



Morphological and optical properties of rough surfaces

Gabriel Fricout

► To cite this version:

Gabriel Fricout. Morphological and optical properties of rough surfaces. domain_other. École Nationale Supérieure des Mines de Paris, 2004. English. NNT: . pastel-00001102

HAL Id: pastel-00001102

<https://pastel.hal.science/pastel-00001102>

Submitted on 15 Apr 2005

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

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



Propriétés Morphologiques et Optiques des Surfaces Rugueuses

THÈSE

présentée à
l'École Nationale Supérieure des Mines de Paris
par

Gabriel Fricout

pour obtenir le titre de

DOCTEUR

en

MORPHOLOGIE MATHÉMATIQUE

Soutenue le 5 Novembre 2004 à Fontainebleau devant le jury composé de :

Mikhaël BALABANE	<i>Rapporteur</i>
Michel JOURLIN	<i>Rapporteur</i>
Pierre-Jean KRAUTH	<i>Examineur</i>
Jean-Michel MATAIGNE	<i>Examineur</i>
Hassan ZAHOUANI	<i>Examineur</i>
Dominique JEULIN	<i>Directeur de thèse</i>

Remerciements

Et voilà, il est temps de mettre un point final à cette entreprise qui aura duré trois années. Trois années au cours desquelles de nombreuses personnes auront contribué plus ou moins directement à la forme définitive de cette thèse.

Parmi elles, il me faut chaleureusement remercier Dominique Jeulin pour sa direction exemplaire. Il a tout d'abord réussi à convaincre un jeune ingénieur (qui il est vrai ne demandait qu'à être convaincu !) de venir passer trois ans de réflexion à Fontainebleau au lieu d'aller affronter de plein fouet le monde du travail, avec les avantages et les inconvénients que cela comporte. Ensuite, que ce soit par la pertinence de ses suggestions ou de ses corrections, son suivi régulier et attentif des travaux en cours ou encore par les rapports humains toujours très agréables qui ont accompagné nos déplacements et nos réunions, il a très certainement marqué cette thèse de l'empreinte de sa rigueur scientifique.

Ce travail doit également beaucoup à Pierre-Jean Krauth qui a su admirablement faire le lien entre le monde industriel et le monde scientifique dont les tenants et les aboutissants sont parfois divergents. Je le remercie en particulier pour son enthousiasme et son dynamisme dans la mise en oeuvre pratique de nos outils et pour avoir su les valoriser au sein du groupe Arcelor. Sur un plan moins professionnel, je le remercie pour sa bonne humeur et son hospitalité qui ont agrémenté de nombreux déplacements dans le grand Est voire dans le grand Nord ! C'est aussi l'occasion de remercier l'ensemble du personnel du département Mesure-Contrôle pour tout le travail fourni autour des données, acquisitions, mesures, prototypage, intégration informatique... Sans eux, les "applications industrielles" de cette thèse seraient restées à l'état de perspectives.

Je remercie d'autre part l'ensemble des personnes qui ont suivi les avancées de ce travail au cours des différentes réunions du groupe surface, ainsi que les membres du Jury qui ont accepté de lire ce manuscrit et d'assister à la soutenance : leur avis et leurs remarques comptent beaucoup pour moi.

Fernand Meyer est aussi l'un de ceux sans qui cette thèse n'aurait probablement pas vu le jour : il a su mettre l'eau à la bouche des optionnaires "Vision et Morphologie", de telle sorte que plusieurs d'entre eux ont été piqués par le virus du traitement d'image !

L'ensemble des chercheurs du centre de morpho (voire des centres adjacents !) a également formé un environnement de travail particulièrement fertile : Béa, Etienne, Serge, Michel, Jean-Claude, Petr, Christian "Lantu", Jean Serra, tous ont, à un moment ou à un autre (et souvent beaucoup plus d'une fois), au détour d'une réunion de centre ou d'une pause café, donné un avis ou une idée qui se retrouve quelque part dans le manuscrit qui suit. Je remercie bien évidemment Marc et Catherine, qui ont réussi à supporter des râleries récurrentes, des vélos dans le bureau (même s'ils étaient parfois bien dissimulés, et même s'il s'agit de magnifiques vélos qui n'enlèvent rien à la déco de la pièce...) et une incompétence notoire pour tout ce qui concerne les papiers administratifs.

Il n'y aurait pas de thèse sans thésards, et ceux-là sont un élément incontournable de trois années à Fontainebleau. Le bureau d'abord constitue un microcosme particulière-

ment intéressant : entre les trucs qui marchent pas, ceux dont on voudrait qu'ils marchent, les projectiles identifiés et moins identifiés, les invasions barbares, et bien sûr le tableau dont la seule présence peut suffire à stimuler trois esprits sur les sujets les plus divers, on arrive malgré tout à travailler, un peu, et même parfois efficacement ! En tout cas, je remercie Romain, Timothée et Mathilde pour avoir passé une thèse pleine d'échanges incessants, de péripéties et de rebondissements (de balle cela va sans dire) de stress parfois, mais rarement le mien, et de franches rigolades, que ce soit au bureau ou en dehors. Il y a aussi les adeptes du grenier, Raffi, le thésard permanent que l'on peut toujours appeler quand plus personne ne répond : quelle que soit l'heure, la soirée repartira, Thomas et ses discussions toujours très pertinentes autour d'une bière ou d'une table de jeu, Costin, avec qui l'occupation la plus simple devient parfois une aventure complexe, et dont le breuvage aux essences importées de Roumanie aura servi de potion magique à un noyau dur d'irréductibles thésards. Au fin fond de nulle part, on trouve aussi Nico, le deuxième rescapé morphal, dont les réceptions sont sans pareils. Bien d'autres personnes ont partagé ces trois années bellifontaines et ont contribué à les rendre inoubliables : Julie avec qui les séances d'escalade sont toujours un plaisir, Cyril, l'incontournable psychologue partenaire de vélo ou de bloc, toujours là pour vous redonner bonne conscience lorsqu'on croit qu'on exagère parce que deux heures de grimpe tous les midis, c'est excessif ! David est resté un lièvre inaccessible dans les côtes en VTT (heureusement qu'il y a aussi des descentes), Shan et Janos ont été les compagnons de nombreux week-end verticaux : la direction est donnée, reste à déterminer le sens. Outre la fréquente controverse (Grimpe ou Vélo ?), je les remercie d'avoir évité de me casser complètement, pendu au bout d'une corde pour l'un et un vélo sous les pieds (c'est le cas le plus favorable) dans des descentes parfois indécentes pour l'autre. Mes frères et sœurs, Bruno, Sandra et Nathalie ont permis des squatts chambériens ou parisiens toujours très accueillants. Et il y a tous les autres, cela va sans dire, dont la liste exhaustive prendrait des pages mais parmi lesquels on peut citer : Sergio, Jaromir, Marc et Miguel, Gaël, Thibaut, Jesus, Mireille, l'équipe de horse-ball de Milly la Forêt, les hardcorps, les minéros, Serwan et Damien ...

Je tiens à remercier mes parents, qui m'ont toujours encouragé et poussé dans les directions que j'avais choisies. Aller au bout de cette thèse en faisant de mon mieux est aussi un moyen de les remercier de tout ce qu'ils ont fait et feront encore pour moi.

Et bien sûr, en gardant le meilleur pour la fin, je voudrais tout particulièrement remercier Laurence d'avoir accepté l'exil bellifontain, la remercier pour sa patience, sa gentillesse et sa bonne humeur. Je suis certain que d'autres s'associeront à moi pour louer ses qualités culinaires prouvées et éprouvées maintes et maintes fois à l'occasion de fréquentes soirées. Je la remercie finalement et tout simplement pour son amour sans faille...

Propriétés morphologiques et optiques des surfaces rugueuses

Gabriel Fricout

Centre de Morphologie Mathématique, Ecole des Mines de Paris

Dans cette thèse, nous nous intéressons à l'aspect général des surfaces rugueuses, et tout particulièrement à celui de surfaces d'acier, revêtues ou non, pouvant se destiner à de très larges domaines d'application. Par *aspect de surface*, nous entendons ici *caractéristiques générales de texture* sans nous focaliser sur des défauts très localisés. Actuellement, le contrôle de l'aspect des surfaces en phase de production s'obtient par échantillonnage régulier du produit et par comparaison *visuelle* avec des échantillons de référence. Non seulement cette méthode ne permet pas un contrôle continu des tôles produites, mais elle introduit également une certaine subjectivité dans l'évaluation de leur apparence.

Dans un premier temps, nous développons un système de caractérisation des surfaces basé sur l'extraction de *descripteurs morphologiques* (segmentation des surfaces, granulométrie, courbe de tortuosité) et statistiques (courbe de covariance, histogramme des pentes, orientation des normales). Nous nous intéressons ensuite à la distribution, et en particulier à la variance, de ces descripteurs lorsque l'on observe plusieurs fois une même surface en plusieurs endroits différents. Ceci nous amènera tout d'abord à utiliser des techniques de géostatistique puis à exploiter des méthodes de sur-échantillonnage de type *bootstrap* dans des situations délicates de dépendance entre variables.

Dans un second temps, nous proposons une méthode statistique basée sur plusieurs étapes d'analyse factorielle permettant de choisir un espace de représentation adapté à la mise en oeuvre d'une classification bayésienne.

Cette méthodologie générale d'étude des surfaces sera ensuite appliquée à une situation réelle : la *classification automatique* de surfaces de type "peau d'orange" en fonction de la qualité de leur aspect. Après plusieurs essais en laboratoire, nous étudions la mise en ligne d'un système d'acquisition d'images de façon à mettre en place un contrôle systématique de l'aspect des surfaces lors de leur production.

Pour améliorer encore les performances de notre système, nous envisageons une possible caractérisation des surfaces par des *arbres d'inondation* et nous développons un algorithme de division d'arbre en sous-arbres afin d'appliquer des méthodes de bootstrap à ces structures particulières.

Enfin, nous nous intéressons au *comportement optique* des surfaces à travers deux problèmes. Nous simulons d'abord, à partir de données topographiques, l'image caméra que l'on obtiendrait dans une configuration optique donnée, puis nous abordons le problème

inverse, qui consiste à calculer la topographie d'une surface à partir d'une seule image caméra.

Tous ces éléments s'inscrivent dans le cadre d'un effort marqué des producteurs d'acier en général, et d'Arcelor en particulier, pour mieux comprendre la notion d'esthétisme pour les surfaces afin de prévoir et même d'anticiper les attentes des utilisateurs en ce domaine. Dans cette optique, les outils que nous avons développés ouvrent des perspectives pour tenter de caractériser au mieux la "surface idéale" pour une application donnée.

Morphological and optical properties of rough surfaces

Gabriel Fricout

Centre de Morphologie Mathématique, Ecole des Mines de Paris

In this PhD, we have focused our interest on studying general aspect of rough surfaces and particularly steel surfaces, that can be coated and designed for any kind of applications. In the following, the words *surface aspect* have to be understood as *general texture characteristics* without taking into account local defects. The control of such an aspect during production process is performed by regular sampling of the product and *visual* comparison with a reference data set. This method does not only forbid a continuous quality control of surface aspect, but also introduces a very subjective part in surfaces appearance evaluation.

Firstly, we develop a set of *morphological features* (based on surface segmentation, tortuosity map, granulometry ...) as well as statistical features (covariance curve, slopes histogram ...), that leads us to surface characterization. The question of features stability is then considered by means of features variance when a large surface is characterized several times using independent observation windows. Some geostatistical tools enable us to address this problem in some particular cases and *bootstrap over-sampling* is also introduced with special care to statistical dependence between variables.

A statistical method based on different steps of factorial analysis is then introduced in order to chose a representation space that is well suited for bayesian classification.

This very generic methodology for analyzing surface texture is then applied to a real problem. We study the *automatic classification* of surfaces with "orange-peel" aspect according to the quality of their appearance. Once successful experiments were performed in laboratory using interferometric microscopy, we design a vision system based on a collimated source and a matricial camera, so that an automatic and systematic quality control of surface aspect can be set up on production line.

In order to improve our classifier performance, the possibility of using *flooding trees* for texture characterization is considered and we develop an algorithm to divide a flooding tree into subtrees so that bootstrap techniques become usable with such a particular data structure.

Finally, the *optical behavior* of surfaces is modeled by the way of ray casting in order to simulate the image obtained by a camera under specific lighting conditions. Reciprocally, we propose an algorithm to compute a surface topography if only a single camera image is known.

All these points join a recent dynamic coming from steel industry in general and Arcelor group in particular, that should lead to a better understanding of esthetic for surfaces, the final aim being the prediction and anticipation of consumer's desires in this particular field. Considering this, our tools open perspectives for defining the "ideal surface" for a given application.

Table des matières

1. Introduction	5
I. Méthodologie de caractérisation et d'analyse	9
2. Descripteurs pour les représentation 2D	11
2.1. Paramètres classiques de rugosité	11
2.1.1. Caractérisation de profils	12
2.1.2. Paramètres surfaciques : les nouveaux standards	13
2.2. Descripteurs basés sur une représentation de type "image"	13
2.2.1. Autocorrélation	14
2.2.2. Erosion-Dilatation	15
2.2.3. Segmentation	17
2.2.4. Tortuosité	22
2.2.5. Vecteurs Normaux	27
2.2.6. Ondelettes	43
3. Stabilité de la description	49
3.1. Origine du problème	49
3.2. Variance d'extension	50
3.2.1. Fonction aléatoire et covariogramme	50
3.2.2. Exemples d'utilisation	51
3.2.3. Limitations pratiques	56
3.3. Bootstrap	57
3.3.1. Origine et principe	57
3.3.2. Descripteurs "favorables" (indépendance des individus)	59
3.3.3. Bootstrap sous dépendance	64
3.4. Conclusion	72
4. Méthode d'analyse	75
4.1. Analyse de données	76
4.1.1. Généralités sur l'analyse factorielle	76
4.1.2. Séparation d'ensembles	80
4.1.3. Différents types d'analyse factorielle	83
4.1.4. Utilisation de l'analyse factorielle	85

Table des matières

4.2.	Classification	87
4.2.1.	Classification bayésienne	88
4.2.2.	Estimation des fonctions de densité	89
4.3.	Méthodologie d'analyse sélectionnée	94
4.3.1.	Apprentissage	95
4.3.2.	Classification	98
4.4.	Bilan	98
II.	Application industriel	103
5.	Classification de surfaces "peau d'orange"	105
5.1.	Base de surfaces utilisée, contexte industriel	106
5.2.	Classification à partir de topographies	107
5.2.1.	Etude des différents descripteurs	107
5.2.2.	Analyse globale	114
5.2.3.	Résultats de classification	115
5.3.	Vers une application en ligne : choix d'un capteur	128
5.3.1.	Système Ondulo	128
5.3.2.	Système SIAS	128
5.3.3.	Système "caméra matricielle"	129
5.4.	Classification à partir d'images	130
5.4.1.	Acquisition et prétraitement	130
5.4.2.	Analyses spécifiques	131
5.5.	Mise au point d'un prototype, premiers essais	149
5.6.	Autres applications	150
III.	Caractérisation indirecte	153
6.	Représentation arborescente	155
6.1.	Construction d'un arbre	156
6.1.1.	Notations	156
6.1.2.	Inondation	156
6.2.	Extraction de descripteurs	160
6.2.1.	Pondération des noeuds et des arêtes	160
6.2.2.	Exploitation de cette pondération	162
6.3.	Divisions en sous-arbres	164
6.3.1.	Critères simples	164
6.3.2.	Algorithme plus complexe	165
6.4.	Utilisation des sous-arbres	168
6.4.1.	Bootstrap	168
6.4.2.	Filtrage	170
6.5.	Conclusion et Perspectives	173

6.5.1.	Prise en compte des noeuds de transition	173
6.5.2.	Amélioration de l'algorithme de sélection	176
6.5.3.	Autres applications	176
7.	Comportement optique des surfaces	179
7.1.	Problème direct	179
7.1.1.	Rappels de radiométrie	180
7.1.2.	Le modèle de Ward	181
7.1.3.	Modélisation du dispositif	181
7.1.4.	Simulation de lobes	188
7.1.5.	Images obtenues	193
7.2.	Problème inverse	193
7.2.1.	Préliminaires	195
7.2.2.	Techniques classiques de "shape from shading"	196
7.2.3.	Le cas des surfaces peau d'orange	197
7.2.4.	Etude de la réponse optique	199
7.2.5.	Reconstruction des pentes puis de la topographie	204
7.2.6.	Résultats	214
7.2.7.	Utilisation de deux incidences	222
7.2.8.	Conclusion et perspectives	224
8.	Conclusion et perspectives	227
8.1.	Cuique suum (A chacun le sien)	227
8.2.	Perspectives	229

Chapitre 1.

Introduction

Si les propriétés physico-chimiques des matériaux sont essentielles pour remplir la fonction qui leur est assignée, leurs propriétés cosmétiques sont à prendre en considération lorsqu'ils sont soumis à la perception sensorielle de tout un chacun. L'acier ne déroge pas à cette règle, qu'il s'agisse d'automobile, d'appareil électroménager ou de bâtiment, la surface d'une tôle se doit d'être esthétique. Dans un contexte industriel où la compétitivité reste l'un des maîtres mots, une question se pose au producteur d'acier : "Comment contrôler la texture des surfaces produites de manière efficace, objective, automatique et en continu sur une ligne de production, de façon à garantir au client un aspect de qualité homogène?". C'est à cette question précise que nous nous sommes efforcés de répondre tout au long de cette thèse. En pratique bien sûr, la production de l'acier n'est qu'une étape vers un produit fini dont les différents revêtements, la peinture, et éventuellement la mise en forme vont conditionner l'aspect. Nous nous plaçons ici au début de cette chaîne, c'est à dire au niveau de la tôle nue, ou éventuellement pré-peinte, mais avant la phase de peinture proprement dite.

L'aspect visuel d'une surface dépend de nombreux paramètres parmi lesquels on compte principalement les propriétés optiques du revêtement et la topographie de la surface au travers de la forme des vallées et des pics, de l'escarpement du relief ou encore de l'agencement des motifs, en résumé de la *texture* de la surface. C'est sur cette dernière que porteront plus spécifiquement nos efforts de caractérisation. Nous nous placerons donc en quelque sorte à propriétés optiques fixées, en supposant que ces dernières restent constantes pour un type de produit et un revêtement défini (cette présomption sera d'ailleurs confirmée dans la suite de notre travail, chapitre 7).

Voici donc précisés l'objectif industriel et le support physique de l'étude. Reste, en guise de préliminaires, à définir plus concrètement ce que l'on entend par les termes "aspect de bonne qualité". En effet, l'esthétique d'une surface se définit essentiellement en fonction de l'application à laquelle on la destine : une tôle prisée dans le monde de l'électroménager ferait probablement frémir n'importe quel constructeur automobile, et réciproquement. Impossible donc de définir un indice absolu de qualité pour la texture d'une surface, d'autant plus qu'un utilisateur (ou un client) éprouve souvent de grandes difficultés à définir en termes quantitatifs les propriétés essentielles d'une tôle de bonne qualité, par contre, il lui est assez facile de classer un ensemble de surfaces en fonction de ses attentes en terme d'aspect. En conséquence, la méthode actuelle d'inspection de l'aspect des surfaces consiste à prélever régulièrement des échantillons sur la ligne de fabrication, puis à les faire observer par un oeil averti qui les compare à des échantillons de référence

définissant d'une part un aspect "cible" et d'autre part des aspects de mauvaise qualité. Dans ces conditions, il paraît naturel d'utiliser un exemple de classification comme donnée d'entrée d'un logiciel de contrôle automatique de texture.

Exploiter correctement un exemple de classification constitué de plusieurs surfaces suppose être capable de mesurer puis de décrire informatiquement chacune d'entre elles. Dans la première partie de notre travail (chapitre 2), nous présentons plusieurs mesures utilisées couramment dans le monde de la sidérurgie, puis nous développons de nouveaux descripteurs obtenus à partir d'une représentation bidimensionnelle des surfaces, c'est à dire la sortie d'un appareil de mesure assimilable à une image (microscope, caméra, ...). Ces descripteurs peuvent être d'ordre statistique ou morphologique et ils enrichissent considérablement le jeu de paramètres disponibles pour caractériser une surface. Malheureusement, une mesure utilisée pour comparer plusieurs objets n'est rien sans la donnée de la précision qui lui est intrinsèquement associée : ce sera l'objet de notre troisième chapitre qui emprunte des développements à la géostatistique et à la théorie du sur-échantillonnage pour calculer le plus précisément et le plus rigoureusement possible l'erreur (ou la variance en termes plus statistiques) liée à chacun de nos descripteurs de texture. A l'issue de ce troisième chapitre, nous serons capables de décrire une surface au moyen d'un grand nombre de valeurs numériques, et nous pourrons donc chercher à utiliser l'exemple de classification dont nous disposons pour le contrôle automatique de la texture de surfaces inconnues. Ce problème sera abordé dans le quatrième chapitre où nous présentons les détails de cette méthodologie générale d'étude basée d'une part sur de l'analyse factorielle pour projeter les surfaces dans un espace de dimension réduite, puis sur un schéma de classification bayésien utilisant les résultats du chapitre 3. Après avoir établi des bases théoriques solides dans nos quatre premiers chapitres, nous suivons la devise de l'Ecole des Mines, "Théorie et Pratique", et nous replaçons ces développements dans un contexte industriel précis, celui de l'étude des surfaces de type "peau d'orange" destinées à l'industrie de l'électroménager. Dans le chapitre 5, nous proposons un exemple complet d'analyse et de classification sur des surfaces réelles étudiées soit avec un microscope interférométrique, soit avec une caméra matricielle, puis nous décrivons le cheminement expérimental qui nous a conduit depuis les premiers essais en laboratoire jusqu'à l'installation en ligne d'un prototype de contrôle automatique de la texture des surfaces prélaquées. Avec ce chapitre, nous répondons à la question originale sous-tendant notre étude : les résultats obtenus en terme de classification sont conformes à la réalité, le système s'est bien comporté au cours des tests de reproductibilité et fournit donc une information objective qui n'est pas influencée par l'opérateur qui le manipule et en interprète les résultats. Enfin, le premier prototype développé représente un pas important vers un système fiable de contrôle en ligne de l'aspect des surfaces.

En parallèle des essais industriels réalisés dans un contexte précis, nous avons continué à nous intéresser à la caractérisation des surfaces de façon à disposer d'une vaste bibliothèque de descripteurs capables de s'adapter à différents types de produit. Dans notre sixième chapitre, nous suggérons d'utiliser une représentation arborescente des surfaces, et nous développons différentes méthodes pour intégrer les descripteurs qui en sont issus dans notre méthodologie globale d'analyse, au travers notamment d'algorithmes de division d'un arbre en sous-arbres.

Enfin, dans un septième chapitre, nous étudions le comportement optique des surfaces "peau d'orange". En effet, lorsqu'une caméra matricielle destinée à faire des acquisitions d'images en ligne a dû être substituée au microscope interférométrique de laboratoire, nous avons pu apprécier la perte d'information conséquente que cela représentait. Nous avons donc cherché à mieux comprendre les phénomènes optiques mis en jeu lors de la photographie d'une topographie, en vue peut-être d'optimiser la configuration optique du système d'acquisition. Pour ce faire, nous avons développé un algorithme de simulation apparenté aux techniques de lancer de rayons qui permet, moyennant quelques hypothèses simplificatrices, de calculer la réponse optique d'une surface de topographie donnée. Après avoir résolu de façon satisfaisante le problème direct, nous avons envisagé le problème inverse parfois connu sous le nom de "shape from shading", c'est à dire le calcul du relief à partir de l'acquisition en ligne. La tâche est difficile car mathématiquement mal posée : il s'agit de compenser la perte effective d'information entre topographie et image caméra par une connaissance a priori sur les surfaces et des traitements numériques adaptés.

L'ensemble de ces travaux a été réalisé dans le cadre d'une étroite collaboration entre Arcelor Research (anciennement IRSID) et le Centre de Morphologie Mathématique de l'Ecole des Mines de Paris. En particulier, l'ensemble des échantillons étudiés ont été fournis par le groupe Arcelor et l'ensemble des mesures ont été réalisées soit en laboratoire à Maizière-lès-Metz au sein du département Mesure-Contrôle soit en usine sur le site Arcelor de Montataire pour les essais d'acquisition en ligne. La validation pratique des algorithmes développés est à notre avis l'un des points forts de cette thèse qui doit donc beaucoup à la réactivité d'Arcelor Research.

Première partie .

Méthodologie de caractérisation et d'analyse

Résumé

Dans cette partie, nous nous intéressons à la caractérisation des surfaces :

- Nous définissons de nouveaux descripteurs associés à une représentation bidimensionnelle d'une surface.
- Nous étudions la précision de ces descripteurs et leur dispersion lorsqu'une surface est observée plusieurs fois à différents endroits.
- Nous utilisons des méthodes d'analyse statistique pour réduire la dimension de l'espace de représentation (le nombre de descripteurs), et nous utilisons une approche bayésienne pour classer automatiquement les surfaces en plusieurs catégories.

Chapitre 2.

Descripteurs pour les représentation 2D

Dans ce chapitre, nous nous intéressons à la caractérisation quantitative des surfaces. Il s'agit d'un préliminaire essentiel à toute analyse statistique pertinente dont l'objectif sera la classification automatique et l'application industrielle le contrôle qualité. Cette caractérisation est un problème difficile, parce qu'une surface est un objet physique complexe pour lequel on dispose en général d'une représentation numérique (carte de topographie discrète, image issue d'une caméra, profil, ...) elle-même complexe. La caractérisation consiste à extraire de ces représentations une partie de l'information qui y est contenue, par exemple un chiffre, dans un but de comparaison d'une surface à l'autre. Si l'information extraite d'une surface est pertinente, elle nous permettra de prédire certaines propriétés (physiques, optiques, tactiles...), et la comparaison d'une surface à l'autre sera significative, à l'opposé, certaines valeurs calculées peuvent ne pas permettre de distinguer deux surfaces de qualité très différentes pour une application donnée. On touche là d'ailleurs l'une des difficultés essentielles de cette approche : la caractérisation des surfaces devra fortement dépendre de l'application à laquelle elles sont destinées. Fort de ces considérations, nous ne chercherons pas ici à établir des descripteurs "universels", mais plutôt à initier une "boîte à outils" dans laquelle il sera possible de choisir telle ou telle variable en fonction du contexte d'utilisation.

Sur le plan terminologique, nous nommerons "variable" une fonction associant une valeur (réelle le plus souvent, mais éventuellement complexe) à une représentation d'une surface, par opposition à un descripteur qui pourra être formé de plusieurs variables obtenues de manière cohérente et ayant des propriétés particulières. Par exemple, un histogramme est un descripteur d'une image formé de plusieurs variables (les fréquences ou effectifs de chaque classe de l'histogramme).

2.1. Paramètres classiques de rugosité

Par paramètre "classique" de rugosité qui est la terminologie couramment adoptée, on entend paramètres "standards", c'est à dire paramètres reconnus internationalement par une norme ISO. Jusqu'à maintenant, l'ensemble de ces paramètres utilisés pour caractériser une surface étaient obtenus à partir de profils, c'est à dire d'un ensemble de valeurs représentant l'altitude d'une surface selon une droite, mais de plus en plus, les progrès en terme d'acquisition et de traitement des données poussent vers une standardisation des paramètres surfaciques, c'est à dire extraits à partir d'une représentation fournissant des altitudes sur une carte bidimensionnelle.

2.1.1. Caractérisation de profils

Supposons que l'on travaille sur un profil défini par une fonction $Z(x)$. On considère classiquement la fonction Z comme la superposition de 3 échelles caractéristiques : la forme, l'ondulation et la rugosité.

La forme (shape) $S(x)$ contient tout ce qui est considéré comme une déformation à grande échelle postérieure à la production de la surface étudiée, il s'agit par exemple de la courbure globale d'un échantillon provenant de mauvais traitements (défaut de stockage, de transport, ...). L'obtention de S à partir de Z n'est pas normalisée, mais on cherche souvent à approximer au mieux Z par des polynômes de degré plus ou moins important suivant la taille des structures que l'on considère comme de la forme. On peut aussi utiliser des splines ou d'autres fonctions approximantes et on voit apparaître ici un certain degré de subjectivité quant au choix de ce qui est considéré comme de la forme.

Le profil P dit brut est simplement obtenu par différence : $P = Z - S$. On parle alors d'extraction de forme.

L'ondulation (waviness) $W(x)$ s'obtient par filtrage passe-bas à partir de P . Il s'agit d'une convolution dont la nature du noyau (gaussien par exemple) et la longueur d'onde de coupure (cut-off) sont choisis par l'utilisateur.

La rugosité R constitue simplement le résidu entre P et W : $R = P - W$.

On peut définir un grand nombre de paramètres de rugosité préfixés par une lettre X qui désigne le profil (W, R ou P) que l'on étudie. Sans prétendre à l'exhaustivité, nous citerons simplement quelques exemples de paramètres parmi les plus couramment utilisés, mais les différentes normes recensant l'ensemble des paramètres "standard" peuvent être consultées : ISO 4287, ISO 12085, ISO 13565.

$$\begin{aligned}
 Xa &= \frac{1}{n} \sum_{k=0}^n |X(k)| \\
 Xq &= \sqrt{\frac{1}{n} \sum_{k=0}^n X(k)^2} \\
 Xp &= \max_k (X(k) - E(X)) \\
 Xv &= \min_k (E(X) - X(k)) \\
 Xz &= Xp + Xv \\
 Xsk &= \frac{1}{Xq^3 \cdot n} \sum_{k=0}^n X(k)^3 \\
 Xku &= \frac{1}{Xq^4 \cdot n} \sum_{k=0}^n X(k)^4 \\
 &\vdots
 \end{aligned}$$

Où E désigne l'espérance mathématique. On peut donc constater l'existence d'une certaine richesse dans l'ensemble de ces paramètres, cependant, on observe assez vite la

2.2. Descripteurs basés sur une représentation de type "image"

limite de la pertinence de ces critères dans de nombreuses applications où les structures bidimensionnelles sont importantes, et en particulier la caractérisation de l'aspect d'une surface.

2.1.2. Paramètres surfaciques : les nouveaux standards

Cette insuffisance des paramètres de descriptions de profils incite de plus en plus les utilisateurs de surfaces à les caractériser par des paramètres issues d'une analyse bidimensionnelle [47, 36, 52, 3, 4]. La définition de standards basés sur des calculs surfaciques commence d'ailleurs à être envisagée : un nouveau groupe de travail (WG16) au sein des comités de normalisation (ISO technical committee 213) a été créé travaillant sur ces paramètres. Ce type de descripteurs bidimensionnels devrait permettre d'obtenir des résultats beaucoup plus pertinents en termes de comparaison de surface, mais la parution officielle de ces travaux n'ayant pas encore eu lieu, il reste assez difficile de s'appuyer sur ces futurs standards. Par ailleurs, les paramètres décrits sont des variables numériques souvent obtenues comme statistique d'une variable aléatoire, et dans le cas qui nous intéresse (le contrôle en ligne), ceux-ci peuvent se révéler peu discriminants sur des ensembles de surfaces du même type. Par exemple, si la statistique définie dans le standard est la moyenne quadratique des hauteurs, des études ont montré que des surfaces de qualité visuelle très différentes pouvaient présenter les mêmes valeurs. Le principal problème réside donc dans le choix de la statistique en fonction du type de surfaces à traiter, et c'est pourquoi, dans un premier temps, nous définissons de nouveaux descripteurs surfaciques formés de plusieurs variables (courbe ou histogramme par exemple) cherchant à capturer le maximum d'information sur la distribution d'une variable aléatoire, tandis que le choix de la statistique à considérer sera effectué par la suite de manière automatique en fonction des données.

2.2. Descripteurs basés sur une représentation de type "image"

Dans cette section, nous présentons des descripteurs pouvant être extraits à partir d'une représentation 2D d'une surface. Ces descripteurs nous seront utiles pour caractériser des topographies obtenues par microscopie interférométrique ou des images issues d'une caméra haute définition. Bien que ces deux types de représentation soient fondamentalement différents, elles se présentent sous la même structure informatique (image), et nous pouvons donc leur appliquer les mêmes algorithmes pour en extraire des descripteurs. Nous verrons malgré tout dans la suite que la signification physique de ces descripteurs est différente dans le cas d'une topographie ou d'une image, de même que leur comportement et leur importance vis à vis de la qualité des surfaces.

Plusieurs types de descripteurs sont décrits dans la suite : certains sont purement statistiques comme la fonction d'autocorrélation d'une image, d'autres font intervenir des fonctions morphologiques (courbe d'érosion-dilatation, segmentation de l'image, courbe de tortuosité). Enfin, les derniers descripteurs font intervenir la théorie de l'interpolation

pour extraire les vecteurs normaux à l'image considérée comme un relief topographique ou encore la théorie des ondelettes très largement utilisée dans le domaine de la caractérisation des textures.

2.2.1. Autocorrélation

La fonction d'autocorrélation permet de savoir dans quelle mesure une image et son translaté d'un vecteur quelconque se ressemblent. En général, l'image initiale et son translaté présentent un fort pourcentage de similarité si le vecteur de translation est plus petit que la taille des structures présentes dans l'image (taille de grain par exemple). Si le vecteur de translation augmente beaucoup, l'image et le translaté n'ont plus rien en commun, sauf en cas de périodicité exacte à l'intérieur de l'image (nous supposons que ce n'est pas le cas), et la fonction de corrélation s'annule. La norme du vecteur au-delà de laquelle la fonction d'autocorrélation s'annule est appelée *portée*.

Mathématiquement parlant, on exprime cette fonction en terme d'espérance. On définit tout d'abord la fonction de covariance pour une fonction aléatoire stationnaire $I(x)$ de moyenne $m = E(I(x))$:

$$C(\vec{h}) = E(I(x).I(x + \vec{h})) - m^2$$

La stationnarité de la fonction I assure que la fonction $C(\vec{h})$ ne dépend pas du point x . En pratique d'ailleurs, on supposera vérifiée l'hypothèse d'ergodicité : l'espérance $E(I(x).I(x + \vec{h}))$ et la moyenne m sont estimées par moyenne spatiale sur la surface de l'image. La fonction d'autocorrélation se définit alors en normalisant la covariance par la variance de l'image :

$$Cr(\vec{h}) = \frac{C(\vec{h})}{\sigma_I^2}$$

σ_I désigne l'écart-type du niveau de gris, calculé sur l'ensemble des points de l'image I (hypothèse ergodique) :

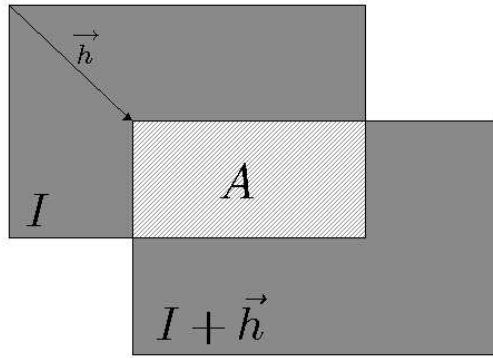
$$\sigma_I = \sqrt{E(I^2) - (E(I))^2}$$

Supposons les images centrées, c'est à dire de moyenne nulle. On peut calculer la fonction C de plusieurs façons différentes. La première consiste à considérer les surfaces comme périodiques et à calculer l'espérance du produit de l'image et de son translaté compte tenu de cette hypothèse. La transformée de Fourier rapide (FFT) de la fonction de covariance est alors le module de la transformée de Fourier rapide de l'image.

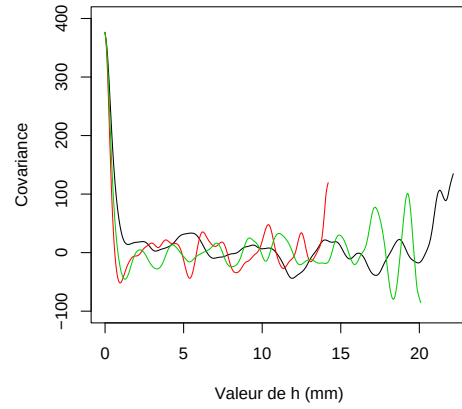
Cependant, cette méthode classique de périodisation introduit un biais dans le calcul de la covariance, qui sera principalement détectable lorsque la question de la "variance d'extension" sera abordée. La deuxième méthode consiste à ne calculer l'espérance du produit que sur la zone de recouvrement disponible entre l'image et son translaté de $\vec{h}$, ce qui n'introduit pas de biais directement mais limite le nombre de points de calcul de l'espérance pour les fortes valeurs de $\vec{h}$ et donc en affaiblit la représentativité statistique (figure 2.1(a)). En pratique, si la taille de l'image est grande devant la portée, l'annulation de la fonction de corrélation intervient alors que la zone d'intégration est encore assez étendue pour être représentative, et c'est donc cette méthode qui sera privilégiée.

2.2. Descripteurs basés sur une représentation de type "image"

Pour construire un descripteur, on utilise en général la fonction d'autocorrélation dans une direction donnée plutôt que la covariance afin de s'affranchir de problèmes de normalisation sur les images. La courbe obtenue est tronquée de façon à ne conserver que les valeurs de h pour lesquelles la surface de recouvrement est suffisante pour assurer la pertinence du calcul de l'espérance (50% de la surface de l'image est une valeur typique). Le résultat est une courbe dont le comportement à l'origine caractérise les structures de très petite taille, la quantité de bruit par exemple (effet de pépite), et dont la décroissance caractérise les structures à plus grande échelle présentes dans l'image et leur éventuelle périodicité.



(a) Schéma de calcul de la fonction de covariance lorsque l'image n'est pas périodisée : on calcule la moyenne du produit $I(x).I(x + \vec{h})$ sur l'ensemble de recouvrement A



(b) Fonction de covariance d'une image dans trois directions (direction X : courbe noire, direction Y : courbe rouge et première diagonale : courbe verte)

FIG. 2.1.: Autocorrélation

2.2.2. Erosion-Dilatation

L'érosion et la dilatation constituent les opérations de base de la morphologie mathématique. La famille de courbes que nous nous proposons d'utiliser rentre dans la catégorie des courbes dites "granulométriques" qui cherchent à caractériser la distribution de la taille des structures présentes dans une image. L'idée d'utiliser ces fonctions comme descripteur d'une surface rugueuse a été originellement présentée dans [1, 3] : il s'agit d'observer l'évolution d'une image lors de l'itération d'opérations morphologiques.

On pose :

$$\begin{aligned} I_0^d &= I_0^e = I \\ I_{n+1}^d &= \delta(I_n^d) \end{aligned}$$

$$\begin{aligned}
 I_{n+1}^e &= \epsilon(I_n^d) \\
 R_n^d &= I_{n+1}^d - I_n^d \\
 R_n^e &= I_n^e - I_{n+1}^e \\
 e_n &= E(R_n^e) \\
 d_n &= E(R_n^d)
 \end{aligned}$$

Les itérés successifs d'érosion et de dilatation convergent en un nombre fini d'itérations (disons N) vers des images constantes égales respectivement au minimum et au maximum de l'image initiale. La courbe ED_p (figure 2.3) qui nous intéresse est formée par les suites de valeurs (e_n) et (d_n) :

$$\begin{aligned}
 \forall p \in [1, N] \quad , \quad ED_p &= e_{N-p} \\
 \forall p \in [N+1, 2N] \quad , \quad ED_p &= d_{p-(N+1)}
 \end{aligned}$$

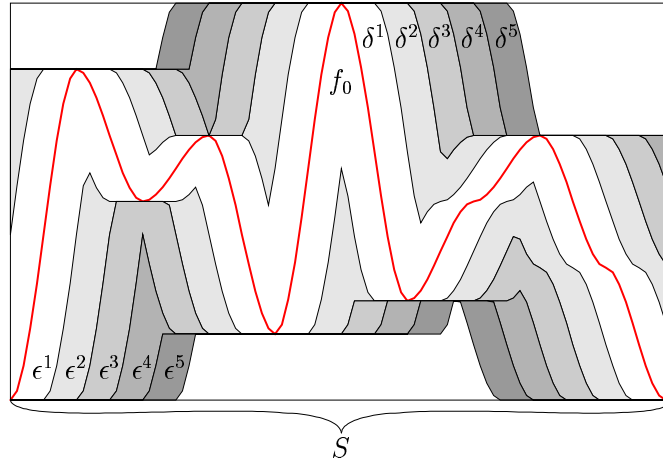


FIG. 2.2.: Courbes d'érosion $(\epsilon^1, \dots, \epsilon^5)$ et de dilatation $(\delta^1, \dots, \delta^5)$ d'un profil f_0 . Les résidus R_i^d et R_i^e correspondent à la différence entre deux courbes successives. Les volumes grisés correspondent à $e_i.S$ ou $d_i.S$ (S correspond à la longueur (surface en 2D) du support de f_0)

Pour mieux comprendre la signification de cette courbe, on peut remarquer que $e_n.S$ (S désigne la surface de l'image) correspond à la proportion du volume compris entre la surface originale et l'image constante égale à son minimum (image obtenue à la N ème étape d'érosion), tandis que $d_n.S$ correspond à la proportion du volume compris entre la surface originale et l'image constante égale à son maximum (image obtenue à la N ème étape de dilatation). La courbe ED_p représente donc une répartition en plusieurs classes du volume compris entre maximum et minimum de la fonction originale (voir figure 2.2). Cette interprétation en termes statistiques sera très utile dans la suite, lorsque l'on cherchera à comparer des surfaces sur la base de ces courbes.

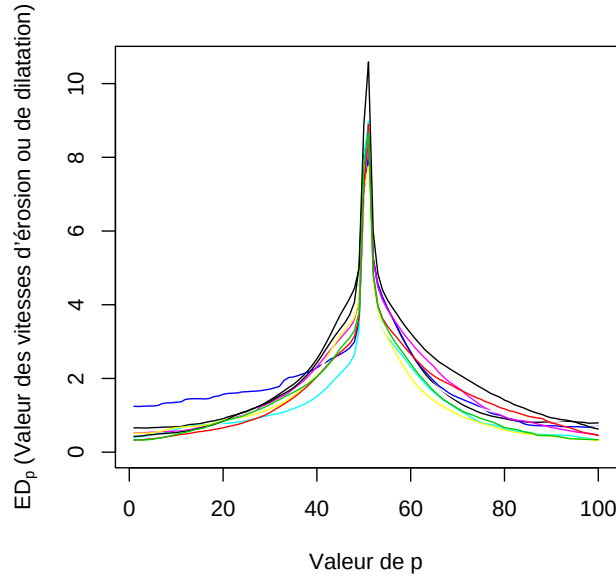


FIG. 2.3.: Exemple de courbes d'érosion-dilatation obtenues sur des surfaces de type peau d'orange (figure 2.4(a)).

2.2.3. Segmentation

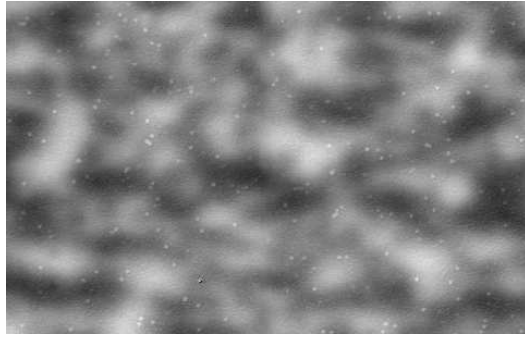
Ces descripteurs ont été développés initialement pour caractériser les surfaces de type peau d'orange (voir figure 2.4(a)) qui présentent un aspect granuleux et vallonné assez marqué. L'objectif est de diviser l'image en zones claires et sombres séparées par des contours cohérents. Pour y parvenir, un schéma classique de morphologie mathématique consiste à travailler sur une ligne de partage des eaux de l'image gradient [14, 11, 10], bien que le bruit et l'irrégularité dans les images imposent un filtrage préalable important.

a. Préfiltrage

L'opération de segmentation commence par un léger préfiltrage (filtrage gaussien puis ouverture-fermeture de petite taille par exemple) pour améliorer la qualité des images et éliminer quelques impuretés comme des tâches de poussières parfois présentes sur certaines topographies (voir figure 2.4).

b. Extractions de marqueurs préliminaires

Une ouverture par h-reconstruction [59, 60, 61, 63] consiste traditionnellement à soustraire une constante λ à l'image de départ, puis à reconstruire l'image ainsi obtenue par dilatation géodésique sous le masque formé par l'image originale. Plus concrètement, les maxima sont étêtés et remplacés par des zones plates, ce qui induit la disparition



(a) Image originale de type "peau d'orange" (Pdo 14)



(b) Image issue de l'opération de préfiltrage

FIG. 2.4.: Préfiltrage

des maxima dont la dynamique est plus faible que λ . L'opération duale (fermeture par h-reconstruction) qui consiste à remplir les vallées et à remplacer les minima par des zones plates peut être obtenue très simplement en appliquant le même algorithme à l'image inversée puis en prenant finalement l'inverse du résultat obtenu. Ces deux opérations peuvent être appliquées successivement pour obtenir un effet plus symétrique sur les maxima et les minima (filtre alterné par h-reconstruction).

L'algorithme de segmentation proprement dit débute donc par un filtre alterné par h-reconstruction de faible amplitude (λ faible) et par une extraction des maxima et minima régionalisés de l'image filtrée (voir figure 2.5(c)). Ces extrema serviront ensuite de marqueurs pour l'obtention de zones cohérentes, le filtrage par reconstruction sert donc dans ce cas à limiter le nombre d'extrema retenus (donc le nombre de zones dans l'image finale).

c. Calcul d'un gradient de l'image originale

On applique ensuite un nouveau filtre alterné par h-reconstruction à l'image de départ avec une valeur de λ beaucoup plus importante, ce qui a pour effet de créer de larges zones plates (figure 2.5(a)), et de simplifier considérablement le gradient de l'image (figure 2.5(b)).

Il reste, comme souvent, à déterminer quelle valeur de λ s'adapte le mieux à nos images : si λ est trop important, certains minima et maxima pertinents vont disparaître, et inversement, si λ est trop faible, le gradient sera bruité, et la segmentation pourra faire apparaître des contours contre nature. Par ailleurs, il est important que ce paramètre puisse s'adapter de manière automatique à des images, certes du même type, mais dont les caractéristiques peuvent varier fortement en fonction des différents paramètres du processus de production. En particulier, le Ra (voir 2.1) et l'écart-type de nos images peuvent être assez différents d'une image sur l'autre. Or, ce qui importe dans le choix de λ , c'est la dynamique des extrema que l'on considère comme pertinents, et cette dynamique est fortement corrélée avec l'écart-type de l'image. De manière heuristique,

2.2. Descripteurs basés sur une représentation de type "image"

on peut dire que les pics et les vallées sont répartis autour de la moyenne de l'image avec un écartement moyen que l'on peut comparer à l'écart-type. De cette remarque provient l'idée de choisir un paramètre de filtrage λ proportionnel à l'écart-type de l'image.

Si on note I_r l'image filtrée qui servira de base au calcul du gradient, on obtient :

$$I_r = \mathcal{R}_I(I - k \cdot \sigma_I, B)$$

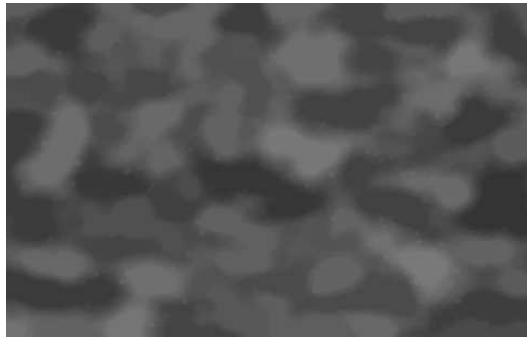
Où k est un paramètre ajusté de manière empirique sur un ensemble d'images représentatif des images à traiter, et $\mathcal{R}_I(J)$ représente l'opération de reconstruction de l'image J sous l'image I . L'influence du choix de k dans le résultat final peut s'observer sur la figure 2.6 : plus k est important, plus les zones plates sont larges et plus le nombre de zones est faible.

d. Ajustement des marqueurs et ligne de partage des eaux

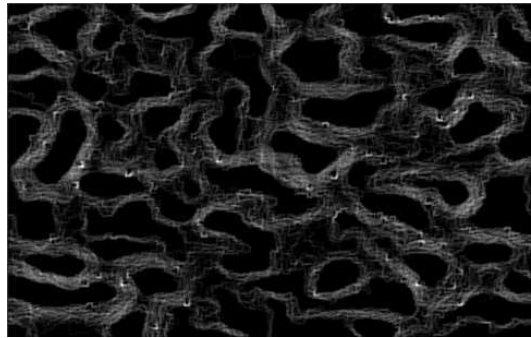
L'image gradient de l'image I_r (figure 2.5(b)) fait donc ressortir les contours les plus marqués de l'image de départ et comporte de larges zones où le gradient est nul (zones qui correspondent aux zones plates de I_r) (figure 2.5(d)). Cependant, appliquer directement une ligne de partage des eaux à cette image conduit à une sur-segmentation de certaines zones de transition (figure 2.6(d)). C'est pourquoi les extrema calculés lors de la première étape de filtrage (faible filtre par reconstruction) vont être introduits comme marqueurs, ce qui permettra de limiter le nombre de zones au nombre d'extrema. Là encore, le résultat n'est pas toujours pertinent. Il se trouve que le nivellement important peut regrouper dans une même zone plate plusieurs extrema de faible dynamique, qui se retrouveront alors séparés artificiellement lors du calcul de la ligne de partage des eaux de l'image gradient (simplement par une ligne médiane à l'intérieur d'une zone plate de l'image filtrée I_r) (figure 2.6(c)). Dès lors, il s'agit de regrouper ces zones artificielles : au lieu d'utiliser les extrema directement, on utilise le résultat de la reconstruction de ces extrema à l'intérieur de l'ensemble formé par les zéros du gradient de l'image I_f . Le résultat devient alors satisfaisant (figure 2.6(a) et 2.6(b)). On peut s'interroger malgré tout sur le choix des marqueurs : pourquoi ne pas utiliser directement les extrema régionaux de l'image I_r ? La raison est la suivante : le filtrage par reconstruction crée certes de larges zones plates, ce qui a pour effet de simplifier et contraster l'image gradient, mais il diminue aussi considérablement le nombre d'extrema : certaines zones plates de I_r ne sont pas des extrema, mais de larges zones de transition. Le résultat de la segmentation en utilisant les extrema de I_r comme marqueur fournit donc un résultat médiocre.

e. Construction de descripteurs

Cette segmentation ne constitue pas un descripteur de la surface en soi, mais elle permet d'en construire plusieurs à partir de la division en zones obtenue. En effet, chaque zone z peut-être labellisée par un paramètre X_z qui correspond par exemple au volume, à la surface ou à la profondeur de la zone considérée. On peut ensuite s'intéresser à la variable aléatoire X et construire un descripteur de la surface à partir de l'histogramme des valeurs de X prises par chacune des zones d'une surface. Cependant, on s'intéressera



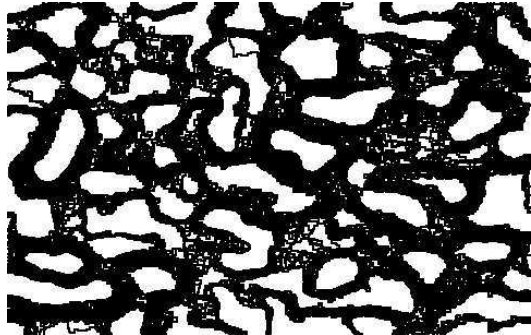
(a) Image issue de l'opération de reconstruction. Le contraste est plus faible que dans l'image initiale, mais de larges zones plates ont été créées et les contours sont plus marqués



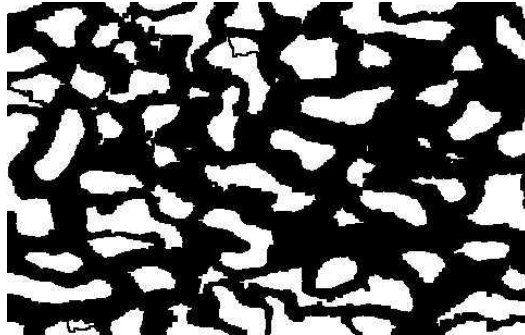
(b) Image gradient de l'image filtrée



(c) Extrêma retenus



(d) Zones plates du gradient



(e) Marqueur final

FIG. 2.5.: Etapes de segmentation

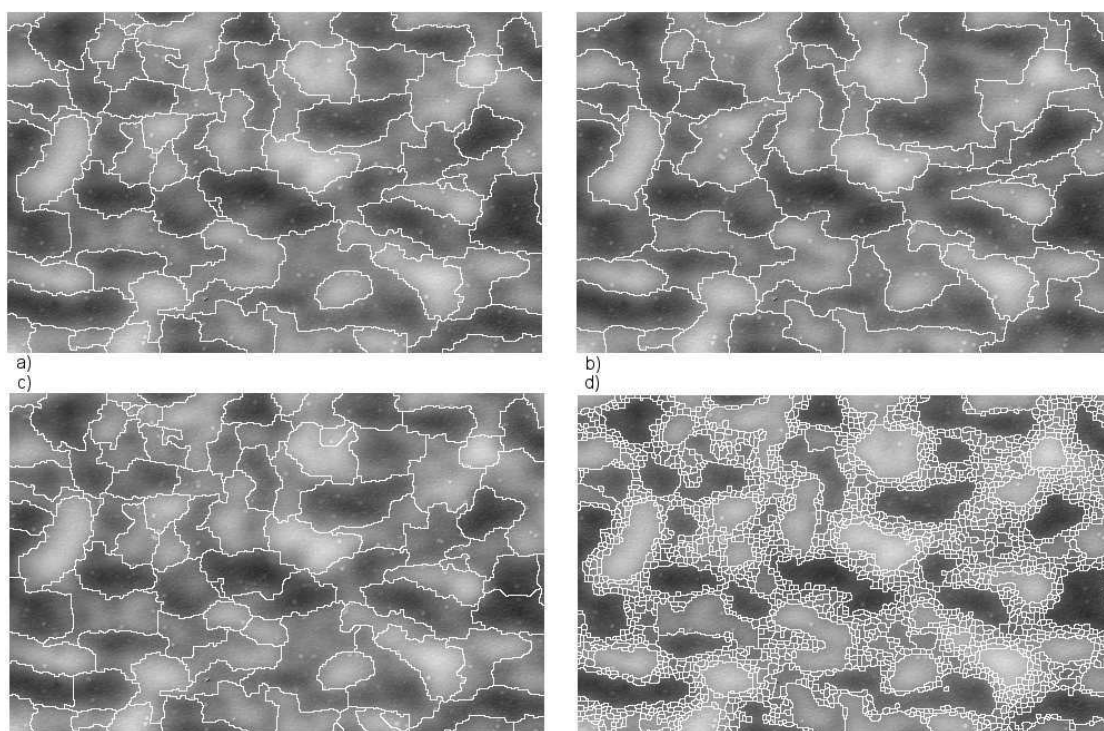


FIG. 2.6.: Comparaison des résultats pour différents paramètres. a : résultat avec $k=1$, b : résultat avec $k=1.5$ (les zones sont plus grossières), c : les marqueurs sont directement les extrema (on remarque certaines divisions peu naturelles), d : résultat sans marqueur dans le calcul du gradient (sursegmentation des zones de transition)

généralement à la probabilité induite sur l'image par la variable aléatoire X et le découpage en zone : on attribue à chaque pixel p une valeur X_p correspondant à la valeur X_z de la zone à laquelle il appartient et on construit ainsi un histogramme des valeurs X_p (cela revient de manière équivalente à attribuer à chaque zone un poids proportionnel à sa surface dans l'histogramme de X_z).

2.2.4. Tortuosité

Dans cette section, on cherche à comparer deux reliefs topographiques (ou deux images assimilées à des reliefs topographiques) sur la base de la propagation d'une fonction distance à l'intérieur de ce relief. L'analogie à l'échelle humaine serait de comparer deux zones montagneuses par la facilité de leur traversée, en effet, suivant l'agencement des monts et des vallées, les chemins franchissant les montagnes comporteront plus ou moins de cols ou d'ascensions, ce qui se traduira par une longueur de chemin et un dénivelé total plus ou moins important. Le calcul de la distribution des distances séparant un ensemble source d'un ensemble d'arrivée est ici envisagé comme un descripteur de la texture d'une surface.

Dans ce qui suit, nous nous intéressons à plusieurs types de distance ou pseudo-distance pouvant être utilisés pour caractériser une surface.

Soit $I(x, y)$ une image que l'on pourra aussi concevoir comme une fonction discrète associant à un pixel p une valeur $I(p)$. Soit $C = p_0 p_1 \dots p_n$ un chemin de l'image. La longueur de C est défini à partir des longueurs élémentaires séparant deux pixels voisins :

$$l(C) = \sum_{i=0}^{n-1} l(p_i p_{i+1})$$

On définit alors les longueurs suivantes entre deux pixels voisins p_1 et p_2 :

- Longueur 2D (longueur de référence) :

$$\begin{aligned} le(p_1 p_2) &= \|p_1 p_2\|_{2D} \\ \|p_1 p_2\|_{2D} &= 1 \end{aligned}$$

La norme utilisée $\|\cdot\|_{2D}$ peut bien sûr varier, mais pour des raisons de simplicité, nous avons choisis ici la norme 1, en supposant équivalentes les échelles utilisées dans les directions X et Y : la longueur d'un chemin est donc dans ce cas directement liée au nombre de pixels le composant. En pratique, cette longueur ne dépend pas de l'image I considérée et sera utilisée à des fins de normalisation de distance.

- Longueur "pondérée" :

$$lp(p_1 p_2) = I(p_2) - \min(I)$$

Le terme $\min(I)$ assure la positivité des valeurs pour tous les pixels de l'image, ce qui assurera le bon fonctionnement de l'algorithme de calcul du plus court chemin. Cette distance "somme des niveaux de gris" a pour particularité d'attribuer une faible longueur aux chemins de fond de vallée, et une forte longueur au chemin de crête. On peut aussi obtenir l'effet opposé en inversant l'image avant le calcul : $I(p_2) - \min(I)$ devient $\max(I) - I(p_2)$.

2.2. Descripteurs basés sur une représentation de type "image"

- Longueur "dénivelé" :

$$ld(p_1, p_2) = |I(p_2) - I(p_1)|$$

Dans ce cas, on associe à un chemin une longueur correspondant au dénivelé total franchi en positif comme en négatif. Les chemins de faible longueur seront donc les chemins restant "à flanc de coteaux".

- Longueur curviligne :

$$lc(p_1 p_2) = \sqrt{|I(p_2) - I(p_1)|^2 + \|p_1 p_2\|_{2D}^2}$$

Il s'agit de la longueur curviligne sur la surface (dans l'espace 3D). On pourrait éventuellement pondérer les deux termes sous la racine carrée pour attribuer plus de poids à la longueur 2D ou à la longueur dénivelé.

Ces définitions permettent de définir la pseudo-distance dans l'espace discret entre un pixel a et un ensemble de pixel S par la relation :

$$dx(a, S) = \inf_{m \in S, C \in chem(a, m)} lx(C)$$

Où $chem(a, m)$ désigne l'ensemble des chemins joignant le pixel a au pixel m . Le x de dx et lx peut ici être remplacé par r, p, d ou c en fonction des définitions précédentes. On définit $dx(a, S)$ de manière identique. Les fonctions dx vérifient toutes l'inégalité triangulaire, ce qui est essentiel pour l'implémentation de l'algorithme du calcul de dx , mais ne sont pas toutes des distances au sens mathématique du terme, par exemple $dd(a, S) = 0 \nRightarrow a \in S$, ou encore $dp(p_1, p_2) \neq dp(p_2, p_1)$. Ces aspects ne sont pas gênants pour le calcul de la fonction distance par file d'attente hiérarchique.

a. Implémentation

On se donne une image I , une longueur lx , un ensemble de pixel S et on cherche ensuite à calculer $dx(p, S)$ pour chaque pixel de l'image. L'implémentation s'effectue au moyen de files d'attentes hiérarchiques [45] selon un algorithme s'apparentant à celui de Moore-Dijkstra. Dans ce qui suit, $\mathcal{V}(p)$ désigne l'ensemble des voisins de p pour la connexité choisie.

Etapes de l'algorithme :

- 1: **Initialisation :**
- 2: **Pour tout** $p \in S$ **Faire**
- 3: empile p à une priorité 0
- 4: **Fin Pour**
- 5: **Boucle principale :**
- 6: **Tant que** la pile est non vide **Faire**
- 7: On dépile le premier pixel p de la file
 ▷ p est le point empilé de priorité la plus faible, le premier dans l'ordre d'empilage en cas d'égalité
- 8: $pr \leftarrow$ la priorité de p
- 9: **Si** p n'a jamais été dépilé **Alors**
- 10: $dx(S, p) \leftarrow pr$
- 11: **Pour** $v \in \mathcal{V}(p)$ **Faire**
- 12: **Si** v n'a jamais été dépilé **Alors**
- 13: empile v avec une priorité $pr + lx(p, v)$
- 14: **Fin Si**
- 15: **Fin Pour**
- 16: **Fin Si**
- 17: **Fin Tant que**

On peut remarquer qu'un pixel p peut être empilé plusieurs fois (au plus, une fois pour chacun de ses voisins et une fois au début), mais on ne prend en compte p qu'une seule fois (la première fois qu'il est dépilé) pour le parcours des voisins. Soit p un pixel de l'image, soit $C_p = p_0 p_1 \dots p_n p$ le chemin reliant p à S de longueur minimum. On peut montrer que $lx(C_p)$ est égale à la première priorité $pr(p)$ avec laquelle p est dépilé.

Preuve de l'algorithme. La preuve s'effectue en plusieurs temps :

- On prouve dans un premier temps que tout les pixels de l'image sont traités, et que l'algorithme se termine
- On prouve ensuite que $pr(p) \leq d(p, S)$.
- On montre enfin par récurrence que $pr(p) \geq d(p, S)$.

1. Terminaison de l'algorithme

La preuve du premier point peut se faire par l'absurde : Supposons qu'un pixel p n'est pas traité, c'est à dire n'est jamais dépilé, et donc n'a jamais été empilé. Considérons un chemin joignant un pixel quelconque $s \in S$ à p : $C = s p_1 \dots p_n p$. Si p n'a jamais été empilé, cela signifie qu'aucun de ses voisins n'a été dépilé, en particulier p_n . p non traité implique donc p_n non traité. En itérant le raisonnement sur l'ensemble des pixels de C , on arrive à s non traité, ce qui est incompatible avec l'étape d'initialisation de l'algorithme. Donc tous les points sont traités, et on peut définir pour tout pixel p de l'image la fonction $pr(p)$ qui attribue à un pixel la première priorité avec laquelle il est dépilé.

2. Preuve de $pr(p) \leq d(p, S)$

2.2. Descripteurs basés sur une représentation de type "image"

La preuve du second point fait intervenir la définition de $pr(p)$. On commence par définir l'ordre de dépillement $o(p) : o(p_1) < o(p_2)$ signifie que p_1 est dépilé avant p_2 . On obtient en particulier les propriétés suivantes :

$$o(p_1) < o(p_2) \Rightarrow pr(p_1) \leq pr(p_2) \quad (2.1)$$

$$pr(p_1) < pr(p_2) \Rightarrow o(p_1) < o(p_2) \quad (2.2)$$

Un point p n'appartenant pas à S est systématiquement empilé par ses voisins tant que p n'a pas encore été dépilé. $pr(p)$ est la plus petite des priorités avec laquelle p est dépilé :

$$pr(p) = \inf_{v \in \mathcal{V}(p), o(v) < o(p)} (pr(v) + lx(v, p))$$

Cependant, pour les points v de $\mathcal{V}(p)$ qui seront dépilés après p , on obtient :

$$pr(v) + lx(v, p) \geq pr(p) + lx(v, p) \geq pr(p)$$

La première inégalité provient de la propriété 2.1, et la seconde de la positivité de la fonction lx . Cette remarque permet d'élargir la définition de pr :

$$pr(p) = \inf_{v \in \mathcal{V}(p)} (pr(v) + lx(v, p))$$

Pour tout chemin $C = sp_0 \dots p_n p$ reliant p à S , on obtient donc :

$$\begin{aligned} pr(p) &\leq pr(p_n) + lx(p_n, p) \\ pr(p) &\leq pr(p_{n-1}) + lx(p_{n-1}, p_n) + lx(p_n, p) \\ &\vdots \\ pr(p) &\leq pr(s) + lx(s, p_1) + \dots + lx(p_{n-1}, p_n) + lx(p_n, p) \\ pr(p) &\leq lx(C) \end{aligned}$$

Ceci est valable pour tous les chemins reliant p à S , donc en particulier pour les chemins réalisant le minimum de $lx(C)$, c'est à dire la distance de p à S . On obtient donc pour tout pixel p de l'image :

$$pr(p) \leq dx(S, p)$$

3. Preuve de $pr(p) \geq d(S, p)$

La dernière partie de la preuve s'effectue par récurrence sur l'ensemble des pixels ordonnés selon l'ordre o .

L'hypothèse (H) de récurrence est la suivante :

Pour un pixel p donné, $pr(p) \geq dx(S, p)$.

Initialisation :

Les premiers pixels dans l'ordre de o sont les pixels de priorité la plus faible possible, c'est à dire 0, qui ont été empilés les premiers, c'est à dire les points de S qui sont empilés pendant la phase d'initialisation. Tous ces pixels $p \in S$ sont dépilés avec

une priorité 0 qui correspond à la distance $dx(S, p) = 0$.

Réurrence :

Supposons (H) vraie pour les k premiers pixels dans l'ordre de o , et considérons p tel que $o(p) = k + 1$. p a été empilé par un voisin v_0 tel que $o(v_0) < o(p)$, en appliquant l'hypothèse de récurrence à v , on peut écrire :

$$\begin{aligned} pr(p) &= pr(v_0) + l(p, v_0) \\ pr(p) &\geq dx(v_0, S) + l(p, v_0) \\ pr(p) &\geq dx(v_0, S) + dx(p, v_0) \\ pr(p) &\geq dx(p, S) \end{aligned}$$

Le passage de la première à la seconde ligne fait explicitement intervenir (H) , le passage de la seconde à la troisième fait intervenir la définition de $dx(p, v_0)$ comme borne inférieure des longueurs de chemin joignant p à v_0 , et enfin, le passage de la troisième à la quatrième ligne provient de l'inégalité triangulaire, qui elle-même découle trivialement de la positivité des longueurs des chemins. L'application du théorème de récurrence permet de conclure pour tout pixel p de l'image :

$$pr(p) \geq dx(S, p)$$

Ceci termine la preuve de la relation $dx(p, S) = pr(p)$ pour l'ensemble des pixels de l'image, et assure donc le bon fonctionnement de l'algorithme.

□

b. Normalisation

Pour pouvoir comparer ces distances d'une image à l'autre (en particulier si les images de départ n'ont pas la même taille ou si les ensembles de départ sont différents), on normalise la distance obtenue par une distance 2D "mise à l'échelle" en la multipliant par la distance moyenne entre deux pixels voisins de l'image. Le résultat de cette opération de normalisation est appelé carte des "tortuosités" : plus la tortuosité est forte, plus le chemin obtenu est long par rapport à la distance classique, donc plus il est sinueux -tortueux- et réciproquement si la tortuosité est faible. Cependant, il est important de noter que le terme "sinueux" ou "tortueux" est pris ici dans un sens assez large, car une forte tortuosité peut-être provoquée par un chemin comportant beaucoup de pixels avec des poids normaux, ou au contraire un chemin comportant peu de pixels, mais avec des poids très élevés.

Par exemple, dans le cas de la distance "somme des dénivelés" ($dd(,)$), un point m obtiendra une tortuosité $td(m)$ obtenue par la relation :

$$\begin{aligned} td(m) &= \frac{dd(m, S)}{k_d \cdot dr(m, S)} \\ k_d &= E(I(p) - I(n), p \in I, n \in \mathcal{V}(p)) \end{aligned}$$

2.2. Descripteurs basés sur une représentation de type "image"

Où E désigne l'espérance mathématique calculée sur l'ensemble de l'image I . Cela revient dans ce cas précis à comparer la distance dd obtenue à la valeur que l'on obtiendrait si la surface était simplement un plan de pente constante (égale au dénivelé moyen entre deux pixels voisins de I) entre l'ensemble de départ et l'ensemble d'arrivée.

Pour la distance pondérée (somme des niveaux de gris) ($dp(,)$), $tp(m)$ est obtenue par :

$$tp(m) = \frac{dp(m, S)}{k_p \cdot dr(m, S)} \quad (2.3)$$

$$k_p = E(I + \min(I)) \quad (2.4)$$

c. Construction d'un descripteur

On construit simplement le descripteur basé sur la carte des tortuosités en choisissant un ensemble d'arrivée dans l'image étudiée et en considérant l'histogramme des tortuosités obtenues sur cet ensemble pour un ensemble de départ donné. Grâce à la normalisation, les résultats obtenus devraient pouvoir s'affranchir quelque peu du choix des ensembles de départ et d'arrivée et de la taille des images, en pratique cependant, il reste plus pertinent de calculer la carte des tortuosités de manière cohérente pour l'ensemble des images étudiées.

d. Exemple de carte des tortuosités

Dans le cas qui nous intéresse (les surfaces Pdo), la source choisie est la ligne supérieure de l'image tandis que l'ensemble d'arrivée en est la ligne inférieure. Les résultats obtenus pour une surface test (figure 2.7(a)) sont visibles sur les figures 2.7(b) et 2.7(c) pour des distances de type pondérée et "somme des dénivelés" respectivement.

2.2.5. Vecteurs Normaux

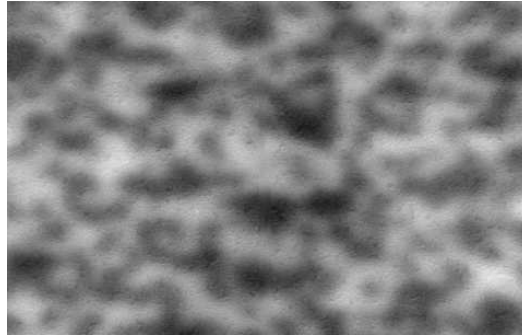
Une surface observée au microscope interférométrique se présente sous la forme d'une carte topographique donnant les altitudes des points de la surface sur une grille régulière. Dans cette section, nous chercherons à estimer le vecteur normal en chaque point de la surface à partir de ces données discrètes. La méthode de calcul repose sur l'interpolation ou l'approximation¹ de la surface sur un voisinage par une surface analytique régulière (figure 2.8(a)) permettant de calculer le vecteur normal par simple dérivation.

On pourra ensuite construire des histogrammes à partir de la carte des vecteurs normaux pour obtenir plusieurs descripteurs de la surface (histogramme des pentes ou histogramme des orientations du vecteur normal dans le plan (X, Y)).

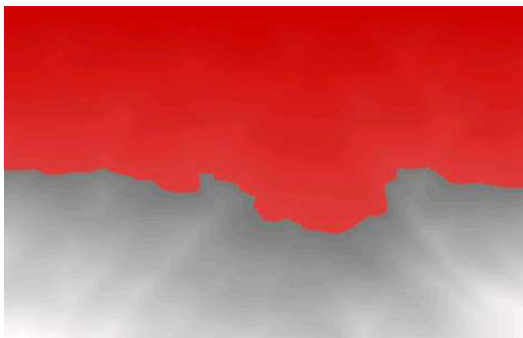
a. Théorie de l'approximation

On considère un voisinage V formé de N points $P_0, P_1, \dots, P_{N-1}$ de $\mathbb{R}^u$ appartenant au support de l'image considérée. On cherche une surface analytique $\mathcal{S}$ interpolant ou ap-

¹ voir par exemple le site web <http://aips2.nrao.edu/docs/notes/notes/>, note 224



(a) Image originale



(b) Exemple de carte de distance pondérée



(c) Exemple de carte de distance "somme des dénivelés"

FIG. 2.7.: Cartes des tortuosités. La ligne rouge représente une ligne de niveau pour la distance considérée sur chacune des figures 2.7(b) et 2.7(c)

2.2. Descripteurs basés sur une représentation de type "image"

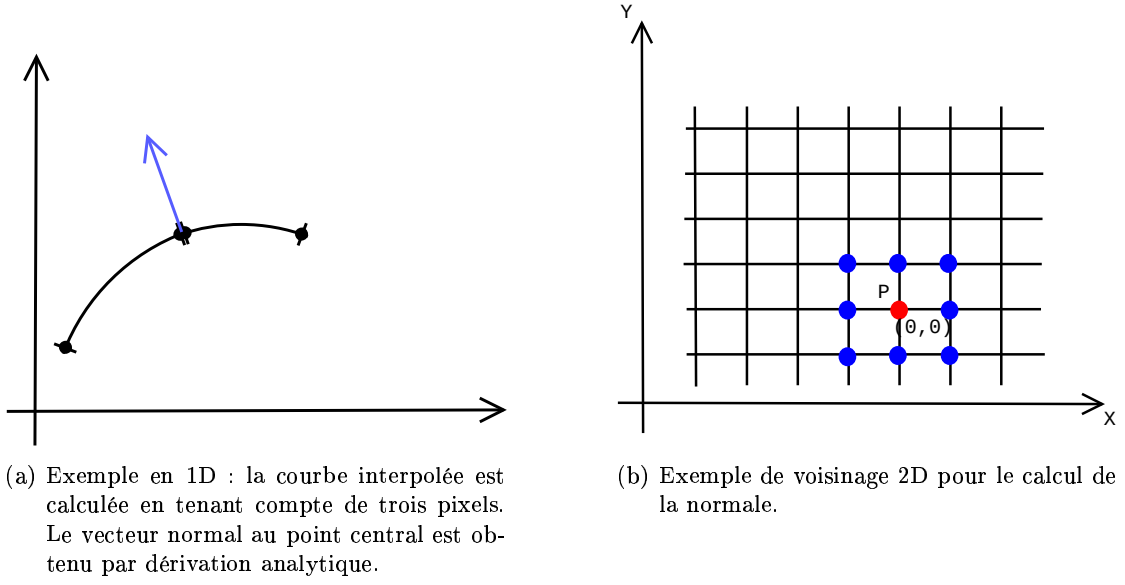


FIG. 2.8.:

proximant au mieux les points de $\mathbb{R}^{u+1}$: $M_0 = (P_0, I(P_0)), M_1 = (P_1, I(P_1)), \dots, M_{N-1} = (P_{N-1}, I(P_{N-1}))$.

Posons tout d'abord quelques notations : On suppose que la surface $\mathcal{S}$ peut s'exprimer comme une contrainte de la forme

$$M \in \mathcal{S} \Leftrightarrow \mathcal{L}(M) = \sum_{k=0}^{n-1} x^k \cdot f_k(M)$$

Où M représente un point de l'espace $\mathbb{R}^{u+1}$. En pratique, f_k , $k = 0..n-1$ et $\mathcal{L}$ sont des fonctions C^1 de $\mathbb{R}^{u+1} \rightarrow \mathbb{R}$. $x^k, k = 0..n-1$ est une série de coefficients réels.

On note alors :

$$\begin{aligned} a_i &= \begin{pmatrix} f_1(M_i) \\ \vdots \\ f_n(M_i) \end{pmatrix} \\ \forall i \in [0, N-1], y_i &= a_i \cdot x = \sum_{k=1}^n x^k \cdot f_k(M_i) \\ l_i &= \mathcal{L}(M_i) \\ l &= \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix} \end{aligned}$$

Chapitre 2. Descripteurs pour les représentation 2D

On cherche ensuite à minimiser la fonction de coût :

$$\chi^2(x) = \sum_{i=0}^{N-1} \left[\frac{l_i - y_i}{\sigma_i} \right]^2$$

En d'autres termes, on cherche le vecteur x de coefficients de façon à ce que les points de V soient les plus proches possibles² de la surface définie par combinaison linéaire à partir des fonctions f_k de la base choisie. Les termes $1/\sigma_i^2$ permettent de pondérer³ l'importance donnée à chacun des points de V .

Pour minimiser la fonction de coût, on cherche à annuler ses dérivées partielles :

$$\forall k \in [0, n-1], \frac{\partial \chi^2}{\partial x^k} = 0 \quad (2.5)$$

Ceci peut s'écrire :

$$\forall k \in [0, n-1], \sum_{i=0}^{n-1} \frac{l_i - y_i}{\sigma_i^2} \cdot \frac{\partial y_i}{\partial x^k} = 0$$

On appelle A la matrice des coefficients a_i^k :

$$A = \begin{pmatrix} a_0^0 & \dots & a_0^{n-1} \\ \vdots & & \vdots \\ a_{N-1}^0 & \dots & a_{N-1}^{n-1} \end{pmatrix}$$

On définit ensuite les matrices Q , C et L par :

$$\begin{aligned} Q &= \begin{pmatrix} \sigma_0^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{N-1}^2 \end{pmatrix} \\ C &= A^t \cdot Q^{-1} \cdot A \\ L &= A^t \cdot Q^{-1} \cdot l \end{aligned}$$

Les équation 2.5 peuvent alors se résumer sous forme matricielle par :

$$C \cdot x = L \quad (2.6)$$

Si la matrice C est inversible, on peut exprimer x et donc obtenir une surface analytique approximant au mieux les points de V au sens de la fonction de coût. En revanche, s'il n'y a pas suffisamment de conditions indépendantes, C n'est pas inversible, et le problème n'a pas de solution unique. On peut alors choisir d'imposer aux coefficients x^k de vérifier des conditions supplémentaires exprimée sous la forme :

$$\Phi_j(x) = 0, \quad j = 0, \dots, p-1$$

²au sens de la fonction de coût

³ σ_i peut être proportionnel à la distance à un point donné par exemple

2.2. Descripteurs basés sur une représentation de type "image"

Où chacun des Φ_j est une contrainte linéaire. En utilisant la méthode des multiplicateurs de Lagrange, on obtient les $n + p$ équations :

$$\sum_{k=0}^{n-1} \frac{\partial \Phi_j}{\partial x^k} dx^k = 0, \quad j = 0, \dots, p-1 \quad (2.7)$$

$$\frac{\partial \chi^2}{\partial x^k} + \sum_{j=0}^{p-1} \lambda_j \frac{\partial \Phi_j}{\partial x^k} = 0, \quad k = 0, \dots, n-1 \quad (2.8)$$

En définissant la matrice B par :

$$B = \begin{pmatrix} \frac{\partial \Phi_0}{\partial x_0} & \cdots & \frac{\partial \Phi_{p-1}}{\partial x_0} \\ \vdots & & \vdots \\ \frac{\partial \Phi_0}{\partial x_{n-1}} & \cdots & \frac{\partial \Phi_{p-1}}{\partial x_{n-1}} \end{pmatrix}$$

On peut résumer les conditions de (2.7) et (2.8) sous forme matricielle par :

$$\begin{pmatrix} C & B \\ B^t & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} L \\ 0 \end{pmatrix} \quad (2.9)$$

Si les contraintes ont été astucieusement choisies, cette matrice est inversible et on peut retrouver x et λ .

On peut aussi remarquer que dans le cas général, $l \neq y$, autrement dit, on trouve une surface approximant au mieux les points M_i . Le cas de l'interpolation exacte, lorsqu'elle est possible pour une base de fonction f_k particulières, peut se traiter de la même façon, puisque dans ce cas, le minimum de la fonction de coût est simplement 0.

b. Calcul du vecteur normal

Dans ce paragraphe, on supposera le cas courant, c'est à dire $P = (P_x, P_y) \in \mathbb{R}^2$ et $M = (P, I(P)) \in \mathbb{R}^3$. Le calcul précédent permet d'exprimer $\mathcal{S}$ grâce à la condition :

$$\begin{aligned} M \in \mathcal{S} &\Leftrightarrow l(M) - \sum_{k=1}^n x^k \cdot f_k(M) = 0 \\ M \in \mathcal{S} &\Leftrightarrow F(M_x, M_y, M_z) = 0 \end{aligned}$$

Dans le cas de nos surfaces, l'échantillonnage fait apparaître M_z comme une fonction de M_x et M_y , on supposera dans la suite que la dérivée partielle de F par rapport à M_z ne s'annule pas et que l'on peut appliquer le théorème des fonctions implicites⁴, c'est à dire trouver une fonction f de classe C^1 telle que :

$$M \in \mathcal{S} \Leftrightarrow M_z = f(M_x, M_y)$$

⁴ $\frac{\partial F}{\partial M_z} \neq 0$ sera d'ailleurs un critère pour le choix des fonctions de base, en général elles-mêmes de la forme $M_z = f_k(M_x, M_y)$

Le vecteur normal peut alors être calculé par la relation :

$$\vec{N}(M) = \left(-\frac{\partial f}{\partial M_x}, -\frac{\partial f}{\partial M_y}, 1\right) \quad (2.10)$$

Le théorème des fonctions implicites permet d'exprimer les dérivées partielles de f en fonction de celle de F , et on obtient finalement :

$$\vec{N}(M) = \left(\frac{\frac{\partial F}{\partial M_x}}{\frac{\partial F}{\partial M_z}}, \frac{\frac{\partial F}{\partial M_y}}{\frac{\partial F}{\partial M_z}}, 1\right)$$

c. Calcul sur une image complète

En pratique, les surfaces observées peuvent être assez vastes (300×400 pixels par exemple), et on cherche à calculer le vecteur normal à la surface en chacun de ses points. Il faut donc appliquer la méthode précédente environ 120 000 fois ! Ceci est bien sûr très lourd en temps de calcul, et il est bon de chercher à optimiser chacune des opérations.

