



HAL
open science

Morphologie Mathématique et traitement d'images

Samy Blusseau, Élodie Puybareau

► **To cite this version:**

Samy Blusseau, Élodie Puybareau. Morphologie Mathématique et traitement d'images. Laurent Fuchs (éd.). Informatique Mathématique Une photographie en 2023, CNRS, Chapitre 3, pp. 87 –131, 2023, CNRS Alpha. hal-04148876

HAL Id: hal-04148876

<https://hal.science/hal-04148876>

Submitted on 3 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapitre 3

Morphologie Mathématique et traitement d'images

Samy Blusseau et Élodie Puybareau

Dans ce chapitre nous présentons la Morphologie Mathématique comme une approche non linéaire du traitement d'images, basée sur des critères de forme et de taille. Nous essaierons de montrer son attrait en mettant en avant l'élégance de sa théorie ainsi que la puissance des outils qu'elle permet de construire : fonction distance, filtres, algorithme du Watershed pour la segmentation et autres représentations hiérarchiques, pour les principaux.

3.1 Introduction

La Morphologie Mathématique a été fondée dans les années 1960 par Georges Matheron et Jean Serra, avant de devenir incontournable dans le traitement numérique des images pendant plus de quatre décennies. Son succès s'explique par une théorie accessible, aussi riche qu'élégante, qui a inspiré et s'est nourrie, dès ses débuts, d'une approche pratique à l'origine d'algorithmes géniaux.

Aujourd'hui en 2023, l'état de l'art en traitement d'images et vision par ordinateur repose presque essentiellement sur l'apprentissage statistique et en particulier le *deep learning*. Le paradigme d'adaptation automatique des algorithmes aux données est devenu si performant et universel qu'il mérite toute l'attention de la communauté pour en accroître la compréhension et le faire progresser. C'est pourquoi il n'est pas surprenant que de jeunes chercheuses et chercheurs, déjà très engagés dans ce domaine foisonnant, méconnaissent voire ignorent des approches dépourvues d'apprentissage.

Pourtant, s'intéresser à ces dernières, et en particulier à la "morpho", peut être bénéfique à plusieurs égards. D'abord, pour l'émerveillement et la stimulation qu'on peut éprouver en les découvrant. En suite, pour les idées nouvelles qu'elle peuvent inspirer dans le contexte scientifique actuel : se ré-appropriier un domaine moins visible offre un espace de recherche plus serein, propice à l'émergence de pistes originales et fécondes y compris dans les domaines plus en vue. Enfin, avoir une "culture morphologique", par exemple, c'est être capable de traiter des images lorsque leur qualité, leur nombre ou leurs annotations sont insuffisantes pour des approches supervisées, ce qui n'est pas rare dans les applications de la vie réelle.

Ce chapitre est donc pensé comme un point d'entrée rapide vers la Morphologie Mathématique, à l'attention de nos collègues, comme nous jeunes mathématicien.nes et informaticien.nes. Sur le contenu, il est largement redondant avec les références dont nous nous sommes inspirés [5, 21, 28], mais beaucoup moins complet. Nous l'espérons néanmoins assez synthétique et susceptible d'éveiller l'intérêt des personnes qui n'ont jamais entendu parler de la morpho ou qui n'ont pas encore pris le temps de s'y pencher. L'essentiel de ce qui est abordé est bien antérieur à 2023, mais c'est un pré-requis nécessaire pour apprécier les travaux récents dans le domaine, que nous esquissons surtout en fin de chapitre.

Enfin, soulignons que des implémentations open source des opérateurs et méthodes présentées existent, et ont été utilisées pour les illustration de ce chapitre : Smil¹ [11] pour la majorité, et Hgra² [22] pour la section 3.6.2.

3.2 Cadre algébrique

Une part importante du traitement d'images s'appuie sur l'algèbre linéaire. Les filtres des réseaux de neurones convolutifs (CNNs), par exemple, sont des opérateurs linéaires d'images, représentés par un noyau ("kernel"), ce qui les rend invariants par translation³. Plus précisément, si l'on voit une image comme une fonction $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$, le résultat $H[f]$ de sa convolution avec un noyau $h : \mathbb{Z}^2 \rightarrow \mathbb{R}$ est défini par

$$\forall x \in \mathbb{Z}^2, H[f](x) = \sum_{y \in \mathbb{Z}^2} f(y) \cdot h(x - y). \quad (3.1)$$

1. <https://smil.cmm.minesparis.psl.eu/doc/>

2. <https://hgra.readthedocs.io/en/stable/index.html>

3. Le terme *équivariant* est plus approprié, car le résultat du filtrage d'une image tradatée est la translation du filtrage de l'image originale, donc la sortie *varie comme* l'entrée. Mais "invariance par translation" est plus commun pour désigner cette propriété en traitement du signal et des images.

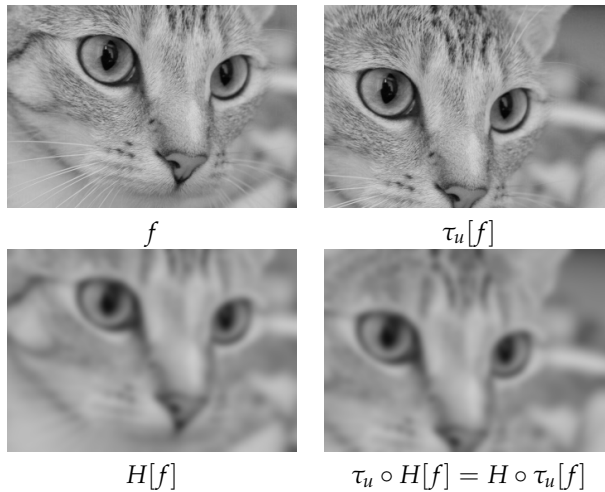


FIGURE 3.1 – Invariance par translation de la convolution d'une image par un noyau (ici un noyau gaussien $h(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2})$, avec $\sigma > 0$).

On vérifie aisément que l'opérateur H est linéaire. Par ailleurs, en notant τ_u l'opérateur de translation par un vecteur $u \in \mathbb{Z}^2$, c'est à dire $\tau_u[f](x) = f(x - u)$, on a bien $H \circ \tau_u[f](x) = H[f](x - u) = \tau_u \circ H[f](x)$, tel qu'illustré par la Figure 3.1. Réciproquement, on peut se convaincre que tout opérateur linéaire et invariant par translation est une convolution par un noyau telle que (3.1) (Théorème de Riesz).

Des résultats très analogues apparaissent si on change d'algèbre pour une algèbre non linéaire, par exemple en remplaçant la somme par un supremum (ou un infimum) dans (3.1), et le produit “ \cdot ” par “ $+$ ”. Bien qu'elle n'ait pas été introduite de cette manière originellement, la Morphologie Mathématique (MM) peut être vue comme l'étude des opérateurs dans ce type d'algèbres non linéaires. Alors que l'algèbre linéaire repose sur la structure d'espace vectoriel, stable par combinaisons linéaires, la structure algébrique de la Morphologie Mathématique est la *treillis complet* (*complete lattice* en anglais), ensemble partiellement ordonné et stable par supremum et infimum. De même que les applications linéaires sont les morphismes qui préservent la structure d'espace vectoriel en commutant avec les combinaisons linéaires, deux types de morphismes de treillis complets émergent naturellement. Ceux qui commutent avec l'infimum, que l'on appelle les *érosions* et ceux qui commutent avec le supremum, que l'on appelle *dilatations*.

tions. Voyons maintenant les définitions plus formellement, en commençant par celle d'image.

3.2.1 Images numériques

Les images deviennent des objets mathématiques dès qu'on les voit comme des fonctions : chaque point de l'image correspond une couleur ou un niveaux de gris, représentés respectivement par un vecteur ou un scalaire. Le traitement algorithmique des images nécessitant leur numérisation, en pratique celles-ci sont des fonctions d'un espace borné et discret, à valeurs bornées et discrètes. Il n'est pas rare néanmoins de les modéliser par des fonctions à valeurs réelles, comme dans l'exemple précédent, pour faciliter les développements théoriques.

Dans ce chapitre on appellera *image* une fonction $f : E \rightarrow \mathcal{V}$, où

- E peut être tout \mathbb{Z}^n ($n = 2$ ou 3) ou bien de façon plus réaliste une grille discrète finie, typiquement en deux dimensions $E = \{0, \dots, M\} \times \{0, \dots, N\}$, $M, N \in \mathbb{N}^*$
- \mathcal{V} est un ensemble fini et totalement ordonné de valeurs.

Les éléments de E sont les *pixels*, contraction de *picture elements*. On parle d'*image binaire* lorsque \mathcal{V} contient seulement deux valeurs, typiquement $\mathcal{V} = \{0, 1\}$, et d'*image à niveaux de gris* lorsque $\text{Card}(\mathcal{V}) > 2$, par exemple $\mathcal{V} = \{0, \dots, 255\}$ pour des images encodées sur huit bits. Par convention, la valeur associée à un pixel est d'autant plus élevée que celui-ci est clair, le noir et le blanc étant encodés par le minimum et le maximum de \mathcal{V} respectivement. Les *images couleur*, ont pour valeurs des vecteurs de dimension supérieure, le plus souvent trois, correspondant aux canaux rouge, vert et bleu, par exemple $\mathcal{V} = \{0, \dots, 255\}^3$.

Images binaires et ensembles Si E est une grille discrète, l'ensemble $\mathcal{P}(E)$ des parties de E est équivalent à l'ensemble des images binaires définies sur E , par la bijection qui à une image f associe $X_f = \{x \in E, f(x) = 1\}$. C'est cette correspondance qui a motivé les définitions de base de la Morphologie Mathématique sur les ensembles avant une généralisation aux images non binaires et aux treillis complets abstraits. Dans la suite, nous utiliserons régulièrement les ensembles pour illustrer des définitions abstraites, et les images binaires pour les visualiser et saisir leur pertinence dans l'analyse des formes.

L'image comme paysage En morphologie mathématique, une image à niveaux de gris est souvent vue comme un "paysage". Concrètement, on peut l'interpréter comme un relief topographique dont les altitudes sont

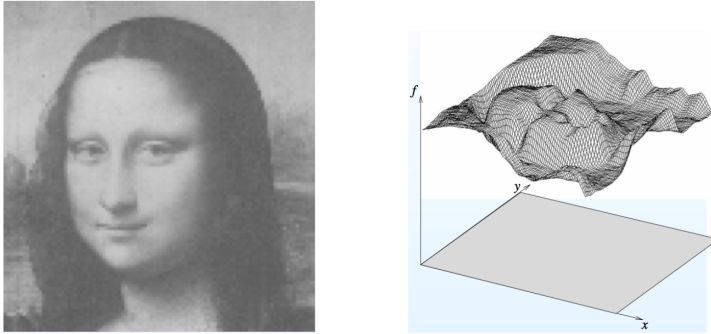


FIGURE 3.2 – Représentation d'une image sous la forme d'un paysage. A gauche : l'image originale. A droite : sa représentation sous la forme d'un relief. Il s'agit d'une représentation 3D dans laquelle l'image est une fonction de x, y et z , où x et y sont les lignes et colonnes, et z la valeur des pixels.

les valeurs des pixels. Ainsi, les parties les plus claires sont les sommets et les plus sombres les vallées. Cette représentation, illustrée Fig. 3.2, permet de sortir des représentations traditionnelles et de leurs limites, ainsi que d'imaginer de nouvelles manières de traiter les images. Elle a inspiré des algorithmes fondamentaux comme celui de la ligne de partage des eaux (*Watershed*), qui sera présenté dans la section 3.6.1.

Comme nous allons le voir, les ensembles d'images de même support sont des treillis complets, structure algébrique de base pour la Morphologie Mathématique.

3.2.2 Treillis complets

Rappelons d'abord la définition d'un ensemble partiellement ordonné.

Définition 1 (Ensemble partiellement ordonné (ou *Poset*)). *Un ensemble partiellement ordonné, ou Poset, est un ensemble E muni d'une relation binaire \leq , appelée ordre partiel sur E , qui vérifie pour tous $a, b, c \in E$*

$$\left\{ \begin{array}{ll} a \leq a & \text{(Réflexivité)} \\ (a \leq b \text{ et } b \leq c) \Rightarrow a \leq c & \text{(Transitivité)} \\ (a \leq b \text{ et } b \leq a) \Rightarrow a = b & \text{(Anti-symétrie).} \end{array} \right.$$

On note alors (E, \leq) ce poset. L'adjectif *partiel* qualifie le fait que deux éléments ne sont pas nécessairement ordonnés.

Exemple 1. 1. $(\mathcal{P}(E), \subseteq)$: les parties d'un ensemble E , munies de la relation d'inclusion, est un poset.

2. Pour $a, b \in \mathbb{R}$ et $n \in \mathbb{N}^*$, l'hypercube $[a, b]^n$ est un poset avec l'ordre partiel \leq_n défini par

$$\forall x, y \in [a, b]^n, x \leq_n y \iff x_i \leq y_i, \forall i \in \{1, \dots, n\}$$

où \leq est l'ordre usuel sur \mathbb{R} .

3. Si E est un ensemble et (\mathcal{V}, \leq) est un poset, alors l'ensemble \mathcal{F} des fonctions de E dans \mathcal{V} est un poset, pour l'ordre partiel $\leq_{\mathcal{F}}$ défini par

$$\forall f, g \in \mathcal{F}, f \leq_{\mathcal{F}} g \iff f(x) \leq g(x), \forall x \in E.$$

Ceci généralise le cas précédent, avec $E = \{1, \dots, n\}$ et $\mathcal{V} = [a, b]$.

4. L'ensemble \mathcal{S}_n des matrices semi-définies positives en dimension $n \in \mathbb{N}^*$ est un poset, avec l'ordre partiel \preceq défini par

$$\forall A, B \in \mathcal{S}_n, A \preceq B \iff \{x \in \mathbb{R}^n, x^T A x \leq 1\} \subseteq \{x \in \mathbb{R}^n, x^T B x \leq 1\}.$$

Il s'agit de l'ordre d'inclusion entre ellipsoïdes représentés par les matrices de \mathcal{S}_n .

L'existence d'un ordre partiel suffit à définir les notions de supremum (plus petit majorant) et infimum (plus grand minorant).

Définition 2 (Supremum, Infimum). Soit (E, \leq) un poset, et $A \subset E$.

Alors $x_A \in E$ est dit un supremum de A si

- c'est un majorant de A : $\forall a \in A, a \leq x_A$
- il est plus petit que tout majorant de A : $\forall z \in E, (\forall a \in A, a \leq z) \Rightarrow x_A \leq z$.

