



**Traitement d'images :
au delà de la deuxième
dimension**

Mémoire présenté à
l'Université Jean Monnet
de Saint-Etienne

par

Etienne Decencière

pour obtenir
l'habilitation à diriger des recherches

le 6 Mars 2008 devant le jury composé de :

Bernard Besserer *Rapporteur*

Isabelle Bloch *Rapporteur*

Isabelle Magnin *Rapporteur*

Fernand Meyer *Examineur*

Michel Jourlin *Directeur de recherche*

Préface

La rédaction de ce mémoire m'offre une bonne occasion pour marquer un arrêt et prendre de la distance par rapport aux projets que j'ai menés à bien pendant les dix dernières années.

Cependant, le choix de son titre n'a pas été simple. La plupart des travaux qui y sont décrits ont été réalisés dans le cadre de projets de recherche en collaboration avec des partenaires industriels. L'objectif principal consistait donc à apporter des solutions originales à des problèmes concrets. Si j'espère avoir obtenu quelques succès dans cette tâche, il faut aussi constater que face à la variété des applications il a été difficile de garder une ligne de recherche constante tout au long de ce parcours. Renouvellement régulier de l'inspiration ou émiettement des efforts de recherche ?

Tout d'abord, la morphologie mathématique constitue une large base de travail, à laquelle j'ai apporté quelques briques supplémentaires à travers différentes études. Cependant, je trouvais ce thème trop large pour constituer le sujet de ce mémoire, et d'autant plus que je n'ai pas hésité à chercher ailleurs les outils dont j'avais besoin pour travailler.

J'avais aussi la possibilité de présenter une sorte de catalogue des projets sur lesquels j'ai travaillé. Cela s'est déjà vu, mais j'ai préféré choisir un fil directeur a posteriori, à savoir l'augmentation du nombre de dimensions des images, qui sert de guide assez commode à travers les différentes applications abordées, et qui synthétise de façon satisfaisante une bonne partie des difficultés rencontrées.

Alors, émiettement de l'effort de recherche ? De mon point de vue, la variété des applications apporte de la richesse. Mais, afin de produire une recherche académique de haut niveau, le chercheur doit non seulement trouver des solutions aux problèmes posés, mais aussi essayer d'identifier, parmi les solutions

envisagées, celles qui pourraient être généralisées et ainsi appliquées à d'autres domaines ; il doit prendre de la hauteur par rapport au cas particulier, et tenter de trouver des réponses fertiles, qui permettraient de répondre à d'autres questions. C'est à ce prix que la recherche transversale et appliquée, comme celle qui est menée en particulier au Centre de Morphologie Mathématique de l'école des mines de Paris, peut prétendre à une certaine reconnaissance académique.

J'espère que cet exercice de synthèse aidera à faire ressortir les résultats les plus intéressants, dont on peut espérer qu'ils constitueront une base consolidée d'où repartir vers des objectifs que cette halte aura aidé à définir.

Etienne Decencière

Hendaye, juillet 2007

Table des matières

1	Introduction	1
1.1	Les images en tant que représentation de la réalité	2
1.2	Point de vue mathématique	3
1.3	Algorithmique et programmation informatique	6
1.4	Visualisation	7
1.5	Plan du mémoire	8
2	De la surface au volume	11
2.1	Caractérisation de surfaces	11
2.2	Simulation du contact hydrodynamique	15
2.3	Exploitation des résultats	18
2.4	Conclusion	19
3	Quelques problèmes liés à la troisième dimension	21
3.1	Premier cas d'étude : visualisation de structures anatomiques	22
3.2	Deuxième cas d'étude : segmentation de cubes sismiques	24
3.3	Leçons tirées des cas d'étude	25
3.4	Des solutions au cas par cas	27
3.5	Segmentation interactive 3d	31
3.6	Conclusion	33
4	Le temps : une dimension comme les autres ?	37
4.1	Lorsque deux dimensions ne suffisent pas : le cas de la restauration de films	39
4.1.1	Détection de défauts locaux	39

4.1.2	Interpolation des défauts	40
4.2	Représentation d'un champ de vecteurs de déplacement	43
4.3	Conclusion	47
5	Et au delà ?	49
5.1	Des besoins réels	49
5.2	Coupes minimales de graphes	51
5.3	Segmentation par marqueurs à l'aide de coupes de graphes	53
5.4	Coupe d'un graphe associé à une segmentation par LPE	56
5.5	Conclusion	59
6	Conclusion	61
A	Curriculum vitæ détaillé d'Etienne Decencière	63
B	Article joint au chapitre 2	65
C	Article joint au chapitre 3	67
D	Article joint au chapitre 4	69
E	Article joint au chapitre 5	71
	Bibliographie	73

Dieu est par définition inétendu, mais il nous est permis, pour la clarté de notre énoncé, de lui supposer un nombre quelconque, plus grand que zéro, de dimensions, bien qu'il n'en ait aucune, si ces dimensions disparaissent dans les deux membres de nos identités. Nous nous contenterons de deux dimensions, afin qu'on se représente aisément des figures de géométrie plane sur une feuille de papier.

Alfred Jarry, Gestes et opinions de Dr Faustroll, pataphysicien
(1898)

1

Introduction

Le traitement d'images constitue une branche des mathématiques appliquées qui ne cesse de trouver de nouvelles applications. Cette progression est possible grâce aussi bien au développement de moyens de calcul de plus en plus puissants, qu'à l'apparition de moyens d'acquisition nouveaux. On constate en particulier une augmentation du nombre de dimensions : initialement à deux dimensions, les images à trois dimensions, ou davantage, deviennent courantes.

Le cas de l'imagerie médicale est emblématique. Le développement et la progressive démocratisation d'appareils médicaux d'acquisition tels que les scanners à rayons X, ou les imageurs par résonance magnétique (IRM), fournissent désormais aux médecins des images à trois dimensions riches en information, et précieuses pour la recherche médicale, le diagnostic et le traitement. Cependant, cette richesse implique aussi une difficulté accrue d'interprétation, au point que le domaine de l'analyse d'images médicales tridimensionnelles est devenu un champ de recherche très actif. D'autres domaines d'activité connaissent les mêmes développements, tels que les sciences des matériaux, ou l'imagerie sismique.

Nous souhaitons dans ce mémoire, en partant de l'expérience des projets de recherche que nous avons menés, mettre en évidence quelques unes des difficultés qui surgissent lorsqu'on traite des images comportant plus de deux dimensions, les analyser, et proposer quelques lignes directrices pour la recherche de solutions.

Remarquons que les valeurs en chaque pixel, initialement binaires, puis scalaires, peuvent aujourd'hui être des vecteurs, ou des matrices. Cette augmentation de la complexité de l'espace des valeurs constitue aussi un défi important pour les traiteurs d'images, mais nous ne l'aborderons pratiquement pas dans cette étude.

1.1 Les images en tant que représentation de la réalité

Le verbe *représenter* connaît une riche palette de sens, similaires mais dont les nuances sont importantes. Nous lisons ainsi dans le dictionnaire de l'ATILF (Analyse et Traitement Informatique de la Langue Française) :

REPRÉSENTER

- I.** - Rendre effectivement présent à la vue, à l'esprit de quelqu'un.
- II.** - Rendre présent sous la forme d'un substitut, en recourant à un artifice ; être cet artifice, ce substitut.
- III.** - Rendre présent par son existence, par sa propre présence ; être le représentant de.
- IV.** - Rendre présent à l'esprit (un objet généralement absent) ; être l'intermédiaire qui permet à l'esprit d'appréhender cet objet.

C'est ce quatrième sens qui nous semble décrire le mieux le rôle que jouent les images, en tant que représentation de la réalité. Elles sont le fruit d'un processus d'acquisition ; beaucoup d'information est perdue au long de la chaîne de création de l'image, de façon volontaire ou pas, et des éléments nouveaux sont rajoutés, tels que le bruit. Un exemple typique est celui des photographies ou vidéos numériques, où une image bidimensionnelle cherche à donner une vue aussi fidèle que possible d'une scène tridimensionnelle. La fidélité en

question est, soit dit en passant, très liée à l'application visée. Pour obtenir une bonne représentation, il faut donc non seulement limiter le bruit, mais aussi, et surtout, faire en sorte que l'information perdue soit aussi peu intéressante que possible pour l'application finale. Le but du traiteur d'images peut alors être vu comme l'amélioration de la qualité de l'image comprise comme moyen d'appréhender la réalité.

De nombreux objets réels sont convenablement modélisés en deux dimensions, et peuvent donc être représentés de façon satisfaisante par une image bidimensionnelle. Tel est le cas, par exemple, d'un support de texte, que ce soit une page, ou un panneau routier. Néanmoins, très souvent les images 2D représentent des scènes qui seraient mieux modélisées avec trois, voire quatre dimensions. Ainsi, la plupart des photos ou des séquences vidéo déjà évoquées sont obtenues en projetant une scène à trois dimensions sur un support à deux dimensions, ce qui introduit de nombreux problèmes correspondant à des pertes d'information, comme les occultations.

Le passage de l'image 2D vers l'image 3D fait donc partie de cette recherche d'une meilleure représentation de la réalité. On perd ainsi moins d'information pendant l'acquisition, et l'image 3D devrait être plus facile à analyser. En contrepartie, de nouvelles sources d'erreurs peuvent apparaître (telles qu'un bruit plus important) mais surtout la quantité d'information est nettement plus importante, ce qui pose de nombreux problèmes pratiques, et les outils mathématiques employés dans un cadre 2D peuvent se révéler insuffisants dans un cadre 3D.

Regardons maintenant ce que cette augmentation du nombre de dimensions implique pour certains cadres théoriques employés en traitement d'images.

1.2 Point de vue mathématique

L'augmentation du nombre de dimensions ne pose pas forcément de nouvelles difficultés mathématiques. Par exemple, un des cadres théoriques les plus couramment employés en traitement d'images, à savoir celui de l'algèbre linéaire, ne pose pas de difficulté formelle lorsque le nombre de dimensions de l'espace de départ augmente. On peut dresser le même constat pour l'approche non linéaire

du traitement d'images ayant rencontré le plus de succès, à savoir la morphologie mathématique. Là encore, l'augmentation du nombre de dimensions ne pose pas de problème.

Le cadre topologique est, quant à lui, étroitement lié au nombre de dimensions. Dans \mathbb{R}^3 existent des structures topologiques inconnues dans \mathbb{R}^2 . La caractérisation de la forme d'un objet devient par conséquent plus complexe avec l'augmentation du nombre de dimensions. Nous verrons, dans le chapitre 3, que ce comportement rend aussi plus complexe le filtrage.

Le moment est venu de préciser que nous nous limiterons, dans ce mémoire, à une approche purement discrète. Toutes les images que nous considérerons seront représentées par des fonctions d'un sous ensemble D de \mathbb{Z}^d , où d est un entier strictement positif, dans un ensemble E (typiquement \mathbb{Z}). Nous choisirons souvent directement \mathbb{Z}^d comme espace de départ, et non pas un produit d'intervalles de \mathbb{Z} , pour simplifier l'exposé. D'une part les notations résultantes sont plus légères, et d'autre part nous évitons ainsi de parler des problèmes de bords, non pas que ces derniers soient négligeables ni toujours faciles à résoudre, mais nous ne souhaitons pas aborder ici ces considérations techniques.

Ainsi, une image 2D classique, pourra être représentée par une fonction de \mathbb{Z}^2 dans \mathbb{Z} , et une séquence d'images 3D, telle qu'une séquence d'images obtenues par résonance magnétique, par une fonction de \mathbb{Z}^4 dans \mathbb{Z} . La dimensions temporelle, comme nous le verrons dans le chapitre 4, joue un rôle particulier.

Les exemples précédents indiquent déjà que, dans ce mémoire, l'ensemble d'arrivée sera très souvent \mathbb{Z} . Nous serons cependant amenés, lorsque nous parlerons des séquences d'images, à dire quelques mots sur les images de vecteurs. Plus généralement, comme nous verrons dans le chapitre 2, le nombre de dimensions d'une image n'est pas forcément une donnée intrinsèque. Ainsi, n'importe quelle image 2D à niveaux de gris peut être modélisée comme une fonction entière de \mathbb{Z}^2 :

$$I : \mathbb{Z}^2 \longrightarrow \mathbb{Z} \tag{1.1}$$

$$(x, y) \longmapsto I(x, y) \tag{1.2}$$

ou comme une fonction binaire de \mathbb{Z}^3 :

$$J : \mathbb{Z}^3 \longrightarrow \{0, 1\} \quad (1.3)$$

$$(x, y, z) \longmapsto J(x, y, z) \quad (1.4)$$

où, pour tout (x, y) , $J(x, y, z)$ vaut 1 si $z \leq I(x, y)$, et sinon 0. $J(., ., z)$ est en fait le seuil de I au niveau z .

De même, dans les séquences vidéo, les images de perfusion médicales, ou les images multi-spectrales, en chaque position on a une suite de valeurs, correspondant à des temps consécutifs dans les deux premiers cas, et à des réponses dans des bandes spectrales différentes dans le troisième cas. Elles peuvent donc être modélisées soit comme des images à deux dimensions à valeurs vectorielles :

$$I : \mathbb{Z}^2 \longrightarrow \mathbb{Z}^d \quad (1.5)$$

$$(x, y) \longmapsto I(x, y) = (I_1(x, y), \dots, I_d(x, y)), \quad (1.6)$$

soit comme des images tridimensionnelles :

$$J : \mathbb{Z}^2 \times \{1, \dots, d\} \longrightarrow \mathbb{Z} \quad (1.7)$$

$$(x, y, z) \longmapsto J(x, y, z) = I_z(x, y). \quad (1.8)$$

Le choix du modèle est souvent guidé par des considérations pratiques. Ainsi, d'un point de vue informatique, une image 2D à niveaux de gris prend moins de place qu'une image 3D binaire. Mais il est aussi important de comprendre que si on incorpore des dimensions supplémentaires dans l'espace de départ, on suppose implicitement que deux points proches (quelle que soit la distance choisie) dans cet espace sont corrélés. Par exemple, dans le cas des images multi-spectrales, deux bandes voisines peuvent être très différentes. Dans ce cas, utiliser une représentation 3D peut être trompeur.

Justement, le choix d'une relation de voisinage est une étape importante dans la chaîne d'analyse. Comme nous venons de dire, elle n'a de sens que si deux points voisins ont, en général, des valeurs similaires. Sur des images 2D, ce choix peut déjà poser quelques difficultés, et d'autant plus lorsqu'on veut que la relation de voisinage ait de bonnes propriétés topologiques. On

imagine alors aisément, et nous illustrerons cette difficulté dans les chapitres 4 et 5, que ce choix peut devenir largement non trivial avec l'augmentation du nombre de dimensions, et en particulier lorsque nous considérerons une dimension temporelle.

Nous avons commencé à aborder dans cette section l'aspect informatique. Regardons d'un peu plus près les conséquences de l'augmentation du nombre de dimensions dans ce domaine.

1.3 Algorithmique et programmation informatique

Constatation qui pourrait sembler surprenante, l'augmentation du nombre de dimensions pose plus de problèmes aux niveaux algorithmique et informatique qu'au niveau mathématique. En effet, si une solution mathématique à un problème donné ne marche plus suite à l'augmentation du nombre de dimensions, alors, bien évidemment, il faudra aussi modifier le programme informatique correspondant. Mais, même si du point de vue mathématique on ne rencontre pas de problème particulier, il en va souvent autrement en matière d'algorithmes ou de programmation.

Concernant les algorithmes, contentons-nous de dire que leur complexité théorique devient capitale lorsque le nombre de dimensions augmente, tout simplement parce que le nombre de valeurs à traiter croît de façon exponentielle avec le nombre de dimensions. Ainsi, il devient souvent indispensable d'améliorer les algorithmes existants, voire d'en concevoir des nouveaux.

Concernant la programmation, commençons par résumer le cas du Centre de Morphologie Mathématique. Au début des années 80, un logiciel, appelé *Xlim*, a été écrit en C pour traiter des images 2D à niveaux de gris. Quelque temps après, un nouveau logiciel, *xl3d*, toujours écrit en C, a été développé pour traiter les images 3D. Il était aussi capable de traiter les images 2D. Ceci dit, *xl3d* ne comportait pas toutes les fonctions disponibles dans *Xlim*, ce qui a motivé la création d'un logiciel, *Xlim3d*, englobant les deux précédents. Cette nouvelle version n'était véritablement qu'un chapeau recouvrant les précédentes. Avec le temps, des compléments ont été apportés pour gérer les images en couleurs ainsi que les images de vecteurs, mais on atteignait là la limite de la structure

de *Xlim3d*, car chaque fois qu'on rajoutait un nouveau type d'image, il fallait écrire une nouvelle version de toutes les fonctions de la librairie de traitement d'images. Par ailleurs, on était limité à trois dimensions. En 2003, on a donc démarré le développement d'une nouvelle librairie, *Morph-M*, cette fois-ci en langage C++, ce qui a permis d'avoir des types d'images abstraits, avec un nombre de dimensions quelconque, à valeurs, elles aussi, abstraites. De façon plus générale, cette nouvelle plateforme de recherche et développement a été conçue en termes de *méta-programmation* [29].

Cet exemple nous montre que l'augmentation du nombre de dimensions oblige souvent à revoir les programmes informatiques de traitement d'images. Garder les anciennes structures peut constituer une cause de complexité inutile et risque d'entraver la mise en œuvre de nouveaux concepts. La méta-programmation permet d'écrire du code suffisamment abstrait pour éviter ces problèmes. Ce code peut, en contrepartie, se révéler moins performant. Des mécanismes de spécialisation permettent alors de l'optimiser pour traiter des cas particuliers.

