

N-35/88/MM

INTRODUCTION
AUX OUTILS MORPHOLOGIQUES
DE SEGMENTATION

S. BEUCHER, L. VINCENT

CMM Fontainebleau

Décembre 1988

ANRT

INTRODUCTION AUX OUTILS MORPHOLOGIQUES DE SEGMENTATION

Serge BEUCHER & Luc VINCENT
Centre de Morphologie Mathématique, Ecole des Mines
35 rue Saint-Honoré, 77305 Fontainebleau, France

ABSTRACT

This paper presents several segmentation tools provided by Mathematical Morphology. Together with that presentation, it describes the basic morphological transformations involved in the required algorithms. It then attempts to derive a general approach of segmentation problems using Mathematical Morphology, which is based on the two concepts of *markers* and *marking function*. This philosophy is illustrated with several examples.

1 INTRODUCTION

Un des problèmes les plus courants en traitement d'images consiste à extraire d'une scène les objets ou les régions à analyser. Cette *segmentation* d'images présente plusieurs facettes, des plus triviales comme le simple seuillage, aux plus complexes, basées sur la géométrie des objets [1], l'algorithmique géométrique (*computational geometry* en anglais) [2], l'analyse de textures [3], la croissance hiérarchique de régions [4]. . .

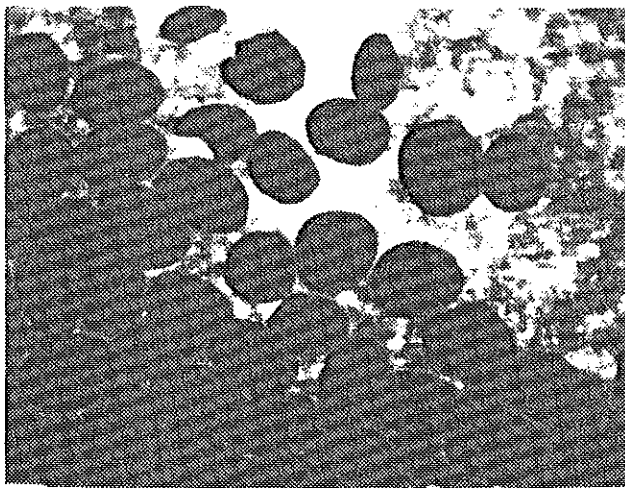
Pour résoudre ce genre de problèmes, la Morphologie Mathématique [1] offre à l'utilisateur un arsenal d'outils s'appliquant à la fois à l'imagerie binaire et numérique. Nous allons dans cet article tenter de décrire ces outils en illustrant leur usage par divers exemples. Cependant, avant cela, il paraît souhaitable de rappeler succinctement les transformations de base de la Morphologie Mathématique (que nous désignerons parfois par MM), tant dans le domaine binaire (MM dite ensembliste) que dans celui des images à teintes de gris (MM dite numérique).

2 LES TRANSFORMATIONS MORPHOLOGIQUES

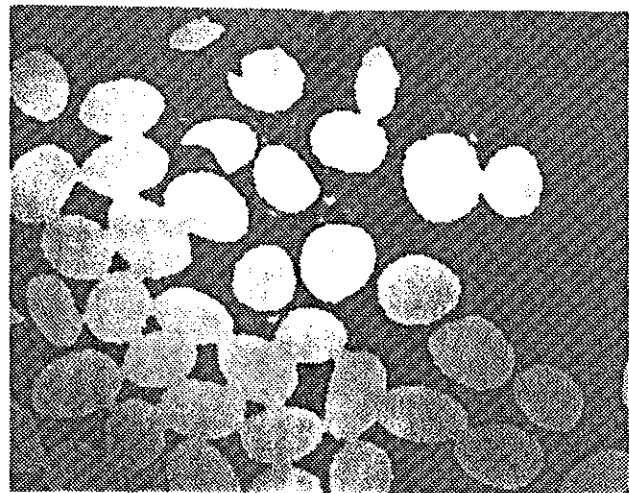
Par Morphologie Mathématique, on entend généralement une méthodologie de traitement d'image basée sur des transformations dites *morphologiques*. Ces transformations agissent sur des ensembles en morphologie binaire et sur des fonctions en morphologie numérique. Le résultat d'une transformation étant de même nature que l'objet sur lequel elle s'applique (un ensemble sera transformé en un

ensemble, une fonction en une fonction), il est alors facile de construire des transformations de plus en plus complexes, soit par itération, soit par concaténation de transformations élémentaires. On définit ainsi des opérations de plus haut niveau qui, malgré une complexité algorithmique de plus en plus grande, restent néanmoins faciles d'emploi car leur finalité est relativement immédiate. Par exemple, la résolution d'un problème comme la mise en évidence de cellules isolées dans une préparation cytologique nécessite l'enchaînement de plusieurs centaines de transformations élémentaires. En dépit de cela, le processus utilisé est facilement compréhensible, et son résultat prévisible.

Pour illustrer l'effet des diverses transformations morphologiques que nous allons être amené à présenter au cours de cet article, un exemple nous servira de fil conducteur: celui de l'image 1a, qui représente des grains de café. Cette image à teintes de gris est en fait quasiment binaire, puisqu'un simple seuillage permet d'obtenir l'image binaire 1b. On voit pourtant que sur cette image, les différents grains se recouvrent en partie; lorsque nous aurons présenté suffisamment d'outils morphologiques, nous serons capable de *segmenter* cette image, c'est à dire de séparer et d'extraire les différents grains.



a. image originale



b. image seuillée

FIG. 1: Grains de café

2.1 Propriétés des transformations morphologiques

Les transformations que nous allons introduire par la suite seront souvent dotées de propriétés intéressantes qu'il nous paraît utile de définir dès à présent. Soit Φ une transformation que nous supposons binaire dans un premier temps. Φ agit donc sur des éléments de $\mathcal{P}(\mathbb{R}^2)$, c'est-à-dire des ensembles de \mathbb{R}^2 ($\Phi : \mathcal{P}(\mathbb{R}^2) \rightarrow \mathcal{P}(\mathbb{R}^2)$). Elle sera dite *extensive* si et seulement si

$$\forall X \in \mathcal{P}(\mathbb{R}^2), \quad X \subset \Phi(X) \tag{1}$$

et *anti-extensive* si et seulement si

$$\forall X \in \mathcal{P}(\mathbb{R}^2), \quad \Phi(X) \subset X. \quad (2)$$

Par ailleurs, Φ est *croissante* si et seulement si elle préserve les relations d'ordre entre éléments de \mathbb{R}^2 , soit

$$\forall (X, Y) \in \mathcal{P}(\mathbb{R}^2)^2, \quad X \subset Y \implies \Phi(X) \subset \Phi(Y). \quad (3)$$

Nous dirons également que Φ est *idempotente* si appliquer Φ plusieurs fois de suite revient à ne l'appliquer qu'une seule fois:

$$\forall X \in \mathcal{P}(\mathbb{R}^2), \quad \Phi(\Phi(X)) = \Phi(X). \quad (4)$$

Enfin, deux transformations Φ et Ψ seront dites *duales* si et seulement si appliquer la première à un ensemble X revient à appliquer la seconde au complémentaire X^C de X :

$$\forall X \in \mathcal{P}(\mathbb{R}^2), \quad \Phi(X)^C = \Psi(X^C). \quad (5)$$

Ces propriétés se transcrivent immédiatement dans le cas numérique, lorsque Φ agit sur des fonctions de \mathbb{R}^2 dans \mathbb{R} . Dans ce cas, la relation d'ordre utilisée est la suivante:

$$\forall (f, g) \in \mathcal{F}(\mathbb{R}^2, \mathbb{R})^2, \quad f \leq g \iff \forall x \in \mathbb{R}^2, f(x) \leq g(x) \quad (6)$$

2.2 Erosion et dilatation

Ces deux opérations sont sans doute les plus classiques de la Morphologie Mathématique [1]. Elles sont à la base d'un très grand nombre de transformations plus élaborées. Avant de les présenter, il nous faut définir ici quelques notations qui nous seront utiles par la suite: choisissons dans \mathbb{R}^2 un point arbitraire O appelé *point origine*. A tout point x appartenant à \mathbb{R}^2 , on peut donc associer le vecteur \vec{Ox} . On utilisera par la suite une notation abrégée où, suivant le contexte, x désigne le point x de \mathbb{R}^2 ou le vecteur \vec{Ox} . soit alors $B \in \mathcal{P}(\mathbb{R}^2)$ un ensemble. Nous noterons \check{B} l'ensemble transposé (Voir figure 2):

$$\check{B} = \{-x, x \in B\}. \quad (7)$$

Par ailleurs, pour $a \in \mathbb{R}^2$, nous noterons B_a le translaté de B par le vecteur a :

$$B_a = \{a + x, x \in B\}. \quad (8)$$

Considérons à présent un ensemble X inclus dans \mathbb{R}^2 . Le dilaté de X par B est l'ensemble des x de \mathbb{R}^2 tels que l'intersection entre X et B_x est non vide (Voir figure 3a). Cette opération est notée $X \oplus \check{B}$:

$$X \oplus \check{B} = \{x \in \mathbb{R}^2, X \cap B_x \neq \emptyset\}. \quad (9)$$

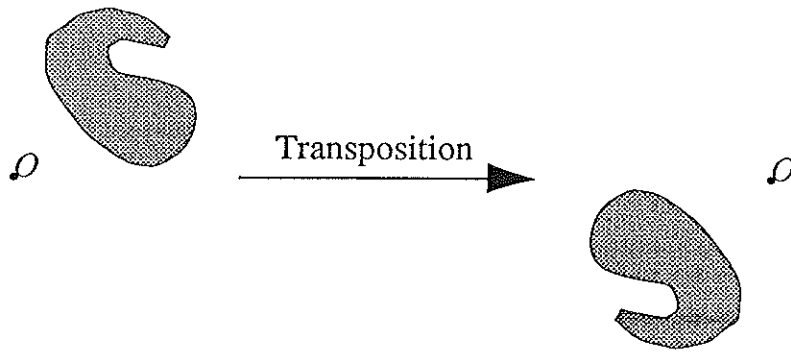


FIG. 2: Transposition d'un ensemble

L'érosion se définit de manière similaire: l'érodé de X par B , noté $X \ominus \check{B}$ est l'ensemble des x de \mathbb{R}^2 tels que B_x est totalement inclus dans X (Voir figure 3b):

$$X \ominus \check{B} = \{x \in \mathbb{R}^2, B_x \in X\}. \quad (10)$$

L'ensemble B , qui joue un rôle particulier dans ces transformations, est appelé *élément structurant*.

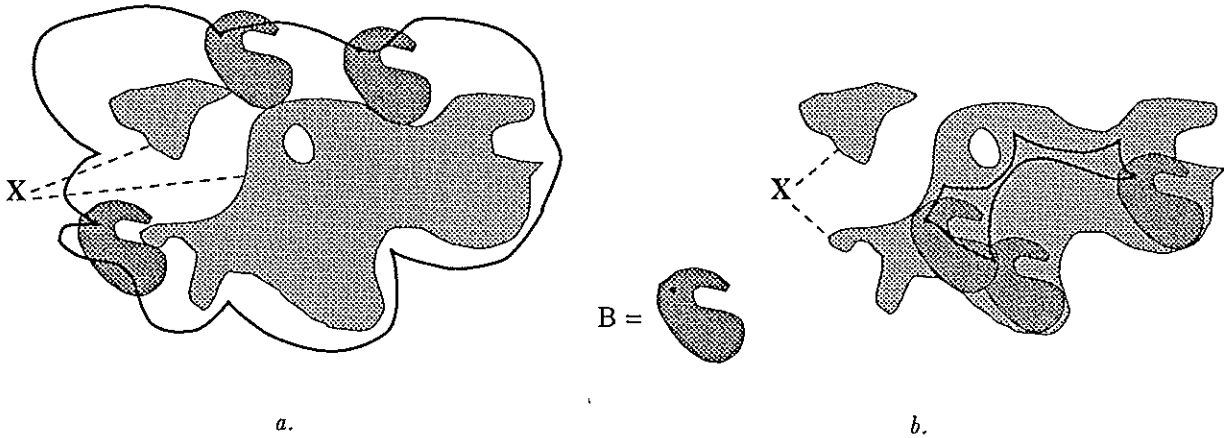
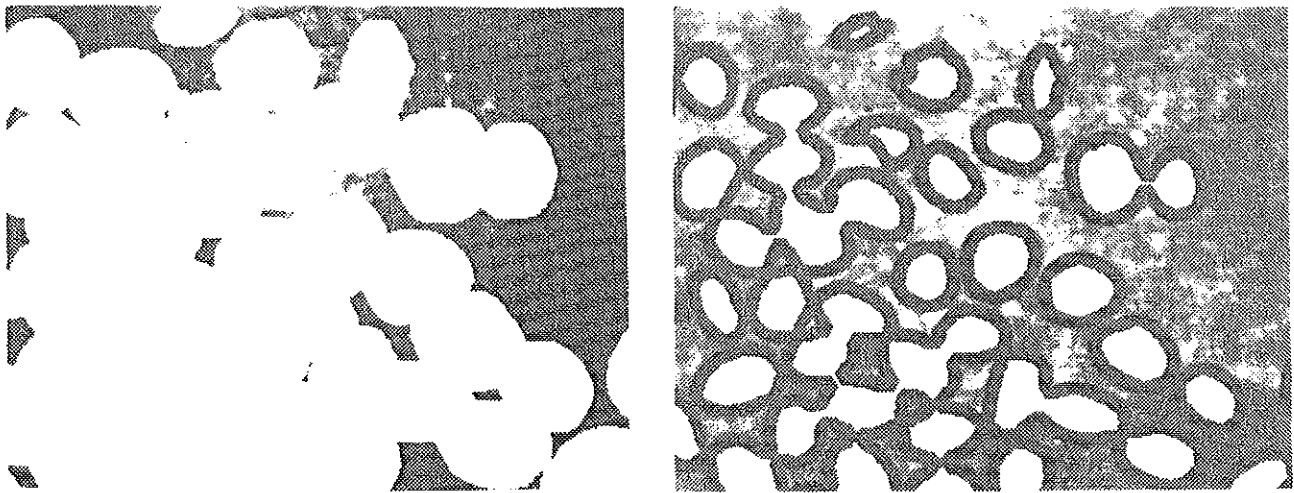


FIG. 3: Dilatation et érosion de X par B

La figure 4 illustre l'effet d'une érosion et d'une dilatation par un hexagone sur nos grains de café.

Dilatations et érosions sont croissantes. De plus, elles sont duales, c'est-à-dire que la dilatation d'un ensemble X est identique au complémentaire de l'érosion de X^C :

$$X \oplus \check{B} = (X^C \ominus \check{B})^C. \quad (11)$$



a. dilatation

b. érosion

FIG. 4: Dilatation et érosion binaires hexagonales

On peut encore dire que lorsqu'on érode un ensemble, on dilate automatiquement son complémentaire. En ce sens, les transformations morphologiques agissent sur tout l'espace et pas simplement sur l'ensemble explicitement indiqué. Les notations \oplus et \ominus utilisées ont été en fait introduites par Minkowski [5]. En effet, dilater (resp. éroder) X par B revient à effectuer l'addition (resp. la soustraction) de Minkowski de X et \tilde{B} [5].

Ces deux notions s'étendent aisément au cas numérique [6]. Il suffit pour cela d'introduire la notion de sous-graphe d'une fonction. Soit f une fonction de \mathbb{R}^2 dans \mathbb{R} ; son sous-graphe $G(f)$ est l'ensemble des points (x, y) de $\mathbb{R}^2 \times \mathbb{R}$ tels que $y \leq f(x)$. On peut donc dilater et éroder cet ensemble par un élément structurant tridimensionnel B . On montre alors que la dilatation (resp. l'érosion) d'un sous-graphe est à nouveau le sous-graphe d'une fonction, appelée dilatée (resp. érodée) de f par B , et notée $f \oplus \tilde{B}$ (resp. $f \ominus \tilde{B}$). La figure 5 montre un exemple de dilatation d'une fonction numérique—dans le cas mono-dimensionnel, c'est-à-dire lorsque $f : \mathbb{R} \rightarrow \mathbb{R}$.

Très souvent en Morphologie Mathématique, on se limite pour des raisons de simplicité algorithmique à des éléments structurants *plans*, c'est-à-dire inclus dans $\mathbb{R}^2 \times \{0\}$. On montre que dans ce cas particulier, la dilatation et l'érosion d'une fonction f par un élément structurant plan B peuvent s'écrire:

$$\forall x \in \mathbb{R}^2 \quad \begin{cases} (f \oplus \tilde{B})(x) = \sup_{y \in B_x} f(y) \\ (f \ominus \tilde{B})(x) = \inf_{y \in B_x} f(y) \end{cases} \quad (12)$$

La figure 6 montre—dans le cas mono-dimensionnel—comment on transforme une fonction en la dilatant ou l'érodant par un élément structurant "plan". On voit en particulier que la dilatation élargit les pics et bouche les vallées alors que l'érosion a l'effet inverse. Quant à la figure 7, elle illustre l'effet de ces transformations sur une image numérique réelle.