Si A admet un supremum, alors il est unique par anti-symétrie de \leq .

De même, $y_A \in E$ est dit un infimum de A si

- c'est un minorant de A : $\forall a \in A, y_A \leq a$
- il est plus grand que tout minorant de A : $\forall z \in E, (\forall a \in A, z \leq a) \Rightarrow z \leq y_A$.

Si A admet un infimum, alors il est unique.

Cela nous amène à la définition de la structure qui nous intéresse, les treillis complets.

Définition 3 (Treillis complet). Un treillis complet est un poset (\mathcal{L}, \leq) tel que toute partie $A \subseteq \mathcal{L}$ admet un supremum, noté $\bigvee A$, et un infimum, noté $\bigwedge A$.

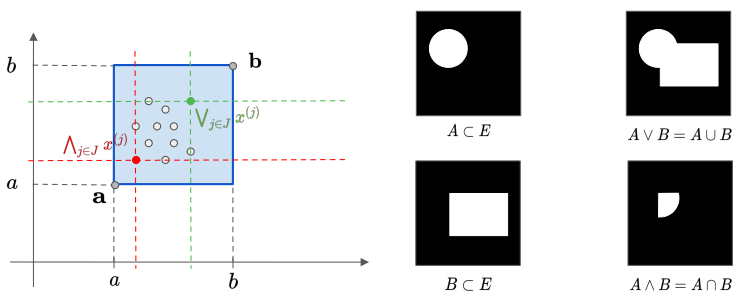


FIGURE 3.3 – Illustration des treillis complets de l'exemple 2 : $[a, b]^n$ et $\mathcal{P}(E)$.

- Exemple 2.**
1. $(\mathcal{P}(E), \subseteq)$ est un treillis complet, avec le supremum et l'infimum définis respectivement par l'union et l'intersection d'ensembles. En effet, pour toute famille $\mathcal{A} = (A_i)_{i \in I}$ de parties de E , il est direct que $\bigvee \mathcal{A} = \bigcup_{i \in I} A_i$ et $\bigwedge \mathcal{A} = \bigcap_{i \in I} A_i$ pour l'ordre \subseteq . En particulier, $\bigvee \mathcal{P}(E) = E$ et $\bigwedge \mathcal{P}(E) = \emptyset$. Ainsi, l'ensemble des images binaires de même support E est un treillis complet.
 2. Si (\mathcal{V}, \leq) est un treillis complet, alors le poset $(\mathcal{F}, \leq_{\mathcal{F}})$ des fonctions de E dans \mathcal{V} , est également un treillis complet, avec $\leq_{\mathcal{F}}$ l'ordre défini dans l'exemple 1. Pour toute famille $\mathcal{E} = (f_i)_{i \in I}$ d'éléments de \mathcal{F} , $\bigvee \mathcal{E}$ et $\bigwedge \mathcal{E}$ sont les fonctions qui à tout $x \in E$ associent respectivement $\bigvee_{i \in I} f_i(x)$ et $\bigwedge_{i \in I} f_i(x)$. Ainsi, l'ensemble des images à niveaux de gris de même support E , est un treillis complet. Pour un ensemble E fini de cardinal n , et des valeurs $\mathcal{V} = [a, b]$, ce treillis est équivalent à $([a, b]^n, \leq_n)$ introduit dans l'exemple précédent et illustré figure 3.3.
 3. (\mathcal{S}_n, \preceq) n'est pas un treillis complet, comme l'illustre la figure 3.4 pour $n = 2$ et les deux ellipses rouge et bleue, définies respectivement par deux matrices A et B . Si le supremum C de A et B existait, l'ellipse correspondante serait nécessairement isotrope, par unicité du supremum et par symétrie de l'exemple. Ce serait donc le cercle jaune, le plus petit possible. Or l'ellipse violette montre l'existence d'un autre majorant de A et B , qui n'est pas comparable avec C , contredisant la définition du supremum.

Les ensembles d'images binaires et à niveaux de gris forment des treillis complets car elles sont à valeur dans des treillis complets qui sont totalement ordonnés. Pour les images couleurs, la définition d'un ordre sur les valeurs prises par les pixels est moins naturel. En particulier, le supremum et l'infimum d'un ensemble de couleurs peuvent ne pas appar-

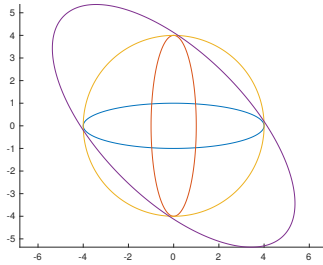


FIGURE 3.4 – Illustration de l'absence de supremum pour deux ellipses, montrant que $(\mathcal{S}_n, \subseteq)$ n'est pas un treillis complet, comme expliqué dans l'exemple 2.

tenir à cet ensemble, auquel cas ils sont qualifiés de *fausses couleurs* et sont perceptuellement gênants. Le problème de la définition d'ordres sur les couleurs est un des sujets de recherche en traitement d'images morphologiques, mais il ne sera pas abordé dans ce chapitre, où nous nous limitons aux images à niveaux de gris.

Comme annoncé plus tôt, nous pouvons maintenant définir les éléments de base des opérateurs morphologiques, les érosions et les dilations.

3.2.3 Dilatations et érosions, morphismes de treillis complets

Définition 4 (Dilatation, Erosion). Soient (\mathcal{L}, \leq) et (\mathcal{L}', \leq') deux treillis complets. On appelle dilatation de \mathcal{L} dans \mathcal{L}' tout opérateur $\delta : \mathcal{L} \rightarrow \mathcal{L}'$ qui commute avec le supremum, c'est à dire tel que

$$\forall (x_i)_{i \in I} \subseteq \mathcal{L}, \quad \delta \left(\bigvee_{i \in I} x_i \right) = \bigvee'_{i \in I} \delta(x_i). \tag{3.2}$$

On appelle érosion de \mathcal{L} dans \mathcal{L}' tout opérateur $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}'$ qui commute avec l'infimum,

$$\forall (x_i)_{i \in I} \subseteq \mathcal{L}, \quad \varepsilon \left(\bigwedge_{i \in I} x_i \right) = \bigwedge'_{i \in I} \varepsilon(x_i). \tag{3.3}$$

Un des exemples les plus importants de dilatation et érosion, à l'origine des développements de la Morphologie Mathématique en traitement d'images, concerne le treillis des parties d'un ensemble. Comme on va le voir, il éclaire ce qu'on entend par *analyse par critères de forme et de taille*.

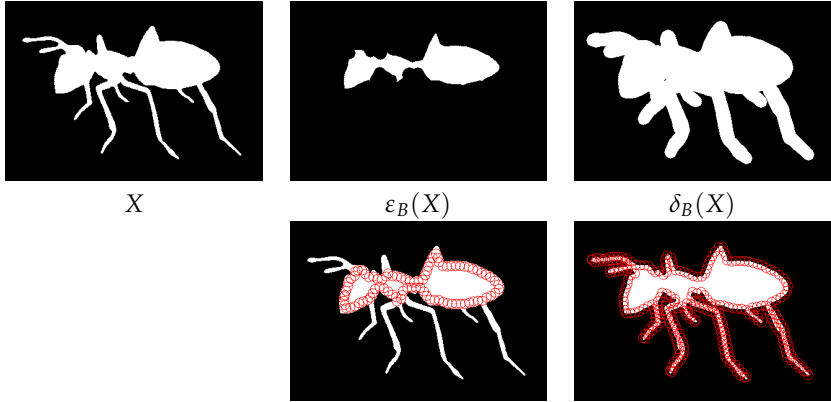


FIGURE 3.5 – Érosion et dilatation d'un ensemble $X \subset \mathbb{Z}^2$, en blanc, par un disque B , comme défini dans l'exemple 3. On voit sur la deuxième ligne que les translations de B déterminantes sont celles situées sur les bords.

Exemple 3. Soient $E = \mathbb{Z}^n$, et $B \subseteq \mathbb{Z}^n$ non vide. Pour tous $A \subseteq E$ et $x \in E$ on notera $A_x := \{x + a, a \in A\}$ et $\check{A} = \{-a, a \in A\}$. On définit alors $\delta_B : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ par

$$\forall X \in \mathcal{P}(E), \quad \delta_B(X) = \bigcup_{x \in X} B_x \quad (3.4)$$

et $\varepsilon_B : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ par

$$\forall X \in \mathcal{P}(E), \quad \varepsilon_B(X) = \{x \in E, B_x \subseteq X\}. \quad (3.5)$$

On vérifie aisément que $\delta_B(\cup_{i \in I} X_i) = \cup_{i \in I} \delta_B(X_i)$ et $\varepsilon_B(\cap_{i \in I} X_i) = \cap_{i \in I} \varepsilon_B(X_i)$, impliquant que δ_B et ε_B sont bien une dilatation et une érosion respectivement, sur le treillis complet $(\mathcal{P}(E), \subseteq)$.

Notons qu'il existe pour ces deux opérateurs des écritures alternatives, qui pourront être utiles par la suite. En particulier, on peut montrer que

$$\delta_B(X) = \bigcup_{b \in B} X_b = \{y \in E, \check{B}_y \cap X \neq \emptyset\} \quad (3.6)$$

et

$$\varepsilon_B(X) = \bigcap_{b \in \check{B}} X_b. \quad (3.7)$$

Le critère de forme et de taille évoqué ci-dessus est contenu dans l'ensemble B , paramètre des opérateurs δ_B et ε_B , et qu'on appellera dans

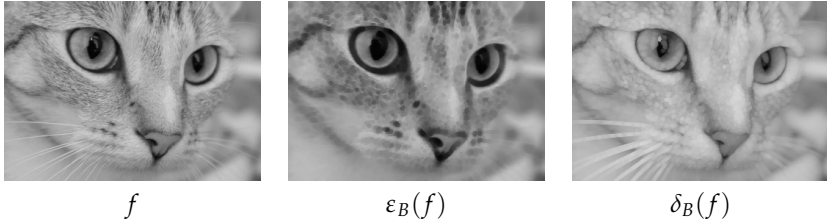


FIGURE 3.6 – Érosion et dilatation d'une fonction f par un disque B , comme défini dans l'exemple 4.

la suite *élément structurant*. On peut déjà constater qu'un ensemble X dans lequel aucun B_x ne serait contenu, quel que soit $x \in E$, est détruit par l'érosion : $\varepsilon_B(X) = \emptyset$. C'est le cas par exemple des pattes de la fourmi dans la figure 3.5, où $E = \mathbb{Z}^2$, X est l'ensemble des points constituant la fourmi et B est un disque fermé centré sur l'origine, de rayon supérieur à l'épaisseur des pattes. L'érosion peut donc être vue comme un test à l'aide de la forme B dont l'issue donne une information sur celle de X . Notons toutefois que même les parties qui contiennent des B_x sont modifiées - ici, le corps de la fourmi. Pour un filtrage plus précis ne modifiant pas les parties répondant au critère de forme et taille, il faut s'intéresser aux filtres morphologiques que sont les *ouvertures* et les *fermetures*, introduites dans la prochaine section. Avant cela, voyons l'exemple canonique de dilatation et érosion sur les images à niveaux de gris.

Exemple 4. Soient $E = \mathbb{Z}^n$, \mathcal{V} un treillis complet, et $B \subseteq \mathbb{Z}^n$ non vide. On note \mathcal{F} l'ensemble des fonctions de E dans \mathcal{V} .

On définit alors $\delta_B : \mathcal{F} \rightarrow \mathcal{F}$ par

$$\forall f \in \mathcal{F}, \forall x \in E, \quad \delta_B(f)(x) = \bigvee_{b \in B} f(x - b) = \bigvee_{y \in \check{B}_x} f(y), \quad (3.8)$$

et $\varepsilon_B : \mathcal{F} \rightarrow \mathcal{F}$ par

$$\forall f \in \mathcal{F}, \forall x \in E, \quad \varepsilon_B(f)(x) = \bigwedge_{b \in B} f(x + b) = \bigwedge_{y \in B_x} f(y). \quad (3.9)$$

Pour les images à niveaux de gris, la dilatation et l'érosion reviennent donc à affecter à chaque pixel x respectivement le maximum et le minimum de la fonction sur un ensemble dépendant de x , le plus souvent un voisinage, au lieu de calculer une moyenne pondérée comme dans le cas du filtre linéaire présenté plus haut (voir figure 3.1). Ceci a pour effet de propager les parties claires, dans le cas de la

dilatation, et les parties sombres dans le cas de l'érosion, comme sur la figure 3.6. Là aussi, dans le cas de l'érosion par un disque, on observe que les objets qui ne répondent pas au critère de forme, comme les moustaches du chat, disparaissent.

Notons que les définitions sur les fonctions (ou images à niveaux de gris) généralisent celles sur les ensembles (ou images binaires), en considérant les images binaires comme des fonctions indicatrices : en effet, $\delta_B(\mathbb{1}_X) = \mathbb{1}_{\delta_B(X)}$ et $\varepsilon_B(\mathbb{1}_X) = \mathbb{1}_{\varepsilon_B(X)}$, où l'on utilise les mêmes notations pour les fonctions et les ensembles.

Enfin, remarquons également que $\delta_B(f)(x) = \bigvee_{y \in E} f(y) + \log(\mathbb{1}_B(x - y))$ et $\varepsilon_B(f)(x) = \bigwedge_{y \in E} f(y) - \log(\mathbb{1}_B(y - x))$, ce qui, d'une part, laisse voir une généralisation à des dilatations et érosions par des fonctions structurantes (ici $\log(\mathbb{1}_B)$) et, d'autre part, met en évidence le parallèle avec la convolution linéaire - on peut parler ici de convolutions max-plus et min-plus. Dans ce chapitre, nous nous limiterons au cas d'opérateurs dépendants d'éléments structurants B .

3.2.4 Ouvertures et fermetures, projecteurs non linéaires

Définition 5 (Ouverture, Fermeture.). Soit (\mathcal{L}, \leq) un treillis complet. Une application $\gamma : \mathcal{L} \rightarrow \mathcal{L}$ est une ouverture si et seulement si elle est

- Croissante : $\forall x, y \in \mathcal{L}, x \leq y \Rightarrow \gamma(x) \leq \gamma(y)$
- Idempotente : $\gamma \circ \gamma = \gamma$
- Anti-extensive : $\forall x \in \mathcal{L}, \gamma(x) \leq x$.