Ainsi, un logiciel générique est un élément clef pour permettre la mise en œuvre rapide de nouveaux opérateurs. Il permet donc d'accélérer le retour d'expérience. En généralisant, il nous semble que pour pouvoir faire aujourd'hui de la recherche en traitement d'images en dehors d'un cadre applicatif précis, et pouvoir traiter toute la variété d'images disponibles, il est essentiel de disposer d'une librairie logicielle suffisamment abstraite pour pouvoir factoriser au mieux le temps de développement et mettre en pratique rapidement de nouvelles idées.

Lorsqu'on s'attaque à des images à trois dimensions, ou davantage, il devient aussi nécessaire de changer de méthode de visualisation.

1.4 Visualisation

L'augmentation du nombre de dimensions entraîne un problème pratique de taille : comment faire pour les visualiser de façon ergonomique ? Le moniteur est en effet un média adapté à l'affichage d'images à deux dimensions, mais commence à montrer ses limites en 3D. Il faut alors un traitement, souvent complexe, et qui fait l'objet d'ailleurs de recherches actives, pour pouvoir

transformer l'image 3D en une image 2D qui résume de façon satisfaisante l'information tridimensionnelle. Il peut s'agir d'une simple coupe, et c'est d'ailleurs souvent la première étape, mais on a du mal à comprendre ainsi des structures complexes. Par exemple, visualiser un réseau vasculaire ou un intestin coupe par coupe est véritablement un casse-tête. Le traitement d'images rejoint alors la visualisation 3D pour apporter des solutions. Nous montrerons dans le chapitre 3 une étude que nous avons menée autour de cette thématique.

La dimension temporelle est souvent plus facile à visualiser. Il suffit de laisser défiler les images, à une vitesse convenable, sur l'écran, et notre système visuel fait le reste. D'ailleurs, cette méthode de visualisation est aussi utilisée pour naviguer dans des images à trois dimensions spatiales.

Notre système visuel justement atteint ses limites au delà de la troisième dimension spatiale. Il nous est alors impossible de visualiser globalement l'image. Ce problème se pose par exemple en imagerie sismique, où les données brutes comportent souvent cinq dimensions (cf. chapitre 5). Notre capacité à concevoir des traitements pertinents est alors singulièrement gênée par notre difficulté à appréhender les données. Peut-on alors continuer à parler de traitement d'images, dans la mesure où ces données, globalement, n'ont pas de sens visuel pour nous ?

1.5 Plan du mémoire

Le corps de ce mémoire est constitué de quatre chapitres.

Dans le chapitre 2 nous exposons notre expérience en matière d'analyse de surfaces topographiques. Ce travail a ensuite mené vers une étude de l'écoulement tridimensionnel entre une surface rugueuse et une surface plane.

Dans le chapitre suivant, nous parlons d'images à trois dimensions spatiales. Plus particulièrement, nous décrivons comment, en partant d'un problème de visualisation, nous nous sommes intéressés au filtrage d'images tridimensionnelles, puis à leur segmentation.

Dans le chapitre 4 il est question de la dimension temporelle. Nous commençons par résumer nos travaux sur la restauration de films anciens, puis nous introduisons une méthode d'estimation et modélisation du déplacement entre

deux images.

Finalement, nous consacrons un chapitre aux travaux en cours et aux perspectives, qui mettent en œuvre des outils à base de graphes.

Do you not remember [...] how when you entered the realm of Lineland, you were compelled to manifest yourself to the King, not as a Square, but as a Line, because that Linear Realm had not Dimensions enough to represent the whole of you, but only a slice or section of you? In precisely the same way, your country of Two Dimensions is not spacious enough to represent me, a being of Three, but can only exhibit a slice or section of me, which is what you call a Circle.

Edwin A. Abbott, Flatland (1884)

2

De la surface au volume

La morphologie mathématique est plus qu'une théorie mathématique : elle est porteuse aussi, en tant que discipline du traitement d'images, d'une certaine façon de faire quand on aborde des problèmes, qu'on pourrait qualifier de naturaliste. La manière de décrire les images illustre cette mentalité. On utilise ainsi des analogies topographiques, le niveau de gris correspondant à une altitude. Les images analysées par les praticiens de la morphologie mathématique deviennent des paysages composés de pics, vallées, cols, plateaux et autres structures propres aux reliefs.

Mais il arrive aussi que les images correspondent vraiment à des mesures de hauteur.

2.1 Caractérisation de surfaces

Entre 1998 et 2006, nous avons travaillé sur une étude portant sur l'influence de la rugosité des chemises de moteurs à explosion sur le frottement, dans le but de concevoir de nouvelles surfaces permettant de réduire la consommation

d'essence, et donc aussi les émissions de dioxyde de carbone, sans pour autant augmenter la consommation d'huile. En effet, le contact entre la chemise, surface recouvrant l'intérieur du cylindre, et les segments du piston, constitue une des principales sources de perte d'énergie par frottement dans un véhicule. Plus précisément, nous nous sommes intéressés au contact entre le segment dit coupe-feu et le piston, illustré par la figure 2.1.

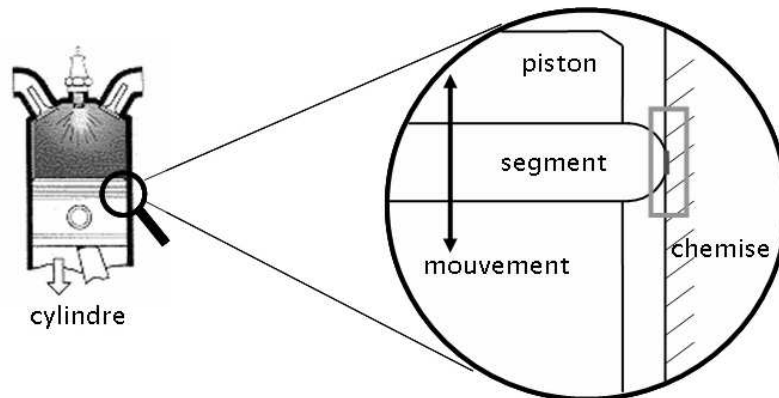


FIG. 2.1: Illustration du contact entre le segment coupe-feu et la chemise du cylindre

La première partie de cette étude a été réalisée en collaboration avec PSA, Renault, ESSO, JPX et l'École nationale supérieure de mécanique et des microtechniques (ENSM), grâce à un financement du programme PREDIT (Programme de recherche et d'innovation dans les transports terrestres). Nous avons commencé par étudier la surface de chemise classiquement utilisée dans les moteurs à explosion des voitures de tourisme. Ces surfaces peuvent être décrites comme des plateaux rugueux traversés par des stries suivant deux directions, symétriques par rapport à l'axe de déplacement du piston, et faisant un angle de 25° avec la direction perpendiculaire à cet axe. C'est deux fois cet angle, soit 50° , appelé angle de striation, qui est utilisé comme caractérisation de la direction des stries. Un exemple d'une telle surface est montré sur la figure 2.2, représentée de deux façons différentes : en niveaux de gris et en vue 3D.

Ce type de surface a été choisi de façon empirique dans les années 40, et continue à être utilisé aujourd'hui dans la très grande majorité des moteurs à explosion de voitures. Au début de notre étude, on ne savait pas vraiment à quoi étaient dues les bonnes performances de ces surfaces, ni quel était le rôle

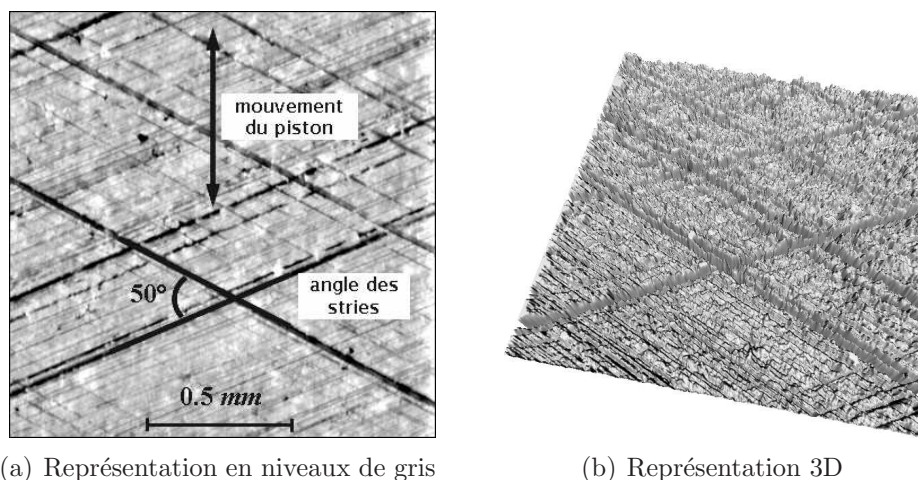


FIG. 2.2: Exemple de chemise type. Dimensions de l'image : 512×512 . Échelle : $3 \mu\text{m} / \text{pixel}$. L'image comporte 256 niveaux de gris, et un niveau de gris correspond à environ $10^{-2} \mu\text{m}$.

des différentes structures.

En partant donc de ces surfaces, nous avons commencé par modifier deux paramètres : l'angle de striation et la rugosité des plateaux. Les surfaces ainsi produites ont été soumises à des tests de frottement grâce à un dispositif expérimental, appelé machine SPC (pour “Segment-Piston-Chemise”), permettant d’approcher les conditions d’utilisation réelles [62], et de mesurer la *puissance moyenne de frottement* ou *PMF*. Sur la surface des chemises, deux relevés de surface étaient réalisés avant l’essai, et deux autres après, le premier en haut de la chemise, au “point mort-haut” (pmh), et le deuxième au milieu de la course du piston, au “point milieu” (pm), à l’aide d’un microscope mécanique. Notons qu’on n’a pas pu, pour des raisons pratiques, faire ces mesures exactement au même endroit de la circonférence de la chemise, avant et après essai. Les paramètres d’acquisition ont été choisis de façon à produire une image 512×512 dont la résolution horizontale était de $3 \mu\text{m}$ par pixel et l’échelle verticale d’environ $10^{-2} \mu\text{m}$ par niveau de gris. Le même type de relevé a été effectué sur le segment coupe-feu, considéré comme le plus important du point de vue du frottement.

Nous avons remarqué assez vite que les spécialistes de la caractérisation de surfaces considèrent que ce genre de relevés sont tridimensionnels. C’est bien

sûr vrai ; comme nous l'avons vu dans l'introduction, ce n'est bien souvent que la façon de les coder qui détermine le nombre de dimensions des images.

Nous avons en charge l'analyse des images ainsi obtenues. Notre but final était d'extraire de ces surfaces des mesures qu'on pourrait corrélérer avec la puissance moyenne de frottement. Nous avons donc commencé par mettre en œuvre des outils classiques de caractérisation globale de mesures, tels que les granulométries, sans trouver de corrélation pertinente.

Nous avons ensuite conçu une méthode de séparation des deux composantes de la texture : les stries et les rayures. Grâce à cet outil, nous avons pu mettre en évidence une diminution de la rugosité des plateaux suite aux essais (cf. l'article [24] joint en annexe B). Ce même outil a permis de montrer que la rugosité des plateaux était corrélée avec le frottement mesuré expérimentalement : une faible rugosité permettait de réduire ce frottement. Il apparaissait donc clairement qu'on avait intérêt à prendre des surfaces avec des rugosités de plateaux faibles.

Par ailleurs, il a été mis en évidence expérimentalement que les surfaces à faible striation étaient clairement les plus intéressantes.

Ainsi, à la fin du projet notre consortium a proposé une surface avec un angle de striation de 15° et des plateaux faiblement rugueux. Des chemises avec cette configuration ont été gravées grâce à un laser. Des test expérimentaux effectués au moyen de la machine SPC on montré qu'on réduisait ainsi de 20% la puissance moyenne de frottement, ce qui permet d'espérer une réduction de 2% des émissions de dioxyde de carbone au niveau du véhicule. Ces résultats sont synthétisés par le graphique de la figure 2.3.

Cependant, aucune autre corrélation n'a pu être obtenue. Nous avons en particulier, à partir de la connaissance de la surface de la chemise et de celle du segment, construit une image binaire 3D correspondant en principe au volume d'huile piégé entre les deux surfaces. En effet, d'après les spécialistes, l'essentiel des pertes par frottement se produit au milieu de la course du piston, alors que le régime de contact entre les deux surfaces est hydrodynamique, par conséquent sans contact entre solides. Dans ce volume, nous avons calculé des longueurs de chemins minimaux pour aller d'un côté à l'autre du segment, pour différents écarts, mais aucune de ces mesures n'a donné de corrélation probante avec les résultats expérimentaux.

La première phase de ce projet s'est donc achevée par un excellent résultat

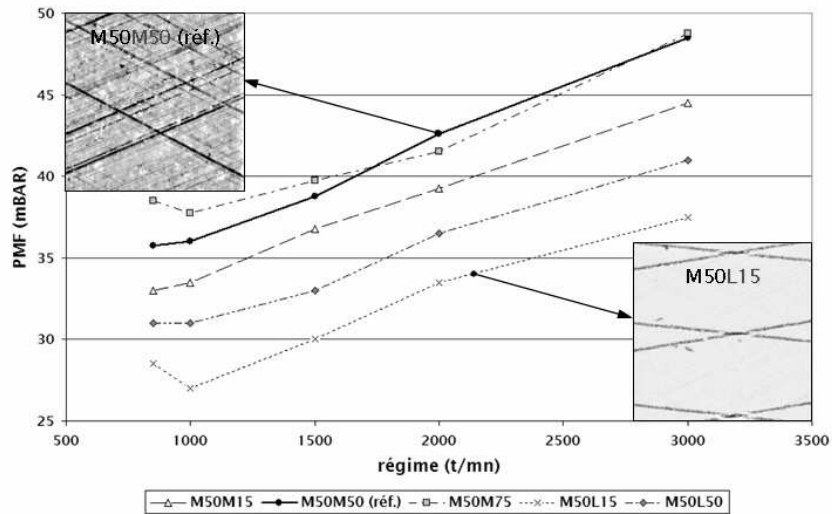


FIG. 2.3: Puissance moyenne de frottement (PMF) en fonction de du régime (en tours par minute) pour différentes surfaces. La surface de référence est désignée par le code M50M50 ; la meilleure surface est la M50L15 (usinage mécanique des plateaux avec un angle de 50° ; usinage laser des stries avec un angle de 15°)

en termes de réduction des pertes par frottement, mais n'a pas permis de comprendre les mécanismes qui étaient en jeu. Il a donc été décidé de lancer une deuxième phase d'étude. Cette fois-ci, il s'agissait pour nous d'aller plus loin dans l'étude du contact 3D. Avec notre approche de traiteurs d'images, nous nous sommes donc lancés dans la simulation d'écoulement entre deux surfaces.

2.2 Simulation du contact hydrodynamique

Cette deuxième phase de l'étude a été financée par l'ADEME. Un nouvel industriel, TOTAL, s'est joint au consortium en remplacement d'ESSO. Un doctorant, Costin Alin Caciui, a été recruté pour effectuer ce travail sous ma direction et celle de Dominique Jeulin. Nous ne ferons ici que survoler cette étude qui, bien qu'extrêmement intéressante et originale, nous fait presque sortir du domaine du traitement d'images. Une description complète de ces travaux est disponible dans le mémoire de thèse [9] et les publications associées [13, 10, 69, 12, 11], dont une est jointe en annexe B.

Nous avons commencé par modéliser le contact entre le segment coupe-feu et la chemise. Vue l'échelle de l'étude, un modèle de contact lubrifié entre une surface rugueuse et une surface plane a été choisi. Nous avons vu plus haut que, d'après les spécialistes, l'essentiel des pertes par frottement se produit en régime hydrodynamique. Nous avons donc supposé qu'il n'y avait pas de contact solide-solide entre les deux surfaces. La surface rugueuse, c'est à dire la chemise, est supposée immobile, et la surface lisse animée d'une vitesse \vec{V} , ce qui nous a permis d'avoir un système stationnaire. L'axe (Ox) est choisi colinéaire à \vec{V} , et l'axe (Oy) parallèle à la surface. Le système est illustré par la figure 2.4. Le domaine intérieur, compris entre les deux surfaces, est appelé Ω .

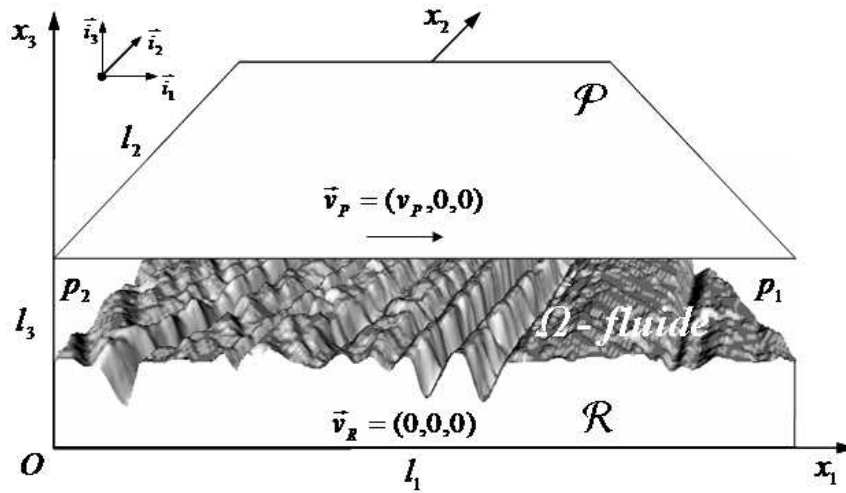


FIG. 2.4: Modélisation du contact entre la chemise et le segment coupe-feu. Les dimensions de la chemise sont $1,5 \times 1,5mm$ et l'écart entre les surfaces est de quelques micromètres

Nous avons adopté les hypothèses suivantes concernant le liquide lubrifiant, et son interaction avec le système :

- le fluide est Newtonien ;
- le fluide est incompressible ;
- nous supposons que la température est partout la même, et constante dans le temps, afin de négliger tout effet thermodynamique. Nous supposons en particulier que la viscosité η et la masse volumique ρ du liquide sont constantes ;

- le fluide adhère aux surfaces solides ;
- les forces à distance sont négligées ;
- finalement, nous supposons qu’aucune réaction chimique ne se produit au sein du système.

