \mathcal{P}(E)$). Hence we may state:

Proposition 6.4 *In the lattice of the increasing operators from $\mathcal{P}(E)$ into itself, the openings and the closings by reconstruction constitute two complete quasi sub-lattices.*

6.5 Leveling

Levelings have been introduced by F. Meyer, in [49], as gray tone connected operators on digital spaces, for the usual digital arcwise connections based on neighbor pixels in square or hexagonal grids. In [42], G. Matheron proposes a generalization to an arbitrary space (hence, without *a priori* connection). In his approach, connection arrives as a final result, and is generated by an extensive dilation. Now in both cases, levelings turn out to be mainly *flat* operators, *i.e.* that treat each grey level independently of the others. This circumstance suggests to try and generalize F. Meyer's approach by focusing on *set* levelings, but re-interpreted in the framework of an arbitrary connection $\mathcal{C}$. J. Serra entered this way of thinking [72], which allowed him to obtain the key theorem 6.11

Independently of these approaches, H.Heijmans has introduced and studied the class of "grain operators" in[25]. Levelings, in the sense of definition 6.9 below, are particular grain operators. However, the "good" properties of these grain operators appear when they derive from *markers based* openings and closings. So we will restrict ourselves to such criteria (for example, we will not accept or reject a particle according to its area).

6.5.1 Adjacency

The notion of adjacency [72], which governs the structure of the levelings below, is defined as follows

Definition 6.5 *Let $\mathcal{C}$ be a connection on $\mathcal{P}(E)$, and let $X, Y \in \mathcal{C}$. Sets X and Y are said to be* adjacent *when $X \cup Y$ is connected, whereas X and Y are disjoint.*

Definition 6.6 *Given two grains $A, M \in \mathcal{C}$, one says that A* touches *M, and one writes $A \parallel M$ when either $A \cap M \neq \varnothing$, or A and M are adjacent. By duality, one says that A* lies in *M when A does not touch M^c; one writes $A \subseteqq M$.*

The duality under complement provides the two following equivalences

$$A \parallel M \Longleftrightarrow A \nsubseteqq M^c \text{ and } A \nparallel M \Longleftrightarrow A \subseteqq M^c \tag{6.4}$$

Note that relation $A \parallel M$ (A touches M) is less demanding than $A \cap M \neq \varnothing$ (A meets M), since it accepts in addition that A and M be adjacent. Similarly, $A \subseteqq M$ (A lies in M), is more severe than $A \subseteq M$, since none of the grains of A and of M must be adjacent to each other.

When $\gamma_x(A) \neq \gamma_y(A)$ for an arbitrary $A \in \mathcal{P}(E)$ one cannot have $\gamma_x(A) \parallel \gamma_y(A)$ since $\gamma_x(A)$ is the largest element of $\mathcal{C}$ included in A. But $\gamma_x(A)$ may not touch some pores Y_i of A and, nevertheless, touch their union $\cup Y_i$ For example, for the "open" connection *iii/* of section 8.2.1, none of the six point pores of the central gulf, in fig.6.9a, is adjacent to the set, whereas their union touches it. The most powerfull connections are those which prevent this perverse effect, i.e. which fullfill the following condition of adjacency prevention:

Definition 6.7 *A connection is* adjacency preventing *when for all elements $M \in \mathcal{C}$ and all families $\{B_i, i \in I\}$ in $\mathcal{C}$, if M is not adjacent to each B_i, then it is not adjacent to their union $\cup B_i$.*

In particular, adjacency prevention governs the strenght of the filters by reconstruction, as proved in proposition 6.8 and in theorem 6.11.

Proposition 6.8 *Let $\mathcal{C}$ be an adjacency preventing connection on $\mathcal{P}(E)$. If γ and φ stand for an opening and a closing by reconstruction based on connection $\mathcal{C}$, then both alternated filters $\gamma\varphi$ and $\varphi\gamma$ are strong.*

proof: We shall prove the proposition for $\varphi\gamma$. We proved in sect.6.2 that $\varphi\gamma$ is an $\wedge$-filter; we have only to show that it is a $\vee$-filter, i.e. that for all $A \in P(E)$, if $x \in E$ is an arbitrary point, then $x \notin \varphi\gamma(A)$ implies $x \notin B = \varphi\gamma[A \cup \varphi\gamma(A)]$. Suppose first that $x \notin A$. Opening γ can only enlarge pore $\gamma_x(A^c)$, and closing φ keep it unchanged (if not, we would not have $x \notin \varphi\gamma(A)$). Hence $\gamma_x(A^c)$ is equal to $\gamma_x[A \cup \varphi\gamma(A)]^c$, and finally $x \notin B$. Suppose now that $x \in A$. The grain $\gamma_x(A)$ touches none of the grains and the pores of A that compose $\varphi\gamma(A)$ (if not, $\gamma_x(A)$ would belong to $\varphi\gamma(A)$, now $x \notin \varphi\gamma(A)$). Then, according to the assumption of the proposition, $\gamma_x(A)$ does not touch $\varphi\gamma(A)$, neither $[A \cup \varphi\gamma(A)] \setminus \gamma_x(A)$, hence $\gamma_x(A) = \gamma_x[A \cup \varphi\gamma(A)]$ and finally $x \notin B$, which achieves the proof. □

Proposition6.8 implies that $\varphi\gamma$ admits a decomposition as $\gamma'\varphi'$, but for a $\gamma'$and a φ' priori *different* from γ and φ. We will now see under which condition these primitives can be the same.

6.5.2 Set Levelings

From now on, we denote by $\gamma_M(A)$ the union of all grains of set A that touch an arbitrary set M, called *marker:*

$$\gamma_M(A) = \cup \{\gamma_x(A)\,,\, x \in E\,,\, \gamma_x(A) \parallel M\}$$

Similarly, the complement of closing $\varphi_{N^c}(A)$ is the union of those pores of A that hit marker N^c,

$$[\varphi_{N^c}(A)]^c = \cup \{\gamma_x(A^c)\,,\, x \in E\,,\, \gamma_x(A^c) \parallel N^c\}\,; \tag{6.5}$$

hence

$$A^c \cap \varphi_{N^c}(A) = \cup\{\gamma_x(A^c)\,,\, x \in E\,,\, \gamma_x(A^c) \subseteqq N\} \tag{6.6}$$

is the union of those pores of A lying in marker N.

Definition 6.9 *Let E be an arbitrary set, and $\mathcal{C}$ be a connection on $\mathcal{P}(E)$. Let γ_M and φ_{N^c} an opening and a closing, both by marker reconstruction, from $\mathcal{P}(E)$ into itself. The leveling $\lambda : \mathcal{P}(E) \to \mathcal{P}(E)$, of primitives γ_M and φ_{N^c} is then defined by the relation*

$$\lambda = \gamma_M \cup \left(\complement \cap \varphi_{N^c}\right) = \varphi_{N^c} \cap \left(\complement \cup \gamma_M\right) \tag{6.7}$$

where $\complement$ stands for the complement operation on $\mathcal{P}(E)$.

When applied to set A, leveling λ yields the two equalities

$$\begin{aligned} A \cap \lambda(A) &= A \cap \gamma_M(A) \\ A^c \cap \lambda(A) &= A^c \cap \varphi_{N^c}(A) \quad (\Longleftrightarrow A \cup \lambda(A) = A \cup \varphi_{N^c}(A)) \end{aligned} \tag{6.8}$$

so that $\lambda(A)$ acts inside A as opening γ_M, and inside A^c as closing φ_{N^c}. System (6.8) also relates to the *activity lattice*, where a mapping ψ on $\mathcal{P}(E)$ is said to be less active than another, ψ', when $\psi'(A)$ modifies more points of A than $\psi(A)$ does, $\forall A \in \mathcal{P}(E)$, (ch.8 in [62]). If Id stands for the identity operator, the activity ordering is as follows

$$\begin{aligned} Id \cap \psi &\supseteq Id \cap \psi' \\ Id \cup \psi &\subseteq Id \cup \psi' \end{aligned}$$

and one notes $\psi \preceq \psi'$. A complete lattice is associated with this ordering, where the supremum and the infimum of a family $\{\psi_i\,,\, i \in I\}$ are given by

$$\begin{aligned} \curlyvee \psi_i &= \left[\complement \cap (\cup \psi_i)\right] \cup [\cap \psi_i] \\ \curlywedge \psi_i &= [Id \cap (\cup \psi_i)] \cup [\cap \psi_i]\ . \end{aligned}$$

When applying this system to the family $\{\gamma_M, \varphi_{N^c}\}$ of the two leveling primitives, we draw from (6.7) that

$$\gamma \curlyvee \varphi = \lambda \qquad \gamma \curlywedge \varphi = Id\,.$$

Conversly, the relation $\gamma \curlyvee \varphi = \lambda$ yields equation (6.7), hence may be considered as an alternative definition for leveling.

An operation whose definition involves the complement $\complement$ risks not to be increasing. But in the present case, we will now see that the condition under which λ is increasing makes it also a strong filter, which means much more.

Lemma 6.10 *Let $\mathcal{C}$ be an adjacency preventing connection on $\mathcal{P}(E)$, let $A, N \in \mathcal{P}(E)$, and let Y be a pore of A. If Y lies in N, then all grains of A which are adjacent to Y meet N. By duality, if a grain X of A does not touch set N, then none of the pores of A adjacent to X is included in N .*

proof: Consider a pore Y of A, with $Y \subseteqq N$, and a grain X of A which is adjacent to Y. Since there exists a point $x \in X$ that is adjacent to Y, and since $Y \not\parallel N^c$ (Eq (6.4)), x belongs necessarily to N; hence $X \cap N \neq \emptyset$.

Take now a grain X of A that does not touch N, i.e. such that $X \subseteqq N^c$. We draw from the first part of the proof that every pore Y of A that is adjacent to X meets N^c, hence is not included in N. $\square$

Theorem 6.11 *Let $\mathcal{C}$ be an adjacency preventing connection on $\mathcal{P}(E)$. Given $M, N \subseteq E$ with $N \subseteq M$, the leveling $\lambda_{M,N} : \mathcal{P}(E) \to \mathcal{P}(E)$ of primitives γ_M and φ_{N^c} is a* strong *connected filter, and admits the double decomposition*

$$\lambda = \gamma_M \ \varphi_{N^c} = \varphi_{N^c}\gamma_M \ .$$

proof: We have to prove that the three following operations are identical:

i/ to take the union of the pores of A lying in N and of the grains of A touching M;

ii/ to take the union A'of the grains of A touching M, and to add it to the pores of A' that lie in N;

iii/ to add to A all its pores lying in N, and to extract from the result the union of all grains touching M.

Indeed, when $N \subseteq M$, the lemma states that all grains of A adjacent to a pore $Y \subseteqq N$ hit N, hence hit also M. On the other hand, a grain $\gamma_x(A)$ of A which is not adjacent to various grains X and pores Y of A, with $Y \subseteqq N$, are neither adjacent to the union of these X and Y (assumption of adjacency prevention), so that the two processings *i/* and *iii/* are identical. The proof is achieved by observing that *i/* is a self-dual procedure, and that *ii/* and *iii/* are dual of each other. □

Remark that, when $N \subseteq M$, the supremum of the two logical conditions $A \not\Uparrow M$ and $A \subseteqq N^c$ is the certainty. Then, according to proposition 8.5 in [25], we find again the increasingness of λ.For extending levelings from sets to numerical functions, we need to consider them as functions of their three arguments A, M and N. Now, is the mapping $\lambda(A, M, N)$ from $[\mathcal{P}(E)]^3$ into $\mathcal{P}(E)$ increasing ?

Corollary 6.2 *The leveling $\lambda : [\mathcal{P}(E)]^3 \to \mathcal{P}(E)$ is increasing if and only if the two operands M and N are ordered by $N \subseteq M$*

proof: We draw from the theorem that, given M and N, with $N \subseteq M$,

$$A \subseteq A' \Longrightarrow \lambda(A, M, N) \subseteq \lambda(A', M, N).$$

On the other hand, given A', when $M \subseteq M'$ and $N \subseteq N'$ more grains of A' are touched and more pores of A' are lying, hence

$$\lambda(A', M, N) \subseteq \lambda(A', M', N')$$

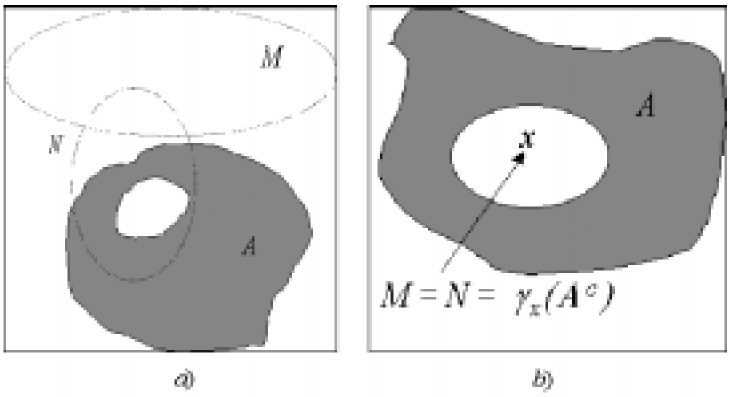


Figure 6.8: a)Non increasingness of λ when $N \nsubseteq M$.Take for A' grain A plus its pore; then A$\subseteq A'$ whereas $\lambda(A') \subseteq \lambda(A)$. b) Take the internal pore of grain A as M and N, then $\lambda(A)$ equals the pore without the grain (flip-flop effect)

which achieves the proof (the only if part is given by the counter example of Fig.6.8a. □

An interesting feature of levelings concerns their possible self-duality. Firstly, we may consider the behaviour, under complement, of the triple mapping $(A, M, N) \rightarrow \lambda(A, M, N)$. We have

$$[\lambda(A^c, M^c, N^c)]^c = [\gamma_{M^c}(A^c)]^c \cap [A \cap [\gamma_N(A)]^c]^c = \varphi_{M^c}(A) \cap [A^c \cup \gamma_N(A)],$$

$$\text{hence } [\lambda(A^c, M^c, N^c)]^c = \gamma_N(A) \cup [A^c \cap \varphi_{M^c}(A)] = \lambda(A, N, M)$$

Therefore self-duality of $\lambda(A, M, N)$ is reached when and only when the two markers N and M are identical. Since, in addition, condition $M \equiv N$ implies the inreasingness of λ, we may state

Proposition 6.12 *The leveling $(A, M, M) \rightarrow \lambda(A, M, M)$ is an increasing self-dual mapping from $\mathcal{P}(E) \times \mathcal{P}(E)$ into $\mathcal{P}(E)$.*

In this approach, we implicitely supposed that the data of A and of M are independent. In practice, it often occurs that marker M derives from a previous tranformation of A itself, $M = \mu(A)$, say. Then the proposition shows that the leveling $\lambda : \mathcal{P}(E) \to \mathcal{P}(E)$, with $\lambda = \lambda(A, \mu(A), \mu(A))$ is self-dual if and only if *mapping μ itself* is already self-dual.