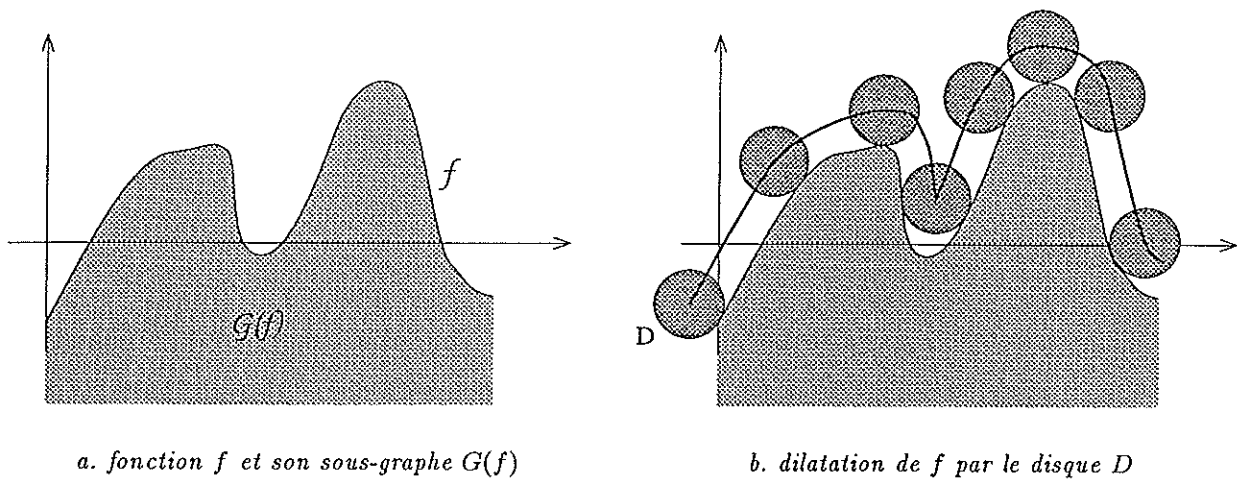


FIG. 5: Dilatation numérique

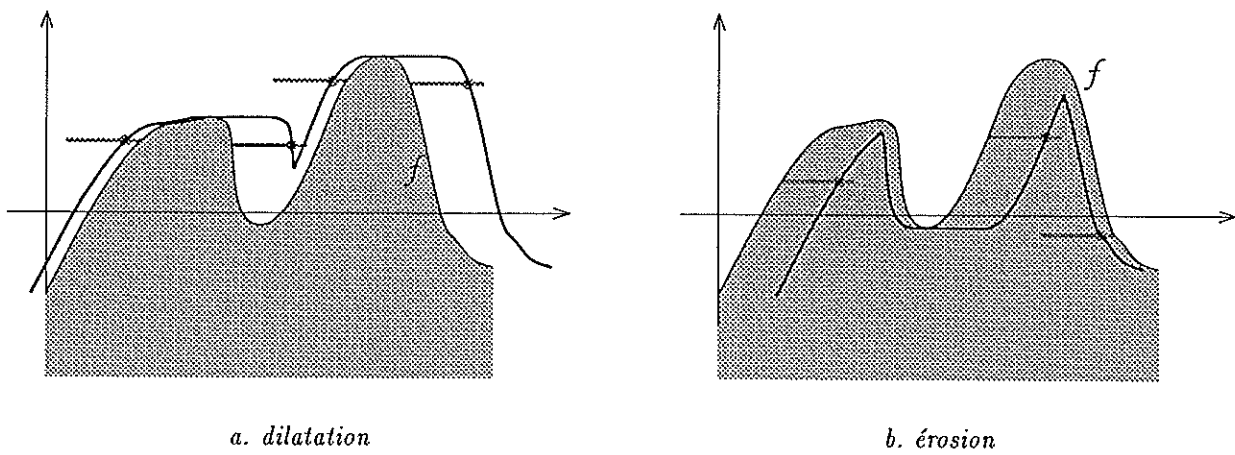


FIG. 6: Dilatation et érosion de fonctions par des éléments structurants plans

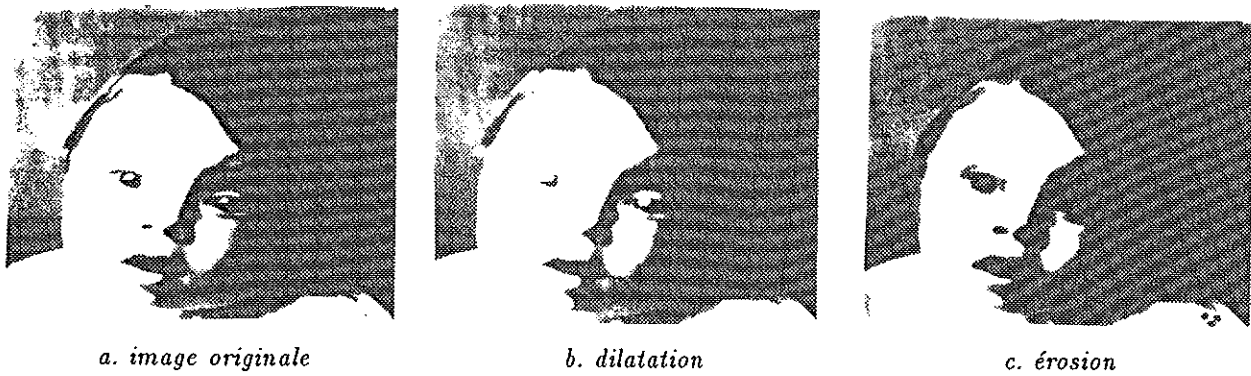


FIG. 7: Dilatation et érosion numériques par un hexagone de taille 2

2.3 Transformation en tout ou rien, amincissements, épaisissements

Considérons à présent un élément structurant B constitué de deux parties disjointes B_1 et B_2 (ayant la même origine). Un tel élément permet de définir une nouvelle transformation appelée *transformation en tout ou rien* (ou *HMT*, de l'anglais *Hit or Miss Transformation*); cette opération, notée \star , est plus générale que l'érosion et est définie de la manière suivante:

$$\forall X \in \mathbb{R}^2, \quad X \star B = (X \ominus \check{B}_1) \cap (X^C \ominus \check{B}_2). \quad (13)$$

La transformation en tout ou rien n'est en général ni croissante, ni extensive, ni idempotente... Elle est surtout utilisée pour définir l'amincissement $X \circ B$ de X par B et l'épaissement $X \odot B$ de X par B :

$$X \circ B = X / (X \star B) \quad (14)$$

$$X \odot B = X \cup (X \star B) \quad (15)$$

L'amincissement est une opération anti-extensive tandis que l'épaissement est extensif. En outre, l'opération duale de l'amincissement par $B = (B_1, B_2)$ est l'épaissement par $B' = (B_2, B_1)$:

$$X \circ B = (X^C \odot B')^C, \quad \text{avec } B = (B_1, B_2) \text{ et } B' = (B_2, B_1). \quad (16)$$

Épaissements et amincissements peuvent s'étendre également au cas numérique. Ils sont utilisés principalement dans des algorithmes générant des squelettes connexes et des squelettes par zones d'influence (Voir § 3.1).

2.4 Ouvertures et fermetures

Comme nous l'avons dit en § 2, les transformations morphologiques et en particulier l'érosion et la dilatation peuvent être itérées. On montre que sous certaines conditions sur lesquelles nous reviendrons (Voir § 2.4.1), dilater n fois un ensemble X (ou une fonction f) revient à la dilater par un élément structurant homothétique. La remarque précédente s'applique également à l'érosion, par dualité. Mais d'un autre côté, on peut aussi concaténer érosion et dilatation; on définit ainsi deux nouvelles transformations: *l'ouverture* et *la fermeture*. Dans le cas binaire par exemple, l'ouverture d'un ensemble X par un élément structurant B , notée $(X)_B$, est définie par

$$(X)_B = (X \ominus \check{B}) \oplus B. \quad (17)$$

Pour ouvrir X par l'élément structurant B , on érode donc d'abord X par B , puis on dilate le résultat par le transposé \check{B} de B (en effet, $\check{\check{B}} = B$...). L'opération d'ouverture ainsi introduite est anti-extensive. La figure 8 montre un exemple d'ouverture binaire par un disque.

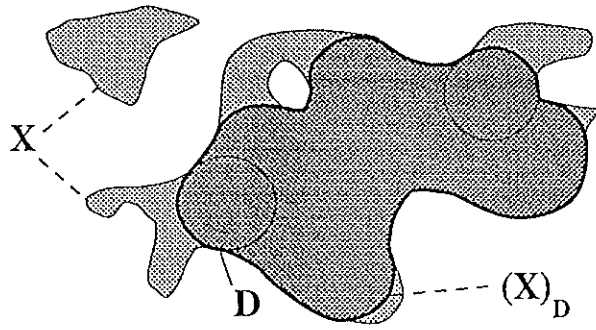


FIG. 8: Ouverture d'un ensemble X par un disque D

De la même façon, on définit la fermeture de X par B , notée $(X)^B$, par la formule suivante:

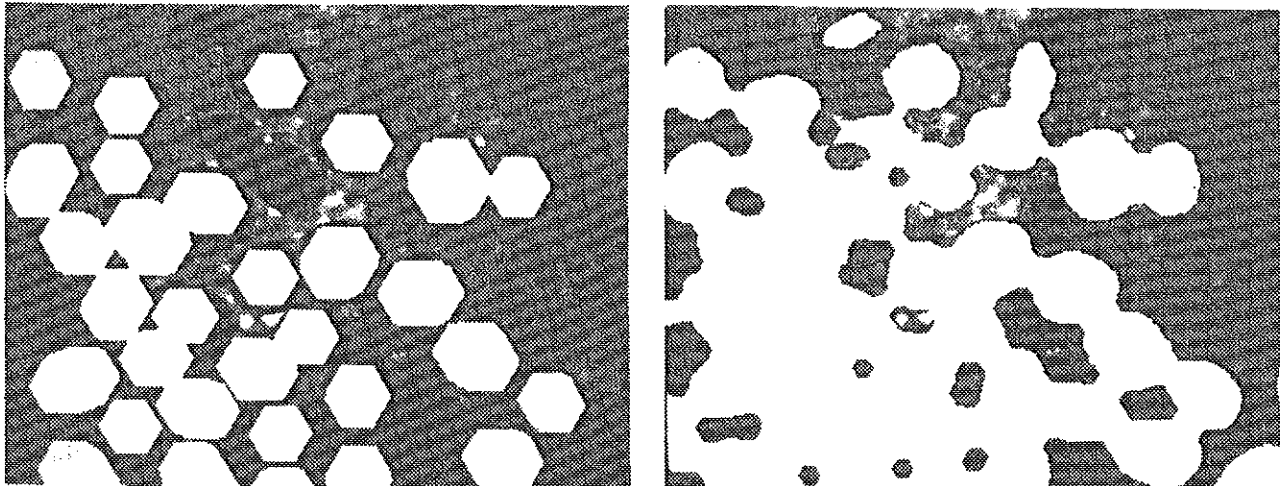
$$(X)^B = (X \oplus \tilde{B}) \ominus B. \quad (18)$$

Cette opération est ici extensive.

Comme le couple érosion-dilatation, ouvertures et fermetures sont des opérations duales:

$$(X)_B = ((X^C)^B)^C \quad (19)$$

La figure 9 illustre l'effet de ces transformations sur notre image binaire de grains de café. On peut



a. ouverture de taille 12

b. fermeture de taille 4

FIG. 9: Ouverture et fermeture binaires hexagonales

remarquer que toutes deux ont tendance à lisser les contours des ensembles, l'ouverture en supprimant les proéminences et la fermeture, en comblant les golfes. En outre, l'ouverture peut déconnecter les ensembles et faire disparaître les composantes connexes trop petites, alors que la fermeture tend à

relier des particules et à boucher leurs trous. Pour les raisons précédentes, ouvertures et fermetures ont deux utilisations majeures en traitement d'images, les *granulométries* et le *filtrage*.

Les définitions que nous avons donné de l'ouverture et de la fermeture s'étendent sans difficulté aucune au cas numérique. La figure 10 montre—dans le cas mono-dimensionnel—comment agissent ces transformations sur des fonctions. On peut entre autres remarquer à nouveau les effets de lissage

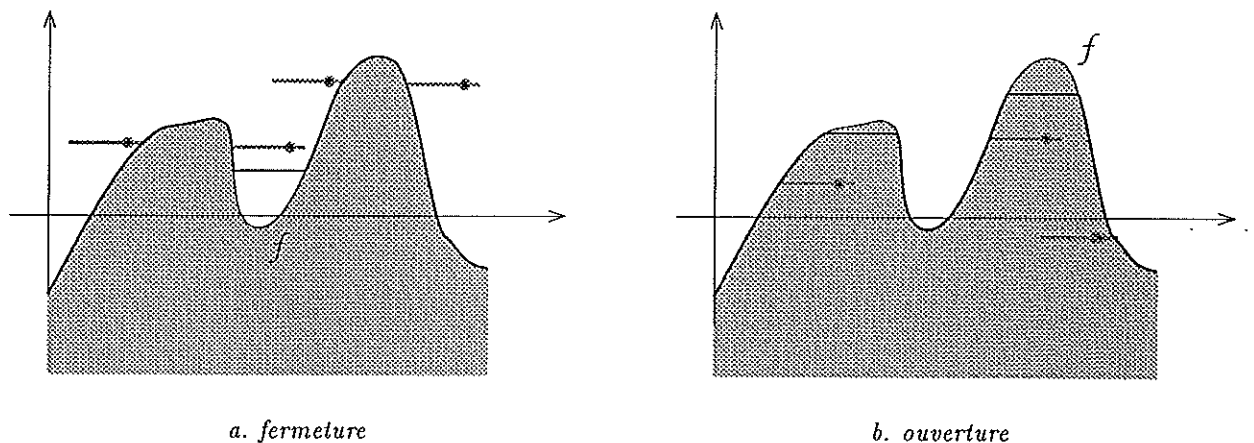


FIG. 10: Ouverture et fermeture de fonctions par des éléments structurants plans

obtenus: l'ouverture tend à supprimer les pics alors que la fermeture bouche les vallées. La figure 11 montre quant à elle le résultat d'une ouverture et d'une fermeture hexagonale sur l'image 7a.

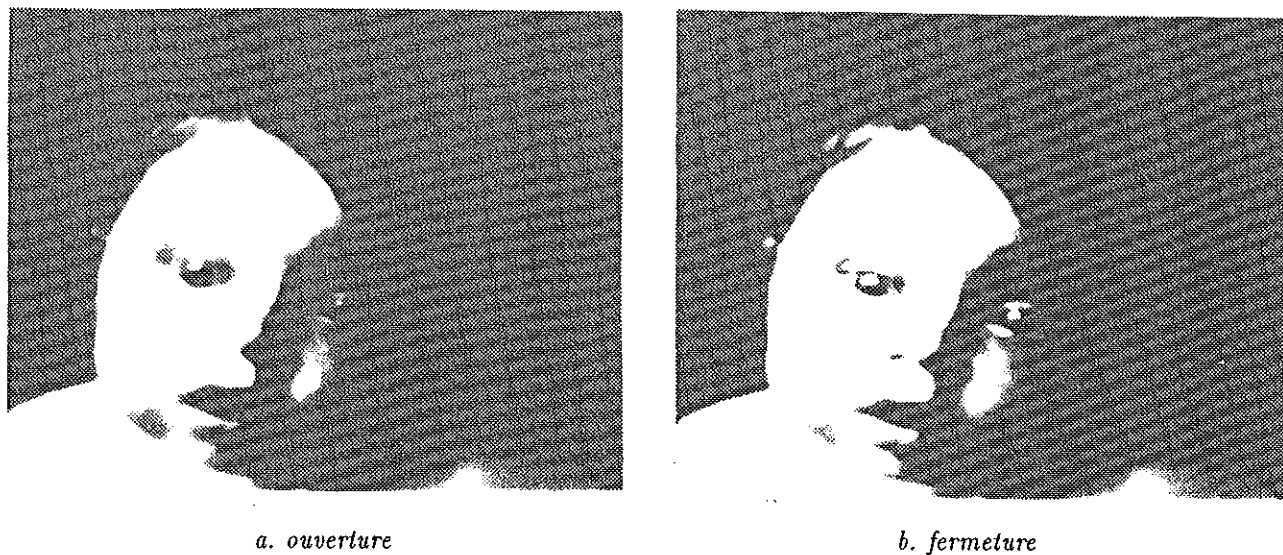


FIG. 11: Ouvertures et fermetures numériques hexagonales

2.4.1 Granulométries

Nous avons pu observer sur les exemples binaires présentés qu'une ouverture faisait disparaître les particules (c'est-à-dire les composantes connexes) plus petites que l'élément structurant utilisé (i.e. ne pouvant pas le contenir). Il semble donc possible de "tamiser" un ensemble en utilisant cette opération avec des éléments structurants de plus en plus grands. Pour formaliser quelque peu ce type de transformation, on introduit la notion de *granulométrie* et de *transformation granulométrique* [7]. Soit $(\phi_\lambda)_{\lambda \in \mathbb{R}^+}$ une famille de transformations dépendant d'un paramètre unique λ . Cette famille constitue une granulométrie si et seulement si elle vérifie les trois propriétés suivantes:

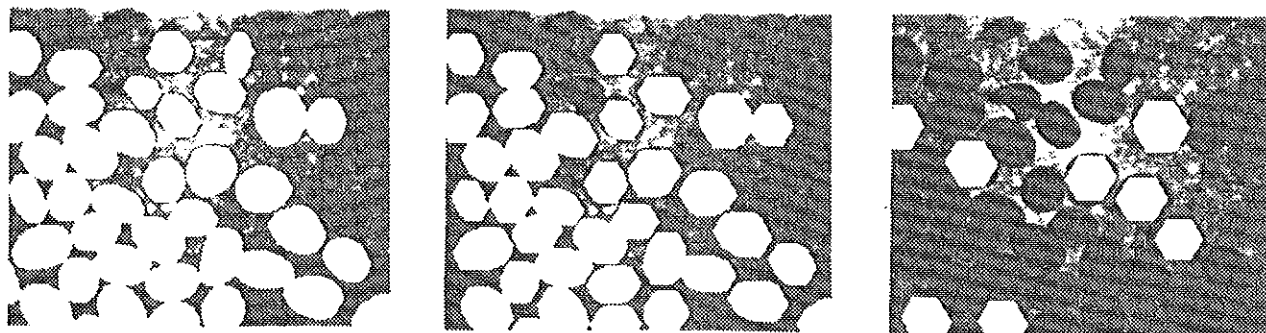
$$\forall \lambda \in \mathbb{R}^+, \quad \phi_\lambda \text{ est croissante,} \quad (20)$$

$$\forall \lambda \in \mathbb{R}^+, \quad \phi_\lambda \text{ est antiextensive,} \quad (21)$$

$$\forall (\lambda, \mu) \in \mathbb{R}^{+2}, \quad \phi_\lambda \phi_\mu = \phi_\mu \phi_\lambda = \phi_{\max(\lambda, \mu)} \quad (22)$$

Remarquons que la propriété 22 implique l'idempotence des ϕ_λ .