Une application $\varphi : \mathcal{L} \rightarrow \mathcal{L}$ est une fermeture si et seulement si elle est

- Croissante : $\forall x, y \in \mathcal{L}, x \leq y \Rightarrow \varphi(x) \leq \varphi(y)$
- Idempotente : $\varphi \circ \varphi = \varphi$
- Extensive : $\forall x \in \mathcal{L}, x \leq \varphi(x)$.

Les premiers exemples d'ouverture et fermeture se construisent justement à l'aide des érosions et dilatations des exemples précédents.

Exemple 5. En reprenant les notations des exemples 3 et 4, soient les opérateurs γ_B et φ_B définis par

$$\gamma_B := \delta_B \circ \varepsilon_B \quad \text{et} \quad \varphi_B := \varepsilon_B \circ \delta_B. \quad (3.10)$$

Alors on vérifie que γ_B est une ouverture et φ_B est une fermeture. On peut s'en convaincre en particulier dans le cas des ensembles. D'abord, la croissance de ces opérateurs est évidente, par composition d'opérateurs croissants. En suite, l'anti-extensivité de γ_B apparaît en remarquant que $\gamma_B(X)$ est une union d'ensembles inclus dans X :

$$\forall X \in \mathcal{P}(E), \quad \gamma_B(X) = \delta_B \circ \varepsilon_B(X) = \bigcup_{y \in \varepsilon_B(X)} B_y = \bigcup_{y \in E, B_y \subseteq X} B_y. \quad (3.11)$$

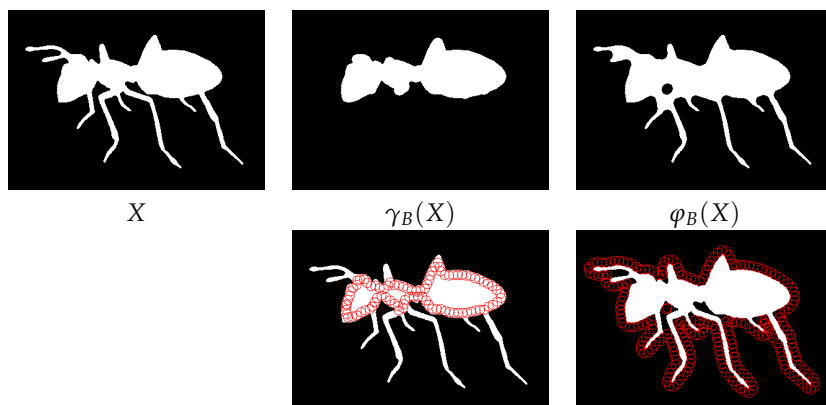


FIGURE 3.7 – Ouverture et fermeture d'un ensemble $X \subset \mathbb{Z}^2$, en blanc, par un disque B , comme défini dans l'exemple 5. La ligne du bas illustre l'interprétation des équations (3.16) et (3.17), et, pour la fermeture φ_B fait aussi appel à la notion de la dualité (voir Section 3.2.5).

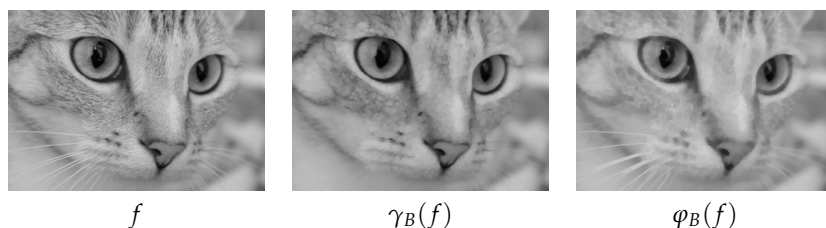


FIGURE 3.8 – Ouverture et fermeture d'une fonction f par un disque B , comme défini dans l'exemple 5.

D'où $\gamma_B(X) \subseteq X$. On obtient l'extensivité de φ_B de façon tout aussi directe, en développant

$$\begin{aligned} \forall X \in \mathcal{P}(E), \quad \varphi_B(X) = \varepsilon_B \circ \delta_B(X) &= \{y \in E, B_y \subseteq \delta_B(X)\} \\ &= \{y \in E, B_y \subseteq \bigcup_{x \in X} B_x\} \end{aligned} \quad (3.12)$$

d'où il apparaît $X \subseteq \varphi_B(X)$.

L'idempotence s'obtient alors en montrant que si $Y = \gamma_B(X)$ alors $Y \subseteq \gamma_B(Y)$ et donc $Y = \gamma_B(Y)$; et de même, si $Y = \varphi_B(X)$ alors $\varphi_B(Y) \subseteq Y$ et donc $Y = \varphi_B(Y)$.

Les démonstrations dans l'exemple qui précède demandent des manipulations d'ensembles un peu fastidieuses, et elles seraient à refaire pour le cas des fonctions. Comme on va le voir tout de suite, une approche plus générale est possible dès lors qu'on identifie un lien particulier, appelé *adjonction*, entre l'érosion et la dilatation qui sont composées.

Définition 6 (Adjonction). Soient (\mathcal{L}, \leq) et (\mathcal{L}', \leq') deux treillis complets, et $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}'$ et $\delta : \mathcal{L}' \rightarrow \mathcal{L}$ deux opérateurs. Alors le couple (ε, δ) est appelé adjonction si et seulement si pour tout $x \in \mathcal{L}'$ et tout $y \in \mathcal{L}$,

$$\delta(x) \leq y \iff x \leq' \varepsilon(y). \quad (3.13)$$

Les opérateurs ε et δ sont dits adjoints l'un de l'autre.

L'adjonction a les propriétés suivantes, qui éclairent l'exemple 5.

Proposition 1 (Propriétés de l'adjonction). Soient (\mathcal{L}, \leq) et (\mathcal{L}', \leq') deux treillis complets et (ε, δ) une adjonction avec $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}'$ et $\delta : \mathcal{L}' \rightarrow \mathcal{L}$. Alors,

1. ε est une érosion et δ est une dilatation

2.

$$\forall x \in \mathcal{L}, \quad \varepsilon(x) = \bigvee' \{y \in \mathcal{L}', \delta(y) \leq x\} \quad (3.14)$$

3.

$$\forall y \in \mathcal{L}', \quad \delta(y) = \bigwedge \{x \in \mathcal{L}, y \leq \varepsilon(x)\} \quad (3.15)$$

4. $\varepsilon\delta\varepsilon = \varepsilon$ et $\delta\varepsilon\delta = \delta$

5. $\gamma := \delta\varepsilon$ est une ouverture et $\varphi := \varepsilon\delta$ est une fermeture.

Pour démontrer ces propriétés, on peut commencer par montrer l'extensivité de $\varphi = \varepsilon\delta$ et l'anti-extensivité de $\gamma = \delta\varepsilon$, qui sont des conséquences directes de l'équation (3.13). La propriété 4 devient alors évidente (en écrivant $\varepsilon\delta\varepsilon = \varepsilon\gamma = \varphi\varepsilon$ et $\delta\varepsilon\delta = \delta\varphi = \gamma\delta$), et donne l'idempotence de γ et φ . Avec la croissance de ε et δ , qui est aussi une conséquence directe de

(3.13), on obtient la propriété 5. Enfin, les propriétés 1 à 3 se déduisent aussi de (3.13) et de l'extensivité et anti-extensivité de φ et γ respectivement.

La proposition 1 affirme qu'à toute adjonction correspond une dilatation et une érosion. Réciproquement, toute dilatation δ admet une unique érosion adjointe ε , définie par l'équation (3.14), et toute érosion ε admet une unique dilatation adjointe δ définie par (3.15). Il y a donc bijection entre l'ensemble des adjonctions sur $(\mathcal{L}, \mathcal{L}')$, l'ensemble des érosions de \mathcal{L} dans \mathcal{L}' et l'ensemble des dilatations de \mathcal{L}' dans \mathcal{L} .

Avec les définitions de l'exemple 5, on vérifie que

$$\forall X, Y \in \mathcal{P}(E), \quad \delta_B(X) \subseteq Y \iff X \subseteq \varepsilon_B(Y),$$

et

$$\forall f, g \in \mathcal{F}, \quad \delta_B(f) \leq g \iff f \leq \varepsilon_B(g),$$

ce qui prouve, dans les deux cas, que $(\varepsilon_B, \delta_B)$ est une adjonction, γ_B une ouverture et φ_B une fermeture. Donnons maintenant une interprétation de l'effet de γ_B , en s'intéressant au cas des ensembles. En utilisant l'expression (3.14) et le fait que la dilatation commute avec le supremum (ici l'union), on obtient

$$\begin{aligned} \gamma_B(X) = \delta_B \varepsilon_B(X) &= \bigcup \{ \delta_B(Y), Y \in \mathcal{P}(E), \delta_B(Y) \subseteq X \} \\ &= \bigcup \{ Y \in \delta_B(\mathcal{P}(E)), Y \subseteq X \}. \end{aligned} \quad (3.16)$$

Ainsi, $\gamma_B(X)$ est un *dilaté*, c'est à dire une union de translatés de B , mais pas n'importe lequel : parmi tous ceux qui sont contenus dans X , c'est le plus grand (voir Figure 3.7). Il s'agit donc de la meilleure approximation "inférieure" de X par un dilaté, puisque tout $Y \in \delta_B(\mathcal{P}(E))$ tel que $Y \subseteq X$, est nécessairement plus "loin" de X que ne l'est $\gamma_B(X) : Y \subseteq \gamma_B(X) \subseteq X$. En ce sens, γ_B peut être vu comme un *projecteur non linéaire sur l'ensemble des dilatés*, et son idempotence paraît alors évidente. En pratique, appliquer γ_B à X laisse intactes les parties qui sont une union de translatés de B (le corps de la fourmi dans la figure 3.7), et détruit tout ce qui s'en éloigne (les pattes de la fourmi, trop fines pour l'élément structurant B). De façon analogue, $\varphi_B(X)$ est la meilleure approximation "supérieure" de X par un érodé,

$$\varphi_B(X) = \bigcap \{ Y \in \varepsilon_B(\mathcal{P}(E)), X \subseteq Y \}. \quad (3.17)$$

Bien entendu, tout ceci peut s'écrire dans le cas des fonctions et, plus généralement, d'une adjonction quelconque. En reprenant les notations de la proposition 1,

$$\forall x \in \mathcal{L}, \quad \gamma(x) = \delta \varepsilon(x) = \bigvee \{ y \in \delta(\mathcal{L}'), y \leq x \} \quad (3.18)$$

et

$$\forall y \in \mathcal{L}', \quad \varphi(y) = \varepsilon\delta(y) = \bigwedge \{x \in \varepsilon(\mathcal{L}), y \leq x\}. \quad (3.19)$$

Ici γ projette sur $\delta(\mathcal{L})$, qui est l'ensemble des points fixes de γ . On appellera cet ensemble *invariants* de γ . De même, l'ensemble des invariants de φ est $\varepsilon(\mathcal{L}')$. On va voir que le principe de projection sur les invariants est vrai pour toutes les ouvertures et fermetures, pas seulement celles construites à partir d'une adjonction.

Définition 7 (Invariants). Soit $\psi : \mathcal{L} \rightarrow \mathcal{L}$ un opérateur idempotent sur un treillis complet \mathcal{L} . On appelle invariants de ψ , noté \mathcal{I}_ψ , l'ensemble

$$\mathcal{I}_\psi := \psi(\mathcal{L}) = \{x \in \mathcal{L}, \psi(x) = x\}. \quad (3.20)$$

Sans surprise, les ouvertures et fermetures sont des projecteurs sur leurs invariants, au sens explicité plus haut.

Proposition 2. Soit $\psi : \mathcal{L} \rightarrow \mathcal{L}$ un opérateur sur le treillis complet \mathcal{L} . Alors ψ est une ouverture si et seulement si il existe un ensemble $\mathcal{J} \subseteq \mathcal{L}$ stable par supremum, i.e.

$$\forall \mathcal{F} \subseteq \mathcal{J}, \quad \bigvee \mathcal{F} \in \mathcal{J}, \quad (3.21)$$

tel que

$$\forall x \in \mathcal{L}, \quad \psi(x) = \bigvee \{y \in \mathcal{J}, y \leq x\}. \quad (3.22)$$

Dans ce cas $\mathcal{I}_\psi = \mathcal{J}$.

On a bien sûr le résultat analogue pour les fermetures en remplaçant supremum par infimum et $y \leq x$ par $x \leq y$ dans (3.22).

Dans tout ce qui précède, on voit qu'échanger supremum et infimum fait apparaître des analogies entre érosions et dilatations, ainsi qu'entre ouvertures et fermetures. Cette dualité est formalisée grâce aux transformations qui inversent l'ordre d'un treillis à un autre, que l'on appelle les *inversions*.

3.2.5 Dualité

Définition 8 (Inversion, involution). Soient (\mathcal{L}, \leq) et (\mathcal{L}', \leq') deux treillis complets. On appelle inversion de \mathcal{L} dans \mathcal{L}' toute fonction $n : \mathcal{L} \rightarrow \mathcal{L}'$ bijective décroissante, i.e. telle que

$$\forall x, y \in \mathcal{L}, \quad x \leq y \Rightarrow n(y) \leq' n(x). \quad (3.23)$$

Si $\mathcal{L}' = \mathcal{L}$ et $n = n^{-1}$, on parle d'involution.

Exemple 6. Dans le cas où $\mathcal{L} = \mathcal{L}' = \mathcal{P}(E)$, l'inversion la plus utilisée est une involution, à savoir la complémentation d'ensemble $n : X \mapsto X^c := E \setminus X$.

Dans le cas où $\mathcal{L} = \mathcal{L}' = \mathcal{F}$, l'ensemble des fonctions de E dans $\mathcal{V} = \{0, \dots, M\}$, l'inversion classique est l'involution $n : f \mapsto M - f$.

Proposition 3. Soit $n : \mathcal{L}' \rightarrow \mathcal{L}$ une inversion. Alors

- $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}'$ est une érosion si et seulement si $\delta = n \circ \varepsilon \circ n$ est une dilatation⁴ de \mathcal{L}' dans \mathcal{L}
- $\gamma : \mathcal{L} \rightarrow \mathcal{L}$ est une ouverture si et seulement si $\varphi = n^{-1} \circ \gamma \circ n$ est une fermeture sur \mathcal{L}' .

On parle alors d'érosion et dilatation (resp. ouverture et fermeture) duales pour l'inversion n .