We conclude this section by exhibiting two examples showing how necessary are some assumptions above. Take for A a single grain with an internal pore, and for $M \equiv N$ the set made by the pore of A in fig.6.8b. Suppose we replace, in definition 6.9, the condition $\gamma_x(A) \parallel M$ by $\gamma_x(A) \cap M = \varnothing$, and $\gamma_x(A^c) \subseteqq N$ by $\gamma_x(A^c) \subseteq N$ Clearly, we have

$$\begin{array}{lll} & \varphi_{M^c}(A) = A \cup M & \Longrightarrow \quad \gamma_M \varphi_{M^c}(A) = A \cup M, \\ \text{but} & \gamma_M(A) = \varnothing & \Longrightarrow \quad \varphi_{M^c}\gamma_M(A) = \varnothing \end{array}$$

whereas $\lambda(A) = M$ is neither $\gamma_M \varphi_{M^c}(A)$ nor $\varphi_{M^c}\gamma_M(A)$. Moreover, the example shows that $\lambda(A \cap \lambda(A)) = \varnothing$ and that $\lambda(A \cup \lambda(A)) = A \cup M$; this implies that λ cannot be decomposed into the product of an opening by a closing or *vice versa* (theorem 6-11, corollary 2 in[40]). Notice also, finally, that in the example of fig.6.8b the border between the grain and its internal pore is preserved, but not the sense of variation. As a matter of fact, such a "flip-flop" effect is due to the case when M contains a pore of A, but misses the surrounding grain(s). It *cannot appear* in the actual levelings of definition 6.9.

The second counter-example concerns adjacency prevention. Let us adopt the "open " connection, and take for A the set 6.9a, and for marker $M = N$ the six point pores of the central gulf.Figures6.9b and 6.9c show the two transforms $\varphi_{M^c}\gamma_M(A)$ and $\gamma_M \ \varphi_{M^c}(A)$ which are obviously different : one cannot drop the adjacency prevention, in theorem 6.11 !

6.5.3 Levelings as function of their markers

For the sake of simplicity, we shall take $M = N$ through this section, although self-duality is not really required here, and write $\lambda_A(M)$ for $\lambda(A, M, M)$.

Theorem 6.13 *Let $\mathcal{C}$ be an adjacency preventing connection on $\mathcal{P}(E)$. Given $A \subseteq E$, the mapping $\lambda_A : \mathcal{P}(E) \to \mathcal{P}(E)$ is a morphological filter from $\mathcal{P}(E)$ into itself.*

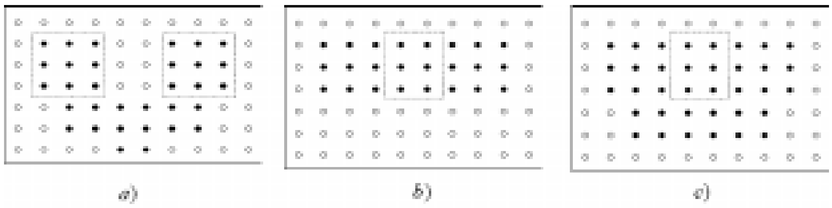


Figure 6.9: a) initial set A. For the "open" connection by the 3x3 square, A is made of 16 grains, namely the two $3x3$ squares plus 14 isolated points ; b) $\varphi_{M^c}\gamma_M(A)$ for M equal to the six pores of the central gulf (surrounded); c) $\gamma_M\ \varphi_{M^c}(A)$ for the same marker.The difference comes from that the "open" connection is not *adjacency preventing.*

proof: For $A, M \in \mathcal{P}(E)$, A given and M variable, $\lambda_A(M)$ is the union of some grains and some pores of A, in such a way that each accepted pore arrived in $\lambda_A(M)$ accompanied by the whole collection of its adjacent grains. So a grains of $\gamma_x(A)$ that does not participate to $\lambda_A(M)$ does not touch any of the A-connected elements (grains or pores) involved in $\lambda_A(M)$; hence, by adjacency prevention, $\gamma_x(A) \not\parallel \lambda_A(M)$. By duality, $\gamma_x(A^c) \nsubseteq M$ implies $\gamma_x(A^c) \nsubseteq \lambda_A(M)$, so that $\lambda_A[\lambda_A(M)] = \lambda_A(M)$. □

The relevant formalism to go further is that of the activity ordering for sets (and no longer for set mappings)[42]. As a matter of fact, any fixed set A generates an ordering denoted by $\preceq_A$, from the two relationships

$$M_1, M_2 \subseteq E \quad \begin{array}{c} M_1 \cap A \supseteq M_2 \cap A \\ \\ M_1 \cap A^c \subseteq M_2 \cap A^c \end{array} \quad \Leftrightarrow \quad M_1 \preceq_A M_2$$

From this ordering derives the so called A-*activity lattice*, where the supremum and the infimum of a family $\{M_i, i \in I\}$ of sets are given by

$$\begin{array}{rcl} \curlyvee_A M_i & = & [A^c \cap (\cup M_i)] \cup [\cap M_i] \\ \curlywedge_A M_i & = & [A \cap (\cup M_i)] \cup [\cap M_i] \end{array}$$

with A itself as the minimum element, and A^c as the maximum one (a system very similar to that presented above about the activity lattice for operators). In this framework, the following theorem holds[42][72]

Theorem 6.14 *Given set A, the leveling $M \to \lambda_A(M)$ from the A-activity lattice of $\mathcal{P}(E)$ into itself is an openings. Moreover, for all $A, M_1, M_2 \in \mathcal{P}(E)$, we have*

$$M_1 \preceq_A M_2 \quad \Rightarrow \lambda_{M_1}\lambda_{M_2}(A) = \lambda_{M_2}\lambda_{M_1}(A) = \lambda_{M_2}(A)$$

This last granulometric type pyramid is specially usefull in practice, for it allows to grade the activity effects of markers: it means that we can directly implement a highly active marker, or, equivalently, reach it by intermediary steps. An example is given in fig.6.10.

6.5.4 Function levelings

Let $\mathcal{T}$ be a discrete axis; denote by $\mathcal{T}^E$ the lattice of all numerical functions $f : E \longrightarrow \mathcal{T}$. An increasing operator Ψ on $\mathcal{T}^E$ is said to be *flat (see sect. 2.5.4.)* if there exists an increasing set operator ψ such that

$$X\left[\Psi(f), t\right] = \psi\left[X(f), t\right] \tag{6.9}$$

where stands for the thresholding of function f at level t, i.e. :

$$X(f, t) = \{x : x \in E, \quad f(x) \geq t\} \tag{6.10}$$

In the discrete cases of digital imagery, relation (6.9) is sufficient to characterize the function operator Ψ associated with an increasing set operator ψ.

Definition 6.15 *Let f, g, h, be three functions from E into T,with $g \leq h$.Then the relation*

$$X\left[\Lambda(f), t\right] = \lambda\left[X(f,t)\right], X(g,t), X(h,t)]$$

defines one and only one leveling $\Lambda(f)$ on $\mathcal{T}^E$.

When connection $\mathcal{C}$ is obtained from the iterations of an elementary dilation δ, of adjoint erosion ε, then a digital algorithm for $\Lambda(f)$ from the data of f, g and h derives from the decomposition theorem 6.11, by computing successively the opening by reconstruction $g_\infty(f)$ and then $\Lambda(f) = h_\infty[g_\infty(f)]$. The first operation is thus given by the limit of the sequence

$$\begin{aligned} & g_n = (f \wedge \delta g_{n-1}) \\ \text{with} \quad & g_1 = (f \wedge \delta g) \end{aligned}$$

and the second one by

$$\begin{aligned} & h_n = [g_\infty(f) \vee \varepsilon h_{n-1}] \\ \text{with} \quad & h_1 = [g_\infty(f) \vee \varepsilon h] \end{aligned}$$

All theorems and propositions of the binary case extend directly to numerical one. Concerning self-duality for example, if 0 and m stand for the two extreme bounds of the gray axis $\mathcal{T}$, we have

$$m - \Lambda(m - f, m - g, m - g) = \Lambda(f, g, g)$$

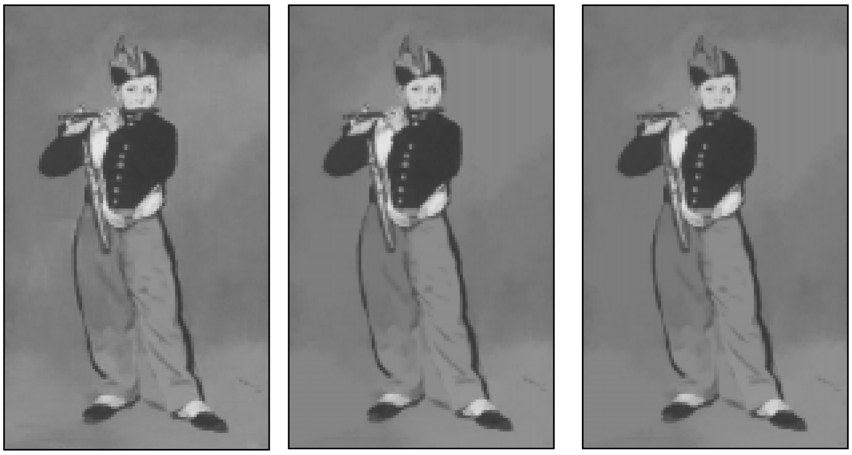
which means that the leveling $f, g \to \Lambda(f, g)$ is always a self-dual mapping. In addition, when one takes for marker g a self-dual mapping (*e.g.* convolution, median operator, etc..), then the leveling Λ, considered as a function of f only, becomes in turn self-dual, and we have

$$m - g(m-f) = g(f) \Rightarrow m - \Lambda[m-f, g(m-f), g(m-f)] = \Lambda[f, g(f), g(f))]$$

In practice, the role of the marker is crucial. In Fig.6.10, the marker is obtained by replacing f by zero on the extended maxima and minima of f, and by leaving f unchanged elsewhere (*extended maxima of* f : do the opening by reconstruction $\gamma_{\text{rec}}(f)$ of f from $f - k$, where k is a positive constant. Then the maxima of $\gamma_{\text{rec}}(f)$ define the so called *extended maxima* of f, and those points x where $f(x) - \gamma_{\text{rec}}(f)(x) = k$ define the (non extended) maxima of f of dynamics $\geqslant k$; the extended minima are obtained by duality).The corresponding levelings are shown in Fig.6.10a and Fig.6.10b, for markers g_{30} and g_{60}, of dynamics 30 and 60 respectively (over 256 gray levels).

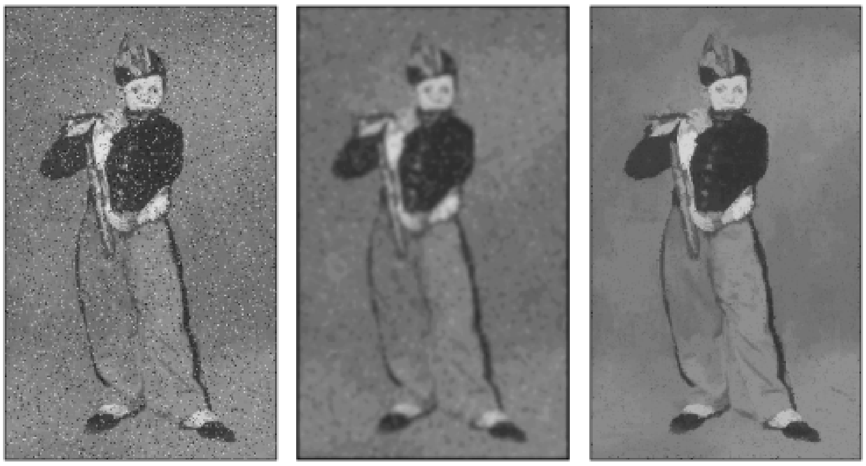
These two markers are self-dual by construction, and satisfy the condition of activity increasingness of theorem 6.14. Their progressive leveling action appears clearly when confronting Fig.6.10a and Fig.6.10b. Notice the relatively correct preservation of some fine details such as buttons, eyes, eyebrows, fingers, etc.. These details are preserved because of their high dynamics.

