Image analysis of cement paste: relation to diffusion transport

Vincent Tariel

Image analysis of cement paste: relation to diffusion transport

THÈSE

présentée et soutenue publiquement le 19 février 2009

pour l'obtention du

Doctorat de l'école Polytechnique

(spécialité physique)

par

Vincent Tariel

Composition du jury

Rapporteurs :	Mr. Dominique BERNARD	Directeur de recherche-CNRS/Université de Bordeaux
	Mr. Mark KNACKSTEDT	ARC QEII Fellow-Australian National University
Examinateurs :	Mr. Michel BORNERT	Professeur-École Polytechnique
	Mr. Dominique JEULIN	Professeur-École des Mines
	Ms. Karen SCRIVENER	Professeur-École Polytechnique Fédérale de Lausanne
Codirecteur de thèse :	Mr Denis DAMIDOT	Professeur-École des Mines de Douai
Directeur de thèse :	Mr Pierre LEVITZ	Directeur de recherche-CNRS/École Polytechnique
Invités :	Mr Emmanuel GALLUCCI	Docteur-École Polytechnique Fédérale de Lausanne
	Mr. Xavier GUILLOT	Docteur-ATILH

Mis en page avec la classe thloria.

Remerciements

Je tiens tout d'abord à remercier les membres du jury de ma thèse. Je remercie Dominique Bernard et Dominique Jeulin pour l'intérêt qu'ils ont apporté à mon manuscrit. Leurs remarques constructives ont permis d'améliorer grandement sa qualité. Je remercie Marck Knackstedt, qui a posé des questions pertinentes portant sur la partie physique de mon travail et cela malgré la fatigue de son voyage. Il est arrivé d'Australie après plus de 30 heures d'avion juste le matin de ma soutenance (il est bien difficile de trouver une date convenant à tous). Je continue l'aventure de la recherche dans son laboratoire à Canberra. Je remercie Karen Scrivener, l'experte mondiale de la chimie du ciment qui a su comprendre mon manuscrit pour en tirer l'information ciment. Je remercie Michel Bornert pour sa curiosité intellectuelle. Ses questions, sans pièges, étaient reliées aux points d'ombres de mon manuscrit. Je remercie Dominique Jeulin, mon ancien directeur de stage, pour m'avoir soutenu dans mon choix d'entreprendre une thèse et pour m'avoir apporté la culture de la rigueur scientifique. Enfin, je remercie Denis Damidot, mon co-directeur de thèse, Xavier Guillot et Angélique Vichot, mes encadrant industriels, pour leur disponibilité et leur soutien tout au long de ma thèse. Je tiens à remercier Pierre Calka avec qui j'ai collaboré au sein du projet ANR mipomodim.

Pour remercier les autres, voici une petite histoire typiquement gauloise ! Nous sommes en 2000 après Jésus-Christ. Toute la Gaule est occupée par le Sarkozisme... Toute? Non ! Un laboratoire peuplé d'irréductibles chercheurs résiste encore et toujours à l'envahisseur. Et la vie n'est pas facile pour les garnisons de technocrates romains des camps retranchés autour de efficacum, evaluatum, controlum et precaratum...

Le laboratoire PMC, Passionné et Militant Chercheurs, farouchement indépendant, vit en paix. Voici quelque gaulois habitant ce village:

- Chimistrix, le druide, cueille des plantes et prépare des potions magique, tel le Sol-Gel rempli de nanoparticules,
- Electrochemistrix, le sculpteur, dépose des ornements nanostructurum sur la surface 111 des épées leur conférant des pouvoirs magiques,
- Photonix, le barde, émet des chants photoniques dans l'optique d'une mélodie magnétique,
- Irrégularitum, le Pécheur, crie cette phrase perdu sur son bateau dans le dédale des golfes et des criques de la côte Bretonne "Quel fractalum !",
- Ozanamix, le chef de la tribu, tient à respecter à ce que chaque gaulois ait sa hutte et à ce que le fruit du travail de chacun soit distribué à tous.

Ce village ne craint qu'une chose: c'est que le ciel lui tombe sur la tête !

Un matin d'automne, je suis entré dans ce village comme apprenti-forgeron. Mon maître-forgeron, était Levitz. Comme à chacun de ses apprentis, il demandait, en plus du travail courant, d'approfondir un sujet. Le mien a été de comprendre le lien entre la morphologie de l'objet et ses propriétés physiques. Ce sujet était aussi large que le tour de taille d'Obélix. Forcément, il y a eu un contour du sujet m'intéressant. Tout de suite, mon maître m'a fait confiance ce qui est étonnant vu que je venais d'une petite école. La plupart des autre apprentis venait des célèbres écoles de Lutèce comme Polytechnicum, Normalium... Armé de la confiance et de la passion, je me jetai dans la mêlée du travail, par Toutatix ! Parfois, il y avait des impasses. Par exemple, pendant une semaine, j'ai passé mon temps à essayer de marteler le métal en vain. Il m'expliqua ensuite qu'il était très dur de travailler le métal si on ne le chauffait pas: le fameux recuit ! Par Bélénox, il aurait pu me le dire tout de suite ! Oui, mais l'apprentissage, c'est peut-être ça. Etant tous les deux de fiers gaulois, nos discussions étaient de tant en tant musclées. Au bout de trois ans, après maintes péripéties, j'étais enfin prêt à affronter un jury composé des plus éminents maîtres-forgerons de l'époque venant de différents horizons.

La maitrise de l'art de la forge est aussi difficile que d'écraser une mouche avec une enclume. Seul, c'est impossible. J'ai eu la chance d'écouter les cours de grands druides pendant le master, Chaire Lafarge, dans la forêt des Carnuttes située après le pont et à côté de la chaussée. Le druide organisateur s'appelle

Coussix. Il nous montra l'étendue de ses pouvoirs, par exemple en versant de l'eau sur des gaines de bois plantées dans la pierre, il arriva à fendre un bloc de pierre. Cette technique est bien différente de celle de notre tailleur de pierre Obélix qui lui se sert uniquement de ses mains ! Je lui ai demandé comment cela était possible. Il m'expliqua simplement les concepts de la poromécanique. Lors de ce rassemblement, je me suis fait de nombreux amis apprentis: Jia, Laurent, Quentin, Matthieu, Julien... Avec ces acolytes, j'ai eu la chance d'aller écouter un/au MIT aux Amériques.
Pour apprendre l'art d'observer un alliage pour en extraire ses caractéristiques, je suis allé chez les helvètes. A l'EPFL, Galluccix et ses apprentis m'ont tout expliqué touuuuut douuuuucement. Heureusement car je ne comprenais pas vite. Ce que je n'ai toujours pas compris c'est pourquoi il m'appelait le geek. Il sont fous ces helvètes !
A la fin de ces trois années, j'ai du graver sur des pierres le résultat de mon travail: "ad vitam aeternam". Ce fastidieux travail s'est fait avec l'aide d'un russe Grebenkovix qui n'arrêtait pas de me parler de ses expériences de grains de pollens à la surface de l'eau d'un récipient. Il comptait le temps que prend une particule de Pollen pour toucher le bord du récipient pour la première fois. A chacune de ces expériences, il criait: "alea jacta est !". Ils sont fous ces russes !
Avec certains apprentis dont Brisardix et N'Guyenix, il y a eu un vrai plaisir de diffuser nos idées sous différents angles. Nous étions en phase sur bien bien des champs.
Après des heures à marteler le métal, un gentleman venant de l'île de Bretagne nommé Lassaillix me préparait un breuvage appelé "tea for two" qui me donnait plus de force que la potion magique de Panoramix. Mais bon, quand j'étais vraiment trop fatigué, il y avait juste les cervoises bues avec mes anciens amis d'école pour me détendre !
Pendant mon apprentissage, à cause d'un mauvais geste lors d'une trempe, je me suis sérieusement brulé. Rien de très grave finalement car ma famille m'a préparé un beau pansement.
Il faut savoir que notre petit village n'est pas constitué uniquement de Gaulois. Il y a un brésilien, Hugo, qui a importé de son lointain pays la coutume de jouer au moins une fois au football par semaine, la communauté maghrébine, Sabrina, Amel, Linda, kamila, qui m'ont ouvert les yeux sur certains préjugés, des personnages provenant du soleil Levant, Vu et Xiaxin, des gens qui parlent la plupart du temps avec les mains, Matéo et Filippo, des personnes de grande taille de l'Europe de l'Est, Théodorat la motivée et Dorine le moldave.
Il y aussi Fred qui a conservé son flegme malgré des difficultés, Damien avec ses passements de jambes, Clotilde qui compte plus qu'une clopinette, Carine qui est plus têtue que moi, Raffi, l'homme critique, Aurélie et Nayely, les grandes sportives des abdominaux de la rigolade, Larbi, qui souhaiterait changer “une grande école, pourquoi pas moi?” par “un job en ayant une formation, pourquoi pas moi?”, Anne qui est toujours prête à te réconforter, Kuk, le codeur fou artistique, Divad, l'organisateur des Oktoberfests, Nicolas, qui est into the wild, Gringos, mon binome de backloop en windsurf, El Guedon, le frisé musicos, les Antis, les artistes de l'aventure, Gayrémyx l'aficionado de club med gym, Géraud, l'auvergnat, Caroline, la force tranquille, la dream-team RMN composée de Dominique, Jean-Pierre, Houria, Gabriel, la team informatique, Julien et Denis, et tous les gaulois et non-gaulois que j'ai oubliés.

A tous ces personnages, j'adresse mes chaleureux remerciements car sans vous cette histoire n'aurait jamais été possible.

Table of contents

Abstract

Since the emergence of imaging techniques, MRI, CT, it is now possible to directly probe the geometrical organization of systems such as bone, cement, paper, glass, rocks. As the physical and mechanical properties depend on the geometrical organization, there is a scientific and industrial interest for the understanding of this relationship by using these imaging techniques. In this context, the purpose of this thesis is the development of a toolkit for digital image analysis of the material geometry, then the application of this toolkit in the study of the evolution of the pore space of cement paste.

In part one, after a discussion on the choice of an imaging technique adapted to a material, we present the two imaging techniques selected, scanning electron microscopy and synchrotron tomography for the analysis of cement paste and the experimental protocol for sample preparation.

In the second part, we propose a generic, efficient and simple methodology of segmentation. Segmentation is the transformation of a grey-level image to an labeled image where each label represents a phase of the material. Generic means that this methodology can be used for a wide range of materials and imaging techniques. Effective means that the segmented structure matches the real structure. Simple means that the calibration is easy. The implementation of the optimized algorithms associated with this methodology is done thanks to the theoretical conceptualization of the region growing.

In the final part, we quantify the morphology and topology of the geometry of the material statistically. Then, we decompose a phase in term of elementary components along two agreements: one morphological and the other topological. Finally, we use the stereological information estimated on the 2D slice to reconstruct a 3D model larger than the representative elementary volume using the optimized algorithm of simulated annealing. The validation of the 3D reconstruction is performed by the comparison of properties of diffusive transport.

Depuis l'émergence des techniques d'imagerie, IRM, tomographie, il est maintenant possible d'observer directement l'organisation géométrique de systèmes tels l'os, le ciment, le papier, le verre, les roches. Comme les propriétés physiques et mécaniques dépendent de l'organisation géométrique, il existe un intérêt scientifique et industriel de comprendre et de définir cette relation de dépendance à l'aide de ces techniques d'imagerie. S'inscrivant dans ce contexte, le but de cette thèse est de développer un ensemble d'outils numériques pour l'analyse d'image de la géométrie d'un matériau, puis d'appliquer ces outils dans l'étude de l'évolution de la porosité de la pâte de ciment.

En première partie, après une discussion sur le choix d'une technique d'imagerie adaptée à un matériau, nous présentons les deux techniques d'imageries sélectionnées, la microscopie électronique à balayage et la tomographie par synchrotron, pour l'analyse de la pâte de ciment et le protocole expérimental pour la préparation des échantillons.

En deuxième partie, nous proposons une méthodologie générique, efficace et simple de segmentation. La segmentation est la transformation de l'image en niveaux de gris en une image labellisée où chaque label représente une phase du matériau. Générique signifie que cette méthodologie est applicable pour une large classe de matériaux et de techniques d'imageries. Efficace spécifie que la structure segmentée correspond à la structure réelle. Simple indique que l'étape de calibration est facile. L'implémentation de l'ensemble des algorithmes optimisés associés à cette méthodologie est rendue possible grâce à la conceptualisation théorique de la croissance de régions.

En dernière partie, nous quantifions statistiquement la morphologie et la topologie de la géométrie

du matériau. Puis, nous décomposons une phase en éléments élémentaires suivant deux conventions: l'une morphologique, l'autre topologique. Enfin, nous utilisons l'information stéréologique estimée sur une coupe 2D pour reconstruire un modèle 3D plus grand que le volume élémentaire représentatif à l'aide de l'algorithme optimisé du recuit simulé. Une validation de la reconstruction 3D est effectuée par un suivi des propriétés de transport diffusif.

Introduction

In material sciences, a porous medium is a solid, often called matrix, permeated by a pore network (voids) filled with a fluid (liquid and/or gas) such as a sponge. Many natural systems such as rocks (e.g. limestone), soils (e.g. sand), biological tissues (e.g. bones, lung), and man made materials such as cements, foams and ceramics can be considered as porous media.
The geometrical organisation of the porous medium affects numerous physico-chemical phenomena such as molecular diffusion, excitation relaxation, reaction kinetics, phase transitions, adsorptions and capillary condensation. For example, one feature of this geometrical organisation, the specific surface area measuring the surface area per unit of volume, is related to the adsorption, heterogeneous catalysis, and reactions on surfaces. In a industrial context, of activated charcoal, synthetic resins, and water purification, the main goal is to design porous media that exhibit a large specific surface area in order to maximise the possible interaction with reactants.
In turn, for some industrial applications, the aim is the conception of porous media in order to minimise transport properties. For cementitious materials, the transport properties are closely related to the durability since they control the invasion dynamics of various molecules inside the porous medium. One direct effect of these potential contaminants is the corrosion of the structural steel by the chloride ions. The mechanical strain, from the dilation induced by the corrosion, is the source of cracks [131, 184].
In porous media, the two main processes for transport are diffusion (due to difference in concentration) and convection (due to a difference inpressure). Molecular diffusion is usually slower. However, convection in a porous medium, with pores smaller than a few microns, is slowered down by the viscosity effects, so that molecular diffusion is predominant and allows the chemical species to move in the porous medium. This is the case of a cement paste with a pores size smaller than a few microns. The aim of this thesis is to understand this transport process.
In porous media, molecular diffusion is affected by the fluid saturation, the electro-chemical gradient, the surface chemistry and the surface reactivity. In this thesis, we assume that the pore network is homogeneously filled either by liquid or gas water, there is no electro-chemical gradient and finally the interface is considered as inert in order to focus on the understanding of the relation between the diffusive transport and the geometrical organisation of the porous medium of cement paste.
To achieve this task, it is crucial to describe correctly the geometrical organisation. A first level of description is reduced to few numbers that characterise the global properties of the geometrical organisation. The most common numbers are the porosity, ϕ, and specific surface area, $\mathcal{S}_v$. In most of cases, the model of the geometrical organisation is based on these two numbers as input parameters. The variety of the cement chemical compositions imposes a wide range of the geometrical organisations of the porous medium. The use of ϕ and $\mathcal{S}_v$ does not take into account many geometrical features, like the connectivity of the pore network, the structural correlation and the hierarchy at different length scales. A more extended analysis has thus to be performed.
In this respect, experimental imaging techniques play and will play an important role in understanding the metric and topology of geometrical organisation of porous medium. A development in material science imaging is actually taking place involving for example X-ray, neutron, electron 2D/3D microscopies. However, several obstacles remain. The image processing needs an accurate and robust segmentation. This step is often badly controlled with no matching between the segmented phases and the real phases, although efficient tools of segmentation already exist, for example the watershed transformation in mathematical morphology. Since most often we have only access to the observation of 2D sections through a 3D sample, an important question is how to do an analysis allowing to quantify statistical information,

like the average pore size, the mean curvature, the pore shape, the surface roughness and the structural correlation, the isotropy, the homogeneity and the class of disorder. Using this extended metric information, another challenge is to reconstruct a realistic 3D configuration from the 2D sections that would enable the topological and diffusion properties of the original medium to be retrieved. Finally, an important question is to know if it it possible to simplify the geometrical description of the 3D pore network using a small number of functions that are well representative for the analysis of transport properties.
The main objective of this PhD work is to improve the processing of experimental images in order to extract quantitative information, then to understand the relation between this quantitative information and the diffusion properties. The principle components of this work are as follows:

- In chapter 1, we rapidly present cement paste as an example of a reactive porous material where the diffusive transport takes place. Two papers will be reviewed to stress the guiding ideas of our work.
- In chapter 2, we describe the two experimental techniques: Scanning Electron Microscopy (SEM) and X-ray tomography that are used to image two cement pastes (Portland and alite) during the setting.
- In chapter 3, we introduce a theoretical framework for a class of algorithms, called Seeded Region Growing by Pixels Aggregation/Dissolution (SRGPAD). This framework allows the fast implementation of advanced and optimised algorithms dedicated to segmentation, decomposition of the porous medium in elementary pores and permutation localisation on the phase boundaries.
- In chapter 4, we apply these algorithms to get a simple, generic and robust method to segment experimental two- or three-dimensional images of materials obtained by X-ray tomography and SEM.
- In chapter 5, we characterise the geometrical organisation at three levels. First, we present the extended metric analysis of 2D or 3D images, then we describe how to extract relevant information about the connectivity of the pore network on 3D images and finally, we decompose a porous medium in term of elementary pores using either a metric and a topological approach. The guiding idea of this decomposition will be to study restricted diffusion inside and between adjacent pores, and to get a coarse graining description of the diffusive transport in an effective network of elementary pores.
- In the chapter 6, we introduce an efficient procedure to reconstruct a representative elementary volume of a multi-phases material using metric information of a bi-dimensional imaging technique such as SEM. This last experimental set up is widely used in the cement industry and the possibility to restore 3D models from 2D observation should provide a gain of efficiency in a industrial context.

Finally, major results of this work as well as futur works are summarised in the conclusion.

Glossary

To simplify the reading, the notations will always be the same.

	Notations used in set theory
E	a vector space
$\mathbb{Z}$	the integers set
Ω	a subset of E
Ω^c	the complement of Ω
$\oplus$	the Minkowski addition, also called dilation, operator: $A \oplus B = \{a + b : a \in A \text{ and } b \in B\}$
A_x	the set A translated by the vector x
$\ominus$	the erosion operator: $A \ominus B = \{a + b : a \in A^c \text{ and } b \in B^s\}^c$
B^s	the symmetric of B, $B^s = \{x : -x \in B\}$.
$\circ$	the opening operator: $A \ominus B = (A \ominus B) \oplus B$
$\bullet$	the closing operator: $A \bullet B = (A \oplus B) \ominus B$
$\oplus_k$	the dilation operator times k: $A \oplus_k B = \overbrace{(.(A \oplus B) \ldots \oplus B)}^{\text{times k}}$
$+$	the disjoint union operator in the set theory: $A + B = \{A \cup B : A \cap B = \emptyset\}$
$\uplus$	the disjoint union operator in the set theory: $A \uplus B = \{A \cup B : A \cap B = \emptyset\}$
$-$	the inclusion restriction operator in the set theory: $A - B = \{A \setminus B : B \subset A = \emptyset\}$
Ω	a domain included in the space E
I	the characteristic function of Ω : $I(x) = 1$ if $x \in \Omega, 0$ otherwise
$\partial\Omega$	the boundary of the domain Ω (the closure of Ω without the interior of Ω)

	Notations used in seeded region growing by pixels aggregation/dissolution
X_i^t	a region (a domain of the space) at time t with the label i
Z_i^t	a zone of influence associated to the region X_i^t: $Z_i^t = (X_i^t \oplus V_i) \setminus (\bigcup_{j \in N_i} X_j^t)$
N_i	a restricted set, a subset of $\mathbb{N}$
V_i	a neighbourhood
$\delta(x, i)$	an ordering attribute function
$\mathcal{C}_{x,y}$	the set of continuous application from $[0, 1]$ to E such as the two extremities are equals to x and y
f	a grey-level image: an application of E to $\mathbb{Z}$
$\wedge$	the operator and in the symbolic logic
$d^{\Omega}(x, y)$	the lower bound of the paths length in Ω linking x and y

	Notations used in geometrical characterization
I_i	the characteristic function of the phase i
W_i	the i-Minkowski functional
ϕ_i	the volume fraction of the phase i
$\alpha_c(t)$	the degree of hydration at time t
$\mathcal{S}_i$	the specific surface area of the phase i
$S_{2,i}$	the 2-point probability function of the phase i (auto-correlation function)
$f_{i,\mu}$	the μ-chord length distribution function of the phase i
l_i	the mean chord length
N_3	the Euler-Poincaré invariant
K	the Gaussian curvature
p_c	the percolation threshold
ϕ_c	the volume fraction of the given phase at the percolation threshold
β_0	the number of connected components of the microstructure
β_1	the number of irreducible cycles
β_2	the number of internal surfaces or the number of connected components of the complementary
α_0	the number of vertexes of the topological graph
α_1	the number of edges of the topological graph
$N_c(v)$	the coordination number of the vertex v
$< N_c >$	the mean coordination number
C	the intensive topological number

	Notations used in annealing simulated algorithm
$S_{2,i,j}$	the 2-point probability functions of the phase i and j ($i = j$ auto, $i \neq j$ cross)
$d(\mathcal{M}, \mathcal{R})$	the distance between the model and the reference
$P(\mathcal{M}\|\mathcal{R})$	the probability of the model $\mathcal{M}$ given $\mathcal{R}$
Q	the selection matrix (the perturbation)
ρ	the acceptance matrix
$N_{i,i}^{\mathcal{I}^t}(d)$	the number of auto-correlation of the phase i, at the distance d, for the partition $\mathcal{I}^t$
$\mathcal{N}_{i,i}^{\mathcal{I}^t}(d)$	the number of chords of the phase i which size is d, for the partition $\mathcal{I}^t$
$p_g^{\vec{e}}(y)$	the phase label at left of the voxel, y, on the direction $\vec{e}$
$p_d^{\vec{e}}(y)$	the phase label at right of the voxel, y, on the direction $\vec{e}$
$l_g^{\vec{e}}(y)$	the length of the chord at left of the voxel, y, on the direction $\vec{e}$
$l_d^{\vec{e}}(y)$	the length of the chord at right of the voxel, y, on the direction $\vec{e}$

1

Bibliography overview: Image analysis versus diffusive transport in cement paste

Contents

1.1 Cement paste

Cement is made by heating limestone with small quantities of other materials (such as clay) to 1450°C in a kiln. The resulting hard substance, called 'clinker', is then ground with a small amount of gypsum into a powder to make 'Ordinary Portland Cement', the most commonly used type of cement (often referred to as OPC). Cement refers to a dry powder substance. Upon addition of water, the cement mixture is referred to as cement paste, with the addition of water and of sand, it is referred to a mortar, and with the addition of of water, of sand and of aggregate (generally a coarse aggregate such as gravel, limestone, or granite) is referred to as concrete. The cements set and harden because of chemical reactions when it is mixed with water. A category of cement, called hydraulic cements, retains strength and stability even under water. The key requirement for this strength and stability is that the hydrates formed by immediate reaction with water be essentially insoluble in water. Most construction cements today are hydraulic, and most of these are based on Portland cement.
The typical constituents of Portland clinker are

name	chemist notation	cement chemist notation	Mass
Tricalcium silicate	$(CaO)_3(SiO_2)$	C3S	45-75%
Dicalcium silicate	$(CaO)_2(SiO_2)$	C2S	7-32%
Tricalcium aluminate	$(CaO)_3\ (Al_2O_3)$	C3A	0-13%
Tetracalcium aluminoferrite	$(CaO)_4\ (Al_2O_3)\ (Fe_2O_3)$	C4AF	0-18%
Gypsum	$(CaSO_4)(H_2O)_2$		2-10%

Upon the addition of water, the hydration reactions of the anhydrous phases begin. These processes can vary widely depending upon the mix used and the conditions of curing of the product, but a typical

concrete sets (i.e. becomes rigid) in about 6 hours, and develops a compressive strength of 8∼MPa in 24 hours. The strength rises to 15∼MPa at 3 days, 23∼MPa at one week, 35∼MPa at 4 weeks, and 41∼MPa at three months. In principle, the strength continues to rise slowly as long as water is available for continued hydration, but cement paste is usually allowed to dry out after a few weeks, and the increase in strength is stopped. Hydration products formed in hardened cement pastes (HCP) are complicated, because many of these products have nearly the same formula (the hyphens in C-S-H indicate a phase of variable composition).

name	chemist notation	cement notation
Calcium Silicate Hydrate	$(CaO)_x\ (SiO_2)_y\ (H_2O)_z$	C-S-H
Calcium hydroxide	$Ca(OH)_2$	CH
Ettringite	$[Ca_3Al(OH)_6.12H_2O]_2.2H_2O$	$C_3A.3CaSO_4.32H_2O$
Aluminate Ferrite trisulfate	contains three anhydrite molecules: $C_3A.3CaSO_4.32H_2O$	AFt
Aluminate Ferrite monusulfate	contains one anhydrite molecule: $C_3A.CaSO_4.A2H_2$	AFm

The hydration reaction can be decomposed into two steps (see figure 1.1):

1. dissolution of the ions in the solution from the anhydrous phases. For example, the dissolution of the calcium and silicate from C3S:

$$Ca_3SiO_{5(s)} + 3H_2O_{(l)} \rightleftarrows 3Ca^{2+}_{(l)} + H_2SiO^{2-}_{4(l)} + 4OH^-_{(l)}$$

   where the subscript (s) means a solid phase, (l) means a liquid phase.

2. precipitation of hydrate phases from the solution. For example, the precipitation of C-S-H and CH from water solution of calcium and silicate ions:

$$\begin{aligned} 1.5Ca^{2+}_{(l)} + H_2SiO^{2-}_{4(l)} + OH^-_{(l)} + H_2O_{(l)} &\rightleftarrows ((CaO)_{1.5}(SiO_2)(H_2O)_{2.5})_{(s)} \\ Ca^{2+}_{(l)} + 2OH^-_{(l)} &\rightleftarrows Ca(OH)_{2(s)} \end{aligned}$$

The dissolution-precipitation reaction and the pattern formation of the porous medium (space distribution of the solid and liquid phases) depend on many parameters:

1. the constituents of the dry powder substance of cement,
2. the ratio of water to cement (W/C) ,
3. the temperature [110],
4. the relative humidity [157],

The porous medium exhibits a geometrical organisation in a wide range of length scales from nanometres to millimetres. Powers and Brownyard [136] distinguished two kinds of pores: gel water pores (under the influence of adsorbing forces) from the nanometre to 0.1 micrometer and capillary water pores ("free" water) form 0.1 micrometer to millimetre. The geometrical features characterising the porous medium are related to a wide range of different mechanical, physical and chemical properties. One important issue, connected to the concrete durability, is the understanding of the transport properties. This is a crucial point in order to control the invasion dynamics of various molecules inside the porous medium. One effect of these possible contaminants is the corrosion of the structural steel induced by the steel depassivation when pH decreased (for example the chemical reaction of portlandite, $Ca(OH)_2$, and calcium silicate hydrate, C-S-H, in the cement matrix with carbon dioxide gas leading to calcite $CaCO_3$ for example). Since iron oxide formation is accompanied by an increase in volume, it should lead to micro-cracking in the surrounding cement paste [184, 131] (see figure 1.2). The two main origins for molecular transport are diffusion (due to a difference a concentration) and convection (due to a difference in pressure). Molecular diffusion is usually slower. However, for a porous medium, having pores smaller than a few microns, since convection is slowed down by the viscosity effects, molecular diffusion is predominant and allows the chemical species to move in the porous medium. As the studied material is cement paste with a pore size smaller than a few microns, our work will be focused on molecular diffusion.

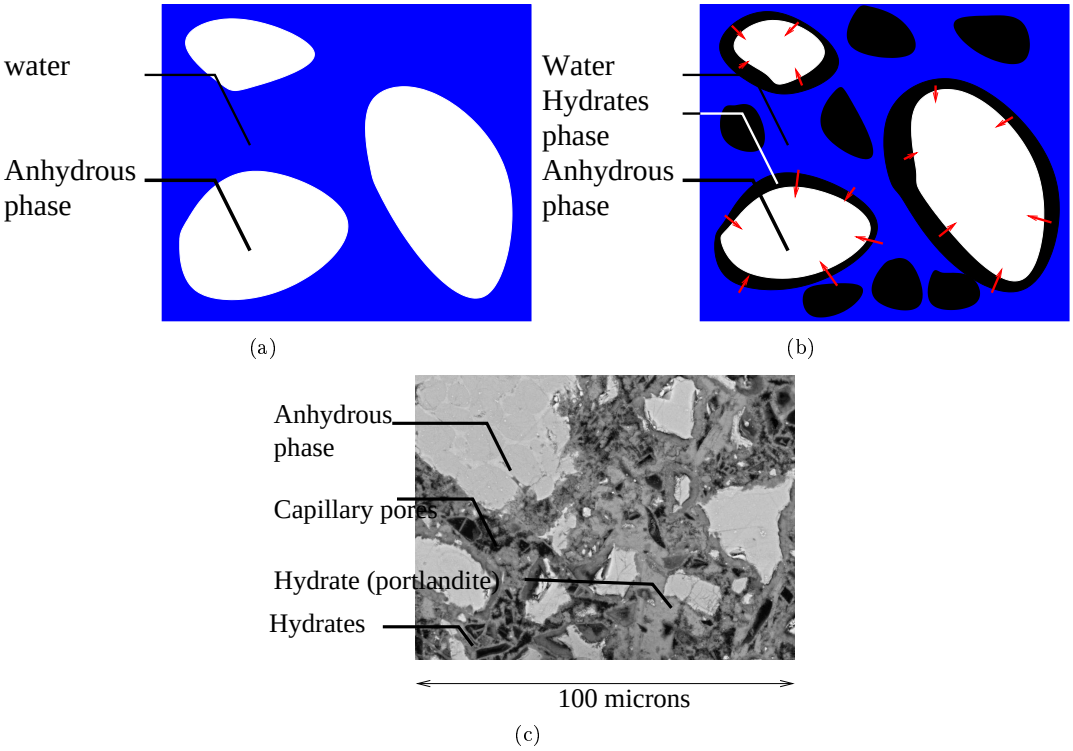


Figure 1.1: Dissolution-precipitation reaction. (a) at time of hydration = 0, (b) at the beginning of the hydration, (c) Portland cement image obtained by SEM at time of hydration = 1 day.

1.2 Diffusion in a confined domain

Let us consider a bath at uniform temperature without fluid motion. At initial state, a black ink drop falls in the bath. Propagation of the colour occurs due to the molecular diffusion. There is a flux, $\vec{j}$, of molecules (pigments) from regions of high concentration (where the drop is fallen) to regions of low concentration. In 1845, Fick introduced a macroscopic law of diffusion, which governs the molecular diffusion:

$$\vec{j} = -D_0 \vec{\nabla} c \quad \text{with } \vec{\nabla} \text{ the gradient operator .} \tag{1.1}$$

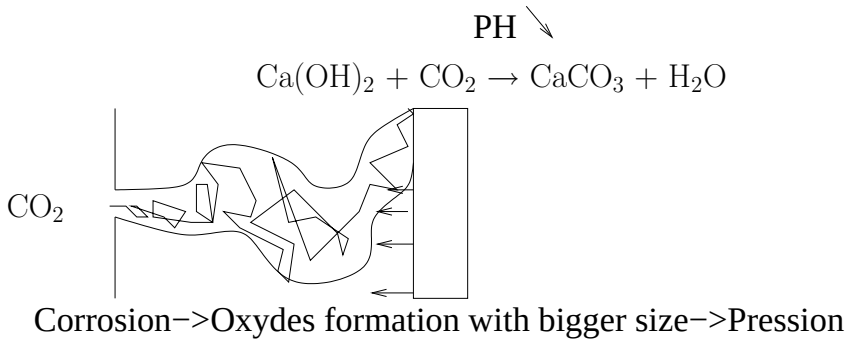
where c is the molecular concentration and D_0 is the free diffusion coefficient depending on the temperature and the interactions between the species and the solution.
Using the mass conservation law,

$$\frac{\partial c}{\partial t} + \nabla \cdot \vec{j} = 0 \text{ with } \nabla\cdot \text{ the divergence operator,} \tag{1.2}$$

we get the second Fick's law:

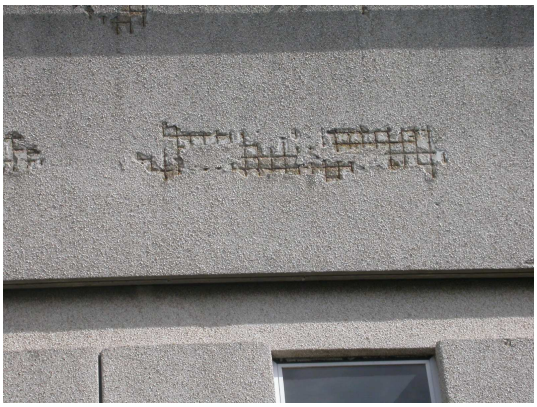
$$\frac{\partial c}{\partial t} = D_0 \nabla^2 c \quad \text{with } \nabla^2 \text{the Laplace operator.} \tag{1.3}$$

By dimensional analysis, we find that the diffusion characteristic length is proportional to $\sqrt{t}$ (for a water molecule at 20^oC, this length is in the order of 100 μm for $t = 1s$) . This slow dynamics is observed in the diffusion of black ink drop in the bath. The diffusion inside a porous medium of cement paste is also slow (see figure 1.3). In a confined domain (the porous medium of cement paste), a good understanding of the diffusive transport will involve:



(a)

(b)



(c)

Figure 1.2: Carbonation. (a) Transport of the carbon dioxide gas inside the porous medium, (b) front of PMC Laboratory (c) zoom on the top of front of PMC Laboratory-cracks due to the corrosion of the structural steel.

- the water saturation inside the porous medium since the water diffusion coefficients in liquid phase and in gas phase are different. In most materials made of cement, the relative water humidity is above 30%, large enough to have both water gas inside the capillary porosity and water liquid inside the gel porosity [13].
- the electro-chemical gradient since, if the diffusing species are ionic, their motion depends also on the electro-chemical gradient ($j^e = -uc\nabla\psi$ where j^e is the flow due to electro-chemical gradient, c is the concentration, u is the ionic mobility, $\nabla\psi$ is the gradient of local potential) [145]. The water inside the cement paste is an ionic solution (pH=13): it is a electrolyte. By imposing an electro-chemical gradient (see figure 1.4), the measurement of the diffusivity of ions species in water-saturated hardened cement pastes [130, 100] can be done using the Nernst-Einstein relation:

  $$u = \frac{DzF}{RT}$$

  where the z the valence of the ionic particle, $F = Ne$ is the Faraday's constant, T is the temperature and R is the gas constant.
- the surface chemistry and the surface reactivity, for example, the diffusing species can be chemically transformed into other species after hitting the surface [136]. In cement paste, there is a possibility of a strong carbonation, the chemical reaction of the interface (portlandite, $Ca(OH)_2$, and calcium silicate hydrate, C-S-H, in the cement matrix) with diffusing species such as carbon dioxide gas leading to calcite, $CaCO_3$.

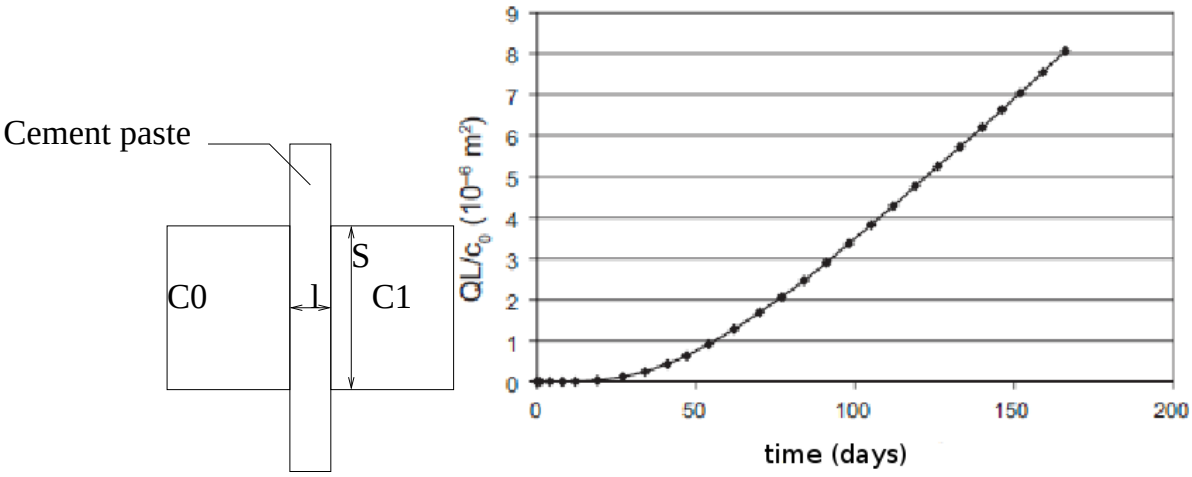


Figure 1.3: Experimental measurement of the diffusion coefficient. In steady state, the constant flow is given by $j_x = \frac{\Delta Q}{S\Delta t}$ where ΔQ is the diffused quantity in the time interval Δt through the section S of the specimen. The effective diffusion coefficient is calculated using the first Fick's law with the relation: $D_e = \frac{j_x}{|c_1 - c_0|} l$ where c_1 and c_0 are the concentrations in the two chambers. For a concrete with a thickness equal to 3 cm, where a constant concentration gradient is applied between the two faces, the steady state flow is obtained after one year [45] (corresponding to the linear part of the curve at long time).

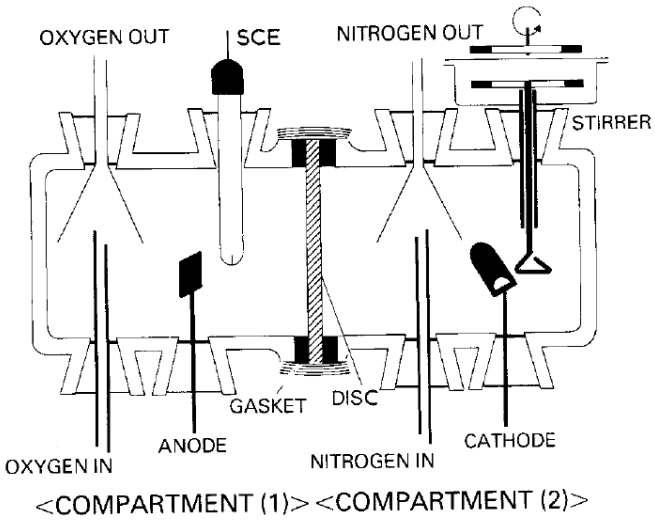


Figure 1.4: Migration of the ionic particles under a electronic field. The advantage of this experiment is that the steady state is rapidly reached, allowing the diffusion coefficient to be measured in a reasonable time (5 days).

- the geometry of the confined pore network.

In this thesis, we focus on the link between the diffusion transport and the geometry of the porous medium of cement paste. We assume that the pore network is homogeneously filled either by liquid or gas water, there is no electro-chemical gradient and finally the interface is considered as inert.
In a homogeneous isotropic porous medium, the effective Fick law is:

$$\frac{\partial c_e}{\partial t} = \phi D_e \nabla^2 c_e. \tag{1.4}$$

where D_e (m^2/s) is the effective diffusion coefficient and c_e is the effective molecular concentration (ϕ is the volume fraction of pore phase). Experimentally [130, 100], we measure the effective diffusion coefficient (see figure 1.5). D_e may be related to its free diffusivity, D_0 (m^2/s), by the following equation:

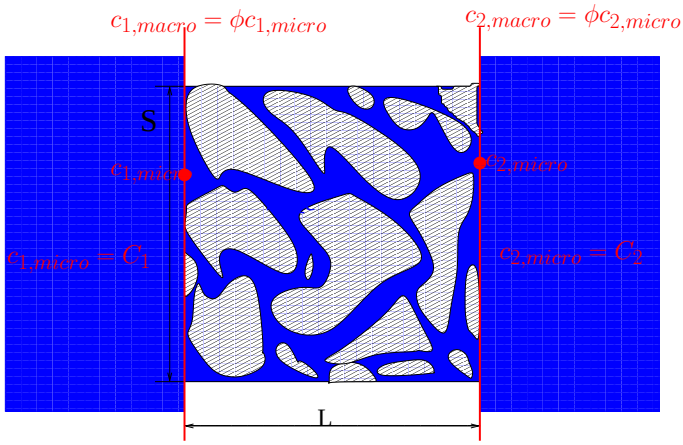


Figure 1.5: Experimental measurement of the diffusion coefficient. In steady state, the constant flow is given by $j_x = \frac{\Delta Q}{\phi S \Delta t}$ where ΔQ is the diffused quantity in the time interval Δt through the section S of the specimen. The microscopic diffusion coefficient is $D_e = \frac{j_x}{|c_{1,e} - c_{2,e}|} L = \frac{j_x}{|C_1 - C_2|} L$ since the boundary conditions impose $C_i = c_{i,e}$.

$$D_e = \frac{1}{\tau} D_0 \tag{1.5}$$

where τ is dimensionless parameter ($\tau \geq 1$) called the tortuosity of the porous medium. In fact, the above relationship attempts to scale D_0 to D_e by including a factor that is the increased transport path length due to the geometrical organisation of the porous medium of cement paste (PMCP) (τ). This physical parameter is different to the geometrical tortuosity. The direct or indirect estimation of the tortuosity can be done by:

- the utilisation of the Archie's empirical law [4] where we have:

$$\tau \sim \frac{1}{\phi^\alpha} \quad \text{with} \quad \frac{1}{2} < \alpha < \frac{3}{2} \tag{1.6}$$

  Although this approach is commonly used in reservoir engineering, it is useless to predict the tortuosity for an unknown material.

- the proposition of a "toys model" of the porous medium based on experimental constraints [105] or the coarse graining description in using a perturbation method [156] or self consistent method [17, 43]. The common input parameters of these methods are the porosity, the specific surface area, and the structural correlation function.

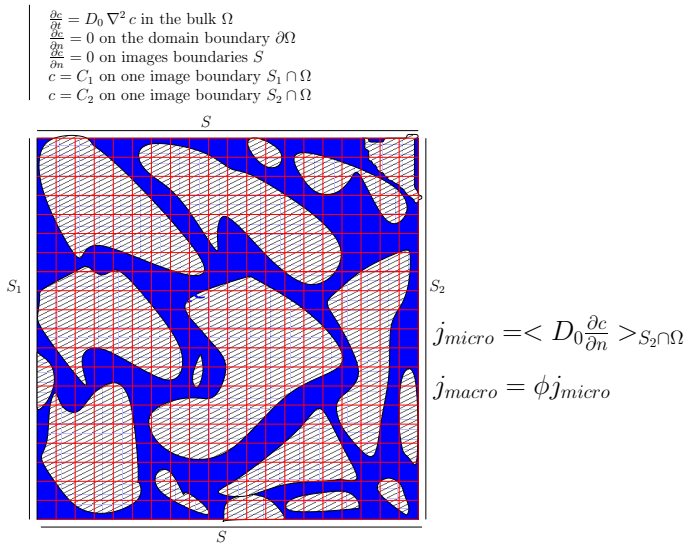


Figure 1.6: Finite difference. The grid is in red.

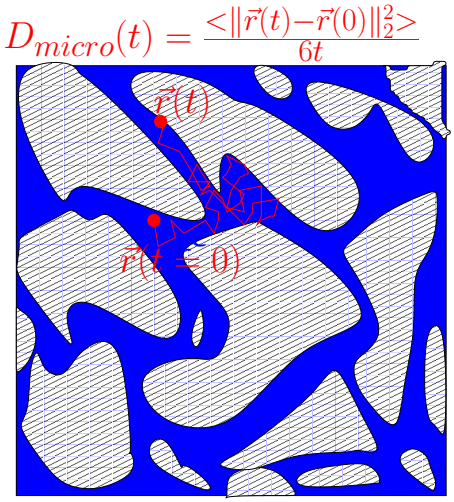


Figure 1.7: For the random walk methods, with the assumption of a self-diffusion and an unbounded domain, we have $D_e = \lim_{t\to\infty} D_e(t)$ and $D_e(t) = \frac{<\|r(t)-r(0)\|_2^2>}{2dt}$ with d the dimension.

- the simulation of the diffusive transport in the three dimensional microstructure. The computational methods are finite difference [3, 6] (see figure 1.6) or random walk methods [151] (see figure 1.7). In order to get the three dimensional microstructure, the solutions are to propose a "toy model" or to use an 3D imaging technique.

In this thesis, we will focus on this last approach with the utilisation of two imaging techniques: Scanning Electron Microscopy (SEM) and X-ray tomography to see the microstructure. In the next section, we comment two papers related to this thesis.

1.3 Link between the geometry of cement paste and transport properties

1.3.1 Estimating transport properties of mortars using image analysis of SEM images

Wong *et al* [180] study the porous medium of two ordinary Portland cement mortars at a water to cement ratio of 0.35 and 0.70 respectively. They correlate two simple metric parameters, the porosity and the specific surface area, obtained by using quantitative scanning electron microscopy, with transport coefficients obtained experimentally. Then, they introduce some models incorporating tortuosity and constrictivity to predict oxygen diffusivity and permeability.
The work is in five parts: selection of the specimens, segmentation, comparison between the metric parameters obtained by image analysis and by other indirect methods, correlation between transport properties and metric parameters, and proposition of a model.

Selection of the samples

Wong *et al* select the samples in order to obtain a wide range of PMCP characteristics:

- Ordinary Portland Cement (OPC) and medium graded (BS 882:1992) siliceous sand were used to prepare two mortar mixtures at different proportions,
- high and low water/cement ratio,
- curing period (time of hydration), 2 days or 28 days,
- conditioning regime: temperature, 20^{o}C, 50^{o}C, 105^{o}C, and relative humidity, 100%, 55%, 10%.

However, as the pore structure of cement paste strongly depends on its chemical composition, they do not explore a wide range of possible PMCP since they selected only one kind of cement paste (OPC).

Segmentation

The quantitative analysis of images needs an accurate segmentation. Segmentation is the transformation of a grey-level image to a labelled image where a label is associated to a phase. The classical segmentation procedure is a threshold segmentation based on:

- simple pre-filtering such as median filter, erosion, dilation, opening, closing,
- thresholding,
- simple post-filtering such as erosion, dilation, opening, closing.

To our knowledge, threshold segmentation using tint information is the only method applied to extract the different phases of cementitious materials [120, 123, 137, 182, 181]. This segmentation procedure is suitable when a good contrast between the phases gives a specific peak for each phase in the histogram. When this is not the case, as in the case of the porosity on the backscattered electron images, first the selection of the threshold using the histogram is difficult and second, the segmented pore structure will not match the real pore structure. In order to do a comparative analysis, Wong *et al* determine the upper threshold grey-level for pores from the inflection point of the cumulative brightness histogram of the BSE image (see figure 1.8). However, this segmentation does not allow a accurate prediction of the physical behaviour or properties of the material since the segmented microstructure does not match the real microstructure.

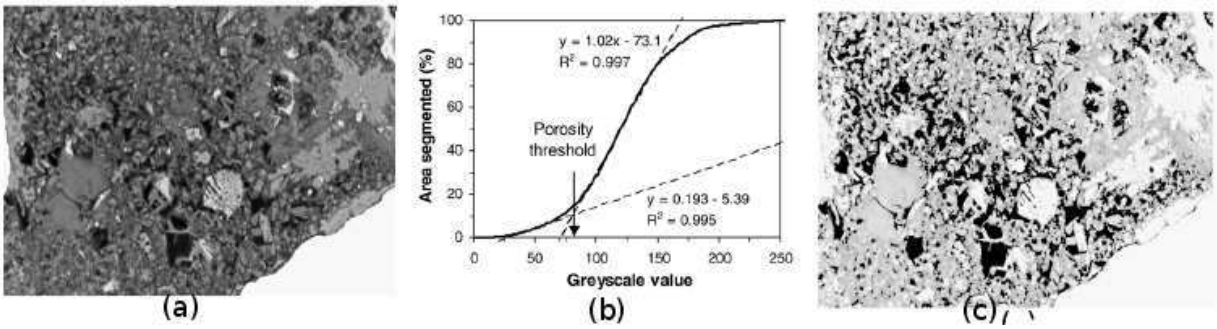


Figure 1.8: (a) Typical BSE image of cement paste. Field of view: 240x180 μm^2; (b) cumulative brightness histogram showing the location of the upper threshold value for porosity; (c) segmented pores (black pixels) superimposed onto the original faded BSE image. Pore area fraction is 17.8

Comparison between the metric parameters obtained by image analysis and by other indirect methods

As noted by Wong *et al*, the comparison between specific surface areas measured by image analysis and by other indirect methods (gas sorption [157], small angle scattering using neutrons [179], and nuclear magnetic resonance relaxation [92]) would lead to a large discrepancy since the indirect methods cover a wider range of pore sizes including gel pores, small capillary pores and large air voids. As capillary pores control the transport properties in cement paste, the metric parameters measured by image analysis are suitable since they are calculated using images at this specific range of length scales.

Correlation transport properties and metric parameters

Oxygen diffusivity is measured by imposing an electro-chemical gradient [130, 100], and oxygen permeability is determined by measuring the steady-state flow rate when a pressure difference is applied between two faces of the sample. Wong *et al* find a strong correlation between the measured pore properties and effective diffusive coefficient (see figure 1.9) emphasising that:

1. **the range of porous medium sizes analysed by SEM is relevant,**
2. **the tortuosity follows the Archie's law: $\frac{1}{\phi^\alpha}$ with $\alpha = 1$.**

Modelling

Wong *et al* introduce some models to predict oxygen diffusivity and oxygen permeability. However, the input parameters of these models are the porosity, the geometrical tortuosity and the constrictivity. The problem is that the geometrical tortuosity parameter cannot be measured by SEM. The study of Béjaoui *et al* [15] shows that the capillaries porosity of Portland are almost equivalent to the capillaries porosity of Portland-Pozzolanic with the same ratio W/C, but for a given W/C, the diffusion coefficient in Portland is twenty times smaller than the diffusion coefficient in Portland-Pozzolanic, despite a larger porosity (see figure 1.10). Therefore, **the tortuosity strongly depends on the chemical composition of the cement.**

Conclusion of this work

To summarise,

1. SEM allows to image PMCP at relevant range of length scale to study transport properties. However, the relative similarly of the electron back scattering emission coefficients in the pores and C-S-H makes the segmentation difficult and in more the partial volume effect cannot be neglected.

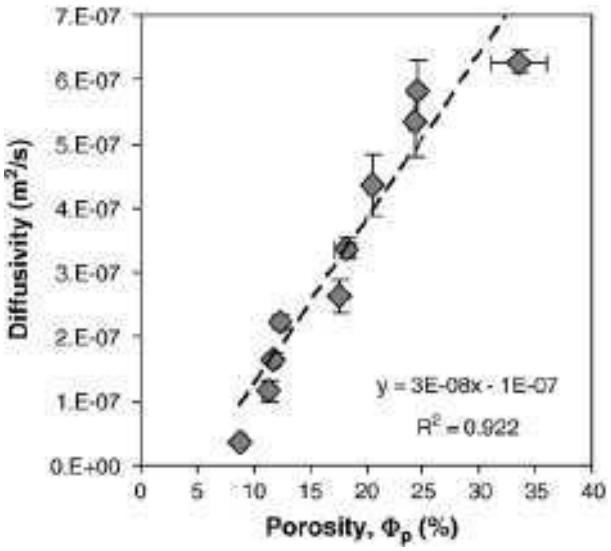


Figure 1.9: Correlation between the porosity and diffusion coefficient for OPC mortar.

2. The geometrical/physical tortuosity strongly depends on the chemical composition of the cement.

3. The geometrical/physical tortuosity cannot be estimated by using SEM images because the images are two-dimensional sections of a three-dimensional microstructure. 3D imaging techniques or an algorithm allowing the 3D reconstruction of multiphase material from a 2D observation are then required to allow the simulation of the diffusive transport in the three dimensional microstructure given the estimation of the tortuosity.

The second article is related to the application of 3D imaging techniques using X-ray microtomography.

1.3.2 X-ray microtomographic studies of pore structure and permeability in Portland cement concrete

Lu *et al* [107] studied the pore structure of conventional concrete mixture and three Pozzolanic-modified concrete mixtures. They related a three-dimensional parameter, called "disconnected pore distance", obtained in X-ray microtomographic images, to standard measures of chloride permeability.

Selection of the samples

Lu *et al* use conventional concrete mixture and three Pozzolanic-modified concrete mixtures cured at 28 days in order to represent those typically used in transportation applications in the Northeast U.S. They explore the wide range of pore structures of concrete, varying the chemical composition.

Quantitative analysis

In order to study the pore connectivity, Lu *et al* analyse the percolation of the pore space in looking at how the connected components hitting one surface are connected to the opposite surface. For all the samples, the pore space is not connected at the resolution of $1\mu m$ for a mature concrete [137, 55, 139]. They observe that the difference between the cement pastes is the depth of the pore network connected with the face. Lu *et al* evaluate this depth with a distance, called disconnected pore distance. Then, they find a correlation between this distance and the chloride permeability (see figure 1.11)

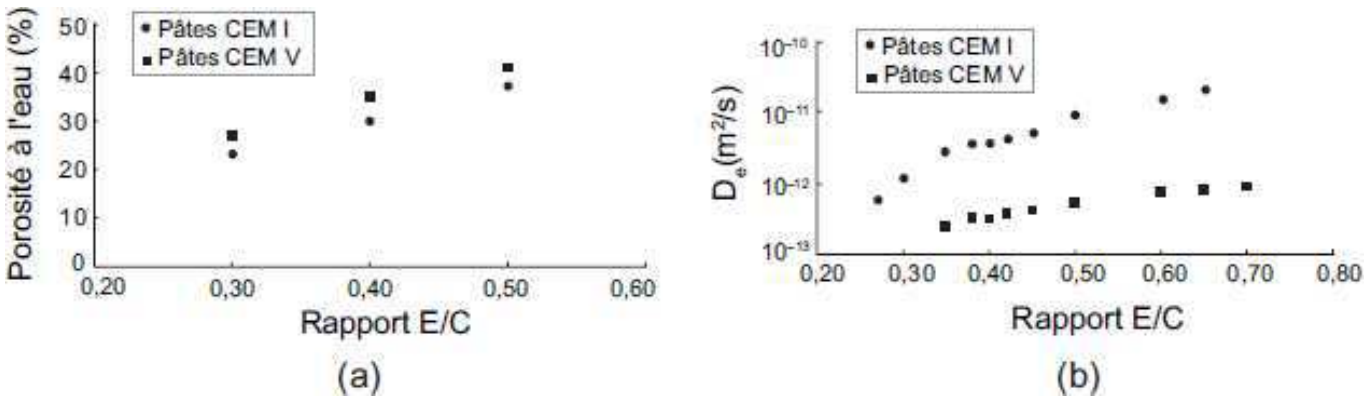


Figure 1.10: (a) capillary porosity as a function of the ratio W/C , (b) diffusion coefficient as a function of the ratio W/C [15].

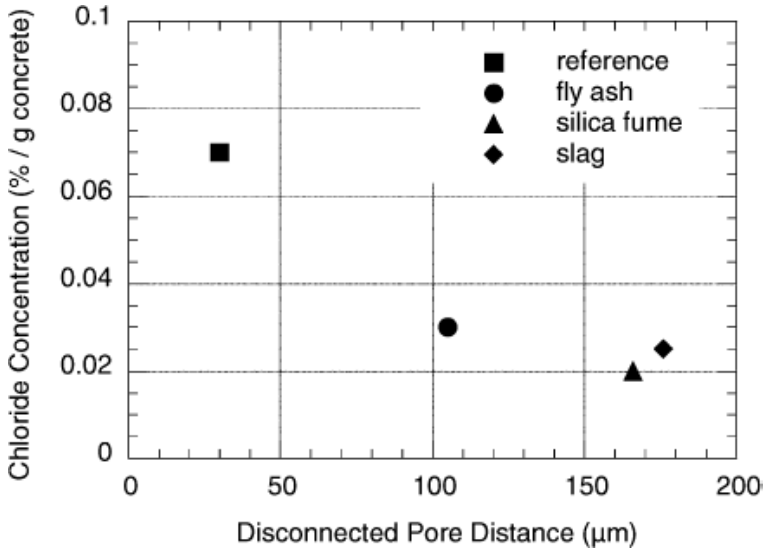


Figure 1.11: Chloride permeability plotted against disconnected pore distance for four different concretes[107]. For details of the link between the chloride permeability and the chloride concentration see [50].

Conclusion of this work

To summarise,

1. the main drawback of X-ray microtomography is its spatial resolution which is at present quite low relative to the sizes of capillary pores controlling transport properties in mature cement pastes [55]. This fact is observed by the absence of connexion in the capillary pores. Nevertheless in the early stages of hydration, the size of the capillary pores is large enough in comparison to the image resolution that allows to quantify the tortuosity. This information can be interfaced with microstructural models, which in turn should allow extrapolation to mature pastes [55].

2. as for SEM imaging, the relative similarity of X-rays attenuation coefficients for pores and C-S-H makes the segmentation difficult,

1.4 Conclusion and discussion

In 2D, SEM allows a qualitative and quantitative analysis of a mature paste in the relevant range of length scales. In the cement paste literature [51, 154, 155, 180], the metric analysis is focused on the volume fraction and the specific surface area of the phases. However, a more extended metric analysis would give important information about parameters such as the average pore size, the mean curvature, the surface roughness, the structural correlation, the isotropy, the homogeneity and the class of disorder (see section 5.2). Also, a challenge will be to reconstruct a realistic 3D configuration from the 2D image that would enable to retrieve the topological and diffusion properties of the original medium (see chapter 6).
In 3D, X-ray microtomography is relevant to image PMCP only in the early ages of pastes. It will be interesting to observe the disconnection of the pore network as a function of the curing times (see section 5.3).
For both imaging techniques, the low grey-level contrast between the phases makes the extraction of the phases difficult and practically impossible by standard thresholding procedures. An advanced segmentation procedure will be developed to capture realistic geometrical features and to accurately predict the material properties (see chapter 4).

2

Imaging methods: SEM and X-ray tomography

Contents

2.1 Introduction

2.1.1 Why 3D imaging?

The geometrical organisation of the porosity of cement paste affects numerous processes such as molecular diffusion, excitation relaxation, reaction kinetics, adsorption and capillary condensation as well as mechanical strength of concrete constructions and buildings. In order to especially understand how the diffusion occurs in this porous media, the requirements are:

- the statistical characterisation of the geometry of the porous microstructure,
- a theory to link this characterisation and the diffusion transport.

The characterisation of the geometry can be done by indirect measurement. For global geometrical features,

- the porosity can be measured using gas porosimetry based on the Boyle's law [35, 64],
- the specific surface area can be measured using Brunauer, Emett and Teller (BET) method [27, 127] or nuclear magnetic relaxation dispersion [12].

For statistical geometrical features (see subsection 5.2.4),

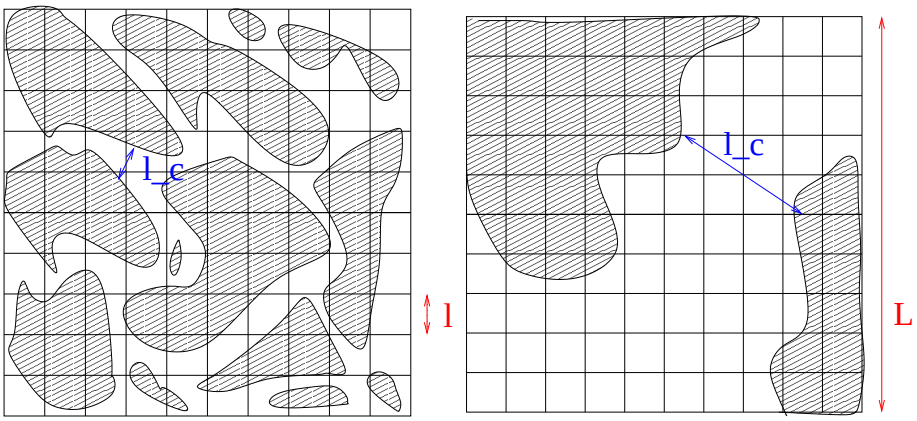
- the two-point probability function can be measured using small angle scattering [39] or direct energy transfer [104],
- the chord length distribution can be measured using small angle scattering. [115, 106, 109].

These metric features obtained by indirect methods are used:

1. to probe a reliable model of the real microstructure [105],
2. to correlate these measurements with a phenomenological law [109].

The initial goal of this thesis is to understand the link between the metric/topological features[1] and the diffusion transport. Since the diffusion transport depends on the topology and the metric of the porous microstructure and these indirect methods give only metric information, a 3D imaging method has been chosen in order to characterise the topology and the metric of the porous microstructure.
One image is a grid of the sample where the value of cell is the average of a physical interaction (e.g. attenuation) of the different solids and fluids contained into it at the moment of acquisition. An imaging technique is defined by two characteristic length scales: the resolution[2] (length of the cell) and the field of view (length of the grid). In order to a have complete spatial information of the microstructure,

- the resolution of the image has to be much smaller than the smallest feature size of the material (see figure 2.1),



(a) $l_c \sim l$ (b) $l_c \sim L$

Figure 2.1: For a complete spatial information of this two-phases material, we need to have $l \ll l_c \ll L$ with l the image resolution, L the image field of view and l_c the feature size of this material .

- the field of view of the image has to be much bigger than the biggest feature size of the material.

[1] The first one deals with average pore size, mean curvature, pore shape, surface roughness, structural correlation between pixels belonging to the solid, the interface or the pore network. . . The second is closely related to the long-range connectivity or percolation of pore network (Gauss curvature of the interface left angle bracket Kright-pointing angle bracket, deformation retract, genus of the interface). Several interesting properties should be analysed at this level such as the number of available paths linking two distinct points of the pore network, the metric distance between two points compared to their shortest (geodesic) distance and the important role of pore throats.

[2] The resolution describes the ability of an imaging technique to resolve detail in the material. Basically, to quantify the resolution, we estimate the minimum distance between two points still visibly resolved. The resolution fixes the limit length scale of the investigation and it is a physical parameter. Therefore, the length of the pixel/voxel size has to be larger than the resolution because there is no spatial information at the sub-scale.

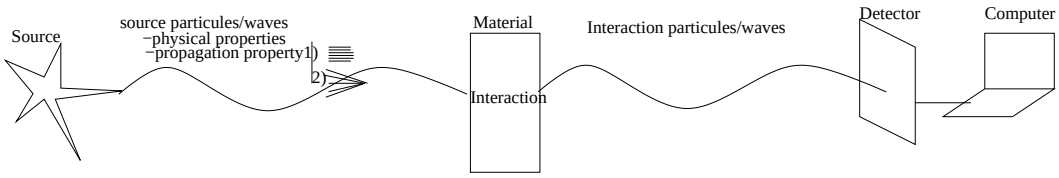


Figure 2.2: General principle of an experiment

The material under investigation is cement paste that is a hierarchical media. The shape of the porous media is organised over a large range of length-scale and, moreover, each length scale organisation is specific. Therefore, it is impossible to have an imaging technique covering this large range of length scale. In the thesis, we focus on the upper length-scale of the porosity of cement paste, called the capillary porosity[3]. To observe the cement paste capillary porosity, ideally, we need a resolution of 0.1 μm (to have one magnitude order of length scale between the resolution and the characteristic length of the capillary porosity) and a field of view of 200 μm (the characteristic length of the Representative Elementary Volume (REV) of capillary porosity is 200 μm, see subsubsection 5.2.4).

2.1.2 Choice of an imaging technique

Schematically, an imaging technique is (see figure 2.2):

1. a source generating waves-particles (e.g. electromagnetic radiation) with some physical properties (e.g. monochromatic),
2. an interaction between the material and the source generating waves-particles,
3. a detector to collect the waves-particles,
4. a computer to generate the mapping of the physical interaction.

There are several 3D imaging techniques:

- Nuclear Magnetic resonance Imaging (MRI) has seen increasing application for characterisation the structure of porous material, in rocks [44], in cement pastes [25], in bones [135]. This technique is based on the application of the strong main magnetic field (source 1) to align the nuclear magnetisation of hydrogen atoms in water on the same direction of this field (interaction 1), then a second electromagnetic field (source 2), that oscillates at radio-frequencies and that is perpendicular to the main field, is then pulsed to push the protons on the normal plan of direction of main field (interaction 2). So, at initial time, the protons rotate perpendicular of the direction of the main field with phase coherence. Then, the protons drift back into alignment with the main field, T1 relaxation, and loses progressively the phase coherence, T2 relaxation. These two physical phenomenons of relaxation emit a detectable radio frequency signal (emission). For the study of the capillary porosity, this method is of limited use since the best resolution of this technique, 40 μm, is much coarser than the biggest feature size of the cement paste porosity, 3-4μm. In cement, this method is used to give profiles of liquid diffusion or saturation, for example the moisture content profile for drying concrete [25].
- The Focused Ion Beam (FIB) nanotomography, based on the serial sectioning procedure with a sequence of erosion by FIB and 2D imaging by Scanning Electron Microscopy (SEM), has been

[3] In order to investigate the whole length-scale range of heterogeneity of cement paste, this thesis is coupled with two others thesis:

- S. Brisard supervised by P.L. Levitz (understanding the nanostructure of C-S-H: macroscopic implications),
- H. Chemni supervised by D. Petit (multi-scale diffusion and moisture scorpion in cement paste materials).

increasing use to characterise the structure of porous material, in ceramics [71], and in cement [70]. The resolution of 50 nm is suitable for the study of the capillary porosity. However, the field of view is 50 μm, much lower the characteristic length scale of the REV of cement paste, 200 μm. Moreover, this method is intrusive, causing irreversible modification of the pore matrix.

- The Transmission Electron Microscopy (TEM) tomography, based on TEM acquisitions at different angles, is a powerful technique to investigate the nano-geometrical organisation [91, 118]. For cement paste, the resolution is suitable for the study of CSH (thesis of S. Brisard). However, in the case of the present thesis, the characteristic length scale to be investigated is way above the nanometer.

- The laser scanning confocal microscopy, based on a sequence of focusing the beam on different layers of the material, has been applied in geosciences [54] in cement [93]. Due to the small size in depth of the image, this technique is suitable to investigate individual particle like the 'Hadley' grains [66] but for the study of porosity of cement paste, this technique is not appropriate.

- The synchrotron-based X-ray Tomographic Microscopy (SRXTM) based on X-ray acquisitions at different angles is a powerful, non-destructive, and high-resolution technique for the observation of many different kinds of materials [52]. The resolution, in order of 1 μm, is just little bit smaller than the characteristic length scale of the capillary porosity, 1 μm. The field of view, 500 μm, is much bigger than the characteristic length scale of the REV, 200 μm estimated in the subsubsection 5.2.4.

Two techniques have been used to image our cement sample:

1. *Synchrotron-based X-ray Tomography.* This is suitable technique for the study of capillary porosity but the resolution is limited. Because of the difficult access to the experimental set-up this technique is inconceivable for the cement industry, as the quantity of cement paste samples to analyse is high.

2. *SEM.* This is widely used in the cement industry to give qualitative and quantitative information. As this technique gives 2D observation, the quantitative analysis is limited. A standing and open question is to link the 3D physical properties with a 2D observation (see chapter 6).

The preparation of samples and the image acquisition have been done with the strong collaboration of the Laboratory of Ecole Polytechnique Fédéral de Lausanne with the assistance of E. Gallucci.

2.2 Scanning Electron Microscopy (SEM)

In this section, we briefly recall the principles of SEM, then we present the application for cement paste, by explaining the sample preparation.

2.2.1 Principle

The source of SEM is a monochromatic electron beam focused on a region of the sample surface. The interaction between this source and the atoms is electrostatic (interaction between the beam electrons and the electron cloud of atoms). The interaction generates secondary and back scattered electrons, characteristic X-rays and light. In our case, 2d-mapping of the microstructure will be built based on the phase contrast of back scattered electrons, since:

1. heavy elements (high atomic number) containing phases or dense phases backscatter electrons more strongly than light elements containing phases or poorly compacted phases due to the monotonic increase in electron backscatter coefficient η as a function of their weighted average atomic number [126, 37],

2. the electron penetration depth is in the range of 100nm [67], which is small enough to have a good resolution and deep enough to not to be affected by the remaining roughness of the polished cross section.

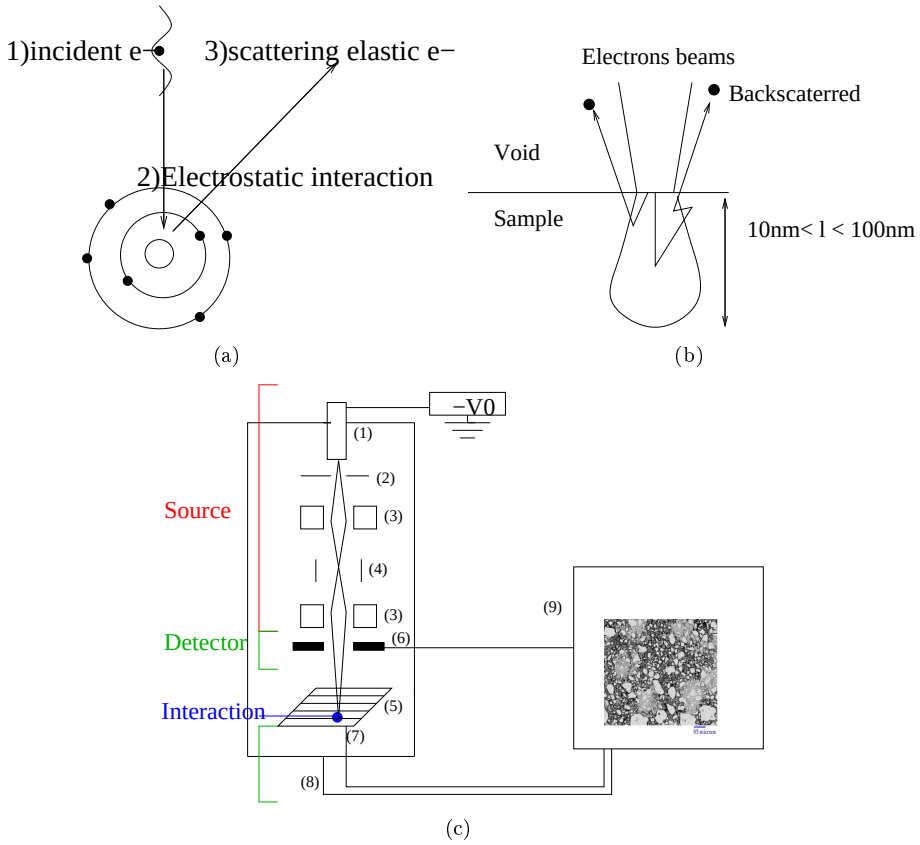


Figure 2.3: SEM. (a): interaction between an electron and an atom; (b): interaction between the incident beam of electrons and the sample; (c): a scheme of SEM (1) Electron gun with a negative voltage. (2) Anode to the mass. (3) Condenser lenses. (4) System of deflexion-scanning which deflect the beam horizontally and vertically (to scan the surface). (5) Sample fixed on a goniometer. (6) Detector of backscattered electrons. (7) I_a, Measure of the adsorbed current. (8) I_t, Measure of the transmitted current. (9) Visualisation of the image.

The detector is a double Everhart-Thornley detector located above the sample in a "doughnut" type arrangement, concentric to the electron beam. To obtain a surface mapping, the electrons beam scans the sample surface[46] (see figure 2.3).

The image contrast is the result of two effects: the topographic contrast and composition/atomic number contrast. In order to quantify the phases organisation, only the composition contrast is relevant. To remove the topographic contrast, the sample is filled in resin and polished. The brightness of each pixel is the average atomic number of the different phases contained into it. For the major phases present in Portland cement, the phases encountered from brightest to darkest pixels are tetracalcium aluminoferrite (C4AF), tricalcium silicate (C3S), dicalcium silicate (C2S), tricalcium aluminate (C3A) and gypsum, and the resin-filled voids. Because the back scattered electrons signal is weak, images are inherently noisy. To reduce the noise, the counting time is long.

2.2.2 State of art

Further pioneering work on the physical principles of the SEM and beam specimen interactions was performed by Manfred von Ardenne in 1937 [174]. In the study of cementitious materials, further pioneering work on quantitative backscattered electron (BSE) imaging in SEM, was performed by Scrivener *et al* in the early 1980 [153]. Since this time, SEM has demonstrated its great potential for the study of cemen-

titious materials for qualitative observations and quantitative measures [154]. For quantitative measures (see section 5.2), we can mention the determination of:

- the volume fraction of the void phase [180],
- the volume fraction of the anhydrous phase which gives the degree of chemical reaction, called the degree of hydration in cement industry [51] (chemical information) ,
- the metric parameters of cracks [124, 159] (micro-mechanical information),

The major drawback of this technique is the fact that only two dimensional sections of a three-dimensional microstructure may be observed. A standing and open question is to link the 3D physical properties with 2D observation. Some authors used a phenomenological approach to link some porosity metric descriptors obtained by SEM and physical properties such as compressive strength [129] and diffusion[4] [180]. Others reconstructed a 3D model using some metric descriptors obtained by SEM [138]. This last approach will be explored on the chapter 6.

2.2.3 Sample preparation

The preparation of polished specimens is relatively tedious, and requires special equipment and skills. Severe artefacts can result from improper preparation. This preparation is known to produce irreversible changes in the pores structure, particularly at small sizes.

Preparation of the slices

After mixing the water with cement, the cement paste is poured into a plastic cylindrical container. For a chosen hydration age, a slice is cut from the mother cement paste cylinder. The slice is immersed into isopropanol. The chemical reaction of hydration is immediately stopped because the chemical potential is changed. After three days, the water inside the slice is substituted by isopropanol. Then, the sample is put in a vacuum chamber during a half-day to substitute the liquid phase by the vapor phase. The capillary pressure of the isopropanol is lower than that of water (the interfacial energy/superficial strain of the isopropanol, $\gamma_{air/liq} = 22,3mN/m$, is lower than the superficial strain of the water, $\gamma_{air/liq} = 72.79$). Therefore, a weak vaccum is sufficient for this process without damaging the microstructure.

Epoxy impregnation

The sample is filled with an epoxy resin to stabilise and reinforce the microstructure. The slices are put under vacuum ($\sim$10Mbar) for half hour (enough time to reach the steady-state pressure between outside and inside the slice). Impregnation by the resin is performed with a small plastic pipe. At this level, the porosity is not filled but the porosity is isolated from the outside by the resin. The vacuum is then suddenly broken given a pressure difference between atmospheric pressure outside the sample and the vaccum pressure inside the sample. This gradient forces the resin to invade the porosity efficiently.

Polishing

Once cured, the resin layer above the sample is removed using a coarse SiC grinding paper (SiC grade : 1200). Polishing is done manually until the outcrop of cement paste. The outcrop is reached when the boundary between the cement paste and the resin is observed by sun reflects. Then, the sample is polished with a series of decreasing diamond suspension grades of 9, 3, 1, and 0.25 μm on a low-speed lap wheel using oil as a polishing lubricant for about 30 minutes for each step with a pressure of 25N. A 5-10 nm thick coating of carbon is evaporated onto the polished surface to eliminate specimen charging during SEM observation.

[4] The measure of diffusion has been done by oxygen permeability [88] but the NMR is a powerful alternative [99, 125].

Imaging by SEM

The results are presented in the table 2.1. In the figure 2.4-j and figure 2.5-j, four phases are mainly observed: :

- the bright grains are remaining unhydrated cores of tricalcium silicate (C3S),
- the light grey clusters including the bright ones are portlandite (CH), one hydrated phase,
- the dark grey phase is the main hydrated phases : calcium silicate hydrate (CSH),
- the black regions are the porosity.

In each column in the figure 2.4, we observed the advancement of the hydration reaction over time. For instance, in the last row, the anhydrous grains are surrounded by a small layer of CSH at the beginning (see first figure 2-4-c) but this layer increases over time (see figures 2-4-e,i). We observe that the size of the cluster of portlandite is much larger for alite than for OPC. The reason is related to the nucleation of the portlandite clusters which is less likely in the case of alite since the reaction of germination is driven by the heterogeneous precipitation. In the figure 2.4-2-4-c, some cracks are observed in the anhydrous grains due to sample preparation.

Cement	Water/cement	time sampled (days)
Alite	0.4	1
	0.4	3
	0.4	7
CEMI	0.5	1
	0.5	3
	0.5	7
CEMII-slag	0.4	1
	0.5	1
	0.4	3

Table 2.1: Samples observed by SEM. For each sample, the pixel size of the BE images are 1, 0.675, 0.25, 0.1, 0.025 μm

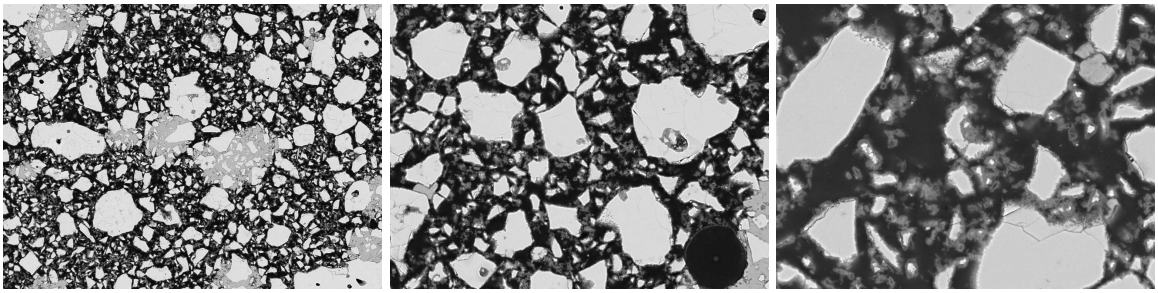
2.2.4 First comments on image analysis

Quantitative analysis of the geometrical features of a complex media needs as a prerequisite to propose an accurate segmentation of the image. The segmentation process consists of converting a grey-level image[5] to a labelled image[6]. In order to do a comparative analysis of the capillary porosity over hydration time, the segmentation has to be robust. For SEM, the major difficulty is the acquisition parameters at not constant over time. Between two successive acquisitions of the same cement paste, the images are different since the acquisition parameters are not stable (for example, the electrons energy obtained by Electron gun is not stable). Therefore, the shape of the grey-level histogram moves between each acquisition (see figure 2.6). In the chapter 4, we will apply a segmentation procedure based on boundary information allowing a reproducible result even with a shift of the grey-level.

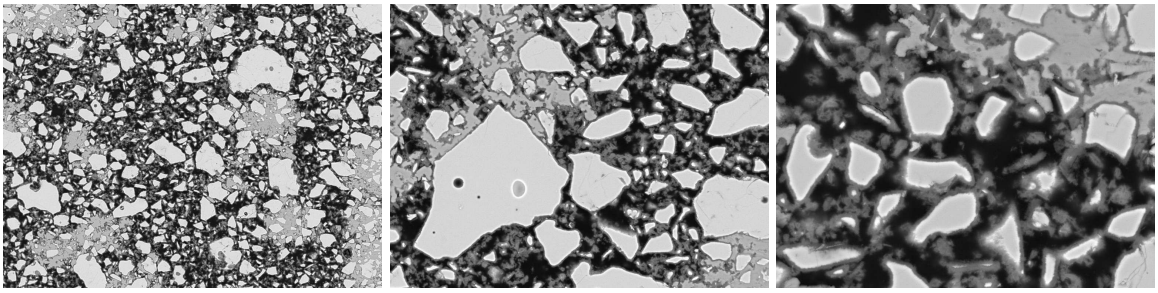
2.3 SRXTM

Synchrotron X-ray Tomographic Microscopy (SRXTM) is a powerful, non-destructive, and high-resolution technique to observe the 3D geometrical organisation of many different types of materials [162, 52, 28]. In this section, we briefly remind the principle of SRXTM, then we present the application of this technique to cement, after we explain the sample preparation and finally we present the results.

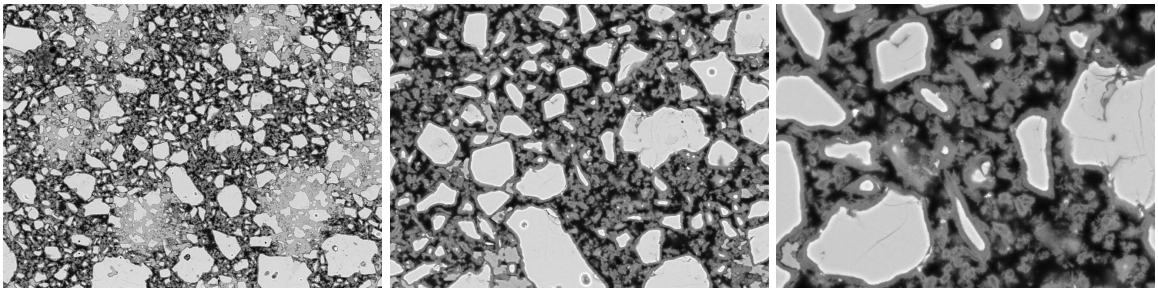
[5] A pixel is coded by a grey-level corresponding to a quantification of the physical phenomenon under investigation.
[6] A pixel is coded by a label corresponding to a phase of the material.