Nous obtenons donc, en tout point du domaine Ω , le système d’équations de Navier-Stokes :

$$\begin{cases} (\vec{v} \cdot \nabla) \vec{v} + \frac{1}{\rho} \nabla p - \nu \Delta \vec{v} = \mathbf{0} \\ \nabla \cdot \vec{v} = 0 \end{cases} \quad (2.1)$$

où ρ est la masse volumique du fluide, ν la viscosité cinématique, v la vitesse et p la pression.

Plusieurs conditions aux limites ont été proposées et mises en œuvre pour les frontières de Ω ne correspondant pas à des surfaces solides, à travers desquelles un écoulement peut se produire. On pourra consulter l’article donné en annexe pour une description complète.

Ce système d’équations, non linéaire, est classique en mécanique des fluides. Étant donnée sa complexité, une solution analytique n’est pas envisageable. Nous avons décidé de discrétiser le système par une approche par différences finies. Les éléments finis risquaient de se révéler trop complexes à mettre en œuvre à cause de la rugosité à modéliser. Au contraire, une approche par différences finies nous permettait de rester en terrain connu, car le résultat de la discrétisation constituait, finalement une image à trois dimensions.

Plusieurs approches pour la résolution du système ont été envisagées. Finalement, deux ont été retenues : une méthode de relaxation, et la méthode de projection de Chorin-Temam, cette dernière permettant de prendre en compte des conditions de cavitation. Par ailleurs, la résolution des équations de Reynolds (obtenues à partir des équations de Navier-Stokes grâce à des hypothèses simplificatrices) par une méthode de relaxation a aussi été implantée.

Étant donnée la complexité du modèle, plusieurs procédures de validation ont été entreprises. Premièrement, on a vérifié numériquement que les modèles étaient en accord avec des résultats connus, comme ceux donnés dans des configurations simples par des solutions analytiques. Deuxièmement, on a confronté les résultats des simulations avec des résultats expérimentaux obtenus grâce d’une part à un tribomètre, et d’autre part à la machine SPC. Le première

étape a été un succès complet. La deuxième étape, beaucoup plus ambitieuse vue la complexité des dispositifs expérimentaux (tribomètre et machine SPC) a donné des résultats satisfaisants.

Par exemple, le graphique de la figure 2.5 illustre la bonne corrélation entre la prédiction de la puissance dissipée par unité de volume, obtenue par simulation, et la puissance moyenne de frottement, mesurée sur les chemises correspondantes sur la machine SPC, usinées pendant le projet PREDIT.

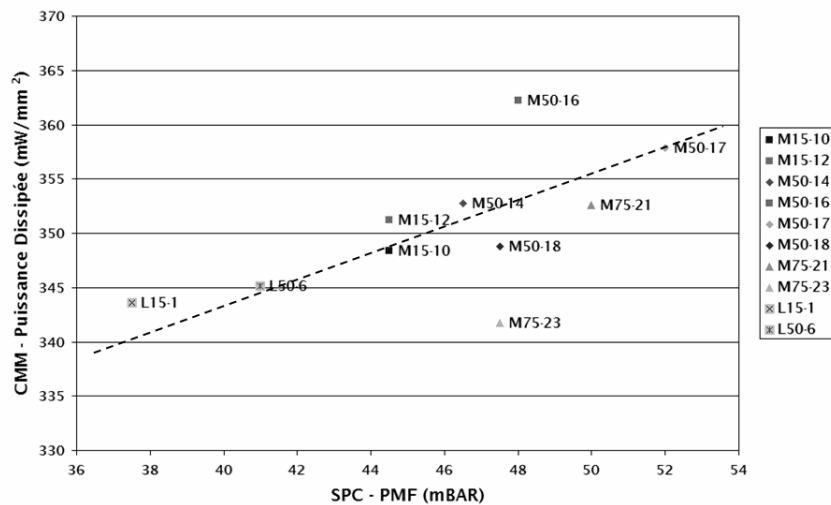


FIG. 2.5: Confrontation des simulations avec les mesures PREDIT SPC (3000 t/mn)

2.3 Exploitation des résultats

Cet outil de simulation nous a servi pour comparer différents types de surfaces entre elles, afin de les classer en fonction de leurs performances. Nous avons ainsi pu non seulement évaluer des surfaces réelles, mais aussi des surfaces de synthèse. Tout naturellement, cette démarche nous a conduits à la construction de plusieurs outils d'optimisation [13].

Il faut remarquer que le modèle proposé permet d'obtenir, en tout point du volume Ω discrétisé, une valeur des différentes grandeurs physiques concernées. On peut ainsi, par exemple, connaître avec précision les régions de sur-pression ou identifier celles où un phénomène de cavitation (dû à une dépression) risque

de se produire. De même, on peut repérer les zones qui constituent les principales sources de dissipations d'énergie. Cette information peut servir à améliorer les outils d'optimisation des surfaces évoqués plus haut. En effet, si nous connaissons les régions du volume où se produisent le plus de pertes d'énergie, nous pouvons alors modifier la surface topographique dans la zone correspondante dans l'espoir de les réduire.

2.4 Conclusion

Il est important de souligner l'originalité de ce travail, situé à l'intersection entre le traitement d'images et la mécanique des fluides. A notre connaissance, il s'agit là de la première étude d'un contact hydrodynamique tridimensionnel prenant en compte une vraie surface texturée. Les tribologistes qui ont eu à évaluer ce travail l'ont très favorablement reçu.

Le traitement d'images fait souvent appel à des modèles physiques. Par exemple, les méthodes de filtrage par diffusion ont comme point de départ de l'équation de diffusion de la chaleur, le modèle d'Ising est utilisé pour restaurer des images [35], et des méthodes de recalage sont inspirées par la mécanique des fluides [61].

La fertilisation dans l'autre sens est, à notre avis, moins courante. Nous pensons en effet que notre culture de traiteurs d'images a largement inspiré les choix qui ont été adoptés. En particulier, la décision de travailler avec des différences finies découle de cette approche, qui nous a poussés à sortir des sentiers battus.

Finalement, ce travail montre comment, en passant d'une analyse 2D à une analyse 3D, il a été possible d'apporter une solution satisfaisante au problème posé.

When I could find voice, I shrieked loud in agony, "Either this is madness or it is Hell." "It is neither, calmly replied the voice of the Sphere, "it is Knowledge; it is Three Dimensions : open your eye once again and try to look steadily."

Edwin A. Abbott, Flatland (1884)

3

Quelques problèmes liés à la troisième dimension

Il y a quelques années, je me trouvais à Lausanne, et disposais de quelques minutes de liberté avant un rendez-vous. J'en ai donc profité pour faire une petite promenade. J'avais très peu de temps, j'ai donc choisi l'option la plus simple : faire le tour du pâté de maisons. Quoi de plus naturel que de toujours tourner à gauche pour revenir, inéluctablement, au point de départ ? Mais assez vite, je me suis aperçu de l'échec de cette stratégie. En effet, Lausanne, installée au bord du lac Léman, est une ville à la géographie accidentée. Le trottoir que j'ai emprunté m'a rapidement mené dans un tunnel, et finalement je me suis retrouvé bien plus bas que mon point de départ. J'ai finalement rebroussé chemin, pour éviter d'être en retard. Serais-je revenu à mon point de départ en persévérant ? Cette petite anecdote illustre le fait que le passage de deux à trois dimensions rend certains problèmes plus complexes. Nous verrons dans ce chapitre quelques autres exemples tirés de problèmes pratiques en traitement d'images.

Commençons par décrire deux études sur des images 3D auxquelles j'ai pris part. Bien que portant sur des sujets différents, dans les deux cas nous retrouverons le même genre de problème. Nous verrons ensuite quelles solutions nous avons apportées.

3.1 Premier cas d'étude : visualisation de structures anatomiques

La première étude a fait l'objet du travail de thèse de Mathilde Boehm [2]. Le sujet portait sur la visualisation de données médicales obtenues par tomodensitométrie (ou scanner par rayons X), et faisait l'objet d'une collaboration avec la société General Electric.

Un problème sérieux se pose lorsqu'on passe à trois dimensions : celui de la visualisation. Notre système visuel a évolué pour fonctionner dans un monde tridimensionnel. Cependant, il traite essentiellement des objets opaques ou transparents, beaucoup plus rarement des objets translucides. Ceci veut dire que nous percevons des surfaces. Or, en pratique, lorsque nous faisons du traitement d'images en 3D nous avons le plus souvent à traiter des données à niveaux de gris, voire plus complexes. L'affichage de ces images devient alors un problème délicat, si nous voulons évaluer visuellement le résultat d'un traitement, ou tout simplement comprendre l'image.

Les moyens d'affichage, tels que l'écran ou le papier imprimé, restent pour l'essentiel bidimensionnels. On utilise donc des techniques de rendu 3D pour produire des images 2D qui nous donnent l'impression de voir une image 3D. Remarquons au passage que l'analogie topographique, souvent appliquée aux images 2D à niveaux de gris afin de mieux comprendre les structures, perd de son intérêt en 3D. En effet, d'une part nous n'arrivons plus à nous représenter visuellement ces structures, qui deviennent en fait des surfaces dans un espace à quatre dimensions, et d'autre part nous rencontrons des structures propres au 3D qu'on ne retrouve pas en topographie.

Notre étude concernait justement l'amélioration du rendu tridimensionnel au moyen d'outils de traitement d'images. Deux approches ont été envisagées : le filtrage, et l'utilisation d'une segmentation hiérarchique. Nous ne décrivons ici

que la première. Nous avons construit une image de synthèse (appelée *fantôme*) comportant différents types de structures, dans le but de reproduire, dans un cas bien contrôlé, les types de structures que l'on peut rencontrer sur une image médicale 3D. Pour rendre plus réaliste ce modèle, nous lui avons rajouté du bruit ou des textures extraites de vraies images médicales. Les images ainsi texturées devenaient difficiles à visualiser, comme le montre la figure 3.1.

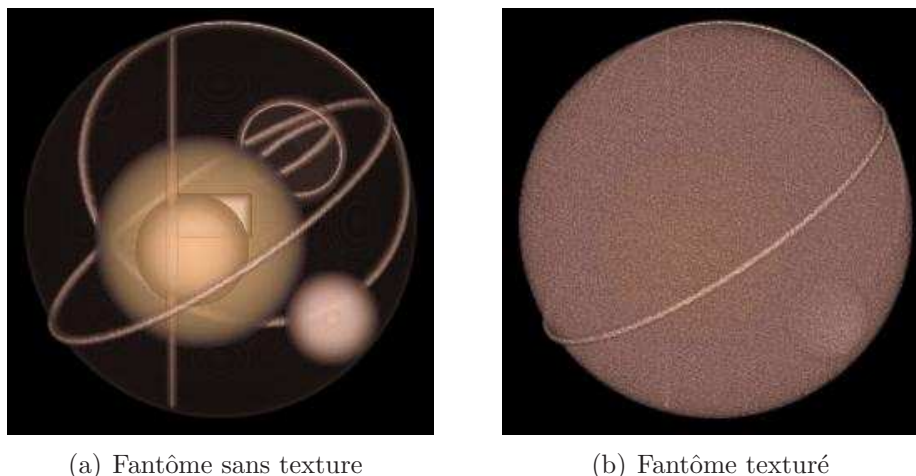


FIG. 3.1: Fantôme construit par Mathilde Boehm. Dimensions de l'image : $256 \times 256 \times 256$. La méthode de rendu 3D utilisée est une implantation faite par Mathilde Boehm de celle utilisée par General Electric dans ses stations de visualisation, méthode qui elle-même est proche de celle proposée par Levoy [50, 51], avec les optimisations proposées par Levoy et Lacroute [52, 44].

En effet, les techniques de visualisation médicales partent du principe que le spécialiste est intéressé par certaines surfaces dans le volume. Ces surfaces correspondent à la frontière des structures anatomiques qui l'intéressent. Les méthodes de visualisation tentent par conséquent, d'une façon ou d'une autre, de faire ressortir les surfaces. Les méthodes actuellement utilisées pour la visualisation de volumes font l'impasse sur le calcul explicite de ces surfaces, et tentent de les faire ressortir en utilisant différents types de gradients. Plusieurs approches s'inspirant de ce principe ont été publiées en 1988. Celle de Levoy [50, 51] a rencontré le plus de succès ; baptisée "ray casting" en Anglais, elle peut être vue comme une adaptation de la méthode dite de lancer de rayons, ou "ray casting", utilisée en synthèse d'images. Or, lorsque l'image présente

des niveaux importants de bruit ou de texture, le gradient parasite résultant a tendance à cacher le gradient des surfaces.

Dans un premier temps, nous avons appliqué des filtres classiques de la morphologie mathématique, connus par leur capacité à préserver les frontières, tout en effaçant le bruit. Par exemple, nous avons employé des ouvertures ou des fermetures par reconstruction ou, plus généralement, des nivellements. Les résultats ont été très décevants. Leur rendu volumique produisait quelque chose de très proche de l'image texturée originale, telle que présentée sur la figure 3.1(b). En effet, le bruit ou la texture se trouvant à l'intérieur des structures étaient bien réduits, mais quand ces structures parasites touchaient les bords des objets, elles étaient naturellement préservées par les nivellements. Le rendu volumique, par conséquent, était très peu satisfaisant.

3.2 Deuxième cas d'étude : segmentation de cubes sismiques

Entre 2004 et 2006, Timothée Faucon a effectué son travail de doctorat, sous ma direction, sur la segmentation de cubes sismiques [32], dans le cadre d'une collaboration avec la société ERM.S. Une étape essentielle de son travail était la détection des surfaces, appelées *horizons*, correspondant à des interfaces géologiques. Elle se présentent comme des nappes fines, de niveau de gris élevé, et globalement horizontales. L'observation sur des coupes 2D verticales donnait l'impression que ces horizons étaient bien continus, et qu'ils devraient, par conséquent, être faciles à détecter grâce à une ligne de partage des eaux. Les premiers essais se sont cependant révélés très décevants. Mis à part quelques uns, bien contrastés, les horizons ressortaient très mal au niveau de la segmentation. Une observation plus minutieuse nous a en effet révélé que la plupart des surfaces sont trouées du fait d'une part du bruit introduit par l'acquisition, et d'autre part des structures géologiques, telles que les failles. La nature globale de la ligne de partage des eaux empêchait par conséquent de les extraire proprement.

Un point qu'il est indispensable de soulever pour comprendre le cadre de cette étude est celui concernant la taille des données. Les cubes sismiques peuvent

comporter facilement des milliards de points. Les tests que nous avons réalisés se sont limités à des *petits* cubes de seulement quelques centaines de millions de pixels. Les méthodes de calcul retenues doivent par conséquent être très performantes, aussi bien en termes de complexité algorithmique, que de besoin en mémoire.

3.3 Leçons tirées des cas d'étude

Intuitivement, les deux problèmes que nous venons d'évoquer sont dus à la même raison : en trois dimensions, le domaine offre plus de possibilités à l'apparition de nouvelles structures. Dans le premier cas, des structures, éventuellement longues, viennent se coller sur les surfaces d'intérêt. Dans le deuxième cas, vue la taille des surfaces, il est fort probable que quelque part un trou y apparaisse.

Un modèle extrêmement simple peut nous aider à comprendre ce qui arrive. Plaçons-nous dans le domaine $D = \{0, \dots, N - 1\}^d$ de \mathbb{Z}^d . Notons (x_1, \dots, x_d) les coordonnées d'un point M de D . Alors l'ensemble des points vérifiant $x_1 = N_1$, avec $0 < N_1 < N - 1$ est un hyper-plan discret H de D . Cet hyper-plan contient N^{d-1} points. Considérons l'ensemble aléatoire E de D donné par la loi uniforme indépendante $P(M \in E) = p$ ($0 \leq p \leq 1$). E peut servir, par exemple, pour modéliser un bruit de type impulsif. Alors la probabilité que l'intersection entre E et H soit nulle est tout simplement $P(E \cap H = \emptyset) = (1 - p)^{N^{d-1}}$.

Il apparaît alors clairement que d , toutes choses égales par ailleurs, joue un rôle décisif. Le modèle est certes simpliste. En particulier, il n'est pas facile de justifier de l'invariance de p par rapport à d , mais il décrit assez bien ce qu'on observe en pratique. Par exemple, si $N = 10$ et $p = 10^{-3}$, nous avons :

d	$P(E \cap H = \emptyset)$
1	0.99
2	0.90
3	0.37

Nous constatons clairement que la probabilité d'avoir une surface non trouée diminue de façon importante avec l'augmentation du nombre de dimensions.

Notons aussi que la ligne de partage des eaux est classiquement appliquée au gradient d'une image (mais pas toujours ; dans l'étude sur la segmentation de cubes sismiques elle était appliquée directement sur les images originales). Or, le calcul de ce gradient tend à effacer les petits trous. Nous devrions donc nous intéresser à la création, par le bruit, de structures plus grandes.

Plaçons-nous cette fois-ci dans l'espace \mathbb{Z}^d , muni de la distance L_1 , notée d_{L_1} . Pour tout couple de points (M, P) , de coordonnées respectives (x_1, \dots, x_d) et (y_1, \dots, y_d) , nous avons :

$$d_{L_1}(M, P) = |x_1 - y_1| + \dots + |x_d - y_d|.$$

Définissons la relation de voisinage \mathcal{V} de la façon suivante : deux points M et P de D sont voisins si et seulement si $d_{L_1} \leq 1$. En deux dimensions, cette relation de voisinage correspond à la 4-connexité, et en trois dimensions à la 6-connexité. A partir de cette relation de voisinage, de façon classique, on peut définir les composantes connexes d'un ensemble. Nous pouvons alors nous intéresser à la taille des composantes connexes de l'ensemble aléatoire E . Ce problème fait l'objet de la théorie de la percolation [37]. On sait que si p est plus grand qu'une *probabilité critique* $p_c(d)$, alors l'ensemble E comportera une unique composante connexe de taille infinie, et qu'au contraire si $p < p_c(d)$ les composantes connexes seront toutes de taille finie.