On montre que si B est un convexe, la famille des ouvertures par les homothétiques $(\lambda B)_{\lambda \in \mathbb{R}^+}$ de ce convexe est une granulométrie (Voir figure 12). Insistons sur le fait que l'élément structurant B



a. ouverture de taille 6

b. ouverture de taille 10

c. ouverture de taille 14

FIG. 12: Granulométrie par ouvertures de l'image des grains de café

utilisé doit impérativement être convexe pour que la propriété 22 soit vérifiée, c'est-à-dire pour que l'opération ait un sens physique. De plus, il n'est absolument pas nécessaire d'introduire la notion de particule pour définir les granulométries morphologiques: la granulométrie n'est pas associée à la notion de composante connexe, à la différence des opérations usuelles de tamisage, qui ne peuvent s'appliquer à des milieux continus.

En particulier, on peut également faire la granulométrie par ouvertures du complémentaire X^C d'un ensemble X (Voir figure 13). Or, on a vu qu'effectuer des ouvertures sur X^C revient à effectuer des fermetures sur X . Plutôt que de manipuler deux types de transformations granulométriques—les

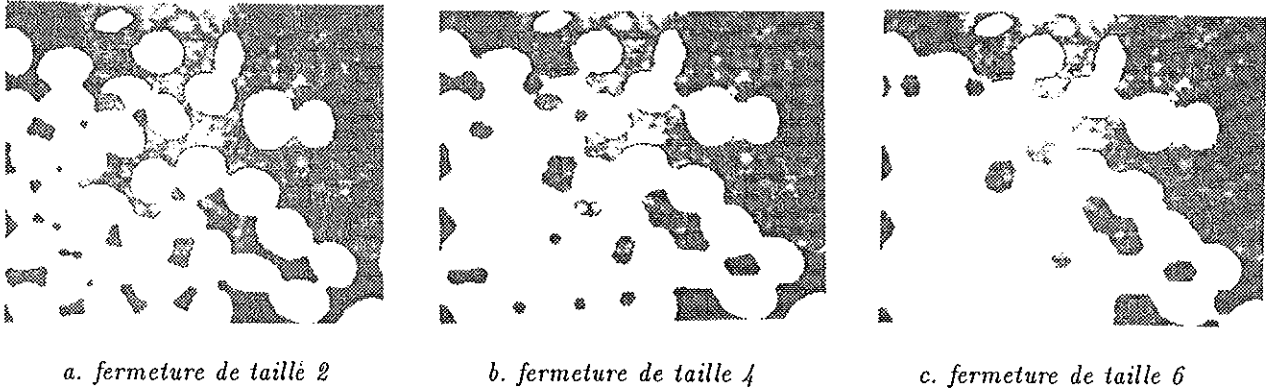


FIG. 13: Granulométrie par fermetures de l'image des grains de café

ouvertures par λB de X et de X^C , on peut généraliser la notion de granulométrie à des images biphasées. Il suffit pour cela de transformer la propriété 21 en:

$$\begin{cases} (i) \Phi_0 = \text{identité} \\ (ii) \forall (\lambda, \mu) \in \mathbb{R}^2, \lambda \geq \mu \implies \forall X \subset \mathbb{R}^2, \Phi_\lambda(X) \subset \Phi_\mu(X) \end{cases} \quad (23)$$

Dans le cas des granulométries par ouvertures, $(\Phi_\lambda)_{\lambda \in \mathbb{R}}$ est constitué de deux types de transformations et on a:

$$\begin{aligned} \Phi_\lambda(X) &= (X)^{-\lambda B} \quad \text{pour } \lambda < 0 \\ \Phi_0(X) &= X \\ \Phi_\lambda(X) &= (X)_{\lambda B} \quad \text{pour } \lambda > 0 \end{aligned} \quad (24)$$

On peut alors associer facilement une mesure à chaque $\Phi_\lambda(X)$ et tracer ainsi des courbes granulométriques [8]. De même, à tout ensemble X , correspond une fonction g_X appelée *fonction granulométrique*. A chaque point x de X , cette fonction associe la taille λ du plus grand ouvert $(X)_{\lambda B}$ de X contenant x . Cette fonction synthétise toute l'information granulométrique (Voir figure 14). Cette notion de granulométrie s'étend sans restriction aucune au cas numérique. Remarquons enfin

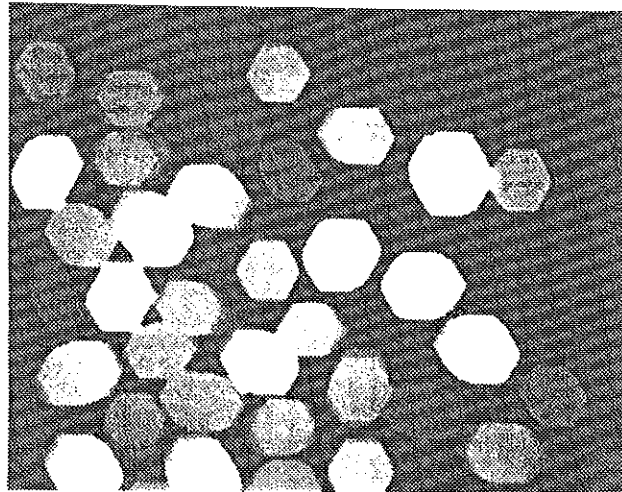


FIG. 14: Fonction granulométrique des grains de café

que parmi tous les convexes B utilisables pour faire des granulométries, les boules et leurs approximations digitales —hexagones (dénotés par $H, 2H, \dots, nH$), carrés, rhombododécaèdres, etc. . . —sont privilégiés.

2.4.2 Filtrage

Nous avons vu que les opération d'ouverture et de fermeture avaient tendance à lisser les contours des objets. Ces propriétés de lissage s'observent aussi dans le cas numérique, où elles sont particulièrement intéressantes. En fait, ouvertures et fermetures sont des cas particuliers de filtres morphologiques, un filtre morphologique étant une transformation Φ dotée des deux propriétés suivantes:

- (i) Φ est croissante
- (ii) Φ est idempotente

L'étude théorique de ces filtres morphologiques permet d'exhiber un certain nombre de transformations, combinaisons d'ouvertures et de fermetures de différentes tailles [9, chap. 5–10]. Elles sont particulièrement intéressantes dans tous les problèmes d'amélioration des images à teintes de gris. Nous ne traiterons pas davantage ici de ces filtres, leur présentation sortirait du cadre de cet article. Nous présenterons simplement un exemple de leur utilisation au § 4.4.

2.5 Boules maximales et squelettes

D'après ce qui précède, on constate que l'ouverture d'un ensemble X par un disque de rayon r est l'ensemble des points de X balayés par ce disque lorsque celui-ci est astreint à rester dans X . De plus, lorsque l'on fait une granulométrie par ouvertures (par des boules $\{\lambda B\}_{\lambda \in \mathbb{R}^+}$), on décompose l'ensemble X en une union de boules de plus en plus grandes. Ainsi, l'ouvert de taille λ de X —i.e. l'ouvert de X par λB —est la partie de X dans laquelle on peut inclure des boules de rayon λ . De même, $(X)_{\mu B}$ est la partie de X dans laquelle on peut inclure des boules de rayon μ . Cependant, si $\lambda < \mu$, les boules de rayon λ qui sont également incluses dans $(X)_{\mu B}$ seront recouvertes par au moins une boule de rayon μ . Elles sont donc redondantes pour la description de X en termes de boules.

Pour mieux formaliser cette représentation intuitive de X , on introduit la notion de *boule maximale*. Une boule B incluse dans X est dite maximale si et seulement si il n'existe aucune autre boule de X la contenant:

$$\forall B' \text{ boule, } B' \subset X, \quad B \subset B' \implies B' = B. \quad (25)$$

On voit que les boules redondantes de rayon λ du paragraphe précédent ne sont justement pas maximales. Il est à présent facile de montrer que tout ensemble X peut s'exprimer comme la réunion de

ses boules maximales:

$$X = \bigcup_{B \text{ boule max de } X} B \quad (26)$$

En fait, on peut encore simplifier la représentation: seuls suffisent les centres des boules maximales et les rayons associés. On appellera squelette de X le lieu des centres des boules maximales [10]:

$$S(X) = \bigcup \{x \in X, \exists r \geq 0 \text{ tel que } B(x, r) \text{ maximale dans } X\}. \quad (27)$$

La figure 15 présente quelques exemples de squelettes et de boules maximales. A tout point du squelette

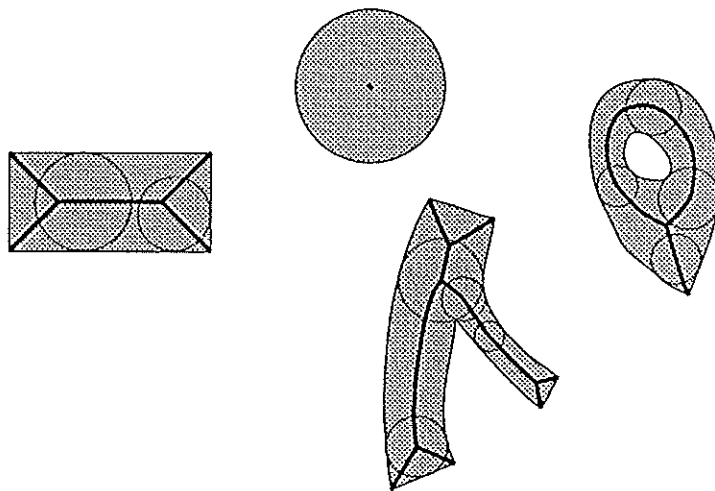


FIG. 15: Quelques exemples de squelettes et de boules maximales

$S(X)$, on peut associer le rayon de la boule maximale correspondante. On définit ainsi un fonction q_x sur $S(X)$, à valeurs dans \mathbb{R}^+ , qui est appelée *fonction d'étanchéité* (*quench function* en anglais) ou encore *fonction d'extinction* (Voir figure 16b):

$$q_x \begin{cases} S(X) & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto r, B(x, r) \text{ maximale} \end{cases} \quad (28)$$

De plus, on montre que le squelette de X peut s'écrire comme l'union des résidus des ouverts successifs de X [11]; dans le cas digital, la formule est la suivante:

$$S(X) = \bigcup_{n=0}^{+\infty} [(X \ominus nB) / (X \ominus nB)_B] \quad (29)$$

En effet, pour $n \geq 0$ donné, $X \ominus nB$ est le lieu des centres des boules de taille n incluses dans X ; $(X \ominus nB)_B$ est lui le lieu des centres de ces boules de taille n qui sont incluses dans une boule de taille

$n + 1$. Ces boules ne sont donc pas maximales. $(X \ominus nB)/(X \ominus nB)_B$ est donc le lieu des centres des boules maximales de taille n , et en faisant l'union pour tous les $n \geq 0$, on obtient le squelette de X .

La fonction d'étanchéité est particulièrement utile pour compacter des images binaires [12, page 49]. En outre, cette représentation d'un ensemble X par le couple $(S(X), q_X)$ permet de définir certaines transformations en associant à q_X une nouvelle fonction $\gamma(q_X)$: partant de X donné par son couple $(S(X), q_X)$, on obtient alors par reconstruction un nouvel ensemble $\gamma(X)$ défini par:

$$\gamma(X) = \bigcup_{x \in S(X)} B(x, \gamma(q_X)(x)) \quad (30)$$

En particulier, on peut obtenir érosion, dilatation et ouverture par ce procédé [12, pages 50–52] (Voir tableau 1). On constatera que la fermeture ne peut pas s'obtenir par ce moyen. D'autre part, le

Transformation	Fonction $\gamma(q_X)$ associée
$X \ominus \lambda B$	$\gamma(q_X)(x) = 0$ si $x \leq \lambda$ $= q_X(x) + \lambda$ sinon
$X \oplus \lambda B$	$\gamma(q_X)(x) = q_X(x) + \lambda$
$(X)_{\lambda B}$	$\gamma(q_X)(x) = 0$ si $x \leq \lambda$ $= q_X(x)$ sinon

TABLEAU 1: Comment obtenir érosion, dilatation et ouverture à partir du couple $(S(X), q_X)$.

squelette de l'ensemble $\gamma(X)$ obtenu par ce type de transformation n'a en général aucun rapport avec le squelette $S(\gamma(X))$. Seul le squelette de l'érodé de X est inclus dans $S(X)$, mais il n'existe déjà plus de relation entre $S(X)$ et $S(X \oplus \lambda B)$.

Paradoxalement, ce squelette par boules maximales, bien que fort efficace dans la représentation des ensembles, est assez peu utilisé en pratique: en effet, il n'est pas connexe (Voir § 3.1), comme l'illustre la figure 16a.

2.6 Gradient et "Chapeau haut-de-forme"

D'autres opérations élémentaires peuvent être définies par concaténation d'érosions, de dilatations, d'ouvertures et de fermetures. En particulier, en MM numérique, la combinaison d'opérateurs morphologiques et arithmétiques permet de générer de nouvelles transformations. Parmi elles, deux sont particulièrement utilisées: le *gradient* et la transformation appelée *chapeau haut-de-forme*.

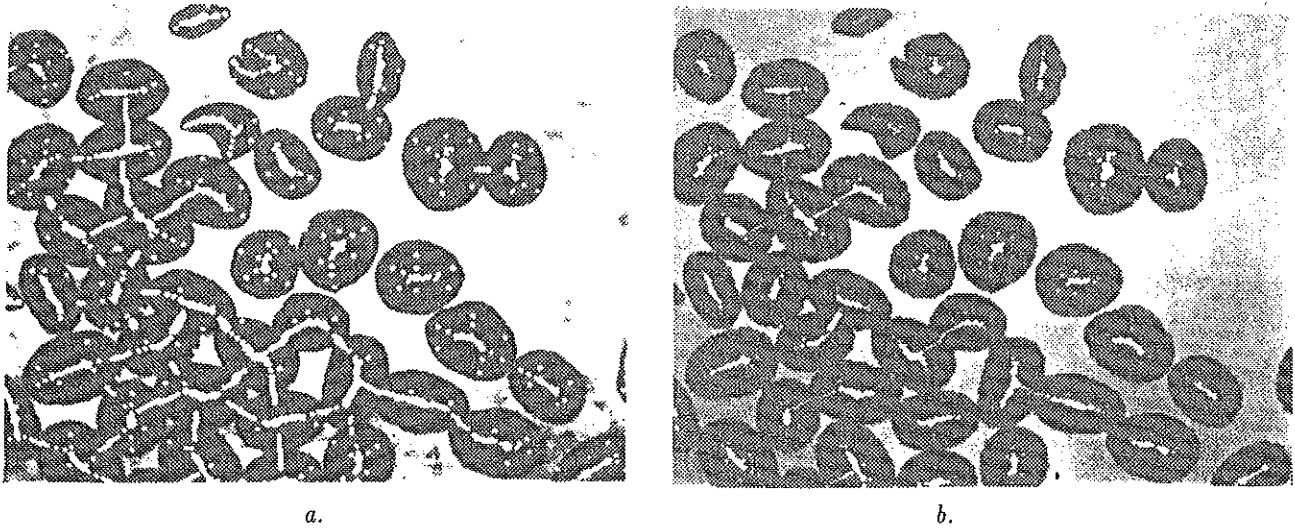


FIG. 16: Squelette par boules maximales (a) et Fonction d'étanchéité (b) sur les grains de café

2.6.1 Le gradient

Le gradient morphologique d'une fonction f est défini par:

$$\text{grad}(f) = \lim_{\epsilon \rightarrow 0} \frac{(f \oplus \epsilon B) - (f \ominus \epsilon B)}{2\epsilon}, \quad (31)$$

où ϵB désigne une boule de rayon ϵ . La digitalisation en trame hexagonale de cette transformation s'écrit:

$$\text{grad}(f) = \frac{(f \oplus H) - (f \ominus H)}{2}, \quad (32)$$

H représentant ici l'hexagone élémentaire introduit précédemment [13]. On montre que ce gradient morphologique est égal au module du gradient de f —supposée continuellement différentiable—, classiquement défini par:

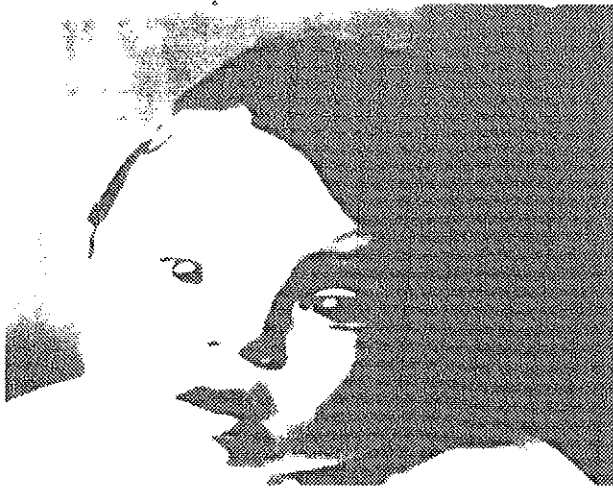
$$|\text{grad}(f)| = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{\frac{1}{2}} \quad (33)$$

De nombreuses variations de cette transformation existent. On peut notamment définir en termes d'opérateurs morphologiques des gradients directionnels et représenter les azimuts des vecteurs gradients.