Pour simplifier, nous nous placerons dans le cas linéaire sans contrainte (relation 2.6), mais le cas contraint (relation 2.9) s'en déduit immédiatement. Nous placerons systématiquement le point P en lequel on cherche à calculer la normale au centre du système de coordonnées $(0, 0)$, et les points de $V(P)$ auront donc toujours les mêmes coordonnées sur les axes X et Y . Lorsque l'on change de point, on translate simplement le repère de travail (figure 2.8(b)). Le seul paramètre variant d'un point P à un autre pour le calcul de la normale sera donc le vecteur $(I(P_0) = Z_0, \dots, I(P_{N-1}) = Z_{n-1})$ des altitudes des points de $V(P)$.

Plusieurs cas de figure peuvent alors se présenter, mais dans le cas le plus favorable, toutes les surfaces de base s'expriment sous la forme $M_z = f_k(M_x, M_y)$ et la matrice C ne dépend pas du vecteur des altitudes qui n'est autre que le vecteur l du paragraphe précédent, et dans ce cas particulier, on peut simplifier les relations du paragraphe a :

$$\begin{aligned} l(M) &= M_z \\ M \in \mathcal{S} \Leftrightarrow M_z &= \sum_{k=0}^{n-1} x^k \cdot f_k(M_x, M_y) \end{aligned}$$

On introduit la matrice R telle que :

$$R = C^{-1} \cdot A^t \cdot Q^{-1}$$

Et d'après (2.6) on obtient alors le vecteur x de coefficients par

$$x = R \cdot l = R \cdot \begin{pmatrix} Z_0 \\ \vdots \\ Z_{n-1} \end{pmatrix} \quad (2.11)$$

On peut alors calculer les coordonnées du vecteur normal au point P de coordonnées $(0, 0)$:

$$\vec{N}(P) = (-v_1^t \cdot x, -v_2^t \cdot x, 1)$$

2.2. Descripteurs basés sur une représentation de type "image"

Où v_1 et v_2 sont les vecteurs des dérivées partielles des fonctions f_k au point M :

$$v_1 = \begin{pmatrix} \frac{\partial f_0}{\partial M_x}(0,0) \\ \vdots \\ \frac{\partial f_{n-1}}{\partial M_x}(0,0) \end{pmatrix} \text{ et } v_2 = \begin{pmatrix} \frac{\partial f_0}{\partial M_y}(0,0) \\ \vdots \\ \frac{\partial f_{n-1}}{\partial M_y}(0,0) \end{pmatrix}$$

En utilisant la relation 2.11, on obtient $\vec{N}$ en fonction de l , c'est à dire du vecteur des altitudes :

$$\vec{N}(P) = (-v_1^t \cdot R \cdot l, -v_2^t \cdot R \cdot l, 1) \quad (2.12)$$

Les deux vecteurs $v_1^t \cdot R$ et $v_2^t \cdot R$ ne dépendent que des valeurs des fonctions f_k et de leurs dérivées partielles aux points de V , et sont donc en particulier indépendants du point P considéré. Le calcul des coordonnées de $\vec{N}(P) = (N_x(P), N_y(P), 1)$ est donc une fonction linéaire des altitudes $Z_i(X, Y)$ des points de $V(P)$, et le calcul en chaque point de l'image de topographie peut donc se traduire par des opérations de convolution de I par des noyaux Kx et Ky :

$$\begin{aligned} N_x(x, y) &= \sum_{u,v} I(x+u, y+v) Kx(u, v) \\ N_y(x, y) &= \sum_{u,v} I(x+u, y+v) Ky(u, v) \end{aligned} \quad (2.13)$$

Les noyaux Kx et Ky sont calculés à partir des valeurs de $v_1^t \cdot R$ et $v_2^t \cdot R$, de façon à ce que $Kx(u, v)$ (respectivement $Ky(u, v)$) soit égal à $(-v_1^t \cdot R)[i]$ (respectivement $(-v_2^t \cdot R)[i]$) si et seulement si le point P_i a pour coordonnées (u, v) dans le repère centré sur P . Si le voisinage V ne contient pas le point de coordonnées (u, v) alors $Kx(u, v) = 0$ (et symétriquement pour Ky). Dans ces conditions, on peut simplifier (2.13) en :

$$N_x(x, y) = Z * \check{K}x \quad (2.14)$$

$$N_y(x, y) = Z * \check{K}y \quad (2.15)$$

Où $\check{K}x$ et $\check{K}y$ désigne les noyaux symétriques de Kx et Ky par rapport à l'origine et $*$ l'opération de convolution.

Cette formulation sous forme de produit de convolution permet essentiellement de bénéficier des algorithmes efficaces de calcul par FFT (Fast Fourier Transform) :

$$\begin{aligned} N_x(x, y) &= FFT^{-1}(FFT(Z) \times \overline{FFT(Kx)}) \\ N_y(x, y) &= FFT^{-1}(FFT(Z) \times \overline{FFT(Ky)}) \end{aligned} \quad (2.16)$$

Où $\bar{\cdot}$ désigne l'opération de conjugaison (provenant du $\check{\cdot}$ de $\check{K}x$ et $\check{K}y$ dans les équations (2.14) et (2.15)).

Dans le cas où les calculs de N_x et N_y ne sont pas linéaires par rapport au vecteur des altitudes, il n'existe pas de méthode systématique pour accélérer sensiblement le calcul comme peut le faire l'utilisation de la FFT : il faut donc s'adapter au cas par cas.

Les développements de cette section, et en particulier les relations 2.16 montrent comment calculer les vecteurs normaux à une image discrète par simple convolution. Nous allons maintenant appliquer ces résultats à des cas particuliers de bases de fonctions.

d. Interpolation par des triangles

La première surface analytique permettant d'interpoler exactement l'ensemble des points connus est obtenue par simple triangulation. La triangulation la plus simple consiste à diviser chaque carré élémentaire de la grille en deux triangles : il y a deux possibilités (diagonale Nord-Ouest/Sud-Est ou Nord-Est/Sud-Ouest) et l'une des deux est choisie de manière arbitraire (figure 2.9(a)).

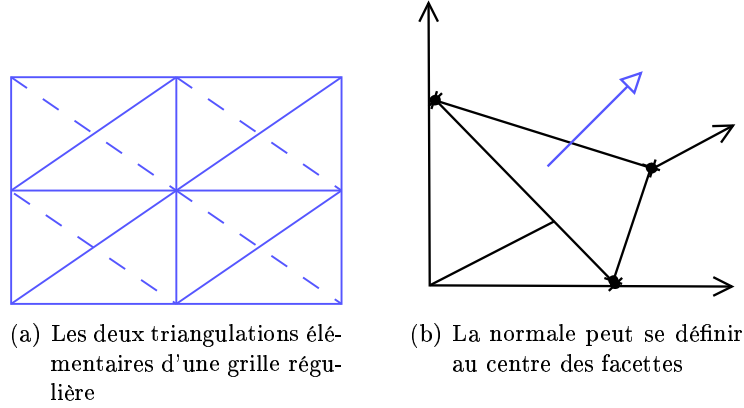


FIG. 2.9.: Triangulation

La surface obtenue a l'inconvénient de ne pas être C^1 car la différentielle n'est pas définie au niveau des arrêtes, mais on peut tout de même obtenir le vecteur normal au centre de chacun des triangles. Chaque facette possède alors une équation de la forme $Z = aX + bY + c$, et on peut dérouler le calcul des matrices décrit dans le paragraphe précédent, même si dans ce cas particulièrement simple, cela peut sembler superflu. Les fonctions de bases peuvent s'exprimer de la manière suivante :

$$\begin{aligned} f_1(x, y) &= x \\ f_2(x, y) &= y \\ f_3(x, y) &= 1 \end{aligned}$$

Si l'on considère les triangles formés à partir des diagonales Sud-Ouest/Nord-Est, alors les coordonnées des points de V considérés sont de la forme :

$$\begin{aligned} M_0 &= (0, 0, Z_0 = I(P_0)) \\ M_1 &= (1, 0, Z_1 = I(P_1)) \\ M_2 &= (0, 1, Z_2 = I(P_2)) \end{aligned}$$

On obtient donc la matrice A de la forme :

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

2.2. Descripteurs basés sur une représentation de type "image"

Puis la matrice C peut être calculée (Dans ce cas précis, nous voulons attribuer un poids identique à chacun des points M_i , et la matrice Q est prise égale à l'identité) :

$$C = A^t.A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

Enfin, on exprime la matrice R :

$$R = C^{-1}.A^t.Q^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Les vecteurs des dérivées partielles au point $M_0 = (0,0)$ s'écrivent :

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ et } v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

On obtient alors :

$$\begin{aligned} -v_1^t.R &= (1, -1, 0) \\ -v_2^t.R &= (1, 0, -1) \end{aligned}$$

Et les coordonnées du vecteur normal sont alors (d'après (2.12)) :

$$\begin{aligned} \vec{N}(P)_X &= v_1^t.R. \begin{pmatrix} Z_0 \\ Z_1 \\ Z_2 \end{pmatrix} = Z_0 - Z_1 \\ \vec{N}(P)_Y &= v_2^t.R. \begin{pmatrix} Z_0 \\ Z_1 \\ Z_2 \end{pmatrix} = Z_0 - Z_2 \\ \vec{N}(P)_Z &= 1 \end{aligned}$$

Ce qui dans ce cas élémentaire s'obtenait aisément par le produit vectoriel des vecteurs $\overrightarrow{M_0M_1}$ et $\overrightarrow{M_0M_2}$, mais il est bon d'appliquer la méthode générale de manière systématique quelle que soit la base de fonctions utilisées.

Les noyaux de convolution utilisés pour effectuer le calcul sur l'ensemble de l'image peuvent alors s'écrire sous la forme suivante⁵ :

$$\begin{aligned} Kx &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \\ Ky &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \end{aligned}$$

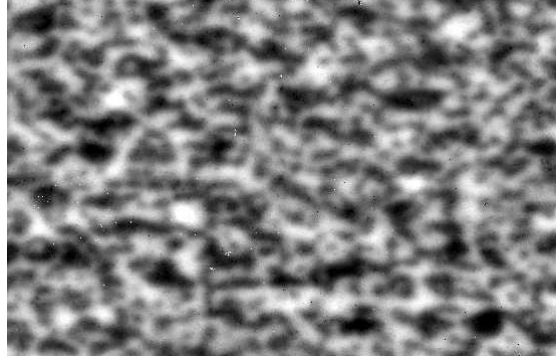


FIG. 2.10.: Exemple de surface de type peau d'orange (Pdo 1)

On se place dans le cas pratique d'une surface de type peau d'orange. (figure 2.10).

On applique alors le calcul des normales au centre des facettes des triangles (diagonale orientée Sud-Ouest/Nord-Est). On obtient ainsi trois composantes, selon l'axe des X (direction de laminage), selon l'axe des Y (direction traverse) et selon l'axe des Z (figure 2.11).

L'utilisation que l'on peut faire de cette carte des normales sera envisagée dans les paragraphes suivants, mais certaines remarques peuvent d'ores et déjà être faites en observant l'image résultante : on retrouve des structures principalement orientées dans le sens de laminage qui induisent des vecteurs normaux principalement orientés dans le sens travers (direction Y). On retrouve également des plateaux formés par la troncature de certains pics lors de la conversion des images depuis le format du microscope : ce sont des zones plates, et la normale est donc dirigée uniquement selon l'axe des Z . On peut aussi noter l'aspect fortement bruité de cette image qui provient des discontinuités de normale d'un triangle à l'autre et de l'absence totale de "lissage" de la surface interpolante qui n'est pas différentiable.

e. Interpolation par des paraboloides

Dans cette section, on cherche à introduire une base de fonctions permettant d'obtenir une carte de normales plus régulière que dans le cas des triangles.

Nous introduisons ici une base de fonctions polynomiales de degré 2 en X , Y et Z . En chaque point, la surface sera donc interpolée par une quadrique d'équation :

$$x^1 X^2 + x^2 Y^2 + x^3 Z^2 + x^4 YZ + x^5 XZ + x^6 XY + x^7 X + x^8 Y + x^9 Z = 0$$

On ne prend pas en compte de terme constant, pour imposer à la quadrique de passer par l'origine, c'est à dire par le point en lequel on calcule la normale. On se placera sur un voisinage de 9 points (c'est à dire un carré de 3 pixels sur 3), et on imposera à la quadrique une condition sur ses coefficients de façon à ne pas annuler tous les termes en Z

⁵Le coefficient placé au centre correspond au point central en lequel on cherche à calculer la normale.

2.2. Descripteurs basés sur une représentation de type "image"

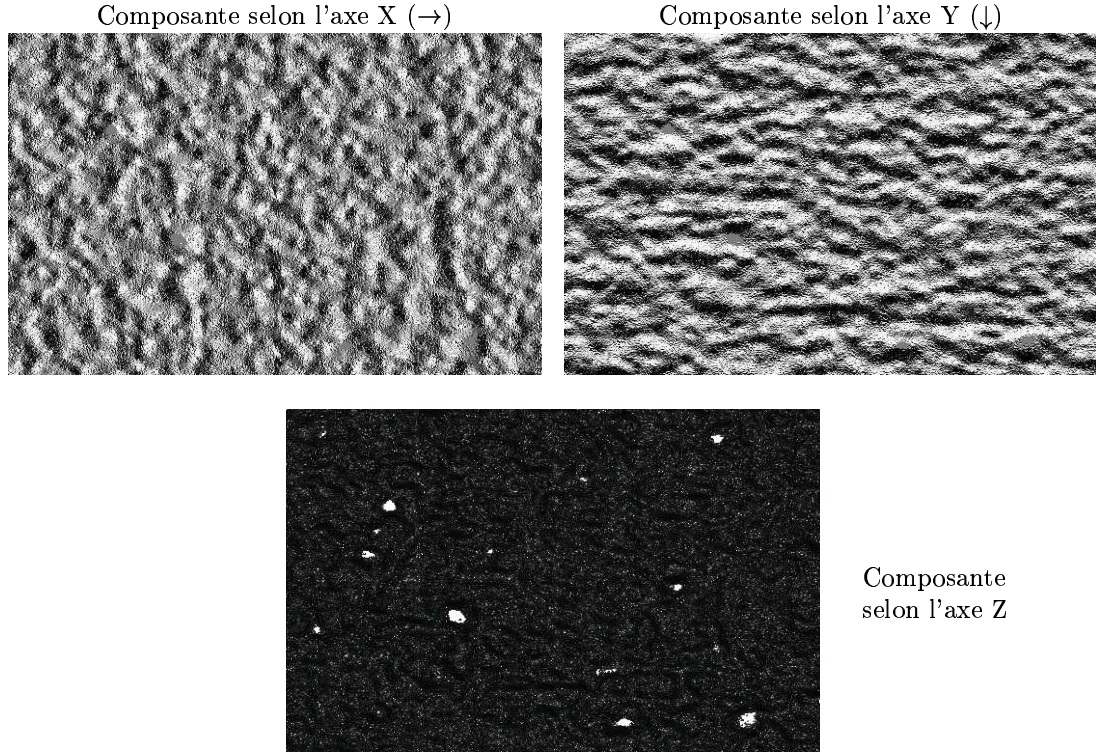
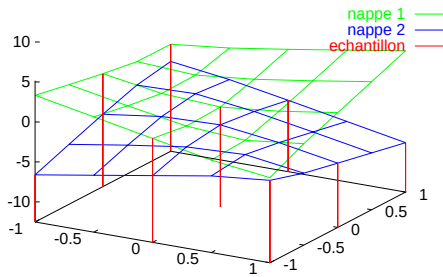


FIG. 2.11.: Composantes du vecteur normal à la surface "Pdo01" obtenues par triangulation

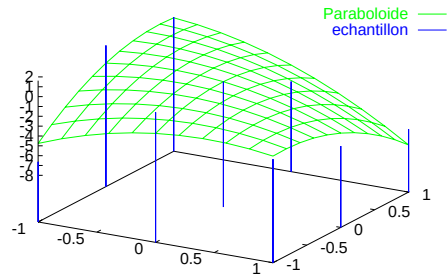
(la surface est à priori une surface pouvant s'exprimer, au moins localement, sous la forme $z = f(x, y)$). Le plus simple pour obtenir une telle condition serait d'imposer au produit $|x^3| + |x^4| + |x^5| + |x^9| > 0$, mais pour pouvoir appliquer les résultats des relations (2.7) et (2.8), il est nécessaire d'exprimer la contrainte sous la forme d'une relation linéaire des coefficients, et la formulation retenue suppose donc que $x^3 + x^4 + x^5 + x^9 = 1$, ce qui limite quelque peu la généralité des quadriques considérées. Dans ce cas particulier, plusieurs problèmes se posent. Tout d'abord, la matrice A dépend du point considéré, puisque les fonctions de Z ne sont pas toutes linéaires, ce qui ne permet pas d'appliquer la méthode efficace de calcul par FFT. D'autre part, les résultats peuvent conduire à des approximations par des quadriques contenant plusieurs nappes (figure (2.12(a))). Devant ces deux problèmes importants, il a été décidé de réduire les classes des quadriques aux seuls paraboloides s'exprimant sous la forme :

$$Z = x^1 X^2 + x^2 Y^2 + x^3 XY + x^4 X + x^5 Y$$

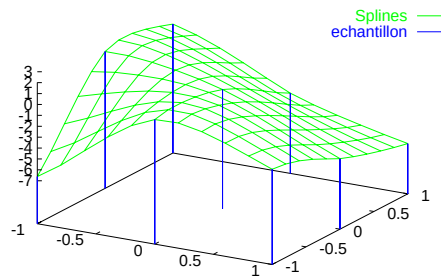
Ceci présente l'avantage de rétablir la linéarité en Z et de prévenir l'apparition de plusieurs nappes sur la surface approximante. Par contre, l'approximation en terme de moindre carrés est moins bonne que précédemment (figure 2.12(b)).



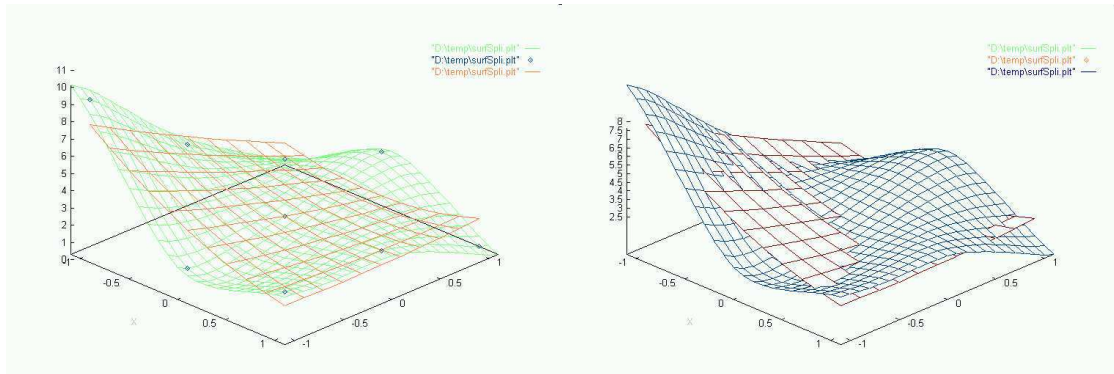
(a) Meilleure approximation de neuf points par une quadrique “générique” : plusieurs nappes se forment.



(b) Meilleure approximation des neuf points de la figure 2.12(a) par un paraboloïde.



(c) Interpolation des neuf points de la figure 2.12(a) par des fonctions de splines



(d) Comparaison des surfaces obtenues en utilisant des splines ou des paraboloïdes comme base de fonctions approximantes

FIG. 2.12.: Différentes fonctions d'interpolation

2.2. Descripteurs basés sur une représentation de type "image"

Dans le cas des paraboloïdes, on peut appliquer l'algorithme de calcul par FFT sur la totalité de l'image "Pdo01" (figure 2.10) pour obtenir les résultats visibles sur la figure 2.13. Les valeurs obtenues sont cohérentes avec la carte calculée en utilisant des triangles, mais l'effet de "lissage" dû aux paraboloïdes est important, et la carte des normales est plus régulière et continue que celle de la figure 2.11.

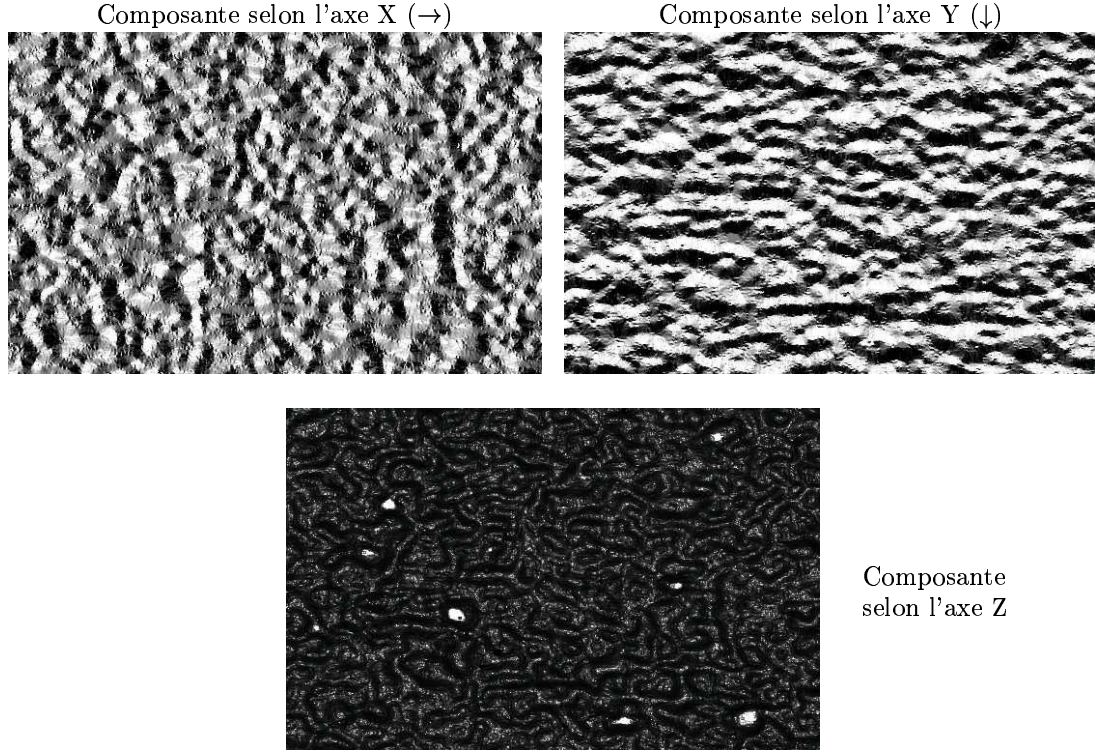


FIG. 2.13.: Composantes du vecteur normal à la surface "Pdo01" obtenues au moyen de paraboloïdes

f. Interpolation par des fonctions radiales

Dans les paragraphes précédents, deux types de bases de fonctions permettant de calculer une carte de normales ont été présentées : la triangulation permet une interpolation exacte des points de l'image pour chaque voisinage considéré, mais induit une carte des normales bruitée tandis que les paraboloïdes permettent de "lisser" l'image et d'obtenir une carte des normales plus régulières, mais les points de l'image ne sont plus interpolés mais seulement approximatés.

Dans ce paragraphe on cherche à utiliser une base de fonctions permettant de trouver un compromis entre ces extrêmes : interpoler exactement la surface tout en régularisant la carte des normales par rapport à la triangulation [18].

On considère un voisinage V de $\mathbb{R}^2$ centré autour d'un point P en lequel on cherche à calculer la normale⁶. V est formé de N points $P_0, \dots, P_{N-1}$ (en général, $P \in V$).

On considère alors la surface définie par :

$$\forall P \in \mathbb{R}^2, Z = p_m(P) + \sum_{k=0}^{N-1} x^k \psi(\|P - P_k\|) \quad (2.17)$$

Où p_m désigne un polynôme de degré au plus m en les coordonnées de P , $\|\cdot\|$ désigne la norme euclidienne dans $\mathbb{R}^2$ et ψ désigne une fonction prédéfinie de $\mathbb{R}^+ \rightarrow \mathbb{R}$. Dans ce qui suit, nous nous sommes intéressés au cas où ψ est une fonction de type "spline" de la forme :

$$\psi(r) = r^2 \ln(r) \quad (2.18)$$

Il s'agit d'une classe de fonctions régulières permettant d'interpoler exactement les points de V tout en assurant à la surface approximante d'être C^1 . Le polynôme p_m est ici de degré au plus 1 ($m=1$), c'est à dire l'équation d'un plan. Une interpolation de ce type consiste à choisir un plan approchant correctement les points choisis, puis à le déformer autour de chacun des points comme si le plan était élastique et pouvait s'étirer pour atteindre les points à interpoler.

Pour assurer l'unicité de la solution, on impose une contrainte sur les coefficients x^k :

$$\forall q \in \pi_1, \sum_{k=0}^{N-1} x^k q(P_k) = 0 \quad (2.19)$$

Où π_1 désigne l'ensemble des polynômes de degré au plus 1. Cette contrainte permet d'imposer la propriété suivante : si l'ensemble des points à interpoler se trouve sur un polynôme de degré 1, c'est à dire un plan, alors la fonction interpolante retenue sera le plan en question[48]. En utilisant les notations de la section précédente, la relation 2.19 est équivalente à introduire 3 conditions supplémentaires définies par les trois fonctions :

$$\begin{aligned} \Phi_1 &= \sum_{k=0}^{N-1} n - 1x_k \\ \Phi_2 &= \sum_{k=0}^{N-1} n - 1x_k \cdot P_k^x \\ \Phi_3 &= \sum_{k=0}^{N-1} n - 1x_k \cdot P_k^y \end{aligned}$$

Notons que dans ce cas précis, $n = N$.

En réintroduisant les notations du paragraphe a, on obtient :

$$\begin{aligned} \Psi &= (\psi_i^k) = (\psi(\|P_i - P_k\|)) \\ \Pi &= \begin{pmatrix} 1 & P_0^X & P_0^Y \\ \vdots & \vdots & \vdots \\ 1 & P_{N-1}^X & P_{N-1}^Y \end{pmatrix} \end{aligned}$$

⁶Le point P est toujours considéré comme l'origine de l'espace

2.2. Descripteurs basés sur une représentation de type "image"

$$\begin{aligned} A &= (\Phi, \Pi) \text{ (matrice par bloc)} \\ C = A^t A &= \begin{pmatrix} \Phi^2 & \Phi\Pi \\ \Pi^t\Phi & \Pi^t\Pi \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \tilde{x} &= (x^0, \dots, x^{N-1})^t \\ c &= (c_0, c_1, c_2)^t \\ x &= (x^0, \dots, x^{N-1}, c_0, c_1, c_2)^t \\ B &= \begin{pmatrix} (\Pi) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} C & B \\ B^t & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ \lambda \end{pmatrix} &= \begin{pmatrix} L \\ 0 \end{pmatrix} \end{aligned}$$

Cependant, il a été signalé que dans ce cas particulier, l'interpolation obtenue était exacte, et on peut montrer que la matrice :

$$\tilde{C} = \begin{pmatrix} \Phi & \Pi \\ \Pi^t & 0 \end{pmatrix}$$

est inversible[48]. On peut donc directement utiliser la relation :

$$\begin{aligned} \tilde{C} \cdot \begin{pmatrix} \tilde{x} \\ c \end{pmatrix} &= \begin{pmatrix} l \\ 0 \end{pmatrix} \\ \begin{pmatrix} \tilde{x} \\ c \end{pmatrix} &= \tilde{C}^{-1} \cdot \begin{pmatrix} l \\ 0 \end{pmatrix} \end{aligned}$$

qui résume à la fois la propriété d'interpolation et la contrainte (2.19). Les résultats obtenus par les deux méthodes (celle de l'équation (2.9) et celle de l'inversion directe) fournissent bien entendu des résultats compatibles, puisque le minimum de la fonction de coût est atteint lorsque l'interpolation exacte est réalisée. On peut ensuite appliquer la méthode de calcul des normales par convolution, comme décrit dans le paragraphe c.

Nous comparons sur un exemple simple de voisinage les différences entre les surfaces obtenues en interpolant un carré de 3×3 points par une quadrique générique (figure 2.12(a)), par un parabolöide (figure 2.12(b)) ou par une superposition de splines (figure 2.12(c)). La figure 2.12(d) illustre l'effet régularisant plus marqué dans le cas du parabolöide que dans celui des splines ainsi que la légère différence de normale entre ces deux surfaces au point central.

On peut aussi effectuer le calcul sur l'ensemble de l'image 2.10. On obtient sur la figure 2.14 un résultat intermédiaire pour ce qui est de la régularité de la carte de normale entre les figures 2.11 et 2.13.

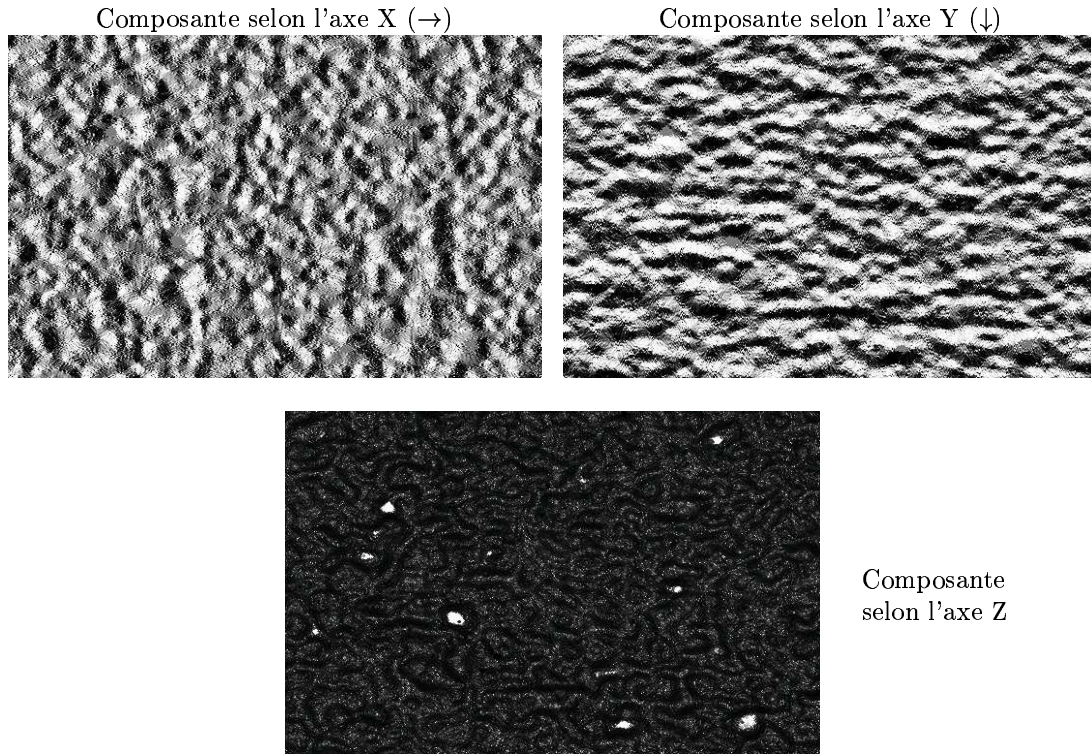


FIG. 2.14.: Composantes du vecteur normal à la surface “Pdo01” obtenues par des fonctions de splines

g. Construction d'un descripteur

Pour construire un descripteur de la surface à partir de la carte des vecteurs normaux, nous nous intéresserons particulièrement à la distribution des pentes et à l'orientation des vecteurs dans le plan XY .

La pente de la surface s'obtient en chaque point par la relation :

$$\delta(P) = \arctan\left(\frac{\sqrt{N_x(P)^2 + N_y(P)^2}}{N_z(P)}\right)$$

A partir de ces valeurs, nous construisons un histogramme qui servira de descripteur et pourra être comparé d'une surface à l'autre.

Pour ce qui est de l'orientation des vecteurs dans le plan XY , il faut se fixer une origine qui sera choisie dans la direction de laminage.

$$\begin{aligned} \cos(\theta(P)) &= \frac{N_x(P)}{\sqrt{N_x(P)^2 + N_y(P)^2}} \\ \sin(\theta(P)) &= \frac{N_y(P)}{\sqrt{N_x(P)^2 + N_y(P)^2}} \end{aligned}$$

2.2. Descripteurs basés sur une représentation de type "image"

Là encore, un histogramme de ces orientations permet de rendre compte de l'asymétrie (figure 2.15) entre sens laminage et sens travers qui peut être plus ou moins marquée d'une surface à l'autre. Ce descripteur permet en quelque sorte de mesurer l'anisotropie de la surface comme nous le verrons par la suite.

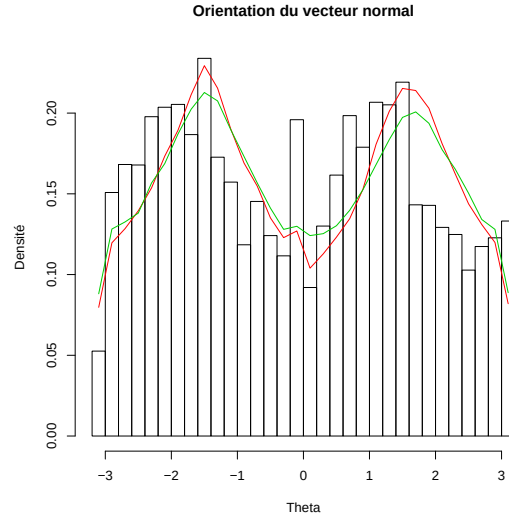


FIG. 2.15.: Asymétrie d'orientation entre sens laminage ($\theta = 0$) et sens travers ($\theta = \pi/2$). La courbe rouge est obtenue au moyen de splines, la verte au moyen de paraboloides, et l'histogramme en noir est obtenu par une triangulation de la surface.

2.2.6. Ondelettes

L'utilisation d'outils de type ondelettes provient essentiellement de la richesse de l'analyse multi-échelles en terme de traitement du signal. En particulier, les rugosités des surfaces observées peuvent entraîner, selon leur échelle caractéristique, des effets très différents en terme d'aspect ou de réponse à une opération de type peinture, et c'est pourquoi la transformation en ondelette apparaît bien adaptée pour en étudier certains aspects.

La notion de courbe de "Ma" pour caractériser des surfaces rugueuses a été initialement développé au LTDS, à l'Ecole Centrale de Lyon, lors de la thèse de Lee[40]. Les bons résultats obtenus pour caractériser des surfaces peintes destinées à l'industrie automobile à partir de ces courbes [68] nous ont amenés, en collaboration avec le LTDS, à les introduire dans notre jeu de descripteurs. Même si nous ne sommes qu'utilisateur de la notion de "Ma", nous avons tenu à décrire ici le moyen de les obtenir. Pour ce faire, nous présentons succinctement les formules de transformée en ondelette et de reconstruction avant de présenter la définition d'une courbe de Ma.

Les développements mathématiques présentés ici sont essentiellement issus du livre de Mallat "a Wavelet Tour of signal processing" [42], et il est conseillé de s'y reporter, ainsi

qu'à la thèse de *Lee* pour de plus amples commentaires sur la transformée en ondelette ou sur la courbe de Ma.

Nous ne traiterons ici que de l'application à un signal monodimensionnel discret (profil) de la théorie de la transformation en ondelettes dite "continue" mais il existe des développements tant dans le domaine de la généralisation du problème à un signal bidimensionnel (une surface) que dans l'amélioration des résultats sur des signaux discrets.

a. Aperçu de la transformée en ondelette continue

L'idée de la transformée en ondelette est d'obtenir une représentation temps/fréquence d'un signal, c'est à dire d'observer au cours du temps l'évolution du contenu fréquentiel d'un signal.

Une ondelette est une fonction $\psi \in L^2(\mathbb{R})$:

$$\text{De moyenne nulle : } \int_{-\infty}^{+\infty} \psi(t) dt = 0$$

$$\text{Normée : } \|\psi\| = \int_{-\infty}^{+\infty} |\psi(t)|^2 dt = 1$$

Elle est aussi centrée au voisinage de $t = 0$, et une famille d'atomes temps/fréquence est alors obtenue par changement d'échelle d'un facteur s et par translation d'un vecteur u .

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$

L'atome reste normé, et on définit alors la transformée en ondelette d'une fonction $f \in L^2(\mathbb{R})$ à un temps u donné et pour un facteur d'échelle s donné par :

$$Wf(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \bar{\psi}\left(\frac{t-u}{s}\right) dt \quad (2.20)$$

où $\bar{z}$ représente le complexe conjugué de z .

Cette transformation peut se réécrire sous la forme d'un produit de convolution, ce qui peut être utile en terme d'implémentation :

$$Wf(u, s) = f * \psi_s^*(u) \quad (2.21)$$

avec $\psi_s^*(t) = \frac{1}{\sqrt{s}} \bar{\psi}\left(\frac{-t}{s}\right)$

Une fois la transformation Wf définie, on se pose en générale la question de la transformée inverse. Deux cas de figure sont possibles suivant que l'ondelette est réelle ou analytique (c'est à dire vérifiant la propriété (2.23) qui implique que Ψ est complexe). Le cas réel est souvent utilisé pour détecter des transitions brutales dans un signal tandis que le cas analytique permet d'étudier l'évolution des fréquences instantanées notamment en introduisant la notion de phase.

Soit $\psi \in L^2(\mathbb{R})$ une ondelette réelle telle que :

$$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty$$

2.2. Descripteurs basés sur une représentation de type "image"

où $\hat{\psi}(\omega)$ désigne la transformée de Fourier de ψ .

Alors pour toute fonctions $f \in L^2(\mathbb{R})$, on obtient la relation :

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u, s) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) du \frac{ds}{s^2} \quad (2.22)$$

Soit ψ une ondelette analytique, c'est à dire une ondelette telle que :

$$\forall \omega < 0, \hat{\psi}(\omega) = 0 \quad (2.23)$$

Si f est un signal réel, on obtient la formule de reconstruction :

$$f(t) = \frac{2}{C_\psi} \mathcal{Re} \left(\int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u, s) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) du \frac{ds}{s^2} \right)$$

Où $\mathcal{Re}(x)$ désigne la partie réelle de x . Cette formule peut également s'écrire sous la forme d'un produit de convolution :

$$f(t) = \frac{2}{C_\psi} \mathcal{Re} \left(\int_0^{+\infty} Wf(\cdot, s) * \psi_s(-t) \frac{ds}{s^2} \right) \quad (2.24)$$

avec $\psi_s(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t}{s}\right)$

b. Application à un signal discret

En pratique, lorsque l'on cherche à analyser un signal informatiquement, il est échantillonné en un certain nombre N de points temporels. Pour simplifier le problème, considérons un signal $f(t)$ défini sur $[0, N]$ échantillonné par pas de 1. On notera $f[n], n = 0, 1, \dots, N-1$ la valeur prise par f aux points d'échantillonnage.

Le calcul de la transformée en ondelette discrétisée se fera pour les facteurs d'échelles $s = a^j$ avec $a = 2^{1/\nu}$. Soit $\psi(t)$ une ondelette dont le support est inclus dans $[-K/2, K/2]$. Pour $2 \leq a^j \leq NK^{-1}$, on définit une ondelette discrète dilatée par a^j :

$$\psi_j[n] = \frac{1}{\sqrt{a^j}} \psi\left(\frac{n}{a^j}\right)$$

Pour éviter les effets de bords, le signal $f[n]$ et les ondelettes $\psi_j[n]$ seront traités comme périodiques. En utilisant la discrétisation de l'intégrale définissant la transformée en ondelette sous forme de convolution (équations 2.20 et 2.21), on obtient l'expression :

$$Wf[n, a^j] = \sum_{m=0}^{N-1} f[m] \bar{\psi}_j[m-n] = f \otimes \psi_j^*[n]$$

Où $\otimes$ désigne l'opération de convolution discrète.

Pour les formules de reconstruction, il suffit de discrétiser les intégrales 2.22 et 2.24, et dans le cas où l'ondelette est analytique et f est un signal réel dont l'énergie dans les basses fréquences est faible, on obtient la formule utilisée en pratique :

$$f[n] \simeq \frac{2 \log_e(a)}{C_\psi} \mathcal{Re} \left(\sum_{j=I}^J Wf[\cdot, a^j] \otimes \psi_j[n] \right) \quad (2.25)$$

c. Courbe de "Ma", construction de descripteurs

La notion de courbe de "Ma" a été introduite par Lee [40] dans sa thèse. Il s'agit de construire une courbe représentative de la structure fréquentielle du signal à partir de sa représentation en ondelette.

La formule de reconstruction discrète 2.25 se présente comme une somme de portion de signaux reconstruits chacun à partir d'un seul facteur d'échelle. L'idée du "Ma" consiste à considérer ces portions de signaux reconstruites pour chaque facteur d'échelle et à en calculer le Ra (moyenne des valeurs absolues des écart à la moyenne d'un profil donné, voir 2.1). On obtient ainsi une valeur correspondant à chaque facteur d'échelle que l'on peut représenter sous la forme d'une courbe appelée courbe de *Ma* (figure 2.17).

$$Ma(j) = Ra\left(\frac{2\log_e(a)}{C_\psi} \mathcal{R}e(Wf[., a^j] \otimes \psi_j)\right)$$

En pratique, pour l'étude des surfaces rugueuses de type "Peau d'orange", on dispose d'images, tandis que les courbes de "Ma" s'extraient à partir de profil. On a donc choisi d'extraire plusieurs profils de la surface, de calculer le "Ma" sur chacun d'eux, et d'attribuer à la surface la courbe correspondant à la moyenne point par point de l'ensemble des courbes obtenues profil par profil (figure 2.16).

Par ailleurs les transformées en ondelettes des figures 2.17 et 2.16 ont été obtenues au moyen d'ondelettes de Morlet, mais la théorie s'applique bien sûr à n'importe quelle ondelette analytique.

Mise à part la courbe de "Ma", d'autres descripteurs peuvent être construits à partir de la transformée en ondelettes d'un profil. On peut par exemple s'intéresser aux distributions des valeurs $Wf(., s)$ à s fixé, et comparer des statistiques (moyenne, histogrammes, etc...) de ces distributions d'une surface à l'autre.

2.2. Descripteurs basés sur une représentation de type "image"

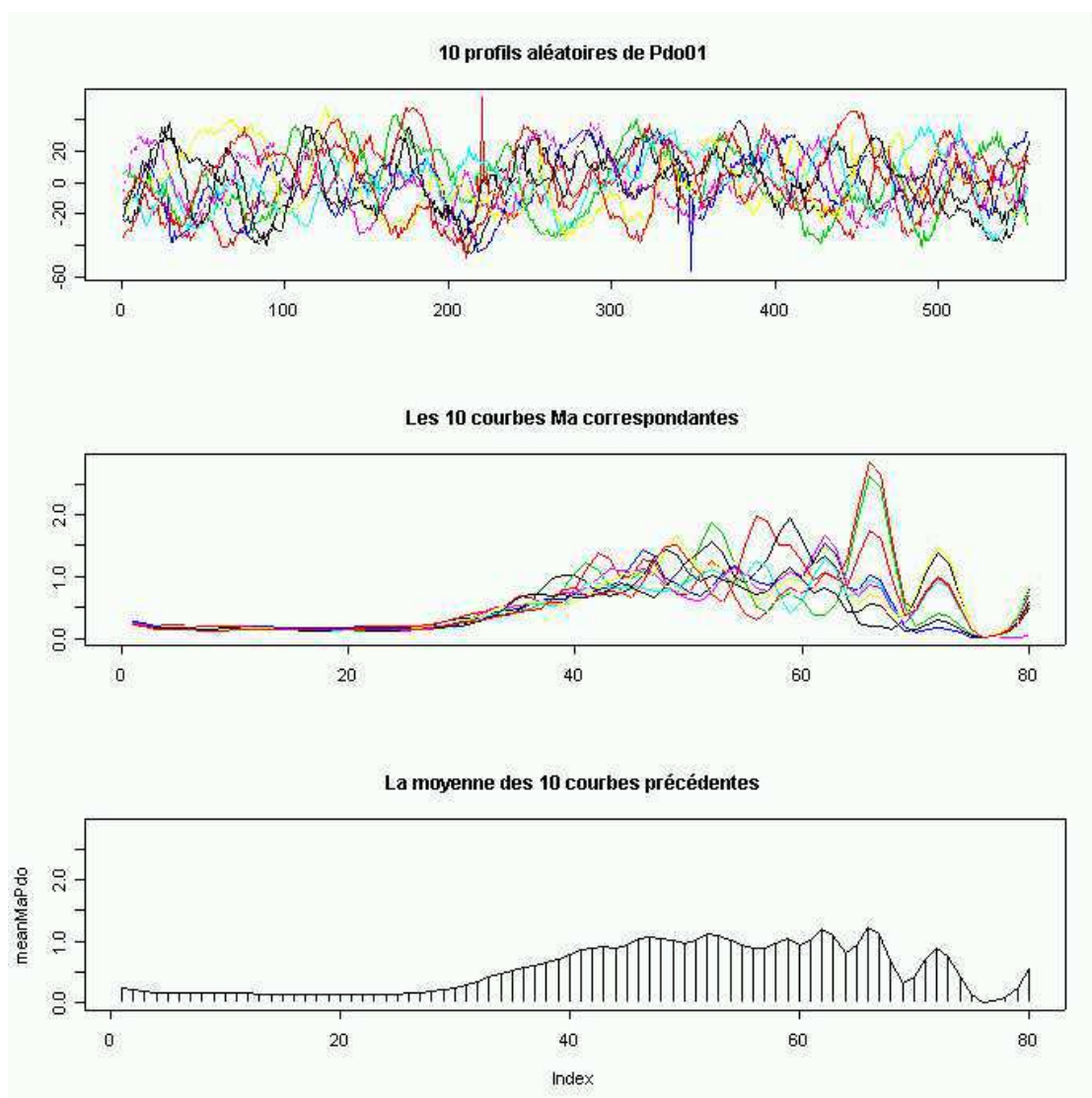


FIG. 2.16.: Extraction d'une courbe de Ma à partir de 10 profils d'une image Peau d'orange

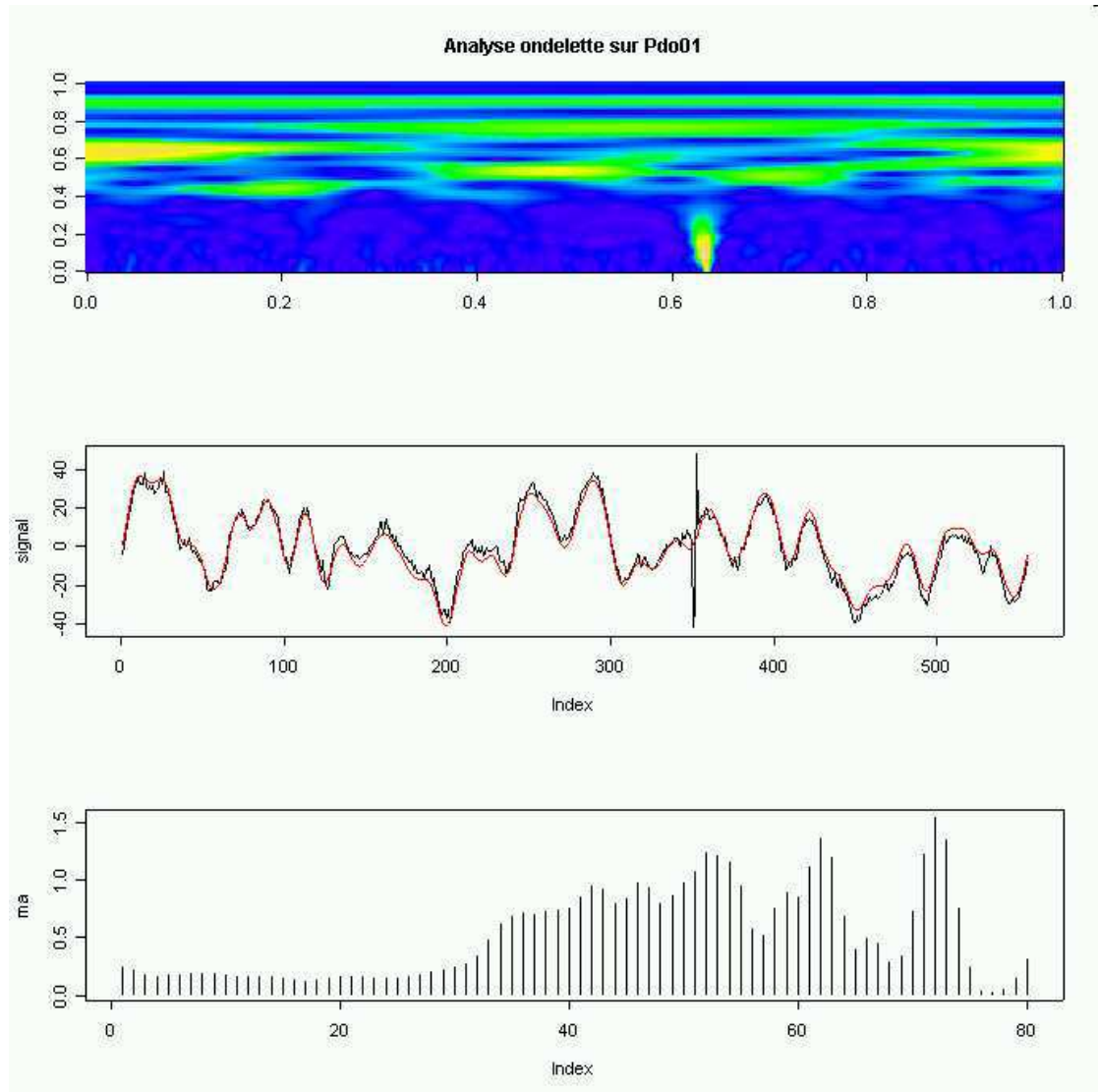


FIG. 2.17.: Exemple sur un profil d'une image de type Peau d'orange. Le signal analysé est en noir sur la figure centrale. Sa transformée en ondelette est présentée sur la figure du haut et la courbe de "Ma" sur la courbe du bas. En rouge, superposée au signal (figure centrale), on observe une courbe correspondant à la reconstruction du signal à partir des facteurs d'échelle supérieurs à 20. Ces facteurs d'échelles correspondent à la zone maximum de la courbe de "Ma".

Chapitre 3.

Stabilité de la description

3.1. Origine du problème

Ce chapitre concernant la stabilité de la description des surfaces constitue l'un des points essentiels de notre méthodologie car il introduit la notion de variance "d'extension" [43] dont l'importance nous conduira à profondément modifier les méthodes classiques de caractérisation des surfaces. Le problème que nous nous proposons de soulever est bien connu en géostatistique, et les développements de Matheron en particulier nous permettront de l'appréhender dans sa version théorique, tandis que des techniques de sur-échantillonnage statistique de type "bootstrap" nous permettront de tenir compte de la variance d'extension sur le plan pratique pour la classification de surfaces industrielles.

Le problème concret qui se pose à nous est le suivant : nous cherchons à caractériser une surface par des descripteurs variés, et nous nous basons pour cela sur une représentation de type "image". Or cette image a le mauvais goût d'être de dimensions limitées, ceci pour des raisons pratiques évidentes, et si l'aire de l'image de travail correspond au champ d'observation d'un microscope ou au nombre de pixels d'un capteur CCD, la surface sous-jacente dont elle est extraite présente des dimensions beaucoup plus importantes, que l'on peut presque considérer comme infinies, compte-tenu de l'échelle d'observation. La question qui se pose alors concerne la représentativité des descripteurs extraits à partir d'un champ d'observation de "petite" taille, et de façon corollaire, la quantification de l'erreur commise en estimant ainsi certains attributs caractéristiques de la surface infinie sous-jacente. L'un des moyens d'appréhender cette erreur consiste à s'intéresser à la variance dite variance d'extension d'un descripteur mesuré sur un champ réduit lorsque la fenêtre d'observation se déplace sur une surface plus large.

Dans un premier temps, nous présentons, à partir de modèles de structures aléatoires [33, 34, 35, 31], quelques résultats géostatistiques permettant d'estimer la variance d'extension en illustrant par un exemple l'intérêt de sa prise en compte dans la comparaison des surfaces. Dans un second temps, nous soulèverons les problèmes pratiques liés à la théorie géostatistique et nous introduirons alors les techniques de bootstrap qui nous permettront de tenir compte de la variance d'extension dans nos méthodes de classification.

3.2. Variance d'extension

3.2.1. Fonction aléatoire et covariogramme

Dans cette section, on considère l'image I définie sur un domaine v comme la réalisation sur v d'une fonction aléatoire sous-jacente $\mathcal{I}$ stationnaire d'ordre 2 définie sur un domaine $\mathcal{D} = \mathbb{R}^2$ dans le cas d'images 2D :

- $\forall x \in \mathbb{R}^2, E(\mathcal{I}(x)) = m < \infty$
- $\forall x \in \mathbb{R}^2, E((\mathcal{I}(x) - m)^2) = \sigma^2 < \infty$
- $\forall x, y \in \mathbb{R}^2, C(x, y) = C(x - y)$

Rappelons que dans ces expressions, $\mathcal{I}(x)$ ne désigne pas un nombre mais une variable aléatoire. $C(x, y)$ désigne ici la covariance entre les variables aléatoire $\mathcal{I}(x)$ et $\mathcal{I}(y)$. La stationnarité de $\mathcal{I}$ fait que cette covariance ne dépend que de la différence $x - y$, et la définition de $C(h)$ est alors équivalente à celle la section 2.2.1 :

$$\forall x \in \mathbb{R}^2, C(h) = E(\mathcal{I}(x + h) \cdot \mathcal{I}(x)) - E(\mathcal{I}(x + h)) \cdot E(\mathcal{I}(x))$$

Nous nous intéressons dans ce qui suit à un descripteur simple formé à partir de la moyenne de la réalisation I :

$$Z(I) = \frac{1}{v} \int_v I(x) dx$$

Dans notre contexte d'étude, le nombre $Z(I)$ n'est autre que la réalisation d'une variable aléatoire Z_v définie par l'intégrale stochastique suivante (dont les propriétés de $\mathcal{I}$ assurent l'existence) :

$$Z_v = \frac{1}{v} \int_v \mathcal{I}(x) dx$$

Ce formalisme permet essentiellement d'introduire, par le biais de $\mathcal{I}$, la notion de variable aléatoire et de variance dans un contexte complètement déterministe. En effet, la surface que l'on considère est bien réelle, et si il existe une fonction aléatoire sous-jacente, nous ne pourrions jamais l'appréhender que par une unique réalisation (la surface réelle observée) : la justification philosophique et mathématique de cette approche est très bien soulignée par Matheron dans ses différents ouvrages [43].

On peut aussi calculer la variance $D^2(Z_v)$ de la variable Z_v . Si l'on considère une surface infinie comme une réalisation de $\mathcal{I}$ sur un domaine $\mathcal{D}$, les propriétés de stationnarité et d'ergodicité permettent d'affirmer que D^2 est aussi la variance de l'intégrale $Z(I)$ considéré comme une variable aléatoire lorsque le domaine v est choisi aléatoirement sur la surface.

$$\begin{aligned} D^2(Z_v) &= C(Z_v, Z_v) \\ &= C\left(\frac{1}{v} \int_v \mathcal{I}(x) dx, \frac{1}{v} \int_v \mathcal{I}(y) dy\right) \end{aligned} \tag{3.1}$$

$$\begin{aligned} &= \frac{1}{v^2} \int_v \int_v C(\mathcal{I}(x), \mathcal{I}(y)) dx dy \\ &= \frac{1}{v^2} \int_v \int_v C(x - y) dx dy \end{aligned} \tag{3.2}$$

Le passage de 3.1 à 3.2 provient simplement de la bilinéarité de la fonction de covariance.

Dans le cas particulier où le domaine d'intégration v est très grand devant la portée, c'est à dire la valeur de h à partir de laquelle $C(h)$ devient négligeable, et où l'intégrale $\kappa(\mathcal{I}) = \int_{\mathbb{R}^2} C(h)dh$ vérifie la relation $0 < \kappa(\mathcal{I}) < \infty$, on peut obtenir l'expression approchée suivante :

$$D^2(Z_v) \simeq \frac{1}{v} \int_{\mathcal{D}} C(h)dh \quad (3.3)$$

Le terme $\kappa(\mathcal{I})$ (κ lorsqu'il n'y a pas d'ambiguïté) est appelée portée intégrale¹.

On pourrait supposer que l'importance de cette variance est faible lorsque l'on s'intéresse à des surfaces réelles, et que la comparaison des moyennes suffit largement à classer des surfaces dans les cas courants... Il n'en est rien, et l'illustration de la section suivante se veut un exemple de la nécessité de prendre en compte cette variance.

3.2.2. Exemples d'utilisation

Pour justifier l'utilisation de la variance d'extension, nous allons effectuer quelques opérations sur des images construites à partir de modèles aléatoires. Ces images ont l'avantage de posséder des propriétés statistiques parfaitement connues, et en particulier un covariogramme et une portée intégrale calculable analytiquement. Il est important de souligner que les modèles de synthèse utilisés, les fonctions de dilution, sont bien loin d'être des cas particuliers pathologiques. En fait, ces modèles sont très utiles pour simuler des matériaux ou des phénomènes réels divers, et leur aspect n'est pas si éloigné de celui des surfaces de type "peau d'orange" : les exemples présentés ici ne sont donc pas des phénomènes de foire, mais bien des exemples réalistes qui pourraient se présenter au détour d'un laminoir.

a. Fonctions de dilution

Les fonctions de dilution [58] sont créées à partir d'un agencement de structures aléatoires et rentrent dans un ensemble plus vaste de modèles basés sur l'implantation de grains primaires, aléatoires ou non, centrés en chacun des événements d'un processus de Poisson spatial².

Processus de Poisson ponctuel :

Il s'agit d'un ensemble $\mathcal{P}$ de points aléatoires qui vérifie les propriétés suivantes :

- Si A est un ensemble de $\mathbb{R}^2$ de surface $|A|$, alors la variable aléatoire $N(A)$ correspondant au nombre de points du processus appartenant à A suit une loi de Poisson de paramètre $\theta|A|$:

$$\forall A \subset \mathbb{R}^2, N(A) \sim \text{Poisson}(\theta|A|)$$

- Si $A_1, A_2, \dots, A_n$ sont n ensembles deux à deux disjoints, alors les variables aléatoires $N(A_1), N(A_2), \dots, N(A_n)$ sont indépendantes.

¹Dans les ouvrages de géostatistique classique, on appelle souvent portée intégrale l'intégrale de la fonction de corrélation, c'est à dire $\frac{\kappa}{\sigma^2}$

²Dans sa version la plus générique, ce modèle peut évoluer au cours du temps, mais nous nous placerons ici dans le cas stationnaire

Chapitre 3. Stabilité de la description

Grains primaires :

Les grains primaires sont une famille $(\mathcal{G}_u)_{u \in \mathcal{D}}$, de fonctions aléatoires définies sur un domaine $\mathcal{D}$ dont le support est presque sûrement compact.

Dans le cas des fonctions de dilution, on plante un grain primaire en un point u en ajoutant la valeur d'une réalisation de $\mathcal{G}_u$. Le résultat final est donc la fonction :

$$\mathcal{F}_d(x) = \sum_{p \in \mathcal{P}} \mathcal{G}_p(x)$$

Dans ce qui suit, $\mathcal{G}_p$ est un disque de niveau de gris constant centré en p et dont le rayon est déterminé en chaque point p par la réalisation d'une même variable aléatoire R . Les propriétés statistiques des fonctions de dilution que nous considérons seront donc déterminées par :

- θ la densité du processus de Poisson spatial. Dans la suite, nous choisirons une densité $\theta = 0.1$, ce qui correspond à environ 26000 grains implantés sur une image de taille 512×512 .
- $f(r)$, la densité de probabilité de la variable aléatoire R caractérisant la distribution des rayons des disques (il pourra s'agir de mesure de Dirac dans le cas discret).
- g le niveau de gris utilisé pour l'ensemble des grains primaires qui sera choisi égal à 1 pour tous les modèles étudiés dans la suite.

Dans ces conditions, on peut établir les résultats suivants :

$$\begin{aligned} S &= \pi R^2 \\ E(\mathcal{F}_d) &= \theta \cdot E(S) \cdot g \\ V(\mathcal{F}_d) &= \theta \cdot E(S) \cdot g^2 \\ C_{\mathcal{F}_d}(h) &= \theta \cdot E(\mathcal{G}(x) \cdot \mathcal{G}(x+h)) \\ \kappa(\mathcal{F}_d) &= \theta \cdot E(S^2) \cdot g^2 \end{aligned} \tag{3.4}$$

$$\tag{3.5}$$

La moyenne et la variance de la fonction aléatoire dépendent de la surface moyenne des grains primaires utilisés, et sont donc ici proportionnelles à la moyenne des carrés des rayons. La portée (valeur de h à partir de laquelle $C(h) = 0$) dépend du diamètre maximum des grains primaires, dans le cas d'une distribution de disques, il s'agit de la borne inférieure d_m pour lesquels $d > d_m \Rightarrow f(d/2) = 0$. Si $f(r)$ n'est pas à support borné, la portée peut être infinie, mais c'est rarement le cas en pratique lorsqu'on travaille sur des distributions discrètes. La portée intégrale κ dépend de la moyenne du carré de la surface des grains, et fait donc intervenir le moment d'ordre 4 de la variable aléatoire R .

b. Application

Dans cette section, nous cherchons à construire deux fonctions aléatoires $\mathcal{I}_1$ et $\mathcal{I}_2$ avec les propriétés suivantes :

- $\mathcal{I}_1$ et $\mathcal{I}_2$ ont des "allures" semblables (et en particulier des variances similaires), en d'autres termes, il n'est pas complètement impossible que ces deux fonctions soient issues de deux phénomènes physiques comparables. Dans le contexte de cette étude, il pourrait s'agir par exemple de deux surfaces obtenues avec des paramètres process différents.
- $\mathcal{I}_1$ et $\mathcal{I}_2$ ont deux moyennes m_1 et m_2 différentes.
- $\mathcal{I}_1$ et $\mathcal{I}_2$ ont des portées intégrales κ_1 et κ_2 significativement différentes ($\kappa_1 < \kappa_2$)

Avec de telles propriétés, on peut obtenir des réalisations de $\mathcal{I}_2$ dont la moyenne m sur un domaine v sera plus proche de m_1 que de m_2 . La seule comparaison de la moyenne d'une réalisation I avec les moyennes de références ne permet donc pas de savoir si l'image observée I est une réalisation de $\mathcal{I}_1$ ou $\mathcal{I}_2$. En revanche, si l'on tient compte des variances d'extension de $\mathcal{I}_1$ et $\mathcal{I}_2$ sur le domaine v , il sera possible d'attribuer à I un label 1 ou 2 de manière plus pertinente, et en tout état de cause, *d'estimer le risque d'erreur* encouru lors de l'attribution de ce label, ce qui est essentiel.

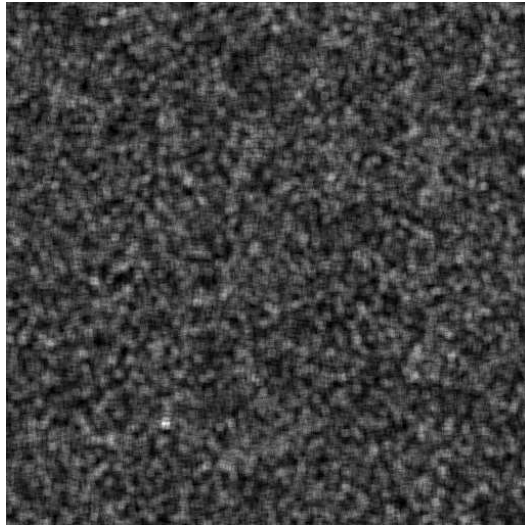
Les deux fonctions aléatoires que nous allons construire sont deux fonctions de dilution obtenues pour des distributions de rayons différentes. La première distribution est un Dirac en $R_0 = 4$ (les disques utilisés sont donc de diamètre constant) (voir figure 3.1(a)), et la seconde suit une distribution exponentielle de paramètre a discrétisée et tronquée pour des rayons supérieurs à 20 (voir figure 3.1(b)) :

$$\begin{aligned} a &= \frac{R_0}{\sqrt{2}} \\ \forall i \in [1, 19], P(R = i) &= e^{-\frac{i-1}{a}} - e^{-\frac{i}{a}} \\ P(R = 20) &= e^{-\frac{19}{a}} \end{aligned}$$

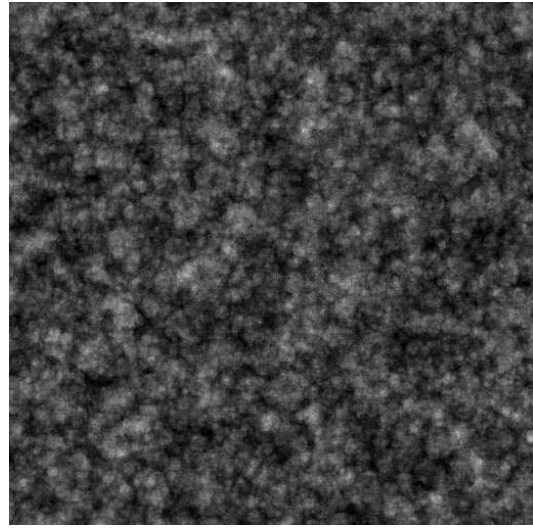
Il est important de noter que la formule 3.4 est valable dans le cas de vrais disques. Dans nos simulations, les grains primaires implantés ne sont que des approximations discrètes de disques, et la formule $S = \pi R^2$ n'est qu'approchée. On peut cependant estimer les caractéristiques statistiques de nos fonctions de dilution discrètes en utilisant les espérances réelles des surfaces de nos grains primaires discrets dans les relations 3.5. On obtient alors les résultats suivants :

$\mathcal{I}_1$	$\mathcal{I}_2$
$E(\mathcal{I}_1) = 6$	$E(\mathcal{I}_2) = 6.03$
$V(\mathcal{I}_1) = 6$	$V(\mathcal{I}_2) = 6.03$
$\kappa_1 = 360$	$\kappa_2 = 1657.303$

En ajoutant une constante à la fonction de dilution choisie pour obtenir $\mathcal{I}_1$, on peut modifier son espérance tout en conservant les autres propriétés statistiques (variance ponctuelle et portée intégrale). On peut alors observer plusieurs réalisations de $\mathcal{I}_1$ et $\mathcal{I}_2$ sur un volume v fixé, et observer les moyennes de chacune d'entre elle. On obtient une variance de ces moyennes proche de la valeur théorique κ/v , et la différence entre κ_1 et κ_2 induit l'effet prévu : certaines réalisations de $\mathcal{I}_2$ ont des moyennes plus proche de $E(\mathcal{I}_1)$ que de $E(\mathcal{I}_2)$. Les résultats pour 100 réalisations de $\mathcal{I}_1$ et $\mathcal{I}_2$ sur des carrés de 70 pixels de côtés sont visibles à la figure 3.2(a).

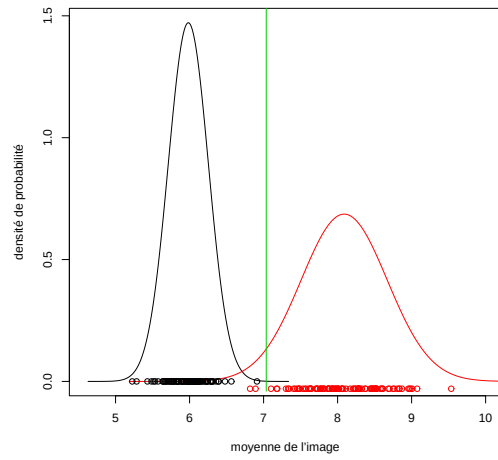


(a) Fonction de dilution obtenue à partir de disques de rayon constant (4 pixels) .

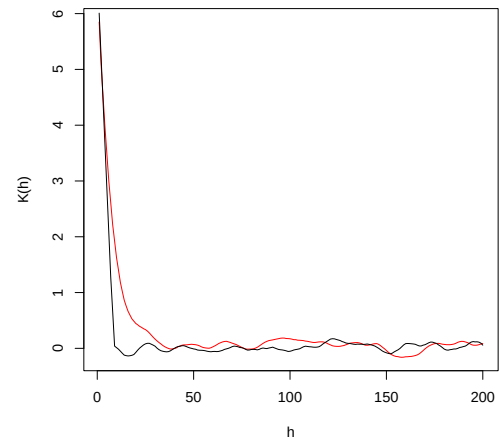


(b) Fonction de dilution obtenue à partir de disque dont le rayon suit une loi exponentielle discrète .

FIG. 3.1.: Fonction aléatoire de dilution



(a) Moyennes obtenus pour 100 réalisations des fonctions $\mathcal{I}_1$ et $\mathcal{I}_2$. Des densités de probabilités gaussiennes de moyenne $E(\mathcal{I}_1)$ et $E(\mathcal{I}_2)$ et de variance κ_1/v et κ_2/v sont représentées à titre indicatif. La milieu entre les deux moyennes théoriques est représenté par une ligne verticale.



(b) Covariogrammes expérimentaux de deux fonctions de dilution (rouge : disques de rayon exponentiel, portée $\simeq 40$, en noir : disque de rayon constant, portée $\simeq 8$)

FIG. 3.2.: Propriétés de deux fonctions de dilution

c. Estimation de la variance d'extension

Dans le cas des modèles proposés, le covariogramme et la portée intégrale peuvent être calculés analytiquement, et la valeur théorique de la variance d'extension pour une taille de champ donnée est donc parfaitement connue. Pour des applications réelles, les fonctions aléatoires ne sont pas connues explicitement, et il faut donc estimer la portée intégrale à partir d'une réalisation sur un domaine v , dont les dimensions se doivent d'être grande devant la portée pour obtenir des résultats pertinents. La méthode naïve consisterait à évaluer tout d'abord la fonction de covariance grâce à la méthode décrite en 2.2.1 puis à l'intégrer. Malheureusement, la valeur estimée $C(h)$ peut fluctuer de manière importante pour les grandes valeurs de h et osciller autour de la valeur 0 pour des valeurs de h proche de la portée (voir figure 3.2(b)). L'intégration numérique du covariogramme devient donc assez instable et les résultats peuvent devenir très hasardeux : sur les modèles précédents, en faisant varier le domaine d'intégration du covariogramme au delà de la portée, on peut obtenir des erreurs de plus de 100% sur la variance, voire même obtenir des valeurs négatives dans certains cas catastrophiques !

L'autre méthode, plus stable, qui est le plus souvent préconisée [43, 44] se base sur la notion de variance d'un ensemble de taille v dans un ensemble de taille V . On considère un ensemble V obtenu par union de N ensembles disjoints équivalents v_i , et on s'intéresse à la variable aléatoire :

$$s^2(v|V) = \frac{1}{N} \sum_{i=1}^N (Z_{v_i} - Z_V)^2 \quad (3.6)$$

En particulier, on note $\sigma^2(v|V) = E(s^2(v|V))$, et on peut montrer la formule suivante :

$$\sigma^2(v|V) = \frac{1}{v^2} \int_v dx \int_v C(y-x) dy - \frac{1}{V^2} \int_V dx \int_V C(y-x) dy$$

Et si à la fois v et V sont de grandes dimensions par rapport à la portée, on obtient la valeur approchée suivante :

$$\sigma^2(v|V) \simeq \kappa \left(\frac{1}{v} - \frac{1}{V} \right) \simeq \kappa \frac{1}{v} \quad (3.7)$$

Le protocole d'estimation de κ consiste donc à diviser l'image de taille V étudiée en plusieurs domaines v_i disjoints de taille v , à calculer la moyenne de l'image sur chacun de ces domaines puis à déterminer $s^2(v|V)$. En répétant ces opérations pour plusieurs valeurs de v et en utilisant la relation 3.7, on peut estimer κ par régression de Z_v en fonction de $1/v - 1/V$. Les résultats obtenus sont beaucoup plus stables que précédemment, mais il faut tout de même mettre un bémol quand à leur précision. Tout d'abord, l'égalité 3.7 est une approximation qui n'est valable que dans le cas où les dimensions de v et V sont grandes devant la portée, ce qui signifie que les domaines v utilisés pour la régression ne doivent pas être trop petits. Si v est au contraire trop grand, l'image ne contiendra qu'un petit nombre de domaines de taille v disjoints, et c'est l'estimation de $s^2(v|V)$ qui pourra être assez éloignée de $\sigma^2(v|V) = E(s^2(v|V))$. Les valeurs de v admissibles pour l'estimation de κ sont donc limitées, et il faut y prêter attention en pratique, car cela peut aussi induire des variations significatives dans la valeur de κ estimée.

Nous avons appliqué cette méthode aux fonctions de dilution de la section précédente (figure 3.3) : On divise une image de 512 pixels de côtés en carrés plus petits (côté compris entre 40 et 100 pixels) on obtient les portées intégrales avec une précision de l'ordre de 10% (droite noire vs droite bleue sur la figure 3.3) pour les disques de rayon constant (et de portée faible $\simeq 8$ pixels) et de 21% (droite rouge vs droite bleue sur la figure 3.3) pour les disques de rayon variable (portée $\simeq 40$ pixels). Dans ce dernier cas, l'erreur d'estimation tombe à 2% (droite verte vs droite bleue sur la figure 3.3) si l'on ne prend en compte que les carrés de côté ≥ 75 , c'est à dire les sous-domaines v pour lesquels l'approximation " $v \gg$ portée" est la plus réaliste. On constate donc assez nettement l'impact du rapport entre taille de champ et portée sur la qualité de l'estimation.

Si la précision obtenue, de l'ordre de 10%, peut paraître médiocre au premier abord, nous verrons dans la suite que la prise en compte de cette valeur dans les méthodes d'analyse multivariable fournit un outil pertinent pour juger de la qualité des résultats et permet d'éviter certains pièges de sur-apprentissage.

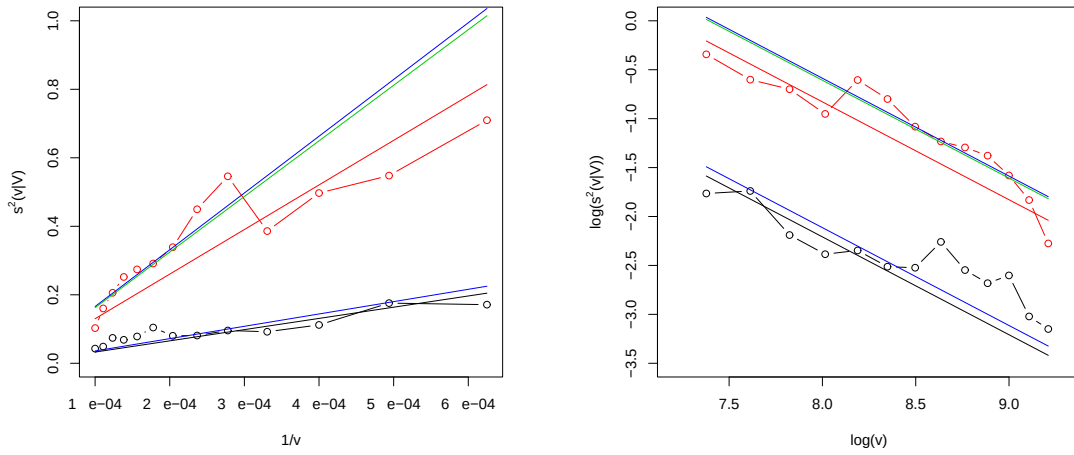


FIG. 3.3.: Estimation de la portée intégrale. Les points et droites représentent les valeurs expérimentales de $s^2(v|V)$ en fonction de $1/v$ (figure de gauche) et la régression linéaire correspondante obtenues pour des fonctions aléatoires de dilution (rouge : disques de rayon variant selon une loi exponentielle. vert : droite interpolée seulement à partir des 6 premiers points rouges. noir : disques de rayon constant). Les droites bleues représentent les droites théoriques obtenue par la relation 3.7 et les valeurs réelles des portées intégrales. La figure de droite représente les mêmes quantités en échelle logarithmique.

3.2.3. Limitations pratiques

La variance d'extension est un outil théorique intéressant parce que s'employant aussi bien dans un contexte continu que dans un contexte discret, mais son utilisation peut

conduire à des difficultés pratiques. Tout d'abord, son cadre d'application est restreint aux descripteurs pouvant s'exprimer comme moyenne d'une fonction sur un ensemble, ce qui n'est pas toujours directement le cas (voir le cas des courbes de tortuosité ou des histogrammes issus d'une segmentation par exemple). Même s'il est parfois possible de se ramener à une configuration favorable par une transformation de l'image étudiée, cela entraîne souvent des calculs assez coûteux : si l'on s'intéresse à un histogramme par exemple, il faudra effectuer un calcul de variance d'extension pour chacune de ses classes.

Ensuite, il est parfois intéressant de combiner plusieurs sources de variance dans un même schéma, ce que la donnée de la variance d'extension ne permet pas facilement. Imaginons que l'on s'intéresse à plusieurs échantillons considérés comme équivalents pour une application donnée, chacun d'entre eux va conduire au calcul d'une variance d'extension, qu'il faudra être capable d'intégrer de manière pertinente à la variance inter-échantillons.

Ces deux raisons principales vont nous amener à utiliser des méthodes de sur-échantillonnage de type bootstrap pour résoudre en pratique le problème de l'intervalle de confiance associé aux surfaces étudiées.

3.3. Bootstrap

3.3.1. Origine et principe

Le bootstrap est une technique de sur-échantillonnage qui a été introduite dans les années 80 par Efron[22] qui cherchait à calculer *informatiquement* des variances ou des intervalles de confiance. Cette méthode tire son nom des aventures du baron de Münchausen qui, empêtré dans un marécage, dût s'en extraire en se tirant par les lanières de ses propres bottes (bootstrap en anglais). Dans un cadre statistique, l'analogie consiste à extraire non seulement une information, mais aussi la précision avec laquelle cette information est connue à partir des mêmes données. Plus récemment, Bigerelle [16] a utilisé le bootstrap pour quantifier la relation entre la brillance d'un échantillon métallique et différents paramètres de rugosité caractérisant la surface.

Dans cette section, on utilisera les notations suivantes :

- F désigne une densité de probabilité représentant la loi d'une variable aléatoire X .
- E désigne l'espérance mathématique associée à F .
- $\theta(F)$ ou θ quand il n'y a pas d'ambiguïté désigne un paramètre d'intérêt associée à F .
- $x = (x_1, x_2, \dots, x_n)$ désigne un n -échantillon obtenu à partir de F par n tirages aléatoires *indépendants*.
- $\hat{F}$ désigne une estimation de F calculée à partir de $(x_1, x_2, \dots, x_n)$.
- $\hat{E}$ désigne l'espérance mathématique associée à $\hat{F}$.
- $\hat{\theta} = \hat{\theta}(x)$ est une statistique associée à x utilisée comme estimateur de θ . En général, $\hat{\theta} = \theta(\hat{F})$.
- $\hat{\Theta}$ désigne la variable aléatoire dont $\hat{\theta}$ est une réalisation conditionnée par le n -échantillon $(x_1, x_2, \dots, x_n)$.

- $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ désigne un échantillon bootstrap comme il sera défini dans la suite, et auquel on associe les quantités F^* , $\theta^* = \hat{\theta}(x^*)$ (en général, $\theta^* = \theta(F^*)$) et Θ^* .

Supposons que l'on cherche à estimer un paramètre $\theta(F)$ à partir d'un échantillon $(x_1, x_2, \dots, x_n)$ obtenu à partir de F par n tirages aléatoires indépendants. Une méthode classique procède en deux étapes :

- On estime une distribution $\hat{F}$ à partir de $(x_1, x_2, \dots, x_n)$. Cette distribution peut être simplement la distribution dite "plug-in" qui attribue à chaque réalisation x_i une probabilité égale à $1/n$:

$$\hat{P}(X < x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{x_i < x}$$

Où $\mathbb{1}_{x_i < x}$ désigne la fonction définie par $\mathbb{1}_{x_i < x} = 1$ si $x > x_i$ et $\mathbb{1}_{x_i < x} = 0$ sinon.

$\hat{F}$ peut aussi être une distribution paramétrique dont les paramètres sont obtenus à partir de $(x_1, x_2, \dots, x_n)$, par exemple une gaussienne de moyenne et variance calculées à partir de l'échantillon observé.

- On calcule $\hat{\theta}$ à partir de $\hat{F}$ et on l'utilise comme estimateur de θ

On s'interroge alors sur la précision de cette estimation, ce qui en termes statistiques revient à s'intéresser au biais et à la variance de la variable aléatoire $\hat{\Theta} - \theta$. Par exemple, si θ est simplement l'espérance mathématique $E(F)$, l'estimateur $\hat{\theta}$ associé à la distribution "plug-in" est simplement la moyenne d'échantillon $\bar{x}$ dont on sait qu'il est sans biais ($E(\hat{\Theta}) = E(\bar{x}) = \theta = E(F)$) et de variance $V(\hat{\Theta}) = \sigma^2/n$ où σ^2 désigne la variance de F (en supposant qu'elle existe).

Dans le cas simple de la moyenne, on peut facilement dérouler des calculs pour obtenir une expression analytique de la variance, malheureusement, pour des estimateurs θ plus complexes, ces calculs sont souvent inextricables, et c'est l'une des raisons de l'introduction du bootstrap dont le principe est de "remplacer" informatiquement la distribution F inconnue par la distribution $\hat{F}$ parfaitement connue.