On ne connaît pas la valeur de $p_c(d)$ pour $d > 1$. On peut en obtenir facilement une valeur approchée par simulation. Ainsi, pour $d = 2$ la probabilité critique vaut environ 0,59. Par ailleurs, la probabilité critique est une fonction strictement décroissante de d , et elle converge vers $(2d)^{-1}$ quand d tend vers l'infini.

Ceci signifie, dans notre contexte, que les structures générées par le bruit auront une portée d'autant plus grande que le nombre de dimensions est élevé. On comprend alors mieux pourquoi, dans le cas de la visualisation de structures anatomiques, les méthodes de filtrage qui marchaient relativement bien en 2D se sont révélées insuffisantes en 3D.

3.4 Des solutions au cas par cas

Pour le premier problème, concernant la visualisation de structures anatomiques, la solution la plus intéressante s'est révélée être le filtrage. D'ailleurs, cette étude a été à l'origine du concept d'*amibe morphologique*, que nous avons proposé avec Lerallut, Meyer et Boehm [48, 47, 49]. Il s'agit d'une mise en œuvre des fonctions structurantes, concept théorique introduit par Serra en 1982 [65], et qui n'avait pratiquement pas connu d'application pratique.

L'idée de base est d'adapter la forme de l'élément structurant en chaque point, en prenant en compte son voisinage. L'élément structurant aura tendance à s'étendre le long des zones plates et, au contraire, à éviter de franchir les barrières représentées par des gradients forts.

Pour mettre en œuvre cette approche, on se donne une distance d_v dans l'espace des valeurs. Par exemple, en niveaux de gris, on prend simplement la valeur absolue de la différence. Ensuite, l'image concernée I étant munie d'une relation de voisinage, on se donne une distance entre points voisins M et P :

$$d_s(M, P) = 1 + \lambda d_v(I(M), I(P))$$

où λ est un réel positif. La distance d_s entre deux points quelconques est obtenue en prenant le minimum des longueurs des chemins les reliant (et en considérant qu'un chemin constitué d'un seul point est de longueur nulle). L'amibe $B_{\lambda,r}(M)$ de paramètre λ et de rayon r , centrée au point M , est l'ensemble des points P dont la distance à M est inférieure à r .

Cette approche est décrite en détail dans [49], qu'on trouvera en annexe C de ce mémoire.

Revenons à la visualisation 3D d'images médicales. En appliquant cette approche à l'image 3D texturée, nous obtenons l'image de la figure 3.2, à comparer avec l'image de la figure 3.1(b).

Quant au deuxième problème, avec Timothée Faucon nous avons aussi commencé par tenter une approche par filtrage, mais finalement nous avons abandonné cette voie, car d'une part la qualité des images était très variable, ce qui imposait de revoir le paramétrage au cas par cas, et d'autre part il semblait important de préserver certaines structures fines, comme l'espace entre les ho-

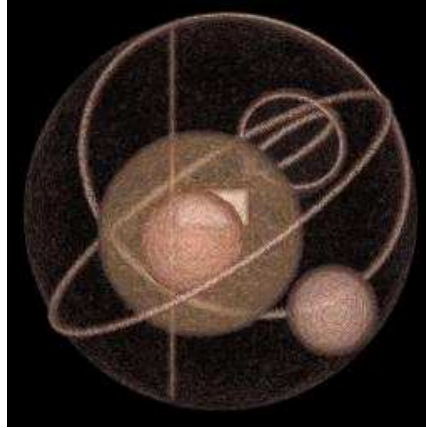


FIG. 3.2: Fantôme texturé, après application d'un filtrage par amibe morphologique

rizons, et les failles, en sachant que, vue la taille des données, nous ne pouvions pas nous permettre d'adopter des filtres trop gourmands en temps de calcul. Nous avons finalement résolu cette difficulté grâce à une approche par cylindres glissants [30, 31, 32].

Dans le cas présent, nous ne pouvons faire abstraction des dimensions de l'image I . Nous considérons donc une image de taille $N_x \times N_y \times N_z$:

$$I : D = \{0, \dots, N_x - 1\} \times \{0, \dots, N_y - 1\} \times \{0, \dots, N_z - 1\} \longrightarrow \mathbb{Z} \quad (3.1)$$

$$(x, y, z) \longmapsto I(x, y, z) \quad (3.2)$$

On définit un cylindre $C(x_0, y_0, r)$ à base carrée de taille $x_0 \times y_0$, de hauteur N_z , sous-domaine de l'ensemble de définition D de l'image I :

$$C(x_0, y_0, r) = \{(x, y, z) \in D / x_0 \leq x < x_0 + r, y_0 \leq y < y_0 + r\}.$$

Notons que dans les images sismiques, le plan défini par les axes (Ox) et (Oy) correspond au plan horizontal. Or, nous supposons que les horizons géologiques, que nous cherchons à détecter, sont proches de l'horizontale. Donc en prenant des valeurs petites de x_0 et y_0 , nous obtenons des cylindres allongés orthogonaux aux horizons sismiques. Dans chacun de ces cylindres, nous appliquons une

ligne de partage des eaux, avec un critère de volume, en ne retenant que les N régions les plus représentatives. Les contours de ces régions, constituant pour l'essentiel des nappes globalement horizontales, sont des morceaux des horizons recherchés.

Pour chaque position de cylindre, nous obtenons donc un ensemble de frontières, codé en tant qu'image binaire, que nous ajoutons dans une image accumulateur. Nous pouvons prendre toutes les positions possibles, ou adopter un *pas* de déplacement pour le cylindre, afin d'alléger le calcul. Si un point appartient à plusieurs frontières, correspondant chacune à la segmentation d'un cylindre, sa valeur dans l'image accumulateur sera alors importante. La procédure est illustrée par la figure 3.3.

Cette approche permet de rendre le calcul des contours plus local. Si un horizon géologique est localement troué, les cylindres qui ne touchent pas le trou produiront une bonne segmentation pour l'horizon concerné. Par ailleurs, cette approche permet de réduire considérablement les besoins en mémoire.

Le travail ne s'est pas arrêté là. Il a fallu ensuite binariser l'image accumulateur, et analyser le résultat pour séparer les différents horizons [31]. En effet, la géologie ne garantit nullement la séparation des différents horizons, et de plus des erreurs d'acquisition ou de traitement peuvent provoquer l'imbrication de certaines surfaces. Cette étape de séparation de surfaces, a été particulièrement riche en enseignements. Nous avons utilisé pour cela des outils de la topologie discrète, en particulier les méthodes de Malandain et al. [54]. Ce genre d'étude illustre bien les difficultés de nature topologique introduites par la troisième dimension.

Le travail de thèse de Timothée Faucon a été jugé comme très satisfaisant par les industriels impliqués dans la collaboration, et a mené au développement d'une nouvelle activité au sein de la société ERM'S.

Dans la section suivante, nous verrons que dans le cadre de la segmentation interactive 3D on rencontre des problèmes similaires.

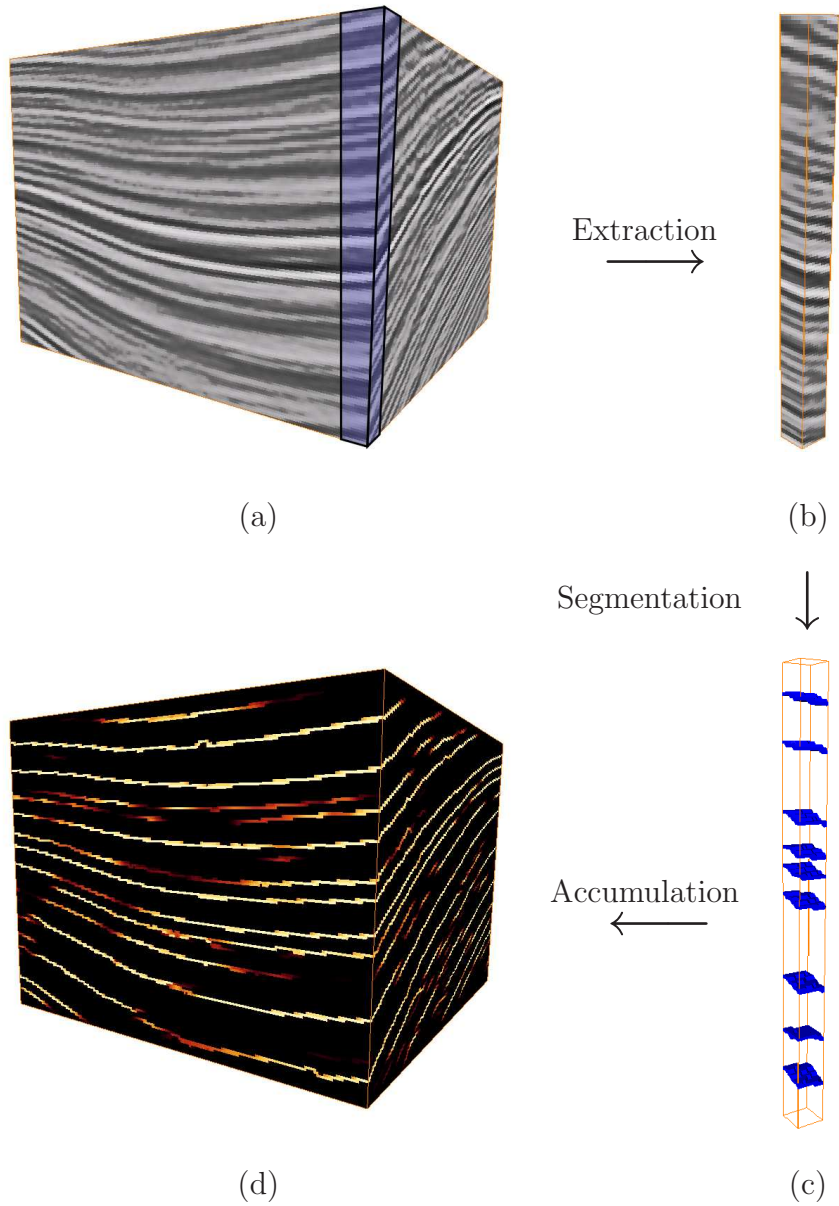


FIG. 3.3: Chaîne de traitement de la segmentation par cylindres. On extrait de l'image d'amplitude (a) un cylindre (b) que l'on segmente en autant de régions que l'on veut (c). On accumule ensuite les frontières des régions ainsi obtenues dans l'image accumulateur (d). Extrait de [32].

3.5 Segmentation interactive 3d

Nous avons vu, dans la section précédente, que le traitement d'images médicales tridimensionnelles est souvent rendu difficile par la présence de bruit, ou de textures. En outre, certains organes peuvent ne pas être bien contrastés. Par exemple, sur une image obtenue au moyen d'un scanner X, le foie et l'estomac apparaissent avec des niveaux de gris très similaires, et par conséquent il est délicat de les séparer.

Or, de nombreuses applications médicales requièrent une segmentation des structures anatomiques. Par exemple, dans la planification en radiothérapie, il est indispensable de contourer au préalable aussi bien la tumeur que les organes à protéger, afin de choisir la trajectoire des faisceaux irradiants, et de calculer la dose de radiation reçue par chaque organe.

Il existe déjà des solutions utilisées au niveau clinique pour segmenter certaines structures, telles que les poumons, automatiquement. Cependant, il semble difficile d'adopter la même approche pour des organes dont les contours sont moins bien contrastés, ou dont la forme n'est pas normale, et presque impossible de le faire avec les tumeurs. Quoi qu'il en soit, aujourd'hui la grande majorité des radiothérapeutes continuent à utiliser des méthodes manuelles, en traçant à la main les contours d'intérêt sur un nombre important de coupes 2D, au point que le détournement d'un organe complexe comme le foie peut prendre jusqu'à une demi-heure.

Suite au succès des techniques de segmentation semi-automatique dans le domaine de la vidéo [55], nous avons décidé d'appliquer la même approche à la segmentation d'images médicales 3D. Très tôt, nous avons commencé à collaborer avec l'Institut Gustave Roussy, et en particulier avec François Bidauld, radiologue du service d'imagerie médicale. Faute de moyens, au début une partie du travail a été effectué par des stagiaires : Sébastien Aymard, Marc Huertas, et Mohamed Mzali. Les premiers résultats se sont révélés prometteurs, ce qui nous a permis d'obtenir le financement nécessaire pour une bourse de thèse, et d'embaucher ainsi Jean Stawiaski en Octobre 2005.

Comme nous l'avons déjà évoqué à maintes reprises, le passage du 2D au 3D pose souvent des problèmes de visualisation. Dans ce projet, nous avons donc investi un effort important dans le développement d'une plateforme de

3 Quelques problèmes liés à la troisième dimension

visualisation. Nous avons bâti une interface classique, divisée en quatre vues de l'image 3D : un rendu volumique, et trois coupes orthogonales (cf. figure 3.4). Sur cette plateforme, nous avons implanté les outils de segmentation interactive.

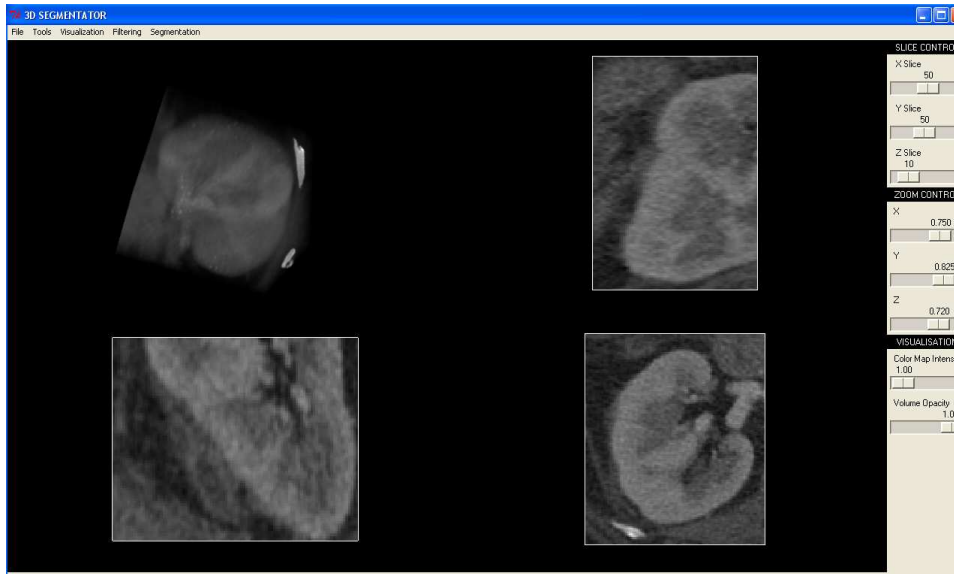


FIG. 3.4: Capture d'écran de l'interface 3D

Nous avons commencé par la mise en place de la segmentation hiérarchique, et les moyens d'interagir avec elle, notamment la navigation entre les différents niveaux de la hiérarchie. On commence par calculer une ligne de partage des eaux hiérarchique, en utilisant un critère d'extinction volumique [58]. Ensuite, l'utilisateur peut choisir le niveau de résolution qu'il souhaite, grâce à un ascenseur, et choisir les régions qui vont constituer le masque de l'objet d'intérêt. Cette méthode donne globalement de bons résultats. Lorsque les structures sont relativement bien contrastées, on arrive souvent à un niveau de la hiérarchie dans lequel l'objet d'intérêt est bien représenté par une seule région. Dans les cas plus complexes, il est nécessaire de procéder en plusieurs étapes. Les fuites, produites par une zone peu contrastée de la frontière, demandent souvent de travailler à un niveau bas de la hiérarchie. En pratique, nous avons constaté que cette approche demande un temps d'apprentissage assez long de la part d'un utilisateur qui ne serait pas familier avec le traitement d'images.

La deuxième approche implantée a été la segmentation par marqueurs [60, 59]. L'utilisateur doit indiquer, par des traits de différentes couleurs tracés sur

les coupes orthogonales, les objets qu'il cherche à extraire. Ainsi, dans le cas le plus simple, il faut marquer l'objet d'intérêt et le fond. La procédure est illustrée par les images de la figure 3.5. Si le résultat n'est pas satisfaisant, on a la possibilité de rajouter des marqueurs pour colmater, par exemple, une fuite. Cette approche s'est révélée tout aussi efficace que la première, et présente l'avantage d'être plus intuitive pour un non spécialiste, même si le marquage du fond n'est pas une étape naturelle.

Il faut noter que les deux approches requièrent un seul calcul de la hiérarchie de partitions, effectué en un temps dépendant linéairement du nombre de pixels de l'image grâce à l'algorithme de ligne de partage des eaux à base de files d'attente hiérarchiques [57].

L'ensemble des outils ont été implantés dans un démonstrateur, qui a été mis à disposition de spécialistes de l'Institut Gustave Roussy, en vue d'une première validation. Par ailleurs, le même outil peut être utilisé pour segmenter des images tridimensionnelles provenant d'autres domaines, comme les microtomographies par rayons X utilisées dans le domaine des matériaux. Notons pour finir qu'à partir de ces mêmes outils, et toujours en collaboration avec l'Institut Gustave Roussy, mais cette fois-ci avec le service de médecine nucléaire, nous avons mené une étude sur la segmentation d'images acquises par tomographie par émission de positons (TEP) [68].

3.6 Conclusion

La définition de la ligne de partage des eaux est indépendante du nombre de dimensions. Lorsqu'elle est calculée à l'aide d'un algorithmes à base de files d'attente hiérarchiques sa complexité est linéaire en fonction du nombre de pixels de l'image, ce qui la classe parmi les méthodes de segmentation les plus rapides, et qui fait d'elle une des méthodes de choix pour le traitement d'images à trois dimensions, ou davantage.