La figure 17 illustre cette transformation. On remarque que les valeurs élevées du module du gradient correspondent aux zones à fort contraste de l'image. Cette propriété sera utilisée par la suite pour détecter les contours d'une image numérique (Voir § 4.3).

2.6.2 Le chapeau haut-de-forme

Le *chapeau haut-de-forme* [12, page 101] est une transformation permettant d'extraire d'une image numérique les zones claires et étroites. Ces zones sont celles qui disparaissent après une ouverture



a. image numérique originale



b. gradient

FIG. 17: Gradient morphologique

numérique. Etant donné un élément structurant B , on définit ainsi le chapeau haut-de-forme HdF_B d'une fonction f par:

$$HdF_B(f) = f - (f)_B \quad (34)$$

Cette transformation extrait les résidus de l'ouverture de f , la "taille" de ces résidus étant inférieure à la taille de l'élément structurant B utilisé. Là encore, B sera choisi préférentiellement parmi les ensembles convexes (disques, segments, etc...). On peut en particulier combiner différents homothétiques d'un même convexe ou utiliser des éléments structurants dans différentes directions afin de sélectionner de façon très précise les régions à extraire.

De manière similaire, on peut extraire les zones sombres et étroites d'une image en définissant le chapeau haut-de-forme conjugué HdF_B^* par:

$$HdF_B^*(f) = (f)^B - f \quad (35)$$

La figure 18 illustre ces opérations.

Très souvent, on filtre également les régions extraites par ces transformations en fonction de leur contraste en effectuant un seuillage des résidus. Par exemple, le seuillage au niveau h d'un chapeau haut-de-forme effectué avec une boule de rayon λ permet d'extraire de l'image les crêtes de hauteur supérieure à h et de largeur inférieure à 2λ (Voir figure 19).

2.7 Extrema d'une fonction

La notion d'*extrema* d'une fonction nous sera également très utile par la suite, c'est pourquoi il nous faut l'introduire dès maintenant. Soit $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction. Le graphe de f peut être regardé



a. chapeau "standard"



b. chapeau conjugué

FIG. 18: Transformations chapeau haut-de-forme hexagonales de l'image 17a

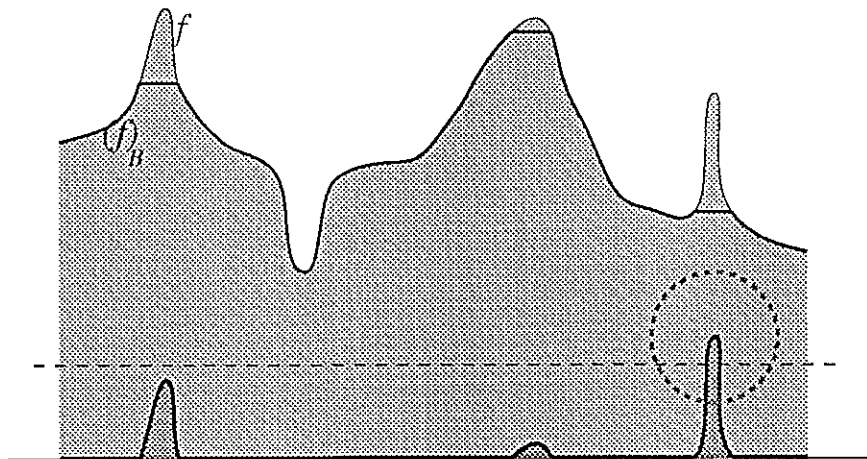


FIG. 19: "Crêtes" extraites par chapeau haut-de-forme suivi d'un seuillage

comme une surface topographique. Un *maximum* de f est alors un sommet de cette surface, c'est-à-dire une zone (connexe, mais pas forcément réduite à un point) d'où il est impossible de rejoindre un point d'altitude supérieure par un chemin jamais descendant (Voir figure 20).

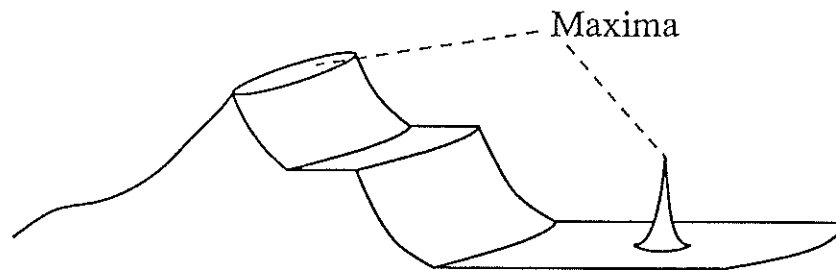


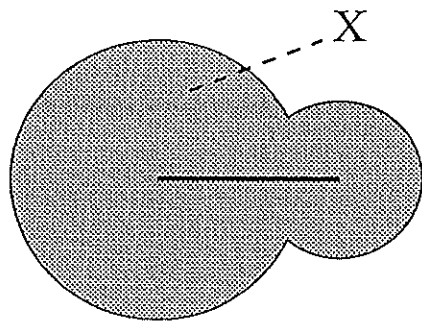
FIG. 20: Maxima d'une fonction

En particulier, si l'on considère les différents seuils de f , un maximum de cette fonction à l'altitude h sera une composante connexe du seuil $S_h(f) = \{x \in \mathbb{R}^2, f(x) \geq h\}$ ne contenant aucune composante connexe de tout seuil $S_{h'}$ où $h' > h$.

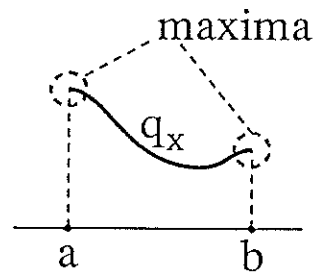
2.8 Premiers pas vers la segmentation: le marquage des objets

Les notions que nous venons d'introduire vont à présent nous servir pour la résolution d'un problème classique de segmentation binaire: la séparation de boules ou d'objets qui se recouvrent partiellement. C'est un cas de figure qui se rencontre fréquemment dans de nombreux domaines d'application du traitement d'images, comme la répartition de billes dans des matériaux frittés [8], la séparation de cellules sur frottis [12], la comptage de particules, etc..., sans oublier le problème de nos grains de café. Résoudre ce problème nécessite dans un premier temps le marquage correct des composantes que l'on veut séparer. Cette démarche est en fait très naturelle: c'est celle qu'adopterait une personne à qui on demanderait de montrer du doigt les différents objets présents dans l'image.

Examinons dans un premier temps un cas d'école, celui où l'ensemble X est constitué de deux composantes connexes relativement circulaires (Voir figure 21a). Comme nous l'avons vu, X peut être représenté par son squelette $S(X)$ et sa fonction d'étanchéité q_X . Pour mettre en évidence ses deux composantes, il paraît naturel de rechercher dans X les régions où les boules maximales sont les plus grosses. On va donc rechercher sur $S(X)$ les points correspondant à ces "plus grosses" boules maximales: ces points seront aussi les maxima de la fonction d'étanchéité q_X (Voir figure 21b).



a. squelette de X



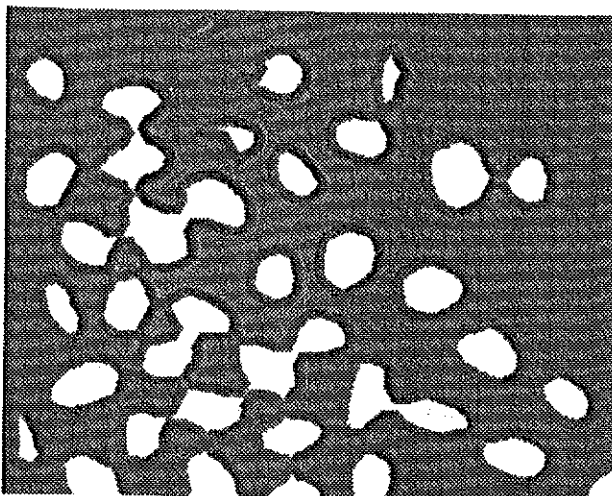
b. q_x et ses maxima

FIG. 21: Squelette et maxima de la fonction d'étanchéité

Malheureusement, il est délicat en pratique de déterminer les maxima d'une fonction définie sur un support non connexe. Nous allons donc introduire une nouvelle fonction $dist_X$, définie sur X et appelée *fonction distance*:

$$dist_X \begin{cases} X & \rightarrow \mathbb{R}^+ \\ x & \mapsto d(x, X^c) \end{cases} \quad (36)$$

A chaque point de X , $dist_X$ associe sa distance au point le plus proche du complémentaire de X (Voir figure 22). Les différentes lignes de niveau de cette fonction correspondent aux frontières des érodés successifs de X . On peut montrer que les maxima de $dist_X$ sont exactement ceux de la fonction d'étanchéité q_x . Dans le cas discret, ces maxima peuvent d'ailleurs être mis en évidence de la manière suivante: soit X_n l'érodé de taille n de X ; X_n est le seuil entre $n + 1$ et $+\infty$ de la fonction distance. Effectuons une érosion élémentaire—de taille 1—de cet érodé. Toute composante connexe C de X_n qui disparaît après cette nouvelle érosion est par définition des maxima (Voir § 2.7) un maximum de la fonction distance. L'ensemble de ces maxima constitue l'érodé ultime de X (Voir figure 22b).



a. fonction distance (clair \Leftrightarrow distance importante)



b. érodé ultime = maxima de la fonction distance

FIG. 22: Fonction distance et érosion ultime

L'exemple des grains de café sur la figure 22 montre bien que l'érosion ultime constitue un marqueur au sens où nous l'avons indiqué plus haut. Chaque composante de l'ensemble de départ—les grains de café—se trouve marquée par cette procédure. De plus, le marquage est biunivoque: aux erreurs de parité dues à la digitalisation près, à chaque grain correspond un et un seul marqueur. En outre, nous avons vu que ces marqueurs peuvent être définis comme les extrema d'une fonction—la fonction distance.

On essaiera de montrer par la suite que cette démarche est assez générale en morphologie mathématique. Toute segmentation est pilotée par une—ou plusieurs—fonctions dites *de marquage*. La mise en évidence des extrema de ces fonctions constitue la première étape de la segmentation: la génération de marqueurs des objets à segmenter. On verra que cette fonction de marquage sert également dans le processus de segmentation lui-même (Voir § 4).

3 TRANSFORMATIONS HOMOTOPIQUES ET GÉODÉSQUES

Une fois les marqueurs déterminés, passons à l'étape de segmentation proprement dite, c'est-à-dire à l'extraction des objets marqués. Ce faisant, il faudra nécessairement utiliser des opérateurs extensifs. Pour préserver le marquage effectué, ces opérateurs devront vérifier les deux critères suivants:

1. Ne pas connecter de marqueurs, conserver leur nombre et leurs relations de voisinage. De tels opérateurs morphologiques existent et sont appelés *transformations homotopiques*.
2. Travailler uniquement sur l'ensemble X à segmenter. Il est en effet inutile que ces opérateurs "débordent" de X , puisque les points de X^c sont sans intérêt. Pour formaliser ce type d'opérateurs morphologiques, on introduit la notion de *géodésie* et de *transformation géodésique*.

3.1 Homotopie et squelettes

Soit $X \subset \mathbb{R}^2$ un ensemble connexe, et x et y deux points de X . Considérons deux chemins différents C_1 et C_2 reliant x à y . C_1 et C_2 sont dits *homotopiques* si et seulement si il est possible de passer de C_1 à C_2 par une suite continue de déformations continues. Par exemple sur la figure 23, les chemins C_2 et C_3 sont homotopiques; ce n'est en revanche pas le cas pour C_1 et C_2 , car il est impossible d'appliquer C_2 sur C_1 par déformations continues à cause du trou existant dans l'ensemble. On montre que la relation d'homotopie entre chemins de X est une relation d'équivalence.

X sera dit *simplement connexe* si et seulement si il n'existe qu'une et une seule classe d'équivalence de chemins entre deux points quelconques de X . Ceci revient à dire que X est *connexe et sans trou*.

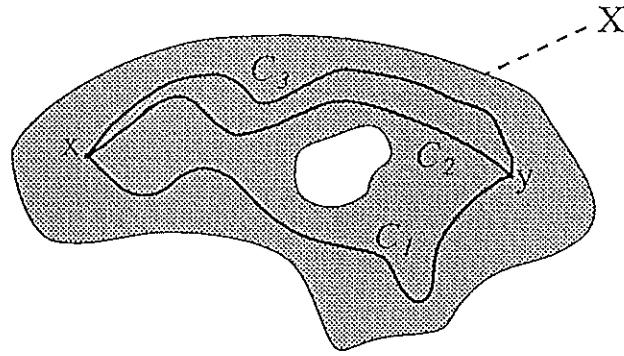


FIG. 23: La paire de chemins (C_2, C_3) est homotopique alors que la paire (C_1, C_2) ne l'est pas

Une transformation $\Phi: \mathcal{P}(\mathbb{R}^2) \rightarrow \mathcal{P}(\mathbb{R}^2)$ sera dite homotopique si et seulement si

$$\forall X \subset \mathbb{R}^2, \exists \Psi: \mathbb{R}^2 \times [0;1] \rightarrow \mathbb{R}^2, \text{ application continue,} \\ \text{telle que } \begin{cases} \Psi(X, 0) = X, \\ \Psi(X, 1) = \Phi(X). \end{cases} \quad (37)$$

De façon plus imagée, une transformation homotopique transforme un ensemble X en un ensemble Y par une suite continue de déformations continues. Ainsi, un ensemble simplement connexe sera transformé en un ensemble simplement connexe, un ensemble avec un trou, en un ensemble avec un trou, etc. . . Plus généralement, si $X \subset \mathbb{R}^2$ est composé de plusieurs composantes connexes $(X_i)_{i \in [1,n]}$, le nombre et les positions relatives de ces X_i seront préservés (une transformation homotopique conserve les classes de chemins à la fois pour X et pour X^c ; voir figure 24).

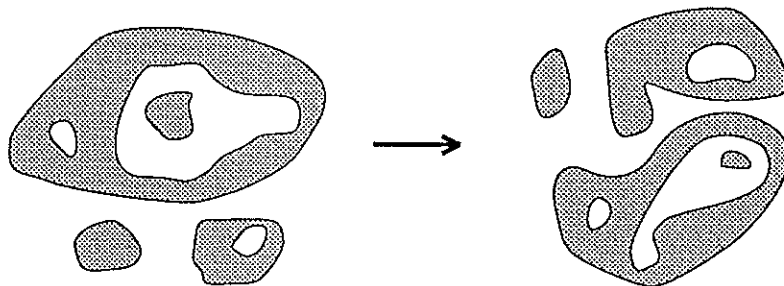


FIG. 24: Un exemple de transformation homotopique

Des exemples classiques de transformation homotopique en MM sont ceux des squelettes dits *connexes* [12, page 75]. Ils peuvent être réalisés par une séquence d'amincissements—pour le squelette

classique—ou d'épaississements—pour l'*exosquelette*, c'est-à-dire le squelette du fond—utilisant des éléments structurants définis sur la "boule élémentaire" (hexagone en trame hexagonale, carré en trame carré). Ces éléments structurants ont été sélectionnés de façon à produire des amincissements et épaississements homotopiques. Ainsi, en trame hexagonale, on montre qu'il existe 3 éléments structurants—aux six rotations près— possédant cette propriété; ils sont dénommés L , M et D [14] (Voir figure 25). L'utilisation de ces éléments structurants montre que seuls L et M permet-

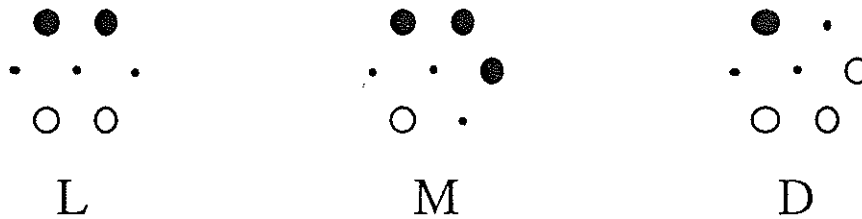


FIG. 25: Éléments structurants utilisés pour construire des squelettes connexes

tent d'obtenir des transformations de type squelette—c'est-à-dire une espèce d'axe médian des objets (Voir figure 26). En fait, on peut montrer qu'il existe des relations étroites entre ces squelettes de type L et M et le squelette par boules maximales décrit en 2.5 [15].