In Fig.(6.11), the leveling is used for noise reduction, from a marker obtained by Gaussian moving average of size 5, namely Fig.(6.11)b, of the initial



a) *b*) *c*)

Figure 6.10: a) Manet' Joueur de fifre b)and c) levelings of a) by extended extrema of dynamics 30 (b) and 60 (c).



a) *b*) *c*)

Figure 6.11: a) noisy version, b) gaussian convolution of a, c) leveling of a) by marker b)

noisy image of Fig.(6.11)a. It results in Fig.(6.11)c where the noise reduction of Fig.(6.11)b is preserved, but where the initial sharpness of the edges is recovered.

A last word. There are two ways for developing a theory in discrete geometry. One can start from some Euclidean notions and adapt them to discrete spaces, or elaborate the whole approach independently of the fact that it may apply to a continous, or a discrete, or a finite, space E. It is this second that was chosen here.

Chapter 7

2-D Geodesy and Segmentation

7.1 Introduction

Almost all the morphological transformations presented in the previous chapters were *increasing* ones. The operations that we shall introduce in this chapter will not necessarily share this characteristic. Indeed, we are no longer concerned with morphological filters and would like to show how morphology can be successfully applied to *segmentation* problems. Let us first say a couple of words on the meaning that we shall attach to the word "segmentation": in our sense, segmenting an image consists in *extracting* its different objects or regions and *contouring* them as precisely as possible. As concerns binary images, a very common segmentation problem is to separate its overlapping objects. Similarly, segmenting a grey-tone image comes most of the time down to dividing it into different regions (generally, one of these regions stands for the *background,* whereas the others correspond to the *objects*). In this chapter, starting from the concept of *marker,* we shall derive a general (and morphological) approach of segmentation problems.

To reach this goal, i.e. to really segment objects from their markers, we will firstly describe the so-called *geodesic operations*, which differ from the Euclidean ones (i.e. the classical operations with hexagons, for instance) in that the underlying space is no longer the whole space, but a given subset X of this space. The last section is devoted to the *watershed* mapping. Defined for grey-tone images, it constitutes the basic morphological tool for segmentation.

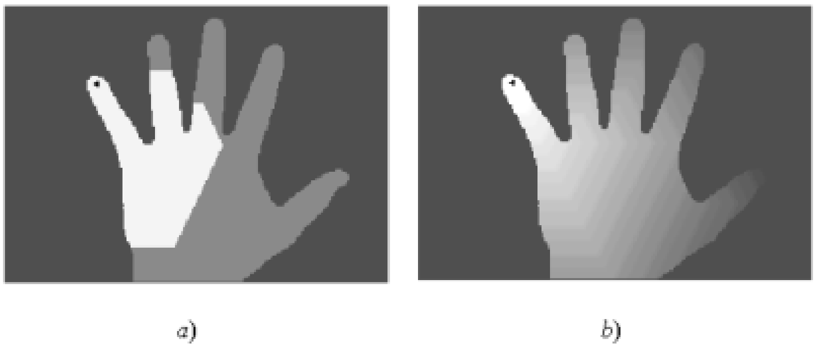


Figure 7.1: a) Geodesic disc ; b) Geodesic distance function in two dimensions

7.2 Geodesic transformations

7.2.1 Choquet's theorem

When a stone is thrown into a lake and generates a disturbance, a wave string is being created and spreads out while going around the possible obstacles, until the most remote points from the middle. The wavefront, circular in the case of a lack of borders, laps the islands and the lake contours and finally covers them completly (see fig.(7.1)).

In order to extract connected objects selected by markers, F. Meyer[43] and J.C. Klein [30] were the first ones to transfer these notions to the mathematical morphology, and the very first formalization, named "geodesic metrics" was established by C. Lantuejoul and S. Beucher [33]. Indeed, in figure 3, the zone of the reference set Z, swept between instants 0 et λ by the wavefront born from point x at the original instant is a disk $\mathrm{B}_\lambda(x)$, smaller than the Euclidean disk with a radius λ and completly containd in Z. When the reference set Z is compact, the induced metrics $\{B_\lambda(x), x \in Z\}$ satisfy the following theorem, from G. Choquet ([10],theorem 11-6)

Theorem 7.1 *Let E be a metric compact space and let A et B be two disjoint closed subsets of E. If there exist rectifiable curves with extremities in A and B respectively, and if λ stands for the lower limit of their lengthes, then there exists a simple arc whose length is λ and whose extremities lie in A and B* *respectively.*

In what follows, we will always suppose that references sets Z are compact, and that for any points x, y selected in a same connected component Z, there is an rectifiable path with a length limited by a $\lambda \max(Z, x)$ and linking these two points. This happens, particularly, when in $\mathbb{R}^n$, the set Z is the topological closure of a bounded open set. Rectifiable arcs, as a precaution, are meant to exclude compact sets such as, for instance, a spiral which winds indefinitly around a circle. We now present geodesic metrics for two dimensions in a digital approach, but we shall treat the 3-D case, in chapter 9, in the Euclidean version.This will allow the reader to compare both styles.

7.2.2 Geodesic distance

Let X be a set in $\mathbb{Z}^2$. We define the geodesic distance between two points p_1 and p_2 of X as the infimum of the length of the paths between p_1 and p_2 in X (if there are such paths at all):

$$d_X(p_1, p_2) = \inf\{l(C_{p_1,p_2}), C_{p_1,p_2} \text{ path between } p_1 \text{ and } p_2 \text{ included in } X\}. \tag{7.1}$$

Remark that this distance is a generalized one, in the sense that we put conventionally $d_X(p_1, p_2) = +\infty$ when p_1 and p_2 are in different connected components of X. The definition of d_X is illustrated in Fig. 7.2

We call *geodesic ball* of radius $n \in \mathbb{Z}^+$ and of center $p \in X$ the set $B_X(p, n)$ defined by:

$$B_X(p, n) = \{p' \in X, d_X(p', p) \leq n\}. \tag{7.2}$$

7.2.3 Geodesic dilations and erosions

Suppose now that X is equipped with its associated geodesic distance d_X. Given a $n \in \mathbb{Z}^+$, we consider the structuring function which associates with each pixel $p \in X$ the geodesic ball $B_X(p, n)$ of radius n centered in p. This allows us to define the *geodesic dilation* of a subset Y of X in the following way:

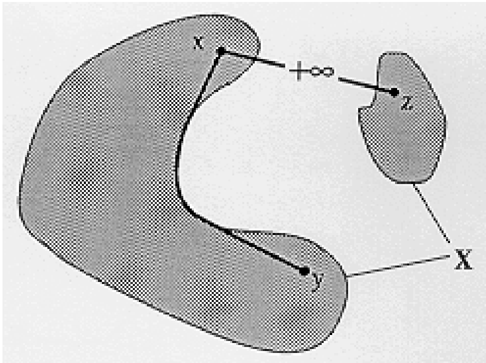


Figure 7.2: Geodesic distance in a set X.

Definition 7.2 *The geodesic dilation $\delta_X^{(n)}(Y)$ of size n of set Y inside set X is given by*

$$\delta_X^{(n)}(Y) = \bigcup_{p \in Y} B_X(p,n) = \{p' \in Y, \exists p \in Y, d_X(p',p) \leq n\}. \tag{7.3}$$

The dual formulation of the geodesic erosion of size n of Y inside X is the following:

$$\varepsilon_X^{(n)}(Y) = \{p \in Y, B_X(p,n) \subseteq Y\} = \{p \in Y, \forall p' \in X/Y, d_X(p,p') > n\}. \tag{7.4}$$

Examples of geodesic dilation and erosion are shown in Fig. 7.3.

As already remarked above, the result of a geodesic operation on a set $Y \subseteq X$ is always included in X, which is our new workspace. As far as implementation is concerned, an elementary geodesic dilation (of size 1) of a set Y inside X is obtained, in the hexagonal case, by intersecting the result of a (Euclidean) dilation of size 1 of Y with the workspace X:

$$\delta_X^{(1)}(Y) = (Y \oplus H) \cap X. \tag{7.5}$$

A geodesic dilation of size n is obtained by iterating n elementary geodesic dilations:

$$\delta_X^{(n)}(Y) = \underbrace{\delta_X^{(1)}(\delta_X^{(1)}(\ldots\delta_X^{(1)}(}_{n \text{ times}}Y))). \tag{7.6}$$

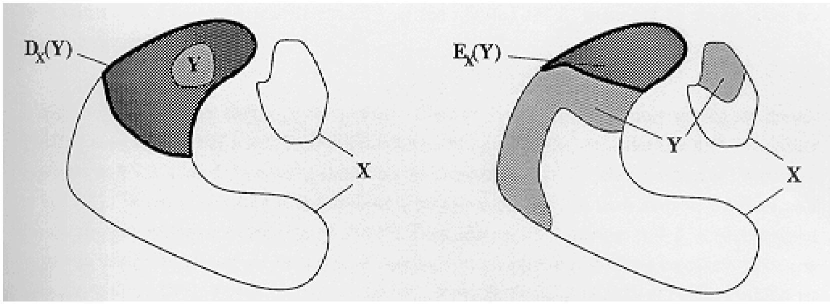


Figure 7.3: Examples of a geodesic dilation and of a geodesic erosion of set Y inside set X.

(By duality, geodesic erosions are easily determined).

7.2.4 Reconstruction and Geodesic SKIZ

One can remark that by performing successive geodesic dilations of a set Y inside a set X, it is impossible to intersect a connected component of X which did not initially contain a connected component of Y. Moreover, in this successive geodesic dilations process, we progressively "reconstruct" the connected components of X that were initially *marked* by Y. This is shown in figure 7.4.

Now, the sets with which we are concerned are finite ones. Therefore, there exists a n_0 such that

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

At step n_0, we have entirely reconstructed all the connected components of X which were initially marked by Y. This operation is naturally called *reconstruction:*

Definition 7.3 *The reconstruction $r_X(Y)$ of the (finite) set X from set Y is given by the following formula:*

$$r_X(Y) = \lim_{n\to+\infty} \delta_X^{(n)}(Y). \tag{7.7}$$

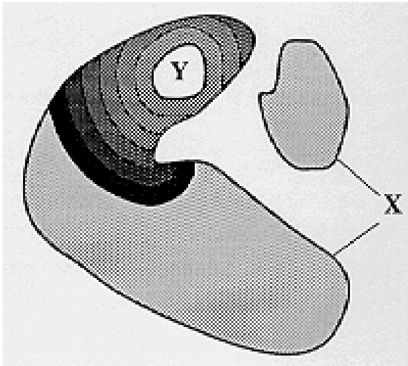


Figure 7.4: Successive geodesic dilations of set Y inside set X.

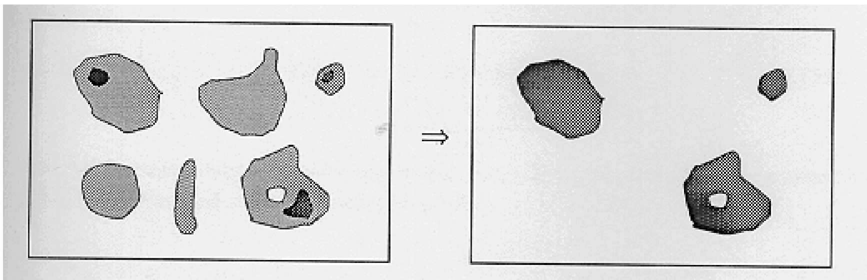


Figure 7.5: Reconstruction of X (light set) from Y (dark set).

Fig. 7.5 illustrates this transformation. It may occur, sometimes, that a given grain X contains more than one marker. In such cases, one uses to share the zones of influene of the various markers Y_i in set X by means of the *geodesic influence zones* of the connected components of set Y inside X .

Let us remind brefly the so-called (non geodesic) *SKeleton by Influence Zones* or *SKIZ.* Given a set $C \subseteq \mathbb{R}^2$ made of N compact connected components $(X_n)_{1 \leq n \leq N}$, with $N < \infty$. The influence zone $Z(X_n)$ of X_n is the locus of those points X^c which are closer to X_n than to any other connected component of X:

$$Z(X_n) = \{p \in \mathbb{R}^2, \forall j \neq n, d(p, X_n) \leq d(p, X_j)\}. \tag{7.8}$$

The major interest of the SKIZ, versus the calssical skeleton lies in a property of robustness, proved by Ch. Lantuejoul [32], and which is stated as follows:

Theorem 7.4 *Let $X = \cup\{K_n, 1 \leq n \leq N < \infty\}$ be a finite union of N compact sets K_n. Let $\{X_i\}$ be a sequence $\mathcal{F}(\mathbb{R}^2)$, where each element X_i is the union of N disjoint compact terms $K_{n,i}$, and $K_{n,i}$ converges towards a K_n (i.e. $K_{n,i} \to K_n$ and $n \neq p \Rightarrow K_n \cap K_p = \emptyset$) Then the SKIZ Sz:$\mathcal{F}(\mathbb{R}^2) \to$:$\mathcal{F}(\mathbb{R}^2)$ is a continuous mapping*

$$X_i \to X \quad \Rightarrow \quad Sz(X_i) \to Sz(X)$$

In more intuitive terms, the skeleton is parasited by all its barbs. By clipping them completely one obtains a continuous operation.This Euclidean relation Eq(7.8) remains valid in the digital planec $\mathbb{Z}^2$.The distance used then is a discrete one, defined on the square or on the hexagonal grid. Most of the time, we use the hexagonal distance d_H: $d_H(p_1, p_2) = n$ if and only if the length of the shortest paths between p_1 and p_2, whose edges are included in the grid, is equal to n. An example of SKIZ is presented in Fig. 7.6. One can show that the skeleton by influence zones is a subset of the skeleton $S(X^C)$ of the *background* of X, i.e. of X^C. In practice, it is often determined by removing the irrelevant edges of $S(X^C)$, which are called *parasitic barbs.* The SKIZ will be very useful for the binary segmentation algorithm presented in section 3 of this chapter.