Il reste néanmoins que les méthodes de segmentation basées sur la ligne de partage des eaux sont assez sensibles aux problèmes de fuites, et d'autant plus lorsque le nombre de dimensions augmente. Ce genre de phénomène peut être modélisé de façon intéressante grâce à la théorie de la percolation. Ceci est

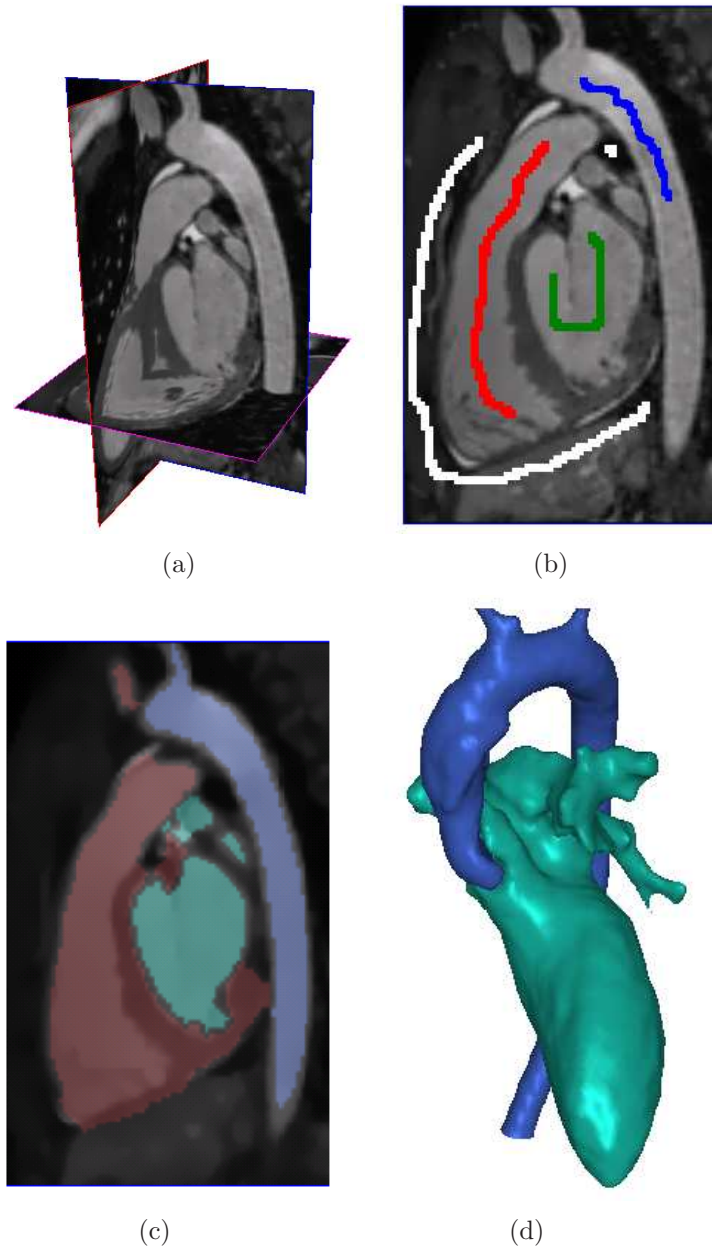


FIG. 3.5: Exemple de segmentation par marqueurs. (a) Image originale, vue en coupes. (b) Marqueurs dessinés sur une coupe. (c) Segmentation résultante, superposée sur l'image originale, vue sur une coupe. (d) Visualisation 3D de deux régions

gênant vis à vis de certaines applications, comme nous venons de le voir. Bien qu'il soit possible de trouver différents types de solutions, comme un filtrage adapté, en fonction justement des applications, nous avons décidé d'investiguer la possibilité de mettre au point une solution plus fondamentale, afin de rendre les méthodes à base de ligne de partage des eaux plus robustes. Ce travail est actuellement en cours. Nous en dirons un mot dans le chapitre 5. Nous verrons aussi que l'approche que nous y proposons permet d'alléger la charge de calcul, facteur souvent limitant en imagerie tridimensionnelle.

Avant de finir, notons que l'imagerie 3D, en apportant plus d'information, permet de supprimer certains problèmes propres à l'imagerie 2D projective. En effet, sur les images 3D nous n'avons plus de déformations introduites par une projection, ni d'occultation de certains objets par d'autres. Nous verrons dans le chapitre suivant que les séquences d'images 2D, qui peuvent aussi être considérées comme des images 3D, posent de ce fait des problèmes différents.

There are really four dimensions, three which we call the three planes of Space, and a fourth, Time. There is, however, a tendency to draw an unreal distinction between the former three dimensions and the latter, because it happens that our consciousness moves intermittently in one direction along the latter from the beginning to the end of our lives.'

H. G. Wells, *The Time Machine* (1895)

4

Le temps : une dimension comme les autres ?

Les premiers traitements portant sur les séquences d'images numériques se faisaient image par image, indépendamment les unes des autres. Ce premier pas, tout naturel étant donné ce qu'on savait déjà faire sur les images fixes, a été assez rapidement suivi par d'autres pas, plus ambitieux, visant à exploiter la redondance temporelle de ces séquences. On a donc adopté des approches dans lesquelles le traitement de l'image courante exploitait l'information de l'image précédente, puis d'autres images voisines. Par exemple, dans le domaine de la compression vidéo, cette approche a permis d'atteindre des taux de compression largement supérieurs à ceux des images fixes. Ceci a été possible parce que, *normalement*, deux images consécutives dans une séquence sont très similaires. Les différences sont dues, pour l'essentiel, au mouvement des objets par rapport à la caméra.

Afin d'exploiter au mieux cette redondance temporelle, et aussi pour analyser le mouvement des objets composant la scène, on a lancé des recherches sur

l'estimation du mouvement dans les séquences d'images. Vue l'importance des enjeux, et les difficultés rencontrées, ce domaine de recherche s'est révélé très actif. Cependant, ces méthodes restent en général coûteuses en temps de calcul. Il faut donc souvent trouver un compromis entre la précision du champ de vecteurs calculé, et la vitesse de traitement.

Plus récemment, des tentatives pour traiter un nombre important d'images simultanément, en les considérant comme une seule image à trois dimensions, ont été entreprises, avec des résultats intéressants [70]. La puissance croissante des moyens de calcul dont nous disposons autorise en effet désormais ce genre d'approche. Tout se passe alors comme si le temps était une troisième dimension spatiale.

Dans la suite de ce chapitre nous allons seulement considérer des séquences d'images bidimensionnelles, telles que les films cinématographiques ou les séquences vidéo. Nous nous bornerons aux images à niveaux de gris. Par conséquent, nous modéliserons une telle séquence comme une fonction de \mathbb{Z}^3 dans \mathbb{Z} :

$$I : \mathbb{Z}^3 \longrightarrow \mathbb{Z} \tag{4.1}$$

$$(x, y, t) \longmapsto I(x, y, t) = I_t(x, y). \tag{4.2}$$

A travers deux études que nous avons réalisées dans ce domaines, nous allons essayer d'apporter des réponses à la question suivante :

Dans quelle mesure peut-on traiter la dimension temporelle comme les dimensions spatiales ?

Cette question n'a de sens que par rapport à certaines applications. Si on cherche à mesurer, par exemple, la vitesse des objets de la scène, alors elle ne se pose pas, car il est indispensable dans ce cas de calculer le déplacement explicitement entre images consécutives. Cependant, pour de nombreuses applications, et en particulier pour celles qui requièrent une segmentation spatio-temporelle des objets de la séquence, elle devient intéressante.

4.1 Lorsque deux dimensions ne suffisent pas : le cas de la restauration de films

En 1994, avec Jean Serra, nous avons commencé à travailler sur la restauration automatique de films cinématographiques, qui était à ce moment-là un domaine pratiquement inexploré. Ces recherches ont démarré grâce au projet européen ESPRIT-NOBLESSE [16, 17, 26, 18, 20, 19, 21], qui a donné lieu à des collaborations en particulier avec des équipes de traitement d'images de l'université de Strathclyde, de l'université de Trieste, et de l'université technologique de Tampere [25, 40, 41]. Dans un deuxième temps, et au vu des résultats, cette étude a mené à un contrat de valorisation avec la société Kodak [22].

Au cours de cette étude, nous avons abordé la correction des défauts suivants :

- correction du pompage (variation anormale de la luminosité des images au cours du temps) ;
- correction des vibrations, ou *jitter* en Anglais ;
- restauration des rayures verticales ; et
- restauration des défauts locaux, provoqués par des poussières ou des moisissures, par exemple.

Le quatrième thème, la restauration des défauts locaux, est celui sur lequel nous avons porté l'essentiel de notre effort de recherche. C'est d'ailleurs celui-là qui a fait l'objet du contrat de collaboration avec la société Kodak.

4.1.1 Détection de défauts locaux

Les défauts locaux apparaissent comme des taches claires ou sombres sur les images. Étant donné le mécanisme de leur formation, ils ne sont pas corrélés dans le temps. On les a aussi appelés *défauts locaux aléatoires* ou *défauts locaux non corrélés dans le temps*.

Quelques chercheurs ont essayé de détecter certains types particuliers de défauts locaux sans exploiter la dimension temporelle. Par exemple, on a tenté de détecter les défauts allongés, fins et sombres, produits par des cheveux adhérent à la pellicule. Cependant, des éléments de la scène peuvent répondre aux mêmes caractéristiques, si bien que ce genre d'approche pâtira toujours, nous le pensons, d'une mauvaise sélectivité.

Inspirés par les travaux de Pardàs et al. [64], nous avons décidé d'utiliser des filtres par reconstruction pour les détecter. On considère une image 3D constituée par plusieurs images de la séquence. L'ouverture est effectuée à l'aide d'un segment parallèle à l'axe du temps, et la reconstruction est opérée spatialement. Ce fonctionnement est illustré par la figure 4.1, sur un exemple binaire. En niveaux de gris, nous calculons la différence entre l'image originale et le filtre, que nous binarisons grâce à un seuil par hystérésis.

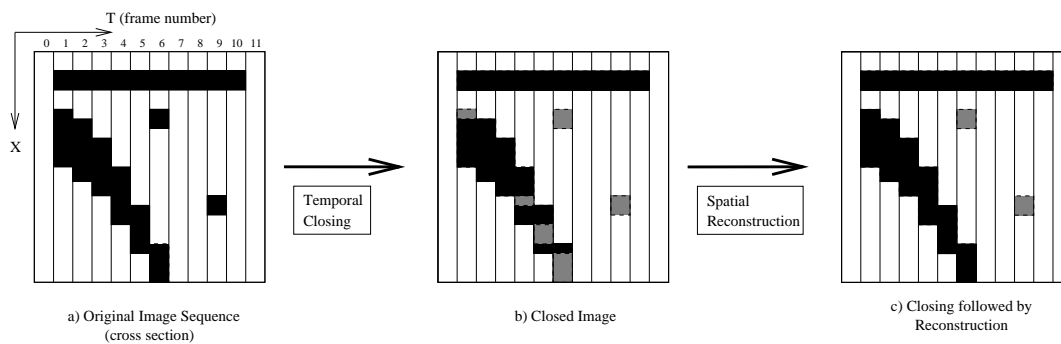


FIG. 4.1: Illustration du fonctionnement de la fermeture temporelle suivie par une reconstruction spatiale sur une coupe d'une image 3D binaire

Cependant, on détecte de cette façon non seulement les défauts aléatoires, mais aussi les objets de la scène dont la vitesse de déplacement est suffisamment importante pour déconnecter temporellement l'objet entre les images consécutives. Pour résoudre ce problème, on peut opter pour une approche classique qui consiste à faire une compensation de mouvement, dans le but de supprimer le déplacement des objets par rapport à une image de référence, avant d'appliquer les opérateurs par reconstruction. Nous avons, de fait, mis en place cette stratégie. Elle donne de bons résultats, mais au détriment du temps de calcul. Or, pour une utilisation industrielle du procédé, il était indispensable de viser un traitement en temps réel. Nous avons donc finalement abandonné la compensation de mouvement globale, pour n'utiliser l'estimation de mouvement que dans l'étape d'interpolation des défauts.

4.1.2 Interpolation des défauts

A l'issue de l'étape de détection des défauts aléatoires, nous obtenons, pour l'image I_t , un masque binaire J_t indiquant les zones qui correspondent à ce que

nous pouvons appelé des *accidents temporels*. Il s'agit de régions appartenant à une image de la séquence, constituant des extréma régionaux, telles que sur les images voisines les pixels correspondant aux mêmes positions ont une valeur de gris plus petite (dans le cas des maxima) ou plus grande (dans l'autre cas) que celle du pixel concerné. Comme nous avons vu, ces accidents temporels peuvent correspondre à des défauts aléatoires, mais aussi à des objets qui bougent vite par rapport à leur taille. D'autres phénomènes peuvent introduire ce genre de structure, tels que le bruit, des flash ou autres changements d'éclairage brutaux.

Nous allons donc chercher à interpoler l'information manquante (dans le cas où l'accident temporel serait bien un défaut) tout en vérifiant si, par hasard, il ne s'agit pas en fait d'un objet en mouvement, et ceci pour chaque accident temporel, c'est à dire pour chaque composante connexe du masque de détection.

Pour cela, nous commençons par définir une zone rectangulaire B autour de chaque accident temporel P , avec une marge de largeur m , tel qu'illustré par la figure 4.2 (plus précisément, nous prenons le plus petit rectangle de côtés parallèles aux axes du repère de \mathbb{Z}^2 , contenant le dilaté de taille m de l'accident temporel, en 4-connexité spatiale).

Le rectangle B (aussi appelé bloc), contient donc P , l'accident temporel en cours d'étude, et une zone non suspecte $P^c = B - J_t$, c'est à dire les points de B qui n'appartiennent pas à des défauts temporels. Remarquons que nous ne prenons pas simplement $B - P$, car d'autres accidents temporels peuvent intersecter le bloc B .

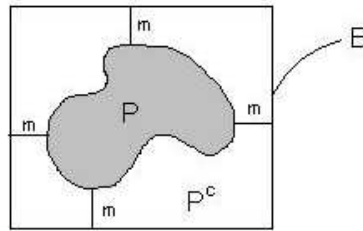


FIG. 4.2: Définition du bloc autour de l'accident temporel

Un déplacement (d_x, d_y) est un élément de \mathbb{Z}^2 . L'erreur interne correspon-

dante est :

$$EI(d_x, d_y) = \frac{1}{S(P)} \sum_{(x,y) \in P} e(I_t(x, y), I_{t-1}(x + d_x, y + d_y)),$$

où la fonction $e(.,.)$ est typiquement une distance dans l'espace d'arrivée. Nous avons utilisé la différence absolue, qui donnait de bons résultats en échange d'un coût calculatoire faible. $S(P)$ donne la mesure de la surface de P , en pixels. Une valeur faible de $EI(d_x, d_y)$ signifie que le déplacement correspondant produit un bon recalage entre le défaut temporel au temps t et une région de l'image $t - 1$ obtenue par translation de P avec le vecteur (d_x, d_y) .

De façon analogue, nous définissons l'erreur externe :

$$EE(d_x, d_y) = \frac{1}{S(P^c)} \sum_{(x,y) \in P^c} e(I_t(x, y), I_{t-1}(x + d_x, y + d_y)),$$

qui donne une mesure de la correspondance du "fond" pour le déplacement concerné.

On se donne un domaine de recherche pour les déplacements autorisés, du genre $-d_{max} \leq d_x, d_y \leq d_{max}$, où d_{max} est un entier positif, et on cherche les déplacements qui minimisent $EI(d_x, d_y)$ et $EE(d_x, d_y)$. Les valeurs minimales respectives sont EI_{min} et EE_{min} .

Ces deux erreurs minimales sont riches en information. Si EI_{min} est petite, alors il semble légitime de penser qu'on a trouvé une bonne correspondance pour la région P , et que par conséquent il s'agit d'une région en mouvement. Au contraire, si cette erreur est grande, l'hypothèse comme quoi l'accident temporel est en fait un défaut sera confortée. Si EE_{min} est petite, alors on peut considérer qu'on a trouvé une bonne correspondance pour P^c sur l'image précédente. Si ce n'est pas le cas, alors il est probable qu'un événement, qu'on n'arrive pas à modéliser, se soit produit. Par exemple, un déplacement très important produisant du flou, ou un changement important d'éclairage, peuvent produire une erreur extérieure importante.

Afin de résumer les remarques ci-dessus, nous avons utilisé le rapport EI_{min}/EE_{max} . Si ce rapport est grand, nous pouvons en pratique supposer que nous avons affaire à un vrai défaut. Dans le cas contraire, nous considérons qu'il s'agit soit

d'un objet en mouvement, soit d'une région mal modélisée. Cependant, d'autres analyses, plus fines, sont envisageables.

Cette approche a donné de très bons résultats. Elle est illustrée par la figure 4.3. En pratique, elle a été combinée avec une estimation de mouvement globale de l'image, destinée à compenser les mouvements de la caméra, et qui peut être calculée efficacement. Nous avons ainsi obtenu un bon compromis entre qualité de la restauration et temps de calcul.