On définit un échantillon bootstrap $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ comme n réalisations indépendantes d'une variable aléatoire $\hat{X}$ de distribution $\hat{F}$. A partir de x^* , on peut estimer une distribution F^* par le même processus qui permet d'estimer $\hat{F}$ à partir de x , puis calculer $\theta^* = \theta(F^*)$. Tout le principe du bootstrap consiste à utiliser la variable aléatoire $\Theta^* - \hat{\theta}$ pour estimer les propriétés de $\hat{\Theta} - \theta$ (espérance et variance par exemple), ce qui revient en quelque sorte à utiliser le principe du "plug-in" pour estimer l'erreur commise en utilisant $\hat{\theta}$ à la place de θ .

Pour pouvoir utiliser le bootstrap, il faut donc être capable de calculer des statistiques de $\Theta^* - \hat{\theta}$. En règle générale, les expressions exactes sont impossibles à calculer dès que $\hat{\theta}$ est une statistique complexe, mais la distribution $\hat{F}$ étant parfaitement connue, il est possible de simuler informatiquement non pas un, mais un grand nombre B d'échantillons bootstrap $x_1^*, \dots, x_B^*$ (chaque échantillon étant lui-même constitué de n individus). Ces échantillons permettent d'obtenir B réalisations $\theta^{1*}, \dots, \theta^{B*}$ de Θ^* , et on peut les utiliser pour estimer les propriétés de $\Theta^* - \hat{\theta}$. Plus B est important, plus les statistiques de $\Theta^* - \hat{\theta}$

seront connues avec précision, mais plus le temps de calcul (tirage aléatoire et calcul de θ^*) sera long et il faut donc trouver un compromis acceptable.

La qualité des résultats obtenus, en particulier l'estimation de la variance $V(\hat{\Theta} - \theta)$ qui est en quelque sorte l'équivalent de la variance d'extension dans un contexte *discret*, dépend essentiellement de trois choses :

- La qualité de l'approximation de F par $\hat{F}$ dans laquelle la taille n de l'échantillon considéré et les suppositions faites a priori dans le cas où $\hat{F}$ est une distribution paramétrique jouent un rôle essentiel.
- La sensibilité du paramètre θ aux petites variations de $\hat{F}$.
- La qualité du tirage aléatoire permettant d'obtenir x^* à partir de $\hat{F}$. Cette opération peut-être extrêmement simple comme très complexe dans le cas où la variable aléatoire X est en fait un vecteur de plusieurs variables non indépendantes.

Dans la suite, nous nous intéressons à l'utilisation et à l'adaptation des méthodes de type bootstrap pour la caractérisation de surfaces rugueuses par les descripteurs statistiques et morphologiques décrits dans le chapitre 2.

3.3.2. Descripteurs "favorables" (indépendance des individus)

Dans un premier temps, nous nous intéressons au cas le plus favorables où les principes du bootstrap peuvent être appliqués tels quels, en particulier le cas où la condition d'indépendance des réalisations $x_1, x_2, \dots, x_n$ au sein de l'échantillon x est vérifiée. Ce sera le cas lorsque les individus considérés sont plusieurs profils indépendants ou plusieurs zones issues d'une segmentation par exemple, mais ça ne le sera plus si comme dans la section précédente, on s'intéresse à des réalisations de fonctions aléatoires dont le covariogramme diffère d'un Dirac en $h = 0$, ce cas sera traité ultérieurement à la lumière des résultats obtenus concernant la variance d'extension.

a. Profils indépendants

Dans ce paragraphe, l'échantillon x consiste en n profils *indépendants* $(x_1, x_2, \dots, x_n)$. De ces profils nous extrayons une statistique θ qui peut être par exemple la moyenne d'un paramètre de rugosité classique type *Ra*, ou la moyenne des courbes de *Ma* (voir section 2.2.6), ou encore toute autre statistique pertinente pour une application donnée.

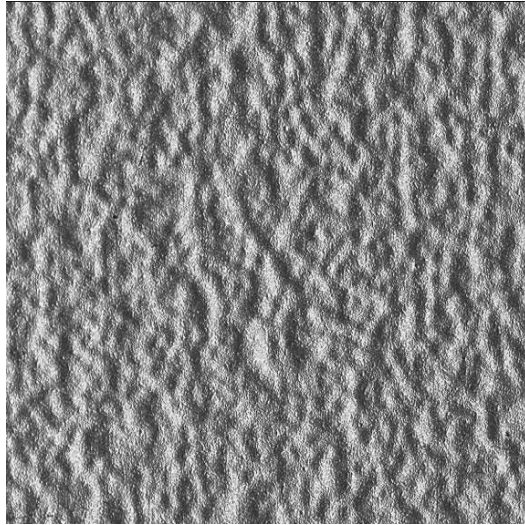
La distribution F des profils est inconnue, et on utilise la distribution "plug-in" $\hat{F}$ pour construire des échantillons bootstrap. Dans ce cas, on obtient x^* en tirant au hasard, de manière uniforme *et avec remise* un n-échantillon parmi les n profils $(x_1, x_2, \dots, x_n)$ (cela signifie que x^* peut contenir plusieurs fois le même profil). On effectue cette opération B fois, et pour chaque échantillon bootstrap x_i^* , on recalcule la statistique d'intérêt θ_i^* . La variance et la moyenne de Θ^* comme estimateur de $\hat{\theta}$ donne alors une bonne estimation de la variance et la moyenne de $\hat{\Theta}$ comme estimateur de θ .

b. Segmentation

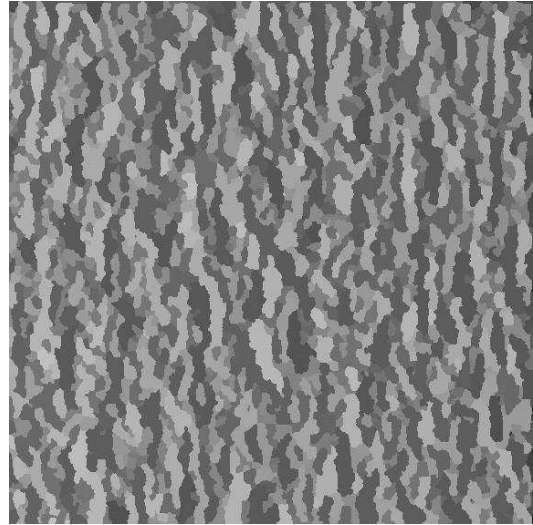
Le cas des histogrammes obtenus à partir d'une segmentation de l'image est particulièrement intéressant, parce qu'il permet d'appliquer la méthode du bootstrap "standard" sur une structure de données à laquelle s'adaptent aussi facilement les techniques d'estimation de la variance d'extension. Imaginons que nous disposions d'une image segmentée en n zones $z_1, \dots, z_n$. Pour chaque zone z_i , on note respectivement s_i et h_i la surface et l'altitude moyenne (ou niveau de gris moyen) sur la zone, que l'on suppose obtenues par tirage d'un couple de variable aléatoire (S_i, H_i) . On construit ensuite un histogramme expérimental formé de K classes $\hat{c}_1, \dots, \hat{c}_K$ contenant chacune la surface totale de l'image appartenant à des zones dont l'altitude moyenne est comprise entre deux valeurs seuils h^k et h^{k+1} . Il s'agit donc d'un histogramme des altitudes des zones pondérées par leur surface. La supposition sous-jacente à l'utilisation des méthodes de bootstrap est l'indépendance des attributs entre deux zones différentes :

$$\forall i, j \in [1, n], i \neq j \Rightarrow (S_i, H_i) \text{ indépendant de } (S_j, H_j)$$

Il s'agit d'une supposition préalable qui n'est pas complètement vérifiée entre deux zones adjacentes, et nous en verrons l'influence par la suite. Si l'indépendance (supposée) entre deux couples de variables est importante pour le fonctionnement du bootstrap, on peut remarquer que l'indépendance entre S_i et H_i à l'intérieur de chaque couple n'est pas essentielle.



(a) Image originale (image caméra)



(b) Image segmentée. Le niveau de gris est le niveau de gris moyen sur la zone.

FIG. 3.4.: Segmentation en zone.

Ce que l'on cherche à obtenir ici est une précision de l'approximation des valeurs réelles $C = (c_1, \dots, c_k)$ par les valeurs observées $\hat{C} = (\hat{c}_1, \dots, \hat{c}_K)$, et pour résumer cette

information, nous nous intéresserons à la variance de la norme du vecteur $\hat{C} - C$:

$$\|\hat{C} - C\|^2 = \sum_{k=1}^K (\hat{c}_i - c_i)^2$$

Nous cherchons à estimer cette variance par bootstrap pour plusieurs taille de champ, ce qui conduit au protocole suivant :

- On définit un champ carré de côté a .
- On évalue le nombre de zones n^* intersectant un carré de côté a par la partie entière de $n.a^2/S_0$ où S_0 désigne la surface de l'image initiale. Cela revient à dire que le nombre de zones intersectant un carré donné est approximativement proportionnel à sa surface.
- On calcule B échantillons bootstrap constitué chacun de n^* zones, c'est à dire de n^* couple (s, h) tirés au hasard avec remise parmi les n couples observés dans l'image initiale.
- Pour chaque échantillon bootstrap, on détermine C_b^* , puis on calcule la variance expérimentale de $\|C^* - \hat{E}(C^*)\|$ au moyen de la formule :

$$\frac{1}{B} \sum_{b=1}^B \|C_b^*\|^2 - \left\| \frac{1}{B} \sum_{b=1}^B C_b^* \right\|^2$$

Nous avons effectués ces calculs à partir de l'image 3.4(b), avec $B = 50$, pour des carrés de côté a compris entre 50 et 550 (pour une image initiale de taille 574×574). Les résultats obtenus sont représentés par une courbe noire sur la figure 3.5, et peuvent être comparés aux points représentés en rouge sur le même graphique qui correspondent à l'évaluation de la variance sur des sous-images selon le processus décrit au paragraphe c (division de la surface en plusieurs sous-domaines disjoints, puis estimation de la variance de la moyenne de l'image sur ces sous-domaines). Pour des raisons de stabilité dans le calcul, les variances expérimentales n'ont été calculées que pour des carrés de côté compris entre 50 et 200 pixels de côté, mais sur cette portion, les résultats expérimentaux, sont en bon accord avec les résultats fournis par la méthode du bootstrap.

c. Compatibilité avec la variance d'extension

Pour évaluer la compatibilité des résultats fournis par le bootstrap et ceux fournis par la théorie de la variance d'extension, nous avons appliqué la méthode du paragraphe 3.2.2c à l'image 3.4(b) afin de comparer les résultats obtenus. Pour ce faire, il est d'abord nécessaire d'exprimer le descripteur "histogramme des hauteurs de zones" comme la moyenne d'une fonction aléatoire sur un ensemble donné. Nous avons choisi ici de construire une image vectorielle $\mathcal{J}$, c'est à dire une image qui associe à chaque point du plan non pas une valeur de gris mais un vecteur de dimension K (le nombre de classe de l'histogramme). La construction se fait comme suit : soit p un pixel de l'image initiale, et z_i la zone à

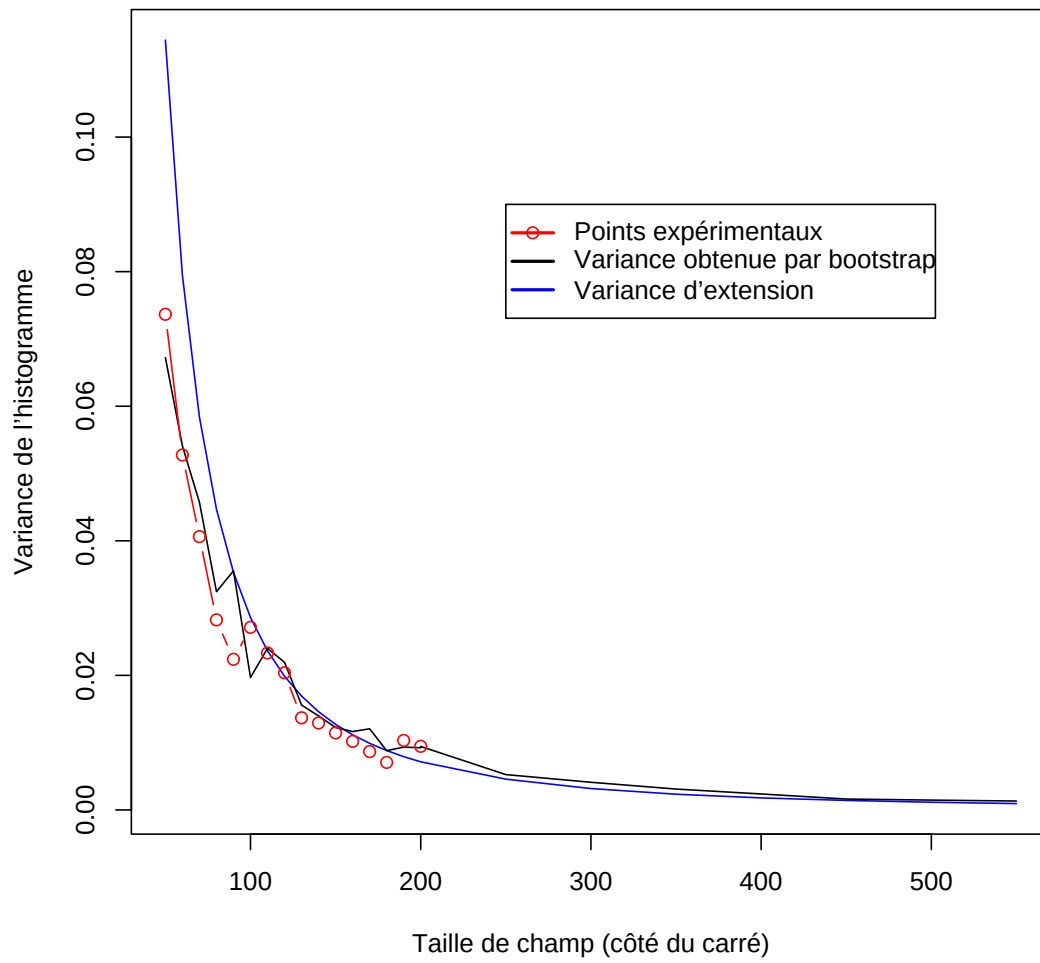


FIG. 3.5.: Estimation de la variance des histogrammes issus d'une segmentation.

laquelle appartient p , alors :

$$\begin{aligned} \mathcal{J}(p) &= (j_1, \dots, j_K) \\ h_i \in [h^k, h^{k+1}[&\Rightarrow j_k = 1 \\ \forall r \in [1..K], r \neq k &, \quad j_r = 0 \end{aligned}$$

Dans ces conditions, la distribution expérimentale $\hat{C} = (\hat{c}_1, \dots, \hat{c}_K)$ est bien égale à la moyenne de $\mathcal{J}$, composante par composante.

On peut alors diviser l'image en N sous-domaines disjoints v_i de taille v et calculer $s^2(v|V)$ au moyen de la formule 3.6 adaptée pour travailler sur la norme des vecteurs Z :

$$s^2(v|V) = \frac{1}{N} \sum_{i=1}^N \|Z_{v_i} - Z_V\|^2$$

Ceci nous permet, d'après la relation 3.7, de calculer une portée intégrale par régression linéaire pour plusieurs taille de sous-domaine v . Les points expérimentaux pour des domaines carrés de côté compris entre 100 et 200 pixels sont utilisés dans la régression. La limite 100 est choisie pour avoir une taille de champ minimum de l'ordre de deux fois la portée de l'image initiale qui est approximativement de 50 pixels. Les points expérimentaux sont représentés en rouge sur la figure 3.5 tandis que la courbe correspondant à la variance d'extension calculée par l'approximation 3.3 en fonction de la taille de champ v est représentée en bleu sur cette même figure. On remarque d'ailleurs la limite de validité de cette approximation pour les faibles tailles de champ.

Cette étude montre que les approches très différentes que sont le bootstrap et la variance d'extension fournissent des résultats tout à fait compatibles dans le cas où ces deux théories sont facilement applicables. On peut au passage souligner les points suivants : la portée intégrale calculée ici, ne nous permet de calculer que la variance de la norme de l'écart entre $\hat{C}$ et C , alors que le bootstrap peut fournir sans coût supplémentaire les variances correspondant à chaque composante (ce qui réclamerait autant de régression linéaire pour le calcul des portées intégrales correspondantes dans la méthode de la variance d'extension). Par contre l'estimation de la variance par bootstrap possède dans ce cas des difficultés intrinsèques plus profondes qu'un problème de précision dans une régression linéaire : l'indépendance des niveaux de gris entre deux zones adjacentes est une simplification de la structure de l'image nécessaire au bon fonctionnement du bootstrap, mais qui conduit, pour les grandes tailles de champ, à des variances systématiquement légèrement plus importantes que celles obtenues par la portée intégrale qui tient compte automatiquement de ces dépendances à travers le covariogramme sous-jacent. De la même façon, le tirage d'échantillons bootstrap est imparfait, car il faudrait en théorie conditionner le tirage des zones par la surface totale à couvrir plutôt que simplement par le nombre de zones. On pourrait aussi stratifier la population des zones en zones entières et zones de bord de champ, car ces deux populations ont des statistiques différentes en terme de surface... Ces approximations mènent à un résultat imparfait, mais la précision obtenue sur la variance est néanmoins tout à fait suffisante pour la plupart de nos applications, et par ailleurs, cette précision est probablement du même ordre de grandeur que celle de la

variance d'extension obtenue par régression linéaire. Finalement, on peut conclure qu'il est légitime et surtout cohérent d'utiliser soit la variance d'extension (contexte continu, ou approximation discrète d'une fonction aléatoire continue), soit le bootstrap (cadre complètement discret) en fonction de la situation, de l'image observée, et de l'objectif final, qui dans le cadre du bootstrap, peut dépasser le calcul d'une simple variance.

3.3.3. Bootstrap sous dépendance

Lorsque la condition d'indépendance des réalisations $x_1, x_2, \dots, x_n$ n'est pas réalisée, la méthode "classique" de bootstrap ne s'applique plus : si l'on tire au hasard avec remise n valeurs à l'intérieur de l'échantillon, on ne rend pas compte des interactions, parfois complexes, existant entre les individus, et on peut obtenir un résultat fortement biaisé. C'est le cas par exemple si les données sont issues d'une acquisition spatiale d'un phénomène physique variant peu entre deux points d'acquisitions.

Obtenir des échantillons bootstrap réalistes à partir de données dépendantes constitue un axe de recherche très actif en statistique, et la méthode classiquement utilisée consiste à effectuer du "bootstrap par bloc", c'est à dire à tirer au hasard avec remise non pas un individu, mais un groupe d'individus $x_i, \dots, x_{i+l}$ de taille $l < n$ issu de l'échantillon initial, de façon à réobtenir un échantillon x^* de taille n conservant une partie des propriétés spatiales de l'échantillon original et fournissant ainsi un candidat plus réaliste qu'un échantillon bootstrap classique à utiliser dans des calculs de variance, de distribution ou d'intervalle de confiance. Notre objectif n'est pas ici de rentrer dans les détails qui sous-tendent la théorie du bootstrap par bloc, ni de faire une étude exhaustive de l'ensemble des possibilités existantes (moving block bootstrap, stationary block bootstrap, circular block bootstrap, tapered block bootstrap ...), mais plutôt d'extraire de la littérature abondante quelques résultats nous permettant d'appliquer ces techniques à des surfaces réelles. La plupart des résultats théoriques présentés dans cette section proviennent d'un ouvrage de Lahiri[39] auquel il est recommandé de se reporter pour de plus amples détails mathématiques, en particulier pour tout ce qui concerne les hypothèses (parfois techniques) et énoncés exacts des théorèmes de convergence mentionnés.

Dans un premier temps nous illustrerons l'emploi du bootstrap par blocs sur un exemple simple, puis nous passerons en revue les différents descripteurs du chapitre 2 non abordés dans le paragraphe 3.3.2 et bootstrapperons chacun d'eux de la façon la plus pertinente possible en validant nos approches sur des exemples issus de fonctions de dilutions.

a. Bootstrap par bloc

Dans un premier temps, nous présentons les résultats dans le cas classique d'une fonction (suite) aléatoire stationnaire $\mathcal{F}_d$ monodimensionnelle discrète dont nous considérons un sous-ensemble formé de n points $x_1, x_2, \dots, x_n$. La fonction de covariance de $\mathcal{F}_d$ sera notée C comme précédemment et sa portée intégrale κ . Dans le cadre de cette section, nous considérons une fonction de dilution dont le grain primaire est un segment de longueur s de niveau de gris constant g . La densité du processus de Poisson correspon-

dant sera noté ρ . La fonction de covariance associée à la fonction de dilution est donnée par la formule :

$$\begin{aligned} C(h) &= \rho(s - \|h\|)g^2 & \text{pour } \|h\| \leq s \\ C(h) &= 0 & \text{pour } \|h\| > s \end{aligned} \quad (3.8)$$

Nous avons choisis d'utiliser la technique du bootstrap par bloc circulaire qui génère des échantillons bootstrap de la façon suivante :

- Pour $k \in [1, \lceil n/l \rceil]$ (où $\lceil p \rceil$ désigne la partie entière supérieure de p), on construit un bloc U_k .
- Le point de départ de chaque bloc est un nombre $d(k)$ tiré au hasard et uniformément entre 1 et n .
- On considère ensuite le bloc $U_k = (x_{d(k)[n]}, x_{d(k)+1[n]}, \dots, x_{d(k)+l[n]})$ où $p[n]$ désigne la valeur p prise modulo n , ce qui revient à périodiser le profil, ou à le considérer comme circulaire.
- On construit ensuite la suite $(U_1, \dots, U_{\lceil n/l \rceil})$ que l'on tronque éventuellement pour fabriquer un échantillon bootstrap x^* de taille n .

Ce procédé de génération des données induit une loi de probabilité $\hat{P}$ sur l'ensemble des échantillons que l'on peut ainsi obtenir. Comme précédemment, une simulation informatique va nous permettre de remplacer la probabilité vraie P ayant conduit à l'échantillon x par la probabilité $\hat{P}$ afin d'obtenir un grand nombre d'échantillons (profils) bootstrap³ $x^{1*}, \dots, x^{B*}$. A partir de ces échantillons bootstrap, on évalue certaines caractéristiques de Θ^* , qui est à $\hat{P}$ ce que $\hat{\Theta}$ est à P . Le procédé est valide (sous certaines conditions techniques de convergence de la fonction de covariance) dans le sens où si n tend vers l'infini, l tend vers l'infini (mais l/n tend vers 0), alors les résultats obtenus pour Θ^* convergent en probabilité vers leur équivalent pour $\hat{\Theta}$. Cependant, les vitesses de convergence sont plus faibles que dans le cas du bootstrap classique, mais sont en général améliorées si l'on s'intéresse à la variable $\Delta^* = \Theta^* - \theta(\hat{P})$ (plutôt que $\Theta^* - \hat{\theta}$) comme estimateur de $\hat{\Theta} - \theta$. Dans le cas où la statistique utilisée est non biaisée ($\theta = E(\hat{\Theta})$), on obtient $\theta(\hat{P}) = \hat{E}(\Theta^*)$, ce qui conduit à $\Delta^* = \Theta^* - \hat{E}(\Theta^*)$ où $\hat{E}(\Theta^*)$ désigne l'espérance de Θ^* sous $\hat{P}$, espérance qui est estimée par la moyenne :

$$\bar{\theta}^* = \frac{1}{B} \sum_{b=1}^B \theta^{b*}$$

Notons maintenant G le moment d'ordre 1 de la fonction de covariance (discrète) associée à P , et D un paramètre lié à sa portée intégrale (discrète) κ :

$$\begin{aligned} G &= \sum_{k=-\infty}^{+\infty} |k|C(k) \\ D &= \frac{4}{3}\kappa^2 \end{aligned}$$

³La différence essentielle par rapport au formalisme du bootstrap classique est le passage d'une loi marginale F décrivant le comportement d'une seule variable aléatoire à une loi de probabilité P décrivant le comportement conjoint de plusieurs variables aléatoires non indépendantes.

Chapitre 3. Stabilité de la description

Dans le cas où la statistique d'intérêt est la moyenne d'échantillon ($\hat{\theta} = \bar{x}$), et où l'on s'intéresse particulièrement à $Var(\sqrt{n}.\bar{x})$, c'est à dire à la portée intégrale κ que l'on estime par bootstrap au moyen de σ_b^{2*} :

$$\sigma_b^{2*} = \frac{n}{B} \sum_{b=1}^B \bar{x}^{b*2} - n. \left(\frac{1}{B} \sum_{b=1}^B \bar{x}^{b*} \right)^2$$

Lahiri démontre les résultats suivants (sous certaines conditions de régularité sur P) :

Lorsque $n \rightarrow \infty$, $l \rightarrow \infty$, $l = o(\sqrt{n})$:

$$\begin{aligned} E(\sigma_b^{2*}) - \kappa &= -\frac{G}{l} + o\left(\frac{1}{l}\right) \\ Var(\sigma_b^{2*}) &= \frac{l}{n}D + o\left(\frac{l}{n}\right) \end{aligned}$$

Ce théorème permet en quelque sorte d'obtenir une taille de bloc l optimale [53] pour l'estimation de κ par σ_b^{2*} minimisant $\frac{l}{n}D + (\frac{G}{l})^2$. Il est important de noter que la recherche de cette taille de bloc optimale fait intervenir les valeurs de G et D inconnue, et qu'il faut donc estimer. Remarquons aussi que la connaissance de D est équivalente à la connaissance de κ qui est le but finale de notre analyse, ce qui peut sembler assez paradoxal : pour défendre cette approche, signalons qu'une évaluation très approximative de D et G peut conduire à une taille de bloc sous-optimale pour une taille d'échantillon donnée, mais ne remettra pas en cause la convergence en probabilité de σ_b^{2*} vers κ sous les conditions du théorème de Lahiri. Autrement dit, on gagne en précision en utilisant le bootstrap après une estimation grossière de D et G (par des tests de nullité sur la fonction de covariance expérimentale ou par des méthodes paramétriques par exemple). D'autre part, ce théorème assure une taille de bloc idéale dans le cas de la moyenne d'échantillon, mais cette taille de bloc idéale sera inconnue lorsque les statistiques d'intérêt seront plus complexes, ou lorsque l'on cherchera à estimer une distribution plutôt qu'une simple variance sur la variable aléatoire $\hat{\Theta}$. La largeur de bloc à utiliser pour une taille d'échantillon n donnée reste donc dans la plupart des cas un paramètre à régler de façon quelque peu heuristique.

Le théorème de Lahiri dans le cas précis de la moyenne d'échantillon permet de tester la "qualité" de la convergence en distribution de l'estimateur bootstrap $\bar{x}^* - \hat{E}(\bar{x}^*)$ vers la variable aléatoire $\bar{x} - E(X)$ pour une valeur de n finie. Nous allons à présent vérifier ces résultats théoriques sur un exemple concret en considérant un échantillon x de taille $n = 1000$ issue d'une fonction aléatoire $\mathcal{F}_d$, générée avec pour grain primaire un segment de taille $s = 15$ pixels de niveau de gris $g = 1$ constant et avec un processus de Poisson de densité $\rho = 0.1$.

La formule 3.8 permet de calculer :

$$\begin{aligned} \theta &= E(X) = 1.5 \\ \kappa &= \rho.s^2 = 22.5 \\ D &= \frac{4}{3}.\rho^2.s^4 = 675 \end{aligned}$$

$$G = \rho \cdot \frac{s^3}{3} = 112.5$$

$$l_{id} = \left\lceil \left(\frac{2 \cdot G^2}{D} \right)^{1/3} \cdot n^{1/3} \right\rceil = 34$$

La taille de bloc est ici choisie de manière idéale pour l'estimation de la variance d'extension. Pour tester les résultats du théorème de Lahiri (calcul de la moyenne et la variance de σ_b^{2*} ainsi que la concordance entre les histogrammes de $\bar{x}^* - \hat{E}(\bar{x}^*)$ et de $\bar{x} - E(X)$), nous allons procéder de la façon suivante :

- On génère 100 réalisations indépendantes d'échantillons de taille $n = 1000$ issue de $\mathcal{F}_d$.
- Pour chacun d'entre eux, on applique la méthode du bootstrap par bloc en générant 500 échantillons bootstrap qui nous permettent d'établir un histogramme h^k ($k \in [1..100]$) de $\Theta^* - \hat{E}(\Theta^*)$ (voir figure 3.6(b)). On calcule ensuite la distance d^k du χ^2 entre h^k et l'histogramme théorique⁴ de $\bar{x} - E(X)$.
On calcule également σ_b^{2*} pour chaque réalisation de $\mathcal{F}_d$ afin de comparer $E(\sigma_b^{2*})$ et $Var(\sigma_b^{2*})$ avec les valeurs asymptotiques fournies par le théorème de Lahiri.
- Si la distribution de $\Theta^* - \hat{E}(\Theta^*)$ est équivalente à celle de $\bar{x} - E(X)$ (ce qui validerait l'utilisation du bootstrap dans ce cas précis de distribution, de taille d'échantillon et de taille de bloc), la distribution des d^k doit être proche d'une loi du χ^2 dont le nombre de degrés de libertés dépend du nombre de classe choisie pour h^k . La dernière étape du test consiste donc à faire un histogramme des d^k pour $k \in [1..100]$, et à le comparer à la distribution du χ^2 adéquate (voir figure 3.6(a)). Dans le cas qui nous occupe, le nombre de classe pour h^k est 12, et la distribution théorique pour les d^k est donc un χ^2 à 11 degrés de liberté.
- On évalue également $E(\sigma_b^{2*})$, $Var(\sigma_b^{2*})$ et $E((\sigma_b^{2*} - \kappa)^2)$ à l'aide des 100 valeurs indépendantes de σ_b^{2*} (voir résultats table 3.1).

Les résultats obtenus sont en bon accord avec le théorème de Lahiri (qui, rappelons-le, fournit des résultats asymptotiques) et suggèrent quelques commentaires. Tout d'abord, la précision obtenue sur la variance d'extension est de l'ordre de 30% en moyenne, ce qui est moins bon que par la méthode d'estimation directe utilisée à la section 3.2.2c, mais le nombre n d'individu considérés est ici beaucoup plus faible (profils de 1000 points contre une image de 512×512 pixels), et les résultats concernant la précision de l'estimation ne sont donc pas directement comparable. D'autre part, l'approche par bootstrap permet également d'estimer les distributions de $\hat{\Theta} - \theta$ (et pas seulement la variance), et la figure 3.6(b) montre la compatibilité de ces estimations avec la distribution théorique.

b. Bootstrap de la fonction de covariance

Dans ce paragraphe et les suivants, nous chercherons à vérifier la pertinence du bootstrap pour l'ensemble des descripteurs étudiés au chapitre 2 ne rentrant pas dans le cadre

⁴qui s'obtient par simple changement de variable à partir de la loi donnant le nombre de grains primaires implantés sur un profil de taille n qui n'est autre qu'une loi de poisson de paramètre $\rho \cdot n$

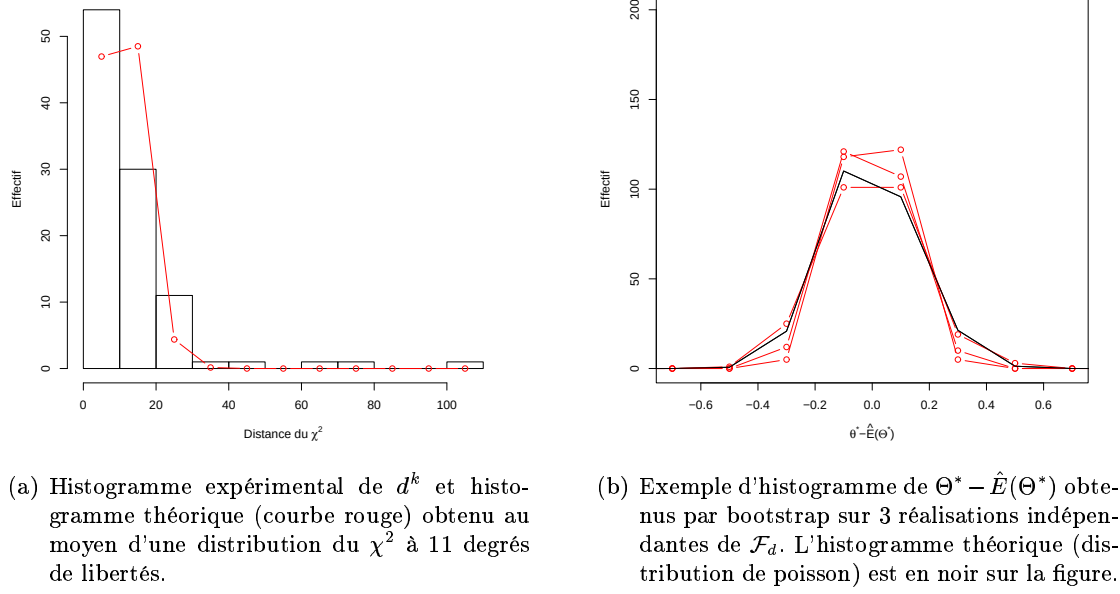


FIG. 3.6.: Estimation par bootstrap de la distribution d'une moyenne d'échantillon

Résultats Expérimentaux	Résultats théoriques (Lahiri)
$E(\sigma_b^{2*}) - \kappa \simeq -4.06$	$E(\sigma_b^{2*}) - \kappa = -3.31$
$Var(\sigma_b^{2*}) \simeq 25.61$	$Var(\sigma_b^{2*}) = 22.95$
$\sqrt{E((\sigma_b^{2*} - \kappa)^2)} \simeq 6.47 \simeq 0.29.\kappa$	$\sqrt{E((\sigma_b^{2*} - \kappa)^2)} \simeq 5.82 \simeq 0.26.\kappa$

TAB. 3.1.: Evaluation de la variance d'extension par bootstrap

du bootstrap classique. Pour ce faire, nous procéderons comme dans l'exemple du paragraphe précédent en utilisant des fonctions de dilutions bidimensionnelles de densité de Poisson égale à 0.04 et dont les grains primaires sont des disques de diamètre 50 pixels. Pour évaluer la distribution théorique de $\hat{\Theta} - \theta$ d'un paramètre d'intérêt observé sur un champ de taille 512×512 , on s'appuiera sur 200 réalisations indépendantes de fonctions de dilutions obtenues avec les mêmes paramètres, car le calcul analytique de cette distribution est souvent impraticable. On choisira ensuite une réalisation parmi ces 200 pour lui appliquer des techniques de bootstrap par bloc et comparer la distribution de $\Theta^* - \hat{E}(\Theta^*)$ obtenue au moyen de B échantillons bootstrap avec la distribution théorique.

Dans le cas où l'on s'intéresse à un descripteur construit à partir de la fonction de covariance expérimentale d'une image dans une direction précise, par exemple la direction X , le paramètre d'intérêt θ est la covariance théorique de la fonction de dilution considérée tandis que le paramètre mesuré est de la forme $\hat{\theta} = (\hat{C}(h))_{h \in [0..h_m]}$, $\hat{C}(h)$ étant la covariance *mesurée* sur l'image (voir 2.2.1). On peut considérer ensuite les covariances mesurées sur chaque colonne de l'image : $\hat{C}_y(h)$ pour $y \in [1..512]$, et on obtient donc $\hat{C}(h) = E_Y(\hat{C}_y(h))$. Nous considérerons ainsi le descripteur final comme la moyenne des fonctions de covariance obtenue pour chaque profil correspondant à une valeur de y , et nous allons appliquer les méthodes de bootstrap par bloc à la fonction aléatoire mono-dimensionnelle qui à une ordonnée y associe un vecteur aléatoire correspondant à la covariance C_y . Dans ce cas, les blocs considérés sont donc constitués d'un ensemble de fonctions de covariance mesurées sur des colonnes adjacentes.

Pour simplifier la représentation des résultats, nous avons représenté par un histogramme (figure 3.9(a)) la distribution des carrés des *distances euclidiennes* séparant la covariance théorique et la covariance mesurée, c'est à dire à la distribution de $\|\hat{\Theta} - \theta\|^2$. La distribution bootstrapée équivalente est représentée en rouge sur la figure 3.9(a). Le bon accord entre ces courbes permet de valider la méthode employée pour estimer la variation de la fonction de covariance au sein d'une population d'images.

c. Bootstrap de l'histogramme d'une image

Dans cette section, nous abordons la notion de bootstrap par bloc dans le cas où la fonction aléatoire considérée n'est plus un profil mais une image. Dans la situation la plus classique, le paramètre d'intérêt θ est la moyenne ou l'histogramme d'une image. Des descripteurs rentrant dans cette catégorie sont par exemple l'histogramme des pentes ou l'histogramme de l'orientation des vecteurs normaux, ou plus simplement l'histogramme des niveaux de gris observés sur une image. Lahiri consacre tout un chapitre à ce cas particulier du bootstrap par bloc en prouvant la cohérence de ces méthodes si les blocs sont choisis comme des sous-images de taille réduite à l'intérieur d'une image plus large. Le principe et les résultats asymptotiques sont voisins du cas mono-dimensionnel : on tire au hasard et avec remise des blocs (bidimensionnels) de taille l de façon à construire une image bootstrap dont on peut extraire un paramètre θ^* .

Nous présentons ici un exemple dans le cas où $\hat{\theta}$ est l'histogramme du niveau de gris d'une fonction de dilution. La distribution théorique, et donc θ , est connue dans le cas des fonctions de dilution : il s'agit d'une loi de Poisson de paramètre $\rho.E(S)$ où ρ est

la densité de Poisson utilisée (ici 0.04) et $E(S)$ est l'espérance de la surface du grain primaire (ici $\pi.25^2$), et comme dans le paragraphe précédent, on s'intéresse à la norme de $\|\hat{\Theta} - \theta\|^2$. Les résultats obtenus sont présentés à la figure 3.9(b), et là encore, on observe un accord correct entre distribution réelle et bootstrappée.

d. Bootstrap de l'histogramme des tortuosités

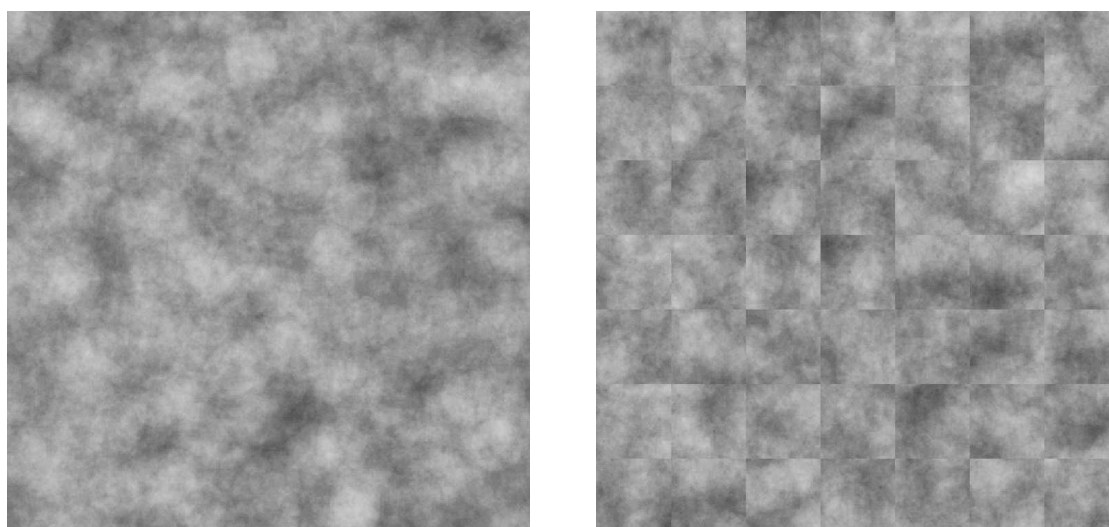
L'histogramme des tortuosités (voir section 2.2.4) pour une image constitue un exemple où deux méthodes de bootstrap par bloc peuvent raisonnablement être envisagées, et où l'une des deux échoue à cause d'une dépendance à trop grande échelle.

La première façon d'aborder ce descripteur consiste à considérer les valeurs des tortuosités sur l'ensemble d'arrivée comme une réalisation d'une fonction aléatoire dont on cherche l'histogramme. Il suffit dans ce cas d'appliquer la méthode monodimensionnelle standard de bootstrap par bloc décrite au paragraphe 3.3.3a et de calculer l'histogramme θ^* pour chacun des échantillons bootstrap. Malheureusement, la distribution de $\|\Theta^* - \hat{E}(\Theta^*)\|^2$ ainsi obtenue et la distribution de $\|\hat{\Theta} - \theta\|^2$ estimée à partir de 200 réalisations sont assez différentes : les échantillons bootstrap se "ressemblent" beaucoup plus que les réalisations réelles d'une fonction de dilution (voir figure 3.9(c)). La raison principale de cet échec est la très forte portée de la fonction aléatoire considérée. La propagation d'une distance à l'intérieur de l'image induit des dépendances entre des points de l'ensemble d'arrivée qui peuvent être assez éloignés les uns des autres, à ceci peut s'ajouter des effets des bords de l'image qui peuvent perturber les valeurs des distances obtenues pour certains points. Le profil des tortuosités observé sur l'ensemble d'arrivée d'une image n'est donc pas suffisamment représentatif pour rendre compte de la véritable variabilité de ce descripteur.

Ce premier échec suggère d'aborder le bootstrap de façon plus "bidimensionnelle", et de revenir à sa forme première : bootstrapper une image équivalente à l'image observée et calculer un paramètre d'intérêt θ^* correspondant à l'histogramme des tortuosités sur cette nouvelle image. Cette approche à l'inconvénient d'être beaucoup plus coûteuse en temps de calcul, mais l'avantage d'être plus précise dans le cas qui nous intéresse. Les figures 3.7(a) et 3.7(b) montrent l'équivalent bootstrap d'une réalisation d'une fonction de dilution pour des blocs de taille 80.

Les figures 3.8(a) et 3.8(b) montrent respectivement des courbes de niveaux de la fonction distance propagée à l'intérieur d'une réalisation d'une fonction de dilution d'une part, et à l'intérieur de son équivalent bootstrappé d'autre part. Même si certains artefacts dus à la découpe en blocs sont visibles, les courbes de niveaux ont des allures semblables : elles présentent en particulier approximativement les mêmes échelles de variations, ce qui est de bon augure pour la cohérence de la distribution bootstrap. La distribution de $\|\Theta^* - \hat{E}(\Theta^*)\|^2$ dans cette configuration est représentée par la courbe rouge de la figure 3.9(c) et se montre en bon accord avec l'histogramme représenté en noir correspondant à $\|\hat{\Theta} - \theta\|^2$.

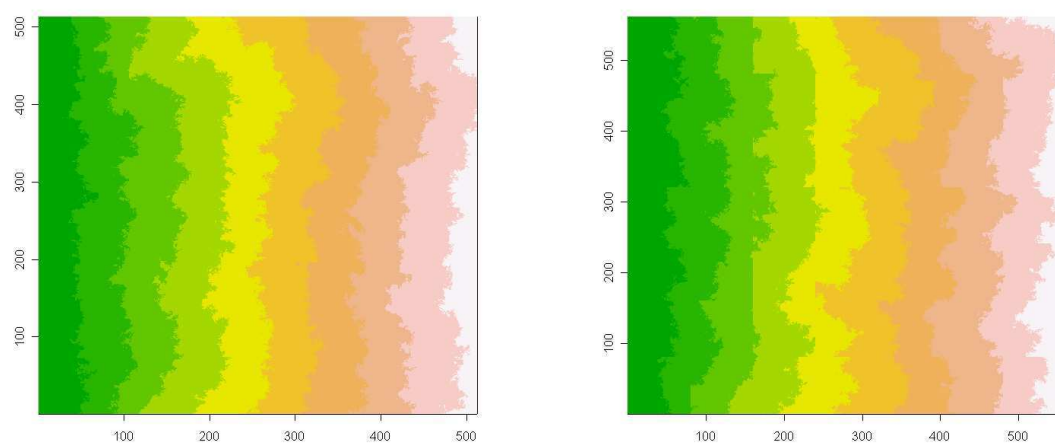
Ce cas de figure intéressant soulève principalement deux réactions. La première concerne la pertinence de l'histogramme des tortuosités en tant que descripteur d'une surface : il est important de garder à l'esprit que l'histogramme des tortuosités issu d'une seule



(a) Fonction de dilution originale .

(b) Equivalent bootstrap .

FIG. 3.7.: Bootstrap d'une fonction aléatoire de dilution



(a) Courbes de niveau de la fonction distance .

(b) Equivalent bootstrap .

FIG. 3.8.: Bootstrap des tortuosités

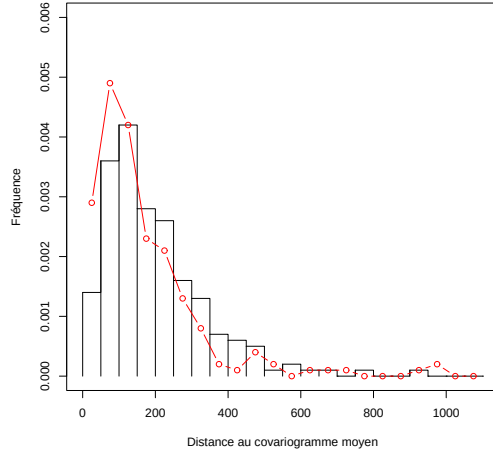
réalisation d'une fonction aléatoire peut ne pas permettre d'appréhender la variabilité réelle possible des valeurs des tortuosités. Il est donc essentiel dans ce cas d'avoir recours au bootstrap (par exemple) pour tenter de reconstruire et de prendre en compte ces variations. D'autre part, cet exemple montre que les méthodes de sur-échantillonnage doivent être manipulées avec beaucoup de soin pour ne pas fournir des résultats erronés. L'utilisation abusive de la première méthode de bootstrap présentée conduirait par exemple à une très forte sous-estimation de la variance d'extension, ce qui peut avoir ensuite des conséquences importantes en terme d'inspection de surface.

e. Bootstrap des courbes d'érosion-dilatation

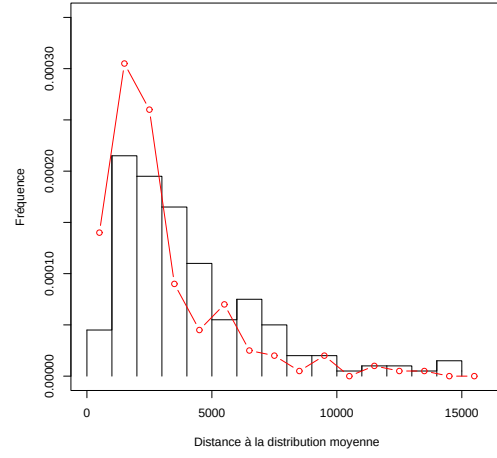
Le principe de bootstrap pour les courbes d'érosion-dilatation est très proche de la version bidimensionnelle du paragraphe précédent : on construit des images bootstrap en tirant au hasard et avec remise des sous-images de taille l à l'intérieur de l'image originale (voir figure 3.7), puis pour chacune de ces images bootstrap, on recalcule une courbe d'érosion-dilatation qui correspond à θ^* . On compare ensuite la distribution de $\Theta^* - \hat{E}(\Theta^*)$ avec $\hat{\Theta} - E(\Theta)$. Comme pour les autres descripteurs vectoriels, nous représentons à la figure 3.9(d) les résultats concernant la distribution de $\|\Theta^* - \hat{E}(\Theta^*)\|^2$. Là encore, le bon accord entre la distribution réelle obtenue par simulation de 200 réalisations indépendantes de fonctions de dilution et la distribution obtenue à partir de 50 échantillons bootstrap provenant de l'une de ces réalisations montre la pertinence et l'intérêt de la méthode dans le cas où le descripteur considéré est une courbe d'érosion-dilatation.

3.4. Conclusion

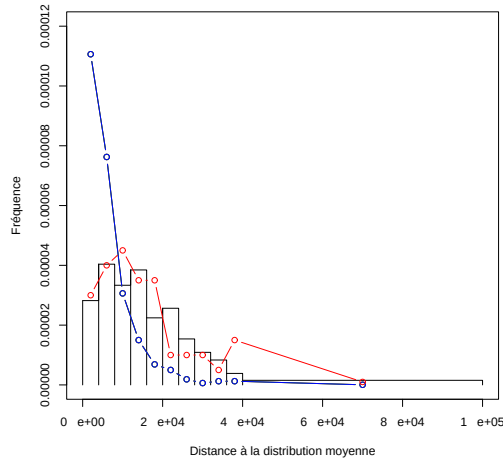
Quels que soient les descripteurs utilisés pour rendre compte des propriétés des surfaces, il est indispensable de s'intéresser à l'erreur qui peut être commise en attribuant à une surface la valeur mesurée sur un échantillon. En termes statistiques, une surface est représentée par une *fonction aléatoire*, un échantillon observé par une *réalisation* de taille finie issu de cette fonction, les descripteurs par un *paramètre* θ caractérisant la fonction aléatoire et la valeur mesurée par une *statistique* $\hat{\theta}$ extraite de la réalisation observée. Le formalisme géostatistique permet d'introduire la notion de variance d'extension, c'est à dire la variance de $\hat{\theta}$ comme estimateur de θ , et ceci dans un contexte continu ou discret. Cependant, l'application de ce formalisme limite fortement l'ensemble des paramètres d'intérêt utilisable à moins de mettre en oeuvre des calculs et transformations assez complexes. Une alternative à cette théorie consiste à utiliser des méthodes de sur-échantillonnage de type bootstrap restreignant le cadre d'application à un formalisme discret, mais élargissant considérablement le champ des descripteurs utilisables. Une étude détaillée nous a permis de vérifier la pertinence de ces méthodes (validité des résultats asymptotiques notamment) pour chacun des descripteurs de surface décrits au chapitre 2, cependant, une analyse mathématique plus poussée permettrait certainement de confirmer la convergence (et éventuellement la vitesse de convergence) des distributions bootstrap vers les distributions réelles pour des descripteurs morphologiques



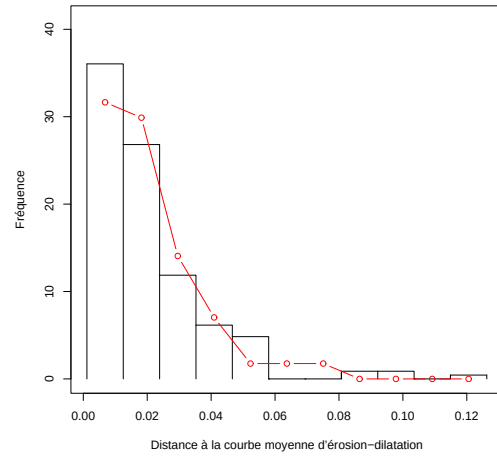
(a) Cas de la fonction de covariance



(b) Cas de l'histogramme d'une image



(c) Cas de l'histogramme des tortuosités. La courbe rouge représente la distribution de $\|\hat{\Theta}^* - \hat{E}(\Theta^*)\|^2$ obtenue en bootstrappant des images. La courbe bleue correspond à la distribution obtenue par bootstrap sur le profil de tortuosité de l'ensemble d'arrivée



(d) Cas des courbes d'érosion-dilatation

FIG. 3.9.: Comparaison des distributions théoriques et bootstrappées pour différents descripteurs. L'histogramme noir représente la distribution de $\|\hat{\Theta} - \theta\|^2$, tandis que la courbe rouge représente la distribution de $\|\hat{\Theta}^* - \hat{E}(\Theta^*)\|^2$

Chapitre 3. Stabilité de la description

(segmentation, érosion-dilatation, tortuosité) suffisamment particulier pour ne pas rentrer trivialement dans le cadre des hypothèses des théorèmes de convergence existant dans la littérature. Malgré ce point négatif, les fonctions aléatoires considérées pour la représentation de surfaces rugueuses ont en générale de "bonnes" propriétés (stationnarité, ergodicité, portée et variance finie) plutôt plus forte que les hypothèses requises pour le bon fonctionnement du bootstrap par bloc, et ceci associé à des tests expérimentaux concluants peut permettre d'espérer de bons comportements théoriques en terme de convergence.

Finalement, l'ensemble des techniques présentées dans ce chapitre nous permet d'associer à chaque échantillon étudié, et pour chacun des descripteurs choisis, non pas une seule valeur, mais une bonne approximation de la *densité de probabilité* des valeurs que l'on pourrait obtenir en observant un grand nombre d'échantillons issus de la même surface.

Chapitre 4.

Méthode d'analyse

Dans les chapitres précédents nous avons progressé depuis la représentation d'une surface (topographie, image, profil) vers un nombre important de descripteurs auxquels les méthodes de bootstrap nous ont permis d'associer une distribution d'erreur. La suite logique de notre démarche consiste à traiter l'ensemble de ces informations pour aboutir à une classification des surfaces en plusieurs catégories, qu'il sera possible d'adapter aux besoins industriels de contrôle qualité. L'objet de ce chapitre est la présentation des outils mathématiques nécessaires à la mise au point d'une telle méthodologie d'analyse des surfaces.

Le premier impératif dont nous devons nous acquitter est la définition d'un espace de représentation pour lequel la relation entre la position d'une surface et une variable qualitative (classification) ou quantitative (régression) est plus facile à établir que dans l'espace original des descripteurs. L'une des caractéristiques clés de cet espace de travail doit être sa faible dimension (une surface doit pouvoir être représentée par un petit nombre de variables) afin de garantir stabilité et rapidité des opérations ultérieures de classification [52]. Le choix de l'analyse factorielle[9] pour construire un tel espace nous a permis d'établir une continuité avec certains travaux déjà menés à Arcelor Research [1, 3, 4] dont les résultats encourageaient à poursuivre dans cette voie. Ces méthodes présentent aussi l'avantage d'être facilement interprétables car linéaires, peu paramétriques, relativement robustes, et surtout, d'être transposables telles qu'elles à une grandes gammes de produits et descripteurs différents, ce qui ne serait pas nécessairement le cas de méthodes plus sophistiquées faisant intervenir des comportements non linéaire par exemple. Nous n'avons certainement pas l'ambition de décrire ici tous les détails théoriques de l'analyse factorielle (voir par exemple pour cela [9]), mais plutôt d'en présenter brièvement les caractéristiques essentielles et les quelques originalités que nous avons pu appliquer dans le cadre de l'étude des surfaces rugueuses. Dans cette optique, certaines notations sont empruntées au cours d'analyse des données destinés aux élèves de l'école des Mines [50].

Après que l'espace de travail a été choisi, nous nous intéressons aux opérations de classification proprement dites, c'est à dire à l'attribution d'un label à une surface en fonction de sa position dans cet espace. De nombreuses méthodes ont été développées pour résoudre ce problème souvent épineux, mais dans le cas qui nous préoccupe, le schéma de classification bayésien est l'un de ceux dont le cadre probabiliste permet le plus facilement la valorisation des données bootstrap que nous considérons comme essentielles, et c'est donc sur cette base que nous appuierons la deuxième étape de notre méthodologie.

Si ce chapitre explicite les fondements théoriques qui nous permettent de franchir le

fossé séparant la représentation des surfaces dans un espace quelconque de leur classification en quelques catégories, il sera avare d'illustrations et d'exemples concrets. L'ensemble des résultats que nous avons pu obtenir est en effet recensé au chapitre 5 auquel on pourra donc se reporter pour une vision plus intuitive, graphique et physique des espaces de projection, nuages d'individus bootstrap et autres variables composites auxquels les prochaines sections feront référence.

4.1. Analyse de données

4.1.1. Généralités sur l'analyse factorielle

a. Notations

On considère l'espace vectoriel $\mathbb{R}^p$ muni du produit scalaire canonique (noté $(\cdot|\cdot)$). $\mathbb{R}^p$ sera considéré dans la suite comme un espace affine euclidien pour lequel on notera $d(x, y)$ la distance entre deux points x et y .

Le tableau de données se présente sous la forme d'une matrice :

$$X_{(n,p)} = \begin{pmatrix} x_1^1 & \dots & x_1^j & \dots & x_1^p \\ \vdots & & \vdots & & \vdots \\ x_i^1 & \dots & x_i^j & \dots & x_i^p \\ \vdots & & \vdots & & \vdots \\ x_n^1 & \dots & x_n^j & \dots & x_n^p \end{pmatrix}$$

Les vecteurs lignes x_i de X sont des points de $\mathbb{R}^p$ que l'on munit de poids $p_1, p_2, \dots, p_n$ tels que :

$$\forall i \in [1..n], p_i \geq 0, \text{ avec } \sum_{i=1}^n p_i = 1$$

On appelle nuage de points (noté N) l'ensemble des couples (x_i, p_i) . On note g le barycentre du nuage, c'est à dire le point tel que :

$$g = \sum_{i=1}^n p_i x_i$$

Soit H un sous-espace affine de $\mathbb{R}^p$, on note alors H_0 le sous-espace vectoriel associé à H , c'est à dire le sous-espace parallèle à H passant par l'origine O de $\mathbb{R}^p$, et H_g le sous-espace affine parallèle à H passant par g .

On note $\hat{x}_i = P_H(x_i)$ la projection orthogonale de x_i sur H et $\hat{g}$ la projection de g sur H .

Le moment d'inertie du nuage par rapport à H sera noté $M^t(H)$:

$$M^t(H) = \sum_{i=1}^n p_i d^2(x_i, \hat{x}_i)$$

On rappelle la formule de Huyghens :

$$M^t(H) = M^t(H_g) + d^2(H, H_g)$$

b. Critères de choix d'un sous-espace

Notre objectif étant de représenter le nuage N en le projetant dans un sous-espace H de $\mathbb{R}^p$ de dimension $q < p$ fixée, il est nécessaire de se donner un critère pour lequel H sera choisi de manière "optimale". Nous fournissons ici 3 critères équivalents :

- Critère de proximité : on cherche le sous-espace H de dimension q le plus "proche des données", c'est à dire minimisant le moment d'inertie du nuage :

$$H = \operatorname{argmin}(M^t(H))$$

La formule de Huyghens impose alors à H de passer par le centre de gravité du nuage.

- Critère de fidélité des distances : on cherche le sous-espace H contenant g pour lequel les distances entre les projetés des points du nuage sont les plus proches possible des distances originales. Sachant que $d(\hat{x}_i, \hat{x}_j) \leq d(x_i, x_j)$, on cherche le sous-espace H contenant g maximisant le carré des distances pondérées entre les projetés des points du nuage :

$$H = \operatorname{argmax}(Q = \sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} d^2(\hat{x}_i, \hat{x}_{i'}))$$

L'équivalence avec le premier critère provient du développement suivant :

$$\begin{aligned} Q &= \sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} d^2(\hat{x}_i, \hat{x}_{i'}) \\ Q &= \sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} (d^2(\hat{x}_i, \hat{g}) + d^2(\hat{x}_{i'}, \hat{g}) - 2 \cdot (\hat{x}_i - \hat{g} | \hat{x}_{i'} - \hat{g})) \\ Q &= 2 \cdot \sum_{i=1}^n p_i d^2(\hat{x}_i, \hat{g}) - 2 \cdot \underbrace{\left(\sum_{i=1}^n p_i \hat{x}_i \right)}_{\hat{g}} \cdot \underbrace{\left(\sum_{i=1}^n p_i \hat{x}_i - \hat{g} \right)}_{\hat{g}} \end{aligned} \quad (4.1)$$

$$Q = 2 \cdot \sum_{i=1}^n p_i (d^2(x_i, \hat{g}) - d^2(\hat{x}_i, x_i)) \quad (4.2)$$

$$Q = 2 \cdot (I_{\hat{g}} - M^t(H)) = 2 \cdot (I_g + d^2(g, \hat{g}) - M^t(H)) \quad (4.3)$$

$$Q = 2 \cdot (I_g - M^t(H_g)) \quad (4.4)$$

Où $I_g = \sum_{i=1}^n p_i d^2(x_i, g)$ désigne l'inertie totale du nuage par rapport à son centre de gravité.

Le passage de 4.1 à 4.2 n'est rien d'autre que le théorème de Pythagore tandis que le passage de 4.3 à 4.4 fait intervenir la relation de Huyghens, et le fait d'imposer $g \in H$ pour ce critère complète l'équivalence entre la maximisation de Q et la minimisation de $M^t(H)$.

- Maximisation de l'inertie du nuage projeté : si l'on appelle $I(H) = \sum_{i=1}^n p_i d^2(\hat{x}_i, g)$ l'inertie "expliquée" par H , la relation 4.1 montre que $g \in H \Rightarrow Q = 2I(H)$. Les deux premiers critères sont donc équivalents à chercher le sous-espace affine H de dimension q contenant g et maximisant $I(H)$.

C'est ce critère qui sera utilisé en pratique pour calculer explicitement H .

c. Expression matricielle

Dans la suite, on supposera que le nuage est centré, c'est à dire que le centre de gravité g coïncide avec l'origine O de $\mathbb{R}^p$, l'expression de $I(H)$ devient donc simplement $\sum_{i=1}^n p_i \|\hat{x}_i\|^2$.

Soit Δu une droite passant par l'origine engendrée par le vecteur unitaire u . On peut exprimer la projection sur Δu d'un point du nuage par :

$$\hat{x}_i = (x_i | u) u = (x_i^t u) u \quad (4.5)$$

L'inertie expliquée par Δu peut donc s'obtenir de la manière suivante :

$$\begin{aligned} I(\Delta u) &= \sum_{i=1}^n p_i \|\hat{x}_i\|^2 = \sum_{i=1}^n p_i (x_i^t u)^2 \\ &= \sum_{i=1}^n p_i (u^t x_i) (x_i^t u) \\ &= u^t \left(\sum_{i=1}^n p_i x_i x_i^t \right) u \\ &= u^t X^t D X u \\ &= u^t V u \end{aligned} \quad (4.6)$$

V est la matrice d'inertie du nuage définie par :

$$V = \sum_{i=1}^n p_i x_i x_i^t = X^t D X$$

D représente la matrice diagonale des poids des individus :

$$D = \text{diag}(p_1, \dots, p_n)$$

Si H_1 et H_2 sont deux sous-espaces orthogonaux contenant l'origine, le théorème de Pythagore permet d'obtenir la relation de composition de l'inertie :

$$I(H_1 \oplus H_2) = I(H_1) + I(H_2)$$

d. Principe de calcul

Soit $k \leq p$; si F_k est le sous-espace maximisant l'inertie expliquée parmi l'ensemble des sous-espaces de dimension k , alors le sous-espace vectoriel de dimension $k + 1$ d'inertie expliquée maximale est $F_{k+1} = F_k \oplus \Delta u$, où Δu est la droite vectorielle orthogonale à F_k d'inertie expliquée maximale.

En conséquence, on recherchera F_k de manière récursive en calculant d'abord $F_1 = \Delta u_1$. Pour cela on cherche u_1 tel que :

$$I(\Delta u_1) = \sup_{u \in \mathbb{R}^p} \{I(\Delta u)\}$$

On cherche ensuite F_2 sous la forme $\Delta u_1 \oplus \Delta u_2$ avec u_2 vérifiant :

$$I(\Delta u_2) = \sup_{u \in \mathbb{R}^p, u \perp u_1} \{I(\Delta u)\}$$

Et ainsi de suite jusqu'à obtenir F_k

Les droites $\Delta u_1, \dots, \Delta u_p$ sont appelées 1er, ..., p-ème axe principal d'inertie du nuage.

D'après (4.6), $I(\Delta u) = u^t V u$. V étant une matrice réelle symétrique (et positive), elle est diagonalisable dans une base orthonormée $(u_1, u_2, \dots, u_p)$ de vecteurs propres de V :

$$\forall k = 1..p, V u_k = \lambda_k u_k \quad (\lambda_k \in \mathbb{R})$$

On numérote les u_k de façon à ce que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$.

Soit u un vecteur unitaire que l'on décompose sur la base des vecteurs propres : $u = \sum_i \alpha_i u_i$. L'inertie du nuage projeté sur Δu s'exprime alors au moyen de la relation 4.6 :

$$\begin{aligned} I(\Delta u) &= u^t V u \\ &= (u^t | \sum_i \alpha_i \lambda_i u_i) \\ &= (\sum_i \alpha_i u_i | \sum_i \alpha_i \lambda_i u_i) \\ &= \sum_i \alpha_i^2 \lambda_i \\ &\leq \lambda_1 \sum_i \alpha_i^2 \\ &\leq \lambda_1 \end{aligned}$$

La quatrième ligne vient de l'orthogonalité de la base des vecteurs propres, la cinquième du choix de l'ordre des vecteurs propres ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$), et enfin la sixième s'explique par le fait que u est unitaire. On obtient donc :

$$\forall u \in \mathbb{R}^p, I(\Delta u) \leq \lambda_1$$

Or $I(\Delta u_1) = \lambda_1$, et $I(\Delta u)$ est donc maximale pour $u = u_1$. On obtient par un raisonnement similaire :

$$F_k = \Delta u_1 \oplus \Delta u_2 \oplus \dots \oplus \Delta u_k \quad \text{et} \quad I(F_k) = \lambda_1 + \lambda_2 + \dots + \lambda_k$$

On appelle part d'inertie expliquée par F_k la quantité :

$$\tau_{F_k} = \frac{I(F_k)}{I_g}$$

e. Facteurs principaux

On appelle α -ème facteur principal le vecteur c^α de $\mathbb{R}^n$ assimilable à un point de l'espace des variables dont les composantes sont les coordonnées c_i^α des points x_i de $\mathbb{R}^p$ sur le α -ème axe principal Δu_α .

Matriciellement, on obtient la relation :

$$c^\alpha = X \cdot u_\alpha$$

En pratique, on peut considérer c^α comme une nouvelle variable obtenue par une combinaison linéaire des variables originales dont les coefficients sont les coordonnées du vecteur propre u_α de la matrice d'inertie.

Il existe deux indicateurs facilitant l'analyse du nuage N projeté sur un sous-espace F_k .

Tout d'abord la contribution du points x_i à l'inertie de l'axe Δ_α s'exprimant :

$$Ctr_\alpha(i) = \frac{p_i(c_i^\alpha)^2}{\lambda_\alpha}$$

Ces quantités représentent la décomposition point par point de l'inertie d'un axe et permettent de déterminer quels sont les individus originaux qui jouent le plus grand rôle dans l'obtention des premiers axes factoriels.

Le cosinus, ou plus exactement le carré du cosinus de l'angle entre un axe propre Δ_α et le point représentatif de x_i permet de juger de la qualité de la représentation du point x_i par l'axe Δ_α .

$$Cos_\alpha^2(i) = \frac{(c_i^\alpha)^2}{\|x_i\|^2} = \cos^2 \theta_\alpha(i)$$

Ces quantités représentent la décomposition axe par axe du carré de la distance d'un point au centre de gravité, et permettent d'en interpréter correctement la position sur les axes principaux Δ_α .

4.1.2. Séparation d'ensembles

Dans la section précédente, nous avons déterminé un sous-espace H de dimension fixé maximisant l'inertie des projetés des points du nuage sur H (second critère). Il peut arriver des situations où l'on cherche à trouver un espace séparant deux ou plusieurs *ensembles* (ou classes) de points (par exemple des surfaces correctes et des surfaces incorrectes). Dans ce cas, les distances que l'on cherche à maximiser sont les distances interclasses, tandis que les distances entre deux points d'une même classe n'ont que peu d'intérêt si l'objectif final est l'attribution à une surface inconnue d'un label correspondant à l'une des classes d'intérêt. Nous allons donc restreindre la sommation des distances

dans le critère Q pour ne prendre en compte que les distances entre deux points de classes différentes.

Nous introduisons les notations suivantes :

- C_k est l'ensemble des indices des points de la classe C_k .
- $k(i)$ est la classe à laquelle appartient le point i .
- $P_k = \sum_{i \in C_k} p_i$ est le poids global de la classe k .
- $g_k = \frac{1}{P_k} \sum_{i \in C_k} p_i x_i$ est le barycentre de la classe k .
- $P^k = 1 - P_k$ est le poids global des points n'appartenant pas à la classe k .
- $g^k = \frac{1}{P^k} \sum_{i \notin C_k} p_i x_i$ est le barycentre des points n'appartenant pas à la classe C^k .
- q_i^k est une pondération attribuée au point x_i défini par :

$$\begin{aligned} i \in C_k &\Rightarrow q_i^k = P^k \cdot p_i \\ i \notin C_k &\Rightarrow q_i^k = P_k \cdot p_i \end{aligned}$$

- On introduit le nouveau critère de fidélité des distances :

$$Q' = \sum_{k=1}^K \sum_{i \in C_k} \sum_{j \notin C_k} p_i p_j d^2(\hat{x}_i, \hat{x}_j)$$

Nous montrons qu'on peut trouver le sous-espace H de dimension k maximisant le critère Q' par la méthode du paragraphe précédent en remplaçant la matrice d'inertie V par une matrice V' définie par :

$$V' = \sum_{i=1}^n 2 \cdot p_i \cdot P^{k(i)} x_i x_i^t + 2 \cdot \sum_{k=1}^K P_k^2 g_k \cdot g_k^t \quad (4.7)$$

Preuve. On supposera que le centre de gravité g du nuage coïncide avec l'origine.

Comme dans le paragraphe c, on cherche à exprimer le critère Q' lorsque le sous-espace H est une droite Δu par une relation du type :

$$Q'(\Delta u) = u^t V' u$$

Où V' est une matrice symétrique réelle.

Une telle expression permet d'utiliser la même construction itérative que dans le paragraphe d pour obtenir F_k .

Partons de l'expression de Q' :

$$\begin{aligned} Q' &= \sum_{k=1}^K \sum_{i \in C_k} \sum_{j \notin C_k} p_i p_j d^2(\hat{x}_i, \hat{x}_j) \\ &= \sum_{k=1}^K \sum_{i \in C_k} \sum_{j \notin C_k} p_i p_j (\|\hat{x}_i\|^2 + \|\hat{x}_j\|^2 - 2 \cdot (\hat{x}_i | \hat{x}_j)) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^K \sum_{i \in C_k} p_i \sum_{j \notin C_k} p_j \|\hat{x}_j\|^2 + \sum_{j \notin C_k} p_j \sum_{i \in C_k} p_i \|\hat{x}_i\|^2 \\
&\quad - 2. \sum_{k=1}^K \left(\sum_{i \in C_k} p_i \hat{x}_i \mid \sum_{j \notin C_k} p_j \hat{x}_j \right) \\
&= \sum_{k=1}^K P_k \sum_{j \notin C_k} p_j \|\hat{x}_j\|^2 + P^k \sum_{i \in C_k} p_i \|\hat{x}_i\|^2 \\
&\quad - 2. \sum_{k=1}^K (P_k \cdot \hat{g}_k \mid P^k \cdot \hat{g}^k) \tag{4.8}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^K \sum_{l=1}^n q_l^k \|\hat{x}_l\|^2 - 2. \sum_{k=1}^K (P_k \cdot \hat{g}_k \mid P^k \cdot \hat{g}^k) \\
&= \sum_{l=1}^n \left(\sum_{k=1}^K q_l^k \right) \|\hat{x}_l\|^2 - 2. \sum_{k=1}^K (P_k \cdot \hat{g}_k \mid P^k \cdot \hat{g}^k) \tag{4.9}
\end{aligned}$$

Considérons maintenant l'associativité des barycentres qui fournit :

$$\begin{aligned}
P_k g_k + P^k g^k &= \sum_{i \in C_k} p_i x_i + \sum_{i \notin C_k} p_i x_i = g = 0 \\
g^k &= -\frac{P_k}{P^k} g_k \\
(P_k \cdot \hat{g}_k \mid P^k \cdot \hat{g}^k) &= -P_k^2 \cdot \|\hat{g}_k\|^2
\end{aligned}$$

Le dernier terme à expliciter reste la somme :

$$\begin{aligned}
\sum_{k=1}^K q_l^k &= q_l^{k(l)} + \sum_{k \in [1..K], k \neq k(l)} q_l^k \\
&= p_l \cdot P^{k(l)} + \sum_{k \in [1..K], k \neq k(l)} p_l \cdot P_k \\
&= p_l \cdot (P^{k(l)} + 1 - P_{k(l)}) \\
&= 2 \cdot p_l \cdot P^{k(l)}
\end{aligned}$$

En remplaçant ces résultats dans 4.9, on obtient :

$$Q' = \sum_{l=1}^n 2 \cdot p_l \cdot P^{k(l)} \|\hat{x}_l\|^2 + 2 \cdot \sum_{k=1}^K P_k^2 \cdot \|\hat{g}_k\|^2$$

Lorsque l'on projette sur une droite Δu , l'expression de $\hat{x}_i$ est donnée par 4.5, et le même type de calcul qu'en 4.6 permet d'obtenir $Q'(\Delta u)$ sous la forme :

$$Q'(\Delta u) = u^t \left(\sum_{i=1}^n 2 \cdot p_i \cdot P^{k(i)} x_i x_i^t + 2 \cdot \sum_{k=1}^K P_k^2 g_k \cdot g_k^t \right) u$$

$$Q'(\Delta u) = u^t V' u$$

V' est ici une matrice symétrique réelle, et les raisonnements du paragraphe d en termes de valeurs et vecteurs propres de V restent donc valable pour déterminer un sous-espace de dimension k maximisant le critère Q' . $\square$

Le critère Q' nous permet en quelque sorte de mettre l'accent sur la séparation de groupes de points sans nous intéresser à l'inertie interne de ces groupes, mais l'utilisation de Q' à la place de Q entraîne la perte de certaines propriétés de l'analyse factorielle classique comme l'équivalence avec les critères de proximité ou de maximum d'inertie ou encore la dualité entre points et variables (non présentée ici) qui peut conduire à des résultats intéressants en terme d'interprétation des axes propres.

4.1.3. Différents types d'analyse factorielle

Les techniques de réduction d'espace basées sur la détermination de facteurs principaux sont généralement qualifiée d'"analyse factorielle", mais il en existe plusieurs variantes suivant le type de données que l'on cherche à analyser. Il peut par exemple s'agir de cas où l'on veut utiliser une distance non-euclidienne entre les points. La méthode couramment utilisée consiste à se ramener au cas euclidien décrit dans le paragraphe précédent par une modification du tableau de données. Nous présentons ici plusieurs de ces variantes qui seront utiles dans l'exploitation de nos descripteurs de surfaces.

a. Changement de produit scalaire

Le cas "standard" du changement de distance consiste à considérer une nouvelle matrice de produit scalaire S pour le calcul des projections et des normes :

$$(x|y) = X^t S Y$$

Où S est une matrice symétrique définie positive. On peut alors la diagonaliser dans une base orthonormée, et se ramener à une base orthogonale dans laquelle la matrice du produit scalaire devient l'identité :

$$\begin{aligned} (u|v) &= U^t S V \\ &= U^t P^t R P V \\ &= U^t (\sqrt{R} P)^t \sqrt{R} P V \\ &= (\sqrt{R} P U)^t \sqrt{R} P V \end{aligned}$$

Où P est la matrice de changement de base diagonalisant S , et R la matrice diagonale des valeurs propres de S (ces valeurs propres sont positives, et on peut donc utiliser la notation $\sqrt{R}$ sans ambiguïtés). On peut donc aisément se ramener au cas où la distance est euclidienne en travaillant sur une nouvelle matrice de données X' , telle que :

$$X' = X.(\sqrt{R}P)^t$$

En effet, X est une matrice dont les lignes $x_1, \dots, x_p$ sont les vecteurs représentant la position des points dans l'espace des descripteurs :

$$X = \begin{pmatrix} x_1^t \\ \vdots \\ x_p^t \end{pmatrix}$$

Lors du changement de repère, on obtient $x'_i = \sqrt{R}P x_i$ soit $(x'_i)^t = x_i^t (\sqrt{R}P)^t$, et donc $X' = X.(\sqrt{R}P)^t$.

Le cas le plus fréquent est celui où un point x_i est défini par n variables quantitatives (que l'on supposera centrées pour faire coïncider O et g). Une distance euclidienne classique fait intervenir de la même façon un écart entre deux valeurs prises sur une variable de forte variance et sur une variable de faible variance : par exemple, si chaque point x est représenté par sa valeur sur deux variables dont la première est de variance 1 et la seconde de variance 100, un écart de 1 unité sur la première variable est probablement plus significatif qu'un écart de 10 unités sur la seconde variable, pourtant, une distance euclidienne classique sera insensible à cet effet. Pour compenser cette différence de variance, on utilise souvent des variables centrées et réduites, ce qui revient à effectuer le calcul précédent dans lequel la matrice R est la matrice diagonale inverse des variances, et la matrice P est l'identité : une telle analyse est appelée Analyse en Composantes Principales (ACP) normée. Un autre exemple classique consiste à utiliser une distance faisant intervenir les corrélations entre variables (distance de Mahalanobis).

b. Analyse factorielle des correspondances (AFC)

Les points que l'on compare ne sont pas toujours des individus auxquels correspondent des valeurs sur plusieurs variables (par exemple poids et taille) : on peut aussi s'intéresser à plusieurs groupes d'individus (ou populations) caractérisés par l'histogramme ou la distribution d'une même variable ν . L'une des distance utilisée pour comparer de tels histogrammes est la distance du χ^2 qui attribue à chaque classe ou modalité un poids inversement proportionnel à une fréquence théorique (qui sera ici la fréquence moyenne sur l'ensemble des populations étudiées). Une telle analyse est appelée "analyse factorielle des correspondances" (AFC).

Dans ce contexte, un point y_i est un vecteur de $\mathbb{R}^p$ représentant une distribution, ou un histogramme de fréquence (c'est à dire normalisé par le nombres d'individus utilisés pour sa construction) ; il possède donc les propriétés :

$$\begin{aligned} \forall i \in [1..n] \quad , \quad y_i^j &\geq 0 \\ \forall i \in [1..n] \quad , \quad \sum_{j=1}^p y_i^j &= 1 \end{aligned} \tag{4.10}$$

On attribue à chaque point y_i un poids p_i souvent proportionnel au nombre d'individus ayant servi à la construction de l'histogramme de ν dans la population i .

On définit la distribution moyenne $\bar{y}$ par :

$$\bar{y} = \sum_{i=1}^n p_i y_i$$

On va alors modifier le tableau Y en normalisant chaque colonne par $1/\sqrt{\bar{y}^j}$. Soit $D_Y = \text{diag}(\bar{y})$ la matrice diagonale contenant les valeurs représentant la distribution moyenne des y_i . Ces valeurs sont supposées strictement positives car une modalité non observée n'apporte aucune information sur les populations étudiées. On construit le tableau X de la façon suivante :

$$X = Y.D_Y^{-1/2}$$

On définit le barycentre $g = \sum_{i=1}^n p_i x_i$, et la pondération par les valeurs de $\bar{y}$ induit :

$$\begin{aligned} d^2(x_i, g) &= \sum_{j=1}^p (x_i^j - g^j)^2 \\ d^2(x_i, g) &= \sum_{j=1}^p \frac{1}{\bar{y}^j} (y_i^j - \bar{y}^j)^2 \\ d^2(x_i, g) &= d_{\chi^2}^2(y_i, \bar{y}) \end{aligned}$$

La distance euclidienne entre un point x_i et g correspond à la distance du χ^2 entre y_i et $\bar{y}$.

L'analyse du tableau X selon la méthode du paragraphe c possède certaines particularités provenant de la propriétés 4.10 et permet de construire des facteurs principaux par combinaison linéaire des fréquences de chaque modalité. Ces facteurs permettent en quelque sorte de mettre en avant les modalités de la variable ν pour lesquelles les populations étudiées se différencient le plus, et de résumer par quelques valeurs pertinentes les différences entre les distributions de ν sur les différentes populations étudiées et la distribution moyenne.

4.1.4. Utilisation de l'analyse factorielle

L'analyse factorielle permet de résumer un grand nombre de variables en un petit nombre de facteurs principaux. Cette démarche de réduction des dimensions de l'espace de représentation est essentielle car la plupart des algorithmes de classification rencontrent des problèmes de stabilité et de temps de calcul si l'on prend en compte trop de variables.

a. Choix d'un espace à partir d'une base de surface

Le première étape d'utilisation de l'analyse factorielle consiste à établir les facteurs principaux, c'est à dire le sous-espace dans lequel on va projeter les données. Pour cela, on dispose d'une "base" d'individus parfaitement connus et supposés représentatifs d'une population plus large que l'on cherche à étudier. Dans le cas du contrôle qualité par

exemple, une base "représentative" est une base dont les individus sont répartis dans toute la gamme de qualité. Elle contient des exemples d'objets contenant les pires défauts comme d'objets de la meilleure qualité en passant par la gamme intermédiaire. On applique alors la méthodologie de la section 4.1.1 sur cette base pour calculer les axes propres, c'est à dire les combinaisons linéaires des descripteurs de départ qui sont le plus adaptés à l'étude de nos objets. En plus d'un espace de projection pertinent, cette étape d'apprentissage donne souvent des indications sur les individus de la base. Elle permet par exemple d'établir un lien entre des connaissances qualitatives sur les individus (en termes de surfaces, on pourrait citer l'aspect visuel structuré ou lisse, homogène ou hétérogène, l'usine de fabrication, certains paramètres process, etc...) et les descripteurs quantitatifs (mesures) dont on dispose pour caractériser ces surfaces.

b. Projection d'individus supplémentaires

Une fois déterminé un sous-espace d'analyse pertinent pour un critère donné, on peut étudier de nouveaux individus extérieurs à la base d'apprentissage en comparant leur projection dans ce sous-espace avec les projections des points de la base. Le critère de fidélité des distances permet en quelque sorte d'assurer la pertinence de ces calculs malgré la dégradation de l'information due à la projection

c. Travers possibles de l'analyse factorielle

Dans de nombreux cas, l'analyse factorielle fournit de très bons résultats, mais s'intéresser aux dangers potentiels de la méthode reste instructif. Ces dangers sont essentiellement de trois types :

Sous-apprentissage : c'est le cas lorsqu'un individu supplémentaire présente des comportements très particuliers vis à vis des descripteurs considérés et qu'il se trouve très "loin" du sous-espace d'étude choisi. Sa projection sur ce sous-espace pourra alors être trompeuse et le faire paraître proche de l'un des individus de la base de manière complètement artificielle. Ce phénomène se détecte souvent en observant les différents indicateurs pour le nouvel individu (Cos^2 , inertie expliquée par chacun des axes propres, etc...)