Indeed, the notions of influence zones and of SKIZ extend directly to the geodesic case: it suffices to use a geodesic distance in Eq.(7.8). We then obtain partitions of the type shown in Fig.(7.7).

7.2.5 Geodesic operations for grey-tone images

At present, all the tools required for solving our binary segmentation problem are available. However, in the grey-tone case, a few more tools will be necessary, namely the geodesic decimal erosions and dilations. Therefore, it seems rather convenient to present these operations just after the corresponding binary ones.

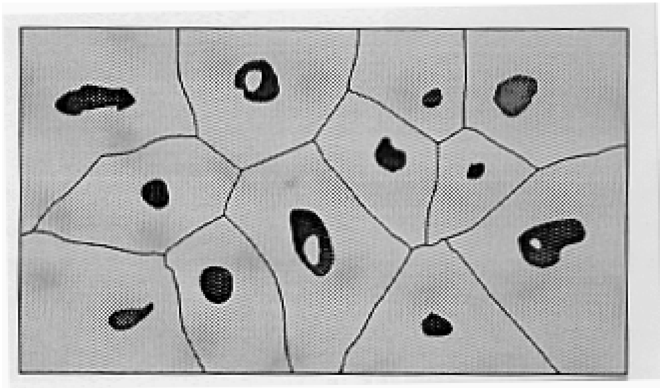


Figure 7.6: Example of skeleton by influence zones.

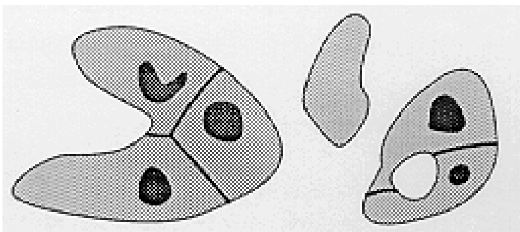


Figure 7.7: Example of geodesic SKIZ.

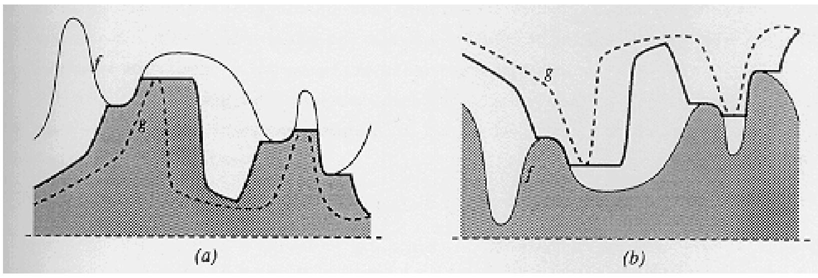


Figure 7.8: Examples of grey-tone geodesic dilation (a) and erosion (b).

The digital formulas (7.5) and (7.6) can easily be extended to the case of grey-tone functions defined on a digital grid. Given two functions f and g in $\mathcal{F}(\mathbb{Z}^2, \mathbb{Z})$, such that $f \leq g$, the geodesic dilations of size 1 and of size $n > 1$ of g with respect to f are defined by:

$$\begin{aligned} \delta_f^{(1)}(g) &= (g \oplus H) \wedge f, \\ \delta_f^{(n)}(g) &= \underbrace{\delta_f^{(1)}(\delta_f^{(1)}(\dots \delta_f^{(1)}(}_{n \text{ times}} g))). \end{aligned} \tag{7.9}$$

By duality, when $g \geq f$, geodesic decimal erosions of g with respect to f are defined in the following way:

$$\varepsilon_f^{(1)}(g) = (g \ominus H) \vee f, \tag{7.10}$$

$$\varepsilon_f^{(n)}(g) = \underbrace{\delta_f^{(1)}(\delta_f^{(1)}(\dots \varepsilon_f^{(1)}(}_{n \text{ times}} g))). \tag{7.11}$$

These two transformations are presented in Fig. 7.8. They will turn out to be extremely useful when associated with *watersheds* (see § 7.4).

Lastly, the concept of reconstruction is easily extended to the decimal case. Given two functions f and g in $\mathcal{F}(\mathbb{Z}^2, \mathbb{Z})$ such that $f \leq g$, the reconstruction of f from g is defined by:

$$r_f(g) = \lim_{n \to +\infty} \delta_f^{(n)}(g). \tag{7.12}$$

This transformation, which is shown in Fig. 7.9 is constantly useful in mathematical morphology. Among other things, it is frequently used in image

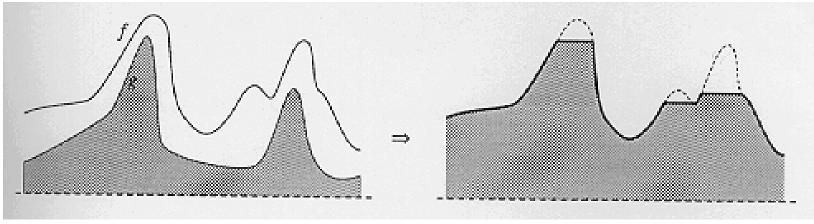


Figure 7.9: Example of decimal reconstruction.

filtering. Indeed, it is very easy to prove that the transformation ψ_n defined by

$$\psi_n(f) = r_f(f \ominus H)$$

is a morphological filter. Moreover, by subtracting $\psi_n(f)$ from the initial function f, we get a powerful tool for extracting the light and thin areas of this grey-tone image.

7.3 Binary segmentation

We will now make use the tools that we have presented at the beginning of this chapter for designing a powerful binary segmentation algorithm. Starting from the markers of our objects, i.e. from the ultimate erosion, our goal is to contour finely these objects. We could consider using the geodesic SKIZ, and defining each object as the geodesic influence zone of its marker inside the initial set. Unfortunately, this is not a satisfactory algorithm. Indeed, as shown in Fig. 7.10, the separating lines thus defined between objects are poorly located. This is due to the fact that we did not take the *altitude* of the markers—i.e. the value that is associated with them by the quench function—into account.

The way for designing a good segmentation procedure—in taking the above altitudes into account—is to use successive geodesic SKIZ. Let n_m be the size of the largest non empty erosion of X, i.e. such that

$$X \ominus n_m H \neq \emptyset \quad \text{and} \quad X \ominus (n_m + 1)H = \emptyset.$$

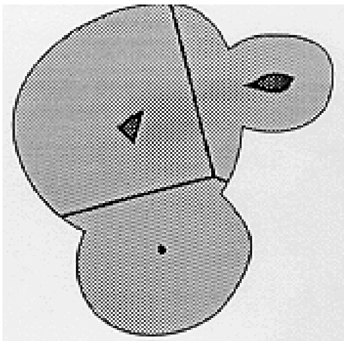


Figure 7.10: Bad segmentation algorithm (geodesic SKIZ of the ultimate erosion of X inside X).

Necessarily, $X \ominus n_m H$ is a subset of the ultimate erosion of X . Denote X_{n_m} this set. Now, consider the erosion of size $n_m - 1$ of X, i.e. $X \ominus (n_m - 1)H$. Obviously, the following inclusion relation holds:

$$X_{n_m} \subseteq X \ominus (n_m - 1)H.$$

Now, let Y be a connected component of $X \ominus (n_m - 1)H$. There are three possible inclusion relations between Y and $Y \cap X_{n_m}$:

1. $Y \cap X_{n_m} = \emptyset$: in this case, Y is another connected component of $\mathrm{Ult}(X)$.
2. $Y \cap X_{n_m} \neq \emptyset$ and is connected: here, Y is used as a new marker.
3. $Y \cap X_{n_m} \neq \emptyset$ and is not connected: in this last case, the new markers that are produced are the geodesic influence zones of $Y \cap X_{n_m}$ inside Y.

These three different cases are shown on Fig. 7.11.

Let $X_{n_m - 1}$ be the set of the markers produced after this step. To summarize what we have just said, $X_{n_m - 1}$ is made of the union of

- the geodesic influence zones of X_{n_m} inside $X \ominus (n_m - 1)H$,
- the connected components of $\mathrm{Ult}(X)$ whose *altitude* is $n_m - 1$.

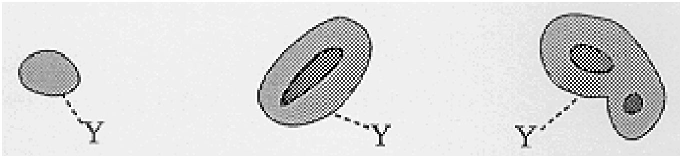


Figure 7.11: The three possible inclusion relations between Y and $Y \cap X_{n_m}$.

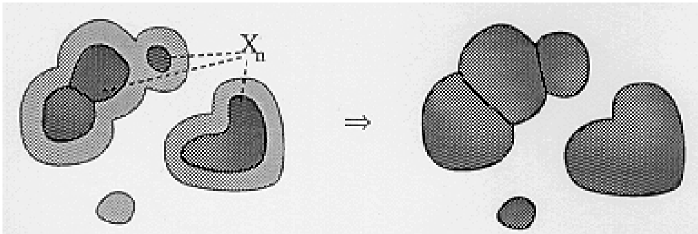


Figure 7.12: How to obtain X_{n-1} from X_n.

This procedure is then iterated at levels $n_m - 2$, $n_m - 3$,etc ... until level 0 is reached. In a more formal way, for every $n \in]0, n_m[$, let us introduce the following notations:

(i) $u_n(X)$ is the set of the connected components of $\mathrm{Ult}(X)$ having altitude n:
$$p \in u_n(X) \iff p \in S(X) \text{ and } q_X(p) = n.$$

(ii) For every set $Y \subseteq X$, $z_X(Y)$ designates the set of the geodesic influence zones of the conected components of Y inside X.

The recurrence formula between levels n and $n-1$ can now be stated:
$$X_{n-1} = z_{X \ominus (n-1)H}(X_n) \cup u_{n-1}(X). \tag{7.13}$$
It is illustrated by Figure 7.12.

The set X_0 that is finally obtained after applying this algorithm produces a good segmentation of X. Fig. 7.13 presents an example of this binary segmentation algorithm. In some cases, it is still unsufficient for obtaining a satisfactory segmentation, and other procedures, making use of an a priori knowledge on the objects to segment, or based on more elaborated notions, such as *critical balls* [6], must be designed.

7.4 Watersheds and segmentation of grey-tone images

7.4.1 Catchment basins, watersheds

At first sight, the elaboration of the algorithm presented in the preceding section seems complicated. Therefore, we now give a more intuitive approach of this procedure. Consider the function $-\mathrm{dist}_X$, where dist_X is the distance function (i.e. at each point $x \in X$, the shortest distance from x to the points of X^c) and regard it as a topographic surface. The *minima* of this topographic surface are located at the different connected components of the ultimate erosion of X. Now, if a drop falls at a point p of X, it will slide along the topographic surface until it finally reaches one of its minima. We define the *catchment basin* $W(m)$ associated with a minimum m of our topographic surface in the following way:

Definition 7.5 *The catchment basin $W(m)$ associated with a minimum m of a function regarded as a topographic surface is the locus of the points p such that a drop falling at p finally reaches m.*

This definition is not very formal, but it has the advantage of being relatively intuitive. In our example, the catchment basins of the function $-\mathrm{dist}_X$ exactly correspond to the regions that were extracted by the algorithm presented in the previous section.

Actually, this notion of catchment basin can be defined for any kind of grey-tone function and the algorithm can be easily adapted to the determination of the basins of any decimal image I: it suffices to replace the successive erosions $X \ominus nH$—which correspond to the different thresholds of the distance function of X—by the successive thresholds of I. The lines separating different basins are called *watersheds* or *dividing lines.* These notions are illustrated in Fig. 7.14.

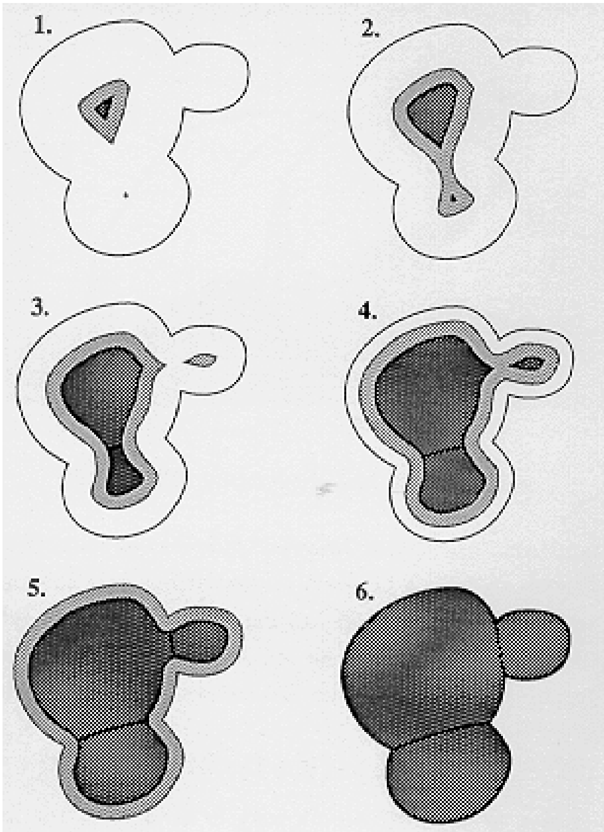


Figure 7.13: A step by step example of the correct binary segmentation algorithm.