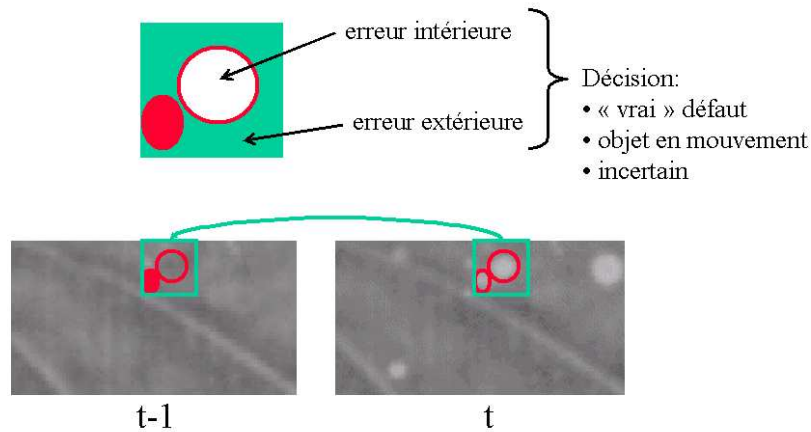


FIG. 4.3: Illustration de la procédure d'interpolation et vérification. En bas à droite : un accident temporel circulaire est traité. Le bloc autour de l'accident est représenté par un carré. Remarquons que ce bloc contient un autre accident temporel, ovale. En bas à gauche : la meilleure correspondance pour le bloc a été trouvée. En haut : illustration de la définition des erreurs intérieure et extérieure, dans le cas présent

4.2 Représentation d'un champ de vecteurs de déplacement

Nous avons décrit dans la section précédente une stratégie de traitement qui visait à ne calculer des vecteurs de déplacement que lorsque cela semblait indispensable. En fait, il arrive souvent dans le traitement de séquences d'images qu'on ne connaisse les vecteurs de déplacement que d'un nombre réduit de points (on parle alors d'un champ *épars* de vecteurs). Ceci peut être dû soit à

un souci d'économie de temps de calcul, soit à une limitation de la méthode d'estimation de mouvement. Dans ce cas, si on a besoin d'avoir un vecteur de déplacement pour tous les points de l'image, c'est à dire si on a besoin d'un champ *dense*, il est indispensable d'utiliser une méthode d'interpolation.

Cependant, les méthodes d'estimation de mouvement sont souvent peu précises. L'estimation par blocs, par exemple, peut produire des vecteurs complètement faux. L'utilisation d'une méthode d'interpolation peut alors provoquer une extension de cette erreur sur une zone plus large.

Nous avons par conséquent développé une méthode originale de modélisation de champs de vecteurs épars, qui permet de produire des interpolations qui limitent l'effet des vecteurs aberrants.

Tout a commencé par une étude réalisée dans le cadre du projet européen Morpheco, qui visait à proposer de nouvelles méthodes de compression de séquences d'images. En collaboration avec Chantal de Fouquet et Fernand Meyer, nous avons eu l'idée d'utiliser un des principaux outils de la géostatistique, le krigeage. En quelques mots, le krigeage [15, 14, 56] est une méthode qui permet, à partir de quelques valeurs ponctuelles connues dans un certain domaine de travail, d'en estimer la valeur en d'autres points. Basé sur une approche probabiliste, il permet d'utiliser différents modèles de la fonction à estimer. Ainsi, en choisissant un modèle de covariance, le géostatisticien peut injecter une certaine information a priori sur le phénomène à modéliser.

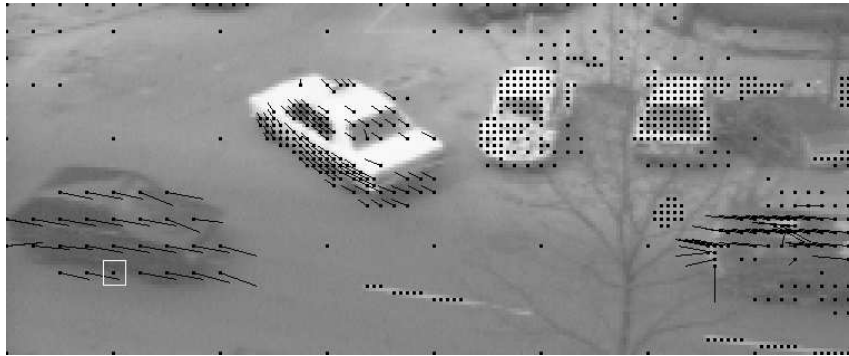
Dans notre cas, nous disposons d'un certain nombre de points $(M_i)_{0 \leq i < n}$ de \mathbb{Z}^2 , où n est un entier strictement positif, en chacun desquels nous connaissons une certaine valeur V_i . Nous avons appelé ces points les *points expérimentaux*. Au début de notre étude les V_i étaient entiers. Assez vite, nous avons commencé à travailler avec des vecteurs, et par conséquent $V_i \in \mathbb{Z}^2$. Il est possible d'utiliser le krigeage pour obtenir une interpolation donnant des valeurs pour tous les points de \mathbb{Z}^2 . Cependant, afin d'imposer une certaine régularité au résultat, nous nous donnons un ensemble de points $(P_i)_{0 \leq i < p}$ de \mathbb{Z}^2 , où p est un entier typiquement inférieur à n . Ces points, que nous appellerons *points d'information*, nous serviront de support de modélisation. Nous calculons alors une valeur V'_i pour chaque P_i de façon à ce que l'interpolation par krigeage en chaque point M_i à partir des (P_i, V'_i) soit aussi proche que possible, au sens des moindres carrés, des V_i . Nous avons aussi utilisé des méthodes robustes [71]

pour réduire l'influence des valeurs aberrantes sur la modélisation. Une description complète de la méthode est donnée dans l'article *Applications of kriging to image sequence coding* [23], joint en annexe D, et qui a obtenu le prix du meilleur article de l'année du journal *Signal Processing : image communication*.

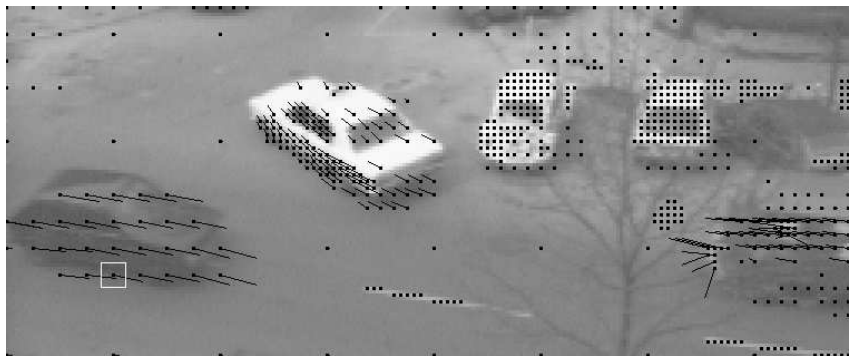
La combinaison de cette approche avec une segmentation de l'image est très riche [20]. En effet, l'hypothèse suivant laquelle le champ de vecteurs varie peu à l'intérieur de chaque région de la segmentation est très souvent vraie. Par conséquent, la modélisation, dans chacune de ces régions, d'un champ éparé avec la méthode que nous venons de décrire est souvent de bonne qualité.

Afin d'illustrer cette approche, nous avons segmenté une image d'une séquence en utilisant la ligne de partage des eaux. Les points expérimentaux sont placés, région par région, sur une grille régulière, dont la densité dépend de la taille de la région. En chacun des points expérimentaux, un vecteur de déplacement par rapport à l'image précédente est calculé en utilisant une estimation par blocs (cf. figure 4.4(a)). On constate que cette méthode d'estimation a donné globalement d'assez bons résultats, sauf en certains points, comme celui entouré d'un carré blanc. Ensuite, dans chaque région de la segmentation, on modélise le champ de vecteurs en utilisant uniquement quatre points d'information, placés aux quatre coins de l'image. Cette modélisation nous permet d'obtenir un champ dense. Sur l'image de la figure 4.4(b) on a représenté les vecteurs de ce champ correspondant aux points expérimentaux. On constate que le vecteur qui était faux a été corrigé. La dernière image (figure 4.4(c)) montre la segmentation qui a été utilisée, avec le résultat de la modélisation du champ de vecteurs.

Nous n'avons pas eu l'occasion de poursuivre l'étude de ce modèle. De nombreuses questions restent sans réponse, comme par exemple la pertinence du choix du nombre et de l'emplacement des points d'information et des points expérimentaux. En particulier en ce qui concerne ces derniers, il semblerait naturel de les placer de façon à faciliter le travail de l'estimateur de déplacement utilisé. Cependant, cette étude montre que l'utilisation d'une segmentation apporte une information précieuse à l'estimation du mouvement. Cette observation n'a rien de très original, vu que cette approche a été exploitée par de nombreux auteurs. Ceci dit, il reste que la morphologie mathématique, grâce à ses puissants outils de segmentation, est particulièrement bien équipée pour



(a)



(b)



(c)

FIG. 4.4: Exemple de krigeage inverse d'un champ de vecteurs.

l'exploiter.

4.3 Conclusion

Alors, peut-on traiter la dimension temporelle comme les dimensions spatiales? Nous avons vu, à travers deux études, que le mouvement est souvent traité à part dans une application de traitement d'images. Ceci est essentiellement dû au fait qu'il perturbe fortement la notion de voisinage. Dès qu'il y a du mouvement dans une séquence (et s'il n'y en a pas, la séquence n'est guère intéressante), deux points, à la même position spatiale sur deux images consécutives risquent fort d'avoir des niveaux de gris différents. En particulier, il suffit alors que la scène contienne des petits objets pour que la connexité temporelle de ceux-ci ne soit plus garantie. En théorie, une fréquence d'échantillonnage temporel suffisamment élevée doit permettre d'éliminer ces problèmes. En pratique, on peut rarement se permettre de le faire. Par conséquent, en règle générale, on ne peut pas traiter le temps comme une dimension spatiale supplémentaire.

Il devient alors nécessaire, d'une façon ou d'une autre, de calculer une estimation de mouvement. On peut alors opérer une compensation de mouvement, qui vise à éliminer les différences introduites par le déplacement entre deux images. Alternativement, on peut aussi bâtir, en utilisant l'information de déplacement, une relation de voisinage adaptée à la séquence, et qui permet de retrouver la continuité temporelle des objets. Laveau [46, 45] propose par exemple des connections temporelles adaptées, qui lui permettent de définir des éléments structurants spatio-temporels suivant le mouvement. A partir de là, il est envisageable de traiter le temps comme les autres dimensions. Le problème devient alors d'effectuer cette étape de façon pertinente; il existe de très nombreux outils d'estimation de mouvement, chacun avec ses caractéristiques, qui peuvent être plus ou moins appropriés pour une application donnée. Nous avons vu, dans le cas de la restauration de films, qu'il est envisageable de limiter le calcul à des régions bien délimitées. Dans le cas de la modélisation d'un champ de vecteurs par krigeage inverse, nous avons aussi vu une façon qui permet de choisir intelligemment les points où se fera l'estimation de mouvement. Ce genre d'approche est, à notre avis, en pratique très intéressante.

Pour finir, disons un mot sur les séquences d'images 3D, comme celles qu'on trouve de plus en plus dans le domaine médical. Nous obtenons alors des images à quatre dimensions, de type $I(x, y, z, t)$. L'augmentation du nombre de pixels entraîne l'augmentation du coût calculatoire, mais en contrepartie, comme nous l'avons déjà noté dans le chapitre précédent, avec ce genre de séquence nous nous affranchissons de toute une série de difficultés qui découlent du fait que les séquences vidéo et assimilées sont obtenues par projection. En effet, dans une séquence d'images 3D nous ne rencontrons pas de problèmes d'occlusion, ni de déformation des objets du fait de la projection. Par conséquent, les modèles de champs de déplacements sont beaucoup plus simples.

In that blessed region of Four Dimensions, shall we linger at the threshold of the Fifth, and not enter therein? Ah, no! Let us rather resolve that our ambition shall soar with our corporal ascent. Then, yielding to our intellectual onset, the gates of the Six Dimension shall fly open; after that a Seventh, and then an Eighth...

Edwin A. Abbott, Flatland (1884)

5

Et au delà ?

Nous voici donc rendus à des séquences temporelles d'images tridimensionnelles, qui sont de ce fait des images à quatre dimensions. Cette inflation du nombre de dimensions va-t-elle se poursuivre? Dans la mesure où ces images représentent déjà les quatre dimensions de notre monde de tous les jours, nous pourrions penser qu'on n'ira pas au delà, mais cette intuition n'est pas exacte.

5.1 Des besoins réels

Au moins dans un domaine on traite déjà depuis un certain temps des données qui sont structurées comme des images à cinq dimensions. Il s'agit de la géophysique. Décrivons de façon très succincte le principe de la technique d'acquisition employée. Une matrice de capteurs (deux dimensions) est posée sur la surface. Un émetteur d'ondes sismiques est déplacé sur la surface (deux dimensions), et à chaque position envoie une onde dans le sous-sol. Chaque capteur enregistre alors un signal temporel (l'écho) à une dimension. Nous obtenons donc en tout cinq dimensions. Ces données brutes sont destinées à être

traitées pour obtenir une image du sous-sol, comme celles sur lesquelles nous avons travaillé avec Timothée Faucon (cf. chapitre 3).

La méthode d'acquisition est illustrée, dans le cas plus simple où les capteurs sont placés sur une droite, par la figure 5.1. Depuis quelques années, ce processus est répété régulièrement pour suivre l'évolution de certains gisements en cours d'exploitation, ce qui rajoute une dimension supplémentaire, temporelle, aux données.

La géophysique est une discipline fascinante et complexe ; l'aperçu que nous venons d'en donner est très réducteur. Nous renvoyons les lecteurs intéressés par le sujet aux nombreux livres sur le sujet, comme celui de Henry [38].

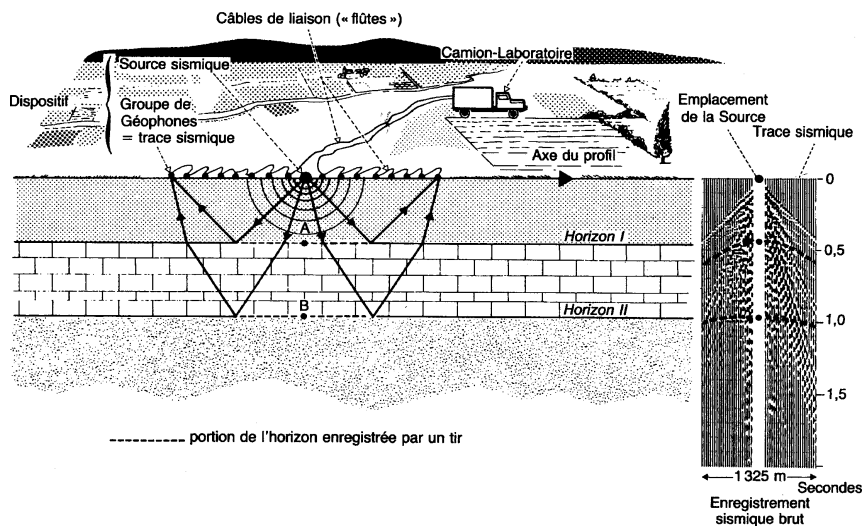


FIG. 5.1: Principe de la méthode sismique et acquisition des données (document de l'Institut Français du Pétrole)

Il est parfaitement envisageable d'appliquer des opérateurs de l'analyse d'images aux données brutes, issues directement du processus d'acquisition. Un des aspects à ne pas négliger est alors le choix d'une relation de voisinage pertinente. Par ailleurs, la visualisation de ces données devient une tâche très délicate, qui requiert de très bien connaître le domaine. Mais avant tout, la taille titanesque de ces images, qui peuvent représenter plusieurs gigaoctets de données, rend leur traitement particulièrement délicat.

Quittons la géophysique pour revenir aux images numériques en général. En

parallèle avec l'augmentation du nombre de dimensions, on constate aussi une nette tendance vers l'augmentation de la résolution. Le domaine de la photographie numérique en offre un éclatant exemple. Aujourd'hui, la plupart des appareils photographiques contiennent des capteurs avec plus de cinq millions de pixels, et les capteurs de dix millions, ou plus, deviennent courants. L'imagerie médicale connaît la même évolution, et désormais des images de scanner de taille $1000 \times 1000 \times 1000$ ne sont plus exceptionnelles.

Dès qu'on voudra appliquer des opérations tant soit un peu complexes à ces masses de données on risque de se retrouver à court de puissance de calcul. Nous atteignons là les limites de la représentation des images en tant que fonctions de \mathbb{Z}^d . Comment pouvons-nous faire pour les traiter de façon efficace ?

Dans tous les cas, le point de départ est le même : l'image. Il faudrait commencer par une simplification efficace de l'information. Par efficace, nous voulons dire non seulement rapide, mais aussi pertinente. Idéalement il faut faire abstraction de ce qui n'est pas utile pour l'application concernée. Cette première étape devrait mener à une autre représentation des données initiales, plus compacte que l'image.

Avec Jean Stawiaski, dont j'encadre le travail de thèse avec Dominique Jeulin (cf. chapitre 3), nous avons décidé d'explorer une piste relativement nouvelle pour apporter un début de solution à ce problème, dans le cas particulier, mais assez courant, de la segmentation. L'idée de base consiste à calculer une segmentation de bas niveau de l'image originale, puis de traiter la partition obtenue à l'aide de graphes.

Commençons par quelques rappels rapides sur les coupes de graphes.

5.2 Coupes minimales de graphes

Soit $G = (X, E, c)$ un graphe non orienté aux arêtes valuées : X est un ensemble fini, contenant au moins deux éléments, appelés *sommets*. E est un ensemble de parties de deux éléments de X , appelés *arêtes*, et notées $e = [x_1, x_2]$. Finalement, c est une fonction de E , à valeurs réelles positives, appelée *capacité*.

Une coupe de G est une partition de X en deux parties S et T . On note

$E(S, T)$ l'ensemble des arêtes reliant un sommet de S à un sommet de T . La capacité de la coupe (S, T) , notée $c(S, T)$, est égale à la somme des capacités des arêtes appartenant à la coupe :

$$c(S, T) = \sum_{e \in E(S, T)} c(e).$$

Soient deux sommets particuliers de X , notés s et t , appelés respectivement *source* et *puits*, et conjointement *terminaux*. Une coupe séparant s et t , dite aussi coupe- (s, t) , est une coupe (S, T) de G telle que s et t appartiennent respectivement à S et à T . Chacune des coupes séparant s et t , qui sont en nombre fini, a une capacité. Celles dont la capacité est minimale sont appelées coupes- (s, t) minimales.