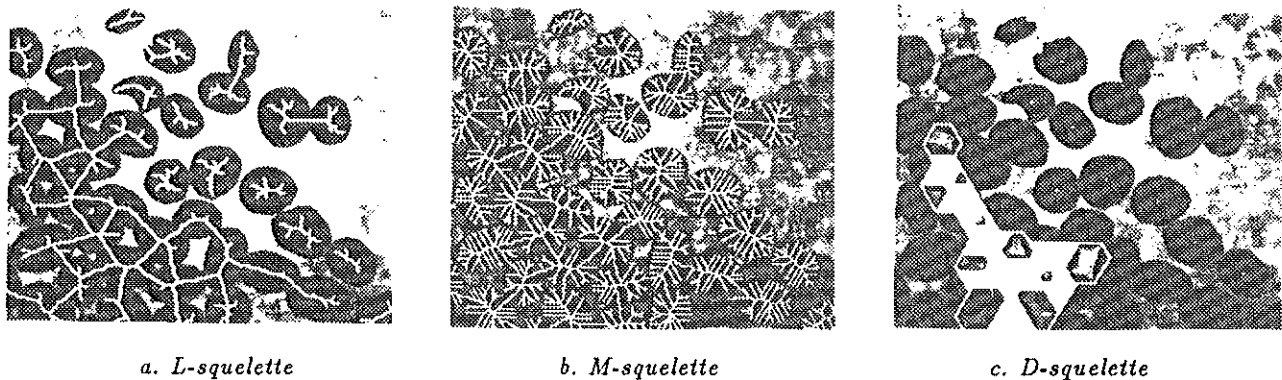


FIG. 26: Squelettes obtenus par amincissements successifs avec les éléments structurants L , M et D

On peut montrer également que ces squelettes connexes sont les lignes de crête de la fonction distance et peuvent être obtenus comme tels [9, chap. 13] (Voir figure 27). Cette approche permet d'ailleurs d'étendre la notion de squelette au cas des images numériques [16]. La figure 28 montre une image de chromosomes et le squelette numérique correspondant (en fait dans ce cas, on a effectué un squelette de l'image inversée, c'est-à-dire qu'on a suivi les "fonds de vallée").

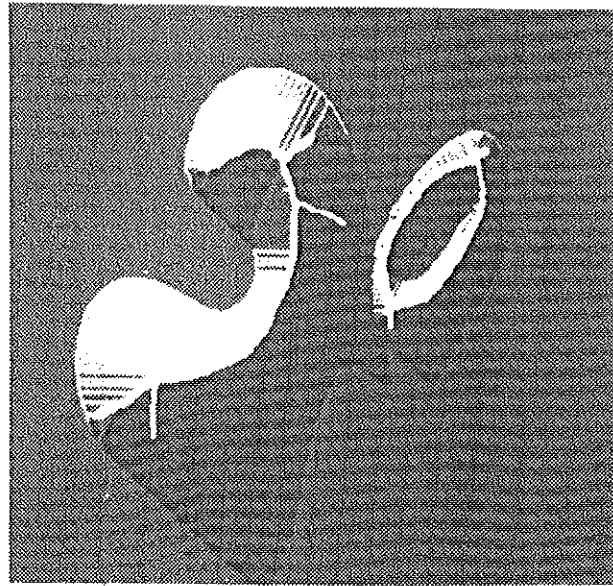


FIG. 27: Le squelette suit les crêtes de la fonction distance (artificiellement ombrée ici)



a. image de chromosomes



b. squelette numérique

FIG. 28: Squelette numérique

Le squelette par épaisissements est quant à lui principalement utilisé pour produire le *squelette par zones d'influence* (ou *Skiz*, de l'anglais *Skeleton by influence zones*). Ce *Skiz* est défini de la manière suivante: soit $X \subset \mathbb{R}^2$ un ensemble formé de n composantes connexes $\{X_i\}_{i \in [1,n]}$. On définit la zone d'influence $z(X_i)$ de X_i comme l'ensemble de points de \mathbb{R}^2 qui sont plus proches de X_i que de toute autre composante connexe de X :

$$z(X_i) = \{x \in \mathbb{R}^2, \forall j \neq i, d(x, X_i) < d(x, X_j)\} \quad (38)$$

Remarquons que la zone d'influence d'une composante connexe X_i de X n'est en général pas homotopique à X_i (Voir figure 29). L'ensemble constitué des frontières des différentes zones d'influence est par définition le squelette par zones d'influence de X [17]; la figure 29 en présente un exemple.

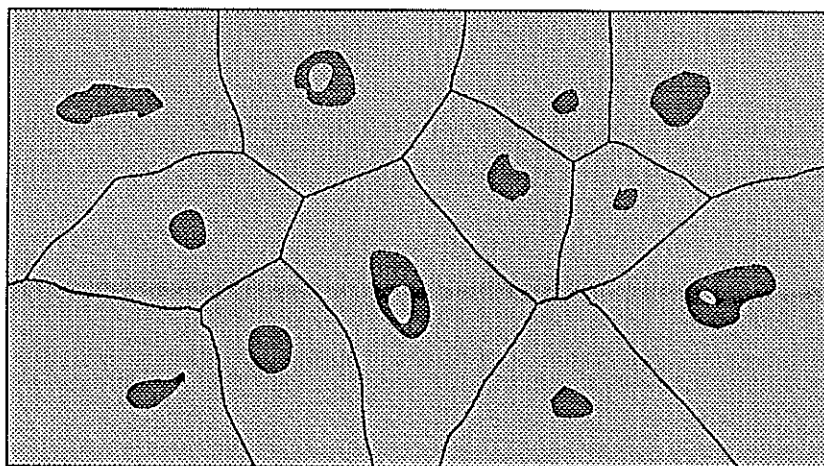


FIG. 29: Un exemple de squelette par zones d'influence

Le *Skiz* est généralement obtenu en ne conservant du squelette par épaisissements de X (*exosquelette*) que les arcs fermés. Pour supprimer les arcs en trop, on utilise un amincissement particulier de $S(X^C)$ appelé *ébarbulage*. Soulignons que l'ensemble produit après cette étape peut encore dans certains cas particuliers être un sous-ensemble du *Skiz* vrai (Voir figure 30).

3.2 Géodésie

3.2.1 Distance géodésique

Le concept de chemin a déjà été utilisé pour définir les transformations homotopiques. Il va être repris ici pour définir une distance associée à un ensemble X de \mathbb{R}^2 . Soit $(x, y) \in X^2$; la distance géodésique

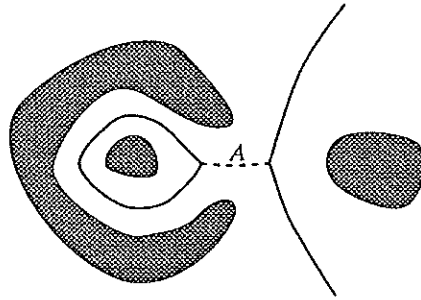


FIG. 30: Les pièges du *Skiz*: l'arête A est de trop!

$d_X(x, y)$ entre x et y est la borne inférieure des longueurs des chemins entre x et y dans X —s'il en existe:

$$d_X(x, y) = \inf\{l(C_{x,y}), C_{x,y} \text{ chemin reliant } x \text{ à } y \text{ dans } X\}. \quad (39)$$

Dans le cas où x et y appartiennent à des composantes connexes de X différentes, on posera par convention $d_X(x, y) = +\infty$. En ce sens, d_X est une distance "généralisée" (Voir figure 31).

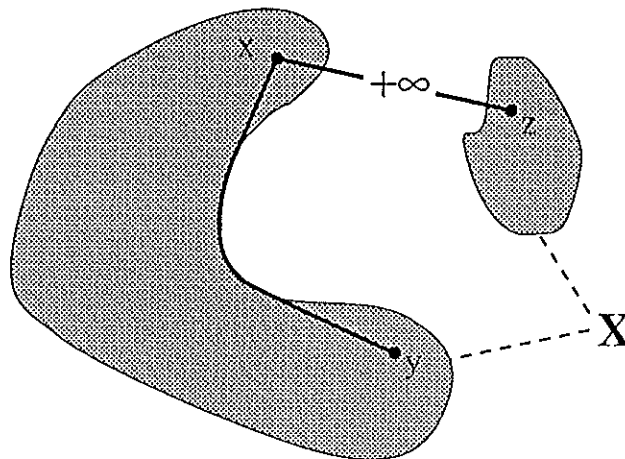


FIG. 31: Distance géodésique

Nous appellerons *boule géodésique* de rayon r centrée en $x \in X$ l'ensemble $B_X(x, r)$ défini par

$$B_X(x, r) = \{y \in X, d_X(x, y) < r\}. \quad (40)$$

3.2.2 Transformations géodésiques

Soit à présent Y un sous-ensemble de X muni de la distance géodésique d_X induite par X . On peut définir l'érosion $E_X^{(r)}(Y)$ et la dilatation $D_X^{(r)}(Y)$ de Y par des boules géodésiques de rayon r donné [18,19] de la manière suivante:

$$E_X^{(r)}(Y) = \{y \in X, B_X(y, r) \subset Y\} \quad (41)$$

$$D_X^{(r)}(Y) = \{y \in X, B_X(y, r) \cap Y \neq \emptyset\} \quad (42)$$

Ces définitions sont tout à fait analogues aux définitions classiques euclidiennes de l'érosion et de la dilatation par des boules. Les différences proviennent essentiellement de deux facteurs:

- X est désormais l'espace de travail (au lieu de \mathbb{R}^2), autrement dit le résultat de ces transformations sera toujours contenu dans X .
- Les éléments structurants utilisés sont définis à l'aide d'une distance et n'agissent pas en tant qu'éléments "rigides" comme dans le cas euclidien. La gamme d'éléments structurants disponibles est ici très pauvre.

La figure 32 montre un exemple de dilatation et d'érosion géodésiques. En pratique, la dilatation

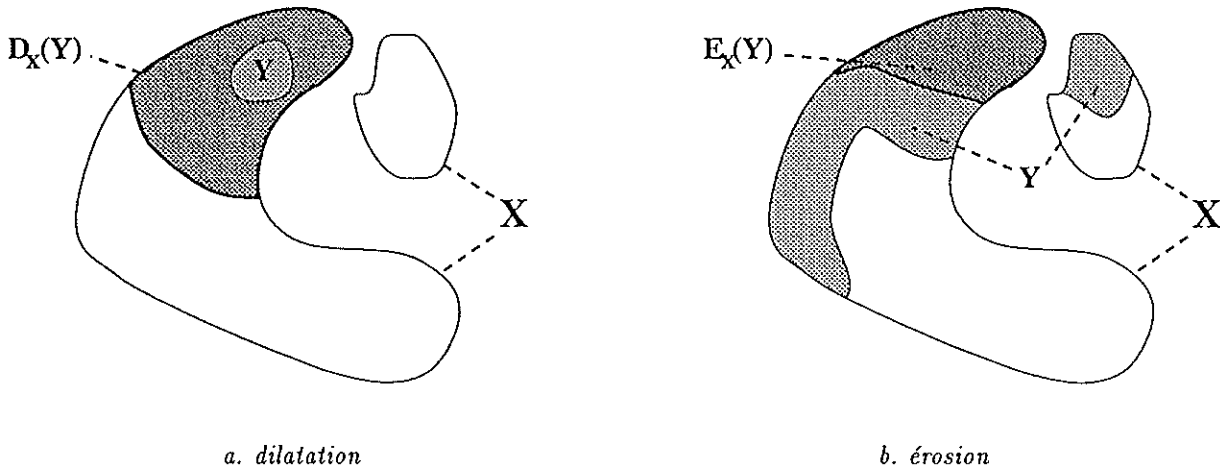


FIG. 32: Opérations géodésiques de base

géodésique de taille n est réalisée sur une image digitale en itérant n fois une dilatation géodésique de taille 1. Cette dernière s'obtient simplement en intersectant le dilaté euclidien de taille 1 de l'ensemble Y avec l'"ensemble-espace" X :

$$D_X^{(1)}(Y) = (Y \oplus H) \cap X \quad (43)$$

$$D_X^{(n)}(Y) = \underbrace{D_X^{(1)}(D_X^{(1)}(\dots D_X^{(1)}(Y)))}_{n \text{ fois}} \quad (44)$$

L'érosion $E_X^{(n)}(Y)$ s'obtient aisément par dualité. Par analogie, les mêmes formules s'étendent immédiatement à des fonctions: ainsi, on peut dilater géodésiquement la fonction $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ par rapport à une fonction majorante $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de la manière suivante:

$$D_f^{(1)}(g) = \inf(g \oplus H, f) \quad (45)$$

$$D_f^{(n)}(g) = \underbrace{D_f^{(1)}(D_f^{(1)}(\dots D_f^{(1)}(g)))}_{n \text{ fois}} \quad (46)$$

Remarquons qu'il est difficile d'exhiber ici la distance géodésique sous-jacente. L'érosion numérique géodésique de f par rapport à g (fonction minorante) est définie de façon analogue:

$$E_g^{(1)}(f) = \sup(f \ominus H, g) \quad (47)$$

$$E_g^{(n)}(f) = \underbrace{E_g^{(1)}(E_g^{(1)}(\dots E_g^{(1)}(f)))}_{n \text{ fois}} \quad (48)$$

Considérons maintenant un ensemble Y inclus dans une composante connexe X_i de X . Si l'on dilate Y géodésiquement dans X , on est sûr de ne pas déborder de X_i quelle que soit la taille de la dilatation. De plus, en dilatant suffisamment Y , on obtient X_i intégralement, c'est-à-dire la composante connexe de X marquée par Y . Ceci s'étend d'ailleurs au cas où Y marque plusieurs composantes connexes de X . La transformation que nous venons d'effectuer s'appelle *reconstruction*, est notée R_X et est définie plus formellement de la manière suivante:

$$R_X(Y) = D_X^{(+\infty)}(Y) = \lim_{r \rightarrow +\infty} D_X^{(r)}(Y). \quad (49)$$

On peut voir sur la figure 33 un exemple de reconstruction binaire. Il est clair que la reconstruction

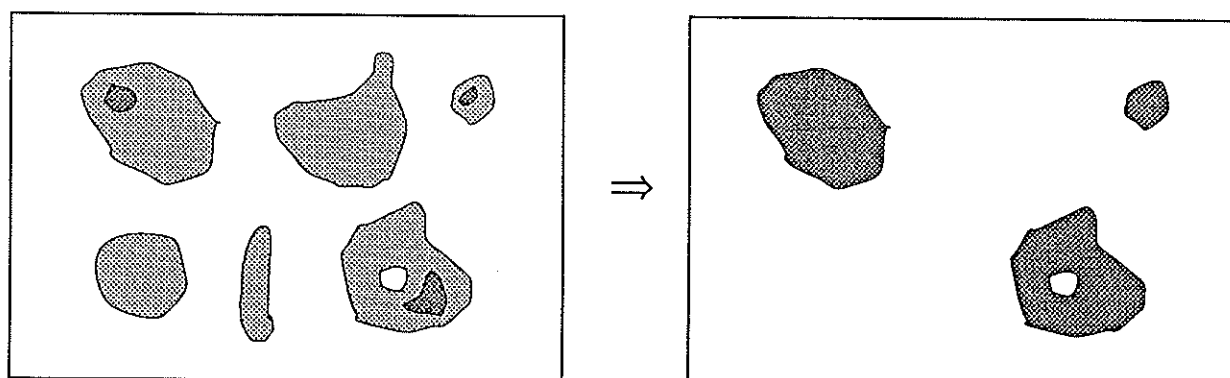


FIG. 33: Reconstruction binaire à partir de marqueurs

va nous être d'un grand secours dans les problèmes de segmentation. En effet comme nous l'avons vu, elle permet d'extraire de façon simple toutes les composantes connexes marquées par un ensemble *marqueur* Y . Cet outil ne nous suffira cependant pas, car le résultat de la reconstruction est inchangé quel que soit le nombre de marqueurs par composante connexe.

Dans le cas numérique, on peut également reconstruire une fonction f à partir d'une "fonction-marqueur" g , majorante ou minorante. Si $g \leq f$, on pose:

$$R_f(g) = D_f^{(+\infty)}(g). \quad (50)$$

Par cette opération, seuls les "dômes" de f marqués par g pourront être reconstruits ou partiellement reconstruits (Voir figure 34). Dans le cas où $g \geq f$, on pose:

$$R_f^*(g) = E_f^{(+\infty)}(g) \quad (51)$$

et g marque ici les "cuvettes" de f qui subsisteront après cette opération de reconstruction.

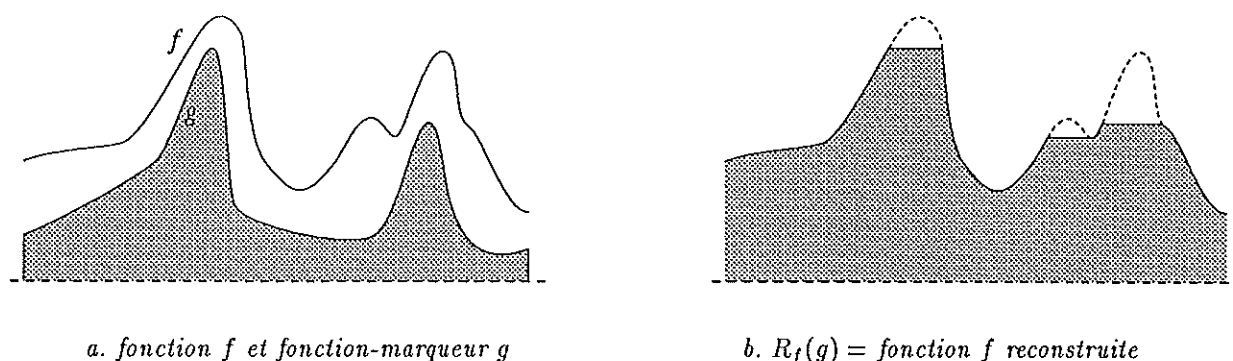


FIG. 34: La reconstruction numérique R_f

Pour exhiber des outils de séparation puissants, il faut à présent adjoindre à ces transformations géodésiques la notion d'homotopie précédemment décrite. Cela se fait de façon relativement simple: on peut définir par exemple des squelettes géodésiques en utilisant une démarche similaire à celle employée dans le cas euclidien [12, page 43]. C'est ainsi que la notion de boule maximale se transpose au cas géodésique. De la même façon, rien n'empêche dans la définition du squelette par zones d'influence de remplacer la distance euclidienne par une distance géodésique. On définira ainsi le *Skiz géodésique* d'un ensemble Y dans un ensemble X (Voir figure 35). On constate en particulier que si Y est constitué de n composantes connexes, on segmentera X en n parties.