Sur-apprentissage : lorsque les descripteurs ou la distance choisie sont mal adaptés à l'étude d'une base, l'analyse devient hyper-sensible à des particularités minimales des individus de la base comme le bruit de mesure par exemple et on observe dans ce cas des écarts artificiels dans l'espace de projection entre des individus en réalité très ressemblants. Il est d'ailleurs fréquent qu'un phénomène de sur-apprentissage sur une base entraîne un phénomène de sous-apprentissage pour les individus supplémentaires. Un très faible pourcentage d'inertie expliqué par le sous-espace choisi constitue l'un des symptômes du sur-apprentissage, et nous verrons dans la suite que l'utilisation d'individus "bootstrap" fournit aussi un bon moyen de détecter ce problème.

Corrélations des descripteurs : si certains descripteurs de départ sont naturellement fortement corrélés tout se passera comme si on leur attribuait un poids plus im-

portant dans l'analyse. Si l'on pousse à l'extrême, c'est ce que l'on observerait si l'on utilisait par mégarde plusieurs fois le même descripteurs dans l'analyse. Il ne s'agit pas forcément d'un problème, mais il faut être conscient de cet effet lors de l'interprétation de certains résultats.

d. Bootstrap et Analyse factorielle

Les individus fictifs obtenus par sur-échantillonnage de type bootstrap sont équivalents sur le plan mathématique aux individus originaux, il est donc tout à fait possible de les utiliser dans une analyse factorielle, et ce, de deux façons différentes.

La première utilisation possible consiste à faire plusieurs calculs de facteurs principaux [7] en substituant les individus bootstrap aux individus originaux dans la matrice X . On obtient ainsi plusieurs facteurs principaux, qui reflètent la variation "naturelle" que l'on observerait si l'on disposait de plusieurs bases d'échantillons statistiquement équivalentes : la variance observée dans les facteurs principaux lorsque l'on fait varier les individus bootstrap est un indicateur très intéressant, car elle quantifie la stabilité d'un sous-espace de projection, et une variance élevée est un symptôme net de sur-apprentissage. Ce type d'exploitation des données bootstrap est très riche d'information mais présente quelques inconvénients : tout d'abord, les vecteurs propres de la matrice d'inertie sont définis au sens près, et cela peut perturber les calculs de variance, ensuite, les calculs à fournir pour estimer correctement cette variance peuvent être assez conséquents car ils impliquent de nombreux calculs de matrice d'inertie et des vecteurs propres associés.

La seconde utilisation possible consiste simplement à projeter les échantillons bootstrap dans le sous-espace d'étude choisi, et à observer la qualité de cette projection ainsi que la dispersion des échantillons dans cet espace. Comme pour la première approche, ces indicateurs nous permettront de détecter, bien que de manière plus qualitative, les phénomènes de sur-apprentissage, mais surtout nous disposons dans ce cas de tous les éléments qui nous permettront de mettre en oeuvre une classification bayésienne.

4.2. Classification

Introduction

Dans cette section, on suppose qu'un individu est caractérisé par un petit nombre de valeurs $x^1, x^2, \dots, x^v$ considérées comme les réalisations de v variables aléatoires $X^1, X^2, \dots, X^v$. Dans un contexte physique réel, les valeurs $x^1, x^2, \dots, x^v$ sont les résultats de v mesures effectuées sur un individu, il peut aussi s'agir de combinaisons linéaires de mesures obtenues comme facteurs principaux d'une analyse factorielle. On se donne d'autre part une variable aléatoire de classement C discrète pouvant prendre plusieurs modalités $c_1, \dots, c_K$. Cette variable représente une partition des individus en plusieurs classes (individus corrects ou incorrects par exemple dans le cas d'un contrôle qualité). L'objectif d'une opération de classification est de déterminer la modalité de la variable de classement c_r correspondant à un individu, connaissant les valeurs des v me-

sures $x^1, x^2, \dots, x^v$ qui lui sont associées. Si les variables aléatoires $X^1, X^2, \dots, X^v$ ont été bien choisies, les mesures effectuées doivent permettre la détermination de c_r avec une faible probabilité d'erreur. La classification bayésienne nous permet de replacer ce problème dans un contexte probabiliste et de déterminer ainsi une stratégie de classification "optimale" tout en évaluant le risque d'erreur associé à un choix de c_r .

4.2.1. Classification bayésienne

a. Notations

Nous commençons par introduire quelques notations :

- La densité d'une variable aléatoire X sera notée $L(X = x)$ ou $L(x)$ si la notation ne présente pas d'ambiguïté.
- Une densité multivariée sera noté $L(X^1 = x^1, X^2 = x^2, \dots)$, ou plus simplement $L(x^1, x^2, \dots)$.
- La densité conditionnelle par rapport à une variable Y sera notée $L(X^1 = x^1, X^2 = x^2, \dots | Y = y)$ ou plus simplement $L(x^1, x^2, \dots | y)$

En ce qui concerne la variable de classement discrète C , on utilisera plutôt les notations $P(c)$ et $P(c|y)$ pour la probabilité (conditionnelle ou non) d'obtenir la modalité c .

b. Principe de l'estimation bayésienne

Etant donnée une réalisation $x^1, x^2, \dots, x^v$ des variables aléatoires caractéristiques d'un individu, on cherche à déterminer la valeur c_{max} la plus probable pour la variable de classement C , c'est à dire la valeur maximisant la probabilité $P(c|x^1, x^2, \dots, x^v)$. On utilise pour cela la relation de Bayes :

$$P(c|x^1, x^2, \dots, x^v) = \frac{L(x^1, x^2, \dots, x^v|c)P(c)}{L(x^1, x^2, \dots, x^v)} \quad (4.11)$$

- La probabilité $P(c)$ caractérise une connaissance a priori sur l'ensemble des modalités de la variable C . Si l'on ne dispose pas de cette information, on peut simplement supposer que toutes les modalités de C sont équiprobables.
- Les densités conditionnelles $L(x^1, x^2, \dots, x^v|c)$ sont supposées connues et proviennent dans notre cas d'une analyse préliminaire sur une base d'échantillons pour lesquels la variable c est donnée.
- La densité de probabilité $L(x^1, x^2, \dots, x^v)$ peut s'obtenir au moyen de la formule de décomposition :

$$L(x^1, x^2, \dots, x^v) = \sum_c L(x^1, x^2, \dots, x^v|c)P(c)$$

Notons que pour la recherche de c_{max} maximisant $P(c|x^1, x^2, \dots, x^v)$, la valeur de $L(x^1, x^2, \dots, x^v)$ n'intervient pas puisqu'elle est la même pour toutes les modalités $c_1, \dots, c_K$ de C .

L'attribution de la classe c_{max} pour un individu caractérisé par $x^1, \dots, x^v$ minimise donc le risque d'erreur donné par :

$$\epsilon = \sum_{c \neq c_{max}} P(c|x^1, x^2, \dots, x^v) = 1 - P(c_{max}|x^1, x^2, \dots, x^v)$$

c. Introduction de fonctions de coût

La démarche précédente fait uniquement intervenir la probabilité $P(c|x^1, x^2, \dots, x^v)$. Cependant, le fait de choisir c_{estim} à la place d'une valeur réelle c_r peut avoir des conséquences plus ou moins importantes. Parfois, la valeur la plus probable c_{max} est un choix désastreux en cas d'erreur, alors qu'une autre valeur c limiterait éventuellement les effets néfastes si $c_r \neq c$. Ce cas de figure est typique des jeux de hasard où l'on peut limiter ses chances dans l'espoir de gagner plus, ou au contraire prendre moins de risques de perdre tout en limitant le gain potentiel. Chacun se doit alors d'établir une stratégie en fonction de son goût du risque et de son portefeuille!

Un tel choix de stratégie n'apparaît pas dans le calcul du paragraphe précédent, mais il peut être introduit à l'aide de la notion de fonction de coût. Soit :

$$\begin{aligned} \Phi : \mathcal{K} \times \mathcal{K} &\rightarrow \mathbb{R} \\ (c_i, c_j) &\rightarrow \Phi(c_i, c_j) \end{aligned}$$

$\mathcal{K}$ représente l'ensemble des modalités de la variable aléatoire C , et $\Phi(c_1, c_2)$ caractérise le coût induit pour le décideur lorsque la modalité c_1 de C est choisie à tort à la place d'une modalité réelle c_2 . En général, $\Phi(c, c)$ est pris nul, car aucun coût ne résulte du choix correct de la valeur de c .

Le calcul du paragraphe précédent peut être modifié de façon à minimiser l'espérance du coût $\psi(c_0)$ induit par le choix de c_0 .

$$E(\psi(c_0)) = \sum_{c \in \mathcal{K}} \Phi(c_0, c) P(c|x^1, x^2, \dots, x^v)$$

On cherche donc à déterminer $c_{min} = \operatorname{argmin}_{c \in \mathcal{K}} E(\psi(c))$

Si tous les $P(c_i, c_j)$ sont identiques lorsque $c_i \neq c_j$ et $\forall i, P(c_i, c_i) = 0$, on retrouve le résultat du paragraphe précédent.

4.2.2. Estimation des fonctions de densité

Les formules de décisions du paragraphe 4.2.1 font explicitement intervenir les fonctions de densités conditionnelles multivariées $L(x^1, x^2, \dots, x^v|c)$ et leur estimation est donc un prérequis nécessaire aux opérations de classification proprement dites.

Cette section est destinée à passer en revue plusieurs méthodes d'évaluation d'une densité à partir d'un échantillon test de n individus $x = (x_1, \dots, x_n)$ caractérisés chacun par v valeurs $x_i^1, \dots, x_i^v$. Le premier paragraphe est consacré à la méthode la plus simple basée sur des histogrammes, le second traite de méthodes paramétriques basées

sur des fonctions de densités de type "mélange" [49] et sur l'algorithme EM (Expectation-Maximisation) [51]. Enfin, nous présentons dans un troisième paragraphe une technique semi-paramétrique basée sur la méthode des noyaux [41].

a. Histogrammes

La première façon d'évaluer une densité est d'utiliser une discrétisation de l'espace en p régions $R_1, R_2, \dots, R_p$: la probabilité pour un point de tomber dans la région R_i est simplement évaluée en construisant un histogramme, c'est à dire en calculant la proportion d'individus de l'échantillon tombant dans la région R_i . Cette méthode a l'avantage de la simplicité et de la rapidité de calcul : avec un nombre de régions limité, on peut même, dans une optique de classification, précalculer pour chacune d'entre elle la modalité de C la plus probable (ou la moins coûteuse), et la classification d'un individu inconnu consiste simplement à déterminer la région R_i correspondant aux valeurs $x^1, \dots, x^v$ mesurées.

Des écueils importants viennent tout de même contrebalancer cette simplicité. En premier lieu, le choix de la partition de l'espace à adopter (nombre de régions, taille des régions) est un problème difficile, en particulier si le nombre de dimensions de l'espace dépasse 1 ou 2 [17][57]. D'autre part, l'estimation d'une densité par cette méthode se révèle souvent instable si le nombre d'individus de l'échantillon n'est pas très important, et d'autant plus que le nombre de dimensions de l'espace considéré augmente. Enfin, cette méthode reste assez grossière, puisqu'elle ne tient pas compte de la position exacte d'un individu, mais simplement de la région de l'espace à laquelle il appartient. Tous ces éléments nous incitent à renoncer à l'utilisation des histogrammes (en vue d'une classification bayésienne) dès que la dimension de l'espace des descripteurs considérés est strictement supérieure à 1.

b. Méthodes paramétriques

Les méthodes paramétriques constituent une classe importante des techniques utilisées pour évaluer des densités de probabilité. On se donne ici un modèle, c'est à dire une classe de fonctions de densité dépendant d'un petit nombre de paramètres que l'on détermine en fonction des données de l'échantillon test.

On représente l'ensemble des paramètres du modèle par une variable θ et l'on exprime la fonction de densité cherchée sous la forme :

$$L(x) = f(x|\theta)$$

(x représente ici le vecteur de coordonnées $x^1, \dots, x^v$)

Le paramètre θ est souvent estimé en utilisant le principe du maximum de vraisemblance, c'est à dire en maximisant $\prod_{i=1}^n f(x_i|\theta)$.

Exemple des mélanges On suppose que la fonction de densité recherchée est un mélange [49] de plusieurs fonctions de densité $p_j(x|\theta_j)$, par exemple des densités gaussiennes de

moyenne et d'écart-type inconnus :

$$\begin{aligned} f(x|\theta, \alpha) &= \sum_{j=1}^m \alpha^j p_j(x|\theta_j) \\ \forall j \in [1..m], \quad \alpha^j &> 0 \\ \sum_{j=1}^m \alpha^j &= 1 \end{aligned}$$

Les poids α^j peuvent être considérés comme la distribution d'une variable aléatoire discrète Z (dite "cachée") caractérisant pour chaque individu la distribution p_j dont il est issu. Dans ce contexte, on cherche à déterminer en fonction des données les poids α_j de chaque fonction p_j ainsi que le vecteur θ représentant l'ensemble des paramètres θ_j . L'algorithme EM[51] est une méthode itérative pour parvenir à ce résultat.

L'algorithme EM On suppose que l'on dispose d'individus tests pour lesquels on observe une variable aléatoire X dont la distribution dépend d'une variable aléatoire Z cachée. La distribution $L(X = x|\theta)$ recherchée dépend d'un paramètre (ou d'un vecteur de paramètres) θ maximisant la fonction de vraisemblance $L(\theta) = \log(L(x_1, \dots, x_n|\theta))$.

A partir d'une première estimation θ_0 de θ , on applique un algorithme dont chaque itération se décompose en deux étapes :

- Etape "E" (expectation) : On cherche une distribution $\tilde{L}^t$ sur Z telle que :

$$\tilde{L}^t(Z = z) = L(z|x, \theta^{t-1})$$

- Etape "M" (maximisation) : On détermine θ^t de façon à maximiser :

$$E_{\tilde{L}^t}[\log L(z, x|\theta)]$$

Plusieurs versions de cet algorithme existent selon la façon d'obtenir $\tilde{L}^t$ ou de maximiser $E_{\tilde{L}^t}[\log L(z, x|\theta)]$ qui peuvent conduire à des vitesses de convergence différentes et des résultats plus ou moins approchés.

Grâce à sa robustesse, cette méthode se révèle une bonne solution lorsque la dimension de l'espace de départ est importante ou que le nombre d'individus de l'échantillon test est limité. Le revers de la médaille provient du choix du modèle à priori qui peut se révéler en désaccord avec les données et induire ultérieurement des erreurs de classification. D'autre part, l'algorithme *EM* reste complexe et nécessite parfois des temps de calculs assez long pour une convergence vers un maximum local de $L(\theta)$ qui n'est pas nécessairement le maximum global.

c. Utilisation de noyaux

Cette dernière méthode se présente en quelque sorte comme un compromis entre stabilité et paramétrisation. La classe de fonctions admissibles est plus large que dans le cas des mélanges et la méthode évite certains problèmes liés à l'utilisation d'histogrammes.

La densité en chaque point de l'espace est ici considérée comme la somme de contributions provenant des n individus de l'échantillon test :

$$L(x) = \frac{1}{n} \sum_{i=1}^n K(x, x_i, h(x_i)) \quad (4.12)$$

Où le noyau K est une fonction de densité admissible pour la variable X . En général K s'exprime sous la forme suivante [41][19][20] :

$$K(x, x_i, h(x_i)) = \frac{1}{h(x_i)^v} K\left(\frac{x - x_i}{h(x_i)}\right)$$

Où v désigne la dimension de l'espace des descripteurs.

Concrètement, la densité est forte en un point x si celui-ci est proche de nombreux individus de l'échantillon initial. Les paramètres $h(x_i)$, $i \in [1..n]$ permettent de régulariser la distribution, il sont en général choisis de façon à être inversement proportionnel à la densité en un point x_i . Si $h(x_i)$ est faible, cela signifie qu'il apporte une très forte contribution aux points situés proche de lui mais que son influence spatiale décroît très vite avec la distance entre x et x_i . Au contraire, si $h(x_i)$ est fort, le point x_i a une faible contribution à la densité calculé en un point x , mais son influence s'exerce à plus grande distance. $h(x_i)$ joue un rôle de filtrage compensant le faible nombre d'individus tombant dans les zones de faible densité, et il sera donc faible lorsque de nombreux individus de l'échantillon test sont concentrés dans une faible zone (forte densité), et fort dans le cas contraire.

On parle d'estimation de densité semi-paramétrique pour ce genre de méthode, car s'il faut choisir a priori un noyau K et les $h(x_i)$, ce choix est beaucoup moins restrictif que dans le cas des mélanges par exemple, et les densités obtenues ont plus de chance de s'adapter correctement à n'importe quel type de distribution. On peut même tenter de limiter encore la quantité d'information à fournir en déterminant les $h(x_i)$ à partir des données.

Les méthodes à noyaux ont d'autres propriétés intéressantes, comme celle de pouvoir estimer non seulement la densité de probabilité en un point quelconque, mais aussi le gradient de cette densité [19][20], ce qui nous rend capable de détecter efficacement les modes d'une distribution. Ce résultat peut se révéler utile dans différentes circonstances, par exemple pour redécouper a posteriori une classe de la variable C en plusieurs sous-classes. Imaginons que C soit composée de deux classes, une classe "correcte" et une classe "incorrecte" et que la densité correspondant à la classe "incorrecte" présente plusieurs modes bien marqués : on peut en déduire que les descripteurs choisis nous permettent de mettre en lumière plusieurs types de défauts, et donc de rediviser la classe "incorrecte" en plusieurs sous-classes.

Evaluation du gradient de la densité et détection des modes Nous nous plaçons dans le cas où :

$$K(x, x_i, h_i) = \frac{1}{h_i^v} k\left(\left\|\frac{x - x_i}{h_i}\right\|^2\right)$$

Et on notera $g(x) = -k'(x)$.

En calculant le gradient de $L(x)$ à partir de la relation 4.12, on obtient l'égalité :

$$\nabla L(x) = \frac{2}{n} \left[\sum_{i=1}^n \left(\frac{1}{h_i} \right)^{v+2} g \left(\left\| \frac{x - x_i}{h_i} \right\|^2 \right) \right] \times \left[\frac{\sum_{i=1}^n \left(\frac{x_i}{h_i} \right)^{v+2} g \left(\left\| \frac{x - x_i}{h_i} \right\|^2 \right)}{\sum_{i=1}^n \left(\frac{1}{h_i} \right)^{v+2} g \left(\left\| \frac{x - x_i}{h_i} \right\|^2 \right)} - x \right] \quad (4.13)$$

Le second facteur du second membre est appelé "mean shift vector", et présente la bonne propriété d'être calculable facilement tout en s'orientant selon le gradient de L . On note :

$$M_v(x) = \frac{\sum_{i=1}^n \left(\frac{x_i}{h_i} \right)^{v+2} g \left(\left\| \frac{x - x_i}{h_i} \right\|^2 \right)}{\sum_{i=1}^n \left(\frac{1}{h_i} \right)^{v+2} g \left(\left\| \frac{x - x_i}{h_i} \right\|^2 \right)} - x \quad (4.14)$$

A partir d'un point y_0 quelconque, on peut calculer $y_1 = y_0 + M_v(y_0)$ puis $y_2 = y_1 + M_v(y_1)$ et itérer le processus. Si k est une fonction monotone, décroissante et convexe, on peut montrer que la suite $(y_r)_{r \in \mathbb{N}}$ converge vers un point en lequel le gradient de la densité L s'annule et que la suite des valeurs $(L(y_r))_{r \in \mathbb{N}}$ est croissante. La limite de $(y_r)_{r \in \mathbb{N}}$ est donc un maximum potentiel de la fonction de densité L .

Choix des $h(x_i)$ Nous présentons ici quelques résultats [19][20] concernant le choix des valeurs $h(x_i)$ en fonction des données.

La méthode la plus simple consiste à choisir une valeur h unique pour l'ensemble des x_i . Il existe des formules basées sur la minimisation du MISE (mean integrated square error) définie par :

$$MISE(\hat{f}, f) = E \int (\hat{f}(x) - f(x))^2 dx$$

Où $\hat{f}$ désigne la densité estimée par la méthode des noyaux et f la véritable densité (inconnue). En développant cette formule en série de Taylor et en supposant que la densité réelle f est une gaussienne de variance σ^2 estimée à partir de l'échantillon test par l'écart-type expérimental $\hat{\sigma}$, on peut obtenir une valeur pour h minimisant le MISE :

$$h_0 = n^{\frac{-1}{v+4}} \hat{\sigma}$$

On peut ensuite utiliser la valeur de h_0 estimée pour calculer une première estimation $\hat{f}_0$ de f (cette approche se généralise bien au cas vectoriel en raisonnant composante par composante avec des noyaux gaussiens multidimensionnels dont la matrice de variance-covariance est diagonale).

On s'aperçoit en général que la qualité de l'estimation d'une densité augmente considérablement si l'on choisit des $h(x_i)$ différents pour chaque x_i . Une méthode efficace consiste à utiliser $\hat{f}_0$ pour affiner le choix des $h(x_i)$ en chaque point au moyen de la formule :

$$h(x_i) = h_0 \left(\frac{\lambda}{\hat{f}_0(x_i)} \right)^{1/2} \quad (4.15)$$

Où λ représente une constante de proportionnalité.

Comaniciu signale[21] que ce choix des $h(x_i)$ diminue considérablement le biais dans l'estimation de f et accélère la vitesse de convergence de $\hat{f}$ vers f lorsque n tend vers l'infini, et cela malgré le choix quelque peu empirique de λ pris égal à la moyenne géométrique des densités $\hat{f}_0(x_i)$.

D'autres méthodes de choix se basent sur de la "validation croisée" ou du bootstrap, mais elles sont souvent coûteuses à la fois en temps de calcul et en volume de données à exploiter.

La figure 4.1 montre l'application de la méthode des noyaux à l'estimation d'une densité (monodimensionnelle) correspondant à une loi gamma de paramètre 2. Un échantillon de 200 individus est utilisé pour calculer $\hat{f}_0(x)$, puis $\hat{f}_1(x)$ et $\hat{f}_2(x)$ en utilisant la relation 4.15 pour modifier les valeurs de $h(x_i)$ en chaque point de l'échantillon. Les courbes montrent nettement l'intérêt d'adapter les valeurs $h(x_i)$ en chaque point de l'échantillon. Le résultat de l'estimation, si il n'est pas parfait, reste très correct, compte-tenu de la très forte asymétrie de la loi gamma (estimée ici avec un noyau K gaussien, donc symétrique) et de la très forte variation de densité au voisinage de 0. D'autre part, la figure présente l'estimation de la densité par une loi normale de moyenne et écart-type estimé à partir de l'échantillon : la très faible qualité de cette approximation souligne l'intérêt de la méthode des noyaux par rapport à une méthode "paramétrique".

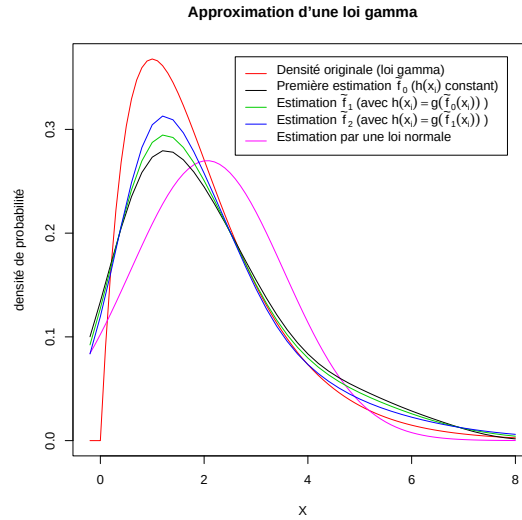


FIG. 4.1.: Estimation d'une loi gamma à partir d'un échantillon de 200 individus

4.3. Méthodologie d'analyse sélectionnée

Dans cette section, nous présentons une méthodologie originale d'étude des surfaces rugueuses basée à la fois sur des descripteurs morphologiques et sur des méthodes d'ana-

lyse des données, de bootstrap et de classification bayésienne. Le résultat se veut le plus générique possible en terme d'objets étudiés et de descripteurs sélectionnés de façon à être applicable dans un cadre le plus large possible.

Notre protocole d'étude requiert une phase d'apprentissage à partir d'un échantillon de surfaces parfaitement connues regroupées en plusieurs catégories ($c_1, c_2, \dots, c_k$). Cette étape fournit les prérequis nécessaire à une phase de classification de surfaces inconnues selon les mêmes catégories ($c_1, c_2, \dots, c_k$).

4.3.1. Apprentissage

La phase d'apprentissage de notre méthodologie aspire à l'automatisme le plus poussé possible, cependant, un oeil averti et critique vis à vis de ces résultats sera la meilleure garantie de succès pour la phase de classification.

a. Analyse des descripteurs

La plupart des descripteurs du chapitre 2 fournissent une courbe ou un histogramme caractéristique d'une surface. Ces courbes contiennent une grande richesse d'information qui dépasse souvent le cadre de l'analyse et peut être réductrice pour la stabilité ultérieure de la classification. C'est pourquoi une première phase d'analyse factorielle permet de les résumer en quelques valeurs adaptées à l'échantillon étudié. Selon la nature des descripteurs et des courbes étudiées, on appliquera l'une ou l'autre des méthodes d'analyse factorielle présentées (Analyse euclidienne classique, ACP normée, AFC ...) de façon à ce qu'à l'issue de cette étape, chaque descripteur se trouve résumé en quelques valeurs (en générale une ou deux). Si les méthodes d'analyse factorielle fournissent un bon moyen d'effectuer cette "compression" de l'information utile *en fonction des données*, elle ne sont certainement pas les seules applicables : la plupart des paramètres dits "classiques" de rugosité par exemple, résument l'information contenue dans un profil en une valeur obtenue par moyenne ou borne supérieure : il s'agit déjà d'une réduction de l'information, mais cette réduction n'est pas guidée par les données que l'on cherche à analyser.

b. Analyse globale

L'analyse de données descripteur par descripteur permet de disposer d'un grand nombre de *valeurs* (et non plus des courbes) caractéristiques d'une surface. Malgré tout, la dimension de cet espace est encore importante, en particulier si le nombre des surfaces de base est peu élevé. Dans la première phase d'analyse, il est en général possible de tenir compte de la cohérence des données (courbes, histogrammes) pour appliquer une analyse factorielle spécifique à chaque descripteur. Dans cette seconde phase d'analyse globale, les descripteurs peuvent être beaucoup plus hétéroclites, et la méthode la plus adaptée à ce type de situation reste bien souvent l'ACP normée. Le résultat est une représentation des surfaces dans un espace de dimension très faible (deux ou trois), mais synthétisant au mieux (au sens d'un critère inertiel) les différences entre les surfaces de l'échantillon test. Il est tout à fait possible de rajouter a posteriori des dimensions à cet espace réduit en utilisant des descripteurs formés de valeurs uniques pouvant apporter

une spécificité à l'analyse et d'obtenir ainsi un espace "mixte" formé de descripteurs composites issus de deux étapes d'analyse factorielle et de descripteurs numériques issus directement d'appareils de mesure.

c. Utilisation du bootstrap

L'utilisation pertinente des méthodes de bootstrap par bloc reste l'un des points essentiels de notre méthodologie : chaque descripteur se doit d'être accompagné d'une méthode adaptée permettant de générer un grand nombre d'individus bootstrap. Ceux-ci peuvent être utilisés au cours de la première phase d'analyse comme décrit dans la section d c'est à dire être projetés dans les différents espaces réduits. Cette projection peut conduire à éliminer certains descripteurs de l'analyse globale si ceux-ci sont visiblement soumis à un phénomène de sur-apprentissage.

La méthode que nous utilisons ensuite pour obtenir un individu bootstrap dans l'analyse globale consiste à bootstrapper chacun des descripteurs pris en compte, et de considérer l'ensemble des valeurs obtenues comme les valeurs caractérisant une seule surface équivalente à la surface d'origine. Ce nouvel individu est d'abord projeté dans les espaces réduits correspondant à la première phase d'analyse puis dans l'espace correspondant à l'analyse globale. Il faut bien souligner que cette méthode de génération d'individus bootstrap *indépendamment* pour chaque descripteur peut introduire une distorsion des relations entre descripteurs existant pour les surfaces réelles. Ces distorsions sont *en pratique* minimales devant les différences entre surfaces de catégories différentes, et correspondent à ce qu'on observerait si chacun des descripteurs était extrait sur des surfaces différentes mais statistiquement équivalentes au lieu d'une seule et même surface (c'est d'ailleurs ce qui se passe dans le cas de mesure destructives). Cette indépendance "forcée" plaide, lorsque c'est possible, en faveur de la génération d'individus bootstrap directement au niveau de l'image, de façon à extraire l'ensemble des descripteurs à partir d'une même image bootstrap et à conserver ainsi les relations de dépendance entre descripteurs.

d. Estimation des densités pour la classification

Soit S_0 une surface projetée en x_0 dans l'espace réduit. On considère l'ensemble des individus bootstrap $(x_1^*, \dots, x_B^*)$ correspondant à S_0 et projetés dans l'espace réduit. Dans ce qui suit X désigne un vecteur aléatoire représentant la position dans l'espace réduit d'une surface issue du même processus aléatoire que S_0 . Soit $\theta_0 = E(X)$ la position moyenne de ces surfaces. Cette quantité étant inconnue, on représente l'information dont on dispose à son sujet par une variable aléatoire Θ . La quantité que l'on cherche à évaluer est la densité $L(X = x)$, et on dispose pour cela :

- de x_0
- de la connaissance (avec une bonne approximation) de $\hat{E}(X^*)$.
- d'un échantillon important d'individus bootstrap permettant d'appréhender $L(x^* - \hat{E}(X^*))$ que la théorie du bootstrap permet d'assimiler à la distribution $L(x - \theta_0)$. Dans la suite, la densité $L(X - \theta_0 = \alpha)$ sera noté $f(\alpha)$ par commodité.

θ_0 étant inconnu, il nous est impossible de déterminer la "vraie" distribution $L(X = x)$, par contre, on peut en former un estimateur raisonnable basé sur la loi de la différence entre deux variables aléatoires indépendantes :

$$\hat{L}(X = x) = L(X - X_0 = x - x_0) \quad (4.16)$$

$$\hat{L}(X = x) = f * \check{f}(x - x_0) \quad (4.17)$$

Où $*$ désigne le produit de convolution et avec $\check{f}(x) = f(-x)$.

Démonstration. Posons dans un premier temps :

$$X = \theta_0 + \Delta$$

Δ représente l'écart entre X et $E(X)$ et suit une loi définie par f .

On définit Δ_0 de manière similaire, et l'équation 4.16 peut se réécrire sous la forme :

$$\hat{L}(X = x) = L(\Delta - \Delta_0 = x - x_0)$$

On emploie alors une formule de balayage classique ajoutée à l'indépendance de variables Δ et Δ_0 pour obtenir :

$$\begin{aligned} L(\Delta - \Delta_0 = x - x_0) &= \int_{\delta} L(\Delta = \delta, \Delta_0 = \delta - (x - x_0)) d\delta \\ L(\Delta - \Delta_0 = x - x_0) &= \int_{\delta} L(\Delta = \delta) L(\Delta_0 = \delta - (x - x_0)) d\delta \\ L(\Delta - \Delta_0 = x - x_0) &= \int_{\delta} f(\delta) f(\delta - (x - x_0)) d\delta \\ L(\Delta - \Delta_0 = x - x_0) &= \int_{\delta} f(\delta) \check{f}(x - x_0 - \delta) d\delta \\ L(\Delta - \Delta_0 = x - x_0) &= f * \check{f}(x - x_0) \end{aligned}$$

□

La donnée des individus bootstrap et de $\hat{E}(X^*)$ permet de construire les nouveaux échantillons bootstrap $\delta^* = (\delta_1^*, \dots, \delta_B^*)$ et $\gamma^* = (\gamma_1^*, \dots, \gamma_B^*)$ définis par :

$$\begin{aligned} \forall b \in [1..B] \quad , \quad \delta_b^* &= x_b^* - \hat{E}(X^*) \\ \forall b \in [1..B] \quad , \quad \gamma_b^* &= \hat{E}X^* - x_b^* \end{aligned}$$

Ces échantillons peuvent être respectivement considérés comme obtenus par tirages aléatoires indépendants à partir de f et $\check{f}$. Un moyen d'obtenir un échantillon de taille B^2 issu de $f * \check{f}$ consiste à construire l'échantillon $\mu^* = (\mu_1^*, \dots, \mu_{B^2}^*)$ défini par :

$$\forall n \in [1..B], \forall b \in [1..B], \mu_{(n-1).B+b}^* = \gamma_n^* + \delta_b^*$$

L'échantillon μ^* permet ensuite d'estimer la densité de probabilité $\hat{L}(X = x)$ qui représente l'estimation de la distribution de la position du projeté dans l'espace réduit d'une surface issue du même processus aléatoire que S_0 .

Supposons maintenant que plusieurs surfaces $S_1, S_2, \dots, S_p$ appartiennent à une même catégorie c_k . Pour chacune d'entre elle, un échantillon bootstrap $\mu^{*i}, i \in [1..p]$ est généré. La réunion de l'ensemble de ces échantillons bootstrap fournit un échantillon $\mu_{c_k}^*$ issu de la distribution de $\hat{L}(x|c_k)$ qui peut donc être estimée en pratique par la méthode des noyaux. En effectuant ces calculs pour toutes les catégories, on obtient l'ensemble des prérequis nécessaire à l'application d'une classification bayésienne dans un espace réduit, ce qui termine la phase d'apprentissage. Dans la suite, pour simplifier les notations pour tout ce qui concerne la classification bayésienne, nous confondrons $\hat{L}(x|c_k)$ et $L(x|c_k)$ qu'il est de toute façon impossible de connaître exactement.

4.3.2. Classification

La seconde étape de notre méthodologie consiste à attribuer une catégorie à une surface S inconnue. Pour ce faire, on commence par extraire de S l'ensemble des descripteurs choisis pour caractériser les surfaces de la base. Pour chacun d'entre eux, on projette S dans l'espace réduit associé issu de la première phase d'analyse factorielle. Dans un second temps, les valeurs ainsi obtenues sont utilisées pour projeter la surface S dans l'espace réduit issu de l'analyse globale. On obtient donc pour S les mêmes descripteurs composites, c'est à dire construits par combinaison linéaire des descripteurs initiaux que pour les surface de base. Ces valeurs peuvent être résumées par un vecteur x auquel sont associées les densités de probabilités $L(x|c_k), k \in [1..K]$ qui, selon le paragraphe 4.2.1 conduisent au calcul de $P(c_k|x), k \in [1..K]$. Une classification bayésienne standard attribue alors à la surface S la catégorie c_S telle que $P(c_S|x) = \max(P(c_k|x), k \in [1..K])$.

On peut aussi utiliser diverses fonctions de coût Φ pour enrichir la classification. Imaginons par exemple que l'on attribue une qualité numérique q_k à chaque classe c_k et que l'on construise la fonction de coût particulière Φ définie par :

$$\begin{aligned} \Phi : \mathcal{K} \times \mathcal{K} &\rightarrow \mathbb{R} \\ (c_i, c_j) &\rightarrow q_j \end{aligned}$$

Dans ces conditions, le coût est le même quelle que soit la catégorie choisie pour S , et vaut :

$$E(q_S) = \sum_{k \in [1..K]} q_k P(c_k|x)$$

Cette valeur fournit en quelque sorte l'espérance de la qualité de la surface en fonction de sa position. La précision de cette quantité s'évalue par la variance :

$$V(q_S) = \left(\sum_{k \in [1..K]} q_k^2 P(c_k|x) \right) - E^2(q_S)$$

4.4. Bilan

Notre méthodologie d'analyse [27] peut se résumer en plusieurs opérations élémentaires : les premières sont liées à l'entraînement du classificateur (choix de l'espace de

représentation, estimation des densités de probabilités pour la classification bayésienne, voir table 4.1); les suivantes sont liées à la classification d'une surface inconnue (voir table 4.2). Dans ce schéma, le rôle essentiel joué par la base d'apprentissage mérite d'être de nouveau souligné. En effet, la base de surface déjà classifiée choisie en début d'analyse conditionne complètement l'ensemble des résultats, il s'agit finalement d'un raisonnement purement inductif à partir d'un exemple de ce que l'on veut obtenir. Un raisonnement plus déductif aurait cherché à calculer a priori le comportement d'une classe de surface, en aurait déduit les descripteurs les plus adaptés et aurait testé ces hypothèses sur des ensembles de surfaces dont le rôle aurait été quelque peu limité vis à vis du schéma global de classification. Une telle approche est rendue très difficile par la nature de la classification à effectuer, qui s'appuie le plus souvent sur des variables de classement assez subjectives comme l'esthétique des surfaces qui dépend grandement de l'application à laquelle elles sont destinées (automobile, électroménager, etc...). Ce contexte particulier justifie en quelque sorte l'approche employée et l'utilisation d'une base qui doit être adaptée à la population et à la variable de classement étudiée. Il peut d'ailleurs arriver que les résultats obtenus par nos méthodes fournissent des indications sur les surfaces traitées et permettent ainsi l'élaboration d'hypothèses plus déductives et plus physiques sur les relations entre la variable de classement et les propriétés des surfaces.

Le second point de notre méthodologie appelant un commentaire est le degré d'automatisme. Le cahier des charges d'un classificateur stipule en général que l'intervention de l'opérateur doit être la plus restreinte possible de façon à limiter les risques de subjectivité. Dans notre cas, l'action de l'utilisateur est requise à plusieurs reprises :

- Il détermine la variable de classement et élabore la base d'apprentissage, il choisit les descripteurs à extraire des surfaces en fonction de la bibliothèque disponible, du type de surface analysée, de l'usage de ces surfaces... Ces éléments constituent en quelque sorte un préliminaire à l'apprentissage.
- L'utilisateur doit aussi fournir le type d'analyse de données à effectuer pour chaque descripteur, la méthode d'obtention d'échantillons bootstrap ainsi que leur nombre. Il précise également le noyau à utiliser pour l'estimation de densité et la fonction de coût à considérer.

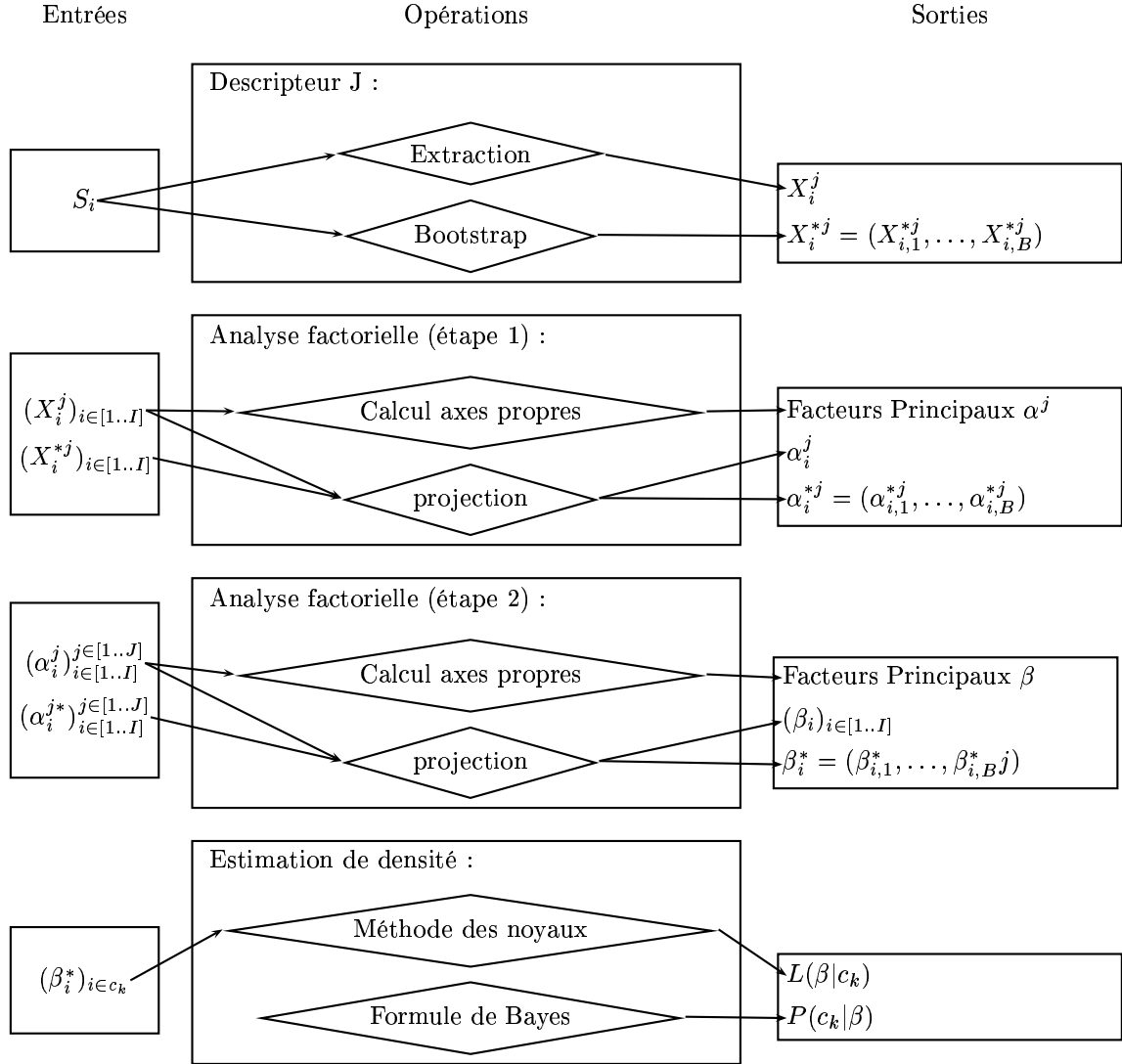
Ces interventions ne sont pas vraiment gênantes, car elles sont *préalables* à toute analyse et permettent au contraire une certaine flexibilité dans la classification et son adaptation à des problèmes assez divers. Dans ce cas de figure surtout, l'utilisateur n'a besoin d'aucune connaissance sur les méthodes employées pour parvenir à la classification finale. L'ensemble du processus peut donc rester une "boîte noire" utilisable par tous.

Idéalement, ces réglages préalables devraient être les seuls requis pour assurer une objectivité totale des résultats. En pratique cependant, on s'aperçoit que les résultats de classification peuvent être nettement améliorés si l'on s'autorise une certaine interactivité *a posteriori*, c'est à dire en fonction des premiers résultats obtenues sur les données :

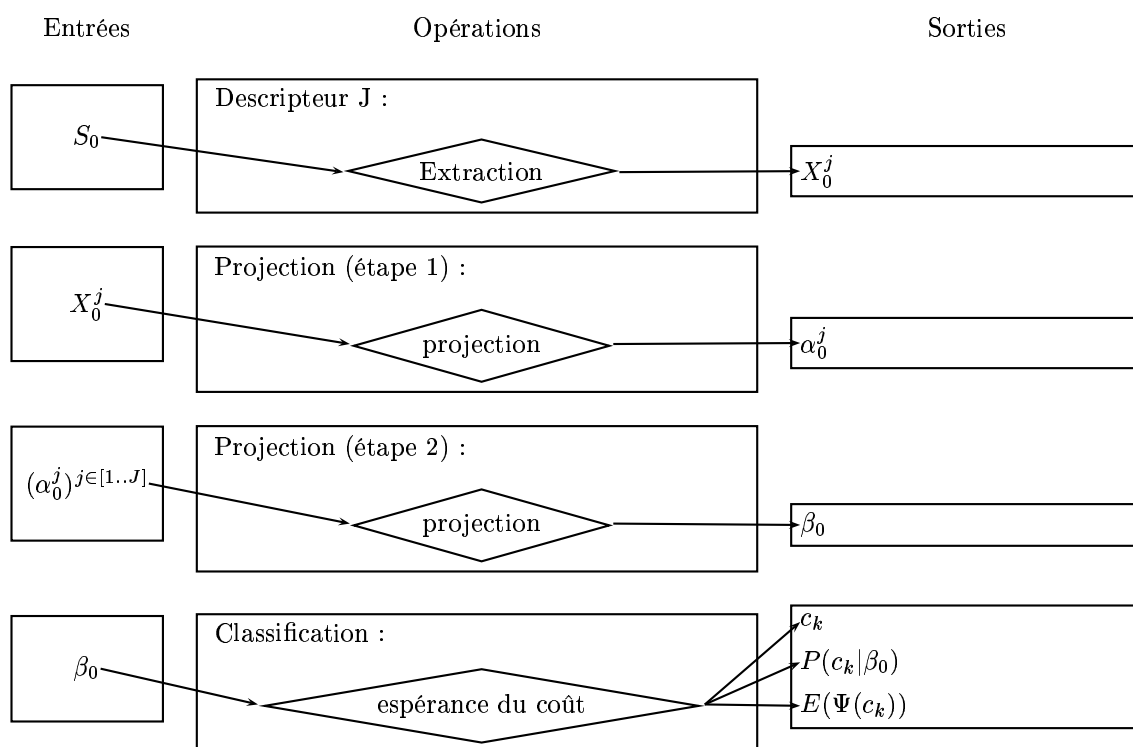
- L'utilisateur peut évaluer la qualité des descripteurs choisis a priori *en fonction* des résultats de la première puis de la seconde étape d'analyse de données. Il peut donc réajuster le panel de descripteurs et recommencer l'analyse et éviter ainsi beaucoup de problèmes de sur-apprentissage.

- En fonction des densités de probabilité calculées dans l'espace réduit, l'utilisateur peut être amené à raffiner ou fusionner les modalités de la variable de classement.

Ces interactions sont plus gênantes que les premières puisqu'elles nécessitent un regard sur les résultats de l'analyse, et requièrent par la même une certaine expertise vis à vis des méthodes employées. Bien sûr, il ne s'agit là que de la phase d'apprentissage, et pas de la phase de classification proprement dite, mais cela reste tout de même un aspect négatif de la méthode. L'automatisation de ces interventions a posteriori offre donc des perspectives en vue de l'amélioration de notre méthodologie.



TAB. 4.1.: Méthodologie d'analyse : Apprentissage



TAB. 4.2.: Méthodologie d'analyse : Classification

Deuxième partie .

Applications industrielles

Résumé

Dans cette partie, nous appliquons à des surfaces réelles les méthode développées précédemment, et nous envisageons un système automatique de contrôle en ligne de la qualité des surfaces.

- Nous classons en laboratoire des surfaces de type "peau d'orange" destinées à l'industrie de l'électroménager.
- Nous choisissons un capteur pour une application en ligne.
- Nous validons ce choix et présentons un premier prototype.

Chapitre 5.

Classification de surfaces "peau d'orange"

Dans les chapitres précédents, nous avons passé en revue l'ensemble des outils de caractérisation, d'analyse et de classification dont la synthèse nous a permis de mettre au point une méthodologie globale d'étude des surfaces rugueuses. Nous présentons maintenant l'aboutissement de ces travaux au travers d'une nécessaire validation expérimentale [25, 28]. On peut même parler de véritable application industrielle, puisque toutes les étapes depuis les premiers tests en laboratoire jusqu'à la mise en place d'un prototype sur ligne de production ont été franchies avec succès. Il faut d'ailleurs préciser que l'ensemble des mesures et essais physiques ont été réalisés au sein d'Arcelor Research, soit sur le site de Maizières-lès-Metz, soit sur l'un des sites de production du groupe Arcelor (usine de Montataire en particulier).

Notre objectif final a toujours été le contrôle automatique *en ligne* de la texture des surfaces produites, mais pour arriver à un tel résultat, il faut d'abord se placer dans des conditions "idéales", c'est à dire chercher à caractériser les surfaces à partir d'une représentation la plus fidèle et la plus informative possible. Dans notre cas, la texture est entièrement déterminée par la topographie de la surface étudiée qui peut être mesurée très précisément par microscopie interférométrique. La première étape de validation consiste donc à vérifier que nos outils nous permettent de classer de manière pertinente un ensemble de surfaces test sur la base de mesures topographiques effectuées en laboratoire, et nous présentons ces résultats "témoins" dans une première partie. Vient ensuite le constat qu'une telle mesure de topographie est impossible en ligne, et qu'il faut donc envisager un nouveau capteur. La recherche d'une technologie adaptée à l'étude non destructive d'un produit défilant à une vitesse de 150m/mn est en soi un problème complexe et plusieurs solutions ont été envisagées avant d'arrêter un système de vision compatible avec ces contraintes industrielles. Nous présentons dans un second temps une partie des considérations et des essais qui ont mené au choix d'un capteur de type "source/caméra matricielle" dont la configuration optique a ensuite été optimisée de façon à obtenir les meilleurs résultats possibles malgré la dégradation de l'information disponible due au passage d'une mesure topographique à une image caméra (qui n'est que le rendu de la surface). Nous effectuons alors une analyse complète menée à partir de telles images obtenues en laboratoire sur le *même jeu de surfaces* que nous avons utilisé pour valider nos outils à partir de mesures topographiques. Le succès de cette étude conditionnait le passage à l'étape suivante, c'est à dire l'élaboration d'un prototype pouvant être testé en situation réelle sur ligne de production : les premières acquisitions et les premiers résultats ainsi obtenus feront l'objet d'une quatrième partie. Enfin, nous présentons suc-

cinctement d'autres contextes, domaines, produits et applications dans lesquels les outils que nous avons développés semblent donner ou promettre de bons résultats.

5.1. Base de surfaces utilisée, contexte industriel

Au cours de l'ensemble des campagnes de tests de classification, nous nous sommes focalisés sur une catégorie de surfaces particulière : les surfaces pré-peintes de type "peau d'orange" destinées au domaine de l'électroménager. Ces surfaces sont volontairement texturées pour ce type de marché, mais l'aspect exact recherché peut varier en fonction des clients (peau d'orange plus ou moins marquée). Les ondulations visibles sur la figure 5.1 sont obtenus par la réticulation d'un primaire, en contrôlant l'épaisseur de cette première couche de peinture.

La méthode actuelle de quantification de la qualité de l'aspect de surface se base sur une inspection visuelle d'échantillons. En règle générale le client et le producteur d'acier établissent un échantillon "cible", un échantillon "très structuré" qui sert de limite haute et un échantillon "plus lisse" qui sert de limite basse. Sur ligne de production, on prélève régulièrement des échantillons dont l'aspect est visuellement comparé à celui des échantillons de référence, ce qui permet de suivre l'évolution de la texture des surfaces produites. Ce système de contrôle permet de tirer la sonnette d'alarme en cas de catastrophe, mais il n'assure pas un suivi continu et complètement objectif de la qualité de l'aspect des surfaces. Une telle procédure peut aussi induire une perte de temps importante lors du démarrage d'un cycle de production, lorsqu'il faut adapter les réglages du processus de réticulation à l'aspect souhaité par le client.

Le choix de ces surfaces particulières a été initiée par une étude parallèle menée dans l'optique de remplacer l'inspection visuelle des échantillons prélevés par l'utilisation d'un capteur de type "wave-scan". Malheureusement, les tests de pertinence et de reproductibilité effectués avec ce capteur n'ont pas donné les résultats escomptés à cause des fortes pentes des surfaces considérées qui se distinguent nettement sur ce point des surfaces destinées à l'industrie automobile pour lesquelles avait initialement été conçu le capteur "wave-scan". Par ailleurs, le capteur "wave-scan" ne permet qu'une mesure limitée impossible à mettre en oeuvre dans un système dynamique. D'une part, il nous a paru intéressant de placer nos essais dans le cadre de cette étude pour laquelle une base de surfaces classées par un expert était disponible. D'autre part, l'échec du "wave-scan" montrait que ce problème de classification constituait pour notre méthode une épreuve intéressante et à fort enjeu industriel.

La base de référence utilisée comporte 25 surfaces, numérotées de 1 à 25 issues de l'usine de La Magona. Parmi ces surfaces, 6 ont été déclarées d'aspect irrecevable par un expert "client" (1,2,13,16,22 et 23), tandis que les autres surfaces ont été déclarées d'aspect correct. Cependant, l'expérience et les avis complémentaires d'autres personnes compétentes à Arcelor Research nous ont amené à raffiner cette classification en déclarant les surfaces 14, 15, 19 et 16 d'aspect "moyen". Ces 4 surfaces se ressemblent beaucoup, notamment si l'on mesure leur Wa qui semble être l'un des paramètres classiques de rugosité se corrélant le mieux avec la qualité estimée des surface. Elles se trouvent en

quelque sorte à la frontière entre l'aspect correct et incorrect.

5.2. Classification à partir de topographies

Dans cette section, nous nous intéressons aux cartographies des surfaces de la base de référence (voir figure 5.1) obtenues grâce au microscope interférométrique d'Arcelor Research (Maizières-lès-Metz). Ces cartographies fournissent des images 8 bits (256 niveaux de gris) de tailles 555×356 pixels représentant un champs d'observation de taille $20,9 \times 15,8$ mm (Recollement de 4×4 acquisitions). La résolution spatiale est donc de l'ordre de 0.040 mm et la résolution verticale de l'ordre de 0.001 mm. Ces conditions métrologiques ont été choisies en fonction de la taille des motifs que l'on cherche à analyser et font suite à une série de tests de répétabilité et de dispersion de la mesure. Dans ce qui suit, les topographies obtenues seront redécoupées en 4 sous-images. La première sous-image de chaque surface s'inclut dans une base d'apprentissage destinée à l'entraînement de notre classificateur, tandis que les trois sous-images restantes seront utilisées pour tester la procédure de classification. Dans un premier temps, nous passons en revue chacun des descripteurs du chapitre 2 pour étudier en détails l'analyse factorielle qui lui est associée, son interprétation en terme de facteurs principaux, et l'utilisation du bootstrap pour détecter d'éventuels phénomènes de sur-apprentissage.

Si la division des topographies est le seul moyen de véritablement appréhender la pertinence de notre méthodologie et de l'utilisation du bootstrap en particulier, il s'ensuit néanmoins une perte de qualité dans les descripteurs disponible et dans la classification. Tout d'abord, la réduction des dimensions des surfaces de la base va entraîner une augmentation de la variance d'extension (qui varie de façon inversement proportionnel au nombre de pixels de l'image), ensuite, certains descripteurs comme les courbes de covariance ou les analyses en ondelettes ne pourront accéder à des longueurs d'ondes trop grandes par rapport à la taille des images étudiées, ce qui entraîne une dégradation de l'information disponible. D'autres essais (non présentés ici) ont été réalisés en utilisant des topographies complètes pour la phase d'apprentissage et la classification de surfaces différentes ne faisant pas partie de la base d'apprentissage et les résultats obtenus sont compatibles (bien que plus précis sur certains points) avec ceux qui seront décrits dans les sections suivantes.

5.2.1. Etude des différents descripteurs

Pour chacun des descripteurs de surface considéré, nous présentons deux séries de graphiques issues de la première étape d'analyse factorielle. La première série contient la projections des 25 surfaces de la base dans le premier plan factoriel, c'est à dire dans un espace de dimension 2. Les échantillons corrects sont représentés en noir, et les échantillons incorrects en rouge. Les pourcentages d'inertie associé aux deux premiers axes propres sont mentionnés sur la figure. Nous présentons ensuite la projection dans le même espace de l'ensemble des 400 échantillons bootstrap ($x_0 + \mu^*$) associés à chacune des surfaces de la base (les échantillons bootstrap associés à une surface incorrecte sont en bleu clair, et les autres en vert). Ces échantillons bootstrap simulent la dispersion

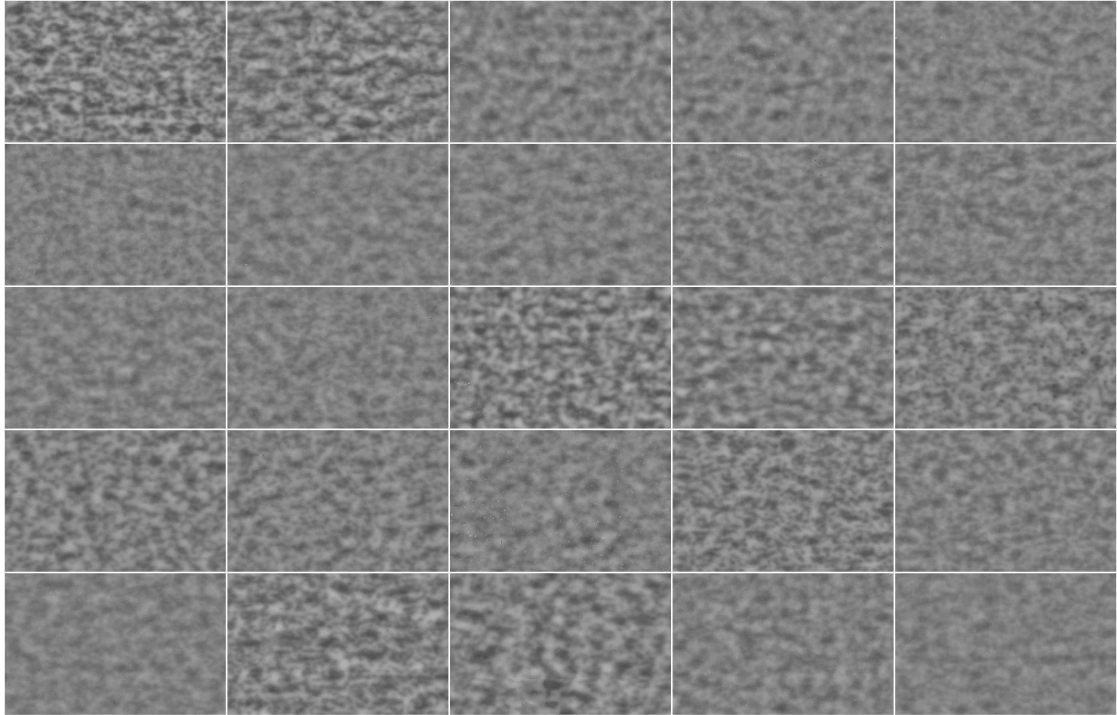


FIG. 5.1.: Topographies de 25 surfaces "peau d'orange". Les surfaces sont ordonnées de la gauche vers la droite puis de haut en bas. Les surfaces 1,2,13,22,23 sont d'aspect "incorrect", les surfaces 14,15,16,19 sont d'aspect "intermédiaire".

naturelle au sein des réalisations d'une même fonction aléatoire. Pour vérifier la pertinence de cette simulation, nous projetons dans le même espace et pour chaque surface originale les 4 sous-surfaces issues de la même cartographie (les 25 sous-surfaces de la base d'apprentissage et les 75 autres sous-surfaces qui serviront à tester la pertinence du classificateur). Les sous-surfaces incorrectes sont toujours représentées en rouge, et se trouvent donc logiquement à l'intérieur du nuage (bleu clair) des échantillons bootstrap incorrects, et de la même façon, les points noirs sont les projections des sous-surfaces correctes et doivent donc être plongé dans le nuage vert des échantillons bootstrap correspondant. Le troisième graphique de la première série représente, pour chaque individu de la base d'apprentissage le cosinus carré de l'angle formé par x_i et $\hat{x}_i$ si l'espace de projection est simplement le premier axe propre (courbe rouge), le second axe propre (courbe bleu), ou le plan formé par les deux premiers axes (courbe noir). Ces valeurs caractérisent la qualité de la projection pour chaque individu.

La deuxième série de graphique montre tout d'abord l'interprétation des deux premiers facteurs principaux. Chaque facteur principal est obtenu comme combinaison linéaire des variables de départ : le premier graphique associe à chacune d'entre elles le coefficient qui lui est attribué dans le calcul du premier facteur principal (courbe rouge) et du second (courbe bleue). Ce graphique permet donc d'interpréter la signification des facteurs principaux, c'est à dire celle des axes de coordonnées de l'espace de projection choisi en fonction du descripteur original. Pour mieux illustrer l'interprétation de ces axes, le deuxième graphique présente les descripteurs de l'espace de départ qui, projetés sur le premier axe propre, fournissent la valeur minimale (en vert) et maximale (en noir). Le troisième graphique donne la même information concernant le second axe propre.

Ces deux séries de graphiques vont nous permettre de détailler et d'illustrer l'usage de l'analyse de donnée pour chaque descripteur.

a. Courbes de covariance

Le premier facteur principal est ici obtenu par une combinaison linéaire des valeurs de la fonction de covariance de l'image initiale qui attribue un fort poids aux premiers points de la courbe (voir figure 5.3(b)). Cela signifie que les individus ayant une forte¹ valeur sur le premier facteur principal sont les surfaces pour lesquelles $C(h)$ est élevé pour les faibles valeurs de h . En particulier, ces surfaces ont une forte variance : les pics et les vallées sont fortement contrastés. La figure 5.3(a) montre que les individus se positionnant sur la "droite" du plan de projection (forte valeur sur l'axe 1) sont principalement les individus incorrects, qui sont donc globalement plus "structurés" que les individus corrects.

Le second facteur principal au contraire met en avant les surfaces présentant des oscillations pour les valeurs de h de l'ordre de 40 pixels ($\simeq 1.5mm$). En réalité, les valeurs de la courbe de covariance pour les fortes valeurs de h sont moins fiables que pour les faibles valeurs. Les oscillations observées pour les valeurs de h supérieure à 30 pixels ($\simeq 1.1mm$) laissent à penser que la portée réelle de la fonction aléatoire est inférieure à 40 pixels

¹Le terme "fort", ici comme dans la suite, est employé par comparaison à *la moyenne* de la quantité considérée pour l'ensemble des surfaces de la base, car leur barycentre est l'origine de l'espace.

($\simeq 1.5mm$) : le second axe propre met donc l'accent sur des fluctuations statistiques particulières aux *réalisations* utilisée dans la base d'apprentissage : il s'agit d'un phénomène de sur-apprentissage. Le troisième graphique de la figure 5.3(b) montre bien la faible différence de covariance entre les surfaces prenant les valeurs extrêmes sur l'axe 2. Les échantillons bootstrap mettent aussi en évidence ce phénomène de sur-apprentissage : les nuages sont très étendus sur l'axe 2, et ne permettent pas de distinguer les surfaces. De la même façon, les faibles cosinus carré entre les individus et leur projection sur l'axe 2 révèlent le même phénomène.

Le résultat de cette analyse sera donc la sélection du premier facteur principal pour une étude ultérieure (seconde étape d'analyse factorielle), mais pas du second.

b. Courbes d'érosion-dilatation

L'analyse sur les courbes d'érosion-dilatation ne fait pas apparaître de phénomènes de sur-apprentissage aussi flagrant que dans le cas des courbes de covariance. La superposition des échantillons bootstrap et des individus réels montre la pertinence de la méthodologie : la dispersion de réalisations indépendantes issues d'une même fonction aléatoire est très comparable à la dispersion des échantillons bootstrap obtenus à partir d'une seule réalisation.

Le premier facteur principal met en valeur les surfaces ayant de faibles valeurs de vitesse d'érosion ou de dilatation pour des éléments structurants de petite taille, mais des fortes vitesses pour les éléments de grande taille, c'est à dire supérieure à 20 pixels soit $0.75mm$ de rayon (voir figure 5.5). C'est le cas de la courbe noire sur le second graphique de la figure 5.4(b). Ici, les surfaces ayant une forte valeur sur l'axe 1 sont plutôt les surfaces correctes, même si la différence est moins nette que dans le cas des courbes de covariance), dont l'une des caractéristique est donc la présence de beaucoup de structures spatiales de grandes dimensions et de peu de structures de faibles dimensions, que ces structures soient positives (pics) ou négatives (vallées). Le second facteur principal est formé par une combinaison linéaire associant des poids dissymétriques aux courbes de vitesse d'érosion et de dilatation : les vitesse d'érosion par des éléments structurants de grande taille ont un fort poids positifs, les vitesse d'érosion par des éléments structurant de taille moyenne et petite (entre 35 pixels ($\simeq 1.3mm$) et 6 pixels ($\simeq 0.22mm$) de rayon) ont des poids négatifs, et les vitesse d'érosion pour les très petits éléments structurant ont des poids positifs. Les vitesses de dilatation ont des poids plus faibles et donc moins d'importance dans le calcul de ce facteur, avec malgré tout une tendance à mettre en valeur les fortes vitesses de dilatation pour de petits et très grand éléments structurants. En résumé, une surface se projetant loin sur le second axe propre présente beaucoup de pics de très petite taille et de grande taille, mais peu de structures de dimensions intermédiaires. Les structures négatives (vallées) jouent peu de rôle dans le positionnement sur l'axe 2.

c. Histogramme des hauteurs de zones

La projection des surfaces sur le premier axe propre de l'analyse traitant les histogrammes des hauteurs de zones permet de bien séparer les surfaces correctes et incor-

rectes. Par contre, la forte dispersion des nuages bootstrap sur l'axe 2 (par rapport à la dispersion des individus de la base d'apprentissage) suggère un phénomène de sur-apprentissage. Ceci se confirme par la forte irrégularité des coefficients de la combinaison linéaire formant le second facteur principal (figure 5.6(b), premier graphique), et par la faible différence entre les histogrammes des surfaces prenant les valeurs extrêmes sur cet axe (figure 5.6(b), troisième graphique). Nous n'interpréterons donc pas le second axe principal de cette analyse, ni ne l'utiliserons dans l'analyse globale.

Le premier facteur principal, en revanche semble très pertinent : les surfaces qui prennent une forte valeur sur cet axe sont les surfaces ayant peu de zones de hauteur moyenne, mais beaucoup de zones de très faible ou de très forte altitude, ce sont donc des surfaces pour lesquels les transitions entre pics et vallées sont bien marquées (voir figure 5.7), et assez naturellement, ce sont les surfaces incorrectes qui ont tendance à prendre une forte valeur sur cet axe.

d. Histogramme des surfaces de zones

Cette analyse est certainement celle qui fournit le plus beau phénomène de sur-apprentissage sur les deux axes : les pourcentages d'inertie sont faibles pour les deux premiers axes propres, les facteurs principaux n'ont pas d'interprétation claire, les cosinus carré sont faibles et les nuages d'individus bootstrap sont fortement dispersés : l'histogramme des surfaces de zones n'est donc pas un descripteur pertinent pour différencier les surfaces de type "peau d'orange" à partir de leur topographie.

e. Histogramme des tortuosités

L'histogramme des tortuosités permet assez bien de distinguer certaines surfaces incorrectes des surfaces correctes, il n'y a pas de phénomènes de sur-apprentissage flagrant. Les histogrammes des tortuosités sont suffisamment différents d'une surface à l'autre pour avoir des supports pour ainsi dire disjoints. L'analyse distingue ici simplement les surfaces en fonction de ce paramètre : le premier axe propre projette avec une forte valeur positive les surface dont le support de l'histogramme des tortuosités est entièrement porté par les classes 7 à 15 correspondant à des tortuosités intermédiaires. Il projette aussi avec une forte valeur négative les surfaces dont le support se trouve entre 1 et 7. Le second axe propre projette positivement les surfaces dont le support est entre 4 et 10 et négativement celle dont le support est entre 1 et 4. On voit d'ailleurs que la surface ayant la plus faible valeur sur ces deux axes (figure 5.10(b)) est la surface 6 dont le support de l'histogramme est entièrement porté par les 4 premières classes (toutes les tortuosités sont faibles). Aucun des deux facteurs principaux ne fait apparaître de phénomène flagrant de sur-apprentissage, mais l'observation du second graphique de la figure 5.9(a) montre que l'axe 1 permet de faire une différence nette entre les surfaces correctes et incorrectes tandis que l'axe 2 superpose l'ensemble des surfaces incorrectes avec des surfaces correctes. Dans un objectif de classification, il paraît donc plus judicieux de ne pas tenir compte de ce second facteur principal dans l'analyse globale.

f. Histogramme des pentes

L'histogramme des pentes est l'un des descripteurs pour lesquels les nuages correspondant aux 4 individus issus d'une même topographie sont le plus disjoints (figure 5.10(a), deuxième graphique). Là encore, la dispersion des échantillons bootstrap reflète assez bien ce phénomène. On observe tout de même un "accident" : un point représentant la réalisation d'une surface réelle ne faisant pas partie de la base d'apprentissage (il s'agit de l'une des sous-surfaces de la topographie 23) se trouve très loin de l'ensemble des individus bootstrap obtenus à partir de la surface correspondante de la base d'apprentissage. Ce phénomène assez déroutant pour la pertinence du bootstrap trouve une explication simple dans l'observation de la topographies originale. La topographie 23, comme les autres provient d'un microscope interférométrique dont le champs d'observation est bien inférieure à la taille $20,9 \times 15,8$ mm annoncée en début de chapitre : pour obtenir cette taille finale, on a recours à un logiciel de "stitching" qui "recolle" l'information provenant d'acquisitions sur des champs différents. Le résultat est en général très bon, et le collage est invisible à l'oeil nu. Il arrive malgré tout que le logiciel "se trompe", et que la bande de recollement soit bien visible. C'est le cas de la surface 23 sur l'une des sous-surface, mais pas sur celle utilisée dans la base d'apprentissage (voir figure 5.2). Cette irrégularité dans le calcul des pentes ne peut être reproduite par bootstrap puisque le phénomène n'est pas présent dans la réalisation utilisée pour la base d'apprentissage, et cela explique la position inattendue de l'une des réalisations dans l'espace de projection.

L'axe 1, d'après le premier graphique de la figure 5.10(b), projette loin positivement les surfaces possédant beaucoup de pentes moyennes et au contraire peu de fortes et de très faibles pentes, tandis que l'axe 2 projette loin positivement les surfaces ayant beaucoup de très fortes pentes et peu de pentes moyennes. Comme pour les courbes de tortuosités, les valeurs prises par les surfaces sur le second facteur principal ne permettent pas de discriminer correctement surfaces correctes et incorrectes ; nous n'utiliserons donc que le premier facteur principal dans l'analyse globale.

g. Histogramme des orientations des vecteurs normaux

Là encore, on détecte assez bien un léger phénomène de sur-apprentissage sur les deux premiers axes propres, surtout, les "pics" dans les coefficients de la combinaison linéaire utilisée pour produire le premier facteur principal sont assez singuliers et méritent qu'on s'y attarde. De tels pics dans les histogrammes caractérisent des surfaces dont la proportion de vecteurs orientés parallèlement à l'un des bords de l'image est forte. Tout suggère qu'il s'agisse là encore de problèmes de "stitching" (pourcentage de recollement insuffisant) : si l'algorithme ne fait pas coïncider parfaitement les différentes acquisitions du microscope, on verra apparaître de légères discontinuités le long des zones de recollement (parallèles à l'image). Ces discontinuités vont donc induire des vecteurs normaux principalement orientés orthogonalement aux lignes de stitching et générer des pics dans l'histogramme des orientations des vecteurs normaux. Les surfaces se projetant avec une forte valeur sur l'axe 1 sont donc les surfaces de la base d'apprentissage pour lesquelles l'algorithme de stitching a été le moins efficace. Cette caractéristique n'est pas vraiment

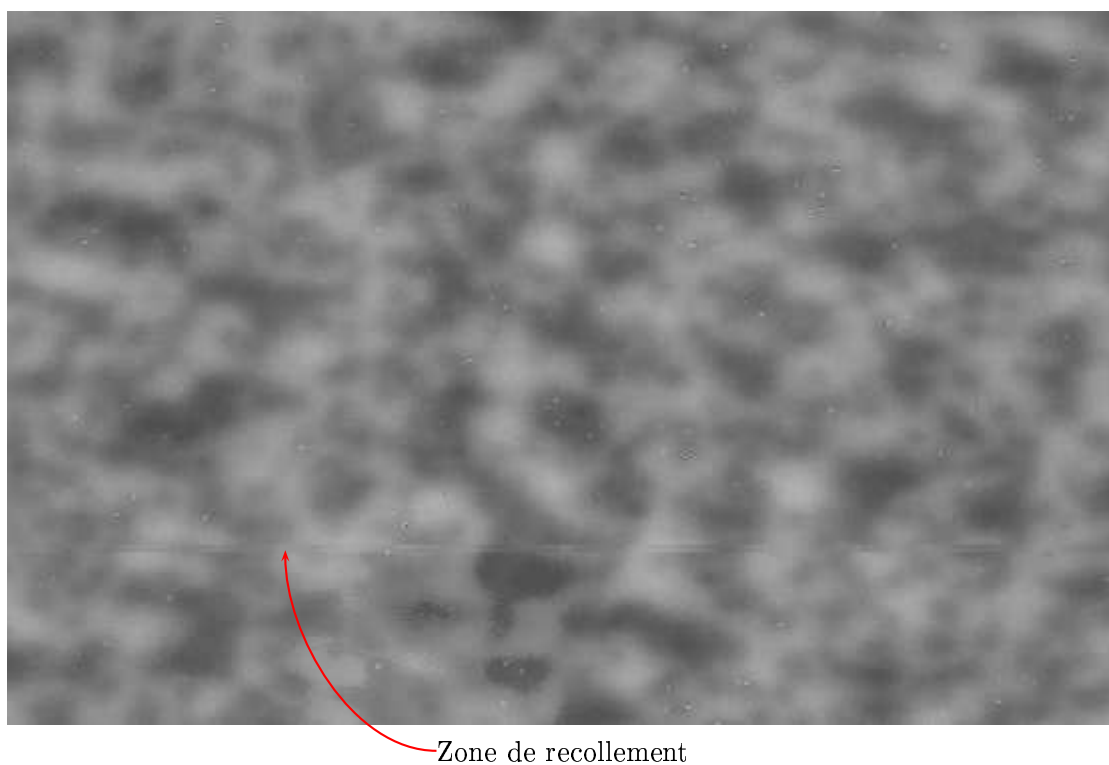


FIG. 5.2.: Topographie de la surface n° 23 : on voit nettement une ligne de "stitching"

pertinente pour la classification des surfaces correctes et incorrectes, ce qui nous amène à ne pas retentir ce descripteur pour l'analyse finale.

h. Courbes de Ma

La courbe de Ma caractérise en quelque sorte la répartition de l'énergie du signal entre différentes longueurs d'onde. Ici, le point $i \in [1..32]$ de la courbe de Ma correspond à un facteur d'échelle de $2^{1+i/8}$, c'est à dire une longueur d'onde réelle de $0.034 \times 2^{1+i/8} mm$. Le premier graphique de la figure 5.12(b) montre qu'une forte valeur sur le premier facteur principal est synonyme d'une forte énergie dans toute la gamme de longueur d'ondes étudiées, tandis qu'une forte valeur sur le second facteur principal caractérise des surfaces ayant une forte énergie (un fort Ra) pour la gamme de longueurs d'onde comprises entre 0.5 et 1mm, mais une faible valeur pour la gamme 0.07/0.42 mm.

Encore une fois, si l'interprétation des axes propres est tout à fait claire et intéressante, la projection sur le plan factoriel final n'introduit pas de division franche entre surfaces correctes et incorrectes, et ces variables ne seront pas conservées dans l'analyse globale. En fait, les analyses de des courbes de Ma sont plus intéressantes si l'on peut obtenir de l'information sur de plus grandes longueurs d'ondes. Dans ce cas précis malheureusement, la division de la topographie en quatre sous-surfaces réduit d'autant les longueurs d'ondes admissibles, puisque le support de l'ondelette utilisée doit être de taille inférieure à celle du profil étudiée.

5.2.2. Analyse globale

La section 5.2 a permis d'étudier en détail les résultats des analyses factorielles spécifiques à chaque descripteur. Nous allons maintenant réutiliser dans une analyse globale le petit nombre de variables (0, 1 ou 2) sélectionnées pour chaque descripteur. On pourrait aussi réinjecter à ce niveau des variables monodimensionnelles (Ra ou Wa par exemple, voir 2.1), qui viendraient simplement s'ajouter aux variables composites issues de descripteurs plus complexes. Dans ce qui suit, nous nous contenterons d'une analyse globale sur le premier facteur principal de l'analyse des courbes de covariance, celui de l'analyse des histogrammes des tortuosités, des histogrammes des hauteurs de zones et des histogrammes des pentes. Nous utiliserons également les deux premiers facteurs principaux de l'analyse des courbes d'érosion-dilatation, ce qui forme un total final de 6 variables. Comme pour les analyses spécifiques, le second graphique de la figure 5.13(a) montre la projection de l'ensemble des échantillons bootstrap sur les deux premiers axes propres de l'analyse globale, ainsi que la projection de l'ensemble des sous-surfaces associées à chaque topographie originale. Cette figure fait apparaître trois groupes assez distincts formés par les surfaces correctes, les surfaces incorrectes, et les surfaces dites intermédiaires "situées entre ces deux groupes". Le sous-espace d'analyse choisi est donc a priori tout à fait compatible avec la classification visuelle des surfaces de la base d'apprentissage. Cette pertinence sera confirmée par les résultats de la classification bayésienne que nous présentons à la section suivante.

Pour ce qui est de l'interprétation des axes, les courbes de la figure 5.13(b) sont plus

difficile à interpréter que leur équivalent pour les analyses spécifiques puisque les descripteurs de l'espace de départ sont déjà des variables composites. On peut néanmoins remarquer que les surfaces se projetant loin sur l'axe 1 sont celle prenant à la fois :

- Une forte valeur sur le premier facteur principal de l'analyse des courbes de covariance (forte variance et covariance pour les faibles valeurs de h).
- Une forte valeur sur le premier facteur principal de l'analyse des histogrammes des tortuosités (support de l'histogramme compris entre des valeurs spécifiques).
- Une forte valeur sur le premier facteur principal de l'analyse des hauteurs de zones (fort contraste entre zones hautes et basses).
- Une faible valeur sur le premier facteur principal de l'analyse des histogrammes des pentes (peu de pentes moyennes, beaucoup de pentes fortes et très faibles).

On peut aussi noter une influence légèrement moindre du premier facteur principal issu de l'analyse des courbes d'érosion-dilatation qui joue malgré tout négativement (faible vitesse d'érosion-dilatation pour des éléments de grande taille et forte valeurs pour les éléments structurant de petite taille). Le second facteur principal de l'analyse des courbes d'érosion-dilatation présente lui un poids pratiquement nul dans la construction du premier facteur principal de l'analyse globale.

Par contre, l'axe 2 de l'analyse globale projette loin les surfaces ayant une faible valeur sur les deux premiers facteurs principaux de l'analyse des courbes d'érosion dilatation, et en particulier le second (peu de pics de très grandes ou très petites dimensions).

5.2.3. Résultats de classification

Dans cette section, nous appliquons les méthodes de classification bayésienne décrites en 4.2 dans l'espace obtenu après l'analyse globale du paragraphe précédent.

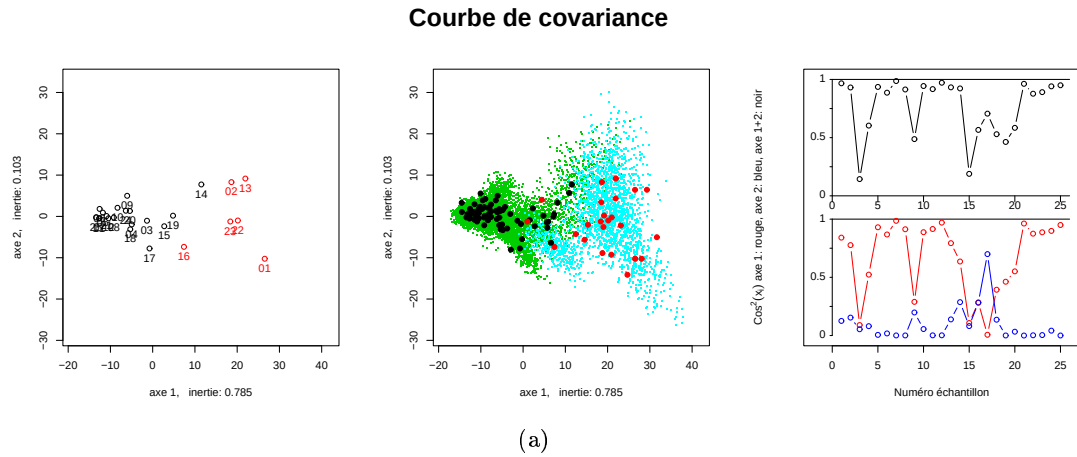
Nos surfaces sont divisées en trois classes : "incorrecte" (surface 1,2,13,22,23), "intermédiaire" (surface 14,15,16,19) et "correcte" (les autres). Chacune de ces classes conduit à une estimation de la densité de probabilité sous-jacente à partir des 400 échantillons bootstrap correspondant à la première sous-surface de chaque topographie originale. Pour toutes les sous-surfaces test, on évalue ensuite les probabilités $P(C = c|X = x)$ (voir le paragraphe correspondant de la section 4.2) pour les trois modalités de C : 1 (incorrecte), 2 (intermédiaire) et 3 (correcte).

a. Classification bayésienne classique

La méthode classique, que nous avons appliquées dans un premier temps, consiste à attribuer à chaque sous-surface la modalité pour laquelle $P(C = c|X = x)$ est maximale. Il est aussi possible de calculer la probabilité d'erreur (égale à $1 - P_{max}$). Ces résultats sont reportés à la figure 5.14(a) : on observe deux fausses classifications, uniquement pour des sous-surfaces normalement intermédiaires (une est classée correcte, et une autre incorrecte), ce qui correspond à un pourcentage d'erreur de 2.6%. La plupart des surfaces sont classées avec une faible probabilité d'erreur, malgré quelques incertitudes (cohérentes avec la position des sous-surfaces dans le graphique 5.13(a)).

b. Estimation de la qualité

Nous utilisons ici la fonction de coût "qualité" décrite au paragraphe 4.3.2. Les valeurs $Q(c)$ associées aux modalités de C sont : 0 pour les surfaces incorrectes, 0.5 pour les surfaces intermédiaires et 1 pour les surfaces correctes. Le calcul des probabilités $P(c|x)$ fournit l'espérance et l'écart-type de la qualité de chaque sous-surface analysée. Les résultats sont présentés à la figure 5.14(b). Cette information est comparable à celle de la figure 5.14(a), avec l'avantage de fournir une information plus continue que la simple donnée d'une classe 1,2 ou 3.