La figure 5.2 illustre les différentes notions introduites ci-dessus.

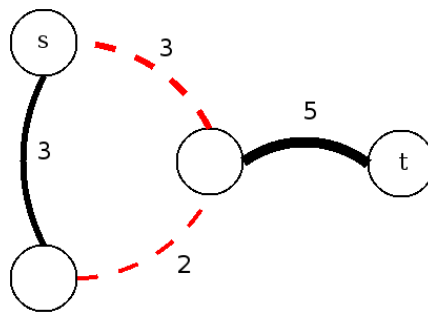


FIG. 5.2: Graphe aux arêtes valuées, et coupe séparant les sommets s et t . Ici, il s'agit d'une coupe minimale, de capacité égale à 5. Cependant, ce n'est pas la seule coupe- (s, t) minimale

Il existe plusieurs algorithmes de calcul de coupes- (s, t) minimales, basés sur le calcul d'un flot maximal entre les terminaux. Citons en particulier Ford et Fulkerson [33], qui ont été les premiers à démontrer que la valeur maximale des flots entre deux sommets d'un graphe est égale à la valeur minimale des coupes séparant ces sommets. Ils ont proposé un algorithme pour calculer cette valeur, qui a été ensuite amélioré par plusieurs auteurs.

Soit n le nombre de sommets du graphe, et m le nombre d'arêtes. Certains algorithmes supposent que les valeurs des arêtes sont uniquement entières ; dans ce cas, soit U la valeur maximale des capacités des arêtes. Edmonds et Karps

[28] ont proposé une variante de l'algorithme originel, avec une complexité en $O(nm^2)$. Les années suivantes on vu la proposition de plusieurs autres algorithmes, de complexités analogues pour nous, traiteurs d'images, pour qui le nombre d'arêtes est typiquement égal au nombre de sommets multiplié par une constante. Par exemple, Dinic [27] a proposé un algorithme en $O(n^2m)$ en 1970. Par la suite, de nouveaux algorithmes, basés sur la notion de pré-flot [39], ont permis d'obtenir de meilleures performances. Ahuja et Orlin [1] ont ainsi proposé en 1986 un algorithme en $O(nm + n^2 \log U)$.

Si la notion de coupe de graphes est relativement ancienne, son application au traitement d'images est beaucoup plus récente. En 1989 Greig et al. [36] ont montré comment utiliser des coupes de graphes pour restaurer des images binaires, mais cette étude est restée largement ignorée jusqu'aux travaux de Boykov, Kolmogorov et leurs collègues, qui ont repris cette approche avec succès [7, 8, 43, 6], pour l'appliquer à plusieurs problèmes de traitement d'images, dont la segmentation [4, 5, 3].

5.3 Segmentation par marqueurs à l'aide de coupes de graphes

Nous expliquons maintenant une façon d'utiliser les coupes de graphes pour effectuer une segmentation avec deux marqueurs. Cet exposé reprend le travail de Boykov et al. [4].

Soit I une image à d dimensions ($d > 0$). Son ensemble de définition D est un produit de d intervalles de \mathbb{Z} , de la forme $\{0, \dots, N_1\} \times \dots \times \{0, \dots, N_d\}$. Nous notons $x = (x_1, \dots, x_d)$ les éléments de D , appelés pixels. Nous considérons une relation de voisinage \mathcal{V} sur l'image I , qui est une relation binaire symétrique anti-réflexive (un pixel n'est jamais en relation avec lui-même) ou, de façon équivalente, un ensemble de parties de deux éléments de D . Nous nous donnons deux sous-ensembles non vides disjoints M_s et M_t de D , respectivement appelés marqueur de l'objet et marqueur du fond. En pratique, dans notre contexte, ces marqueurs correspondent à des pixels désignés par l'utilisateur comme des exemples de points appartenant à l'objet et au fond, où l'objet désigne tout simplement la région que l'utilisateur souhaite détourner, et

le fond son complémentaire. Remarquons qu'en dehors de celles citées, aucune contrainte ne pèse sur les marqueurs. En particulier, ils peuvent être constitués de plusieurs composantes connexes.

A partir de cette image nous bâtissons un graphe $G = (X, E, c)$. Les pixels de l'image, et deux terminaux s et t (quelconques, mais n'appartenant pas à D) constituent l'ensemble des sommets du graphe :

$$X = D \cup \{s, t\}.$$

Deux pixels quelconques de l'image, en tant que sommets du graphe, sont reliés par une arête du graphe si, et seulement si, ils sont voisins (c'est à dire en relation par \mathcal{V}). Le terminal s est uniquement relié à tous les pixels du marqueur de l'objet M_s et, de façon analogue, le terminal t est uniquement relié à tous les pixels du marqueur du fond, M_t . Vu que nous avons supposé les marqueurs disjoints, un pixel ne peut être relié au plus qu'à un terminal. Ceci nous définit E :

$$E = V \cup \{[s, x]/x \in M_s\} \cup \{[x, t]/t \in M_t\}.$$

Concernant la définition de la fonction c donnant les valeurs ou capacités des arêtes, contentons-nous pour l'instant de préciser qu'il faut que les valeurs associées aux arêtes reliant des pixels aux terminaux soient suffisamment grandes pour qu'aucune coupe minimale ne passe par celles-ci. Soit (S, T) une coupe (s, t) minimale de G . Alors, grâce à la condition ci-dessus sur c , les marqueurs M_s et M_t seront respectivement inclus dans S et T . De plus, vu que (S, T) est une partition de X , S privé de s , et T privé de t constituent une partition de D : c'est la segmentation de l'image obtenue au moyen de cette coupe de graphe.

Plusieurs stratégies ont été proposées pour que cette partition corresponde à une bonne segmentation de l'image. En particulier, si on veut que les contours suivent des zones de gradient important, il faut que les valeurs de c soient élevées pour les arêtes reliant des pixels dont les niveaux de gris sont similaires et, au contraire, faibles pour les arêtes entre pixels dont les niveaux de gris sont éloignées. C'est cette approche qui a été utilisée pour segmenter l'image donnée par la figure 5.3 pour obtenir la première segmentation (5.3(b)). On peut faire

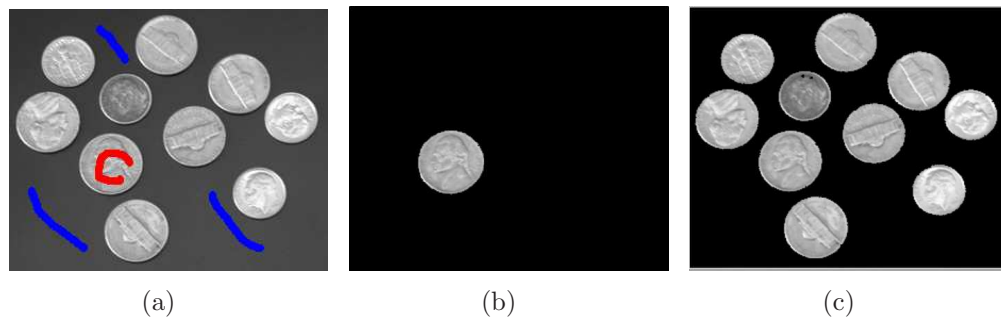


FIG. 5.3: A gauche : image originale (300×246), avec marqueurs. Au centre : segmentation obtenue avec l'approche décrite dans la section 5.3. A droite : segmentation sans contraintes. Les méthodes complètes pour obtenir ces segmentations sont disponibles dans Stawiaski et Decencière [66]

aussi intervenir des critères de longueur [5].

En tant que telle, la segmentation par coupes de graphes, lorsqu'elle est bien utilisée, permet d'éviter les problèmes de fuites qui pénalisent la segmentation par ligne de partage des eaux (LPE) dans certaines applications. En contrepartie, son paramétrage est plus délicat, et elle présente un biais vers les coupes contenant peu d'arêtes. Remarquons aussi que ces méthodes permettent de minimiser certaines classes d'énergies définies sur les images [42, 43, 34]. Notons par ailleurs que la mise en œuvre que nous venons de présenter n'est pas la seule envisageable. Nous avons en particulier utilisé aussi des approches qui exploitent les marqueurs pour extraire de l'information sur les niveaux de gris, sans pour autant que chaque composante connexe d'une partie de la segmentation contienne forcément un marqueur [66], tel qu'illustré par la figure 5.3(c).

Dans tous les cas, la complexité calculatoire des méthodes à base de graphes, et les besoins importants en termes de mémoire, font que cette approche devient délicate lorsqu'on s'attaque à des images de dimension élevée.

5.4 Coupe d'un graphe associé à une segmentation par LPE

L'idée que nous exploitons ici consiste à associer le graphe non pas à l'image initiale, mais à une première segmentation de l'image. Cette représentation a déjà été utilisée par les approches dites de fusion de régions pour la segmentation. Plus récemment, elle a été exploitée pour faire des coupes- (s, t) par Li et al. [53]. C'est cette voie que nous avons décidé de développer.

Reprenons l'image I de la section précédente, et construisons le graphe des régions $G_R = (S_R, E_R, c_R)$.

Nous commençons donc par produire une segmentation fine de l'image. La segmentation par LPE est particulièrement appropriée pour effectuer cette première étape. D'une part, comme nous l'avons déjà évoqué à plusieurs reprises, les algorithmes de calcul les plus performants de la LPE ont une complexité linéaire en fonction du nombre de pixels, donc sont extrêmement rapides. D'autre part, on peut, pendant l'inondation, construire le graphe de voisinage de la partition finale, et calculer une première série de mesures sur les régions résultantes, comme les valeurs d'extinction volumiques, tout en gardant la complexité linéaire. En pratique, nous opérons une LPE sur le gradient de l'image ; la segmentation résultante est dite *de bas niveau*.

Soient $\{R_i\}_{1 \leq i \leq k}$ les k régions de la partition. Avec s et t , elles constitueront les sommets de notre graphe :

$$S_R = \{R_i\}_{1 \leq i \leq k} \cup \{s, t\}.$$

Ces régions héritent de la relation de voisinage de l'image initiale I : deux régions sont voisines si au moins un pixel de l'une est voisin d'un pixel de l'autre. Appelons \mathcal{V}_R la relation de voisinage entre régions. Comme précédemment, les arêtes entre terminaux et autres sommets seront déterminées par les marqueurs. Ici nous rencontrons une difficulté : suite à la première segmentation, une même région peut rencontrer les deux marqueurs M_s et M_t . Si ce cas se présente, ce qui n'est guère courant en pratique vue la taille des régions de la segmentation fine, nous divisons la région concernée en régions n'intersectant chacune au plus qu'un marqueur. Il ne reste plus qu'à lier s et t respectivement aux régions

rencontrées par M_s et M_t :

$$E_R = V_R \cup \{[s, R]/R \cap M_s \neq \emptyset\} \cup \{[R, t]/R \cap M_t \neq \emptyset\}.$$

Il ne reste plus qu'à définir la capacité c des arêtes. Écartons rapidement les arêtes incidentes aux terminaux : comme avant, il faut que leurs capacités soient suffisamment élevées pour qu'elles ne risquent pas de se retrouver dans une coupe minimale.

La définition des capacités entre régions offre beaucoup plus de possibilités que dans le cas des pixels. En effet, le contenu de chaque région permet de construire des mesures de similarité beaucoup plus fines. Ainsi, on peut faire appel aux différentes méthodes de comparaison d'histogrammes pour définir des capacités intéressantes. Callaghan [63] donne une revue de ces mesures de similarité.

La figure 5.4 montre, sur un exemple simple, le résultat de la segmentation par LPE, coupe de graphe au niveau des pixels, et coupe du graphe de voisinage d'une partition fine produite par une LPE à partir de tous les minima du gradient de l'image. Nous constatons que la LPE a tendance à fuir, du fait du bruit important, qui perturbe fortement le gradient. Bien évidemment, un connaisseur ne l'utiliserait pas sans un filtrage préalable, mais nous cherchons à illustrer ici la différence de comportement. Les deux coupes de graphes donnent des résultats qualitativement très proches. Dans cet exemple, les coupes de graphes exploitent uniquement la différence de niveau de gris moyen entre les pixels ou les régions. Les détails du calcul des capacités sont disponibles dans Stawiaski et Decencière [66].

La figure 5.5 montre un exemple plus complexe, concernant une IRM du cœur. Nous avons segmenté cette image 3D d'une part avec une LPE, et d'autre part avec une coupe du graphe à partir d'une LPE. Les mêmes marqueurs (non représentés) ont été utilisés dans les deux cas. Comme précédemment, les capacités du graphe ont été calculées à partir des différences de niveau de gris moyen entre régions voisines. Là encore, la LPE tend à produire des fuites, qui n'apparaissent pas avec la coupe de graphes. Une coupe de graphes au niveau des pixels aurait produit un résultat tout aussi bon, mais au dépens d'un temps de calcul plus important et d'une consommation de mémoire supérieure.

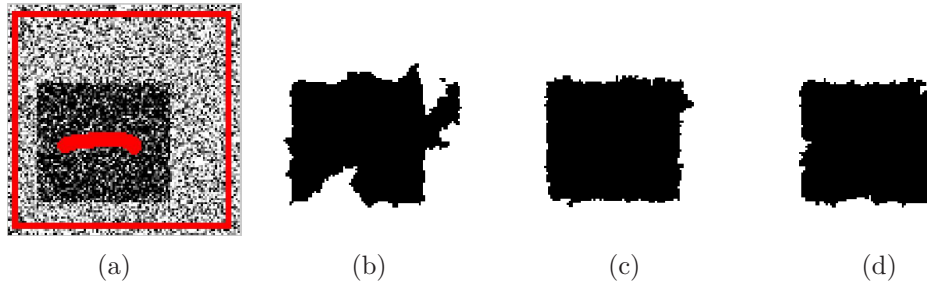


FIG. 5.4: Comparaison de segmentations. De gauche à droite : image originale avec marqueurs ; segmentation par LPE du gradient de l'image originale ; segmentation par coupe de graphe au niveau des pixels ; segmentation par coupe du graphe de voisinage de la segmentation fine produite par la LPE

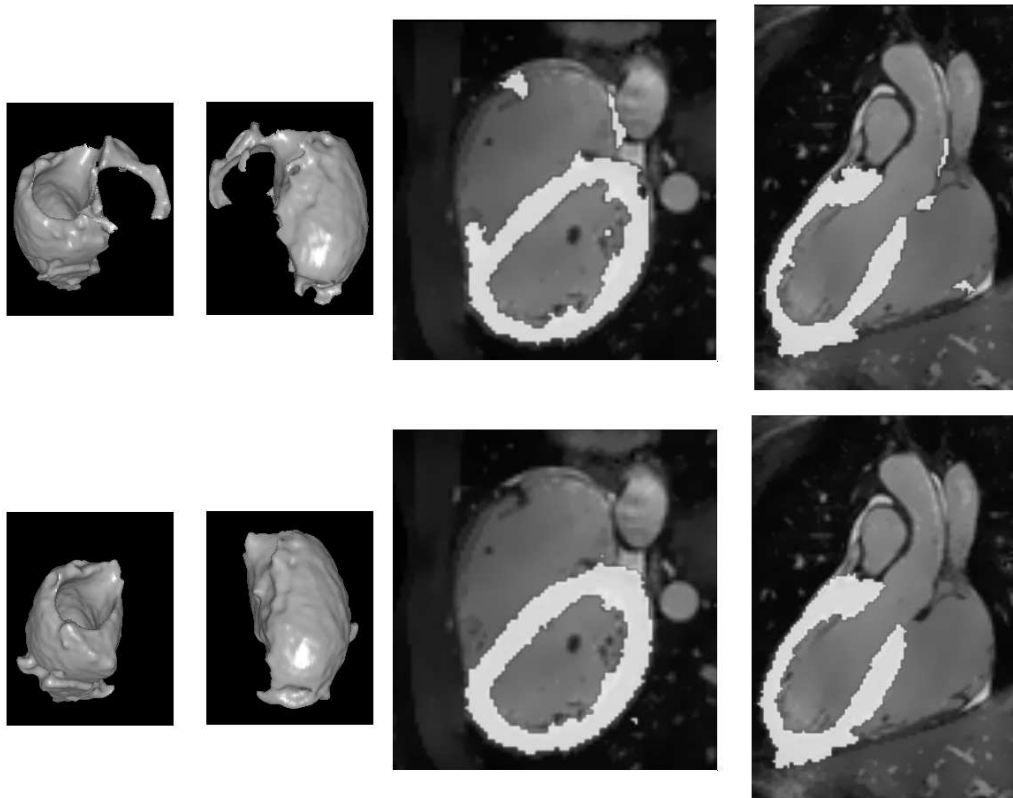


FIG. 5.5: Comparaison de segmentations d'une IRM du cœur ($90 \times 128 \times 90$) par LPE, et coupe de graphes à partir de la LPE ; les mêmes marqueurs ont été utilisés dans les deux cas. Première ligne : résultat obtenu avec la LPE. Deuxième ligne : résultat obtenu avec la coupe du graphe de voisinage de la segmentation de bas niveau [66]

Par ailleurs, la frontière même entre régions voisines peut fournir aussi de l'information intéressante. Leur mesure est le premier critère qui nous vient à l'esprit. Nous l'avons exploité avec succès pour segmenter des images médicales [67] (cf. article en annexe E). Mais il est aussi envisageable d'exploiter d'autres critères géométriques. Peut-on faire intervenir la courbure, ou la régularité des contours, de façon explicite dans ce cadre ?

5.5 Conclusion

La représentation de données sous la forme d'un tableau de d dimensions pose de sérieux problèmes pratiques quand le nombre de dimensions devient élevé, et ceci malgré l'amélioration de la puissance de calcul des ordinateurs. L'augmentation de la résolution, ainsi que la complexité de l'espace des valeurs, accentuent ces problèmes. Il devient alors nécessaire de procéder à une simplification performante des données, et de les représenter de façon plus compacte.