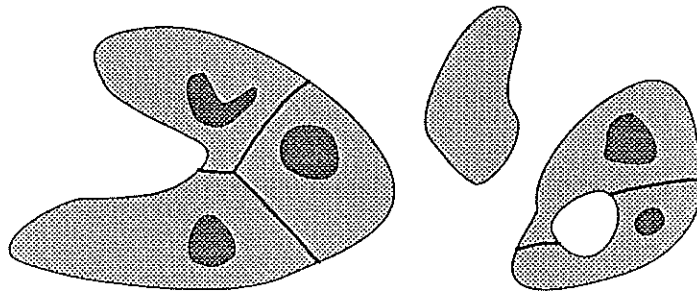
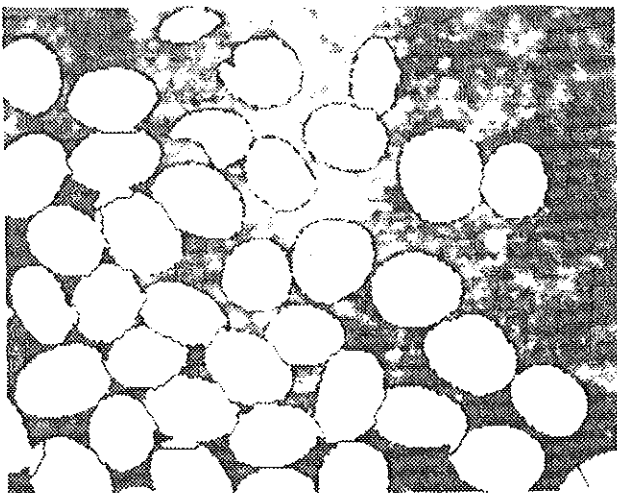


FIG. 35: *Skiz géodésique*

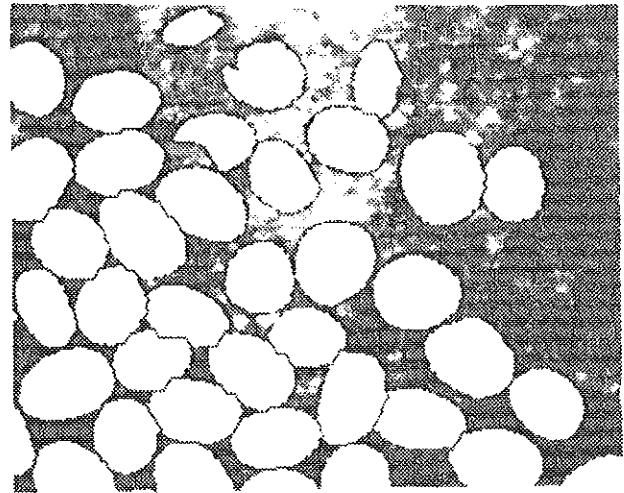
4 DU MARQUAGE À LA SEGMENTATION

4.1 Un algorithme de segmentation binaire

Revenons à présent à notre problème de segmentation de grains. Dans une première étape, ceux-ci ont été marqués par érodés ultimes. On pourrait à présent envisager d'utiliser le squelette par zones d'influence géodésique vu au § 3.2 pour séparer ces grains. Chacun d'entre eux serait alors défini comme la zone d'influence géodésique de son marqueur (Voir figure 36a). Cependant, le résultat obtenu n'est pas satisfaisant, car les lignes de séparation que l'on exhibe entre les grains sont mal positionnées. Ceci est dû au fait que, par construction, la ligne de séparation aura tendance à se placer à mi-distance des marqueurs. Or, on sait que tous ces marqueurs ne correspondent pas à une même taille d'érosion (Voir § 2.8). Lorsque deux grains adjacents ont des tailles très différentes, la ligne de séparation devra être plus proche du marqueur correspondant au grain le plus petit (Voir figure 37).



a. *par Skiz géodésique*



b. *par Ligne de partage des eaux*

FIG. 36: Segmentation des grains de café

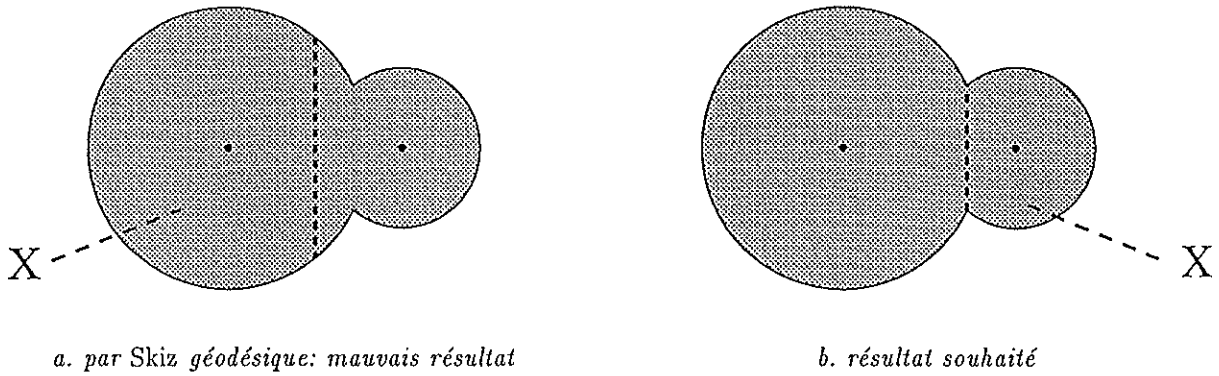


FIG. 37: Séparation de boules qui se recouvrent

Pour résoudre ce problème, il va falloir tenir compte de la valeur associée aux composantes de l'érodé ultime par la fonction d'étanchéité q_X . Ceci peut être fait en utilisant la procédure suivante, que nous allons présenter dans le cas digital: soit n_{\max} la taille de la plus grande érosion non vide de X :

$$X \ominus n_{\max}H \neq \emptyset \text{ et } X \ominus (n_{\max} + 1)H = \emptyset. \quad (52)$$

$X \ominus n_{\max}H$ est nécessairement un sous ensemble de l'érodé ultime. Soit $X_{n_{\max}}$ cet ensemble. Considérons à présent l'érodé de X de taille immédiatement inférieure, soit $X \ominus (n_{\max} - 1)H$. On a la relation d'inclusion suivante:

$$X_{n_{\max}} \subset X \ominus (n_{\max} - 1)H. \quad (53)$$

Soit Y une composante connexe de $X \ominus (n_{\max} - 1)H$. Il y a plusieurs relations d'inclusion possibles entre Y et $Y \cap X_{n_{\max}}$ (Voir figure 38):

1. $Y \cap X_{n_{\max}} = \emptyset$: dans ce cas, Y est une nouvelle composante connexe de l'érodé ultime de X .
2. $Y \cap X_{n_{\max}} \neq \emptyset$ et est connexe: on va alors pouvoir utiliser Y comme nouveau marqueur.
3. $Y \cap X_{n_{\max}} \neq \emptyset$ et est non connexe: dans ce dernier cas, les nouveaux marqueurs utilisés sont les zones d'influence dans Y des composantes connexes de $Y \cap X_{n_{\max}}$.

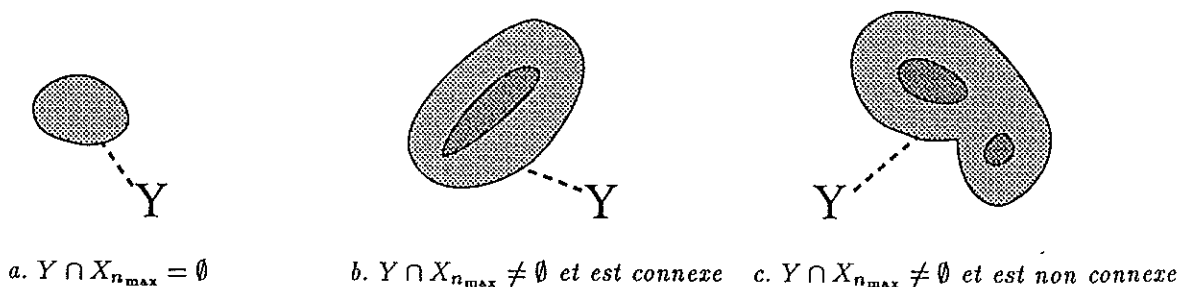


FIG. 38: Les trois relations d'inclusion possibles entre Y et $Y \cap X_{n_{\max}}$

Notons $X_{n_{\max}-1}$ l'ensemble des marqueurs exhibés au niveau $n_{\max} - 1$. Comme ce qui vaut pour une composante connexe vaut pour toutes, $X_{n_{\max}-1}$ sera en résumé constitué des zones d'influence géodésiques de $X_{n_{\max}}$ dans $X \ominus (n_{\max} - 1)H$ auxquelles viennent s'ajouter les composantes connexes de l'érodé ultime au niveau $n_{\max} - 1$ (Voir figure 39).

Cette procédure de reconstruction peut être réitérée aux niveaux $n_{\max} - 2, n_{\max} - 3, \text{etc.} \dots$ jusqu'au niveau 0. De façon plus formelle, soit $n \in]0; n_{\max}]$; pour tout $X \subset \mathbb{R}^2$, nous introduisons les notations suivantes:

- $u_n(X)$ = ensemble constitué des composantes connexes de l'érodé ultime de X au niveau n :

$$x \in u_n(X) \iff x \in S(X) \text{ et } q_X(x) = n. \quad (54)$$

- $\forall Y \subset X, z_X(Y)$ = ensemble des zones d'influences géodésiques des composantes connexes de Y dans X .

La formule de récurrence entre les niveaux n et $n - 1$ s'écrit alors:

$$X_{n-1} = z_{X \ominus (n-1)H}(X_n) \cup u_{n-1}(X) \quad (55)$$

et est illustrée par la figure 39. L'ensemble X_0 obtenu à l'aide de cette procédure produit une segmentation plus correcte qu'un brutal *Skiz* géodésique (Voir figure 36b). En effet, les différentes composantes connexes de l'érodé ultime sont intervenues ici dans un ordre qui respecte la valeur que leur associe la fonction d'étanchéité.

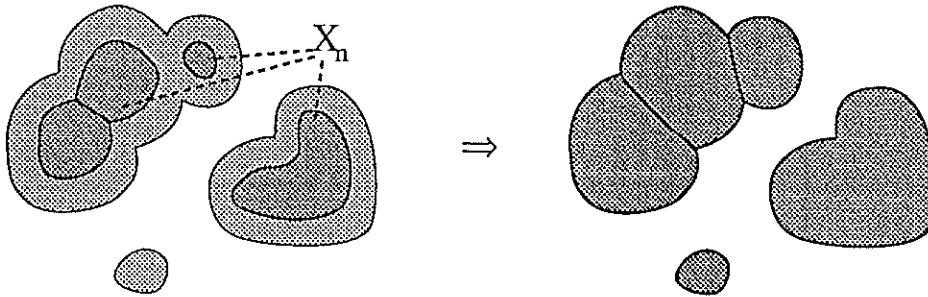


FIG. 39: Passage de X_n à X_{n-1}

4.2 Ligne de partage des eaux (LPE)

L'algorithme précédent peut paraître d'une élaboration complexe. C'est pourquoi nous allons en donner ici une interprétation plus intuitive en utilisant à nouveau la fonction distance $dist_X$ associée à X . Pour

des raisons de cohérence avec la suite, nous considérerons la fonction $f = -dist_X$ et la regarderons comme une surface topographique. Les composantes connexes de l'érodé ultime de X sont exactement égales aux minima de f .

Supposons que l'on trouve cette surface topographique aux emplacements des minima. Plongeons-la alors lentement dans un lac. L'eau va passer par les trous en commençant par ceux qui percent les minima les plus profonds et va ensuite progressivement inonder la surface. Complétons l'expérience en donnant pour mission à un observateur d'ériger une digue en tout point où les eaux provenant de deux minima disjoints pourraient se rejoindre. A la fin de la procédure de plongement, lorsque la surface topographique est intégralement immergée, la digue construite par notre zélé observateur formera une *ligne de partage des eaux* provenant des minima (Voir figure 40). Les différents plans d'eau séparés par cette ligne recouvriront des zones de f appelées *bassins versants* associés à chacun des minima.

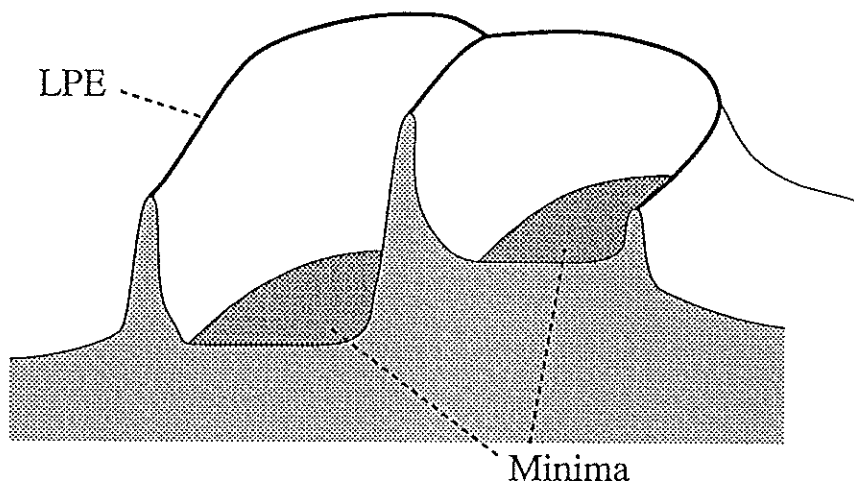


FIG. 40: La ligne de partage des eaux

On voit immédiatement en examinant les plans d'eau obtenus à chacune des étapes de l'immersion que leur ensemble est exactement un des X_n utilisés précédemment (Voir § 4.1). En d'autres termes, la procédure de plongement est identique à l'algorithme de séparation que nous avons mis en œuvre et la ligne de partage des eaux obtenue est donc égale à la ligne de séparation définie par cet algorithme.

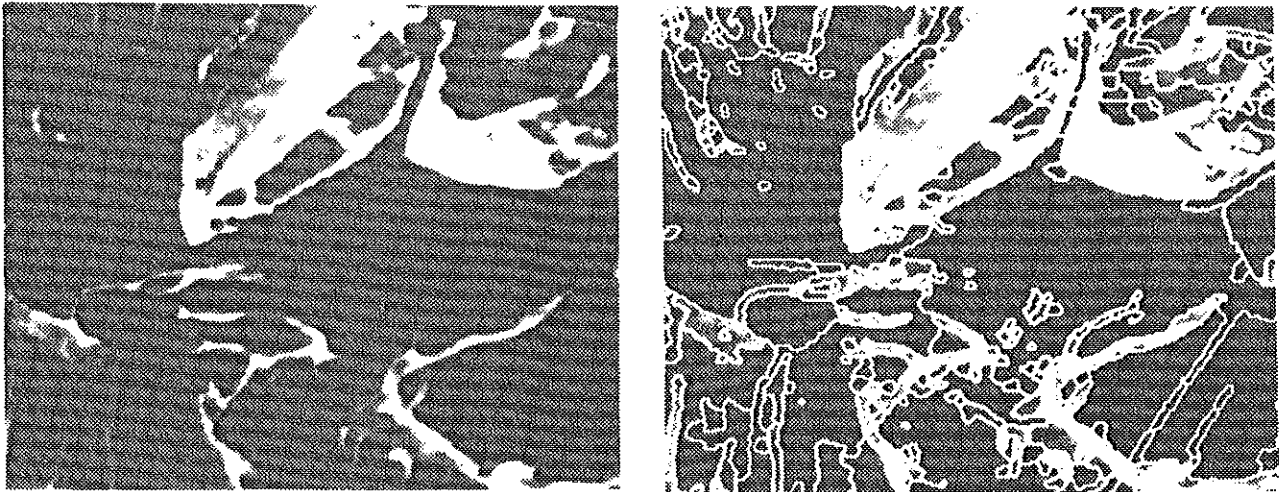
Plus généralement, la LPE peut être définie pour toute fonction numérique $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (Voir [20,21]). En considérant le graphe de f comme une surface topographique et en mettant en évidence les minima de f , rien ne s'oppose à ce que la même procédure de plongement soit appliquée. A chaque minimum de f correspondra donc un bassin versant. En outre, la procédure opératoire utilisée peut

être transposée pour le calcul de la LPE de f . Il suffit pour cela de remplacer les érodés $X \ominus nH$ par les seuils successifs de f .

4.3 Ligne de partage des eaux et segmentation d'images numériques

La notion de LPE d'une fonction va nous servir ici de guide pour la segmentation des images à teintes de gris. Très souvent dans ce cas, la segmentation est associée à un problème de détection de contours. Cependant, cette notion de contour n'est pas toujours bien définie et là encore, lorsqu'on demande à un observateur de préciser les objets à segmenter, il sera capable de les désigner sans pour autant les *contourer* exactement. Autrement dit, une forme dans une image à teintes de gris sera le plus souvent caractérisée par le fait qu'il lui correspond une zone de gris homogène plutôt qu'un contour parfaitement délimité et contrasté. Les zones de gris relativement homogènes de l'image constituent donc là encore des marqueurs des objets à extraire.