Courbe de covariance: Interprétation des axes

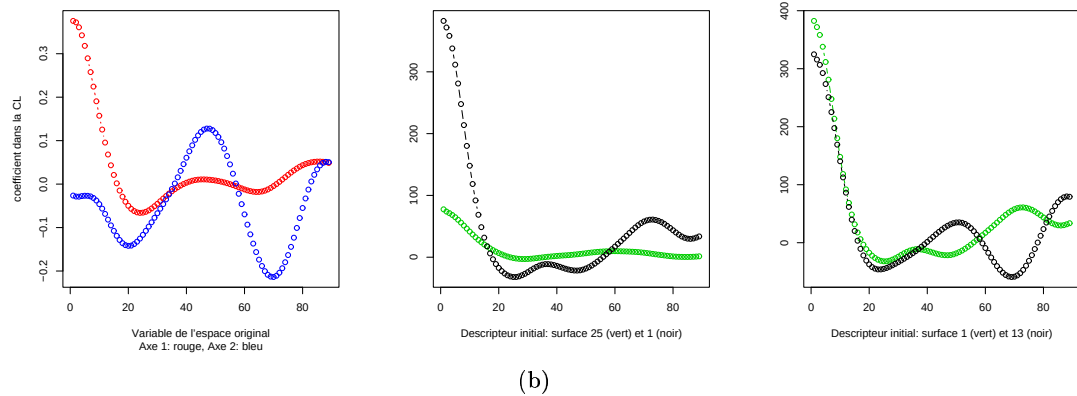
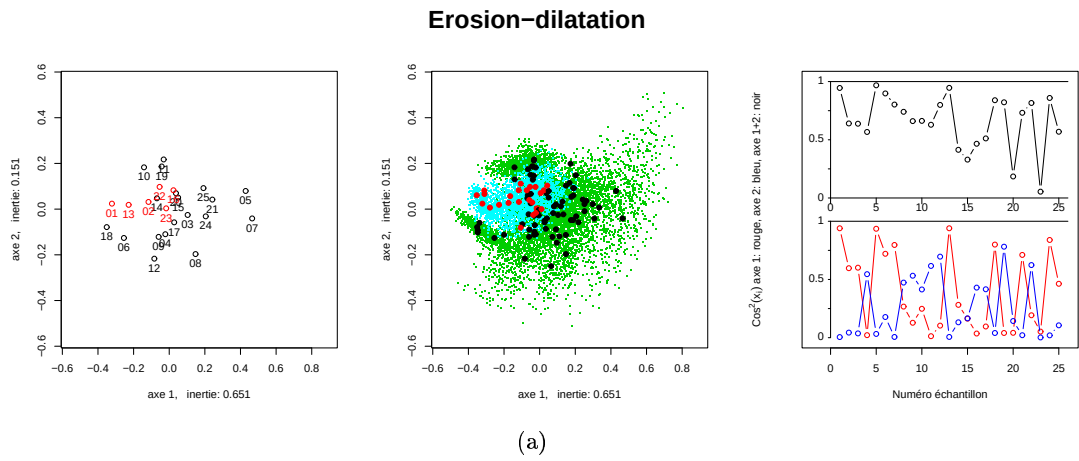


FIG. 5.3.: Covariance



Erosion-dilatation: Interprétation des axes

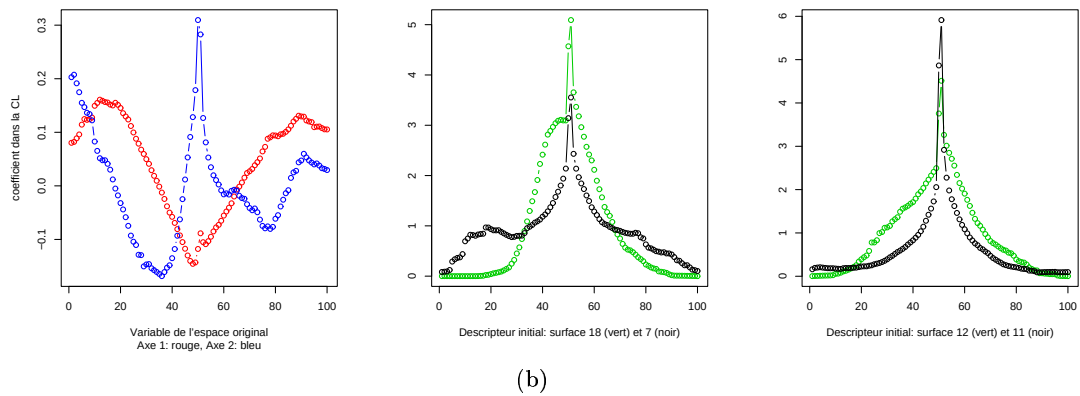
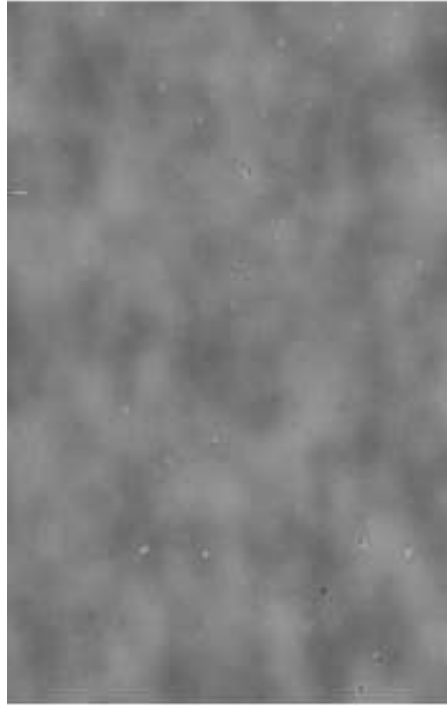


FIG. 5.4.: Erosion-dilatation

Pdo 07



Pdo 18

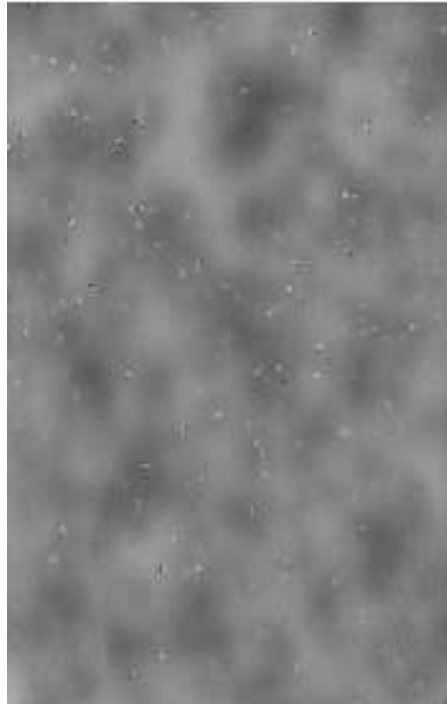
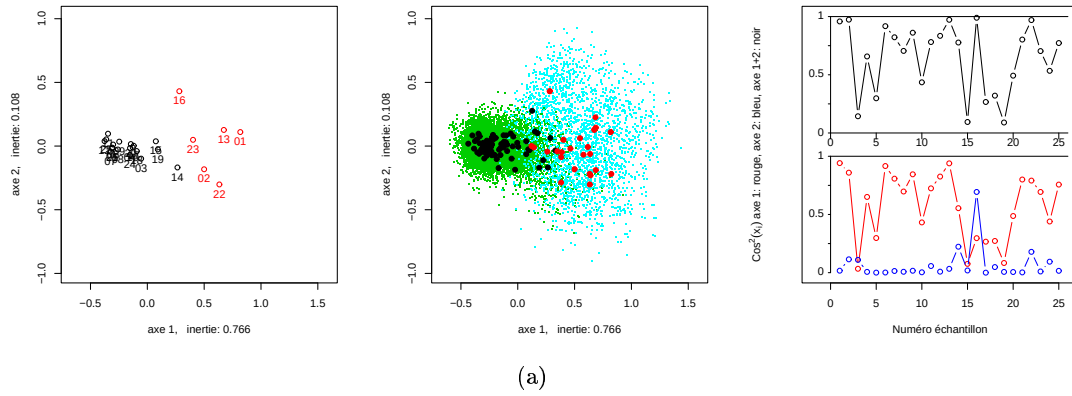


FIG. 5.5.: Images extrêmes pour l'axe 1 des courbes d'érosion-dilatation. La surface 07 possède plus de structures de très grande taille (qui ne disparaîtront qu'avec des éléments structurants de taille $\simeq 20$ pixels), en particulier des pics, mais aussi des vallées. La surface 18 par comparaison possède plus de structures de petites dimensions spatiales.

Histogramme des hauteurs de zones



Histogramme des hauteurs de zones: Interprétation des axes

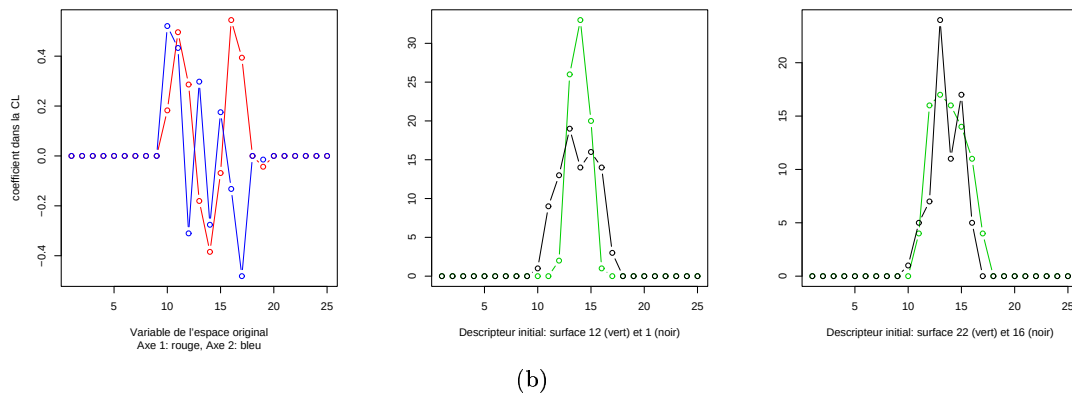
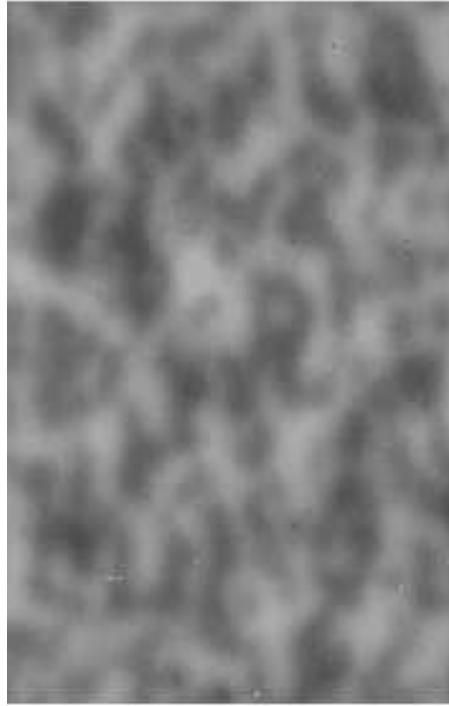


FIG. 5.6.: Hauteurs de zones

Pdo 01



Pdo 12

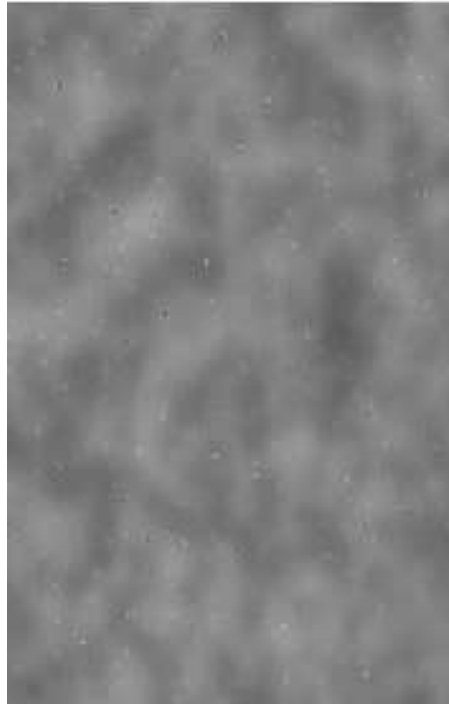
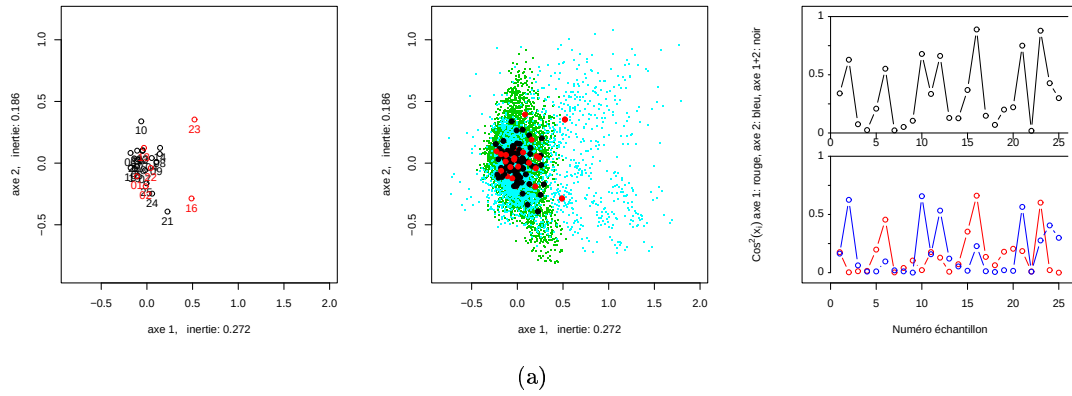


FIG. 5.7.: Images extrêmes pour l'axe 1 des hauteurs de zones. La surface 1 prend une forte valeur sur cet axe, et on devine que les zones extraites de cette image seront plus contrastées (beaucoup de zones de forte ou de faible valeur moyenne) que celles de la surface 12.

Histogramme des surfaces de zones



Histogramme des surfaces de zones: Interprétation des axes

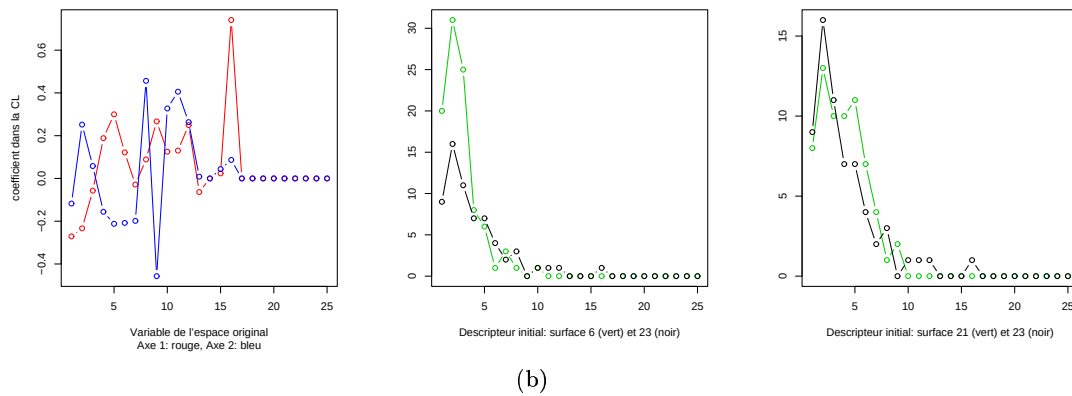


FIG. 5.8.: Surfaces de zones

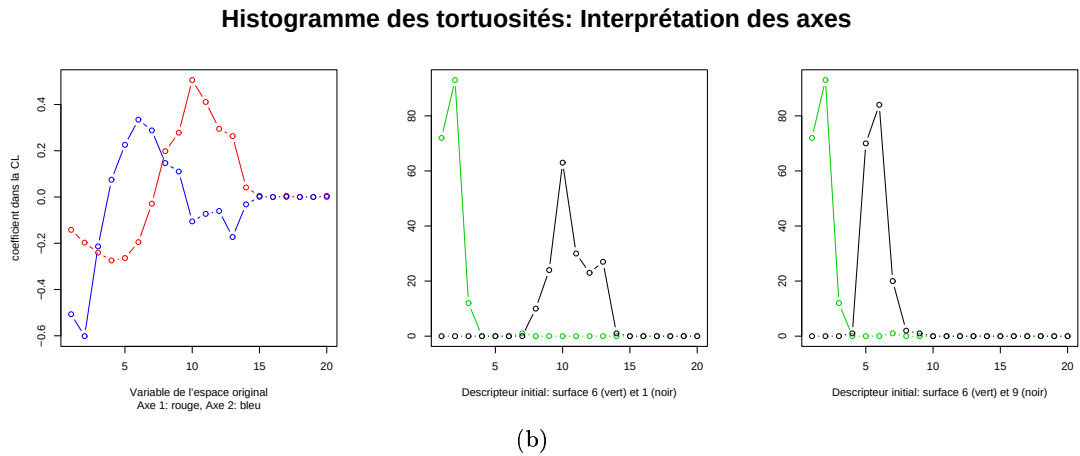
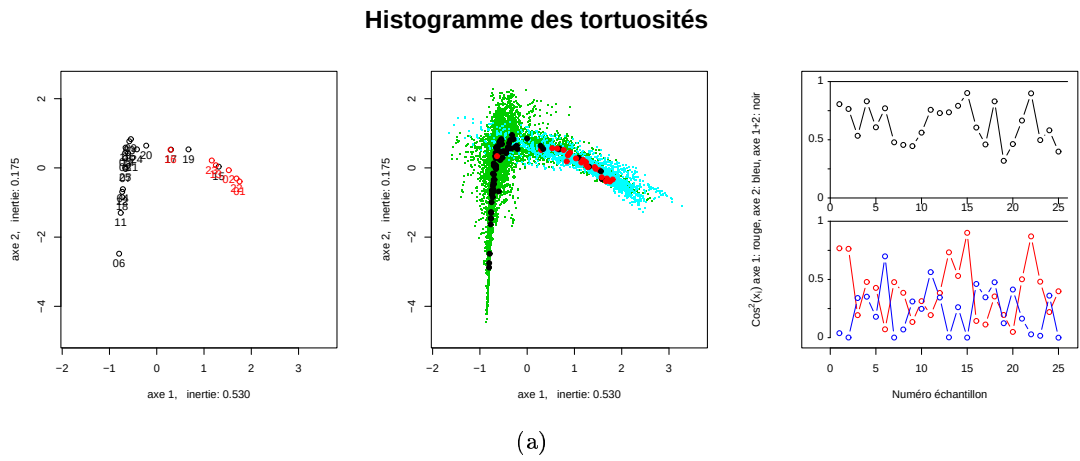


FIG. 5.9.: Tortuosités

5.2. Classification à partir de topographies

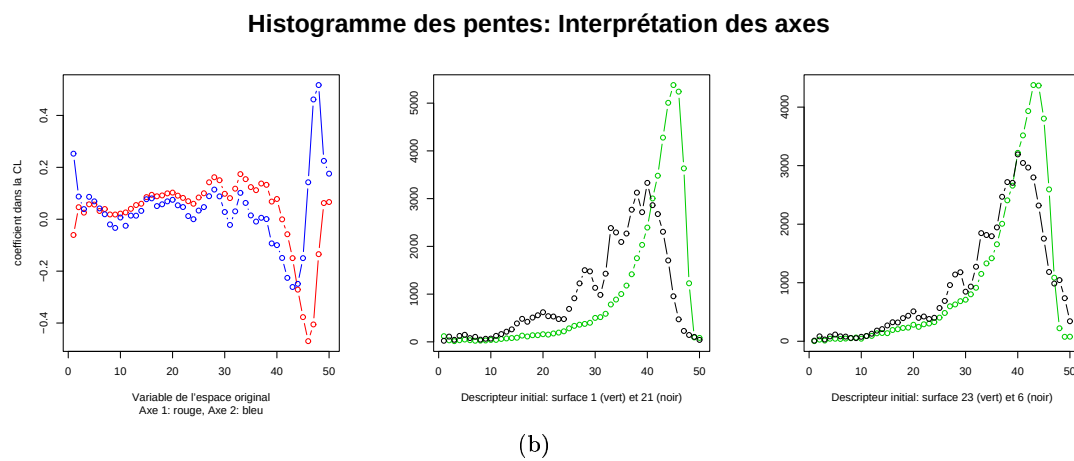
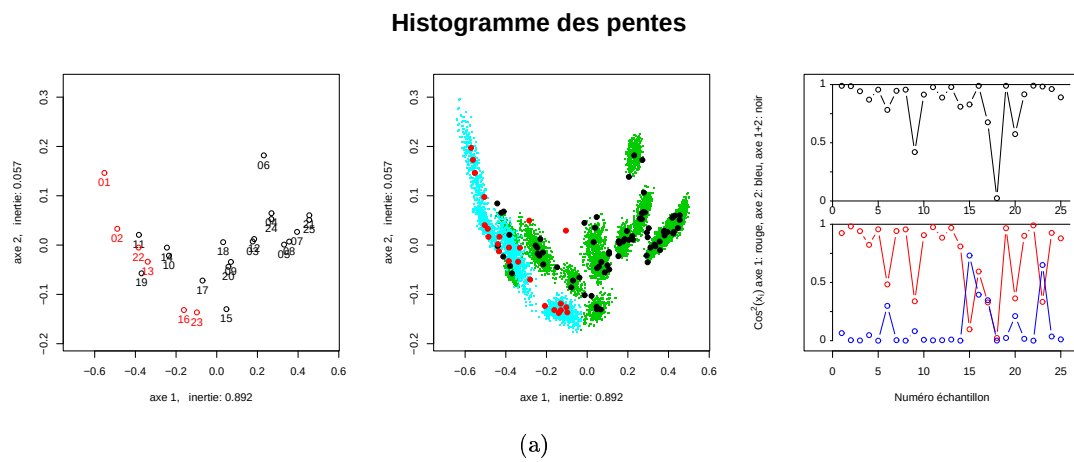
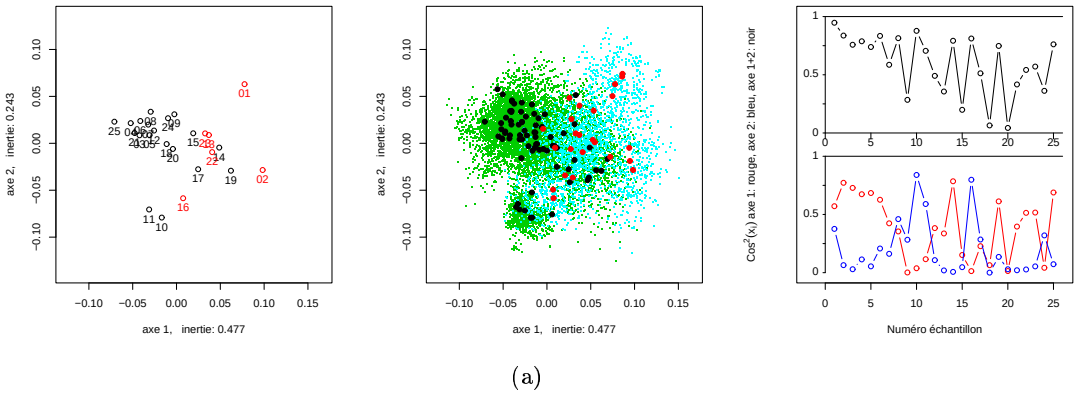


FIG. 5.10.: Pentes

Histogramme des orientations de normales



Histogramme des orientations de normales: Interprétation des axes

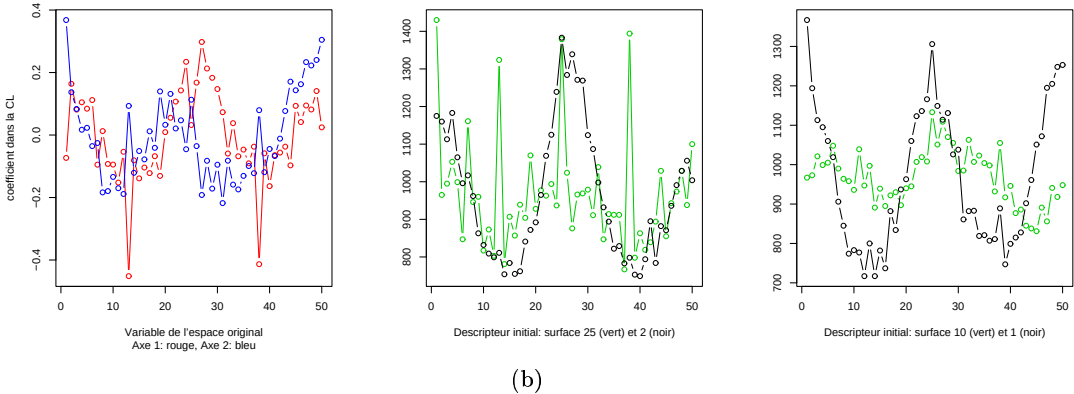
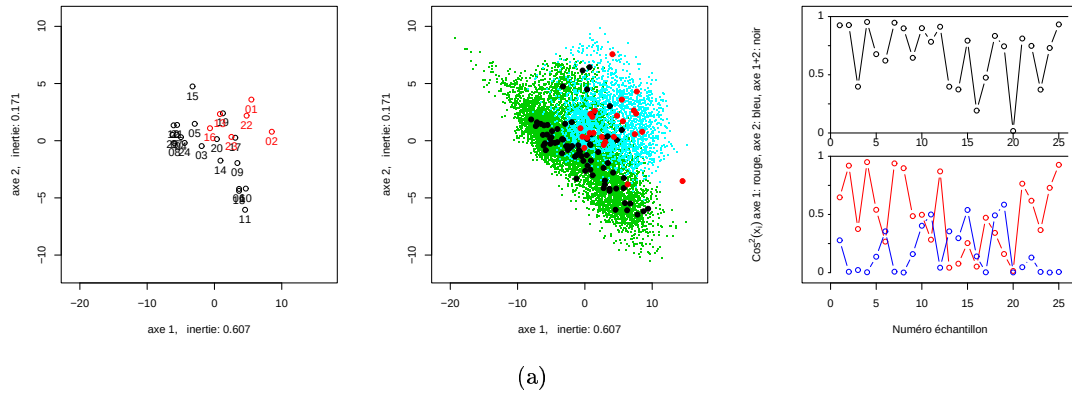


FIG. 5.11.: Orientation des normales

Analyse en ondelette par Ma



Analyse en ondelette par Ma: Interprétation des axes

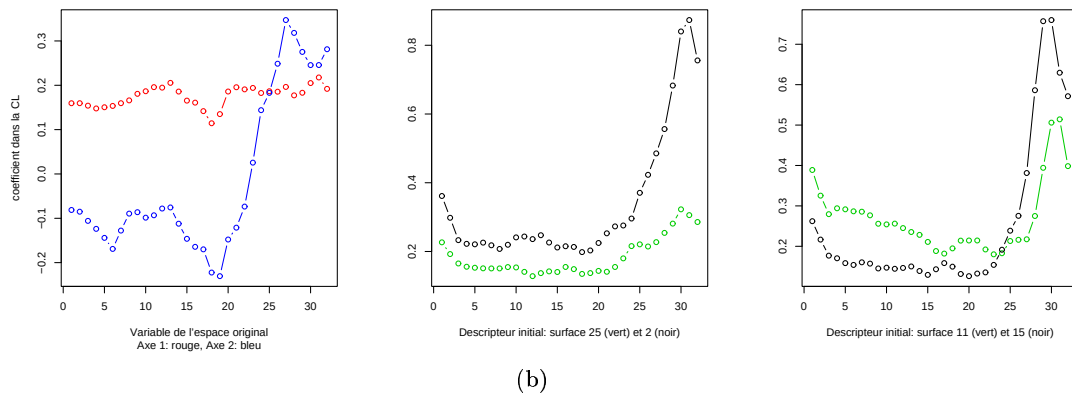
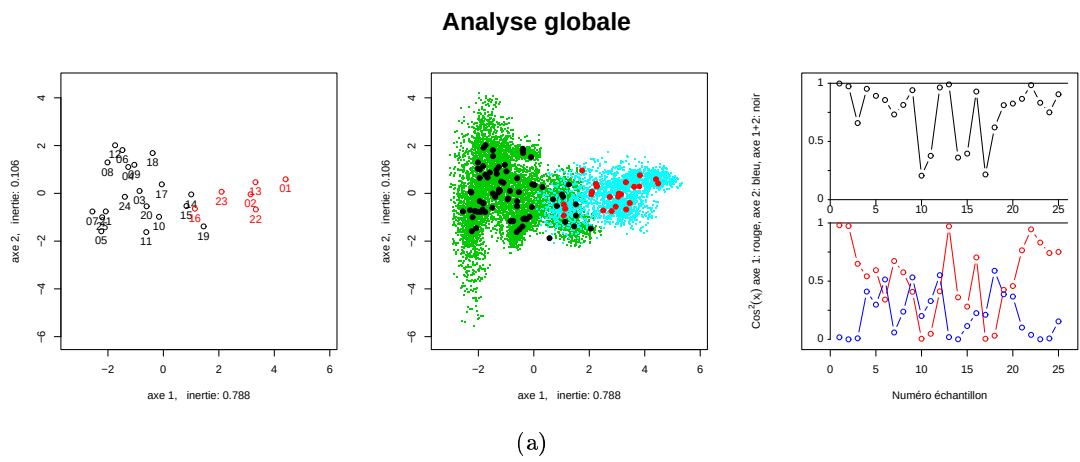


FIG. 5.12.: Courbes de Ma



Analyse globale: interprétation des axes

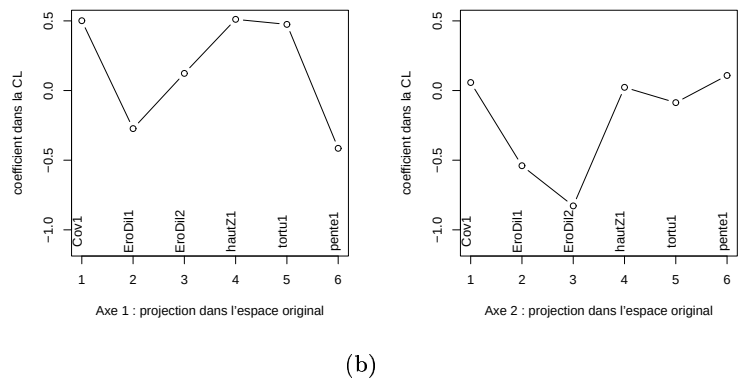
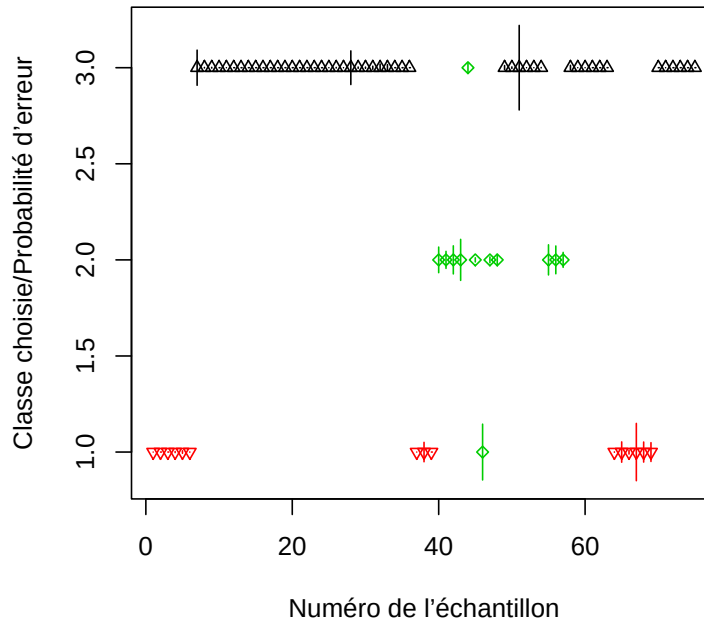
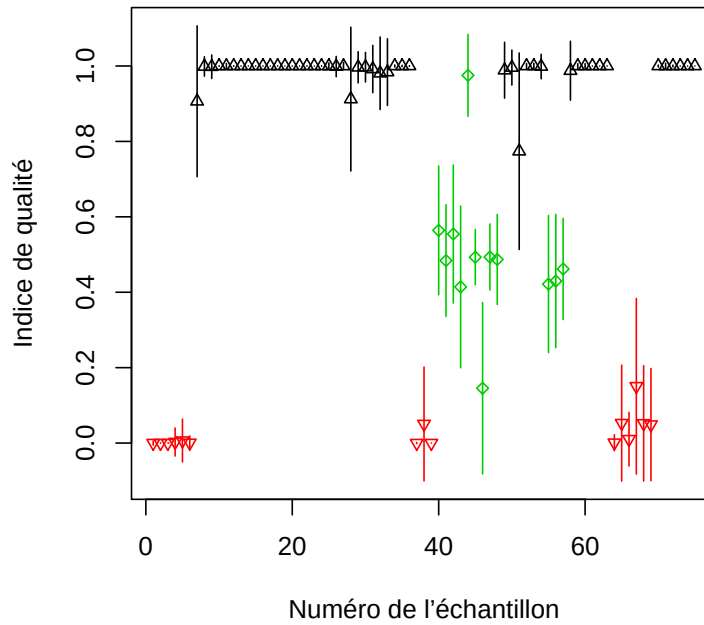


FIG. 5.13.: Analyse globale sur topographies



- (a) Les points rouges représentent les sous-surfaces issues d'une topographie incorrecte, les verts sont les sous-surfaces intermédiaires et les noirs correctes. L'ordonnée représente la classe attribuée à chaque sous-surface par le classificateur bayésien, et la longueur du segment associé à chaque point représente la probabilité d'erreur.



- (b) Pour chaque sous surface, on estime l'espérance de la qualité (position verticale) et son écart-type (longueur du segment associé à chaque point).

FIG. 5.14.: Résultat de classifications à partir de topographies

5.3. Vers une application en ligne : choix d'un capteur

Les premiers essais réalisés à partir de topographies ont permis de valider notre méthodologie globale par de bons taux de classification. Il faut quand même souligner à nouveau que la mesure topographique représente le maximum d'information que l'on puisse obtenir sur la texture des surfaces étudiées, et il est donc heureux que nos méthodes de caractérisation fonctionnent bien dans ce cas de figure : à vaincre sans péril... Malheureusement, la mesure précise d'une carte topographique d'un produit défilant à une vitesse de l'ordre de 150m/mn n'est pas un problème simple, c'est même un problème suffisamment complexe pour qu'il soit plus ardu de le résoudre que d'envisager l'utilisation d'un capteur dégradant l'information topographique, quitte à raffiner les traitements numériques.

Dans le choix du capteur à utiliser pour rendre compte plus ou moins directement de la texture des surfaces, nous avons choisi d'imposer certaines contraintes. Tout d'abord, la représentation des surfaces issues de ce capteur doit être bidimensionnelle (c'est à dire une image), ce qui à notre sens est tout à fait nécessaire pour rendre compte de l'aspect des produits considérés. D'autre part, les données obtenues doivent être suffisamment proches d'une carte topographique pour que les descripteurs employés dans ce cas de figure restent utilisables.

5.3.1. Système Ondulo

Le système Ondulo est un système de projection de franges développé par la société Techlab. Le principe très simplifié de ce capteur réside dans l'observation par *réflexion sur la surface* d'une grille régulière projetée sur un écran. Un algorithme de calcul évalue les pentes de la surface à partir des déformations de la grille, puis en déduit un champ de courbure qui est classiquement utilisé pour la détection de défauts locaux. Le système semble très pratique à utiliser en pied de ligne, mais les tests d'acquisition que nous avons pu faire ont rendu incertains les résultats d'une application en ligne. En fait, le système se base idéalement sur la projection successive de plusieurs grilles légèrement décalées, ce qui est impossible (ou du moins très délicat) à obtenir sur un produit en mouvement, et la précision des résultats s'en trouve fortement dégradée par rapport à un système de laboratoire.

5.3.2. Système SIAS

Le système SIAS (Système d'inspection automatique des surfaces) est un dispositif d'acquisition vidéo basé sur une caméra *linéaire* et sur le défilement du produit. La caméra linéaire photographie à un instant t une ligne de la surface orthogonale au sens de défilement du produit. Le mouvement de la surface fait qu'à l'instant $t + \delta t$, la caméra photographie une ligne légèrement décalée par rapport à la première. Un réglage adapté de l'intervalle d'acquisition de la caméra par rapport à la vitesse de défilement du produit permet de reconstruire une image de la surface. Une telle image se prête très bien à une analyse par nos méthodes de classification, et les résultats obtenus en

laboratoire (Arcelor Research, Maizières-lès-Metz) sur la base des surfaces de type peau d'orange sont corrects. Ce système a déjà été testé en ligne pour le contrôle de défauts ponctuels, mais il reste coûteux et fragile, et le réglage angulaire de la configuration "source d'éclairage/caméra" est assez délicat, ce qui peut avoir parfois tendance à fausser l'analyse de texture. Par ailleurs, le système SIAS fait l'acquisition de lignes sur toute la largeur du produit, mais l'aspect de celui-ci reste en général assez homogène selon la direction orthogonale au sens de défilement, et il n'y a donc que peu d'intérêt à caractériser l'ensemble de la bande. On peut en général se limiter à des mesures de taille plus limitée, c'est à dire à des images plus localisées.

5.3.3. Système "caméra matricielle"

La complexité de l'application ligne du système SIAS vient de l'indispensable asservissement entre la vitesse de la ligne et la vitesse d'acquisition de la caméra. Malgré tout, l'expérience SIAS a montré qu'un système basé sur la réponse optique des surfaces devrait permettre leur classification avec une bonne sensibilité. Pour alléger le dispositif SIAS, on pense assez naturellement à remplacer les acquisitions successives d'une caméra linéaire par l'acquisition unique d'une caméra matricielle. Pour utiliser un tel dispositif, il faut prendre en compte certaines contraintes. Tout d'abord, la résolution de la caméra doit être adaptée aux échelles caractéristiques des phénomènes observés (de l'ordre de la dizaine de microns). Ensuite, les surfaces observées sont très spéculaires, et une caméra de bonne résolution ne sera capable d'obtenir une image intéressante qu'en présence d'une source lumineuse adaptée : les différents essais ont montré que les meilleures images étaient obtenues avec un éclairage collimaté éloigné de quelques degrés par rapport à la direction spéculaire imposée par l'angle entre la caméra et la normale à la surface, de façon à ne pas saturer la caméra en conservant malgré tout une bonne sensibilité. Finalement, le défilement du produit nécessite d'avoir une caméra suffisamment sensible pour pouvoir réduire au maximum le temps d'obturation et éviter ainsi les effets de flou.

L'utilisation d'une photographie au lieu d'une carte topographique constitue une véritable perte d'information : dans la configuration choisie, l'image obtenue dépend essentiellement des pentes dans la direction d'observation : plusieurs topographies différentes peuvent donc théoriquement générer la même image caméra. Par contre, on peut imaginer qu'une image caméra est plus proche de l'information recueillie par l'oeil humain qu'une carte topographique, et qu'à ce titre la classification de l'aspect des surfaces sera facilitée. Ce n'est que partiellement vrai, puisque un observateur humain aura la capacité de faire varier les angles d'observation et d'incidence de la lumière en pivotant légèrement l'échantillon observé ou simplement en modifiant la position de son oeil. Il faudrait donc une information constituée de plusieurs images sous différentes incidences pour tenter de reproduire le capteur "oeil humain". Un tel système n'est pas inenvisageable, mais le bon sens suggère d'essayer de classer les surfaces à partir d'une seule incidence et d'enrichir éventuellement cette information si celle-ci n'est pas discriminante. En pratique, les tests effectués se sont montrés suffisamment concluants pour valider l'utilisation d'une seule caméra matricielle dans un premier prototype ligne d'analyse de la texture des surfaces.

5.4. Classification à partir d'images

Dans cette section, nous nous intéressons aux données acquises en laboratoire par caméra matricielle, et nous cherchons à montrer la faisabilité de la classification des surfaces par le biais de ce capteur.

5.4.1. Acquisition et prétraitement

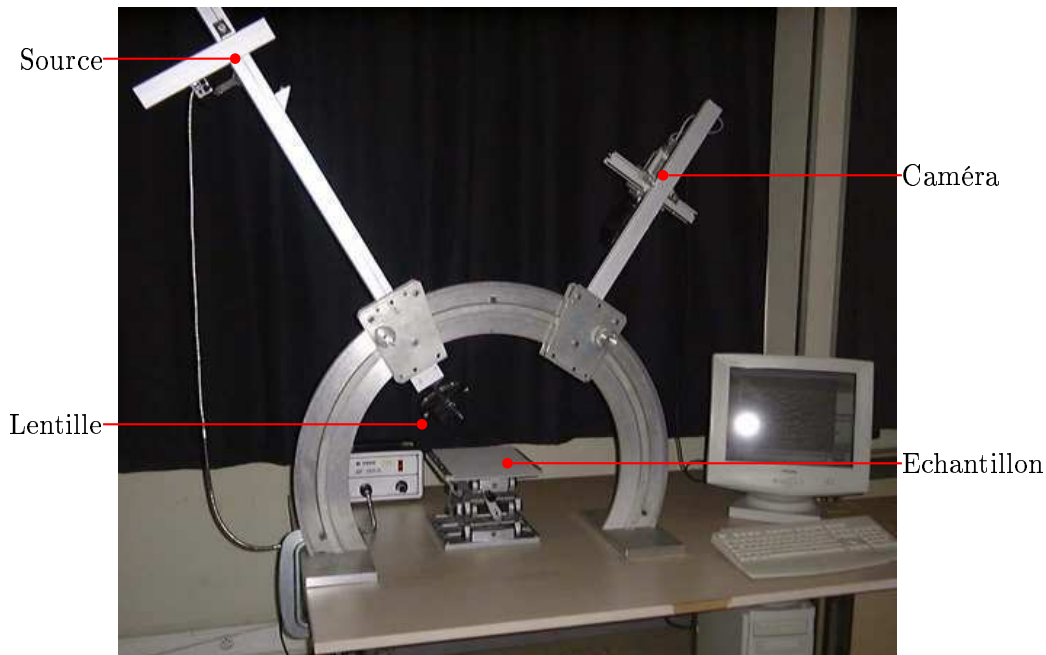


FIG. 5.15.: Banc d'imagerie utilisé en laboratoire

Comme nous l'avons signalé, le choix de la configuration surface/source/caméra demande quelques réglages, et nous présentons ici l'une des configurations pour lesquelles les images obtenues étaient les plus à même de permettre une classification. Nous avons utilisé une caméra vidéo standard de 768×576 pixels avec un objectif de 105mm de focale, une source halogène en lumière blanche d'une puissance de 500W collimatée par une lentille convergente biconcave d'un mètre de focale. L'angle d'incidence de la lumière par rapport à la normale moyenne de l'échantillon est $\theta_i = 20^\circ$, tandis que l'angle de reprise de la caméra est de $\theta_r = 18^\circ$, soit un décalage de 2° par rapport au spéculaire.

Nous avons travaillé sur les mêmes échantillons physiques que dans la section 5.2, et nous avons voulu reproduire le même protocole expérimental : chaque surface a été observée 5 fois indépendamment, c'est à dire positionnée de 5 façons différentes sous la caméra (en respectant toutefois la direction de laminage), de sorte que les 5 images obtenues pour chaque surface sont assez représentatives de la variabilité que l'on pourrait obtenir en mesurant la même bobine à quelques secondes d'intervalle lors du défilement du produit. La seule différence par rapport à la section précédente est l'absence des

surfaces 15, 17 et 22 pour lesquelles les images n'ont pas pu être obtenues. Les premières images correspondant à chaque surface sont visibles à la figure 5.16.

Nous verrons dans la suite que le passage d'une topographie à une image caméra s'accompagne d'une perte de sensibilité dans la classification qui est dû à un appauvrissement intrinsèque de l'information contenue dans l'image analysée, à noter que cette dégradation provient véritablement de la réponse optique de la surface, et pas d'une baisse de résolution, car celle de la caméra est comparable à celle du microscope interférométrique utilisé en mode "stitching" (soit $1 \text{ pixel} \simeq 40 \mu\text{m}$). Par contre, les manipulations conduisant à l'acquisition des images caméra sont plus simples que l'utilisation d'un microscope interférométrique, et il est facile de disposer pour chaque surface physique de plusieurs images indépendantes complètes et donc d'un matériel expérimental conséquent pour tester la qualité de la classification sans avoir à diviser les images en sous-images. La quantité d'information disponible (grande taille) dans les images caméra compense donc quelque peu (mais pas complètement) sa perte de qualité par rapport aux topographies.

Paradoxalement, les acquisitions en laboratoire présentent des difficultés intrinsèques qui ne se présenteront pas nécessairement lors d'une application ligne. L'une d'elle vient des défauts de planéités qui sont fréquents sur des échantillons découpés, qui ont été beaucoup manipulés. La caméra est très sensible à ces défauts, et un écart de quelques dixièmes de degré peut entraîner de forte variation de niveaux de gris d'une image à l'autre. Il apparaît donc judicieux de renormaliser la moyennes des images, de façon à ce que ces effets de planéités n'aient pas trop d'influence sur les résultats de classification. On peut justifier cette normalisation, si l'on considère que dans la plage angulaire considérée (quelques dixièmes de degrés), la réponse optique de la surface (brdf) est quasi-linéaire. Dans ces conditions, modifier la normale moyenne, c'est à dire l'orientation globale de l'échantillon, conduit à ajouter ou retirer une constante à l'image obtenue par la caméra. Ceci n'est bien sûr vrai qu'au premier ordre, donc de façon approximative, mais les résultats de classification se trouvent nettement améliorés par une telle renormalisation.

Par ailleurs, on observe beaucoup de structures de très hautes fréquences sur les images (bruit, rugosité à petite échelle) qui ne sont pas utilisables pour la différenciation des surfaces correctes et incorrectes. Nous avons donc appliqué un léger préfiltrage gaussien avant le traitement, de façon à faciliter les opérations de segmentation et à assurer la pertinence du calcul des vecteurs normaux.

Après ces préliminaires, le déroulement des opérations de classification est identique au cas des topographies.

5.4.2. Analyses spécifiques

Nous présentons ici les mêmes graphiques que dans le cas de l'analyse des topographies (figure 5.17, 5.18, 5.19, 5.20, 5.22, 5.23, 5.24, 5.25).

La signification des facteurs principaux de chaque analyse change par rapport à l'étude des topographies, de même que le choix des variables à conserver dans l'analyse globale, mais la méthode d'interprétation reste identique. Nous nous contentons donc de résumer les principales conclusions à tirer de l'ensemble des analyses spécifiques :

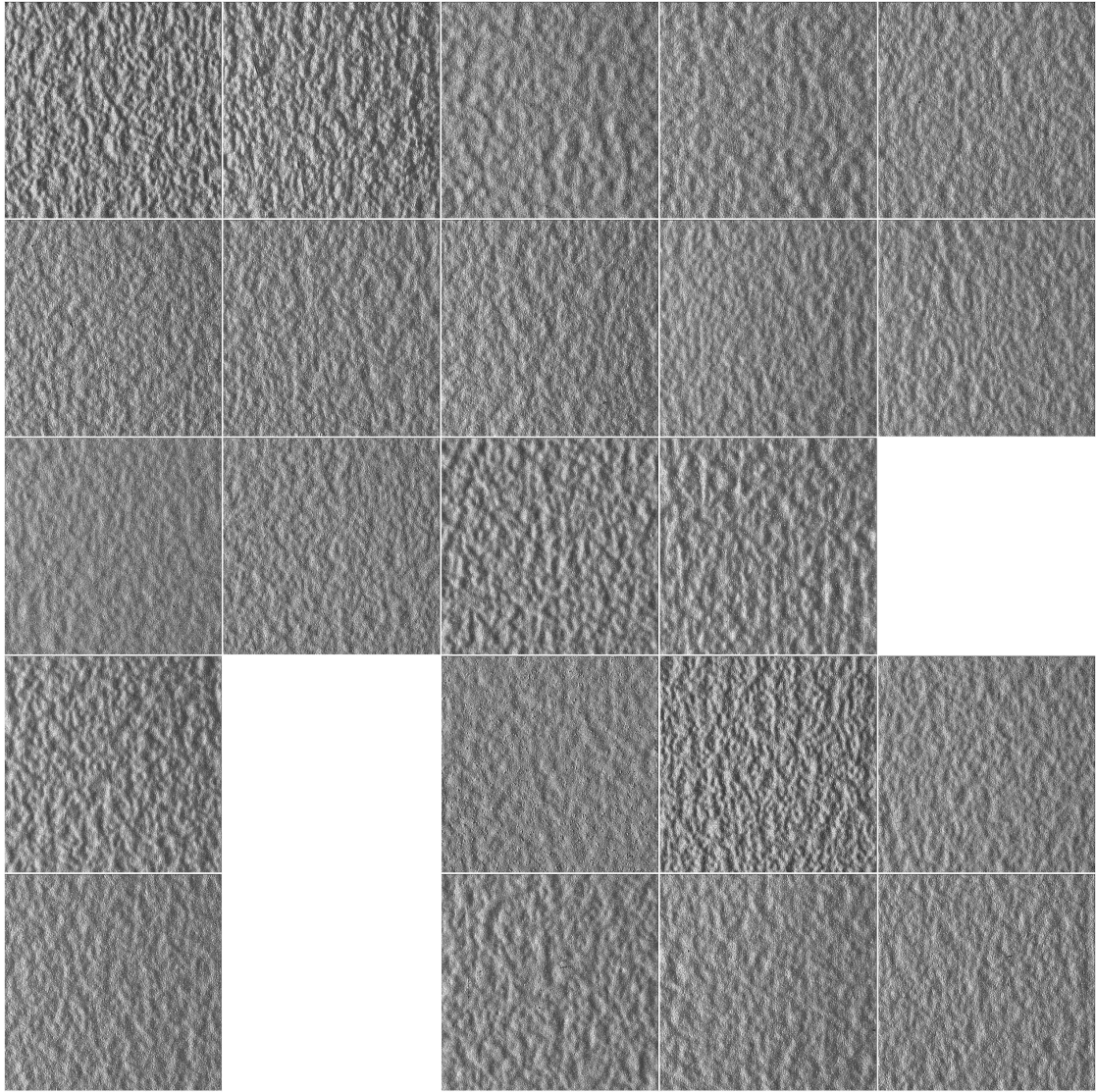


FIG. 5.16.: Image caméra de 25 surfaces "peau d'orange"

Covariance : On retient l'axe 1 pour l'analyse globale (présence d'une covariance négative pour $h \simeq 15$ pixels). Les oscillations caractérisées par le second axe propre sont dues à un phénomène de sur-apprentissage.

Erosion-dilatation : On conserve les deux premiers facteurs principaux pour l'analyse globale.

Hauteurs de zones : L'axe 1 caractérise un fort contraste entre les zones, l'axe 2 une dissymétrie entre zones hautes et basses : on conserve ces deux facteurs principaux.

Surfaces de zones : On garde uniquement l'axe 1 (peu de très petites zones beaucoup de moyennes et quelques grandes, cf figure 5.21).

Tortuosités : On conserve l'axe 1 (beaucoup de fortes tortuosités, peu de faibles), mais pas l'axe 2, car il ne permet pas de distinguer surfaces correctes et incorrectes (cf figure 5.22(a)).

Pentes : On ne garde aucun axe. Là encore, il n'y a pas de véritable phénomène de sur-apprentissage, car les axes correspondent à des caractéristiques réelles des surfaces, mais ces caractéristiques ne sont pas discriminantes des classes correctes et incorrectes.

Orientation : Pas d'axes propres sélectionnés : sur-apprentissage.

Ondelettes : On conserve les deux axes. Le premier caractérise des surfaces ayant peu d'énergie pour la bande de longueurs d'onde 0.12/0.42 mm et beaucoup d'énergie dans la gamme 0.5/1 mm (voir figure 5.26). Le second axe propre caractérise les surfaces ayant beaucoup d'énergie dans la gamme 0.09/0.7 mm.

a. Analyse globale

L'analyse globale permet de projeter les surfaces dans un espace où surfaces correctes et incorrectes sont assez facilement séparables (voir figure 5.27. La grande différence par rapport au cas des topographies est la perte de sensibilité : si la séparation entre les classes correcte et incorrecte est pertinente, il n'y a plus de classe "intermédiaire" nette : les surfaces 14, 19 et 16 sont se projettent au milieu des échantillons incorrects, tandis que la surfaces 23 (qui était la surface incorrecte la plus proche des surfaces intermédiaires) se trouve nettement plus roche des surfaces correctes. Nous serons obligées d'adapter la phase de classification à ce manque de discrimination entre surfaces incorrectes et surfaces "intermédiaires".

b. Pertinence du bootstrap

La question de la validité des méthodes du bootstrap dans le cas des images peut légitimement être soulevée. En effet, l'acquisition et les prétraitements effectués à cause de la non planéité des échantillons pourraient parfaitement induire d'une image à l'autre des variations statistiques dont le seul bootstrap ne pourrait rendre compte. Très concrètement, le problème vient de la superposition de deux fonctions aléatoires : celle générant la topographie de la surface d'une part et celle caractérisant l'observation de cette topographie d'autre part (planéité de l'échantillon observé, variation de l'éclairage ambiant

d'une acquisition sur l'autre, bruit de la caméra etc...). Le bootstrap permet de rendre compte des variations dans les descripteurs étudiés dues à la première fonction aléatoire, mais pas de celles dues à la seconde, qui ne transparaissent pas sur une seule image : toute la question est donc de savoir si cette seconde dispersion est significativement plus faible que la première ou non : si c'est le cas, alors les distributions estimées à partir des échantillons bootstrap seront pertinentes, sinon, les variances seront sous-estimées.

La seule façon de répondre à cette question est de comparer la distribution des échantillons bootstrap à celle des échantillons réels issus des mêmes fonctions aléatoires. Nous disposons pour chaque surface test de cinq réalisations indépendantes, et dans la plupart des cas (notamment dans l'analyse globale), les projections des réalisations réelles issues d'une même surface sont incluses dans le nuage d'échantillons bootstrap générées par l'une d'entre elle, ce qui tend à valider l'utilisation du bootstrap. Les deux premiers graphiques de la figure 5.23(a) montrent même un exemple assez intéressant : les distributions bootstrap associées aux surfaces 11 et 19 (par exemple) sont très différentes en terme de variance (notamment sur l'axe 1), mais la superposition des réalisations réelles montre la pertinence respective de ces deux distributions bootstrap : des lois aussi différentes et anisotropes pour les surfaces 11 et 19 auraient probablement été assez difficile à modéliser sans avoir recours à des méthodes de sur-échantillonnage. Les distributions des échantillons bootstrap semblent donc toujours en adéquation avec les distributions d'échantillons réels, même si un oeil attentif constatera que le rapport des variance calculées sur un nuage bootstrap d'une part (400 individus) et sur des réalisations réelles d'autre part (5 individus) est plutôt plus faible (et moins réaliste) que dans le cas des topographies, certaines réalisations étant même à la limite des nuages d'échantillons bootstrap correspondants.

En ligne, le problème de planéité devrait être moins important qu'en laboratoire, à cause de la tension exercée par les rouleaux sur la bande d'acier. On conserve donc de bonnes chances de pouvoir continuer à utiliser le bootstrap tout à fait légitimement. Une autre solution à ce problème serait de modéliser d'une façon ou d'une autre les variations des descripteurs dues à l'acquisition des images, et de développer une méthode de génération d'échantillons hybride entre des techniques de bootstrap pur s'appuyant uniquement sur les données observées et des techniques paramétriques s'appuyant sur l'élaboration d'un modèle simulant la variabilité dans l'acquisition.

c. Classification à partir des images

Si l'on applique exactement le même algorithme que pour les topographies (division en trois classes et classification bayésienne classique ou par espérance de la qualité), on obtient les résultats présentés à la figure 5.28. On obtient 10 erreurs de classification (6 échantillons incorrects sont classés "intermédiaires", un est classé "correct", et 3 échantillons "intermédiaires" sont classés "corrects") soit 11% de mauvaises classifications. Les écarts-type de la qualité attribuée à chaque surface montrent aussi que la division entre surface incorrecte et intermédiaire n'est pas nette. Un simple regard sur la projection des surfaces dans le sous-espace déterminé par l'analyse globale permettait de prévoir ce résultat moyen. Le recours qui s'offre à nous est de considérer que les surfaces 14, 16 et

19 sont finalement incorrectes, de classer alors les surfaces en deux catégories "correctes" et "incorrectes" et d'estimer leur qualité en attribuant la valeur 0 aux surfaces 14, 16 et 19. Les résultats de cette seconde opération de classifications sont visibles figure 5.29. On n'observe plus qu'une véritable mauvaise classification (l'une des images de la surface 23 est classée correcte au lieu d'incorrecte), soit un pourcentage d'erreur de 1.1%. Cependant, la figure 5.29(b) montre que deux des images de la surface 23 ont des qualités estimées à 0.4 et 0.8 au lieu de 0.

La sensibilité obtenue à partir des images caméra n'est pas aussi bonne qu'à partir des images topographiques : celles-ci permettaient la distinction de trois classes parmi les surfaces de la base de test tandis que les images caméra n'en distinguent plus que deux. Cela étant dit, si ces résultats ne sont pas parfaits, ils restent très encourageant : les traitements appliqués sur les images caméra et sur les topographies sont ici rigoureusement identiques, à l'exception du léger prétraitement des images caméra, et il est probable que la modification de certains descripteurs (voire même l'introduction de nouvelles variables) permettrait de mieux tirer parti de l'information contenue dans les images caméra. Quoi qu'il en soit, même en l'état, ces résultats sont suffisamment bons pour valider la méthode et le choix du capteur puis entamer le processus de prototypage pour des essais en ligne.

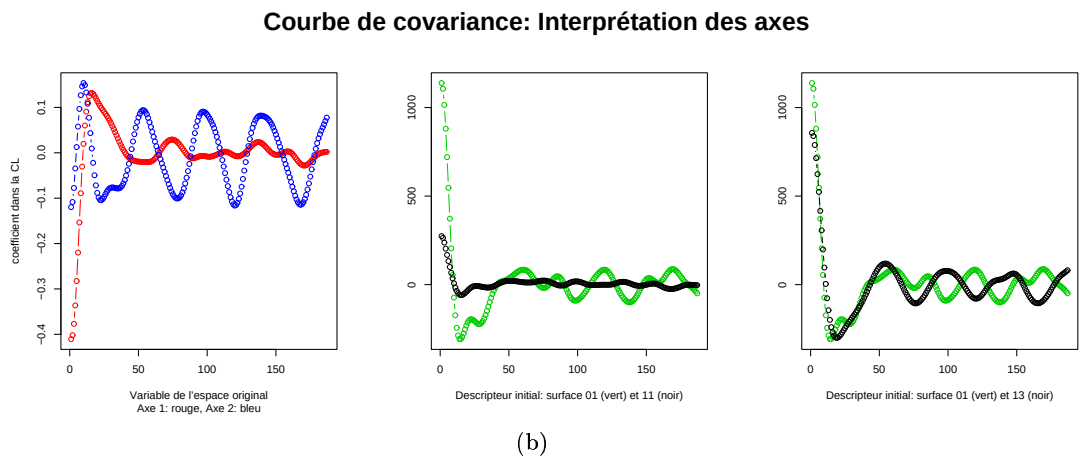
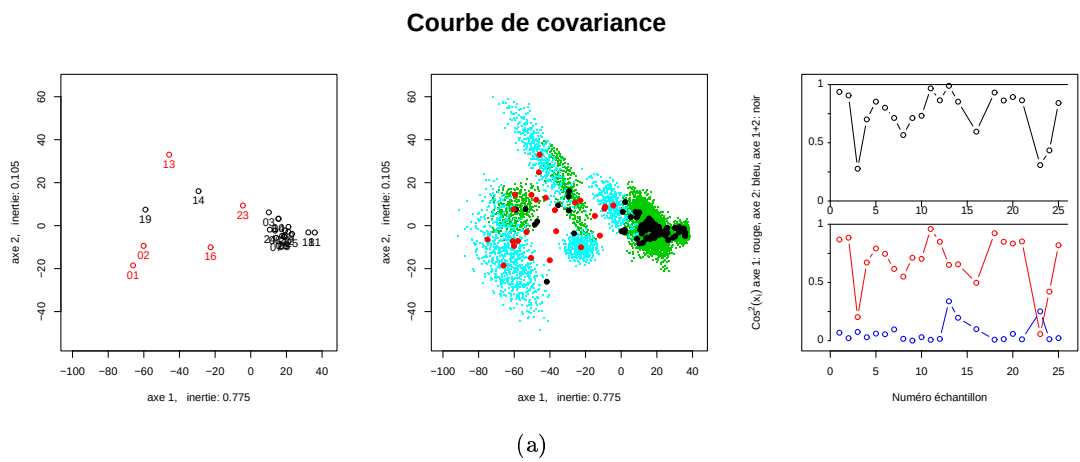
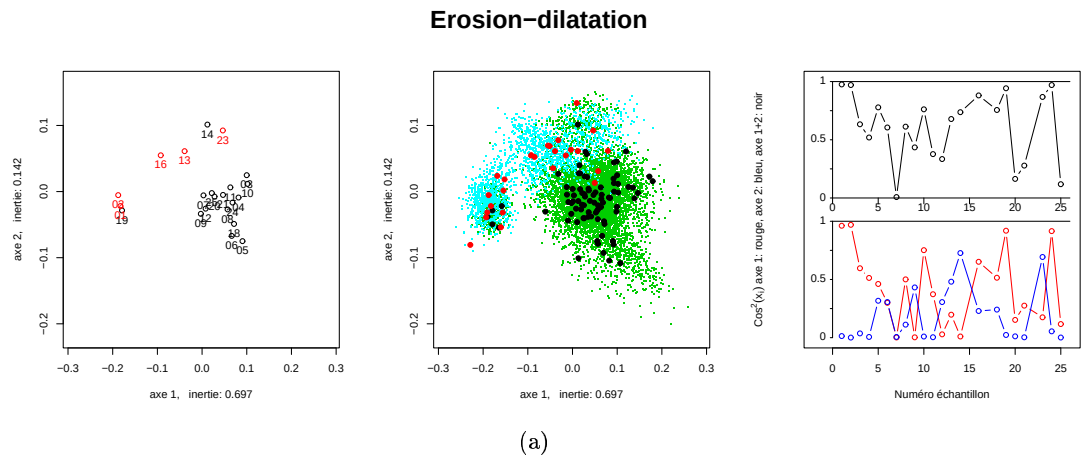


FIG. 5.17.: Courbes de covariance



Erosion-dilatation: Interprétation des axes

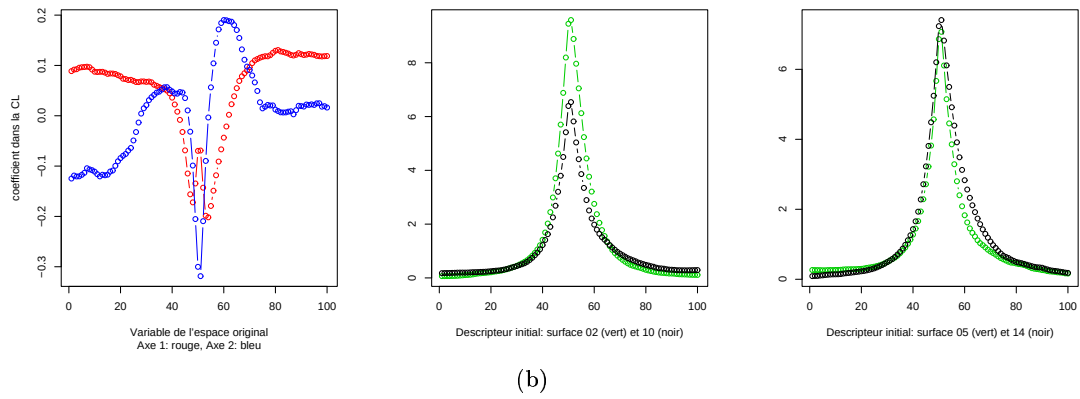


FIG. 5.18.: Courbes d'érosion-dilatation

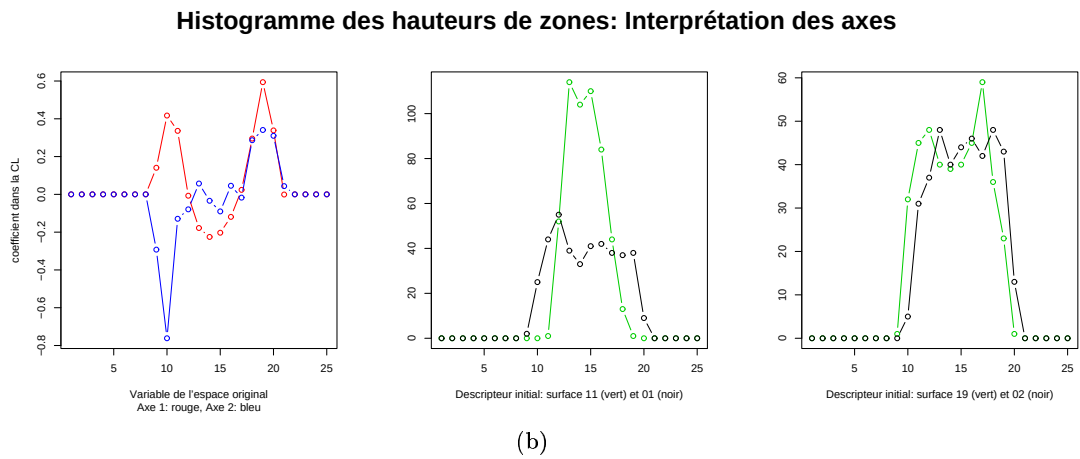
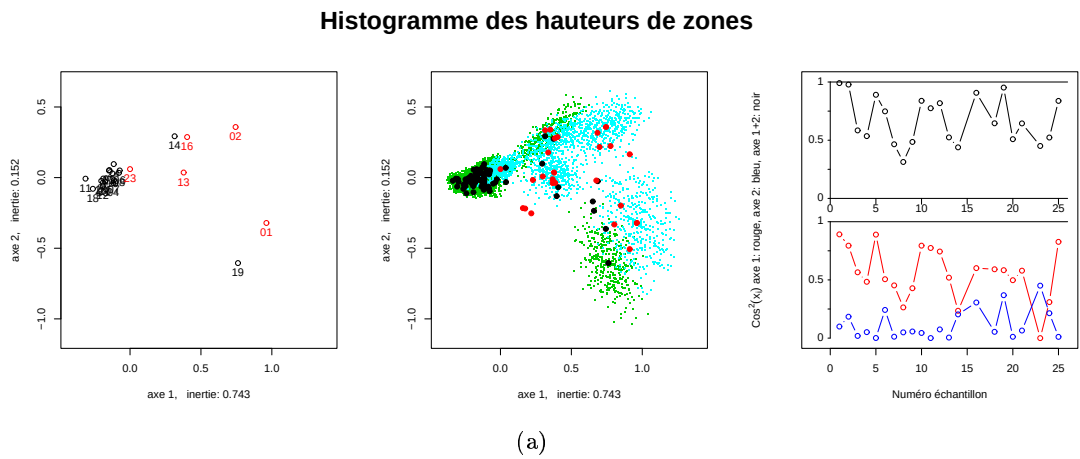


FIG. 5.19.: Histogramme des hauteurs de zones

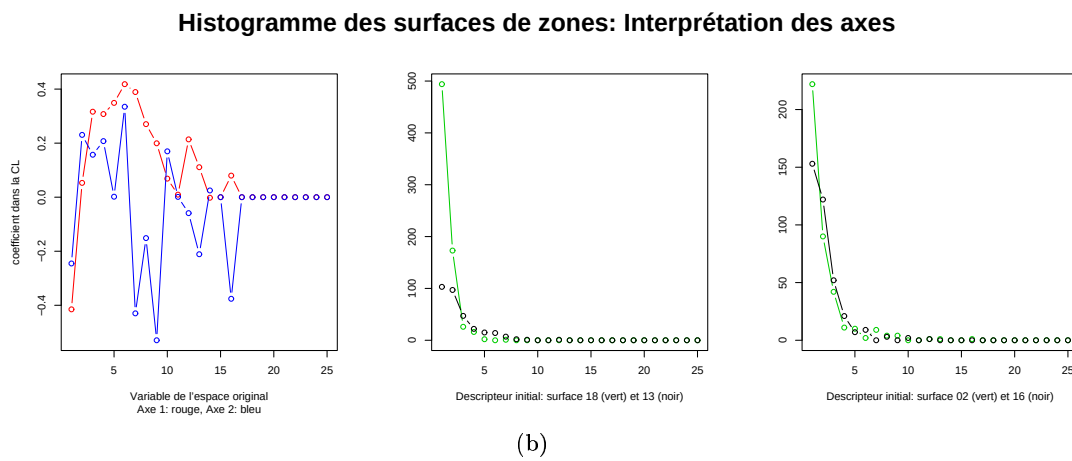
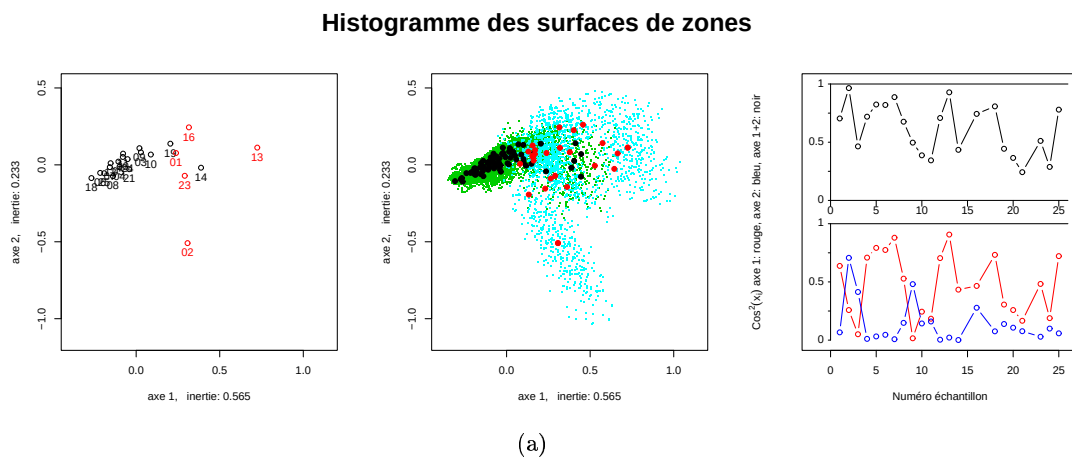
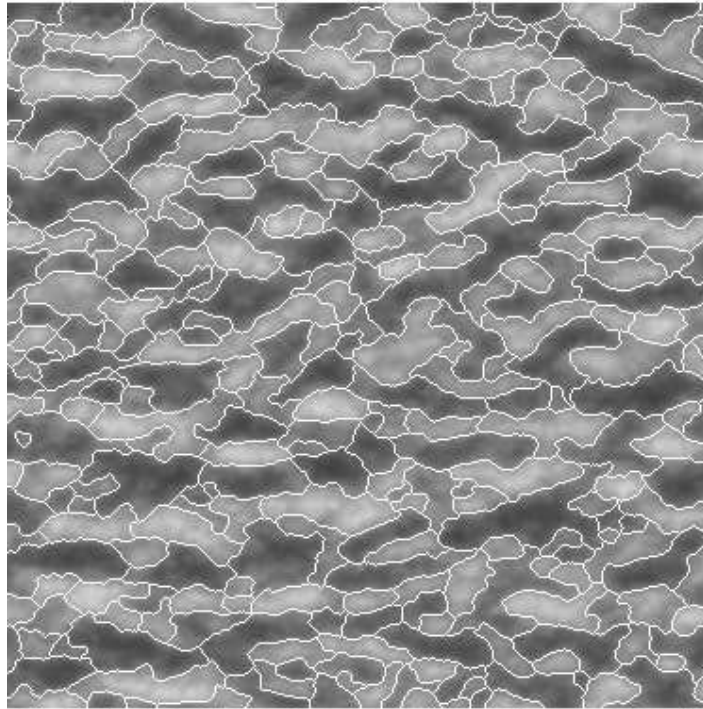


FIG. 5.20.: Histogramme des surfaces de zones

Pdo 13



Pdo 18

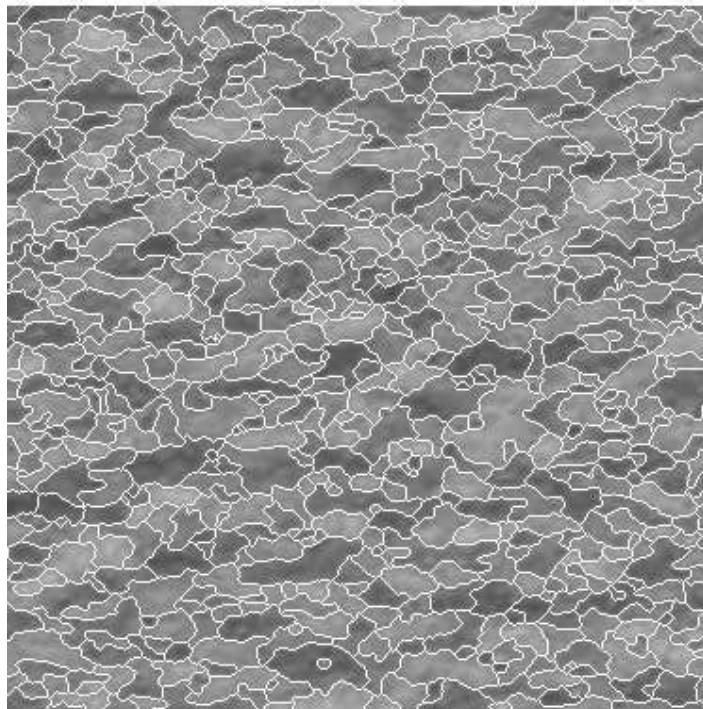
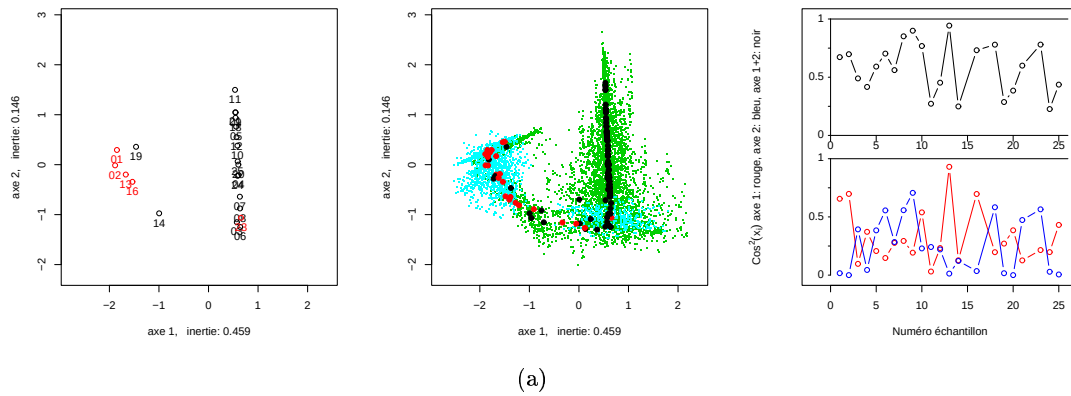


FIG. 5.21.: Images extrêmes pour l'axe 1 des surfaces de zones. L'image 13 possède moins de petites zones et plus de très grandes que l'image 18

Histogramme des tortuosités



Histogramme des tortuosités: Interprétation des axes

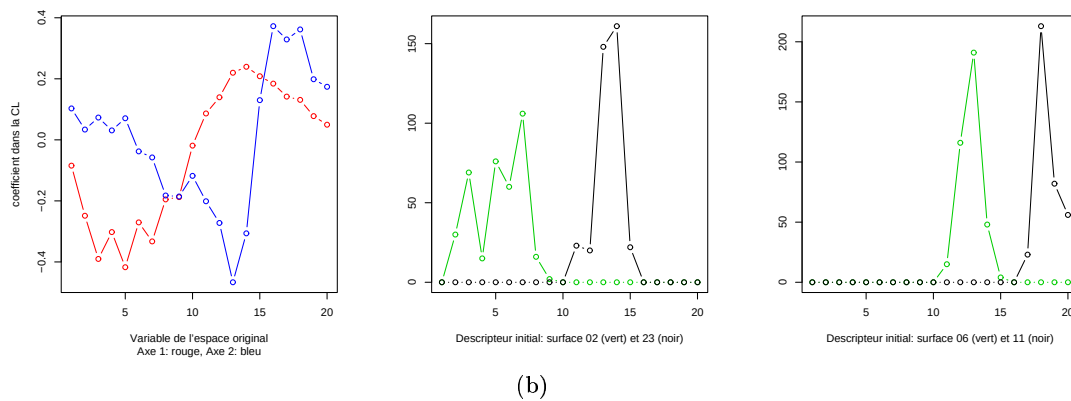


FIG. 5.22.: Histogramme des tortuosités

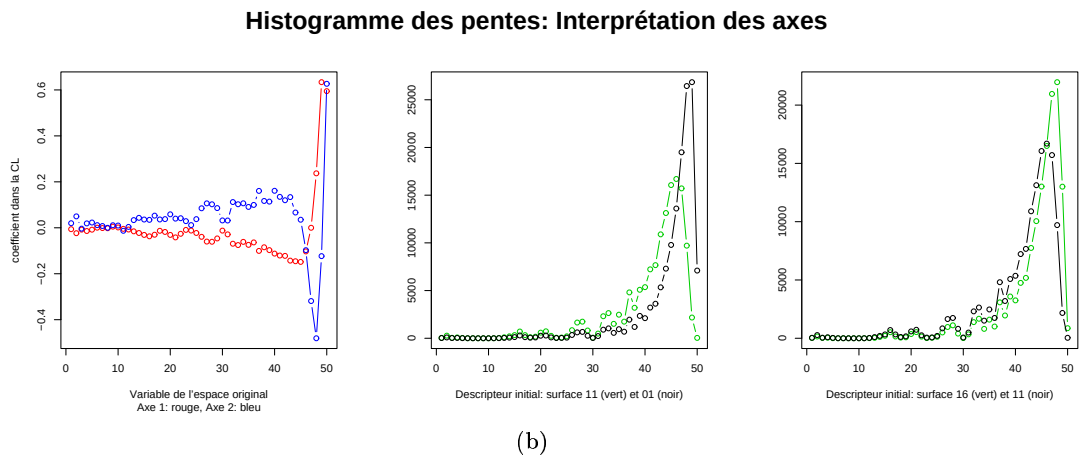
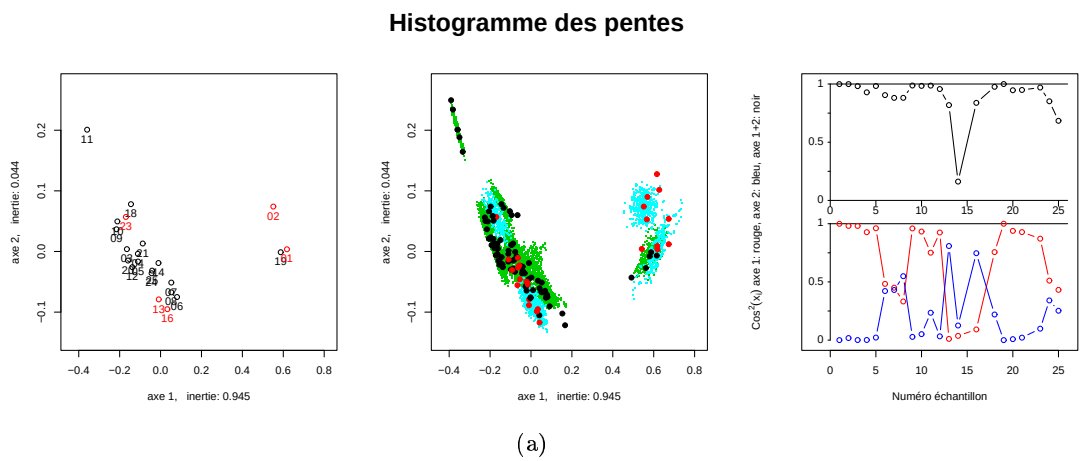
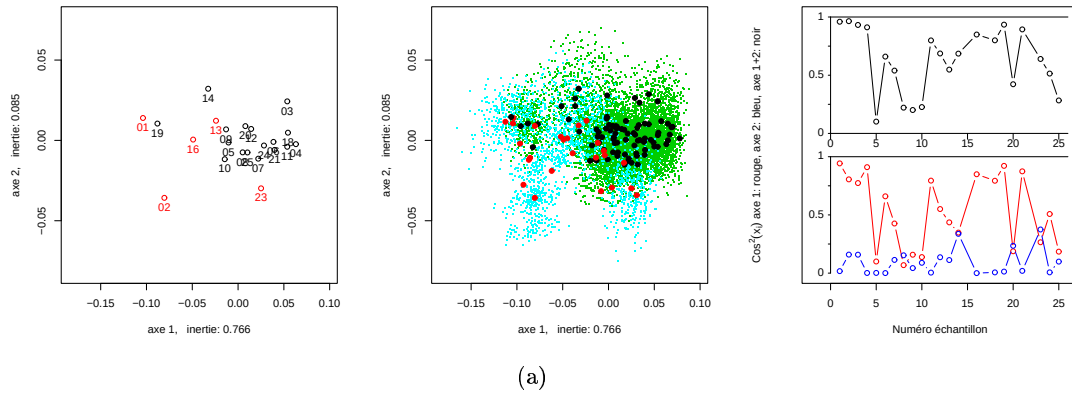


FIG. 5.23.: Histogramme des pentes

Histogramme des orientations de normales



Histogramme des orientations de normales: Interprétation des axes

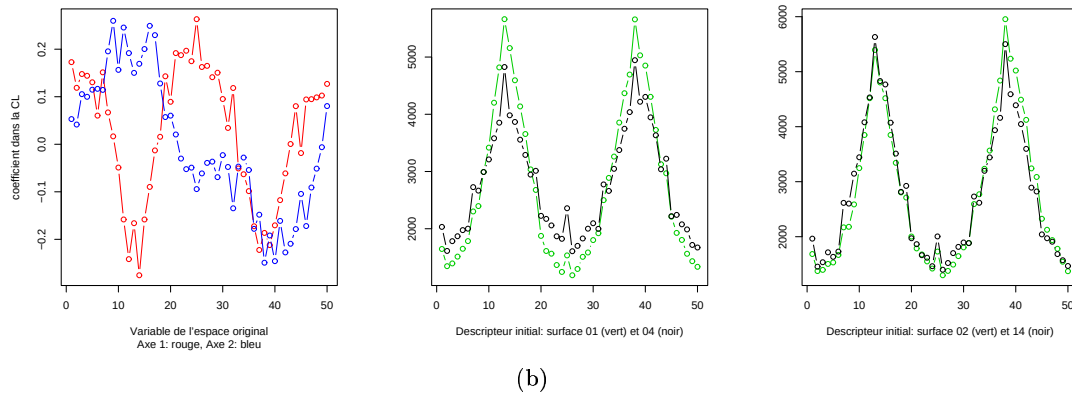


FIG. 5.24.: Histogramme des orientations des vecteurs normaux

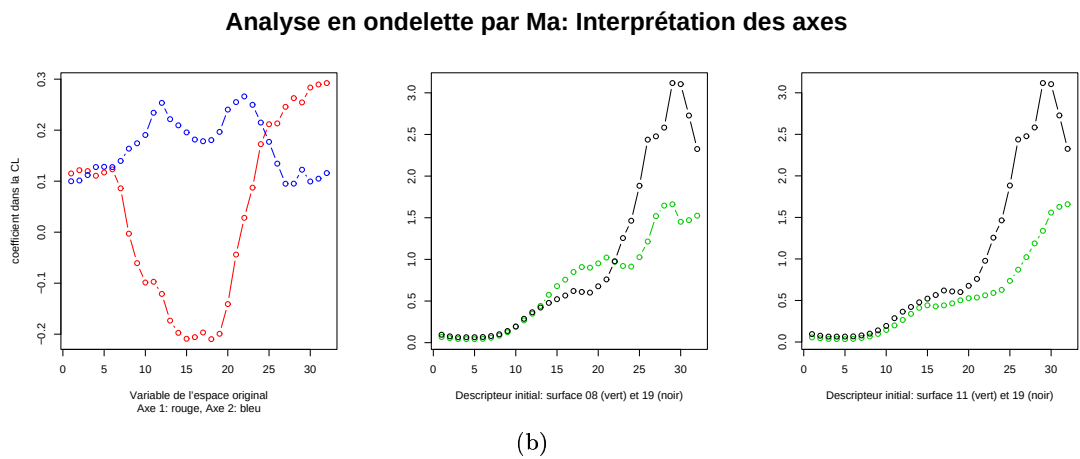
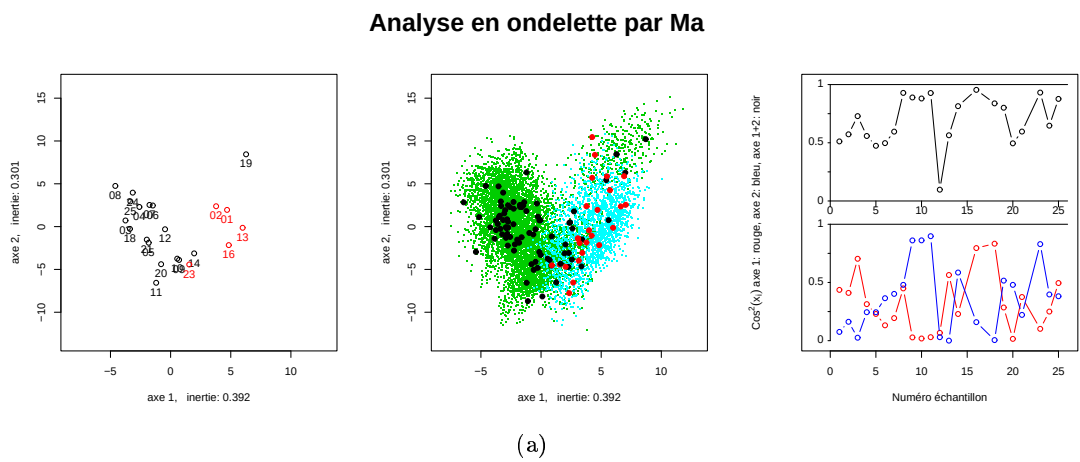
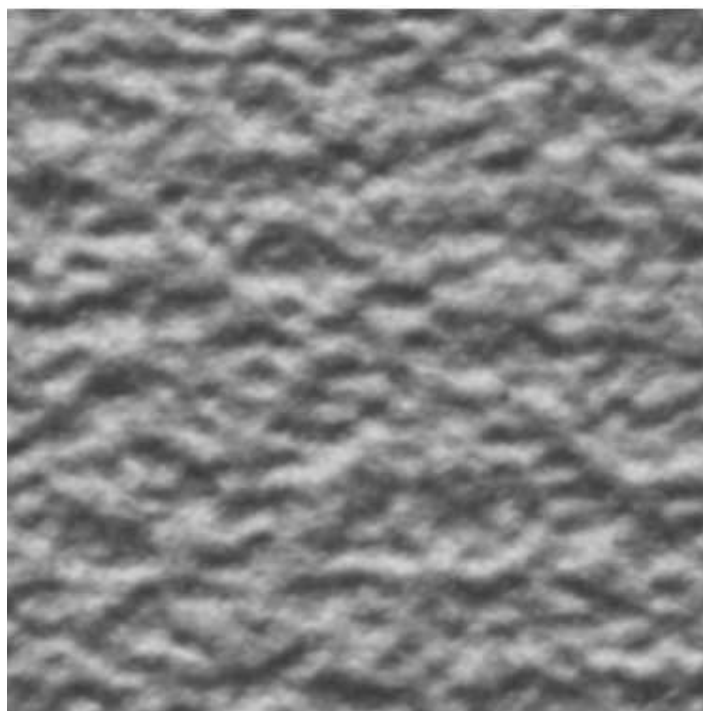


FIG. 5.25.: Courbes de Ma

Pdo 19



Pdo 08

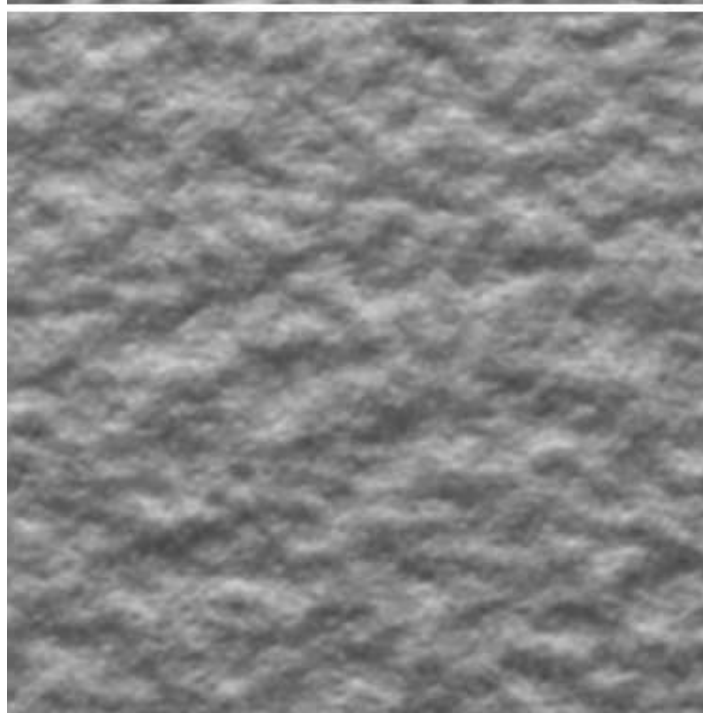
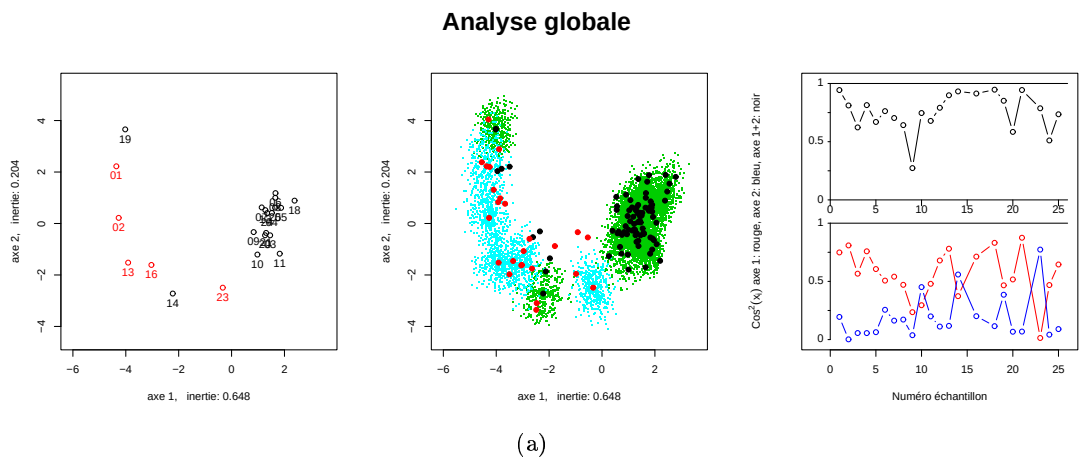


FIG. 5.26.: Images extrêmes pour l'axe 1 des courbes de Ma. L'image 19 présente de fortes amplitudes dans les grandes longueurs d'onde, tandis que l'image 8 (faible valeur sur l'axe 1) possède plus de fortes amplitudes pour des longueurs d'onde plus faibles. On distingue nettement un "grain" caractéristique de faible longueur d'onde pour cette image.