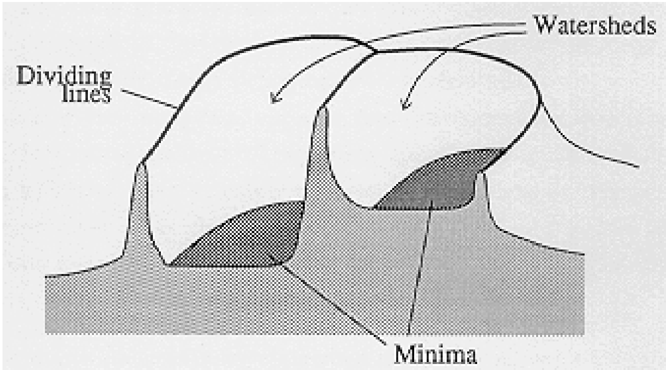


Figure 7.14: Catchment basins and watersheds.

The watersheds constitute an extremely powerful tool for segmenting grey-tone images [4]. Indeed, grey-tone segmentation mostly comes down to a contour detection problem, which can be approached by watersheds: contours can be defined in grey-tone images as regions where the grey values are varying very fast, i.e. as *crest-lines of the gradient.* Notice that the gradient of a decimal image I can have several morphological definitions, the most common among them being:

$$\mathrm{grad}(I) = (I \oplus H) - (I \ominus H).$$

The determination of the crest lines of $\mathrm{grad}(I)$ can be done by means of the watersheds transformation. Finally, we define the contours of a grey-tone image I as the **dividing lines of its gradient.**

7.4.2 Geodesic watersheds

The watersheds of the gradient build a very general approach of contour detection. However, the resulting images are most of the time **over-segmented,** i.e. the relevant contours are swamped by a mass of irrelevant ones. This is often due to the fact that the images under study are noisy. Moreover, this approach is unsatisfactory in the sense that it does not make use of any markers when in fact we know that the first step of every segmentation is the marking of the objects to be segmented.

This initial marking step actually makes use of an external knowledge about the image or the collection of images under study. We may well want to extract only one type of objects among all those that are present. To achieve this, we shall first *mark* our objects—i.e. make use of the particular knowledge available on the problem—by means of a procedure that can be completely different from a problem to another. The question that arises then is the following: is it possible, starting from these markers, to detect the precise contours of our objects and to avoid at the same time the appearence of irrelevant contour arcs? The response is yes. By using markers, we will not remove the irrelevant contour arcs of the watersheds of the gradient, but we will avoid the over-segmentation by **modifying the gradient function** on which the watersheds are computed.

Let now I be a grey-tone image and suppose that the desired markers have been extracted. Denote $M \subset \mathbb{Z}^2$ this set of markers. They must correspond exactly to the minima of the function $\theta(I)$ on which we plan to compute watersheds. Moreover, the second requirement is that this function must be as close as possible to the gradient function $\mathrm{grad}(I)$. It is only under this condition that its dividing lines will be properly located. Therefore, starting from $\mathrm{grad}(I)$, the construction of $\theta(I)$ is done in two steps:

- Impose as minima the previously extracted markers (i.e. the set M).
- Suppress the undesirable minima.

In step 1, we simply construct the function f defined by:

$$\forall p, f(p) = \begin{cases} c & \text{when } p \in M, \\ \mathrm{grad}(I)(p) & \text{otherwise,} \end{cases}$$

with c being an arbitrary constant, strictly minorating $\mathrm{grad}(I)$.

In the step 2, we have to suppress the unwanted minima of f, without forgetting to fill their associated basins! To do so, we first construct the following function g:

$$\forall p, g(p) = \begin{cases} c & \text{when } p \in M, \\ A & \text{otherwise,} \end{cases}$$

with A being an arbitrary constant majorating $\mathrm{grad}(I)$. Then, we iteratively erode g geodesically "over" f until stability is reached. These two operations

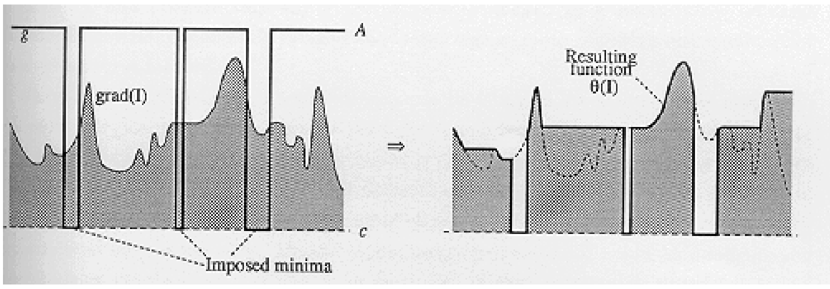


Figure 7.15: Construction of the function whose watersheds correspond to the desired contours.

are illustrated in Fig. 7.15. The resulting function, $\theta(I)$, is such that its watersheds correspond exactly to the desired contours. Note that this transformation ($\mathrm{grad}(I) \longmapsto \theta(I)$) is actually a $\vee$-filter, since it is the composition product of a closing by an opening!he whole procedure presented above may be referred to as *geodesic watersheds segmentation.* It is extremely powerful in a number of complex segmentation cases, since the only problem (which can be itself very complicated!) comes down to detecting the markers of the objects to extract.

7.4.3 Variants of the watershed approach

The general algorithm detailed in the preceding section is based on the detection of markers of the objects present in the grey-tone image I under study. In simple cases, these markers are merely the minima of the gradient of I. Indeed, an object O often corresponds to a region of I which is relatively homogeneous compared to its neighborhood. Therefore, this region is nothing but a minimum of the gradient $\mathrm{grad}(I)$. The associated basin "extends" this region until it is bounded by crest lines of the gradient, i.e. by the actual contours of O.

For more complex problems, taking the minima of the gradient image of I is far from providing a good set of markers. As explained above, this may be due to noise or to the fact that the desired objects constitute only a small subset of the objects present in the image. In such cases, special marking procedures have to be designed before applying the segmentation algorithm of § 7.4.2.

But in some other cases, it is impossible to find markers of the regions or of the objects to extract, since these objects or regions are not themselves well defined. This kind of situation often occurs, for instance, with remote sensing images, where the large variety of zones (fields, roads, houses, towns, lakes,etc ... under different lighting conditions) makes it almost impossible to design good marking procedures. Therefore, other segmentation algorithms have to be applied. One of the possibilities is to use the following approach:

- Computation of the watersheds of $\text{grad}(I)$. This results in an awfully over-segmented image.

- Removal of the irrelevant contours.

This kind of method is part of a more general class of algorithms called *region growing algorithms:* Starting from the watersheds image $W(\text{grad}(I))$, we can assign to each of its different basins a value characterizing them (e.g. the mean value of the corresponding pixels in I), and produce this way a sort of *mosaic image.* In a second step, adjacent basins may be progressively merged into larger regions (thus removing contours) until a given criterium is fulfilled.

Many different criteria for merging regions can be found in literature. When associated with morphological treatments, some of them may result in particularly good segmentations [5]. Another way of approaching this problem is to regard the mosaic image of the catchment basins as a planar graph, whose vertices are the different regions and whose edges are each pair of adjacent regions. This kind of object being a lattice, it can be processed by mathematical morphology [78]. Morphological merging procedures, based on gradients and watersheds on graphs [77], seem then to provide very efficient segmentation methods.

Chapter 8

3-D grids and Operators

8.1 Introduction

In image processing, 3-D treatments appeared during the 80's for both analysis and synthesis purposes. In the present paper, we concentrate on analysis of images, or more precisely, of stacks of binary images. These piles of sections are nowadays currently produced macroscopically (e.g. NMR), or at microscopical scales (e.g. confocal microscopes). They produce experimental data on 3-D rasters which tend to be cubic. Downstream, these computerized data are binarized by some techniques we will not consider here. These binary data constitute, by definition, sets in $\mathbb{Z}^3$, as well as estimations of sets of $\mathbb{R}^3$. How to access them ? How to extend to the 3-D space the usual 2-D notions of sizes, directions, distances, connectivity, homotopy, etc. ? This is what we would like to develop hereafter. A survey of literature shows that in 3D morphology, the two places that have been producing the most substantial series of results, and for a long time, are the pattern recognition section, at Delft University of Technology (see in particular P.W. Verbeek [76], J.C. Mullikin [51], Jonker [29]) and the Centre de Morphologie Mathématique, at the Ecole des Mines de Paris (see in particular Serra [60], Meyer [47] [18], Gratin [19], Gesbert et al.[16]).

8.2 Three dimensional grids

By grid, we do not only mean a regular distribution of points in the the 3-D space, but also a definition of the elementary edges, faces, and polyhedra associated with these points. The three crystallographic grids we find below derive from the cube, and are constructed as follows

i/ cubic grid, which is generated by translations of a unit cube made of 8 vertices ;

ii/ the centred cubic grid (cc grid) where the centres of the cubes are added to the vertices of the previous grid ;

iii/ the face-centred cubic grid (fcc grid) where the centres of the faces are added to the vertices of the cubic grid.

A comprehensive comparison of these grids can be found in F. Meyer's study [47].

8.2.1 Interplane distances

In the last two grids, the vertices generate square grids in the horizontal planes, and in vertical projection the vertices of plane No n occupy the centres of the squares in plane No $n-1$. We shall say that these horizontal plane are *staggered*. If a stands for the spacing between voxels in the horizontal planes, then the interplane vertical spacing is equal to $a/2$ in the cc case, and to $a\sqrt{2}/2$ in the fcc one.

8.2.2 First neighbors

Every vertex has

- 6 first neighbors in the cubic case
- 8 first neighbors in the cc case
- 12 first neighbors in the fcc case

Geometrically speaking, when point x is located at the centre of the $3 \times 3 \times 3$ cube, its projections

- on the faces of the cube provide the cubic neighbors

- on the vertices the cc-neighbors
- and on the edges the fcc-neighbors

Fig. (8.1) illustrates this point. One can see, also, that the first neighbors generate the smallest isotropic centred polyhedron of the grid, i.e. a 7-voxel tetrahedron (cubic case) a 9-voxel cube (cc-grid) a 13-voxel cube-octahedron (fcc grid). Denote them by the generic symbol B, and the n^{th} iteration of B by B_n, i.e.

$$B_n = B \oplus B \quad ... \quad \oplus B \qquad n \text{ times },$$

with $B_0 =$ Identity. From the implication $n \geq p \Rightarrow B_n \geq B_p \quad n, p$ non negative integers, from the equality $B_n B_p = B_{n+p}$, and from the symmetry of B we draw (proposition 2.4 in Serra [63]) that the 3-D raster of points turns out to be a metric space (in three different ways, according to the grid), where the smallest isotropic centred polyhedron is the unit ball.

8.3 Elementary edges, faces, and polyhedra

In order to complete the definition of the grids, we will introduce now elementary edges, faces and polyhedra. Edges are necessary to define paths, hence connectivity. Faces and polyhedra are required to introduce notions such as Euler-Poincare number for example, or more generally, to introduce the graph approach.

8.3.1 Cubic grid

As elementary edges, the best candidates are obviously the closest neighbors (in the Euclidean sense), i.e. those of fig. (8.1). However, they are not so numerous, in the cubic and in the cc case, in particular, which leads to poor connections. For example, in the cubic grid, the extremities of the various diagonals are not connected, we meet here a circumstance similar to that which led to the 8 and 4-connectivities in the 2-D grid. For the same reason, the authors who focused on the cubic grid, such as A. Rosenfeld [31], at the beginning of the 80's, introduced the 26- and the 6-connectivity on the cubic grid. When the foreground X is 26-connected, then the background X^c is

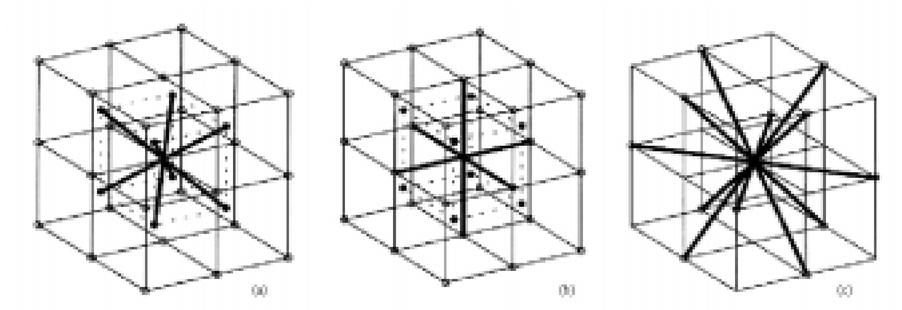


Figure 8.1: The three 3-D grids that derive from the cubic symmetry.