Or, qui dit zone relativement homogène d'une fonction dit région à faible variation des niveaux de gris, donc région à *faible gradient*. Les marqueurs de nos objets vont donc correspondre aux *minima* du module du gradient de f (Voir § 2.7). Délimiter les objets consistera alors à calculer la ligne de partage des eaux du gradient. Chaque objet représentera en effet une "zone en creux", c'est-à-dire un bassin versant de la fonction gradient. La figure 41 montre un exemple de détection de contours par ligne de partage des eaux du gradient sur une image de fractures métalliques [22].



a. image originale: fractures métalliques

b. LPE du gradient

FIG. 41: Détection de contours par ligne de partage des eaux du gradient

On voit donc que là encore, les minima d'une fonction—le gradient—ont servi de marqueurs des zones à segmenter. La LPE de cette fonction permet alors d'extraire les contours de ces zones. On retrouve une démarche similaire à celle employée dans le cas binaire, et qui consiste à définir une

fonction dite “de marquage”; cette fonction fournit à la fois marqueurs et contours dans le processus de segmentation. L’unique problème est donc de définir une *bonne* fonction de marquage, c’est-à-dire une fonction qui synthétise les caractéristiques physiques des objets à segmenter; la fonction distance a été choisie pour la segmentation binaire de grains car elle est étroitement liée à la notion de boules maximales incluses dans l’ensemble. De façon similaire, le gradient est utilisé pour segmenter des images à teintes de gris parce que son module correspond à la notion physique de contraste. Cependant, nous allons voir qu’en pratique, rares sont les cas de figure où la fonction de marquage peut être mise en évidence simplement.

4.4 Du bon usage des outils de segmentation

Nous allons illustrer ce qui vient d’être énoncé en appliquant nos outils de segmentation à des images de taches d’électrophorèses bidimensionnelles [23] (Voir planche I, fig. *a*). Le problème est ici bien défini: il s’agit de contourner les taches sombres présentes sur l’image afin d’en calculer d’une part les coordonnées et d’autre part la taille et la densité.

Si, reprenant la méthode exposée, nous effectuons simplement la ligne de partage des eaux du gradient de cette image (Voir § 4.3), le résultat est quelque peu décevant (Voir planche II, fig. *a–b*). Le gradient présente en effet une très grande quantité de minima dûs essentiellement au fait que l’image de départ est bruitée. On peut cependant remarquer que toutes les taches sont marquées par ces minima et donc que sur la LPE, tous les contours qui semblent à l’œil pertinents sont présents (Voir planche II, fig. *b*). Cette ligne de partage des eaux est simplement extrêmement sur-segmentée.

Face à ce résultat, deux attitudes sont possibles:

- Eliminer la sur-segmentation en supprimant de la LPE les éléments d’arcs non significatifs.
- Eviter que la sur-segmentation ne se produise en remettant en question le choix des marqueurs et de la fonction de marquage.

Nous ne nous appesantirons pas sur la première possibilité: c’est le plus souvent un pis-aller—sauf exception—, dû au fait que pour certains problèmes de segmentation, il n’est pas toujours facile de préciser les objets à extraire. La sur-segmentation provient essentiellement du fait qu’il n’y a pas adéquation entre les objets à contourner et les marqueurs utilisés. En résumé, la qualité de la segmentation est directement liée à la fonction de marquage.

La deuxième solution pour éviter la sur-segmentation passe par l’élaboration d’une fonction de marquage appropriée. Les extrema de cette fonction devront marquer exactement les objets que l’on veut extraire. Diverses possibilités existent pour construire ces fonctions de marquage: on peut par exemple simplement filtrer l’image gradient (Voir § 2.4.2). Cette possibilité, qui est plus une amélioration

d'une fonction de marquage existante qu'une réelle construction de cette fonction peut dans bien des cas s'avérer suffisante. Ce n'est malheureusement pas vrai pour les images d'électrophorèses qui nous occupent. Dans ce cas, on va donc vraiment construire une fonction de marquage. Pour ce faire, la première étape—comme nous l'avons exposé—va consister à marquer les régions que l'on veut extraire, c'est à dire les taches, *mais aussi le fond*. A chaque région devra correspondre un et un seul marqueur.

Pour produire ces marqueurs, il faudra utiliser des outils morphologiques autres que le gradient, des outils plus adaptés au problème posé. Ici, les taches constituant la partie sombre de l'image, elles peuvent être interprétées comme des minima de teinte de gris. Une simple extraction des minima de l'image originale produit cependant là encore beaucoup trop de minima (Voir planche I, fig. b). Mais à la différence de l'image gradient, filtrer cette image originale (par un filtre morphologique appelé *Filtre Alterné Séquentiel* [9, chap. 10]) suffit ici à produire une image dont les minima constituent des marqueurs corrects des taches (Voir planche I, fig. c-d).

Quant au marqueur du fond, il sera constitué de la ligne de partage des eaux de l'image filtrée. En effet, cette LPE séparera les taches en passant par les zones les plus claires de l'image (Voir planche I, fig. e).

A présent, nous disposons des marqueurs cherchés (Voir planche I, fig. f), mais cela ne suffit pas: nous avons aussi besoin d'une fonction de marquage dont les minima correspondent à ces marqueurs. En outre, cette fonction devra être suffisamment "proche" de l'image gradient pour que sa LPE soit bien située sur les contours des taches. C'est pourquoi, pour construire notre fonction de marquage, nous allons partir du gradient morphologique (Voir planche II, fig. a) et effectuer deux opérations:

1. Imposer comme minima ceux qui ont été choisis précédemment (par la suite, nous noterons M l'ensemble des points de \mathbb{R}^2 appartenant à ces marqueurs-minima).
2. Supprimer les minima du gradient non désirés.

Pour ce faire, nous allons dans une première étape affecter aux points de la fonction gradient correspondant aux minima souhaités une valeur négative arbitraire c . Ensuite, nous allons éliminer les minima non désirés en comblant les bassins versants qui leurs sont associés. La technique morphologique pour réaliser cette deuxième opération utilise la reconstruction géodésique numérique R^* introduite au § 3.2. La fonction marqueur g employée est définie par:

$$\forall x \in \mathbb{R}^2, \begin{cases} x \in M & \implies g(x) = c \\ x \notin M & \implies g(x) = A \end{cases} \quad (56)$$

où A est une constante arbitraire qui majore la fonction gradient. Quant à la fonction f à reconstruire, c'est tout simplement l'inf de g et du gradient (Voir figure 42). On remarquera sur la figure 42b que la fonction reconstruite correspond exactement à ce que l'on cherche. En particulier, sa LPE est un

sous-ensemble de la LPE de la fonction gradient initiale. De plus, les lignes conservées sont situées sur les zones de l'image où le gradient est le plus fort.

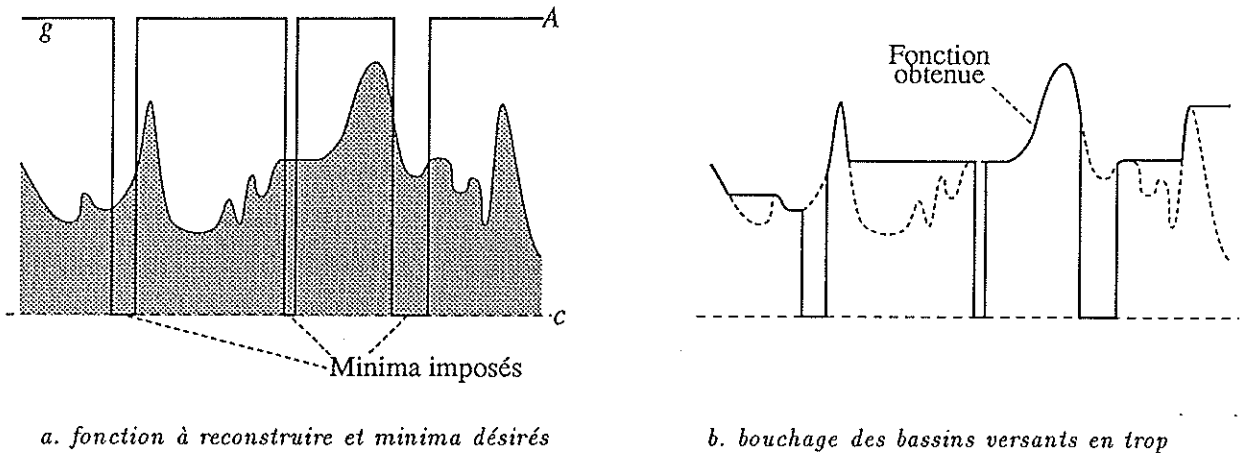


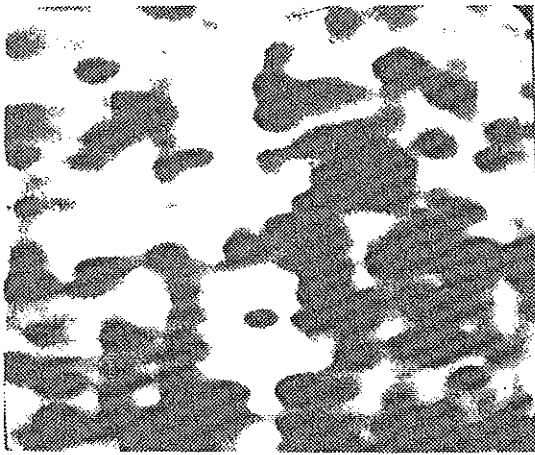
FIG. 42: Ligne de partage des eaux "géodésique"

Les figures *c* à *e* de la planche II illustrent cette procédure de segmentation, ainsi que le résultat obtenu. La figure *f* de cette même planche montre, elle, les marqueurs des taches et les contours correspondants. Compte tenu des marqueurs choisis et de l'algorithme de segmentation mis en œuvre, ce résultat est conforme à ce que l'on attendait: chaque tache a un contour unique et celui-ci est situé sur les points d'inflexion de la fonction luminance initiale. Il est possible par des procédures plus raffinées d'améliorer encore le contourage [23]. On peut par exemple remarquer que les contours obtenus par la méthode précédente délimitent pour chaque tache un masque de mesure un peu trop petit. Cependant, en tenant compte des valeurs du gradient le long d'un contour donné, on peut étendre ce masque sur la périphérie des taches. L'algorithme utilisé consiste à dilater chaque point du contour en raison inverse du module de son gradient.

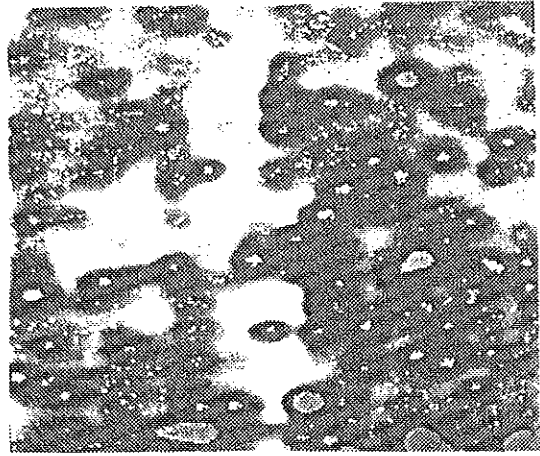
4.5 Choix des fonctions de marquage: variations et nouveaux développements

4.5.1 Les difficultés d'un bon marquage

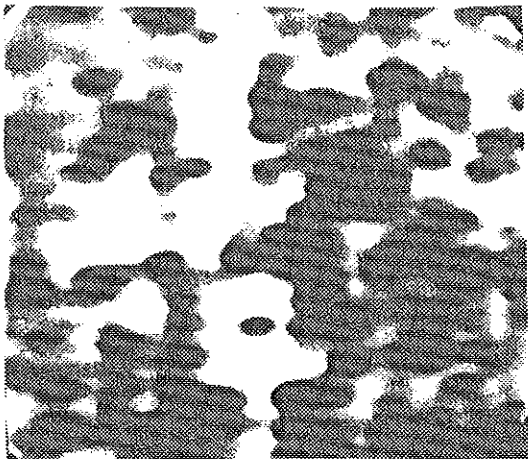
La description de l'algorithme pour extraire les contours des taches d'électrophorèse a montré qu'il n'était pas toujours simple de définir des marqueurs. Affirmer que les marqueurs des objets correspondent aux minima de la fonction gradient est souvent trop brutal. De la même façon, le marquage par érosion ultime dans la segmentation binaire a ses limites [24]. Il n'est efficace que si les composantes de l'ensemble X sont d'une part "relativement circulaires" et d'autre part, si leur imbrication n'est



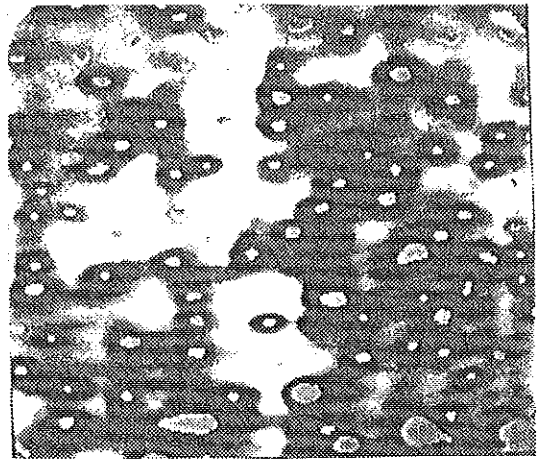
a. Image originale



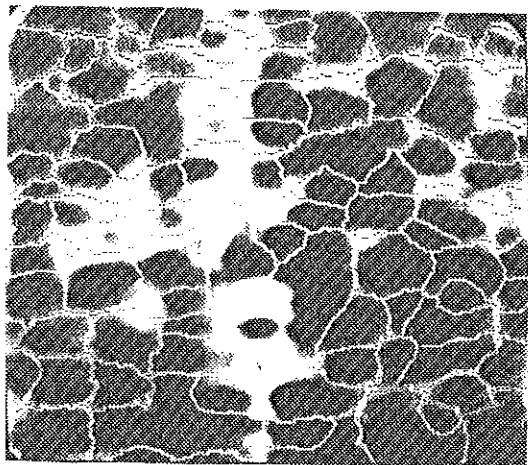
b. Minima de l'image originale



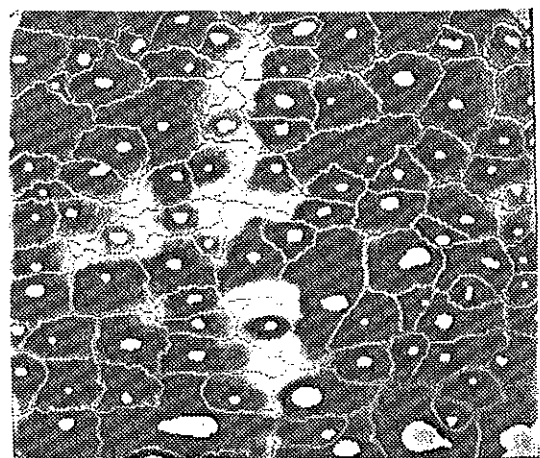
c. Image filtrée (FAS)



d. Minima de l'image filtrée

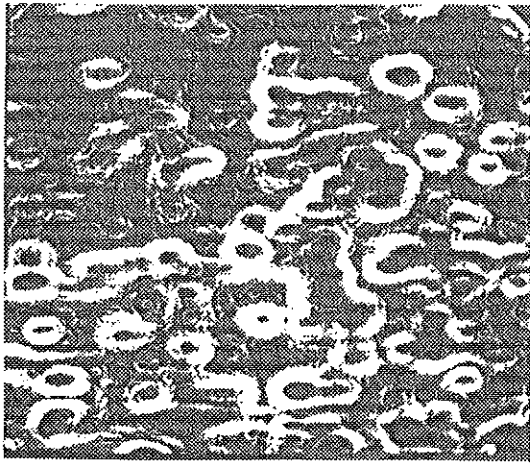


e. LPE de l'image filtrée

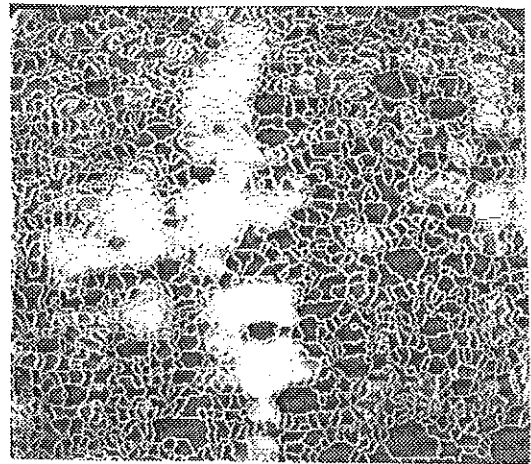


f. LPE et minima

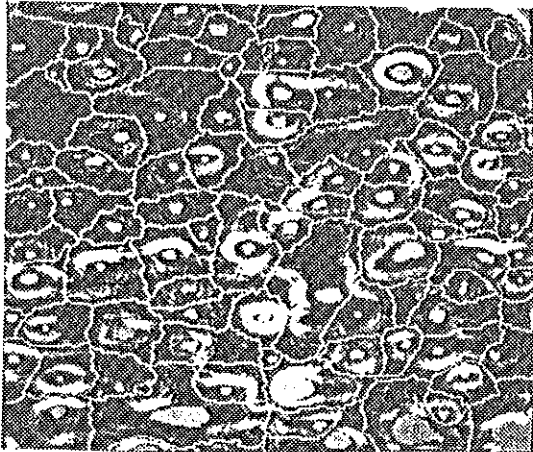
PLANCHE I: Segmentation de taches d'électrophorèse