Alors que l'adjonction associe systématiquement une érosion à une dilatation, la proposition précédente indique qu'un autre appariement, par dualité, est possible dès lors qu'une inversion existe. On va voir dans l'exemple qui suit que ces deux appariements, par adjonction et par dualité, sont en général différents.

Exemple 7. Reprenons les cas où $\mathcal{L} = \mathcal{L}' = \mathcal{P}(E)$, avec l'inversion de l'exemple 6 (complémentation). Quelle est alors l'érosion duale $n \circ \delta_B \circ n$ de la dilatation δ_B définie dans l'exemple 3 ? À l'aide des écritures alternatives des équations (3.6) et (3.7), on obtient

$$(\delta_B(X^c))^c = \left(\bigcup_{b \in B} X_b^c \right)^c = \bigcap_{b \in B} X_b = \varepsilon_{\check{B}}(X). \quad (3.24)$$

Ainsi l'érosion duale de δ_B est $\varepsilon_{\check{B}}$, l'érosion par l'élément structurant symétrique \check{B} , différente de l'érosion adjointe ε_B . Comme $\check{\check{B}} = B$, on déduit que la dilatation duale de ε_B est $\delta_{\check{B}}$. Enfin, ici $n = n^{-1}$, d'où la fermeture duale de γ_B est $n \circ \delta_B \varepsilon_B \circ n = n \circ \delta_B \circ n \circ n \circ \varepsilon_B \circ n = \varepsilon_{\check{B}} \delta_{\check{B}} = \varphi_{\check{B}}$ et l'ouverture duale de φ_B est $\gamma_{\check{B}}$. Dans la figure 3.7, B est symétrique, donc $B = \check{B}$. Ainsi, $\varphi_B(X)$ étant le plus petit érodé contenant X , c'est par dualité le plus petit ensemble contenant X et dont le complémentaire est un dilaté, i.e. une union de translatés de B . C'est ce qui est illustré sur l'image de droite de la ligne du bas. La conséquence est que toute partie noire dans laquelle aucun disque B_x n'est inclus, est "fermée". Les autres sont inchangées.

Notons qu'on obtient exactement les mêmes liens de dualité dans le cas des fonctions à valeur dans $\{0, \dots, M\}$ et de l'involution $n : f \mapsto M - f$. L'érosion duale de δ_B est encore $\varepsilon_{\check{B}}$, et ainsi de suite.

4. De façon plus générale, on pourrait prendre $\delta = p \circ \varepsilon \circ n$, où $p : \mathcal{L}' \rightarrow \mathcal{L}$ serait une autre inversion, mais ce n'est pas utilisé en pratique.

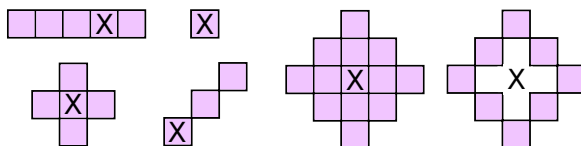


FIGURE 3.9 – Exemples d'éléments structurants. L'origine est modélisée par la croix.

3.2.6 Éléments structurants et connexité

Élément structurant Comme évoqué en début de chapitre, une des idées de base de la morphologie mathématique est de comparer ce qu'on veut traiter avec un objet de géométrie connue : l'*élément structurant*. Il s'agit d'un sous-ensemble de \mathbb{Z}^n ($n = 2$ ou 3). Les éléments structurants peuvent avoir des formes et des tailles variées, contenir l'origine ou non (ce qui est plus rare en pratique), comme on le voit sur la figure 3.9. Ils font partie des piliers de la morphologie mathématique dans la mesure ou le choix de l'élément structurant détermine la façon d'analyser les formes dans une image.

Connexité Les pixels de E peuvent être vus comme les noeuds d'un graphe, et un élément structurant B définit ainsi une relation d'adjacence entre pixels, l'ensemble des *voisins* d'un pixel $x \in E$ étant $B_x = x + B = \{x + b, b \in B\}$. Lorsque cette relation d'adjacence est symétrique, c'est à dire lorsque B l'est, on peut dire que deux pixels voisins sont *connectés*. On parle alors de relation de *connexité*, qui est une relation d'équivalence.

En pratique on caractérise souvent la connexité par le nombre de voisins qu'elle induit pour un pixel. En deux dimensions, on va parler de connexité 4 et 8. En connexité 4, les pixels sont voisins s'ils ont une arête en commun ; en connexité 8, s'ils ont une arête ou un angle en commun, comme illustré par les figures 3.10 (a) et (c). En trois dimensions, on considère les connexités 6, 18 et 26. La connexité 6 correspond aux voxels liés par leurs *faces*, la 18 par leurs *faces* ou leurs *arêtes*, et la 26 est le *cube englobant* le voxel.

Dans un sous ensemble $X \subseteq E$, les classes d'équivalence induites par la relation de connexité sont appelées *composantes connexes*. Le choix de la connexité est très important car il va déterminer, pour une forme donnée, le nombre de ses composantes connexes comme illustré Figure 3.10.

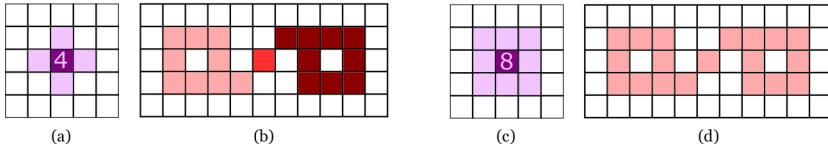


FIGURE 3.10 – Exemple de connexité 4 et 8. (a) Voisins du pixel central en connexité 4. (b) L'ensemble des pixels colorés contient trois composantes connexes en connexité 4. (c) Voisins du pixel central en connexité 8. (d) Le même ensemble que précédemment compte une seule composante connexe en connexité 8.

3.2.7 Invariance par translation et éléments structurants

Nous avons évoqué, en début de cette partie, l'équivalence entre filtrage linéaire invariant par translation et convolution par un noyau. Nous pouvons à présent préciser l'analogie avec les opérateurs morphologiques vus dans les exemples précédents. En particulier, on peut se convaincre qu'une dilatation $\delta : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ est de type δ_B pour un certain élément structurant $B \subseteq E$, si et seulement si elle est invariante par translation, c'est à dire $\delta(X_u) = \delta(X)_u$ pour tous $X \subseteq E$ et $u \in E$. Dans ce cas, $B = \delta(\{0\})$. L'élément structurant B joue ici le rôle du noyau représentant δ . Ceci est également vrai pour les érosions invariantes par translation.

De même, dans le cas des fonctions, les opérateurs δ_B et ε_B sont invariants par translations verticales, c'est à dire $\delta_B(\tau_u[f]) = \tau_u \circ \delta_B(f)$, et horizontales, $\delta_B(\lambda + f) = \lambda + \delta_B(f)$ ⁵. De plus, $\delta_B(\mathbb{1}_{\{0\}}) = \mathbb{1}_B$. On peut se convaincre que, réciproquement, ces trois conditions sont suffisantes pour qu'une dilatation sur \mathcal{F} soit du type δ_B .

3.3 Premières applications avec les résidus

Dans la section précédente, nous avons posé les bases de la morphologie mathématique, en particulier avec les premières définitions d'opérateurs très simples : érosions, dilatations, ouvertures et fermetures par des éléments structurants. Avant d'aller plus loin, nous montrons ici que de simples combinaisons de ces opérateurs par différences, suffisent déjà à la définition d'outils très utiles en traitement d'images. C'est le résultat des différences entre opérateurs qu'on appelle communément *résidus*.

5. Ici il faudrait s'assurer que $\lambda + f$ et $\lambda + \delta_B(f)$ sont encore dans le treillis \mathcal{V} , c'est pourquoi la condition exacte est : $\forall \lambda \leq 0, \delta_B((\lambda + f) \vee 0) = (\lambda + \delta_B(f)) \vee 0$



FIGURE 3.11 – Gradient morphologique d’une image à niveaux de gris.

3.3.1 Le gradient morphologique

Le gradient morphologique, tout comme la norme du gradient classique, permet de souligner les fortes transitions d’une image à niveaux de gris, entre régions claires et régions sombres. Ces transitions incluent notamment les contours des objets, d’où leur intérêt. Le gradient morphologique d’une image f est défini par

$$G_B(f) = \delta_B(f) - \varepsilon_B(f) \quad (3.25)$$

où B est un élément structurant symétrique, contenant l’origine, et le plus souvent “élémentaire”, c’est à dire ne contenant qu’un pixel et ses pixels adjacents dans la grille - typiquement, en deux dimensions, la croix de la connexité 4 ou le carré de la connexité 8. La dilatation $\delta_B(f)$ et l’érosion $\varepsilon_B(f)$ sont alors respectivement le maximum et le minimum de f sur un petit voisinage B . Leur différence est donc faible sur une région à peu près constante, mais forte à l’endroit des transitions. Ceci est visible sur la figure 3.11.

3.3.2 Le Top-Hat

Une transformation très utilisée en morphologie mathématique est le « chapeau haut-de-forme » ou « top-hat ». Il s’agit d’une opération qui extrait des petits éléments au sein des images, en calculant la différence entre l’image d’entrée et son ouverture, ou entre la fermeture de l’image et l’image d’entrée (l’image la plus sombre est soustraite à la plus claire). La figure 3.12 illustre cette procédure sur un signal 1D. Il existe deux types de top-hat : le top-hat blanc (T_w) permet de récupérer les petites structures les plus claires tandis que le top-hat noir (T_b) permet de récupérer les petites structures les plus sombres. Si γ et φ sont respectivement une ouverture et

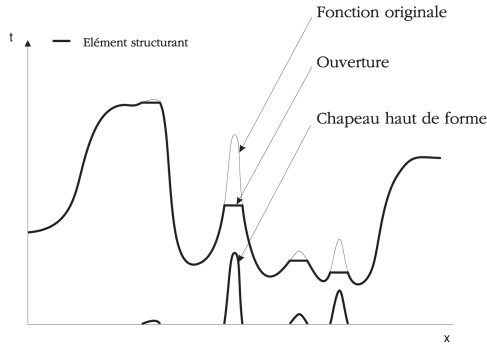


FIGURE 3.12 – Exemple de top-hat sur un signal 1D (figure reprise de [21]).

une fermeture, leurs top-hats correspondants sont définis par

$$T_w(f) = f - \gamma(f) \quad (3.26)$$

$$T_b(f) = \varphi(f) - f \quad (3.27)$$

où f est une image, binaire ou non. Dans le cas binaire, la soustraction arithmétique est équivalente à la soustraction d'ensemble (soit encore, l'intersection avec le complémentaire).

Si l'ouverture et la fermeture sont paramétrés par un élément structurant B , le top-hat sélectionne les éléments qui sont plus petits que B , et plus clairs que le reste de l'image dans le cas du top-hat blanc, ou plus sombres dans le cas du top-hat noir. Le choix de la taille de l'élément structurant est donc fondamental : s'il est trop petit, on manque des éléments que l'on cherche à récupérer, et si il est trop grand, on en récupère trop. L'étude granulométrique (voir Section 3.5.3) ou la connaissance à priori de la taille des objets recherchés sont une grande aide pour l'optimisation de cette taille. De façon duale à la sélection de petits objets, le top-hat permet par exemple d'éliminer les variations lentes de l'image, tel qu'un fond non uniforme, comme illustré par la Figure 3.13. Sur le plan algébrique, le top-hat est non croissant, mais le top-hat blanc est idempotent (contrairement au top-hat noir). Une présentation plus détaillée du top-hat et accompagnée de nombreux exemples d'application est proposée dans l'ouvrage de Pierre Soille [28].

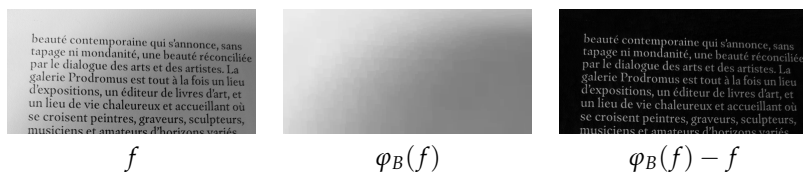


FIGURE 3.13 – Suppression de fond non-uniforme par top-hat noir, à l’aide d’un élément structurant carré de taille supérieure à celle des caractères mais assez petite pour ne pas capturer de variations de luminosité.



FIGURE 3.14 – Une image binaire et sa fonction distance.

3.3.3 La fonction distance

La fonction distance est un incontournable de l’analyse d’images binaires. Elle associe à chaque pixel d’un ensemble X , c’est à dire à chaque pixel blanc de l’image, sa distance au complémentaire de X , c’est à dire au pixel noir le plus proche. Cette distance est entendue au sens de la longueur du plus court chemin entre les deux pixels, pour la relation de connexité définie par un élément structurant B , symétrique et contenant l’origine (voir Section 3.2.6). La force de la fonction distance est qu’elle traduit en relief (c’est à dire, avec de nouvelles valeurs non binaires) une information sur la forme d’une composante binaire. En particulier, une composante connexe constituée de deux formes convexes qui se touchent, se traduit par deux maxima de la fonction distance, marquant les deux parties convexes. C’est ce qui apparaît sur les figures 3.14 et 3.15. Le “col” qui sépare ces deux maxima peut être pris comme séparatrice de ces deux composantes, dans l’algorithme du watershed (voir Section 3.6.1).

La fonction distance peut se calculer très simplement grâce à une suite d’érosions par B . En effet, suite à une première érosion de X , le résidu $X \setminus \varepsilon_B(X)$, soit l’ensemble des pixels ayant changé de valeur (de blanc à noir), est exactement l’ensemble des pixels à distance un du complémentaire de

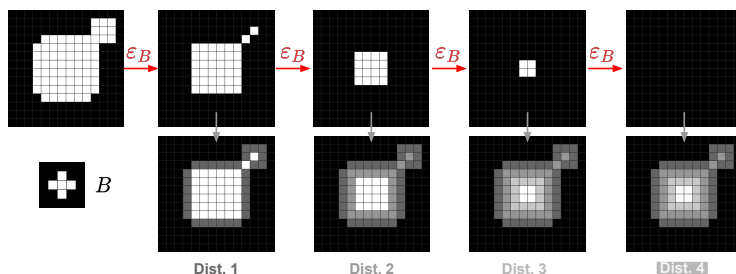


FIGURE 3.15 – Calcul de la fonction distance à partir d'érosions successives.

X. Ceux devenant noir à l'érosion successive sont à distance deux, et ainsi de suite. Cet algorithme est illustré Figure 3.15.