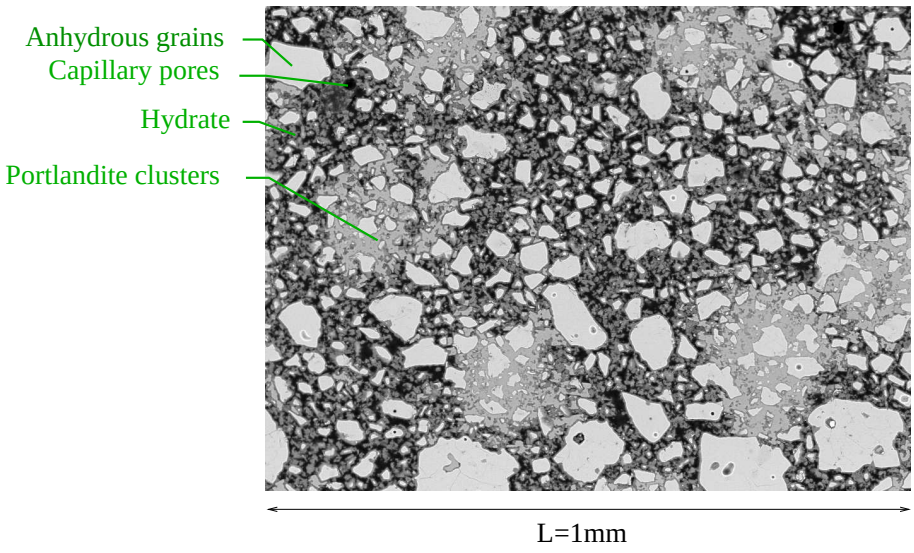
(a) l=0.675 μm, h = 1 day (b) l=0.25 μm, h = 1 day (c) l=0.1 μm, h = 1 day



(d) l=0.675 μm, h = 3 day (e) l=0.25 μm, h = 3 day (f) l=0.1 μm, h = 3 day

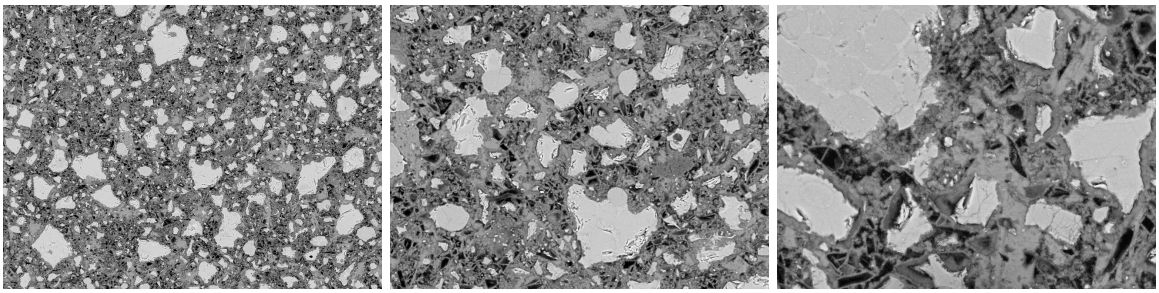


(g) l=0.675 μm, h = 7 day (h) l=0.25 μm, h = 7 day (i) l=0.1 μm, h = 7 day

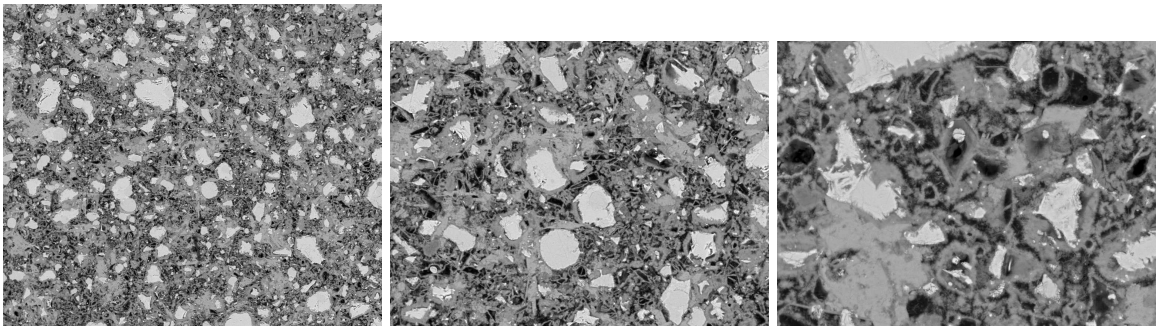


(j) l=0.675 μm, h = 7 day

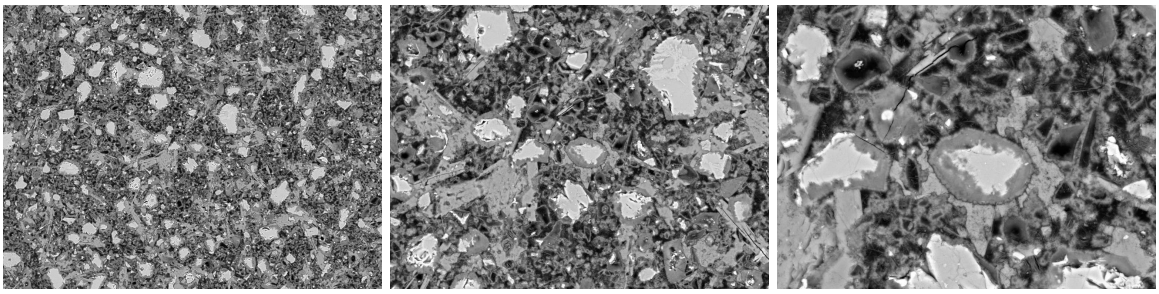
Figure 2.4: The cement paste is **alite** with W/C=0.4 observed by SEM. l = pixel size and h = hydration time.



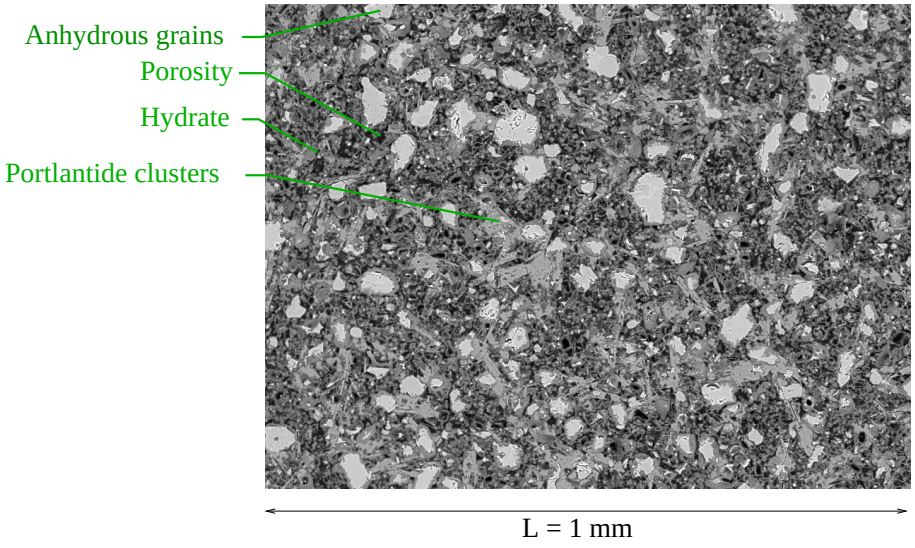
(a) l=0.675 μm, h = 1 day (b) l=0.25 μm, h = 1 day (c) l=0.1 μm, h = 1 day



(d) l=0.675 μm, h = 3 day (e) l=0.25 μm, h = 3 day (f) l=0.1 μm, h = 3 day



(g) l=0.675 μm, h = 7 day (h) l=0.25 μm, h = 7 day (i) l=0.1 μm, h = 7 day



(j) l=0.675 μm, h = 7 day

Figure 2.5: The cement paste is **OPC** with W/C=0.4 observed by SEM. l = pixel size and h = hydration time.

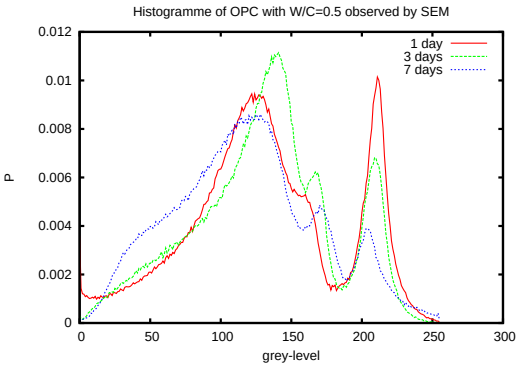


Figure 2.6: Histogram of OPC with W/C=0.5 at different times. If the acquisition parameters are still constants, the peaks of the histograms have always located at the same value. Here, it is not the case. Therefore, the acquisition parameters are not constants.

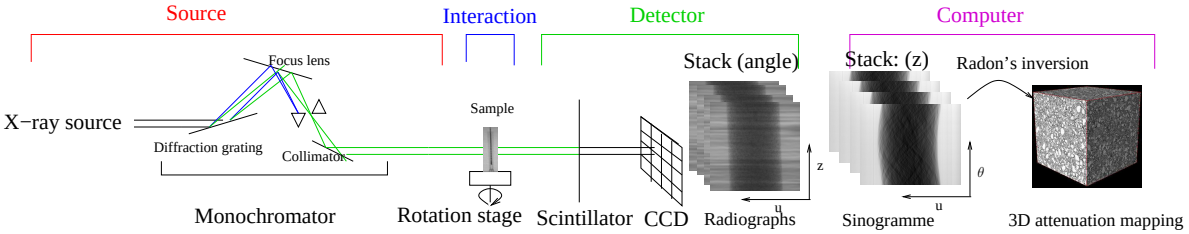


Figure 2.7: A scheme of SRXTM

2.3.1 Principle

There exists two kinds of X-ray sources: laboratory or synchrotron. For the study of the capillary porosity of cement pastes, synchrotron has been chosen because the highest resolution of laboratory-based X-ray Tomographic Microscopes (10 micrometers) is bigger than the feature size of the cement paste capillary porosity (3-4 micrometers). A synchrotron is a toroidal vacuum tube (a ring) surrounded by electromagnets in which electrons travel in the centre of the ring. The orbital motion is the result of a centripetal acceleration. During this acceleration, electrons lose energy in the form of photons, called synchrotron radiation. In the photon beam lines, end-stations (small laboratories) use this radiation for various applications. In tomography application, a monochromator filters the frequency spectrum of the photon beam to produce a bright, almost monochromatic[7], parallel photon beam (X-ray frequency).
The figure 2.7 shows the principle of X-ray tomography. Next, we will explain the different stages.

Interaction between the monochromatic photons beam and the material

Intersection between a X-ray and a cloud of identical atoms:
Let us considered a X-ray going through a cube containing a cloud of identical atoms represented by spheres. The absorption coefficient, μ, is defined in such a way that μL is the probability that a random line, with a normal direction to a side of the cube, will intersect with the spheres[8] included in the cube (see figure 2.8). Therefore, the probability is the ratio of the projected spheres surface area to the surface area of the side under the assumption of diluted spheres:

[7] in our application, the energy was 20 KeV.
[8] the spheres are diluted

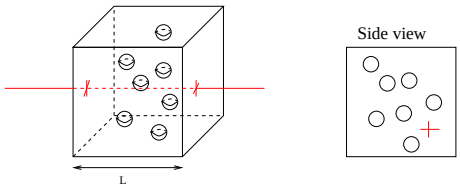


Figure 2.8: probability of interaction between a X-ray and a cloud of atoms contained in the cube

$$\begin{aligned} \mu L &= \frac{\text{sum of the cross sections}}{\text{surface area of the side}} \\ \mu L &= \frac{N\sigma}{L^2} \\ \mu &= \frac{N\sigma}{L^3} \\ \mu &= \frac{N\sigma}{V} \end{aligned}$$

where σ is the cross absorption section (1 barn $= 10^{-24}\text{cm}^2$) and N is the number of atoms contained in the cube. Since the density, ρ, is equal to $\frac{NA}{V\mathcal{N}}$ where A is the atom mass (g) and $\mathcal{N}$ is the Avogadro number, we obtain:

$$\mu = \rho\frac{\sigma\mathcal{N}}{A}$$

Tabulation:
The absorption coefficient depends on the density and, therefore, the physical state. For the purpose of tabulation, the absorption coefficient is divided by the density:

$$\frac{\mu}{\rho} = \frac{\sigma\mathcal{N}}{A}$$

The mass absorption coefficient μ/ρ (cm^2/g) is more or less independent from the physical state of the element allowing the elements tabulation.
Cross section absorption coefficient for one atom:
With wavelengths of the order of 1 Å, the interaction between the electromagnetic field and the atom affects the electronic system of the latter. Two main effects must be distinguished:

- *Photoelectric absorption.* The absorbed radiation is used to eject a lower-level from the atom, the photo-electron carrying away any excess energy in the form of kinetic energy. τ is the linear photoelectric absorption.

- *Scattering.* The X-ray photon is deflected from its original direction of propagation, with or without loss of energy, by collision with an electron or atom.

The cross section absorption coefficient is the sum of the cross section of these two interactions.
Intersection between a X-ray and a cloud of different atoms :
Let us consider a X-ray going through a cube containing a cloud of different atoms represented by spheres. The absorption coefficient is an additive measure (see figure 2.9). Therefore, the absorption coefficient is the sum of the absorption coefficients of each element, μ_i:

$$\mu = \sum_i \mu_i$$

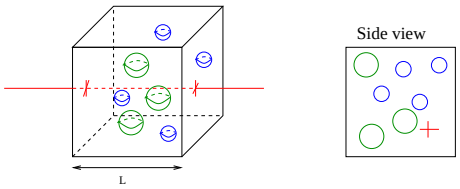


Figure 2.9: The additive attenuation comes from the additive surface for diluted sphere

The mass absorption coefficient is

$$\begin{aligned}\frac{\mu}{\rho} &= \sum_i \frac{\mu_i}{\rho} \\ \frac{\mu}{\rho} &= \sum_i \frac{\rho_i}{\rho}\frac{\mu_i}{\rho_i} \\ \frac{\mu}{\rho} &= \sum_i g_i(\mu/\rho)_i\end{aligned}$$

where g_i is the mass fraction contributed by the element i and $(\mu/\rho)_i$ is the mass absorption coefficient of the element i.

Radiograph:

The attenuation of a photon beam[9], ΔI, going through a cube, is:

$$-\Delta I = \mu I L \tag{2.1}$$
$$\tag{2.2}$$

For a heterogeneous material, taking an elementary cube, with a flow parallel to the direction y, the absorption law is:

$$-dI(x,y,z) = \mu(x,y,z)I(x,y,z)dy \tag{2.3}$$
$$\frac{dI(x,y,z)}{I(x,y,z)} = -\mu(x,y,z)dy \tag{2.4}$$

Integrating the eq 2.4 along a beam path parallel to y-axis, we get the transmitted intensity through a path known as the Beer-Lambert law:

$$ln(I_t/I_0) = -\int_{\text{path}} \mu(x,y,z)dy \tag{2.5}$$
$$I_t = I_0 e^{-\int_{\text{path}} \mu(x,y,z)dy} \tag{2.6}$$

In our experiment, the initial beam is parallel to a direction y. If a 2D detector is localised perpendicular to the y-axis and behind the sample, a radiograph is obtained where the value of each pixel is the measure of I_t. A blank image measures the incident intensity, I_0. The logarithm of the ratio of the measure of I_0 to I_t, is the integral of the attenuation along the path line using the eq 2.6 (see figure 2.10):

$$p(x,z) = -\int_{\text{path}} \mu(x,y,z)dy$$

From the radiographic images to the μ-mapping:

In order to obtain the μ-mapping, a set of radiographic images is recorded for different angular positions of the sample. Let p_θ be the radiograph for the angular position θ. We have (see figure 2.11):

[9] A sum of individual rays.

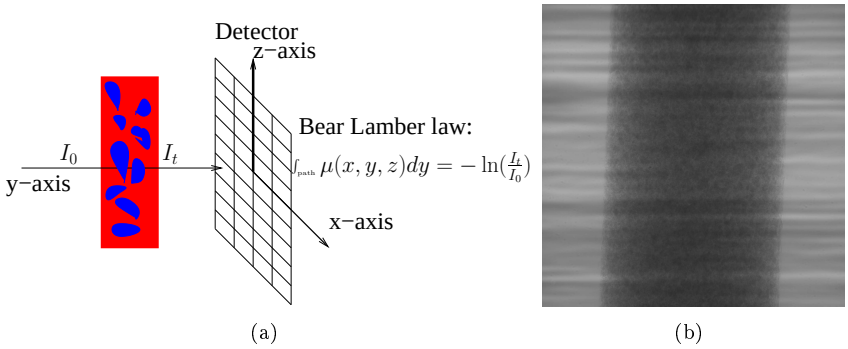


Figure 2.10: (a): Principle of radiograph; (b): the X-ray beam attenuation of an OPC with a W/C=0.4 and an a ageing equal to 1 day.

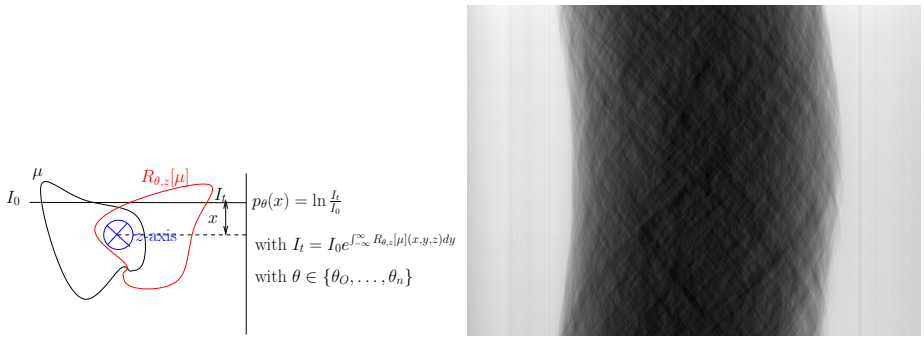


Figure 2.11: (a): acquisition of a radiograph with a angle θ; (b): sinogram of $p_\theta(x,z)$ of an OPC with a W/C=0.4 and an hydration time=1 day, **in axis** x, **in coordinate the angle** θ and **at** z **fixed**. Due to its definition, the sinogram is the addition of sinusoid as shown in figure b.

$$p_\theta(x,z) = \int_{-\infty}^{+\infty} R_{\theta,z}[\mu](x,y,z)dy \quad (2.7)$$

$$p_\theta(x,z) = \int_{-\infty}^{+\infty} \mu(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z)dy \quad (2.8)$$

where $R_{\theta,z}$ is the rotation operator of angle θ about the z-axis[10].
Slice by slice:
For a height fixed, z_{cst}, we will reconstruct a 2D slice using a sinogram. A sinogram is $p_\theta(x,z)$ where x and θ are the variables (with x in axis and θ in coordinate) and z is fixed (see figure 2.11). Given a sinogram, in 1917 Radon has proved that this inverse problem has a solution. The filtered back projection algorithm is usually used for this purpose [75] (see figure 2.12). To reconstruct the 3D-μ-mapping, we stack the 2D-μ-mapping done at different height (see figure 2.17). In our experiment[11], the typical resolution is 1 μm.

2.3.2 State of the art for cementitious materials

Tomographic microscopy has been used by many authors to quantify transport properties according to diffusion [150, 142] and permeability [161, 8]. In order to obtain representative quantitative data correlated with others experiments, the development of advanced computational algorithms[12] is required [146]. Many groups, for example in Japan [122, 123] or in Australia [89, 90, 6], have this capability.
The application of this technique to the study of cementitious materials can be seen from two angles. The first one is the study of global features of porosity in relation with a physical phenomenon:

- Burlion *et al* [29] quantified the evolution of porosity profile during an accelerated leaching process (see figure 2.13),
- Hu *et al* [74] studied the moisture gradient due to the surface drying. Since the mass of the sample is constant[13] except the variation due to the evaporation of the water caused by surface drying, the global X-ray transmission allows the global weight loss due to water evaporation to be measured, and, therefore, the drying profile to be assessed (see figure 2.14).

The second angle is the relation between a geometrical organisation and a physical phenomenon.

- At the length scale of millimetre, the study of the shape of aggregates in concrete has enabled the access of many concrete properties, especially the rheology of fresh concrete and early-age mechanical properties [56]. In order to characterise the aggregates on the segmented images obtained by X-ray tomography, Garboczi used spherical harmonic functions to interpolate each aggregate (see figure 2.15). Therefore, the classification the shape of aggregates in different categories has been possible.
- At a lower length scale 100μm, Landis *et al* [96] studied the formation of cracks under mechanical compression. Fracture energy was calculated using a linear elastic fracture mechanics approach and correlated with the measures of surface area of the internal cracks. This last measure is obtained by the analysis of X-ray tomography images (see figure 2.16).
- At the micro-length-scale, the microstructural analysis of the porosity has been done by many authors [94, 55, 137, 107]. Lu *et al* have found a good correlation between a parameter "disconnected pore distance" and standard measures of chloride permeability.

[10] An axis of rotation is defined by a direction and a translation vector $v = (v_x, v_y, v_z)$. Experimentally, the translation vector is unknown. Since the determination of v_y is required, this property is used: $p_\theta(x,z) = v_y - p_{\theta+\pi}(x,z)$ to find it.
[11] To see the matter more deeply with a resolution in the order of $0.1\mu m$, the nanotomography, based on the focusing of the beam, is a very promising technique.
[12] The typical algorithms chain is: segmentation and simulation of a physical property.
[13] "Rien ne se perd, rien ne se crée, tout se transforme", the law of matter conservation-Lavoisier.

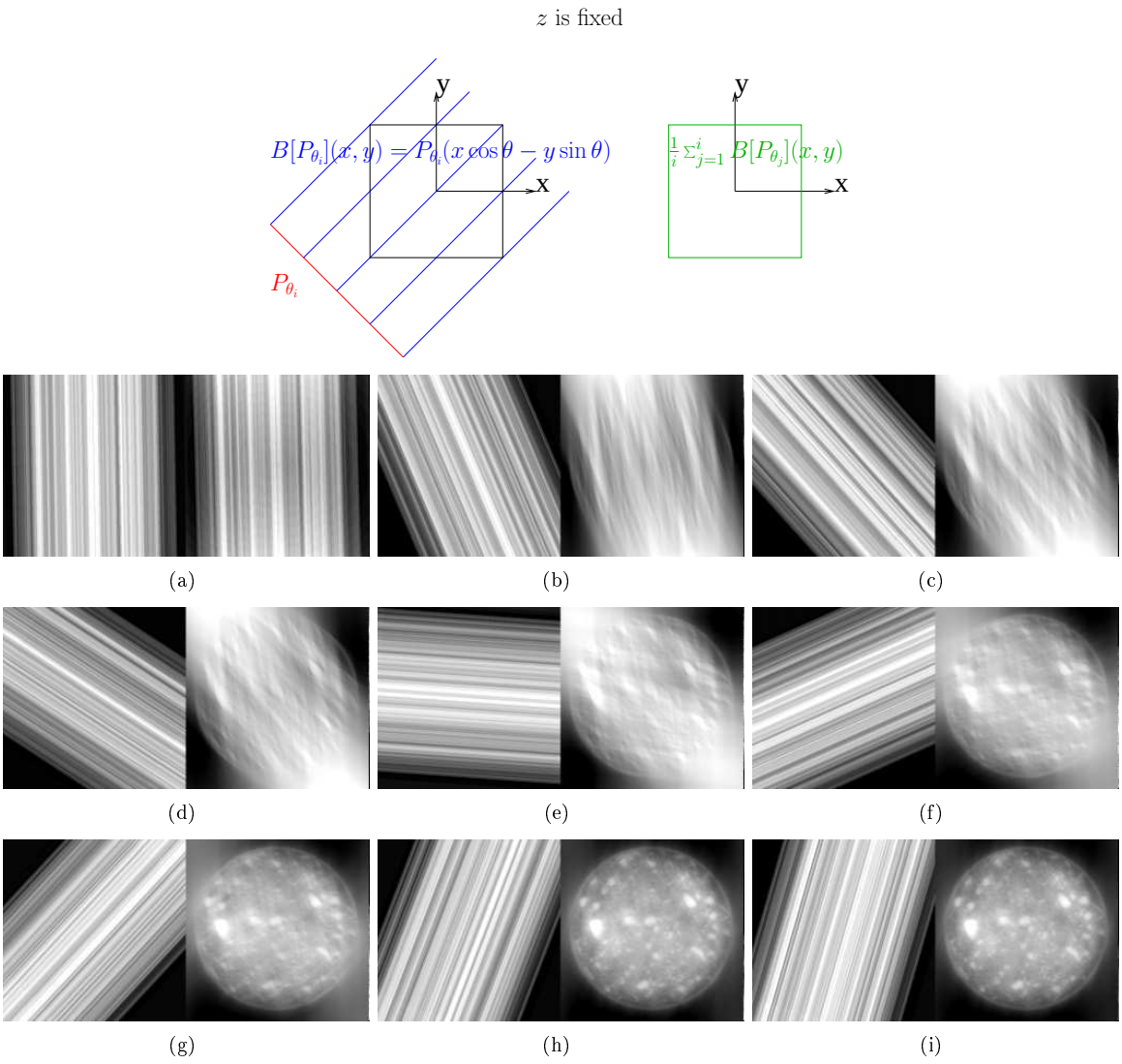


Figure 2.12: Principle of back projection algorithm where $B[p_{\theta_i}]$ is one back projection at the given angle θ_i and $\frac{1}{i}\sum_{j=1}^{i} B[p_{\theta_j}]$ is the cumulative of the back projections form j=1 to i. For each figure couple: first figure $B[p_{\theta_i}]$ and second figure $\frac{1}{i}\sum_{j=1}^{i} B[p_{\theta_j}]$. From (a) to (i): we progressively increases the subscript i. At this end, we observe the anhydrous grains of the cement paste. Courtesy by E. Gallucci.

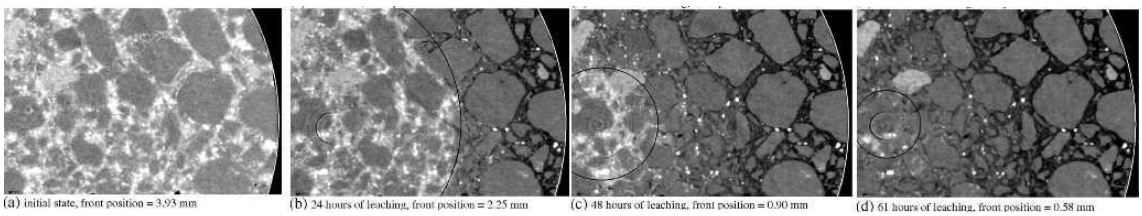


Figure 2.13: Evolution of the front evolution during an accelerated leaching process.

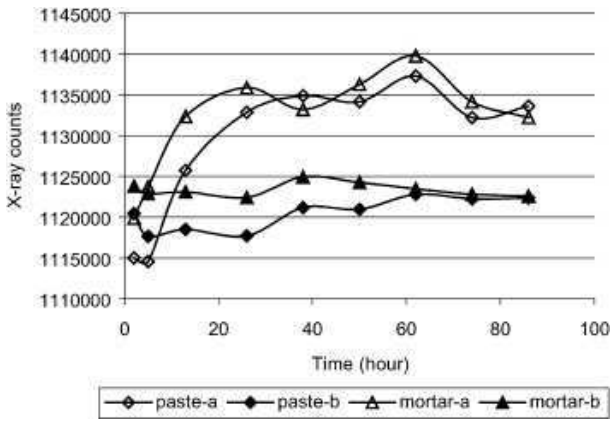


Figure 2.14: X-ray count = X-ray transmission. Mortars reveal a slightly lower evaporation rate since the aggregate increases the length of the transport route because of a larger tortuosity.

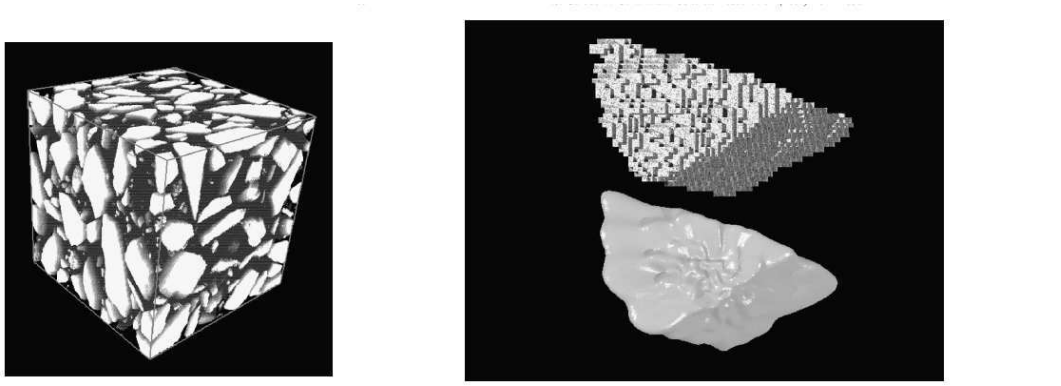


Fig. 6. A 270^3 voxel piece of a tomographic image of a concrete material

Fig. 12. Comparing the digital image taken directly from the original X-ray tomograph (mottled gray) (digital volume = 9175 voxels), with the spherical harmonic expansion reconstruction (shiny gray). A slight surface texture has been added to the tomographic image in order to see the digital detail more clearly.

Figure 2.15: Left figure: aggregate after the segmentation of the image obtained by X-ray tomography. Right figure: Comparison between one aggregate taken directly from the original X-ray tomography (mottled grey) (digital volume = 9175 voxel), and its spherical harmonic expansion reconstruction (shiny grey).

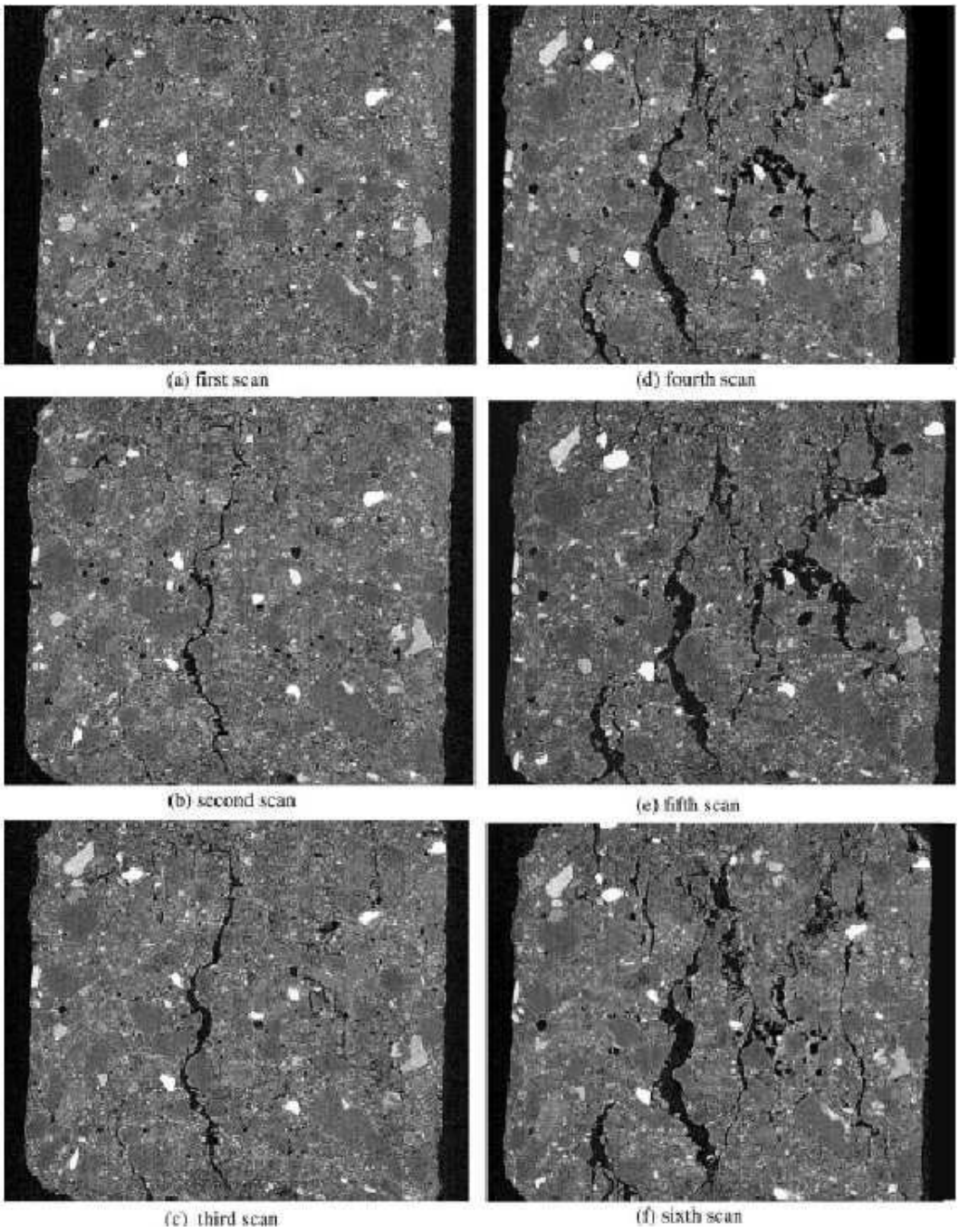


Figure 2.16: Vertical sections through specimen at five different levels of damage: (a) first scan; (b) second scan; (c) third scan; (d) fourth scan; (e) fifth scan; (f) sixth scan.

2.3.3 Preparation

Cement and distilled water were mixed by hand for 5 min and then injected into thin cylindrical plastic tubes with a diameter of 1 mm and a wall of thickness of 30 μm. The plastic cylinder tubes have two advantages. First, the absorption of plastic is low and second the cylinder shape is a suitable geometry[14] for tomography acquisition. After one day, the cylindrical plastic are immersed into water[15] to avoid the surface drying. Two series of experiments[16] were performed at the Swiss Light Source (SLS) in Villigen (Switzerland) on the TOMCAT beam line [162]. For the first series, samples were prepared in advance and kept unsealed into water in order to have the desired age at the same date. For the second series, in order to prevent the variations due to sample preparation, the same capillaries of two different cements were studied over time (see table 2.2 and figure 2.17). Reconstructed slices were computed using the filtered back projection algorithm in use at the SLS.

Cement	Water/cement	time sampled (hours)
series 1		
Alite	0.4	27, 55
Alite	0.5	81
OPC	0.4	24, 61, 95, 150
OPC	0.5	34, 83, 150
OPC+slag	0.4	80, 150
OPC+slag	0.5	35, 48, 150
OPC+limestone	0.4	26.79,150
OPC+limestone	0.5	32.150
series 2		
Alite1	0.4	24.4, 31, 34.5, 38.5
Alite2	0.4	32, 40, 45, 51, 55, 60, 67
OPC	0.4	26.75, 31.33, 34, 37.75, 40, 45.7, 52.3, 55.75, 60, 66, 72

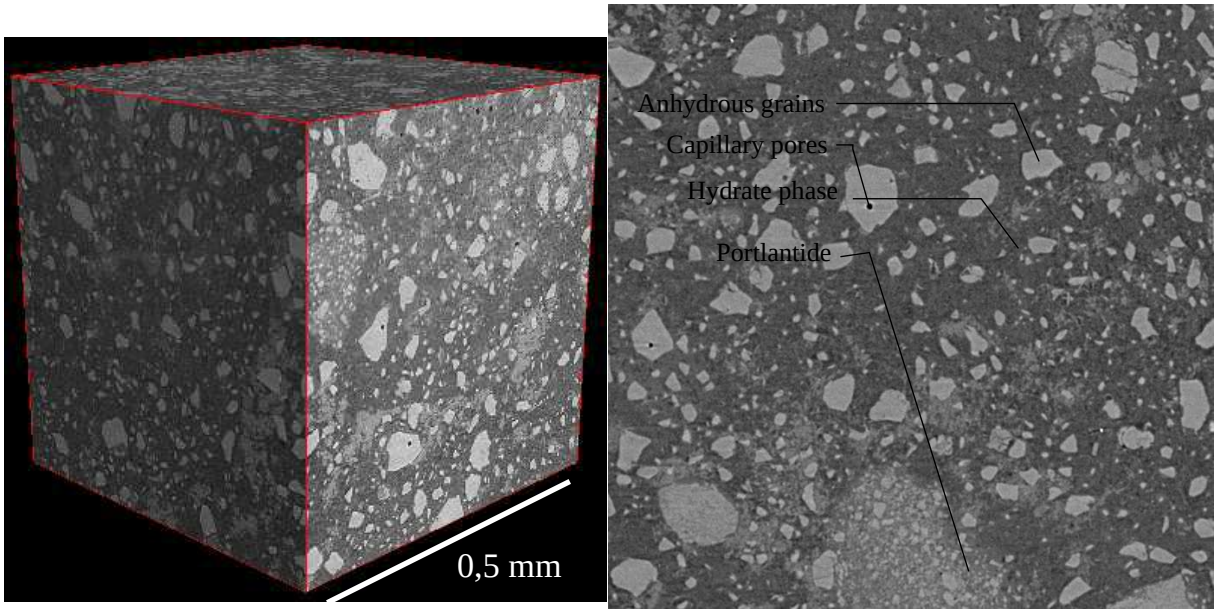
Table 2.2: For the series 2, the same capillaries of different cements were studied over time

The grey level distribution is similar for SEM and for the X-ray tomography:

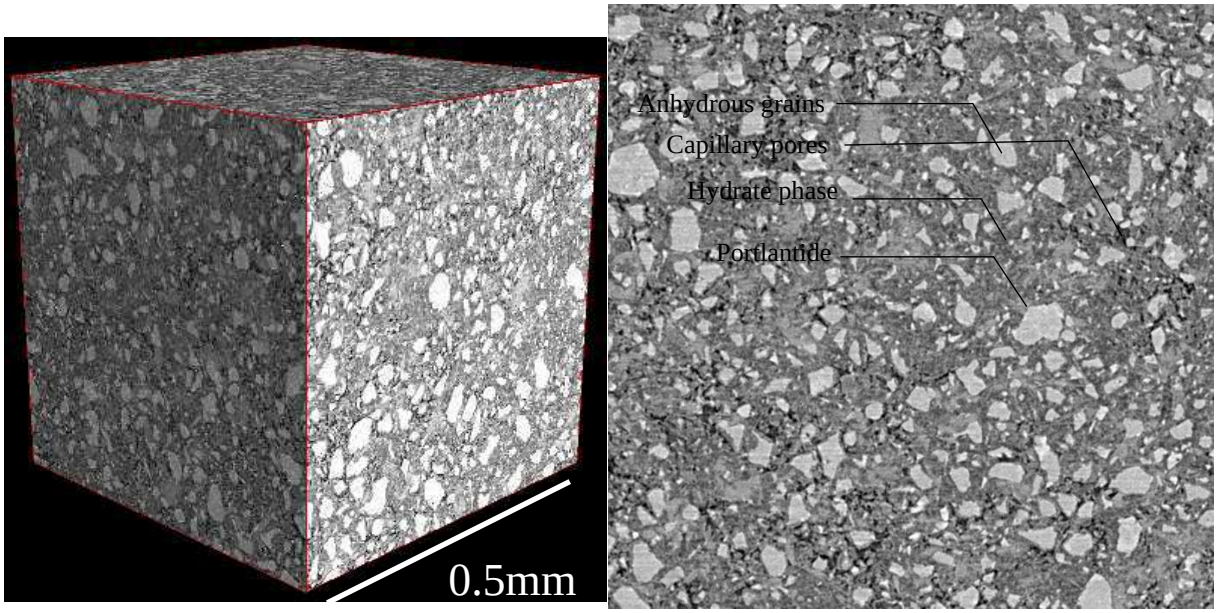
1. the bright grains are remaining unhydrated cores,
2. the light grey clusters are portlandite,
3. the dark grey phase is a mix of hydrated phases,
4. the blackest phase is the porosity

We observe in the alite image that the porosity in tomography image is much smaller than the porosity in SEM image (see figure 2.4-a). The first explanation is that the reaction of hydration/precipitation depends on the temperature. The samples, observed by SEM, were stored at constant temperature of 16^{o}C whereas the samples, observed by tomography, were prepared and immediately transported to the SLS during a warm day. However, this effect cannot only explain the large discrepancy between these two results. Another explanation is the shrink of the porosity during the preparation of the sample for SEM images.

[14] If the axis of rotation coincides with the axe of cylinder, the sample thickness is invariant about the radiograph angle.
[15] We use the mixing water in order to not disturb the reaction of dissolution/precipitation.
[16] In order to enhance the contrast between the porosity and the solid phases, the void has been filled for some samples with a liquid alloy called Wood's metal[155, 178]. However, due to the high absorption of the metal (http://physics.nist.gov/PhysRefData/XrayMassCoef/tab3.html), the exit beam intensity has been too low even for the maximum of energy given that not allows an accurate reconstruction



(a)



(b)

Figure 2.17: (a): alite with W/C=0.4 and t=32h, the size is 700*700*700 voxel with a pixel size of 0.7 μm^3, (b): OPC with W/C=0.4 and t=24h, the size is 700*700*700 voxel with a pixel size of 0.7 μm^3.

2.3.4 First comments on image analysis

The volume of the tomography image is 700*700*700 voxel where each voxel represents a cube of size 0.675 μm^3 of the sample. Several preliminary comments can be made:

1. as the image size obtained by tomography is large, the algorithms associated to the image processing must be efficient. The next chapter will explain how to implement efficient algorithms,
2. as the intensity SLS synchrotron beam line is stable, the results are reproducible[17],
3. as the contrast is weak between the solid phases, the classical threshold procedure is not suitable for the segmentation. For an accurate segmentation, boundary information will be use in the chapter 4,
4. as the resolution is in the same order of the capillary pore size, the capillary pores observation is only possible in the early-ages.

2.4 Conclusion

We have obtained images of cement paste using two imaging techniques: SEM and SRXTM. SEM images are spatially resolved for the investigation of capillary porosity but this is bi-dimensional imaging technique. The chapter 6 will propose an easy method to reconstruct 3D model from 2D slice. The qualitative comparison between SEM and SRXTM for the same sample lead us to ask this question: do you really observe the natural pore network since the preparation of SEM sample is invasive? With SRXTM, we observe the tri-dimensional pore network without any destruction but, due the limited spatial resolution, a relevant information can be obtained only on the early-age of the cement paste. A question remains, can you extrapolate the results obtained by SRXTM in the early-age cement paste for mature cement paste?
Analysis of the 3d geometrical features of a complex media needs, as a prerequisite, an accurate segmentation of the image. Due to the properties of the images (low contrast for tomography and instability of acquisition parameters for SEM), the threshold segmentation will not yield a segmented microstructure consistent with real microstructure. This is why a segmentation procedure, using advanced algorithms, based on boundary information, will be developed in the chapter 4. Since the size of tomography images is large, the advanced algorithms have to be efficient. In order to speed up the implementation of efficient and advanced algorithms, we have conceptualised a class of algorithms, called seeded region growing by pixels aggregation/dissolution. The next chapter will develop this point.

[17] The Filtered Back Projection algorithm in used in SLS requires the selection of a range of attenuation. This range has always been the same to assure the reproductively

3

Seeded region growing by pixels aggregation/dissolution

Contents

3.1 Introduction

3.1.1 Which field of Image Processing?

A grey-level image is a grid of a sample where the value of each cell (pixel/voxel) is the average of a physical interaction (e.g. attenuation of X-rays for tomography or emission of back scattered electrons for SEM) between different solid and fluid phases contained in the cell and the excitation source, at the moment of acquisition. A labeled image is a grid of the sample where the label of each cell corresponds to a phase of the material. Phase partition is the transformation of the grey-level image to the labeled image. The phase partition is segmented in three levels:

- *Pre-filtering:* to increase the signal-to-noise ratio without removing some structural features like the presence of cracks.
- *Segmentation:* to convert the grey-level image to a labelled image.
- *Post-filtering:* to remove some numerical artefacts.

The classical phase partition is the threshold method (see chapter 3). This method is suitable when the imaging device yields enough "contrast-to-noise ratio" for the material in order to have a specific peak for each phase in the histogram. When it is not the case, as tomography images of cement pastes hampered by a low signal-to-noise ratio, an advanced phase partition method has to be used in order to extract the different phase accurately. In this thesis, we focus on the second level of phase partition, the segmentation. The literature of computer science reveals three principle fields for this purpose[18]:

- partial difference equation such as level set [128],
- stochastic processes such as Bayesian statistics [26],
- automate/Seeded Region Growing by pixels Aggregation:Dissolution (SRGPAD)[2, 23].

The algorithms of SRGPAD are efficient since their complexity is $O(n)$ where n is the number of pixels/voxel of the image. Therefore, this class of algorithms allows the treatment of the large tomography size which typical size is one giga voxel in a reasonable time. In this thesis, for the purposes of segmentation, decomposition in elementary pores, 3D reconstruction of multiphase porous media obtained by SEM, geometrical characterisation, a wide range of algorithms of SRGPAD is used: distance function, dynamic filter, global minima, geodesic dilation, watershed transformation, permutation localisation on the phase interface, geometrical tortuosity, hole filling, ...
There are many commercial (Aphelion®,Morph-M) or non commercial (ImageJ) libraries available implementing SRGPAD algorithms. At the beginning of this thesis in 2005, these libraries were implemented for 32 bits processor architecture that limited addressing to Random Access Memory (RAM) to 3.2 Gb. Since advanced algorithms requires dynamic structures having a high memory cost and the typical size of a tomography image is 1 Gb, the demand of random access memory exceeds largely 3.2 Gb. This point led us to the creation of a home-made library implementing these algorithms for 64 bits processor architecture.

3.1.2 Why conceptualising?

A direct implementation of optimised advanced algorithms is relatively long and tedious, and requires special skills. Therefore, in first order, the development of a library containing many algorithms of SRGPAD is a too long work for a PhD student. In order to speed-up the implementation of algorithms of SRGPA, a theoretical work is done to extract the generic concepts in this class of algorithm. Then, we translate these generic al concepts in term of generic objects in object-oriented programming with an optimized implementation. Finally, these generic objects yield implementation of optimised algorithms of SRGPAD with less than 40 lines of codes and, thus, allow the implementation of the library containing many algorithms of SRGPAD in a reasonable time.
In this chapter, we introduce a theoretical framework for a class of algorithms, called Seeded Region Growing by Pixels Aggregation/Dissolution (SRGPAD). In the first section, we present an example of SRGPAD algorithm in order to exhibit the key ideas behind this conceptualisation. In the second section, the theoretical concepts of SRGPAD are introduced. In the third section, we use these concepts to implement a wide range of algorithms. In the last section, we deal with the different conventions to

[18]These three frameworks are also used to analyse pattern formation. The science of pattern formation deals with the visible, (statistically) ordered outcomes of self-organisation and the common principles behind similar patterns. In physics, the analysis of pattern-forming systems often consists of finding a model in using a
- a Partial Difference Equation (PDE). An approximate solutions of PDE can be estimated using finite difference method, finite element method [73], level set [128], phase field [30, 83].
- a Cellular Automata for the study of phases transitions[32]
- a stochastic process for the study of the evolution sea-coasts through damped erosion [147]

manage the collision between two regions.
In this chapter, each image is a subset of the space E. E is a n-dimensional discrete space $\mathbb{Z}^n$, consisting of lattice points with integer coordinates. The elements of a n-dimensional image array are called pixels if n=2 and voxel if n=3. A point will be called pixel although the algorithms can be developed whatever the space dimension.

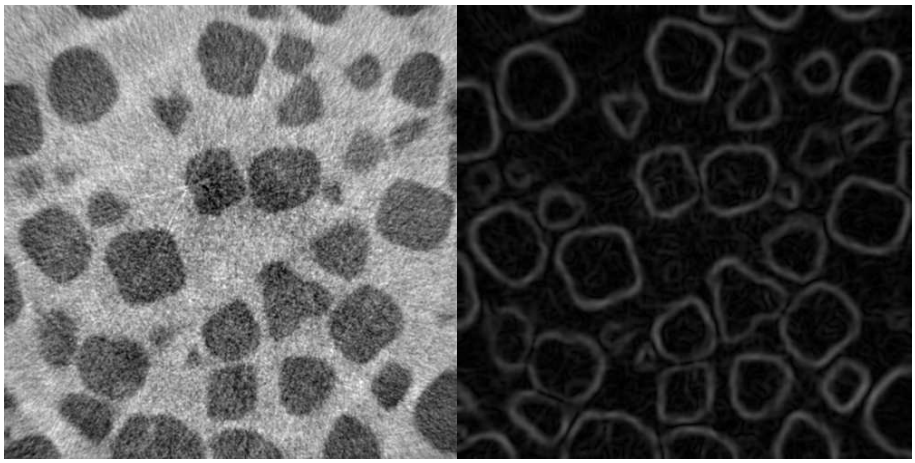


Figure 3.1: Left figure: one object feature is the texture and one texture feature is the tint. Right figure: the object boundary discontinuity becomes the lines of ridge in the gradient image.

3.2 Example of SRGPAD algorithm: the watershed transformation

Many fields in computer science, stereovision [82], mathematical morphology [158], use algorithms which principle is Seeded Region Growing by Pixels Aggregation/Dissolution (SRGPAD). This method consists in initialising each region with a seed, then iterating pixels aggregation/dissolution on regions until getting a convergence [2, 69]. The general purpose of this field is to define a metric divided into two distinct categories [10]: the object feature like the tint [2, 133] and the object boundary discontinuity [23, 49, 164, 108] (see figure 3.1). For instance, the watershed watershed developed in mathematical morphology [23] is a segmentation tool and an algorithm of SRGPAD using the boundary information.

3.2.1 Presentation of the watershed transformation

Any grey-level image can be considered as a topographic surface and all boundaries as sharp variations of the grey level. When a gradient is applied to an image, boundaries are enhanced (see the second image in the figure 3.1). When the topographic surface obtained from the gradient is flooded from its seeds, the waterfronts meet on watershed lines in 2D, and on watershed surfaces in 3D. The investigated volume is partitioned by the watershed basins (see figure 3.2). Efficient implementations of the watershed transformation based on immersion simulation, were proposed by many authors [24, 173] but each implementation remains specific to the algorithm. In this section, we introduce some generic objects to implement rapidly this algorithm. But, these generic objects will also allow the implementation of any SRGPAD algorithm.

3.2.2 Selection of a pair

A SRGPAD algorithm is not global but local. The aggregation/dissolution is pixel by pixel. A time is defined in order to follow the evolution of the basins. At each aggregation or dissolution of a pixel, x, on a basin, i, the time, t, is incremented: $t \leftarrow t+1$. With this notation for the set of basins: $(X_0^t, \ldots, X_i^t, \ldots, X_n^t)$ where X_i^t is the localisation of the basin[19] with the label i at time t, we have:

$$\forall j \in (0, \ldots, n) : X_j^{t+1} = \begin{cases} X_j^t \pm \{x\} & \text{if } i = j \\ X_j^t & \text{otherwise} \end{cases}$$

[19] $X_i^t \subset E$

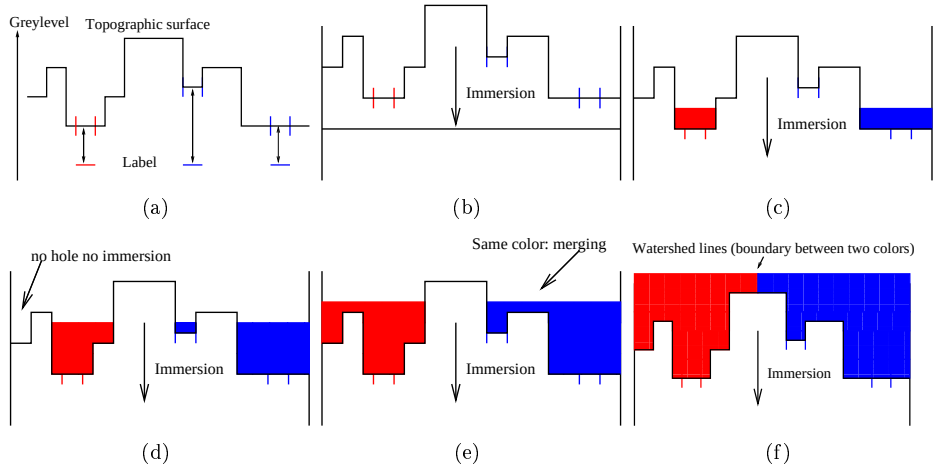


Figure 3.2: Watershed transformation requires two images: the topographic surface (a grey-level image) and the seeds image. The process is: (a) association of each seed to a hole (b) immersion (c) the water enters in the topographic surface by the holes and the basins take the colour of the hole, (d) a part of the topographic surface is not merged although its level is under the level of the immersion, (e) fusion of two basins of same colour, (f) creation of a dam when two basins have different colours: *A video is available at http://pmc.polytechnique.fr/~vta/water.mpeg.*

where the addition, $+$, between two sets A and B, is the disjoint union:

$$A + B = \{A \cup B : A \cap B = \emptyset\}$$

and the subtraction, $-$, between two sets A and B, is the inclusion restriction:

$$A - B = \{A \setminus B : B \subset A\}$$

Therefore, at each step of time, the algorithm has to select a pair (x, i) where x is the pixel and i is the label. This selection is done following three rules:

1. The value given by a metric[20], $\delta(x, i)$, respects a condition, ψ. For the watershed transformation, when the level of immersion is equal to l, the basins can grow on the pixels, x, where the level of the topographic surface, $f(x)$, is under this level of immersion. Therefore, the metric is associated to the topographic surface, $\delta(x, i) = f(x)$, and the condition [21] is: $\delta(x, i) \leq l$.

2. The pixel x belongs to the "basin neighbourhood' (homogeneous aggregation). To localize where the aggregation can take place, for each basin, a Zone of Influence (ZI), Z_i^t, is defined. For the watershed transformation, the ZI of each basin is the outer boundary of the basin in excluding the other basins. The second condition is: $x \in Z_i^t$

3. Several pairs can respect these two previous conditions. In order to select only one of these pairs, we add a condition of the entering time, $\mathcal{T}$. For the the watershed transformation, we select the pair that respects for the first time the two previous conditions and still respects their until now.

3.2.3 Algorithmic implementation of the watershed transformation

In the SRGPAD algorithm, a basin is called a region. The algorithm of watershed transformation is (see figure 3.3):

[20] called the ordering attribute function
[21] $\psi(\delta(x, i)) = 1$ for $\delta(x, i) \leq l, 0$ otherwise

1. Input: topographic surface, f, and seeds, $(s_i)_{0 \leq i \leq n}$.
2. Sequential initialisation of the regions with the seeds, (localisation of the regions on the seeds $X_i^{t=0} = s_i$),
3. //**Growing process :**
4. For the $l = 0$ to the maximum level of the topographic surface,
   (a) While there is a pair (x, i) such that $\delta(x, i) \leq l$ and $x \in Z_i^t$
      - Selection of the pair (x, i) that respects for the first time the two conditions ($\delta(x, i) \leq l$ and $x \in Z_i^t$) and still respects their until now.
      - Aggregation of the pixel, x, on the region, X_i^t ($X_i^{t+1} = X_i^t + \{x\}$)

   (b) End while
5. Return the regions

This is an algorithmic implementation of the watershed transformation. In the next subsection, the implementation is presented in using object-oriented programming allowing the implementation using modern languages as C++ or java.

3.2.4 Implementation of the watershed transformation using generic classes

In the previous implementation, we have seen that:

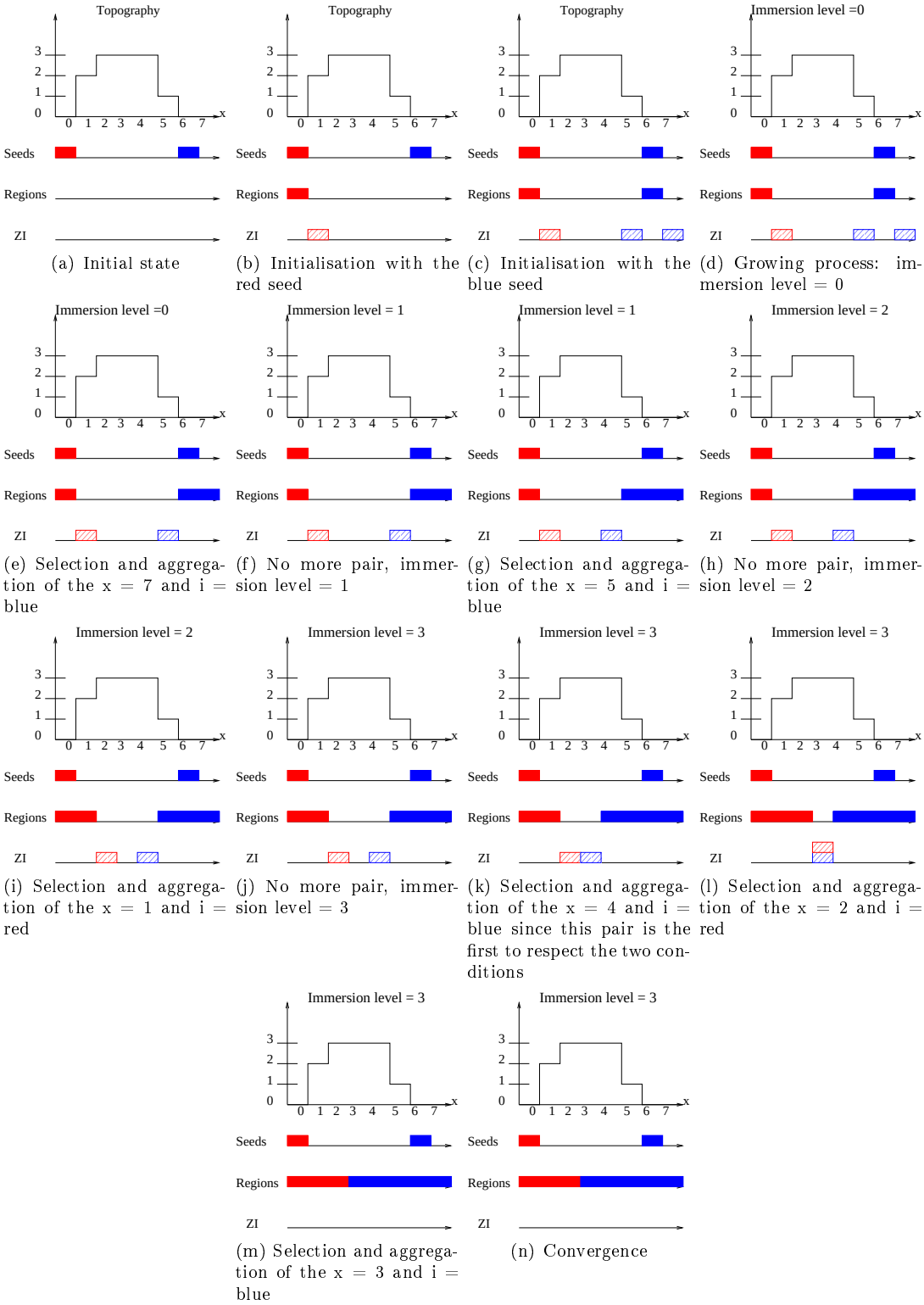
- the regions localise the basins,
- the Zones of Influences (ZI), associated at each region, localise the "region neighbourhood' where the aggregation can take place given the second condition of the pair selection,
- the metric (the topographic surface for the watershed transformation) given the first condition of pair selection,
- the entering time allowing the selection of the pair that respects for the first time the two previous conditions and still respects their until now.

For each of this point, we will define a class.

Class: regions

The first class is the set of regions, $(X_0^t, \ldots, X_i^t, \ldots, X_n^t)$. This class has four accessible methods for the implementation of the SRGPAD algorithm:

1. int add_region(), addition of a region to the set of regions in returning the label of the added region.
2. void initial_seed(set s, int i), initialisation of the region i with the seed s.
3. void growth(pixel x, int i), aggregation of the pixel x on the region i: $X_i^{t+1} = X_i^t + \{x\}$ (see figure 3.4)
4. void degrowth(pixel x, int i), dissolution of the pixel x on the region i: $X_i^{t+1} = X_i^t - \{x\}$



(a) Initial state

(b) Initialisation with the red seed

(c) Initialisation with the blue seed

(d) Growing process: immersion level = 0

(e) Selection and aggregation of the x = 7 and i = blue

(f) No more pair, immersion level = 1

(g) Selection and aggregation of the x = 5 and i = blue

(h) No more pair, immersion level = 2

(i) Selection and aggregation of the x = 1 and i = red

(j) No more pair, immersion level = 3

(k) Selection and aggregation of the x = 4 and i = blue since this pair is the first to respect the two conditions

(l) Selection and aggregation of the x = 2 and i = red

(m) Selection and aggregation of the x = 3 and i = blue

(n) Convergence

Figure 3.3: The watershed transformation with non-periodic boundary condition.

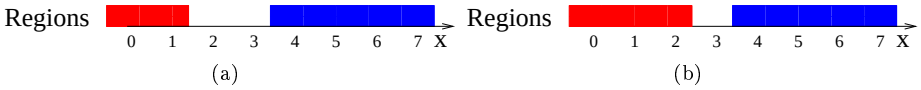


Figure 3.4: Aggregation of pixel x=2 on the red region. (a) before, (b) after

Class: ZI

The second class is the set of the zone of influences, $(ZI_0^t, \ldots, ZI_i^t, \ldots, ZI_n^t)$. A zone of influence localises where the growing process can take place for the associated region. This class has one accessible method for the implementation of the SRGPAD algorithm [22]:

1. void add_ZI(int i, "some parameters"), addition of the a ZI to the region, i. "some parameters" precise the localisation of the zone of influence (see next section). For the watershed transformation, the ZI is the outer boundary of this basin in excluding the other basins.

Class: metric

The third class defines the metric, $\delta(x, i)$. This class has a specefic constructor with a functor as parameter in order to define the metric[23]. For the watershed transformation, the functor is $max(f(.), l)$ where f is the topographic surface and l the level of immersion.

Class: system of queues

To manage the pixel by pixel organisation, a fourth class, called System of Queues (SQ), stores all pairs (x, i) (pixel, x belonging to ZI i) in a data structure. It gives the possibility to select a pair (x, i) depending on three conditions:

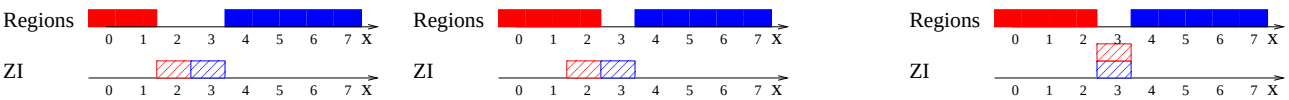
- the pixel x belongs to the ZI i ($x \in Z_i^t$).
- the metric, δ respect a condition ($\psi(\delta(x, i)) = 1$)
- the entering time.

This method has one constructor with two parameters. The first parameter defined the kind of queue and the second one the number of queue. For the the watershed transformation, the condition of the entering time is to select the pair that respects for the first time the two previous conditions and still respects their until now. The convenience data structure is composed by n queues, $Q_n^t = \{q_0^t, \ldots, q_n^t\}$, because the strategy is:

- to store each pair (x, i) whose metric is the same in the same queue (a queue is associated to a quantification of the metric and n is the metric range),

[22] Since the ZI depends on the regions, after each growth or degrowth of a region, the ZI have to be actualised. Therefore, in the implementation of the methods, growth(pixel x, int i), or degrowth(pixel x, int i) of the class Region, one of these method is called (see figure below):

1. void actualise_growth(pixel x, int i), actualise the zones of influence depending on the aggregation of pixel x on the region i.
2. void actualise_degrowth(pixel x, int i), actualise the zones of influence depending on the dissolution of pixel x on the region i. This method is called by the method degrowth(pixel x, int i) of the class Region.



Initial state at time t | Aggregation of the pixel x=2 on the red region | Actualisation of the ZI

[23] In the computer, for each different metric, a sub-class is implemented where the polymorphic method, int metric_value(pixel x, int i), returning $\delta(x, i)$ is defined.

- to organise in First In First Out (FIFO) "what comes in first is handled first, what comes in next waits until the first is finished".

This class has two accessible methods for the implementation of the SRGPAD:

1. Boolean empty(metric m) return false if there exists a pair (x, i) such that $x \in Z_i^t$ and $m = \delta(x, i)$, otherwise true.
2. pair=(x,i) pop_pair(metric m), return pair (x, i) such that $x \in Z_i^t$, $m = \delta(x, i)$ and a condition of the entering time.

[24]

Class: Population

Due to the interactions between the objects of these different classes, a master class, call Population, is defined in order to link their. This class has only a constructor defined by: Population(Region & r,ZI & zi, Metric & m, SQ & sq).

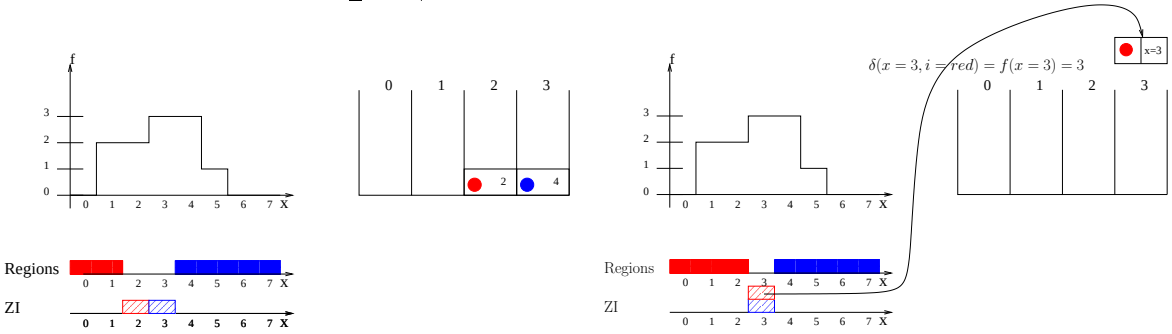
Implementation of the watershed transformation

Using these generic classes, the watershed transformation implementation is simple (see figure 3.5 and algorithm 1). The major difficulty is the implementation and the optimisation of these generic classes. But, if this task is achieved, the implementation of any SRGPAD algorithm will requires less than twenty lines of codes and will already give an optimised algorithm.

Implementation of the generic classes

In footnote of the presentation of these generic classes, I describe which methods to implement in order to allow the interaction between the different classes. This interaction is: after the growth of a region on a pixel, first, we have the actualization of the zones of influence. Then, during this actualisation, at each growth of the zone of influence, j, on the pixel y, then the pair (y, i) is pushed in the system of queue and the index of the queue is given by the metric. Numerically, the implementation of the classes, Region and Metric, is easy. The class region has an instance, label space, given the label[25] of the region for any pixel. The next section will be focused on the description of the two others classes allowing their implementation.

[24] At each addition of a pixel, x to a ZI, i, the pair (x, i) is stored in system of queue. Therefore, this class has one method, void push_pair(pixel x, int i), to store the pair (x, i) in the data structure. The method metric_value(pixel x, int i) of the class metric gives the index of the queue.



Initial state

store the pair $(x = 2, i = red)$

[25] if the label is equal to -1, there is no region

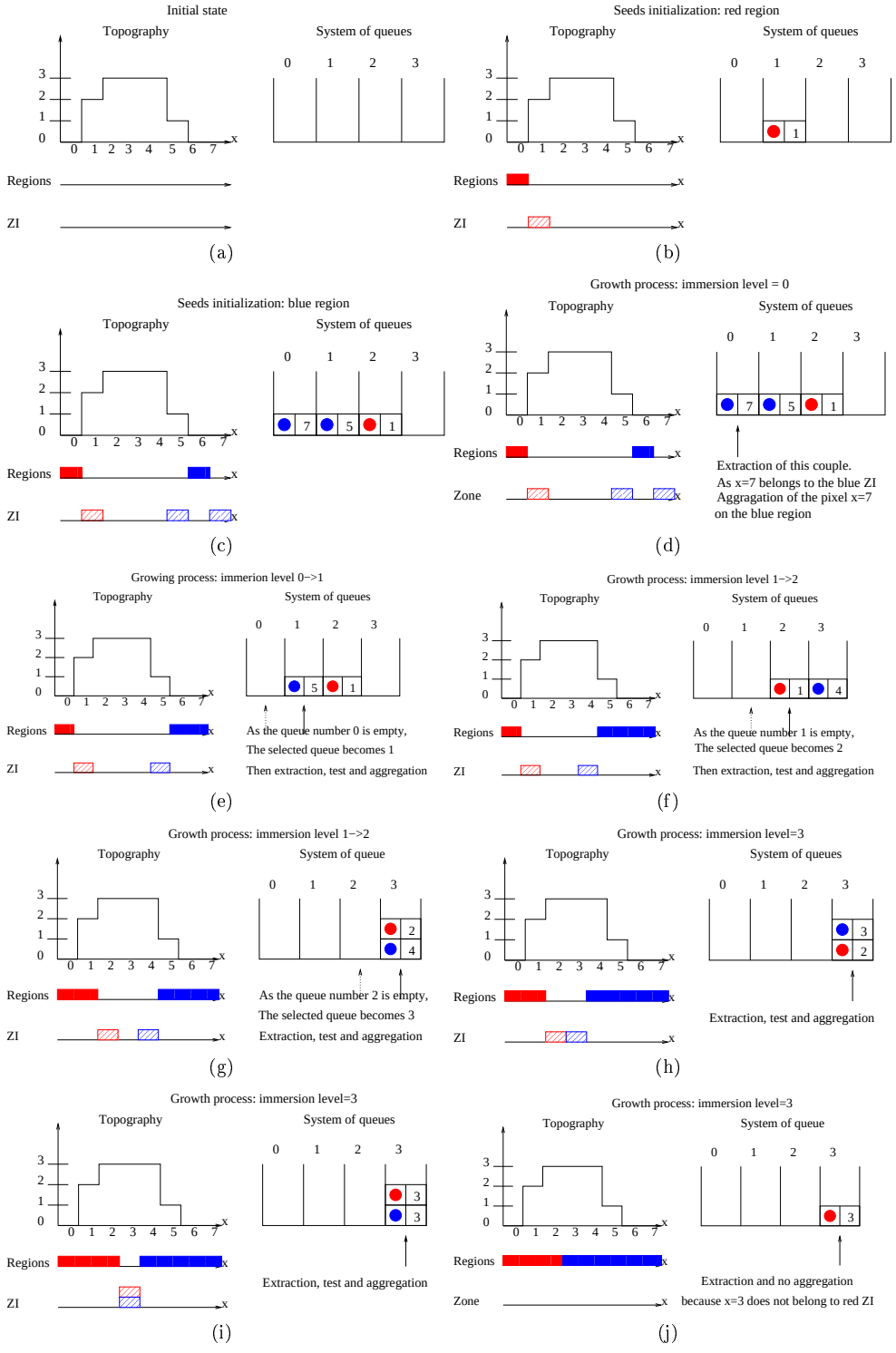


Figure 3.5: The watershed transformation with non-periodic boundary condition. initialisation two seeds: x=0 for the red region and x=6 for the blue region; (d,e,f,g,h,i,j): growth process; (d) the growth process begins on the pixels with the lowest elevation: the queue number 0 with the aggregation of the pixel x=7 on the blue region, (e): 1) after this aggregation, the ZI have be actualised, 2) as the queue number 0 is empty, the queue number 1 is selected, and so one, (j) there is no aggregation of the pixel x=3 on the red region because this pixel does not belong to the red ZI.

Algorithm 1 the watershed transformation

Require: f, $(s_i)_{0 \leq i \leq n}$ *//the topographic surface, the seeds*
  //**initialisation**
  int immersion_level =0;
  Region r;
  ZI zi;
  Metric m (max(f,immersion_level); *//*$\delta(x,i) = max(f(x)$,*immersion_level)*
  SQ sq(FIFO, max level of the topography);*//n FIFO queue such as the pixels with the corresponding label of the ZI are stored depending on the elevation on the topographic surface*
  Population pop (r,zi,m,sq);
  for all for j = 0 to n **do**
    int region_label = r.add_region();
    zi.add_ZI(region_label, "some parameters") ;
    r.inti_seed(region_label, s_i);
  end for
  //**Growth process:**
  for all immersion_level = 0 to the max level of the topography **do**
    while sq.empty(immersion_level)==false **do**
      (x,i)=sq.pop_pair(immersion_level);
      r.growth(x, i);
    end while
  end for
20: **return** r;

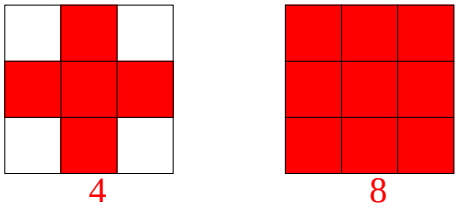


Figure 3.6: Neighbourhood in 2-dimensional discrete space

3.3 The framework of SRGPAD

In this section, we will introduce a mathematical definition of the Zone of Influence (ZI). Then, we will describe an efficient procedure to actualise the ZI after a aggregation/dissolution of a pixel on a region. After, in order to manage the pixel by pixel organisation, we will introduce a System of queues (SQ) that stores all pairs (x, i) (x belongs to the ZI i) in a data structure and allows the selection of a pair following three conditions. Finally, we will compare the efficient of the watershed transformation implemented in this framework and the implementation in the library Aphelion. s. To localize efficiently these fluctuations, a procedure is defined such that some operations have to be computed onvincent@vincent-laptop: /Desktop$vincent@vincent-laptop:\ /Desktop$ly on the aggregation/dissolution neighbourhood.

3.3.1 ZI

An object, called Zone of Influence (ZI), is associated at each region to localise a zone where the propagation can take place. in this subsection, we will introduce its mathematical definition progressively.

Minkowski addition

The Minkowski addition, also known as dilation, of two subsets X_i^t and V of a linear space is the result of the addition of every element of X_i^t to every element of V, i.e. the set (see figure 3.7-b):

$$X_i^t \oplus V = \{x + v : x \in X_i^t, v \in V\}$$

In 2-dimensional discrete space, the set V, called the neighbourhood, is usually defined as (see figure 3.6):

1. 4-neighbourhood $V = \{(0,0),(1,0),(-1,0),(0,1),(0,-1)\}$,
2. 8-neighbourhood $V = \{(0,0),(1,0),(-1,0),(0,1),(0,-1),(1,1),(1,-1),(-1,1),(-1,-1)\}$,

Boundary of a set

The boundary of a set, X_i^t, is the Minkowski addition of this set with the neighbourhood set excluding itself (see figure 3.7-c):

$$(X_i^t \oplus V) \setminus X_i^t = \{x + v : x \in X_i^t, v \in V, x + v \notin X_i^t\}$$

Zone of influence

For the watershed transformation, the ZI of a set, X_i^t, is the Minkowski addition of this set with the neighbourhood set excluding all the regions (see figure 3.7-d):

$$(X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j^t) = \{x + v : x \in X_i^t, v \in V, (\forall j \in \mathbb{N} : x + v \notin X_j^t)\}$$

In general, for a given set $N_i \subset \mathbb{N}$, called restricted set, and a neighbourhood set[26], V_i, we define the ZI

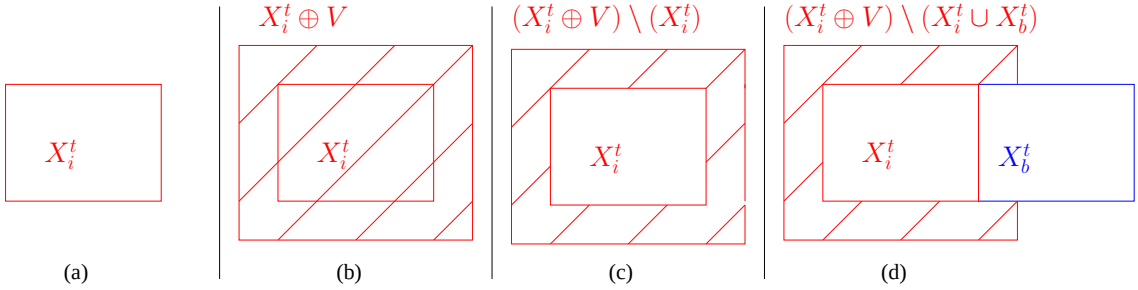


Figure 3.7: Construction of the ZI, Z_i^t. (a): A region X_i^t; (b): Minkowski addition of the region, X_i^t, by the neighbourhood set, V ; (c): Minkowski addition of the dilatation of X_i^t by the neighbourhood set, V excluding X_i^t; (d): Minkowski addition of the dilatation of X_i^t by the neighbourhood set, V excluding the union of X_i and of X_b.

associated to a region$i$ as:

$$Z_i^t \quad = \quad (X_i^t \oplus V_i) \setminus (\bigcup_{j \in N_i} X_j^t)$$

The restricted set is usually defined as:

- $N_i = \{i\}$, then Z_i^t is the outer boundary of the region;
- $N_i = \mathbb{N}$, then Z_i^t is the outer boundary of the region excluding all the regions.

In the method add_ZI(int i, "some parameters") of the class ZI, "some parameters" are the restricted set and the neighbourhood set (see appendix A.1 for the numerical implementation of these two sets).

3.3.2 Actualisation

As the ZI depend on regions, after each aggregation/dissolution of a pixel on a region, the ZI have to be actualised. For the actualisation, the simplest way is to check for each ZI the possible modification. This strategy has a high numerical cost with a complexity equal to $O(n)$ where n is the number of ZI/regions. Therefore, the computational cost will be too high. A solution is to reformulate the link between ZI and the regions: the ZI at time t+1 depends on ZI at time t plus or minus a set defined in the neighbourhood of the aggregation/dissolution. This reformulation gives an actualisation with a constant complexity. This subsection is quite technical and it is not necessary for the understanding of the whole section (it can be skipped at first reading).

Decomposition in two parts

At each step of time, one and only one basin, i, grows by aggregation of the pixel x or by dissolution of the pixel x:

$$\forall j \in (0, \ldots, n) : X_j^{t+1} = \begin{cases} X_j^t \pm \{x\} & \text{for } j = i \\ X_j^t & \text{otherwise} \end{cases}$$

We see two parts in the expression of ZI: $Z_i^t = \overbrace{(X_{i,m}^t \oplus V_i)}^{\text{ownself}} \setminus (\overbrace{\bigcup_{j \in N_i} X_{j,o}^t}^{\text{other}})$. The "ownself part" depends only on the associated region. The "other part" depends on all regions. Therefore, we decompose the actualisation in two stages. First, we actualise the "ownself part": $X_{i,m}^{t+1} = X_{i,m}^t \pm \{x\}$ then the "other part": $X_{i,o}^{t+1} = X_{i,o}^t \pm \{x\}$ (see figure 3.8). All the proofs are in the appendix A.2.

[26] this set is equal to V or $\emptyset$

Ownself part In this part, we have only the aggregation or dissolution of the pixel x on $X_{i,m}^t$ between times t and $t+1$.
If j is not equal to i ($j \neq i$), Z_j^{t+1} is equal to Z_j^t because Z_j^t does not depend on $X_{i,m}^t$.
If j is equal to i:

- for aggregation $X_{i,m}^{t+1} = X_{i,m}^t + \{x\}$, we have:

$$Z_i^{t+1} = Z_i^t + (\{x\} \oplus V_i) \setminus Z_i^t \setminus (\bigcup_{j \in N_i} X_j^t)$$

The numerical implementation is:

for all $\forall y \in \{x\} \oplus V_i$ **do**
  if $y \notin Z_i^t$ **then**
    if $y \notin \bigcup_{j \in N_i} X_j^t$ **then**
      $Z_i^{t+1} = Z_i^t + x$;
    end if
  end if
end for

- for dissolution $X_{i,m}^{t+1} = X_{i,m}^t - \{x\}$, we have:

$$Z_i^{t+1} = Z_i^t - ((\{x\} \oplus V_i) \setminus (X_i^{t+1} \oplus V_i)) \setminus (Z_i^t)^c$$

The numerical implementation is:

for all $\forall x \in \{x\} \oplus V_i$ **do**
  if $y \in Z_i^t$ **then**
    if $(y \oplus V_i^{-1}) \cap X_i^{t+1} = \emptyset$ **then**
      $Z_i^{t+1} = Z_i^t - y$;
    end if
  end if
end for

Other part In this part, we have only the aggregation or dissolution of the pixel x on $X_{i,o}^t$ between times t and $t+1$.