Nous avons proposé ici une mise en œuvre possible de cette stratégie, en vue de faire une segmentation : une simplification à base d'une première segmentation par ligne de partage des eaux, puis une représentation à l'aide d'un graphe de la partition résultante. La LPE est particulièrement adaptée pour effectuer ce travail, vue sa complexité linéaire et la pertinence des résultats qu'elle produit. De son côté, le graphe permet de coder de façon compacte la partition, et offre la possibilité d'utiliser des méthodes relativement performantes de traitement, à savoir les coupes (s, t) minimales. De plus, l'utilisation des graphes permet de faire abstraction du nombre de dimensions de l'espace initial.

Par ailleurs, nous pensons que notre approche ouvre de nouvelles perspectives pour les méthodes à base de coupes de graphes, en permettant de calculer des capacités qui exploitent une information beaucoup plus riche. Ce travail est en cours, mais semble d'ores et déjà prometteur.

6

Conclusion

Finalement, les difficultés apportées par l'augmentation du nombre de dimensions ne se trouvent pas toujours là où on pourrait les attendre. Détecter des défauts sur des séquences de films anciens est bien plus simple si nous disposons de la dimension temporelle, et la détection de failles est moins difficile sur un cube sismique que sur une coupe. A notre avis, la modélisation des objets que nous cherchons à filtrer, détecter ou mesurer est toujours plus facile lorsque nous disposons de davantage de dimensions. Ceci explique d'ailleurs l'engouement pour les techniques d'acquisition plus riches, aussi bien dans le domaine médical que dans celui des matériaux.

Lorsque les modèles sont intrinsèquement bidimensionnels, tels que ceux utilisés par les méthodes de contours actifs paramétriques, le passage en trois dimensions, et au delà, n'est pas immédiat, mais ne pose pas forcément de problèmes importants. Cependant, la topologie devient plus complexe, et par conséquent les méthodes qui l'exploitent doivent être revues plus profondément. La séparation de surfaces imbriquées, par exemple, est bien plus délicate que la séparation de courbes en 2D. Cette complexité topologique supplémentaire

est de même nature que les structures plus complexes que le bruit ou la texture peuvent créer en 3D. La théorie de la percolation aide à modéliser, et donc à comprendre, ce phénomène.

A l'inverse, d'après notre expérience, les problèmes informatiques deviennent nettement plus importants au delà de la deuxième dimension. Étant donnée l'augmentation des volumes de données, les contraintes en mémoire et en puissance de calcul deviennent sévères. Ceci peut nous obliger à coder de façon plus efficace les différents opérateurs ; à utiliser du matériel dédié, comme les processeurs des cartes graphiques ; ou encore à chercher des algorithmes plus performants.

Plus généralement, il peut s'avérer souhaitable de chercher des représentations plus synthétiques des données, comme une segmentation fine de l'image, exploitée sous forme de graphe de voisinage. Dans de nombreux cas, ce n'est que par une approche de ce type qu'on pourra atteindre les performances attendues par l'utilisateur.

Le cas de la dimension temporelle est, à notre avis, particulier. Si, par commodité, on peut être amené à la représenter comme une dimension spatiale supplémentaire, elle demande en pratique un traitement à part - ou du moins un pré-traitement adéquat - permettant d'opérer une compensation de mouvement, ou la construction d'une relation de voisinage adaptée.

La visualisation des données est un aspect qu'il ne faut jamais négliger. S'il est déjà important en 2D, il devient complexe en 3D et en 3D+t. Il existe déjà des solutions sous la forme de logiciels libres ou commerciaux, mais quelle que soit la solution retenue, il ne faut pas hésiter à investir dans les moyens nécessaires pour qu'elle soit à la hauteur des besoins.

D'ailleurs, lorsqu'on dépasse le nombre de dimensions que notre système visuel peut appréhender, nous pouvons aisément imaginer que le problème de la visualisation devient particulièrement délicat. On peut même se demander dans quelle mesure, dans ce cas, on peut encore parler de traitement d'images, vu que les données n'ont plus vraiment, pour nous, un sens visuel.



**Curriculum vitæ détaillé d'Etienne
Decencière**

B

Article joint au chapitre 2

C. Caciu et E. Decenière. Numerical analysis of a 3D hydrodynamic contact, dans *International Journal for numerical methods in fluids*, vol. 51, num. 12, pages 1355-1377, 2005



Article joint au chapitre **3**

R. Lerallut, E. Decencière et F. Meyer. Image filtering using morphological amoebas, dans *Image and Vision Computing*, vol. 25, num. 4, pages 395-404, 2007

D

Article joint au chapitre 4

E. Decencière, C. de Fouquet et F. Meyer. Applications of kriging to image sequence coding, dans *Signal Processing : Image Communication*, Vol. 13, N. 3, pages 227-249, 1998



Article joint au chapitre **5**

J. Stawiaski, E. Decencière et F. Bidault. Computing Approximate Geodesics and Minimal Surfaces using Watershed and Graph-Cuts, dans International Symposium on Mathematical Morphology (ISMM'2007), Brésil, 2007

Bibliographie

- [1] R.K. Ahuja and J.B. Orlin. A fast and simple algorithm for the maximum flow problem. Tech. rep. 1905–87, Massachusetts Institute of Technology, 1987. [53](#)
- [2] Mathilde Boehm. *Contribution à l'amélioration du rendu volumique de données médicales 3D*. PhD thesis, Ecole des Mines de Paris, Paris, 2004. [22](#)
- [3] Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D segmentation. *International Journal of Computer Vision*, 70(2) :109–131, 2006. [53](#)
- [4] Y. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *International Conference on Computer Vision*, pages 105–112, Vancouver, Canada, 2001. [53](#)
- [5] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *International Conference on Computer Vision*, volume 1, pages 26–33, Nice, France, 2003. [53](#), [55](#)
- [6] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 :1124–1137, 2004. [53](#)
- [7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. In *Proceedings of the International Conference on Computer Vision (ICCV)*, volume 1, pages 377–384, 1999. [53](#)
- [8] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23 :1222–1239, 2001. [53](#)
- [9] Costin A. Caciuc. *Analyse et optimisation des surfaces des chemises de moteurs thermiques*. PhD thesis, Ecole des Mines de Paris, Paris, May 2006. [15](#)
- [10] Costin A. Caciuc and Etienne Decencière. Numerical analysis of a 3D hydrodynamic contact. *International Journal for numerical methods in fluids*, 51 :1355–1377, 2006. [15](#)

- [11] Costin A. Caciuc, Etienne Decencière, and Dominique Jeulin. Numerical analysis of the consequences of rugosity modifications in 3D hydrodynamic contacts. *STLE Tribology Transactions (soumis juillet 2006)*. 15
- [12] Costin A. Caciuc, Etienne Decencière, and Dominique Jeulin. Parametric optimisation of periodic textured surfaces for friction reduction in combustion engines. *STLE Tribology Transactions (soumis juillet 2006)*. 15
- [13] Costin A. Caciuc, Etienne Decencière, and Dominique Jeulin. Stochastic shape optimisation. In C. Ronse, L. Najman, and Etienne Decencière, editors, *Mathematical morphology : 40 years on (Proceedings of ISMM'2005)*, pages 333–342, Paris, France, April 2005. 15, 18
- [14] P. Chauvet. *Aide-mémoire de géostatistique linéaire*. Ecole des Mines de Paris, 1993. 44
- [15] P. Chauvet. *Processing data with a spatial support : Geostatistics and its methods*. Ecole des Mines de Paris, 1993. 44
- [16] Etienne Decencière. Motion picture restoration using morphological tools. In P. Maragos, R.W. Schafer, and M.A. Butt, editors, *Mathematical Morphology and its applications to signal processing (Proceedings ISMM'96)*, pages 361–368, Atlanta (GA), USA, May 1996. Kluwer Academic Publishers. 39
- [17] Etienne Decencière. Restoration of old motion pictures. *Microscopy, Microanalysis, Microstructures*, 7(5/6) :311–316, October-December 1996. 39
- [18] Etienne Decencière. *Restauration automatique de films anciens*. PhD thesis, Ecole Nationale Supérieure des Mines de Paris, December 1997. 39
- [19] Etienne Decencière. An automatic system for restoring image sequences. In Stephen Marshall, N. Harvey, and D. Shah, editors, *Noblesse workshop on non-linear model based image analysis*, pages 77–82, Glasgow, UK, July 1998. Springer. 39
- [20] Etienne Decencière. Segmentation-based motion compensation. In H.J.A.M. Heijmans and J.B.T.M. Roerding, editors, *Mathematical Morphology and its applications to signal processing (Proceedings ISMM'98)*, Amsterdam, The Netherlands, June 1998. Kluwer Academic Publishers. 39, 45
- [21] Etienne Decencière. Mathematical morphology and motion picture restoration. In C. Kotropoulos and I. Pitas, editors, *Nonlinear model-based image/video processing and analysis*, chapter 4. John Wiley and Sons, 2001. 39
- [22] Etienne Decencière and Serge Beucher. Contract CT1008 between ARMINES and KODAK Industrie. final report. Technical report N-24/99/MM, Ecole des Mines de Paris, Paris, May 1999. 39

-
- [23] Etienne Decencière, Chantal de Fouquet, and Fernand Meyer. Applications of kriging to image sequence coding. *Signal Processing : Image Communication*, 13(3) :227–249, December 1998. 45
- [24] Etienne Decencière and Dominique Jeulin. Morphological decomposition of the surface of an internal combustion engine cylinder in view of characterizing wear. *Wear*, 249 :482–488, 2001. 14
- [25] Etienne Decencière, Stephen Marshall, and Jean Serra. Application of the morphological geodesic reconstruction to image sequence analysis. *IEE Proceedings : vision, image and signal processing*, 144(6), December 1997. 39
- [26] Etienne Decencière and Jean Serra. Detection of local defects in old motion pictures. In *VII National Symposium on Pattern Recognition and Image Analysis*, pages 145–150, Barcelona, Spain, April 1997. 39
- [27] E.A. Dinic. Algorithm for solution of a problem of maximum flow in networks with power estimation. *Sov. MathK Dokl*, 11 :1277–1280, 1970. 53
- [28] J. Edmonds and R.M. Karp. Theoretical improvements in algorithmic efficiency for network flow problems. *J. ACM*, 19(2) :248–264, 1972. 53
- [29] Raffi Enficiaud. *Algorithmes multidimensionnels et multispectraux en morphologie mathématique : approche par méta-programmation*. PhD thesis, Ecole des Mines de Paris, Paris, 2007. 7
- [30] T. Faucon, Etienne Decencière, and C. Magneron. Morphological segmentation applied to 3D seismic data. In C. Ronse, L. Najman, and Etienne Decencière, editors, *Mathematical morphology : 40 years on (Proceedings of ISMM'2005)*, pages 475–484, Paris, France, April 2005. 28
- [31] T. Faucon, Etienne Decencière, and C. Magneron. Application of surface topological segmentation to seismic imaging. In *Discrete Geometry for Computer Imagery (DGCI 2006)*, pages 506–517, Szeged, Hungary, October 2006. 28, 29
- [32] Timothée Faucon. *Segmentation morphologique et topologique de cubes sismiques*. PhD thesis, Ecole des Mines de Paris, Paris, 2007. 24, 28, 30
- [33] L.R. Ford and D.R. Fulkerson. *Flows in networks*. Princeton University Press, 1962. 52
- [34] D. Freedman and P. Drineas. Energy minimization via graph cuts : settling what is possible. In *CVPR*, pages 939–946, 20025. 55
- [35] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6) :721–741, November 1984. 19

- [36] D. Greig, B. Porteous, and A. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B*, 51(2) :271–279, 1989. 53
- [37] Geoffrey Grimmett. *Percolation, second edition*. Springer, 1999. 26
- [38] G. Henry. *Sismique réflexion : principes et développements*. Technip, 1997. 50
- [39] A.V. Karzanov. Determining the maximal flow in a network by the method of preflows. *Sov. Math. Dokl.*, 15 :434–437, 1974. 53
- [40] L. Khriji, M. Gabbouj, G. Ramponi, and Etienne Decenci re. Old movie restoration using rational spatial interpolators. In *The 6th IEEE International Conference on Electronics, Circuits and Systems*, Paphos, Cyprus, September 1999. 39
- [41] L. Khriji, M. Gabbouj, G. Ramponi, and Etienne Decenci re. Rational spatial interpolators for old movie restoration. In *1999 IEEE International Conference on Image Processing*, Kobe, Japan, October 1999. 39
- [42] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? In *ECCV*, 2002. 55
- [43] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2) :147–159, 2004. 53, 55
- [44] Philippe Lacroute and Marc Levoy. Fast volume rendering using a shear-warp factorization of the viewing transformation. In *SIGGRAPH 1994 Proceedings*, pages 451–457, 1994. 23
- [45] N. Laveau and C. Bernard. Structuring elements following the optical flow. In C. Ronse, L. Najman, and Etienne Decenci re, editors, *Mathematical morphology : 40 years on (Proceedings of ISMM'2005)*, pages 43–52, Paris, France, April 2005. 47
- [46] Nicolas Laveau. *Mouvement et vid o : estimation, compression, et filtrage morphologique*. PhD thesis, Ecole des Mines de Paris, Paris, 2005. 47
- [47] Romain Lerallut, Mathilde Boehm, Etienne Decenci re, and Fernand Meyer. Noise reduction in 3D images using morphological amoebas. In *Proceedings of IEEE International Conference on Image Processing*, Genoa, Italy, 2005. 27
- [48] Romain Lerallut, Etienne Decenci re, and Fernand Meyer. Image filtering using morphological amoebas. In C. Ronse, L. Najman, and Etienne Decenci re, editors, *Mathematical morphology : 40 years on (Proceedings of ISMM'2005)*, pages 13–22, Paris, France, April 2005. 27
- [49] Romain Lerallut, Etienne Decenci re, and Fernand Meyer. Image filtering using morphological amoebas. *Image and Vision Computing*, 25(4) :395–404, April 2007. 27

-
- [50] Marc Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(5) :29–37, May 1988. 23
- [51] Marc Levoy. Efficient ray tracing of volume data. *ACM Transactions on Graphics*, 9(3) :245–261, July 1990. 23
- [52] Marc Levoy. Volume rendering by adaptative refinement. *The Visual Computer*, 6(1) :2–7, July 1990. 23
- [53] Y. Li, J. Sun, C. Tang, and H. Shum. Lazy snapping. In *SIGGRAPH 2004 Proceedings*, volume 23, pages 303–308, 2004. 56
- [54] G. Malandain, G. Bertrand, and N. Ayache. Topological segmentation of discrete surfaces. *International Journal of Computer Vision*, 10(2) :183–197, 1993. 29
- [55] Beatriz Marcotegui, P. Correia, F. Marqués, R. Mech, R. Rosa, M. Wollborn, and Francisca Zanoguera. A video object generator tool allowing friendly user interaction. In *IEEE International Conference on Image Processing (ICIP'99)*, Kobe, Japan, October 1999. 31
- [56] G. Matheron. *La théorie des variables régionalisées et ses applications*. Ecole des Mines, Paris, 1970. 44
- [57] Fernand Meyer. Un algorithme optimal pour la ligne de partage des eaux. In *8ème congrès de reconnaissance des formes et intelligence artificielle*, volume 2, pages 847–857, Lyon, France, November 1991. 33
- [58] Fernand Meyer. Morphological multiscale and interactive segmentation. In *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, Antalya, Turkey, June 1999. 32
- [59] Fernand Meyer. An overview of morphological segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(7) :1089–1118, 2001. 32
- [60] Fernand Meyer and Serge Beucher. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1) :21–46, September 1990. 32
- [61] M.I. Miller, G.E. Christense, and R.D. Rabitt. Deformable templates using large deformation kinematics. *IEEE Transactions on Medical Imaging*, 1996. 19
- [62] G. Monteil and C. Lebeaud. Réduction du frottement au contact segment piston chemise : contribution à l'amélioration du rendement des moteurs. *Revue de la Société des Ingénieurs*, 719, 1998. 13
- [63] Robert J. O'Callaghan and David R. Bull. Combined morphological-spectral unsupervised image segmentation. *IEEE Transactions on Image Processing*, 14 :49–62, 2005. 57

- [64] M. Pardàs, Jean Serra, and L. Torrès. Connectivity filters for image sequences. In *Proc. SPIE : Image Algebra and Morphological Image Processing III*, volume 1769, pages 318–329, San Diego, USA, July 1992. 40
- [65] Jean Serra. *Image Analysis and Mathematical Morphology - Volume I*. Academic Press, 1982. 27
- [66] Jean Stawiaski and Etienne Decencière. Combining graph cuts and watershed segmentation. *Image Analysis and Stereology (soumis juillet 2007)*. 55, 57, 58
- [67] Jean Stawiaski, Etienne Decencière, and Francois Bidault. Computing approximate geodesics and minimal surfaces using watershed and graph-cuts. In *International Symposium on Mathematical Morphology (ISMM'2007)*, Brazil, 2007. 59
- [68] P. Tylsky, G. Bonniaud, Etienne Decencière, J. Stawiaski, D. Lefkopoulos, and M. Ricard. FDG PET images segmentation using morphological watershed : a phantom study. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, San Diego (CA), USA, October 2006. 33
- [69] C. Vincent, Costin A. Caciuc, G. Monteil, Etienne Decencière, and Dominique Jeulin. Friction analysis in laser surface textured lubricated contacts. In *STLE Annual Meeting*, Philadelphia, USA, 2007. 15
- [70] Francisca Zanoquera. *Segmentation interactive d'images fixes et séquences vidéo basée sur des hiérarchies de partitions*. PhD thesis, Ecole des Mines de Paris, Paris, 2001. 38
- [71] Z. Zhang. Parameter estimation techniques : a tutorial with application to conic fitting. *International Journal of Image and Vision Computing*, 15(1) :59–76, January 1997. 44