Bien d'autres représentations très puissantes peuvent être calculées à l'aide d'opérateurs simples et de leurs résidus, telles que le squelette morphologique, pour n'en citer qu'une. Nous laissons les personnes intéressées consulter les ouvrages plus complets déjà cités, pour nous concentrer à présent sur une grande classe d'opérateurs morphologiques, dits *géodésiques*.

3.4 Opérateurs géodésiques

Un des grands avantages du traitement morphologique d'images est le respect des contours des objets. Pourtant, nous avons vu que l'utilisation d'éléments structurants tend à modifier les frontières : la forme de l'élément structurant devient visible. Par exemple, une ouverture par un disque va arrondir les points anguleux de l'objet filtré. Les opérateurs géodésiques permettent justement de reconstruire les contours initiaux des objets qui n'ont pas été complètement détruits.

3.4.1 Le cas des ensembles

On suppose ici que $E = \mathbb{Z}^2$, et que $B \subseteq E$ est un élément structurant symétrique ($\check{B} = B$) et contenant l'origine ($0 \in B$).

Définition 9 (Dilatation géodésique sur les ensembles). Soit $X \subseteq E$. La dilatation géodésique par B dans X , notée $\delta_{B,X}$ est définie par

$$\delta_{B,X} : \begin{cases} \mathcal{P}(X) & \rightarrow \mathcal{P}(X) \\ Y & \mapsto \delta_B(Y) \cap X \end{cases} \quad (3.28)$$

où δ_B est la dilatation sur $\mathcal{P}(E)$ par B , comme définie par l'équation (3.4).

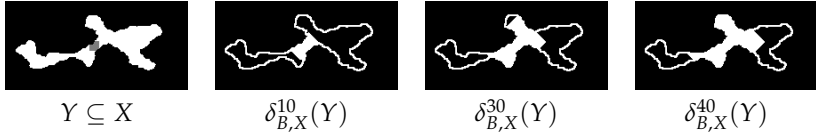


FIGURE 3.16 – Plusieurs itérations de la dilatation géodésique de Y (en gris dans l’image de gauche) dans X (en blanc), avec un élément structurant B en forme de croix (connexité 4).

La dilatation géodésique consiste à dilater “à l’intérieur” de X : si $\delta_B(Y)$ dépasse les frontières de X , on ne garde que le résultat qui intersecte X . L’ensemble Y souvent appelé *marqueur*, et X *masque* ou *référence*.

Il s’agit bien d’une dilatation sur le treillis $\mathcal{P}(X)$. Elle admet donc une érosion adjointe $\varepsilon_{B,X}$ qui est égale à son érosion duale pour la complémentation dans X , puisque B est symétrique.

Définition 10 (Erosion géodésique sur les ensembles). *Pour $X \subseteq E$, on définit l’érosion géodésique par B dans X , notée $\varepsilon_{B,X}$ comme l’érosion adjointe de $\delta_{B,X}$ (Equation (3.28)) :*

$$\varepsilon_{B,X} : \begin{cases} \mathcal{P}(X) & \rightarrow \mathcal{P}(X) \\ Y & \mapsto \{y \in X, \delta_{B,X}(\{y\}) \subseteq Y\} = \varepsilon_B(Y \cup X^c) \cap X \end{cases} \quad (3.29)$$

où ε_B est l’érosion sur $\mathcal{P}(E)$ par B , comme définie par l’équation (3.5). Comme B est symétrique, on a aussi

$$\varepsilon_{B,X}(Y) = (\delta_{B,X}(Y^{c_X}))^{c_X} \quad (3.30)$$

où la complémentation est prise dans X , c’est à dire $Z^{c_X} = Z^c \cap X$.

Bien entendu, on peut itérer ces opérateurs. La dilatation géodésique itérée n fois, pour $n \in \mathbb{N}^*$, est ainsi

$$\delta_{B,X}^n(Y) = \delta_{B,X} \circ \dots \circ \delta_{B,X}(Y) = \delta_B(\delta_{B,X}^{n-1}(Y)) \cap X \quad (3.31)$$

avec $\delta_{B,X}^0 = id$. Si l’on interprète B comme l’encodage d’une relation d’adjacence (c’est à dire B_x comme l’ensemble des voisins de $x \in E$), alors $\delta_{B,X}^n(Y)$ est l’ensemble des points de X que l’on peut atteindre depuis Y via un chemin *dans* X de longueur au plus n (voir Figure 3.16). C’est cette contrainte d’inclusion des chemins dans X qui justifie l’adjectif *géodésique*.

Comme B contient l’origine, δ_B est extensive et $\delta_{B,X}$ aussi. Ainsi, pour tout $Y \subseteq X$, la suite $(\delta_{B,X}^n(Y))_{n \in \mathbb{N}^*}$ est croissante et sa limite est appelée *reconstruction géodésique de X* .

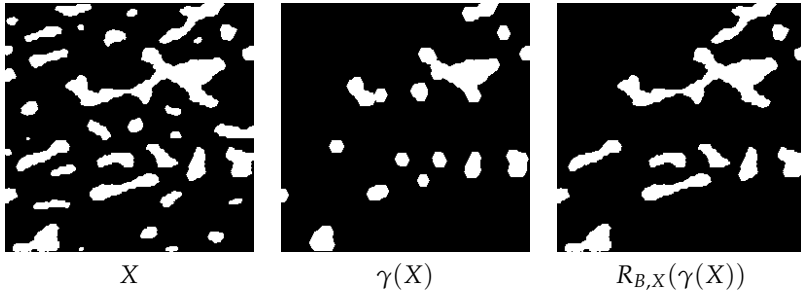


FIGURE 3.17 – Ouverture par un élément structurant hexagonal suivi de reconstruction géodésique en connexité 4 (ouverture par reconstruction). Les composantes connexes de X qui intersectent $\gamma(X)$ sont entièrement reconstruites, tandis que les autres sont entièrement supprimées.

Définition 11 (Reconstruction géodésique). Soient $X \subseteq E$ et $Y \subseteq X$. La reconstruction géodésique de X par dilatation à partir de Y est définie par

$$R_{B,X}(Y) := \bigcup_{n \geq 1} \delta_{B,X}^n(Y). \quad (3.32)$$

La propriété essentielle d'une telle reconstruction est la suivante : toute composante connexe de X (au sens de la connexité induite par B) qui est intersectée par Y , est entièrement reconstruite. À l'inverse, toute composante connexe de X non intersectée par Y est "détruite", au sens où elle est absente de la reconstruction (voir Figure 3.17). Formellement, pour C une composante connexe de X ,

$$\begin{cases} Y \cap C \neq \emptyset \Rightarrow C \subseteq R_{B,X}(Y) \\ Y \cap C = \emptyset \Rightarrow C \cap R_{B,X}(Y) = \emptyset. \end{cases} \quad (3.33)$$

Ainsi, $R_{B,X}(Y)$ est la réunion des composantes connexes de X qui intersectent Y . Cette propriété est très utilisée pour définir des ouvertures moins "destructrices" que celles vues précédemment : les ouvertures par reconstruction, comme celle illustrée par la figure 3.17.

Définition 12 (Ouverture par reconstruction). On appelle ouverture par reconstruction sur $\mathcal{P}(E)$ tout opérateur du type

$$X \in \mathcal{P}(E) \mapsto R_{B,X}(\gamma(X)) \quad (3.34)$$

où γ est une ouverture sur $\mathcal{P}(E)$.

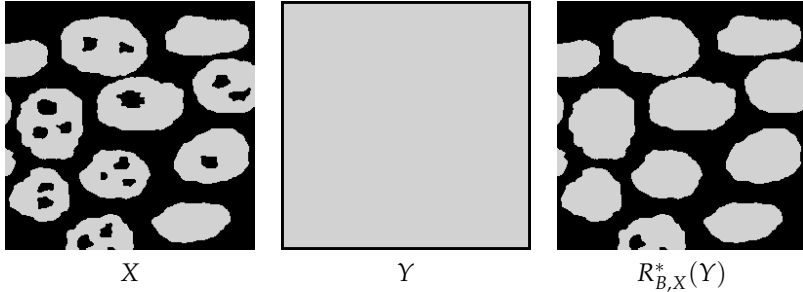


FIGURE 3.18 – Reconstruction duale de X à partir de Y en connexité 8 (élément structurant B carré 3×3). Ici Y est l'ensemble des pixels de l'image privé de ses bords. L'effet est de boucher les "trous" qui n'intersectent pas les bords. Ce filtre est appelé "fill-holes".

La croissance et l'anti-extensivité d'un tel opérateur sont évidentes. Pour vérifier qu'une ouverture par reconstruction est bien une ouverture, on peut se convaincre de son idempotence en posant $Y = R_{B,X}(\gamma(X))$ et en montrant que $R_{B,Y}(\gamma(Y)) = Y$. Comme $\gamma(X) \subseteq Y \subseteq X$, il vient $\gamma(Y) = \gamma(X)$ par croissance et idempotence de γ . De plus, par définition, Y est la réunion des composantes connexes de X qui intersectent $\gamma(X)$. Ainsi toute composante connexe de Y intersecte $\gamma(Y)$, d'où après reconstruction $R_{B,Y}(\gamma(Y)) = Y$.

Naturellement, on peut définir une reconstruction duale de X par érosion à partir de $Y \supseteq X$:

$$R_{B,X}^*(Y) := (R_{B,X^c}(Y^c))^c = \bigcap_{n \geq 1} \tilde{\varepsilon}_{B,X^c}^n(Y) \quad (3.35)$$

où $\tilde{\varepsilon}_{B,X^c}(Y) = (\delta_{B,X^c}(Y^c))^c = \varepsilon_B(Y) \cup X$. Un exemple très utile d'application, appelé filtre "fill-holes", est illustré par la figure 3.18. Enfin, d'après la proposition 3, pour toute fermeture φ , l'opérateur

$$X \in \mathcal{P}(E) \mapsto R_{B,X}^*(\varphi(X)) \quad (3.36)$$

est une fermeture, appelée *fermeture par reconstruction*. On peut noter au passage que $\tilde{\varepsilon}_{B,X^c}^n(Y) = \varepsilon_{B,X^c}^n(Y \cap X^c) \cup X$, si bien que $R_{B,X}^*(Y) = \left(\bigcap_{n \geq 1} \varepsilon_{B,X^c}^n(Y \cap X^c) \right) \cup X$ où ε_{B,X^c} est l'érosion géodésique par B dans X^c , comme définie par l'équation (3.29).

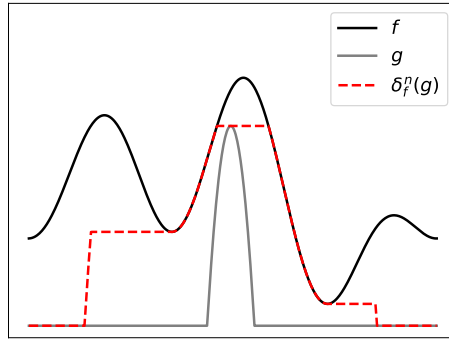


FIGURE 3.19 – Itération de la dilatation géodésique d'une fonction g sous la fonction f , en une dimension. L'élément structurant B est un petit intervalle symétrique contenant l'origine.

3.4.2 Le cas des fonctions

On considère maintenant le treillis \mathcal{F} des fonctions définies sur $E = \mathbb{Z}^2$ et à valeurs dans $[0, M]$, $M > 0$, muni de l'involution $f \in \mathcal{F} \mapsto \bar{f} = M - f$. Comme précédemment, $B \subseteq E$ est un élément structurant symétrique et contenant l'origine. On retrouve alors des définitions exactement analogues au cas des ensembles. Seule l'interprétation du résultat des opérateurs géodésique change, du fait qu'à la connexité (existence de chemins) viennent s'ajouter les valeurs prises par les fonctions le long de ces chemins.

Ainsi pour $f \in \mathcal{F}$, on définit la dilatation géodésique par B sous f par

$$\forall g \leq f, \delta_{B,f}(g) = \delta_B(g) \wedge f, \tag{3.37}$$

opérateur que l'on peut itérer

$$\delta_{B,f}^n(g) = \delta_{B,f} \circ \dots \circ \delta_{B,f}(g) = \delta_B(\delta_{B,\bar{f}}^{n-1}(g)) \wedge f, \tag{3.38}$$

et ce jusqu'à idempotence pour obtenir la reconstruction géodésique de f par dilatation à partir de $g \leq f$:

$$R_{B,f}(g) := \bigvee_{n \geq 1} \delta_{B,f}^n(g). \tag{3.39}$$

La reconstruction duale à partir de $g \geq f$ est, comme dans (3.35),

$$R_{B,f}^*(g) := \overline{\left(R_{B,\bar{f}}(\bar{g}) \right)} = \bigwedge_{n \geq 1} \tilde{\varepsilon}_{B,\bar{f}}^n(g) \tag{3.40}$$

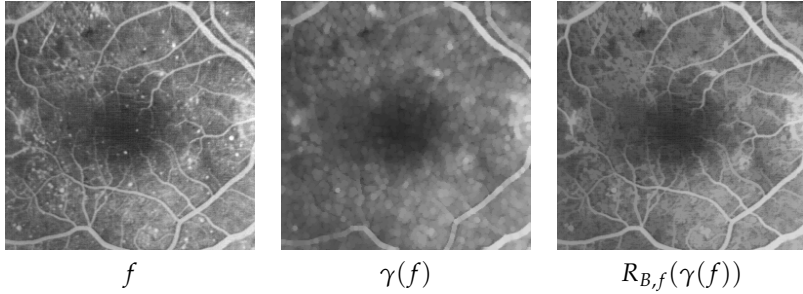


FIGURE 3.20 – Ouverture par un élément structurant hexagonal suivi de reconstruction géodésique en connexité 4 (ouverture par reconstruction).

où $\tilde{\varepsilon}_{B,\bar{f}}(g) = \varepsilon_B(g) \vee f$. De là, on a encore l'ouverture et la fermeture par reconstruction $R_{B,f}(\gamma(f))$ et $R_{B,f}^*(\varphi(f))$, pour toute ouverture γ et fermeture φ .

Comme dit plus haut la principale différence avec le cas des ensembles est l'interprétation des opérateurs géodésiques itérés. Dans le cas des ensembles, $x \in \delta_{B,X}^n(Y)$ si et seulement si il existe un chemin dans X de longueur au plus n , entre Y et x . Dans le cas des fonctions, en notant $h = \delta_{B,f}^n(g)$ et \mathcal{C}_x^n l'ensemble des chemins de longueur au plus n partant de x , alors

$$h(x) = \max_{c \in \mathcal{C}_x^n} \left(\min_{y \in c} f(y) \wedge \max_{y \in c} g(y) \right). \quad (3.41)$$

Ainsi, les maxima de g se propagent jusqu'à ce qu'ils coïncident avec f , après quoi c'est un minimum local de f qui va se propager, comme illustré par la figure 3.19.