Analyse globale: interprétation des axes

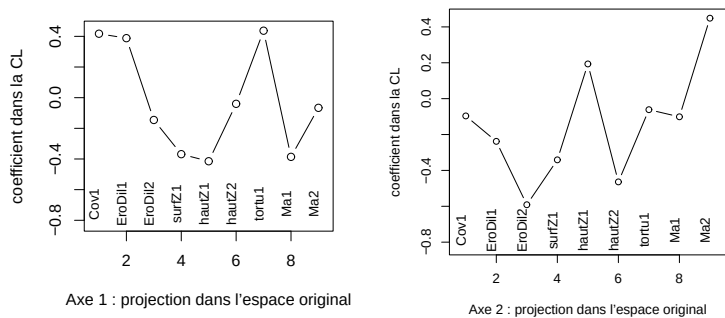
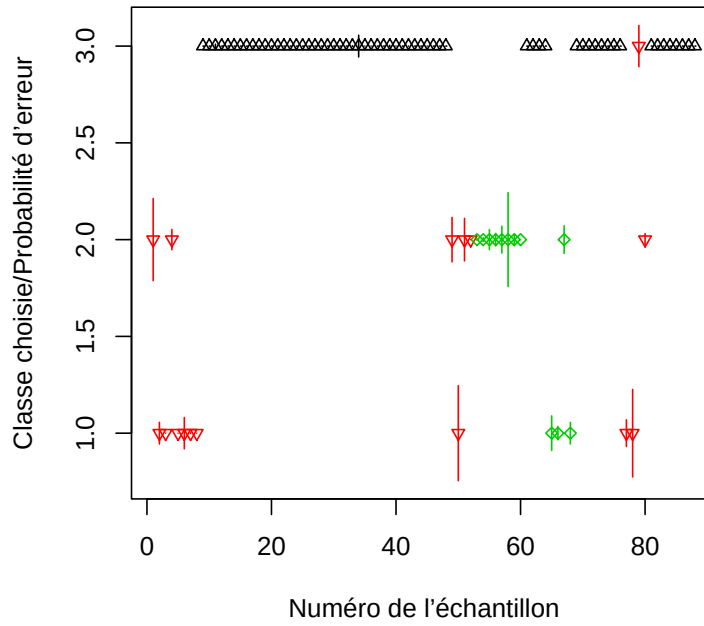
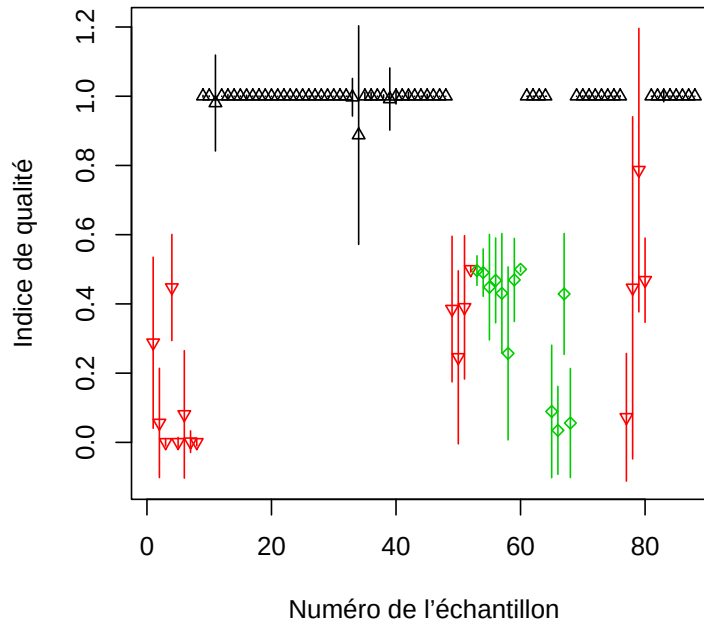


FIG. 5.27.: Analyse globale sur images caméra

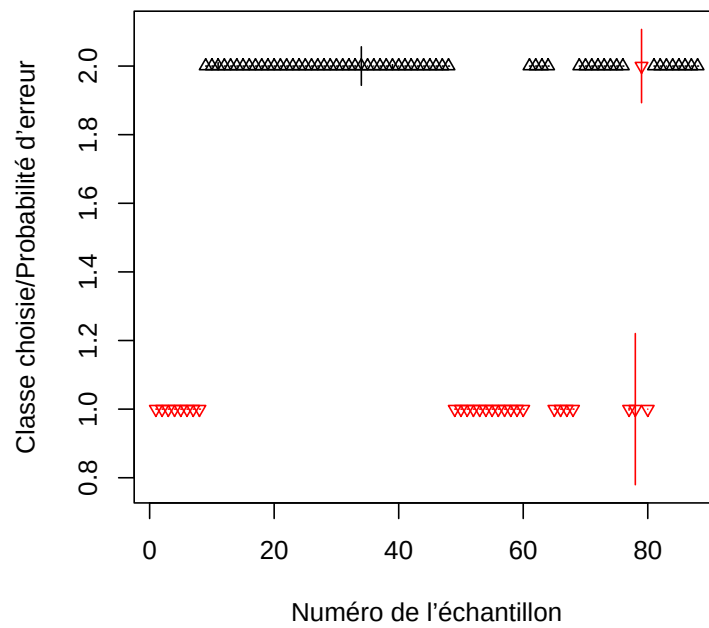


- (a) Les points rouges représentent les sous-surfaces issues d'une topographie incorrecte, les verts sont les sous-surfaces intermédiaires et les noirs correctes. L'ordonnée représente la classe attribuée à chaque sous-surface par le classificateur bayésien, et la longueur du segment associé à chaque point représente la probabilité d'erreur.

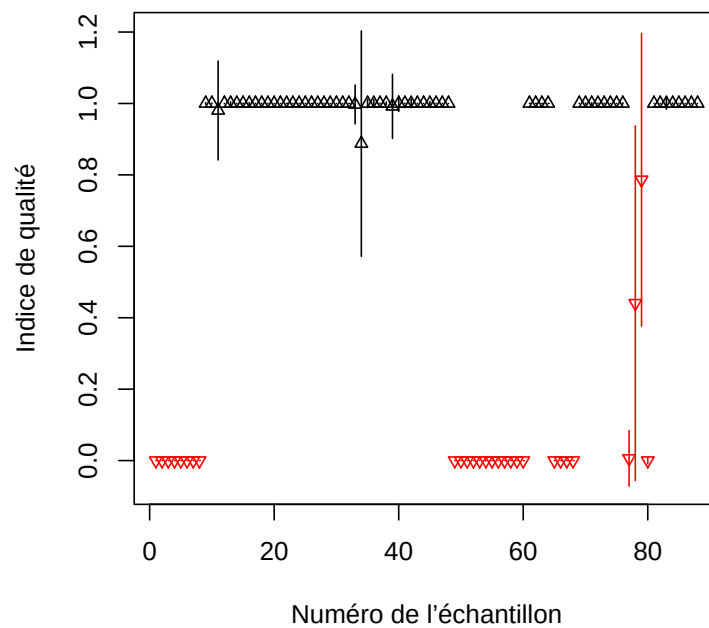


- (b) Pour chaque sous surface, on estime l'espérance de la qualité (position verticale) et son écart-type (longueur du segment associé à chaque point).

FIG. 5.28.: Résultat de classifications à partir de topographies (3 classes)



- (a) Les points rouges représentent les sous-surfaces issues d'une topographie incorrecte, les verts sont les sous-surfaces intermédiaires et les noirs correctes. L'ordonnée représente la classe attribuée à chaque sous-surface par le classificateur bayésien, et la longueur du segment associé à chaque point représente la probabilité d'erreur.



- (b) Pour chaque sous surface, on estime l'espérance de la qualité (position verticale) et son écart-type (longueur du segment associé à chaque point).

FIG. 5.29.: Résultat de classifications à partir de topographies (2 classes)

5.5. Mise au point d'un prototype, premiers essais

Les deux principaux obstacles dans la conception d'un premier prototype sur ligne sont d'une part de choisir une usine pilote, et surtout de choisir un emplacement accessible sur la ligne pour lequel l'intégration d'un dispositif encombrant et fragile est possible, et d'autre part d'adapter le matériel, en particulier la caméra aux contraintes industrielles. La caméra retenue par Arcelor Research est une caméra Sony (XC-HR50) permettant de limiter le temps d'exposition jusqu'à $1/100000$ s. La matrice 659×494 pixels associée à un objectif Nikon de focale 105mm fournit une résolution de $40\mu m$. Ces caractéristiques garantissent une image nette, même si la vitesse de la ligne atteint les 180m/mn. Le système d'éclairage est une lampe de 50W (lumière blanche) collimatée par une lentille de focale 500mm. L'ensemble est fixé sur un support mécanique de façon à ce que l'angle d'incidence de la lumière soit 20° , et l'angle de reprise 16° . Le prototype intègre également un système de protection pour parer à d'éventuels chocs si le produit se mettait à onduler avec une forte amplitude (voir figures 5.30 et 5.31). Ce système a d'abord été testé

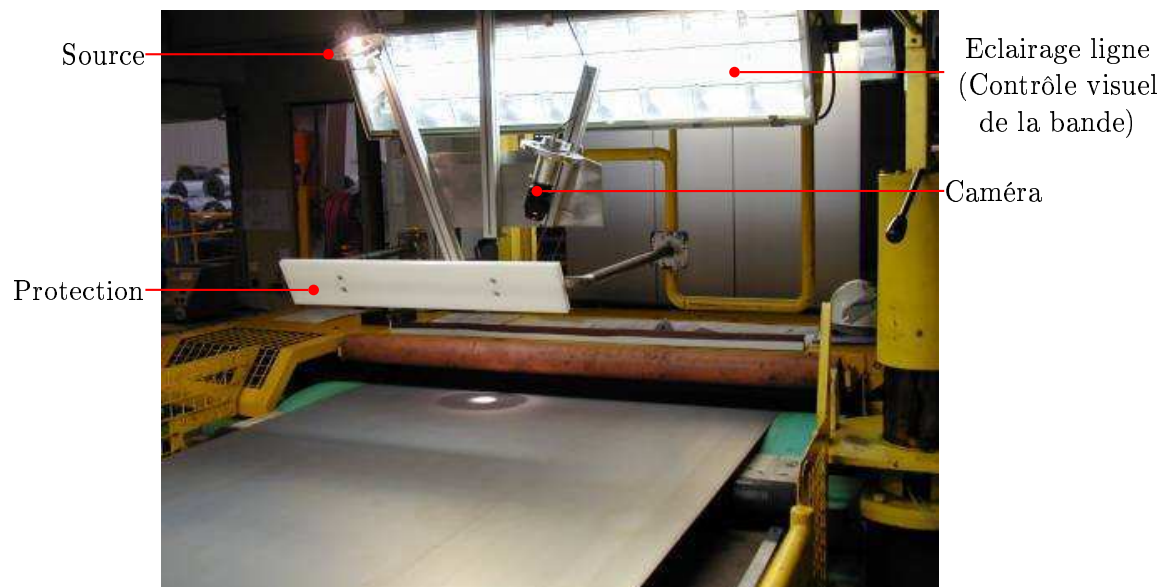


FIG. 5.30.: Prototype d'acquisition en ligne d'images caméra

en laboratoire à Maizières-lès-Metz, avec de nombreux essais de répétabilité (plusieurs mesures au même endroit sur la surface) et de reproductibilité (plusieurs mesures en différentes positions des échantillons tests) avec un temps d'exposition de l'ordre de $10\mu s$ nécessaire à la netteté des images en ligne. Toutes ces expériences ont eu pour support de validation la projection des images dans le sous-espace issu de l'analyse globale, et tous les résultats ont été positifs. Une campagne de tests a ensuite été menée par Arcelor Research [62] de façon à optimiser la configuration angulaire et optique (choix des lentilles, de l'éclairage, etc...) du système.

Nous avons pu installer notre système sur la ligne de laquage de l'usine de Montataire

(voir figures 5.30 et 5.31), et plusieurs tests ont alors été effectués pour vérifier la stabilité de l'acquisition lorsque la vitesse de la ligne augmente : pour des vitesses de ligne allant de 10m/mn à 180m/mn, la qualité des images reste parfaitement constante. Les premières acquisitions en situation "réelle", c'est à dire en période de production, ont donc pu être réalisées, et les analyses de ces données sont en cours. Les résultats seront notamment comparés aux résultats de l'inspection visuelle menée en parallèle.

Fort de ces résultats encourageants, l'objectif affiché par Arcelor Research est de pouvoir proposer dès que possible un "package" d'analyse de la texture des surfaces industriellement exploitable.

5.6. Autres applications

Si tout au long de notre travail, nous avons mis l'accent sur les surfaces de types peau d'orange, c'est essentiellement parce qu'une base de surfaces était disponible en raison d'études parallèles menées sur la caractérisation de leur aspect par un capteur de type wave-scan. Ensuite, lors du développement de nouveaux outils, il nous a paru judicieux de conserver la même série de surfaces test de façon à pouvoir observer et quantifier facilement les effets de telle ou telle modification des techniques d'analyse employées. Le risque d'une telle approche se situe dans la trop grande spécialisation des techniques développées qui pourraient alors ne s'appliquer efficacement qu'à une seule série de surfaces. Tout au long de notre travail, nous avons cherché à éviter cet écueil en cherchant à être le plus générique et le plus transparent possible dans la méthodologie développée. Même si nous n'avons pas mené de campagne de tests aussi poussée que pour les surfaces peau d'orange, nous avons quand même effectué quelques essais pour d'autres types de surfaces (verniss structurés par exemple), et les résultats obtenus sans aucune modification des paramètres d'analyse étaient intéressants.

Outre le contrôle qualité proprement dit, notre méthode d'analyse a aussi été utilisée pour des études plus prospectives. Après le passage dans un bain de zinc, le produit est soumis à un jet de gaz issu d'une tuyère destiné à accélérer l'opération de séchage. Si ce jet est trop fort, on voit apparaître des motifs d'ondulation sur le revêtement de zinc, et des études sont en cours pour mieux comprendre ce phénomène. Dans ce cadre, nos outils permettent de caractériser des images représentant l'épaisseur du film de zinc en fonction de la vitesse du produit et de la pression du jet de gaz.



FIG. 5.31.: Prototype d'acquisition en ligne d'images caméra

Troisième partie .

Caractérisation indirecte : Structure d'arbre et Propriétés Optiques

Résumé

Nous présentons ici des caractérisations indirectes des surfaces, pouvant conduire à de nouveaux descripteurs ou au calcul de paramètres topographiques à partir d'une réponse optique :

- Nous utilisons un arbre d'inondation pour représenter puis caractériser une surface.
- Nous développons des algorithmes pour appliquer des techniques de bootstrap à ces descripteurs particuliers.
- Nous modélisons le comportement optique d'une surface et nous simulons l'acquisition d'une image par une caméra.
- Nous envisageons le problème inverse pour reconstruire la topographie (carte des altitudes) d'une surface à partir d'une seule image acquise par une caméra.

Chapitre 6.

Représentation arborescente

Dans notre processus d'étude des surfaces, plusieurs étapes de caractérisation se succèdent, chacune d'entre elles se base sur une représentation particulière résumant diverses informations au sujet de la surface. La base de la pyramide, c'est à dire l'objet que l'on cherche à étudier par ces différentes représentations n'est autre qu'une large bande de produit que l'on assimile à une fonction aléatoire. Le premier degré d'étude consiste à considérer un échantillon physique de taille limitée, c'est à dire une réalisation particulière de la fonction aléatoire. Cet échantillon possède des caractéristiques (topographiques, optiques, physiques ...) que l'on suppose proches de celles de la fonction aléatoire sous-jacente. Cet échantillon est ensuite observé au moyen d'un appareil de mesure (microscope interférométrique, caméra, système Ondulo, etc) dont la sortie est une nouvelle représentation qui dépend d'un certain nombre de propriétés physiques (pas toutes) de l'échantillon. On peut ensuite effectuer un traitement numérique (calcul d'un histogramme par exemple) pour obtenir une représentation synthétique de la sortie de l'appareil de mesure. Une procédure d'analyse des données permet encore de résumer l'information issue de plusieurs traitements numériques en une ou deux variables réelles, puis un algorithme de classification condense toute notre connaissance sur une surface en une modalité d'une variable discrète. Finalement, toutes ces opérations résultent de la même volonté de simplification ou de compression de l'information disponible. Chacune des représentations utilisées possède des propriétés spécifiques qui s'adaptent à différents problèmes.

Dans ce chapitre, nous proposons d'utiliser une représentation intermédiaire des surfaces qui s'intercale entre l'image de sortie d'un capteur et l'extraction de descripteurs utilisables dans une phase d'analyse statistique : il s'agit de créer *une structure arborescente* à partir d'une image et de l'exploiter ensuite dans un but de classification [26]. Les arbres que nous utilisons sont obtenus à partir d'un processus *d'inondation* d'une image de la surface. Ils représentent la manière dont s'agencent et fusionnent les zones sombres de l'image. Ces arbres d'inondations sont très souvent utilisés en morphologie mathématique pour construire des hiérarchies de partitions à partir d'une ligne de partage des eaux (watershed) [29, 14, 11, 10]. Plusieurs niveau de segmentation sont alors disponible, et l'utilisateur peut choisir celui qui le convient le mieux en fonction de divers critères (nombre de région, taille des régions, etc ...). Les structures d'arbres que nous étudions sont aussi conceptuellement proches des arbres de classification utilisés pour diviser un ensemble de points (ici, les pixels de l'image) en plusieurs catégories en fonction d'un critère d'agrégation (variance intra-classe et inter-classe par exemple).

En fait, plusieurs solutions sont possibles pour obtenir des arbres à partir d'une image, et notre objectif ici n'est pas tellement de défendre l'une d'entre elle, ni même d'en faire une étude comparative, mais plutôt de définir un cadre de travail général dans lequel on puisse adapter une représentation arborescente à notre méthodologie globale d'étude des surfaces rugueuses, ce qui implique notamment l'adaptation des techniques de bootstrap à l'étude des arbres. Ce chapitre se veut donc une étude préalable à l'utilisation d'arbres pour la caractérisation et la classification de textures, mais les étapes de validation et d'optimisation des différents descripteurs possibles issus de cette représentation pour un problème industriel précis n'ont pas été approfondies.

Dans une première partie, nous décrivons l'algorithme de construction de l'arbre par inondation que nous avons utilisé, et nous citons d'autres mécanismes de construction possibles menant à des arbres différents. Nous cherchons ensuite à extraire des descripteurs à partir de ces arbres et à les utiliser pour comparer des surfaces. Les difficultés rencontrées pour appliquer les techniques de sur-échantillonnage nous conduisent ensuite à chercher des algorithmes de division d'un arbre en sous-arbres, dont l'exploitation facilite considérablement le calcul d'échantillons bootstrap, et peut aussi conduire à des méthodes de segmentation de l'image analysée. Enfin, ces travaux étant très largement prospectifs, nous présentons quelques axes de recherche qu'il nous paraîtrait intéressant d'approfondir.

6.1. Construction d'un arbre

6.1.1. Notations

Posons tout d'abord quelques notations : un arbre T est constitué d'un ensemble de noeuds $\mathcal{N}$ apparenté à $[1..n]$ (n est le nombre de noeuds de l'arbre) et d'un ensemble d'arêtes $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$: il existe une arête reliant le noeud i au noeud j si et seulement si $(i, j) \in \mathcal{E}$. Un arbre est non cyclique, c'est à dire qu'il n'existe pas de noeud n_0 pour lequel on pourrait trouver un ensemble de noeuds $n_1 n_2 \dots n_r$ tous différents tels que $n_r = n_0$ et $\forall i \in [0..r-1], (n_i, n_{i+1}) \in \mathcal{E}$.

L'arbre sera supposé orienté et ordonné par rapport à une racine notée R , et on note $f(i)$ le père du noeud i ($f(R)$ n'est pas défini). On désigne par $\mathcal{C}(i)$ l'ensemble des fils du noeud i ($\mathcal{C}(i) \subset \mathcal{P}(\mathcal{N})$).

Le sous-arbre dont i est la racine sera noté $S(i)$. On introduit la relation d'inclusion sur les sous-arbres fournissant une relation d'ordre partielle sur les noeuds : on note $i \prec j$ si et seulement si $S(i) \subset S(j)$. On note $i \not\prec j$ si cette relation n'est pas vérifiée.

6.1.2. Inondation

Nous reprenons dans ce paragraphe l'ensemble des notations utilisée dans le paragraphe concernant les tortuosités.

Soit I une image définie sur D à valeur dans $\mathbb{R}^+$. On s'intéresse aux minima régionaux de I qui sont des ensembles de pixels de même valeur, c'est à dire des zones plates

(éventuellement réduites à un point), que l'on notera $M_1, M_2, \dots, M_K$:

$$\forall p \in D / \bigcup_{i=1}^K M_i, \forall k \in [1..K], \forall p_0 \in M_k, \forall C \in \text{chem}(p_0, p), \exists q \in C : I(q) > I(p_0)$$

On s'intéresse ensuite à une longueur définie par :

$$\forall C = p_0 p_1 \dots p_r, l(C) = \sup_{i \in [0..r]} (I(p_i))$$

Cette longueur induit une (pseudo-)distance ultramétrique (car $d(p, p) \neq 0$) entre les pixels :

$$d(p_1, p_2) = \inf_{C \in \text{chem}(p_1, p_2)} (l(C))$$

Le principe de l'inondation est le suivant : on considère l'image comme un relief topographique (même si il s'agit d'une image caméra) en associant le niveau de gris de l'image à une altitude. On imagine de l'eau s'infiltrant depuis le sol et montant dans le relief à vitesse constante. Cette eau va commencer par jaillir au niveau du minimum régional le plus bas en formant un petit lac. Au fur et à mesure que l'eau monte, plusieurs minima régionaux vont être atteints et les lacs associés à chacun d'eux vont grossir progressivement. De temps à autre deux lacs fusionnent lorsque l'eau dépasse le point col séparant les deux vallées associées. Le phénomène se poursuit jusqu'à ce qu'il n'y ait plus qu'un seul lac. Pour construire un arbre, on considère les minima régionaux comme des feuilles, et à la première fusion de lac, on crée un noeud interne à l'arbre, et deux arêtes reliant ce noeud aux minima régionaux ayant fusionné. On poursuit ensuite la construction en ajoutant un noeud et deux arêtes à chaque fusion de lac (voir figure 6.1).

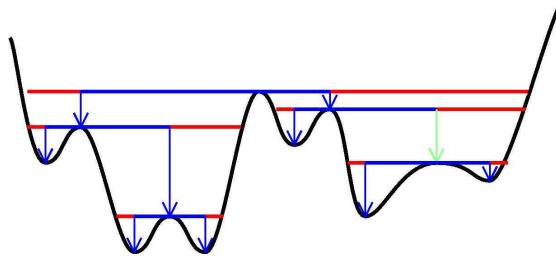


FIG. 6.1.: Construction d'un arbre par inondation

Cette vision imagée de la construction peut se retranscrire en termes de boules pour la distance définie précédemment. On désigne par $\mathcal{B}(p_0, \lambda)$ l'ensemble des points tels que $d(p_0, p) \leq \lambda$. On remarque qu'une boule peut-être l'ensemble vide si $I(p_0) > \lambda$. Les propriétés de $d(x, y)$ font que deux boules de même rayon sont soit disjointes, soit confondues (propriétés d'une distance ultramétrique). Cette dernière remarque nous permet d'élargir sans ambiguïté la notation $\mathcal{B}(p_k, \lambda)$ à $\mathcal{B}(M_k, \lambda)$ (que l'on notera B_λ^k si $p_k \in M_k$).

On pose :

$$\mathcal{N}^0 = [1..K]$$

$$\begin{aligned}\mathcal{E}^0 &= \emptyset \\ \mathcal{A}^0 &= [1..K]\end{aligned}$$

On considère $h_0 = \min_{k \in [1..K]}(I(p), p \in M_k)$ (qui est aussi la valeur du minimum de l'image). D'après les propriétés des minima régionaux, les boules $B_{h_0}^1, \dots, B_{h_0}^K$ qui ne sont pas vides sont deux à deux disjointes (dès lors qu'il y a plusieurs minima). On considère alors la valeur h_1 telle que :

$$h_1 = \inf(h \geq h_0 / \exists i, j \in \mathcal{A}^0 : B_h^i = B_h^j \neq \emptyset)$$

La valeur h_1 correspond à la hauteur d'eau pour laquelle les deux premiers lacs fusionnent. On construit alors :

$$\begin{aligned}\mathcal{N}^1 &= \mathcal{N}^0 \cup \{K+1\} \\ \mathcal{E}^1 &= \mathcal{E}^0 \cup \{(K+1, i), (K+1, j)\} \\ \mathcal{A}^1 &= (\mathcal{A}^0 / \{i, j\}) \cup \{K+1\} \\ \forall \lambda \geq h_1, B_\lambda^{K+1} &= B_\lambda^i = B_\lambda^j\end{aligned}$$

En supposant déterminé $h_1 \leq \dots \leq h_k$, $\mathcal{N}^1, \dots, \mathcal{N}^k, \mathcal{E}^1, \dots, \mathcal{E}^k, \mathcal{A}^1, \dots, \mathcal{A}^k$, en supposant également définis les boules $B_\lambda^{K+1}, \dots, B_\lambda^{K+k}$ pour $\lambda \geq h_k$, le tout pour $k < K-1$, alors on peut construire h_{k+1} de telle façon que :

$$h_{k+1} = \inf(h \geq h_k / \exists i, j \in \mathcal{A}^k : B_h^i = B_h^j \neq \emptyset)$$

Et on actualise ensuite les ensembles :

$$\begin{aligned}\mathcal{N}^{k+1} &= [1..K+k+1] \\ \mathcal{E}^{k+1} &= \mathcal{E}^k \cup \{(K+k+1, i), (K+k+1, j)\} \\ \mathcal{A}^{k+1} &= (\mathcal{A}^k / \{i, j\}) \cup \{K+k+1\} \\ \forall \lambda \geq h_k, B_\lambda^{K+k+1} &= B_\lambda^i = B_\lambda^j\end{aligned}$$

Cette construction par récurrence se poursuit tant que $\mathcal{A}^k$ possède deux éléments ou plus, c'est à dire tant que $k < K$. Cela nous permet de définir un ensemble $\mathcal{N}^{K-1} = [1..2K-1]$ et un ensemble d'arêtes $\mathcal{E}^{K-1}$ dont la réunion définit un arbre de racine $R = 2K-1$. Cet arbre est binaire : chaque noeud possède soit 0 soit 2 fils. En fait, là réside la seule ambiguïté de cette construction : lorsque plus de deux boules fusionnent à la même étape, nous avons implicitement (dans le $\exists i, j$) choisi deux noeuds arbitrairement et nous les avons fusionnés. Nous aurions pu choisir de construire un arbre non binaire en construisant à chaque étape un ensemble G^k :

$$\forall k \in [1..K-1], G^k = \{i \in \mathcal{A}^{k-1} / \exists j \in \mathcal{A}^k : B_{h_k}^i = B_{h_k}^j \neq \emptyset\}$$

On définit ensuite une classe d'équivalence $\mathcal{R}$ sur G^k par :

$$\forall i, j \in G^k, i \mathcal{R} j \Leftrightarrow B_{h_k}^i = B_{h_k}^j$$

Pour chaque classe d'équivalence C_s , $s \in [0..q]$ de $\mathcal{R}$, on peut introduire un noeud $K+k+s$ dans l'arbre dont les fils sont tous les points de C_s , et si certaines classes d'équivalence comportent plus de deux points, l'arbre n'est plus binaire. Ces nouveaux noeuds *remplacent* plusieurs noeuds de la construction binaire classique.

Une telle approche assurerait une construction de l'arbre sans ambiguïtés (c'est à dire sans choix arbitraire). En pratique cependant, l'algorithme utilisé pour la construction de l'arbre se base sur la propagation d'une distance à partir de tous les minima régionaux plutôt qu'en recherchant effectivement les valeurs h_k en testant toutes les boules possibles. Cette propagation utilise des files d'attentes hiérarchiques et un algorithme similaire à celui du calcul des tortuosités dans lequel l'ensemble source aurait plusieurs composantes connexes (ces composantes sont les minima régionaux). Un label différent est attribué à chacune d'entre elles, et lors de la propagation de la distance (voir 2.2.4), on calcule non seulement la distance séparant chaque pixel de l'ensemble source, mais aussi *le label* de la composante connexe dont il est le plus proche. Dès qu'un pixel est à égale distance de deux composantes connexes *différentes* de l'ensemble source, on ajoute un noeud à l'arbre, on fusionne les labels correspondant à ces composantes et on continue la propagation. Dans cet algorithme, il est beaucoup plus simple de traiter séquentiellement les fusions multiples en fonction de l'ordre dans lequel les pixels sont visités plutôt que de construire un noeud ayant strictement plus de deux fils. Bien que nous conservions dans la suite cette méthode de construction, les algorithmes de traitement que nous utiliserons agissant sur les arbres eux-mêmes ne nécessitent généralement pas la binarité de l'arbre, et dans ce cas, leur adaptation au cas non binaire est le plus souvent triviale. La structure arborescente pourrait aussi être modifiée a posteriori pour lever les ambiguïtés de la construction binaire, mais comme nous l'avons dit, nous sommes plus intéressés ici par le traitement de l'arbre que par le choix de l'arbre en lui-même, et nous conserverons donc des arbres binaires pour les exemples présentés dans la suite.

En créant un arbre d'inondation (parfois appelé dendrogramme), on traite différemment les structures négatives (vallées) de l'image et les structures positives (pics). Il est bien entendu possible d'appliquer un algorithme symétrique en partant des maxima régionaux de l'image plutôt que des minima si l'on s'intéresse plus particulièrement aux structures positives.

D'autres approches ont été proposées pour construire un arbre dont le comportement vis à vis des minima et des maxima de l'image serait équivalent, c'est le cas par exemple du dendrone symétrique introduit dans [63], ou de "l'arbre de changement" (change tree) décrit dans ([38]). L'idée de cette dernière approche est d'observer les changements de topologie des courbes de niveaux de l'image en fonction de l'altitude (du niveau de gris) : les extrema sont des courbes de niveaux qui apparaissent ou disparaissent (ce seront les feuilles de l'arbre), et les points cols sont des fusions ou des divisions de lignes de niveaux (ce seront les noeuds internes). Ces différentes approches ont le mérite d'intégrer plus d'information à l'intérieur d'un même arbre et cette dernière a été reprise par Scott [36] dans un projet de certification ISO pour la caractérisation des surfaces.

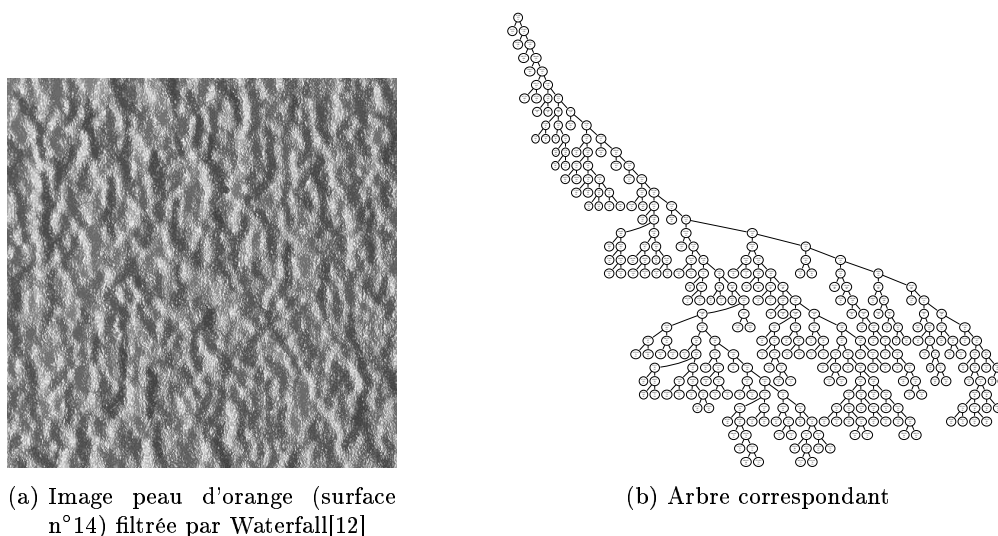


FIG. 6.2.: Arbre associé à une image de type peau d'orange.

6.2. Extraction de descripteurs

Une fois l'arbre construit, il s'agit de l'exploiter pour caractériser la surface dont il est issu, et il faut pour cela rendre possible la comparaison de deux ou plusieurs arbres. Une approche possible consiste à choisir des descripteurs d'arbre, tout comme nous avons choisi des descripteurs de carte topographique ou d'image caméra, et de les utiliser dans le même schéma d'analyse.

6.2.1. Pondération des noeuds et des arêtes

L'arbre que nous avons construit jusqu'à présent, quelle que soit sa nature exacte, contient des informations exclusivement topologiques sur la surface. Il est par exemple parfaitement insensible à un changement d'échelle ou à une déformation ne modifiant pas l'agencement des minima. Les caractéristiques de la surface résumées par un tel arbre peuvent donner naissance à un premier jeu de descripteurs, mais il est des applications pour lesquelles la seule topologie n'est pas suffisante et où d'autres renseignements sur la surface sont requis. Nous allons donc nous efforcer d'introduire des informations dans l'arbre par le biais de poids numériques attribués aux noeuds ou aux arêtes.

Si l'on se cantonne à l'information topologique, un noeud peut être valué par de nombreux indices, parmi lesquels on peut citer les plus couramment utilisés :

- sa génération, c'est à dire le nombre de noeud le séparant de la racine.
- le nombre de noeuds ou le nombre de génération du sous-arbre dont il est la racine
- un facteur de symétrie caractérisant la différence entre le nombre de noeud du sous-arbre induit par son fils droit et par son fils gauche.

Pour obtenir un descripteur d'arbre, on peut par exemple :

- Choisir les variables d'un noeud particulier (la racine) : nombre de noeuds de l'arbre (nombre de générations dans l'arbre...)
- Faire la moyenne sur l'ensemble des noeuds (facteur de symétrie global...)
- Faire un histogramme pour l'ensemble des noeuds (nombre de noeuds par génération, distribution des facteurs de symétrie...)

Ces indicateurs informent sur la nature des fusions : la surface est-elle un grand bassin principal dans lequel se fondent successivement de petits bassins "simples", c'est à dire ne contenant qu'un seul extremum régional (arbre très "longiligne" avec beaucoup de générations et peu de noeuds), ou au contraire est-elle le résultat de fusions très "parallèles" au cours desquelles de petits bassins grossissent indépendamment pour ne fusionner ensemble qu'à la fin du processus (arbre binaire complet).

Pour intégrer maintenant une information numérique quantitative sur la surface, nous revenons à une interprétation plus physique des arêtes et des noeuds de l'arbre en terme d'inondation : chaque noeud représente la fusion de deux lacs, et peut être associé à un point col de la surface, de la même façon, chaque arête peut être associée à un lac. Grâce à ces associations, on peut introduire des poids w_i et $w_{(i,j)}$ définis pour les noeuds et les arêtes de l'arbre. Ces valeurs représentent des caractéristiques des ensembles de D correspondant, par exemple la surface, le volume ou la profondeur d'un bassin pour une arête et l'altitude d'un point col pour un noeud. Là encore, de nombreuses possibilités sont envisageables (facteur de forme du bassin, allongement selon une direction, dimension fractale, etc...) selon la nature des surfaces ou les applications auxquelles elles sont destinées.

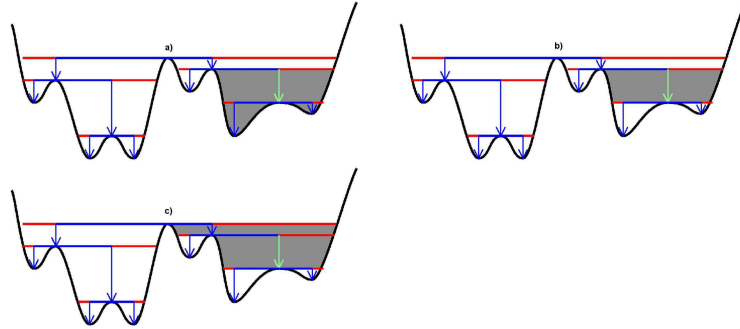


FIG. 6.3.: Différentes stratégies de pondération basées sur le volume des lacs. a) Volume du lac (construction standard). b) Incrément de volume. c) Volume cumulé depuis la racine.

Ces descripteurs numériques viennent s'ajouter aux descripteurs topologiques, et on obtient alors pour chaque noeud et pour chaque arête un poids vectoriel formé de plusieurs attributs. Pour lesquels de nombreuses combinaisons de descripteur deviennent possibles (histogrammes mono- ou bivariés, moyennes ...). Dans le cas de l'arbre il est presque toujours possible de se ramener au cas où seuls les noeuds sont évalués, soit en utilisant l'association canonique existant entre une arête et le noeud fils qui lui correspond, soit en fusionnant les informations contenues dans l'ensemble des arêtes filles associées à un

noeud. Nous supposons donc implicitement dans la suite que seuls les noeuds sont pondérés, et pas les arêtes.

6.2.2. Exploitation de cette pondération

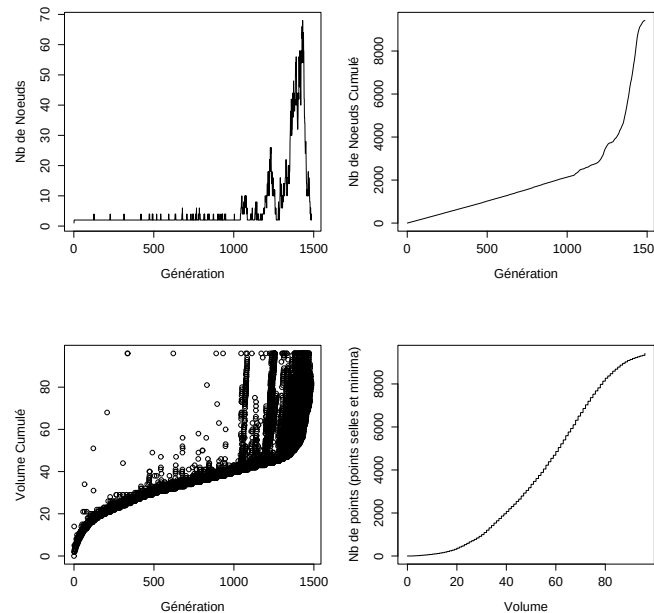


FIG. 6.4.: Exploitation des poids : construction de l'histogramme du nombre de noeud par génération, et de sa version cumulée. Représentation pour chaque noeud du volume cumulé depuis la racine, Représentation du nombre de noeuds de l'arbre en fonction du volume cumulé depuis la racine.

La pondération des noeuds nous permet d'extraire un grand nombre de descripteurs d'arbre (histogrammes des poids, valeurs moyennes des poids des noeuds, poids de la racine,... Voir figure 6.4), ce qui pourrait laisser croire que le travail sur les arbres est terminé : il suffit de choisir les descripteurs adaptés à un problème précis et de les insérer dans l'analyse factorielle au même titre que les variables d'images. Il n'en est rien, car nous faisons aussi appel à des méthodes de bootstrap nous permettant d'évaluer les distributions des descripteurs choisis, et cet aspect de l'étude des surfaces est à notre avis tout à fait essentiel. Or la façon d'obtenir des échantillons bootstrap à partir d'arbres est une opération qui n'apparaît pas complètement évidente, voire même assez complexe vu la forte dépendance de la plupart des variables considérées pour deux noeuds proches dans l'arbre. La méthode naïve consistant à tirer des noeuds au hasard avec remise au sein d'un même arbre conduit d'ailleurs à des résultats désastreux.

Un autre problème se pose lors de l'exploitation des arbres : la taille de l'image peut faire varier considérablement la taille de l'arbre obtenu, et la plupart des descripteurs que

nous considérons varient aussi beaucoup dans ce cas (volumes et surfaces des lacs plus important pour les noeuds proche de la racine, mais altitude comparable par exemple). Or si la fonction aléatoire sous-jacente ne change pas, certaines propriétés communes devraient se retrouver d'un arbre sur l'autre, même si la taille de l'image a variée.

Pour parer à ces difficultés, nous avons choisi d'utiliser une approche inhérente à la notion de texture, c'est à dire au fait que nous regardons la réalisation d'une fonction aléatoire stationnaire et ergodique. Si l'on considère des surfaces de taille comparable (et suffisante), les deux arbres correspondants devraient se ressembler, au sens où il sont issus d'un même processus aléatoire. De la même façon, on devrait retrouver à l'intérieur d'une image plusieurs sous-images se ressemblant et conduisant chacune à des arbres ressemblants. L'idée (intuitive) sous-jacente consiste à dire que l'arbre obtenu à partir d'une surface n'est rien d'autre qu'un agencement de *sous-arbres indépendants* issus d'un même processus aléatoire. En terme de vallées, cela revient à considérer deux phénomènes distincts : jusqu'à une "certaine taille" (ce critère reste à définir), les vallées, même constituées de plusieurs minima régionaux, sont issues du même événement, les vallées plus larges (au dessus de cette taille caractéristique) sont issues de la fusion de deux événements indépendants. Un exemple simple consiste à imaginer une projection d'acide sur une surface : chaque goutte va entraîner un phénomène de corrosion indépendant, c'est à dire une vallée, dont la profondeur, le nombre de minima régionaux et la forme générale seront liés à une seule goutte. Ces vallées s'organisent ensuite entre elles selon un processus dépendant de la densité des gouttes.

La notion d'agencement de sous-arbres indépendants est une modélisation, qui n'est sans doute qu'approchée en réalité, mais qui a de gros avantages pratiques, puisqu'elle permet de résoudre les deux problèmes évoqués au début de cette section. Si l'on arrive à diviser un arbre en plusieurs sous-arbres pertinents, on obtient un échantillon de plusieurs individus indépendants, ce qui nous place dans le cas le plus simple de bootstrap où l'on peut tirer au hasard avec remise un nouvel échantillon parmi les individus disponibles (voir section 3.3.1). Les descripteurs considérés sont alors des moyennes ou des histogrammes de variables (éventuellement vectorielles) calculés sur chacun des sous-arbres. Par ailleurs, la taille de l'image jouera éventuellement sur le nombre de sous-arbres extraits, donc sur la variance calculée par les méthodes de bootstrap, mais les descripteurs seront calculés à partir de sous-arbres de taille voisine, et resteront donc comparables d'une image à l'autre, même si leurs dimensions ne sont pas identiques.

On pourrait aussi appliquer le même genre d'approche que pour certains descripteurs d'image, c'est à dire reconstruire une image bootstrap en tirant des blocs au hasard dans l'image initiale. Tout d'abord, cette méthode introduit des effets de bords gênant pour la construction de l'arbre qui est une structure très globale et pas locale comme certains des descripteurs d'images considérés. Ensuite la division de l'arbre lui-même, plutôt que la division de l'image, permet de conserver les structures d'intérêt sans distorsion, car les vallées ne sont en général pas des blocs carrés à l'intérieur de l'image originale.

6.3. Divisions en sous-arbres

Dans cette section, nous explicitons plusieurs méthodes permettant de sélectionner les sous-arbres à conserver. Soit $T = (\mathcal{N}, \mathcal{E})$ un arbre. Les sous-arbres sont assimilés aux noeuds de l'arbre : on cherche donc un ensemble $A \subset \mathcal{N}$ tel que :

$$\forall i, j \in A, i \not\prec j \quad (6.1)$$

En effet, l'objectif est d'obtenir des sous-arbres que l'on puisse raisonnablement considérer comme indépendants : il est pour cela nécessaire qu'ils soient disjoints.

6.3.1. Critères simples

La manière la plus simple de procéder est de fixer effectivement un critère de taille sur les sous-arbres à sélectionner. Soit p un paramètre croissant défini sur les noeuds de l'arbre :

$$\begin{aligned} p : \mathcal{N} &\rightarrow \mathbb{R} \\ n &\rightarrow p(n) \end{aligned}$$

est dit croissant si et seulement si :

$$\forall i, j \in \mathcal{N}, i \prec j \Rightarrow p(i) \leq p(j)$$

Un critère κ est une variable booléenne définie sur les noeuds de $\mathcal{N}$. Dans ce qui suit, nous utiliserons le critère (croissant) κ_t défini par :

$$\begin{aligned} \forall n \in \mathcal{N} \quad , \quad \kappa_t(n) &= 1 \text{ si } p(n) \geq t \\ \kappa_t(n) &= 0 \text{ sinon} \end{aligned}$$

A partir de p et κ , on définit un ensemble de sous-arbres disjoints :

$$A_1 = \{n \in \mathcal{N} / \kappa_t(n) = 1 \text{ \& } (\forall m \in \mathcal{N}, m \prec n \Rightarrow \kappa_t(m) = 0)\} \quad (6.2)$$

La deuxième condition permet d'assurer la propriété 6.1.

Cette méthode permet par exemple de sélectionner les plus petits sous-arbres (pour $\prec$) dont le nombre de noeuds est supérieur ou égal à une valeur seuil (t) fixée. Cela permet aussi de sélectionner les plus petits sous-arbres pour lesquels les lacs correspondants ont une surface ou un volume minimum fixé.

Une seconde stratégie consiste à considérer le critère (non croissant) $\kappa_{s,t}$ défini par :

$$\begin{aligned} \forall n \in \mathcal{N} \quad , \quad \kappa_{s,t}(n) &= 1 \text{ si } s \geq p(n) \geq t \\ \kappa_{s,t}(n) &= 0 \text{ sinon} \end{aligned}$$

On peut construire l'ensemble de sous-arbres disjoints :

$$A_2 = \{n \in \mathcal{N} / \kappa_{s,t}(n) = 1 \text{ \& } (\forall m \in \mathcal{N}, n \prec m \Rightarrow \kappa_{s,t}(m) = 0)\}$$

On sélectionne dans ce cas les plus grands sous-arbres vérifiant le critère $\kappa_{s,t}$ (voir figure 6.5). La limite supérieure imposée sur p assure de ne pas obtenir systématiquement l'ensemble trivial $A = \{R\}$ contenant uniquement la racine de l'arbre. On peut par exemple choisir comme sous-arbres pertinents les plus grands sous-arbres dont la surface est comprise entre deux bornes fixées.

Le principal problème de cette approche est le choix du critère qui contraint fortement le résultat final. Cela n'est pas anodin, car seuls les sous-arbres choisis seront pris en compte dans la suite : les noeuds $n \notin A$ tels que $\exists a, b \in A : a \prec n \& b \prec n$ sont considérés comme des noeuds de liaisons entre des phénomènes indépendants que nous n'analysons pas dans un premier temps (nous reviendrons sur ce point dans les perspectives), et les noeuds de l'arbre qui ne vérifient pas cette condition et qui ne font pas partie de l'un des sous-arbres sélectionnés sont considérés comme du bruit. Si l'on ne tient pas compte du bruit, on peut dire que les différences entre les surfaces se trouvent soit dans les phénomènes de liaison, soit dans les "motifs" correspondants aux sous-arbres choisis (soit dans les deux !). L'approche que nous envisageons ici privilégie implicitement la seconde hypothèse, puisque seuls les sous-arbres de A sont utilisés pour caractériser les surfaces. Or dans le mode de sélection par critère, les sous-arbres de A sont naturellement assez proches d'une surface à l'autre puisqu'ils partagent tous une caractéristique commune, et cette approche peut donc avoir tendance à concentrer la différence entre surfaces à l'intérieur des noeuds de liaisons. C'est la raison pour laquelle nous avons par la suite cherché à atténuer la contrainte forte imposée aux sous-arbres par le choix du critère.

6.3.2. Algorithme plus complexe

Lorsque l'on choisit un critère, on part du principe que les sous-arbres pertinents ressemblent à un modèle défini, disons pour une variable p . Le fait que les sous-arbres pertinents doivent se ressembler, pour une certaine définition du terme "ressembler", semble être une supposition raisonnable, mais utiliser la transitivité de la relation "ressembler à", c'est à dire considérer que deux sous-arbres se ressemblent *parce qu'ils ressemblent à un modèle précis* peut conduire aux problèmes que nous avons évoqués. L'idée simple qui sous-tend l'algorithme que nous allons présenter est donc de choisir les sous-arbres qui se ressemblent le plus tout en étant le plus différent possible des sous-arbres rejetés, le tout sans être obligé de préciser à quoi ces sous-arbres ressemblent.

La première difficulté consiste à définir le terme "ressembler". En pratique, les poids des noeuds (noté $w(n)$) sont des vecteurs synthétisant certaines propriétés du sous-arbre associé, et nous considérerons une forme de variance de ces vecteurs calculés sur l'ensemble des sous-arbres d'un ensemble A : les sous-arbres de A se ressemblent d'autant plus que cette variance est faible. Soit $A \subset \mathcal{N}$ un ensemble (fini) de noeuds, soit $\#A$ le nombre de noeuds de A :

$$\begin{aligned} M(A) &= \frac{1}{\#A} \sum_{n \in A} w(n) \\ V(A) &= \frac{1}{\#A} \sum_{n \in A} \|w(n) - M(A)\|^2 \end{aligned}$$

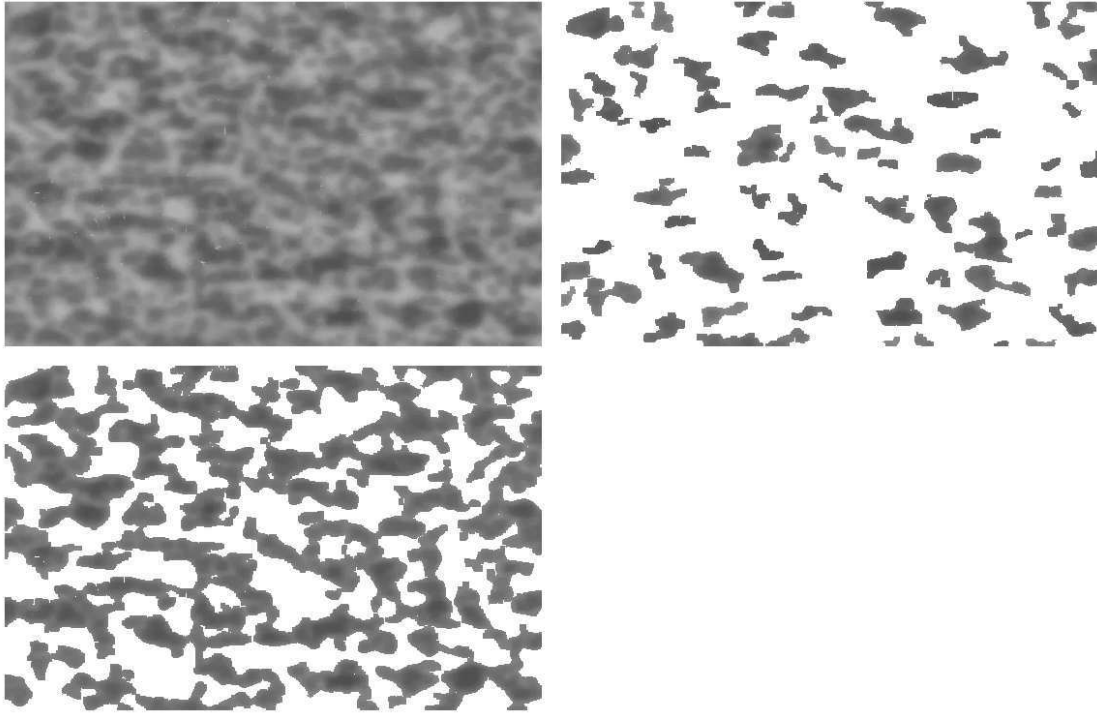


FIG. 6.5.: Exemple de division basée sur le degré d'un noeud. L'image initiale est en haut à gauche. En haut à droite, on ne conserve le niveau de gris original que pour les sous-arbres comportant exactement 2 feuilles. En bas à gauche, on ne conserve le niveau de gris original que pour l'ensemble des sous-arbres contenant entre 3 et 50 feuilles.

V définit ainsi une mesure de la dissimilarité des éléments à l'intérieur de A (nous cherchons à minimiser cette quantité).

Nous voulons aussi mesurer la similarité entre les noeuds de A et les noeuds n'appartenant pas à A (c'est à dire les noeuds du complémentaire de A noté $\bar{A}$). Plusieurs essais algorithmiques nous ont conduit de façon expérimentale à choisir comme mesure de cette similarité la quantité :

$$\Lambda(A, \bar{A}) = \min_{n \in \bar{A}} (\|w(n) - M(A)\|)$$

Nous cherchons à maximiser cette quantité.

Ce problème est comparable à celui d'une classification hiérarchique où l'on cherche à créer des groupes d'individus en maximisant la variance inter-groupe tout en minimisant la variance intra-groupes. La différence dans le cas des arbres provient de la condition 6.1 qui empêche de choisir deux sous-arbres reliés par une relation d'inclusion.

En pratique, nous ne sommes pas parvenus à développer un algorithme efficace menant à un ensemble vérifiant les propriétés voulues de façon exacte. Nous avons donc choisi de

minimiser une fonction de coût adéquate de façon itérative et randomisée pour parvenir à un résultat s'en approchant le plus possible.

La fonction de coût que nous proposons d'utiliser est la suivante :

$$\Phi(A) = \alpha_1.V(A) - \alpha_2.\Lambda(A, \bar{A})$$

Où α_1 et α_2 sont des constantes positives.

Nous initialisons un ensemble A_0 contenant toutes les feuilles de l'arbre (c'est à dire les minima régionaux dans le cas standard). Supposons l'ensemble A_k défini. Nous calculons alors $\Phi(A_k)$, puis nous choisissons au hasard un ordre pour les éléments de A_k . Nous traitons alors le premier noeud n en testant l'influence de plusieurs opérations sur la fonction de coût :

- n peut être simplement retiré de la liste des sous-arbres pertinents : $A_{k+1} = A_k/\{n\}$.
- Soit $f(n)$ le père de n . Si tous les noeuds m tels que $m \prec f(n)$ et $m \not\prec n$ n'appartiennent pas à A , on peut *remplacer* n par $f(n)$ dans la liste des sous-arbres sélectionnés : $A_{k+1} = A_k/\{n\} \cup \{f(n)\}$
- Soit $f(n)$ le père de n . Si plusieurs fils de $f(n)$ appartiennent à A , alors on peut remplacer *tous ces noeuds* par $f(n)$ dans A : $A_{k+1} = A_k/(\mathcal{C}(f(n)) \cap A_k) \cup \{f(n)\}$

Ces trois opérations assurent que A_{k+1} vérifie la propriété 6.1 pourvu que A_k la vérifie (la propriété est trivialement vérifiée pour A_0). Si pour l'une de ces opérations $\Phi(A_{k+1}) \leq \Phi(A_k)$, on conserve les modifications, et on itère le processus à partir du choix de l'ordre de traitement des noeuds de A_{k+1} . Si au contraire aucune de ces modifications n'améliore la fonction de coût, on passe au noeud suivant de A_k et on procède aux mêmes essais. Si tous les noeuds de A sont testés sans succès, on conserve la configuration obtenue, et les sous-arbres choisis sont ceux dont la racine est dans A .

La convergence de cet algorithme vers une forme de minimum de la fonction de coût est assurée. En effet, lorsque l'on remplace un ou deux sous-arbres de A_k par un autre sous-arbre, la génération de ce dernier est strictement inférieure à celle du ou des sous-arbres remplacés, on ne fait donc que "monter" dans l'arbre sans jamais redescendre, il n'y a pas de phénomène "oscillatoire" qui pourrait contrecarrer la convergence de l'algorithme. Le résultat obtenu cependant, n'est pas toujours satisfaisant : comme les feuilles de l'arbre se "ressemblent" beaucoup, et sont assez différentes des autres noeuds, l'ensemble final retenu est souvent très proche de l'ensemble A_0 . Pour obtenir une solution acceptable sans changer la phase d'initialisation, il faudrait passer par des états intermédiaires pour lesquels la fonction de coût augmente avant de décroître à nouveau, et de telles opérations sont proscrites par l'algorithme. Pour corriger ce défaut, nous avons introduit des termes de régularisation dans la fonction de coût faisant intervenir la surface globale de l'image de départ couverte par les sous-arbres retenus que l'on cherche à maximiser et leur nombre que l'on cherche à approcher le plus possible d'une cible Nb .

La nouvelle fonction de coût devient :

$$\Phi'(A) = \alpha_1.V(A^0) - \alpha_2.\Lambda(A, \bar{A}_0) + \alpha_3.\mathcal{S}(A) - \alpha_4.(\#A - Nb).$$

Où $\mathcal{S}(A)$ représente la surface globale de l'image couverte par les sous-arbre de A , α_3 et α_4 sont des constantes positives.

Les résultats obtenus sont très dépendants des poids $\alpha_1, \dots, \alpha_4$ et des variables choisies pour les poids w (pour les premiers essais, les poids que nous avons choisis sont basés sur le nombre de noeuds du sous-arbres et le niveau de gris de sa racine) (voir figure 6.6, 6.7, 6.8). Par contre, lorsque ces paramètres sont choisis correctement, les sous-arbres sélectionnés sont pertinents et rendent mieux compte des différences entre surfaces qu'en utilisant un critère.

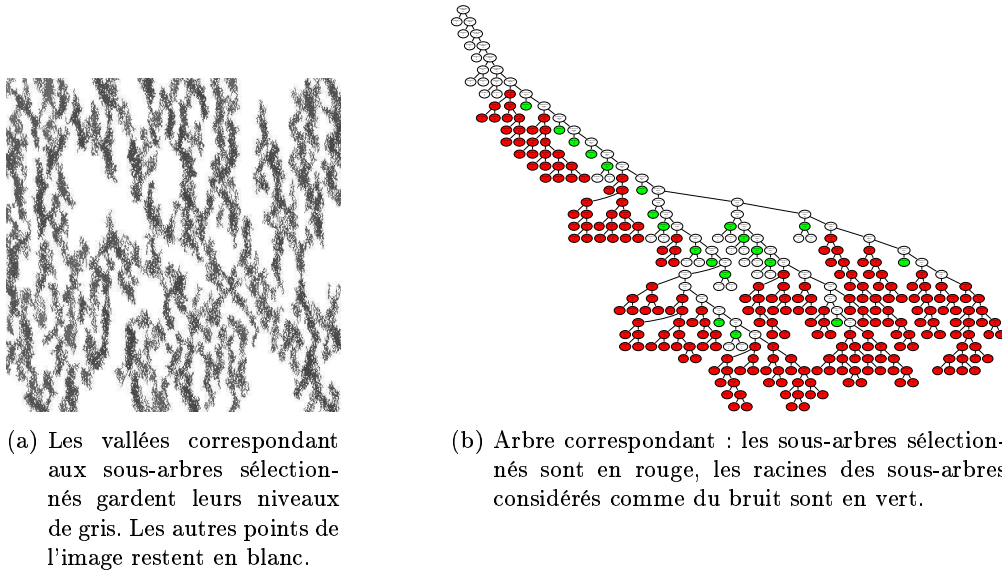


FIG. 6.6.: Résultat de l'algorithme pour $\alpha_1 = 20$, $\alpha_2 = 5$, $\alpha_3 = 0.01$, $\alpha_4 = 5$, $Nb = 15$

6.4. Utilisation des sous-arbres

6.4.1. Bootstrap

Si nous avons mis en oeuvre cette division en sous-arbres, c'est principalement pour pouvoir appliquer plus facilement des techniques de bootstrap, et donc pour pouvoir utiliser les variables d'arbres dans notre méthodologie globale d'analyse. La méthode que nous avons suggérée consiste à extraire des descripteurs quelconques à partir des sous-arbres sélectionnés pour une surface, puis d'en faire la moyenne ou un histogramme pour obtenir le descripteur final θ de la surface originale, à utiliser dans notre schéma d'analyse statistique. Reste à produire les échantillons bootstrap nous permettant d'estimer la variance de ce descripteur final lorsque l'on considère plusieurs surfaces du même type (c'est à dire issues de la même fonction aléatoire). Comme nous considérons les p sous-arbres extraits comme indépendants, il est facile d'appliquer la méthode classique de bootstrap consistant à tirer au hasard *avec remise* p sous-arbres parmi les p disponibles et de recalculer alors un descripteur θ^* . On obtient ainsi B répliques du descripteur d'intérêt, et on peut réutiliser ces informations dans les phases d'analyse factorielle puis

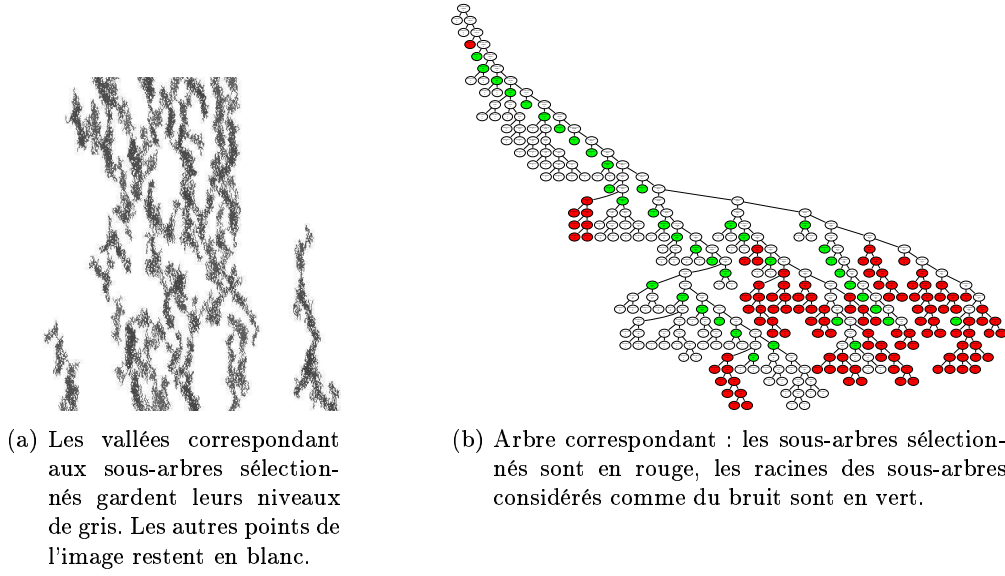


FIG. 6.7.: Résultat de l'algorithme pour $\alpha_1 = 30$, $\alpha_2 = 5$, $\alpha_3 = 0.001$, $\alpha_4 = 10$, $Nb = 20$

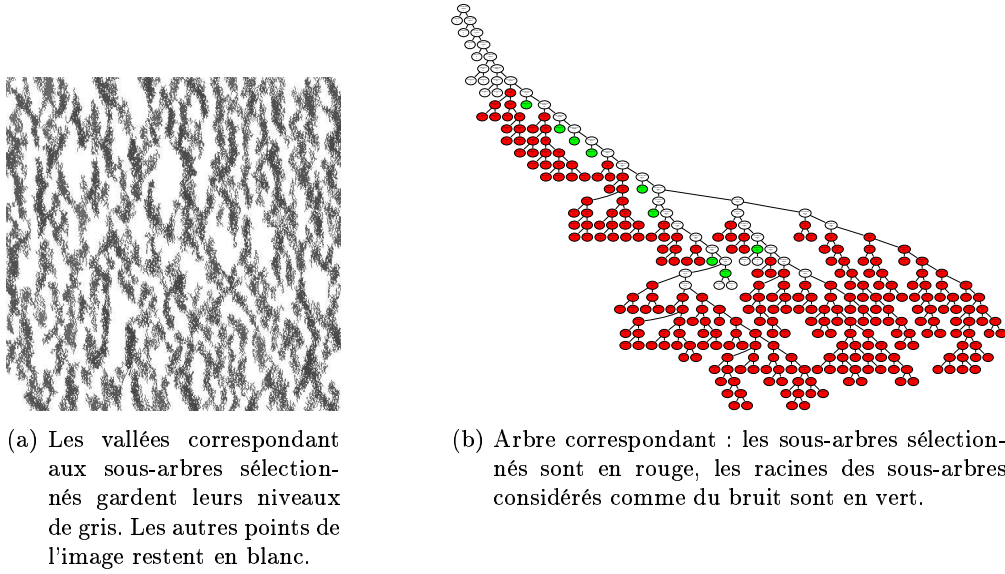


FIG. 6.8.: Résultat de l'algorithme pour $\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 0.05$, $\alpha_4 = 5$, $Nb = 20$

de classification.

A titre d'exemple, nous avons mis en place une analyse basée sur les arbres d'inondation des images caméra utilisées pour la validation de notre méthodologie. Chacun des arbres est décomposé en sous-arbres par l'algorithme du paragraphe 6.3.2, avec les mêmes paramètres que ceux de la figure 6.6. Nous nous intéressons aux descripteurs suivants calculés pour chaque images :

- Distribution (sur l'ensemble des sous-arbres) des niveaux de gris de la racine.
- Distribution du nombre total de noeuds des sous-arbres.
- Distribution de la surface des sous-arbres.
- Moyenne sur l'ensemble des sous-arbres de l'histogramme du nombre de noeuds par génération.

Comme pour les analyses des images caméra, on procède à une analyse factorielle pour chacun de ces descripteurs. On génère des individus bootstrap que l'on projette sur les différents plans factoriels, puis on réutilise les deux premiers axes propres de chaque analyse locale dans une analyse globale. Nous disposons de 5 images par surface, que nous allons pouvoir utiliser pour vérifier la pertinence des distributions bootstrap obtenues à partir des sous-arbres. Nous présentons ici (voir figure 6.9 et 6.10) les projections sur le premier plan factorielle de l'analyse globale de l'ensemble des individus bootstrap et des 5 images réelles correspondant à chaque surface incorrecte (car ce sont les surfaces pour lesquelles les distributions bootstrap ont la variance la plus forte). La projection de l'ensemble des individus bootstrap et de toutes les images réelles (le code de couleur reste identique à celui de la section 5.2.1) est également visible à la figure 6.9. Tous ces graphiques tendent à montrer que l'estimation de la distribution correspondant à un type de surface est moins précise (voir l'exemple de la surface 1) que dans le cas de l'étude des images, a fortiori que dans le cas des topographies, mais qu'elle est tout de même cohérente : les variances des distributions bootstrap sont visuellement du même ordre de grandeur bien que légèrement plus faibles que celles des distributions réelles. Dans certains cas, on retrouve même assez bien des propriétés particulières de certaines distributions : les surfaces 23 et 13 par exemple, ont des distributions fortement asymétriques, ce qui transparaît bien dans la distribution bootstrap.