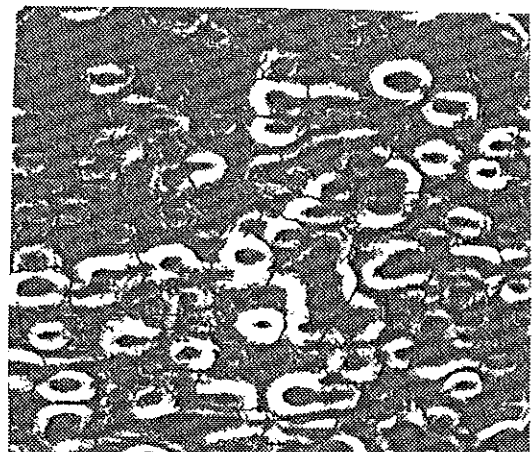
a. Gradient morphologique



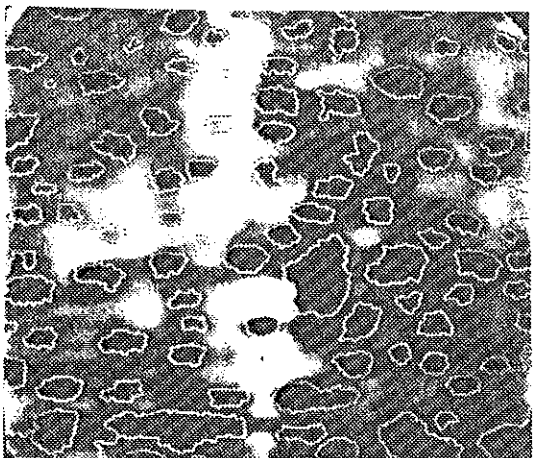
b. LPE du gradient



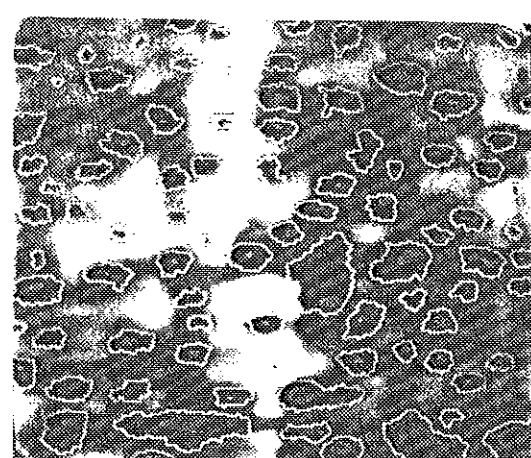
c. Gradient + Minima imposés



d. Gradient modifié



e. LPE du gradient modifié



f. Marqueurs et contours des taches

PLANCHE II: Segmentation de taches d'électrophorèse (suite)

pas trop importante. Par exemple, dans le cas où X est composé de deux disques imbriqués D_1 et D_2 , l'érodé ultime marque ces deux disques uniquement si leurs centres sont situés de part et d'autre de l'axe radical (Voir figure 43). La figure 44 illustre un cas où l'érosion ultime et la LPE de la fonction distance se sont avérées insuffisantes pour séparer des grains de sel.

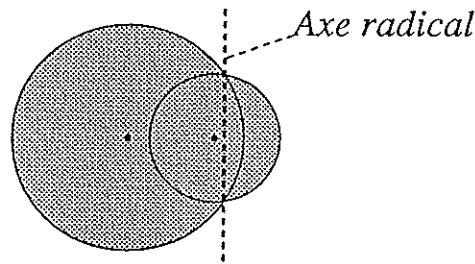
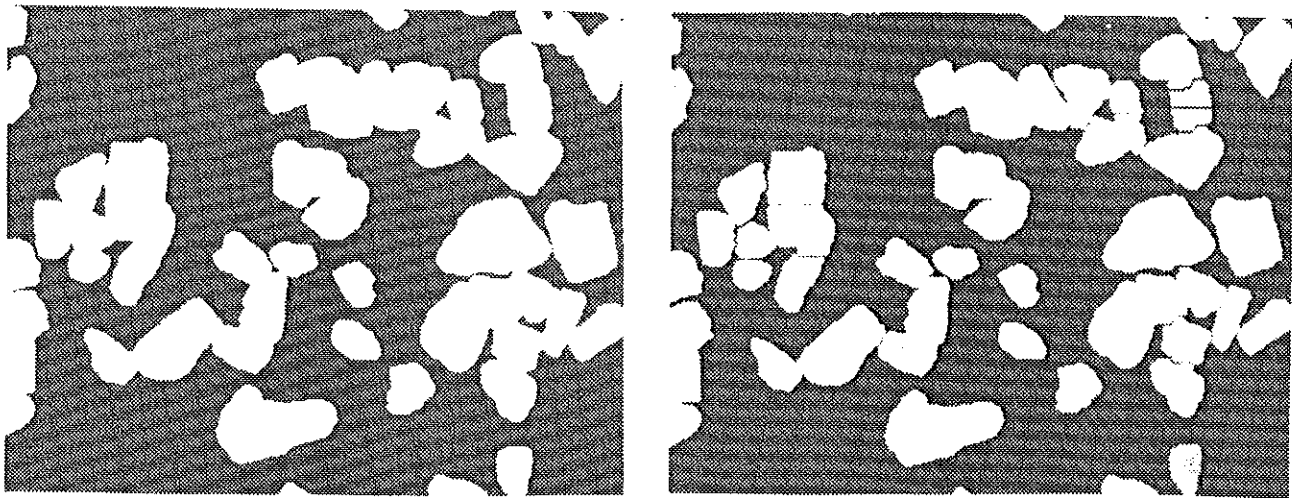


FIG. 43: Ces deux disques sont trop imbriqués pour être marqués par érosion ultime



a. image originale (grains de sel)

b. "séparation" par LPE de la fonction distance

FIG. 44: Les insuffisances de la segmentation par érosion ultime

Il existe cependant des outils de marquage plus raffinés. Parmi eux, citons les transformations appelées *bissectrices conditionnelles* [12, page 55], où le marquage est lié non pas aux maxima de la fonction d'étanchéité, mais aux extrema de sa dérivée. De même, lorsque l'ensemble à segmenter est constitué de particules allongées, il peut être intéressant de marquer les extrémités de celles-ci; ceci peut être fait soit par des techniques à base d'érosions ultimes géodésiques, soit en utilisant les maxima de la *fonction de propagation* [25]. Les méthodes précédentes sont plus complexes à mettre en œuvre

pour des problèmes de segmentation car il n'est pas toujours aisé d'exhiber les fonctions de marquages associées à de tels marqueurs...

4.6 Récents développements

En morphologie binaire, de nombreux problèmes de segmentation sont encore mal résolus. Parmi eux, citons ceux consistant à segmenter des assemblages quelconques d'objets convexes, ou des fibres qui s'intersectent, etc... Souvent dans ces cas, le marquage des objets à extraire nécessite une connaissance à priori de leur géométrie. Sans chercher à résoudre des problèmes aussi complexes, on peut donner quelques indications sur les améliorations possibles des procédures précédemment décrites pour les cas où on a affaire à des assemblages de particules convexes fortement imbriquées. On peut par exemple introduire la notion de *boules critiques* et de *fonction critique*. Les boules critiques forment un sous-ensemble minimal de boules maximales nécessaires à la reconstruction de l'ensemble initial. Utilisées conjointement avec la fonction critique, qui associe à chaque point de l'ensemble la densité de boules maximales qui le recouvrent (Voir figure 45), elles permettent une séparation plus efficace que celle par érodé ultime (Voir figure 46).

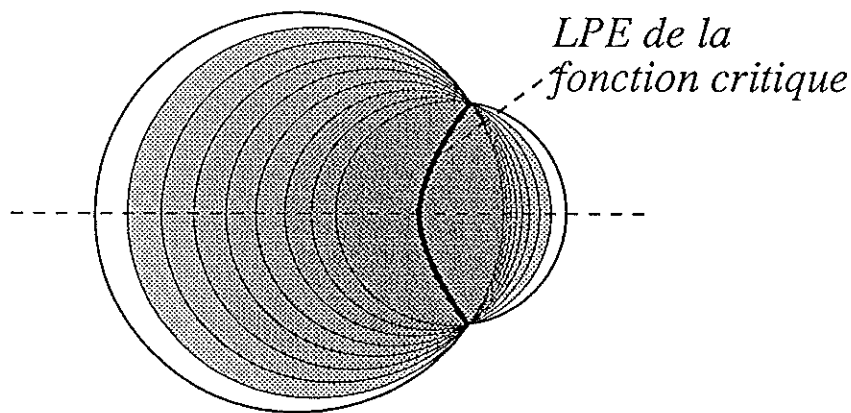
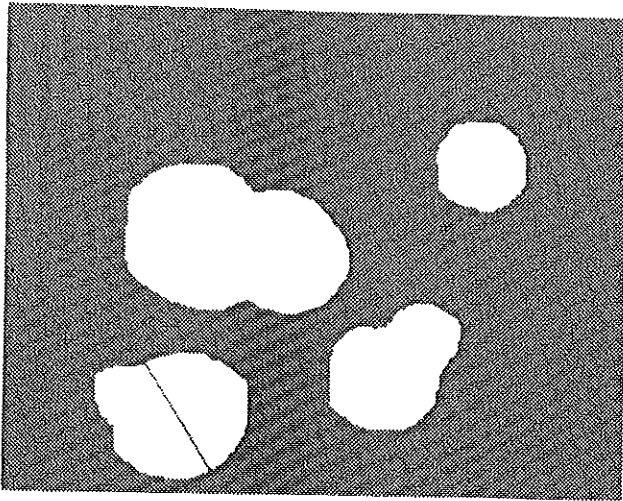
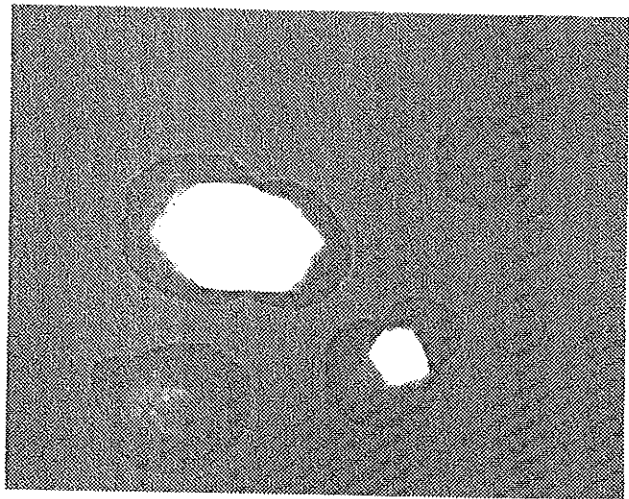


FIG. 45: Séparation de boules très interpénétrées par LPE de la fonction critique

En ce qui concerne la segmentation des images à teintes de gris, le problème majeur vient de la sur-segmentation de la ligne de partage des eaux du gradient. On a vu—et illustré par un exemple—que des techniques élaborées de sélection des minima peuvent permettre d'éviter que cette sur-segmentation ne se produise. Mais on peut aussi l'éliminer après coup par des techniques de hiérarchisation des contours et des bassins versants: la LPE permet de simplifier l'image à segmenter en la réduisant à une mosaïque de plages uniformes (Voir figure 47), mosaïque pouvant être interprétée comme un



a. "segmentation" par LPE de la fonction distance



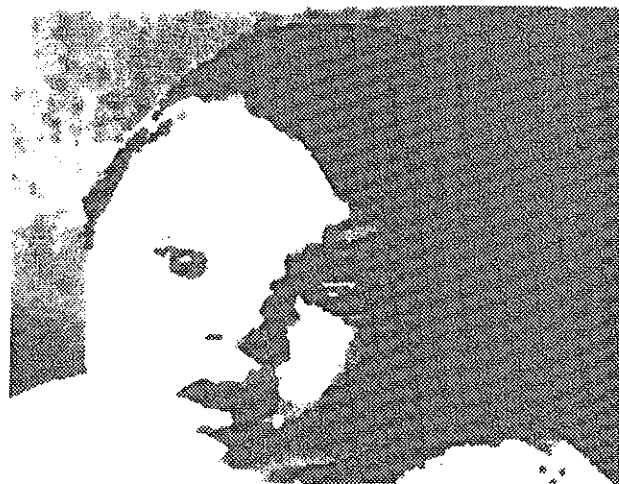
b. fonction critique

FIG. 46: La segmentation de cette image binaire par LPE de sa fonction distance ne donne rien alors que les lignes de crête de la fonction critique "séparent" ses composantes correctement.

graphe. Il est alors possible sur ce graphe de définir et effectuer des transformations morphologiques



a. image originale



b. la mosaïque des bassins versants

FIG. 47: Simplification d'image par extraction des bassins versants

[26,27], et par ce biais, de déterminer les contours les plus significatifs.

5 CONCLUSION

Nous avons tenté de dégager tout au long de cet article une philosophie générale de la segmentation en morphologie mathématique. Marqueurs et fonctions de marquage sont les concepts clés de cette approche. Les exemples présentés, de part leur caractère didactique, ne permettent pas d'illustrer de façon exhaustive les infinies variations possibles autour de ces notions. Marqueurs et fonctions de

marquage formalisent assez bien les deux étapes par lesquelles passe nécessairement la résolution de tout problème de segmentation:

1. *Que faut-il segmenter?*
2. *Comment faut-il s'y prendre?*

RÉFÉRENCES

- [1] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, Londres, 1982.
- [2] F.P. Preparata & M.I. Shamos, *Computational Geometry, an introduction*, Springer Verlag, 1985.
- [3] A. Gagalowicz, *Vers un modèle de texture*, Thèse de doctorat d'état, Université Paris VI, 1983.
- [4] O. Monga, *Segmentation d'images par croissance hiérarchique de régions*, Thèse de doctorat, Univ. de Paris Sud, 1988.
- [5] H. Minkowski, *Allgemein Lehrsätze über konvexe Polyeder*, Nach. Ges. Wiss. Göttingen, 198-219, 1897.
- [6] S. Beucher & J. Serra, *Shapes and patterns of microstructures considered as grey-tone functions*, Proc. 3rd European Symposium on stereology, Ljubljana (Yougoslavie), 22-26 juin 1981.
- [7] G. Matheron, *Éléments pour une théorie des milieux poreux*, Masson, Paris, 1967.
- [8] M. Coster & J-L. Chermant, *Précis d'analyse d'images*, Ed. du CNRS, 1985.
- [9] J. Serra, *Image Analysis and Mathematical Morphology, part II: Theoretical Advances*, J. Serra ed., Academic Press, Londres, 1988.
- [10] L. Calabi & W.E. Harnett, *Shape recognition, prairie fires, convex deficiencies and skeletons*, Parke Mathematical Laboratories Inc., Carlisle MA (USA), Scientific report 1, février 1966.
- [11] G. Matheron, *Quelques propriétés topologiques du squelette*, Note interne, Centre de Morpho. Math., Ecole des Mines, Paris, 1978.
- [12] F. Meyer, *Cytologie quantitative et Morphologie Mathématique*, Thèse de doctorat, Ecole des Mines, Paris, 1979.
- [13] S. Beucher & F. Meyer, *Méthodes d'analyse de contrastes à l'analyseur de textures*, Proc. congrès AFCET-IRIA de reconnaissance des formes et traitement des images, Chatenay-Malabry, Février 1978.
- [14] M. Golay, *Hexagonal pattern transforms*, IEEE trans. on Comp., C18, 8, août 1969.
- [15] S. Beucher, *Squelettes connexes et non connexes*, Note interne, Centre de Morpho. Math., Ecole des Mines, Paris, 1989, à paraître.
- [16] F. Meyer, *Skeletons and perceptual graphs*, Signal Processing, numéro spécial de Morpho. Math., mars 1989, à paraître.

- [17] C. Lantuejoul, *La squelettisation et son application aux mesures topologiques des mosaïques polycristallines*, Thèse de doctorat, Ecole des Mines, Paris, 1978.
- [18] C. Lantuejoul & S. Beucher, *On the use of geodesic metric in image analysis*, J. of microscopy, 121, 39–49, 1981.
- [19] C. Lantuejoul & F. Maisonneuve, *Geodesic methods in image analysis*, Pattern Recogn., 17, 117–187, 1984.
- [20] S. Beucher, *Use of watersheds in contour detection*, Proc. Int. Workshop on image processing, real-time edge and motion detection/estimation, Rennes, 17–21 septembre 1979.
- [21] S. Beucher, *Watersheds of functions and picture segmentation*, Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, 1928–1931, Paris, mai 1982.
- [22] S. Beucher, M. Blanc & T. Hersant, *Analyse quantitative de clichés stéréoscopiques*, J. Microsc. Spectrosc. Electron., 7, 105–106, 1982.
- [23] S. Beucher, *Analyse automatique des gels d'électrophorèse bi-dimensionnels et Morphologie Mathématique*, Note interne, Centre de Morpho. Math., Ecole des Mines, Paris, 1982.
- [24] M. Benali, *Du choix des mesures dans les procédures de reconnaissance des formes et d'analyse de texture*, Thèse de doctorat, Ecole des Mines, Paris, 1986.
- [25] M. Schmitt, *Des algorithmes morphologiques à l'intelligence artificielle*, Thèse de doctorat, Ecole des Mines, Paris, février 1989, à paraître.
- [26] L. Vincent, *Mathematical Morphology on Graphs*, Proc. SPIE Int. Congress. "Visual Communications and Image Processing 88", 3rd in a series, Cambridge MA (USA), 1988.
- [27] L. Vincent, *Graphs and Mathematical Morphology*, Signal Processing, numéro spécial de Morpho. Math., mars 1989, à paraître.