3.5 Filtres morphologiques

Les filtres sont des opérateurs centraux en Morphologie Mathématique. À l'image de tous les concepts présentés dans ce chapitre, ils sont à la fois l'objet d'une théorie sophistiquée et de nombreux usages pratiques.

Définition 13 (Filtres morphologiques et leurs invariants). *Soit \mathcal{L} un treillis complet. Un filtre sur \mathcal{L} est un opérateur $\psi : \mathcal{L} \rightarrow \mathcal{L}$, croissant et idempotent :*

- $\forall x, y \in \mathcal{L}, x \leq y \Rightarrow \psi(x) \leq \psi(y)$
- $\psi \circ \psi = \psi$.

On appelle invariants de ψ , noté \mathcal{I}_ψ , l'ensemble des points fixes du filtre, i.e. $\mathcal{I}_\psi := \psi(\mathcal{L})$.

Comme cela a été déjà souligné dans la littérature [29], la propriété d'idempotence rappelle celle des filtres linéaires passe-bande, mais la plupart des filtres linéaires ne sont pas idempotents (voir l'exemple du flou gaussien de la figure 3.1).

On remarque immédiatement que les ouvertures et les fermetures sont des filtres. À ce stade on sait donc construire des filtres simples, notamment à partir d'érosions et dilatations. Voyons comment la combinaison de filtres produit de nouveaux filtres plus élaborés. Une première combinaison très commode produisant de nouveaux filtres est le supremum d'ouvertures ou l'infimum de fermetures.

3.5.1 Supremum d'ouvertures, infimum de fermetures

Proposition 4 (Supremum d'ouvertures, infimum de fermetures.). *Soit I un ensemble, $(\gamma_i)_{i \in I}$ une famille d'ouvertures et $(\varphi_i)_{i \in I}$ une famille de fermetures. Alors,*

- $\gamma := \bigvee_{i \in I} \gamma_i$ est une ouverture
- $\varphi := \bigwedge_{i \in I} \varphi_i$ est une fermeture.

Exemple 8. *Un exemple classique de supremum d'ouvertures est celui illustré par la figure 3.21. Chaque ouverture de la famille considérée est une ouverture par un segment orienté. Une telle ouverture va préserver les formes qui peuvent contenir les translatsés de ce segment, et donc supprimer celles qui sont de taille inférieure ou qui sont allongées dans une direction différente. Le supremum sur cette famille d'ouverture joue le rôle d'un "ou" logique, préservant toute forme allongée dans au moins une des directions considérées (typiquement des vaisseaux sanguins), et supprimant toute forme dans laquelle aucun des segments n'est inclus, quel que soit son orientation (typiquement des formes isotropes et de diamètre inférieur à la longueur du segment).*

3.5.2 Filtres alternés et alternés séquencés

Qu'en est-il de la composition de filtres? Si elle ne produit pas un nouveau filtre en général, c'est le cas lorsque les filtres sont ordonnés, comme le montre le théorème suivant.

Théorème 1 ([13]⁶). *Soit $(\psi_n)_{n \in \mathbb{N}}$ une famille ordonnée de filtres sur le treillis complet \mathcal{L} , c'est à dire telle que pour tous $p, q \in \mathbb{N}$, $\psi_p \leq \psi_q$ ou $\psi_q \leq \psi_p$. Alors toute composition de ces filtres est un filtre, i.e. : pour tout $k \in \mathbb{N}^*$, et pour tout $(i_1, i_2, \dots, i_k) \in \mathbb{N}^k$, l'opérateur $\eta = \psi_{i_1} \circ \psi_{i_2} \circ \dots \circ \psi_{i_k}$ est un filtre sur \mathcal{L} .*

6. Le résultat original est un peu plus général, mais nous nous restreignons à cette formulation.

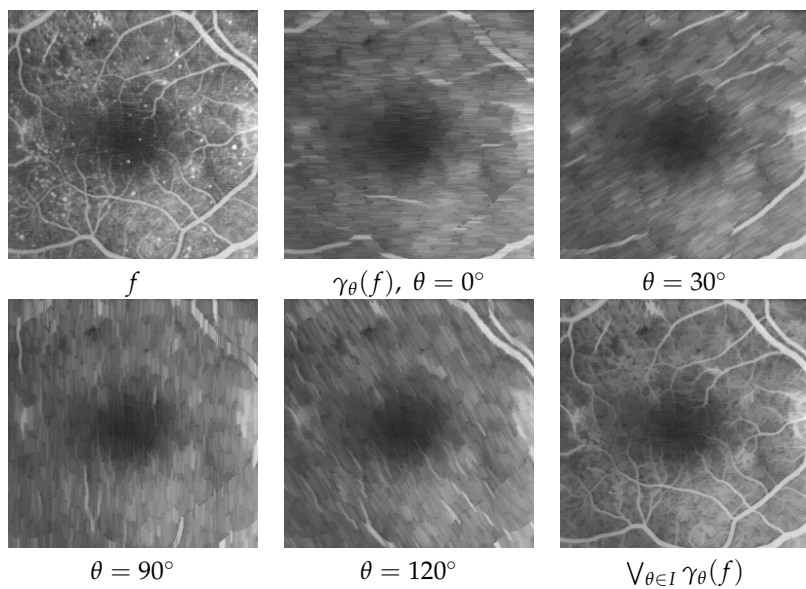


FIGURE 3.21 – Supremum d'ouvertures $\gamma_\theta(f)$ d'une image f par des segments longueur 5 pixels et orientés selon un angle θ . L'ensemble des orientations sur lequel le supremum est pris est $I = \{0^\circ, 15^\circ, \dots, 165^\circ\}$.

Une démonstration simple de ce résultat (comme proposée dans [13]) repose sur le fait que les ψ_{i_j} admettent deux éléments remarquables : l'un majorant tous les autres, et un minorant tous les autres. En utilisant la croissance et l'idempotence des ψ_{i_j} , on obtient $\eta^2 \leq \eta$ et $\eta \leq \eta^2$, ce qui prouve l'idempotence de η . Sa croissance est directe par composition de fonctions croissantes.

Ce résultat s'applique en particulier à une paire de filtres ordonnés, tels qu'une ouverture ψ et une fermeture ξ . Une conséquence importante est qu'à partir de deux filtres ordonnés, on peut construire par composition exactement quatre nouveaux filtres, à savoir $\xi\psi$, $\psi\xi$, $\psi\xi\psi$ et $\xi\psi\xi$, mais toute autre composition redonne l'un de ces quatre filtres. Ces filtres sont appelés *alternés*.

Une version moins générale du théorème 1 avait été proposée précédemment [26, 27], conduisant à la définition d'une classe de filtres, appelés *filtres alternés séquencés*, très utiles en pratique, notamment pour le débruitage.

Définition 14 (Filtres alternés séquencés). Soient $(\xi_n)_{n \in \mathbb{N}^*}$ et $(\psi_n)_{n \in \mathbb{N}^*}$ des suites respectivement croissante ($\xi_n \leq \xi_{n+1}$) et décroissante ($\psi_{n+1} \leq \psi_n$) de filtres sur \mathcal{L} , telles que $\psi_1 \leq \xi_1$. Alors, pour tout $k \in \mathbb{N}^*$, les filtres

$$M_k = \psi_k \xi_k \dots \psi_2 \xi_2 \psi_1 \xi_1 \quad \text{et} \quad N_k = \xi_k \psi_k \dots \xi_2 \psi_2 \xi_1 \psi_1 \quad (3.42)$$

sont appelés *filtres alternés séquencés*.

Exemple 9. La force des filtres alternés séquencés est leur capacité à débruiter sans perte une image. L'élimination du bruit se fait de manière progressive, ce qui permet de conserver un maximum d'information contrairement à une ouverture et fermeture avec un élément structurant plus grand qui élimineraient le bruit, mais aussi de l'information. Ce phénomène est illustré Fig.3.22.

L'idempotence des filtres implique que certaines compositions de filtres ordonnés sont redondantes avec d'autres plus simples, comme nous l'avons déjà vu dans le cas de deux filtres. En particulier deux types de familles de filtres ordonnés rendent inutile toute composition : les familles décroissantes d'ouvertures et les familles croissantes de fermetures.

3.5.3 Le cas particulier des granulométries

Définition 15 (Granulométrie, anti-granulométrie). Une granulométrie sur \mathcal{L} est une famille décroissante d'ouvertures $(\gamma_\lambda)_{\lambda \in I}$, $I \subseteq \mathbb{R}^+$, i.e. telle que pour tous $\lambda, \mu \in I$, $\lambda \leq \mu \Rightarrow \gamma_\mu \leq \gamma_\lambda$.

De façon analogue, une anti-granulométrie est une famille croissante de fermetures $(\varphi_\lambda)_{\lambda \in I}$, $I \subseteq \mathbb{R}^+$.

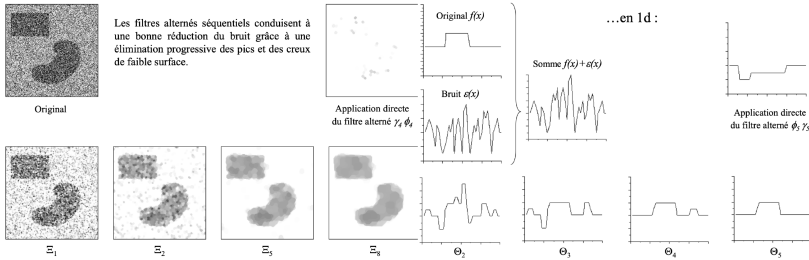


FIGURE 3.22 – Exemples de filtres alternés séquentiels et leur effet sur le bruit.

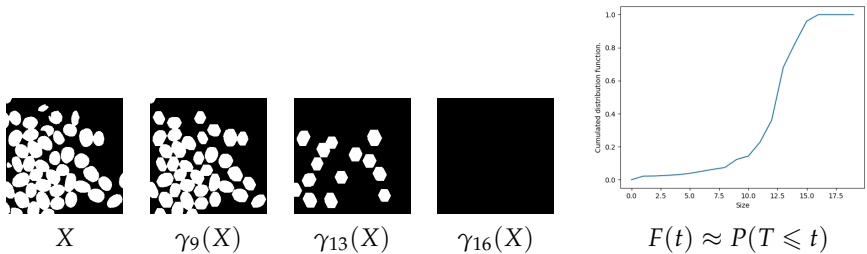


FIGURE 3.23 – Exemple de l’analyse empirique de la distribution de taille par granulométrie. Pour chaque k , γ_k est l’ouverture par l’élément structurant $B_k = \delta_B^{k-1}(B)$, où B est hexagonal de taille 1.

On a le résultat suivant, qui en plus de donner une définition alternative montre qu’il n’y a pas d’intérêt à composer les éléments d’une (anti-)granulométrie.

Proposition 5. Soient $I \subseteq \mathbb{R}^+$, $(\gamma_\lambda)_{\lambda \in I}$ une famille d’ouvertures et $(\varphi_\lambda)_{\lambda \in I}$ une famille de fermetures. Alors $(\gamma_\lambda)_{\lambda \in I}$ est une granulométrie si et seulement si

$$\forall \lambda, \mu \in I, \quad \gamma_\lambda \circ \gamma_\mu = \gamma_{\max(\lambda, \mu)} \tag{3.43}$$

et $(\varphi_\lambda)_{\lambda \in I}$ est une anti-granulométrie si et seulement si

$$\forall \lambda, \mu \in I, \quad \varphi_\lambda \circ \varphi_\mu = \varphi_{\max(\lambda, \mu)}. \tag{3.44}$$

Bien qu’une (anti-)granulométrie n’a pas d’intérêt du point de vue de la composition, elle permet en pratique d’analyser la distribution de taille des objets dans une image, comme dans l’exemple qui suit.

Exemple 10. En appliquant des ouvertures par des éléments structurants de tailles croissantes, on obscurcit de plus en plus l'image. Les formes claires subsistent tant qu'elles peuvent contenir l'élément structurant utilisé à la taille considérée, puis disparaissent, comme illustré par la figure 3.23. En mesurant pour chaque taille d'élément structurant la fraction de pixels "détruits", c'est à dire passant de blanc à noir, on produit une estimation de la fonction de répartition de la taille T d'un objet convexe dans l'image.

3.6 Segmentation d'images et représentations hiérarchiques

La segmentation d'une image est une partition de ses pixels, c'est à dire une collection de sous-ensembles disjoints, appelés *segments*, dont l'union est le support de l'image. Le but de l'analyste d'images est bien entendu d'obtenir des segments qui correspondent à des objets, mais cet objectif n'est presque jamais atteint. D'une part, car le problème est mal posé, ou du moins dépend-il de l'application : quelle est la bonne segmentation d'un personnage ? Un segment pour tout son corps ou un segment pour chacun de ses membres ? D'autre part, même lorsque la segmentation attendue est bien définie, le problème contient de nombreuses difficultés pour une méthode automatique. Une approche naturelle est donc de produire non pas une unique segmentation $S = \cup_{1 \leq i \leq n} s_i$, mais une suite croissante de partitions $S_k = \cup_{1 \leq i \leq n_k} s_{k,i}$, $1 \leq k \leq N$, de la plus fine qui attribue un segment par pixel, à la plus grossière qui ne contient qu'un segment, l'ensemble de tous les pixels. La croissance est entendue au sens de l'ordre partiel $S_k \preceq S_{k+1}$ entre partitions, défini par :

$$\forall i \in \{1, \dots, n_k\}, \exists j \in \{1, \dots, n_{k+1}\}, s_{k,i} \subseteq s_{k+1,j}. \quad (3.45)$$

La suite $(S_k)_{1 \leq k \leq N}$ est appelée une *hiérarchie*. C'est une représentation multi-échelles de l'image, dont on peut extraire une segmentation unique en choisissant des segments à différents niveaux de la hiérarchie, selon l'application. Nous présentons ci-dessous quelques représentations hiérarchiques populaires en segmentation d'image, et qui ont été développées par la communauté de *morphologues*.

3.6.1 Watershed

L'algorithme du *Watershed*, ou *ligne de partage des eaux*, a été proposé à la fin des années 1970 [3] et est devenu une référence pour la segmentation d'images.