- for aggregation $X_{i,o}^{t+1} = X_{i,o}^t + \{x\}$, we have:

$$\forall j \in (0, \ldots, n) : Z_j^{t+1} = \begin{cases} Z_j^t - (\{x\} \setminus (Z_j^t)^c) \text{ if } (i \in N_j) \wedge (V_j \neq \emptyset) \\ Z_j^t \text{ else} \end{cases}$$

where the symbol $\wedge$ means the **and** in the symbolic logic. The numerical implementation is:

for all $j \in (0, \ldots, n) : x \in Z_j^t$ **do**
  if $i \in N_j \wedge V_j \neq \emptyset$ **then**
    $Z_j^{t+1} = Z_j^t - x$;
  end if
end for

- for dissolution $X_{i,o}^{t+1} = X_{i,o}^t - \{x\}$, we have:

$$\forall j \in (0, \ldots, n) : Z_j^{t+1} = \begin{cases} Z_j^t + ((((\{x\} \setminus (X_j^{t+1} \oplus V)^c) \setminus (\bigcup_{k \in N_j} X_k^t)) \setminus Z_j^t) \text{ if } (i \in N_j) \wedge (V_j \neq \emptyset) \\ Z_j^t \text{ else} \end{cases}$$

This numerical implementation is based on the assumption that for all j, V_j is equal to $\emptyset$ or V:

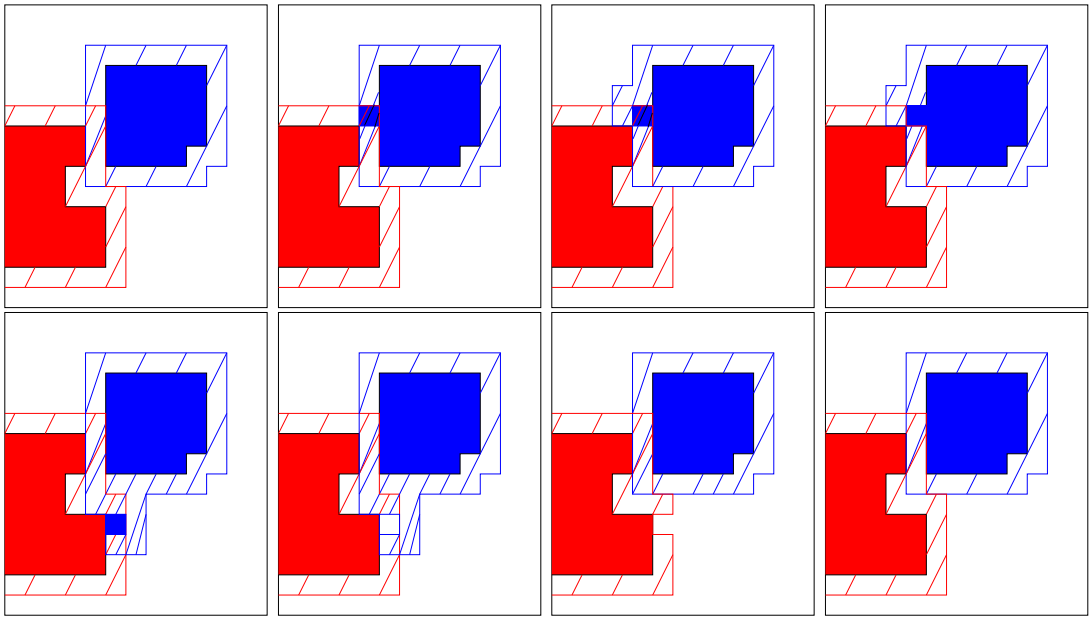


Figure 3.8: Actualisation of the ZI after a aggregation/dissolution of a pixel on a region. Two regions, X_r^t and X_b^t, and two ZI, $Z_r^t = (X_r^t \oplus V_r) \setminus (X_r^t \cup X_b^t)$ and $Z_b^t = (X_b^t \oplus V_b) \setminus (X_r^t \cup X_b^t)$. Both rows show the ZI actualisation after a aggregation for the first row and dissolution for the second row. The first figure is the initial state, the second figure is the aggregation or dissolution of a pixel on the blue region, the third figure is the actualisation of the "ownself part" of ZI, the last figure is the actualisation of the "other part" of ZI.

for all $j : (x \oplus V^{-1}) \cap X_j^t \neq \emptyset$ **do**
  if $i \notin \bigcup_{j \in N_i} X_j^{t+1}$ **then**
    if $i \notin Z_j^t$ **then**
      if $i \in N_j \wedge V_j \neq \emptyset$ **then**
        $Z_j^{t+1} = Z_j^t + x$;
      end if
    end if
  end if
end for

The complexity of this actualisation is time constant $O(1)$ if two data structures are implemented given the label of region and the list of zones of influence located on any pixel of the space. This mathematical definition of the actualisation gives the implementation of the methods of the class ZI directly. The next subsection will describe the class system of queues.

3.3.3 Organisation

To manage the pixel by pixel organisation, the class, System of Queues (SQ), stores all pairs (x, i), where the pixel x belongs to the zone of influence i, and allows the selection of a pair (x, i) following three conditions:

- x belongs to Z_i^t,
- the value of the metric [10], δ, is equal to a certain value (for watershed transformation, a pixel is chosen depending on its elevation on the topographic surface),

- a condition of the entering time (for watershed transformation, we choose the oldest pair respecting these previous two conditions),

At each aggregation/dissolution, we have some pixel added and removed to the ZI:

- The set B^t is the pairs to add to the SQ between times t and $t+1$:

$$B^t = \{(x,i) : x \notin Z_i^t \wedge x \in Z_i^{t+1}\}$$

- The set C^t is the pairs to remove to the SQ between times t and $t+1$:

$$C^t = \{(x,i) : x \in Z_i^t \wedge x \notin Z_i^{t+1}\}$$

The SQ is like a store. There are:

- the products, B^t, to place to the store,
- the products, C^t, to remove from the store,
- a customer that chooses product by product in the store following a label and a condition on the entering time.

To get the best efficiency in the store organisation, the three following points must be respected:

Data structure

A data structure is a way to store data in a computer. A data structure depends on the operations to be performed, using as few resources in execution time and in memory space, as possible. In this thesis, we have used two data structures:

- n queues, $Q_n^t = \{q_0^t, \ldots, q_n^t\}$, because the strategy is:
  - to store each pair (x,i) whose metric is the same in the same queue (a queue is associated to a quantification of the metric and n is the metric range),
  - to organise in First In First Out (FIFO) "what comes in first is handled first, what comes in next waits until the first is finished".
- n vectors, $V_n^t = \{v_0^t, \ldots, v_n^t\}$, because the strategy is:
  - to store each pair (x,i) whose metric is the same in the same vector (a vector is associated to a quantification of the metric and n is the metric range),
  - to organise in random access to handle randomly a pair[27].

We have implemented these two data structure: queue and random access using the Standard Template Library (STL) of the C++ library.

Store

The ordering attribute function, δ, assigns each pair (x,i) of B^t in the appropriate queue. After addition of a pixel, x, is to a zone of influence i, the method push_pair(x, i) of the class SQ is called. This method uses the method metric_value(pixel x, int i) of the class Metric to assign the pair (x,i) in the appropriate queue (see figure 3.9). If the method returns OUT then the pair is not assigned to the system of queues.

[27] A vector is a dynamic array $v = ((x_0,i_0),\ldots(x_p,i_p))$. When the method pop is called, we drawn randomly an integer, I , in $(0,\ldots,p)$. Therefore, we return the pair (x_I,i_I) and we permute the last pair of the array with the pair (x_I,i_I) and we decrease the size of the array by one.

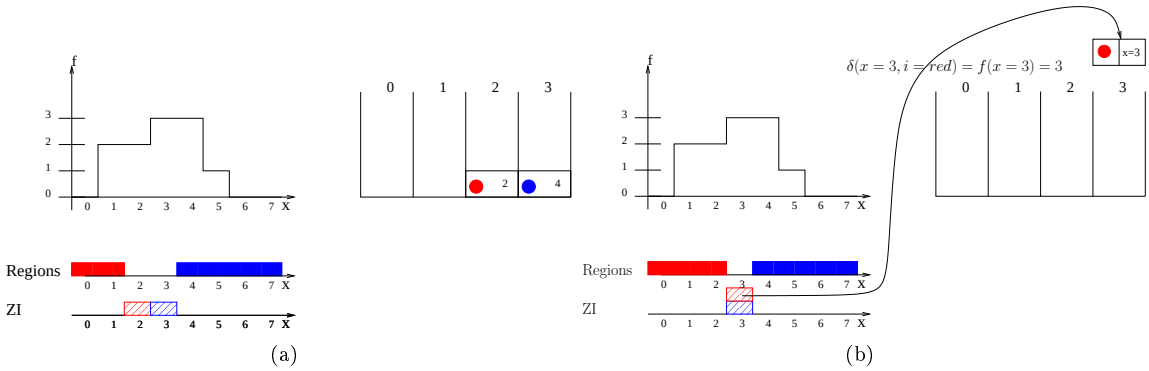


Figure 3.9: Each time, a pixel, x, is added to the ZI i, the pair (x, i) is stored in the queue $\delta(x, i)$. In this example, the pair $(x = 2, i =$red$)$ is assigned in the queue number $\delta(x, i) = f(x) = 3$. (a) before, (b) after

Customer

Between time t and $t + 1$, all pairs belonging to C^t have to be removed from the set of queues. The numerical cost of this strategy is significant. Another solution is to do nothing until a pair (x, i) is extracted from a queue j at time t. Then, if x still belongs to the zone of influence i, the pair (x, i) is selected, otherwise this pair is deleted[28] (see figure 3.10). Using this description, the implementation

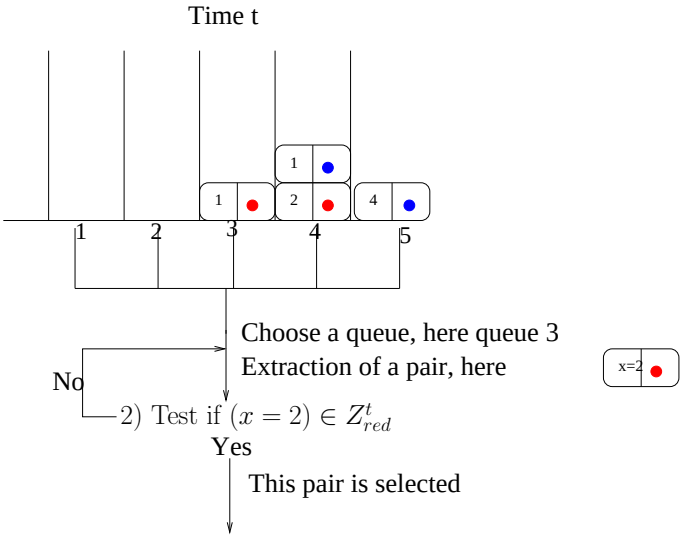


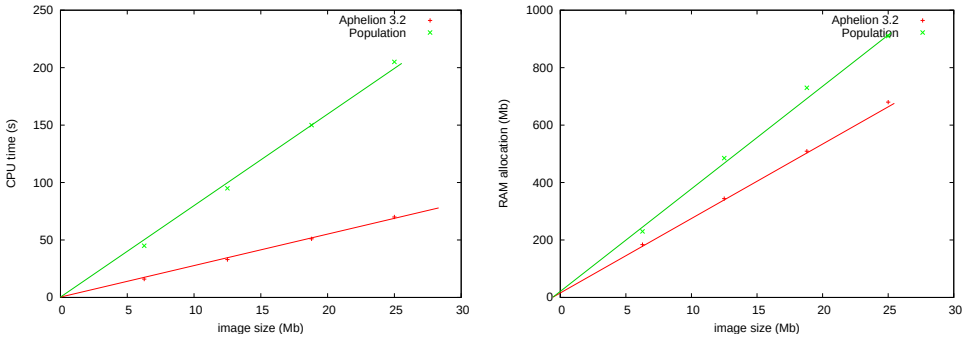
Figure 3.10: How to act? 1)Choose a queue 2)Extract a pair (x, i) 3)Test if $x \in Z_i^{t'}$ 4) Yes, action, 5) No, come back to the first step

of the methods of the class SQ can be done. At this level, all the generic classes are defined allowing their implementation. Since I have done this implementation, the next subsection will compare the efficient of the watershed transformation between the implementation in the library Aphelion and my own implementation.

[28] (when a pair (x, i) enters in the SQ at time t', x belongs to $Z_i^{t'}$, but at time $t > t'$, it cannot be the case any more.

3.3.4 Comparison

In this subsection, we will compare with the classical commercial library of Mathematical Morphology, named Aphelion. The comparison of CPU time and RAM allocation will be done using the watershed transformation with the minima as seeds (see subsection 3.4.2). The figure 3.11 shows that the library



(a) CPU time versus image size

(b) RAM allocation versus image size

Figure 3.11: A voxel is coded in one byte.

aphelion

1. is three times speeder in term of CPU time
2. consumes twice less RAM allocation

than the library Population. The Aphelion Library works more efficiently and use fewer resources than the Population Library. Optimization can occur mainly at two levels:

- **Design Level:** At the highest level, an algorithm can be implemented in many ways. The choice affects the utilisation of the resources of CPU time and RAM allocation. In the library Aphelion, the watershed algorithm is based on immersion simulations proposed by L. Vincent [173] and F. Meyer [24] in the early 90's. In this thesis, the watershed algorithm is based on the conceptualisation of SRGPAD.
- **Source code level:** In the implementation, various scopes can affect the executable computer program. The optimisation of these scopes give a real gain of efficient. Since this Computational task has not been performed in my implementation, the library Aphelion is more efficient. The optimization of the Population library at the Source code level could be resolved by the utilization of the template meta-programming approach in C++ [47]. It is a work in progress

In the next section, we will use the generic classes for the implementation of a wide range of algorithms.

3.4 Some SRGPAD algorithms

In the previous section, we have described some generic classes, called the library Population, dedicated to the implementation of SRGPAD algorithms.We will implement many optimised algorithms for the purpose of this thesis easily: Voronoï tessellation, extraction of the connected components, regional minima, distance function, watershed transformation and geodesic reconstruction. All these algorithms have been applied on 3D image with a size of 700*700*700=$3.43 \cdot 10^8$ pixels and the computational time was less than 3 hours using a conventional computer (Intel(R) Xeon(R) CPU 3.00GH). We will present the algorithms using only one queue in the System of Queue (SQ), then the algorithms using several queues.

3.4.1 One queue

In this section, we will present some algorithms using a single queue during the growing process.

Simulated Voronoï tessellation

Consider Φ a Poisson point process in a metric space M. The cells

$$C(x) = \{y \in M : d(y,x) \leq d(y,x'), x' \in \Phi\}, x \in \Phi,$$

constitute the so-called Poisson-Voronoï tessellation of M where d is a distance of M. Presented by Gilbert in 1962 [57], this statistical model is appropriate for random crystal growth. In the discrete space E, the implementation for a distance[29] associated to norm 1 or ∞[30] is done using the library Population. Starting form the affectation of each region with a point of the Poisson point process, an isotopic growing process at constant velocity is simulated. The ordering attribute function is $\delta(x,i) = 0$ for all x and i. The growing process is (see algorithm 2 and figure 3.12):

- initialisation of the regions/ZI by the seeds
- while the selected queue label 0 is not empty
  - extract the pair (x,i) from the queue label 0
  - aggregation of x on the region i
- return regions

To prove that this growing process gives a correct Poisson-Voronoï tessellation of E, this following property is used:

$$\forall x, y \in E : d(x,y) = \min_{z \in V(x)}((d(z,y)+1)$$

The generation of a Poisson point process is done using the Boost software. This implementation is not restricted to the Poisson-Voronoï tessellation since:

- each seed can be a domain of E (second row in the figure 3.12),

[29] The norm 1 of a point $x = (x_0, \ldots, x_n)$, $\|x\|_1$, is:

$$\|x\|_1 = \sum_{i=0}^{n} |x_i|$$

The norm ∞ of a point $x = (x_0, \ldots, x_n)$, $\|x\|_\infty$, is:

$$\|x\|_\infty = \max(|x_0|, \ldots, |x_n|)$$

The distance associated with the norm n is:

$$d(x,y) = \|x - y\|_n$$

[30] For the Euclidian distance, see [171].

- the growing process can be restricted to a domain $\Omega = \{\forall x \in E : I(x) \neq 0\}$ if the ordering attribute function is: $\delta(x,i) = 0$ if $I(x) \neq 0, OUT$ otherwise (third row in the figure 3.12).

Algorithm 2 Geodesic dilation

Require: S , V // *the seeds, the neighbourhood*
  //**initialisation**
  Region r;
  Restricted N = $\mathbb{N}$;
  ZI zi;
  Metric m (0); //$\delta(x,i) = 0$,
  SQ sq(FIFO, 1);//*1 FIFO queue such as all pixels with the corresponding labels of the ZI are stored in the same queue*
  Population pop (r,zi,m,sq);
  for all for j = 0 to n **do**
    int region_label = r.add_region();
    zi.add_ZI(region_label,V,$\mathbb{N}$) ;// $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j^t)$
    r.initial_seed(region_label, s_i);
  end for
  //**Growth process:**
  while sq.empty(0)==false **do**
    (x,i)=sq.pop_pair(0);
    r.growth(x, i);
  end while
  return r;

Connected components extraction

Let $\mathcal{C}_{x,y}$ be the set of continuous application from $[0,1]$ to E such that the two extremities are equal to x and y ($\forall \gamma \in \mathcal{C}_{x,y} : \gamma(0) = x$ and $\gamma(1) = y$).
$(c_i)_{0 \leq i \leq n}$ is the connected components of Ω if:

$$\biguplus_{0 \leq i \leq n} c_i = \Omega$$
$$\forall i \in (0,\ldots,n) \forall x,y \in c_i \quad \exists \gamma \in \mathcal{C}_{x,y} \quad (\forall t \in [0,1] : \gamma(t) \in \Omega)$$
$$\forall i \neq j \forall (x,y) \in (c_i,c_j) \quad \forall \gamma \in \mathcal{C}_{x,y} \quad (\exists t \in [0,1] : \gamma(t) \notin \Omega)$$

The second line means that all the points belonging to the same connected component are linked by a continuous path that is included in the domain. The third line means that any two points belonging to different connected components are not linked by a continuous path that is included in the domain (see figure 3.13). The connected components give information about the critical percolation concentration, percolation probabilities, and cluster size distributions [72]. Using the library Population, an algorithm can be implemented to extract the set of connected components. The principle is: to scan the image and each time, a connected component is touched, this connected component is removed from Ω using a growing process (see algorithm 3 and figure 3.14). Let I be the characteristic function of $\Omega = \{\forall x \in E : I(x) \neq 0\}$. The ordering attribute function is: $\delta(x,i) = 0$ if $I(x) \neq 0, OUT$ otherwise .

- scan the image ($\forall x \in E$)
  - if $I(x) \neq 0$
    - create a region/ZI initialised by the seed $\{x\}$
    - while the queue label 0 is not empty

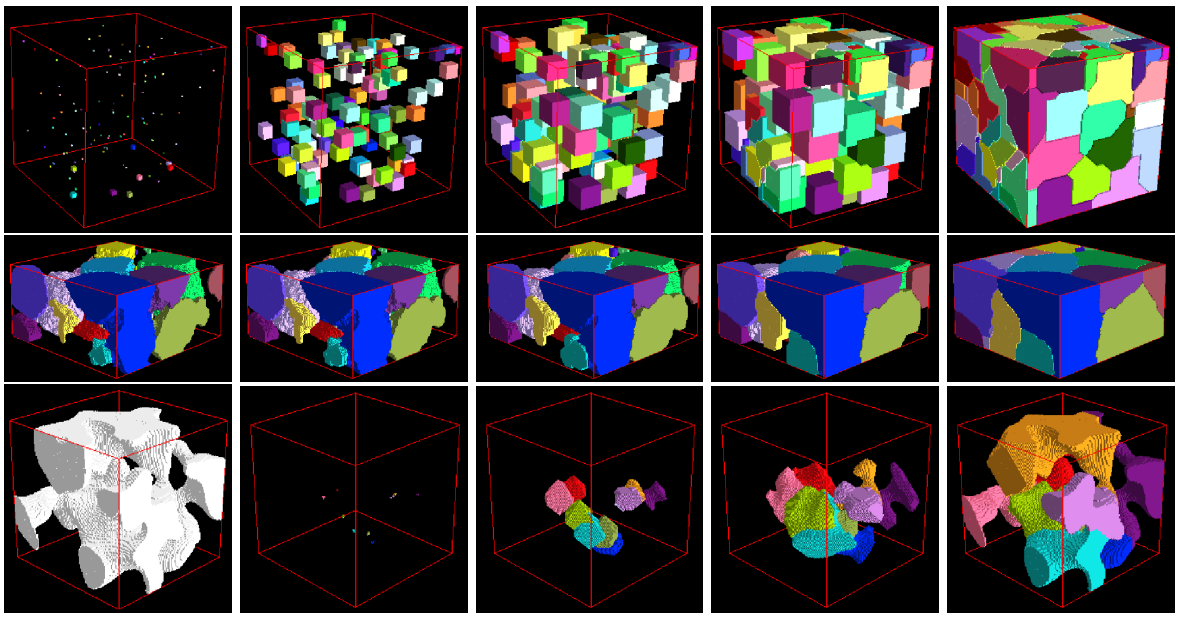


Figure 3.12: Illustration of the Voronoi-growing process. For the first row, the seeds are generated following a Poisson point process, *A video is available at http://pmc.polytechnique.fr/∼vta/geodesic_invariant_cube.mpeg.* For the second row, each seed is a set of voxel (in the subsection 4.4.1, we will see how to extract individually each grain of a grains phase. In order to characterise the connectivity of the grains phase, we apply a Voronoi-growing process with the regions initialised by a grain of the grains phase. At the end of the growing process, a network can be defined by the neighbourhood of the regions and it gives the connectivity of the grains phase.), *A video is available at http://pmc.polytechnique.fr/∼vta/geodesic_invariant_cube_grain.mpeg.* For the last row, the growing process is restricted by the first figure, *A video is available at http://pmc.polytechnique.fr/∼vta/geodesic.mpeg.*

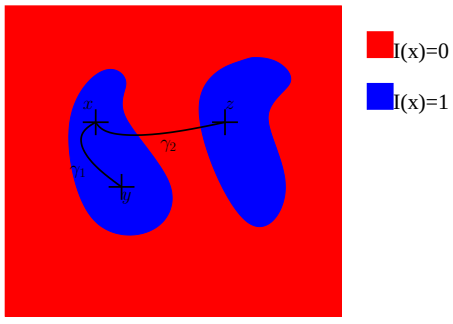


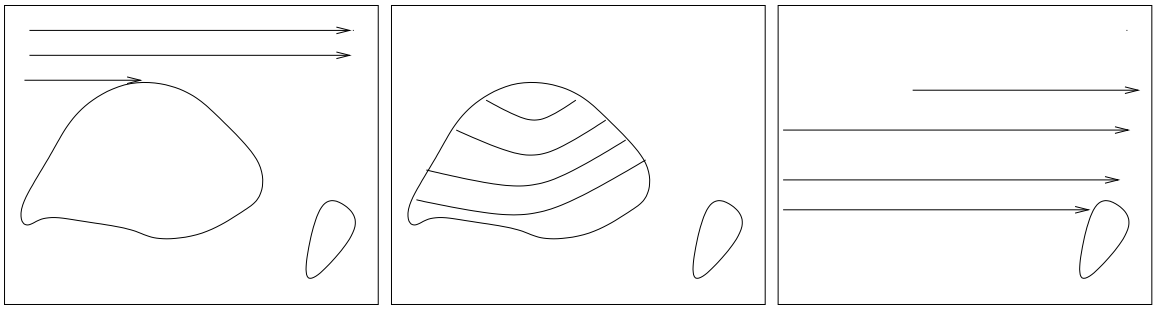
Figure 3.13: Illustration of the connected components extraction. Whatever the continuous path between the point x and z, the path goes through the red region. There exists a path between the point x and y such the path belongs only to the blue region.

  - extract (y, i) from the queue label 0
  - aggregation of y on the region i
  - $I(y) = 0$

- return regions

This algorithm allows (see figure 3.15):

- to remove all the connected components touching the boundary,
- to fill the hole[31],
- to keep only the cluster which area is maximum.



(a) (b) (c)

Figure 3.14: Illustration of the extraction of the connected components (a) : scan the image until $I(x) \neq 0$, (b): growing process starting form x ($I(x) \neq 0$) such that at each growth the characteristic function of Ω is modified $I(x) \neq 0 \rightarrow I(x) = 0$, (c): at the end of the growth, the connected component has been extracted and removed from Ω and the scanning continues until $I(x) \neq 0$.

Regional minima

Let $\mathcal{C}_{x,y}$ be the set of continuous application from $[0,1]$ to E such that the two extremities are equal to x and y ($\forall \gamma \in \mathcal{C}_{x,y} : \gamma(0) = x$ and $\gamma(1) = y$).
Let f be a grey-level image (a topographic surface).
$S = (s_i)_{0 \leq i \leq n}$ is the decomposition of (E, f) in level connected components if:

$$\biguplus_{0 \leq i \leq n} s_i = E,$$

$$\begin{array}{rll} \forall i \in (0, \ldots, n) & \forall x, y \in s_i & \exists \gamma \in \mathcal{C}_{x,y} \ (\forall t \in [0,1] : f(\gamma(t)) = f(x)); \\ \forall i \neq j & \forall x, y \in s_i & \forall \gamma \in \mathcal{C}_{x,y} \ (\exists t \in [0,1] : f(\gamma(t)) \neq f(x)). \end{array}$$

The second line means that all the points belonging to the same level connected component are linked by a continuous path which level is constant. The third line means that any two points belonging to different connected components are not linked by a continuous path which level is constant (see figure 3.16).

[31] To fill the hole, the procedure is the following chain of algorithms:

1. inversion,
2. extraction of the connected components,
3. deleting the connected components no touching the image boundary,
4. binarization and inversion.

Algorithm 3 Extraction of the connected components

Require: I, V //*The binary image, the neighbourhood*
  // ***initialisation***
  Region r;
  Restricted N = $\mathbb{N}$;
  ZI zi;
  Metric m (0,I); //$\delta(x,i) = 0$ *if* $I(x) \neq 0, OUT$ *otherwise*
  SQ sq(FIFO, 1);//*1 FIFO queue such as all pixels with the corresponding labels of the ZI are stored in the same queue*
  Population pop (r,zi,m,sq);
  //***Scan the image***
  for all $\forall x \in E$ **do**
    if $I(x) \neq 0$//*Test if a connected component is touched* **then**
      int region_label = r.add_region();
      zi.add_ZI(region_label,V,$\mathbb{N}$) ;// $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j^t)$
      r.initial_seed(region_label, $\{x\}$);
      //**Growth process:**
      while sq.empty(0)==false **do**
        (y,i)=sq.pop_pair(0);
        r.growth(y,i);
        $I(y) = 0$;
      end while
20: **end if**
  end for
  return r;

In this decomposition, an element s of S is a regional minimum if the levels of points belonging to the outer boundary of s are greater than the level of points belonging to s (see figure 3.16):

$$\forall (x,y) \in (s, (s \oplus V) \setminus s) \quad f(x) < f(y).$$

Using the library Population, a growing process is defined to extract the regional minima. This growing process is: to scan the image ($\forall x \in E$) until there is not a region on x yet($r[x] == NO\_REGION$) , then to start the growing process initialised by the seed equal to $\{x\}$. Let $level = f(x)$ be the level of the growing region. The ordering attribute function is defined as:

$$\delta(y,i) = 0 \text{ if } f(y) \leq level, OUT \text{ otherwise}$$

For this algorithm, the ZI is defined as: $Z_i^t = (X_i^t \oplus V) \setminus X_i^t$. The ZI is localised on the outer boundary region even if there are still some region to check the condition: $\forall (x,y) \in (s, (s \oplus V) \setminus s) : f(x) < f(y)$. The growing process is (see algorithm 4 and figure 3.16):

- scan the image ($\forall x \in E$)
  - if $r[x] == NO\_REGION$
    - create a region/ZI initialised by the seed $\{x\}$
    - $level = f(x)$
    - while the queue label 0 is not empty
      - extract (y,i) from the queue label 0
      - if $f(y) == level$
      - then aggregation of y on the region i
      - otherwise
      - this region/ZI is not a regional minimum
- return regions that are regional minima

Algorithm 4 Regional minima

Require: f, V *//The grey-level image, the neighbourhood*

```
// initialisation
int level;
Region r;
Restricted N = {i};
ZI zi;
Metric m (0,f,level); //δ(y,i) = 0 if f(y) ≤ level, OUT otherwise
SQ sq( FIFO, 1);//1 FIFO queue such as all pixels with the corresponding labels of the ZI are stored in the same queue
Population pop (r,zi,m,sq);
Set set; //Container: self-balancing binary search tree.
//Scan the image
for all ∀x ∈ E do
  if r[x] == NO_REGION//Test if there is no a region on x then
    level = f(x);
    int region_label = r.add_region();
    zi.add_ZI(region_label,V,ℕ) ;// Z_i^t = (X_i^t ⊕ V) \ (X_i^t)
    r.initial_seed(region_label, {x});
    bool regional_minima=true;
    //Growth process:
    while sq.empty(0)==false do
      (y,i)=sq.pop_pair(0);
      if f(y) < level then
        regional_minima=false;
      else
        r.growth(y,i );
      end if
    end while
    if regional_minima==true then
      set.insert(ref_tr);
    end if
  end if
end for
return (r,set);
```

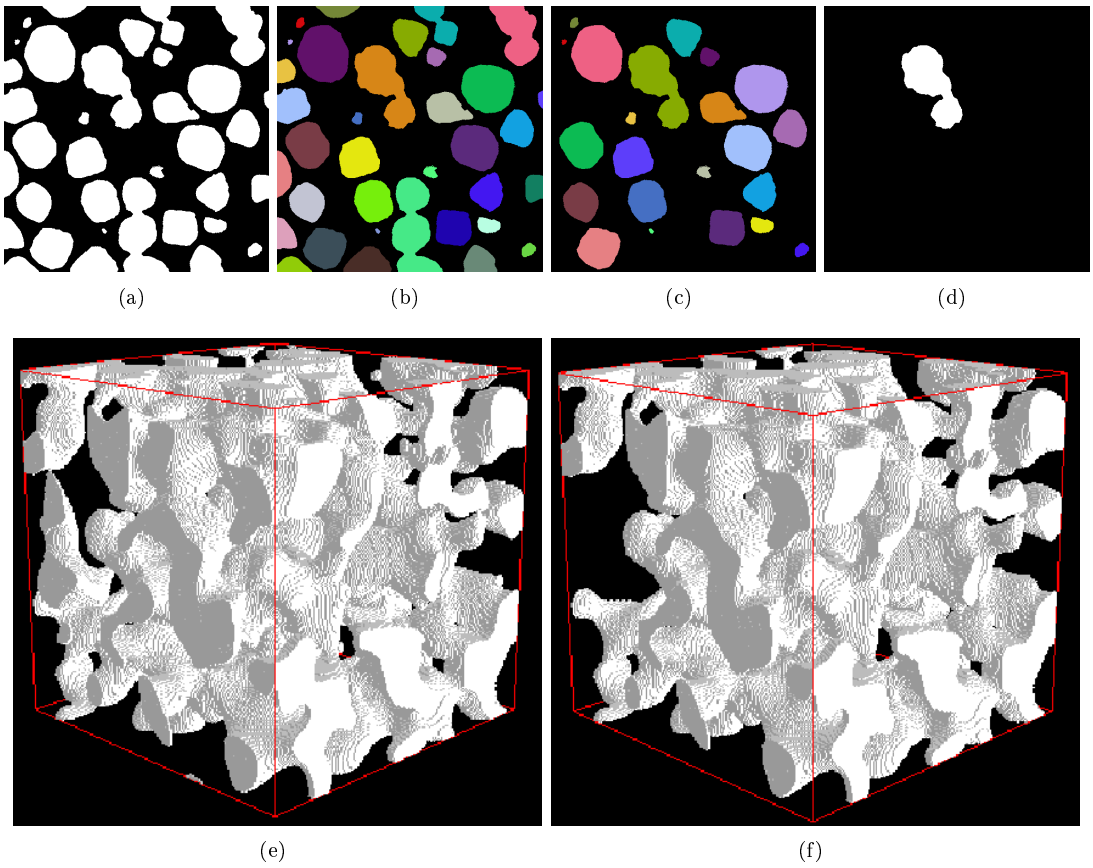


Figure 3.15: (a): the initial binary image; (b): extraction of the connected components; (c): the connected components touching the boundary of the image are removed; (d): the max cluster of the previous image; (e): the initial image; (f): maximum cluster of percolation after the selection of the component whose the area is maximum in the connected components.

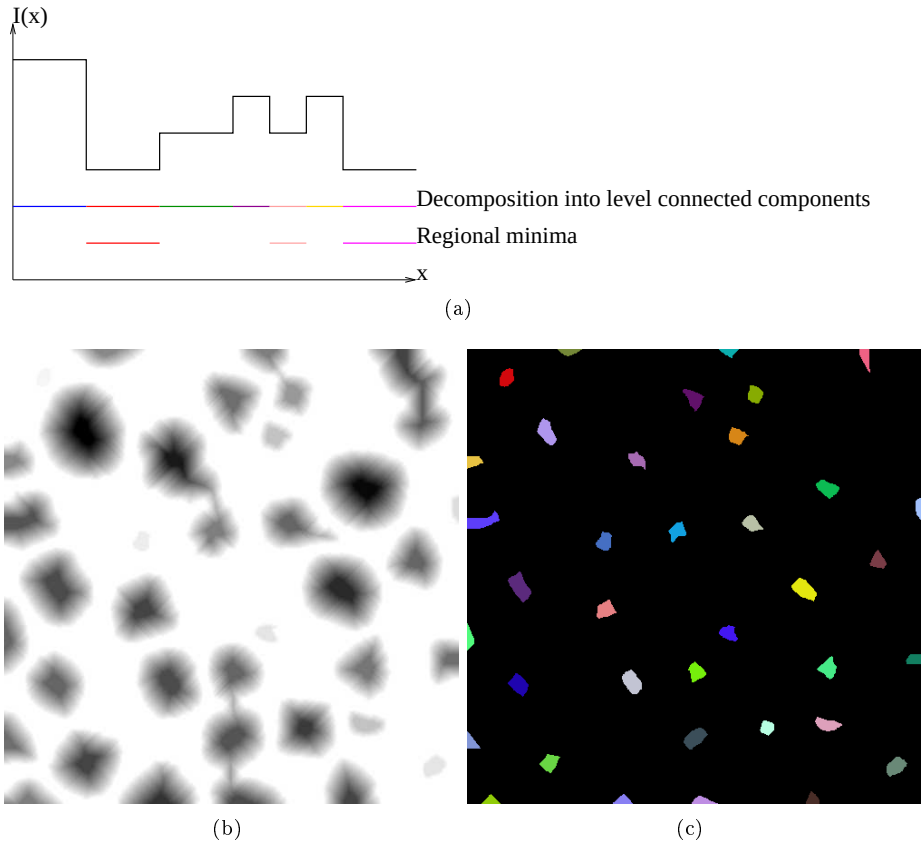


Figure 3.16: (a): principle of minima; (b): a grey-level image; (c): the regional minima of this image

3.4.2 n queues

In this section, we present algorithms which SQ contains several queues.

Distance function: flip-flop queue

Let Ω be a domain of E and x and y two points of Ω. We call geodesic distance $d^{\Omega}(x,y)$ in Ω the lower bound of the paths length in Ω linking x and y:

$$d^{\Omega}(x,y) = \min_{\forall \gamma \in \mathcal{C}^{\Omega}_{x,y}} \text{Length}(\gamma).$$

where $\mathcal{C}^{\Omega}_{x,y}$ be the set of continuous application from $[0,1]$ to Ω such that the two extremities are equal to x and y.
Let S be a set. We call the geodesic distance $d^{\Omega}(S,y)$, the lower bound of all geodesic distance $d^{\Omega}(x,y)$ such that x belongs to S.

$$d^{\Omega}(S,y) = \min_{\forall x \in S} d^{\Omega}(x,y).$$

We have:

$$d^{\Omega}(x,y) = \min_{\forall z \in (\{x\} \oplus B_{r=1}) \cap \Omega} (d^{\Omega}(x,z) + d^{\Omega}(z,y)).$$

In discrete space, as $d^{\Omega}(x,z)$ is equal to 1, we get:

$$d^{\Omega}(x,y) = \min_{\forall z \in (\{x\} \oplus B_{r=1}) \cap \Omega} (d^{\Omega}(z,y)) + 1. \tag{3.1}$$

In order to calculate the distance function, we use the property of the equation 3.1: all pixels at distance $d+1$ are neighbourhood of at least one pixel at distance d. Using the library Population, a SRGPAD algorithm is implemented to calculate the distance function $d^{\Omega}(A,y)$ for all y in Ω. Let I be the characteristic function of Ω.

n queue implementation

First, we initialise the single region by the seed equal to A, the growing process takes place on pixels in the neighbourhood of A. Since the pixels belonging to the set A are at distance 0 ($d^{\Omega}(A, x \in A) = 0$), the pixels belonging to this growing process, D_1, are at distance 1. Then, the growing process takes place on pixels in the neighbourhood of D_1. The pixels belonging to this growing process, D_2, are at distance 2 and so on. To operate this growing process, the ordering attribute function is: $\delta(x,i) = d+1$ if $I(x) \neq 0, OUT$ otherwise. We initialise the variable distance, d at 0 and the single region with the seed A. Since the the ordering attribute function is: $\delta(x,i) = d+1$ if $I(x) \neq 0, OUT$ otherwise, at the initialisation, the pairs are stored in the queue label 1. When the initialisation is finished, we increment the variable d ($d = 1$). We extract the pair (x,i) from the queue label $d = 1$ until this queue is empty. At each extraction, we aggregate the pixel x on the single region and the distance function is assigned at distance $d = 1$ on x. When the queue label $d = 1$ is empty, we increment the variable d, and so on. This recursion is done while the system of queues is not empty.
The growing process is:

- int d=0
- initialise of the single region/ZI by the seed $s = A$
- while the system of queues is not empty
  - d = d + 1
  - while the queue label d is not empty
    - extract (y,i) from the queue label d
    - aggregation of x on the region i

        * dist[x]=d

- return dist

The number of queues is equal to the maximum of the distance function. The problem of this implementation is that this number is unknown before the growing process. To overcome this problem, a solution is to use a flip-flop queue.

Flip-flop queue implementation

In the last implementation, during the growing process, there are only two queues in the SQ not empty at the step d: the queue label d where the pairs are extracted and the queue label $d+1$ where the pairs are stored. Using this property, the pairs are now extracted from the queue label $flip$ and stored in the queue label $flop$. The ordering attribute function, $\delta(x,i)$, is equal to $flop$ if $I(x) \neq 0, OUT$ otherwise. The growing process becomes (see figure 3.17 and 3.18 and algorithm 5):

- int d=0
- initialise of the single region/ZI by the $s = A$
- while the system of queues is not empty
  - d = d + 1
  - switch(flip,flop)
  - while the queue label flop is not empty
    * extract (y,i) from the queue label flop
    * aggregation of x on the region i
    * dist[x]=d
- return dist

The watershed transformation

The watershed transformation was introduced in the subsubsection 3.2.4. The application of the watershed transformation will be presented in the chapter 4.
Let f be the topographic surface. The ordering attribute function is: $\delta(x,i) = \max(level, f(x))$ (see algorithm 6). This growing process is not limited to the watershed transformation on E. The growing process can be restricted to a domain $\Omega = \{\forall x \in E : I(x) \neq 0\}$ if the ordering attribute function is: $\delta(x,i) = \max(level, f(x))$ if $I(x) \neq 0, OUT$ otherwise (see figure 3.19).

Geodesic reconstruction

The geodesic reconstruction is an efficient tool in Mathematical Morphology [158, 20]. Given a function f and a function g with $f \geq g$, the geodesic erosion is defined as:

$$R_g^*(f) = E_g^\infty(f)$$

where $E_g^\infty(f)$ is the infinitely iterated geodesic erosion such as $E_g^{t+1}(f) = \sup(E_g^t(f) \ominus V, g)$ with $E_g^0(f) = f$.
Introduced by Grimaud [59], the geodesic reconstruction is called a dynamic filter when the function f is equal to the function g plus a constant h: $f(*) = g(*) + h$. The dynamic filter belongs to the category of vertical filters that fills the valleys with depth lower than h (see figure 3.20).
Introduced by Beucher [20], the geodesic reconstruction is called a homotopic transformation when the function f is equal to g on the seeds, $(s_0, \ldots, s_n)$, and '∞' on the complement of the seeds [21] (see figure 3.21). The homotopic transformation is used in the watershed transformation implemented by Vincent [173] in order to keep only the most significant contours in the areas of interest between the

Algorithm 5 Distance function

Require: I, $s = A$, V *//the characteristic function, the seed equal to A, the neighbourhood*

*// **initialisation***
Image Img_dist;
int flip=0, flop =1;
Region r;
Restricted N = $\mathbb{N}$;
ZI zi;
Metric m (0,$flip$); *//*$\delta(x,i) = flip$ *if* $I(x) \neq 0, OUT$ *otherwise*
SQ sq(FIFO, 1);*//1 FIFO queue such as all pixels with the corresponding labels of the ZI are stored in the same queue*
Population pop (r,zi,m,sq);
int region_label = r.add_region();
zi.add_ZI(region_label,V,$\mathbb{N}$) *;//* $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j)$
r.initial_seed(region_label, s); *;//* $X_i^{t=0} = s$
int dist=0;
*//**Growth process***
while sq.all_empty()==false **do**
  switch(flip,flop);
  dist++;
  while sq.empty(flop)==false **do**
    (x, i)=sq.pop(flop);
    r.growth(x, i);
    Img_dist(x)= dist;
  end while
end while
return Img_dist;

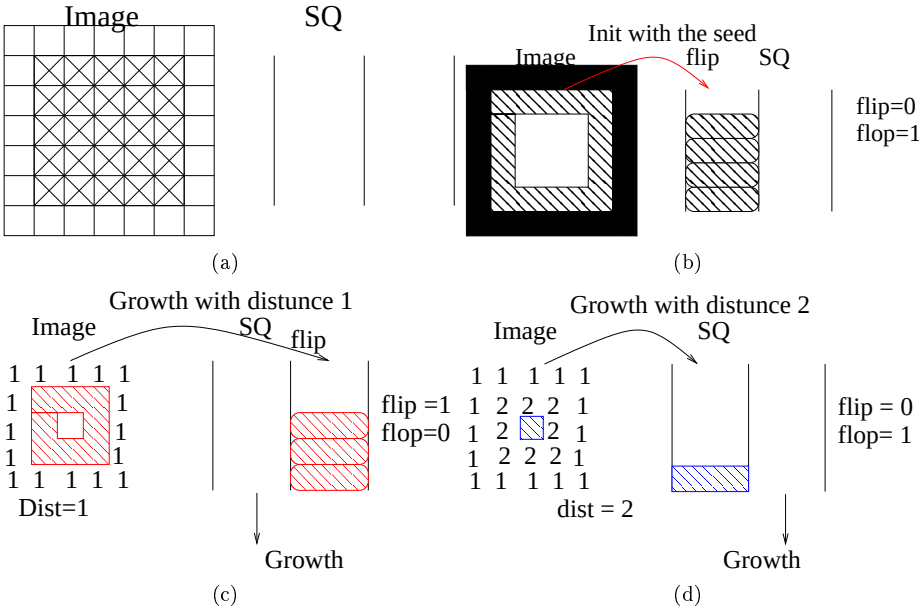


Figure 3.17: Illustration for the distance function algorithm. (a): the seed is the white pixels; (b): after the initialisation of the region/ZI by the seed. The pixels belonging to the ZI are stored in the flip queue; (c): after the growing process at the distance 1, the queue number 0 is empty and the queue number 1 stored the pixels at the distance 2. As the queue number 0 is empty, flip and flop will be switched and the growing process will take place now in the pixels at distance 2; (d): the same as the previous figure with the distance 2.

markers. In our implementation of the watershed transformation, the homotopic transformation is done during the growing process.

The classical implementation of the geodesic reconstruction is to use directly the formula $E_g^{t+1}(f) = \sup(E_g^t(f) \ominus V, g)$ with $E_g^0(f) = f$. Numerically, the recursion is stopped when there is convergence, $E_g^{t+1}(f) = E_g^t(f)$. The implementation is simple but the complexity is $O(n \cdot k)$, where n is the number of pixels of the image and k is the index of the convergence condition.

An alternative to the previous algorithm is a SRGPAD algorithm by a merging procedure. First, the regional minima $(S_i)_{0 \leq i < q}$ of g are extracted. For the convenience, each S_i is reduced to a single pixel x_i thrown randomly in S_i corresponding to a seed. Then, a watershed procedure is applied on the topographic surface g with the difference that the creation of region/ZI is done during the merging procedure. At the immersion level equal to *level*, a number of regions/ZI is created for all x_i such that $f(x_i)$ is equal to *level* and if there is not yet a region on the pixel x_i. At every growth x of a region, the immersion *level* is attributed to the dynamic function on x, $E_g^{\infty}(f)(x) = level$ (see figure 3.22 and algorithm 7). The complexity of this algorithm is $O(n)$ where it is the number of pixels of the image. The application is shown on the figure 3.23.

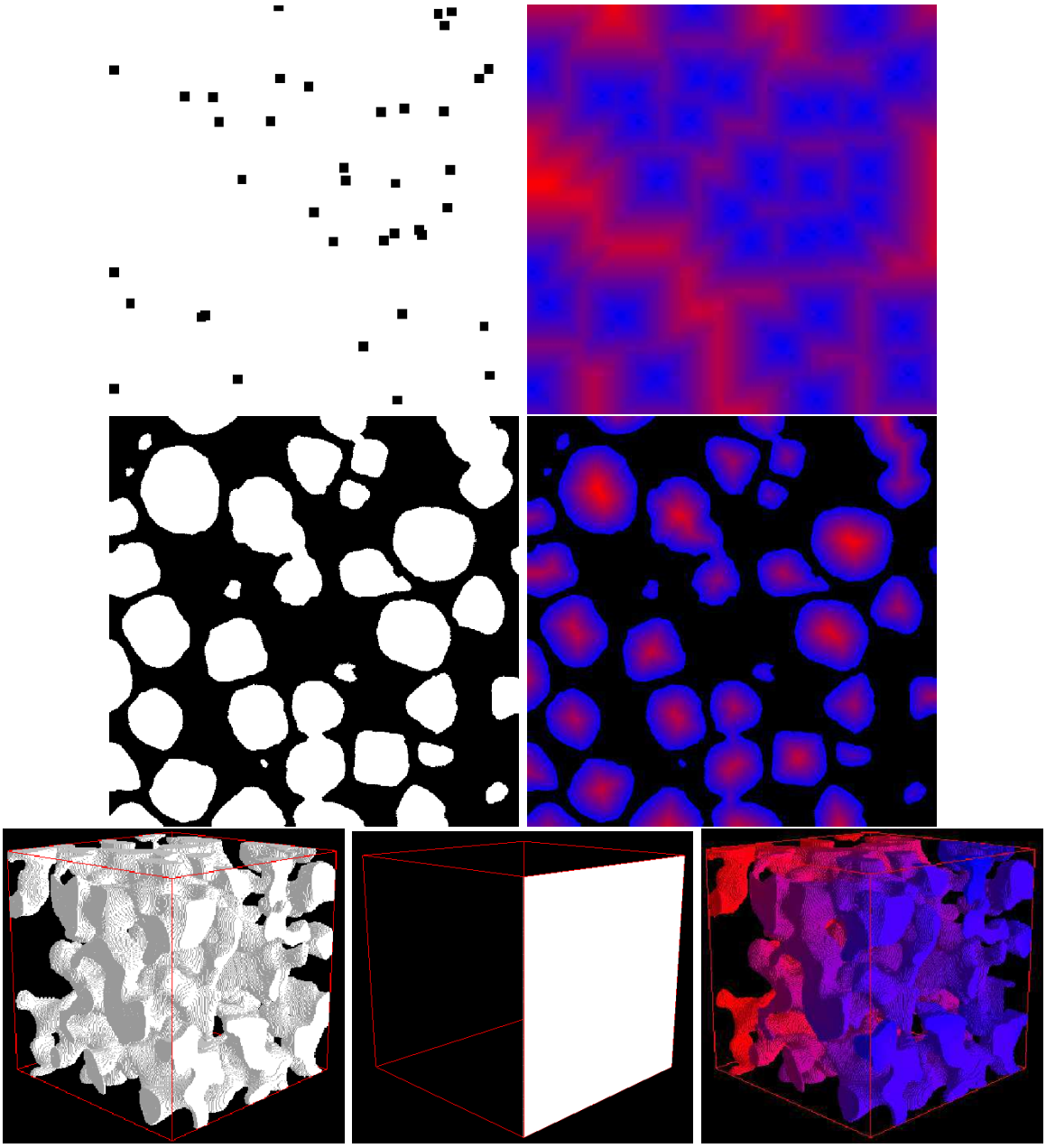


Figure 3.18: Illustration for the distance function. First row: the left image is a realisation of a random points process, $(x_0, \ldots, x_n)$ with $\lambda = 0.005$ (for the visualisation convenience, the realisation has been dilated). Let Ω be the complement of the union of these points. the right image is the distance function $d^{\Omega}(\Omega^c, y)$ for all y belonging to Ω . Second row: Same as first row except that the set Ω is the set of white pixels. Third row: Left figure Ω, middle figure, the seed, s, right figure, $d^{\Omega}(s, y)$ for all y belonging to Ω. This distance function is used to calculate the geometrical tortuosity.

Algorithm 6 The watershed transformation

Require: f, $(s_i)_{0\leq i\leq n}$, V *//The topographic surface, the seeds, the neighbourhood*
//initialisation
int immersion_level =0;
Region r;
ZI zi;
Restricted N = $\mathbb{N}$;
Metric m (I,immersion_level); *//*$\delta(x,i) = max(f(x)$,*immersion_level)*
SQ sq(FIFO, max level of the topography);*//n FIFO queue such as the pixels with the corresponding label of the ZI are stored depending on the elevation on the topographic surface*
Population pop (r,zi,m,sq);
for all for j = 0 to n **do**
  int region_label = r.add_region();
  zi.add_ZI(region_label,V,$\mathbb{N}$) *;//* $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j\in\mathbb{N}} X_j)$
  r.initial_seed(region_label, s_i);
end for
//Growth process :
for all immersion_level = 0 to the max level of the topography **do**
  while sq.empty(immersion_level)==false **do**
    (x,i)=sq.pop_pair(immersion_level);
    r.growth(x, i);
  end while
end for
return r;

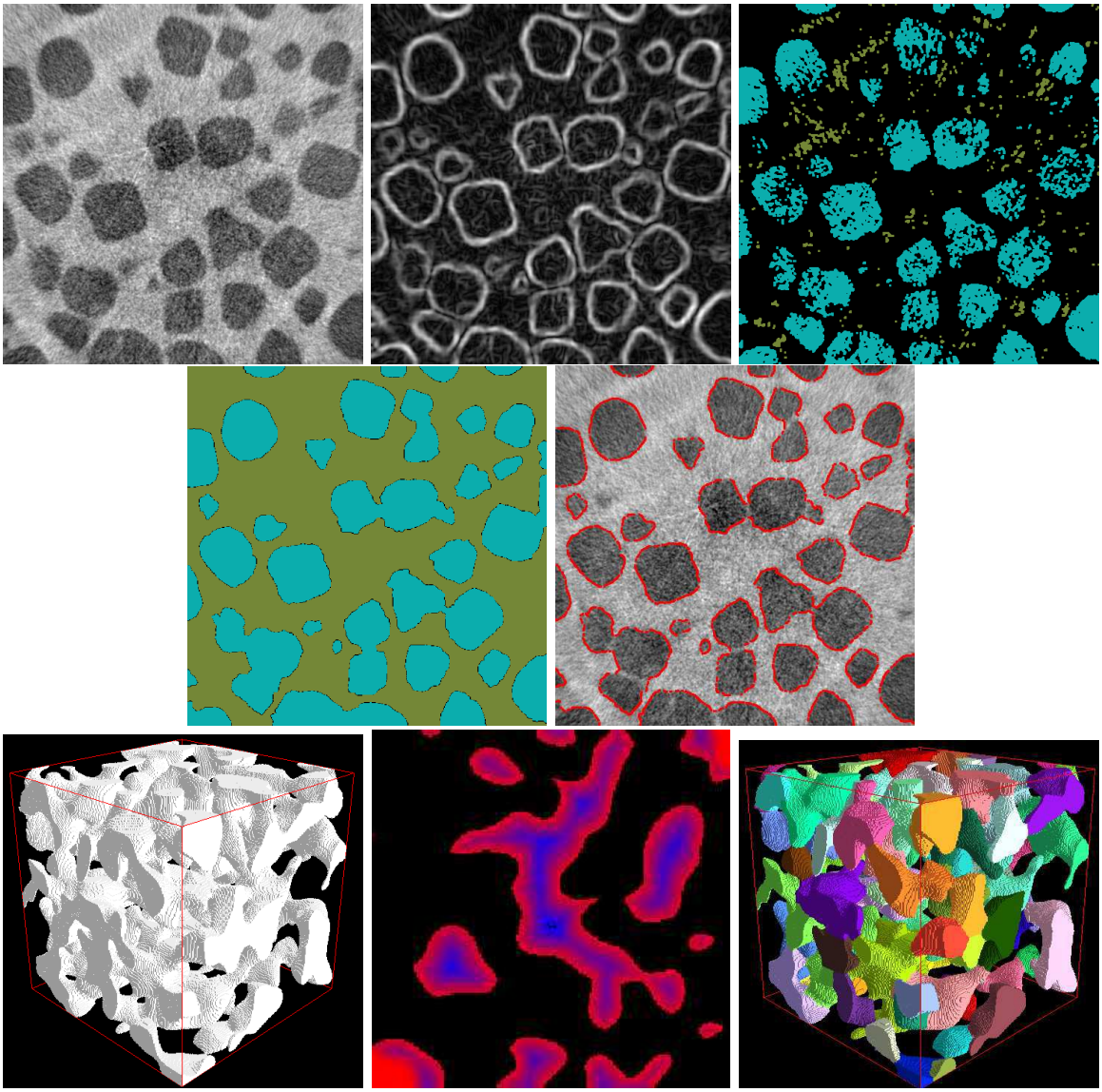


Figure 3.19: The first row: the first image is the initial image, the second image is the application of a Deriche's gradient[41] on the first image, the third image is the visualisation of two seeds: one localised on the grains, the other on the complement of grains. The second row: the first image is the basins due to the watershed transformation on the topographic surface equal to the previous gradient image using the two seeds, the second image is the foreground of the boundary region on the initial image. There is a good match with the visual segmentation. The third row: the first image is the initial image, the second image is a 2D slice of the opposite distance function of the initial image, the third image is the basins due to the watershed transformation restricted by the initial image on the topographic surface equal to the opposite distance function of the initial image with appropriate seeds.

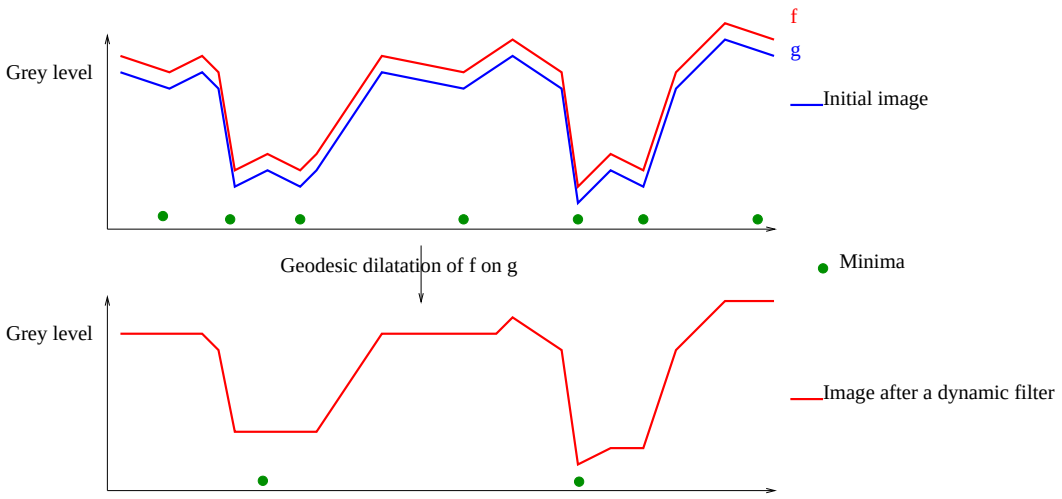


Figure 3.20: The dynamic filter. Before the application of the dynamic filter, there are many minima (green bullets). After the application of the dynamic filter, there are only two minima.

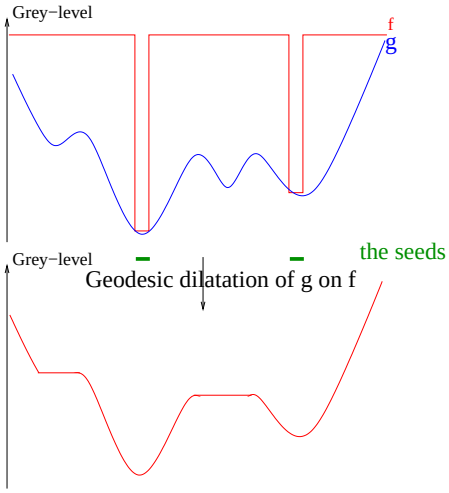


Figure 3.21: Homotopic transformation. The initial image g with seeds. For all x, the function $f(x)$ is equal to $g(x)$ if x belongs to the seeds and ∞ otherwise.

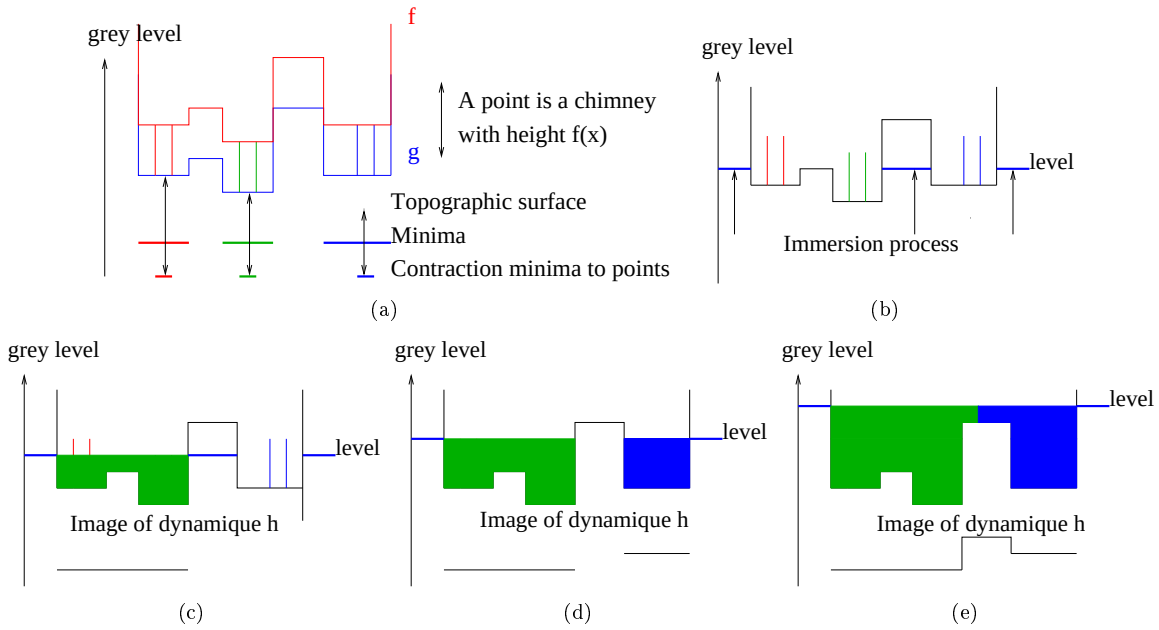


Figure 3.22: Dynamic filter. The process is: (a) extraction of the regional minima of f, contraction of these regional minima to single pixels x_i, association of these pixels, x_i, with chimneys with height $f(x_i) - g(x_i)$; (b) immersion process: the water enters in the topographic surface by the chimneys if there is not a region yet; (c) the basin takes the colour of the chimney and at every growth x of a region, the dynamic image takes the immersion level in x; (d) the red chimney does not create region/ZI because the green region is already here. Note that there are 3 minima in the initial image and only two after the dynamic filter.

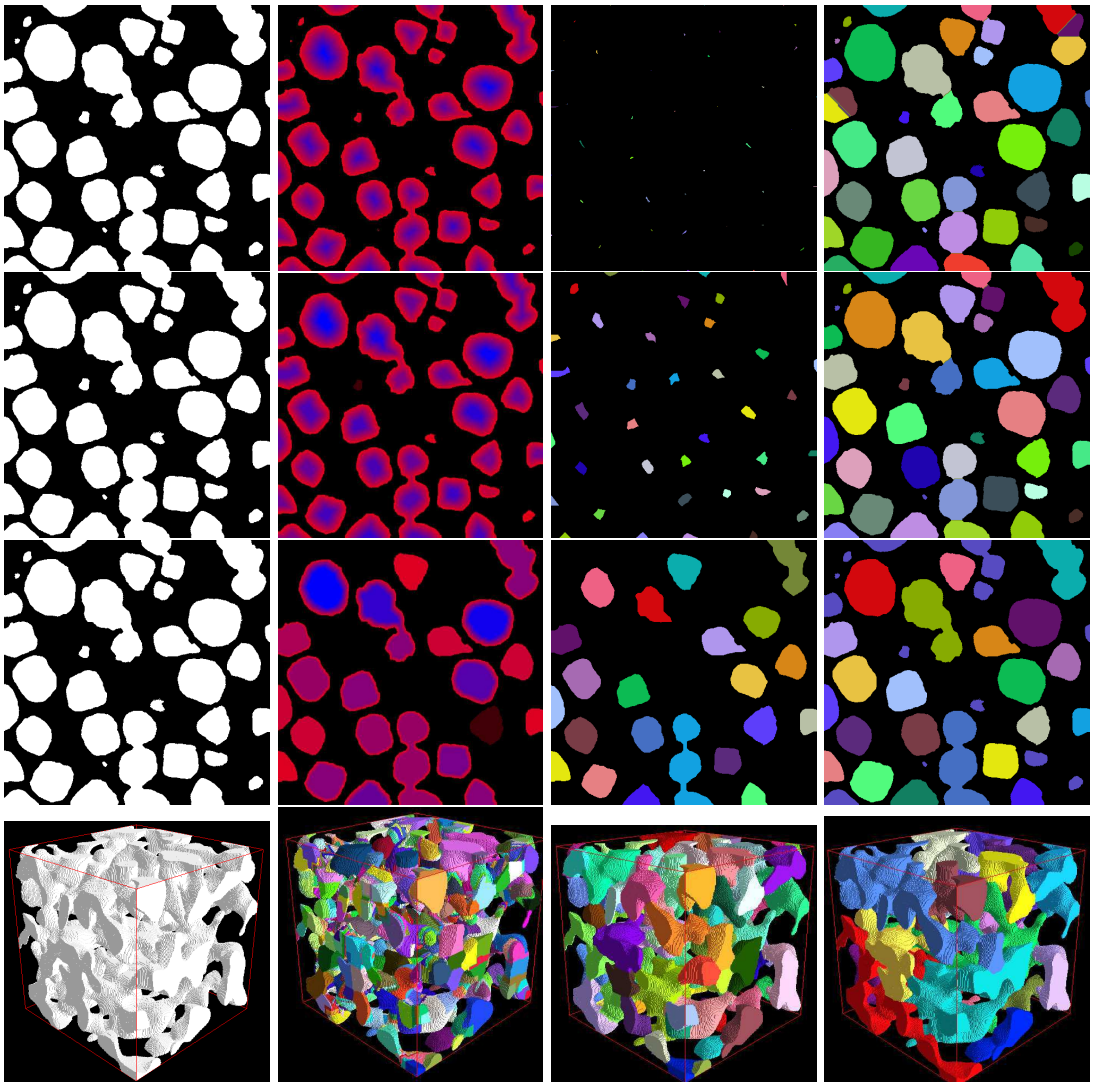


Figure 3.23: For the first three rows, the first image is the initial image, the second image is the inverse of the distance function of the initial image after the application of a dynamic filter, (h=0 for the first row, h=3 for the second row, h=10 for the third row), the third image is the regional minima of the second image, the fourth image is the watershed transformation restricted by the first image on the topographic surface equal to the second image using the seeds equals to the third image. The fourth row is the same process but in 3D. The first image is the initial image. h=0 for the second image, h=3 for the third, h=10 for the fourth.

Algorithm 7 Geodesic reconstruction

Require: f, g V, h //*the two images (for the dynamic filter $f = g + h$, the neighbourhood*

```
    //initialisation
    Image GR(f );
    int immersion_level =0;
    Region r;
    ZI zi;
    Restricted N = ℕ;
    Metric m (I,immersion_level); //δ(x, i) = max(g(x),immersion_level)
    SQ sq( FIFO, max level of the topography);//n FIFO queue such as the pixels with the corresponding
    label of the ZI are stored depending on the elevation on the topographic surface
    Population pop (r,zi,m,sq);
    (S_i)_{0≤i≤q}= regional_minima(f);
    (x_i)_{0≤i≤q}= rand_pixel((S_i)_{0≤i<q});
    //Growth process:
    for immersion_level = 0 to the max level of the topography do
      for i=0 to q do
        // Creation of region/ZI if two conditions
        if (immersion_level == f(x_i)) and r[x_i]==NO_REGION then
          int region_label = r.add_region();
          zi.add_ZI(region_label, V,N) ;
          r.initial_seed(region_label, {x});
20:       GR(x_i)=immersion_level;
        end if
      end for
      while sq.empty(immersion_level)==false do
        (x,i)=sq.pop_pair(immersion_level);
        r.growth( x, i );
        GR(x)=level;
      end while
    end for
    return GR;
```

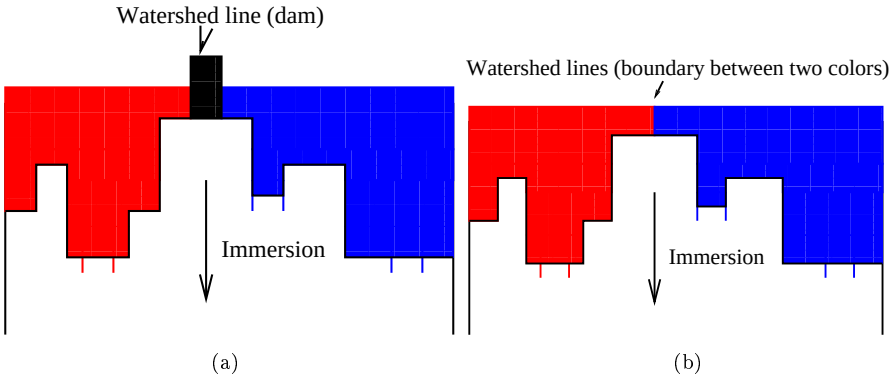


Figure 3.24: (a): A dam (boundary region) is used to separate the basins, (b) nothing is used to separate the basins.

3.5 How to manage the region collisions

In the two previous sections, we have conceptualised the SRGPAD and implemented many algorithms. Each implementation using this library is quick and provides efficient algorithms. At the end of the growing process, the regions are a partition of the space or a domain. In a classical growing process, two conventions for the partition are: with or without a boundary region for the separation of the regions (see figure 3.24). The SRGPAD conceptualisation allows the implemention of these both conventions. But with these convention, the partition depends on the seeded region initialisation order (SRIO) [22, 114, 175] such that, if the order of initialisation is changed, the localisation of the inner border of each region will also change. To overcome this problem, we will localise the boundary region on a set of pixels, called ambiguous points. This set is called ambiguous points since the determination, at which regions they belong, is impossible. We define a growing process assigning:

- unambiguous points to the appropriate regions,
- ambiguous points to the boundary region.

Using this growing process, the localisation of final partition is independent of the SRIO. This section is decomposed in two subsections: the classical growing processes and growing process independent of the SRIO.

3.5.1 Classical growing processes

This section presents two classical growing processes. For the first, there is no boundary region to separate growing regions. For the second, there is a boundary region to separate growing regions. The geodesic dilation [149] is used like an example but this approach can be used for the most of SRGPAD algorithms if the algorithm can be reduced to a succession of geodesic dilation [121]. This section is decomposed in two parts: definition of two distinct partitions and definition of two growing processes.

Two distinct partitions

A segmentation of a set Ω is simple partition of Ω into subsets X_i , $i = 1, \ldots, m$, for some m if :

1. $\Omega = \cup_{i=1}^{m} X_i$
2. $\forall i \neq j : X_i \cap X_j = \emptyset$.

A segmentation of a set Ω is V-boundary partition[32] of Ω into subsets X_i $i = 1, \ldots, m$, for some m, and X_b if:

[32] V-boundary-partition is also a simple-partition.

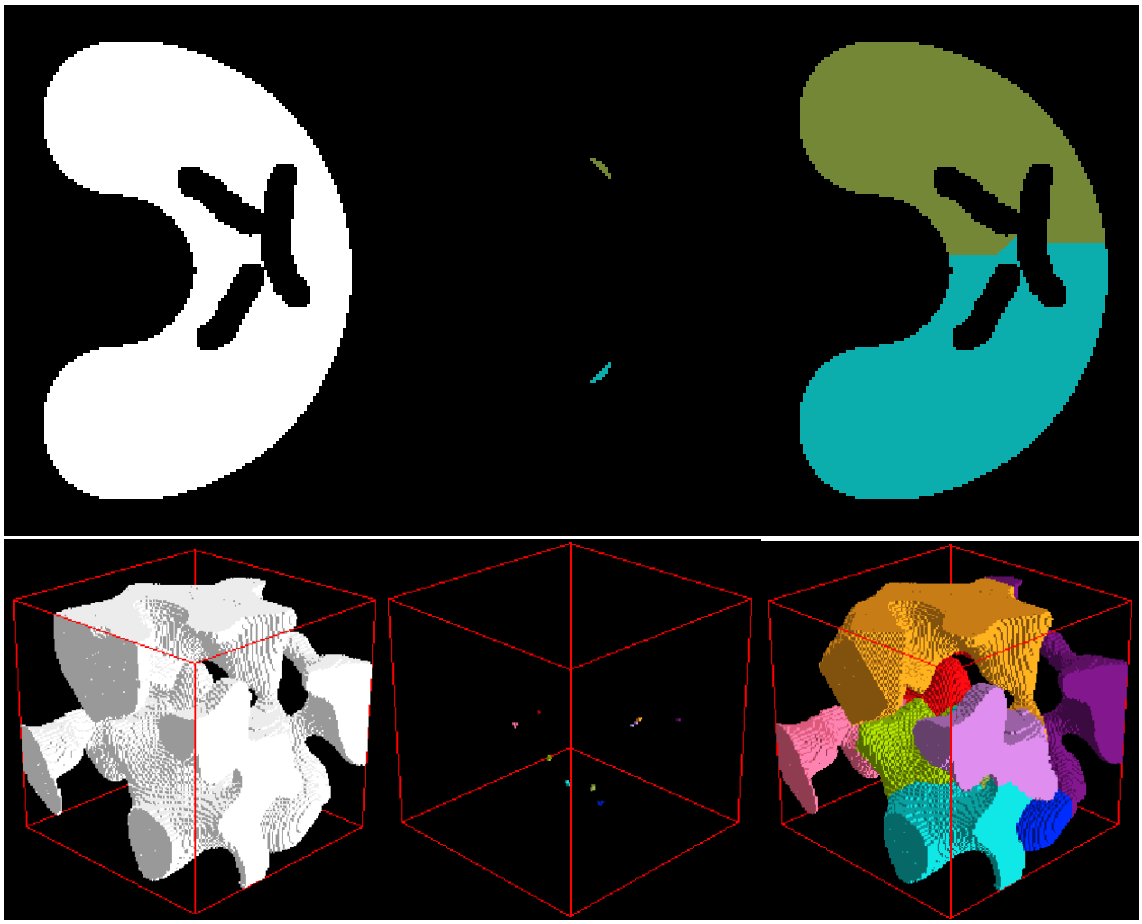


Figure 3.25: Illustration of simple partition. For both rows, the first image is the initial image representing the set Ω , the second image is the seeds and the last image is the simple-partition after the geodesic dilation with a ordering attribute function equal to: $\delta(x,i) = 0$ if $I(x) \neq 0, OUT$ otherwise . The first row is in 2D and the second in 3D. For both cases, the regions are simple-partitions of Ω

1. $\Omega = (\cup_{i=1}^{m} X_i) \cup X_b$
2. $\forall i \neq j : (X_i \oplus V) \cap X_j = \emptyset$
3. $X_b \ominus V = \emptyset$.

The second condition states that the boundary region separates the other regions and the third condition states that the thickness boundary region is equal to 1.

Simple partition

To get a simple-partition using the SRGPAD, the zone of influence (ZI) at each region is localised on the outer boundary region excluding all other regions: $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j)$. During the growing process, when a pair (x,i) is extracted from the SQ, there is a simple growth: r.growth(x, i)). At the end of the growing process, the regions $X_i^{t=\infty}$ $i = 1, \ldots, m$ are a simple-partition of Ω (see figure 3.25).

V-boundary partition

To get a V-boundary using the SRGPAD, a boundary region, X_b, is added in such a way that its ZI is always empty. For all the regions except the boundary region, their ZI are localised on the outer boundary region excluding all the regions: $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j)$. In the implementation, the call of the method, r.growth(x, i), is substituted by

- if there is more than two ZI on x, then growth on x of the boundary region,
- otherwise growth on x of the region i

At the end of the growing process, the regions $X_i^{t=\infty}$ $i = 1, \ldots, m$, and X_b are a V-boundary-partition of Ω. The algorithm 8 is an example (see figure 3.26).

Algorithm 8 Geodesic dilation with a boundary

Require: S , V *// the seeds, the neighbourhood*
```
    //initialisation
    Region r;
    Restricted N = ℕ;
    ZI zi;
    Metric m (0); //δ(x,i) = 0,
    SQ sq( FIFO, 1);//1 FIFO queue such as all pixels with the corresponding labels of the ZI are placed
    in the same queue
    Population pop (r,zi,m,sq);
    //create a passive region
    Neighbourhood V_void = ∅
    int region_boundary = r.add_region();
    zi.add_ZI(region_boundary,V_void ,ℕ) ;// Z_i^t = ∅
    for all for j = 0 to n do
      int region_label = r.add_region();
      zi.add_ZI(region_label,V,ℕ) ;// Z_i^t = (X_i^t ⊕ V) \ (⋃_{j∈ℕ} X_j^t)
      r.init_seed(region_label, s_i);
    end for
    //Growth process:
    while sq.empty(0)==false do
      (x,i)=sq.pop_pair(0);
20:   if sq[x].size()>=2 then
        r.growth( x, region_boundary );
      else
        r.growth( x, i );
      end if
    end while
    return r;
```

The partition depends on SRIO

Whatever the growing process, the final partition depend on the SRIO (see figure 3.27 and figure 3.28). If the order of initialisation is changed, the localisation of the inner border of each region will also change. The next subsection will present a growing process independant of the SRIO.

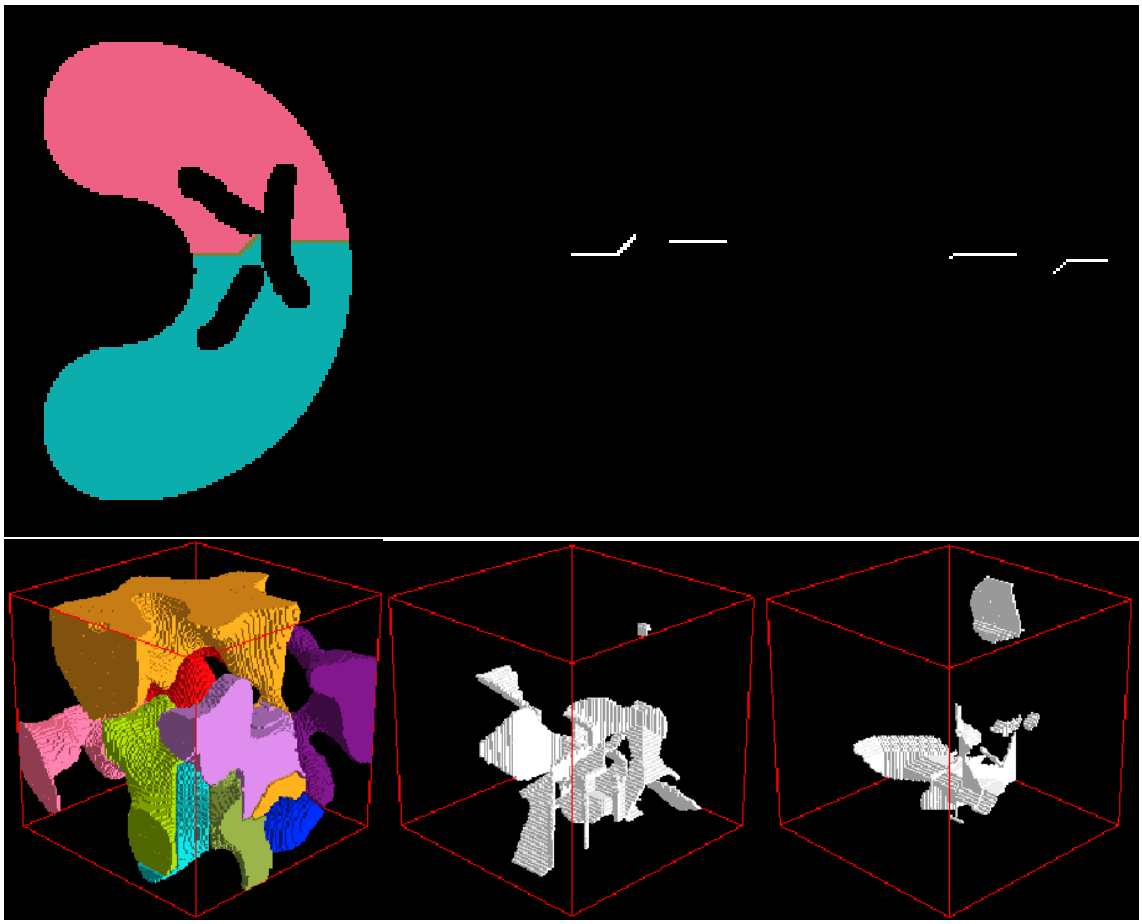


Figure 3.26: Illustration of V-boundary partition. For both rows, the first image is V-boundary-partition obtained by the geodesic dilation with a boundary, the second figure and third figure are the visualisation of the boundary region depending on the chosen neighbourhood. For the second figure, it is the 8-neighbourhood in 2D and 26-neighbourhood in 3D and for the third figure; it is 4-neighbourhood in 2D and 6-neighbourhood in 3D.

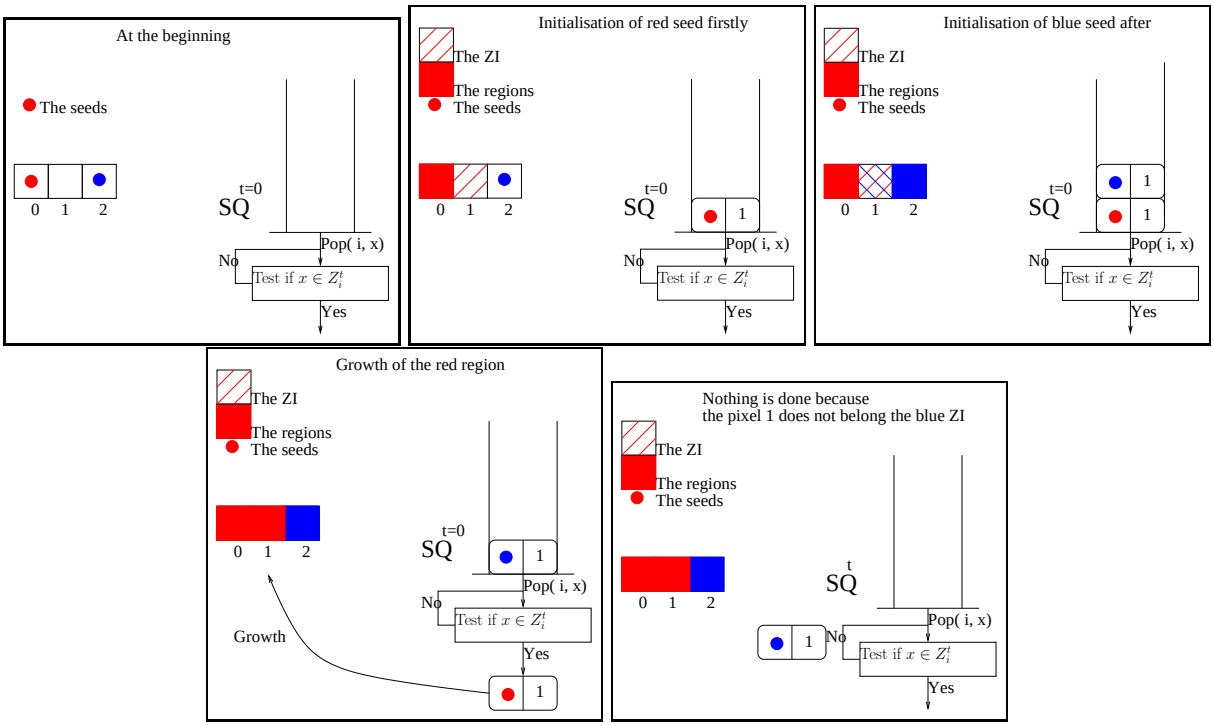


Figure 3.27: The geodesic dilation without a boundary region separating the growing regions. The point 1 is an ambiguous point for this growing process because it belongs to the region that was initialised first (here the red region).