6-connected and vice-versa. In other words, a voxel $x \in X$ admits, as edge partners, all those voxels $y \in X$ that pertain to the cube $C : 3 \times 3 \times 3$ centred at x. Coming back to fig.(8.1b), we now have to take into account not only the centres of the cube faces, but also the 12 middle points of its edges, and its 8 vertices. Such an extension of the connectivity for X is possible only when the connectivity on X^c remains restricted to the six closest neighbors. If not, we should run the risk of over crossings of diagonals of $1's$ and of $0's$, so the faces should be undefined.This dissymmetrical connectivity brings into play a second digital metric, where cube $3 \times 3 \times 3$ is the unit ball. In particular, the boundary of set X^c is

$$\delta X^c = X^c \backslash X^c \ominus C$$

whereas the boundary of set X is defined via the unit tetrahedra T:

$$\delta X = X \backslash X \ominus C$$

We draw from this last equation that $\delta X \ominus T = \emptyset$, and from the previous one that $\delta X^c \ominus C = \emptyset$. The boundary of X is thinner, but it may comprise zones of a thickness 2, and of course lines or fine tubes.

Note also that, unlike tetrahedron T, cube C admits a Steiner decomposition into three orthogonal segments of three voxels length each. Consequently, the dilation $X \oplus nC$ is obtained as the product of three linear dilations of size $2n$ in the three directions of the grid.

8.3.2 cc grid

The cc grid call very similar comments, but now with staggered horizontal planes. The low number of the first neighbors (i.e. 8) of each voxel suggests to add the second neighbors, in number of six (see fig.8.1). This results in the unit rhombododecahedron R shown in fig.(8.2), which exhibits 15 vertices (including the centre), 12 rhomb faces, identical up to a rotation, and 24 edges whose common length is the first neighbor distance.

Just as previously, with the cubic grid, the adjunction of 2nd neighbors complicate the situation, for they cannot be added simultaneously to the $1's$ and $0's$. This results in a 14-connectivity for the grains versus a 8-connectivity for the pores. By comparison with the cubic case, the connectivity contrast between foreground and background is reduced, but it remains.

Again, as previously, a new metric is provided, namely that of the rhombododecahedron. In this metric, the isotropic dilations can be decomposed into segment dilations, since R admits a Steiner decomposition into the four diagonals of the cube $(2, 2, 2)$, i.e.

$$R = \begin{pmatrix} 1 & . \\ . & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & . \\ . & 1 \end{pmatrix} \oplus \begin{pmatrix} . & -1 \\ 0 & . \end{pmatrix} \oplus \begin{pmatrix} . & 0 \\ -1 & . \end{pmatrix} \tag{8.1}$$

where -1, 0, 1 indicate the level of the plane, and where the origin is always assigned to the point of plane 0 [10].

8.3.3 fcc grid

With the fcc grid, things become simpler [6][8]. We still are in a grid where the odd horizontal planes have been shifted by $(a/2, a/2, 0)$ from the cubic spacing, but now each voxel x admits 14 nearest neighbors, at a distance $a\sqrt{2}/2$. They form the unit cube-octahedron D, of figure (8.2), centred at point x. Geometrically speaking, such a high number of first neighbors means that the shape of D is a better approximation of the Euclidean sphere, than those of the cube C and the rhombododecahedron R.

As far as connections are concerned, it becomes cumbersome to resort to 2nd neighbors. Therefore there no longer is a risk of diagonal overcrossing. The existence of an edge no longer depends on the phase under study but exclusively on the intersection between grid and sets: two neighbors $1's$ define an edge in set X, two neighbors $0's$ an edge in set X^c.

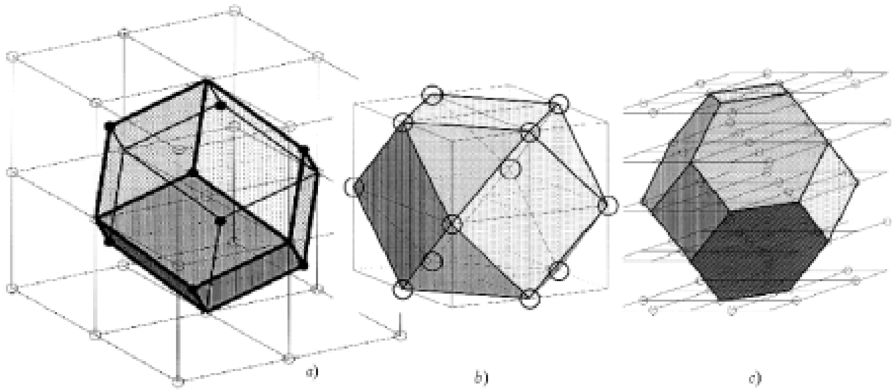


Figure 8.2: a): The rhombo-dodecahedron; b): The cube-octhedron; c) The tetrakai-decahedron.

Finally X and X^c are treated by the same balls D_n, but the latter cannot be decomposed into Minkowski sum of segments, unlike C and R.

8.3.4 Digital fcc grids, virtual staggering

How to produce a stack of staggered square grids, or, equivalently, how to produce a digital unit cube-octahedron ? An easy way is to favor the diagonal horizontal directions, as in Eq. **??**. The staggering structure is created automatically, since each of the two diagonal subgrid appears, alternatively, in the successive horizontal planes. The negative counterpart is that half of the voxels only are taken into account. For example, the dilation of Eq. **??** produces neither the central points at levels +1 and -1, nor the middle points of the sides at level zero. We may always add these points, in order to complete the basic cube-octahedron, but then

i/ We increase the elementary size from 13 up to 19 voxels, hence we become less accurate in delineating boundaries, ultimate erosions, skeletons, etc.

ii/ We lose the advantage of a unique type of edges, which governs homotopy and connectedness properties.

iii/ We do not know what to do with the amount of information carried by the non used voxels.

$$\textit{upper and lower planes} \begin{pmatrix} 1 & 1 & . \\ 1 & 1 & . \\ . & . & . \end{pmatrix} \quad \textit{central plane} \begin{pmatrix} . & 1 & . \\ 1 & 1 & 1 \\ . & 1 & . \end{pmatrix} \quad a)$$

$$\textit{upper and lower planes} \begin{pmatrix} . & . & . \\ . & 1 & 1 \\ . & 1 & 1 \end{pmatrix} \quad \textit{central plane} \begin{pmatrix} . & 1 & . \\ 1 & 1 & 1 \\ . & 1 & . \end{pmatrix} \quad b)$$

Figure 8.3: Decomposition of the unit cube-octahedron on the cubic grid in order to simulate the staggered structure (a: odd central plane, b: even central plane).

An alternative solution should consist in interpolating one horizontal grid every two planes. This would add a computational step, but above all, it seems "fiddled": how to weight the four horizontal neighbors, versus the two vertical ones ? How to display the resulting grid ? etc. Therefore, we propose neither to move nor to remove or even modify, any voxel of the cubic initial data, and to consider each even plane, *as it is*, as being staggered. According as the central plane is odd or even, we then obtain one of the two elementary polyhedra of fig. 8.3.

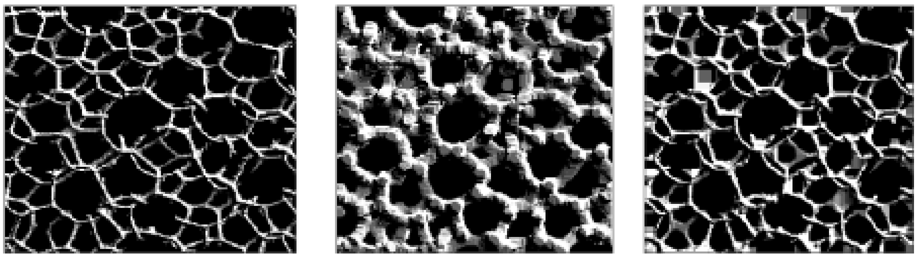
Such a *virtual staggering* is similar to that used in $\mathbb{Z}^2$, when one generates a hexagonal grid from a square raster. In both cases, the irregularity of the unit polyhedron (resp. polygon) is self-compensated by iteration. In other words, the mappings which bring into play sequences of successive sizes, such as distance functions, medial axes, granulometries, sequential alternated filters, etc. are treated by means of *actual* digital cube-octahedra (resp. hexagons), (see fig. 8.4).

8.3.5 Comparison of the grids

As a conclusion, three reasons argue in favor of the fcc grid, namely

1/ the shape of the cube-octahedron D provides a better approximation of the unit Euclidean sphere, than C or R (isotropic dilations, skeletons, distance functions, etc. will seem more "Euclidean") ;

2/ D is more condensed : 13 points on 3 consecutive planes (D) are more economic than 15 points on 5 planes (R), or 27 on 3 planes. D leads to



a) b) c)

Figure 8.4: expanded polystyrene (foam); b) Cube-octahedral dilation of size 5 of a); c) Cube-octahedral closing of size 12 of a).

thinner boundaries, to finer ultimate erosions, etc. and requires less logical tests in its implementation.

3/ In the fcc grid, the connectivity is based on the first neighbors only, which allows a common approach for grains and for pores (in cubic grid, when one decides to attribute *a priori* more than four possible neighbors to the $1's$ than to the $0's$, a rather severe assumption is made, which holds, paradoxically, on the convexity of the *pores.* Most often, both grains and pores exhibit concave and convex portions, and the 26/6-connectivity assumption is just irrelevant).

Facing these advantages, the weakness of the fcc grid is the staggered organization of its successive horizontal planes. However, is it really a drawback ?

8.4 Increasing operations and their residues

As soon as spheres and lines (in a set of directions) are digitally defined, it becomes easy to implement isotropic and linear dilations and erosions, hence openings, closings, granulometries, and all usual morphological filters.The example of figure8.4 illustrates this point. Similarly, the residuals associated with distance function, i.e. skeletons (in the sense of "erosions\openings"), conditional bisectors, and ultimate erosions derive directly.

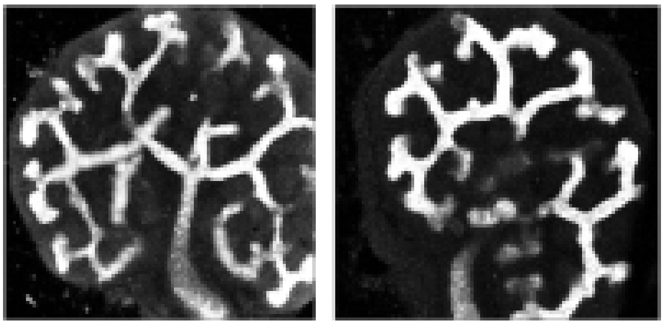


Figure 8.5: a) Kidney under study (supremum of the sections); b) other kidney specimen.

8.5 Confocal microscopy of embryonic kidney

In February 1999, Dr. John Bertram[1], nephrologist, presented his current work at CMM. The subject of his research is the embryonic development of the kidney studied in animals such as the rat. He takes advantage of the property of embryonic kidney to develop in vitro, which enables him to study the organ evolution by confocal microscopy without animal destruction.[3]. The data of the example which follows comes from Dr Bertram's laboratory.

We can see in fig.(8.5) an image of each kidney after binarization, showing that the structure develops in the form of a tree. The expected morphological description goes far beyond the preliminary study. It bears on the geometry of the tree, and involves two objects :

- extremities : where are they located? how are they arranged in space?
- branches : where are they located? according to which hierarchy and length?

Confocal microscopy results in a highly anisotropic sample. Each series contains 29 sections 30 μ thick ; in which the orientation is roughly perpendicular to the trunk.

On each section, the pixels are arranged according to a square grid, whose

[1]Dpt of Anatomy, faculty of Medecine, Univ. de Melbourne, Parleville, Victoria 3052, Australia

spacing is about 4μ. The digital volume element (voxel) looks like a cylinder with a square base, which is seven times as high as it is wide. Each branch extremity is surrounded by nephrons, whose number is indicative of the future capacity of the fully-grown kidney. The nephrons, which cannot be seen here, will become visible through a double staining. Then, we will have to study the relationship between the shape of the tree and the number of nephrons it can receive.

8.6 3-D Geodesic wavefronts

8.6.1 Wavefronts and tree diagrams

Let Z be a compact set in $\mathbb{R}^n$ and $x \in Z$, be a point in Z. The vawefront from a point x at distance λ is the geodesic sphere $F(\lambda, x)$ of (geodesic) radius λ and centered in x, where geodesy is generated by field Z.We propose to study the evolution of the connected components number of the wavefront $F(\lambda, x)$ when, as λ increases, the compact space Z is swept. The two types of branching, division or confluence, supposedly remain in finite number when $\lambda \in [0, \lambda_{\max}]$, so that for any branching at $\lambda = \lambda_0 < \lambda_{\max}$, it is always possible to find an open interval $]\lambda_{1,}\lambda_2[$ containing λ_0, and inside which there are no other branching. The number of branches which may gather in λ_0 is supposed to be finite. Finally, as the branching may take the two dual shapes (division or confluence) when λ increases, it is conventionally agreed in the proof below that the passage $\lambda_1 \to \lambda_2$ corresponds to a division

Therefore, we are led to the situation described in figure(8.6), where point x is in black, the ball $\overset{\circ}{B}(\lambda_0, x)$ in light grey, its complement $K(\lambda_0)$ in Z in dark grey, and where the white wavefront indicates the precise moment of the branching. So, the compact set

$$K(\lambda) = Z \setminus \overset{\circ}{B}(\lambda, x)$$

has a unique connected component, when $\lambda < \lambda_0$, and more when $\lambda > \lambda_0$. In order to determine what happens when $\lambda = \lambda_0$, we first observe that for compact sets, we have $\cap \{K(\lambda), \lambda < \lambda_0\} = K(\lambda_0)$.