Une manière simple de comprendre la ligne de partage des eaux est de (re)partir de l'idée de considérer l'image comme un relief topographique, avec des vallées (pixels sombres) et des montagnes (pixels clairs). On peut alors définir une transformation topographique, la ligne de partage des eaux, qui peut se voir de deux manières.

1) Ruissellement : La première consiste à « faire tomber » une goutte d'eau sur le relief dans l'image. Cette goutte va s'écouler jusqu'à atteindre une vallée, c'est-à-dire un minimum local. Pour chaque minimum local M , on peut associer un ensemble de points P tels qu'une goutte qui démarre en P s'écoule pour rejoindre M . Cet ensemble est appelé « bassin versant ». Les points qui se situent à la frontière de deux bassins forment la ligne de partage des eaux. L'ensemble des bassins versants constitue une partition de l'image.

2) Inondation : La deuxième manière de voir la ligne de partage des eaux est non pas de faire tomber de l'eau dans des bassins, mais d'inonder ces derniers. Si on inonde d'eau ces « bassins », le niveau monte jusqu'à rencontrer l'eau provenant d'un autre bassin (soit par rencontre parfaite, soit par débordement). Les points situés à cette jonction sont ce qu'on appelle la « ligne de partage des eaux ».

Dans tous les cas, une ligne de partage des eaux est un contour fermé (le bord de l'image constitue également un contour) : pour aller dans un autre bassin, il faut forcément franchir cette ligne, on ne peut pas la contourner.

La plupart des bibliothèques de traitement d'image possèdent une implémentation du Watershed. Toutes ne fonctionnent néanmoins pas de la même manière : certains calculeront automatiquement les minima locaux, d'autres demanderont à l'utilisateur de spécifier des marqueurs qui feront office de minima locaux d'intérêt. Le résultat est une segmentation où chaque minimum a un label différent (sauf certaines implémentations qui acceptent plusieurs marqueurs ayant le même label).

En pratique, on n'applique pas directement l'algorithme du Watershed directement sur une image à segmenter. Afin que la ligne de partage des eaux corresponde autant que possible à des contours d'objets, et les bassins versants aux objets, l'algorithme est plutôt appliqué au gradient morphologique de l'image. Par ailleurs, il convient de filtrer cette image auparavant, afin de la lisser et de diminuer les potentiels minima locaux. En effet, comme à chaque minimum local correspond un segment, cette simplification par filtrage permet de limiter la *sur-segmentation*, c'est à dire un nombre bien plus important de segments que d'objets dans l'image. C'est là que peuvent intervenir toutes les méthodes de filtrage morphologique à disposition, comme par exemple sur la figure 3.24.

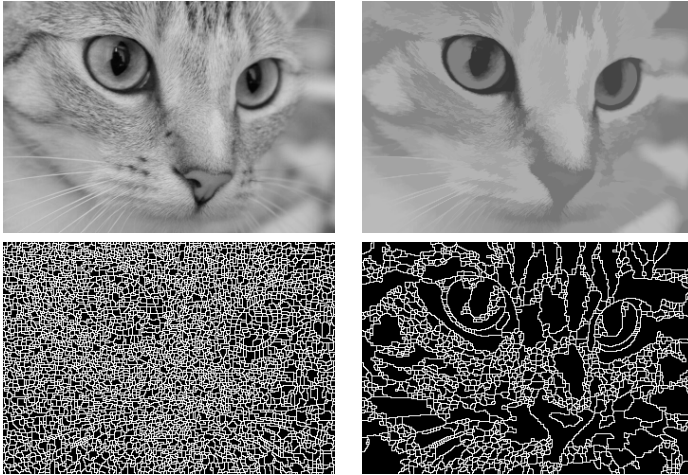


FIGURE 3.24 – Application de l'algorithme du Watershed au gradient de l'image initiale (colonne de gauche) et à celui de l'image filtrée par un filtre alterné séquentiel (à droite).

L'intérêt de la procédure par inondation est qu'elle permet de construire une hiérarchie de segmentations. En effet, on peut par exemple inonder tout le paysage de façon uniforme (un seul niveau d'eau pour tout le paysage) et, lorsque les lacs de deux bassins se rencontrent, l'union des deux bassins constitue un nouveau segment de la partition suivante, et ainsi de suite. De cette façon, on hiérarchise les régions en fonction de la hauteur de la crête la plus basse qui les sépare du bassin voisin (les régions qui fusionnent le plus tôt sont celles séparées par la crête la plus basse, où l'altitude est mesurée par rapport au minimum de ces régions). En utilisant d'autres méthodes d'inondation, d'autres critères peuvent être utilisés pour hiérarchiser les bassins (surface, volume), donnant lieu à d'autres hiérarchies : ces critères sont appelées *valeurs d'extinctions* [31]. Cet algorithme induit un ordre d'importance sur l'ensemble des segments de la hiérarchie, ce qui permet de retenir la partition composée des K segments les plus importants, comme sur la figure 3.25.

3.6.2 Max-trees et min-trees

Les max-trees et min-trees [25] sont deux représentations duales construites à partir des ensembles de niveaux d'une image. L'ensemble de

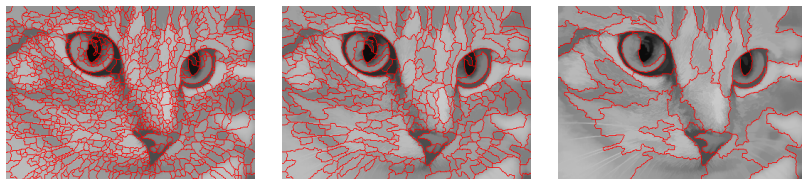


FIGURE 3.25 – Les 1000, 300 et 50 régions les plus importantes d'une hiérarchie de segmentations obtenue par la méthode des valeurs d'extinctions sur une image filtrée. Le critère d'extinction ici est le volume des bassins.

niveau supérieur à $t \in \mathcal{V}$ d'une image f est défini par

$$\chi_t^+(f) = \{x \in E, f(x) \geq t\}. \quad (3.46)$$

Dans l'arbre max-tree, les noeuds sont les composantes connexes de ces ensembles de niveaux, et un noeud \mathcal{N}_1 est le fils d'un noeud \mathcal{N}_2 si $\mathcal{N}_1 \subseteq \mathcal{N}_2$. Ainsi, la racine de l'arbre est l'ensemble des pixels de l'image, et les feuilles sont ses maxima. Le min-tree est défini de façon analogue à partir des ensembles de niveaux inférieurs $\chi_t^-(f) = \{x \in E, f(x) \leq t\}$, ou par dualité en remarquant que $\text{min-tree}(f) = \text{max-tree}(-f)$. La racine de l'arbre est l'ensemble des pixels de l'image, et les feuilles sont ses minima.

Le filtrage d'image à partir de ces structures d'arbre consiste à élaguer certaines de ses branches avant de reconstruire une image avec les ensembles de niveau restants,

$$g(x) = \bigvee \{t, x \in \chi_t^+\} = \bigwedge \{t, x \in \chi_t^-\}. \quad (3.47)$$

L'élagage du max-tree induit une suppression de certains maxima régionaux, celui du min-tree des minima régionaux. La décision de supprimer une branche est basée sur des attributs, ou *critères*, des noeuds qui la composent, tels que la taille, le contraste, la régularité des contours, ou des mesures statistiques sur les valeurs prises par les pixels. De façon remarquable, comme pour les opérateurs géodésiques, ce type de filtrage élimine des contours (des lignes de niveaux) mais n'en déplace pas, ni n'en crée de nouveaux. C'est la propriété des *filtres connectés*, qui traite de façon identique l'ensemble des pixels d'une composante connexe dans un ensemble de niveaux (voir par exemple Figure 3.26). Les max-tree et min-tree sont des structures sur lesquelles reposent les implémentations efficaces des filtres connectés, tels que l'ouverture surfacique, qui consiste à supprimer les composantes connexes de taille inférieure à un certain seuil. Une description plus détaillée des usages possibles de ces structures de données

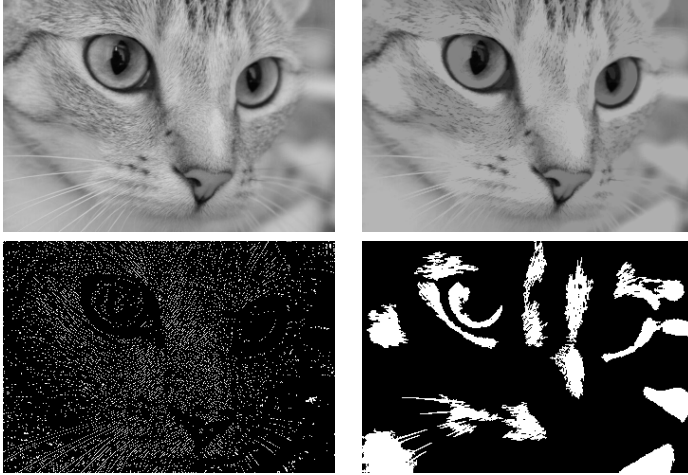


FIGURE 3.26 – Filtrage d'une image par élimination des noeuds de volume inférieure à 50 dans le *max-tree* correspondant, et maxima régionaux des images originale et filtrée respectivement.

est proposée par exemple dans [25, 21]. Notons également que d'autres représentations par arbres ont été définies sur des principes similaires : arbres des composantes, arbres binaires de partition [24], arbres des formes [1], pour les plus populaires.

3.7 Morphologie Mathématique et Deep Learning

En 2023 et depuis une dizaine d'années, les méthodes à l'état de l'art en analyse et traitement d'images utilisent largement l'apprentissage profond (deep learning). Cette section a pour but de montrer que les progrès exceptionnels du *learning from data*, n'excluent pas de mettre à profit également les méthodes jusqu'ici dépourvues d'apprentissage, telles que la Morphologie Mathématique. Cette dernière a en effet l'avantage d'être bien comprise sur le plan théorique et pratique, et d'offrir une grande richesse d'outils mobilisables hors ou dans les réseaux de neurones. Enfin, l'apprentissage peut être un levier pour de nouvelles avancées en *morpho*.

3.7.1 Avant et après l'apprentissage

La Morphologie Mathématique permet de préparer les données pour aider à l'apprentissage, avec des performances accrues. En particulier, un pré-traitement morphologique peut apporter des informations supplémentaires au réseau de neurones afin de mieux segmenter des petites structures [35, 8] ou des structures nécessitant un contexte global [10]. Pour la détection des hyper-intensités de la matière blanche dans [35], en plus des images de deux modalités IRM (T1 et FLAIR), le résultat d'un top-hat blanc de l'image FLAIR est passé comme troisième canal d'entrée d'un réseau VGG (voir Figure 3.27). Cet ajout de pré-traitement a permis d'augmenter le score de détection de 0.39 à 0.61. Dans [8], c'est encore grâce à un top-hat qu'a été améliorée l'interprétabilité de la détection, par un CNN⁷, de déchirures du ménisque sur des IRM du genou. Dans [10] enfin, on voit sur des coupes histologiques de la peau qu'il est souvent impossible de distinguer l'épiderme vivant de l'épiderme mort seulement à partir des informations locales accessibles dans le champs réceptif d'un CNN (voir Figure 3.27). Un pré-traitement géodésique de type *fill-holes* a néanmoins permis d'encoder localement les informations topologiques éloignées, améliorant considérablement la segmentation par le CNN. Un post-traitement morphologique a en suite été appliqué pour corriger les erreurs restantes.

Réciproquement, l'apprentissage profond peut être préparatoire au traitement morphologique. On peut par exemple apprendre des représentations adaptées à ces traitements. C'est ce qui est proposé dans [6], où l'entraînement d'un auto-encodeur permet de décomposer les données d'entrée sur un dictionnaire de représentations "par parties" parcimonieuses, de façon à ce que le calcul des opérateurs de bases sur les atomes du dictionnaire suffise à approcher ces mêmes opérateurs sur l'ensemble des données.

3.7.2 Couches morphologiques dans les réseaux

Comme on l'a vu, les dilatations et érosions par des éléments structurants sont analogues aux convolutions par des noyaux. En changeant simplement d'algèbre, on peut donc définir des couches morphologiques au sein des réseaux de neurones. La Figure 3.28 (a) et la Table 3.1 résument cette idée en comparant les différentes combinaisons entre valeurs d'entrée

7. Convolutional Neural Network, ou réseau de neurones convolutif.

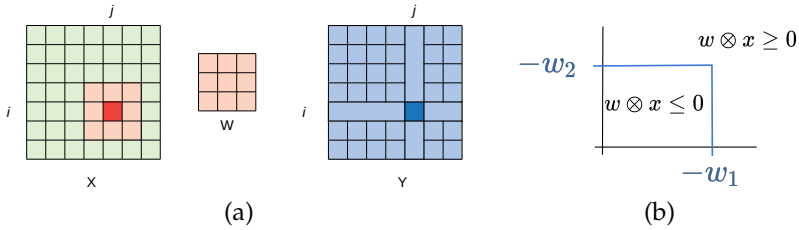


FIGURE 3.28 – (a) Opérateur invariant par translation représenté par un noyau w , qui peut être implémenté comme une couche neuronale convolutive. (b) Séparation du plan par l’hyperplan tropical d’équation $w \otimes x := \bigvee_{1 \leq i \leq 2} w_i + x_i = 0$.

Neurone classique	Dilatation	Érosion
$y = \sum_{i=1}^n w_i \cdot x_i$	$y = \bigvee_{i=1}^n w_i + x_i$	$y = \bigwedge_{i=1}^n w_i + x_i$

TABLE 3.1 – Comparaison des opérations réalisées par différents neurones, pour une entrée $x \in \mathbb{R}^n$ et un vecteur de paramètres $w \in \mathbb{R}^n$.

soit de fixer ces paramètres et entraîner le reste de l’architecture en tenant compte de ces opérateurs dans la chaîne algorithmique (par exemple, comme pré ou post-traitement).

Comme la plupart des optimiseurs performants des bibliothèques de deep learning reposent sur des variantes du gradient stochastique, le problème se pose de pouvoir utiliser ces algorithmes pour l’apprentissage de couches morphologiques alors que les opérateurs max et min les rendent non différentiables. Une première approche consiste à utiliser des approximations différentiables des opérateurs morphologiques, tels que la P -convolution inspirée de la moyenne contre-harmonique [19], ou les couches \mathcal{L} -Morph et \mathcal{S} -Morph basées sur la moyenne de Lehmer et l’ α -softmax respectivement [16, 14]. En plus de leur différentiabilité, ces opérateurs ont l’avantage de pouvoir approcher asymptotiquement dilatation et érosion, en fonction de la valeur d’un paramètre (P pour la P -convolution et la \mathcal{L} -Morph, α pour la \mathcal{S} -Morph).