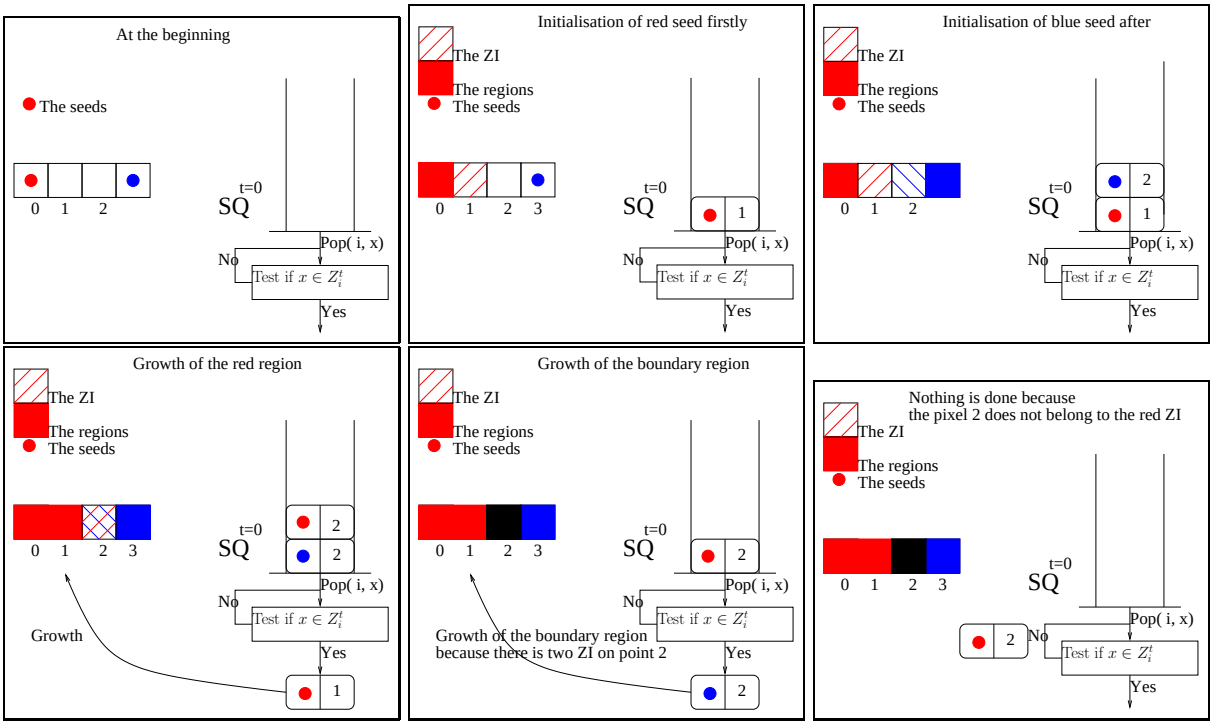


Figure 3.28: The geodesic dilation without a boundary region separating the growing regions The point 1 and 2 are ambiguous pixels in this growing process because they belong to different regions depending on SRIO. In this presented case, the point 1 belongs to the red region and the point 2 belongs to the boundary region but if the blue region was initialised first, the point 1 would belong to the boundary region and the point 2 would belong to the blue region.

3.5.2 Partition independent of the seeded region initialisation order

Why the dependence?

Let $S = (s_i)_{1 \leq i \leq n}$ be a set of subset of Ω. In a metric space M, the cells are

$$C(s_i) = \{y \in M : d^{\Omega}(y, s_i) < d^{\Omega}(y, s_j), s_j \in S\}, s_i \in S,$$

The $C(s_i)_{1 \leq i \leq n}$ is not a partition of Ω because $\cup_{i=1}^{n} C(s_i)) \neq \Omega$. In fact, it is possible to demonstrate that $(\cup_{i=1}^{n} z_{\Omega}(s_i)) \uplus A = \Omega$ [33]. The set A, called ambiguous points, is

$$A = \{\forall x \in \Omega : \exists i \neq j, \big(d^{\Omega}(x, s_i) = d^{\Omega}(x, s_j) \text{ and } (\forall k \in (1, \ldots, n) : d^{\Omega}(x, s_i) \leq d^{\Omega}(x, s_k)\big)\}$$

The set A, called the ambiguous points, contains all the points of Ω for which the geodesic distance of $d^{\Omega}(x, s_i)$ and $d^{\Omega}(x, s_j)$ are equal and $d^{\Omega}(x, s_i)$ and $d^{\Omega}(x, s_j)$ are the lower bound of all the geodesic distance with other elements of S. In the previous implementations of the geodesic dilation, the ambiguous points are distributed depending on the SRIO (see figure 3.27 and 3.28). The next paragraph presents an implementation for which the boundary region is the set of ambiguous points.

Boundary as ambiguous points

We suppose in this paragraph that the seeded initialisation follows this order $0, 1, \ldots, n$. In order to get a boundary localised on the ambiguous points, a boundary region is added such that its ZI is always empty. For all the regions except the boundary region, their ZI are localised on the outer boundary region excluding all the regions: $Z_i^t = (X_i^t \oplus V) \setminus (\bigcup_{j \in \mathbb{N}} X_j)$. When a pair (x, i) is extracted from the SQ, one runs (see figure 3.29,3.30):

1. r.growth(x, boundary region)) if there is more than two ZI in x and if i = min_elements(zi[x]),
2. r.growth(x, i)) otherwise

This partition is independent of the SRIO but is not a V-boundary-partition (see figure 3.29). Since the space is a discrete space, the boundary region is not a closed surface in 3D or a closed line in 2D (some holes on the boundary region). Hence, the partition obtained by this growing process is a simple partition.

How to choose the convention?

Depending on the application, a convention is choose. For instance,

1. for the splitting of grains (see subsection 4.4.1), we use a V-boundary partition in order to separate the grains by a tinny gap,
2. for the study of the evolution of the cement paste microstructure, we choose a partition without a border to avoid the affection of the boundary reagion, but, the SRIO follows at each segmentation the same order.

3.6 Conclusion

In the conceptualisation part, we defined four classes and one procedure to create a library, called Population, dedicated to the implementation of the SRGPAD algorithms. The zone of influence, associated with each region, localises a zone on the outer boundary region. The system of queues organises the pixel by pixel aggregation/dissolution around a concept of store. A procedure was defined to actualise at constant time the zones of influence after a aggregation/dissolution of a pixel of a region. This procedure and these objects permitted to the library Population to be numerically efficient and to implement algorithms

[33] The symbol $\uplus$ means the disjoint union:$B \uplus C = \{B \cup C : B \cap C = \emptyset\}$

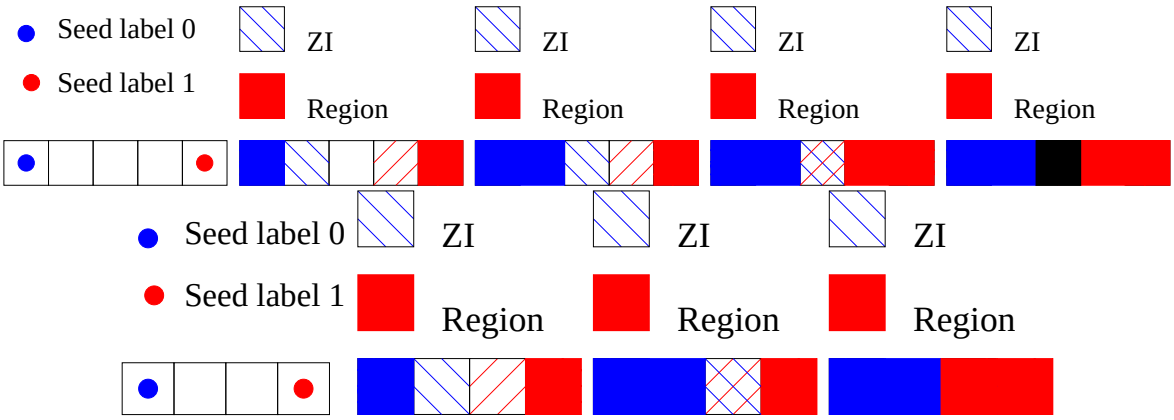


Figure 3.29: Illustration of the independence of the SRIO if the boundary is the ambiguous points. The first row is the case of one ambiguous point. A classical growing process goes on until there appear two ZI in the same point, x. The min_elements(zi[x]) returns 0 because there are two ZI with label 0 and 1. The boundary region grows because the pair extracted from the queue has a label 0 equal to min_elements(pop.Z()[x]). The second row is the case without ambiguous point. A classical growth goes on until there appear two ZI in the same point, x. The min_elements(pop.Z()[x]) returns 0 because there are two ZI of label 0 and 1. As the pair extracted from the queue has a label 1 not equal to min_elements(zi[x]), there is the region growth of label 1

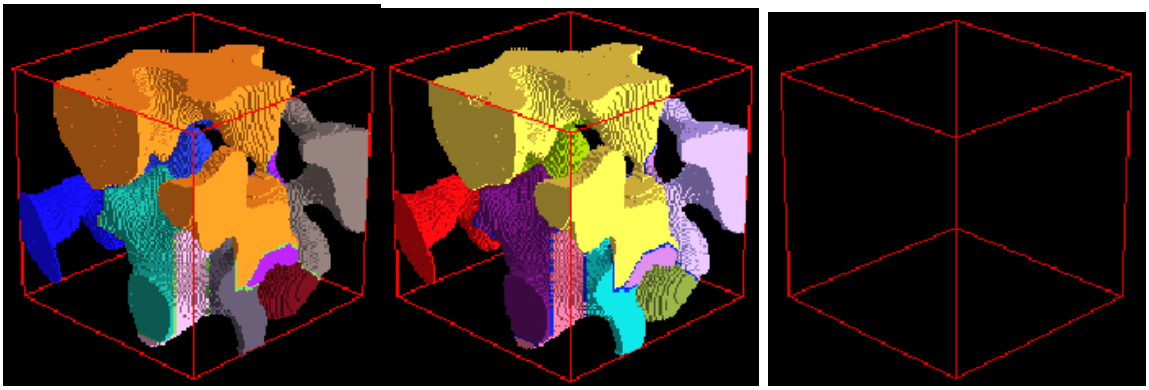


Figure 3.30: Illustration of the independence of SRIO in the case where the boundary is the ambiguous points. The two first images are the geodesic dilation with a boundary region localised on the ambiguous points such as the SRIO is different. The third image represents the XOR the boundary regions difference between the two previous images. It is empty image since the regions localisation is independent of the SRIO.

rapidly.
In the algorithmic section, we implemented various SRGPAD algorithms. Each implementation is simple and efficient using the library Population. When the growing process is done at constant velocity without memory (simulated Voronoï tessellation, domain to clusters, regional minima), a single queue is sufficient to implement these algorithms. When the growing process depends on the topographic surface (watershed transformation and dynamic filter) or when information has to be kept during the growing process (distance function), one needs several queues to implement these algorithms.
In discrete space, the boundary is not clearly defined. Using the SRGPAD, we have proposed two growing processes to obtain a simple or V-boundary partition. These growing processes have an uncertainly on the boundary region localisation. To overcome this problem, we have defined a set of ambiguous points

for which it is impossible to know to which regions they belong. Knowing that, we have defined a growing process with a boundary region localised on these ambiguous points. The partition associated to this growth process is independent of the SRIO but it is only a simple partition since there are some holes on the boundary region.
In the three next chapters, we will apply these algorithms for segmentation, decomposition in elementary pores and 3D reconstruction.

4

Segmentation

Contents

Introduction

A grey-level image is a grid of a sample where the value of each cell (pixel/voxel) is the average of a physical interaction (e.g. attenuation of X-rays for tomography or emission of back scattered electrons for SEM) between different solid and fluid phases contained in the cell and the excitation source, at the moment of acquisition. A labeled image is a grid of the sample where the label of each cell corresponds to a phase of the material. Segmentation is the transformation of the grey-level image to the labeled image. The best segmentation tool is the human eye (see figure 4.1) and, whatever the algorithm, the computer-assisted segmentation will almost always be less efficient[34]. The aim of this chapter is to propose a *simple*, *generic* and *robust* method to segment experimental two- or three-dimensional images of samples

[34] During my PhD, I met many colleagues unfamiliar with Image Processing who asked me to segment the images obtained by various technique like atomic force microscopy. They do not often understand that this task is difficult since their eyes segmented the images easily.

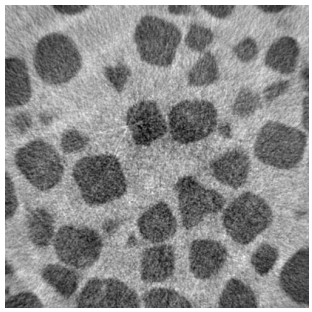


Figure 4.1: Our eyes can draw easily a line between the two phases, but numerically, it is not so easy [165]. reproduced with permission from Alain Fanget (DGA).

obtained by SEM and X-ray tomography. *Simple* means that this method can be used by anyone whether or not they are a specialist in image processing. *Generic* means that this method remains valid for a wide range of materials. *Robust* means that the extraction is less sensitive to a perturbation of the segmented parameters.
The methodology is:

- If there is a sufficient contrast-to-noise ratio between the different phases, a classical threshold procedure followed by a succession of morphological filters is applied,
- If not, and if the boundary needs to be located precisely, a watershed transformation controlled by seeds is applied. The seed localization is the basis for number of algorithms involving the tint propagation [2], the active contour methods [160], the watershed transformation [20], etc. For our purpose, each seed localization has to respect two constraints: inclusion (the seed is included in the phase) and hitting (the seed intersects all the connected components of the phase). Using this seed localisation, two different chain-linking methods are proposed:
  - If a seed can be localised for each phase, a one-step method is applied to extract all phases within the same transformation.
  - If a seed cannot be localised for one phase, a step-by-step method is applied to extract a phase at each step (the last step is skipped since the last phase is the complementary of the sum of the other phases).

This segmentation has been tested for various complex porous media and granular materials, and allowed the prediction of various properties (diffusion, electrical conductivity, deformation field). These numerical results have been validated by a good agreement with experimental data.
Whatever the algorithm, some artefacts are inherent to segmentation:

1. in tomography, some solids are completely surrounded by the liquid phase (levitation),
2. in SEM, some holes in the grain phase have been induced by the polishing process.

A hole-filling algorithm is applied to correct these artefacts.
For many applications, the extraction of each grain of the grain phase, called grain partition, has to be performed. The basic algorithm is the extraction of connected components in the grain phase. However, this algorithm will not yield a consistent grain partition since many grains close to each other are connected on the segmented grain phase. The method, introduced by L. Vincent in 1993 [172], results will lead to an efficient grain partition.
In this chapter, we describe in detail the above mentioned algorithms, with a focus on the segmentation of three-dimensional images of materials obtained by X-ray tomography.

4.1 Materials and methods

4.1.1 Image characteristics

Noise

Each voxel of a microtomography image is associated with a cube included in the sample under investigation. To a first order, its grey-value (its tint) is the space average of the linear X-ray absorption coefficient of the different solid and fluid phases contained in the cube. But, since tomographic reconstruction often amplifies the noise of the projections and generates artefacts, there are extra-terms which give imprecise images that are generally weak quality for threshold segmentation. Due to the large variety of materials[35] and the imprecise images, a generic, simple and robust segmentation procedure has been developed.

Partial-volume effect

Because each pixel/voxel in a microscopy image represents the attenuation properties of a specific material volume, if this volume contains different phases then the resulting value represents an "average' of their properties [85]. This is called the partial-volume effect. During segmentation, each pixel/voxel is affected to a phase. Therefore, if a pixel/voxel contains more than one phase,-the partial-volume effect-, its affectation will be ambiguous, whatever the algorithm. Next, we will quantify this effect for a boolean model related to the characteristic size of the boolean model and the image resolution.

Boolean model We consider, as germs ϕ, the stationary Poisson point process in $\mathbb{R}^2$ with an intensity measure $\lambda\mu(.)(\lambda > 0)$ with μ the Lebesgue measure:

$$\phi = \{X_i, i \in I\}$$

and as grains K_0, the closed ball of radius r using the norm-∞.
Then, we consider Boolean Model such as the random set in $\mathbb{R}^2$ is

$$M = \cup_{i \in I}(X_i \oplus K_0)$$

and the Boolean random field, I, is:

$$I(x) = \mathbf{1}_{\cup_{i \in I}(X_i \oplus K_0)}$$

where $\mathbf{1}$ is the indicator function.

Discretisation using a tessellation To represent the resolution limitation, we transfer the continuous model into discrete counterpart. A tessellation or tiling of the plane is a collection of plane figures that fills the plane with no overlaps and no gaps. Here, we consider a squares regular tessellation (square grid) defined as follows:

$$\{x^r_{i,j}; i, j \in \mathbb{N}\}$$

with $x^r_{i,j} = [ir, jr[[(i+1)r, (j+1)r[$ and r the resolution.
The intensity of each pixel of the regular tessellation is defined by the intersection surface between the random set of the boolean model and the pixel normalised by the surface of the pixel:

$$\begin{array}{ccc} I : \mathbb{N}^2 & \longrightarrow & [0,1] \\ (i,j) & & \dfrac{\mu(x^r_{i,j} \cap M)}{\mu(x^r_{i,j})}. \end{array}$$

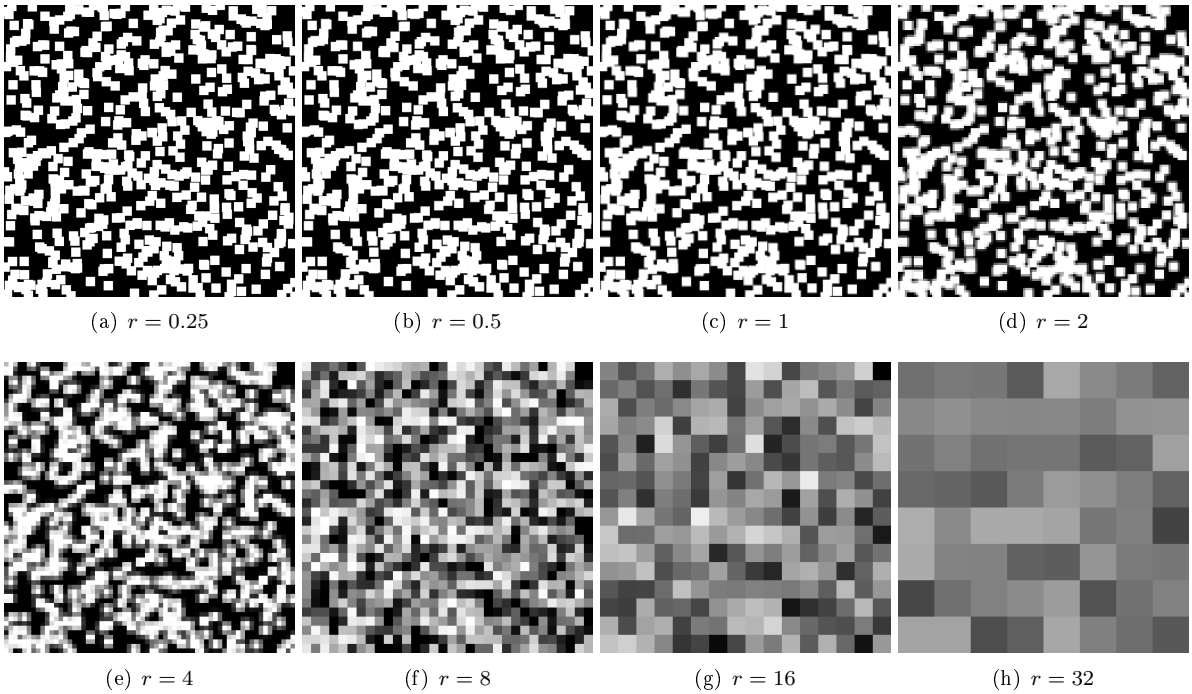
A quantification of the resolution error of the discrete model can be the probability, $P(0 < I(i,j) < 1)$. This is the probability that a random pixel contains two phases since its value is different to 0 and 1. Therefore, this probability quantifies the occurrence of the partial-volume effect. When r tends to 0, this probability tends to 0 and the discrete model tends to the continuous model. When r tends to ∞, this probability tends to 1 and the value of each pixel tends to $\lambda\mu(K_0)$ given the coarse graining description.

[35] Materials are different in their chemical composition and in their geometrical organisation.

Application In this example, the intensity of the boolean model is 0.01 and the grain radius is 4 ($\lambda\mu(K_0) = 0.64$). The characteristic size of the boolean model is in order of 4. The figure 4.2 shows the discretization of boolean model for different resolutions. As illustrated in table 4.1, the partial volume effect can only be neglected (error$<$2%) if the ratio between the characteristic size and resolution is superior to 32. For a ratio equal to 4, the error is superior to 7.5%. The extrapolation of this model on real images leads to the following conclusion: if the resolution of the imaging technique is in the same order of the characteristic size of phase under investigation, segmentation will not yield an accurate phase microstructure whatever the method. This is the case of the capillary porosity of mature paste.

Resolution	Characteristic size/Resolution	$P(I(i,j)=0)$	$P(0<I(i,j)<1)$	$P(I(i,j)=1$
0.125	32	0.513593	0.0200	0.466231
0.25	16	0.486389	0.0407038	0.47290
0.5	8	0.451744	0.0775146	0.470741
1	4	0.361465	0.150116	0.488419
2	2	0.251465	0.237793	0.510742
4	1	0.0939941	0.271729	0.634277
8	0.5	0.00878906	0.55957	0.431641
16	0.25	0	0.839844	0.160156
32	0.125	0	1	0

Table 4.1: The grain radius is equal to 4.



(a) $r = 0.25$ (b) $r = 0.5$ (c) $r = 1$ (d) $r = 2$

(e) $r = 4$ (f) $r = 8$ (g) $r = 16$ (h) $r = 32$

Figure 4.2: The image size is equal to $256r \times 256r$ with r the pixel size.

4.1.2 Materials

For the material A, the data come from a mechanical triaxial test on a sand specimen realised under a synchrotron microtomography (ESRF, ID15A) to follow the structural evolution of the granular media. Digital Image Correlation is used to observe and detect the strain localisation mechanisms at the grain scale [62, 101]. This work is funded by the French project ANR-05-BLANC-0192 (see figure 4.3.a).

For the material B, the data come from pyrotechnical specimen realised under a laboratory microtomography designed by Skyscan®. Finite elements are used to compute the stresses and strains as well as other fields like the thermal flux and the temperature distribution in the material at a micro scale [80]. This work is funded by les Mines de Paris and le Centre d'Etudes de Gramat, DGA (see figure 4.3.b).
The material C is a cement paste presented in the chapter 2 (see figure 4.3.c).
For the material D, the data come from geological rock, limestone, realised under a laboratory microtomography designed by Phoenix X-ray to understand the effects of the porous structure on the resistivity index curves [63] (see figure 4.3.d).

4.1.3 Computational requirements

The segmentation procedure has to be efficient in terms of computational time because the size of the images obtained by X-ray tomography is very large. In the previous chapter, it was shown possible to improve the efficiency of the algorithms used for this purpose. Typically, for the largest image of this article, the material D with a size $700 * 700 * 700 = 0.348 \cdot 10^8$ voxel, the segmentation requires less than 6 hours of executing time with an Intel(R) Xeon(R) CPU 3.00GH and the allocation of RAM is 16 Gb. This short executing time allows the segmentation to be performed on conventional computers. The algorithms are developed and implemented on the open source software, called Population, soon available on the web. For every image, the grey-level is coded on one byte (0-255). A median filter is applied to minimise the ring artefact [16] and to smooth the noise while still in keeping the sharpness of the boundary. For the visualisation convenience, the results are sometimes presented in 2D but the method has been applied in 3D images for all materials.

4.2 Threshold segmentation using tint information

To the best of our knowledge, threshold segmentation using tint information is the only method applied to extract the different phases for cementitious materials [120, 123, 137, 182, 181]. These articles focus on how to find the right grey-level range for thresholding.

4.2.1 Threshold

Threshold operator

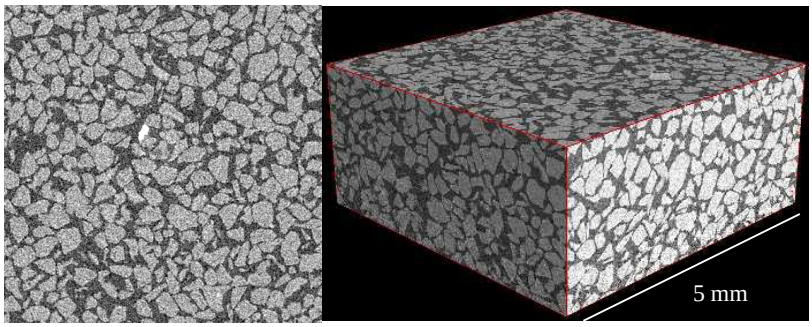
Given that each phase has a specific tint, the threshold operation uses this information to extract the phases. The threshold operation requires the selection of a range of grey-levels. The label '1' is assigned to each voxel of the image when the grey-level belongs to this range, and the label '0' otherwise. The grey-level range selection is usually based on the information contained in the grey level histogram of the image (see figure 4.4). The automatic or manual grey-level range is selected in order to best separate the mode in the histogram [36, 132, 143, 176]. One requirement is that the mode is populated almost always by the given phase.

n times

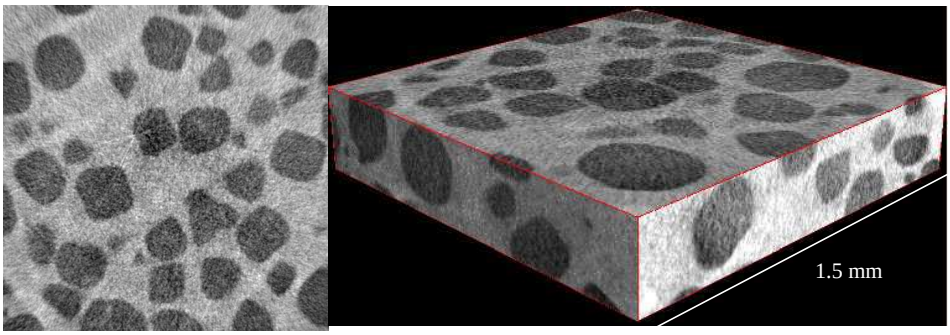
For a n-phases material, n modes have to be present in the histogram in order to operate the threshold segmentation. Under this assumption, n modes give $n-1$ valleys, with value $v_1, \ldots, n_n$ and $v_i < v_{i+1}$. Manually, the threshold operator is applied n times with these grey-level range selections: $[0, v_1], [v_1, v_2], \ldots, [v_n, 255]$.

Results

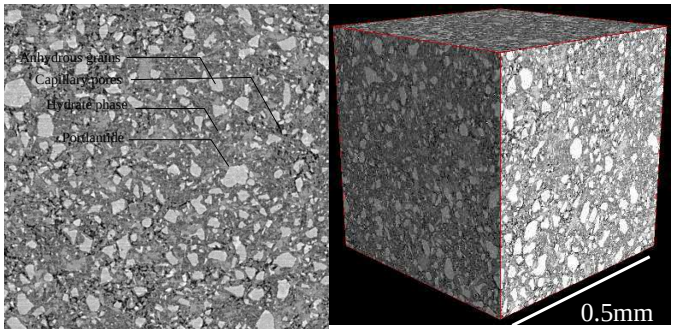
For the two phases of the materials A and B, the figure 4.4-a-b shows two modes in the histogram,-a specific mode for each phase-. For both, the value of the valley is 125. Therefore, to extract the white grains in the material A, the grey-level range is $[125, 255]$ and to extract the black grains in the material B, the grey-level range is $[0, 125]$. Despite some holes in the numerical grains and some isolated islands



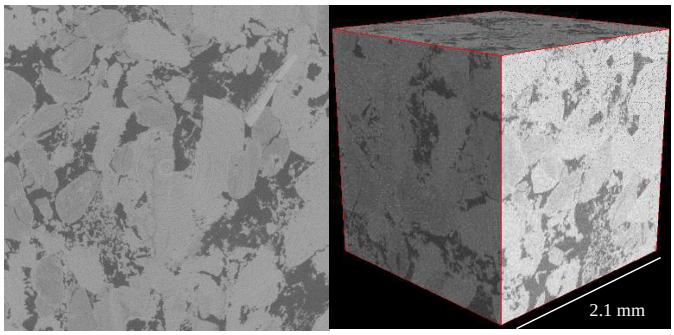
(a)



(b)



(c)



(d)

Figure 4.3: A wide range of materials. (a) material A, size=450x450x200, resolution=14 microns; (b) material B, size=500*500*100, resolution=3 microns; (c) material C, size=700x700x700, resolution=0.675 microns, there are four phases: the anhydrous grains in light, the portlandite (some isolate clusters) in light grey, the hydrate in grey and the capillary porosity in black; (d) material D, size=700x700x700 resolution=3 microns, there are three phases: the void in dark and two grain classes (one lighter and the other with a medium average grey level).

outside the grains, the numerical segmentation matches the visual segmentation. The aim of the next subsection is to remove these both effects by Morphological filtering.

4.2.2 Morphological filtering

Presentation

The four basic operators of the mathematical morphology [158] are:

1. erosion of the set A by the set B, called structural element, is defined by[36]:

$$A \ominus B = \{\forall z \in E : B_z \subset A\}$$

   where B_z is the translation of B by the vector z (see figure 4.5),

2. dilation of object A by the structural element B, also called the Minkowski addition, is defined by:

$$A \oplus B = \{\forall z \in E : (B^s)_z \cap A \neq \emptyset\}$$

   where B^s denotes the symmetric of B, that is, $B^s = \{x : -x \in B\}$ (see figure 4.6),

3. opening of A by B is obtained by the erosion of A by B, followed by dilation of the resulting structure by B:

$$A \circ B = (A \ominus B) \oplus B,$$

4. closing of A by B is obtained by the dilation of A by B, followed by erosion of the resulting structure by B:

$$A \bullet B = (A \oplus B) \ominus B,$$

Application

After the application of the threshold operator, two classical artefacts are to be processed in the binary image (see figure 4.4-b):

- some holes,
- some isolated islands.

To remove isolated islands without changing the size of the clusters, the opening filter is applied [37]. To fill the holes without changing the size of the clusters, the closing filter is applied. Let A be the binary image after the threshold application. The filtering is only: $(A \circ B) \bullet B$ (see figure 4.7). The figure 4.8 shows a agreement between the visual segmentation and the numerical segmentation but the numerical boundary is not closely located on the visual boundary and some grains close to each others become connected.

[36] For visualisation, the inner boundary is defined as: $\partial A = A \setminus (A \ominus B)$

[37] As all the materials are isotropic, the structural element is chosen isometric. The structural element is associated to the 26-connectivity in the cubic grid

$$B = \{(-1,-1,-1),(0,-1,-1),(1,-1,-1),(-1,0,-1),(0,0,-1),(1,0,-1),(-1,1,-1),(0,1,-1),(1,1,-1),$$
$$(-1,-1,0),(0,-1,0),(1,-1,0),(-1,0,0),(0,0,0),(1,0,0),(-1,1,0),(0,1,0),(1,1,0),$$
$$(-1,-1,1),(0,-1,1),(1,-1,1),(-1,0,1),(0,0,1),(1,0,1),(-1,1,1),(0,1,1),(1,1,1)\}$$

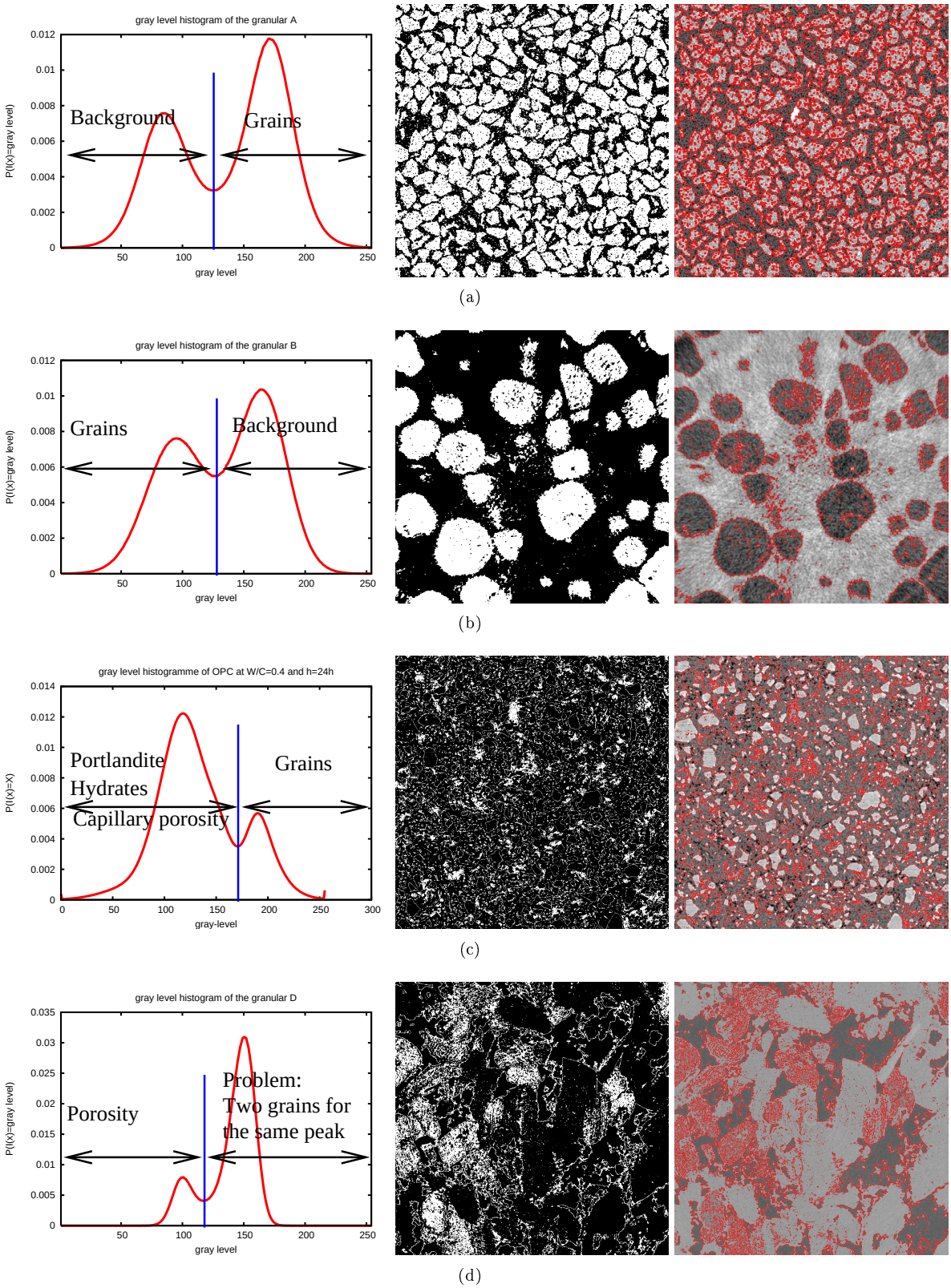


Figure 4.4: Each row is associated with a material, the first image is the histogram, the second figure is the binary image after thresholding, and the third figure is the visualisation of the numerical boundary on the initial image.

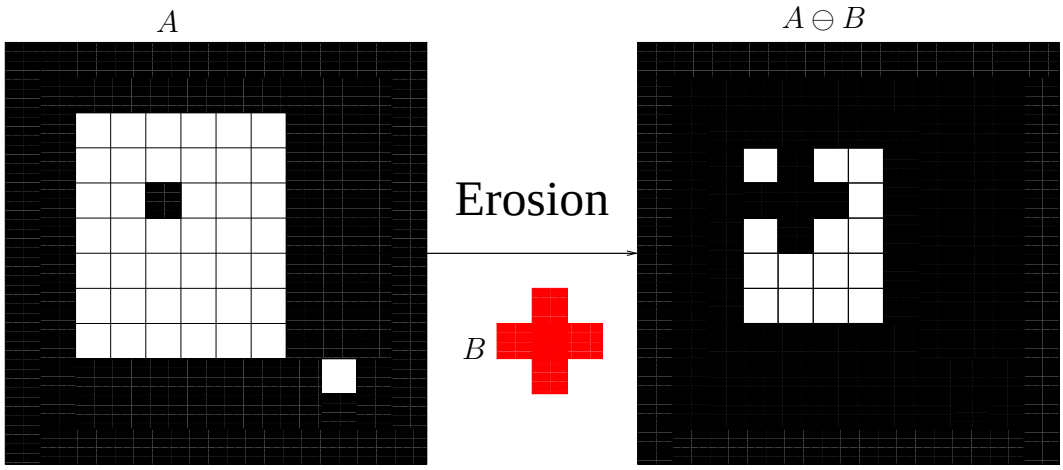


Figure 4.5: Erosion. The set A is the set of white pixels and the structural element B is: $\{(0,0),(-1,0),(1,0),(0,-1),(0,1)\}$ representing by the red cross. The set, $A \ominus B$, is each pixel, z, of the images such that **all** the pixels are whites at on-up-down-left-right of z: $B_z \subset A$.

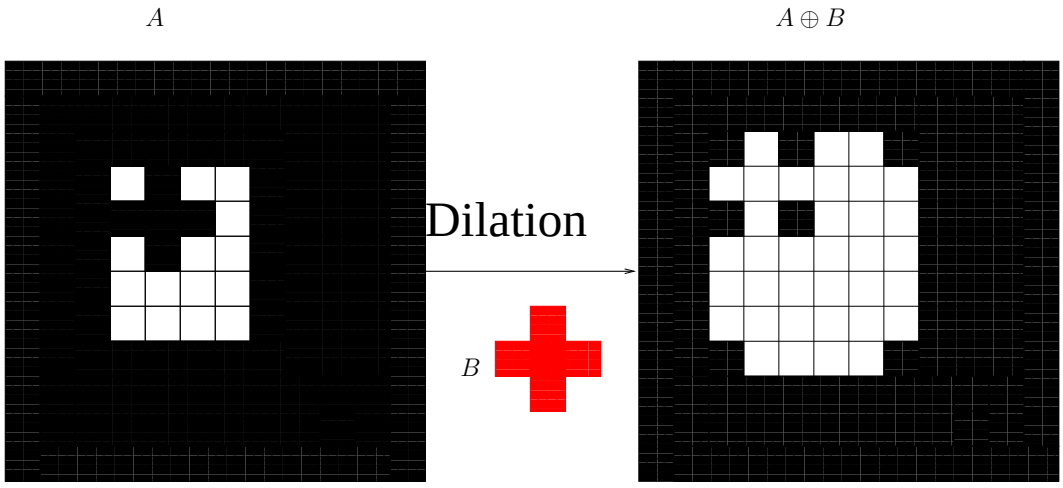


Figure 4.6: Dilation. The set A is the set of white pixels and the structural element, B, is: $\{(0,0),(-1,0),(1,0),(0,-1),(0,1)\}$ representing by the red cross. The set, $A \oplus B$, is each pixel, z, of the images such that **at least one** pixel is white at on-up-down-left-right of z, $B_z \cap A \neq \emptyset$.

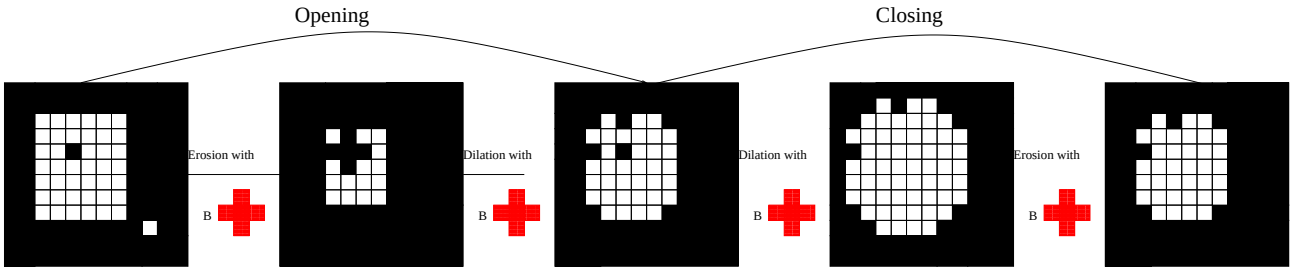


Figure 4.7: Opening then closing. At the beginning, the set A has one hole and one isolated island. After the application of the opening, the isolated island has been removed. After the application of the closing, the hole has been removed.

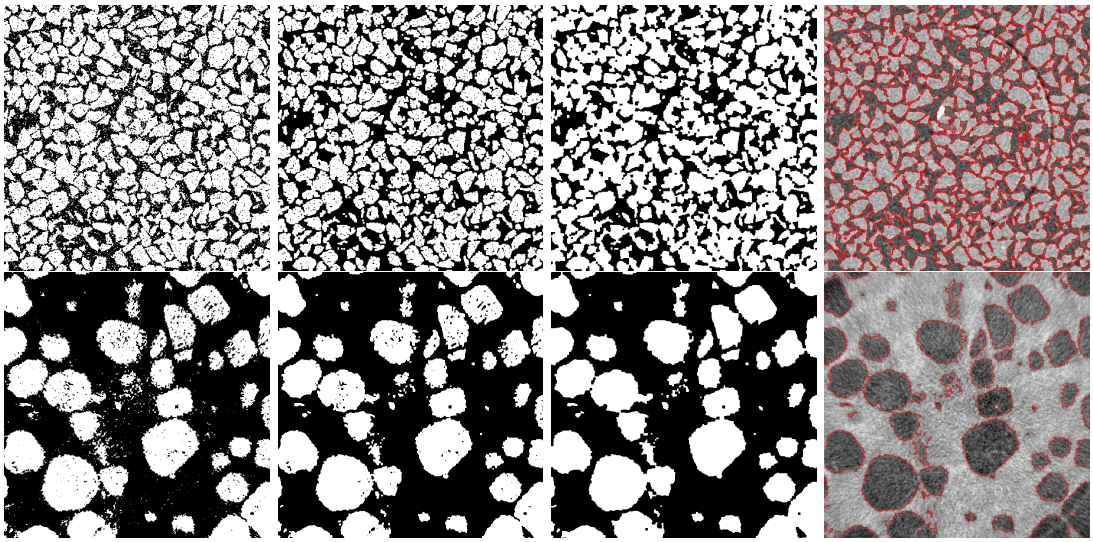


Figure 4.8: For both rows, the binary image is the result of a threshold, the second image is obtained by the application of opening on the first image in order to remove the isolated islands, the third image is obtained by the application of closing on the second image to fill the holes and the last image is the numerical boundary visualisation on the initial image.

4.2.3 Limitation of the threshold segmentation

For the material C, the histogram presents only two modes in the histogram (see figure 4.4.c). The right mode is almost only populated by the anhydrous grain phase, but the left mode is populated by three phases: capillary porosity, hydrates and portlandite since the contrast-to-noise ratio is weak. With the manual grey-level range $[140, 170]$ (see figure 4.4-c), the weak contrast to noise ratio leads to many holes on the portlandite clusters (missed pixels) and many isolated islands outside the clusters (added pixels). Moreover, the boundary of the anhydrous is extracted due to a halo artefact. Due to both these artefacts, an accurate segmentation is impossible even with a post-filtration. This is the same for the material D (see the histogram in the figure 4.4.d)).
Wong *et al* have introduced an automatic method for the selection of the grey-level range when a mode is populated by two phases and, as noted by the authors, this method is very sensitive to the grey-level range [181]. But, moreover, whatever the grey-level range, the segmented microstructures obtained with the threshold operation, even after a post-filtration, will not match the real microstructure since the contrast-to-noise ratio is weak (see figure 4.9). Therefore, the simulation of physical phenomenon like vapour diffusivity, air permeability or electrical conductivity in this segmented microstructure will not yield

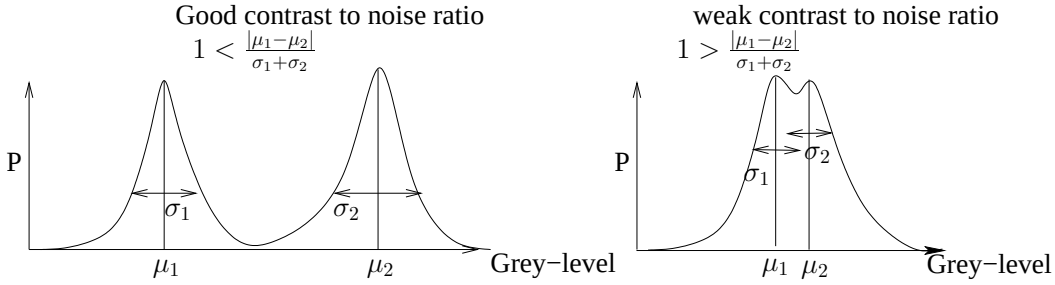


Figure 4.9: As a model, we assume that the grey-level distribution of each phase follows a Gaussian distribution defined by a mean, μ, and a variance, σ^2 (σ is the standard deviation). In the case of two phases, when the contrast-to-noise ratio is weak, the surface intersection between the two distributions is superior to 4% percent of the sum of the surface of the both distribution ($1 - \text{erf}(2/\sqrt{2}) \simeq 0,0455$ where erf is the error function). Therefore, if the threshold value is chosen on the valley, the segmentation error is superior to 4% such that the segmented error is the ratio between the wrong affected pixels/voxels and all affected pixels/voxels. This error has to be added to the partial-volume effect error. Threshold segmentation generates two kinds of wrong affected pixels/voxels:
1)those missed of the segmented phase given the holes on the segmented image,
2)those added by the other phase given the small islands on the segmented image,
For the materials C and D, whatever the threshold range, the segmented error will be high since there is no even some specific peaks associated with the different phases.

direct measurements. **When the contrast-to-noise ratio is weak, threshold segmentation gives a false microstructure that does not allow the correct prediction of the physical behaviour and properties of the material.** The subsection 4.3 describes a method to handle this task.

4.3 Watershed transformation using boundary information

The threshold segmentation is suitable only if there is a sufficient contrast-to-noise ratio. Otherwise, one solution is to improve the contrast-to-noise ratio by filtering the grey-level in order to reduce the noise. In this article [53], the authors introduced a novel technique for noise reduction based on non-linear anisotropic diffusion. This method leads to a significant gain in the signal-to-noise ratio. This allows the visualization and the analysis of Electron tomography images initially hampered by an extremely low signal-to-noise ratio. However, the calibration of input parameters is quite tedious and depends on the material and the device. Generally, it seems difficult to find a generic filter that reduces the noise for different microscopies since the nature of noise depends on the microscopy. To overcome this, we will introduce a simple, generic and robust method based on a growing process algorithm with seeds as prior.

4.3.1 Seeds-controlled watershed

An efficient segmentation procedure developed in mathematical morphology is the watershed segmentation [23], usually implemented as a flooding process from seeds.
Watershed transformation : Any grey-level image can be considered as a topographic surface and all boundaries as sharp variations of the grey level. When a gradient is applied to an image, boundaries are enhanced. When the topographic surface obtained from the gradient is flooded from its seeds, the waterfronts meet on watershed lines in 2D, and on watershed surfaces in 3D. The investigated volume is partitioned by the watershed basins (see figure 4.10).
Seeds : the result depends strongly on the seeds localisation. For example, if a seed is associated with each minimum of the gradient image, the watershed transformation with these seeds will produce the well-known over-segmentation [165]. There are over minima due to noise and local irregularities in the gradient image (see figure 4.11). To avoid this problem, the image is usually filtered by a combination of

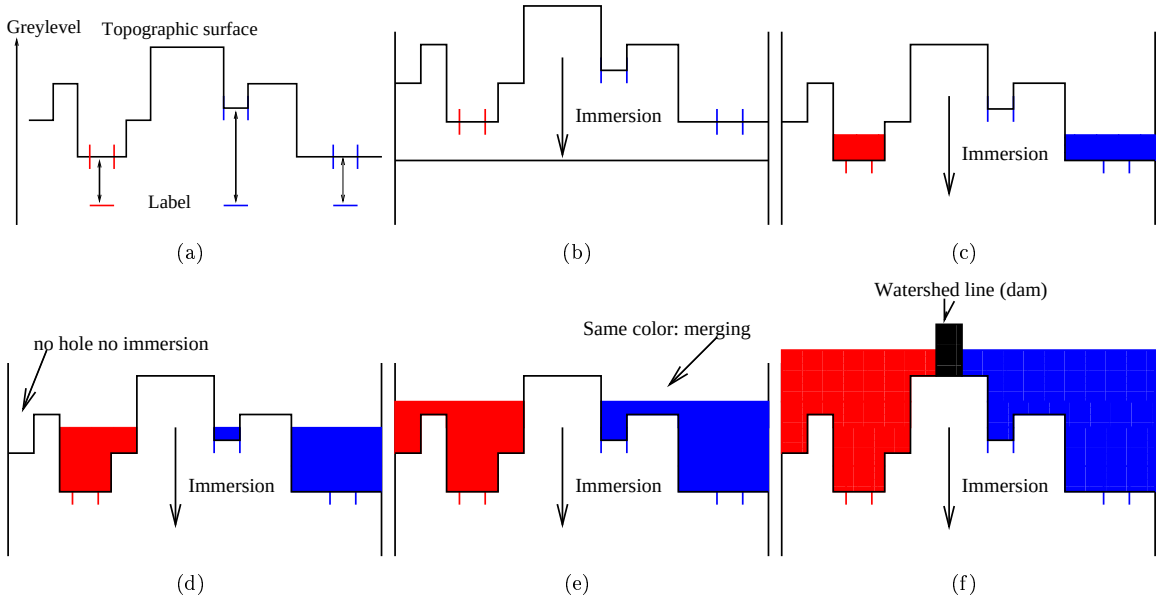


Figure 4.10: Watershed transformation requires two images: the topographic surface (a grey-level image) and the seeds image. The process is: (a) association of each seed to a hole (b) immersion (c) the water enters in the topographic surface by the holes and the basins take the colour of the hole, (d) a part of the topographic surface is not merged although its level is under the level of the immersion, (e) fusion of two basins of same colour, (f) creation of a dam when two basins have different colours: *A video is available at http://pmc.polytechnique.fr/∼vta/water.mpeg.*

a vertical filter (like the dynamic filter [59]) and a horizontal filter (like an alternate sequential filter or a Gaussian filter) in order to individualise each connected component of each phase with a single seed (see figure 4.12) . This individualisation step is complex even with the introduction of the pyramid segmentation [165]. A simple method will be introduced using two approaches: one-step method or step-by-step method. Both these approaches depends on the appropriate seed localization [20, 160].

4.3.2 Seed inside a phase

A seed is a set of pixels/voxels, not necessarily 1-connected. Any localisation of a seed s_i for a phase p_i has to respect these two constraints:

1. the seed is only included in its associated phase,

$$s_i \subset p_i$$

2. the seed intersects[38] each connected component of the phase. Let $(c_{0,i}, \dots c_{n,i})$ be the set of connected components of the phase p_i, then

$$\forall j \in (0, \dots n) : s_i \cap c_{j,i} \neq \emptyset$$

Using the tint information, a 'soft' threshold on the initial image followed by an opening with the structural element B_k [39] is sufficient for this purpose. As shown in the figure 4.13, this filtering removes the

[38] In Mathematical Morphology, the appropriate word is: hit

[39] The opening filter of size k is

$$A \bullet B_k = (A \oplus B_k) \ominus B_k$$

$$\text{with} \qquad B_k = \overbrace{(B \oplus B) \dots \oplus B}^{\text{times k}}.$$

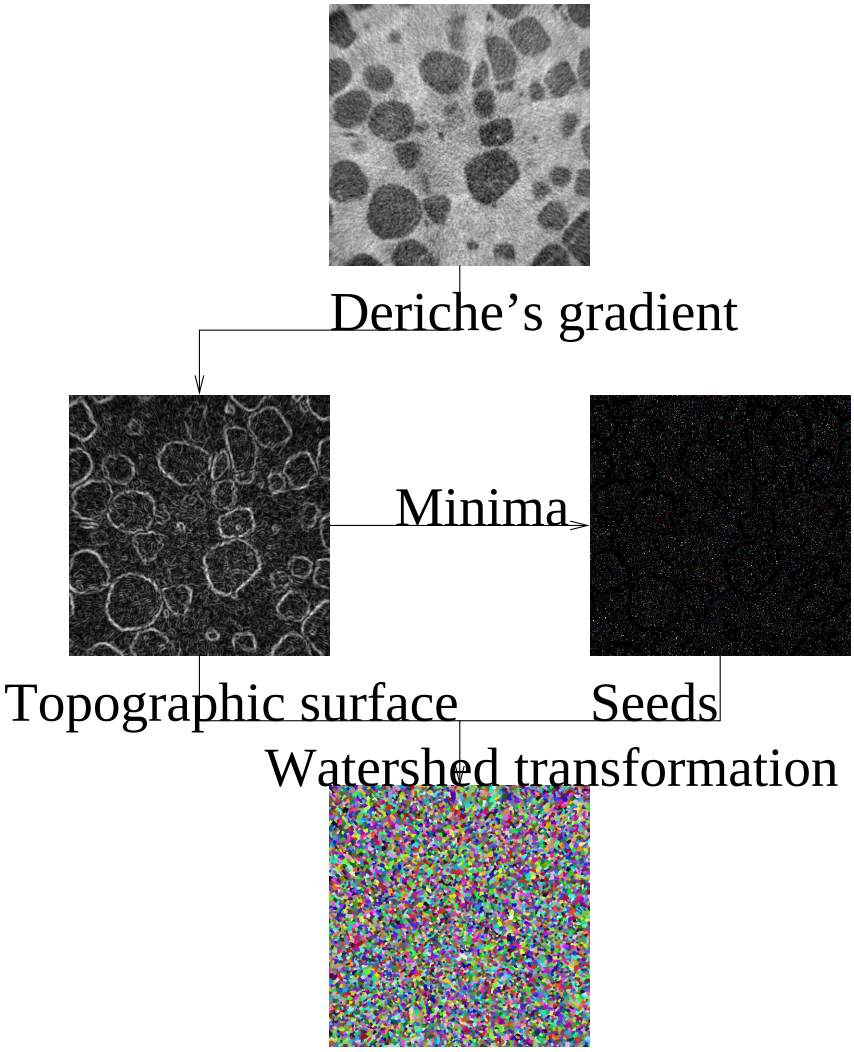


Figure 4.11: Illustration of the over-segmentation. Many basins in the segmented images.

isolated pixels outside the phase and the halo artifact. This filtering can be performed since the constraints of the seed localisation are weak in comparison with the phase-matching for threshold segmentation. The next two subsections explore the possibility to localise a seed inside a phase.

4.3.3 One-step method

The one-step method consists of (see figure 4.14):

- the localisation of a seed inside each phase (n phases = n seeds),
- the application of the watershed transformation[40] to the gradient image with these seeds as prior.

Each basin associated with a seed corresponds to a phase. If each seed localisation respects the constraints, the watershed transformation results in a efficient segmentation of each phase (see figure 4.15

[40] In section 3.5, we have defined different conventions to manage the separation between the growing regions. For this purpose, we use the watershed transformation without a border region in order to avoid the problem of the affectation of the voxels belonging to the dams.

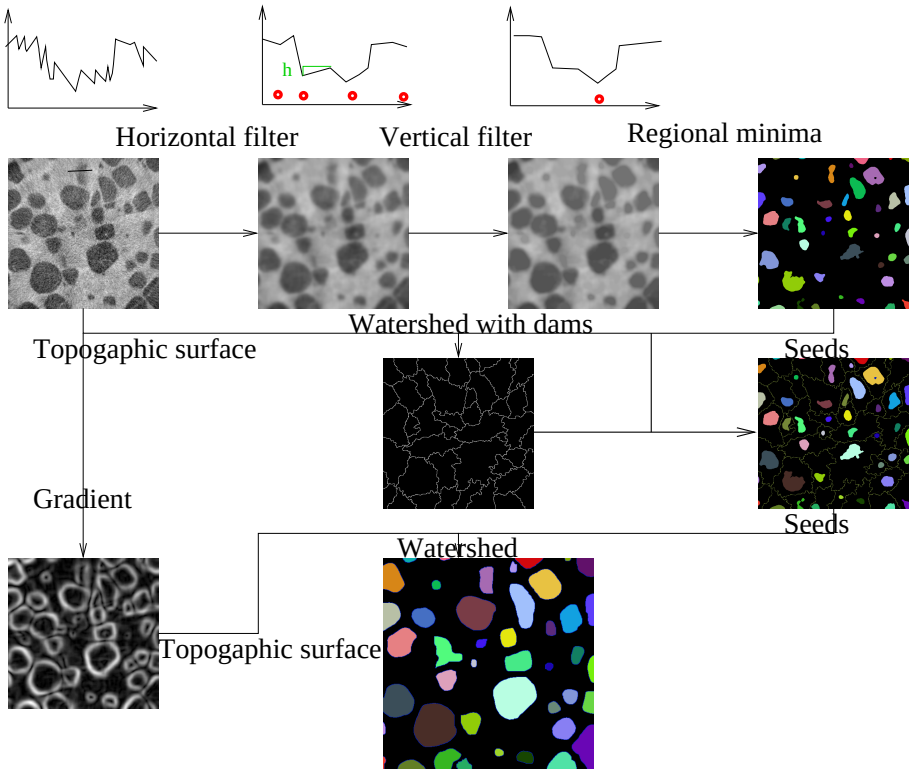


Figure 4.12: Illustration of the swamping segmentation. On the segmented image (the last image), some numerical grains split the real grains and some real grains are missed. To overcome both these artefacts, Tariel and Jeulin [165] have introduced the pyramid segmentation.

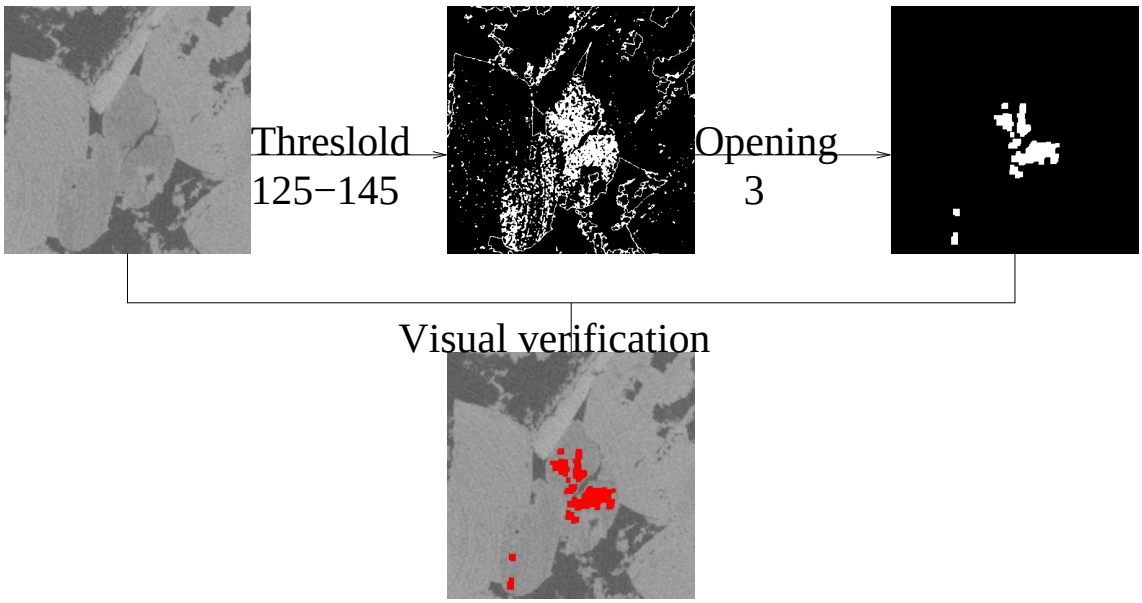


Figure 4.13: Threshold followed by an opening to localise a seed inside a phase.

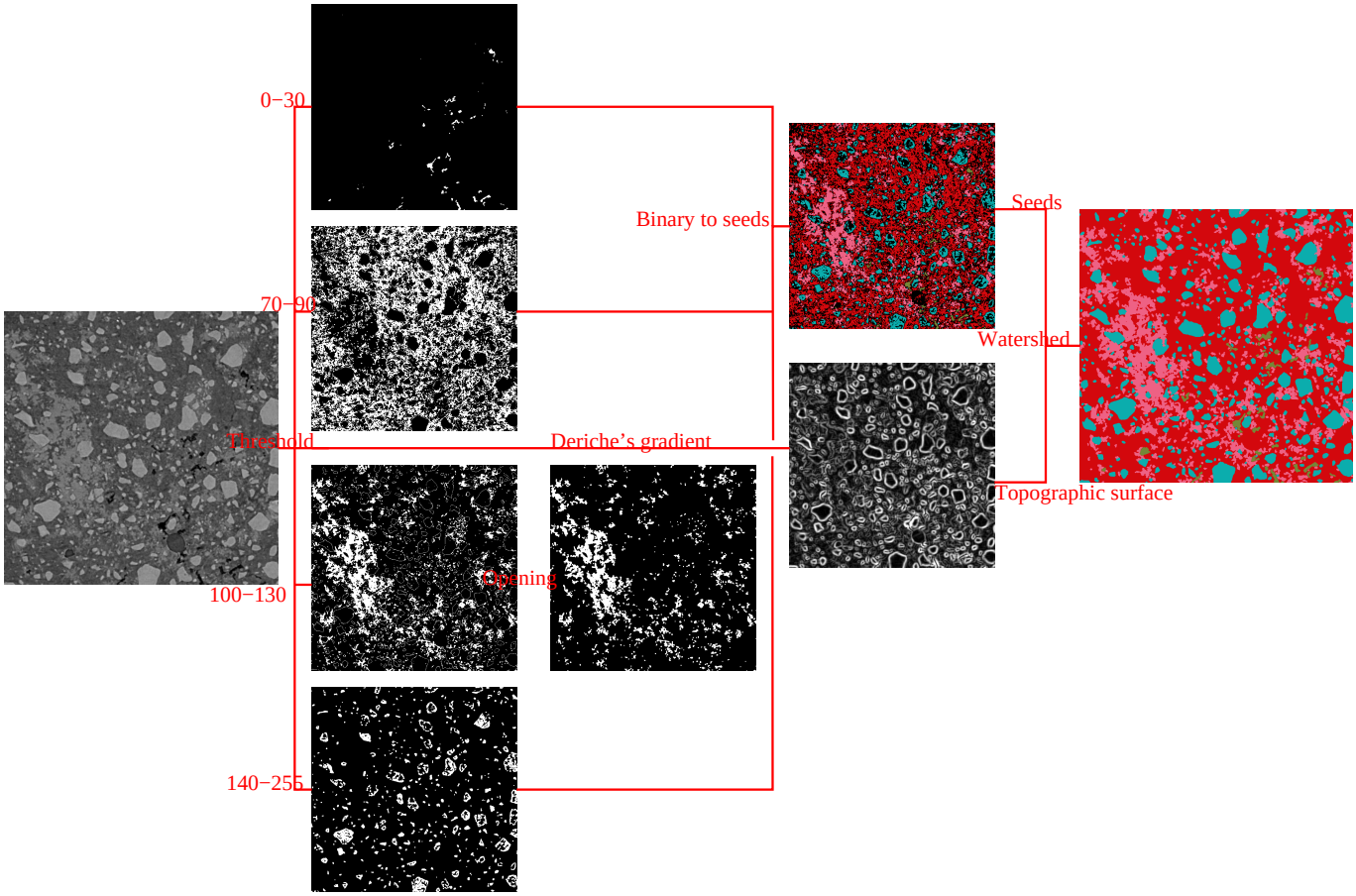


Figure 4.14: The one-step method supposes the localisation of a seed inside each phase. The numerical segmentation matches the visual segmentation.

and table 4.2) with a suitable topography surface (see subsection 4.3.6 for the gradient choice). The major advantage of this approach is that **it is not necessary for a seed localisation to match the phase unlike for threshold segmentation**.

4.3.4 Step-by-step method

For the material D, the constraints of the seed localisation cannot be respected for the light phase. To make its segmentation possible, a step-by-step method is applied. At each step, two seeds have to be localised: one in the phase, another in the phase's complementary. Starting from the simplest phase to extract, we proceed by extracting the next simplest phase step by step. Since the last phase is the complementary of the addition of the extracted phase, its extraction is trivial. We do not have to localise a seed in this phase. This is the main advantage of this method which allows the segmentation of the material D (see figure 4.16).

4.3.5 Application

For the calibration, we begin the segmentation with a slice of the 3D image to find a suitable set of parameters rapidly. Then, we apply the segmentation to the whole 3D image with this set of parameters. In

phase	threshold range	opening size
Material A		
white phase (grains)	150-255	1
black phase (matrix)	0-100	0
Material B		
black phase (grains)	0-80	1
white phase (matrix)	160-255	1
Material C		
whitest phase (anhydrous grains)	140-255	0
light-grey phase (portlandite)	100-130	1
grey phase (hydrate)	70-90	0
black phase (capillary porosity)	0-30	0

Table 4.2: The parameters for the one-step method

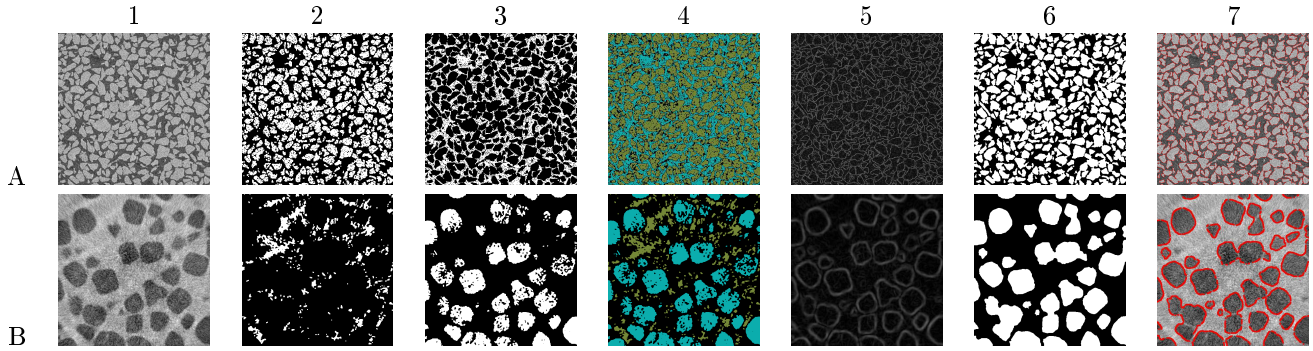
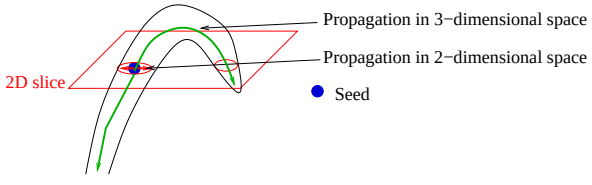


Figure 4.15: Visualisation of each step for the materials A and B: (1) the initial image, (2-3) the seed localisation inside both phases, (4) the seeds image, (5) the gradient image, (6) the catchment basins after the application of the watershed transformation to the gradient image with the previous seeds, (7) visualisation of the boundary of the catchment basin on the initial images. There is a good match with the visual segmentation.

all the studied materials, 3D segmentation seems to work better than 2D segmentation[41](see figure 4.17). In appendix A.4, we present the application for SEM images. For all material, the gradient operator is Deriche operator[41]. In the next subsection, we will compare different gradient operators.

[41]The watershed transformation is a growing process (see chapter 3). In 3D, more paths of propagation are available



than in 2D.

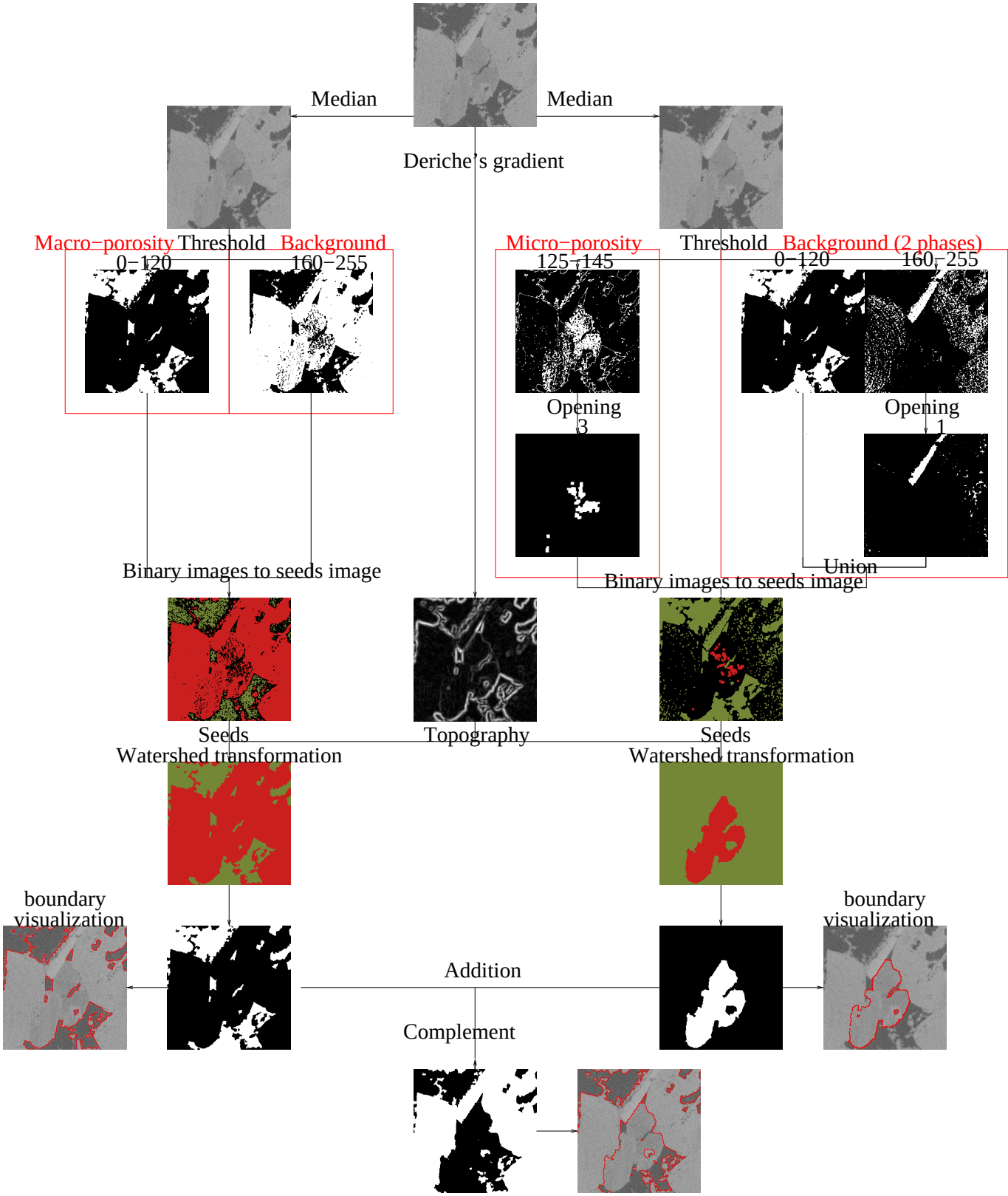
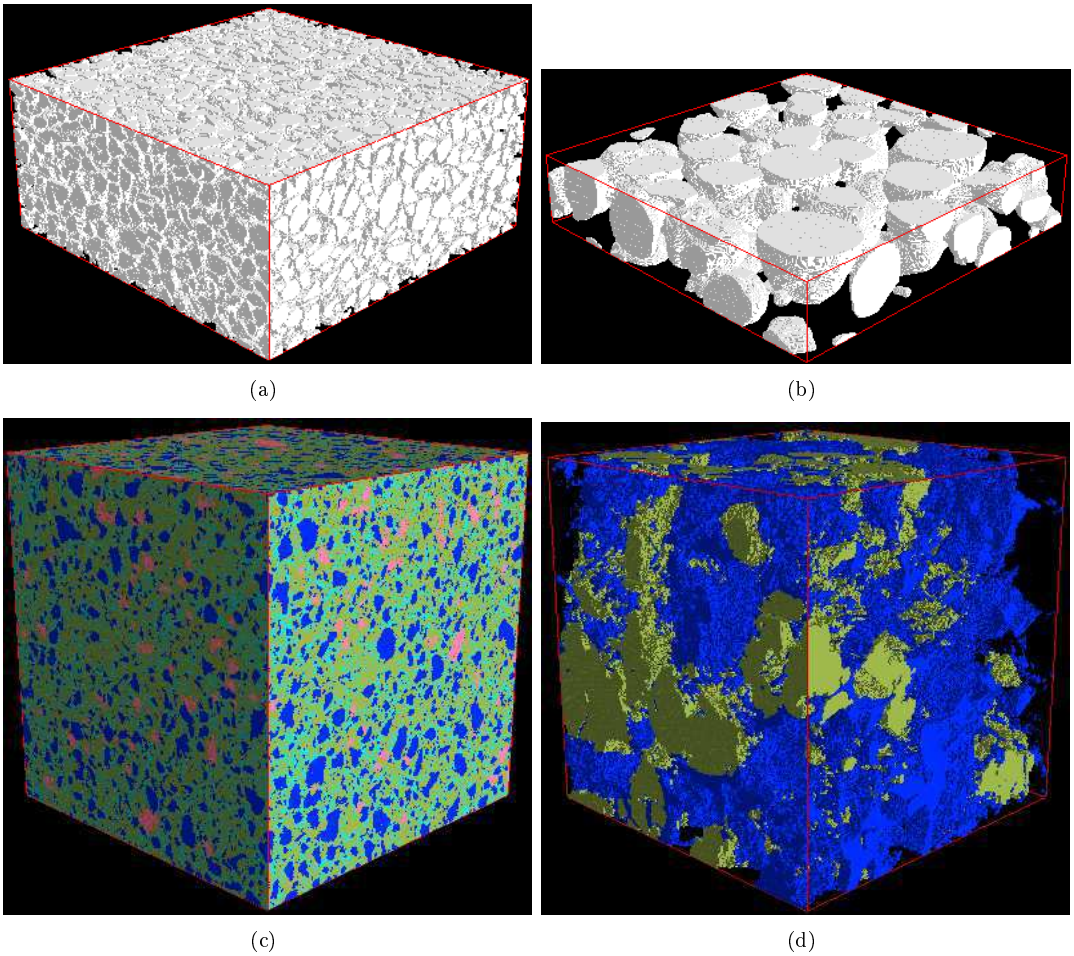


Figure 4.16: Extraction of three phases for the material D using a step-by-step method.



(a) (b)

(c) (d)

Figure 4.17: (a) the white grains for the material A, (b) the black grains (drawn in white in the figure) for the material B, (c) in blue the anhydrous grains, in pink the portlandite phase, in green the hydrate phase and in green/blue the capillary porosity for the material C. (d) in blue the porosity (the black phase) and in yellow the black grains for the material D.

4.3.6 Choice of the gradient operator

The watershed transformation belongs of the category of algorithms that detects the edges. The edges in the gradient image are the phase boundaries of the material in the original since a phase boundary is a sharp changes of grey-level. The efficient of the watershed transformation depends on the edges enhancement by the gradient operator. In the ideal case, the result of applying a gradient operator to an image lead to a set of connected curves in 2D and connected map in 3D that indicate the phases boundaries. Unfortunately, it is not possible to obtain such ideal edges from tomography/SEM images. A variety of computations are available which determine the magnitude of contrast changes and their orientation. Extensive literature exists documenting the available operators. Fast and simple edge detection can be performed by filters such as the popular Sobel gradient [102] which needs the convolution of a small kernel (3×3 pixels) over the image or the Beucher gradient which is the subtraction of the dilated image by the eroded image. Alternatively, more computationally intensive contour detection techniques are available such as the Deriche gradient [41]. In this subsection, we will test the segmentation quality of these three classical gradient operators using the watershed transformation as edges detector.

Sobel operator

The Sobel operator is based on the convolution the image with two 3×3 kernels[42] to calculate the derivation for horizontal changes and for vertical change. If we define I as the source image, G_x and G_y, the horizontal and vertical derivation, are computed as follows:

$$\mathbf{G}_y[I] = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I \quad \text{and} \quad \mathbf{G}_x[I] = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * I$$

where $*$ denotes the 2-dimensional convolution operator.
At each point in the image, the resulting gradient magnitude image is, using the Euclidean norm:

$$\mathbf{G}[I] = \sqrt{\mathbf{G}_x[I]^2 + \mathbf{G}_y[I]^2}$$

The 3×3 Sobel operator acts locally on the image and only detects edges at small scales. As the convolution with a small bounded kernel has small computational cost, this algorithm is very efficient and the implementation is simple. It can be extended to the n-dimensional space[43].

Beucher operator

The Beucher operator belongs to the field of Mathematical Morphology. The erosion and dilation for grey-level image is [158]:

- Dilation of the image I by the set B, called structural element, is defined by:

$$[I \oplus B](x) = \max_{\forall z \in B} I(x - z).$$

[42] small, separable, and integer kernels in horizontal and vertical direction
[43] In n-dimensional space, the i-derived approximation is:

$$\mathbf{G}_i[I] = \overbrace{\underline{g} \otimes \ldots \otimes \underline{g}}^{i-1 \text{ times}} \otimes \underline{d} \otimes \overbrace{\underline{g} \otimes \ldots \otimes \underline{g}}^{n-(i+1) \text{ times}} * I$$

with $\otimes$ the tensor product, $\underline{g} = \frac{1}{4}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ the Gaussian vector and $\underline{d} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ the derivation vector.
At each point in the image, the resulting gradient magnitude image is:

$$\mathbf{G}[I] = \sqrt{\sum_i \mathbf{G}_i[I]^2}$$

- Erosion of the image I by the set B is defined by:

$$[I \ominus B](x) = \min_{\forall z \in B} I(x - z).$$

Morphological gradient of the image I by the set B is defined by:

$$\mathbf{G}[I](x) = \frac{1}{2}([I \oplus B](x) - [I \ominus B](x)).$$

The scale parameter, the radius r of the structural element B, allows the filtering of the high frequency noise. As the algorithm complexity is $O(Nr^n)$ with N the number of pixels/voxels of the image and n the dimension of the space, the computational cost is not independent of the scale parameter. However, this algorithm is quite efficient for practical utilisation since the scale parameter is most often inferior to three. Finally, as the dilation and the erosion are defined in the n-dimensional space, the morphological gradient is also defined in the n-dimensional space .

Deriche operator

The Deriche operator is the extension of the optimal edge detector proposed by J. Canny [31] to a recursive filter. In the theory of edge detection algorithm, Canny defines three criteria for an "optimal" edge detector:

- *good detection* - the algorithm should mark as many phase boundaries in the image as possible even if the grey level variation on the phase boundaries is weak.
- *good localisation* - the lines of ridge[44] should be as close as possible to the phase boundaries in the real image.
- *minimal response* - a given phase boundary in the image should only be marked once, and where possible, image noise should not create false boundaries.

On an image model, using the calculus of variations, Canny proved that the optimal function is described by the sum of four exponential terms, but can be approximated by the first derivative of a Gaussian, the Canny filter. For a reduction of the computational time, Deriche proposed an another good approximation, ψ:

$$\psi(x) \quad = \quad cx \exp(-\alpha|x|)$$

with

$$c = -\frac{[1 - \exp(-\alpha)]^2}{\exp(-\alpha)}$$

The scale parameter α represents the inverse of the standard deviation of the Canny filter ($\alpha = \frac{\pi}{\sigma}$). In one dimension, the resulting gradient magnitude image is:

$$\mathbf{G}[I] = \psi * I$$

where $*$ denotes the convolution operator.
In discrete space, this gradient operator can be efficiently implemented by two recursive filters moving in opposite directions:

$$g^+(m) = aI(m-1) + b_1 g^+(m-1) + b_2 g^+(m-2) \quad \text{for } m = 2, ..., N-1$$

$$g^-(m) = -aI(m+1) + b_1 g^-(m+1) + b_2 g^-(m+2) \quad \text{for } m = N-3, ..., 0$$

$$\mathbf{G}[I](m) = g^-(m) + g^+(m) \quad \text{for } m = 0, ..., N-1$$

[44] The gradient image can be seen as a topographic surface

with $a = ce^{-\alpha}$, $b_1 = 2e^{-\alpha}$, $b_2 = -e^{-2\alpha}$ $c = \frac{(1-e^{-\alpha})^2}{e^{-\alpha}}$. The boundary values are:

$$\begin{aligned} g^+(0) &= \frac{a}{1-b_1-b_2} I(0) \\ g^+(1) &= aI(0) + (b1+b2)g^+(0) \\ g^-(N-1) &= \frac{a}{1-b_1-b_2} I(N-1) \\ g^+(N-2) &= aI(N-1) + (b1+b2)g^+(N-1) \end{aligned}$$

The Deriche output can be adjusted with the α scale parameter to filter out high frequency noise and to avoid the pixel effect[45] by giving an intense, smooth and continuous line of ridges. Due to its recursive nature, the computational time is independent of the scale parameter. Therefore, this algorithm is very efficient. It can be extended to the n-dimensional space [46].

Gradient as topography in the watershed transformation

A gradient operator can generate two artefacts :

- false edges if there are some inhomogeneity in the phase,
- some holes on the edges if there are some weak sharp changes on the phase boundary.