6.4.2. Filtrage

La division en sous-arbre est aussi une méthode classique de filtrage ou de segmentation permettant de sélectionner des structures de l'image pour les simplifier. Le cas le plus courant consiste à modifier l'image de départ en bouchant certains lacs assimilés à du bruit. C'est le cas du filtrage par waterfall par exemple [12] (voir figure 6.12), qui peut s'obtenir (au moins de façon approchée) en fusionnant toutes les feuilles de l'arbre avec leur noeud père (voir figure 6.11). De la même façon, les fermetures par critères [64] consistent à boucher tous les sous-arbres ne vérifiant pas un critère correspondant à la définition 6.2. L'opération de filtrage consiste alors à construire l'arbre correspondant à une image, puis à choisir les sous-arbres que l'on veut filtrer, et enfin, à actualiser l'image

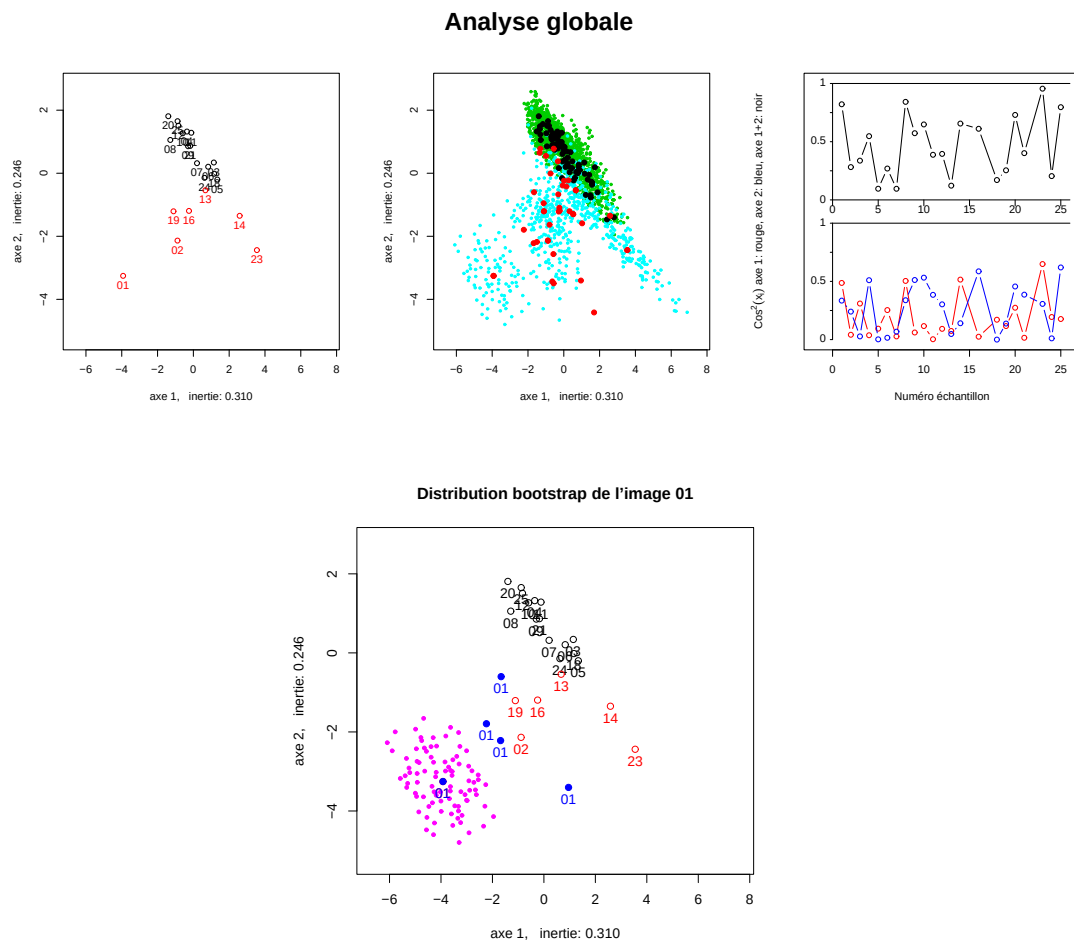


FIG. 6.9.: Analyse global à partir de sous-arbres

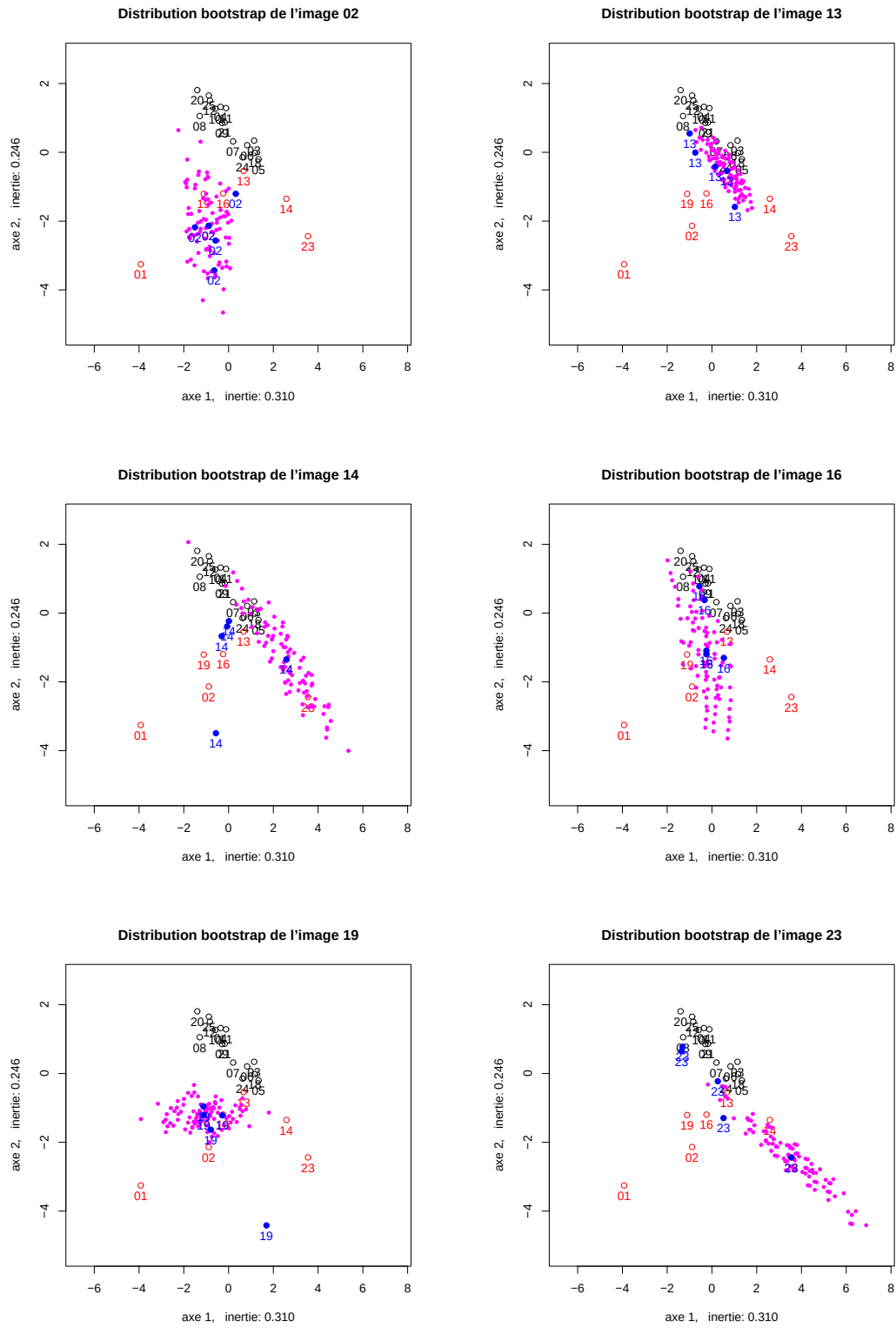


FIG. 6.10.: Distribution Bootstrap à partir de sous-arbres.

en remplaçant les vallées associées à ces sous-arbres par un niveau de gris constant (égal en général à la valeur du maximum du niveau de gris sur la vallée).

L'intérêt de notre algorithme de sélection est la flexibilité plus grande dans le choix des sous-arbres, puisque l'on n'impose pas directement de critère de ressemblance. En contrepartie, il faut tout de même adapter les coefficients choisis pour chacun des termes de la fonction de coût, ainsi que les poids des noeuds pour que le filtrage produise un résultat intéressant.

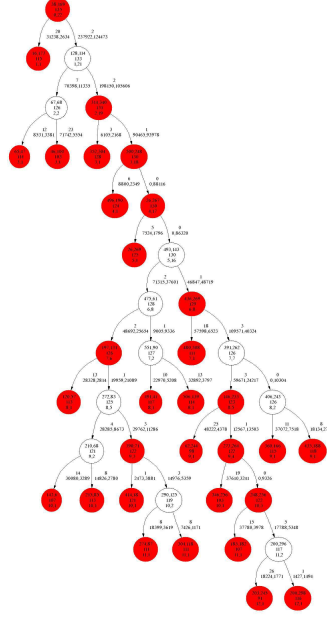


FIG. 6.11.: Exemple de modification d'un arbre par waterfall : les noeuds rouges seront éliminés et les lacs correspondant bouchés pour simplifier l'image.

6.5. Conclusion et Perspectives

Dans ce chapitre, nous avons passé en revue différentes approches pour étudier des surfaces par l'intermédiaire d'une représentation arborescente. Les premiers résultats obtenus laissent deviner la grande richesse de cette approche qui peut certainement être étendue à des processus de branchements très différents des arbres d'inondations considérés ici. De nombreuses pistes, certaines très théoriques, d'autres plus pratiques, restent à explorer pour tirer mieux encore profit de ces structures, et nous en suggérons quelques unes dans les paragraphes qui suivent.

6.5.1. Prise en compte des noeuds de transition

Dans ce qui précède, nous avons modélisé l'arbre d'inondation par trois types de noeuds :

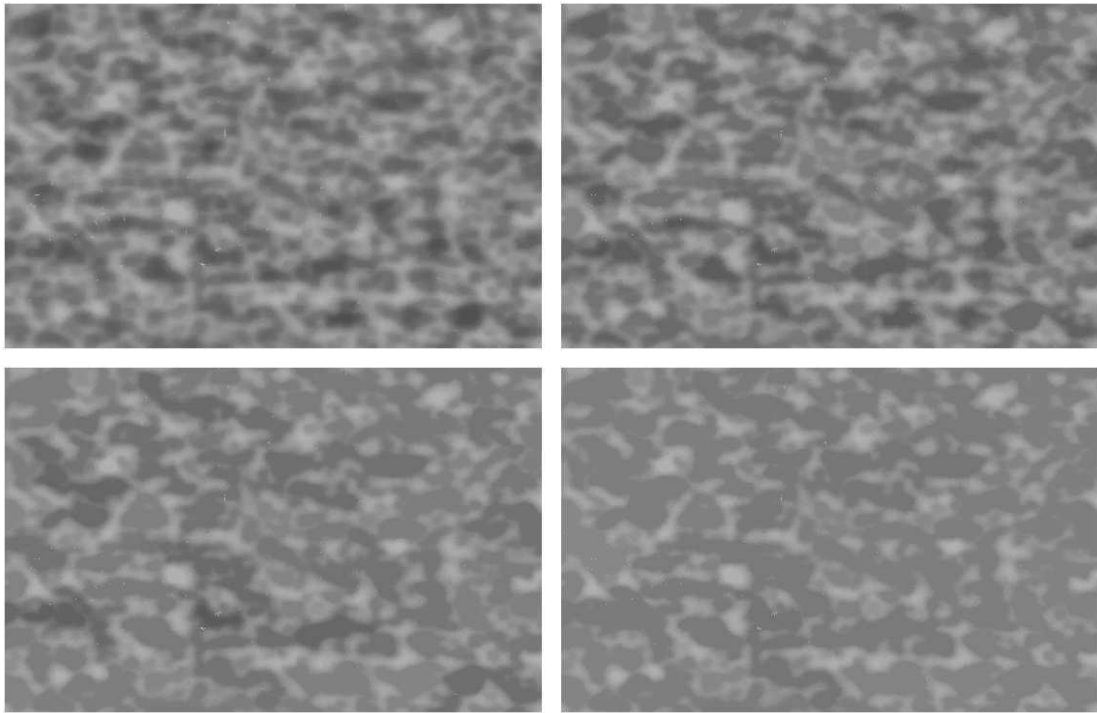


FIG. 6.12.: Exemple d'étapes successives de waterfall sur une image de type Peau d'orange. Dans l'ordre d'itération du processus : Haut gauche, haut droit, bas gauche, bas droit

- Les noeuds appartenant à un sous-arbre dit pertinent. Il s'agit en quelque sorte de "motifs" aléatoires possédant des propriétés communes.
- Les noeuds de liaison entre ces sous-arbres qui assurent la jonction de plusieurs motifs indépendants.
- Les autres noeuds qui sont considérés comme du bruit.

Jusqu'à présent, nous avons seulement analysé l'information contenue dans les motifs, et laissé de côté les autres types de noeuds. Nous proposons ici une méthode permettant de prendre aussi en considération l'information contenue dans les noeuds de transition.

La première étape consiste à transformer l'arbre d'inondation de la façon suivante :

- Chaque noeud racine d'un sous-arbre sélectionné est remplacé par une feuille.
- Chaque noeud racine d'un sous-arbre non sélectionné sans intersection avec un sous-arbre sélectionné est supprimé.
- Une phase de simplification fusionne l'ensemble des noeuds n'ayant plus qu'un seul fils à l'issue de l'étape de suppression.

Après ces modifications, les noeuds de l'arbre obtenu décrivent uniquement les jonctions entre motifs indépendants. On peut faire la supposition que cette indépendance

s'étend aux noeuds de liaison, ce qui suggère de les modéliser par un arbre markovien. Un tel arbre peut s'interpréter comme la représentation d'un processus de vie et de mort : une particule (la racine de l'arbre) possède une durée de vie qui suit classiquement la loi (exponentielle) d'un processus de Poisson : à chaque instant δt la particule a une probabilité $\theta \cdot \delta t$ de disparaître. Lorsque survient cette disparition, la particule peut donner naissance à deux particules filles avec une probabilité p : si cet événement se produit, les deux particules filles suivent *indépendamment* la même loi d'évolution que la particule originale. Chaque naissance est symbolisée par un noeud de l'arbre ; les arêtes représentent la relation de filiation et les labels des noeuds renseignent sur la durée de vie de la particule correspondante. Un tel modèle dépend donc de deux paramètres, le coefficient de Poisson θ , et la probabilité de division p . Pour approfondir les aspects théoriques de ces processus de branchement, étudier les cas où la durée de vie suit une loi plus complexe qu'une loi exponentielle, ou trouver des exemples d'applications pratiques de ces processus de branchement, on pourra se reporter par exemple à [6, 23, 8, 15, 37, 24, 32].

Dans notre cas, on peut revenir au modèle topographique en imaginant un marcheur se trouvant sur un point col rejoignant deux vallées dans lesquelles il peut descendre. Lorsqu'il en a choisit une, il doit descendre un certain dénivelé (variable aléatoire de loi connue, par exemple exponentielle) avant de rencontrer un nouveau point de relief particulier : il peut s'agir soit d'un nouveau col (avec une probabilité p) soit d'un lac (symbolisé par une feuille de l'arbre modifié, c'est à dire un motif ou un sous-arbre pertinent). L'exploration de tous les points col du relief coïncide avec le parcours de l'arbre modifié que nous avons décrit.

L'intérêt de ce schéma serait d'utiliser une méthode de sur-échantillonnage mixte, basée à la fois sur la simulation d'un arbre markovien et sur un tirage aléatoire de type bootstrap. Soit T l'arbre issu d'une image. On extrait par une méthode quelconque $S_1, S_2, \dots, S_K$, sous-arbres pertinents et ressemblants de T . On simplifie l'arbre T en un arbre T' comme décrit précédemment (les sous-arbres $S_1, S_2, \dots, S_K$ sont remplacés par des feuilles). Nous supposons que cet arbre T' est une réalisation d'un processus de branchement markovien dont nous pouvons estimer les paramètres en fonction du type de pondération choisie pour les noeuds.

Pour construire un échantillon statistiquement équivalent à l'arbre original, on commence par simuler une nouvelle réalisation du processus de branchement markovien, et pour chacune de ses feuilles, on tire au hasard et avec remise un sous-arbres parmi ceux qui ont été initialement sélectionnés dans l'arbre global. On peut générer par cette méthode de nombreux arbres équivalents au premier et les utiliser comme des individus bootstrap, même si ils sont obtenus en combinant à la fois des techniques paramétriques (modèle d'arbre markovien) et des techniques de bootstrap proprement dites.

Il est possible que cette nouvelle méthode de sur-échantillonnage améliore l'adéquation entre distribution bootstrap et distribution réelle. Cela ouvre également d'autres perspectives sur des problèmes de convergence. Par exemple, que se passe-t-il lorsque la taille de la surface tend vers l'infini, est-ce que le modèle envisagé reste pertinent ? Arrive-t-on à simuler, à partir d'une image de faible dimension, des arbres qui pourraient être obtenus à partir d'une surface plus large ? Un autre problème de convergence se pose lorsque la taille de l'image reste fixe, mais que le pas d'échantillonnage diminue :

observe-t-on une forme de comportement fractal au sein des sous-arbres pertinents, qui seraient eux-mêmes un agencement aléatoire de phénomènes indépendants à une autre échelle ?

6.5.2. Amélioration de l'algorithme de sélection

Idéalement, on pourrait imaginer l'existence d'un algorithme permettant de trouver un ensemble A réalisant le minimum d'une fonction de coût sans terme de régularisation (donc peu paramétrique). Si cette solution exacte implique une recherche exhaustive parmi toutes les possibilités, il est possible qu'elle soit impraticable et dans ce cas, un algorithme "sous-optimal" comme celui que nous avons suggéré est une bonne solution, ce qui ne veut pas dire qu'il ne puisse pas être amélioré, par exemple en raffinant le groupement des noeuds en trois catégories au lieu de deux (sous-arbres pertinents, noeuds de liaison et autres noeuds), en modifiant la fonction de coût, ou encore en ajoutant des possibilités d'évolution à tester à chaque itération pour les éléments de l'ensemble A_k (tout en assurant la convergence du processus).

L'un des points cruciaux de cet algorithme est aussi la comparaison des sous-arbres. Pour l'instant, nous nous basons sur le poids des noeuds, c'est à dire un vecteur, mais il est probable qu'on puisse faire mieux. Peut-être que des méthodes à noyaux de type "support vector machine" permettrait de faire intervenir à bon escient un produit scalaire directement dans l'espace des arbres. D'autres méthodes de comparaison d'arbres existent dont certaines, dites spectrales, se basent sur le calcul des valeurs propres de la matrice d'incidence (c'est à dire une matrice dont le terme (i, j) vaut 1 si il existe une arête entre i et j , et 0 sinon). Tous ces raffinements dans la comparaison des arbres conduisent potentiellement à une amélioration de la pertinence et de la stabilité de la division en sous-arbres, et par conséquent à l'amélioration de la précision des distributions bootstrap et de la classification ultérieure.

6.5.3. Autres applications

Nous nous sommes ici focalisés sur l'utilisation d'arbres d'inondation pour la caractérisation des textures, éventuellement pour du filtrage ou de la segmentation, mais ces représentations peuvent avoir d'autres intérêts pour des études plus physiques. En effet, la structure d'arbre et la division en sous-arbres permettent de différencier plusieurs échelles caractéristiques conduisant chacune à des traitements différents. Imaginons par exemple que l'on veuille simuler, à partir d'une carte topographique, le comportement optique d'une surface par lancer de rayons, c'est à dire en suivant le trajet inverse de rayons lumineux depuis la caméra jusqu'à la source, ce qui peut impliquer plusieurs réflexions sur les rugosités de la surface. Pour accélérer certains calculs, ou pour compenser un manque de données sur la topographie à petite échelle, on remplace souvent le calcul de plusieurs réflexions par un comportement aléatoire des rayons caractérisé par une fonction de réflectance bidirectionnelle (voir section 7.1). L'usage de sous-arbres peut permettre de déterminer quand utiliser les lois statistiques pour accélérer le calcul (lorsqu'un rayon rentre dans une vallée pour laquelle le sous-arbre correspondant a été sélectionné),

et quand utiliser l'information topographique disponible pour simuler effectivement les multi-réflexions (lorsque le rayon touche une portion de la surface qui n'appartient pas aux sous-arbres sélectionnés). Nous nous intéresserons dans la suite à la simulation du comportement optique des surfaces de type "peau d'orange", mais les faibles pentes de ces surfaces nous conduiront à négliger purement et simplement les multi-réflexions, et nous n'aurons donc pas l'usage d'une représentation arborescente dans ce cadre.

En résumé, la division en sous-arbres nous permet de modéliser différemment des phénomènes physiques (comportement optique, écoulement d'un fluide, etc...) en fonction de certaines caractéristiques des vallées, ce qui est en quelque sorte l'analogue d'un traitement multi-échelles dans lequel le critère de taille peut être assez complexe et plus paramétrable qu'une simple longueur d'onde caractéristique.

Chapitre 7.

Comportement optique des surfaces

Dans ce chapitre, nous abordons un domaine différent de l'étude des surfaces rugueuses, il ne s'agit plus de caractérisation ni de classification mais de simulation et de reconstruction. Nous nous intéressons dans un premier temps à l'image caméra obtenue par un dispositif similaire à celui que nous avons utilisé pour l'acquisition d'images en ligne. Nous avons constaté une perte d'information et de sensibilité de classification lors du passage d'une représentation topographique à une représentation optique (image caméra). Nous cherchons ici à étudier plus en détails les mécanismes qui conduisent de la topographie d'une surface à l'image enregistrée par la caméra, plus précisément nous cherchons dans un premier temps à simuler le plus exactement possible l'acquisition d'une image par la caméra, et nous appliquons pour cela des méthodes comparables à un "lancer de rayons" qui sont de plus en plus utilisées, non seulement dans le monde de l'imagerie virtuelle, mais aussi pour des applications où le réalisme physique des images générées est important [46].

Dans un second temps, nous nous intéressons au problème inverse, c'est à dire au calcul de la topographie originale à partir d'une image caméra. Ce problème est délicat car mathématiquement mal posé et nous présentons ici une méthode de résolution apparentée aux techniques de "shape from shading". D'autres méthodes de reconstruction utilisent une segmentation préalable d'une "paire stéréographique", suivie d'un appariement d'amers [13]. Dans le cas présent, nous opérons directement sur une seule image à niveau de gris.

Les développements et résultats que nous présentons dans ce chapitre ne sont pas encore arrivés à un stade industriel qui nous permettrait d'utiliser ces outils de simulation pour remplacer des tests physiques, ou pour calculer efficacement des cartes topographiques en ligne. Malgré tout, il s'agit de débuts prometteurs qui ont aussi l'avantage de conduire à une meilleure compréhension des phénomènes optiques sur lesquels sont finalement basés les classifications à partir d'images. D'autre part, ces résultats entrouvrent aussi une perspective de recherche de nouveaux types de surfaces vérifiant certaines propriétés optiques induites par leur topographie.

7.1. Problème direct

Pour résoudre le problème de simulation d'images caméra, nous nous placerons dans un schéma de radiométrie et nous raisonnerons donc en terme de flux d'énergie lumineuse au niveau de la surface. Cette approche suggérée par A. Aubert [1, 2, 5] possède l'avantage

de mener à des calculs simples ne faisant pas intervenir de phénomènes complexes tels que diffraction ou interférences lumineuses et nous négligerons par ailleurs la dépendance en longueur d'onde de certains phénomènes d'absorption ou de réflexion. Ces approximations coïncident avec nos conditions expérimentales pour lesquelles la source utilisée émet une lumière blanche incohérente dont les longueurs d'onde prépondérantes sont petites devant les rugosités de la surface à l'échelle considérée.

7.1.1. Rappels de radiométrie

L'unité de base que nous utilisons est le Watt (W) qui correspond à un flux d'énergie lumineuse émis (ϕ_e) par une surface ou incident (ϕ_i) à une surface par unité de temps.

Tous les phénomènes spatiaux pourront être discrétisés en des termes surfaciques (dS) ou angulaires (angle solide $d\omega$) car nous ne considérons pas d'effets volumiques (absorption par un milieu de propagation par exemple). On rappelle la relation suivante exprimant l'angle solide sous lequel une surface élémentaire dS de normale $\vec{n}$ est vue depuis un point M distant de r dans une direction $\vec{d}$:

$$d\omega = \frac{dS}{r^2} \vec{d} \cdot \vec{n} \quad (7.1)$$

Où $\vec{d} \cdot \vec{n}$ désigne le produit scalaire usuel entre les deux vecteurs (les vecteurs étant ici supposés unitaires, il s'agit du cosinus de l'angle formé par ces vecteurs).

Dans la suite, on pourra désigner un angle solide par le vecteur $\vec{\omega}$ dont la direction $\vec{\omega}$ est celle de l'axe du cône considéré et dont la norme correspond à la valeur de l'angle solide.

En pratique, pour l'ensemble des calculs à effectuer, on utilise principalement la notion de luminance qui est le flux énergétique émis (L_e), réfléchi (L_r) ou reçu (L_i) par une surface dans une direction donnée, par unité de surface et d'angle solide projeté (en $Wm^{-2}sr^{-1}$) :

$$L(x, \vec{\omega}) = \frac{d^2\phi(\vec{\omega})}{dS \vec{n} \cdot d\vec{\omega}} \quad (7.2)$$

Si l'on considère un élément de surface dS de normale $\vec{n}$, on peut définir la fonction de réflectance bidirectionnelle (brdf)¹ qui permet de quantifier l'énergie lumineuse se réfléchissant dans une direction donnée (ω_r) depuis une autre direction (ω_i) :

$$f(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) = \frac{L_r(x, \vec{\omega}_r)}{L_i(x, \vec{\omega}_i) \vec{n} \cdot d\vec{\omega}_i} \quad (7.3)$$

On suppose dans cette expression que la surface dS ne reçoit de l'énergie que depuis la direction $\vec{\omega}_i$ et qu'il n'y a pas d'émission propre de la surface, cependant, si la surface reçoit de l'énergie depuis l'ensemble des directions de l'espace, on obtient la relation intégrale :

$$L_r(x, \vec{\omega}_r) = \int_{\Omega_i} f(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) L_i(x, \vec{\omega}_i) \cos(\theta_i) d\omega_i \quad (7.4)$$

Où $\cos(\theta_i) = \vec{\omega}_i \cdot \vec{n}$, $\vec{n}$ étant la normale à la surface au point x .

¹ Bidirectional reflectance distribution function

7.1.2. Le modèle de Ward

Pour effectuer une simulation de rendu réaliste, il est important de modéliser correctement le comportement optique de la surface, supposée plane, en terme de brdf. Le modèle choisi est le modèle de Ward utilisé notamment dans le logiciel "radiance" [66, 65]. Considérant une surface dS de normale $\vec{n}$, on se donne un paramètre ρ (qui caractérise l'écart-type de l'histogramme des pentes à une échelle subfacette). La brdf du modèle de Ward s'exprime alors par :

$$f(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) = \frac{1}{\sqrt{(\vec{n} \cdot \vec{\omega}_i)(\vec{n} \cdot \vec{\omega}_r)}} \cdot \frac{\exp(-(\frac{\|\vec{\omega}_i + \vec{\omega}_r\|^2}{(\vec{\omega}_i + \vec{\omega}_r) \cdot \vec{n}} - 1)/\rho^2)}{4 \cdot \pi \cdot \rho^2} \quad (7.5)$$

On remarquera que cette fonction est parfaitement isotrope dans toutes les directions de l'espace, il existe également une version anisotrope de ce modèle faisant intervenir deux paramètres ρ_x et ρ_y , mais des tests de réponse optique sur des surfaces réelles ont montré que l'hypothèse d'isotropie était vérifiée de manière satisfaisante.

Notons que ce modèle de brdf intègre les effets optiques de la rugosité à une échelle inférieure à l'échelle d'observation. Le calage de cette fonction aux données expérimentales est détaillé dans la section 7.1.4. Pour la résolution du problème direct ("Shading from shape") à partir des données topographiques disponibles, nous utiliserons cette fonction comme loi de comportement de la surface. Physiquement, cette modélisation conduit à négliger la composante diffuse de la réflexion de la lumière. Cette approximation se justifie par la forte spécularité des surfaces étudiées : la composante diffuse est faible devant la composante spéculaire, et d'autre part, sur la plage angulaire proche du spéculaire ($\omega_r \simeq \omega_i$) pour laquelle f est forte, la composante diffuse peut raisonnablement être supposée constante. Nous négligerons aussi dans la suite les effets des multi-réflexions (les facettes de la surface étudiée ne seront jamais considérées comme des sources lumineuses) ainsi que les effets de masquage ou d'ombrage, ce qui se justifie pleinement par les faibles pentes (de l'ordre de 0.2°) des surfaces étudiées à l'échelle considérée.

7.1.3. Modélisation du dispositif

Le montage envisagé pour servir de base à un capteur d'acquisition d'images en ligne est composé d'une source en lumière blanche collimatée par une lentille convergente et d'une caméra matricielle haute définition observant une portion de la surface uniformément éclairée dans une direction éloignée de quelques degrés du spéculaire. Les angles d'éclairement et de reprise du flux lumineux par la caméra sont choisis avec précision et font partie des paramètres de réglage du dispositif.

La caméra est modélisée par une lentille mince L convergente circulaire parfaite. Dans ces conditions, on peut observer un plan objet P_o parallèle à la lentille en positionnant les capteurs lumineux (la matrice ccd par exemple) dans le plan image P_i du plan P_o : l'ajustement de cette position constitue la mise au point. Une fois cette opération effectuée, toute l'énergie lumineuse émise depuis un point V du plan P_o en direction d'un point de la lentille est concentrée en un même point M du plan P_i . Au contraire, l'énergie lumineuse d'un point extérieur au plan P_o sera répartie de manière plus ou moins diffuse

sur une portion du plan P_i , autrement dit, l'image d'un objet extérieur au plan P_o sera floue. L'intensité du flou dépend essentiellement de la surface utile de la lentille, c'est à dire de la taille du diaphragme utilisé : plus celui-ci est ouvert, plus le phénomène de flou est accentué. Pour obtenir un rendu réaliste d'une surface, on se doit donc de tenir compte de la taille du diaphragme utilisé.

Par ailleurs, l'image obtenue par la caméra est très dépendante de la résolution de celle-ci, c'est à dire de la taille du capteur élémentaire produisant un pixel sur l'image finale. En pratique, cette résolution est exprimée en $mm/pixels$: on peut par exemple fournir les dimensions réelles h_x et h_y de l'objet observé ainsi que les nombres de pixels N_x et N_y utilisés pour le représenter.

La caméra de notre dispositif est donc définie par (voir figure 7.1(a)) :

- le centre optique O , la normale $\vec{D}$ et le rayon r_L de la lentille L constituant l'objectif.
- deux vecteurs $\vec{U}$ et $\vec{V}$ tels que $(\vec{U}, \vec{V})$ constitue une base du plan objet P_o . Ces vecteurs représentent donc les directions réelles correspondant aux deux axes de l'image obtenue par la caméra (changer $\vec{U}$ et $\vec{V}$ revient donc à pivoter la caméra selon l'axe défini par O et $\vec{D}$). Par ailleurs il est important de noter que le plan objet est parallèle à la lentille et donc orthogonal à $\vec{D}$, et par conséquent $\vec{U} \cdot \vec{D} = 0$ et $\vec{V} \cdot \vec{D} = 0$.
- la distance d entre le plan objet et la lentille. En pratique, on considère le point $Q = O + d \cdot \vec{D}$ qui appartient au plan objet. L'image de ce point sera donc le centre de l'image obtenue par la caméra.

Pour caractériser la surface $\mathcal{S}$ étudiée, on se donne tout d'abord un point P de la surface (en général Q appartient à la surface et est utilisé comme point de référence) et une normale moyenne $\vec{N}_0$. On peut alors se donner deux vecteurs $\vec{X}$ et $\vec{Y}$ formant une base du plan ainsi défini. Les données topographiques disponibles consistent en un ensemble de valeurs $\delta(x, y)$ qui permettent de savoir que le point $P + x \cdot \vec{X} + y \cdot \vec{Y} + \delta(x, y) \cdot \vec{N}_0$ est un point de $\mathcal{S}$. Les informations $\delta(x, y)$ sont réparties sur une grille régulière et en pratique, ces valeurs sont très faible (de l'ordre du micron) devant les autres données du problème (d en particulier est de l'ordre de 75cm). D'autre part, la donnée véritablement importante pour utiliser le modèle optique relatif à la brdf est la normale à la surface en un point R . Les valeurs de ces normales peuvent être calculées par dérivation en interpolant la surface $\mathcal{S}$ au voisinage d'un point R par une surface analytique (voir section 2.2.5). La surface $\mathcal{S}$ sera donc considérée comme un plan défini par $(P, \vec{N}_0)$ pour tous les calculs géométriques (intersection avec une droite -un rayon- par exemple); en revanche, les normales calculées $\vec{n}(x, y)$ seront prises en compte dans les calculs de comportement optique (voir figure 7.1(a)).

La source lumineuse sera considérée comme un plan infini placé loin de la surface et émettant de l'énergie dans une direction unique $\vec{D}_s$ correspondant à la normale au plan infini. Cela revient à modéliser sa luminance par un dirac dans la direction $\vec{D}_s$. Sur le plan physique, nous considérons donc que la source lumineuse est parfaitement collimatée (en pratique, elle est placée au foyer objet d'une lentille convergente), et que la zone éclairée est très grande devant la zone observée par la caméra.

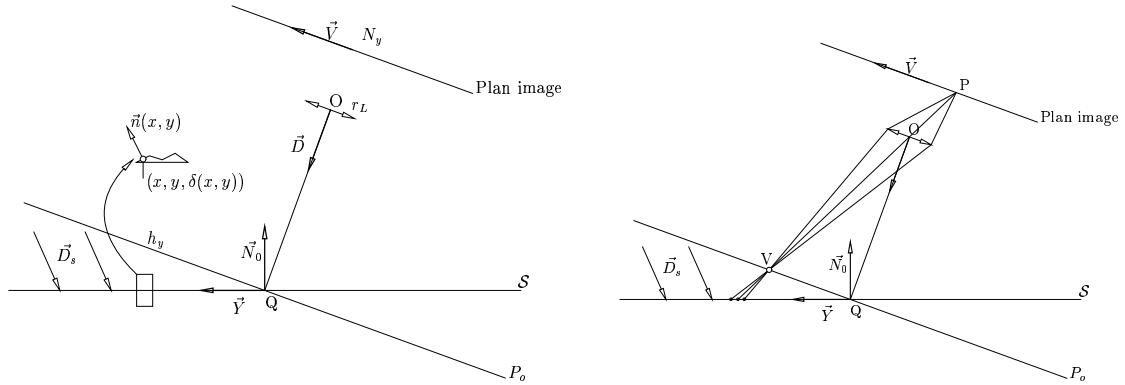


FIG. 7.1.: Modélisation et notations

a. Processus de simulation

A partir des données précédentes, on cherche à calculer le flux $\phi(k_u, k_v)$ parvenant sur la cellule sensible fournissant la valeur du pixel de coordonnées (k_u, k_v) dans l'image résultante. D'après la modélisation choisie pour la caméra, on sait que ce flux correspond à l'ensemble du flux lumineux émis en direction de la lentille par les points du plan objet P_o (voir figure 7.3) situés dans un rectangle ΔS_{V_0} de dimensions $(h_x/N_x, h_y/N_y)$ centré en un point V_0 défini par :

$$V_0 = Q - \left(\frac{h_x}{2} - \frac{h_x}{2N_x} + k_u \cdot \frac{hx}{N_x}\right) \cdot \vec{U} + \left(\frac{h_y}{2} - \frac{h_y}{2N_y} + k_v \cdot \frac{hy}{N_y}\right) \cdot \vec{V} \quad (7.6)$$

k_u et k_v étant des entiers compris respectivement entre 0 et $N_x - 1$ ou $N_y - 1$.

Si on appelle $\Omega_L(V)$ l'angle solide sous lequel est vue la lentille depuis un point V , on cherche à exprimer la quantité $\phi(k_u, k_v)$ par la double intégrale suivante :

$$\phi(k_u, k_v) = \int_{V \in \Delta S_{V_0}} dS_V \int_{\Omega_L(V)} L_e^*(V, \vec{\omega}_r)(\vec{\omega}_r \cdot \vec{n}_V) d\omega \quad (7.7)$$

Où $L_e^*(V, \vec{\omega}_r)$ est une luminance "virtuelle" que nous exprimerons par la suite.

Dans la suite, on note $\theta_r(V)$ l'angle formé par la normale à la surface considérée $\vec{n}_V$ au point V et le vecteur $\vec{\omega}_r$ correspondant à la direction de réflexion de la lumière. De même, on notera θ_i l'angle formé par la normale à la surface au point V et la direction $\omega_i(\vec{V})$ correspondant à la direction de la source considérée. On omettra cependant la dépendance en V lorsqu'il n'y aura pas d'ambiguïté.

On cherche à montrer dans la suite que l'équation 7.7 est valable en remplaçant $L_e^*(V, \vec{\omega}_r)$ par $L_r(R, \vec{\omega}_r)$ où R est le point de $\mathcal{S}$ tel que $R\vec{V} = \lambda \vec{\omega}_r$. Il nous faut considérer deux cas selon que R est derrière ou devant le plan objet.

En réalité, le plan objet P_o est un plan virtuel sans véritable réalité physique, il est simplement commode de le considérer à cause des propriétés de convergence de la lentille. La luminance émise depuis V dans une direction donnée est donc égale à la luminance incidente en V dans la direction opposée :

$$L_e(V, \omega_r) = L_i(V, -\omega_r) \quad (7.8)$$

Dans le premier cas, la surface $\mathcal{S}$ ne fait pas écran entre le plan objet et la lentille, la relation d'intégration est donc valable telle qu'elle avec $L_e^*(V, \vec{\omega}_r) = L_e(V, \vec{\omega}_r)$ et il suffit d'exprimer $L_i(V, -\vec{\omega}_r)$ en fonction de $L_r(R, \vec{\omega}_r)$ pour obtenir la relation souhaitée.

En faisant un bilan énergétique entre les points V et R , on peut exprimer le fait que le flux reçu par une surface élémentaire dS_V depuis un angle solide élémentaire $d\vec{\omega}_i$ est égal au flux émis par une surface dS_R centrée en un point R et vue sous un angle solide $d\vec{\omega}_i$ depuis V , ce flux étant émis selon un angle solide permettant de voir exactement dS_V depuis R (voir figure 7.2(a)). On obtient alors la relation suivante :

$$L_i(V, \vec{\omega}_i) \cos(\theta_i) dS_V d\omega_V = L_r(R, -\vec{\omega}_i) \cos(\theta_r) dS_R d\omega_R \quad (7.9)$$

Or d'après la relation 7.1 :

$$\begin{aligned} d\omega_V &= \cos(\theta_r) \frac{dS_R}{RV^2} \\ d\omega_R &= \cos(\theta_i) \frac{dS_V}{RV^2} \end{aligned} \quad (7.10)$$

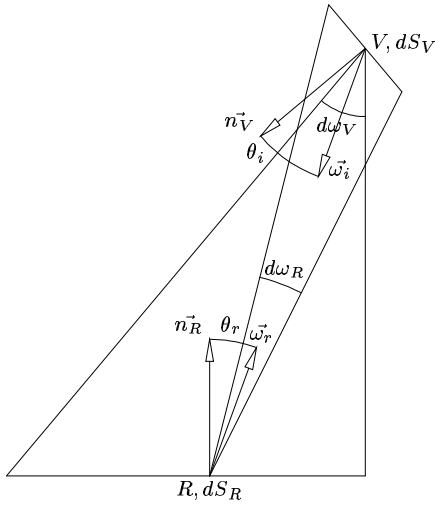
Et en remplaçant dans 7.9 puis en tenant compte de 7.8, on obtient la relation :

$$L_e(V, -\vec{\omega}_i) = L_i(V, \vec{\omega}_i) = L_r(R, -\vec{\omega}_i) \quad (7.11)$$

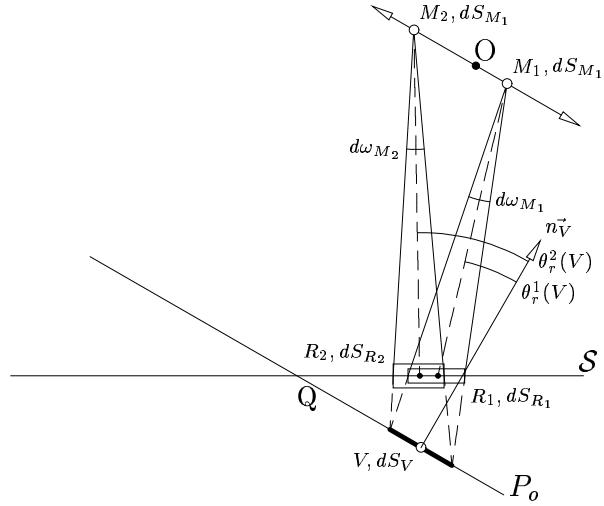
Et dans le cas où la surface ne fait pas écran entre le plan objet P_o et la caméra, on obtient finalement :

$$L_e^*(V, \vec{\omega}_r) = L_e(V, \vec{\omega}_r) = L_r(R, \omega_r) \quad (7.12)$$

Dans le cas où la surface $\mathcal{S}$ fait écran entre la caméra et la source, le raisonnement précédent ne peut être appliqué directement : on ne peut plus en effet remplacer $L_e^*(V, \vec{\omega})$ par la luminance $L_e(V, \vec{\omega})$ puisque celle-ci est en réalité nulle (le point V ne reçoit aucun flux lumineux car il est masqué par la surface $\mathcal{S}$). Pour obtenir le résultat, il faut intégrer le flux reçu par la lentille, et non le flux émis par le plan objet. En pratique, on cherche donc la luminance virtuelle $L_e^*(V, \vec{\omega})$ pour laquelle l'intégrale 7.7 fournit le même résultat que l'intégrale du flux incident sur toute la lentille provenant d'une direction donnée par le point V . Pour obtenir ce résultat, nous raisonnerons de nouveau en terme de conservation de flux lumineux.



(a) Schéma de conservation du flux lumineux : le flux reçu par dS_V dans l'angle solide $d\omega_V$ est le même que le flux émis par dS_R dans l'angle solide $d\omega_R$, ce qui permet d'aboutir à la conservation de la luminance : $L_i(V, \vec{\omega}_i) = L_r(R, \vec{\omega}_r)$.



(b) Schéma du cas où la surface $\mathcal{S}$ est placée entre la caméra et le plan objet. On cherche à évaluer le flux surfacique incident en un point du plan image (sans distinctions d'angle solide) : ce flux est l'intégrale sur la surface de la lentille de la luminance dans la direction du point V : le schéma montre bien la dépendance en M de dS_R et $d\omega_M$. dS_V , lui, est le même pour tous les points de la lentille.

FIG. 7.2.: Bilan radiatif

Chapitre 7. Comportement optique des surfaces

On se place cette fois-ci sur une surface élémentaire dS_M centrée en un point M de la lentille. On considère la luminance reçue par M depuis la direction du point V tel qu'il a été défini en 7.6, et d'après 7.10 et la conservation de l'énergie, on obtient :

$$L_i(M, \omega_{\vec{M}V}) \cos(\theta_i(M)) dS_M d\omega_M = L_r(R, \omega_{\vec{R}M}) \cos(\theta_r(R)) dS_R d\omega_R \quad (7.13)$$

Où R est le point de la surface $\mathcal{S}$ situé sur le segment MV . Il est important de noter que l'angle solide élémentaire $d\omega_M$ considéré dans cette équation est une fonction du point M imposé par les caractéristiques de la lentille. En pratique, cet angle solide est lié à la résolution de la caméra : il s'agira de l'angle solide sous lequel est vue une portion de surface élémentaire du plan objet. En revanche dS_R est directement imposé par $d\omega_M$ puisqu'il s'agit de la portion de la surface $\mathcal{S}$ qui est vue depuis M sous l'angle solide $d\omega_M$ (voir la figure 7.2(b)). L'élément de surface élémentaire dS_M nous permettra d'intégrer sur la surface de la lentille, et c'est ce paramètre qui conditionne $d\omega_R$, car il s'agit de l'angle solide sous lequel on voit dS_M depuis R .

Toujours d'après les relations 7.10 sur les angles solides, on peut obtenir :

$$\begin{aligned} \cos(\theta_i(M)) dS_M d\omega_M &= \cos(\theta_r(R)) dS_R d\omega_R \\ \cos(\theta_i(M)) dS_M d\omega_M &= \cos(\theta_r(V)) dS_V d\omega_V \end{aligned} \quad (7.14)$$

Par soucis de clarté, on omet dans ces expressions et dans celles qui suivent la dépendance en M du point R , mais il convient de se souvenir que R est en fait $R = R(M)$. V étant fixe lors de l'intégration sur la surface de la lentille, on obtient :

$$\begin{aligned} \int_{M \in L} L_i(M, \omega_{\vec{M}V}) \cos(\theta_i(M)) dS_M d\omega_M &= \int_{M \in L} L_r(R, \omega_{\vec{V}M}) \cos(\theta_r(R)) dS_R d\omega_R \\ &= \int_{M \in L} L_r(R, \omega_{\vec{V}M}) \cos(\theta_r(V)) dS_V d\omega_V \\ &= dS_V \int_{\vec{\omega} \in \Omega_L} L_r(R(V, \vec{\omega}), \vec{\omega}) (\vec{\omega} \cdot \vec{n}_V) d\omega_V \end{aligned} \quad (7.15)$$

Ce qui permet d'identifier dans cette dernière équation le terme $L_r(R, \omega_{\vec{V}M})$ avec le terme $L_e^*(V, \vec{\omega}_r)$ de l'équation 7.7.

On peut noter que la démonstration précédente serait aussi valable dans le cas où le plan objet se trouve entre la caméra et la surface, même si il n'était alors pas nécessaire de faire intervenir la notion de "luminance virtuelle". On peut aussi remarquer que les relations 7.15 permettent de transformer une différentiation sur un angle solide élémentaire $d\omega_M$ en une différenciation sur une portion de surface dS_V du plan objet. L'expression de gauche dans la relation 7.15 permet de calculer la valeur du flux lumineux reçu par un pixel de l'image finale en intégrant d'une part sur la surface de la lentille l'ensemble du flux lumineux issu d'un même point V , et d'autre part en intégrant pour chacun des points de la lentille sur un angle solide qui permet de tenir compte de la surface non nulle des capteurs de flux élémentaires (taille des capteurs CCD par exemple). Cependant l'angle solide élémentaire $d\omega_M$ qui intervient dans cette seconde intégration dépend du point M considéré sur la lentille, ce qui rend l'intégration assez peu commode. Le changement de variable des relations 7.15 permettent de calculer le flux lumineux reçu

par un pixel en transformant l'intégration sur un angle solide $d\omega_M$ pour un point de la lentille en une intégration sur une portion de surface dS_V du plan objet. Les propriétés de la lentille font que cet élément dS_V ne dépend pas du point M considéré sur la lentille, ce qui simplifie l'intégration.

b. Intervention de la brdf

D'après la section précédente, le flux lumineux reçu en un pixel de coordonnées (k_u, k_v) vaut :

$$\begin{aligned}
 \phi(k_u, k_v) &= \int_{\Delta S_V} dS_V \int_{\Omega_L(V)} L_e^*(V, \vec{\omega}_r) (\vec{\omega}_r \cdot \vec{n}_V) d\omega \\
 &= \int_{\Delta S_V} dS_V \int_{\Omega_L(V)} L_r(R, \vec{\omega}_r) (\vec{\omega}_r \cdot \vec{n}_V) d\omega \\
 &= \int_{\Delta S_V} dS_V \int_{\Omega_L(V)} d\omega (\vec{\omega}_r \cdot \vec{n}_V) \\
 &\quad \int_{\Omega_i(R)} L_i(R, \omega_i) f(R, \vec{\omega}_i \rightarrow \vec{\omega}_r) \cos(\theta_i) d\omega_i
 \end{aligned} \tag{7.16}$$

On notera que le point R dans ces expressions dépend à la fois du point V et du vecteur $\vec{\omega}_r \in \Omega_L$.

Si on néglige les phénomènes de masquage, d'ombrage et de multi-réflexion (c'est à dire le flux lumineux parvenant en un point de la surface après une ou plusieurs réflexions (ou diffusion) sur d'autres facettes de la surface), on peut considérer que le seul flux lumineux incident en un point de la surface provient de la source telle que nous l'avons modélisée, c'est à dire un dirac dans la direction $\vec{D}_s$. On peut alors simplifier la dernière intégrale des équations 7.16 :

$$\phi(k_u, k_v) = \int_{\Delta S_V} dS_V \int_{\Omega_L(V)} E_0 \cdot f(R, \vec{D}_s \rightarrow \vec{\omega}_r) (\vec{D}_s \cdot \vec{n}_R) (\vec{\omega}_r \cdot \vec{n}_V) d\omega \tag{7.17}$$

Où E_0 est la luminance caractérisant la puissance surfacique de la source.

Comme précédemment, on peut transformer l'intégration sur l'angle solide Ω_L sous lequel est vu la lentille depuis un point V en une intégration sur la surface de la lentille, de façon à faciliter la discrétisation ultérieure des intégrales.

$$\begin{aligned}
 \phi(k_u, k_v) &= \int_{\Delta S_{V_0}} dS_V \int_{\Omega_L(V)} E_0 \cdot f(R, \vec{D}_s \rightarrow \vec{\omega}_r) (\vec{D}_s \cdot \vec{n}_R) (\vec{\omega}_r \cdot \vec{n}_V) d\omega \\
 &= \int_{\Delta S_{V_0}} dS_V \int_{M \in L} E_0 \cdot f(R, \vec{D}_s \rightarrow \omega_{\vec{R}M}) (\vec{D}_s \cdot \vec{n}_R) (\omega_{\vec{R}M} \cdot \vec{n}_V) d\omega_M \\
 &= \int_{\Delta S_{V_0}} dS_V \int_{M \in L} E_0 \cdot f(R, \vec{D}_s \rightarrow \omega_{\vec{R}M}) (\vec{D}_s \cdot \vec{n}_R) (\omega_{\vec{R}M} \cdot \vec{n}_V) dS_M \cdot \frac{(\omega_{\vec{M}R} \cdot \vec{D})}{VM^2}
 \end{aligned} \tag{7.18}$$

La lentille étant parallèle au plan objet, la normale à P_o est la même que la normale $\vec{D}$ à la lentille, les produits scalaires $(\omega_{\vec{R}M} \cdot \vec{n}_V)$ et $(\omega_{\vec{M}R} \cdot \vec{D})$ sont égaux. Ces quantités seront notées $\cos(\theta_V)$ dans la suite.

Finalement, en exprimant VM^2 en fonction de d et $\cos(\theta_V)$ (voir figure 7.1(b) par exemple), on obtient l'expression :

$$\phi(k_u, k_v) = \int_{\Delta S_{V_0}} dS_V \int_{M \in L} E_0 \cdot f(R, \vec{D}_s \rightarrow \omega_{RM}) (\vec{D}_s \cdot \vec{n}_R) \cos(\theta_V)^4 \frac{dS_M}{d^2} \quad (7.19)$$

c. Discrétisation de l'intégrale

La méthode choisie pour évaluer le flux lumineux en un pixel de l'image finale consiste donc à lancer plusieurs rayons et à ajouter leur contribution, ce qui revient à discrétiser l'intégrale double 7.19 en deux sommes. La première consiste à lancer des rayons depuis plusieurs points différents de la lentille en direction d'un même point V du plan objet, et la seconde consiste à faire varier la position du point V sur une surface du plan objet dont la taille dépend de la résolution choisie (voir figure 7.3).

On introduit donc deux paramètres supplémentaires dans la simulation : nv et nl représentant respectivement le nombre de rayons utilisés pour discrétiser l'intégrale sur la surface ΔS_{V_0} et sur la surface de la lentille.

$$\phi(k_u, k_v) = E_0 \cdot \sum_{p=1}^{nv} \frac{\Delta S_{V_0}}{nv} \sum_{q=1}^{nl} \frac{\pi \cdot r_L^2}{nl} f(R_p, \vec{D}_s \rightarrow \vec{\omega}_{R_p M_q}) (\vec{D}_s \cdot \vec{n}_{R_p}) \frac{\cos(\theta_{V_p})^4}{d^2} \quad (7.20)$$

Où V_p , $p \in [1..nv]$ et M_q , $q \in [1..nl]$ sont des points répartis respectivement sur la surface $\Delta S_{V_0(k_u, k_v)}$ et sur la lentille, tandis que R_p , $p \in [1..nv]$ est le point de $\mathcal{S}$ tel que :

$$\exists \lambda \in \mathbb{R} / \overrightarrow{R_p V_p} = \lambda \cdot \overrightarrow{V_p M_q}$$

7.1.4. Simulation de lobes

L'un des problèmes expérimentaux restant à résoudre est l'évaluation du paramètre ρ du modèle de Ward à utiliser pour une surface donnée. Expérimentalement, le protocole est assez simple : on simule une image avec un paramètre ρ_0 , on la compare avec une image réelle obtenue par la caméra matricielle, on modifie ρ si les images ne sont pas semblables, et on itère le processus. En pratique, les images réelles et simulées sont assez difficiles à comparer quantitativement de façon à ajuster ρ efficacement. Un descripteur optique naturel pour évaluer la qualité d'un modèle de brdf est un lobe optique, c'est à dire une courbe donnant l'intensité lumineuse *moyenne* réfléchie par la surface en fonction de la direction d'observation (pour une direction d'incidence fixée). Ces courbes peuvent être mesurées expérimentalement d'une part, et être simulées d'autre part en calculant autant d'images que directions d'observation dans le lobe. Ces images simulées sont ensuite moyennées et les valeurs ainsi obtenues sont à comparer avec le lobe réel après une éventuelle phase de normalisation. En effet, le calcul ne tient pas compte de la composante diffuse supposée constante sur la plage angulaire d'intérêt, il faut donc simplement soustraire cette composante diffuse du lobe expérimental mesuré (figure 7.4(a)). D'autre part, la puissance utile de la source physique collimatée utilisée pour éclairer la surface

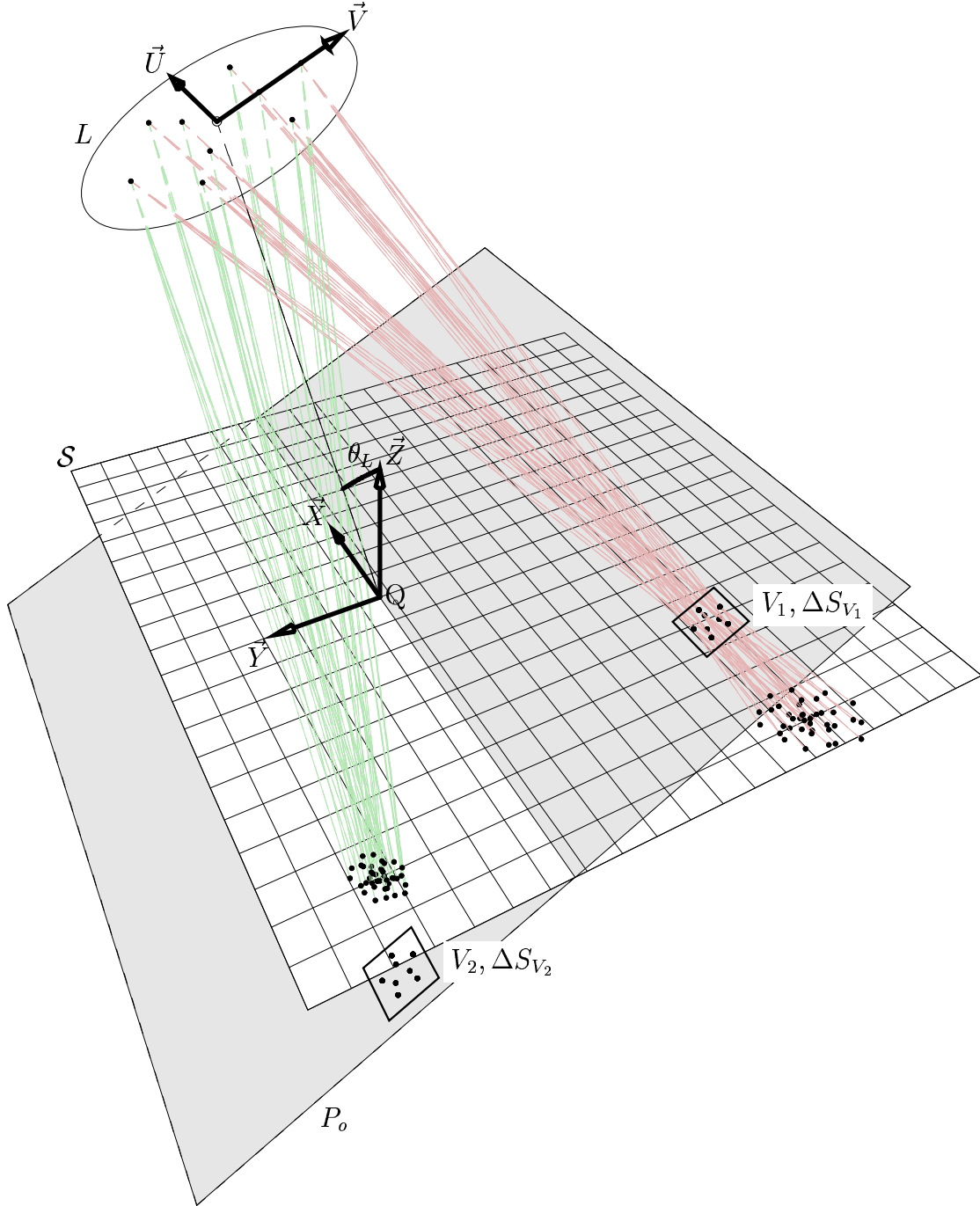
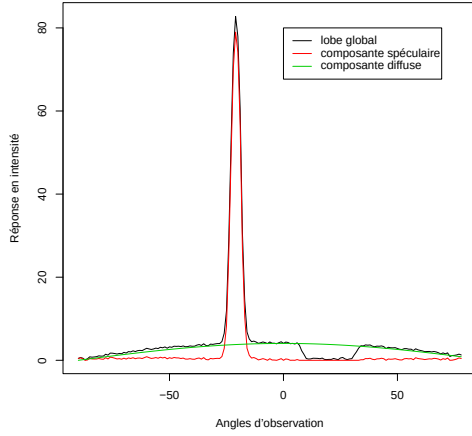


FIG. 7.3.: Modèle de lancer de rayons adopté : on discrétise les intégrales en choisissant plusieurs points sur la surface de la lentille et plusieurs points sur le plan objet dans une surface dépendant de la résolution de l'image finale. On lance des rayons depuis chacun des points choisis sur la lentille vers chacun des points choisis sur le plan objet. On évalue la luminance correspondant à chaque rayon au point d'intersection du rayon et de la surface $\mathcal{S}$.

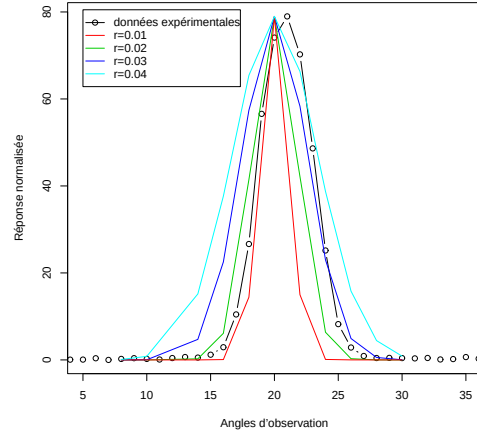
est souvent connue avec peu de précision, et il est préférable de ne pas tenir compte de cette constante en ajustant les valeurs maximales des lobes simulés et expérimentaux, et finalement, l'angle d'incidence de la lumière dans le dispositif expérimental utilisé n'est pas connu avec une grande précision, ce qui nous amène parfois à ajuster les maxima des lobes non seulement en valeur mais aussi en position.

Les calculs de lobes effectués sont réalisés avec des paramètres de définition et de précision assez faibles (un seul rayon lancé par pixel) de façon à pouvoir rapidement obtenir des résultats pour plusieurs valeurs du paramètre de rugosité : la comparaison avec le lobe expérimental permettra de choisir ce paramètre du modèle de manière réaliste. La configuration expérimentale comporte une source incidente à 20° de la normale ; c'est donc ce paramètre qui sera retenu pour les tests. Des lobes sont obtenus pour $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, $\rho = 0.04$, la figure 7.4(b) montre que le lobe calculé le plus proche du lobe expérimental en terme de largeur du pic spéculaire est obtenu pour $\rho = 0.02$, il semble cependant que les maxima ne coïncident pas exactement et un nouveau calcul pour $\theta = 21^\circ$ permet de corriger ce léger décalage. Le résultat final, même imparfait, semble au moins réaliste (figure 7.4(c)).

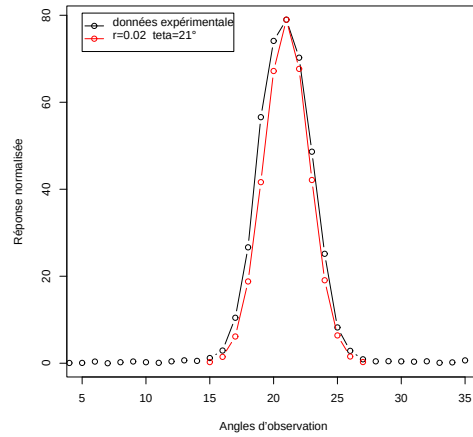
Ces ajustements fournissant une valeur de ρ plausible sont obtenus pour une surface de type "peau d'orange" (Pdo 10) et un angle d'incidence de 20° . L'une des questions que l'on peut se poser concerne la stabilité de ce paramètre ρ estimé pour d'autres surfaces du même type, et pour d'autres valeurs de l'angle d'incidence. Nous fournissons ici les lobes simulés et mesurés pour 4 surfaces de type peau d'orange, et pour 2 angles d'incidences (20° et 60°). Les résultats sont présentés à la figure 7.5. L'ensemble des lobes simulés sont assez proches des lobes réels, ce qui valide l'utilisation du modèle de Ward et la méthode d'estimation de ρ . Malgré tout, il reste des imperfections dans les comparaisons des lobes simulés et des lobes réels, et une amélioration sensible de ces résultats passe très certainement par une estimation plus fine de ρ grâce à un dispositif de mesure de lobe plus précis sur le plan angulaire que celui que nous avons utilisé ici.



(a) Lobe expérimental obtenu pour une surface de type "peau d'orange". Séparation des composantes spéculaire et diffuse par ouverture et fermeture.



(b) Comparaison entre le lobe expérimental et des lobes obtenus par le calcul pour plusieurs valeurs du paramètre ρ .



(c) Meilleur ajustement du paramètre ρ (2 décimales) : comparaison du lobe expérimental et du lobe calculé

FIG. 7.4.: Ajustement du modèle de Ward

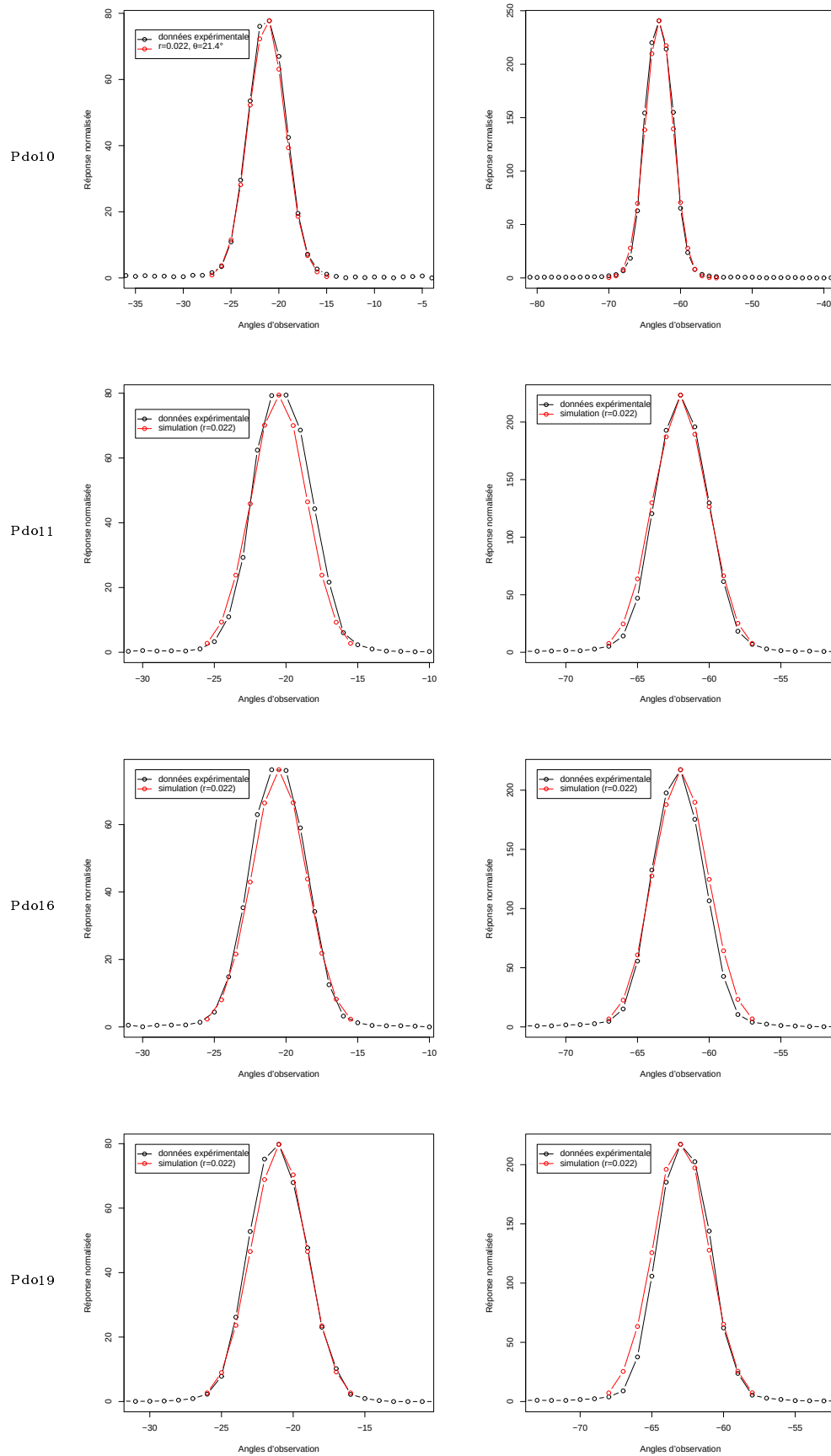


FIG. 7.5.: Simulation de lobes

7.1.5. Images obtenues

Voici des exemples de simulations compatibles en terme de configuration et de définition avec les acquisitions de la caméra matricielle destinées aux tests de classification. Le paramètre de rugosité choisi correspond à la valeur estimée à la section précédente ($\rho = 0.02$). Deux simulations sont effectuées : la première (figure 7.7(a)), grossière, ne tient pas compte de la profondeur de champ induite par l'ouverture du diaphragme de la caméra (pas d'intégration sur la surface de la lentille) et calcule par ailleurs la valeurs d'un pixel en utilisant un seul rayon (pas d'intégration sur une surface élémentaire du plan objet), tandis que la seconde simulation (figure 7.7(b)) utilise 25 rayons par pixels (5 positions différentes visées sur la plan objet et 5 positions différentes sur la lentille). Ces deux images sont à comparer avec la figure 7.7(c) qui représente une image réelle obtenue pour la même configuration géométrique.

Nous montrons aussi deux images simulées à partir de la topographie Pdo 19 (figure 7.6) en utilisant deux valeurs de ρ proches mais différentes (0.02 et 0.022). Ces images montrent qu'une faible différence de ρ peut entraîner des différences significatives sur les images simulées, et cela souligne l'importance de l'estimation de ce paramètre avec une meilleure précision que celle à laquelle nos mesures de lobes nous donne actuellement accès.

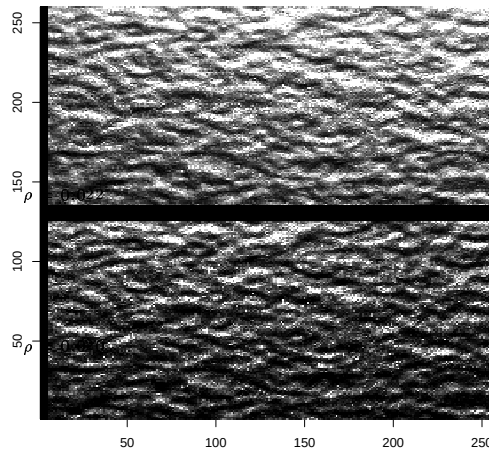
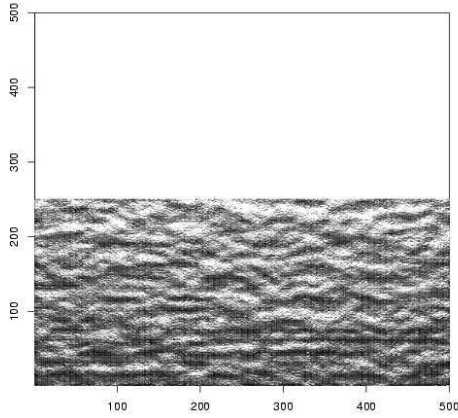


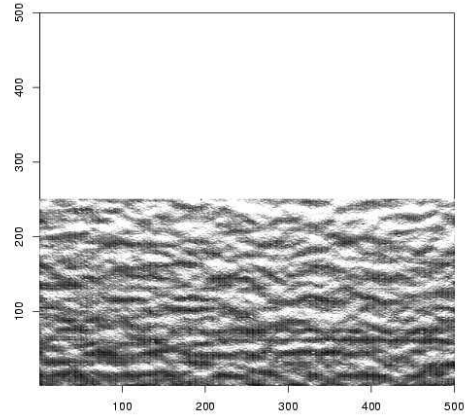
FIG. 7.6.: Pdo19 : simulations

7.2. Problème inverse

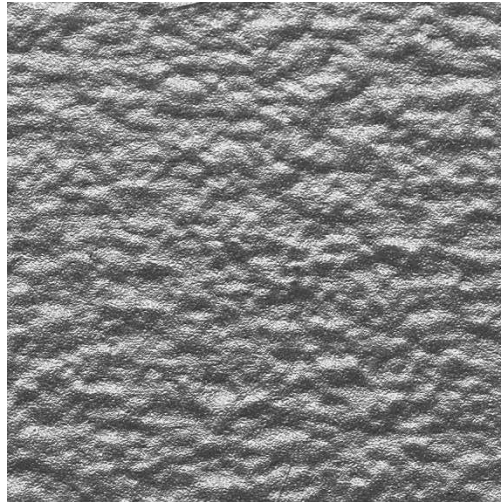
Le problème inverse que nous nous proposons de traiter consiste à déterminer la topographie d'une surface à partir d'une image caméra obtenue dans des conditions similaires (bien que simplifiées) à celle de la section précédente. Les applications industrielles d'un



(a) Image obtenue avec un seul rayon par pixel. En particulier, il n'y a pas d'effet de flou dû au diaphragme (la lentille est ici supposée ponctuelle).



(b) Image obtenue avec plusieurs rayons (25) par pixels. On peut noter l'effet de dérive du niveau de gris observable également sur l'image réelle et l'effet de flou sur les bords du champs (ici le diamètre du diaphragme est de 15 mm).



(c) Image réelle obtenu par caméra matricielle.

FIG. 7.7.: Simulations par lancer de rayons

tel outil peuvent être multiples : tout d'abord, cela permet en quelque sorte de contourner la perte d'information survenant entre une cartographie (utilisable en laboratoire pour la caractérisation des surfaces) et une image caméra que l'on peut obtenir sur ligne de production. Un tel dispositif permettrait de mesurer directement en ligne certains paramètres de rugosité (le "Ra" ou le "Wa" par exemple). D'un point de vue plus prospectif, la capacité à reconstruire une topographie à partir d'une image peut permettre de calculer une topographie "idéale" dont le rendu optique (l'image caméra) vérifierait certaines propriétés adaptées à telle ou telle utilisation.

Ce genre de problème est assez souvent désigné sous les termes "Shape from Shading" (sfs) et se retrouve dans une abondante littérature [69]. Les applications les plus courantes consistent à reconstruire un modèle 3D d'une scène à partir d'une photographie (forme d'un visage, d'une sculpture, d'un bâtiment..) [56], et dans ce cadre, les surfaces considérées sont le plus souvent lambertiennes. Dans certains cas, ces méthodes ont été appliquées à l'inspection de surfaces (lambertiennes) en utilisant plusieurs sources lumineuses et une caméra pour reconstruire la topographie [52].

La solution à un problème de type sfs n'est en général pas unique et il faut par conséquent faire un choix (ou accepter le résultat d'un algorithme) en sachant que le modèle 3D réel représenté en photographie peut-être différent, et parfois significativement, du modèle reconstruit. L'un des aspects du problème consiste donc à se donner des critères de choix dans les algorithmes permettant de converger vers une solution plutôt qu'une autre, en fonction de connaissances à priori sur la surface recherchée.

Dans ce qui suit, nous présenterons tout d'abord quelques unes des méthodes les plus populaires de "Shape from shading". Nous présenterons ensuite un cadre de travail spécifique correspondant à des surfaces de type peau d'orange et nous présenterons finalement quelques résultats préliminaires sur ces même surfaces.

7.2.1. Préliminaires

Avant de chercher à résoudre un problème de type "Shape from shading", certains préliminaires sont requis, et en particulier la connaissance de la configuration des éclairages et systèmes optiques ayant permis l'acquisition de l'image ainsi que le comportement optique de la surface en terme de brdf.

Dans ce qui suit, on considérera deux modèles principaux de brdf : celui de Ward [65] décrit dans la section 7.1.2, et celui d'une surface lambertienne qui diffuse de manière totalement isotrope tout ou partie du flux lumineux incident, indépendamment de la direction d'incidence :

$$f(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) = \frac{\epsilon}{\pi} = cte$$

Où $\epsilon \leq 1$ représente la proportion globale de lumière réfléchiée par la surface.

Dans le cas où la seule source lumineuse considérée est située à l'infini dans une direction donnée (faisceau collimaté provenant d'une direction $\vec{s}_i$), la luminance réfléchiée est indépendante de la direction et peut s'exprimer simplement sous la forme :

$$L_r(x, \vec{\omega}_r) = L_r(x) = E_i \cdot \frac{\epsilon}{\pi} \vec{n} \cdot \vec{s}_i$$

Bien que nous nous replacerons par la suite dans le cadre d'un modèle d'acquisition optique similaire à celui du problème direct précédemment traité, nous considérerons dans un premier temps le modèle d'acquisition le plus classique (dit *orthographique*) : dans le cadre de ce modèle, l'image est obtenue sur un plan $\vec{X}, \vec{Y}$, et l'intensité $I(x, y)$ en un point de l'image est égale à la luminance réfléchie par le point de la scène se projetant orthogonalement en (x, y) sur le plan image. Dans ce cas, on utilise en général un paramétrage de la surface correspondant à la scène sous la forme d'une fonction $h(x, y)$ représentant l'altitude de la surface au point (x, y) . Dans le cadre d'un modèle lambertien, on peut simplement exprimer l'intensité obtenue par :

$$I(x, y) = L_r(x, y, h(x, y)) = E_i \cdot \frac{\epsilon}{\pi} \vec{n}_{x, y, h(x, y)} \cdot \vec{s}_i \quad (7.21)$$

On peut déjà constater que cette intensité lumineuse dépend essentiellement de la normale en un point de la surface, et la procédure du "Shape from shading" procède généralement en deux étapes : la première permet de retrouver la carte des normales à la surface tandis que la seconde consiste à intégrer cette carte des normales pour retrouver l'ensemble des valeurs $h(x, y)$.

Dans la suite nous adopterons la terminologie et les notations suivantes :

- La carte des vecteurs normaux à la surface sera notée $\mathcal{N}$
- On désignera par les termes "image originale" la donnée d'entrée du problème : il s'agit de l'image caméra (réelle ou simulée) obtenue à partir d'une topographie réelle (dite surface originale).
- Les termes "image résultante" désignent au contraire l'image simulée à partir d'une topographie reconstruite. Le "Shape from shading" cherche à produire une surface topographique telle que les images résultantes et originales soient identiques.

7.2.2. Techniques classiques de "shape from shading"

Le problème de l'équation 7.21 est en général mal posé, c'est à dire que plusieurs cartes de normales peuvent induire la même image résultante. Plusieurs méthodes existent pour tenter de lever cette ambiguïté. Certaines utilisent la propagation de grandeur caractéristique à partir de points particuliers de l'image, d'autres cherchent des solutions particulières (solution de viscosité) des équations de Hamilton-Jacobi [54], mais les méthodes les plus à même d'être utilisées dans le cadre de nos surfaces sont basées sur des méthodes de calcul variationnel [30] dont nous allons présenter ici deux exemples.

a. Algorithme de Horne et Brooks

Cet algorithme [30] s'appuie sur la minimisation d'une fonctionnelle basée sur la carte des normales $\mathcal{N}$. Plusieurs termes peuvent être pris en compte, dont au moins un fait intervenir la correspondance entre l'image résultant et l'image originale. D'autres termes introduisent des connaissances a priori sur la surface. Par exemple, on impose classiquement à la surface d'être "régulière" et de ne pas varier trop brutalement. On peut alors chercher à minimiser l'intégrale des dérivées secondes. On utilise ensuite des techniques

de calcul des variations pour trouver la carte des normales minimisant la fonction de coût choisie.

A titre d'exemple, on peut introduire la fonction W définie dans le cadre de l'équation 7.21 (surface lambertienne, projection orthographique) :

$$W(\mathcal{N}) = \int \int \underbrace{(I(x, y) - \vec{n} \cdot \vec{s}_i)^2}_{\text{Eclairement}} + \lambda \cdot \underbrace{\left(\left\| \frac{\partial \vec{n}}{\partial x} \right\|^2 + \left\| \frac{\partial \vec{n}}{\partial y} \right\|^2 \right)}_{\text{Régularité}} dx dy \quad (7.22)$$

On suppose dans cette fonction que les termes constants multiplicateurs de la relation 7.21 ont été normalisés ou simplifiés.

On discrétise ensuite l'intégrale 7.22 (les dérivations deviennent des différences et les intégrales des sommes), et on procède par descente de gradient² à partir d'un état initial $\mathcal{N}_0$ pour parvenir à un minimum local de la fonction W .

Le problème principal de cette approche dans laquelle la contrainte d'éclairement est simplement introduite dans une fonction de coût réside dans le possible non respect de l'image originale. En effet, en fonction de la valeur de λ (ou $\lambda_1, \lambda_2, \dots, \lambda_n$ si plus de deux termes sont introduits dans la fonction de coût), le terme de régularité peut prendre le pas sur le terme d'éclairement, et la carte de normales issue du processus de minimisation peut produire une image résultante assez différente de l'image originale.

b. Méthode de Worthington et Hancock

Pour s'affranchir de ce principal problème, Worthington et Hancock [67] ont proposé une modification de la descente de gradient classique afin de maintenir la cohérence entre image originale et résultante à chaque étape du processus de minimisation.

L'idée consiste tout d'abord à choisir une carte des normales initiales qui vérifie la contrainte d'éclairement parfaitement :

$$\int \int (I(x, y) - \vec{n} \cdot \vec{s}_i)^2 dx dy = 0$$

On peut constater qu'en chaque point (x, y) de la surface, $\vec{n} \cdot \vec{s}_i = I(x, y)$, ce qui signifie géométriquement que la normale est portée par un cône circulaire autour de $\vec{s}_i$.

Ensuite, on minimise une fonction de coût W par descente de gradient, mais au lieu d'actualiser la carte des normales à chaque itération en utilisant simplement le gradient de la fonction de coût (qui peut avoir pour effet de faire sortir la normale du cône admissible), on reprojette la normale actualisée en chaque point sur le cône admissible, et on impose ainsi de manière forte le respect de la contrainte d'éclairement.

7.2.3. Le cas des surfaces peau d'orange

Dans cette section, nous proposons une adaptation des techniques de "Shape from shading" à une configuration optique particulière et à des surfaces très spéculaires (donc non-lambertiennes).

²les aspects plus techniques de ces techniques seront abordés par la suite (voir section b).

Dans la section traitant de la simulation de rendu par lancer de rayons, nous avons décrit un dispositif d'éclairage et de reprise simplifiant le comportement optique de la surface (pas de multi-réflexions en particulier), mais prenant en compte certains aspects du dispositif d'acquisition vidéo (largeur de diaphragme, résolution ...). Dans le cadre de ce problème inverse, nous simplifierons encore ces hypothèses : outre l'absence de multi-réflexions, nous supposerons également qu'à chaque point de l'image observée correspond un unique point de la surface ayant influencé son niveau de gris. Cela revient à supposer que le diaphragme est réduit à un point (pas d'effets de flou sur les bords de champs) et que la résolution de la topographie que l'on cherche à reconstruire est compatible avec la résolution de l'image (pas de moyenne spatiale de la réponse optique sur plusieurs points voisins). Chaque point (p, q) de l'image originale d'intensité $I(p, q)$ est donc en correspondance avec un unique point R de $\mathcal{S}$ de coordonnées (x, y) . Dans la suite, on confondra (p, q) et (x, y) , et on notera simplement $I(x, y)$, cependant, la projection que nous utilisons ici n'est plus orthographique (voir la section 7.1 pour un schéma précis du dispositif optique simulé).

Dans ce cadre particulier, l'expression 7.20 exprimant l'intensité lumineuse en un point de l'image peut être simplifiée :

$$\phi(x, y) = E_0^* \cdot f(R, \vec{D}_s \rightarrow \vec{\omega}_{RO}) (\vec{D}_s \cdot \vec{n}_R) \frac{\cos(\theta_V)^4}{d^2} \quad (7.23)$$

Où la fonction f correspond à la brdf de Ward décrite précédemment et E_0^* est une constante multiplicative.

Dans ce qui suit, nous supposerons parfaitement connus la configuration optique (angle d'incidence, de reprise, distance focale ...) ainsi que la constante E_0^* . En pratique il s'agit bien entendu d'un cas irréalisable : la puissance de la source, légèrement fluctuante dans le temps ne peut être connue avec une précision extrême, de même que les angles d'incidence et de reprise³. Il faudra donc par la suite trouver des méthodes permettant de s'affranchir de ces limitations, mais dans un premier temps, nous nous proposons d'étudier le cas de figure le plus simple possible.

Il est assez difficile d'étudier la performance d'un algorithme destiné à retrouver une topographie à partir d'une image caméra dans le cas qui nous occupe. En effet, pour la précision qui nous intéresse (de l'ordre de quelques microns), il faudrait pouvoir observer une surface au microscope interférométrique et prendre une image caméra de la même surface au micron près dans les conditions idéales décrites dans la section précédente.

Pour développer et tester nos algorithmes, il semble beaucoup plus aisé de travailler sur des topographies réelles (pour rester le plus proche possible de notre contexte d'étude), à partir desquelles des images caméra auront été simulées dans des conditions idéales. Ces images simulées serviront de base pour tenter de reconstruire une topographie qui pourra alors facilement être comparée à la topographie initiale.

³dans notre cas, "l'exactitude" sur ces angles revient à les connaître avec une grande précision par rapport à nos écart de normales, de l'ordre de $0,2^\circ$

7.2.4. Etude de la réponse optique

Dans le cadre de l'équation 7.21, nous avons constaté que la normale devait se positionner sur un cône afin de respecter la contrainte d'éclairement.

Dans le cas où la surface n'est plus lambertienne, il est nécessaire d'étudier plus précisément la brdf du modèle utilisé et la réponse optique qui s'en déduit pour déterminer le lieu des normales correspondant à un éclairage donné : il s'agit donc d'étudier les courbes de niveau de la réponse optique.

a. Changement de référentiel

Le modèle de Ward peut être vu assez simplement de manière géométrique (figure 7.8) en fonction des directions d'éclairement et de reprise :

$$f(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) = \frac{1}{\sqrt{\cos(\theta_i) \cdot \cos(\theta_r)}} \cdot \frac{\exp(-\tan^2 \delta / \rho^2)}{4 \cdot \pi \cdot \rho^2}$$

Nous supposons fixée la valeur de ρ comme décrit à la section 7.1.4. Considérons main-

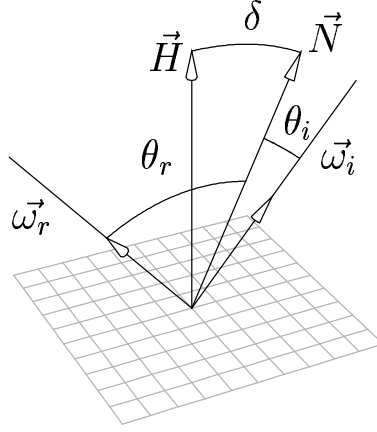


FIG. 7.8.: Modèle de Ward : interprétation géométrique.

tenant la fonction Φ de la relation 7.23 que l'on peut réécrire en tant que fonction de la normale en un point (x, y) :

$$\Phi_{x,y}(\vec{n}) = C_{x,y} \sqrt{\frac{\cos(\theta_i)}{\cos(\theta_r)}} \exp(-\tan^2 \delta / \rho^2) \quad (7.24)$$

Dans l'étude qui suit, on s'intéresse au comportement de la réponse optique en tant que fonction de la normale en un point ; c'est pourquoi on a regroupé en une valeur $C_{x,y}$ l'ensemble des constantes ne dépendant pas de la normale (mais pouvant dépendre de la position (x, y) du point considéré sur la surface). Dans la suite les indices x,y seront omis si il n'en résulte pas d'ambiguïté.