The compact $K(\lambda_0)$ is composed of only one connected component. Otherwise, they would be separated by a minimum distance d ; but this is in-

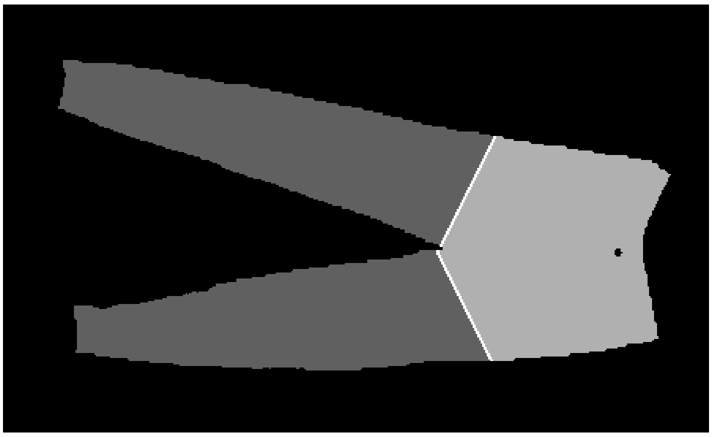


Figure 8.6: Example of branching

compatible with the fact that, for any dilation of size ε, with $0 < \varepsilon < d$, the geodesic dilate of $K(\lambda_0)$ becomes connected. Therefore, the front $F(\lambda_0, x)$ itself is connected, as otherwise, to switch from one of its components to another one, it would be necessary to cross a $K(\lambda)$ with $\lambda > \lambda_0$, but these $K(\lambda)$ are not connected anymore.

When Z has several branchings, the same description applies for each branch, upwards or downwards from the propagation from point x, which consequently partitions the set Z into a series of successive pieces.

The case of the X branching has also to be considered. It occurs when at least two branches stop at the critical front, and at least two of them start from there. In this case, the intermediary connected region is reduced to the front in λ_0, for, if it was larger, we would come back to the previous case; and if the front was not taken into accout, we would no longer have a critical element, but only separated branches. By gathering these results, we can state :

Proposition 8.1 *Let Z be a compact of* $\mathbb{R}^n$. *If, for any point* $x \in Z$, *the wavefront* $F(\lambda, x)$ *emanating from* x *admits a finite number of connected components, with a finite variation, then, as radius* λ *varies,* $F(\lambda, x)$ *partitions Z into a finite number of connected sections, corresponding to open intervals of* λ, *and separated by connected components of the front which are* *located at the critical points of the branchings.*

Clearly, the mapping $x \to P(x)$ which associates with any point $x \in Z$ the tree diagram characterized by the proposition, depends on the choice of point x, even if, when considering the common meaning of a tree, the partition remains almost the same for all the points selected low enough in the trunk. Besides, in this case, the tree may be defined as a partition for which there is no confluence for a suitably selected origin x (i.e. in the trunk).

Note that we are talking about connectivity here, and not about homotopy: in $\mathbb{R}^3$ particularly, the sections may show closed pores or toric holes.

8.6.2 The ultimate elements of the wavefronts

This section takes up a classical C. Lantuejoul's and S. Beucher's result [33], but presents it differently. When using geodesics, it becomes possible to associate any point $x \in Z, Z \in \mathbb{R}^n$, with the point or points $y \in Z$ which are the furthest away from x. Indeed, let $\overset{\circ}{B}(\lambda, x)$ be the geodesic open ball of radius λ and centre x, and $\lambda_0$be the upper limit of the λ such that $\overset{\circ}{B}(\lambda, x)$ be strictly contained in Z. As the non empty compact sets $\left\{Z\backslash \overset{\circ}{B}(\lambda, x), \lambda < \lambda_0\right\}$ decrease and that $\mathbb{R}^n$ is a separated space, the intersection

$$x < \lambda_0 \cap \left[Z\backslash \overset{\circ}{B}(\lambda, x)\right] \tag{8.2}$$

is itself a non empty compact set, whose points are all at the maximum distance λ_0 from x. This intersection is named "geodesic ultimate eroded set", and $\overset{\circ}{B}(\lambda_0, x)$ is the "geodesic ultimate dilated set" of point x.

The existence of extreme points may also be considered in a regional framework, and not a global one anymore. We must suppose that, Z and x being given, it is possible to find a $\mu(Z, x) \leq \lambda_0(Z, x)$ such that each connected component of $Z \setminus \overset{\circ}{B}(\lambda, x), \mu \leqslant \lambda \leqslant \lambda_0$ decreases without subdividing. Then, the previous analysis should simply be applied to sets

$$K_i \cap \left[Z\backslash \overset{\circ}{B}(\lambda, x)\right] \qquad \mu \leqslant \lambda \leqslant \lambda_0$$

where the $K_i, i \in I$ refers to the connected components of $Z \setminus \overset{\circ}{B}(\mu, x)$. Therefore, we obtain the farthest connected components from point x, such as, for instance, the fingers tips for x taken around the middle of the wrist.

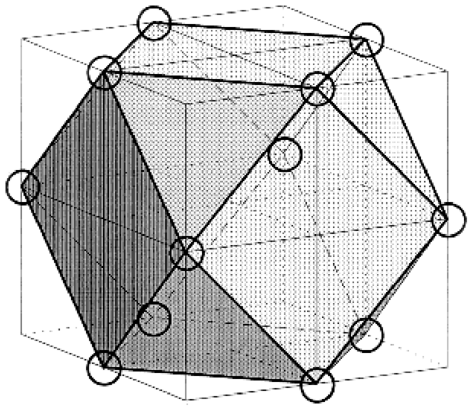


Figure 8.7: Cube-octahedron

Both algorithm families about geodesics correspond to both our points of view. Invasion by geodesic balls led to all the particles reconstruction variants (deletion of the grains crossing the field border, hole filling, individual analysis, etc ...) and the search for extreme residues led to the ultimate eroded points, to the objects limits and to the length of a connected component (as a supremum of the distances between pairs of extreme points).

8.6.3 3-D Digital wavefronts

The digitization of geodesic operations may cause errors, but limited ones ; indeed, it is advisable to choose, as a circle or unit sphere, the closest shapes to their Euclidean homologues. Therefore, in 2D the hexagon, whose six vertices are equidistant from the center is better than the square, and, for the same reason, the cube-octahedron is better than the cube in 3D.

This $\mathbb{Z}^3$ ball is very easy to build, when a numerical data network in square grid [70] is available. It suffices to shift all even planes by half a diagonal of the unit cube (any diagonal, but always the same one). In practice, data are of course not moved, but only structuring elements. For example, the substitute for the 13 vertices of the regular cube-octahedron of fig.(8.7) is calculated by dilating the central point according to the staggered unit cube-octahedron presented in the previous chapter (which differs whether the center lies in an even plane or in an odd one). The wavefront emanating from this central point starts with the point 12 neighbours ; when the interplane equals $a/\sqrt{2}$ ($a =$ square grid spacing of the horizontal planes), the structure becomes completly isotropic and the 12 neighbours are equidistant from the center. This will be our assumption (section 4) about the shinbone, but this hypothesis is not essential, and, in any case, cannot be ventured for the study about embryonic kidneys (section 3)

The switch from the unit ball $C(x)$ of $\mathbb{Z}^3$ (octahedron, prism or cube) to its geodesic version $B_1(x)$ inside a mask Z is

$$B_1(x) = C\,(x) \cap Z$$

and the geodesic ball $B_x\,(x)$ of the size is obtained by n iterations of the previous one :

$$B_n\,(x) = B_1\,[B_{n-1}\,(x)] \cap Z$$

The corresponding wavefront, or geodesic *sphere* equals

$$F_n\,(x) = B_{n+1}\,(x) \setminus B_n\,(x)$$

8.7 Use of the tree diagram for embryonic kidneys

In order to illustrate the above matter, we propose to segment the first one of the two kidneys of fig. 8.5. The analysis contains four steps :

1/ set construction from the initial data ;
2/ geodesic distance function of a marker in the set;
3/ extremities;
4/ branches.

8.7.1 Binarization

This simple operation only requires a thresholding between 60 and 255, followed with the fill-in of the bi-dimensional internal pores. Still, the main connected component has to be extracted. In order to do this, we take as marker x one point at the beginning of the trunk. The reconstruction shows that the kidney tree diagram is broken around the middle in two separated

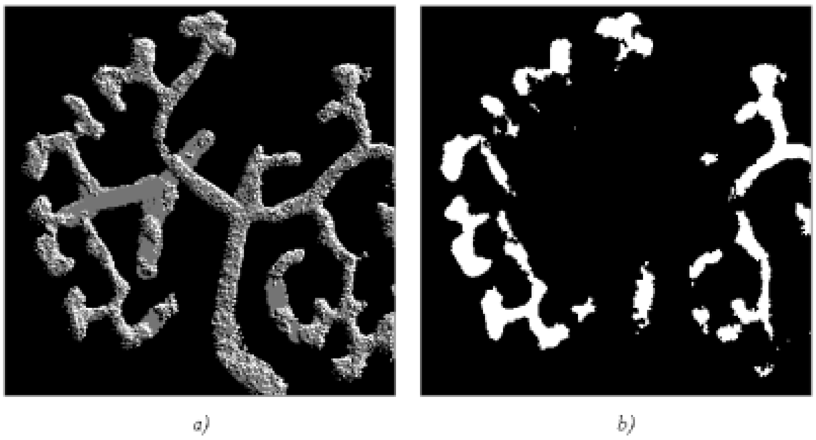


Figure 8.8: a) Perspective view of the binarized kidney; b) confocal section n]0xb014

parts. This is caused by the inaccuracy of confocal microscopy. In order to put it right, both parts have been reconnected by a small closing (see fig.(8.8a)).

8.7.2 geodesic distance function

The geodesic distance function starts from marker x at the base of the kidney and progresses inside the tree according to unit cube-octahedra (fig.8.9).

8.7.3 Extremities

The extremities are nothing but the region maxima of the previous geodesic function. These ultimate eroded points are shown on 8.10a, where lots of quite insignificant but very small real maxima can been observed. They are removed by a small surface opening (fig.8.10b). When using this algorithm in routinely, it would better to start with a regularization of the set under study by means of an isotropic tridimensional opening of size 1 or 2, providing that it does not break the connectivity.

Figure 8.9: Geodesic distance function from the anchorage point (negative view of the supremum of the sections)

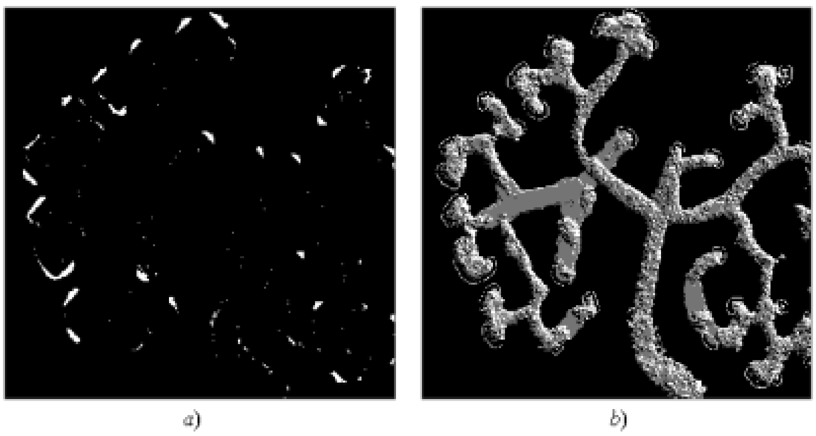


Figure 8.10: a) All extremities of the Kidney ; b) Filtered extremities.

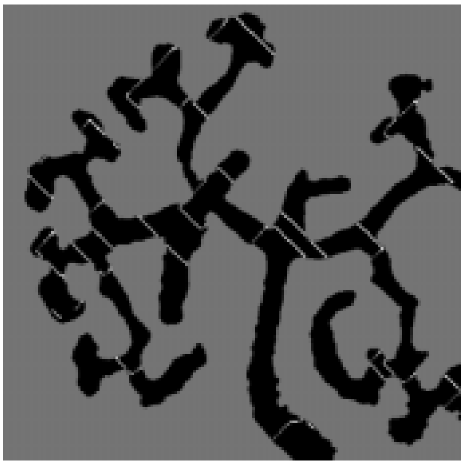


Figure 8.11: Projection of the 3D branchings

8.7.4 Branchings

The extraction of branchings, which is conceptually simple, may nevertheless lead to a appreciable computing time. Considering the quite visible structure of the projected tree, the algorithm used below is slightly less precise, but faster and easier to implement.

In a first step, bidimensional branchings on the tree projection are investigated, then, we get back to the 3D space by building vertical cylinders whose bases are located at the 2D branchings, and slightly dilated (size 2). Finally, we take the intersection between these cylinders and the 3D tree. The operation leads to fig.8.11.