[45] Angularity is inherent of the discrete space

[46] This gradient operator is the first derivative of a smoothing filter,$\Psi(x)$. Thus, we have:

$$\begin{aligned} \frac{d\Psi(x)}{dx} &= \psi(x) \\ \Psi(x) &= k(\alpha|x|+1)\exp(-\alpha|x|) \end{aligned}$$

with

$$k = \frac{(1-e^{-\alpha})^2}{1+2\alpha e^{-\alpha} - e^{-2\alpha}}$$

In n-dimensional space, the i-derivative is the convolution product of one derivation in the i-direction with the $n-1$ smoothing on the other directions:

$$\mathbf{G}_i[I](x_0,\ldots,x_{n-1}) = (\Phi(x_0) * \ldots * \Phi(x_{i-1}) * \phi(x_i) * \Phi(x_{i+1}) * \Phi(x_{n-1})) * I$$

with $(x_0,\ldots,x_{n-1})$ are the coordinates of the point x.
At each point in the image, the resulting gradient image magnitude is:

$$\mathbf{G}[I] = \sqrt{\sum_i \mathbf{G}_i[I]^2}$$

In discrete space, this smoothing filter can be efficiently implemented by two recursive filters moving in opposite directions:

$$s^+(m) = a_0 I(m) + a_1 I(m-1) + b_1 s^+(m-1) + b_2 s^+(m-2) \quad \text{for } m = 0,...,N-1$$

$$s^-(m) = a_2 I(m+1) + a_3 I(m+2) + b_1 s^-(m+1) + b_2 s^-(m+2) \quad \text{for } m = N-1,...,0$$

$$\mathbf{S}[I](m) = s^-(m) + s^+(m) \quad \text{for } m = 0,...,N-1$$

with $a_0 = k$, $a_1 = k(\alpha-1)e^{-\alpha}$, $a_2 = k(\alpha+1)e^{-\alpha}$, $a_3 = -ke^{-2\alpha}$, $b_1 = 2e^{-\alpha}$, $b_2 = -e^{-2\alpha}$, $k = \frac{(1-e^{-\alpha})^2}{1+2\alpha e^{-\alpha}-e^{-2\alpha}}$. The boundary values are:

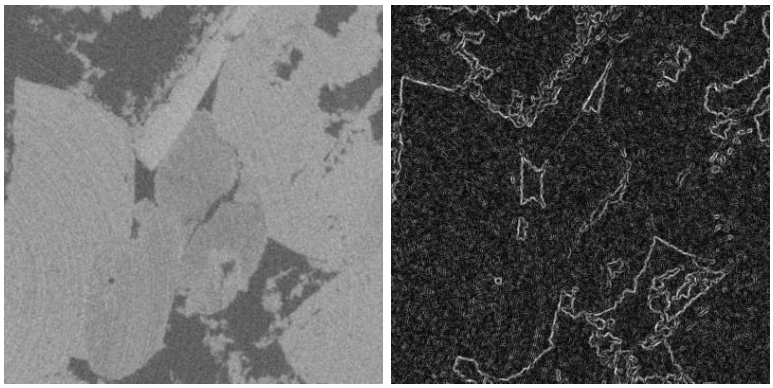
$$\begin{aligned} g^+(0) &= \frac{a_0+a_1}{1-b_1-b_2} I(0) \\ g^+(1) &= a_0 I(1) + a_1 I(0) + (b1+b2)g^+(0) \\ g^-(N-1) &= \frac{a_0+a_1}{1-b_1-b_2} I(N-1) \\ g^+(N-2) &= (a_0+a_1)I(N-1) + (b1+b2)g^+(N-1) \end{aligned}$$

The i-derivative can be implemented by n times two recursive filters using this iterative process:

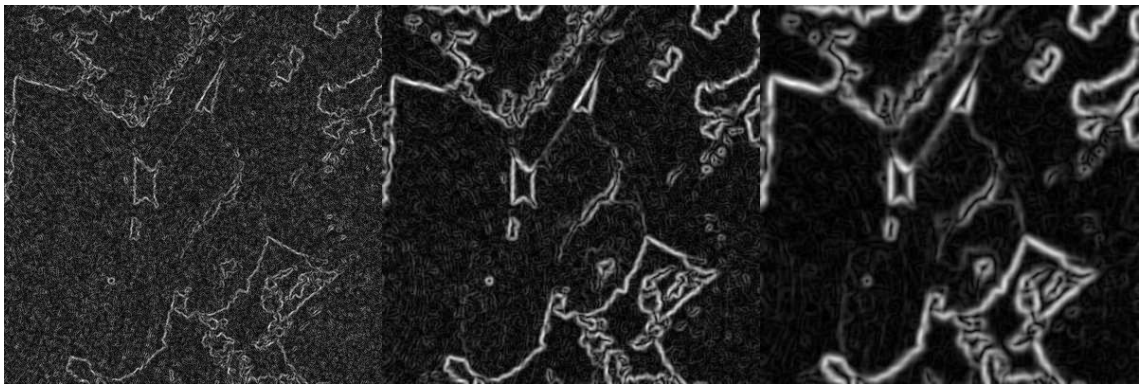
- Initialisation: $\mathbf{R}^0 = I$,
- Until $j < n$, $\mathbf{R}^{j+1} = \mathbf{B}[R^j]$ with $\mathbf{B}^{j+1} = \mathbf{G}$ for $j = i$ and $\mathbf{B}^{j+1} = \mathbf{S}$ otherwise.

For $n = 2$, $\mathbf{G}_x$ is calculated by the derivation in the x-direction following by the smoothing in the y-direction.

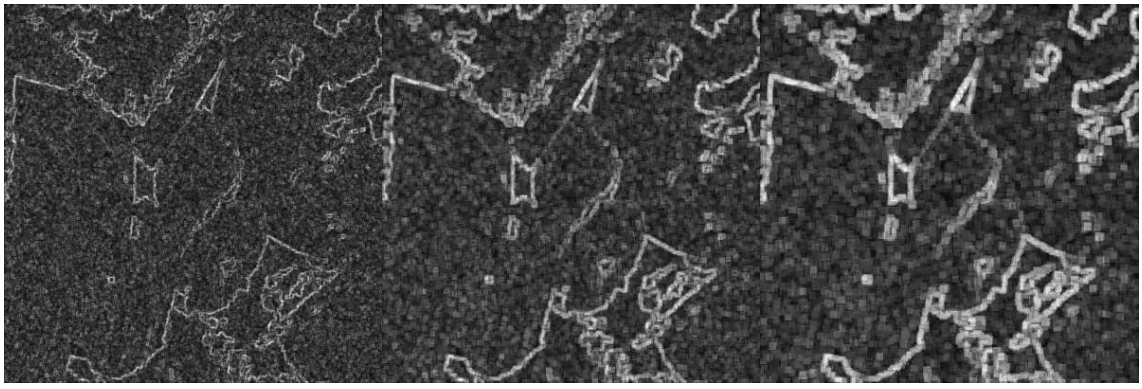
As shown in figure 4.18, the image gradient depends strongly on the operator. With a scale parameter $\alpha = 5$ and $r = 1$, the resulting images of Deriche operator and Beucher operator seem like the resulting image of Sobel operator. With a scale parameter $\alpha = 0.5$ and $r = 3$, Deriche operator blurs the line of ridges corresponding to the spatial inhomogeneity in the phases and reveals the line of ridges of the phase boundary unlike for Beucher operator or Sobel operator. Since the watershed transformation affects



(a) Original image (b) Sobel gradient



(c) Deriche gradient: $\alpha = 5$ (d) Deriche gradient: $\alpha = 1$ (e) Deriche gradient: $\alpha = 0.5$



(f) Beucher gradient: $r = 1$ (g) Beucher gradient: $r = 2$ (h) Beucher gradient: $r = 3$

Figure 4.18: Illustration of the different gradient operators.

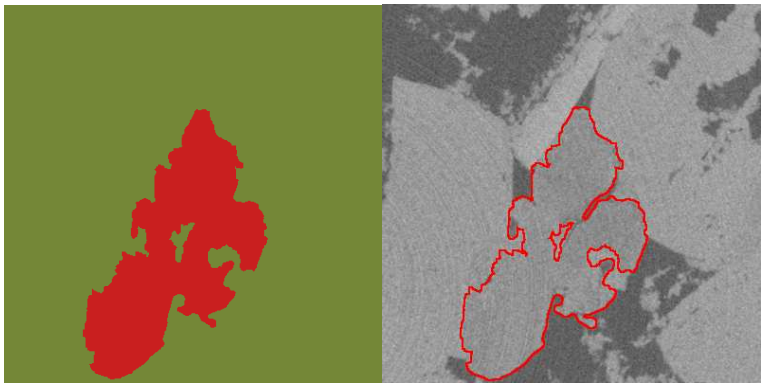
points on the catchment basins with a local law during the growing process[47], both these artefacts can produce a bad localisation of the numerical boundary even with appropriate seeds. The gradient choice

[47] it is not global minimisation of a given functional

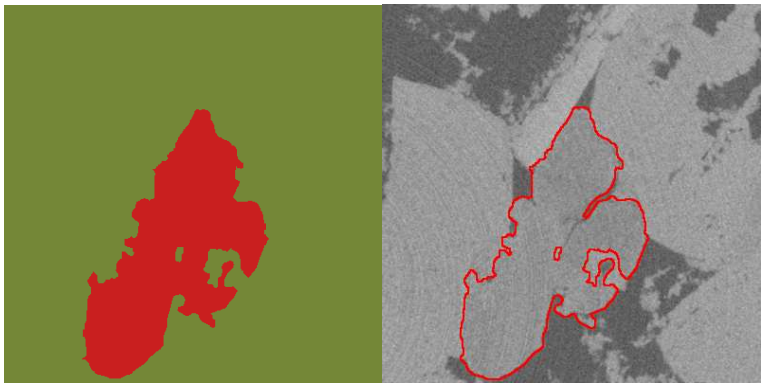
is crucial in the watershed transformation as illustrated in figure 4.19. We observe that Deriche operator leads to the best localisation of the numerical phases boundary. The next subsection evaluates why the

(a) Seeds

(b) Catchment basins with Sobel gradient



(c) Catchment basins with Beucher gradient



(d) Catchment basins with Deriche gradient

Figure 4.19: Application of the watershed transformation with the seeds of the image (a) on the different gradient images.

one step watershed transformation is more robust than threshold segmentation.

4.3.7 Robustness

Generally, the efficient of the segmentation is only evaluated by the accuracy. But, for pratical utilisation, an important parameter is the facility of the calibration of the segmentation procedure. This subsection

will prove that the calibration of the one-step is easy.
The definition of the segmentation robustness can be the stability of the segmentation against perturbation of segmentation parameters. A segmented image localises the phases of the material. The geometrical organisation of each phase is characterised using metric/topological descriptors (see chapter 5). The "stability" means that the metric/topological descriptors of the segmented image are not affected (or shows very little fluctuation) by the perturbation of the segmentation parameters. In this subsection, the characterization is restricted to two functional: chord length distribution and 2-point probability function.
This subsubsection has two paragraphs: morphological analysis and robustness evaluation.

Metric analysis

The chord length distribution function and the two-point correlation function give a statistical analysis of the geometrical organisation of a phase [33, 106, 134]. Their determination gives information about volume fraction, average granular size, surface roughness and structural correlation (for a more extended explanation see section 5.2)
1) A chord is a segment belonging to the phase and having its two extremities at the phase interface. As shown in Fig. 4.20, chords are obtained by tracing random and homogeneously distributed straight lines through the microstructure. The chord length distribution function is the probability of getting a chord length between h et $h + dh$, belonging to the phase, $f_p(h)$.
2) Let us draw an interval of length h randomly in the material : the two-point probability functions is the

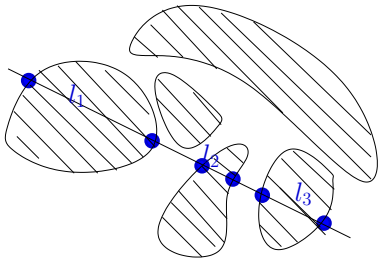


Figure 4.20: A chord trough a material composed by two phases

probability of having both extremities of this interval belonging to the phase $S_{2,p}(h)$[48] (see figure 4.21).

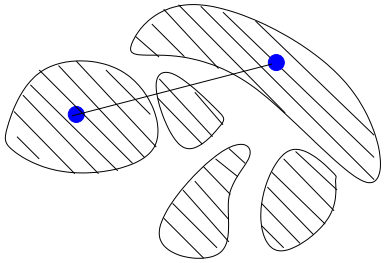


Figure 4.21: A yardstick drawn randomly in a material composed by two phases

Robustness evaluation

Let $S_\Lambda(I)$ be the resulting binary image after the application of the threshold segmentation or the one-step method for a two phase material on the initial image I with the parameter set Λ.
To check the stability, the method is:

[48] also noted $C(h)$ for covariance, or correlation function or covariance

1. to select a parameter, λ, in the parameter set, Λ,
2. to increase step by step the chosen parameter, $\lambda_i = \lambda_{i-1} + \epsilon$, starting from a low value until a high value,
3. for each step,
   - to apply the segmentation procedure to get the binary image $S_{\Lambda_i}(I)$ associated with the parameter λ_i,
   - to calculate the distance between $S_{\Lambda_i}(I)$ and $S_{\Lambda_{i-1}}(I)$. The distance is defined as the "area" between the two plots of 2-point probability function or chord length distribution for $S_{\Lambda_i}(I)$ and $S_{\Lambda_{i-1}}(I)$. A small distance means that the segmentation is stable and vice-versa.

Figure 4.22 shows that:

- the one-step method is more stable than threshold segmentation,
- the global minima for the threshold graph is 125. This value is equal to the value of the valley on the histogram. Therefore, this graph could allow the automatic determination of the threshold parameter,
- the global minima (see figure 4.23).) for the one-step graph is 85. This value is closed to the manual determination of the treshold value for the seed localisation. Therefore, this graph could allow the automatic determination of the threshold parameter,
- as the global minima is a hole for the threshold graph, if the parameter is not exactly 90, the segmentation will be different. For the one-step graph, as the global minima is surrounded by a table, a wide range of values lead to the same segmentation. This has been confirmed by the experiments: the search of the input parameters is easy. Thus, the main advantage of the one-step method is that the calibration is easy since the appropriate input parameters belongs to a range (not a single value as is use with threshold segmentation).

In conclusion, the one-step method is more stable than threshold segmentation. However, this result is not significant since the threshold operator without an advanced filtering is the simplest segmentation method. The relevant results are that the calibration of the one-step method is easy and that an automatic calibration is possible. The next subsection will prove the accuracy of the one-step method.

4.3.8 Noise stability

This methodology tests the noise stability of the one-step method:

- Objective image: it is a boolean random field (see subsection 4.1.1) which is often used for the description heterogeneous material [5, 78].

$$I(x) = \begin{cases} a \text{ for } x \in M \\ b \text{ else} \end{cases}$$

  with $M = \cup_{i \in I}(Xi \oplus K_0)$, $\{X_i\}$ a Poisson point process and K_0 a constant ball.

- Addition of a noise: it is a Gaussian white noise[49], N, with a standard deviation σ to the original image:

$$Y = I + N.$$

  The contrast-to-noise ratio is defined by the dimensionless quantity: $|a - b|/\sigma$.

[49] A process N_x is denoted by Gaussian white noise if:

- N_x is an independent white noise,
- N_x follows a normal law $\mathcal{N}(0, \sigma^2)$

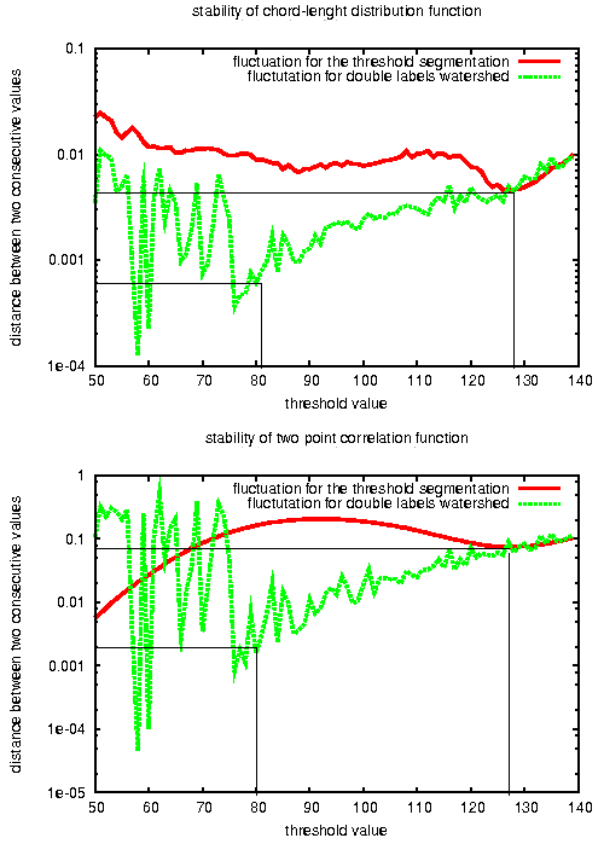


Figure 4.22: Stability of the segmentation methods for the material B. For both distances, the one-step method is more stable by one decade than threshold segmentation.

- Segmentation: it is the application of the one-step method using the same input parameters whatever the contrast-to-noise ratio. The resulting segmented set is denoted by S_Y.
- Distance: it is an error function defined by Tariel *et al* [165]:

$$error(A,B) = \frac{\mu((A \cup B) \setminus (A \cap B))}{\mu(A) + \mu(B)}.$$

  In our purpose, A represents the initial boolean random set, M, and B represents the resulting segmented set, S_Y.

The results are presented in figure 4.24. Qualitatively, we observe a good match between the boolean random set and the segmented image even when the contrast-to-noise ratio is low. Quantitatively, the error values confirm it (see table 4.3). In conclusion, for the one-step method,

1. the calibration of the segmentation parameters is independent of the contrast-to-noise ratio,
2. the segmented microstructure matches the model even with a low contrast-to-noise ratio.

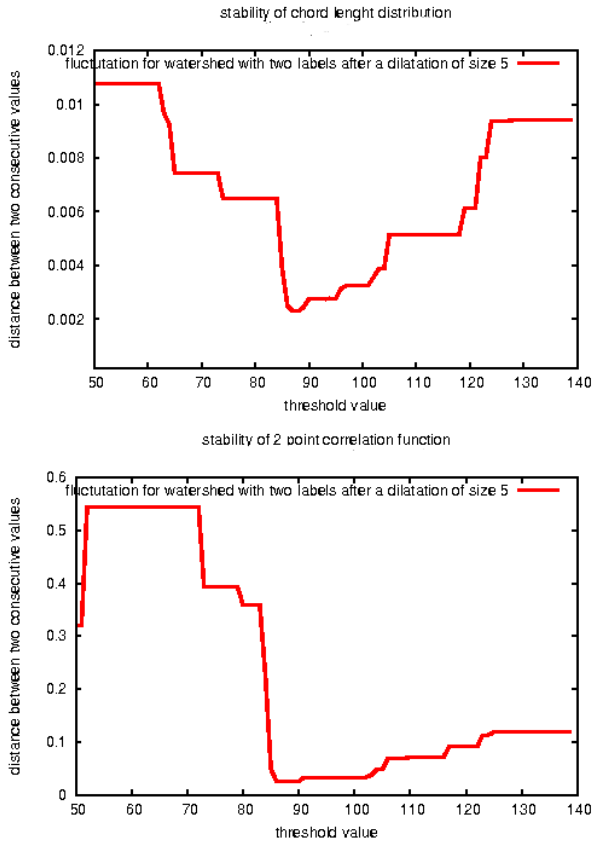


Figure 4.23: After a dilatation with a structural element of size 5 on the graph, the minimum on both curves is 87. This automatic value is closed to the manual value 90.

$\frac{\lvert a-b\rvert}{\sigma}$	$\mu(S_Y)$	$\mu((S_Y \cup M) \setminus (S_Y \cap M))$	$error(S_Y, M)$
2.5	0.4915	0.0160	0.162
1.25	0.4901	0.0242	0.0245
0.83	0.488	0.0342	0.0348
0.55	0.472	0.0600	0.0620

Table 4.3: The Lebesgue measure of the boolean random set is equal to 0.4951.

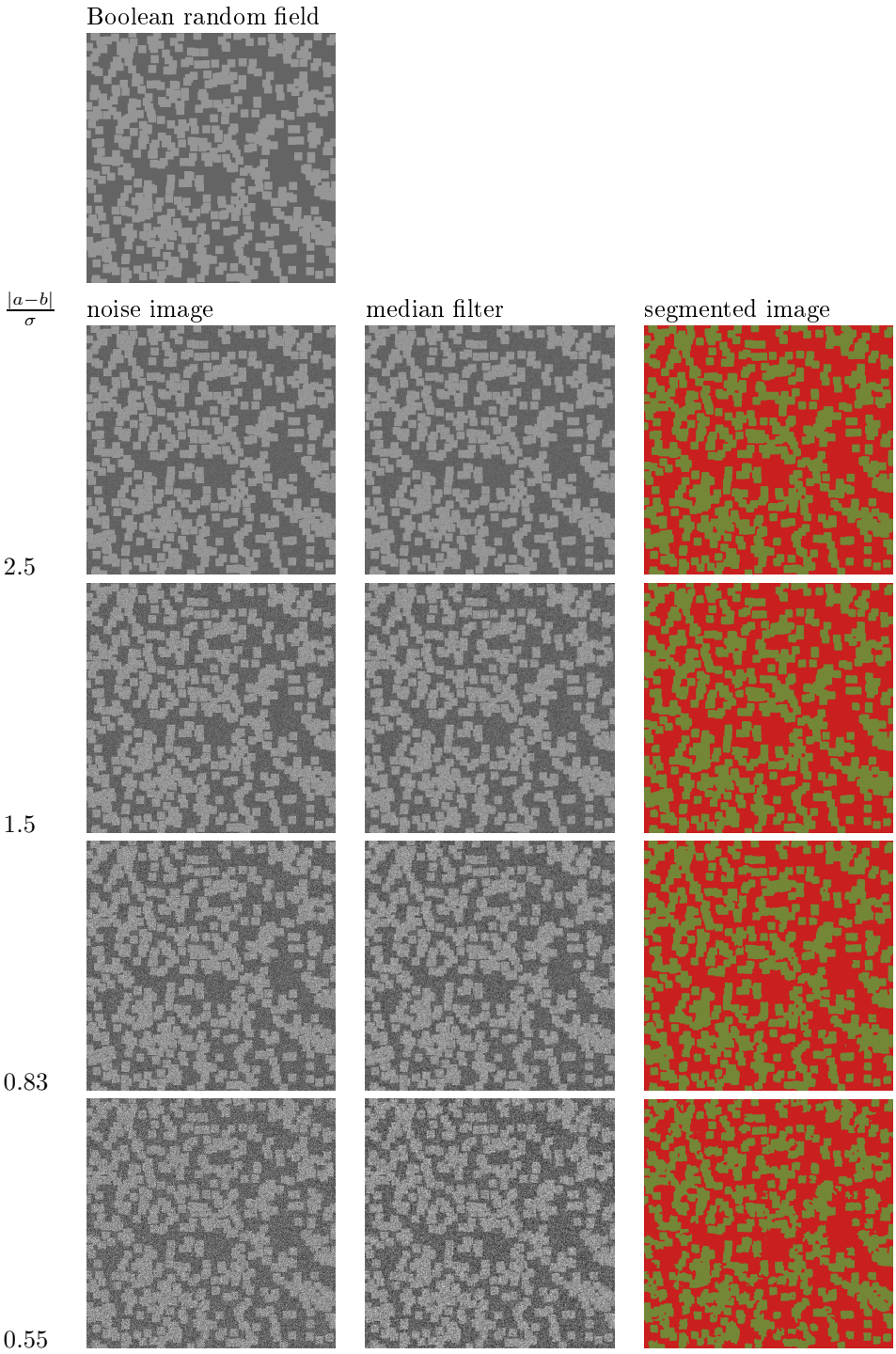


Figure 4.24: For the boolean model, the image size is 1024*1024 with a density of 0.000625 and a ball radius of 16. The pixel value is equal to 150 inside the random set and 100 outside. Therefore, the contrast is equal to 50. The standard deviation, σ is respectively $20, 40, 60$ and 80. The segmentation parameters are chosen constant. For the median filter, the structural element is associated to the 8-neighbourhood. For the seeds selection, we operate a threshold (0-100) and (150-255) following by an opening using the 8-neighbourhood. The Deriche parameter is equal to 0.5.

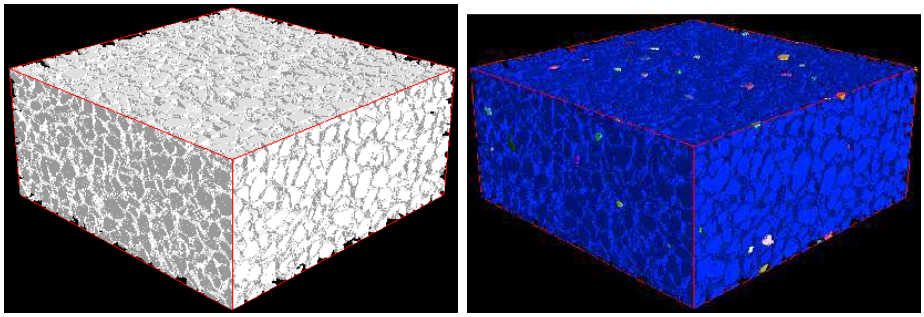


Figure 4.25: First figure: segmented image; second figure: extraction of the connected components. Each colour represents a connected component. There is a large cluster, in blue, containing many grains.

4.4 Artefacts correction

Segmentation generates some artefacts:

1. some grains close to each others become connected,
2. some artificial holes appear.

This section describes two procedures to correct their.

4.4.1 Grain splitting

Whatever the method, segmentation exhibits one difficulty: some grains close to each other become connected. For many applications, the extraction of each grain of the grain phase, named grain partition, has to be performed. If a classical cluster procedure [72] is applied on the segmented binary image, these connected grains are numerically considered as a single grain (see figure 4.25). To overcome this, we apply a procedure introduced by L. Vincent in 1993 [172]. Let Ω be the domain of the grain phase on the segmented image. The algorithm principle is the application of the watershed transformation controlled by seeds restricted by the binary image, I, to the topographic surface. The topographic surface is the opposite of the distance function of I (see figure 4.26).

Opposite of the distance function: The distance function $Dist^{\Omega}$ gives the smallest distance between a voxel x belonging to Ω and any voxel y belonging to the complementary of Ω.

$$Dist^{\Omega}(x) = \begin{cases} \min_{\forall y \in \Omega^c} d(x,y) & \text{for } x \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

The opposite of a image J is: $K(*) = -J(*)$.
A property of the opposite distance function is that the lines of ridge are located on the narrow lines/surfaces between the grains. The watershed transformation, dedicated to the extraction of line of ridges, is also used for this purpose. However, in this application,

1. the watershed transformation is restricted on the domain I (see subsubsection 3.4.2),
2. each grain has to be marked by a single seed,
3. the topographic is the opposite distance function

Appropriate seeds: First, the seeds can be the regional minima (see subsubsection 3.4.1 for the regional minima definition) of the opposite distance function of the binary image. However, this merging produces an important over-extraction of grains due to local irregularities of the binary image shape, generating a set of uncontrolled and unwanted seeds. To avoid this problem, the opposite distance function of the

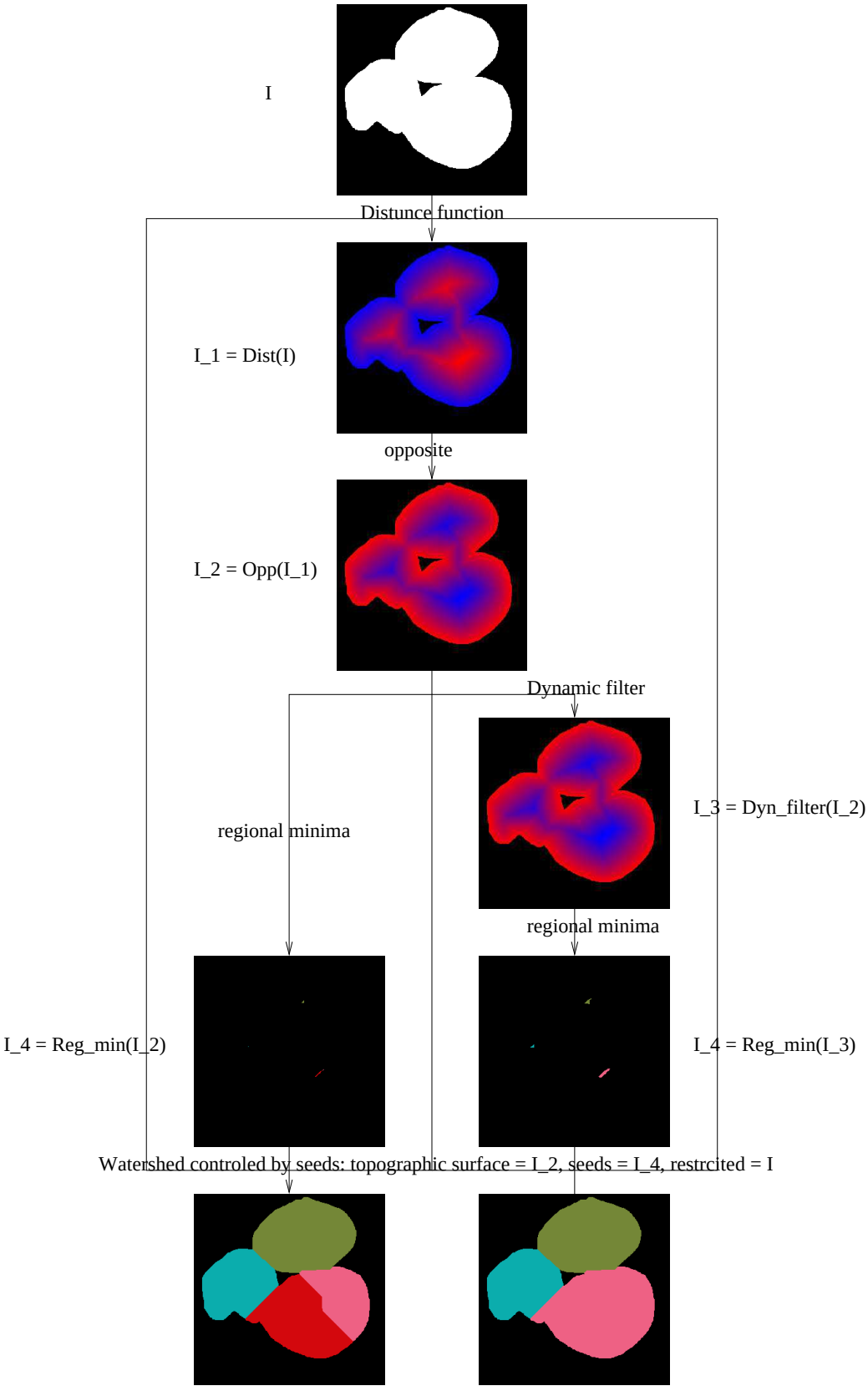


Figure 4.26: Grains extraction: the watershed lines/surfaces are located on the narrow lines/surfaces because the narrow lines/surfaces are located on the lines of ridge of the opposite distance function of I.

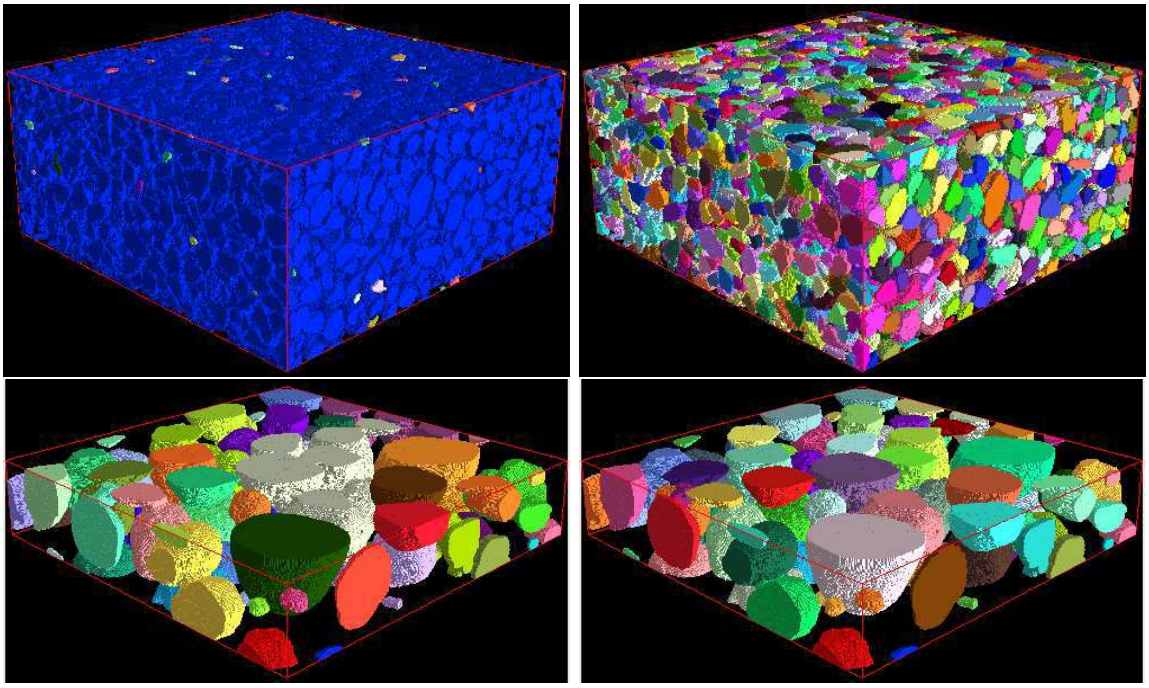


Figure 4.27: Each colour represents a grain. For each row, the first image is the visualisation of the clusters on the segmented binary image and the second image is the catchment basins after the splitting procedure. For both granular materials, the dynamic parameter is equal to 1.

binary image is filtered. A vertical filter is used in order to individualise each grain with a single seed. The vertical filter uses the notion of dynamic (see subsubsection 3.4.2 for the dynamic filter definition) with a parameter h [59]. For the granular materials A and B, this splitting procedure yields a consistent grain partition as illustrated in figure 4.27.

4.4.2 Hole filling

During the preparation of samples for SEM, the surface is polished. This process generates some cracks inside the grains. After segmentation, these cracks are not affected to the grain phase since the tint of the crack is very different to the tint of the anhydrous grains. In tomography images, connected components of the solid phase are numerically surrounded by porosity (levitation). We observe around 5 of such "holes". In order to correct these artefacts, a hole filling procedure is applied (see figure 4.28). This artifact correction is important:

- for the characterization based on the chord length distribution function since this functional is very sensitive to the presence of holes in the phase[50],
- for the contraction at constant topology (see subsection 5.3.3) since this method can only be applied without internal surfaces.

Conclusion

X-ray tomography/SEM images involve specific problems of image segmentation, due to the weak contrast-to-noise ratio and the limitation of the spatial resolution. If the resolution of the imaging technique is in the same order of the characteristic size of phase under investigation, segmentation will not yield an

[50] It is not the case of the 2-point probability function.

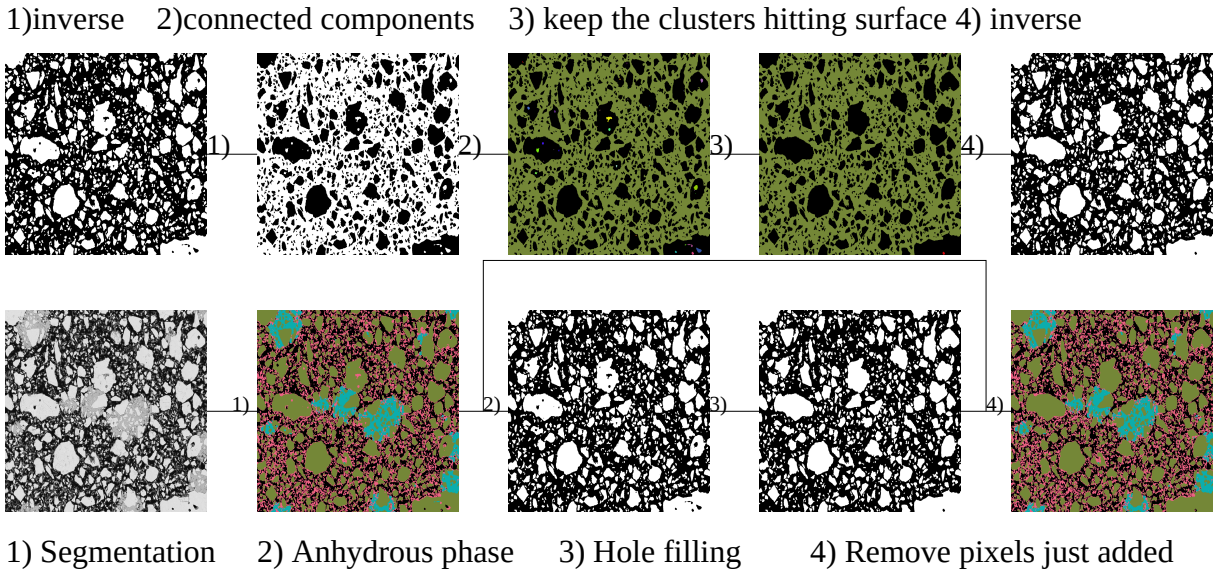


Figure 4.28: The first row is the decomposition of hole filling by elementary operators in such a way: -inverse-extraction of the connected components-keep the connected components touching the image boundaries-inverse. The second raw is the application to anhydrous grains in a SEM image: first figure: initial image -(1)-segmented image -(2)- selection of the anhydrous phase -(3)- hole filling procedure on the anhydrous phase -(4)- removed the pixels/voxels belonging to holes of anhydrous grains initially belonging to other phases.

accurate microstructure whatever the method. We have seen that many problems can appear during the segmentation. In the first section, we presented the classical threshold segmentation given a good result only if the contrast-to-noise ratio is high. When this is not the case, the one-step method or the step-by-step method, was applied. These methods are, simple because the input parameters belongs to a wide range of possible values and does not depend on the contrast-to-noise ratio; generic because they can be applied for different materials obtained with different microscopies and robust because, even with a weak contrast-to-noise ratio, segmentation yields an accurate microstructure.
In the last section, we introduced two procedures to correct two common artefacts:

- a split procedure to separate the connected grains[51],
- a hole filling procedure.

[51] this approach will be also used to decompose the pore network into elementary pores.

5

Geometrical characterization

Contents

After segmentation, the different phases are located in space though a space partition. One challenging problem deals with the ability to describe the geometrical organisation of these phases. Quantitative knowledge of Porous Medium of Cement Paste (PMCP) is important in order to understand the role of geometric confinement in adsorption, condensation, transport, reaction processes and finally, the connection with the microstructure and durability of the material. In this chapter, we focus on the extraction of the geometrical features of PMCP in order to observe their evolution with the hydration time. The chapter is segmented in four sections:

- introduction,
- metric analysis,
- topological analysis,
- decomposition in elementary pores.

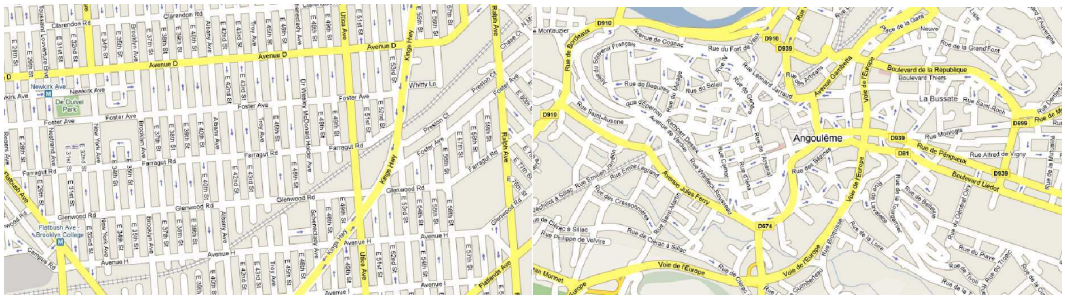


Figure 5.1: Left figure: New York city had been organised to optimise the traffic; right figure: Angouleme city had never been organised according to a global plan.

5.1 Introduction

5.1.1 The geometrical organisation is complex at one scale

To understand the complexity of the PMCP geometrical organisation, an example related to the car traffic is given. Let us consider the car traffic in two towns: New York and Angouleme, a medieval town in South West of France (see figure 5.1). Obviously, the car traffic flow in New-York can be higher than in Angouleme although the fraction of surface area of road is quite similar for the both cities. Qualitatively, in Angouleme, the car traffic is less efficient since its road network has:

1. some topological characteristics as
   - many dead-end roads,
   - no "direct road" between one point and another point,
   - some crossroads where many important roads merge.
2. some metric characteristic as
   - some narrow roads.

This example shows that the properties of car traffic is related to its geometrical organisation. Obviously, the geometrical organisation of road network of Angouleme is disorder. The geometrical organisation of PMCP is also disorder. In order to especially understand how the diffusion occurs in this porous media, the requirement is the statistical characterisation of the geometry of the porous microstructure. Therefore, this chapter will dedicated to the quantitative analysis of the pore network at one scale. The next subsection will introduce the concept of multi-scaling.

5.1.2 The geometrical organisation is specific at each scale

Now, let us consider the car traffic in the road network in France. The road network is segmented depending on the road size: highways, low capacity road,... (see figure 5.2). Each road size is associated to a scale. In order to understand the car traffic, a geometrical analysis has to be done:

- at each scale. For example, in France, at the upper scale (highway), it is a spider network centred at Paris,
- between the scales since there a coupling between them. Go to mappy and choose two villages far enough. Generally, the itinerary respects the principle of up-scaling-downscaling, starting from low scale road (lane) to join a high scale road (highway) and finishing with a low scale road.

The geometrical organisation of PMCP is also:

1. organised through a wide range of length scales from the nanometre to the millimetre,

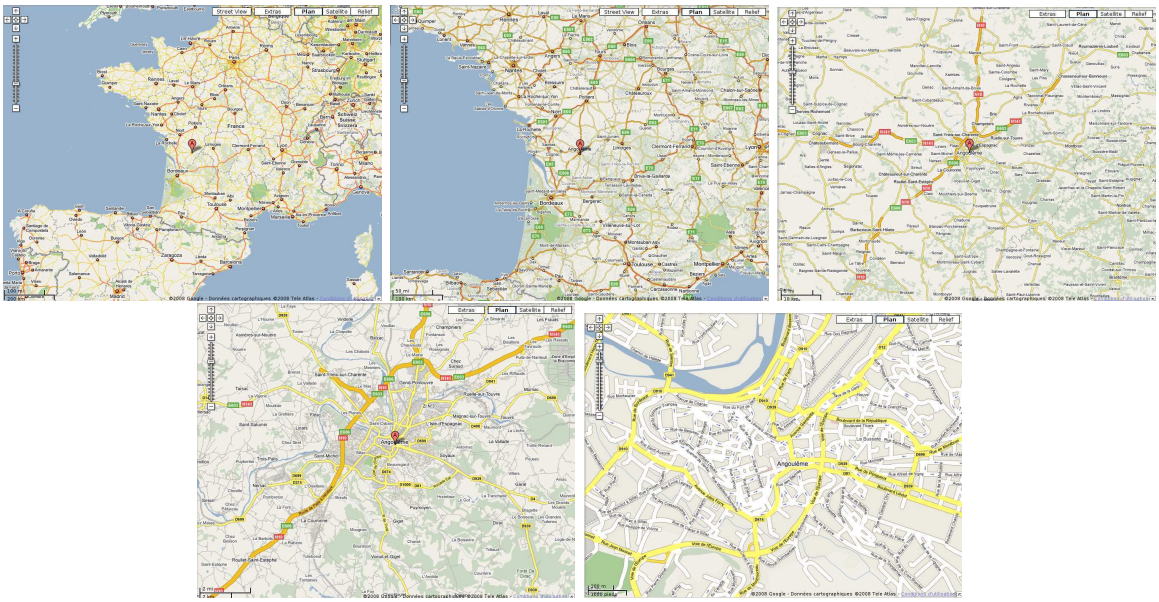


Figure 5.2: Multi scaling road network

2. specific at each length scale without statistical length scale invariance,

3. complex at each length scale.

Moreover, the geometrical organisation is hierarchical since there is no scale separation. At each increasing of the magnification, new features of the PMCP are revealed[52]. In this thesis, we focus on the upper length scale of the PMCP, called the capillary porosity. In order to investigate the whole range of length-scale heterogeneity of a cement paste, this thesis is coupled with two other theses:

- S. Brisard supervised by P. Levitz (understanding the nanostructure of C-S-H: macroscopic implications),

- H. Chemni supervised by D. Petit (multi-scale diffusion and moisture sorption in cement paste materials).

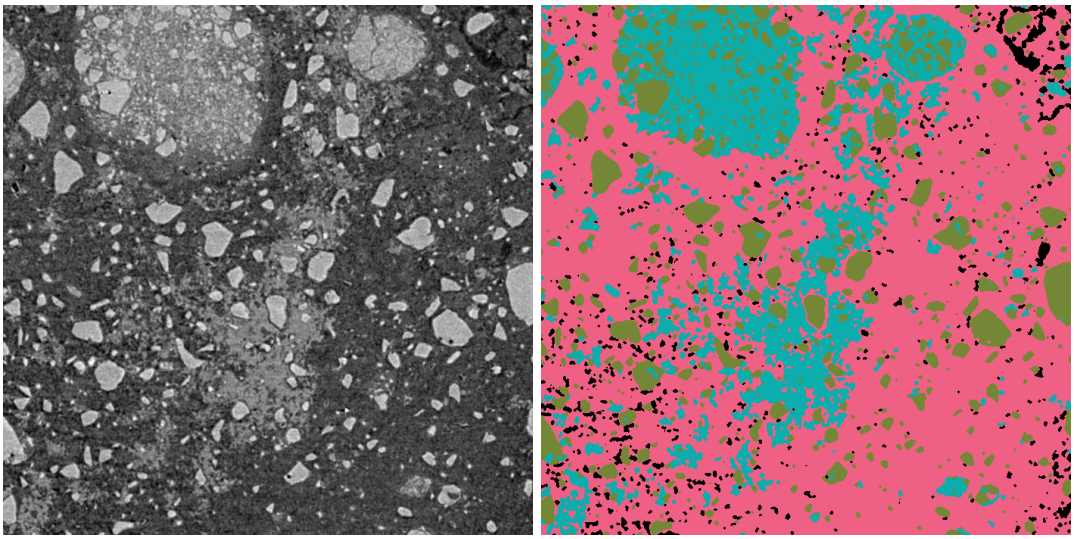
5.1.3 Basic concepts of image analysis

Two length scales

The observation of the the geometrical organisation of PMCP has be done using two imaging techniques. Each experiment imposes two length scales, the resolution, noted l, and the field of view[53], noted L. The resolution, l, sets the lower scale of the investigation and the field of view the upper scale (see figure 5.3). An important property is the spatial homogeneity [95] (translational invariance) allowing that the estimation of a property does not depend where the image is taken. In order to have this property, the field of view has to be larger than the characteristic size of the Representative Elementary Volume (REV). In the literature of cement paste, this volume is estimated around 100 μm^3 [137]. The both imaging techniques, SEM and X-ray tomography, have a field of view much larger than this size.
Another two important properties is the spatial stability. The estimation of a property stays constant when the resolution decreases. This involves that:

[52] For the road network, it is multi-scale since there is a scale separation due to the quantification of the road size

[53] For tomography at SLS, the resolution size is: $l = 0.7\mu$m and as the size of the image is 700^3 voxels, the field of view is: $L = 0.5mm$



(a) (b)

Figure 5.3: A 2D slice of the tomography image of alite at time of hydration, 31h, and W/C=0.4. First image: the acquisition, second image: after the segmentation. There are large clusters of portlandite (blue phase) which size is equal to half of the length of the field of view. Therefore, the volume of field is surely under the representative elementary volume. We observe few porosity (black phase). There is still porosity not revealed at this resolution in the hydrate phase (pink and blue phases).

- a geometrical organisation is observed. For a mature paste, we do not observe the geometrical organisation any more since the upper length scale of PMCP is lower than the tomography resolution (see figure 5.3).

- the observed geometrical organisation drives the physical phenomenon. For instance, a pore network with macro-pores and micro-pores such as the macro-pores are highly connected, the estimation of the transport property at the macro length scale will yield consistent results in agreement with direct experiments since the transport in the micro-pores can be neglected. When a physical phenomenon depends on both scales, micro- and macro-pores, an effective microscopic property can be affected at each pixel/voxel containing the micro-pores in order to take into account the coupling between both scales.

At the beginning of the hydration time, we will see in the section 5.3 that the capillary porosity is highly connected. Therefore, at this stage, the estimation of the diffusion properties is relevant. However, rapidly, at the resolution of the tomography, we do not observe the pore network. At this stage, tomography gives no information about the pore network of cement paste. Therefore, another imaging technique has to be used to observe the smaller capillary pores that control transport for mature pastes.

Dimension

For a 3D disorder system, like cement paste, a section, of a three-dimensional microstructure, allows only a limited geometrical analysis since topology is intrinsically a 3d information. Obviously, diffusive transport depends on the connectivity of the pore network. Therefore, tomography, providing 3D images, allows a full metric and topological analysis whereas SEM, providing 2D images, is limited for the understanding of how the diffusion occurs in porous media.

5.1.4 Which class of geometrical feature?

In this thesis, an extended analysis is done at one scale: the capillary porosity at the resolution of 0.7μm with a field of view of 0.5 mm. The geometrical organisation is described using two descriptor classes: metric and topological. By analogy of the road network, the metric descriptors give information about the throats, the distribution of road size, the distribution of mean curvature whereas the topological descriptors give information about the connectivity of the network such as the existence of path between two towns. For both classes, the analysis is segmented in two levels:

1. scalar analysis providing average numbers such as the porosity,
2. statistical analysis involving some statistical distribution such as the chord length distribution function.

These descriptors do not allow to link between the geometrical features and diffusion transport properties because their do not take into account the metric and the topology analysis simultaneously. A possible way is to decompose the microstructure in terms of elementary pores. This decomposition permits a metric description of the elementary pores and a topological description of the network of the elementary pores. Our aim is then to understand the local transport in elementary pores by using their metric features and to up-scale our analysis of the transport from the elementary pores to the microstructure, using the connectivity properties of the network of the elementary pores.

5.1.5 Which cement paste?

In order to compare the geometrical analysis done with SEM and X-ray tomography, we have acquired at the same resolution $0.675\mu m$, at the same field of view 0.5 mm, two cement pastes: Ordinary Portland Concrete (OPC) and alite (see chapter 2).

5.1.6 Notation

After segmentation, we get a partition of the space defined by the domains $(\Omega_0, \ldots, \Omega_n)$ of the different phases. Let $\mathcal{I}$ be the characteristic function defined by:

$$\mathcal{I}(x) = i \text{ if } x \in \Omega_i$$

For each phase i, a characteristics function is defined as:

$$I_i(x) = \begin{cases} 1 \text{ if } x \in \Omega_i \\ 0 \text{ otherwise} \end{cases}$$

A characteristic function is denoted, I, when the phase is not specified. The space and the image are denoted with the same notation, E.
For a set A, we define the Lebesgue measure,[54] $\mu(A)$ as

1. the surface of A, $S(A)$, in 2-dimensional space,
2. the volume of A, $V(A)$, in 3-dimensional space.

[54] A function from the power set of E, $\mathcal{P}(E)$, the set of all subsets of E to $\mathbb{R}^+$

5.2 Metric analysis

The aim of the metric characterization is to give information about forms, shapes and patterns like the average pore size, the mean curvature, the pore shape, the surface roughness, the structural correlation between pixels belonging either to the solid, to the interface or to the pore network. All the measure presented in this section are accessible either in 2D and in 3D for an isotropic material. This section is segmented in two subsections:

1. Minkowski functionals,
2. statistical analysis.

5.2.1 Minkowski functionals

To quantify the evolution of the pore network, we use three fundamental numbers: the volume fraction ϕ, the specific surface area, $\mathcal{S}$, and the Euler number, N_3. In the more general context of integral geometry, these three numbers are defined for an d-dimensional Euclidean space through $d+1$ Minkowski functionals (see Table 5.1). The Minkowski functionals, W_i, satisfy a number of criteria, the so-called

notation	geometric quantity	Minkowski functionals of number i	$W_i(A)$
$V, \phi = V_V$	volume, volume fraction	0	$V(A) = \phi(A)\mu(E)$
S, S_V	surface, specific surface area	1	$S(A) = S_V(A)\mu(E)$
H	integral surface mean curvature	2	$\frac{1}{3}\left(\frac{1}{2}\int_{\partial A}(\frac{1}{R_1}+\frac{1}{R_2})ds\right)$
K	integral surface Gaussian curvature	3	$\frac{1}{3}\left(\int_{\partial A}(\frac{1}{R_1}\frac{1}{R_2})ds\right)$

Table 5.1: Minkowski functional in 3-dimensional Euclidean space for a set A included in E. $1/R_1$ and $1/R_2$ are the principal curvatures.

Hadwiger conditions:

1. Motion (translation and rotation) invariance, $W_i(A) = W_i(gA)$ with $g =$ rotation or translation ,
2. Homogeneity condition, $\forall\lambda \in \mathbb{R} : W_i(\lambda A) = \lambda^{d-i} W_i(A)$,
3. Additive, $W_i(A \cup B) = W_i(A) + W_i(B) - W_i(A \cap B)$,
4. Convex continuity, the application W_i is continuous for the set of convex bodies.

Integral geometry shows how the sections of the 3D sample allow us to recover its Minkowski functionals. This very attractive property, due to Crofton, open up the a fruitful path for the study of stereology[55] [158, 112, 79, 113] (see table 5.2). The first two Minkowski functionals will be used in the two next subsubsections. Since the last Minkowski functional is a topological number without stereological relation, it will be presented in the subsection 5.3.1.

5.2.2 Volume fraction

Methods

The volume fraction of the different phases, ϕ_i also called the 1-point probability function, is one of the most important geometrical property since it influences physical properties.
The evaluatation of the volume fraction of the void phase can be done in using the following techniques:

[55] The stereology was developed by the French geologist A.E. Delesse in 1845 who has proposed a quantitative analysis of a polished plane section of a rock. Assuming that rock under image analysis investigation is homogeneous, he has proved that [40]:

$$\phi^{3d} = \phi^{2d}$$

where ϕ^{3d} is the volume fraction of a given phase estimated in a 3d sample and ϕ^{2d} is the surface fraction of the same phase estimated on a 2D cut of the sample.

Space of dim. n	Stereological relations			
$n=3$	Volume fraction $\phi = V_V$	Specific surface area S_V	Mean curvature	$\frac{1}{3}$ Gauss curvature
	$\parallel$	$\parallel$	$\parallel$	
$n=2$	Surface fraction A_A	$\frac{4}{\pi}$ Perimeter per surface U_A	$2\pi N_c$ per surface	
	$\parallel$	$\parallel$		
$n=1$	Length fraction L_L	N_c per length		

Table 5.2: Stereological relations. N_c is the number of connected components.

1. the mercury intrusion porosimetry [84, 1],
2. the water vapour adsorption, and nitrogen adsorption using Gurvitch's law in the saturation regime [35, 64].

The drawbacks are that these methods are intrusives (during drying, the geometrical organisation is affected) and that these methods cover the whole range of pore sizes including the small pores, capillary pores and large air voids. In order to focus the analysis of one pore size, major assumptions on the pores geometry have to be done like cylindrical and ink-bottle pores [48].
In the case of cement paste, the volume fraction of the anhydrous grains is typically evaluated using:

1. the loss-on-ignition[56] [9, 61],
2. the analysis of X-Ray-Powder Diffraction Patterns by the Rietveld method [177, 38].

By image analysis, it is simple to evaluate the volume fraction of the different phases in 2D and in 3D according to:

$$\phi_i = < I_i(\vec{r}) > \tag{5.1}$$

$$= \frac{\int_E I_i(\vec{r}) d\vec{r}}{\int_E d\vec{r}} \tag{5.2}$$

$$= \frac{\text{sum pixels/voxels belonging to the phase}}{\text{sum pixels/voxels of the image}} \tag{5.3}$$

Obviously, the sum of the various volume fractions is one. For a homogeneous sample, as explained in the previous paragraph, the volume fraction calculated in a volume is equal to the surface fraction calculated on a slice of this previous volume.
An important feature in cement paste, the degree of hydration, can be estimated using the volume fraction of the anhydrous phase. After mixing cement with water, the dissolution of the anhydrous grains and the precipitation of the hydrate phase begin. The degree of chemical reaction of the anhydrous grains at time t, called degree of hydration, $\alpha_c(t)$, is defined as:

$$\alpha_c(t) = \frac{\phi_c(t=0) - \phi_c(t)}{\phi_c(t=0)} \tag{5.4}$$

where $\phi_c(t)$ is the volume fraction of the anhydrous grains at time t in cement paste. The reaction of dissolution/precipitation depends on many parameters such as, the time of hydration, the ratio the water to cement ratio, the temperature, the chemical composition of the cement. Therefore, for a comparative analysis, the degree of hydration makes more sense than the time of hydration.
The volume fraction of anhydrous grains at time t, $\phi_c(t)$, is calculated on the SEM/tomography images whereas the volume fraction of the anhydrous phase at time t=0 is calculated from the original mix design (water to cement ratio and density of the cement): [57]

$$\phi_c(t=0) = \frac{1}{1 + \frac{\rho_c}{\rho_w} r(w/c)} \tag{5.6}$$

[56] http://en.wikipedia.org/wiki/Loss_on_ignition
[57] Let us consider the mixture of anhydrous cement and water at time t=0. We have:

$$\phi_w(t=0) + \phi_c(t=0) = 1 \tag{5.5}$$

where ρ_c and ρ_w are the densities of cement and water and $r(w/c)$ is the water to cement mass ratio.

Experimental results

The figure 5.4 represents the degree of hydration as a function of the time of hydration. This figure shows that, at the same time of hydration, the degree of hydration obtained by tomography is significantly higher than the degree of hydration obtained by SEM for both cements (the difference is higher for alite than OPC). The reason is that the reaction of hydration/precipitation depends on the temperature [58]. The samples, observed by SEM, have been stored in a room at constant temperature of 16^oC whereas the samples, observed by tomography, have been prepared and immediately carried to the SLS during a warm day 30^oC.
To make possible the comparison, we always express the geometrical feature as a function of the degree of hydration and not the time of hydration. For OPC cement, there is a common range of the degree of hydration between tomography and SEM that allows a comparative analysis between both techniques. For alite cement, there is no a common range of the degree of hydration. Therefore, a comparative analysis cannot be done.

The figure 5.5 represents the volume fraction as a function of the degree of hydration. The increase of porosity for the last point obtained by SEM for OPC is in opposition with the decrease of porosity due to the precipitation. Qualitatively (see figure 5.6), we can make the same comment. Since cement paste is not a homogeneous medium, the slice cut at 7 days is an heterogeneity. Therefore, we will not take into account the last point obtained by SEM. For the remaining points, the volume fraction obtained by SEM and X-ray tomography for OPC are coherent. As the ratio W/C of OPC for X-ray tomography is smaller the ratio W/C of OPC for SEM, the fraction of porosity is smaller.

5.2.3 Specific surface area

Methods

The specific surface area, the ratio of the surface to the volume, has a particular importance in the case of adsorption, heterogeneous catalysis, and reactions on surfaces.
To evaluate the specific surface area, it is current to use the following techniques:

1. the adsorption using Brunauer, Emett and Teller (BET) method [27, 127],
2. the small-angle neutron scattering in the Porod regime [39, 179, 167]

For a volume V of cement paste (anhydrous cement+water), we have:

$$\begin{aligned} V\phi_w(t=0)\rho_w &= m_w \\ V\phi_c(t=0)\rho_c &= m_c \end{aligned}$$

Taking the ratio of these two equalities, we get:

$$\frac{\phi_w(t=0)\rho_w}{\phi_c(t=0)\rho_c} = \frac{m_w}{m_c}$$

As $\frac{m_w}{m_c} = r(w/c)$, noted W/C in the literature of cement paste, we obtain:

$$\phi_w(t=0) = \phi_c(t=0)\frac{\rho_c}{\rho_w} r(w/c)$$

Replacing $\phi_w(0)$ in the equation 5.5, we obtain:

$$\phi_c(t=0) = \frac{1}{1+\frac{\rho_c}{\rho_w} r(w/c)}$$

A simple experiment to evaluate the density of cement is to mix water with a mass m_w and cement with a mass m_c in a volume V and to use this relation:

$$\rho_c = \frac{m_c}{1-\frac{m_w}{V\rho_w}}$$

Knowing that $\rho_w = 1kg.m^{-3}$ and $\rho_c \simeq 3.15kg.m^{-3}$ for the all the cements, we can evaluate the volume fraction of anhydrous cement at the initial state.

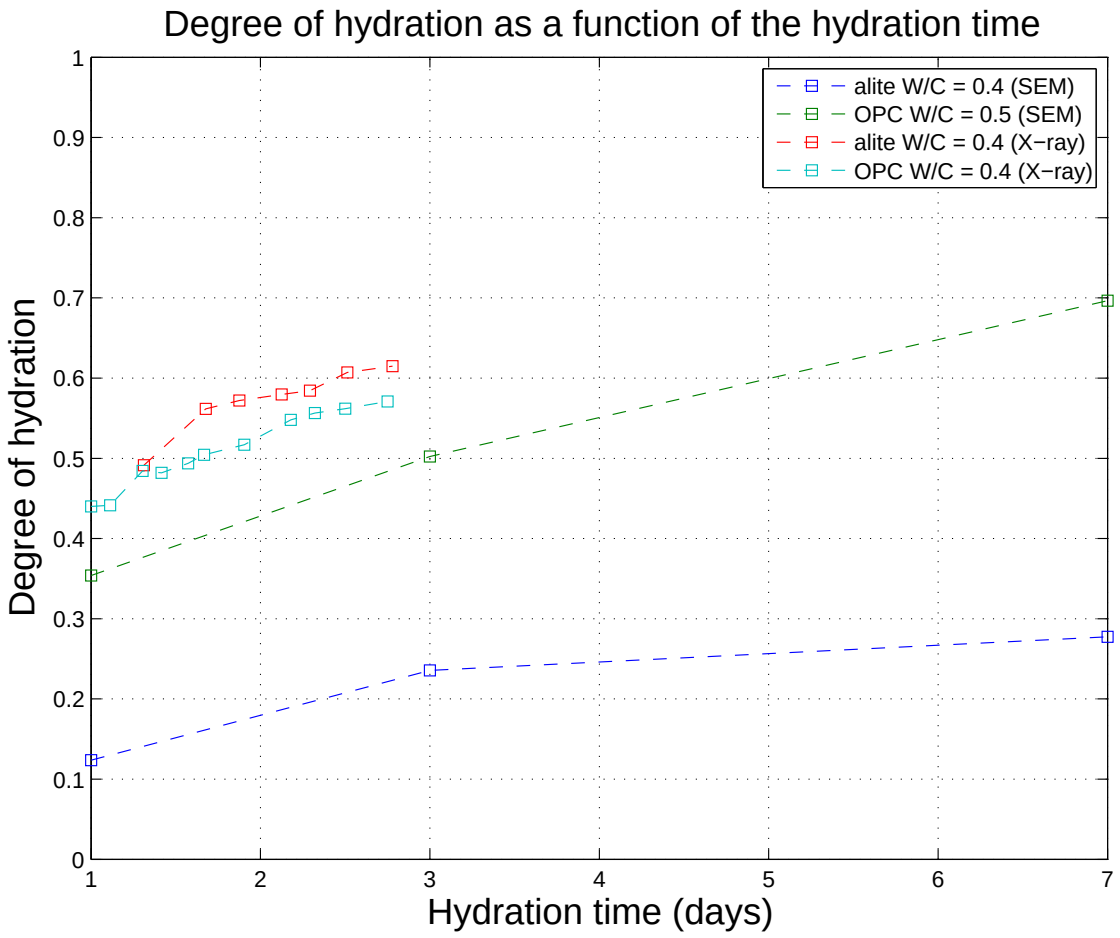


Figure 5.4: Degree of hydration for alite/OPC obtained by SEM/X-ray

By image analysis, the value obtained for specific surface area depends on the method of evaluation (see figure 5.7). Numerically, the specific surface area is estimated with: :

$$\mathcal{S}_i(\Omega_i) = < |\partial\Omega_i| > \tag{5.7}$$

$$= \frac{1}{2}\frac{\sum_{x\in E}\sum_{\forall y\in B_{1,N_1}(x)\setminus\{x\}} I(x)I(y)}{\mu(E)} \tag{5.8}$$

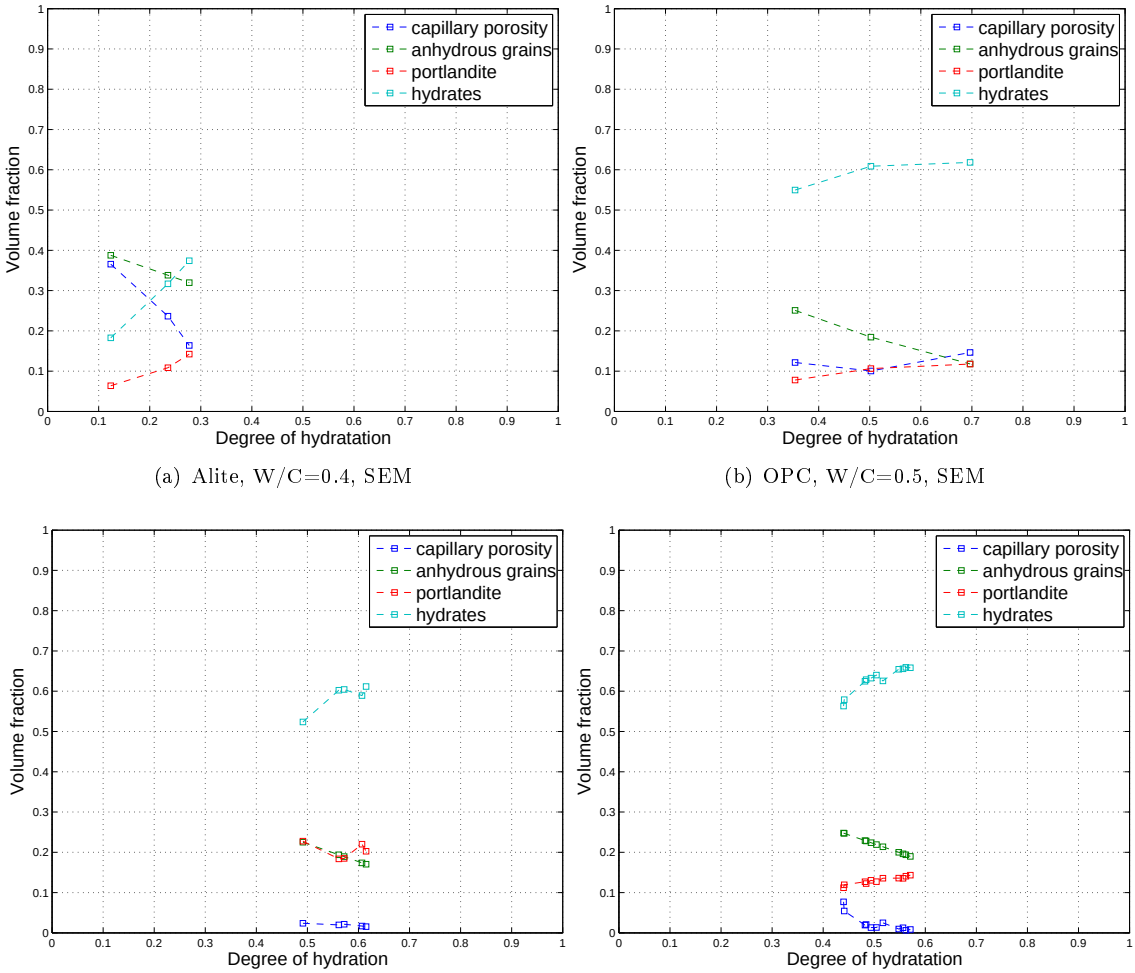
where $B_{1,N_1}(x)\setminus\{x\}$,is the 4-neighbourhood in 2D and the 6-neighbourhood in 3D (see the algorithm in the appendix A.5).
In euclidean space, the specific surface area $\mathcal{S}_i^{d=3}(\Omega_i)$ is equal to the perimeter per surface, $\mathcal{S}_i^{d=2}$ time $\frac{4}{\pi}$:

$$\mathcal{S}_i^{d=3}(\Omega_i) = \frac{4}{\pi}\mathcal{S}_i^{d=2}(\Omega_i \cap \mathcal{C}) \tag{5.9}$$

where $\mathcal{C}$ is a plan of E ($\Omega_i \cap \mathcal{C}$ represents a 2D cut of the sample.). As the specific surface area is evaluated in discrete space, the stereological relation is:

$$\frac{\mathcal{S}_i^{d=3}(\Omega_i)}{3} = \frac{\mathcal{S}_i^{d=2}(\Omega_i \cap \mathcal{C})}{2} \tag{5.10}$$

In order to compare the measures done with SEM and with tomography, we use the relation of stereological specific surface defined in discrete space, $\frac{\mathcal{S}_i^{d=3}(\Omega_i)}{d}$.



(a) Alite, W/C=0.4, SEM

(b) OPC, W/C=0.5, SEM

(c) Alite, W/C=0.4, X-ray tomography

(d) OPC, W/C=0.4, X-ray tomography

Figure 5.5: Volume fraction as function of the degree of hydration

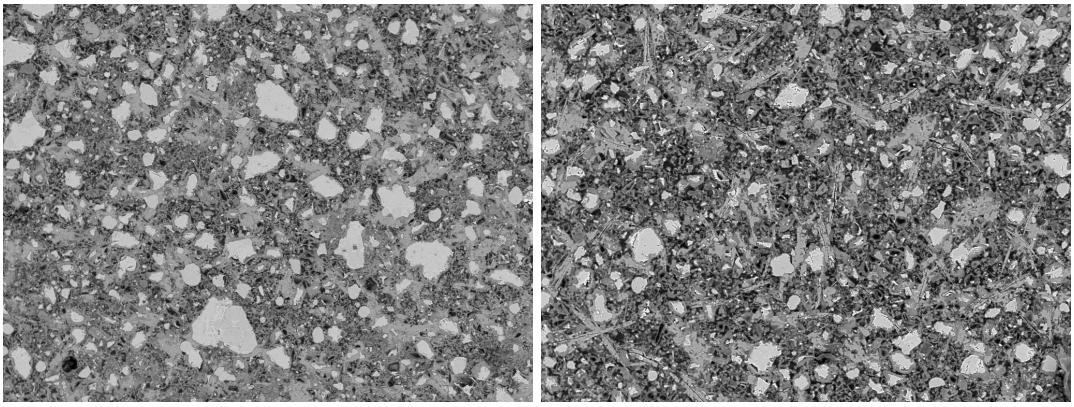
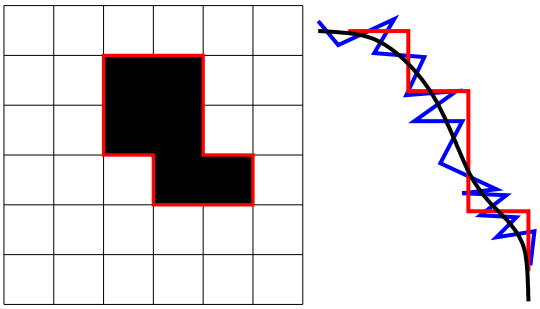


Figure 5.6: on the left: OPC after a setting of 3 days, on the right: same OPC after a setting of 7 days.



(a) Numerical boundary in red (b) Numerical and real boundaries

Figure 5.7: Figure (a): The perimeter in 2-dimensional discrete space is equal to 12 l where l is the image resolution. Figure (b): the organisation of the surface is blurred by the pixel-effect. The boundary on the segmented image is represented in red. The real surface can be a smooth surface in black or a rough surface in blue.

Experimental results

For the portlandite phase and the anhydrous grains, we observe an agreement between the results obtained by SEM and by tomography for OPC (see figure 5.8). For the other phases, hydrate and porosity, this is not the case. Porosity and specific surface area are merely two numbers characterising the geometric properties of a material. Obviously, these two numbers are not sufficient to characterise the morphology of a complex medium like Porous Medium of Cement Paste (PMCP). To do an extended analysis, a stereological statistical analysis will be done using two correlation functions: 2-point probability function and chord length distribution function in the next two subsections.

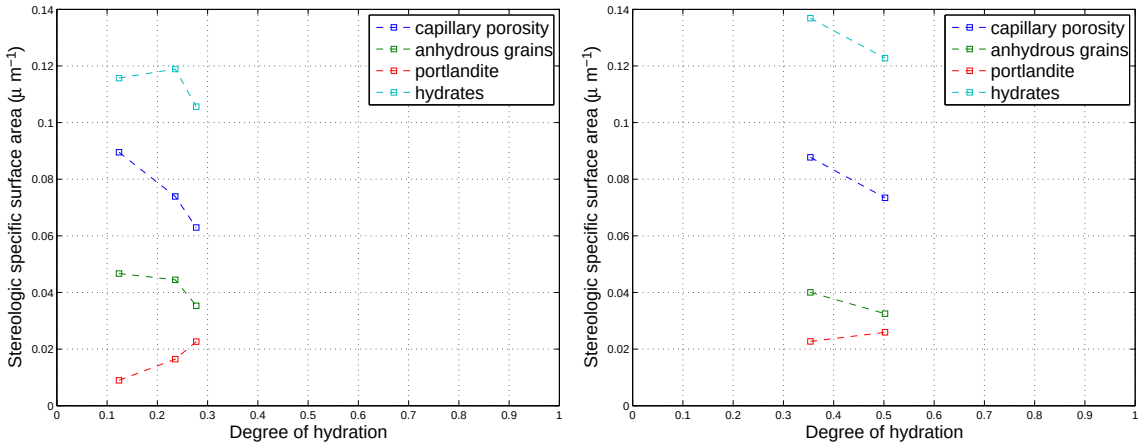
5.2.4 2-point probability function/Covariance

Let draw an interval of length r randomly in the material. The two-point probability functions/covariance gives the probability of having both extremities of this interval belonging to the phase i, $S_{2,i}(\vec{r})$ (see figure 5.9):

$$S_{2,i}(\vec{r}) = P(\vec{r} \in \Omega_i, \vec{r'} + \vec{r} \in \Omega_i) \quad (5.11)$$

$$= < I_i(\vec{r'}) I_i(\vec{r'} + \vec{r}) >_{\vec{r'} \in E} \quad (5.12)$$

This structural correlation quantifies how the "memory" of an initial state is progressively lost when a point is moved away [103, 79]. This quantity has been extensively studied theoretically [33, 169]. Small-angle X-ray or neutron scattering is used for the experimental evaluation [39]. The evaluation of the



(a) Alite, W/C=0.4, SEM

(b) OPC, W/C=0.5, SEM

(c) Alite, W/C=0.4, X-ray tomography

(d) OPC, W/C=0.4, X-ray tomography

Figure 5.8: Specific surface area as function of degree of hydration.

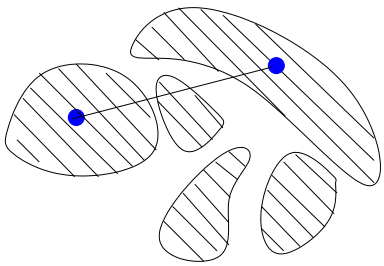


Figure 5.9: 2-point probability function. An interval of length r drawn randomly in a bi-phase material

2-point probability function by Fourier transform is performed using Matlab© [58] (see appendix A.5.2). 2-point probability function allows the evaluation of the material isotropy. The material is isotropic if the 2-point probability function is invariant by rotation:

$$S_2(\vec{r}) \quad = \quad S_2(R\vec{r}) \text{ with } R \text{ a rotation}$$

Practically, the isotropy is characterised by a series of rings in the figure of 2-point probability function (see figure 5.10). Due to isotropy of cement paste, 2-point probability function can be plotted as a function of the norm $l = \|\vec{r}\|_2$:

$$S_2(l) \quad = \quad < S_2(\vec{r}) >_{\|\vec{r}\|_2=l} \tag{5.23}$$

2-point probability function allows the evaluation of the characteristic length, l_c, of the representative elementary volume. The "memory" of an initial state is progressively lost when a point is moved away. When $\|\vec{r'}\|$ tends to ∞, we have, as long as there is no long-range order,[59]:

$$\lim_{\|\vec{r'}\|\to\infty} P(\vec{r} \in \Omega_i, \vec{r} + \vec{r'} \in \Omega_i) \quad = \quad \lim_{\|\vec{r'}\|\to\infty} P(\vec{r} \in \Omega_i) P(\vec{r} + \vec{r'} \in \Omega_i) \text{ as independant events} \tag{5.24}$$

$$= \quad \phi_i^2 \text{ as 1-point probability function equals to } \phi \tag{5.25}$$

Let l be the distance when the graph tends to ϕ_i^2. By convention, the length of REV is [60] defined as 3 times this length, l (see figure 5.11). The length of REV is 200 μm for OPC and 600 μm for alite since there are very large clusters of portlandite (see figure 5.3). As the length of the field of view in X-ray tomography/SEM is 500 μm, the observation is done in a representative elementary volume for OPC, but not for alite.

The 2-point probability function graph allows the estimation of the specific surface area, $\mathcal{S}_1^{d=2}$:

[58] Let $\mathcal{F}[I]$ be the Fourier transform of I:

$$\mathcal{F}[I](\vec{q}) \quad = \quad \int_E e^{-i\vec{r}\vec{q}} I(\vec{r}) d\vec{r} \tag{5.13}$$

Let $\mathcal{I}$ be the signal intensity:

$$\mathcal{I}(\vec{q}) \quad = \quad |\mathcal{F}[I](\vec{q})|^2 \tag{5.14}$$

$$= \quad \mathcal{F}[I](\vec{q})(\mathcal{F}[I](\vec{q}))^* \tag{5.15}$$

$$= \quad \int_E e^{-i\vec{r}\vec{q}} I(\vec{r}) d\vec{r} \int_E e^{i\vec{r'}\vec{q}} I(\vec{r'}) d\vec{r'} \tag{5.16}$$

$$= \quad \int_E \int_E e^{-i(\vec{r}-\vec{r'})\vec{q}} I(\vec{r}) I(\vec{r'}) d\vec{r'} d\vec{r}. \tag{5.17}$$

Let $(\vec{r'}, \vec{r}) = (\vec{r'}, \vec{u} + \vec{r'})$ be the change of variable which Jacobian is equal to 1. We get:

$$\mathcal{I}(\vec{q}) \quad = \quad \int_E \int_E e^{-i\vec{u}\vec{q}} I(\vec{u} + \vec{r'}) I(\vec{r'}) d\vec{r'} d\vec{u} \tag{5.18}$$

$$= \quad \int_E e^{-i\vec{u}\vec{q}} \int_E I(\vec{u} + \vec{r'}) I(\vec{r'}) d\vec{r'} d\vec{u} \tag{5.19}$$

$$= \quad \int_E e^{-i\vec{u}\vec{q}} s_2(\vec{u}) d\vec{u} \tag{5.20}$$

where $s_2(\vec{r}) = S_2(\vec{r})V(E)$. So $\mathcal{I}(\vec{q})$ is the Fourier transform of the correlation function.

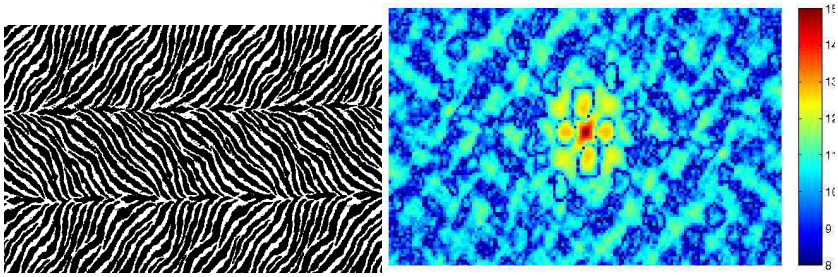
$$\mathcal{I}(\vec{q}) \quad = \quad \mathcal{F}[s_2](\vec{q}) \tag{5.21}$$

$$s_2(\vec{r}) \quad = \quad \mathcal{F}^{-1}[\mathcal{I}](\vec{r}) \tag{5.22}$$

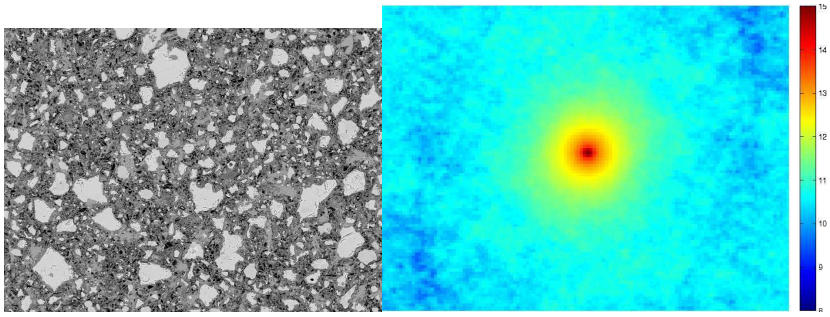
In other words, we use this property $\mathcal{F}[f \star g](\vec{q}) = \mathcal{F}[f](\vec{q})\mathcal{F}[g](\vec{q})$ to substitute g by f and in using $s_2 = f \star f$.

[59] for an non periodic medium

[60] In order to estimate a given physical/geometrical property of porous medium, we investigate the sample in a given volume for a set of positions. If the volume is too small, we will observe a fluctuation between each measure. If we increase the volume, the measures will remain stable progressively. When the stability is reached, then the investigation volume is the Representative Elementary Volume (REV). The REV depends on the material and also the physical/geometrical property under investigation. Numerically, the determination of length of REV is similar. We estimate a given physical/geometrical property on a set of sub volumes randomly located in the initial volume. If the measures remain stable, then the investigation volume is a representative elementary volume [98]. Last point, the length of representative elementary surface is different than the length of representative elementary volume.



(a) Zebra texture (b) 2-point probability function of the zebra texture



(c) OPC, W/C=0.5, SEM, 1 day (d) 2-point probability function of OPC

Figure 5.10: Test of the isotropy of a material in using the 2-point probability function. To observe the difference, we use an image of zebra texture. First row: there are some correlation peaks following two directions in the 2-point probability function. Therefore, the zebra texture is anisotropic. This anisotropy information is used by Redon *et al* [141] to get the orientation of fibres inside a reinforce of concrete matrix. Second row: for the OPC image, the correlation function shows invariance by rotation, so OPC is isotropic like alite.