Une seconde approche repose sur le constat que les couches morphologiques, à l’instar des opérateurs max et min, sont en fait différentiables presque partout. De plus, là où elles ne le sont pas, des directions de descente et de montée existent. Considérons par exemple la fonction $f_w(x) = \bigvee_{i=1}^n w_i + x_i$ définie sur \mathbb{R}^n et paramétrée par $w \in \mathbb{R}^n$. En notant, pour un $x \in \mathbb{R}^n$ fixé et tout $1 \leq j \leq n$, l’ensemble $E_j(x) = \{w \in \mathbb{R}^n, f_w(x) = x_j + w_j\}$, il apparaît que $f_w(x)$ est différentiable par rapport

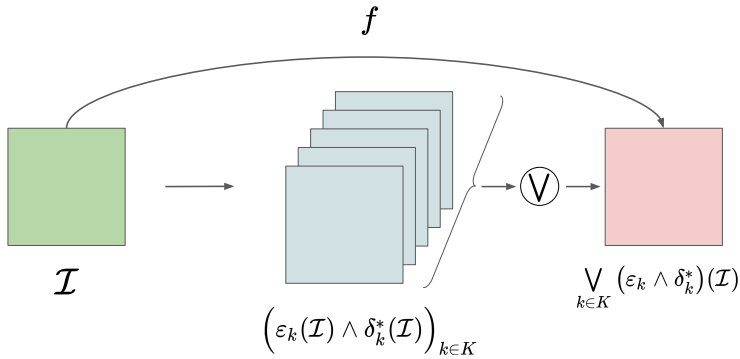


FIGURE 3.29 – Représentation d'un opérateur quelconque f entre treillis d'images, par une famille d'opérateurs morphologiques élémentaires : érosions et anti-dilatations. En principe cette représentation peut-être apprise sous forme de réseaux de neurones morphologiques.

à w sur l'intérieur de chaque $E_j(x)$, de gradient $\nabla_w f_w(x) = e_j$, le vecteur composé de zéros et d'un seul "1" en position j . Les lieux où cette fonction est non différentiable en ses paramètres, sont les intersections de plusieurs E_j , de type $\bigcap_{j \in I} E_j$. Dans ce cas, le vecteur $-\sum_{j \in I} e_j$ est la direction de plus grande descente, et chaque e_j tel que $j \in I$ est une direction de montée en $f_w(x)$. En pratique, de nombreux exemples de couches morphologiques "exactes" optimisées avec succès par descente de gradient, existent dans la littérature [37, 12, 20, 33]. Une librairie nommée *Morpholayers* et basée sur Tensorflow a notamment été développée pour tirer profit de la différentiation automatique dans l'optimisation des couches morphologiques [32].

3.7.3 Représentation des opérateurs de treillis

Un des arguments théoriques en faveur des réseaux de neurones profonds est le théorème d'approximation universelle [15], qui donne espoir de pouvoir représenter un très large ensemble de fonctions par de tels réseaux. Ce résultat est purement théorique (car asymptotique), mais la pratique a fourni des preuves encore plus convaincantes, notamment en traitement d'images.

Si l'on remplace les couches neuronales classiques par des couches morphologiques de type max-plus et min-plus, ce résultat n'est plus valable.

Toutefois, plusieurs résultats fondamentaux existent sur la représentation d'opérateurs entre treillis complets, qui peuvent se traduire en termes de réseaux de neurones morphologiques. Le plus simple d'entre eux est que toute fonction croissante $\psi : \mathcal{L} \rightarrow \mathcal{L}'$ entre deux treillis complets peut s'écrire comme le supremum d'une famille d'érosions, et comme l'infimum d'une famille de dilatations [26]. En effet,

$$\forall x \in \mathcal{L}, \quad \psi(x) = \bigvee'_{y \in \mathcal{L}} \varepsilon_y(x) = \bigwedge'_{y \in \mathcal{L}} \delta_y(x) \quad (3.48)$$

avec $\varepsilon_y(x) = \psi(y)$ si $y \leq x$ et $\varepsilon_y(x) = \wedge \mathcal{L}'$ sinon, et de façon analogue $\delta_y(x) = \psi(y)$ si $x \leq y$ et $\delta_y(x) = \vee \mathcal{L}'$ sinon.

Bien entendu ces décompositions en "échelons" sont peu intéressantes mais elles sont un point de départ pour en construire de plus parcimonieuses. Il a été démontré par exemple [17] que pour tout opérateur $\psi : \mathcal{P}(\mathbb{Z}^2) \rightarrow \mathcal{P}(\mathbb{Z}^2)$ croissant et invariant par translation, il existe une base minimale \mathcal{B} d'éléments structurants tels que

$$\psi = \bigvee_{B \in \mathcal{B}} \varepsilon_B \quad (3.49)$$

et un résultat dual existe également (infimum de dilatations). Cette famille est bien plus restreinte que $\mathcal{P}(E)$, même si elle peut être infinie et même non dénombrable. Dans le cas du filtre médian pour une fenêtre F de cardinal $2n + 1$, cette famille est finie : il s'agit de tous les sous-ensembles de cardinal $n + 1$, donc un total de $\binom{2n+1}{n+1}$ éléments structurants [18].

Le résultat le plus général à ce jour [2] affirme que tout opérateur $\psi : \mathcal{L} \rightarrow \mathcal{L}'$ entre deux treillis complets peut se décomposer comme suit

$$\psi = \bigvee_{k \in K} (\varepsilon_k \wedge \delta_k^*) \quad (3.50)$$

où K est un ensemble, chaque ε_k est une érosion et chaque δ_k^* est une anti-dilatation, c'est à dire tel que $\delta_k^*(\vee_i x_i) = \wedge_i \delta_k^*(x_i)$. Là encore cette décomposition peut être triviale, comme (3.48), en remplaçant les "échelons" par des "impulsions". En effet

$$\psi = \bigvee'_{y \in \mathcal{L}} (\varepsilon_y \wedge \delta_y^*) \quad (3.51)$$

où ε_y est l'érosion définie plus haut et pour tout $x \in \mathcal{L}$, $\delta_y^*(x) = \psi(y)$ si $x \leq y$ et $\delta_y^*(x) = \wedge \mathcal{L}'$ sinon. Mais la construction proposée par [2] montre qu'une décomposition bien plus parcimonieuse existe, bien que potentiellement infinie et en général inconnue.

Dans la mesure où des érosions et des anti-dilatations peuvent être implémentées comme des couches de réseaux de neurones morphologiques, on peut espérer apprendre de telles représentations d'opérateurs d'images à partir des données, par optimisation des paramètres du réseau. Cette piste fascinante et ambitieuse reste encore à explorer.

Bibliographie

- [1] C. BALLESTER, V. CASELLES et P. MONASSE : The tree of shapes of an image. *ESAIM : Control, Optimisation and Calculus of Variations*, 9:1–18, 2003.
- [2] G. J. F. BANON et J. BARRERA : Decomposition of mappings between complete lattices by mathematical morphology, Part I. General lattices. *Signal Processing*, 30(3):299–327, 1993.
- [3] S. BEUCHER : Use of watersheds in contour detection. *Actes de Proc. Int. Workshop on Image Processing, Sept. 1979*, pp. 17–21, 1979.
- [4] I. BLOCH, S. BLUSSEAU, R. PINO PÉREZ, É. PUYBAREAU et G. TOCHON : On some associations between mathematical morphology and artificial intelligence. *Actes de J. LINDBLAD, F. MALMBERG et N. SLADOJE, édés : Discrete Geometry and Mathematical Morphology*, pp. 457–469. Springer International Publishing, 2021.
- [5] I. BLOCH, H. HEIJMANS et C. RONSE : Mathematical morphology. *Dans Handbook of Spatial Logics*, pp. 857–944. Springer, 2007.
- [6] S. BLUSSEAU, B. PONCHON, S. VELASCO-FORERO, J. ANGULO et I. BLOCH : Approximating morphological operators with part-based representations learned by asymmetric auto-encoders. *Mathematical Morphology - Theory and Applications*, 4(1):64–86, 2020.
- [7] S. BLUSSEAU, S. VELASCO-FORERO, J. ANGULO et I. BLOCH : Morphological adjunctions represented by matrices in max-plus algebra for signal and image processing. *Actes de É. BAUDRIER, B. NAEGEL, A. KRÄHENBÜHL et M. TAJINE, édés : Discrete Geometry and Mathematical Morphology*, pp. 206–218. Springer International Publishing, 2022.
- [8] V. COUTEAUX, S. SI-MOHAMED, O. NEMPONT, T. LEFEVRE, A. POPOFF, G. PIZAINÉ, N. VILLAIN, I. BLOCH, A. COTTEN et L. BOUSSEL : Automatic knee meniscus tear detection and orientation classification with mask-rnn. *Diagnostic and Interventional Imaging*, 100(4):235 – 242, 2019.

- [9] J. L. DAVIDSON et G. X. RITTER : Theory of morphological neural networks. *Actes de Digital Optical Computing II*, vol. 1215, pp. 378–389, 1990.
- [10] E. DECENCIÈRE, S. VELASCO-FORERO, F. MIN, J. CHEN, H. BURDIN, G. GAUTHIER, B. LAÏ, T. BORNSCHLOEGL et T. BALDEWECK : Dealing with topological information within a fully convolutional neural network. *Actes de J. BLANC-TALON, D. HELBERT, W. PHILIPS, D. POPESCU et P. SCHEUNDERS, édés : Advanced Concepts for Intelligent Vision Systems*, pp. 462–471. Springer International Publishing, 2018.
- [11] M. FAESSEL et M. BILODEAU : Smil : simple morphological image library. *Séminaire Performance et Généricité, LRDE*, 2014.
- [12] G. FRANCHI, A. FEHRI et A. YAO : Deep morphological networks. *Pattern Recognition*, 102:107246, 2020.
- [13] H. J. HEIJMANS : Composing morphological filters. *IEEE transactions on image processing*, 6(5):713–723, 1997.
- [14] R. HERMARY, G. TOCHON, É. PUYBAREAU, A. KIRSZENBERG et J. ANGULO : Learning grayscale mathematical morphology with smooth morphological layers. *Journal of Mathematical Imaging and Vision*, 64(7):736–753, 2022.
- [15] K. HORNIK : Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257, jan. 1991.
- [16] A. KIRSZENBERG, G. TOCHON, É. PUYBAREAU et J. ANGULO : Going beyond p-convolutions to learn grayscale morphological operators. *Actes de DGMM*, pp. 470–482. Springer, 2021. <https://arxiv.org/pdf/2102.10038.pdf>.
- [17] P. MARAGOS : A representation theory for morphological image and signal processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):586–599, 1989.
- [18] P. MARAGOS et R. SCHAFFER : Morphological filters—part ii : Their relations to median, order-statistic, and stack filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(8):1170–1184, 1987.
- [19] J. MASCI, J. ANGULO et J. SCHMIDHUBER : A learning framework for morphological operators using counter-harmonic mean. *Actes de ISMM*, pp. 329–340. Springer, 2013.
- [20] R. MONDAL, M. S. DEY et B. CHANDA : Image restoration by learning morphological opening-closing network. *Mathematical Morphology-Theory and Applications*, 4(1):87–107, 2020.

- [21] L. NAJMAN et H. TALBOT : *Mathematical morphology : from theory to applications*. John Wiley & Sons, 2013.
- [22] B. PERRET, G. CHIERCHIA, J. COUSTY, S. J. F. GUIMARAES, Y. KENMUCHI et L. NAJMAN : Higua : Hierarchical graph analysis. *SoftwareX*, 10:100335, 2019.
- [23] G. X. RITTER et P. SUSSNER : An introduction to morphological neural networks. *Actes de 13th International Conference on Pattern Recognition*, vol. 4, pp. 709–717, 1996.
- [24] P. SALEMBIER et L. GARRIDO : Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE transactions on Image Processing*, 9(4):561–576, 2000.
- [25] P. SALEMBIER, A. OLIVERAS et L. GARRIDO : Antiextensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, 7(4):555–570, 1998.
- [26] J. SERRA, éd. *Image Analysis and Mathematical Morphology, Part II : Theoretical advances*. Academic Press, London, 1988.
- [27] J. SERRA et L. VINCENT : An overview of morphological filtering. *Circuits, Systems and Signal Processing*, 11(1):47–108, 1992.
- [28] P. SOILLE *et al.* : *Morphological image analysis : principles and applications*, vol. 2. Springer, 1999.
- [29] S. R. STERNBERG : Grayscale morphology. *Computer vision, graphics, and image processing*, 35(3):333–355, 1986.
- [30] P. SUSSNER : Morphological perceptron learning. *Actes de Proceedings of the 1998 IEEE International Symposium on Intelligent Control (ISIC) held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA) Intell*, pp. 477–482. IEEE, 1998.
- [31] C. VACHIER et F. MEYER : Extinction value : a new measurement of persistence. *Actes de IEEE Workshop on nonlinear signal and image processing*, vol. 1, pp. 254–257. Neos Marmaras Greece, 1995.
- [32] S. VELASCO-FORERO : Morpholayers, 2020. URL <https://people.cmm.minesparis.psl.eu/users/velasco/morpholayers>.
- [33] S. VELASCO-FORERO, A. RHIM et J. ANGULO : Fixed point layers for geodesic morphological operations. *Actes de British Machine Vision Conference*, 2022.
- [34] S. S. WILSON : Morphological networks. *Actes de Visual Communications and Image Processing IV*, vol. 1199, pp. 483–496, 1989.

- [35] Y. XU, T. GÉRAUD, É. PUYBAREAU, I. BLOCH et J. CHAZALON : White matter hyperintensities segmentation in a few seconds using fully convolutional network and transfer learning. *Actes de A. CRIMI, S. BAKAS, H. KUIJF, B. MENZE et M. REYES, eds : Brainlesion : Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 501–514. Springer International Publishing, 2018.
- [36] E. ZAMORA et H. SOSSA : Dendrite morphological neurons trained by stochastic gradient descent. *Neurocomputing*, 260:420–431, 2017.
- [37] Y. ZHANG, S. BLUSSEAU, S. VELASCO-FORERO, I. BLOCH et J. ANGULO : Max-plus operators applied to filter selection and model pruning in neural networks. *Actes de ISMM*, pp. 310–322. Springer, 2019.