8.7.5 Results

In all, starting from the connected kidney tree, we got to its segmentation into disjoint branches separated by thin branchings. Some branches contain one or more, of the tree extremities. From such a segmentation, it now becomes possible to replace the object under study by a "tree" in the meaning of graph theory, where the edges can be weighted geometrical characteristics (volume, length, location of its center, possible end points ... etc).

Bibliography

[1] **C. Arcelli and G. Saniti di Baja**, "Width-independent fast thinning algorithms", *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 7, pp. 463-474, 1985.

[2] **G. Bertrand**, "Simple points, topological numbers and geodesic neighborhoods in cubic grids", *Pattern Rec. Letters*, vol. 15, pp. 1003-1012, 1994.

[3] **Clark A.T., Young R.J. and Bertram J.F**. In vitro studies on the roles of transforming growth factor-beta1 in rat metaphrenic development, *Kidney Int* 59 pp 1641-1653 (2001).

[4] **S. Beucher & C. Lantuejoul,** Use of watersheds in contour detection, *Proc. Int. Workshop on image processing, real-time edge and motion detection/estimation,* Rennes (France), September 1979.

[5] **S. Beucher,** Analyse de diagraphies électriques par morphologie mathématique, int. report, CMM, School of Mines, Paris, January 1989.

[6] **S. Beucher & L. Vincent,** Introduction aux outils morphologiques de segmentation*, J. de Spectroscopie et de Microscopie Electronique*, 43 p., September 1989.

[7] **P. Bhanu Prasad, J.P. Jernot, M. Coster, J.L. Chermant**, "Analyse morphologique tridimensionnelle des matériaux condensés", *in Proc. PIXIM,* 1988, pp. 31-42.

[8] **G. Birkhoff,** *Lattice theory,* 3^{rd} edition, A.M.S. Colloq. publ., vol. 25, 1983.

[9] **T..S. Blyth & M.F. Janowitz,** *Residuation theory* Pergamon Press, Oxford, 1972

[10] **G. Choquet** *Topologie* Academic Press, N.Y. 1966.

[11] **M. Coster & J-L. Chermant,** *Précis d'analyse d'images,* CNRS ed., Paris, 1985.

[12] **J.Crespo**, **J. Serra, R.W. Schafer** Theoretical aspects of morphological filters by reconstruction. *Signal Processing,* 1995, Vol. 47, No 2, pp. 201-225.

[13] **J.Crespo, R.W. Schafer,** Locality and adjacency stability constraints for morphological connected operators. *Journal of Mathematical Imaging and Vision* 7, 1 (1997), 85-102.

[14] **Fricout G**. Kidney and Skeletonization Int. report CMM, Ecole des Mines de Paris nov 2000.

[15] **Fricout G., Cullen-Mc Ewen L.A., Harper I.A., Jeulin D., Bertram J.F.** A quantitative method for analysing 3D branching in embryonic kidneys:

[16] S**. Gesbert, C.V. Howard, D. Jeulin and F. Meyer,** "The use of basic morphological operations for 3D biological image analysis", *Trans. Roy. Microsc. Soc.*, vol. 1, pp. 293-296, 1990.

[17] **G.Gierz et Al.** *A Compendium on Continuous Lattices*. Springer, 1980.

[18] **C. Gratin and F. Meyer**, "Mathematical Morphology in three dimensions", *Acta Stereol.* vol. 11, pp. 551-558, 1991.

[19] **C. Gratin**, *De la représentation des images en traitement morphologique d'images tridimensionnelles,* PhD thesis, Ecole des Mines de Paris, 1993.

[20] **H. Hadwiger**, *Vorsesungen über Inhalt, Oberfläche und Isoperimetrie,* Springer, 1957.

[21] **P.R. Halmos,** *Lectures on Boolean Algebras.* D. Van Norstand Company, Inc., Princeton, 1963.

[22] **Haralick R.M. and Shapiro L.G.** *Computer and robot vision.* Vol. I. Addison Wesley, 1992, pp. 191-198.

[23] **H. Heijmans & C. Ronse,** The algebraic basis of mathematical morphology, part one: erosions and dilations, *Computer Vision, Graphics and Image Processing 50* pp 84-102 (1990).

[24] **H.J.A.M. Heijmans,** Morphological image operators. *Advances in Electronics and Electron Physics,* 1994, suppl. 24, Vol. 50, Hawkes P. ed., Boston: Ac. Press.

[25] **Heijmans, M.J.A.M.** Connected morphological operators. Tech. Rep. CWI No PNA-R9708, April 1997.

[26] **G. Heygster,** Rank filters in digital image processing, *Computer Graphics & Image Processing* 19 (2), pp 148–164, June 1982.

[27] **V.C. Howard and K. Sandau**, "Measuring the surface area of a cell by the method of the spatial grid with a CSLM", *J. of Micr.,* vol. 165, pp. 183-188, 1992.

[28] **James and James**, *Mathematics Dictionary,* Van Nostrand, 1982.

[29] **P.P. Jonker**, "Parallel processing in computer vision and collision free path planning of autonomous systems", in *26th ISATA*, 1993.

[30] **Klein J-C**. *Conception et réalisation d'une unité logique pour $\mathcal{L}$'analyse quantitative d'images.* Thesis, University of Nancy, 1976.

[31] **T. Kong and A. Rosenfeld**, "Digital topology: Introduction and survey", *Comp. Vis. Image Proc.*, vol. 48, pp. 357-393, 1984.

[32] **Ch. Lantuejoul** "Skeletonization in quantitative metallography". *Issues in Digital Image Processing* (R.M. Haralick and J.C. Simon Edts) Sijthoof and Noordoff, 1980.

[33] **C. Lantuejoul & S. Beucher,** On the use of geodesic metric in image analysis, *J. of microscopy,* 121, pp 39–49, 1981.

[34] **P. Maragos & R.W. Schafer,** Morphological filters—part I: their set-theoretic analysis and relations to linear shift-invariant filters, *IEEE transactions on acoustics, speech and signal processing*, Vol. ASSP-35 (8), pp 1153–1169, August 1987.

[35] **P. Maragos & R.W. Schafer,** Morphological filters—part II: their relations to median, order-statistics, and stack filters, *IEEE transactions on acoustics, speech and signal processing,* Vol. ASSP-35 (8), pp 1170–1184, August 1987.

[36] **Marcotegui B. and Meyer F.** Morphological segmentation of image sequences. In *Mathematical Morphology and its applications to image processing,* Serra J. and Soille P. eds., Kluwer, 1994, pp. 101-108.

[37] **G. Matheron,** *Eléments pour une théorie des milieux poreux,* Masson, Paris, 1967.

[38] **G. Matheron,** *Random sets and integral geometry,* Wiley, New-York, 1975.

[39] **G. Matheron** Filters and Lattices. Ch 6 in *Image Analysis and Mathematical Morphology.* Vol. 2. Serra J. ed. London: Acad. Press, (1988), pp.115-140.

[40] **G. Matheron and J. Serra** Strong filters and connectivity.Ch 7 in *Image Analysis and Mathematical Morphology.* Vol. 2. Serra J. ed. London: Acad. Press (1988), pp. 141-157.

[41] **G. Matheron** Treillis compacts et treillis coprimaires. Tech. rep. N-23/90/G, Ecole des Mines de Paris, Part 1, 1990, part 2, 1996.

[42] **G. Matheron** Les nivellements. Tech. Rep. Ecole des Mines de Paris, No N-07/97/MM, Feb. 1997.

[43] **F. Meyer** Rapport de stage note interne CMM, (1976).

[44] **F. Meyer,** *Cytologie quantitative et morphologie mathématique,* Ph.D. Thesis, School of Mines, Paris, 1979.

[45] **F. Meyer & J. Serra,** Contrasts and activity lattice, *Signal Processing, special issue on Math. Morph.,* vol. 16 (4), pp 303–317, April 1989.

[46] **F. Meyer & J. Serra,** Filters: from theory to practice, *Acta Stereologica*, 8/2, pp 503-508, 1989.

[47] **F. Meyer,** "Mathematical Morphology: from two dimensions to three dimensions", *J. of Micr.,* vol. 165, pp. 5-29, 1992.

[48] **F. Meyer,** Minimum spanning forests for morphological segmentation. In *Mathematical Morphology and its applications to image and signal processing,* Maragos P. et al. eds., Kluwer,1996.

[49] **F. Meyer,** The levelings. In *Mathematical Morphology and its applications to image and signal processing,* H. Heijmans and J. Roerdink eds., Kluwer, 1998.

[50] **E.H. Moore,** Introduction to a form of general Analysis, *AMS colloq. pub*. vol.2 New Haven 1910.

[51] **J.C. Mullikin**, *Discrete and Continuous Methods for Three Dimensional Image Analysis*, Delft: Univ. Press, 1993.

[52] **J. Mukkerjee and B.N. Chatterji**, "Segmentation of three-dimensional surfaces", *Pattern Rec. Letters*, vol. 11, pp. 215-223, 1990.

[53] **Y. Pomeau and J. Serra**, "Contacts in random packings of spheres", *J. of Micr.*, vol. 138, pp. 179-187, 1985.

[54] **K. Preston & M. Duff,** *Modern Cellular Automata,* Plenum, New York, 1985.

[55] **C. Ronse,** Erosion of narrow image features by combining local rank and max filters, *Proc.* 2^{nd} *Int. Conf. on Image Processing and its Applications*, London, pp 77–81, 1986.

[56] **C. Ronse** Set theoretical algebraic approaches to connectivity in continuous or digital spaces. *JMIV*, 1998, Vol. 8, pp. 41-58..

[57] **C. Ronse, J.Serra** Geodesy and connectivity in grids (to be published in *Fundamenta Informaticae*).

[58] **Salembier P. and Oliveras A.** Practical extensions of connected operators. In *Mathematical Morphology and its applications to image and signal processing,* Maragos P. et al., eds. Kluwer, 1996, pp. 97-110.

[59] **J. Serra,** *Introduction à la Morphologie Mathématique*, *Cahiers du CMM*, Ecole des Mines de Paris, 1969.

[60] **J. Serra,** *Image analysis and mathematical morphology,* Academic Press, London, 1982.

[61] **J. Serra**, "Descriptors of flatness and roughness", *J. of Micr.*, vol. 134, pp. 227-243, 1984.

[62] **J. Serra,** *Image analysis and mathematical morphology, part II: theoretical advances,* J. Serra ed., Academic Press, London, 1988.

[63] **J. Serra**, "Mathematical Morphology for Boolean lattices", chapter 2 in *Image Analysis and Mathematical Morphology*, vol. 2, J. Serra (ed.), London: Acad. Press, 1988

[64] **H.J.A.M. Heijmans and J. Serra,** Convergence, Continuity, and Iteration in Mathematical Morphology *Journal of Visual Communication and Image Representation*, Vol. 3, N°1, March 1992, 84-102.

[65] **J. Serra,** Toggle mappings, in "*from Pixels to Features*", J.C. Simon ed., pp 61–72, North Holland, 1989.

[66] **J. Serra, L. Vincent** An Overview of Morphological filtering , *IEEE Trans. on Circuits, Systems and Signal Processing*, Vol. 11, N°1, pp.47-108,1992 .

[67] **J. Serra,** Lecture Notes on Morphological Operators int. report CMM, 150 p., School of Mines, Paris, June 2001[C06/01/MM].

[68] **J. Serra,** Equicontinuous functions, a model for mathematical morphology. In *Non linear algebra and morphological image processing SPIE,* Vol. 1769, San Diego, July 1992, pp. 252-263.

[69] **Serra J. and Salembier P.** Connected operators and pyramids. In *Non linear algebra and morphological image processing SPIE,* Vol. 2030, , San Diego*,* July 1993, pp. 65-76.

[70] **J. Serra** Cube, cube-octahedron or rhombododecahedron as bases for 3-D shape descriptions", *Advances in Visual Form Analysis, C. Arcelli and Al. (eds.)* World Scientific 1997, 502-519.

[71] **J. Serra** Connectivity on complete lattices *JMIV*, 1998, Vol. 9, pp. 231-251

[72] **J. Serra** Connectivity for sets and functions *fundamenta Informaticae, 41 (2000) 147-186*

[73] **J. Serra** Morphological descriptions using three-dimensional wavefronts (to be presented at *ISS'*2001,Bordeaux, France sept. 2001)

[74] **M. Staub** (personnal communication).

[75] **S.R. Sternberg,** Morphology for grey-tone functions, *Computer Vision, Graphics and Image Processing* 35, pp 333–355, 1986.

[76] **P.W. Verbeek, H.A. Vrooman and L.J. van Vliet**, "Low level image processing by max-min filters", *Signal Process.,* vol. 15, pp. 249-258, 1988.

[77] **L. Vincent,** ElMathematical morphology for graphs applied to image description and segmentation, *electronic Imaging West*, Pasedena CA (USA), pp 313–318, April 1989.

[78] **L. Vincent,** Graphs and mathematical morphology, *Signal Processing, special issue on Math. Morph.,* vol. 16 (4), pp 365–388, April 1989.

[79] **P.D. Wendt, E.J. Coyle & N.C.Jr. Gallagher,** Stack filters, *IEEE transactions on acoustics, speech and signal processing*, vol. ASSP-34 (4), pp. 898–911, August 1986.

[80] **O. Yli-Harja, J. Astola & Y. Neuvo,** Analysis of the properties of median and weighted median filters using threshold logic and stack filter representation*,* submitted to ASSP-EDICS 4.4., 50 p., 1989 development of a technique and preliminary data.