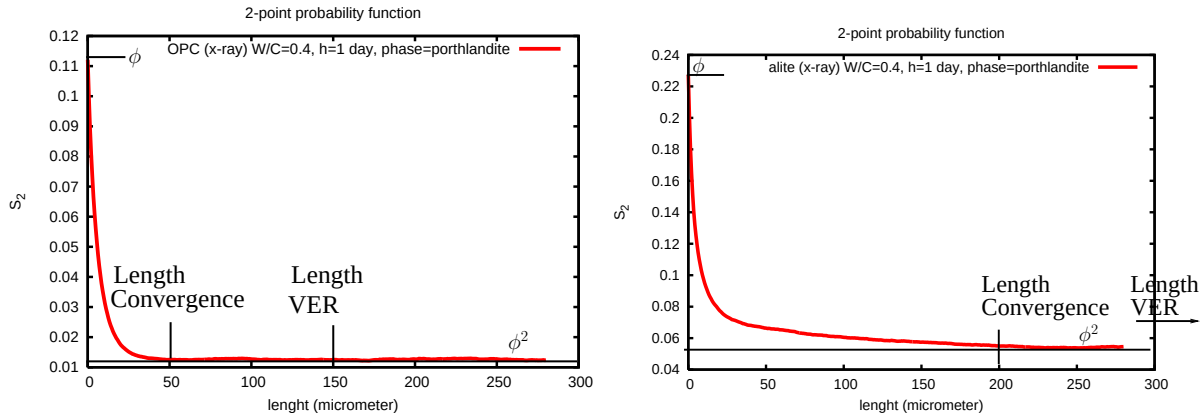


Figure 5.11: Characteristic length of the REV. For OPC, the field of view is above the length of the REV. For alite, the field of view is not above the length of the REV due to the existence of large portlandite clusters.

$$\left.\frac{dS_2(l)}{dl}\right|_{l=0} = -\frac{\mathcal{S}_1^{d=2}}{4}. \tag{5.26}$$

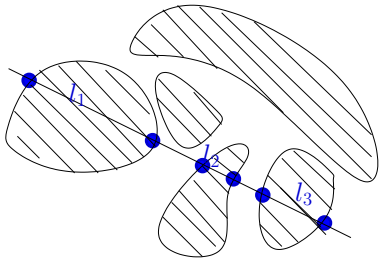


Figure 5.12: A random line through a granular material composed by two phases allows to extract chords belonging either to the solid phase or to the pore network.

We find a good agreement between the direct calculus of the specific surface area and the calculation using the derivative of 2-point probability function at the origin. For vycor glass [105] (7930, lot 742098, Corning) of size 500*500*500, the direct calculation gives: 0.108 and the slope at the origin of 2-point probability function gives 0.106.

5.2.5 Chord length distribution function

Methods

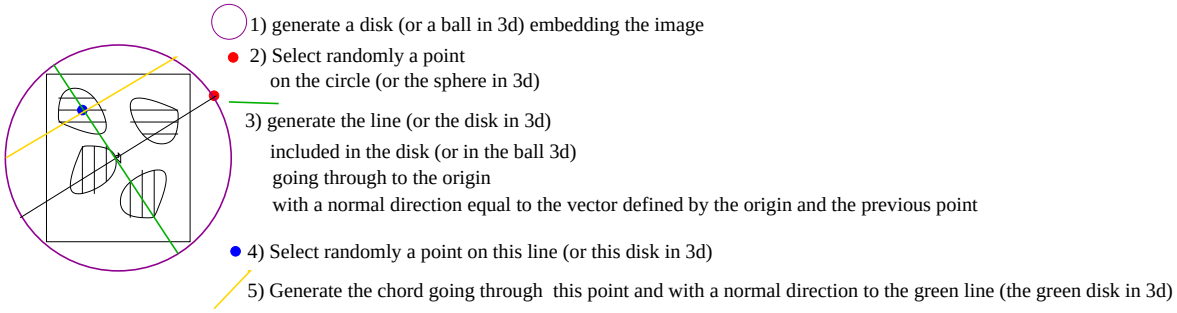
μ-chord A chord is a segment belonging to a phase and having its two extremities on the phase interface. As shown in Fig. 5.12, μ-chord are obtained by tracing random and homogeneous distributed straight lines through the microstructure. The chord-length distribution function gives the probability $f_{i,\mu}(r)$ of getting a chord length between r et $r + dr$, belonging either to the phase.
The chord-length distribution function is a fingerprint of the local morphology of the considered phase. This quantity has been extensively studied theoretically [115, 106, 168, 111]. Small-angle X-ray scattering can be used to evaluate it experimentally [115, 106]. Numerically, the definition of the distribution of straight lines is not evident. Miles and Davy [117] give a precise definition of the random sampling of straight lines, called Isotrope Uniform Randomness (see figure 5.13).

Sharp edges As demonstrated by Mering and Tchoubar [115], the chord-length distribution function is a measure of the existence of sharp edges of the considered phase since $f_{i,\mu}(0) \neq 0$ means that the phase has sharp edges. As a voxel is a cube with sharp edges (see figure 5.14), the value at the origin of the chord-length distribution function will detect this numerical artefact (see figure 5.13-c).

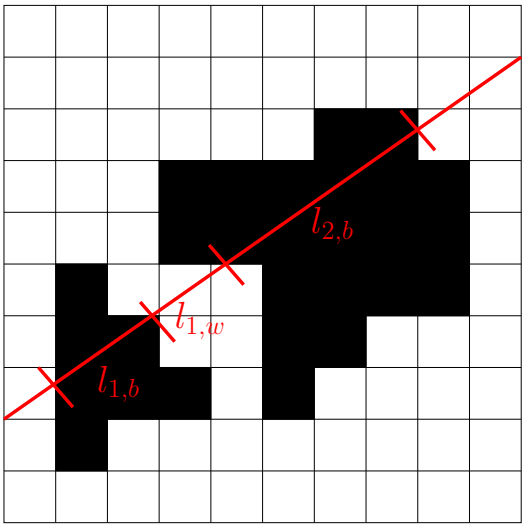
Nd-chord This drawback leads us to consider Nd-chord which straight lines are sampled following a random direction in the orthogonal base of the space and going through a random point of the microstructure (sees figure 5.15).

Catalogue the phase depending on the shape of the distribution As explained by Levitz *et al* [106], a phase can be catalogued in three kinds of disorder depending on the distribution of chord length:

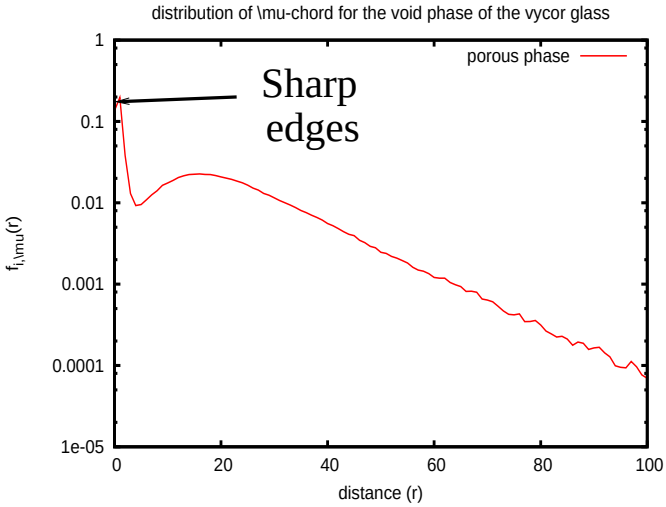
1. LONG-RANGE DEBYE RANDOMNESS. Rapidly, the distribution follows negative exponential (see figure 5.16-a).
2. "CORRELATED" DISORDER. Just after the peak distribution, the distribution follows negative exponential (see figure 5.15). The small-angle scattering of some porous solids shows a peak corresponding to the existence of a relatively well defined correlation length.
3. COMPLEX STRUCTURES WITH LENGTH SCALE INVARIANCE. At large distance, the phase chord distribution exhibits a $1/r^n$ form. In log-log scale, a tail is observed in large distance.



(a) Random sampling of straight lines



(b) Chord lengths for a straight line



(c) Chord length distribution function for vycor

Figure 5.13: Random sampling using the μ-chord distribution. Application for the void phase of vycor glass.

Mean chord length: a way to capture pore size The mean chord length, l_i, is:

$$l_i = \int_0^\infty l f_i(l) dl \tag{5.27}$$

The mean chord length is a characteristic length scale of the given phase[61]. Since the chord length distribution function is stereological [170], this length scale is accessible either in 2D and in 3D. Let f_p^μ be the chord-length distribution function of the phase and $f_{p^c}^\mu$ be the chord-length distribution function of the complement of the phase.
The volume fraction of the phase, ϕ_p, the surface specific area, $\mathcal{S}_p$ and the star volume, V_p^* are formulated by:

$$\phi_p = \frac{l_p}{l_{p^c} + l_p} \tag{5.28}$$

$$\mathcal{S}_p = \frac{4}{l_{p^c} + l_p} \tag{5.29}$$

$$V_p^* = \frac{\pi}{4 l_p} \int_0^\infty l^4 f_p(l) dl \tag{5.30}$$

[61] For mono-disperse sphere packing of radius, r, we have $l_i = \frac{4}{3} r$

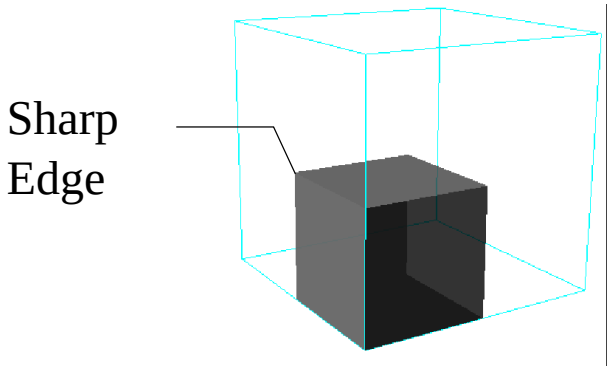
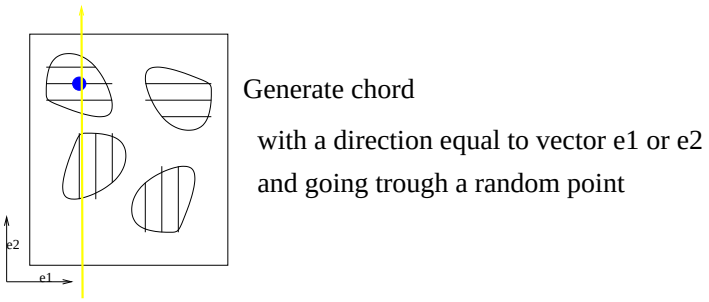
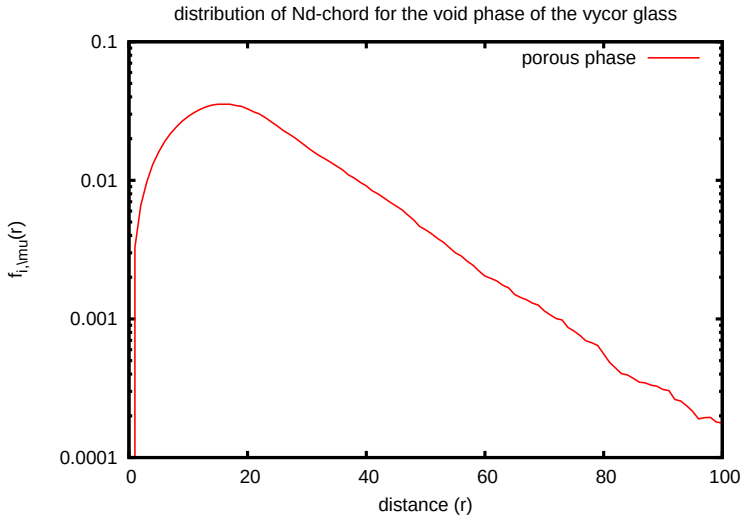


Figure 5.14: A voxel has sharp edge.



(a) Random sampling of straight line



(b) Chord length distribution function for vycor

Figure 5.15: (a): random sample of a straight line; figure 2: Nd-chord distribution of void phase of the vycor glass. We have $f_{i,\mu}(0) = 0$

For vycor glass [105] (7930, lot 742098, Corning)) of size 500*500*500, the direct calculation of volume fraction of the pore phase gives: 0.300 and the calculation using the mean chord length gives 0.310. The the direct calculation of surface specific area gives of the pore phase gives 0.108 and the calculation using the mean chord length gives $0,0775$.

Lineal path function In mathematical morphology, another statistical function is used, lineal path function. The chord length distribution is related to lineal path function [111], $P(l \subset \Omega_i)$, given the probability that a yardstick, l, drawn randomly in the material belongs to the phase i, by this formula [60, 168]:

$$f_i(l) = \frac{l_i}{\phi_i} \frac{d^2 P(l)}{dl^2} \tag{5.31}$$

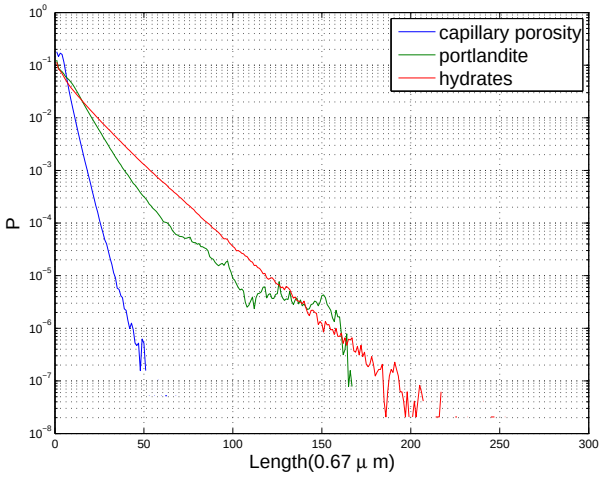
Experimental results

In the case of cement paste, the evolution of the chord length distribution of portlandite, porosity and hydrate phases is a negative exponential without a peak of correlation. It is the signature of a long-range Debye randomness (figure 5.16-a). The figure 5.16-a figure 5.16-b exhibit the same distribution calculated on SEM image and tomography image. This measure is stereological and it can be done directly in a 2D slice. The figure 5.16-c shows that the chord length distribution is independent of the resolution. The disorder of the pore network is similar in this range length scale (from $0.1\mu m$ to 0.5 mm) and it is not a length scale invariance. The figure 5.16-d shows that the chord length distribution of the anhydrous phase follows a power law. It is the signature of length scale invariance. An explanation will be that the poly-dispersion of grains follows a power law (see annexe A.6).

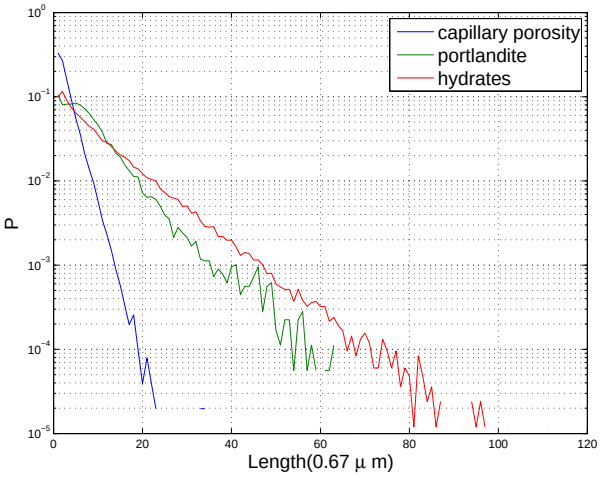
The figure 5.17 shows the evolution of mean chord length as a function of the degree of hydration. The evolution of mean chord length of anhydrous phases (green graphs) seems to follow two regimes. At the beginning of hydration, the mean chord length of anhydrous phase decreases and, then, the value stays constant although the degree of hydration increases. An explanation is that, at the beginning, the dissolution of the anhydrous grains is the same whatever the size of the grains, then this kinetic becomes faster for small grains than for big grains.

5.2.6 Volume distribution of connected components: non-stereological

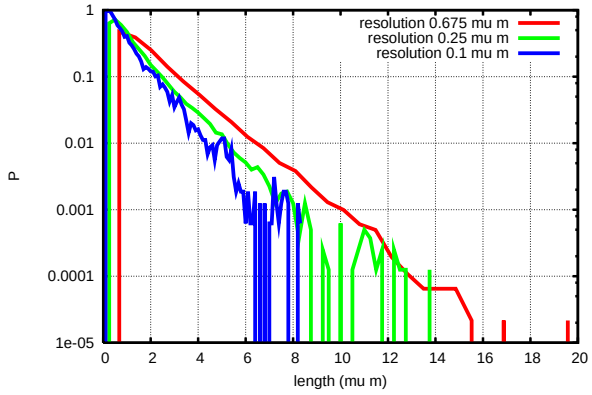
This measures is only accessible in using a 3d imaging technique. In a recent work, Landis *et al* [97] observe the evolution of the distribution of volume and surface area of connected components of cracks porosity in order to give some relationships between cracking and the stress states. To confirm the power law observed by using chord length distribution of the anhydrous grains, we calculate the volume distribution of connected components of anhydrous grains (see figure 5.18). We find a power law often observed during the crush of a material [140]. In order to understand the dissolution of anhydrous grains, we plot the volume distribution of connected components as a function of the degree of hydration (see figure 5.18). We observe distribution invariance about the degree of hydration in the range between [0.42,0.58]. This observation should be confirmed for a larger range of degree of hydration.



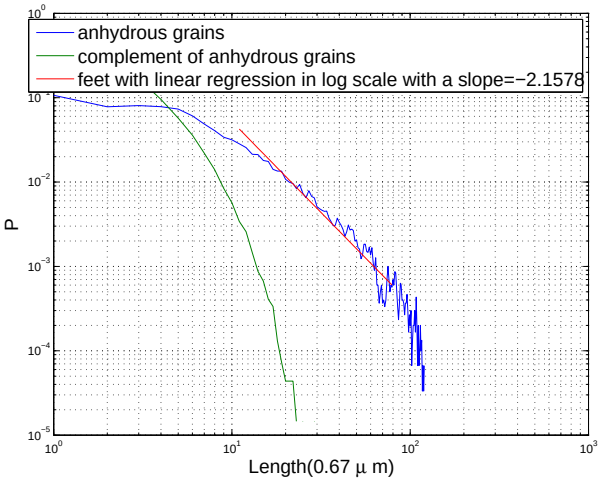
(a) OPC, W/C=0.5, SEM, $\alpha = 0.502$



(b) OPC, W/C=0.4, X-ray tomography, $\alpha = 0.440$

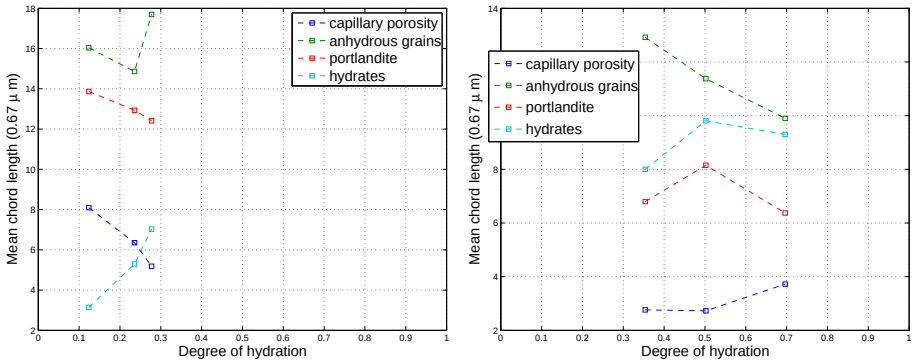


(c) OPC, W/C=0.5, SEM, $\alpha = 0.354$



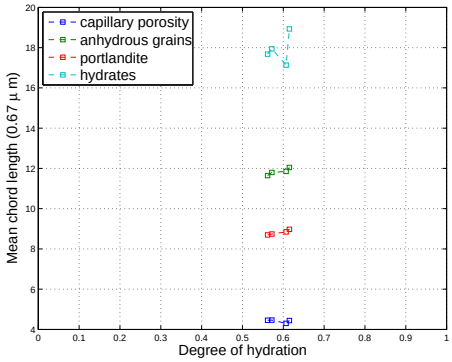
(d) OPC, W/C=0.5, X-ray tomography, $\alpha = 0.440$, log-log scale

Figure 5.16: Chord length distribution.

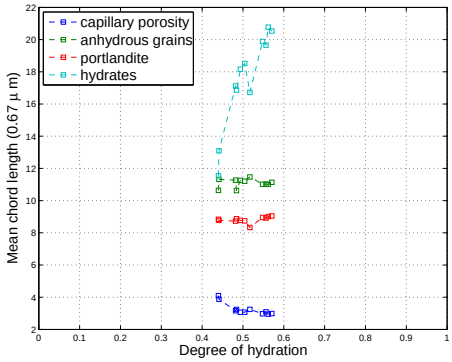


(a) Alite, W/C=0.4, SEM

(b) OPC, W/C=0.5, SEM

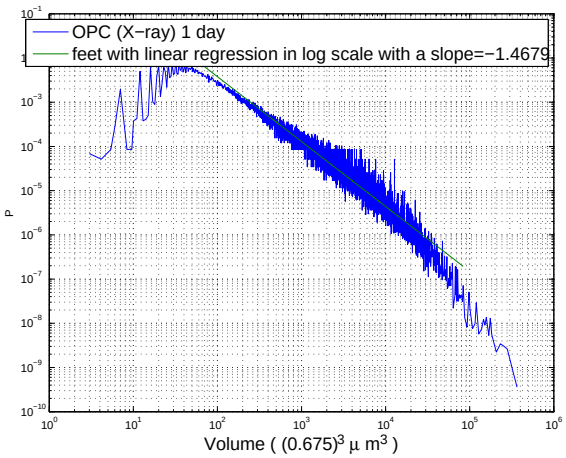


(c) Alite, W/C=0.4, X-ray tomography

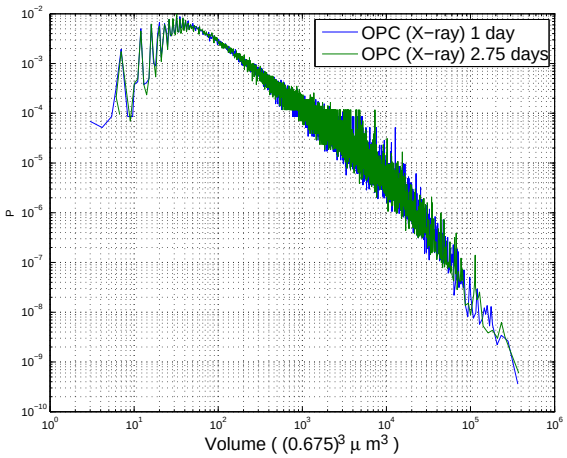


(d) OPC, W/C=0.4, X-ray tomography

Figure 5.17: Mean chord length as function of degree of hydration.



(a) OPC, W/C=0.4, X-ray tomography, $\alpha = 0.422$



(b) OPC, W/C=0.4, X-ray tomography, $\alpha = 0.422$ and 0.580

Figure 5.18: left figure: the distribution of volume of grains of anhydrous phase follows a power law with a cut off at 10^5 voxels. This length scale cut can be due to the sieve during the crush of clinker. Right figure: the distribution of volume of grains of anhydrous phase at degree of hydration=0.42 and at degree of hydration=0.58. The distribution is independent of the degree of hydration in this range of degree of hydration.

5.3 Topological characterisation

The branch of mathematics, called topology, began with the investigation of certain questions in geometry. At the beginning, Euler introduced the basic notions of topology in solving an enigma called seven Bridges of Konigsberg. To give some intuition about the topology, let us consider once again the analogy of car traffic. Starting from Poitiers, a car driver wants to go to Bordeaux. He has the choice between two mains roads (see figure 5.19): the first one is a highway going through Niort/Saintes and the second one is a road going through Angoulême. In a normal day, the fast path is the high way. However, during a day of intense car traffic, to avoid the traffic jam on the highway, a solution is to take the path of the road. This last path is always taken by the truck driver because this path is shorter in length. The topology is related to the connectivity of the network and it plays a crucial role to understand the transport.
X-ray tomography allows the observation of the larger capillary pores of cement paste. At the beginning

Figure 5.19: Road map of the beautiful region of Charente

of the hydration time, these pores are highly connected. Therefore, the global transport is driven by the transport at this scale. However, when the hydration progresses, the larger capillary pores become weakly connected. Thus, there is a coupling between the transport at this scale and the transport at smaller scales leading to an intermittent transport. One aim of this section is to define the transition between the transport at one scale and the intermittent transport.
In this section, the analysis will be done with some scalars and with some statistical function. The system in evolution will be the capillary pores of OPC with W/C=0.5 observed by X-ray tomography at different hydration times.

- at time t=35h, the porosity is equals to 0.205 with a degree of hydration equals to 0.48.
- at time t=83h, the porosity is equals to 0.176 with a degree of hydration equals to 0.62.
- at time t=150h, the porosity is equal to 0.079 with a degree of hydration equals to 0.71.

5.3.1 Minkowski functional: Euler-Poincaré/Gaussian curvature

Methods

The 3-Minkowski functional, $W_3(A)$, where A is the object, is defined by:

$$3W_3(A) = 4\pi N_3(A) \quad (5.32)$$

$$= \int_{\partial\Omega} K(s)ds \quad (5.33)$$

where N_3 is the Euler-Poincaré invariant (see figure 5.20) and K is the Gaussian curvature (see figure 5.21).

The Gaussian curvature has important role physically. For instance, in the case of foam, the gauss

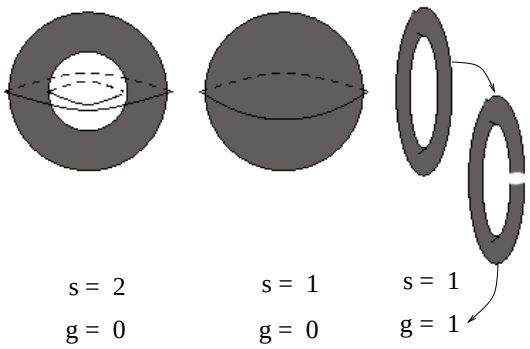


Figure 5.20: We have $N_3(S) = N(S) - \sum_{s\in S} g(s)$ where $N(S)$ is the number of the isolated surface and $g(s)$ is the genus of the surface s (the number of tore in a connected sum decomposition of the surface; intuitively, the number of "handles"). The first set is a hollow sphere with two isolated surfaces, $N(S) = 2$, with a genus equal to 0 for both surfaces; the second set is a sphere with one isolated surface; the last set is a tore with one isolated surface. This surface can be cut one time with a closed line without separation of the surface so the genus is equal to 1. A good property of Euler-Poincaré invariant (homotopy invariant) is the additive which genus does not have.

curvature is related to the bending energy of the interface of surfactant film [68, 76].
The estimation of the Euler-Poincaré/Gaussian curvature can be done in using its additive property [158, 135]:

$$N_3(A \cup B) = N_3(A) + N_3(B) - N_3(A \cap B)$$

Another solution is the utilisation of the relation with the Gaussian curvature (see equation 5.33) presented in the next subsubsection.

Gaussian curvature in a cubic space If the discrete space is cubic, a method to evaluate the Gaussian curvature [76, 7] is to list all patterns of a unit cell (see figure 5.22). For each pattern, we calculate the local integral curvature using the equivalence between the integral curvature and the Gaussian map of the surface, a theorem of differential geometry (see figure 5.23). At this stage, any configuration of a unit cell can be associated to one pattern with the same local integral curvature. This allows the creation of a look-up table (see figure 5.22). Since the global integral curvature is the sum of local integral curvatures, the sum of all unit cell curvatures gives the global curvature of the microstructure.

Experimental results

In the figure 5.24, the graph shows the evolution of the Euler-Poincaré characteristic for the capillary pore network of OPC. A topological conjecture is if $N3 < 0$, then the pore network is connected and the opposite. This conjecture seems to be true for OPC (see subsection 5.3.2)

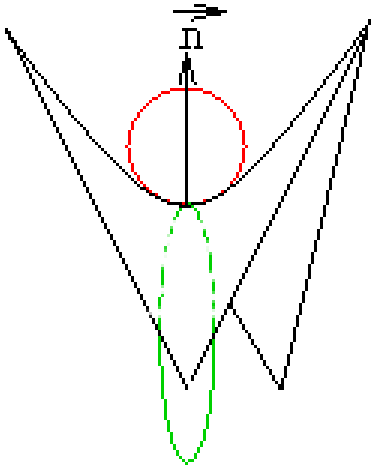


Figure 5.21: Curvature. We have $K(s) = \frac{1}{k_1 k_2}$. k_1 and k_2 are the principal curvatures such as in all the possible tangent vectors, k_1 and k_2 are the maximum and minimum values of the normal curvature at point s. In this image k_1 is the curvature of the green circle and k_2 of the red circle. In this case, k_1 and k_2 have opposite signs. Therefore, the Gaussian curvature is negative at this point.

Limitation

As N_3 is an extensive parameter, the increase of the field of view by a factor λ yields the multiplication of N_3 by a factor λ^3. In order to get a measure independent of the field of view, it is usual to normalise N_3 by the field of view, V. However, this measure cannot allow the comparison between different material because it is not stable by a homothetic transformation (see figure 5.25). In the subsection 5.3.3, we will introduce a new parameter satisfying both properties: intensive and stable by a homothetic transformation.

5.3.2 Percolation

Methods

Coffee A coffee percolator has the property that the hot water can go through[62] the coffee grains extracting the coffee grounds since the void space percolates. There is a path between two opposite face of the coffee percolator, a long-range connectivity. In the science of mathematics/physics, an interesting question is the emergence of a long-range connectivity, the percolation threshold.

Percolation theory A long range connectivity is the existence of a connected component with an infinite size. For random lattice models, such as 3d-simple-cubic lattice model with the occupation probability, p, the critical value, p_c for emergence of infinite connectivity is closed to 0.195, for the site percolation [148, 11]. This model makes the assumption that there is no correlation between two sites to be occupied[63]. In the case of real pore network, this assumption is not yet valid. Therefore, the porous microstructure cannot be understood with this model.

[62] The water goes through the coffee percolator because of:

- a difference of pressure between a chamber at pressure inferior than the atmospheric pressure (the water boils) and another chamber at atmospheric pressure;
- the gravity.

[63] For this model, the 2-point probability function, S_2, is equal to p if $r = 0$ and p^2 if $r \neq 0$

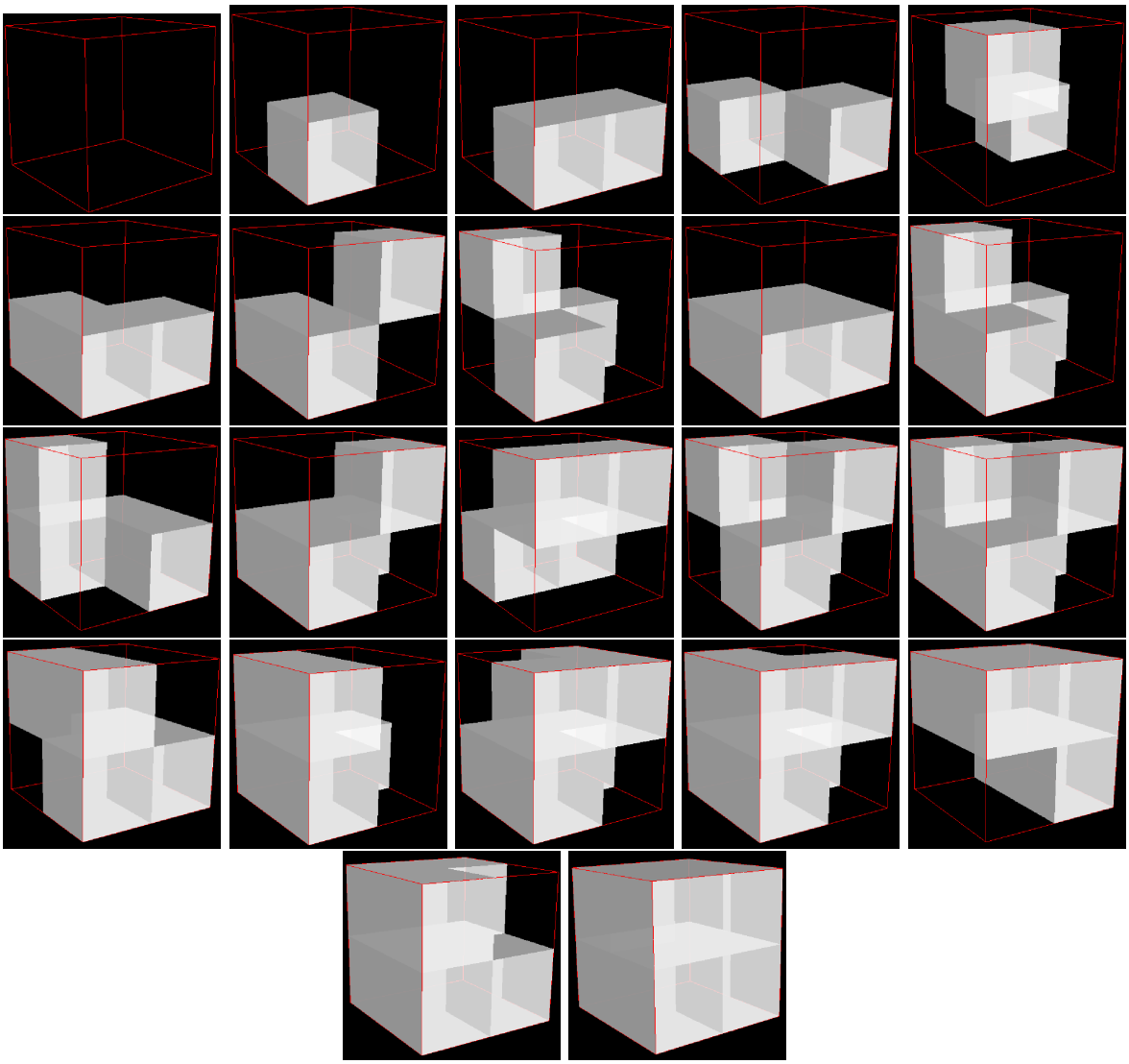


Figure 5.22: 22 patterns. **Property :** any configuration in a unit cell of size 2*2*2 is isotropic with one and only one of these patterns and has the same Gaussian curvature of the associated pattern.

Domain of the space In order to define the infinite connected component, the space has not to be bounded. As the image is a bounded domain, the notion of infinite connected component is replaced by the notion of connected component hitting each face of the 3d image. The equivalence between the infinite connectivity and the hitting of the 6-faces is "valid" when the field of view is larger than REV. For OPC, this property has been verified in the subsubsection 5.2.4. Numerically, first, the connected components are extracted using the procedure introduced in the subsubsection 3.4.1. Then, the connected component with the maximum volume, called max cluster, is kept. If the max cluster hits each face, then the pore network percolates, otherwise not.

C	N	Number	K
0	1	0	0
1	8	1, 2, 4, 8, 16, 32, 64, 128	$\frac{\pi}{2}$
2	12	3, 5, 10, 12, 17, 34, 48, 68, 80, 136, 160, 192	0
3	12	6, 9, 18, 20, 33, 40, 65, 72, 96, 130, 132, 144	-π
4	4	24, 36, 66, 129	-3π
5	24	7, 11, 13, 14, 19, 21, 35, 42, 49, 50, 69, 76, 81, 84, 112, 138, 140, 162, 168, 176, 196, 200, 208, 224	$-\frac{\pi}{2}$
6	24	25, 26, 28, 37, 38, 44, 52, 56, 67, 70, 74, 82, 88, 98, 100, 131, 133, 137, 145, 152, 161, 164, 193, 194	$-\frac{3\pi}{2}$
7	8	22, 41, 73, 97, 104, 134, 146, 148	$-\frac{\pi}{2}$
8	6	15, 51, 85, 170, 204, 240	0
9	8	23, 43, 77, 113, 142, 178, 212, 232	$-\pi$
10	24	27, 29, 39, 46, 53, 58, 71, 78, 92, 114, 116, 139, 141, 163, 172, 177, 184, 197, 202, 209, 216, 226, 228	$-\pi$
11	24	30, 45, 54, 57, 75, 86, 89, 99, 101, 106, 108, 120, 135, 147, 149, 154, 156, 166, 169, 180, 198, 201, 210, 225	0
12	6	60, 90, 102, 153, 165, 195	0
13	2	105, 150	2π
14	8	107, 109, 121, 151, 158, 182, 214, 233	$\frac{3\pi}{2}$
15	24	61, 62, 91, 94, 103, 110, 118, 122, 124, 155, 157, 167, 173, 181, 185, 188, 199, 203, 211, 217, 218, 227, 229, 230	$\frac{\pi}{2}$
16	24	31, 47, 55, 59, 79, 87, 93, 115, 117, 143, 171, 174, 179, 186, 205, 206, 213, 220, 234, 236, 241, 242, 244, 248	$-\frac{\pi}{2}$
17	4	126, 189, 219, 231	π
18	12	111, 123, 125, 159, 183, 190, 215, 222, 235, 237, 246, 249	π
19	12	63, 95, 119, 175, 187, 207, 221, 238, 243, 245, 250, 252	0
20	8	127, 191, 223, 239, 247, 251, 253, 254	$\frac{\pi}{2}$
21	1	255	0

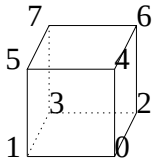


Figure 5.23: Look up table given the Gaussian curvature. The first column (C) contains the label of each pattern of the figure 5.22 (in the same order). Each number in the second column (N) is the number of configurations which are isotopic with the associated pattern. The third column is the configuration label of units cell such that the label is defined by: $n = \sum_{i=0}^{7} 2^i.\text{phase}_i$, phase $_i \in \{0, 1\}$ (see the upper figure). The fourth column is the gauss curvature of this pattern.

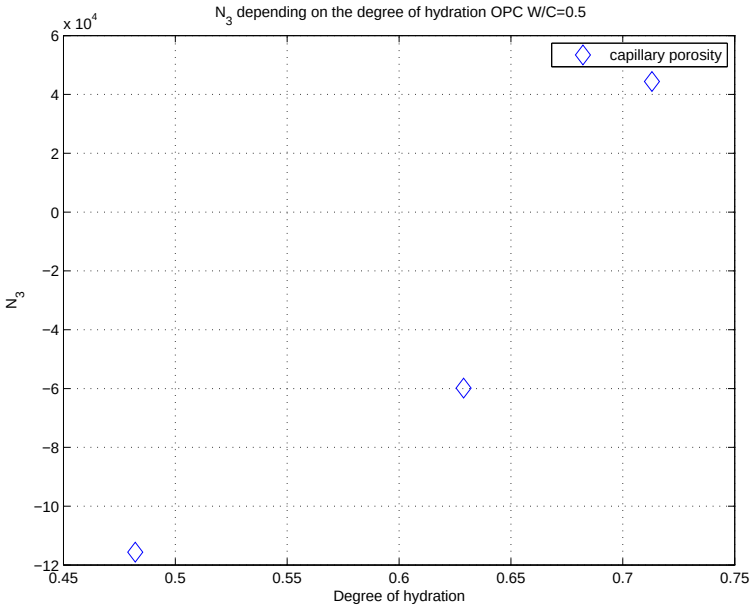


Figure 5.24: Result for OPC at W/C=0.5

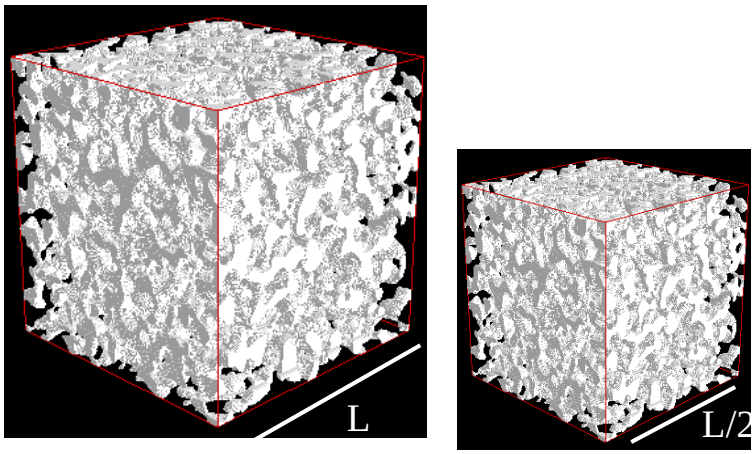


Figure 5.25: The same material except that the characteristic size has been divided by 2. We expect that the topological measure will remain constant for both microstructures. But, with N_3 normalised by the field of view, there will be a difference of factor 8.

Probability to belong to the max cluster of percolation In order to predict the disconnectedness, a second interesting indicator is the probability to belong to the max cluster of percolation, p_∞:

$$\begin{aligned} p_\infty \sim & \quad 1, & p >> p_c \\ p_\infty \sim & \quad (\phi - \phi_c)^\beta, & p > p_c \\ p_\infty \sim & \quad 0, & p < p_c \end{aligned}$$

where ϕ is the volume fraction of the given phase, ϕ_c is the volume fraction of the given phase at the percolation threshold, p_c.

Experimental results

In the figure 5.26, we observe the disconnectedness of capillary pores at time=150h. With X-ray tomography, we observe only the larger capillary pores. This disconnectedness involves that the capillary pore is weakly connected. At this stage, there is an intermittent transport between the pore network at the micro-scale and at a sub-micro-scale. The pores, developed in hydrate component at the sub-micro scale, start to be the main path of diffusion. We also observe that the pore network is still connected at time=83h with a porosity equals to 0.176. Therefore, the critical threshold percolation is above 0.176 and different to the threshold percolation of the random lattice model.

In the figure 5.27, a manual interpolation gives a threshold percolation equal to 0.14. Since there is

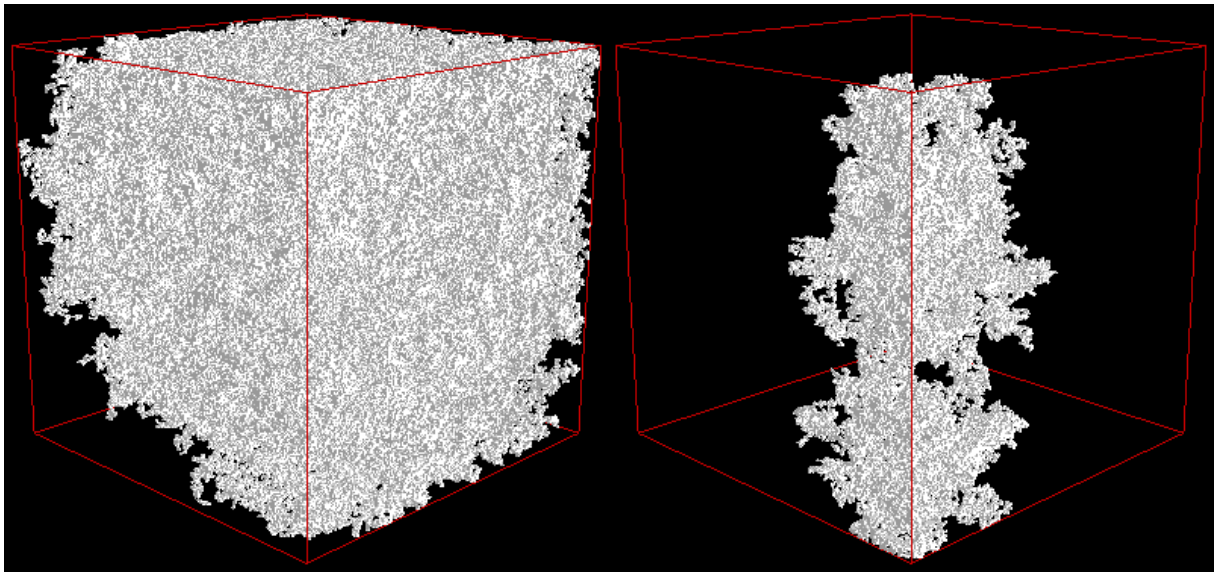


Figure 5.26: Left image: max cluster at time=83h. Right image: max cluster at time=150h. At left, the max cluster hits each face of the cube whereas at right not.

only three points for the interpolation, this value is just an indication and should be validate with more points.

5.3.3 Topological graph

For the statistical analysis of the topology, we will use an old procedure introduced by Euler in 1736 to solve the mathematical enigma of Seven Bridges of Konigsberg (see figure 5.28). This procedure is segmented in three parts [135, 14, 161, 158]:

1. from the microstructure to the thinning skeleton,
2. from the thinning skeleton to the topological graph,
3. analysis of the topological graph.

This approach involves the absence of internal surfaces. In the case of the pore network of OPC, the internal surfaces, the "cluster" solids surrounded by void, have been removed (see subsubsection 4.4.2).

From the microstructure to the thinning skeleton

The thinning skeleton is obtained by progressively narrowing the microstructure starting from the surface at constant topology. This shrinking must satisfy both topological invariance of the three first Betti

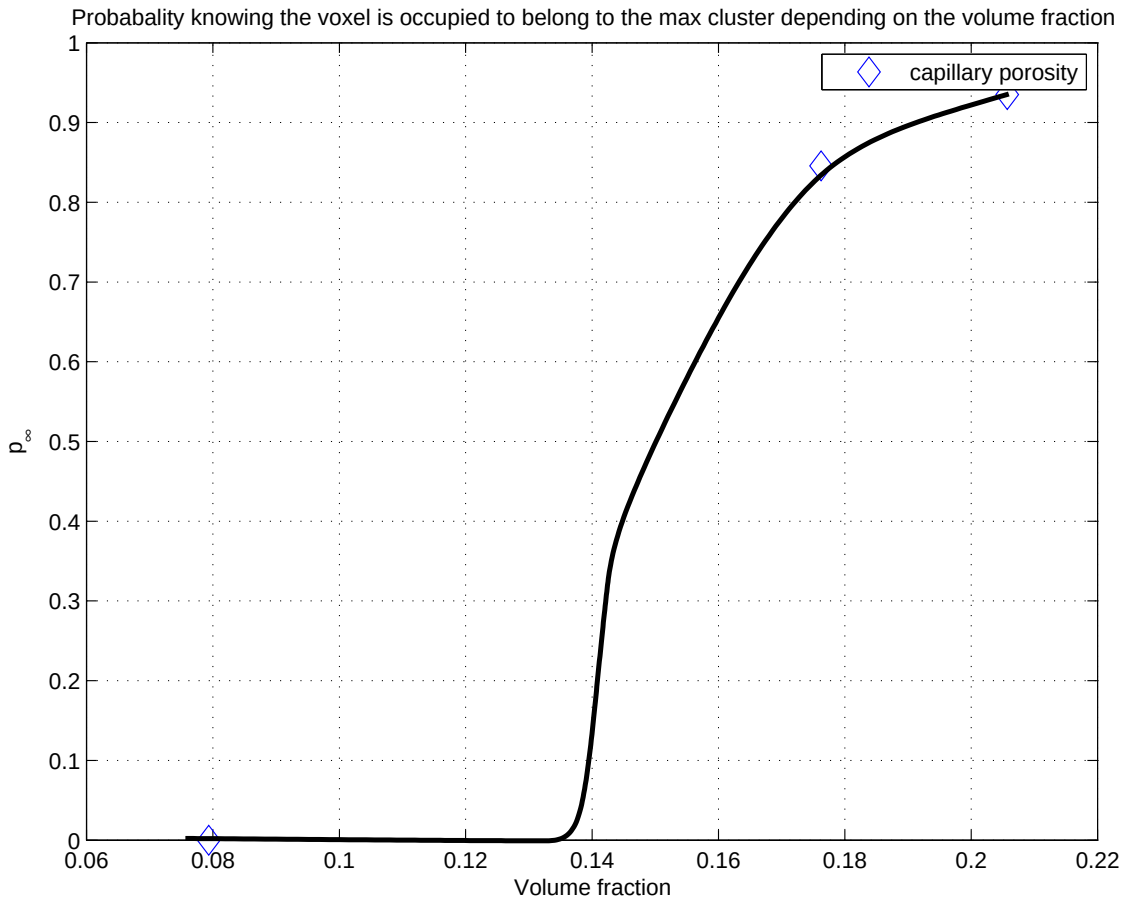


Figure 5.27: Probability to belong to the max cluster of percolation. The black plot is a manual fit.

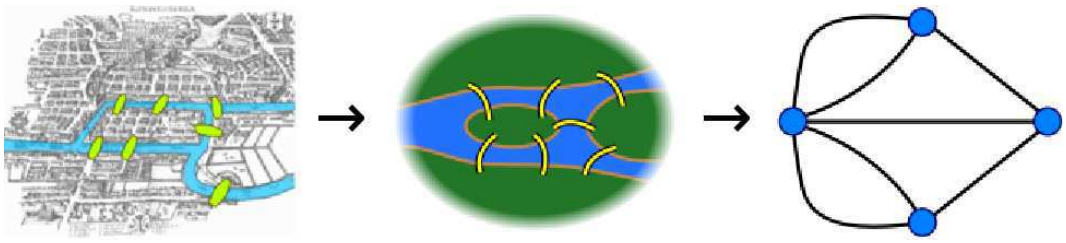


Figure 5.28: The problem was to find a walk through the city that would cross each bridge once and only once. Since only the connection information is relevant, the shape of a pictorial representation of a graph may be distorted in any way without changing the graph itself. Only the existence (or lack) of an edge between each pair of nodes is significant. For example, it does not matter whether the edges drawn are straight or curved, or whether one node is to the left or right of another.

numbers and geometrical constraints in order to be located in the centre of the microstructure. This process involves:

- a boolean function to answer this question, "Let X a discrete microstructure and $X_x = X - \{x\}$ the microstructure X without the voxel x. Do X_x and X have the same topology?"

- a procedure to contract isometrically the original microstructure to the "thinning" skeleton.

boolean function Let X a microstructure and x a voxel belonging to X. Let X_x, the set X without the voxel x, $X_x = X \setminus \{x\}$. To answer the question: "Do X_x and X have the same topology?", Barret & Yust [14] use the local conservation of the Betti numbers.
The Betti numbers characterize the void/solid microstructure topologically:

1. β_0, the number of connected components of the microstructure,
2. β_1, the number of irreducible cycles,
3. β_2, the number of internal surfaces or the number of connected components of the complementary of the microstructure.

The local conservation converts the global conservation of the Betti numbers for the whole microstructure to the local conservation of the Betti numbers for a subset of X of size 3*3*3 located in the 26-neighbourhood of x, $N_{26}(x)$):

$$\forall i = 0, 1, 2 : \beta_i(X \cap N_{26}(x)) = \beta_i(X_x \cap N_{26}(x))$$

Because of the Jordan's theorem, the evaluation of the number of connected components, n, is done with 26-neighbourhood for the microstructure, $\beta_0(A) = n_{26}(A)$ and with the 6-neighbourhood for complementary of the microstructure, $\beta_0(A) = n_6(A)$. A refers to $X \cap N_{26}(x)$ or $X_x \cap N_{26}(x)$. The algorithm of Hoshen & Kopelman [72] or the algorithm developed in the subsubsection 3.4.1 allows the evaluation of the number of connected components.
The evaluation of β_1 is done using the relation between the Euler-Poincaré characteristic, N_3, and the Betti numbers (see subsection 5.3.1 for the evaluation of Euler-Poincaré characteristic):

$$N_3 = \beta_0 - \beta_1 + \beta_2$$

To speed up the computational time, we use a look-up table given 1 if $\forall i = 0, 1, 2 : \beta_i(X \cap N_{26}(x)) = \beta_i(X_x \cap N_{26}(x))$ and 0 otherwise. We have evaluated for all the configuration of a cube of size 3*3*3 if the configuration with and without the center voxel have the same Betti number.

Procedure of contraction The narrowing process must satisfy some geometrical constraints in order to fit as exactly as possible the "thinning" skeleton. The natural algorithm is a step by step direction process [135] (figure 5.29).
This "thinning" skeleton depends on the order of direction steps (see figure 5.30) A solution to overcome this problem is to use the connected ordered topological space approach introduced by E. Khalimsky [86, 19].

From the thinning skeleton to the topological graph

The contraction at constant topology gives a "thinning" skeleton. The next stage is the transformation of the "thinning" skeleton (a set of voxels) to a graph skeleton (a set of vertexes connected by edges) at constant topology. This transformation is segmented in three steps (see figure 5.31):

1. affectation of an edge/vertex label at each voxel belonging to the "thinning" skeleton,
2. extraction of the connected components for the edge labelled image and the vertex labelled image,
3. linking the edge and vertex connected components.

From "thinning" skeleton to vertex/edge images A voxel gets the road label if there is only two voxels belonging to the thining skeleton in its neighbourhood. A voxel gets the crossroad label if there is more than two voxels belonging to the thining skeleton in its neighbourhood. We denoted by $\mathcal{V}$, the vertex labelled image and by $\mathcal{E}$ the edge labelled image (see figure 5.32).

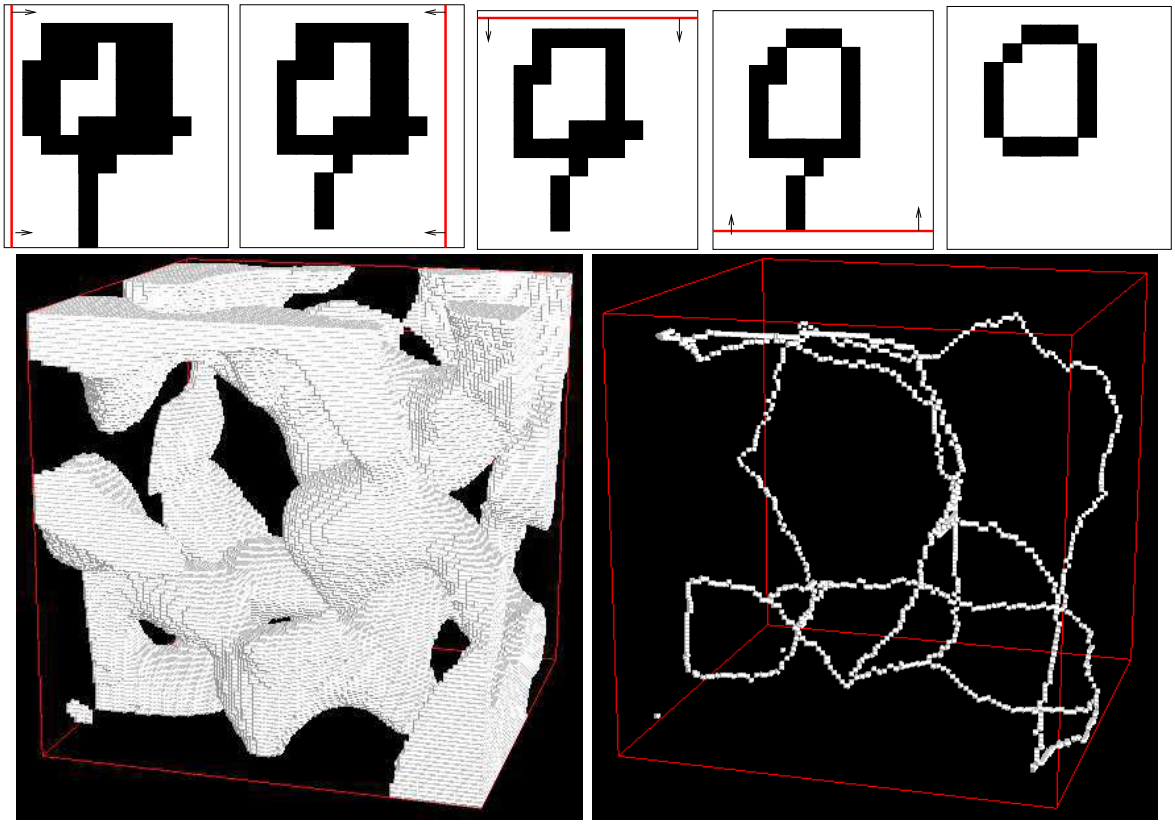


Figure 5.29: Each iteration of the thinning algorithm is decomposed in six steps. Each step corresponds to one of the six main directions (d): up, down, north, south, east, and west. A voxel x of the component X is called a d-border voxel if its six-connected neighbours in the d-direction belongs to the complementary component, X^c. At each iteration, only d-border voxels are possibly removed. A video is available at http://pmc.polytechnique.fr/∼vta/squelette.mpeg. Two last images are before and after the retraction.

From the vertex/edge labelled images to the vertexes/edges connected components For both labelled images, $\mathcal{V}$ and $\mathcal{E}$, we extract the connected components to get two sets, $\mathbb{V} = (v_0, \ldots, v_n)$ and $\mathbb{E} = (e_0, \ldots, e_p)$ (see figure 5.33).

From the edge/vertex connected components to the graph Let $\Gamma = (E, A, \gamma)$ be the topological non-oriented graph of the initial microstructure, where E is a set of vertexes, and A is a set of edges and γ is mapping from A into (E, E) (for example, $\gamma(a) = \{x_1, x_2\}$ is the link between the vertexes x_1 and x_2 done by the edge a). The aim of this part is to construct this graph such that the topology between the graph Γ and the microstructure is the same. To achieve this goal, we will link the vertex connected components and the edge connected components in two steps:

1. the tore graph,

2. the connected graph.

1) The tore graph Γ_t
An edge connected component, e, is defined as tore if there is no vertex voxels on its 26-neighbourhood:

$$(e \oplus N_{26}) \cap (\cup_{0<i\leq n} v_i) = \emptyset$$

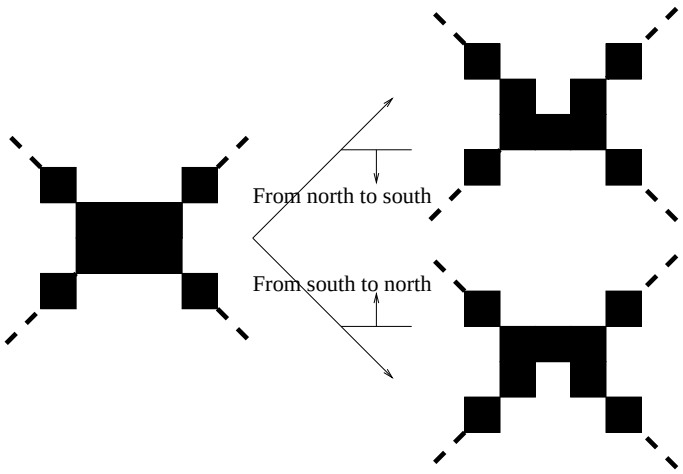


Figure 5.30: Two different "thinning" skeleton depending on the direction step.

In figure 5.31, there is only one edge tore localised at the left. The tore graph has $2 * k$ vertexes and $2 * k$ edges where k is the number of edge tore. The γ function is defined by:

$$\gamma(e_i) = \{2 * (i\%2), 2 * (i\%2) + 1\}$$

The vertexes with the label 1 and 2 are linked by the edges with the label 1 and 2 and so on.
2) The connected graph Γ_c
After removing the tores, each vertex/edge connected component is associated a vertex/edge on the graph Γ_c with the same corresponding label:

$$E = \{v_0^g, \ldots, v_k^g\}$$
$$A = \{e_0^g, \ldots, e_m^g\}$$

For each edge connected component, e_i, we calculate its neighbourhood vertexes:

$$N(e_i) = \{\forall j \in (0, \ldots, k) : (e_i \oplus B_{1,N_\infty}) \cap v_j \neq \emptyset\}$$

$N(e_i)$ gives one or two integers. If two integers, j, k, the edge, e_i, links the vertexes, v_j, v_j. If one integer, j, the edge, e_i, does a loop at the vertex, v_j. Therefore, the gamma function is defined by:

$$\forall i \in (0, \ldots, m) : \gamma(e_i^g) = \begin{cases} N(e_i) \text{ if } card(N(e_i)) = 2 \\ (N(e_i), N(e_i)) \text{ otherwise} \end{cases}$$

The final graph is the union of the previous graphs (see figure 5.34).
3) local analysis
This final graph is not the topological graph since each connected component vertex can have a complex topology. For each vertex connected component, we calculate the Euler-Poincaré invariant, N_3^c. The delta Euler-Poincaré invariant, ΔN_3 is the subtraction of N_3^c by 1 (1 is the value of Euler-Poincaré invariant for a single voxel).

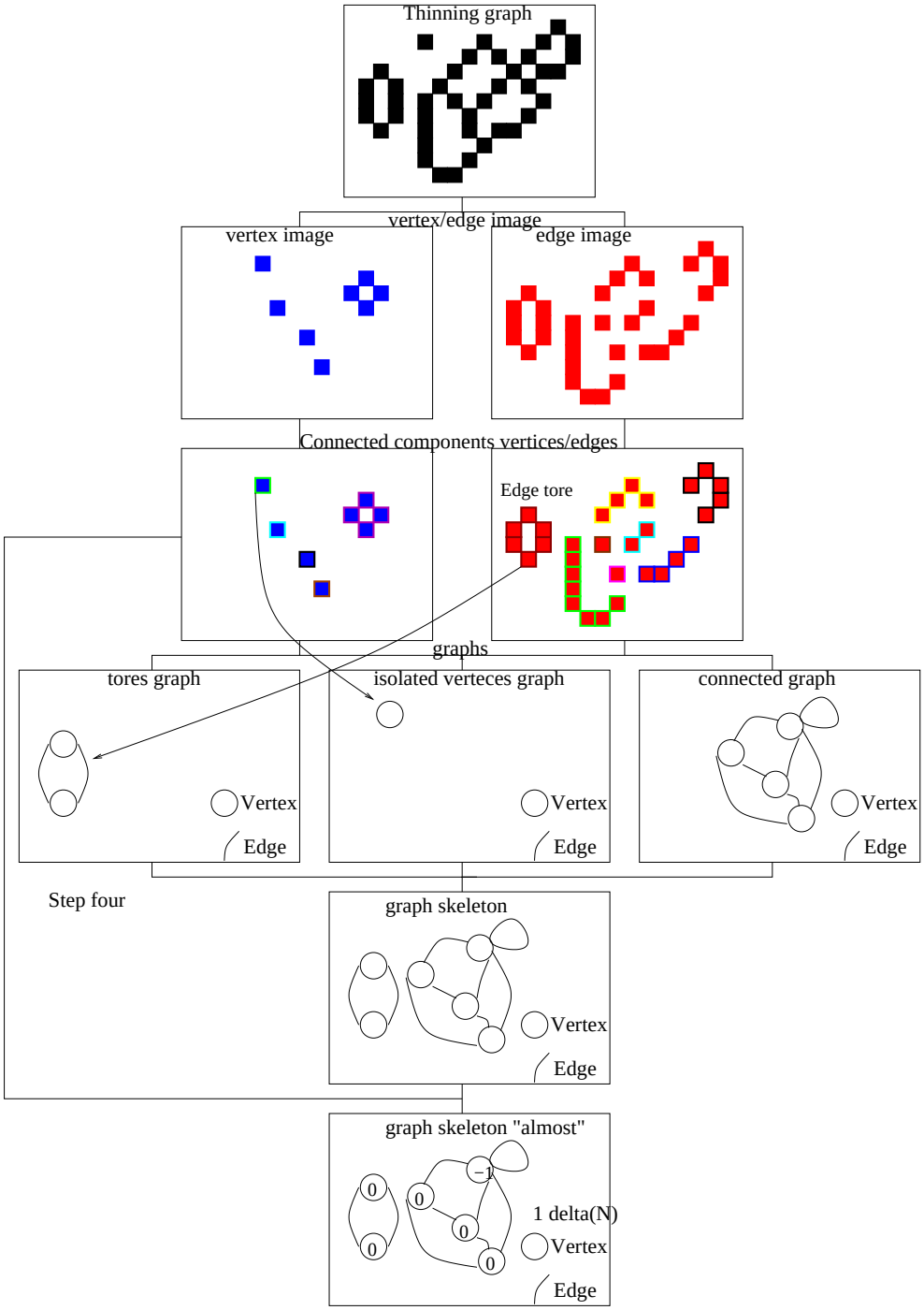


Figure 5.31: From the thinning skeleton to the skeleton graph. First step: labelling each voxel of the thinning graph; second step: extraction of the vertexes and edges connected components; third step: creation of the three graphs; last step: union of these previous graphs.

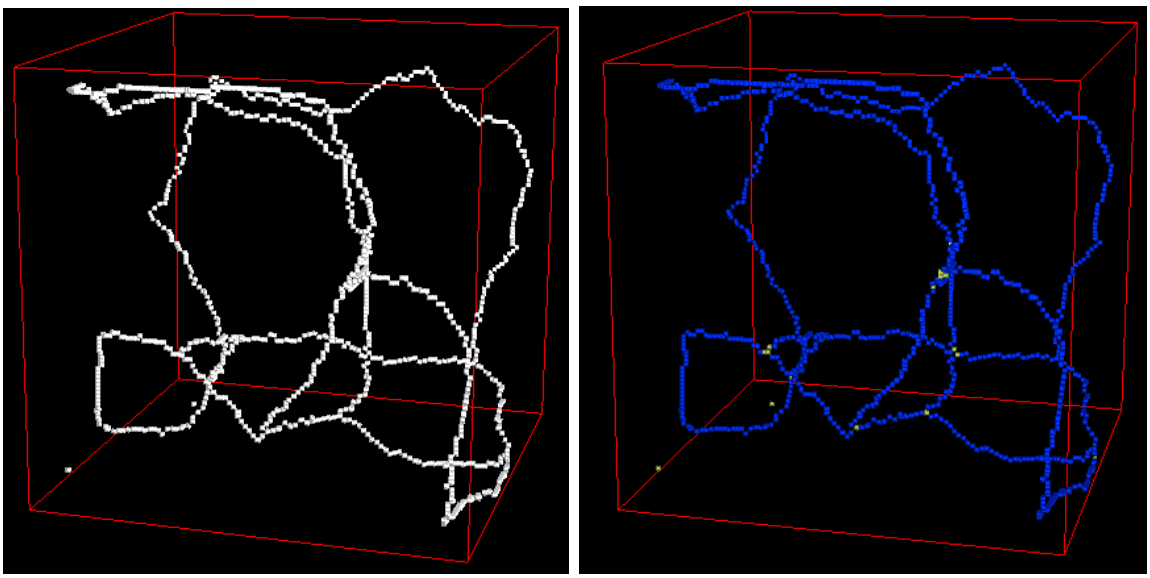


Figure 5.32: From the thinning skeleton to the edge/vertex label image. The blue voxels are the roads and the yellow voxels are the crossroads.

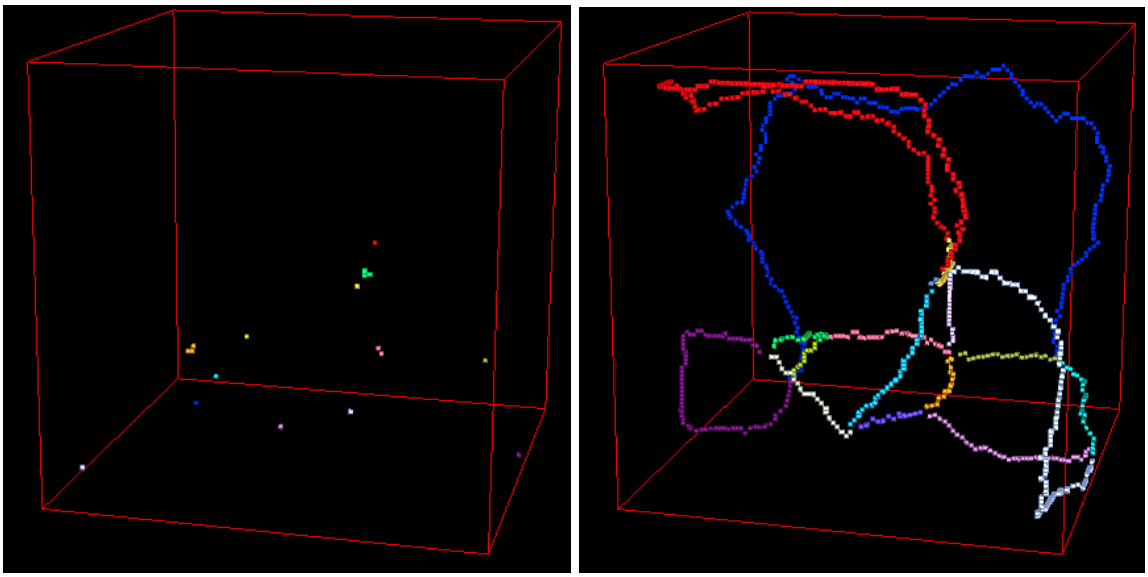


Figure 5.33: From the vertex/edge images to the vertexes/edges connected components. The first image is the connected components for the vertex labelled image and the second for the edge labelled image.

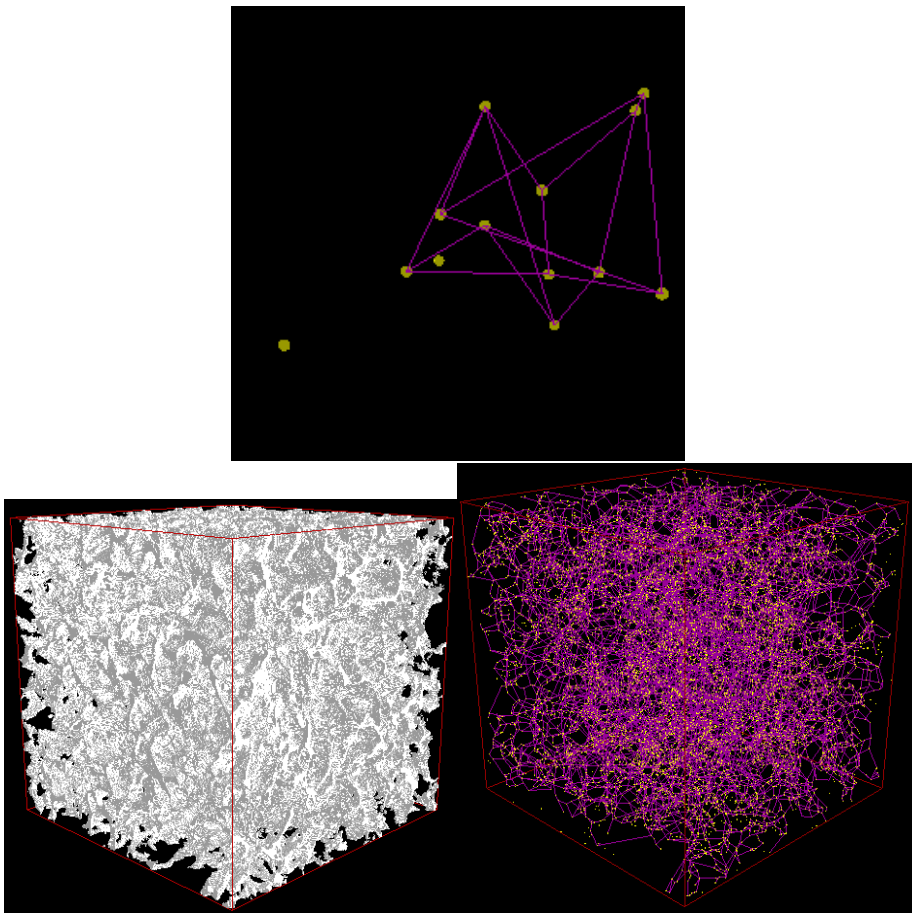


Figure 5.34: Top figure: graph of the initial microstructure. Bottom figure: same process with a material with a representative size.

5.3.4 Analysis of the topological graph

As seen in the previous subsection, the extraction of the topological graph is difficult. But, this graph allows a statistical analysis of the topology. In this subsection, we will present the probability function of coordination number. Then, we will introduce a new topological number having some good properties.

Probability function of coordination number

Methods Let us consider a car driver starting from a town and wanting to go to another town at a given direction, $\vec{d}$. Starting from this town, there are a number of possible roads. The car driver will take the road such that the road direction, $\vec{d}_m$, is the closest of the objective direction $\vec{d}$. If the number of road is high, the probability that the angle between both directions is low is high. In other case, it is the opposite and the car driver will have to come by a roundabout way. Therefore, depending on the connectivity of road network, the existence of a direct path is more or less probable. Thus, the connectivity plays an important role for the understanding of the transport properties.
In graph theory, a town is a vertex and the number of possible roads starting from this town is the coordination number of the vertex. The statistical analysis of the graph allows the determination of the the probability function of coordination number, $P(N_c(V) = n)$ given the probability for a random vertex, V, with a coordination number equal to n. In the case of internet network and in more general case of small words, the probability function of coordination number has a power law form (see figure 5.35). The implication of the power law form is that the internet network is breakable easyly. If the link between two internet backbones highly connected is put down, the internet network can be segmented in two distinct parts. You can imagine the deep consequence

Experimental results The figure 5.36 shows the distribution of the coordination number of the topological graph for OPC. Whatever the hydration time, the shape of the distribution is a negative exponential with the same slope and with a shift between the distribution as the probability of the number of coordination equal to 0 increases with the time of hydration. Therefore, the disconnectedness of the pore network will be the consequence of the proliferation of small isolated components with a $N_3 = 1$ isolating the big clusters. Moreover, the connectivity in the big clusters remains constant since the slope is constant. In the next future, with a model material, a spinodal decomposition generating by a correlated Gaussian random field technique, we will investigate this phenomenon in a wide range of porosity.

Connectivity number

Methods To get a relevant number from this probability function, we calculate the mean of the previous probability function, $< N_c >$. This number is related to the connectivity of the pore network: $< N_c >$ is large, the network is highly connected and the opposite. However, the Euler-Poincaré has an important property that this number does not have. For a statistical network coming from real material, the change of sign, positive to negative, of the the Euler-Poincaré is related to the percolation of the network. To get a connectivity number having this property, Levitz has defined this number:

$$C = -\frac{< N_c > -2}{2}$$

This number is equal to the Euler-Poincare divided by the number of vertex:

$$-\frac{< N_c > -2}{2} = \frac{N_3}{\alpha_0}$$

Proof The topological numbers of the graph and the topological numbers of the voluminous microstructure are related by these relations:

$$\begin{aligned} N_3 &= \beta_0 - \beta_1 \\ N_3 &= \alpha_0 - \alpha_1 \end{aligned}$$

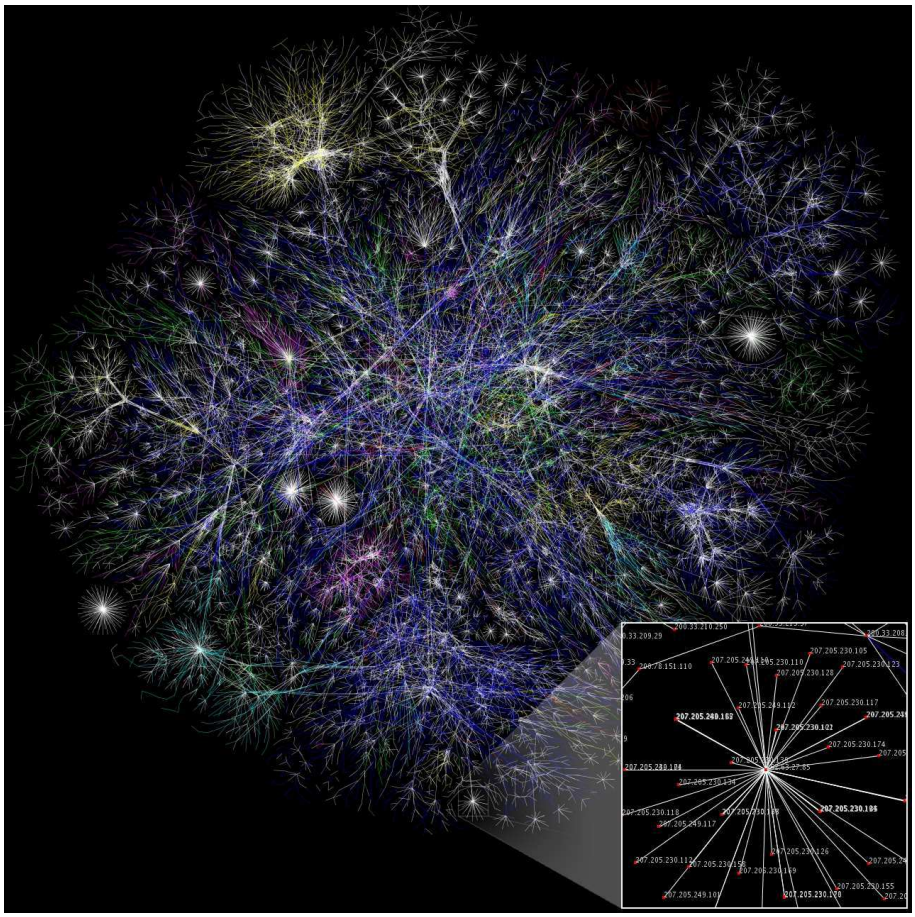


Figure 5.35: Each line is drawn between two nodes, representing two IP addresses. This is a small look at the backbone of the Internet. From wikipedia.

where α_0 is the number of vertexes and α_1 is the number of edges.
The relation between the sum of the coordination numbers and α_1 is:

$$\alpha_0 = \sum_{v \in V} 1 \tag{5.34}$$

$$\alpha_1 = \frac{1}{2} \sum_{v \in V} N_c(v) \tag{5.35}$$

where V is the set of vertexes of the topological graph. Since each edge is shared by two vertexes, we have a pre-factor of $1/2$.
The mean of the distribution of the coordination number, $< N_c >$, is:

$$< N_c >= \sum_{i \in \mathbb{N}} i P(N_c(v) = i)$$

The probability, $P(N_c(v) = i)$, is the ratio of the number of vertexes with a coordination number equal to i over the total number of vertexes:

$$P(N_c(v) = i) = \frac{\sum_{v \in V} \mathbb{1}_{N_c(v)=i}}{\sum_{v \in V} 1}$$

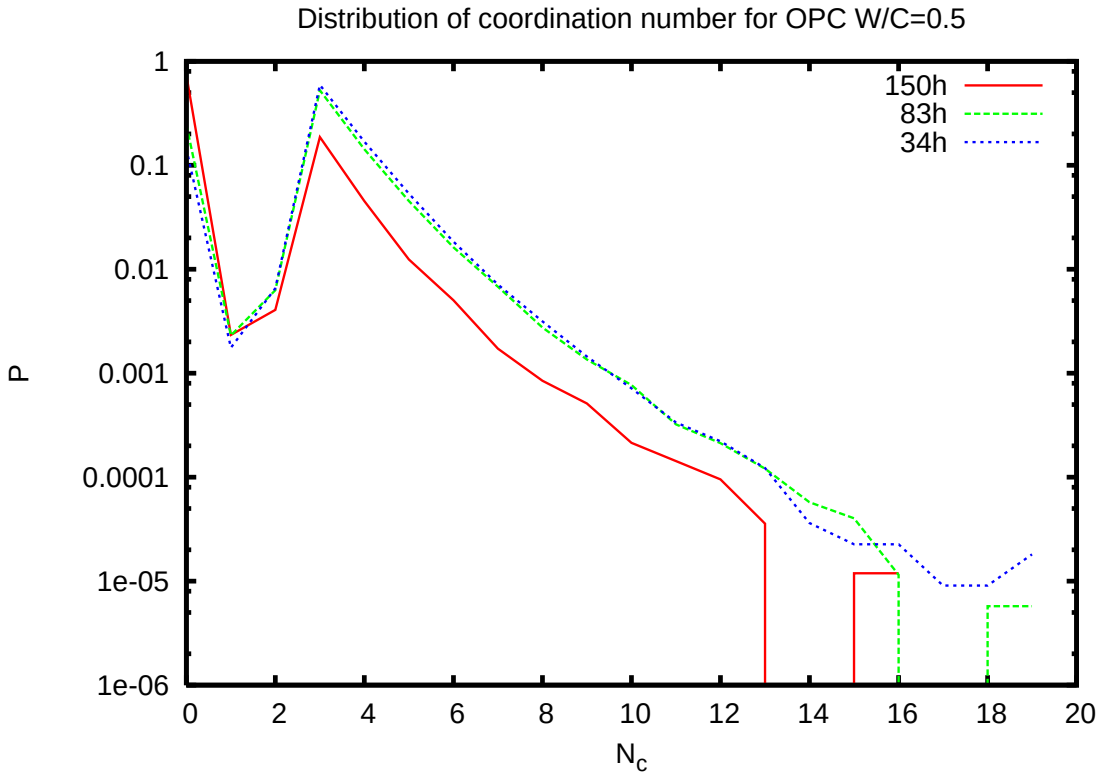


Figure 5.36: The graph shows a linear decay starting from the value equal to 3. Its form is an negative exponential because of the log-scale in coordinate.

Combining the two previous relations, we get:

$$\begin{aligned} < N_c > &= \frac{\sum_{i\in(0,\ldots,n)} i \sum_{v\in V} \mathbb{1}_{N_c(v)=i}}{\sum_{v\in V} 1} \\ < N_c > &= \frac{\sum_{v\in V} \sum_{i\in(0,\ldots,n)} i \mathbb{1}_{N_c(v)=i}}{\sum_{v\in V} 1} \end{aligned}$$

As $\sum_{i\in(0,\ldots,n)} i\mathbb{1}_{N_c(v)=i} = N_c(v)$, we obtain

$$< N_c > = \frac{\sum_{v\in V} N_c(v)}{\sum_{v\in V} 1}$$

Using the relation 5.34, we obtain:

$$\begin{aligned} < N_c > &= \frac{2\alpha_1}{\alpha_0} \\ < N_c > -2 &= 2\frac{\alpha_1 - \alpha_0}{\alpha_0} \\ \frac{< N_c > -2}{2} &= -\frac{N_3}{\alpha_0} \end{aligned}$$

Numerically, to obtain the correspondence, the formula is:

$$\frac{< N_c > -2}{2} = -\frac{N_3 + \Delta N}{\alpha_{0,link} + \alpha_{0,isolate} + 2\alpha_{0,tore}}$$

Properties of this topological number This number C has four important properties:

1. the change of sign is related to the percolation,
2. intensive, its measure does not depends of the field of view[64],
3. invariance by homothetic transformation, its measure does not depend of the characteristic scale of the material,
4. small range, its measure is limited between $]-\infty, 1]$. In practical utilisation, the minimal value is not inferior to -2.

Due to these properties, this number allows a topological comparison between different materials, solid network of bone, pore network of vycor, etc. Moreover, for diffusion transport, when this number is low, $C \leq 1$, the transport seems to drive by the morphological feature, and when the this number is closed to 0, $C \leq 0 - \epsilon$, some critical exponents between this number and the coefficient of diffusion appear. This is a work in progress.

Experimental results The results are presented in the table 5.3. In comparison of Euler-Poincaré number, the values of the connectivity number are limited between $[-1, 1]$.

time	N_3	$< N_c >$	C	percolation
34h	-115643	2.95	0.475	yes
83h	-59851	2.59	0.295	yes
150h	44388	1.78	-0.11	no

Table 5.3: Topological numbers for OPC with W/C=0.5

[64] when the field of view is larger than the REV

5.4 Decomposition in elementary pores

In order to understand various physical phenomenon as mercury intrusion, diffusion, permeability, a possible way is to decompose the microstructure in term of elementary pores. At the scale of the elementary pores, the physical phenomenon can be linked with morphological features. This first level allows to dress the network of the elementary pores with some local properties of the physical phenomenon. At the scale of the microstructure, a coarse graining description can be reached using the connectivity properties of the dressed network of the elementary pores. The decomposition in term of elementary pores depends on the physical phenomenon under investigation. For instance, permeability or mercury intrusion are driven by the presence of throats. Therefore, we will propose a a morphological decomposition such that the microstructure is cut on the throats. On the other hand, diffusion strongly depends on the connectivity of the pore network. Thus, we will propose a topological decomposition such that the network of elementary pores is the same as the topological network previously described. For the morphological decomposition, we will apply the procedure introduced by L. Vincent in 1993 [172]. For the topological decomposition, we apply the procedure introduced by L. Pothuaud *et al* in 2000 [135]. Both approaches have two stages:

1. localisation of a seed inside each elementary pore,
2. seeded region growing to associate a region at each elementary pore.

The outline of this section is as follows: first, the materials are presented, then the morphological/topological decomposition are explained, and two kinds of pore network are defined. At the end, the comparison between the two decompositions is done according to some geometrical features.

5.4.1 Materials

In order to understand the difference between the two decompositions, we have selected three disorder pore network of these materials (see figure 5.37):

- a vycor pore glass,
- a limestone,
- a sand specimen.

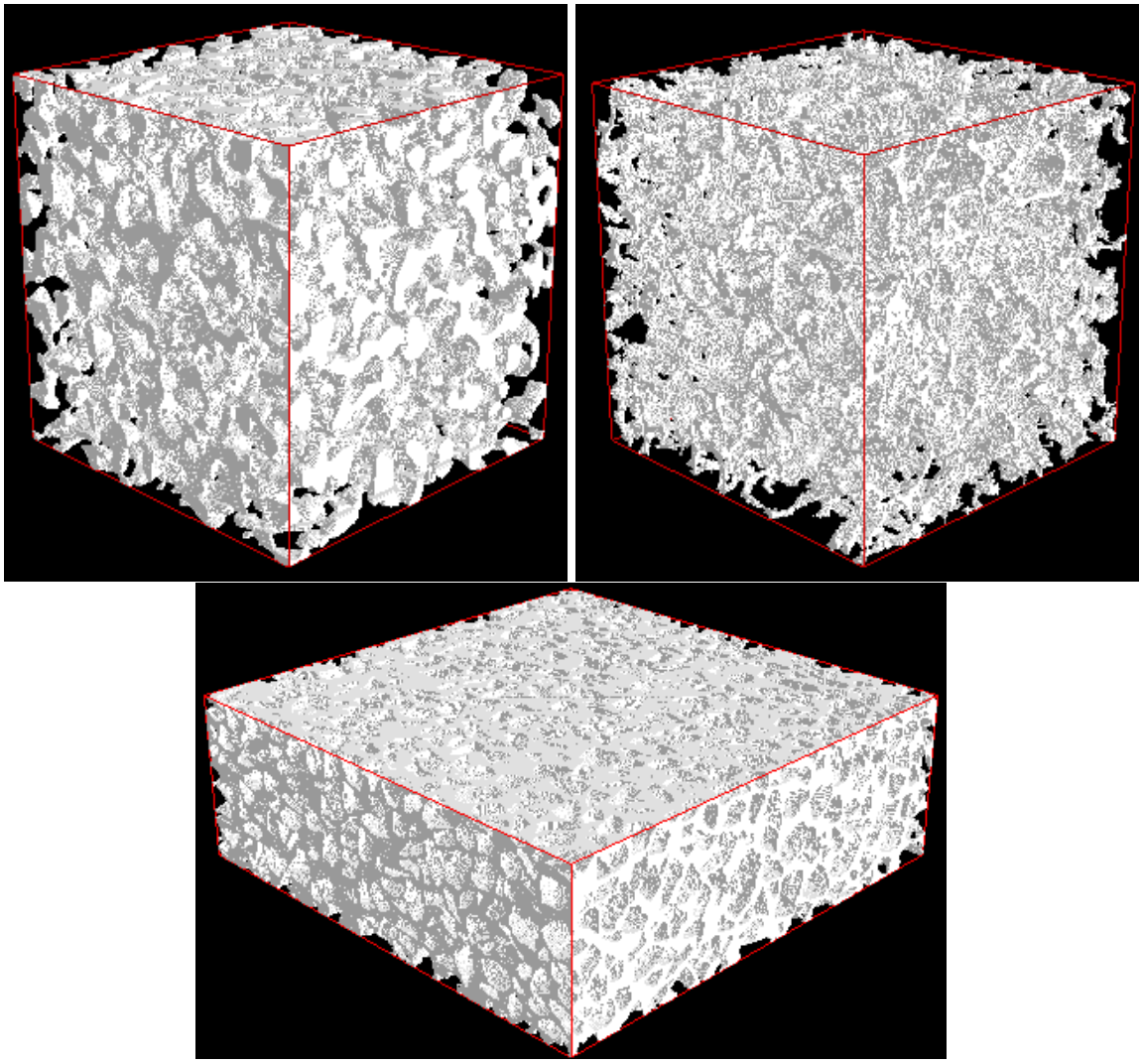


Figure 5.37: Top left: the material A is the vycor pore glass [105]; top figure: the material A is the porosity of a limestone, size=500*500*500, resolution=3 microns; bottom: the material C is the porosity of a sand specimen, size:450x450x150, resolution=3 microns.

5.4.2 Two conventions for the decomposition

The definition of the elementary pores is not obvious. For instance, some people says: "the microstructure is the elementary pore". In our approach, the decomposition of elementary pores is a tool to understand the link between the geometrical features of the disorder pore network and the properties of a physical phenomenon. To be an efficient tool, the decomposition has to respect some constraints. The utilisation of this tool is segmented in two levels. In the first level, a statistical relation has to be found between the geometrical features of the elementary pore and the properties of the physical phenomenon in the elementary pores. Therefore, the microstructure, as a single elementary pore, is not suitable since, in the general case of disorder pore network, only some bounds of the physical properties can be found using the geometrical features. In the second level, an up-scale from the physical phenomenon in the elementary pore to the physical phenomenon in the pore network (microstructure), has to found using the connectivity properties of the graph of the elementary pores. Thus, the relevant geometrical properties must be captured on the graph[65].

Morphological decomposition

As already known, throats slow down the convective transport. In the subsection 4.4.1, we have given a procedure to split the connected grains on the narrow surfaces. We use this method for the purpose of the decomposition of elementary pores (see figure 5.38). The advantage of this decomposition is that each surface between two adjacent pores is a throat.

Topological decomposition

The paper of Pothuaud *et al* [135] presents a method to decompose a bone micro architecture in term of tubercular. We use this procedure to decompose the porous media in term of elementary pores. The advantage of this decomposition is that the topological features are kept during the decomposition.

Seeds like vertex/edge connected components of the thinning skeleton of the microstructure In the subsubsection 5.3.3, we gave a procedure to extract the vertex/edge connected components. For the topological decomposition, we have the choice between two conventions:

1. each vertex connected component is a seed (crossroad pores),
2. each edge connected component is a seed (road pores).

Geodesic influence zone The topological pores are the geodesic influence zones (see subsubsection 3.4.1 for the algorithm) of the previous seeds . The topology is preserved by this decomposition but the surface between two adjacent pores is not located on the throats[66] (see figure 5.39).

[65] Each voxel belonging to the microstructure is an elementary pore. With this decomposition, the relevant properties, as the connectivity or the existence of narrow surfaces, are lost.

[66] To locate the surface between two pores on the throats, the idea is to use the watershed on the topographic surface equal to the opposite distance function of the microstructure. However, as the seeds, associated to the vertexes of the topological graph of the microstructure, are not located on the minima of topographic surface, the watershed transformation with these seeds leads to some bbasinswhich volume is equal to 1 voxel. The reason is that some seeds are located on the line of ridges of the topographic surface. The geodesic influence zone is used rather than the watershed transformation to avoid this problem.

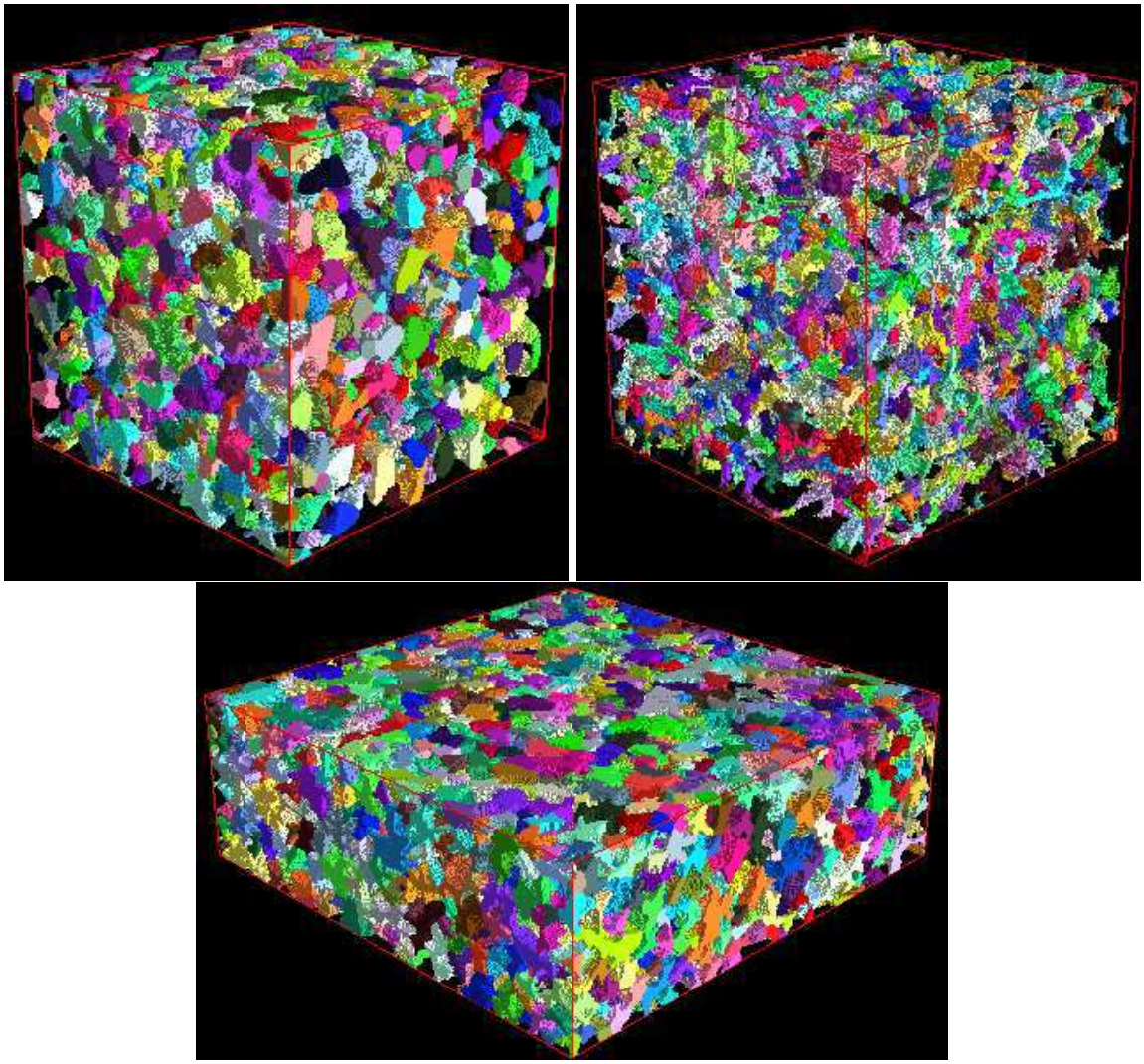


Figure 5.38: Application of the morphological extraction. Each colour represents an elementary pore. The dynamic parameter is always 1.

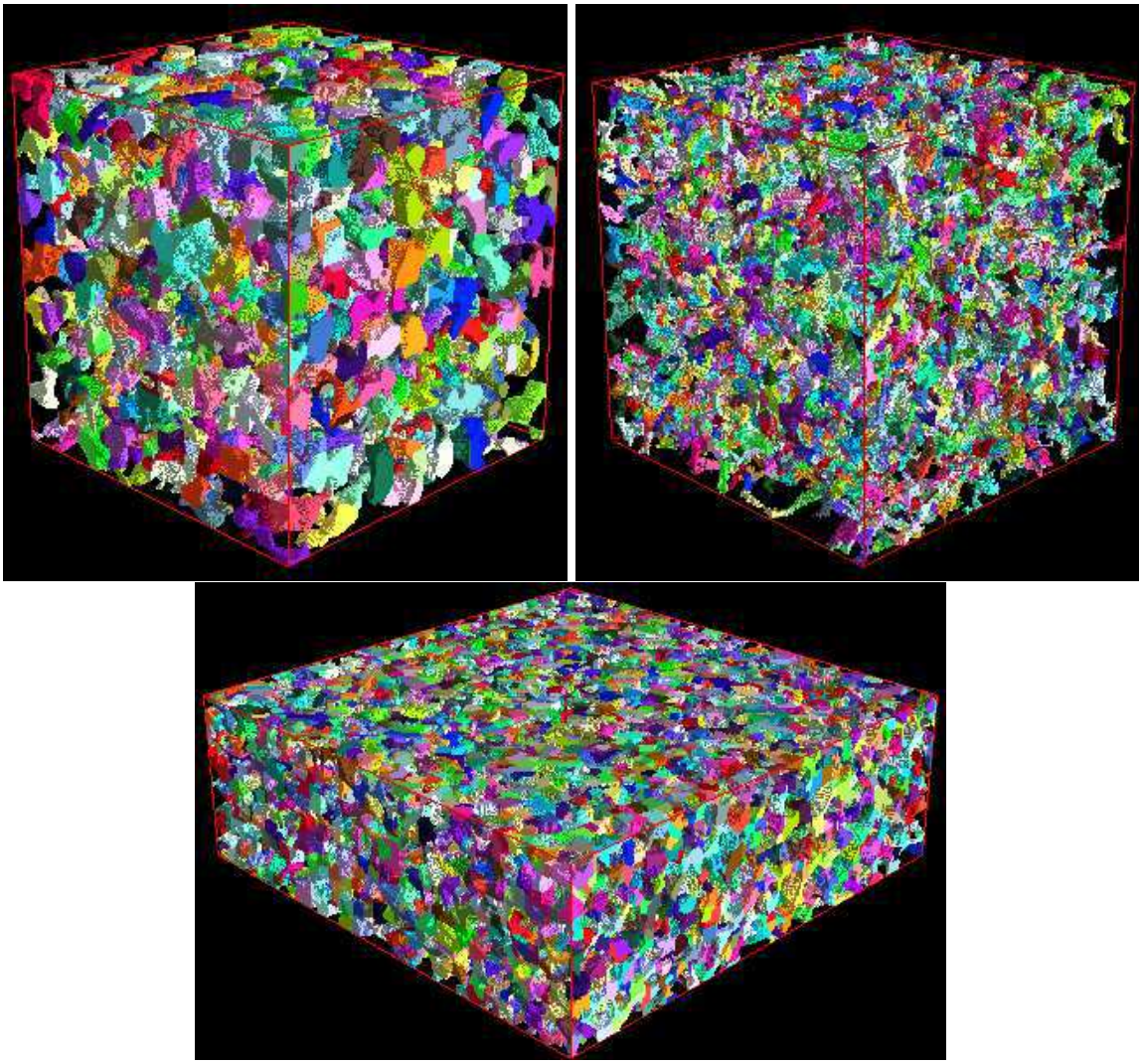


Figure 5.39: Application of the topological decomposition using the crossroad pores. Each colour represents an elementary pore.

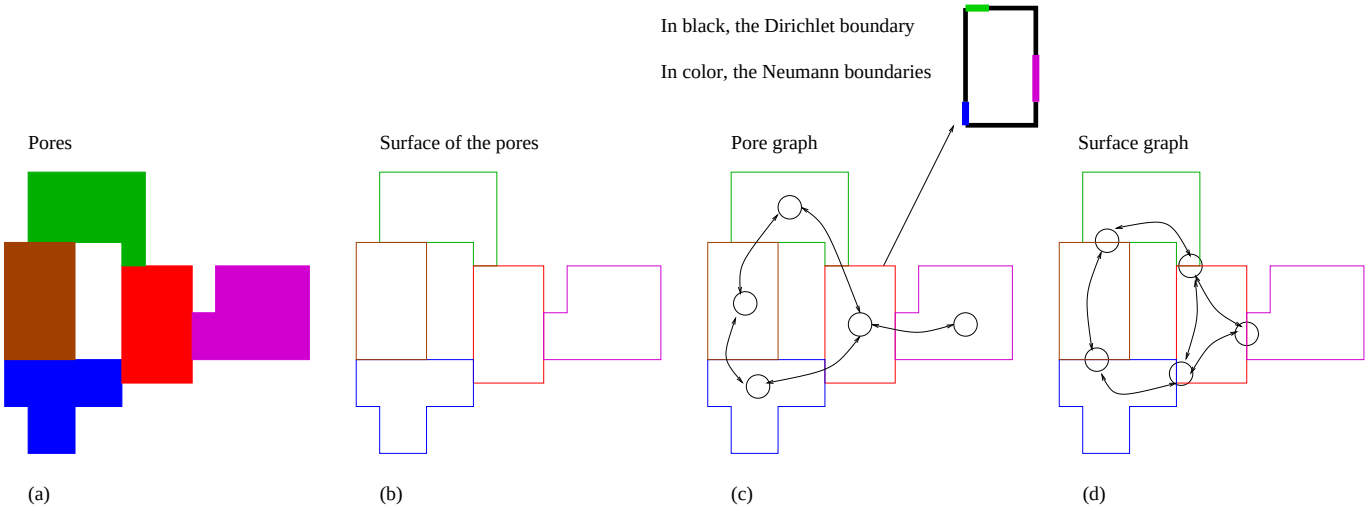


Figure 5.40: (a): the pores; (b): the boundaries of the pores; (c): the pore graph; (d): the surface graph.

Graph associated to the decomposition

The elementary pores, $(P_i)_{i\in(0,\dots,n)}$, are a partition of the microstructure Ω. In this subsubsection, we present two approaches to dress this partition with a graph:

1. by the volume of the pores,
2. by the adjacent surfaces of the pores.

Volume graph 1)A pore is associated to a vertex.
2)The neighbourhood of the pore, i is defined by the Γ_p function:

$$\Gamma_p(i) = \{\forall j \in (0,\dots,n) \setminus i : \partial P_i \cap \partial P_j \neq \emptyset\}$$

3)Let j in $\Gamma_p(i)$. The adjacent surface or Dirichlet surface is the surface between the pores, i and j, $S^D_{i,j}$, by:

$$S^D_{i,j} = \partial P_i \cap \partial P_j$$

where ∂P_i is the boundary of the pore, P_i.
4) The Neumann surface of the pore i is defined as the whole surface minus all the Dirichlet surfaces:

$$S^N_i = \partial P_i \setminus (\cup_{j\in\Gamma_p(i)} S^D_{i,j})$$

The figure 5.40 -c shows a pore graph.

Surface graph 1)A surface between two adjacent pores is associated to a vertex. A vertex is defined by a couple of integer, i, j such as the intersection between the surfaces of pores i and j is not empty.
2) The neighbourhood of the vertex i, j is defined by the Γ_s function::

$$\Gamma_s(i,j) = (\Gamma_p(j) \times \Gamma_p(i)) \setminus (i,j)$$

The figure 5.40-d shows a surface graph.

5.4.3 Characterisation

The aim of this subsection is the comparison between the topological and morphological decompositions using some geometrical features of the pores.

Volume histogram The first important characteristic of the pore is its volume. The figure 5.41 shows the volume distribution. Whatever the material, we observe a large dispersion of volume distribution for the morphological decomposition. Therefore, the two decomposition leads to a different partition. In order to use this decomposition as a tool to understand a physical phenomenon, the choice of the decomposition is crucial since the partition of pore space strongly depends on the convention: morphological or topological.

Volume depending on the Feret's diameter Let P be a pore in the orthonormal vector space E, $(e_1, \ldots, e_n)$. The Feret's vector of P is the size of the minimal box embedding the pore (see figure 5.42):

$$F(P) = \begin{pmatrix} \max_{\forall x \in P}(< x|e_1 >) - \min_{\forall x \in P}(< x|e_1 >) \\ \vdots \\ \max_{\forall x \in P}(< x|e_n >) - \min_{\forall x \in P}(< x|e_n >) \end{pmatrix}$$

There are two possible definitions of the Feret's diameter:

$$\begin{aligned} D_1(F(P)) &= (\prod_i < F(P)|e_i >)^{1/n} \\ D_2(F(P)) &= \frac{\sum_i < F(P)|e_i >}{n} \end{aligned}$$

The results are presented in the figure 5.43. Before plotting, we expected to obtain a proportional relation between the length associated to volume of the pore, $V^{1/3}$, and the Feret's diameter. However, the linear regression in log-log scale does not give a slope equal to 1 but a smaller value. The reason is that larger is the pore, larger is the space of the box not belonging to the pore. This slope depends on the material and the type of decomposition.

Proportion of Dirichlet surface Throats slow down the transport. The key idea behind the morphological decomposition is to cut precisely the microstructure on the throats. To validate this decomposition, we plot the distribution of the ratio between Dirichlet surface area (the surface area shared with other pores) and to the total surface area of the pore. The results are presented in the figure 5.44. Whatever the material, we observe that the mean of the distribution for morphological decomposition is lower than for topological decomposition. Therefore, the topological decomposition has a major drawback that one key feature is lost during the decomposition.

Coordination number The key idea behind the topological decomposition is to keep the topological properties. The figure 5.45 presents the results. Whatever the material, the graph for the morphological decomposition is similar to the graph for the topological decomposition. The morphological decomposition seems to catch the topological connectivity of the microstructure.

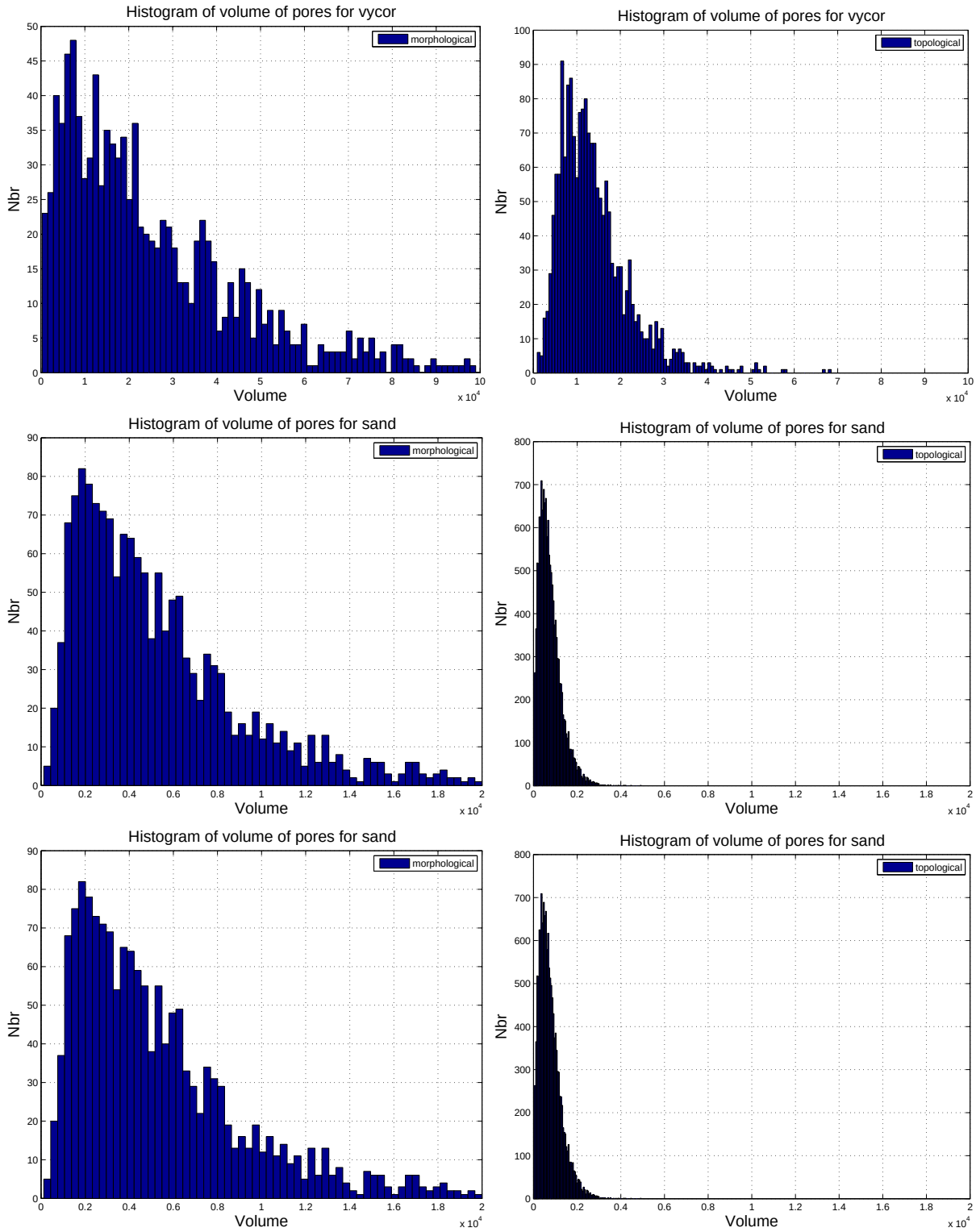


Figure 5.41: Comparison of volume histogram between the morphological and topological decomposition. Raw 1: vycor, raw 2: limestone, raw 3: sand.

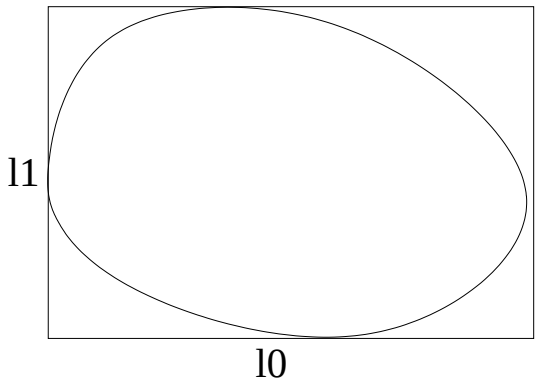


Figure 5.42: The Feret's vector is the vector of the minimal box embedding the pore.

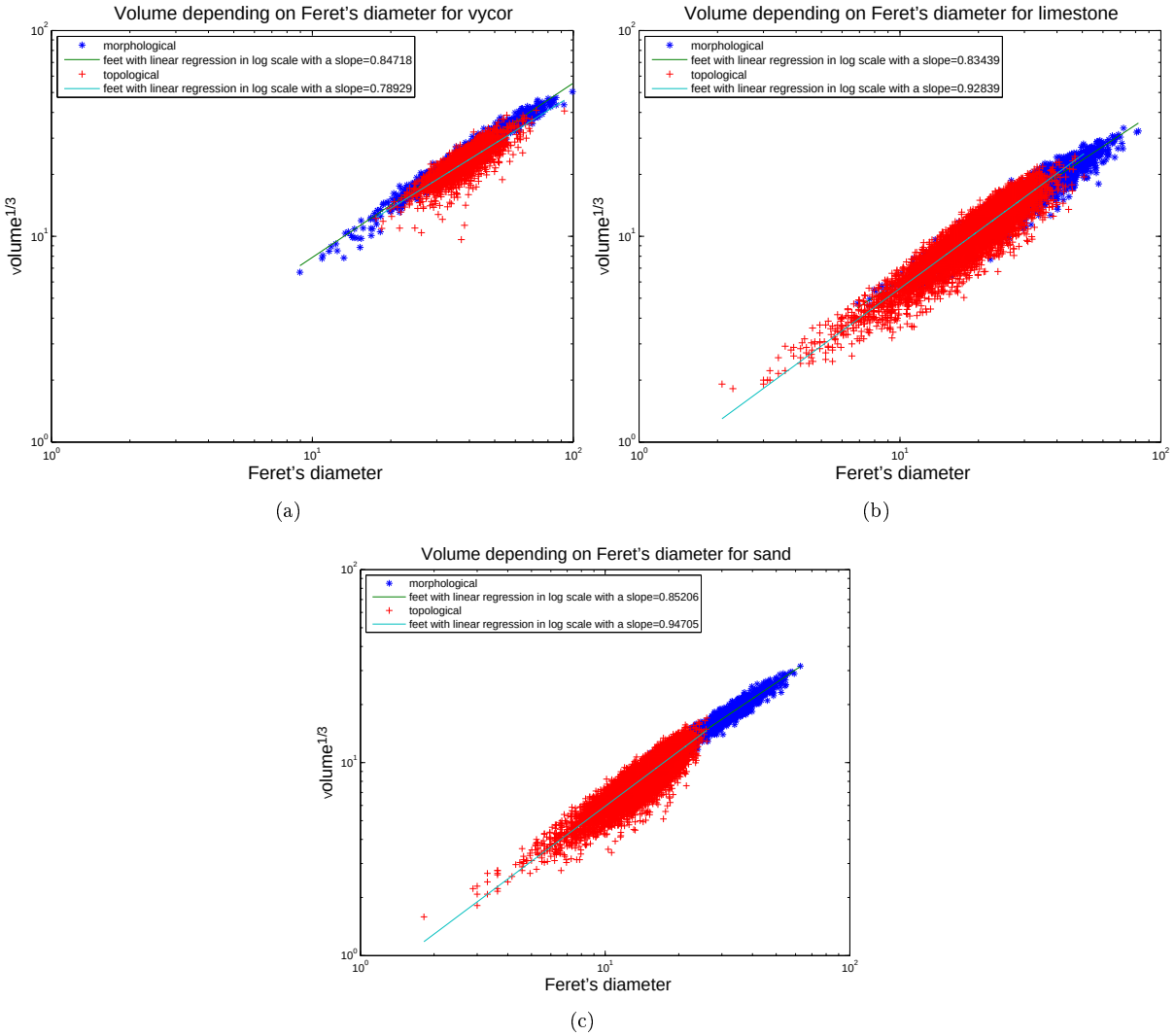


Figure 5.43: Volume versus Feret's diameter in log-log scale. (a): vycor, (b): limestone, (c): sand.

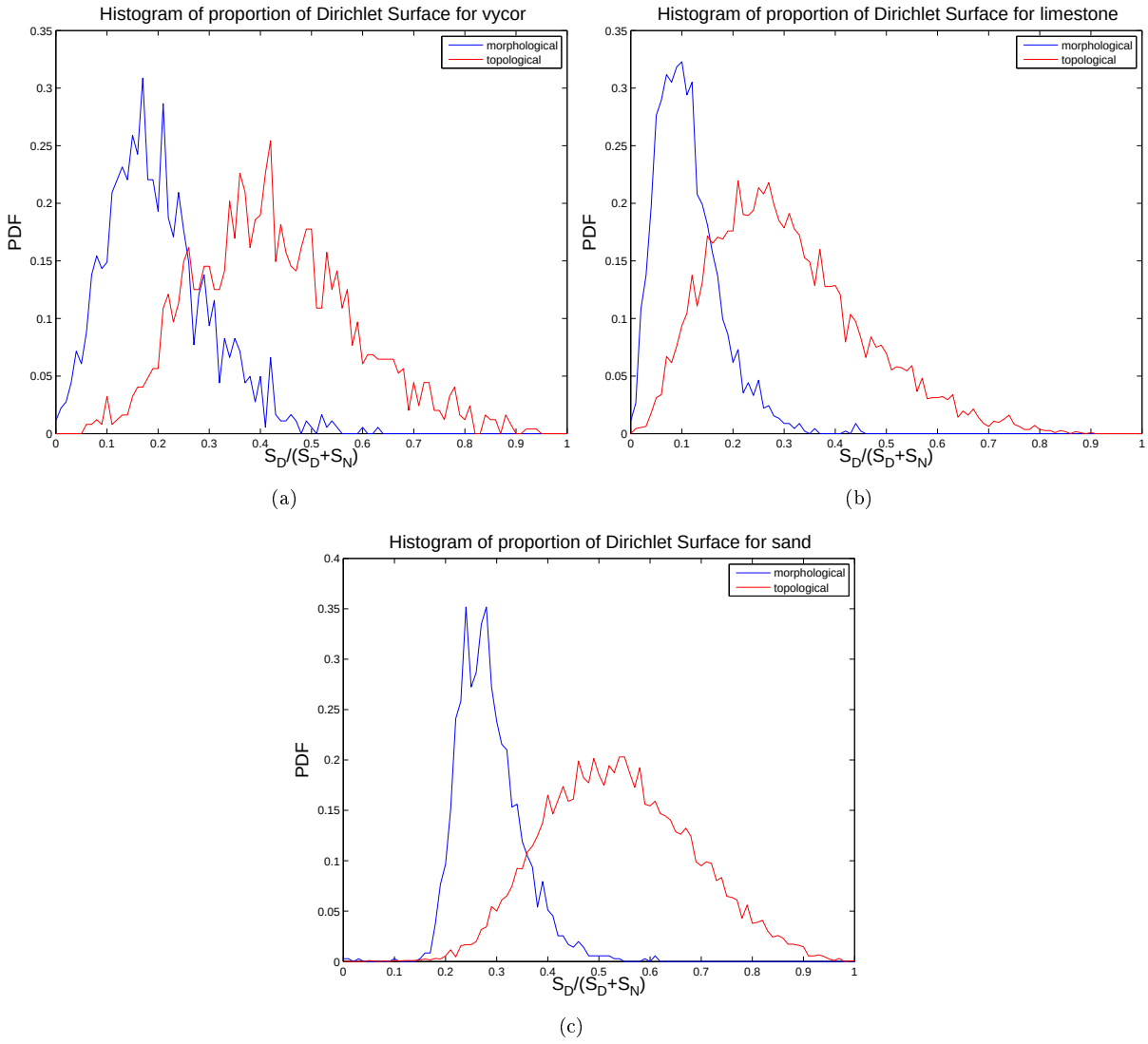


Figure 5.44: Proportion of Dirichlet surface. (a): vycor, (b): limestone, (c): sand.

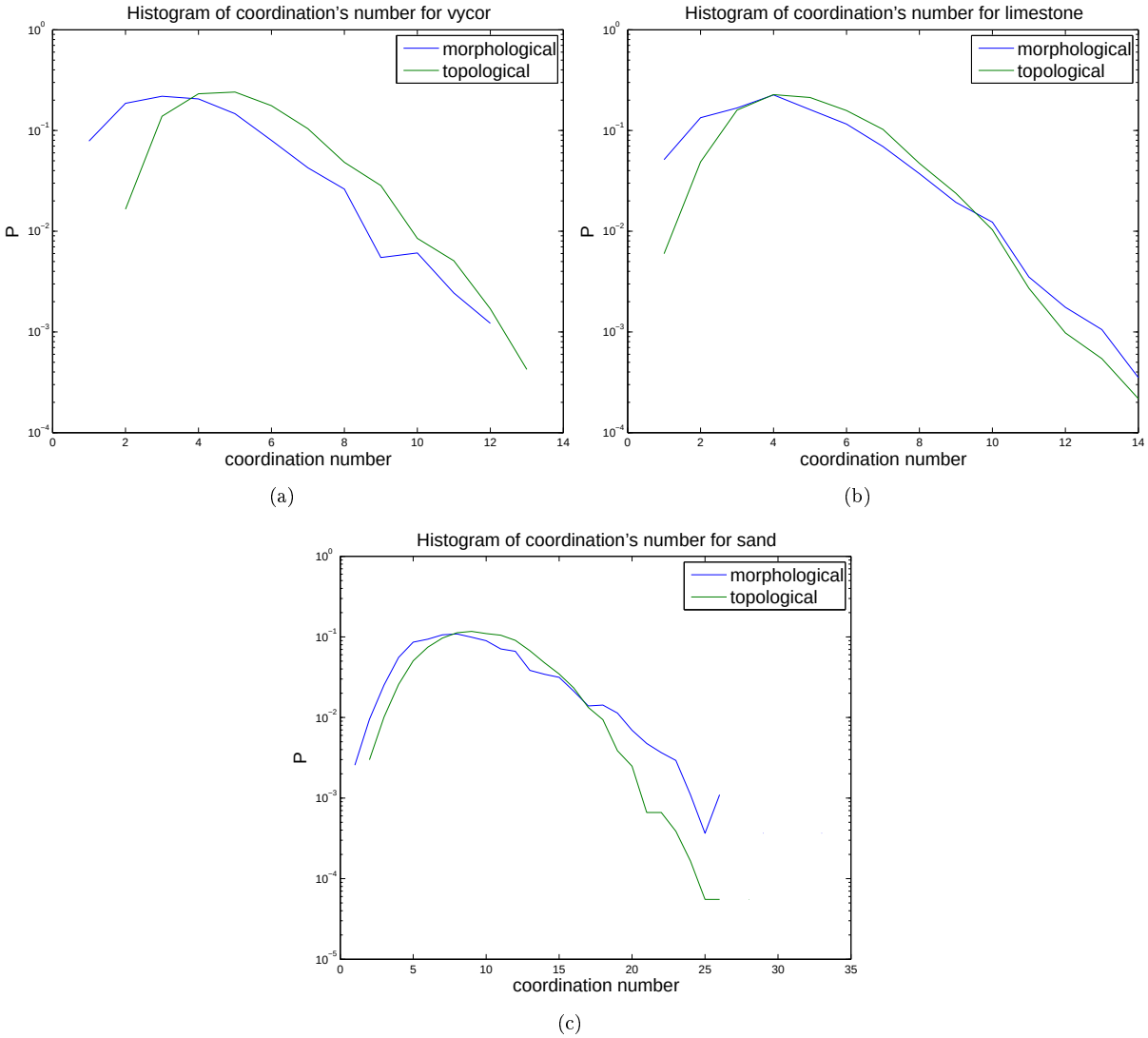


Figure 5.45: Distribution of coordination number. (a): vycor, (b): limestone, (c): sand.

Conclusion

In this chapter, we have segmented the geometrical analysis in two classes and for each class, the analysis has be done using some scalars and some statistical functionals:

1. Stereological metric (accessible in 2D and 3D)
   (a) Scalar analysis
      i. The graph of degree of hydration versus time of hydration has confirmed that the reaction of dissolution/precipitation depends on the temperature. For alite, the influence of temperature is stronger.
      ii. For OPC, there is a good agreement between the volume fraction of phases depending of the degree of hydration observed by tomography and by SEM. This coherence emphasises the quality of segmentation.
      iii. For OPC, there is also a good agreement between the specific surface area of the anhydrous grain phase and portlandite phase depending of the degree of hydration observed by tomography and by SEM.
   (b) Statistical functional analysis
      i. The 2-point probability function is independent of the angle that confirms the isotropy of the material,
      ii. The length, associated to the convergence of the 2-point probability function, is associated to the length of the representative elementary volume. This length is equal to 200 μm for OPC and 600 μm for alite. This difference is due to the presence of large clusters of portlandite in alite.
      iii. the phase can be classified in three kinds of disorder depending on the distribution of chord length:
         - For the portlandite/pore/hydrate phases, it is Debye randomness. It is important information since this kind of disorder can be reconstructed easily.
         - To validate this class of disorder for the pore network, the distribution of chord length has been plotted at different resolution, $0.675\mu m$, $0.25\mu m$ and $0.1\mu m$. We find that the local disorder of the capillary pore seems independent to the chosen resolution.
         - For the anhydrous phase, the distribution follows a power law. It is the signature of power law poly-dispersion of the granular phase. To validate this observation, we measure the volume distribution of the connected components of the anhydrous grains on the tomography images (this measure is not stereological). Also, we find a power law distribution with a cross over. The power law distribution can be due to the crush of a material and the kink is probably due to the sieve size introduced during the crush of clinker. In the same figure, we plot the volume distribution of the connected components of the anhydrous grains at two different degrees of hydration 0.55 and 0.65. The distributions are the same. Therefore, the volume distribution does not depend on the degree of hydration. The first consequence is that the characteristic length scale of this phase is constant in range of degree of hydration.
2. Topology (accessible only in 3D),
   (a) Scalar analysis
   As noticed by Gallucci et *al* [55], "the main drawback of the Synchrotron X-ray Tomographic Microscopy is its spatial resolution which is at present quite low relative to the sizes of capillary pores controlling transport properties in mature pastes." This assertion is confirmed by the observation of the disconnection of the pore network at time=150h (degree of hydration equals 0.713045). This disconnection means that the capillary porosity is weakly connected implying a coupling between the pore network at a sub-scale. The porosity, developed in hydrate phase at the sub-micro scale, starts to be the main path for diffusion. In the next future, the nano-tomography with a resolution of $0.1\mu m$, will be suitable to observe the evolution of capillary porosity controlling transport properties in mature pastes.

(b) statistical functional analysis.
In order to characterise the connectivity of the network, we operate the transformation of the microstructure to the topological graph. The mean coordination number, $< N_c(v) >$ is related to the Euler-Poincaré invariant and the number of vertexes:

$$-\frac{< N_c > -2}{2} = \frac{N_3}{\alpha_0}$$

This number $-\frac{N_3}{\alpha_0}$ has two important properties: intensive and invariant by a homotopic transformation. Therefore, it allows the comparative analysis of the topology of different materials, bone, cement, wood, etc.

To have a tool to understand the link between the geometrical properties and a physical phenomenon, we decompose the microstructure in term of elementary pores following two conventions: morphological and topological. Both decompositions are based on the seed localisation:

- seeds as minima of the opposite distance function of the microstructure after a dynamic filtering for the morphological decomposition,
- seeds as vertexes or edges of the thinning skeleton of the microstructure for the topological decomposition,

and a growing process:

- watershed transformation on the topographic surface equal to the opposite distance function of the microstructure for the morphological decomposition,
- geodesic influence for the topological decomposition,

The advantages of the morphological decomposition are:

- the surfaces between adjacent pores are located on the throat as confirmed by the plot of the distribution of the ratio of the Dirichlet surface area to the total surface area,
- the connectivity of the elementary pore network is similar to that the topological pore network.

The major drawback of the morphological decomposition is that the volume distribution presents a large dispersion. The advantages of the topological decomposition is the topology is preserved.

6

Testing 2D→3D reconstruction of multi-phase porous media obtained by SEM

Contents

The purpose of this chapter is to reconstruct a representative model of the real microstructure using the segmented image of cement paste obtained by SEM. As this thesis is financed by a CIFRE grant (industrial research training grant), the industrial partner (ATILH) has fixed one constraint: the method must be simple.

6.1 Introduction

The imaging technique in cement industry is SEM allowing the observation of a 2D serial section. Experimentally, a 3D image may be built up from a set of 2D serial section. However, this task is tedious and time consuming. Therefore, this method cannot be used in an industrial process. A phenomenological approach can be used in order to build a comportment law between some geometrical properties and a physical/mechanical property for instance compressive strength [129], diffusion [180]. The calibration is done by the measure of the geometrical properties in the 2D section and by the direct measure of physical/mechanical property experimentally. But, when a new material is designed, a great discrepancy of the geometrical organisation between the new microstructure and the microstructures used to build the comportment law, can be found. Therefore, the new material will not follow the compartmental laws.

Thus, the cement industry has to to spend time to make the laboratory tests. This chapter will describe a method for generating a 3D-image from a 2D-image allowing the prediction of physical properties numerically. This process aims to convert a number of real physical tests by numerical prediction in order to save time and money for the cement industry.
There is two ways for the generation of 3D-images. Briefly,

1. the first one, called morphogenesis, is based on the simulation of the physical and chemical process. In cement, the hybrid VCCTL (Virtual Cement and Concrete Testing Laboratory) provided at NIST allows the simulation of hydration of cement-based. At each step of the simulation, a hydration cycle is segmented in three steps: dissolution, random-walk of the mobile agents, precipitation. The simulation of dissolution and precipitation is based on Cellular Automata, a deterministic approach with fixed rule [42]. Therefore, the chemical kinetics are not taken into account. For cement, containing more than 15 different phases, this simplification has to be done in order to make the simulation in a reasonable time. Nowadays, phase-field modeling method has become an important and extremely versatile technique for simulating microstructure evolution at the mesoscale [119]. Historically, this model comes from Cahn and Hiliard [30] for the description of the spinodal decomposition using diffusive interface in a thermodynamical formulation. A comparative analysis between the measures done on the numerical model and the measures done experimentally allows the quantitative validation.

2. the second one deals with the capability of the generation of a representative 3-D microstructure using geometrical information. For this purpose, we can distinguish two schools: mathematical and physical. For the mathematical school, further pioneering work on point processes, random sets and random function models was performed by Georges Matheron [112]. Most of the papers related of these models are focused on the mathematical aspect. However, these models are often suitable for the description and the simulation of heterogeneous materials [77, 163]. In the other hand, physicists aim to find some suitable models for the description of the structure of a given media:

   - the structure of a liquid may be regarded as a 'heap' of molecules and several properties of liquids have been shown to have their geometrical counterparts in a random close packing of sphere [152].
   - the spinodal decomposition is a phase separation occuring uniformly throughout the material. Cahn [30] was proved that the correlated Gaussian random fields [18, 81, 103, 166] is the solution of the thermodynamical formulation.

   All these models form a catalogue and, depending on the media, one or a combination of these models has to be chosen in order to match the geometrical organisation of the media. The utilisation of image analysis allows the calibration and the validation of the selected model. In the purpose of this thesis, since the geometrical organisation of cement paste is complex, it seems difficult to find a suitable model.

Simulated annealing reconstruction [65, 87, 116, 183] is an optimisation technique which involves minimisation of an objective function through perturbations in the spatial organisation of the microstructure. The objective function is defined by the stereological morphological distribution functions in order to reconstruct a 3D model having the same morphology as the 2D slice. This model does not involve any hypothesis of the geometrical organisation. It is the major advantage of this model. But, the major assumption of this model is that all the information of the geometrical organisation is contained in the morphology of the media. For many media, this assumption is not verified. For instance, a granular phase has no irreducible cycles and no holes. This is a strong topological information involving that the Betty numbers, β_0 and β_1 are equals to zero. Thus, we will find a great discrepancy of the geometrical organisation between the model reconstructed with the simulated annealing method and the real grain phase. In cement paste, the granular phase, the anhydrous grains, plays an important role for the mechanical properties. For the diffusion transport, the purpose of this thesis, the geometrical organisation between the hydrate phase and the void phase is the more important. These two phases does not have any topological constraints. Thus, simulated annealing reconstruction should lead to consistent results in relation

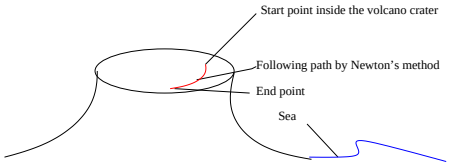


Figure 6.1: Gradient slope

with the diffusion transport. As this method requires a huge computational cost for the reconstruction of a model larger than the representative elementary volume, this chapter will presented two ways of optimisation. We will describe how to actualise the objective function after a perturbation efficiently and how to localize the perturbation [144]. For the validation of this method for cement paste, we will reconstruct some 3D models in using the objective functions defined with the real 3D microstructures obtained by X-ray tomography. Then, we will compare the diffusion/topological properties between the models and the real microstructures.

6.2 Metropolis algorithm for the 3D reconstruction

Let us consider a walker on the mountain wanting to reach the sea. The naive strategy would be to follow the direction where the slope is minimum to come down. This determinism strategy is a gradient slope, called Newton's method. However, when the topographic surface is not convex, the walker can stay a local minima (see figure 6.1). To overcome this problem, another strategy is the Metropolis's algorithm. At each point, a random direction is chosen. If the slope is negative, the walker follows this direction. Otherwise the walker follows this direction depending on a probability that is low if the slope is very positive and inversely. With this strategy, the walker will find the sea. But, the time to join the sea is unknown. The walker is like a human. At the beginning, he accepts to climb (positive slope) easily but, progressively, he will accept less and less. This strategy is the simulated annealing. The advantages are exploration of the phase space and constant time of convergence. In the context of the 3D reconstruction of a multi-phase material, we will define:

1. the phase space,
2. the objective function,
3. the probability space (the topographic surface),
4. the Metropolis algorithm,
5. the annealing simulated,
6. the perturbation (the footstep),
7. the actualisation of the stereological morphological functions.

6.2.1 Phase space

The algorithm works at constant volume fractions. The phase space, $\mathrm{E}(\mathcal{R})$, is the set of models which are the same volume fraction as the reference microstructure, $\mathcal{R}$. The size of this phase space is huge. For two phase microstructure, we have $|\mathrm{E}(\mathcal{R})| = C_{|E|}^{\phi|E|}$ where ϕ is the volume fraction, $|E|$ is the number of voxels of the image and C_n^p is the binomial coefficient ($C_n^p = \frac{n!}{p!(n-p)!}$). With a size of only 100 voxels and a porosity of 0.5, we have $C_{100}^{50} > 10^{93}$ of possible states. This number is upper than the number of atoms in the known universe. It is impossible to explore all the states of the phase space. Therefore, we need a strategy to reach the global minima without exploring the whole space phase.

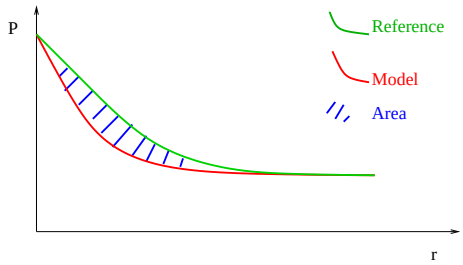


Figure 6.2: The objective is to minimise the area between the two graphs for each stereological metric functions.

6.2.2 Objective function

Simulated annealing reconstruction is an optimisation technique, which involves minimisation of an objective function. In the case of the 3D reconstruction, the objective function is the distance between stereological metric functions measured on the reference and on the model. In this thesis, we use the stereological metric functions introduced in the section 5.2:

- the 2-point probability functions, $S_{2,i,j}$ where i and j are the phase labels[67].
- the chord length distribution functions, $f_{\mu,i}$.

As the cement paste is isotropic (the subsection 5.2.4), these functions are calculated depending only on a distance without a direction. Let $\mathcal{M}$ be a model (it belongs to the set $\mathrm{E}(\mathcal{R})$) and $\mathcal{R}$, the microstructure reference . The distance between the model and the reference is defined by:

$$d(\mathcal{M},\mathcal{G}) = \alpha_0\Big(\sum_{i\in Phase}\|f_{\mu,i}(\mathcal{M})-f_{\mu,i}(\mathcal{G})\|\Big)+\alpha_1\Big(\sum_{i,j\in Phase}\|S_{2,i,j}(\mathcal{M})-S_{2,i,j}(\mathcal{G})\|\Big)$$

where the $\|f-g\|$ is the "area" between the function f and g (see figure 6.2).

6.2.3 Probability space

The Metropolis algorithm is based on a probabilistic approach. At each perturbation, we need to know what is the probability, P, to have the model, $\mathcal{M}$, given the reference, $\mathcal{R}$. The Boltzmann distribution allows the definition of a probability space:

$$P(\mathcal{M}|\mathcal{R}) = \frac{1}{Z}\exp\left(-\frac{E(\mathcal{M},\mathcal{R})}{T}\right)$$

[67] For a bi-phase material, the auto-correlation of one phase can be expressed as a function of the other phase.

$$\begin{aligned} S^{\mathcal{I}}_{1,1}(r) &= \frac{1}{\mu(E)}\int_E \mathbb{1}_{\mathcal{I}(r')=1}\mathbb{1}_{\mathcal{I}(r'+r)=1}dr' \\ S^{\mathcal{I}}_{1,1}(r) &= \frac{1}{\mu(E)}\int_E (1-\mathbb{1}_{\mathcal{I}(r')=0})(1-\mathbb{1}_{\mathcal{I}(r'+r)=0})dr' \\ S^{\mathcal{I}}_{1,1}(r) &= 1-2\phi_0+S^{\mathcal{I}}_{0,0}(r) \end{aligned}$$

The cross-correlation is: $S^{\mathcal{I}}_{1,1}(r)=\phi_1-S^{\mathcal{I}}_{0,1}(r)$. Hence, the information in one 2-point probability function contained the information of the others. However, for a three-phase material, we have:

$$\begin{aligned} S^{\mathcal{I}}_{2,2}(r) &= \frac{1}{\mu(E)}\int_E \mathbb{1}_{\mathcal{I}(r')=2}\mathbb{1}_{\mathcal{I}(r'+r)=2}dr' \\ S^{\mathcal{I}}_{2,2}(r) &= \frac{1}{\mu(E)}\int_E (1-\mathbb{1}_{\mathcal{I}(r')=1}-\mathbb{1}_{\mathcal{I}(r')=0})(1-\mathbb{1}_{\mathcal{I}(r'+r)=1}-\mathbb{1}_{\mathcal{I}(r'+r)=0})dr' \\ S^{\mathcal{I}}_{2,2}(r) &= 1-2(\phi_1+\phi_0)+2S^{\mathcal{I}}_{0,1}(r)+S^{\mathcal{I}}_{1,1}(r)+S^{\mathcal{I}}_{0,0}(r) \end{aligned}$$

Hence, each 2-point probability function contains a specific information. Therefore, for microstructure containing more than two phases, all the 2-correlations have to be taken into account in the objective function.

where Z is the canonical partition function[68], T is the temperature and E, the energy, is equal to the distance just previously defined. In this formalism, the problem is to find a model, $\mathcal{M}$, maximising this probability. As the phase space is too big to test all the states, an exploration procedure is defined.

6.2.4 Metropolis algorithm

The Metropolis-Hastings algorithm is a Markov chain. A Markov chain is a step by step stochastic process such that the state at step $t+1$ depends only on the state at step t. A stochastic process is the fact that the future states will be reached through a probabilistic process instead of a deterministic one.
Let $\mathcal{M}^t$ be a model at step t.
The recurrence part of the Metropolis algorithm is:

1. choose $\mathcal{G}$ with a probability equal to $Q(\mathcal{M}^t, \mathcal{G})$,
   (a) accept $\mathcal{G}$ with probability $\rho(\mathcal{I}_n, \mathcal{G})$ $(\mathcal{M}^{t+1} = \mathcal{G})$
   (b) refuse $\mathcal{G}$ $(\mathcal{M}^{t+1} = \mathcal{M}^t)$

where

- Q is the selection matrix (the perturbation),
- ρ is the acceptance matrix[69]:

$$\begin{aligned} \rho(\mathcal{M}^t, \mathcal{G}) &= \min(1, \frac{P(\mathcal{G}|\mathcal{R})}{P(\mathcal{M}^t|\mathcal{R})}) \\ \rho(\mathcal{M}^t, \mathcal{G}) &= \min(1, \exp\left(-\frac{d(\mathcal{G}-\mathcal{R}) - d(\mathcal{M}^t - \mathcal{R})}{T}\right) \end{aligned}$$

Mathematically, the transition matrix of this Markov chain is :

$$P(\mathcal{F}, \mathcal{G}) = \begin{cases} Q(\mathcal{F}, \mathcal{G})\rho(\mathcal{F}, \mathcal{G}) \text{ if } \mathcal{F} \neq \mathcal{G} \\ 1 - \sum_{\mathcal{F} \neq \mathcal{G}} P(\mathcal{F}, \mathcal{G}) \text{ else} \end{cases}$$

6.2.5 Simulated annealing algorithm

The Metropolis algorithm works at constant temperature. In practice, we observe two cases: the temperature is low, so "only" the transitions decreasing the energy are accepted. Therefore, it is the same as the Newton's slope with the problem to reach only a local minimum. In the other case, the temperature is high, so "all" the transitions are accepted. Therefore, it is like a random walk in the phase space without an exploring strategy.
The simulated annealing algorithm does not work at constant temperature. At the beginning, the temperature is high in order to explore the phase space and progressively, the temperature decreases in order to reach a minima in a reasonable time. Mathematically, when t tends to infinity, a global minimum is reached if the temperature follows a law in $\frac{1}{\ln(t)}$. Numerically, with this law, the temperature decreases too slowly and, thus, we come back to the case of the Metropolis algorithm. Therefore, is is usual to choose a temperature dependence running as $\frac{1}{t}$.

6.2.6 Perturbation

The perturbation is the permutation of two voxels x and y such that the partition, $\mathcal{G}$, is

$$\mathcal{G}(z) = \begin{cases} \mathcal{M}^t(z) \text{if } z \neq x \wedge z \neq y \\ \mathcal{M}^t(y) \text{if } z = x \\ \mathcal{M}^t(x) \text{if } z = y \end{cases}$$

[68] $Z = \sum_{\mathcal{I}} P(\mathcal{I}|\mathcal{R})$
[69] The canonical partition function vanishes that is the trick of the metropolis algorithm.

The localisation of the pair of the permutation voxels is chosen following two conventions: random on the model or random on the phase interfaces of the model. The classical permutation is to select randomly two voxels belonging to two different phases. As seen in the chapter characterization, the resolution of the image has to be chosen much smaller than the smallest feature size of the material in order to describe accurately the microstructure. In this hypothesis, whatever the geometrical shape, there are some clusters meaning that the majority of voxels are surrounded by voxel of the same phase. When the two voxels are selected randomly (selection matrix), in most of case, this permutation generate isolated voxels. The microstructure due to this permutation will be rejected because this permutation goes in the opposite way of the aggregation. So, many permutation will be rejected. To minimise the ratio of rejection, an idea is to localise the permutation in order to aggregate the phases. A solution is to localise the permutation on the interface of the phases. The localisation on the permutation on the interface of the phases has been introduced by Rozman and Utz [144]. The framework, introduced in the chapter 3, allows the implementation of this approach minimising the time needed to localise a voxel on the interface.
Let $(X_0^t, \ldots, X_n^t)$ be the domains of the phase at time t and ϕ_i the volume fraction of the phase i.
The selection of the permutation voxels, x and y, is:

1. choose a phase, i, with a probability equal to ϕ_i,
2. choose a voxel, x, on the outer boundary of the phase, Z_i^t (see subsection 3.3.1)

$$Z_i^t = (X_i^t \oplus V) \setminus X_i^t$$

   Let j be the other phase ($\mathcal{M}^t(x) = j$).

3. choose a voxel y on the outer boundary of the phase j, Z_j^t while the phase on y is not i ($\mathcal{M}^t(y) = i$) [70].

6.2.7 Actualisation of distance function depending on a permutation

As the model evolves at each perturbation, we have to update the stereological functions calculated on the model after each perturbation. To speed-up this actualisation, these functions will be only calculated on 3 directions. Numerically, we store all the correlations and all the chords of the model calculated on the 3 directions in tables. These tables are actualised after each permutation. A permutation is two times the modification of the state of a voxel. The modification of the voxel y is:

$$\mathcal{I}^{t+1}(x) = \begin{cases} \mathcal{I}^t(x) \text{if } x \neq y \\ j \neq \mathcal{I}^t(x) \text{if } x = y \end{cases}$$

In order to avoid some boundary effects, we impose a periodic condition[71].

Actualisation of 2-point probability function

To update the 2-point probability function, the utilisation of Fast Fourier Transform (FFT) is an efficient procedure [34] requiring $O(N \ln N + N)$ operation where N is the number of voxels on the image. An alternative is to formulate the 2-point probability function after the perturbation depending on the 2-point probability function before the perturbation plus an additional term [144]. This method requires $O(N)$ operations if the correlation is calculated on all direction and $O(N^{1/3})$ if the correlation is only calculated on the 3 orthogonal directions.

[70] To assure the conservation of the phase volume fractions, we have to choose randomly a voxel y belonging to: $Z_j^t \cap X_i^t = \left((X_j^t \oplus V) \setminus X_j^t\right) \cap X_i^t$

[71] The value on each voxel, x, is calculate using this relation:

$$\mathcal{I}(x) = \mathcal{I}(x \% s^{\mathcal{M}})$$

where $s^{\mathcal{I}}$ is the size of the image and % is the modulo operator.

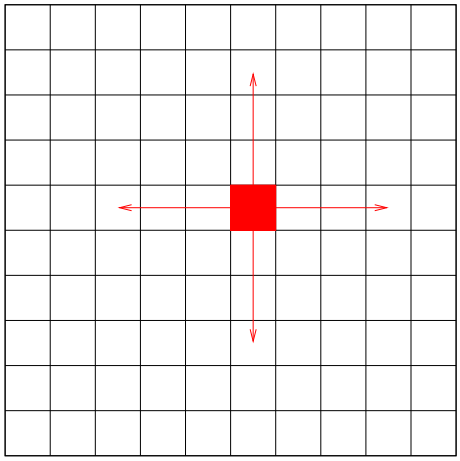


Figure 6.3: To update the auto-cross correlation functions, only basic operations have to be done on the red arrows.

In this subsection, we explain the actualisation for the auto-correlation but the application for the cross-correlation is straightforward. The number of auto-correlation of the phase i, at the distance d, on the three direction, $\vec{e_0}, \vec{e_1}, \vec{e_2}$, for the partition $\mathcal{I}^t$, is defined by this relation:

$$N_{i,i}^{\mathcal{I}^t}(d) = \overbrace{\sum_{x\in E}}^{\text{all the voxels}} \overbrace{\sum_{\vec{e}\in(\vec{e_0},\vec{e_1},\vec{e_2})}}^{\text{the three main directions}} \mathbb{1}_{\mathcal{I}^t(x)=i}\mathbb{1}_{\mathcal{I}^t(x+d\vec{e})=i}$$

We formulate the auto-correlation at time $t+1$, $N_{i,i}^{\mathcal{I}^{t+1}}(d)$, depending on the the auto-correlation at time t, $N_{i,i}^{\mathcal{I}^t}(d)$:

$$\begin{aligned}
N_{i,i}^{\mathcal{I}^{t+1}}(d) &= \sum_{\vec{e}\in(\vec{e_0},\vec{e_1},\vec{e_2})}\sum_{x\in E}\mathbb{1}_{\mathcal{I}^{t+1}(x)=i}\mathbb{1}_{\mathcal{I}^{t+1}(x+d\vec{e})=i}\\
N_{i,i}^{\mathcal{I}^{t+1}}(d) &= \sum_{\vec{e}\in(\vec{e_0},\vec{e_1},\vec{e_2})}\Big(\sum_{x\in E\setminus\{y,y-d\vec{e}\}}\mathbb{1}_{\mathcal{I}^{t+1}(x)=i}\mathbb{1}_{\mathcal{I}^{t+1}(x+d\vec{e})=i}\\
&\quad +\mathbb{1}_{\mathcal{I}^{t+1}(y)=i}\mathbb{1}_{\mathcal{I}^{t+1}(y+d\vec{e})=i}+\mathbb{1}_{\mathcal{I}^{t+1}(y-d\vec{e})=i}\mathbb{1}_{\mathcal{I}^{t+1}(y)=i}\Big)\\
N_{i,i}^{\mathcal{I}^{t+1}}(d) &= N_{i,i}^{\mathcal{I}^t}(d)+\sum_{\vec{e}\in(\vec{e_0},\vec{e_1},\vec{e_2})}\Big(\\
&\quad (\mathbb{1}_{\mathcal{I}^{t+1}(y)=i}\mathbb{1}_{\mathcal{I}^{t+1}(y+d\vec{e})=i}+\mathbb{1}_{\mathcal{I}^{t+1}(y-d\vec{e})=i}\mathbb{1}_{\mathcal{I}^{t+1}(y)=i})\\
&\quad -(\mathbb{1}_{\mathcal{I}^{t}(y)=i}\mathbb{1}_{\mathcal{I}^{t}(y+d\vec{e})=i}+\mathbb{1}_{\mathcal{I}^{t}(y-d\vec{e})=i}\mathbb{1}_{\mathcal{I}^{t}(y)=i})\\
&\quad \Big)
\end{aligned}$$

Hence, the auto-correlation at time $t+1$ is the auto-correlation at time t with the addition of an extra term. The calculation of the extra term is only two simple tests on the three mains directions. As the update has to be done for the distance range of $[0:N^{1/3}]$, it requires $O(N^{1/3})$ operations.

Actualisation of chord length distribution

The number of chords of the phase i which size is d, for the partition $\mathcal{I}^t$, is defined by this relation (see figure 6.4):

$$\mathcal{N}_i^{\mathcal{I}^{t=0}}(d) = \sum_{x\in E}\sum_{\vec{e}\in(\vec{e_0},\vec{e_1},\vec{e_2})} \overbrace{\mathbb{1}_{\mathcal{I}^t(x)\neq i}}^{\text{other phase}} \overbrace{\mathbb{1}_{\mathcal{I}^t(x+1e)=i}\cdots\mathbb{1}_{\mathcal{I}^t(x+de)=i}}^{\text{same phase of length } d} \overbrace{\mathbb{1}_{\mathcal{I}^t(x+(d+1)\vec{e})\neq i}}^{\text{other phase}}$$

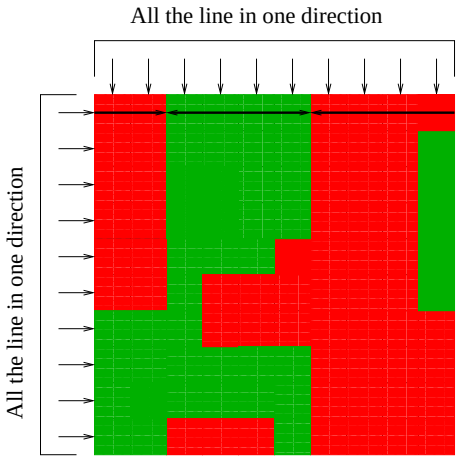


Figure 6.4: On each line (the rows and the columns) of the images, we find all the chords. On the first raw of this figure, there are two chords: one red of distance 6 and one green of distance 4, so we add 1 to $\mathcal{N}^{\mathcal{I}^t}_{red}(6)$ and 1 to $\mathcal{N}^{\mathcal{I}^t}_{green}(4)$

First, we find the phase label at left and at right of the voxel, y, on each direction:

$$\begin{aligned} p^{\vec{e}}_l(y) &= \mathcal{I}^{t+1}(y - 1\vec{e}) \\ p^{\vec{e}}_r(y) &= \mathcal{I}^{t+1}(y + 1\vec{e}) \end{aligned}$$

Second, we find the length of the chord at left and at right of the voxel y, on each direction:

$$\begin{aligned} l^{\vec{e}}_l(y) &= \arg\max_{d\in\mathbb{N}^+}(\mathbb{1}_{\mathcal{I}^t(x-2e)=\mathcal{I}^t(x-e)}\dots\mathbb{1}_{\mathcal{I}^t(x-(d+2)e)=\mathcal{I}^t(x-e)}) + 1 \\ l^{\vec{e}}_r(y) &= \arg\max_{d\in\mathbb{N}^+}(\mathbb{1}_{\mathcal{I}^t(x+2e)=\mathcal{I}^t(x+e)}\dots\mathbb{1}_{\mathcal{I}^t(x+(d+2)e)=\mathcal{I}^t(x+e)}) + 1 \end{aligned}$$

We can prove that:

$$\begin{aligned} \mathcal{N}^{\mathcal{I}^{t+1}}_i(d) &= \mathcal{N}^{\mathcal{I}^t}_i(d) + \sum_{\vec{e}\in(\vec{e_0},\vec{e_1},\vec{e_2})} (\\ \mathbb{1}_{\mathcal{I}^{t+1}(y)=i} & \quad (\mathbb{1}_{p^{\vec{e}}_l=i}\mathbb{1}_{p^{\vec{e}}_r=i}\mathbb{1}_{l_l+l_r+1=d} + \mathbb{1}_{p^{\vec{e}}_l=i}\mathbb{1}_{p^{\vec{e}}_r\neq i}\mathbb{1}_{l_l+1=d} + \mathbb{1}_{p^{\vec{e}}_l\neq i}\mathbb{1}_{p^{\vec{e}}_r=i}\mathbb{1}_{l_r+1=d} + \mathbb{1}_{p^{\vec{e}}_l\neq i}\mathbb{1}_{p^{\vec{e}}_r\neq i}\mathbb{1}_{1=d}) \\ -\mathbb{1}_{\mathcal{I}^t(y)=i} & \quad (\mathbb{1}_{p^{\vec{e}}_l=i}\mathbb{1}_{p^{\vec{e}}_r=i}\mathbb{1}_{l_l+l_r+1=d} + \mathbb{1}_{p^{\vec{e}}_l=i}\mathbb{1}_{p^{\vec{e}}_r\neq i}\mathbb{1}_{l_l+1=d} + \mathbb{1}_{p^{\vec{e}}_l\neq i}\mathbb{1}_{p^{\vec{e}}_r=i}\mathbb{1}_{l_r+1=d} + \mathbb{1}_{p^{\vec{e}}_l\neq i}\mathbb{1}_{p^{\vec{e}}_r\neq i}\mathbb{1}_{1=d}) \\ & \quad) \end{aligned}$$

Hence, the number of chords at time $t+1$ is the number of chords at time t with the addition of an extra term (see figure 6.5).

6.2.8 Numerically

Starting with a random partition with the same volume fractions as the reference microstructure, we do an annealing simulated with an objective function defined with the 2-point probability function (cross and auto). After the convergence, we add the chord length distribution function to the the objective function and we operate a new time the annealing simulated.

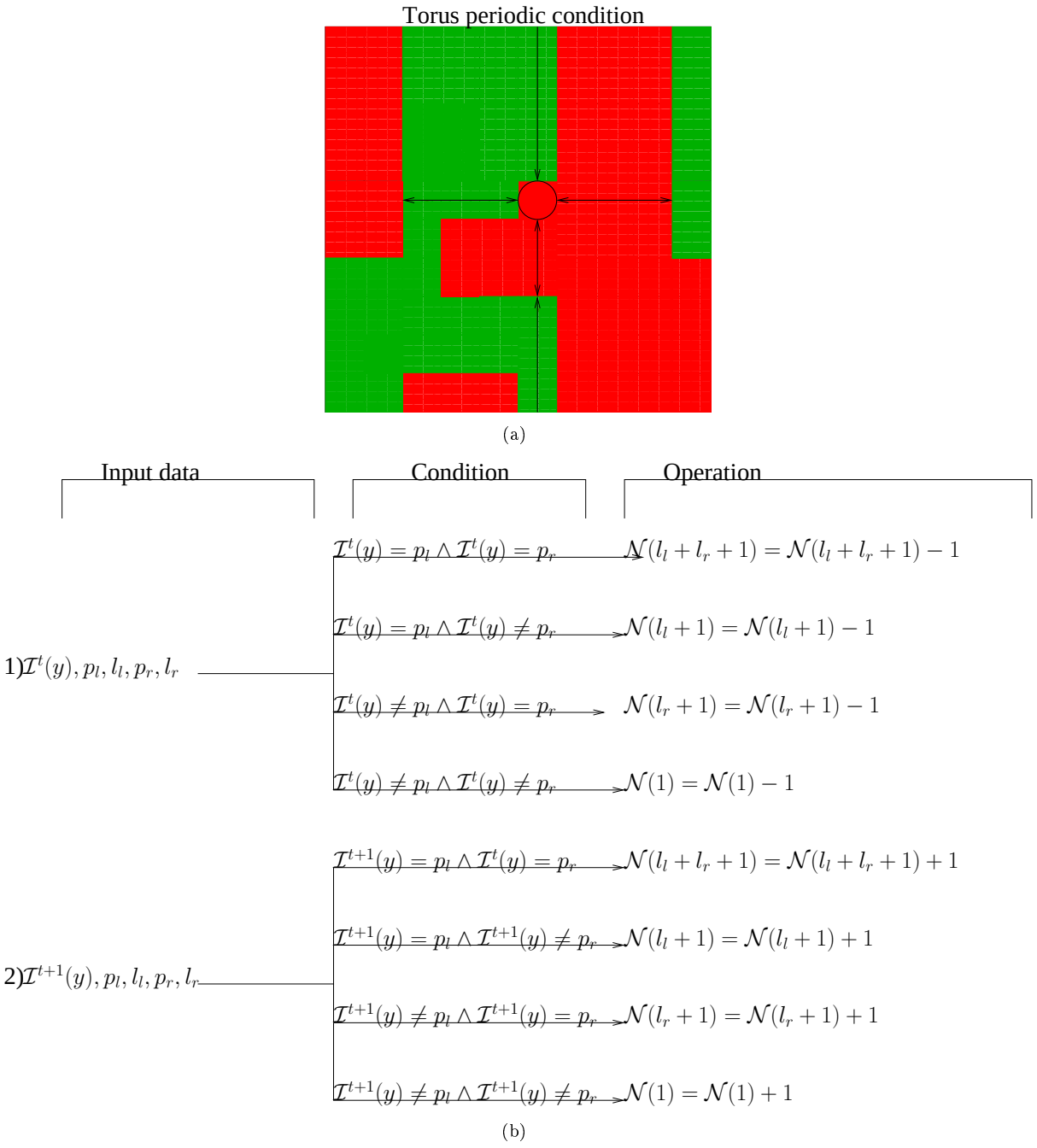


Figure 6.5: (a) To update the number of chord of each phase with a change of state of black bullet, for each direction, first we have to find the two chords starting at the left and at right of the voxel y, p_l, l_l, p_r, l_r where $\mathcal{I}^t(y)$ is the state before and $\mathcal{I}^{t+1}(y)$ is the state after. (b) Depending on conditions, the number of chord of the phase $\mathcal{I}^t(y)$ and $\mathcal{I}^{t+1}(y)$, $\mathcal{N}_{\mathcal{I}^t(y)}$ and $\mathcal{N}_{\mathcal{I}^{t+1}(y)}$ is updated.

6.3 Test and validation using 3D tomography images

Starting from SEM images of cement paste, the 3D reconstruction using simulated annealing stays on open question in two points. The first point is the possibility to reconstruct a model with a volume larger than the representative elementary volume in a reasonable time. For OPC, the REV is 200 μm^3 (see subsubsection 5.2.4). As the resolution of SEM images is 0.675 μm, the volume of 3D reconstruction has to be upper than 300^3 voxels. The second point is the representativeness of the 3D reconstruction. Of course, the 3D reconstruction has the same morphology as the real 3D microstructure since the objective function is defined by two morphological functions. But, this does not imply that the topology and the the physical properties are the same. In the section, we will deal with these two points.

6.3.1 Time of convergence: temperature and perturbation

Gradient versus simulated annealing

In order to speed-up the algorithm, the temperature can be fixed at 0 kelvin, called gradient method. In this case, the transitions increasing the energy are rejected. If the topographic surface[72] is not convex, this approach has to be avoid. As the knowledge of the convexity is a very difficult task, we test the both approaches numerically. The results are presented in the figure 6.6. In the beginning, the gradient

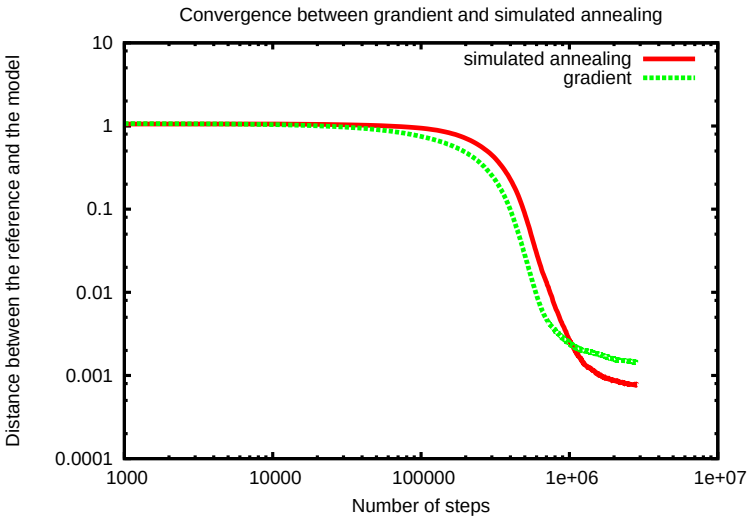


Figure 6.6: Reconstruction of a 3D image of size 30*30*30 (time=30 seconds). We plot the distance between the model and the reference depending on the number of steps.

is faster than the simulated annealing, but the time of convergence is similar for both algorithms. We observe that the energy with simulated annealing method is lower than the energy with the gradient method at the convergence. The reason is that the topographic surface is not convex. There are some local minima. For these both reasons, the simulated annealing method is selected rather than gradient method.

Localisation of the permutation on the phase interfaces

The figure 6.7 illustrates the convergence behaviour with or without the localisation of the permutation on the interface. The decrease of the energy is significantly more rapid with the permutations localisation. After two days, there is the convergence with the permutations localisation whereas there is no convergence without this localisation. Without the localisation, the acceptance ration versus the number

[72] the application of the phase space to the probability

of steps follows a power law. It is a major disadvantaged since the acceptance ratio will be rapidly very low. Therefore, the simulated annealing method with the localisation of the permutation on the interface is selected.

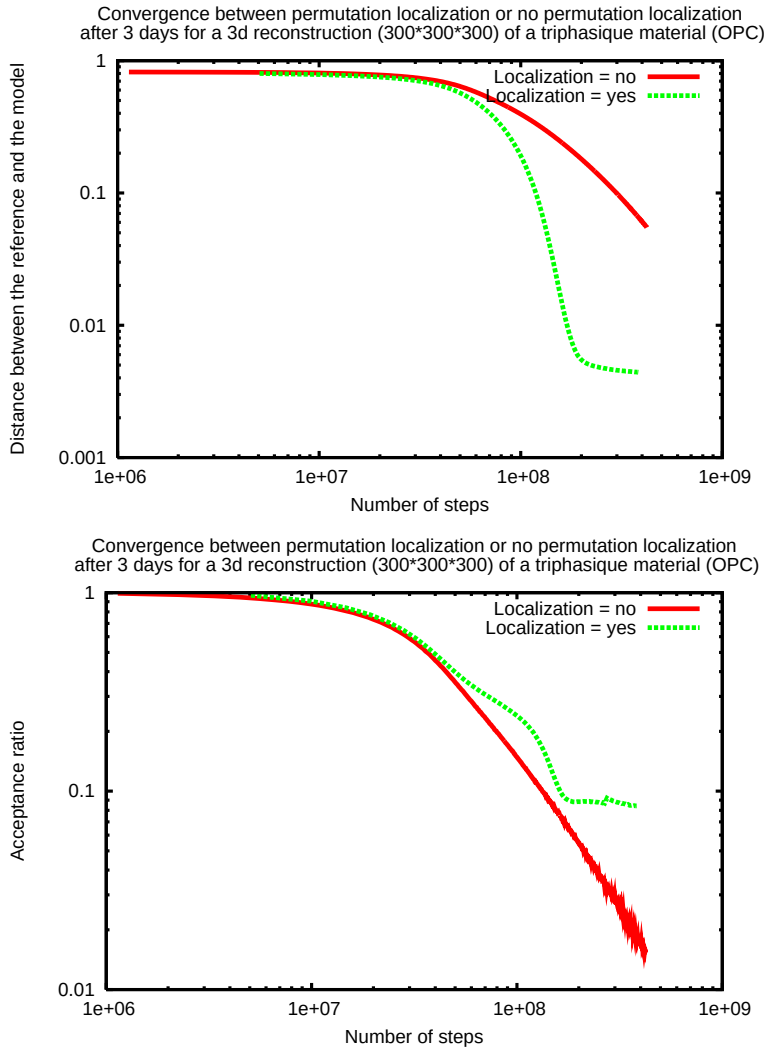


Figure 6.7: Reconstruction of a 3D image of size 300*300*300 (time=2 days). Figure 1: distance between the model and the reference versus the number of steps. it is a log-log plot. Even after 2 weeks, there is no convergence if the permutation is not localised. Figure 2: the acceptance ratio versus the number of steps.

Reconstruction of 3D models with real 3D microstructure as reference

The real 3D microstructures as reference come from the first run to the SLS of OPC at W/C=0.5:

- at time t=35h, with porosity equals to 0.205 and a degree of hydration equals to 0.482
- at time t=83h, with porosity equals to 0.176 and a degree of hydration equals to 0.628
- at time t=150h, with porosity equals to 0.079 and a degree of hydration equals to 0.713

For these three cement pastes, we have reconstructed a 3D model with a size equal to 300^3 voxels such that the objective function is defined using these microstructures (see figure 6.8).

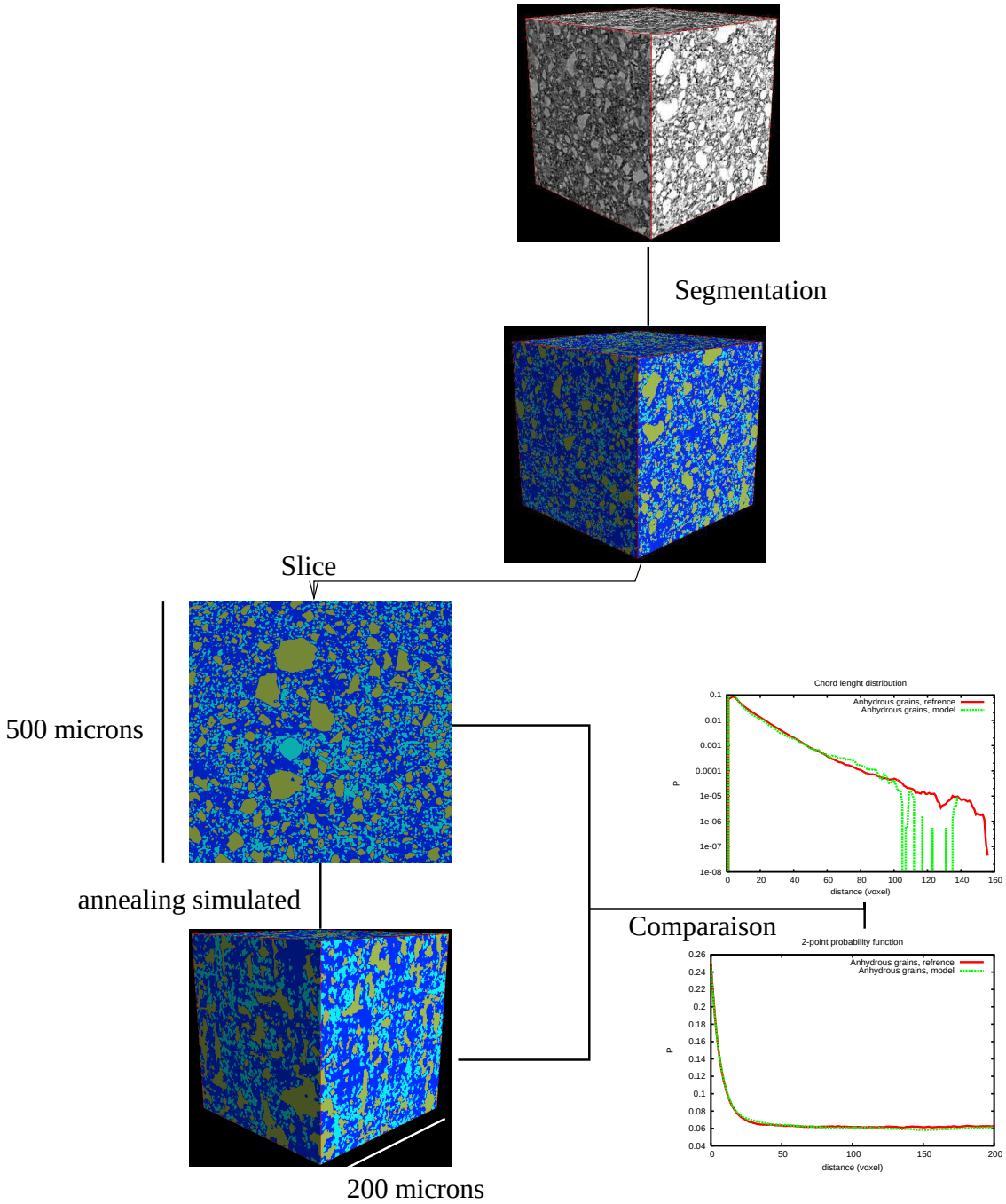


Figure 6.8: Validation using the 3D images obtained by X-ray tomography.

6.3.2 Topological and diffusion validation

Coefficient of Diffusion

The figure 6.9 illustrates the comparison between the model and the reference for the factor $D_0/D(t)$. There is a good agreement between the model and the reference. For the first two cement pastes, we observe the convergence at the same level[73]. For the last cement paste, due to the disconnectedness of the porosity, there is the convergence to a law[74] varying linearly with t for the real microstructure and the model. The slope is quite the same between both systems.

Topological properties

The table 6.1 illustrates the good agreement between the model and the reference for the topological parameters introduced in the section 5.3.

	degree of hydration	$\frac{N_3}{V}$	$< N_c >$	C	percolation
model	0.482	-2.83*10^{-4}	2.95	0.475	yes
reference	0.482	-3.37*10^{-4}	2.95	0.475	yes
model	0.628	-1.02*10^{-4}	2.37	0.189	yes
reference	0.628	-1.74*10^{-4}	2.59	0.295	yes
model	0.713	1.36*10^{-4}	0.652	-0.673	no
reference	0.713	1.29*10^{-4}	0.873	-0.5635	no

Table 6.1: Topological numbers for OPC with W/C=0.5

6.4 Conclusion and discussion

Simulated annealing algorithm was used to reconstruct a 3D multiphase material in order to allow the investigation of the diffusive transport by SEM. An optimised implementation of this algorithm has been reached using the localisation of the perturbation on the interfaces and a fast procedure of actualisation of the objective function after a perturbation. The reconstructed volume can be greater than the REV (for a computational time of 1 week, a volume of 400*400*400 can be reconstructed). For the capillary pores of cement paste, there is a good agreement on the intrinsically 3D information (topology/diffusion property) between the model and the real microstructure. In this chapter, we focus on the pore phase for the validation of this algorithm. However in the case of prediction of mechanical properties, the arrangement of each phase plays an important role. The arrangement of the granular phase (anhydrous grains) between the real microstructure and the model is very different (see figure 6.10) given a bad prediction of mechanical properties. Clearly, the strategy to reconstruct 3D "toy models" from 2D images works for some physical properties and not for all the spectrum of pore texture. For capillary pores of cement paste, simulated annealing algorithm appears to be efficient.

[73] This level is reach when the diffusion time is enough long to explore a space large greater than the REV. The value of this value correspond to the macroscopic tortuosity.

[74] Due to the disconnectedness of the porosity, the quadratic variation is bounded: $< r^2(t) >< A$. As the tortuosity is proportional to $\frac{t}{<r^2(t)>}$ and the bounding of the quadratic variation, the tortuosity follows a law in t.

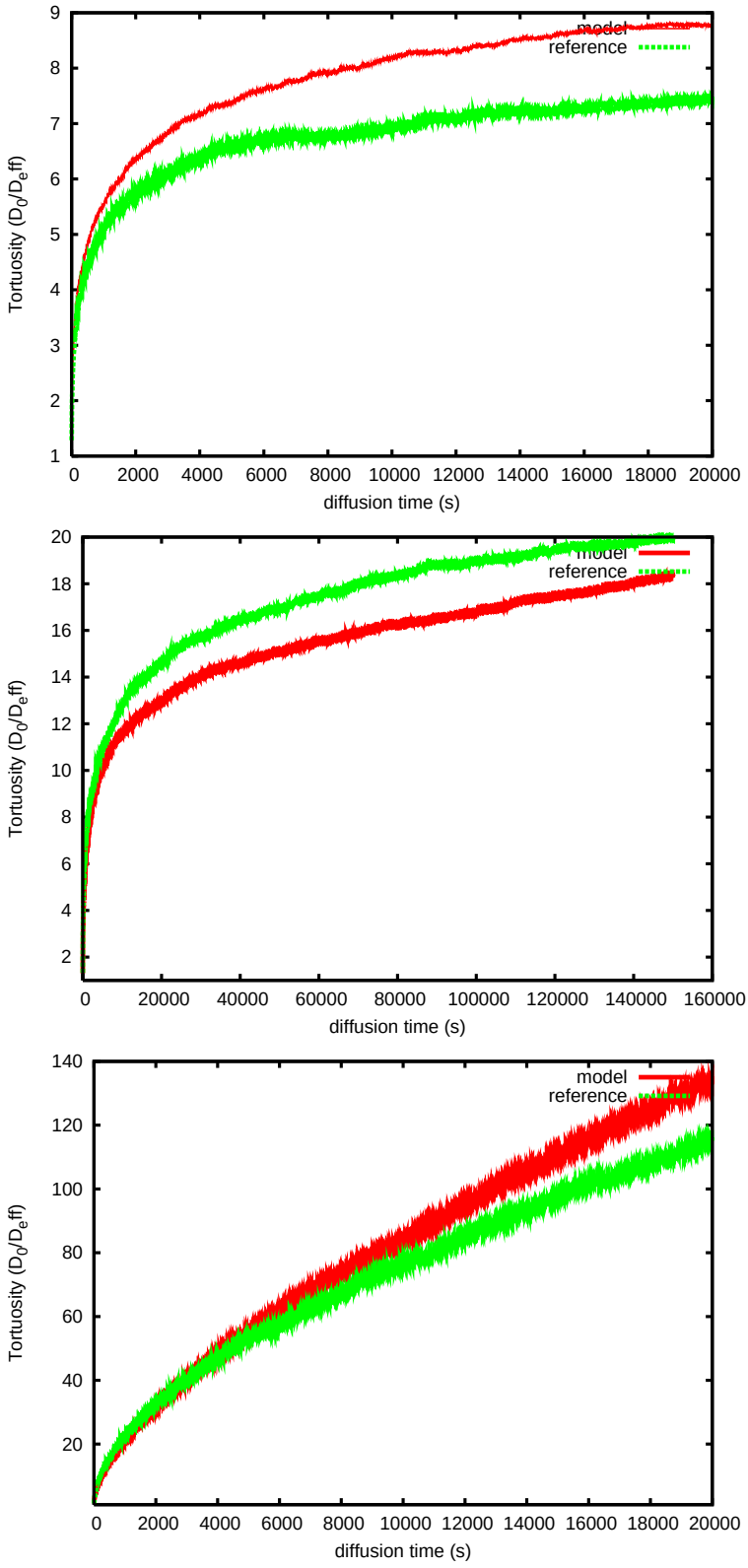


Figure 6.9: Comparison of the factor $D_0/D(t)$ between the model of size 300*300*300 and the reference obtained by X-ray tomography. Figure 1: degree of hydration=0.482; figure 2: 0.628 and figure 3: 0.713045.

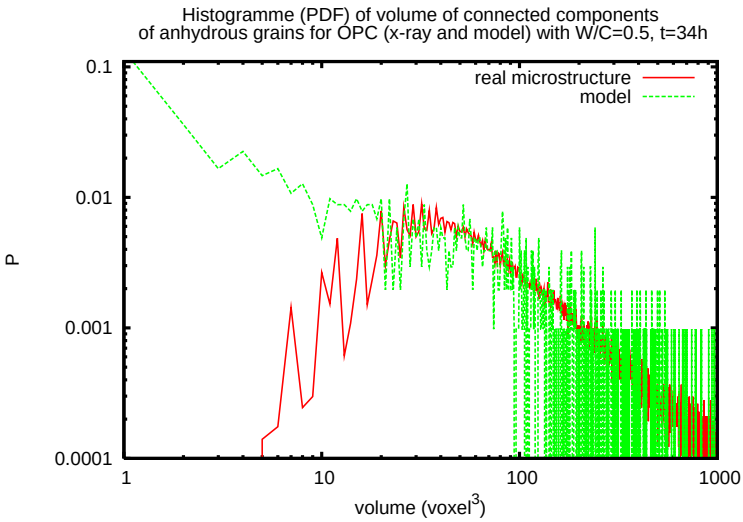


Figure 6.10: The Volume distribution of the connected components of grain phase, anhydrous phase, is very different between the model and the real microstructure.

Conclusion

The main motivation of this PhD work is to improve the processing of experimental images in order to extract quantitative information on geometrical and transport properties of porous media. This task was achieved in three successive stages.

First, experiments were performed to image cement pastes. Though cement paste is a hierarchical porous media, a limited range of length scales corresponding to the capillary pores controls transport. To observe the capillary pores, two imaging techniques are used: Scanning Electron Microscopy (SEM) and Synchrotron-based X-ray Tomography. The samples were two cement pastes, alite and Portland, with different curing time in order to observe a wide range of geometrical organizations. The drawback of both techniques is a weak contrast to noise ratio that makes the extraction of the phases by standard thresholding procedures difficult.

Second, a part of this thesis was devoted to achieve a reliable segmentation of the experimental images. We have introduced a *simple*, *generic* and *robust* method to segment experimental two- or three-dimensional images of samples obtained by SEM and X-ray tomography. *Simple* means that this method can be used by anyone whether or not they are a specialist in image processing. *Generic* means that this method remains valid for a wide range of materials. *Robust* means that the extraction is less sensitive to a perturbation of the segmented parameters. For this purpose, we have used the following advanced algorithms: the watershed transformation, the distance function, the dynamic filter, the hole filling.... Due to the large size of the tomography images, each of theses algorithms has to be optimised that the treatment can occur within a reasonable time. The direct implementation of these optimised advanced algorithms is relatively long, tedious and requires specific skills.

To speed up the implementation of the previous algorithms, we have conceptualised a theoretical framework for a class of algorithms, called Seeded Region Growing by Pixels Aggregation/Dissolution (SRGPAD). This work has led to the creation of a library, called Population, dedicated to the implementation of SRGPAD algorithms. The use of this library resulted in the fast implementation of any SRGPAD algorithm. Furthermore any algorithm implemented by using this library is already optimised.

Thirdly, the resulting segmented images were processed to extract metric and/or topological quantitative information concerning porous media. For the metric analysis, we have focused on stereological measurements. Stereology guarantees the link between measurements done in a 2D section of the 3D sample and measurements done on the whole sample for a homogeneous and isotropic material.

In the cement industry, metric analysis is focused on the evaluation of global geometrical properties. Therefore, we have estimated the most commonly available numbers: the volume fractions of the phases[75] and the specific surface area. The volume fraction of the anhydrous phase gives the degree of hydration of the sample. It allows comparative studies between SEM and X-ray tomography. The volume fraction of the pore phase, called porosity, is directly related to the diffusive property which follows one type of Archie's law. However, many important features are not contained in these numbers. Hence, a statistical analysis has been performed with the 2-point probability functions and the chord length distribution functions. The properties of the 2-point probability function have confirmed the isotropy of the material and the characteristic length of the representative elementary volume equal to 200 μm for portland

[75] anhydrous grain, hydrates, portlandite and pore space.

and 500 μm for alite. The properties of the chord length distribution function have confirmed that the disorder of capillary pore is a simple long-range Debye randomness (not a complex structure with length scale invariance. The average of chord length distribution function gives the characteristic length of the given phase. We observed that this characteristic length of anhydrous grains remains constant during the hydration process although anhydrous grains are dissolved.

The above metric characterization is applicable in two and three dimensions. In turn, the topological characterization is processed only on 3D images, in this case the tomography images. The primary topological feature is the percolation of the pore network. Diffusion is only possible on a percolating network. For Portland cement, no percolation of the pore network found for 150 hours of hydration time and for a resolution of $0.675\mu m$ while diffusion still occured. Imaging techniques are limited to a range of length scales. Diffusion still occurs since the pore network still percolates at a smaller length scale than the tomography resolution. Therefore, nano-tomography will play an important role in the observation of the smaller capillary pores that control transport for mature pastes. Even for early-age paste, which is still with relevant scales of tomography, the percolation does not quantify the degree of connectivity of the pore network. To takle this task, we estimated a new topological number introduced by Levitz. As this number is intensive, it allows the topological comparison of the degree of connectivity between different porous media.

Initially, we intended to relate the geometrical characterization and the diffusion properties in two ways. We verified on several sets of data that the information contained in a 2D section of a 3D sample of cement paste allows the reconstruction of a 3D model that accounts for the topological and diffusion properties of the original medium. In order to obtain a reliable diffusion coefficient, the simulation has to be performed on a representative elementary volume (REV). The optimisation, using voxel permutation at the phases boundary, allows the reconstruction of a model which volume is larger than the REV.

There are two area for future research. Firstly, industrial applications mainly focus on the ageing of materials. In this framework, is is crucial to ensure the durability of cements materials. This property is related to diffusive properties which directly depend on the geometrical organisation of the mature cement pastes. As previously discussed, the extraction of the physical tortuosity by imaging techniques is only reliable as long as the pore size remains bigger than the resolution range of the imaging device in order to oberve the connexion of the capillary porosity. From this perspective, the extraction of tortuosity based on synchrotron x-ray tomography is only reliable at the early-age of cements pastes. Taking an image at early-age of this system in evolution is now possible since the acquisition time is inferior than 10 minutes at most of the synchrotron facilities. The image analysis approach should:

- be validated by comparing the experimental and computed diffusion coefficients,
- explore the phenemolical law of the effective diffusive coefficient in relation to the porosity for different chemical composition of cement pastes.

Provided these two points are observed, we could estimate the diffusion coefficient of mature paste in extrapoling the function of diffusion coefficient in relation to porosity of capillary pores (calculated on SEM image).

An important theoretical question is to know whether it is possible to simplify the geometrical description of the 3D pore network using a small number that are representative of transport properties. The decomposition of elementary pores was achieved in this thesis. Further work will attempt to link some first passage statistics inside and between the pores with simple metric features. The first passage statistics will be used to assign transition times and direction probabilities between vertexes of the pore graph. This will provide a coarser graining description of the diffusive transport in an effective network of elementary pores. Ultimately we hope to reveal an analytical law describing the diffusive transport through some geometrical features. Its validity will be proved by the statistical properties of the pore graph.

Finally, my original library covering the algorithms used in this thesis can be used on a wide variety of image analysis applications in material science. For this purpose I intend to make it more user-friendly

by developing a graphical user interface. In order to make it widely available, it will be released as an open source software.

A

Annexes

A.1 restricted set and neighborhood set

The definition of the set ZI requires two sets: restricted set and neighborhood set.

A.1.1 neighborhood set

Numerically, the definition og the neighborhood set uses two classical methods:

- an image, I, of odd size[76] with c the center of the image[77] such as V is defined by:

$$V = \{\forall x \in E : I(x - c) \neq 0\}$$

For example, $I = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ gives $V = \{(-1,-1),(0,0),(1,1)\}$

- a ball in order to define a neighborhood independly of the dimension. For example, $V = B(r = 1, N_1)$, is (2n)-neighborhood in n-dimensional discrete space (4-neightborhood in 2D) where $B(r = 1, N_1)$ is the closed ball of radius 1 and N_1 is the norm 1 centered at point 0.

A.1.2 restricted set

Numerically to define the restricted set, the method uses two fields:

- a_w (all or without), a boolean
- L, a list of integer

such as

$$N = \begin{cases} \mathbb{N}_n \setminus L \text{ if } a\_w = true \\ L \text{ otherwise} \end{cases}$$

. This is a mathematical definition of ZI. In the next subsection, we will define a numerical process to actualize the ZI depending on aggregation/dissolution on a region.

A.2 Proof of the acualization

A.2.1 Growth of the myself region

If there is only the growth of the myself region: $X_{i,m}^{t+1/2} = X_{i,m}^t + A^t$ then

$$Z_i^{t+1/2} = Z_i^t + (A^t \oplus V_i) \setminus Z_i^t \setminus (\bigcup_{j \in N_i} X_{j,o}^t) \tag{A.1}$$

Proof:

$$Z_i^{t+1/2} = (X_{i,m}^{t+1/2} \oplus V_i) \setminus (\bigcup_{j \in N_i} X_{j,o}^{t+1/2})$$

As we have only the growth of the myself region: $X_{i,m}^{t+1/2} = X_{i,m}^t + A^t$ then

$$Z_i^{t+1/2} = ((X_{i,m}^t + A^t) \oplus V_i) \setminus (\bigcup_{j \in N_i} X_{j,o}^t)$$

[76] Let $S = (s_0, \ldots, s_n)$ be the size of the image. The size of the image is impair if for all j in $(0, \ldots, n)$, s_j is impair.
[77] Let $S = (s_0, \ldots, s_n)$ be the size of the image. The center of the image is $(s_0\%2 + 1, \ldots, s_n\%2 + 1)$

As $(A \cup B) \oplus V = (A \oplus V) \cup (B \oplus V)$ so with $R_i^t = \bigcup_{j \in N_i} X_{j,o}^t$

$$Z_i^{t+1/2} = ((X_{i,m}^t \oplus V_i) \cup (A^t \oplus V_i)) \setminus R_i^t$$

We know that $A \cup B = A + (B \setminus A)$:

$$Z_i^{t+1/2} = ((X_{i,m}^t \oplus V_i) + (A^t \oplus V_i \setminus (X_{i,m}^t \oplus V_i))) \setminus R_i^t$$

and $(A + B) \setminus C = (A \setminus C) + (B \setminus C)$, then

$$Z_i^{t+1/2} = (X_{i,m}^t \oplus V_i) \setminus R_i^t + (A^t \oplus V_i \setminus (X_{i,m}^t \oplus V_i)) \setminus R_i^t$$

We have $Z_i^t = (X_{i,m}^t \oplus V_i) \setminus R_i^t$

$$Z_i^{t+1/2} = Z_i^t + (A^t \oplus V_i) \setminus (X_{i,m}^t \oplus V_i) \setminus R_i^t$$

As $(A \setminus B) \setminus C = (A \setminus (B \setminus C)) \setminus C$ then with $Z_i^t = (X_{i,m}^t \oplus V_i) \setminus R_i^t$

$$Z_i^{t+1/2} = Z_i^{t+1/2} + (A^t \oplus V_i) \setminus Z_i^t \setminus R_i$$

A.2.2 Degrowth of the myself region

If there is only the degrowth of the myself region: $X_{i,m}^{t+1/2} = X_{i,m}^t + A^t$ then

$$Z_i^{t+1/2} = Z_i^t - ((A^t \oplus V_i) \setminus (X_{i,m}^{t+1/2} \oplus V_i)) \setminus (Z_i^t)^c \tag{A.2}$$

Lemme:

$$A \oplus V = ((A \cup B) \oplus V) \setminus ((B \oplus V) \setminus (A \oplus V))$$

First, demonstrate that $A = (A \cup B) \setminus (B \setminus A)$

$$\begin{aligned} A &= A \cup \emptyset \\ A &= A \cup (B \cap B^c) \\ A &= (A \cup B) \cap (A \cup B^c) \\ A &= (A \cup B) \cap (B^c \cup A) \\ A &= (A \cup B) \setminus (B^c \cup A)^c \\ A &= (A \cup B) \setminus (B \cap A^c) \\ A &= (A \cup B) \setminus (B \setminus A) \end{aligned}$$

Switching A by $A \oplus V$ and B by $B \oplus V$ in this last formul, we get:

$$A \oplus V = ((A \oplus V) \cup (B \oplus V)) \setminus ((B \oplus V) \setminus (A \oplus V))$$

As $(A \oplus V) \cup (B \oplus V) = (A \cup B) \oplus V$, we find

$$A \oplus V = ((A \cup B) \oplus V) \setminus ((B \oplus V) \setminus (A \oplus V))$$

Proof:

$$Z_i^{t+1/2} = (X_{i,m}^{t+1/2} \oplus V_i) \setminus (\bigcup_{j \in N_i} X_{j,o}^{t+1/2})$$

As we have only the degrowth of the myself region, $X_{i,m}^{t+1/2} = X_{i,m}^t - A^t$ with $R_i^t = \bigcup_{j \in N_i} X_{j,o}^t$ then

$$Z_i^{t+1/2} = ((X_{i,m}^t - A^t) \oplus V_i) \setminus R_i^t$$

By the lemme, switching A by $(X_{i,m}^t - A^t)$ and B by A^t, we have

$$\begin{aligned} Z_i^{t+1/2} &= (A_1 \setminus A_2) \setminus R_i^t \\ \text{with} \quad & A_1 = ((X_{i,m}^t - A^t) \cup A^t) \oplus V_i \\ \text{with} \quad & A_2 = (A^t \oplus V_i) \setminus ((X_{i,m}^t - A^t) \oplus V_i) \end{aligned}$$

Like $((X_{i,m}^t - A^t) \cup A^t = X_{i,m}^t$ so we obtain $A_1 = X_{i,m}^t \oplus V_i$.
As $X_{i,m}^t - A^t = X_{i,m}^{t+1/2}$ so $A_2 = (A^t \oplus V_i) \setminus (X_{i,m}^{t+1} \oplus V_i)$. We have

$$\begin{aligned} Z_i^{t+1/2} &= (A_1 \setminus A_2) \setminus R_i^t \\ \text{with} \quad & A_1 = X_{i,m}^t \oplus V_i \\ \text{with} \quad & A_2 = (A^t \oplus V_i) \setminus (X_{i,m}^{t+1/2} \oplus V_i) \end{aligned}$$

As $(A \setminus B) \setminus C = (A \setminus C) \setminus B$, substituting A by A_1, B by A_2 and C by R_i^t in this last formula. As $A_1 \setminus R_i^t = Z_i^t$, we have

$$\begin{aligned} Z_i^{t+1/2} = & \quad Z_i^t \setminus A_2 \\ \text{with} \quad & A_2 = (A^t \oplus V_i) \setminus (X_{i,m}^{t+1/2} \oplus V_i) \end{aligned}$$

As $A \setminus B = A - (B \setminus A^c)$, we have:

$$Z_i^{t+1/2} = Z_i^t - ((A^t \oplus V_i) \setminus (X_{i,m}^{t+1/2} \oplus V_i)) \setminus (Z_i^t)^c$$

A.2.3 Growth of the other region

If there is only the growth of the other region: $X_{i,o}^{t+1} = X_{i,o}^{t+1/2} + A^t$ then

$$\forall j \in \mathbb{N}_n \begin{cases} Z_j^{t+1} = Z_j^{t+1/2} & \\ -A^t \setminus (Z_j^{t+1/2})^c & \text{if } i \in N_j \wedge V_j \neq \emptyset \\ Z_j^{t+1} = Z_j^{t+1/2} & \text{else} \end{cases} \tag{A.3}$$

Proof:
Let $V_j = \emptyset$, then $\forall t : Z_j^t = \emptyset$, in particulary $Z_j^{t+1} = Z_j^{t+1/2}$. Let assume now that $V_j = V$.

$$Z_j^{t+1} = (X_{j,m}^{t+1} \oplus V) \setminus \cup_{k \in N_j} X_{k,o}^{t+1}$$

By the commutativity of union, we have:

$$\begin{aligned} Z_j^{t+1} &= (X_{j,m}^{t+1} \oplus V) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1} \cup X_k^{t+1}) && \text{if } i \in N_j \\ Z_j^{t+1} &= Z_j^t && \text{else} \end{aligned}$$

We suppose $i \in N_j$. We have only the growth of the other region: $X_{i,o}^{t+1} = X_{i,o}^{t+1/2} + A^t$

$$Z_j^{t+1} = (X_{j,m}^{t+1/2} \oplus V) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2} \cup (X_k^{t+1/2} + A^t))$$

$A + B = A \cup B$ and by the commutativity of union, we have:

$$Z_j^{t+1} = (X_{j,m}^{t+1/2} \oplus V) \setminus (\cup_{k \in N_j} X_{k,o}^{t+1/2} \cup A^t)$$

As $A \setminus (B \cup C) = A \setminus B \setminus C$ and $Z_j^{t+1/2} = (X_{j,m}^{t+1/2} \oplus V) \setminus (\cup_{k \in N_j} X_{k,o}^{t+1/2})$, we have

$$Z_j^{t+1} = Z_j^{t+1/2} \setminus A^t$$

As $A \setminus B = A - B \setminus A^c$, thus

$$Z_j^{t+1} = Z_j^{t+1/2} - A^t \setminus (Z_j^{t+1/2})^c$$

A.2.4 Degrowth of the other region

If there is only the growth of the other region:$X_{i,o}^{t+1} = X_{i,o}^{t+1/2} - A^t$ then

$$\forall j \in \mathbb{N}_n \begin{cases} Z_j^{t+1} = Z_j^{t+1/2} + \\ (((A^t \setminus (X_{m,j}^{t+1/2} \oplus V_j)^c) \setminus (\bigcup_{k \in N_j} X_{k,o}^{t+1})) \setminus Z_j^{t+1/2}) \\ \qquad \text{if } (i \in N_j) \wedge (V_j \neq \emptyset) \\ Z_j^{t+1} = Z_j^{t+1/2} \text{ else} \end{cases}$$

Proof:
Let $V_j = \emptyset$, then $\forall t : Z_j^t = \emptyset$, in particulary $Z_j^{t+1} = Z_j^{t+1/2}$. Let assume now that $V_j = V$. Let j in $\mathbb{N}_n$

$$Z_j^{t+1} = (X_{j,m}^{t+1} \oplus V) \setminus \cup_{k \in N_j} X_{k,o}^{t+1}$$

By the commutativity of union, we have:

$$\begin{aligned} Z_j^{t+1} &= (X_{j,m}^{t+1} \oplus V) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1} \cup X_k^{t+1}) && \text{si } i \in N_j \\ Z_j^{t+1} &= Z_j^{t+1/2} && \text{if} i \notin N_j \end{aligned}$$

We suppose $i \in N_j$. We have only the growth of the other region: $X_{i,o}^{t+1} = X_{i,o}^{t+1/2} - A^t$

$$Z_j^{t+1} = (X_{j,m}^{t+1/2} \oplus V) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2} \cup (X_{i,o}^{t+1/2} - A^t))$$

As $A - B = A \cap B^c$, we have

$$Z_j^{t+1} = (X_{j,m}^{t+1/2} \oplus V) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2} \cup (X_{i,o}^{t+1/2} \cap (A^t)^c))$$

As $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, thus

$$\begin{aligned} Z_j^{t+1} = & (X_{j,m}^{t+1/2} \oplus V) \\ & \setminus (\cup_{k \in N_j} X_{k,o}^{t+1/2} \cap (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2} \cup (A^t)^c)) \end{aligned}$$

As $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$, thus:

$$\begin{aligned} Z_j^{t+1} = & Z_j^{t+1/2} \cup A_1 \\ \text{with} \quad & A_1 = (X_{j,m}^{t+1/2} \oplus V) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2} \cup (A^t)^c) \end{aligned}$$

As $A \setminus (B \cup C) = (B^c \cap A) \setminus C$, we have

$$\begin{aligned} A_1 &= (A^t \cap (X_{j,m}^{t+1/2} \oplus V)) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2}) \\ A_1 &= (A^t \setminus (X_{j,m}^{t+1/2} \oplus V)^c) \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2}) \\ A_1 &= (A^t \setminus \cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2}) \setminus ((X_{j,m}^{t+1/2} \oplus V)^c) \end{aligned}$$

As $X_{i,o}^{t+1} = X_{i,o}^{t+1/2} - A^t$, so $A^t \cap X_{i,o}^{t+1/2} = \emptyset$ and $A^t = A^t \setminus X_{i,o}^{t+1}$. As $(A \setminus B) \setminus C = A \setminus (B \cup C)$, we have

$$\begin{aligned} A_1 = & \quad (A^t \setminus (\cup_{k \in N_j \setminus i} X_{k,o}^{t+1/2} \cup X_{i,o}^{t+1})) \setminus ((X_{j,m}^{t+1/2} \oplus V)^c) \\ A_1 = & \quad (A^t \setminus (\cup_{k \in N_j} X_{k,o}^{t+1}) \setminus ((X_{j,m}^{t+1/2} \oplus V)^c) \\ A_1 = & \quad (A^t \setminus (X_{j,m}^{t+1/2} \oplus V)^c \setminus (\cup_{k \in N_j} X_{k,o}^{t+1}) \end{aligned}$$

The last step is: $Z_j^{t+1/2} \cup A_1 = Z_j^{t+1/2} + A_1 \setminus Z_j^{t+1/2}$.

A.3 Watershed transformation using meta-programmation approach in modern language

The real implementation of the algorithm in my library Population:

```
template<typename Space>
class Affect_watershed: public Affect<typename Space::Pixel >
{
private:
  Space * _img;
  int * _level;
public:
  Affect_watershed(Space & img, int & level)
      :_img(&img),_level(&level)
    {}
//Number of queue at the initialization (for watershed, it is I.max_range())
  virtual int card()
    {
      return numeric_limits<typename Space::Type>::max()+1 ;
    }
//Number of queue to add when a tribe is created (for watershed, it is 0)
  virtual int new_tribe()
    {
      return 0;
    }
//ordering attribute function, \delta (x, i)= max(I(x),level)
  virtual int  affect(const  typename Space::Pixel & x , int  )
    {
      return max(static_cast<int>((* _img)[x]),*_level) ;
    }
};
template<typename Space_topo,typename Space_label,  typename Neight>
Pop::Space_region watershed_withoutborder(const Space_topo & in,const  Space_label & label,const  Neigh
{
//initialization
      int level=0;
      Affect_watershed< Space_topo > f(in,level);//Creation of the object SQ
      Zi zi(in.size(),f);//Creation of the object ZI
      Pop pop(in.size(),zi);//Creation of the object Population
      pop.neighborhood(V);//Definition of the neighborhood
      restricted r; //Definition of the resttricted set
      r.with_without=false;
      r.my_self    =false;
      Tribe tr(r);
```

```
      label_to_tribe(label ,  pop, tr);//Initialization of the region using the seeds
      typename Zi::maillon_la_Z m;
//growing process
      for(;level< numeric_limits<=typename T::Type>::max();level++)
      {
        zi.select_queue(level);
        while(zi.empty()==false)
        {
          m= zi.pop();//m=(x,i)
          pop.groth(m.first,m.second);
        }
      }
      return pop.X();
}
```

A.4 Application for SEM images of cement paste

In the figure A.1 and A.2, we have applied the one step segmentation for:

1. OPC with W/C=0.5, alite with W/C=0.4,
2. at hydration time: 1, 3, 7 days;
3. at the resolution: 0.675, 0.25, 0.1μm

A.5 Algorithms

A.5.1 Specific surface area

The algorithm is:

Require: $\mathcal{I}$//*The label image*
  // **Count the number of surface for each phase**
  Neightborhood $V = B_{r=1,N_1}(\cdot) \setminus \{\cdot\}$;
  Vector v_count($\mathcal{I}$.nbr_label());
  for all $\forall x \in E$ **do**
    for all $\forall Y \in V(x)$ **do**
      if $\mathcal{I}[x] \neq \mathcal{I}[y]$ **then**
        v_count$[\mathcal{I}[x]]$++;
      end if
    end for
  end for
  // **Normalized by the Lebesgue measure of the image domain (pixels/voxels number)**
  for i = 0 to v_count.size() **do**
    v_count $[\mathcal{I}[x]]$= v_count $[\mathcal{I}[x]]/\mu(E)$;
  end for
  return v_count;

A.5.2 2-point probability functio using Matlab

```
function [] = mafonction(name)
I = imread(name);
F = fft2(double(I));
```

```
F_s=fftshift(F);
F3 = ifft2(abs(F_s));
F2 = log(abs(F3));
F2=fftshift(F2);
S = imshow(F2,[8 15],'InitialMagnification','fit'); colormap(jet); colorbar
imwrite (S, [name '_fourier.png']);
```

A.5.3 Graph of 2-point probability function

There are two possibilities to calculate the graph of the 2-point probability function. Both are built around this formula:

$$S_{2,i}(\vec{r}) \quad = \quad < I_i(\vec{r'})I_i(\vec{r'}+\vec{r}) >_{\vec{r'}\in E}$$

Assuming the material is isotrope, we calculate:

$$S_{2,i}(d) \quad = \quad << I_i(\vec{r'})I_i(\vec{r'}+\vec{r}) >_{\vec{r'}\in E} >_{\|\vec{r}\|=d}$$

The average $<>_{\|\vec{r}\|=d}$ is done only on the orthogonal directions. Numerically, we can estimate the 2-point probability function on the orthogonal directions by calculated all the correlations (see algorithm 9) or by repeated random sampling (method of Monte-Carlo ,see algorithm 10)

Algorithm 9 2-point probability function using the d-mains direction

Require: $\mathcal{I}$ (The label image), d_{max} (max distance)
matrix m_hit(d_{max}+1,$\mathcal{I}$.*nbr_phase*); *matrix of size* $(n,p) = (d_{max}+1, \mathcal{I}.nbr\_phase)$
matrix m_count(d_{max}+1,$\mathcal{I}$.*nbr_phase*);*matrix of size* $(n,p) = (d_{max}+1, \mathcal{I}.nbr\_phase)$
. $\forall x \in E$ **do**
. $\forall d \in (0,\ldots,d_{max})$ **do**
. $\forall e_i \in 1$ **do**
$y = x + e_i d$
m_count[$\mathcal{I}[x]$][d]++;
if $\mathcal{I}[x] == \mathcal{I}[y]$ **then**
m_hit[$\mathcal{I}[x]$][d]++;
end if
end for
end for
end for
matrix m_corr(d_{max}+1,$\mathcal{I}$.*nbr_phase*); *matrix of 2-point probability function of each phase*
for i = 0 to $\mathcal{I}$.*nbr_phase* **do**
for j = 0 to d_{max} **do**
m_corr[i][j] = m_hit[i][j] /m_count[i][j] ;
end for
end for
return m_corr;

A.6 Signature of a polydispersion of grains following a power law

In the section, we prove that a polydispersion of grains following a power law gives the same power law for the chord length distrubution function of the granular phase.
Let M be a granular phase, a collection of grains, with $f_M(r)$, its chord length distribution function such as all the grains have the same characteristic size equal to 1.
Let λM the granular phase after the homothetic transformation with parameter λ. Its chord chord length distribution function is $f_{\lambda M} = f_M(r/\lambda)/\lambda$.

Algorithm 10 2-point probability function using the d-mains direction and a Monte-Carlo algorithm

Require: $\mathcal{I}$ (The label image), d_{max} (max distance), n_{max} (number of tests)
matrix m_hit(d_{max}+1,$\mathcal{I}.nbr\_phase$); *matrix of size* $(n,p) = (d_{max}+1, \mathcal{I}.nbr\_phase)$
matrix m_count(d_{max}+1,$\mathcal{I}.nbr\_phase$);*matrix of size* $(n,p) = (d_{max}+1, \mathcal{I}.nbr\_phase)$
for i = 0 to n_{max} **do**
x= $\mathcal{I}.rand()$
. $\forall d \in (0, \ldots, d_{max})$ **do**
. $\forall e_i \in 1$ **do**
$y = x + e_i d$
m_count[$\mathcal{I}[x]$][d]++;
if $\mathcal{I}[x] == \mathcal{I}[y]$ **then**
m_hit[$\mathcal{I}[x]$][d]++;
end if
end for
end for
end for
matrix m_corr(d_{max}+1,$\mathcal{I}.nbr\_phase$); *matrix of 2-point probability function of each phase*
for i = 0 to $\mathcal{I}.nbr\_phase$ **do**
for j = 0 to d_{max} **do**
m_corr[i][j] = m_hit[i][j] /m_count[i][j] ;
end for
end for
return m_corr;

To take into account the polydispersion of the granular phase, we define a probability function, $P(\lambda)d\lambda$, given the probability to have a grain of size betweeen λ and $\lambda + d\lambda$.
The chord length distribution function is:

$$f_{P,M}(r) = \int_0^\infty f_{\lambda M}(r) P(\lambda) d\lambda$$

Given that $f_{\lambda M} = f_M(r/\lambda)/\lambda$, we get:

$$f_{P,M}(r) = \int_0^\infty \frac{f_M(r/\lambda)}{\lambda} P(\lambda) d\lambda$$

Let us consider that the probability function follows a power law distribution: $1/\lambda^n$.

$$f_{P,M}(r) = \int_0^\infty \frac{f_M(r/\lambda)}{\lambda^{n+1}} d\lambda$$

Let $r/\lambda = u$ be the change of variable. We get:

$$f_{P,M}(r) = \int_0^\infty \frac{f_M(u) u^{n+1}}{r^{n+1}} r \frac{du}{u^2}$$

$$f_{P,M}(r) = \frac{1}{r^n} \int_0^\infty f_M(u) u^{n-1} du$$

$$f_{P,M}(r) \propto \frac{1}{r^n}$$

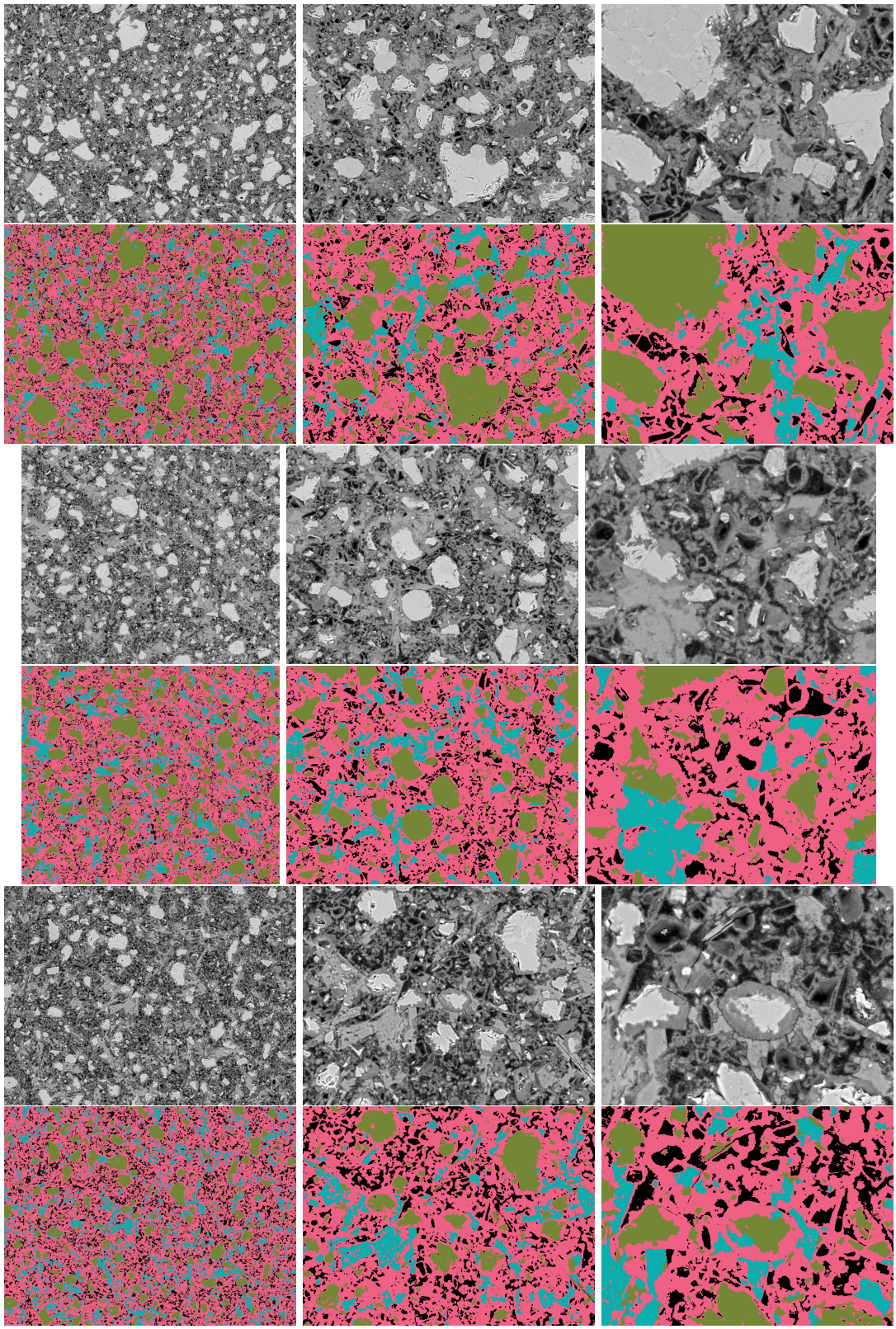


Figure A.1: OPC. For each raw, the resolution is 0.675 μm image 1, 0.25 μm image 2, 0.1 μm image 3. First raw=1 day, second raw= 3 days and last raw = 7 days. Visually, we observe that there is more porosity at 7 days than at 3 days although normally the hydrate takes the place of the porosity during the hydration. A possible reason of this anomaly is that the cement paste is a heterogeneous media.

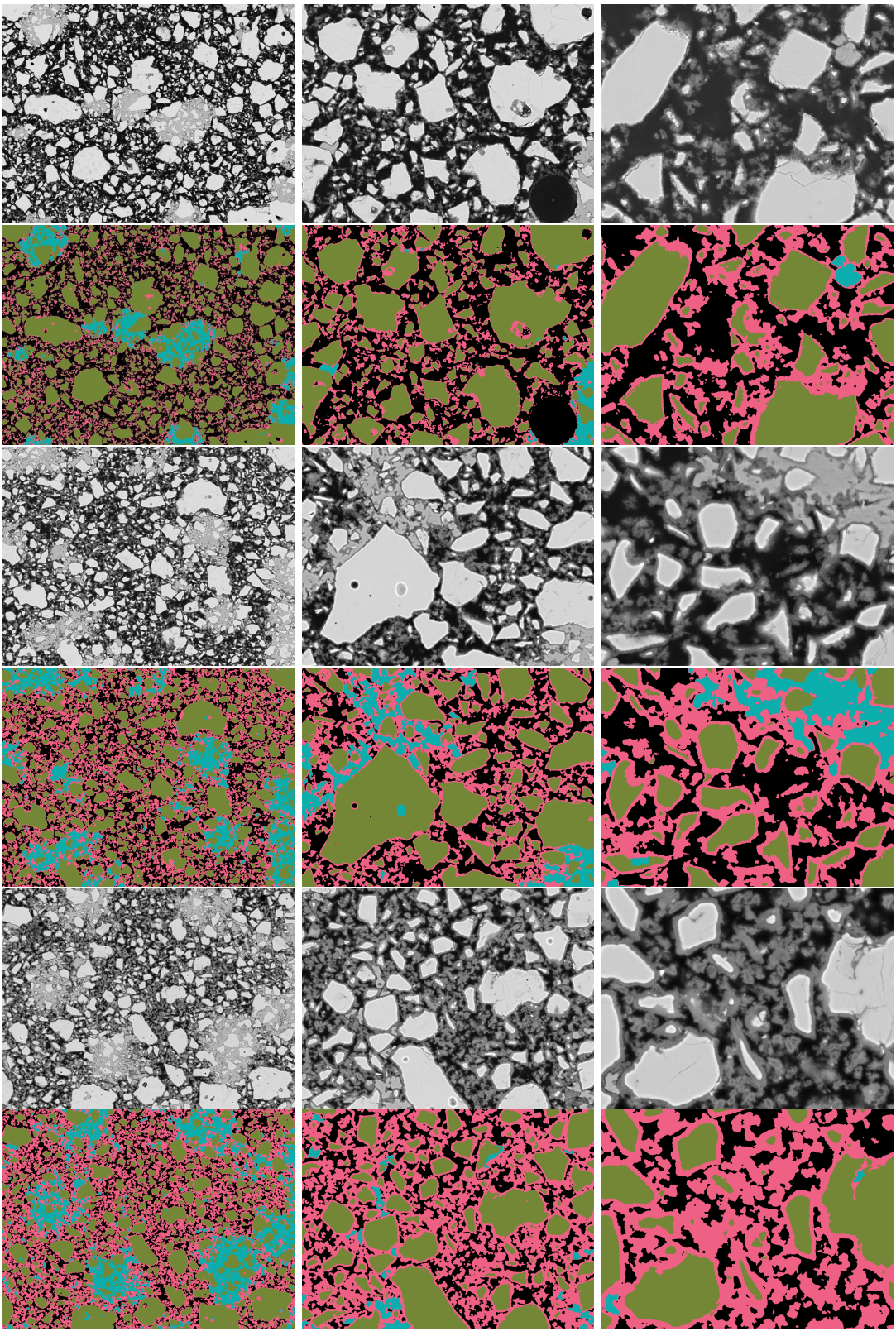


Figure A.2: alite. For each raw, the resolution is 0.675 μm image 1, 0.25 μm image 2, 0.1 μm image 3. First raw=1 day, second raw= 3 days and last sraw = 7 days.

Bibliography

[1] A. B. Abell, K. L. Willis, and D. A. Lange. Mercury intrusion porosimetry and image analysis of cement-based materials. *Journal Of Colloid And Interface Science*, 211(1):39–44, March 1999.

[2] R. Adams and L. Bisschof. Seeded region growing. *Ieee Transactions On Pattern Analysis And Machine Intelligence*, 16(6):641–647, June 1994.

[3] PM. Adler, CG Jacquin, and JF. Thovert. The formation factor of reconstructed porous media. *Water Resour. Res.*, 28(6):1571–1576, 1992.

[4] G Archie. The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of AIME,*, 146:54–63, 1942.

[5] C. H. Arns, M. A. Knackstedt, and K. R. Mecke. Reconstructing complex materials via effective grain shapes. *Physical Review Letters*, 91(21):215506, November 2003.

[6] C. H. Arns, M. A. Knackstedt, W. V. Pinczewski, and W. B. Lindquist. Accurate estimation of transport properties from microtomographic images. *Geophysical Research Letters*, 28(17):3361–3364, September 2001.

[7] C. H. Arns, M. A. Knackstedt, W. V. Pinczewski, and K. R. Mecke. Euler-poincare characteristics of classes of disordered media. *Physical Review E*, 6303(3):031112, March 2001.

[8] F. M. Auzerais, J. Dunsmuir, B. B. Ferreol, N. Martys, J. Olson, T. S. Ramakrishnan, D. H. Rothman, and L. M.b Schwartz. Transport in sandstone: A study based on three dimensional microtomography. *Geophysical Research Letters*, 23(7):705–708, April 1996.

[9] D. F. Ball. Loss-on-ignition as estimate of organic matter + organic carbon in non-calcareous soils. *Journal Of Soil Science*, 15(1):84–&, 1964.

[10] D.H. Ballard and C. Brown. *Computer Vision*. Berlin, Germany: Springer Verlag, 1982.

[11] H. G. Ballesteros, L. A. Fernandez, V. Martin-Mayor, A. M. Sudupe, G. Parisi, and J. J. Ruiz-Lorenzo. Scaling corrections: site percolation and ising model in three dimensions. *Journal Of Physics A-Mathematical And General*, 32(1):1–13, January 1999.

[12] F. Barberon, J. P. Korb, D. Petit, V. Morin, and E. Bermejo. Probing the surface area of a cement-based material by nuclear magnetic relaxation dispersion. *Physical Review Letters*, 90(11):116103, March 2003.

[13] V. Baroghel-Bouny, M. Mainguy, T. Lassabatere, and O. Coussy. Characterization and identification of equilibrium and transfer moisture properties for ordinary and high-performance cementitious materials. *Cement And Concrete Research*, 29(8):1225–1238, August 1999.

[14] L. K. Barrett and C. S. Yust. Some fundamental ideas in topology and their application to problems in metallurgy. *Scripta Metallurgica*, 3(7):461–&, 1969.

[15] S. Béjaoui, B. Bary, S. Nitsche, D. Chaudanson, and C. Blanc. Experimental and modeling studies of the link between microstructure and effective diffusivity of cement pastes. *Revue européenne de génie civil*, 10, n°9:1073–1106, 2006.

[16] D. P. Bentz, D. A. Quenard, H. M. Kunzel, J. Baruchel, F. Peyrin, N. S. Martys, and E. J. Garboczi. Microstructure and transport properties of porous building materials. ii: Three-dimensional x-ray tomographic studies. *Materials And Structures*, 33(227):147–153, April 2000.

[17] A. L. Berdichevsky and Z. Cai. Preform permeability predictions by self-consistent method and finite-element simulation. *Polymer Composites*, 14(2):132–143, April 1993.

[18] N. F. Berk. Scattering properties of the leveled-wave model of random morphologies. *Physical Review A*, 44(8):5069–5079, October 1991.

[19] G. Bertrand and M. Couprie. A model for digital topology. *Discrete Geometry For Computer Imagery*, 1568:229–241, 1999.

[20] S. Beucher. The watershed transformation applied to image segmentation. *Conference on Signal and Image Processing in Microscopy and Microanalysis*, pages 299–314, 1991.

[21] S. Beucher. Geodesic reconstruction, saddle zones & hierarchical segmentation. *Image Analysis Stereology*, 20:137–141, 2001.

[22] S. Beucher. Algorithmes sans biais de ligne de partage des eaux. *Note interne CMM*, 2004.

[23] S. Beucher and C. Lantuejoul. Use of watersheds in contour detection. In *real-time edge and motion detection*. International workshop on image processing, 1979.

[24] S. Beucher and F. Meyer. *the morphological approach to segmentation: the watershed transformation*, chapter Mathematical Morphology in Image Processing. Marcel Dekker, New-York, 1993.

[25] S. D. Beyea, B. J. Balcom, T. W. Bremner, P. J. Prado, D. P. Green, R. L. Armstrong, and P. E. Grattan-Bellew. Magnetic resonance imaging and moisture content profiles of drying concrete. *Cement And Concrete Research*, 28(3):453–463, March 1998.

[26] C.M. Bishop. *Pattern Recognition And Machine Learning*. 2008.

[27] S. Brunauer, P.H. Emmet, and E Teller. Adsorption of gases in multi-molecular layes. *J. Am. Chem. Soc.*, 60:309 – 319, 1938.

[28] J. Buffiere, J.-Y. Maire, and E. Merle. *X-Ray Tomography in Material Science.*. HERMES Science Publications ,Paris, 2000.

[29] N. Burlion, D. Bernard, and D. Chen. X-ray microtomography: Application to microstructure analysis of a cementitious material during leaching process. *Cement And Concrete Research*, 36(2):346–357, February 2006.

[30] J. W. Cahn and J. E. Hilliard. Free energy of a nonuniform system .1. interfacial free energy. *Journal Of Chemical Physics*, 28(2):258–267, 1958.

[31] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:679–714, 1986.

[32] B. Chopard and M. Droz. Cellular automata approach to non-equilibrium phase-transitions in a surface-reaction model - static and dynamic properties. *Journal Of Physics A-Mathematical And General*, 21(1):205–211, January 1988.

[33] S. Ciccariello, G. Cocco, A. Benedetti, and S. Enzo. Correlation-functions of amorphous multiphase systems. *Physical Review B*, 23(12):6474–6485, 1981.

[34] D. Cule and S. Torquato. Generating random media from limited microstructural information via stochastic optimization. *Journal Of Applied Physics*, 86(6):3428–3437, September 1999.

[35] M. Daimon and D. M. Roy. Rheological properties of cement mixes .1. methods, preliminary experiments, and adsorption studies. *Cement And Concrete Research*, 8(6):753–764, 1978.

[36] L. S. Davis, A. Rosenfeld, and J. S. Weszka. Region extraction by averaging and thresholding. *Ieee Transactions On Systems Man And Cybernetics*, SMC5(3):383–388, 1975.

[37] E. De Dale, E. Newbury, Patrick, and David C. J. *Advanced scanning electron microscopy and X-ray microanalysis.* springer, 1986.

[38] A. G. De La Torre, S. Bruque, J. Campo, and M. A. G. Aranda. The superstructure of c3s from synchrotron and neutron powder diffraction and its role in quantitative phase analyses. *Cement And Concrete Research*, 32(9):1347–1356, September 2002.

[39] P. Debye, H. R. Anderson, and H. Brumberger. Scattering by an inhomogeneous solid .2. the correlation function and its application. *Journal Of Applied Physics*, 28(6):679–683, 1957.

[40] M.A. Delesse. Procédé mécanique pour déterminer la composition des roches. *CR Acad. Sci.*, 1847.

[41] R. Deriche. Using canny criteria to derive a recursively implemented optimal edge detector. *International Journal Of Computer Vision*, 1(2):167–187, 1987.

[42] D. Deserable, P. Dupont, M. Hellou, and S. Kamali-Bernard. Cellular automata models for complex matter, 2007.

[43] L. Dormieux, D. Kondo, and F. J. Ulm. *microporo mechanics.* wiley, 2006.

[44] D. A. Doughty and L. Tomutsa. Multinuclear nmr microscopy of two-phase fluid systems in porous rock. *Magnetic Resonance Imaging*, 14(7-8):869–873, 1996.

[45] La durabilité des bétons. Presses de l'école nationale des ponts er chaussées, 2008.

[46] J.P. Eberhart. *Analyse structurale et chimique des matériaux.* DUNOD, 1997.

[47] R. Enficiaud. *Multi-dimensional and multi-spectral algorithms in the field of Mathematical Morphology : The meta-programming approach.* PhD thesis, Mines de Paris, 2008.

[48] R. M. Espinosa and L. Franke. Inkbottle pore-method: Prediction of hygroscopic water content in hardened cement paste at variable climatic conditions. *Cement And Concrete Research*, 36(10):1954–1968, October 2006.

[49] J. P. Fan, D. K. Y. Yau, A. K. Elmagarmid, and W. G. Aref. Automatic image segmentation by integrating color-edge extraction and seeded region growing. *Ieee Transactions On Image Processing*, 10(10):1454–1466, October 2001.

[50] R.F. Feldman, G. W. Chan, R. J. Brousseau, and Tumidajski P. J. Investigation of the rapid chloride permeability test. *Materials Journal*, 91:246–255, 1994.

[51] X. Feng, E. J. Garboczi, D. P. Bentz, P. E. Stutzman, and T. O. Mason. Estimation of the degree of hydration of blended cement pastes by a scanning electron microscope point-counting procedure. *Cement And Concrete Research*, 34(10):1787–1793, October 2004.

[52] B. P. Flannery, H. W. Deckman, W. G. Roberge, and K. L. Damico. 3-dimensional x-ray microtomography. *Science*, 237(4821):1439–1444, September 1987.

[53] A. S. Frangakis and R. Hegerl. Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion. *Journal Of Structural Biology*, 135(3):239–250, September 2001.

[54] J. T. Fredrich. 3d imaging of porous media using laser scanning confocal microscopy with application to microscale transport processes. *Physics And Chemistry Of The Earth Part A-Solid Earth And Geodesy*, 24(7):551–561, 1999.

[55] E. Gallucci, K. Scrivener, A. Groso, M. Stamponani, and Margaritondo G. 3d experimental investigation of the microstructure of cement pastes using synchrotron x-ray microtomography. *Cement and Concrete Research*, 2006.

[56] E. J. Garboczi. Three-dimensional mathematical analysis of particle shape using x-ray tomography and spherical harmonics: Application to aggregates used in concrete. *Cement And Concrete Research*, 32(10):1621–1638, October 2002.

[57] E. N. Gilbert. Random subdivisions of space into crystals. *Annals Of Mathematical Statistics*, 33(3):958–&, 1962.

[58] S. A. Grant, G. E. Boitnott, C. J. Korhonen, and R. S. Sletten. Effect of temperature on hydration kinetics and polymerization of tricalcium silicate in stirred suspensions of cao-saturated solutions. *Cement And Concrete Research*, 36(4):671–677, April 2006.

[59] M. Grimaud. A new measure of constrast: dynamics. In *Proc. SPIE Vol. 1769, pp. 292-305, Image Algebra and Morphological Processing III*, 1992.

[60] A. Guiner and G. Fournet. *Small Angle Scattering of X-Rays*. John Wiley & Sons Inc, 1955.

[61] C. Haecker, E. J. Garboczi, J. W. Bullard, R. B. Bohn, Z. Sun, S. P. Shah, and T. Voigt. Modeling the linear elastic properties of portland cement paste. *Cement And Concrete Research*, 35(10):1948–1960, October 2005.

[62] S.A. Hall, N. Lenoir, G. Viggiani, J. Desrues, P. BÃ©suelle, and Di Michiel M. Evolution of strain localisation in granular media: observation through in-situ xray micro-tomography and 3d-volumetric digital image correlation. *in preparation*, 2008.

[63] M. Han, V. Tariel, S. Youssef, E. Rosenberg, M. Fleury1, and Levitz P. The effect of the porous structure on resistivity index curves. an experimental and numerical study, 2008.

[64] T.C. Hansen. Physical structure of hardened cement paste. a classical approach. *Materials and Structures*, 19:423–436, 1986.

[65] R. D. Hazlett. Statistical characterization and stochastic modeling of pore networks in relation to fluid flow. *Mathematical Geology*, 29(6):801–822, August 1997.

[66] M. K. Head, H. S. Wong, and N. R. Buenfeld. Characterisation of 'hadley' grains by confocal microscopy, August 2006.

[67] K.F.J. Heinrich. *Electron beam x-ray microanalysis*. Van Nostrand Reinhold Co., 1981.

[68] W. Helfrich. *Z. Naturforsch.*, 28:693–703, 1973.

[69] S. A. Hojjatoleslami and J. Kittler. Region growing: A new approach. *Ieee Transactions On Image Processing*, 7(7):1079–1084, July 1998.

[70] L. Holzer, P. Gasser, and B. Muench. Quantification of capillary pores and hadley grains in cement paste using fib-nanotomography, 2006.

[71] L. Holzer, F. Indutnyi, P. H. Gasser, B. Munch, and M. Wegmann. Three-dimensional analysis of porous batio3 ceramics using fib nanotomography. *Journal Of Microscopy-Oxford*, 216:84–95, October 2004.

[72] J. Hoshen and R. Kopelman. Percolation and cluster distribution .1. cluster multiple labeling technique and critical concentration algorithm. *Physical Review B*, 14(8):3438–3445, 1976.

[73] A. Hrennikoff. Solution of problems of elasticity by the fram–work method. *J. Appl/ Mech*, 8:619 715, 1941.

[74] J. Hu and P. Stroeven. X-ray absorption study of drying cement paste and mortar. *Cement And Concrete Research*, 33(3):397–403, March 2003.

[75] R. H. Huesman, G. T. Gullberg, W. L. Greenberg, and Budinger. donner algorithms for reconstruction tomography. *Laurence Berckeley Laboratory*, 1977.

[76] S. T. Hyde, I. S. Barnes, and B. W. Ninham. Curvature energy of surfactant interfaces confined to the plaquettes of a cubic lattice. *Langmuir*, 6(6):1055–1062, June 1990.

[77] D. Jeulin. Modèles morphologiques de structures aléatoires et de changement d'échelle. *Thèse de Doctorat d'état*, 1991.

[78] D. Jeulin. Random texture models for material structures. *Statistics and Computing archive*, Volume 10:Issue 2, 2000.

[79] D. Jeulin. *Homogénéisation en mécanique des matériaux, tome 1 : Matériaux aléatoires élastiques et milieux périodiques*,, chapter Caractérisation morphologique et modèles de structures aléatoires, pages pp. 95–132. 2001.

[80] D. Jeulin, T. Clemenceau, M. Faessel, V. Tariel, G. Contesse, and Fanget A. Morphological analysis of 3d images of pyrotechnical granular materials obtained by microtomography. In *17 DYMAT Technical Meeting Cambridge*, 2007.

[81] M.Y. Joshi. *A Class of Stochastic Models for Porous Media.* PhD thesis, University of Kansas, 1974.

[82] T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window - theory and experiment. *Ieee Transactions On Pattern Analysis And Machine Intelligence*, 16(9):920–932, September 1994.

[83] A. Karma and W. J. Rappel. Quantitative phase-field modeling of dendritic growth in two and three dimensions. *Physical Review E*, 57(4):4323–4349, April 1998.

[84] A. J. Katz and A. H. Thompson. Prediction of rock electrical-conductivity from mercury injection measurements. *Journal Of Geophysical Research-Solid Earth And Planets*, 92(B1):599–607, January 1987.

[85] R. A. Ketcham and W. D. Carlson. Acquisition, optimization and interpretation of x-ray computed tomographic imagery: applications to the geosciences. *Computers & Geosciences*, 27(4):381–400, May 2001.

[86] E. Khalimsky, R. Kopperman, and P. R. Meyer. Computer-graphics and connected topologies on finite ordered sets. *Topology And Its Applications*, 36(1):1–17, July 1990.

[87] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[88] L.J. Klinkenberg. The permeability of porous media to liquids and gases. *Drilling and Production Practice*, 1941.

[89] M. A. Knackstedt, C. H. Arns, M. Saadatfar, T. J. Senden, A. Limaye, A. Sakellariou, A. P. Sheppard, R. M. Sok, W. Schrof, and H. Steininger. Elastic and transport properties of cellular solids derived from three-dimensional tomographic images. *Proceedings Of The Royal Society A-Mathematical Physical And Engineering Sciences*, 462(2073):2833–2862, September 2006.

[90] M. A. Knackstedt, C. H. Arns, T. J. Senden, and K. Gross. Structure and properties of clinical coralline implants measured via 3d imaging and analysis. *Biomaterials*, 27(13):2776–2786, May 2006.

[91] M. Koguchi, H. Kakibayashi, R. Tsuneta, M. Yamaoka, T. Niino, N. Tanaka, K. Kase, and M. Iwaki. Three-dimensional stem for observing nanostructures. *Journal Of Electron Microscopy*, 50(3):235–241, 2001.

[92] J. P. Korb, L. Monteilhet, P. J. McDonald, and J. Mitchell. Microstructure and texture of hydrated cement-based materials: A proton field cycling relaxometry approach. *Cement And Concrete Research*, 37(3):295–302, March 2007.

[93] K. E. Kurtis, N. H. El-Ashkar, C. L. Collins, and N. N. Nalk. Examining cement-based materials by laser scanning confocal microscopy. *Cement & Concrete Composites*, 25(7):695–701, October 2003.

[94] K. Kurumisawa and K. Tanaka. Three-dimensional visualization of pore structure in hardened cement paste by the gallium intrusion technique. *Cement And Concrete Research*, 36(2):330–336, February 2006.

[95] L. Landau and E. Lifchitz. *Mecanique*. Librairie du globe, 1989.

[96] E. N. Landis, E. N. Nagy, and D. T. Keane. Microstructure and fracture in three dimensions. *Engineering Fracture Mechanics*, 70(7-8):911–925, May 2003.

[97] E. N. Landis, T. Zhang, E. N. Nagy, G. Nagy, and W. R. Franklin. Cracking, damage and fracture in four dimensions. *Materials And Structures*, 40(4):357–364, May 2007.

[98] C. Lantuejoul. Ergodicity and integral range. *Journal Of Microscopy-Oxford*, 161:387–403, March 1991.

[99] L. L. LATOUR, R. L. KLEINBERG, P. P. MITRA, and C. H. SOTAK. Pore-size distributions and tortuosity in heterogeneous porous-media. *Journal Of Magnetic Resonance Series A*, 112(1):83–91, January 1995.

[100] C.D. Lawrence. Transport of oxygen through concrete. *F.P. Glasser (Ed), The chemistry and Chemically-Related Poroperties of Cement, British Ceramic Society Proceedings*, 35:288–293, 1984.

[101] N. Lenoir, Y. Pannier, S.A. Hall, Bornert M., P. Bésuelle, J. Desrues, and G. Viggiani. Combining x-ray ct and 3d digital image correlation for studying strain localization in granular materials. In *8th International Workshop on Bifurcations and Degradations in Geomaterials, Lake Louise, Alberta, Canada.*, 2008.

[102] M.D. Levine. Vision in man and machine. *McGraw-Hil*, 1985.

[103] P. Levitz. *Handbook of Porous Media*, chapter 2: Statistical modeling of Pore network, page p 37. Wiley-VCH, 2002.

[104] P. Levitz, J. M. Drake, and J. Klafter. Direct energy-transfer in pores - geometrical cross-overs and apparent dimensionality. *Chemical Physics Letters*, 148(6):557–561, July 1988.

[105] P. Levitz, G. Ehret, S. K. Sinha, and J. M. Drake. Porous vycor glass - the microstructure as probed by electron-microscopy, direct energy-transfer, small-angle scattering, and molecular adsorption. *Journal Of Chemical Physics*, 95(8):6151–6161, October 1991.

[106] P. Levitz and D. Tchoubar. Disordered porous solids - from chord distributions to small-angle scattering. *Journal De Physique I*, 2(6):771–790, June 1992.

[107] S. Lu, E. N. Landis, and D. T. Keane. X-ray microtomographic studies of pore structure and permeability in portland cement concrete. *Materials And Structures*, 39(6):611–620, July 2006.

[108] W. Y. Ma and B. S. Manjunath. Edgeflow: A technique for boundary detection and image segmentation. *Ieee Transactions On Image Processing*, 9(8):1375–1388, August 2000.

[109] R. Maggion. *Etude de l'évolution microtexturale de pâtes de silicate tericalcique hydraté.* PhD thesis, Université d'Orléans, 1992.

[110] Y. Maltais and J. Marchand. Influence of curing temperature on cement hydration and mechanical strength development of fly ash mortars. *Cement And Concrete Research*, 27(7):1009–1020, July 1997.

[111] G. Matheron. *Elements pour une theorie des milieux poreux.* Masson et Cie, 1967.

[112] G. Matheron. *Random Sets and Integral Geometry.* John Wiley, Ney York, 1975.

[113] K. R. Mecke, T. Buchert, and H. Wagner. Robust morphological measures for large-scale structure in the universe. *Astronomy And Astrophysics*, 288(3):697–704, August 1994.

[114] A. Mehnert and P. Jackway. An improved seeded region growing algorithm. *Pattern Recognition Letters*, 18(10):1065–1071, October 1997.

[115] J. Mering and D. Tchoubar. Interprétation de la diffusion centrale des rayons x par les systÃ¨mes poreux. i. *J. Appl. Cryst.*, 1:153–165, 1968.

[116] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal Of Chemical Physics*, 21(6):1087–1092, 1953.

[117] R. E. Miles and P. Davy. Precise and general conditions for validity of a comprehensive set of stereological fundamental formulas. *Journal Of Microscopy-Oxford*, 107(AUG):211–226, 1976.

[118] G. Mobus and B. J. Inkson. Nanoscale tomography in materials science. *Materials Today*, 10(12):18–25, December 2007.

[119] N. Moelans, B. Blanpain, and P. Wollants. An introduction to phase-field modeling of microstructure evolution. *Calphad-Computer Coupling Of Phase Diagrams And Thermochemistry*, 32(2):268–294, June 2008.

[120] M. Mouret, E. Ringot, and A. Bascoul. Image analysis: a tool for the characterisation of hydration of cement in concrete - metrological aspects of magnification on measurement. *Cement & Concrete Composites*, 23(2-3):201–206, April 2001.

[121] L. Najman and M. Schmitt. Geodesic saliency of watershed contours and hierarchical segmentation. *Ieee Transactions On Pattern Analysis And Machine Intelligence*, 18(12):1163–1173, December 1996.

[122] Y. Nakashima and S. Kamiya. Mathematica programs for the analysis of three-dimensional pore connectivity and anisotropic tortuosity of porous rocks using x-ray computed tomography image data. *Journal Of Nuclear Science And Technology*, 44(9):1233–1247, September 2007.

[123] Y. Nakashima, T. Nakano, K. Nakamura, K. Uesugi, A. Tsuchiyama, and S. Ikeda. Three-dimensional diffusion of non-sorbing species in porous sandstone: computer simulation based on x-ray microtomography using synchrotron. *Journal Of Contaminant Hydrology*, 74(1-4):253–264, October 2004.

[124] K. M. Nemati, P. J. M. Monteiro, and K. L. Scrivener. Analysis of compressive stress-induced cracks in concrete. *Aci Materials Journal*, 95(5):617–630, September 1998.

[125] N. Nestle, P. Galvosas, and J. Karger. Liquid-phase self-diffusion in hydrating cement pastes - results from nmr studies and perspectives for further research. *Cement And Concrete Research*, 37(3):398–413, March 2007.

[126] H. Niedrig. Electron backscattering from thin-films. *Journal Of Applied Physics*, 53(4):R15–R49, 1982.

[127] I. Odler. The bet-specific surface area of hydrated portland cement and related materials. *Cement And Concrete Research*, 33(12):2049–2056, December 2003.

[128] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed - algorithms based on hamilton-jacobi formulations. *Journal Of Computational Physics*, 79(1):12–49, November 1988.

[129] A. U. Ozturk. Microstructural investigation on the relationship between microstructure and compressive strength of cement mortars. *Computational Materials Science*, 41(2):127–133, December 2007.

[130] C. L. Page and P. Lambert. Kinetics of oxygen diffusion in hardened cement pastes. *Journal Of Materials Science*, 22(3):942–946, March 1987.

[131] C. L. Page and K. W. J. Treadaway. Aspects of the electrochemistry of steel in concrete. *Nature*, 297(5862):109–115, 1982.

[132] D. P. Panda and A. Rosenfeld. Image segmentation by pixel classification in (gray level, edge value) space. *Ieee Transactions On Computers*, 27(9):875–879, 1978.

[133] R. Pohle and K.D. Toennies. Segmentation of medical images using adaptive region growing, 2001.

[134] G. Porod. Small angle x-ray scattering. In *Syracuse*, 1965.

[135] L. Pothuaud, P. Porion, E. Lespessailles, C. L. Benhamou, and P. Levitz. A new method for three-dimensional skeleton graph analysis of porous media: application to trabecular bone microarchitecture. *Journal Of Microscopy-Oxford*, 199:149–161, August 2000.

[136] T.C. Powers. Study of the physical properties of hardened portland cement paste. *J. Am. Concr. Inst.*, 22, 1947.

[137] M. A. B. Promentilla, T. Sugiyama, T. Hitomi, and N. Takeda. Characterizing the 3d pore structure of hardened cement paste with synchrotron microtomography. *Journal Of Advanced Concrete Technology*, 6(2):273–286, June 2008.

[138] D. A. Quenard, K. Xu, H. M. Kunzel, D. P. Bentz, and N. S. Martys. Microstructure and transport properties of porous building materials. *Materials And Structures*, 31(209):317–324, June 1998.

[139] U. Rattanasak and K. Kendall. Pore structure of cement/pozzolan composites by x-ray microtomography. *Cement And Concrete Research*, 35(4):637–640, April 2005.

[140] S. K. Ray. Melt-clast interaction and power-law size distribution of clasts in pseudotachylytes. *Journal Of Structural Geology*, 26(10):1831–1843, 2004.

[141] C. Redon, L. Chermant, J. L. Chermant, and M. Coster. Assessment of fibre orientation in reinforced concrete using fourier image transform. *Journal Of Microscopy-Oxford*, 191:258–265, September 1998.

[142] M. D. Rintoul, S. Torquato, C. Yeong, D. T. Keane, S. Erramilli, Y. N. Jun, D. M. Dabbs, and I. A. Aksay. Structure and transport properties of a porous magnetic gel via x-ray microtomography. *Physical Review E*, 54(3):2663–2669, September 1996.

[143] A. Rosenfeld and Kak A.C. *Digital Picture Processing*. Academic Press, New York, 1982.

[144] M. G. Rozman and M. Utz. Efficient reconstruction of multiphase morphologies from correlation functions. *Physical Review E*, 6306(6):066701, June 2001.

[145] H. Saito and A. Deguchi. Leaching tests on different mortars using accelerated electrochemical method. *Cement And Concrete Research*, 30(11):1815–1825, November 2000.

[146] A. Sakellariou, C. H. Arns, A. P. Sheppard, R. M. Sok, H. Averdunk, A. Limaye, A. C. Jones, T. J. Senden, and M. A. Knackstedt. Developing a virtual materials laboratory. *Materials Today*, 10(12):44–51, December 2007.

[147] B. Sapoval, A. Baldassarri, and A. Gabrielli. Self-stabilized fractality of seacoasts through damped erosion. *Physical Review Letters*, 93(9):098501, August 2004.

[148] H. Scher and R. Zallen. Critical density in percolation processes. *Journal Of Chemical Physics*, 53(9):3759–&, 1970.

[149] M. Schmitt. Geodesic arcs in non-euclidean metrics: Application to the propagation function. *Revue &Intelligence Artificielle*, 3, no.2:43–76, 1989.

[150] L. M. Schwartz, F. Auzerais, J. Dunsmuir, N. Martys, D. P. Bentz, and S. Torquato. Transport and diffusion in 3-dimensional composite media. *Physica A*, 207(1-3):28–36, June 1994.

[151] L. M. Schwartz, N. Martys, D. P. Bentz, E. J. Garboczi, and S. Torquato. Cross-property relations and permeability estimation in model porous-media. *Physical Review E*, 48(6):4584–4591, December 1993.

[152] G. D. Scott and D. M. Kilgour. Density of random close packing of spheres. *Journal Of Physics D-Applied Physics*, 2(6):863–&, 1969.

[153] K. Scrivener and Pratt P.L. Characterisation of portland cement hydration by electron optical techniques. In *Electron microscopy of materials. Proc Mat Res Soc Symp*, 1983.

[154] K. L. Scrivener. Backscattered electron imaging of cementitious microstructures: understanding and quantification. *Cement & Concrete Composites*, 26(8):935–945, November 2004.

[155] K. L. Scrivener and K. M. Nemati. The percolation of pore space in the cement paste aggregate interfacial zone of concrete. *Cement And Concrete Research*, 26(1):35–40, January 1996.

[156] P. N. Sen, L. M. Schwartz, P. P. Mitra, and B. I. Halperin. Surface relaxation and the long-time diffusion-coefficient in porous-media - periodic geometries. *Physical Review B*, 49(1):215–225, January 1994.

[157] J. Sercombe, R. Vidal, and F. Adenot. Diffusion des gaz dans les ciments, mécanismes et paramètres principaux. In *Transfert, Lille*, 2006.

[158] J. Serra. *Image Analysis and Mathematical Morphology - Vol. I . 610 p.* Ac. Press, London, 1982.

[159] G. Sezer, K. Ramyar, B. Karasu, B. Goktepe, and A. Sezer. Image analysis of sulfate attack on hardened cement paste. *Materials & Design (short communication)*, 29:224–231, 2008.

[160] A. P. Sheppard, R. M. Sok, and H. Averdunk. Techniques for image enhancement and segmentation of tomographic images of porous materials. *Physica A-Statistical Mechanics And Its Applications*, 339(1-2):145–151, August 2004.

[161] P. Spanne, J. F. Thovert, C. J. Jacquin, W. B. Lindquist, K. W. Jones, and P. M. Adler. Synchrotron computed microtomography of porous-media - topology and transports. *Physical Review Letters*, 73(14):2001–2004, October 1994.

[162] M. Stampanoni, G. Borchert, P. Wyss, R. Abela, B. Patterson, S. Hunt, D. Vermeulen, and P. Ruegsegger. High resolution x-ray detector for synchrotron-based microtomography. *Nuclear Instruments & Methods In Physics Research Section A-Accelerators Spectrometers Detectors And Associated Equipment*, 491(1-2):291–301, September 2002.

[163] D. Stoyan, W.S. Kendal, and K. Mecke. *Stochastic Geometry and its ApplicationsD. Stoyan, l, J. Mecke, , ,)*. John Wiley & Sons, 1995.

[164] M. Tabb and N. Ahuja. Multiscale image segmentation by integrated edge and region detection. *Ieee Transactions On Image Processing*, 6(5):642–655, May 1997.

[165] V. Tariel, D. Jeulin, A. Fanget, and G. Contesse. 3d multi-scale segmentation of granular materials. *Image Analysis & Stereology*, 2008.

[166] M. Teubner. Level surfaces of gaussian random-fields and microemulsions. *Europhysics Letters*, 14(5):403–408, March 1991.

[167] J. J. Thomas, H. M. Jennings, and A. J. Allen. The surface area of cement paste as measured by neutron scattering: Evidence for two c-s-h morphologies. *Cement And Concrete Research*, 28(6):897–905, June 1998.

[168] S. Torquato and Lu. Chord-length distribution function for 2-phase random-media. *Physical Review E*, 47(4):2950–2953, April 1993.

[169] S. Torquato and G. Stell. Microstructure of 2-phase random-media .1. the n-point probability functions. *Journal Of Chemical Physics*, 77(4):2071–2077, 1982.

[170] E.E. Underwood. *Quantitative Stereology*. Addison, 1970.

[171] L. Vincent. Exact euclidean distance function by chain propagations. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, pages 520–525, 1991.

[172] L. Vincent. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *Ieee Transactions On Image Processing*, 2(2):176–201, April 1993.

[173] L. Vincent and P. Soille. Watersheds in digital spaces - an efficient algorithm based on immersion simulations. *Ieee Transactions On Pattern Analysis And Machine Intelligence*, 13(6):583–598, 1991.

[174] M. von Ardenne. Das elektronen-rastermikroskop. theoretische grundlagen. *Zeitschrift fur Physik*, 108:553 572, 1938.

[175] S. Y. Wan and W. E. Higgins. Symmetric region growing. *Ieee Transactions On Image Processing*, 12(9):1007–1015, September 2003.

[176] J. S. Weszka. Survey of threshold selection techniques. *Computer Graphics And Image Processing*, 7(2):259–265, 1978.

[177] D. B. Wiles and R. A. Young. A new computer-program for rietveld analysis of x-ray-powder diffraction patterns. *Journal Of Applied Crystallography*, 14(APR):149–151, 1981.

[178] K. L. Willis, A. B. Abell, and D. A. Lange. Image-based characterization of cement pore structure using wood's metal intrusion. *Cement And Concrete Research*, 28(12):1695–1705, December 1998.

[179] D. N. Winslow and S. Diamond. Specific surface of hardened portland-cement paste as determined by small-angle x-ray-scattering. *Journal Of The American Ceramic Society*, 57(5):193–197, 1974.

[180] H. S. Wong, N. R. Buenfeld, and M. K. Head. Estimating transport properties of mortars using image analysis on backscattered electron images. *Cement And Concrete Research*, 36(8):1556–1566, August 2006.

[181] H. S. Wong, M. K. Head, and N. R. Buenfeld. Pore segmentation of cement-based materials from backscattered electron images. *Cement And Concrete Research*, 36(6):1083–1090, June 2006.

[182] R. Yang and N. R. Buenfeld. Binary segmentation of aggregate in sem image analysis of concrete. *Cement And Concrete Research*, 31(3):437–441, March 2001.

[183] C. L. Y. Yeong and S. Torquato. Reconstructing random media. ii. three-dimensional media from two-dimensiomal cuts. *Physical Review E*, 58(1):224–233, July 1998.

[184] P. B. Zdenek. Physical model for steel corrosion in concrete sea structures-application. *Journal of the Structural Division*, 105, No. 6:1155–1166, 1979.