Nous allons maintenant chercher à déterminer les courbes de niveau de la fonction Φ , et pour étudier plus facilement cette fonction nous proposons un changement de repère⁴ (voir figure 7.9) à partir des définitions suivantes :

$$\vec{H} = -\frac{\vec{D} + \vec{D}_s}{\|\vec{D} + \vec{D}_s\|}$$

$$\vec{U} = \frac{\vec{D} - \vec{D}_s}{\|\vec{D} - \vec{D}_s\|}$$

$$\vec{V} = \vec{H} \times \vec{U}$$

$$\vec{D} = si.\vec{U} - ci.\vec{H}$$

$$\vec{N} = n_u.\vec{U} + n_v.\vec{V} + \vec{H}$$

$$\vec{N}' = f.\vec{X} + g.\vec{Y} + \vec{Z} \quad (\vec{N}' \parallel \vec{N})$$

$f_{x,y}$ Pente selon X au point considéré.

$g_{x,y}$ Pente selon Y au point considéré.

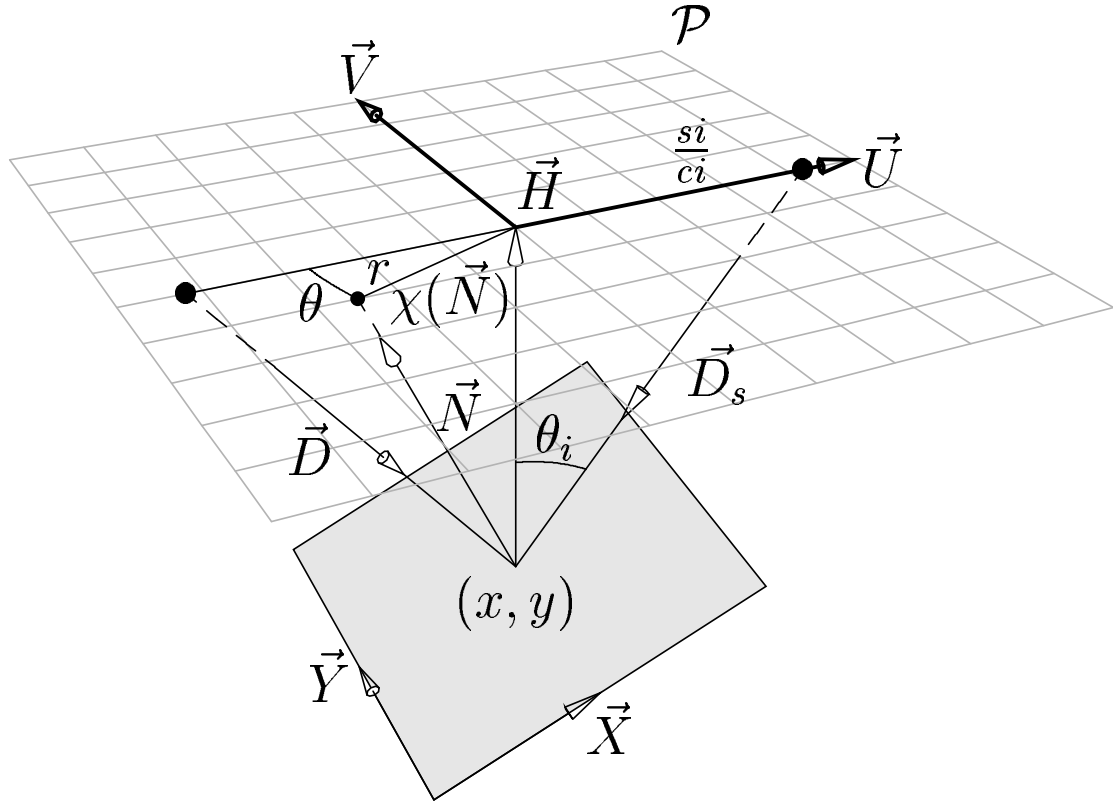
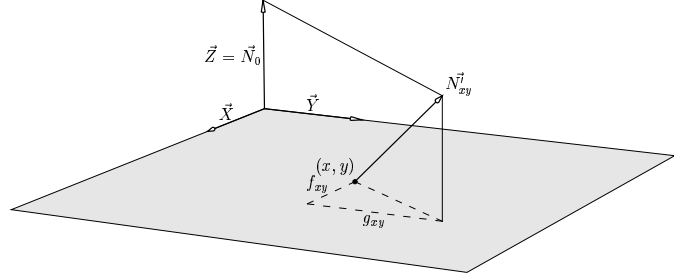


FIG. 7.9.: Changement de repère. Représentation du plan $\mathcal{P}$

⁴repère qui sera défini indépendamment en chaque point (x, y) de la surface.

Dans ces définitions, les vecteur $\vec{D}$ et $\vec{D}_s$ sont les vecteurs représentant la direction de reprise et d'incidence respectivement (Ces vecteurs pointent vers la surface $\mathcal{S}$). La notation $\times$ désigne le produit vectoriel, et le triplet $(\vec{U}, \vec{V}, \vec{H})$ forme donc une base orthonormée de l'espace. si et ci désigne respectivement le sinus et le cosinus de l'angle formé par les vecteurs $-\vec{D}_s$ et $\vec{H}$. $\vec{N}$ et $\vec{N}'$ forment deux représentations de la normale à la surface (de norme en générale différente) et sont donc colinéaires.

On définit le plan $\mathcal{P}$ dans le repère $(R, \vec{U}, \vec{V}, \vec{H})$ (où R est le point de la surface de coordonnées (x, y)) par :

$$P \in \mathcal{P} \Leftrightarrow \vec{R}\vec{P} \cdot \vec{H} = 1$$

On introduit la transformation χ associant à un vecteur $\vec{N}$ le point d'intersection entre le plan $\mathcal{P}$ et la droite défini par le point R et le vecteur $\vec{N}$:

$$\begin{aligned} \forall \vec{N} \in \mathbb{R}^3, \vec{N} \cdot \vec{H} \neq 0 &\Rightarrow \chi(\vec{N}) \in \mathcal{P} \\ \forall \vec{N} \in \mathbb{R}^3, \vec{N} \cdot \vec{H} \neq 0 &\Rightarrow \exists p / \chi(\vec{N}) = R + p \cdot \vec{N} \end{aligned}$$

$\chi(\vec{N})$ pourra dans la suite être considéré soit comme un point de $\mathcal{P}$, soit comme le vecteur du plan $\mathcal{P}$ défini par $\chi(\vec{N}) - \chi(\vec{H})$

Si l'on considère l'ensemble des normales telles que $\vec{N} \cdot \vec{H} > 0$ (normale sortante), alors χ est une bijection, et on peut l'inverser, c'est à dire retrouver les coordonnées de la normale à la surface à partir de $(n_u, n_v) = \chi(\vec{N}')$.

Pour étudier la fonction Φ , on utilise la transformation χ pour projeter les normales dans le plan $\mathcal{P}$. On introduit notamment les variables θ et r correspondant aux coordonnées polaires du vecteur $\chi(\vec{N}) = n_u \cdot \vec{U} + n_v \cdot \vec{V}$:

$$\begin{aligned} r &= \sqrt{n_u^2 + n_v^2} \\ \theta &= (-\vec{U}, \widehat{\chi(\vec{N})}) \end{aligned}$$

b. Etude de Φ

Dans ce référentiel et avec les notations précédentes, on peut exprimer la fonction Φ en fonction de n_u et n_v (figure 7.10), ou de r et θ :

$$\begin{aligned} \Phi(n_u, n_v) &= C \cdot \sqrt{\frac{ci + n_u \cdot si}{ci - n_u \cdot si}} \cdot \exp\left(-\frac{n_u^2 + n_v^2}{\rho^2}\right) \\ \Phi(r, \theta) &= C \cdot \sqrt{\frac{ci + r \cdot \cos(\theta) \cdot si}{ci - r \cdot \cos(\theta) \cdot si}} \cdot \exp\left(-\frac{r^2}{\rho^2}\right) \end{aligned}$$

On peut également calculer les dérivées partielles de Φ par rapport à n_u et n_v :

$$\begin{aligned} \frac{\partial \phi_{xy}}{\partial n_u} &= \sqrt{\frac{ci + n_u \cdot si}{ci - n_u \cdot si}} \cdot \exp\left(-\frac{n_u^2 + n_v^2}{\rho^2}\right) \cdot \left(\frac{ci \cdot si}{ci^2 - n_u^2 \cdot si^2} - \frac{2n_u}{\rho^2}\right) \\ \frac{\partial \phi_{xy}}{\partial n_v} &= \sqrt{\frac{ci + n_u \cdot si}{ci - n_u \cdot si}} \cdot \exp\left(-\frac{n_u^2 + n_v^2}{\rho^2}\right) \cdot \left(-\frac{2n_v}{\rho^2}\right) \end{aligned}$$

Les premières remarques que l'on peut faire après ces calculs sont les suivantes :

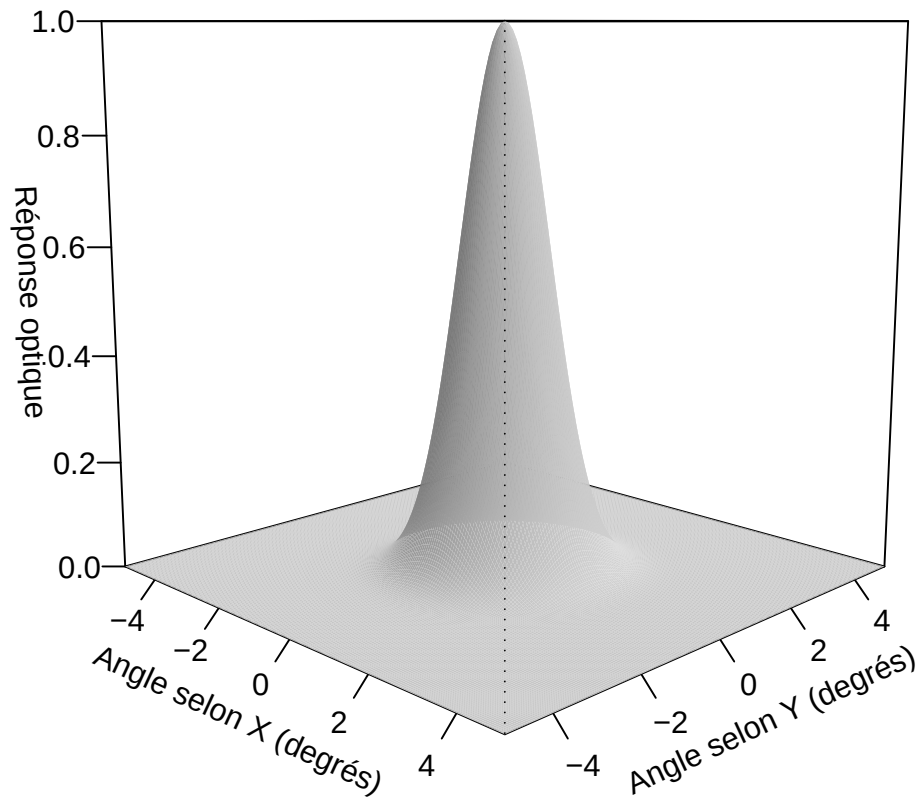


FIG. 7.10.: Représentation 3d de la fonction Φ , en fonction des *angles* entre la normale à la surface $\vec{N}$ et le vecteur $\vec{H}$

- Le maximum M de cette fonction n'est PAS le point $(0,0)$ mais un point situé quelque part sur l'axe des u en un point d'abscisse u_M racine du polynôme :

$$p(x) = x^3 - x \cdot \left(\frac{c_i}{s_i}\right)^2 + \frac{\rho^2}{2} \cdot \frac{c_i}{s_i}$$

- Les lignes de niveaux (dont les tangentes sont orthogonales au gradient de Φ) ne sont PAS des cercles comme dans le cas lambertien.

Ces deux effets proviennent du terme $\frac{ci \cdot si}{ci^2 - n_u^2 \cdot si^2}$ qui dissymétrise les dérivées partielles selon n_u et n_v .

c. Zone d'intérêt

Dans le cadre du "shape from shading", toute normale située sur une courbe de niveau $\Phi(n_u, n_v) = I(x, y)$ de la fonction de réponse optique est admissible. Il est donc nécessaire de limiter au maximum l'ensemble des solutions possibles pour la normale en un point en restreignant l'ensemble des normales candidates. La première réduction consiste à se placer au voisinage de la normale $\vec{N}_0$ moyenne à la surface (en général, il s'agit du vecteur $\vec{Z}$ du système de coordonnées). Dans le plan $\mathcal{P}$, cela revient à considérer un voisinage du point $\chi(\vec{N}_0)$ dont la taille dépend des pentes maximales de la surface.

Par ailleurs, la configuration optique particulière choisie introduit des particularités fortes sur la position de $\chi(\vec{N}_0)$. En effet, dans le cas qui nous occupe, le dispositif éclairement/caméra est orienté selon l'axe des $\vec{X}$ de façon à ce que la normale moyenne $\vec{N}_0$ soit décalées de quelques degrés vers la direction d'observation par rapport à la direction spéculaire $\vec{H}$. Ceci implique que $\chi(\vec{N}_0)$ se situe dans $\mathcal{P}$ du côté négatif de l'axe des $\vec{U}$. Si le point de la surface $\mathcal{S}$ considéré est décalé selon l'axe des $\vec{Y}$ par rapport à l'axe optique de la caméra, $\chi(\vec{N}_0)$ sera aussi positionné légèrement en dehors de l'axe des $\vec{U}$. Ceci peut paraître anecdotique, mais présentera une grande importance dans la suite, essentiellement parce que dans le voisinage de $\chi(\vec{N}_0)$ tel qu'il vient d'être défini, avec les valeurs typiques des pentes de nos surfaces et les valeurs considérées pour ρ , le terme dissymétrique $\frac{ci \cdot si}{ci^2 - n_u^2 \cdot si^2}$ devient fortement négligeable devant le terme $2 \cdot n_u / \rho^2$. Avec les valeurs typiques suivantes :

$$\begin{aligned} \widehat{(\vec{H}, \chi(\vec{D}_s))} &\simeq 22^\circ \\ \delta &\simeq 2^\circ \\ \rho &\simeq 0.02 \\ n_u &\simeq 0.03 \end{aligned}$$

On obtient :

$$\begin{aligned} \frac{ci \cdot si}{ci^2 - n_u^2 \cdot si^2} &\simeq 0.9 \\ \frac{2 \cdot n_u}{\rho^2} &\simeq 150 \end{aligned}$$

$$\frac{2.n_u}{\rho^2} \gg \frac{ci.si}{ci^2 - n_u^2.si^2}$$

On peut alors considérer que le gradient de Φ est radial en tout point, ce qui signifie que les courbes de niveau sont comparables à des cercles centrés sur l'origine $\chi(\vec{H})$ (voir figure 7.11).

Une autre remarque peut être faite concernant la position de $\chi(\vec{N}_0)$. Le décalage sur l'axe des $\vec{U}$ est beaucoup plus important que le décalage sur l'axe des $\vec{V}$, et cela introduit une dissymétrie assez importante entre les deux directions correspondantes, qui sont elles-mêmes très proches en pratique des directions $\vec{X}$ et $\vec{Y}$ sur la surface de départ. Ceci se traduira par la suite par une plus grande précision de reconstruction dans une direction que dans l'autre.

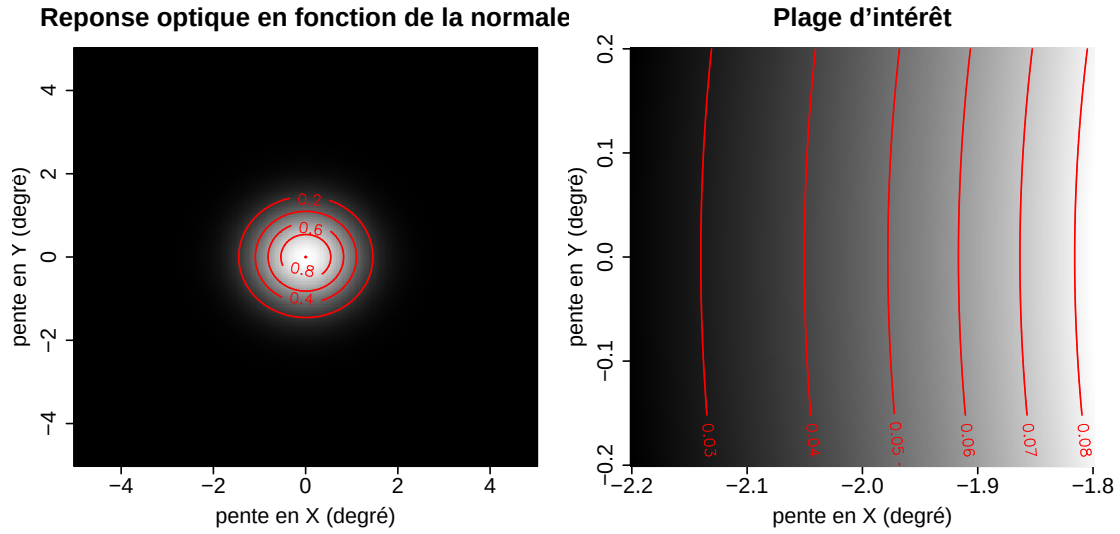


FIG. 7.11.: Courbe de niveau de $\Phi(n_u, n_v)$ en fonction des *angles* entre la normale à la surface $\vec{N}$ et le vecteur $\vec{H}$: vue globale et zoom sur la zone d'intérêt

7.2.5. Reconstruction des pentes puis de la topographie

La surface $\mathcal{S}$ que nous cherchons est définie par une fonction de la forme $z = h(x, y)$. La section précédente a montré que la réponse optique d'un point (x, y) dépend du vecteur normal à la surface en ce point, c'est à dire du gradient $\nabla_{x,y}h$. La reconstruction de la surface $\mathcal{S}$ va donc débuter par une étape permettant de reconstruire ∇h en tout point de $\mathcal{S}$.

L'algorithme de reconstruction des pentes repose sur deux étapes : la première (phase d'initialisation) fournit une première carte de normale compatible avec la condition d'éclairement tandis que la deuxième étape (phase de minimisation) modifie cette carte de façon à minimiser une fonction de coût mais sans altérer la compatibilité de la surface avec la condition d'éclairement.

a. Initialisation de l'algorithme

L'étude de la fonction Φ et de ces courbes de niveau a montré que pour un point donné (x, y) auquel correspond un éclaircissement $I(x, y)$, la transformée de la normale $\chi(\vec{N}_{x,y})$ se trouvait sur un cercle correspondant à la ligne de niveau $\Phi(n_u, n_v) = I(x, y)$. Ce cercle centré sur l'origine se définit exclusivement par son rayon. La première étape de notre algorithme va nous permettre d'établir pour chaque point de la surface (x, y) :

- Les vecteurs $\vec{U}_{x,y}$, $\vec{V}_{x,y}$, $\vec{H}_{x,y}$.
- Le rayon $r_{x,y}$ obtenu par la recherche numérique du zéro de la fonction :

$$f(r) = \Phi(r \cdot (\chi(\vec{N}_0) \cdot \vec{U}), r \cdot (\chi(\vec{N}_0) \cdot \vec{V})) - I(x, y)$$

Cela revient à établir le rayon r compatible avec la condition d'éclaircissement en se plaçant sur la droite reliant l'origine de $\mathcal{P}$ et le point $\chi(\vec{N}_0)$ (voir figure 7.12).

- Un angle $\theta_{x,y}^0$. Le point de $\mathcal{P}$ défini en coordonnées polaires par $(r_{x,y}, \theta^0)$ est alors associé à une normale induisant une réponse optique compatible avec l'image originale.

Cette méthode d'initialisation associée aux remarques sur la position particulière du point $\chi(\vec{N}_0)$ dans le plan $\mathcal{P}$ permet d'établir certaines caractéristiques de la carte de normales ainsi reconstruite. Tout d'abord, la droite reliant l'origine de $\mathcal{P}$ et le point $\chi(\vec{N}_0)$ est, comme signalé précédemment, très proche de l'axe des $\vec{X}$ sur la surface initiale. De plus, à l'échelle considérée (imposée par les pentes maximales de la surface de l'ordre de 0.5°), les courbes de niveau circulaires sont assez grossièrement proches de droites orthogonales à l'axe $(\chi(\vec{H}), \chi(\vec{N}_0))$ (voir figure 7.11). Ceci implique que deux points quelconques de la portion de cercle situé dans la zone admissible voisine de $\chi(\vec{N}_0)$ représenteront deux normales ayant des coordonnées proches selon l'axe des $\vec{X}$ mais des coordonnées selon l'axe des $\vec{Y}$ qui peuvent être très différentes.

Très concrètement, cela signifie que la donnée de l'image originale nous permet de connaître avec une assez bonne précision l'ensemble des dérivées partielles de la surface selon le vecteur $\vec{X}$, mais les dérivées partielles selon $\vec{Y}$ seront a contrario assez imprécises.

L'angle θ^0 peut être choisi de multiples façons différentes, par exemple en choisissant la normale correspondant au point situé sur la droite joignant l'origine $\chi(\vec{H})$ à $\chi(\vec{N}_0)$. La remarque précédente suggère qu'il peut être astucieux de choisir le point du cercle imposant à la dérivée partielle selon $\vec{Y}$ d'être nulle, et c'est ce que nous ferons dans la suite (voir figure 7.12).

En terme de complexité, cette étape de calcul comporte pour chaque point de la surface (x, y) :

- Le calcul du changement de repère propre à chaque point.
- La recherche du zéro d'une fonction dépendant de $I(x, y)$.
- Eventuellement le calcul plus ou moins complexe d'une valeur $\theta_{x,y}^0$. Dans le cas où l'on annule la dérivée partielle de $h(x, y)$ selon $\vec{Y}$, ce calcul correspond à l'intersection d'une droite et d'un cercle.

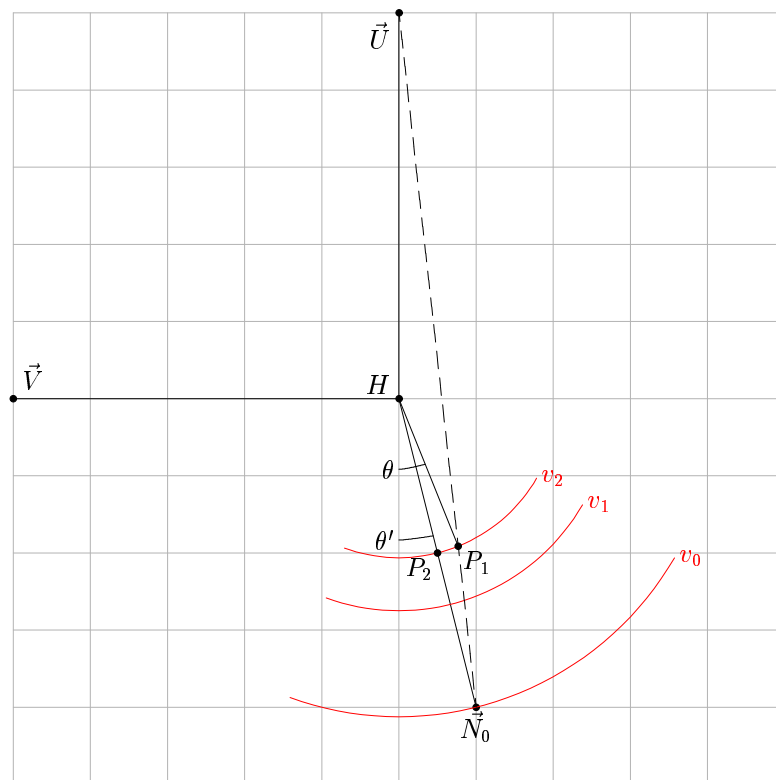


FIG. 7.12.: Phase d'initialisation : On trouve $r_{x,y}$ par minimisation d'une fonction d'une variable. On choisit Θ de façon arbitraire (P_1 : point du cercle pour lequel $g_{x,y} = 0$, P_2 : point du cercle situé sur la droite $(\chi(\vec{H}), \chi(\vec{N}))$).

A l'issue de cette étape, on obtient d'une part des valeurs qui resteront fixes dans la suite de l'algorithme ($r_{x,y}$, vecteurs $\vec{U}_{x,y}$, $\vec{V}_{x,y}$, $\vec{H}_{x,y}$) et qui servent uniquement à passer des dérivées partielles de h selon $\vec{X}$ et $\vec{Y}$ à l'espace dans lequel les courbes de niveau de Φ s'exprime simplement, et d'autre part la carte des angles ($\theta_{x,y}^0$) qui sera, elle, modifiée dans l'étape de minimisation. Le choix de ces angles comme variable de travail permet de modifier les normales sans modifier la réponse optique de la surface.

b. Etape de minimisation

Dans cette section, on cherche à adapter la carte des angles de façon à minimiser une fonctionnelle qui nous rapproche le plus possible de la surface originale.

La fonction de coût que l'on cherche à minimiser sera de la forme :

$$W = \sum_{i=1}^p \lambda_i \cdot W^i$$

Chaque W^i est une fonction dont l'entrée est la carte des angles $\Theta = (\theta_{x,y})$, c'est à dire un vecteur ayant pour taille le nombre de pixels considérés sur la surface de l'image. Les coefficients λ_i permettent de modifier l'importance attribuée à l'un ou l'autre terme, mais leur choix optimal reste un problème complexe.

En règle générale les fonctions W^i introduisent des conditions sur ∇h par l'intermédiaire des cartes ($f_{x,y}$) et ($g_{x,y}$). Dans ce cas :

$$W^i(\Theta) = W^i(f(\Theta), g(\Theta))$$

W^i est souvent obtenu comme intégrale d'une fonction $W_{a,b}^i$. Ces termes peuvent par exemple être calculés en chacun des points (x, y) de la surface.

Descente de gradient L'algorithme de minimisation de W commence à l'étape 0 avec la carte Θ^0 obtenue après l'étape d'initialisation. On va ensuite itérer plusieurs opérations de descente de gradient pour obtenir des cartes $\Theta^1, \Theta^2, \dots, \Theta^n$ convergeant vers un minimum local de la fonction W .

W est choisie de façon à être différentiable sur son espace de définition, et au premier ordre en Θ , on obtient :

$$W(\Theta^k + \delta\Theta) = W(\Theta^k) + \nabla_{\Theta^k} W \cdot \delta\Theta + o(\|\delta\Theta\|)$$

On définit Θ^{k+1} par :

$$\Theta^{k+1} = \Theta^k - \alpha^k \cdot \nabla_{\Theta^k} W$$

En remplaçant dans la formule précédente, on obtient :

$$W(\Theta^{k+1}) = W(\Theta^k) - \alpha^k \cdot \|\nabla_{\Theta^k} W\|^2 + o(\|\delta\Theta\|)$$

Si $\alpha^k > 0$ n'est pas trop grand, on diminue la valeur de W . En pratique, pour améliorer l'efficacité de la procédure, on choisit α^k en minimisant numériquement une fonction d'une seule variable :

$$\alpha^k = \arg \min_{\alpha > 0} (W(\Theta^k - \alpha \cdot \nabla_{\Theta^k} W))$$

En pratique, cet algorithme converge, mais souvent assez lentement. Des méthodes ont donc été développées pour diminuer le nombre d'opérations nécessaires à la convergence. Ces méthodes dites de gradient conjugué ne seront pas détaillées ici [55], mais sont implémentées dans les codes de calcul utilisés.

Dans les sections suivantes, plusieurs exemples de fonctions de coût seront explicités ainsi que le calcul des gradients correspondants, puisque le gradient de W s'obtient simplement par la formule :

$$\nabla W = \sum_{i=1}^p \lambda_i \cdot \nabla W^i$$

Cas général de la convolution Dans la plupart des cas, les fonctions de coût pourront s'exprimer sous la forme suivante :

Soit $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction dérivable. On définit W par :

$$W = \int \int \Psi(a \cdot (f * Kf)(x, y) + b \cdot (g * Kg)(x, y)) dx dy \quad (7.25)$$

Où :

- a et b sont des constantes.
- Kf et Kg sont deux fonctions convoluables avec f et g (en pratique, Kf et Kg sont de support fini).
- $*$ désigne le produit de convolution (continu ou discret suivant le contexte).

Bien sûr, pour appliquer notre descente de gradient, il faut établir une version discrète de ces relations et surtout calculer les dérivées partielles de W en fonction de $f_{x,y}$ et $g_{x,y}$ d'abord, et en fonction de $\theta_{x,y}$ ensuite.

L'expression discrète de l'équation 7.25 est simplement :

$$W = \sum_{x,y} \Psi(a \cdot (f * Kf)_{x,y} + b \cdot (g * Kg)_{x,y}) \quad (7.26)$$

Remarque : on peut négliger les éléments différentiels discrets correspondant aux échelles utilisées en $\vec{X}$ et $\vec{Y}$ car ils apparaîtraient simplement comme des coefficients multiplicateurs ne modifiant pas les minima locaux de W .

Dans ce contexte discret, pour simplifier les notations et symétriser les rôles de f et g , on peut introduire le tableau à trois indices $A_{x,y,z}$ et le noyau $K_{x,y,z}$:

$$\begin{aligned} \forall x, y \quad A_{x,y,0} &= f_{x,y} \\ \forall x, y \quad A_{x,y,1} &= g_{x,y} \\ \forall x, y \quad K_{x,y,0} &= b \cdot K_{g,x,y} \\ \forall x, y \quad K_{x,y,1} &= a \cdot K_{f,x,y} \end{aligned} \quad (7.27)$$

Remarques :

- Dans toutes les opérations discrètes, les images sont supposées périodiques, ici, la période est 2 dans la direction Z et correspond au nombre de points considérés sur la surface dans les direction X et Y.
- L'apparente inversion de Kg et Kf signifie simplement que les valeurs correspondant aux termes de la somme 7.26 se retrouveront avec $(A * K)(x, y, 1)$, et non $(A * K)(x, y, 0)$. Ceci permettra de gérer plus simplement les effets de bord dans la suite.

Avec ces nouvelles définitions, la relation 7.26 devient :

$$W = \sum_{x,y} \Psi((A * K)(x, y, 1)) \quad (7.28)$$

Par soucis de symétrie et de généralité, nous considérerons la relation plus générale suivante :

$$W = \sum_{x,y,z} \Psi((A * K)(x, y, z)) \quad (7.29)$$

Nous expliciterons dans la suite comment gérer les effets de bord dus à la périodisation des surfaces discrètes.

Pour exprimer le gradient de W , il suffit de calculer la dérivée partielle de W en fonction de tous les $A_{x,y,z}$:

$$\begin{aligned} \frac{\partial W}{\partial A_{i,j,k}} &= \sum_{x,y,z} \frac{\partial}{\partial A_{i,j,k}} (\Psi(A * K(x, y, z))) \\ \frac{\partial W}{\partial A_{i,j,k}} &= \sum_{x,y,z} \Psi'(A * K(x, y, z)) \frac{\partial}{\partial A_{i,j,k}} (A * K(x, y, z)) \\ \frac{\partial W}{\partial A_{i,j,k}} &= \sum_{x,y,z} \Psi'(A * K(x, y, z)) \frac{\partial}{\partial A_{i,j,k}} \left(\sum_{p,q,r} A_{x-p,y-q,z-r} \cdot K(p, q, r) \right) \\ \frac{\partial W}{\partial A_{i,j,k}} &= \sum_{x,y,z} \Psi'(A * K(x, y, z)) K(x - i, y - j, z - k) \\ \frac{\partial W}{\partial A_{i,j,k}} &= (\Psi'(A * K) * \check{K})_{i,j,k} \end{aligned} \quad (7.30)$$

Où $\check{K}$ correspond au noyau symétrique de K : $\check{K}(x) = K(-x)$.

Cette expression simple permet de calculer efficacement (par exemple par FFT) le gradient de W par rapport à A . Cependant, le fait de périodiser les images lors des convolutions peut induire des effets de bord indésirables. La solution consiste à restreindre la triple sommation $\sum_{x,y,z}$ dans le calcul de W à l'ensemble des points pour lesquels on peut calculer le produit de convolution avec K sans avoir à utiliser la périodisation de l'image. Si d désigne la taille de A et d_k celle de K ($d_k \geq d$), on peut connaître le produit $A * K$ sur un ensemble de taille $d - d_k + 1$. En réalité, on obtient donc :

$$W = \sum_{x=d_k-1}^{d-1} \sum_{y=d_k-1}^{d-1} \sum_{z=d_k-1}^{d-1} \Psi(A * K(x, y, z)) \quad (7.31)$$

Si $d = dk$, la sommation ne se fait plus que pour une valeur de l'indice correspondant à $d - 1$. C'est le cas en particulier lorsque A et K sont défini au moyen des expressions 7.27 : pour le troisième indice, la dimension est de 2 (un tableau pour f et un tableau pour g) à la fois pour A et K . En ne sommant que pour les indices $z = 1$, la sommation en z des relations 7.28 et 7.31 sont identiques (Le nombre des termes pour les sommation en x et y diffère cependant).

Dans ces conditions, on obtient l'expression suivante pour la dérivée partielle de W :

$$\begin{aligned} \frac{\partial W}{\partial A_{i,j,k}} &= \sum_{x=d_k-1}^{d-1} \sum_{y=d_k-1}^{d-1} \sum_{z=d_k-1}^{d-1} \Psi'(A * K(x, y, z)) K(x - i, y - j, z - k) \\ \frac{\partial W}{\partial A_{i,j,k}} &= \sum_{x=d_k-1}^{d-1} \sum_{y=d_k-1}^{d-1} \sum_{z=d_k-1}^{d-1} \Psi'(A * K(x, y, z)) K(x - i, y - j, z - k) \\ &\quad + \sum_{x=0}^{dk-2} \sum_{y=0}^{dk-2} \sum_{z=0}^{dk-2} 0.K(x - i, y - j, z - k) \\ \frac{\partial W}{\partial A_{i,j,k}} &= (\Psi^*(A * K) * \check{K})_{i,j,k} \end{aligned} \quad (7.32)$$

Où $\Psi^*(A * K)$ désigne un tableau de taille d (alors que $A * K$ n'est défini sans effets de bord que sur un tableau de taille $d - d_k + 1$) :

$$\begin{aligned} \forall x \geq d_k - 1, y \geq d_k - 1, z \geq d_k - 1 \quad , \quad & \Psi^*(A * K)_{x,y,z} = \Psi'((A * K)_{x,y,z}) \\ \text{sinon} \quad & \Psi^*(A * K)_{x,y,z} = 0 \end{aligned}$$

Il suffit donc de replonger la portion de $\Psi'(A * K)$ définie sans effet de bord dans un tableau de taille d en complétant par des 0 pour obtenir le résultat souhaité.

Le calcul de ce gradient n'est finalement qu'un intermédiaire pour calculer le gradient de W en fonction de Θ . Pour cela, il suffit d'appliquer la matrice jacobienne de la transformation $\theta \rightarrow (f(\theta), g(\theta))$:

$$\frac{\partial W}{\partial \theta_{i,j}} = \frac{\partial W}{\partial f_{i,j}} \cdot \frac{\partial f_{i,j}}{\partial \theta_{i,j}} + \frac{\partial W}{\partial g_{i,j}} \cdot \frac{\partial g_{i,j}}{\partial \theta_{i,j}} \quad (7.33)$$

Explicitons les termes $\frac{\partial f_{i,j}}{\partial \theta_{i,j}}$ et $\frac{\partial g_{i,j}}{\partial \theta_{i,j}}$:

$$\begin{aligned} Nx &= r(\cos(\theta).(\vec{U}.\vec{X}) + \sin(\theta).(\vec{V}.\vec{X})) + \vec{H}.\vec{X} \\ Ny &= r(\cos(\theta).(\vec{U}.\vec{Y}) + \sin(\theta).(\vec{V}.\vec{Y})) + \vec{H}.\vec{Y} \\ Nz &= r(\cos(\theta).(\vec{U}.\vec{Z}) + \sin(\theta).(\vec{V}.\vec{Z})) + \vec{H}.\vec{Z} \\ \frac{\partial Nx}{\partial \theta} &= r(-\sin(\theta).(\vec{U}.\vec{X}) + \cos(\theta).(\vec{V}.\vec{X})) \\ \frac{\partial Ny}{\partial \theta} &= r(-\sin(\theta).(\vec{U}.\vec{Y}) + \cos(\theta).(\vec{V}.\vec{Y})) \\ \frac{\partial Nz}{\partial \theta} &= r(-\sin(\theta).(\vec{U}.\vec{Z}) + \cos(\theta).(\vec{V}.\vec{Z})) \end{aligned}$$

$$\begin{aligned}
f(\theta) &= \frac{Nx}{Nz} \\
g(\theta) &= \frac{Ny}{Nz} \\
\frac{\partial f}{\partial \theta} &= \frac{Nz \cdot \frac{\partial Nx}{\partial \theta} - Nx \cdot \frac{\partial Nz}{\partial \theta}}{Nz^2} \\
\frac{\partial g}{\partial \theta} &= \frac{Nz \cdot \frac{\partial Ny}{\partial \theta} - Ny \cdot \frac{\partial Nz}{\partial \theta}}{Nz^2}
\end{aligned} \tag{7.34}$$

Ces formules assez générales permettent de calculer le gradient d'une fonction de poids par rapport à Θ , lorsque celle-ci s'exprime sous la forme d'une somme sur les points de l'image d'une fonction d'un produit de convolution des fonctions f et g par un noyau quelconque (expression 7.26). Dans la suite, nous présentons certains noyaux K^i menant à la construction des termes W^i de notre fonction de coût.

Condition de Schwarz Etant donné que la condition d'éclairement ne nous permet pas de connaître avec précision les pentes selon $\vec{Y}$, il est nécessaire de trouver un "levier" dans la fonction de coût qui nous permette de les reconstruire. Puisque les valeurs cherchées sont les dérivées de la fonction h selon deux directions orthogonales, on peut légitimement supposer que ces dérivées vérifient la condition de Schwarz, c'est à dire :

$$\frac{\partial^2 h}{\partial x \partial y} = \frac{\partial^2 h}{\partial y \partial x}$$

En termes de pentes, cela signifie, en reprenant les notations précédentes :

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}$$

La première contrainte introduite dans la fonction de coût sera obtenue de la façon suivante :

$$\begin{aligned}
\forall x, y, W_{x,y}^1 &= \left(\frac{\partial f}{\partial y}(x, y) - \frac{\partial g}{\partial x}(x, y) \right)^2 \\
W^1 &= \int \int W_{x,y}^1 dx dy
\end{aligned}$$

La discrétisation de ces expressions nécessite le calcul des dérivées de $f(x, y)$ et $g(x, y)$ selon x et y . La première idée qui vient à l'esprit consiste à considérer simplement les différences des valeurs de f et g entre points voisins :

$$\begin{aligned}
\frac{\partial f}{\partial y} &\simeq \frac{f_{x,y+1} - f_{x,y}}{s_y} \\
\frac{\partial g}{\partial x} &\simeq \frac{g_{x+1,y} - g_{x,y}}{s_x}
\end{aligned}$$

Où s_x et s_y désigne les facteurs d'échelles dans les direction $\vec{X}$ et $\vec{Y}$, c'est à dire la distance physique séparant deux points discrets de la surface.

Cependant, cette approche ne donne pas de très bons résultats en pratique : une topographie originale, mesurée au microscope interférométrique ne vérifie pas cette condition lorsque l'on calcule ses normales par des splines (f). La raison en est que les variations de normales d'un point à un autre sont trop importantes pour pouvoir appliquer l'hypothèse sous-jacente de linéarité. La convergence de l'algorithme vers la surface originale en devient donc très difficile. Cela est du reste visible sur l'image caméra issue d'une surface de type peau d'orange (figure 7.7(c)) : on observe une importante composante de haute fréquence qui traduit une assez forte variation de normale d'un point à un autre. Si l'on atténue ces composantes (en utilisant une convolution par une gaussienne par exemple), alors les composantes de moyennes et basses fréquences vérifient très bien le théorème de Schwarz. Pour améliorer le résultat il faut donc considérer les dérivées de f et g convoluée par des gaussiennes plutôt que les simples différences de valeurs entre deux points voisins.

La convoluée d'une fonction par une fonction G admissible vérifie l'agréable propriété :

$$\frac{\partial(f * G)}{\partial x} = \frac{\partial f}{\partial x} * G = f * \frac{\partial G}{\partial x}$$

$G(x, y)$ désigne maintenant une gaussienne de moyenne nulle et de matrice de covariance quelconque (dépendant de l'échelle des images considérées). Le choix de la taille de la gaussienne sera explicité de manière quantitative dans la section b traitant des résultats obtenus.

On transforme donc la définition 7.31 en :

$$\begin{aligned} \forall x, y, W_{x,y}^1 &= (f * \frac{\partial G}{\partial y}(x, y) - g * \frac{\partial G}{\partial x}(x, y))^2 \\ W^1 &= \int \int W_{x,y}^1 dx dy \end{aligned}$$

Il s'agit donc d'un cas particulier du paragraphe b pour lequel on peut appliquer les relations 7.26, 7.32 et 7.34 avec :

$$\begin{aligned} Kf &= \frac{\partial G}{\partial y} \\ Kg &= \frac{\partial G}{\partial x} \\ a &= 1 \\ b &= -1 \\ \Psi(x) &= x^2 \end{aligned}$$

Remarque : Tout noyau permettant d'approcher de "façon lissée" les dérivées partielles de f et g selon x et y peuvent remplacer les noyaux $\frac{\partial G}{\partial x}$ et $\frac{\partial G}{\partial y}$ sans qu'il soit nécessaire de modifier le raisonnement.

Le terme W^1 permet théoriquement d'obtenir des cartes f et g compatibles et donc intégrables sans ambiguïté. Cependant, la carte g , à cause de la configuration particulière décrite à la section c, n'est alors reconstruite qu'à travers ses dérivées partielles selon

x de façon à respecter la condition de Schwarz. Cela signifie que les valeurs de g ne seront obtenues qu'à un profil près $g(., y)$, représentant les constantes d'intégration qui peuvent être différentes pour toutes les ordonnées. Cette différence de profil constituera la principale source d'erreur (qui peut cependant se superposer avec des erreurs plus faibles) entre les pentes originales et les pentes reconstruites.

L'objectif des termes de la fonction de coût présentés dans la suite sera d'imposer des contraintes supplémentaires sur les cartes f et g de façon à restreindre l'ensemble des profils d'erreur admissibles et également à limiter les autres erreurs afin de s'approcher au maximum de la surface réelle.

Moyenne nulle L'une des particularités de nos surfaces est de représenter des textures stationnaires, de moyenne nulle et dont les vecteurs normaux s'orientent symétriquement autour de la normale moyenne $\vec{N}_0$. On peut donc imposer aux cartes f et g d'être de moyenne nulle :

$$W^2 = E(f)^2 + E(g)^2 \quad (7.35)$$

Là encore, la moyenne étant une opération linéaire, ceci est compatible avec les opérations de convolution de la section b, avec $\chi(x) = x^2$ et un noyau constant et normalisé (pour avoir $\sum_{x,y} K_{x,y} = 1$), de la taille des cartes f et g . L'ensemble des sommations de la relation 7.31 se réduit alors à un point pour éviter les effets de bord. Il est également possible de s'intéresser aux moyennes locales sur un voisinage plus ou moins large : il suffit de considérer un noyau de convolution constant (normalisé) de taille réduite. Dans ces conditions, la contrainte est plus restrictive car on sous-entend implicitement que le voisinage utilisé est statistiquement représentatif de la surface.

Non corrélation avec les variables X et Y . Si la contrainte du paragraphe précédent favorise les cartes f et g de moyenne nulle, cela n'évite pas nécessairement toutes les dérives. En effet, les pentes peuvent être de moyenne nulle tout en étant fortement corrélées avec les variables X ou Y ($f_{x,y} < 0$ pour x faible et $f_{x,y} > 0$ pour x fort par exemple). Or notre texture est supposée stationnaire, et on peut donc chercher à minimiser la (pseudo-)corrélation entre f ou g et x ou y .

$$W^3 = \left(\frac{\sum_{x,y} x \cdot f_{x,y}}{\sqrt{\sum_{x,y} x^2}} \right)^2 + \left(\frac{\sum_{x,y} x \cdot g_{x,y}}{\sqrt{\sum_{x,y} x^2}} \right)^2 + \left(\frac{\sum_{x,y} y \cdot f_{x,y}}{\sqrt{\sum_{x,y} y^2}} \right)^2 + \left(\frac{\sum_{x,y} y \cdot g_{x,y}}{\sqrt{\sum_{x,y} y^2}} \right)^2$$

Terme de régularité Pour tenter d'obtenir une surface avec une certaine régularité, on peut minimiser l'intégrale de la dérivée seconde :

$$\int \int \frac{\partial f^2}{\partial x} + \frac{\partial f^2}{\partial y} + \frac{\partial g^2}{\partial x} + \frac{\partial g^2}{\partial y} dx dy \quad (7.36)$$

Là encore, les dérivations s'obtiennent par convolution, et le gradient de la version discrète de cette expression peut se calculer au moyen des formules 7.26, 7.32 et 7.34 séparément pour chacun des termes.

c. Intégration de l'image des pentes

L'algorithme d'intégration utilisé ici se base sur la minimisation d'une fonction de coût $C(T)$ où T représente une topographie candidate [30, 52]. La fonction de coût s'exprime simplement sous la forme :

$$C(T) = \int \int \left(\frac{\partial T}{\partial x}(x, y) - f_{x,y} \right)^2 + \left(\frac{\partial T}{\partial y}(x, y) - g_{x,y} \right)^2 dx dy$$

On détermine dans un premier temps une topographie initiale T^0 en intégrant une colonne puis toutes les lignes. On itère ensuite un processus de minimisation par descente de gradient.

La méthode d'intégration proposée ici reste assez simpliste, et peut probablement être améliorée dans le futur.

7.2.6. Résultats

Dans cette section, nous présentons quelques résultats préliminaires sur des surfaces de type "peau d'orange" illustrant les méthodes développées dans les sections précédentes.

Nous travaillerons ici sur des topographies de 150×150 pixels représentant une surface d'environ $6 \times 6 \text{ mm}^2$. Les simulations de rendu sont effectuées de façon simplifiée en calculant la réponse optique de chaque pixel de l'image. Pour améliorer la stabilité numérique et la visualisation des résultats, nous présentons ici des résultats obtenus sur une "surface d'ondulation", c'est à dire sur une topographie filtrée par une gaussienne, avec une longueur d'onde de coupure de 0.5 mm (figure 7.13).

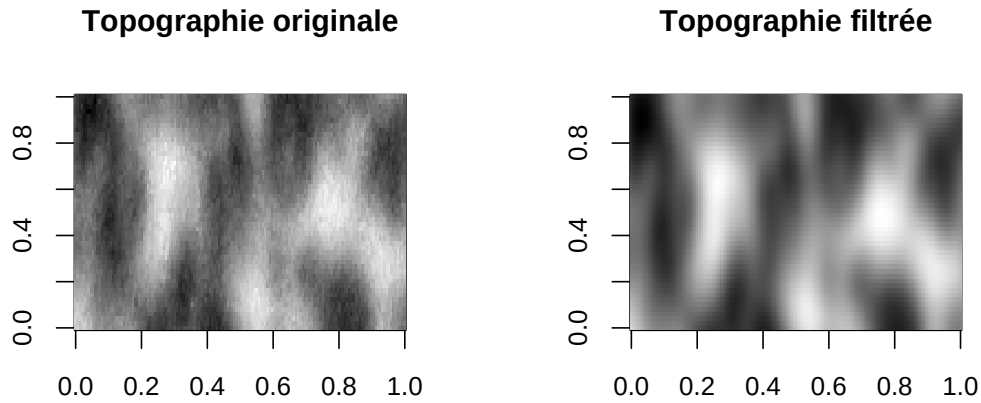


FIG. 7.13.: Filtrage d'une topographie par un noyau gaussien.

a. Phase d'initialisation

Partant de la topographie filtrée de la figure 7.13, on simule l'image originale (figure 7.14) puis on applique la phase d'initialisation de façon à obtenir des cartes f et g utilisables dans un processus de minimisation.

La figure 7.14 montre le résultat de cette phase d'initialisation : le très faible écart relatif sur les images originale et résultante montre que la condition d'éclairement est pour ainsi dire parfaitement réalisée :

$$\begin{aligned} E\left(\frac{|I_{or}(x,y) - I_r(x,y)|}{|I_{or}(x,y)|}\right) &\simeq 4.5 \cdot 10^{-6} \\ Max\left(\frac{|I_{or}(x,y) - I_r(x,y)|}{|I_{or}(x,y)|}\right) &\simeq 4.4 \cdot 10^{-5} \end{aligned}$$

Par ailleurs, on peut vérifier l'hypothèse de circularité des lignes de niveau en comparant les rayons originaux $r_{x,y}^{or}$ correspondant à la projection des véritables normales $\vec{N}_{x,y}$ sur le plan $\mathcal{P}_{x,y}$ en chaque point aux rayons $r_{x,y}$ obtenus après la phase d'initialisation en projetant sur $\mathcal{P}_{x,y}$ les normales reconstruites, pour lesquelles $g_{x,y} = 0$ en tout point. Si l'hypothèse de circularité est juste, les normales originale et résultante doivent se projeter sur un même cercle centré sur l'origine, et les rayons $r_{x,y}^{or}$ et $r_{x,y}$ doivent donc être très proches. En pratique, cette hypothèse se vérifie très bien, et on obtient pour l'exemple présenté :

$$\begin{aligned} E\left(\frac{|r_{x,y}^{or} - r_{x,y}|}{|r_{x,y}^{or}|}\right) &\simeq 7.2 \cdot 10^{-6} \\ Max\left(\frac{|r_{x,y}^{or} - r_{x,y}|}{|r_{x,y}^{or}|}\right) &\simeq 8.6 \cdot 10^{-5} \end{aligned}$$

On peut ensuite s'intéresser, à titre d'information, à l'erreur commise sur les angles $\theta_{x,y}$ en choisissant $g_{x,y} = 0$ en tout point (x,y) : l'erreur moyenne $E(|\theta_{x,y} - \theta_{x,y}^{or}|)$ est de 3° , pour une erreur maximale de 13.5° .

b. Phase de minimisation

Cette phase de minimisation constitue la phase la plus importante et la plus délicate de l'algorithme. La section a montre en effet que la phase d'initialisation remplissait parfaitement son rôle (se placer sur les bonnes courbes de niveau de Φ en chaque point), et que les approximations effectuées dans notre contexte d'étude étaient pertinentes. Quoi qu'il arrive, à la fin de l'étape de minimisation, on obtiendra une carte Θ induisant une carte des normales $\mathcal{N}$ qui vérifiera presque exactement la contrainte d'éclairement. La question est-donc de savoir si il est possible d'obtenir une carte des normales :

- Admissible et intégrable, c'est à dire vérifiant la condition de Schwarz.
- Proche de la carte des normales originales.

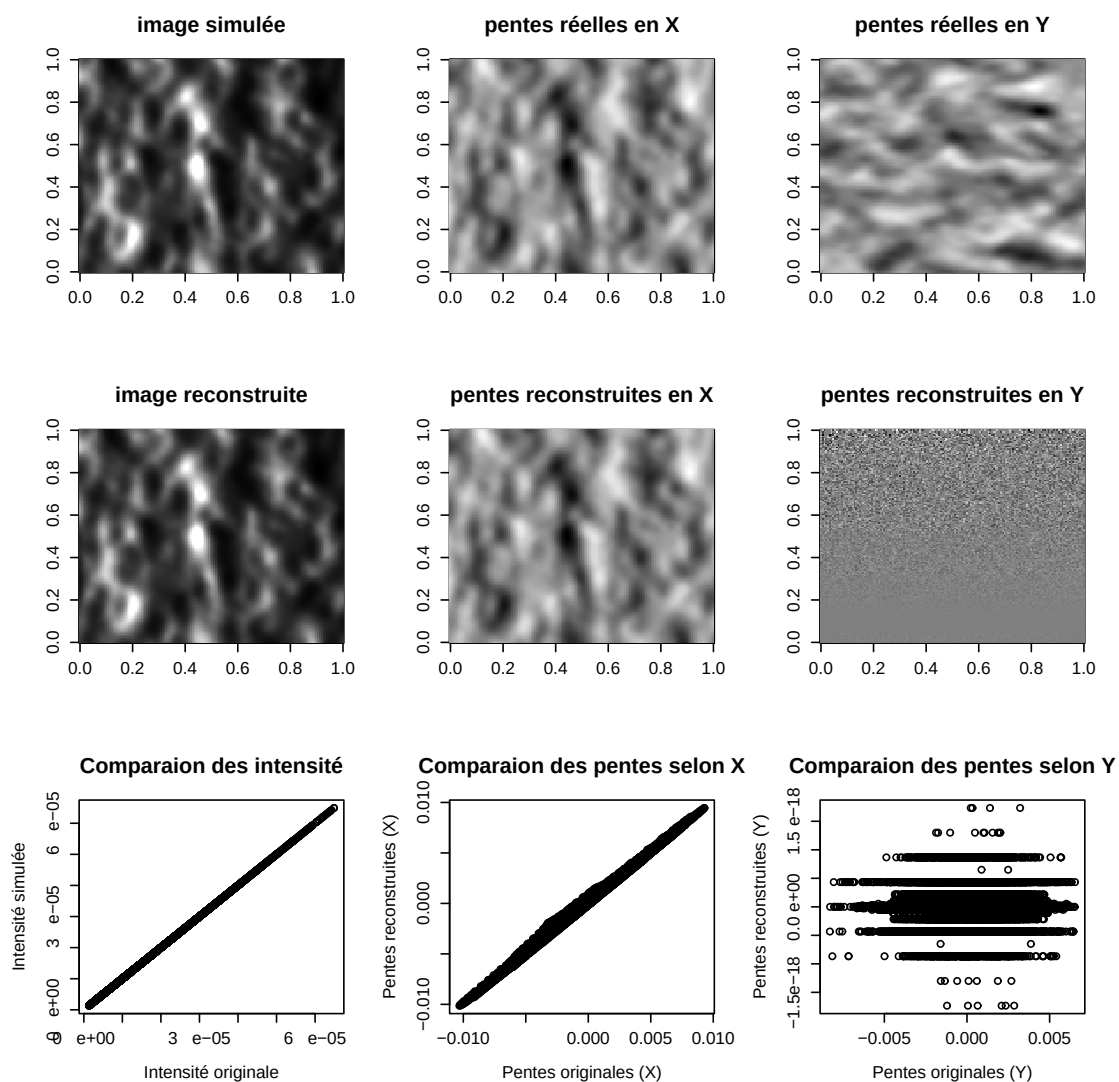


FIG. 7.14.: Résultat de la phase d'initialisation. En haut, image originale et pentes de la topographie originale. Au milieu, image résultante et pentes reconstruites après la phase d'initialisation. En bas, corrélation des images : les images originale et résultante sont pratiquement identiques, les pentes en X sont très semblables mais les pentes résultante en Y sont nulles (aux erreurs numériques près).

La phase de minimisation est une recherche itérative d'un minimum : il est donc nécessaire de se fixer un critère d'arrêt, et pour ces essais nous fixerons simplement et arbitrairement le nombre d'itérations avant de pouvoir fixer un critère plus précis en terme de tolérance sur la norme du gradient par exemple.

Il paraît indispensable de faire intervenir la condition de Schwarz dans notre fonction de coût, et pour ces premiers essais, nous choisirons simplement $W = W^1$.

Choix des noyaux de dérivation Telle que nous l'avons présentée à la section b, cette fonction de coût utilise des noyaux "dérivées de gaussiennes" pour effectuer la dérivation "lissée" de la surface. Comme nous l'avons signalé alors, il est bien entendu essentiel, si l'on veut avoir une chance de converger vers la topographie originale, que celle-ci vérifie la condition de Schwarz pour le noyau choisi. Or les fortes variations dans la topographie de la surface non lissée (figure 7.14) rendent l'approximation linéaire de la dérivation classique inappropriée. Nous allons donc chercher à quantifier la manière dont les cartes f et g vérifient la condition de Schwarz pour plusieurs noyaux de dérivation de façon à choisir le plus approprié.

Dans ce qui suit nous noterons :

$$\begin{aligned} fy &= \frac{\partial f}{\partial y} = f * K_y \\ gx &= \frac{\partial g}{\partial x} = g * K_x \end{aligned}$$

La condition de Schwarz sera vérifiée en pratique si $(gx - fy)^2 \simeq 0$ en tout point de la surface sur laquelle la convolution peut être calculée sans effet de bord. Pour quantifier ce " $\simeq 0$ ", on se base sur l'espérance $E((gx - fy)^2)$ qui doit être la plus faible possible. Un indicateur intéressant est obtenu en comparant cette espérance à la situation où les cartes fy et gx sont indépendantes (et donc ne vérifient pas la condition de Schwarz).

Si la condition d'indépendance est respectée, on obtient :

$$\begin{aligned} E((gx - fy)^2) &= E(gx^2) + E(fy^2) - 2E(gx.fy) \\ E((gx - fy)^2) &= E(gx^2) + E(fy^2) - 2E(gx).E(fy) \end{aligned}$$

On pose donc :

$$E_{ind} = E(gx^2) + E(fy^2) - 2E(gx).E(fy)$$

et on forme l'estimateur :

$$e = \frac{E_{ind} - E((gx - fy)^2)}{E_{ind}}$$

Cet estimateur vaut 0 si gx et fy sont indépendants, c'est à dire ne vérifie pas la condition de Schwarz et vaut 1 si la condition de Schwarz est vérifiée parfaitement, c'est à dire si $E((gx - fy)^2) = 0$. En pratique, on peut considérer cet estimateur comme une fonction des noyaux K_x et K_y utilisés pour la dérivation, et on va alors chercher des noyaux pour lesquels $e(K_x, K_y)$ est suffisamment proche de 1, sans toutefois que les noyaux soient de trop grande taille, afin de maximiser la taille de l'image sur laquelle on peut calculer les dérivées sans effet de bord ($d - dk + 1$ avec les notations du paragraphe b).

Voyons les résultats des tests effectués pour les noyaux de dérivation classiques D_x et D_y obtenus par simple différence entre deux pixels consécutifs :

$$e(D_x, D_y) \simeq 0.37$$

Comme annoncé dans la section b, cette valeur justifie d'utiliser des noyaux plus régularisants.

On peut maintenant s'intéresser aux dérivées de filtres gaussiens pour différentes longueurs d'onde de coupure. Les tests réalisés montrent qu'une longueur d'onde de coupure aux alentours de 0.4 mm permet de vérifier la condition de Schwarz à 99,99%.

Dans le cas où la topographie étudiée est déjà obtenue par filtrage gaussien de la topographie originale avec une longueur d'onde de coupure à 0.5 mm, comme ce sera le cas dans la suite, le choix du filtre importe moins, car les fluctuations de normale sont plus faible, et l'hypothèse de linéarité est plus juste. En fait, dans ce cas, si l'on considère l'associativité et la commutativité du produit de convolution, l'opération revient à composer deux filtres gaussiens avant dérivation (le filtre de prétraitement est composé avec la gaussienne dont on calcule la dérivée pour calculer le noyau de dérivation).

Dans ce cas, on obtient $e(D_x, D_y) \simeq 0.98$, et lorsqu'on utilise un noyau gaussien, une longueur d'onde de coupure de l'ordre de 0.2 mm (que nous utiliserons par la suite) permet d'obtenir la même valeur de 99,99% pour la vérification de la condition de Schwarz.

Application On peut maintenant comparer la valeur de la fonction de coût W^1 pour les pentes de la topographie originale, $W^1(f_{or}, g_{or}) \simeq 10^{-4}$, avec la valeur $W^1(f_0, g_0) \simeq 1$ pour les pentes issues du processus d'initialisation.

Après 30 itérations du processus de minimisation, on obtient le résultat visible à la figure 7.15. La valeur prise par la fonction de coût est alors de $W^1(f, g) \simeq 10^{-3}$

Après la phase de minimisation, on peut intégrer les pentes f et g obtenues pour obtenir une topographie T . On peut également recalculer les cartes f_r et g_r à partir de T (qui seront légèrement différentes des pentes f et g car l'intégration n'est pas exacte). On obtient alors le résultat présenté à la figure 7.16. Les remarques que l'on peut faire concernent tout d'abord la bonne vérification de la condition d'éclairement : la méthode décrite précédemment fournit une topographie dont l'image obtenue dans des conditions d'acquisition parfaites serait identique à celle obtenue à partir de la topographie originale. En revanche, la topographie résultante et la topographie originale ne sont pas identiques et diffèrent essentiellement d'un profil selon Y (voir figure 7.17).

Dans le cas étudié ici, on peut noter que les valeurs moyennes des pentes reconstruites f et g , ainsi que celle des erreurs sur la topographie finale sont nulles : l'introduction du critère de nullité des pentes dans la fonction de coût n'apporterait donc rien de plus à la reconstruction... Nous avons cependant remarqué que leur usage pouvait améliorer quelques peu les résultats dans le cas où l'on s'intéresse aux topographies non lissées.

c. Amélioration des résultats

Pour améliorer quelques peu ces résultats, nous avons modifié légèrement la procédure d'initialisation de l'image des pentes. Partant du constat que la condition de Schwarz nous

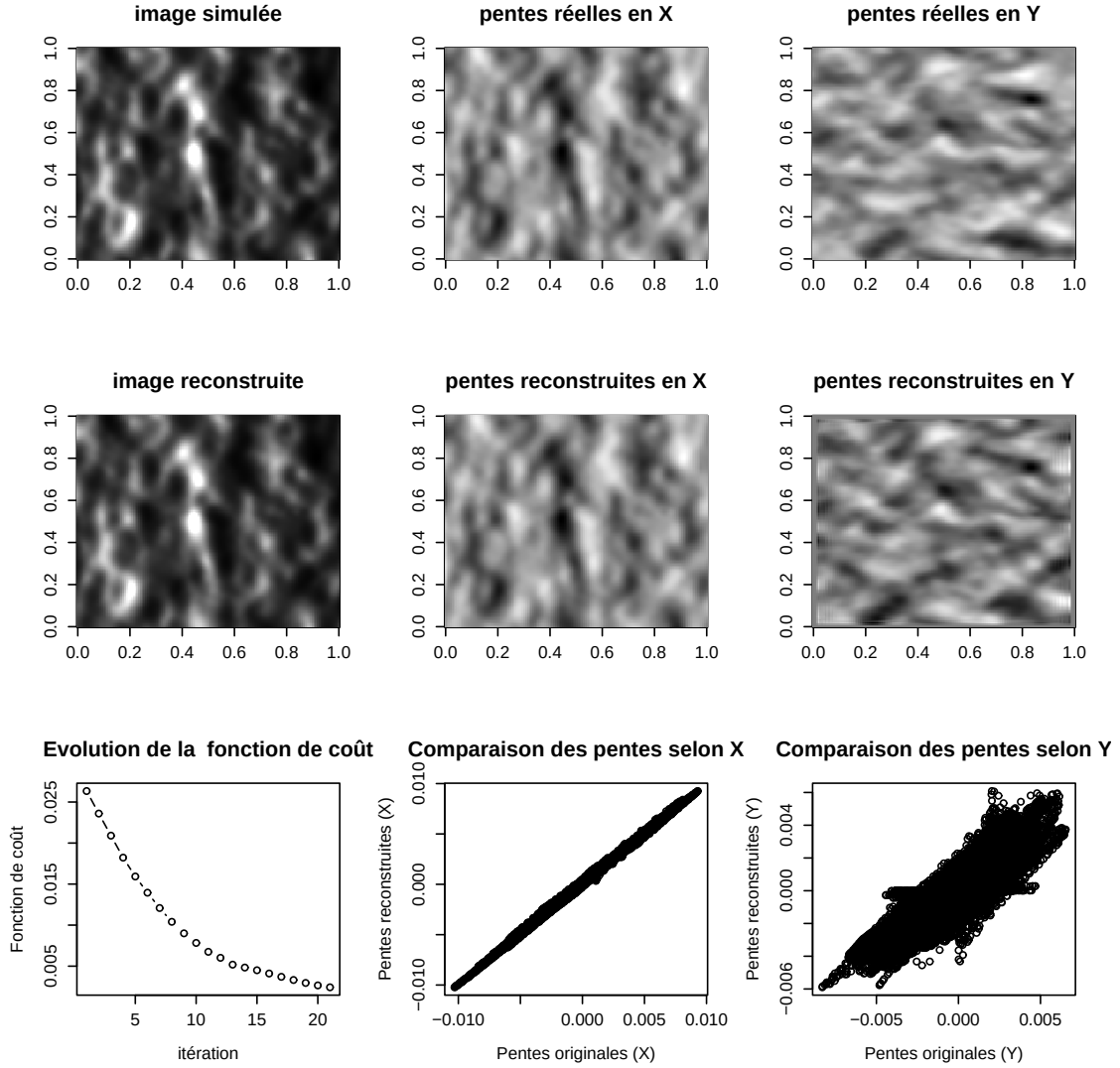


FIG. 7.15.: Résultat de la phase de minimisation après 30 itérations, en n'utilisant que le terme correspondant à la condition de Schwarz dans la fonction de coût. Seules les 20 dernières itérations sont représentées dans le graphe d'évolution de la fonction de coût.

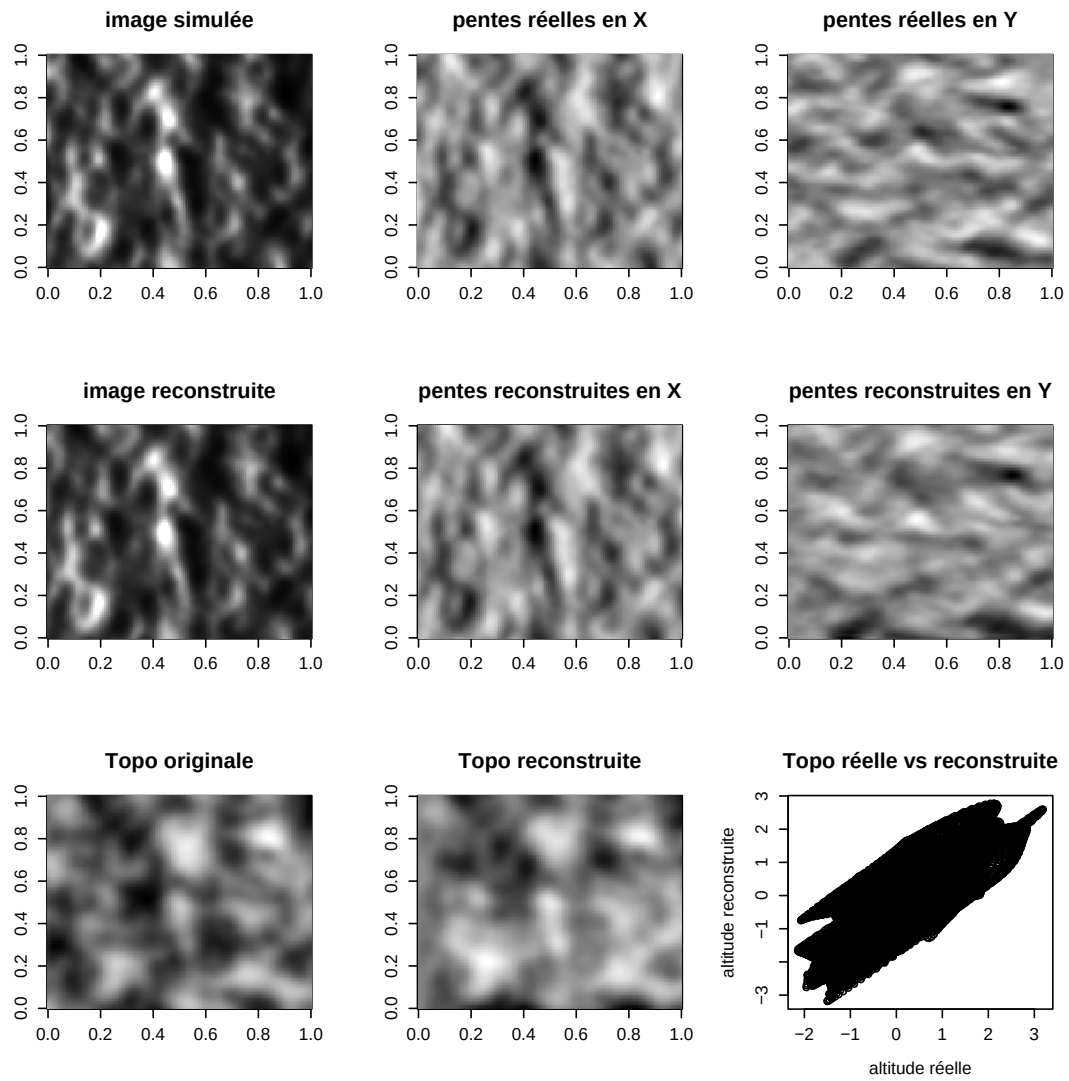
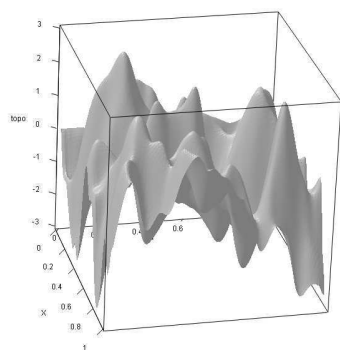
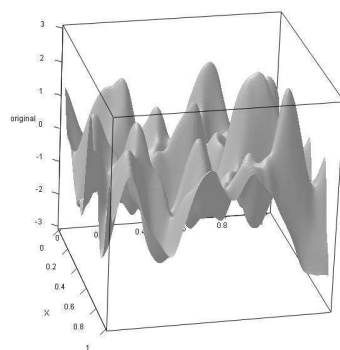


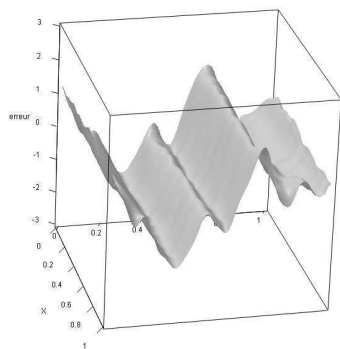
FIG. 7.16.: Résultats obtenus après la phase d'intégration de la carte des normales.



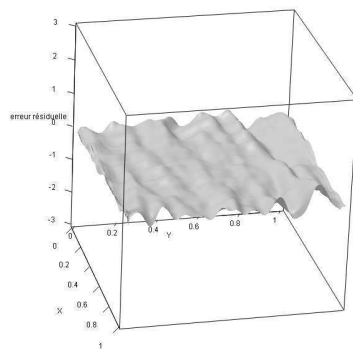
(a) Vue 3d de la topographie originale



(b) Vue 3d de la topographie reconstruite



(c) Différence entre topographie originale et reconstruite. L'erreur est approximativement un profil selon l'axe Y



(d) Erreur résiduelle après correction du profil moyen d'erreur.

FIG. 7.17.: Comparaison entre topographie reconstruite et topographie originale.

permet de reconstruire les pentes selon Y à un profil près dans cette même direction, nous avons cherché à déterminer des points "d'ancrage" pour lesquels la valeur de g pourrait être estimée sans trop d'erreur, ce qui permettrait de contraindre le profil d'erreur, et d'en limiter ainsi l'amplitude.

Nos surfaces ont en général des pentes assez faibles, et on peut même considérer qu'elles sont bornées, donc si la pente selon X (connue avec une bonne précision) est forte, cela signifie que la pente selon Y a des chances d'être faible. Ensuite, la connaissance des pentes selon X nous permet de calculer la dérivée $\frac{\partial g}{\partial x}$ (grâce au théorème de Schwarz). On peut donc intégrer ligne par ligne et connaître une première approximation de g , à une constante près sur chaque ligne, c'est à dire à un profil près :

$$\forall y_0, g_{x,y_0} = C_{y_0} + \int_0^x \frac{\partial g}{\partial x}(t, y_0) dt$$

On fixe ensuite un seuil sur les pentes en x , et on ajuste ensuite les constantes d'intégration de façon à ce que pour tous les points (x, y) tels que $\|f_{x,y}\| > s$, la pente $\|g_{x,y}\|$ soit le plus faible possible. Soit $A_{y_0} = \{x / \|f_{x,y_0}\| > s\}$, alors :

$$A_{y_0} \neq \emptyset \Rightarrow C_{y_0} = -\frac{1}{\text{card}(A_{y_0})} \sum_{x \in A_{y_0}} \int_0^x \frac{\partial g}{\partial x}(t, y_0) dt$$

Cela nous fournit les constantes d'intégration pour chaque ligne contenant un ou plusieurs points d'ancrage, ensuite, on estime les autres constantes d'intégrations par interpolation linéaire. Soit y_0 tel que $A_{y_0} = \emptyset$, et soit $y_i = \sup\{y / y < y_0 \& A_y \neq \emptyset\}$ et $y_s = \inf\{y / y > y_0 \& A_y \neq \emptyset\}$, alors :

$$C_{y_0} = (y_0 - y_i) * \frac{C_{y_s} - C_{y_i}}{y_s - y_i}$$

De cette façon, on obtient une carte g_{init} , que l'on peut réintroduire dans le calcul d'initialisation de la carte des angles Θ . Pour tout point (x, y) , $\theta_{x,y}$ est choisi de façon à ce que $g_{x,y} = g_{init}(x, y)$. Cela modifie bien évidemment les valeurs de $f_{x,y}$ calculée, mais peu, pour les raisons de configuration particulière déjà mentionnée. Dans ces conditions, on obtient une carte Θ_0 compatible avec la condition d'éclairement et choisie de telle sorte que les pentes de la surfaces ne soient pas trop fortes. On peut ensuite utiliser le processus de minimisation déjà décrit en imposant aux normales de vérifier la condition de Schwarz ainsi que l'absence de corrélation avec les variables X et Y (au cours des essais effectués, c'est ce choix de la fonction de coût qui a fourni les meilleurs résultats). La topographie reconstruite par cette méthode comparée à la topographie originale est visible à la figure 7.18. Les résultats font toujours apparaître une différence sensible par rapport à la topographie originale, cette différence est toujours principalement un profil selon l'axe des y , mais ce profil est de plus faible amplitude qu'à la figure 7.16.

7.2.7. Utilisation de deux incidences

La forte asymétrie entre les directions X et Y incite, pour améliorer les résultats, à utiliser deux incidences, c'est à dire deux images obtenues pour des directions d'observation orthogonales. Dans ce cas de figure, on observe donc deux réponses optiques pour

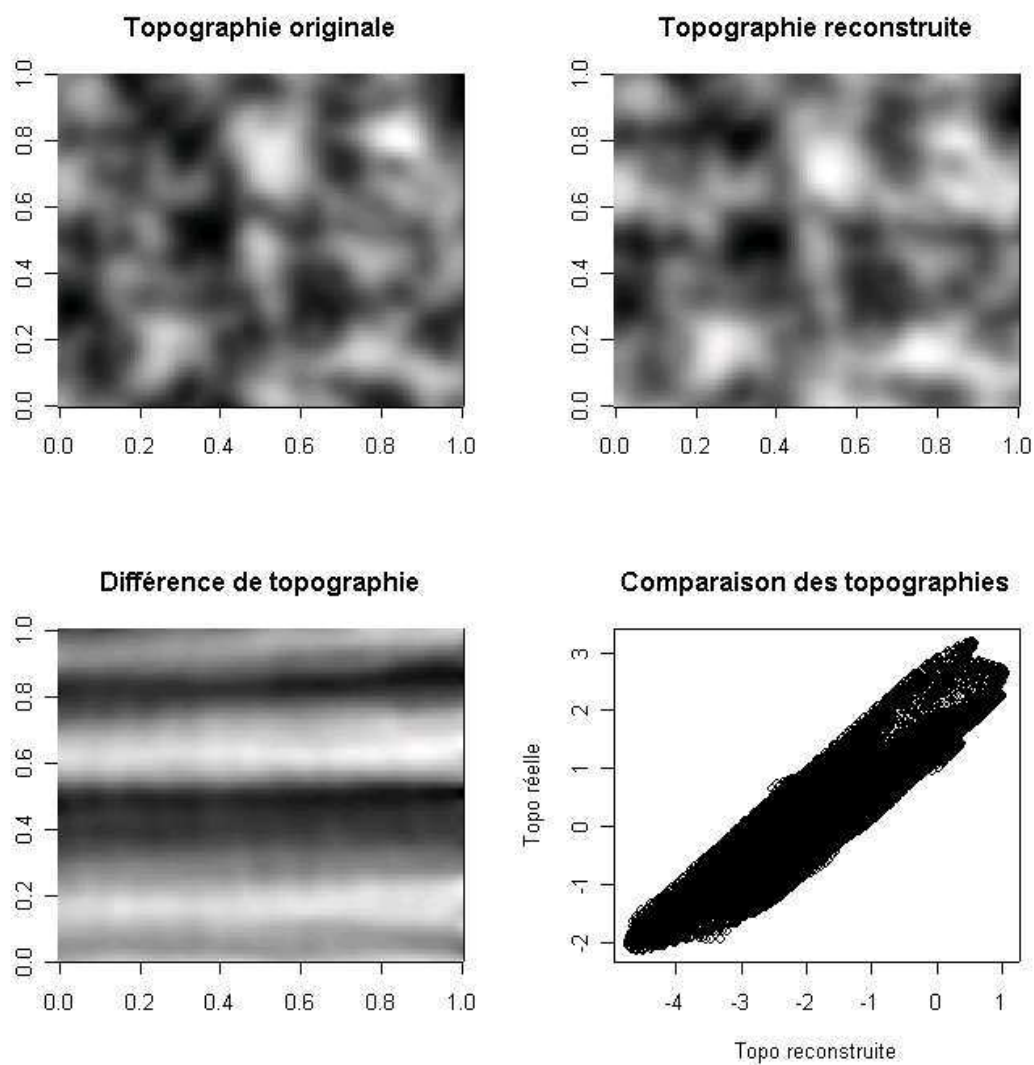


FIG. 7.18.: Résultat de l'algorithme de reconstruction lorsque l'étape d'initialisation fait intervenir des "points d'ancrage".

chaque facette de la surface. D'après ce qui précède, la première fournit f avec une assez bonne précision, tandis que l'autre fournit g avec une assez bonne précision. Il est même probable que le calcul analytique de f et g soit possible en combinant les équations des arcs de cercle obtenus dans les espaces correspondant aux deux directions d'observation (voir section 7.2.4). Pourtant, nous avons préféré utiliser le même schéma de minimisation avec deux incidences qu'avec une, de façon à valider les termes de la fonction de coût utilisée. Nous allons procéder exactement de la même façon qu'à la section précédente en utilisant une seule image. Puis le résultat obtenu sera utilisé pour initialiser l'image des pentes pour une deuxième phase de minimisation utilisant la seconde incidence. On observe une bonne convergence des résultats après seulement quelques itérations de ce procédé. La figure 7.19 montre que les pentes reconstruites par cette méthode sont pratiquement identiques aux pentes originales, de même que la topographie reconstruite est très proche de la topographie originale.

Si ces résultats semblent très bons, il faut là encore préciser que les conditions d'acquisition sont idéales puisqu'elles ont été simulées. En particulier, le recalage des images issues des deux incidences se fait parfaitement : chaque pixel de l'une correspond très exactement à un pixel de l'autre. En pratique, ce problème de recalage vient s'ajouter aux autres problèmes que nous n'avons pas traités et qui viendraient compliquer la tâche de la reconstruction (configuration optique imparfaitement connue, puissance de la source inconnue, ouverture de diaphragme non nulle ...).

7.2.8. Conclusion et perspectives

Les techniques de simulation d'images caméra nous ont permis de rendre compte de phénomènes optiques complexes tels que le flou dû à l'ouverture de diaphragme de la caméra et les résultats obtenus dans ce cadre sont réalistes. L'application industrielle principale et directe de ces développements serait le remplacement des tests physiques de configuration optique d'un banc "source/caméra" par des simulations informatiques systématiques avec recherche de paramètres optimaux pour la classification. Pour parvenir à un tel résultat, il sera important d'approfondir l'étude de la stabilité du paramètre ρ du modèle de Ward, ainsi que celle de la sensibilité des images simulées à ce même paramètre. De la même façon, une comparaison plus fine entre images réelles et simulées pour une même surface est requise pour véritablement quantifier le réalisme des simulations.

Les résultats de reconstruction topographique ont également été obtenus dans des conditions théoriques idéales et irréalistes sur bien des points. Les deux applications immédiates de ces développements sont d'une part l'amélioration de la classification des surfaces à partir d'images caméra et d'autre part la caractérisation d'une topographie possédant des propriétés de rendu optique définies. Nous sommes encore assez loin de ces objectifs, mais l'utilisation de nos algorithmes de reconstruction permet tout de même d'accéder à certaines propriétés topographiques d'une surface (reconstruite à un profil près) qui pourraient être utilisées comme descripteurs dans une optique de classification. Nous avons aussi montré que l'utilisation de plusieurs incidences (plusieurs images obtenues dans des configurations optiques différentes) permet de reconstruire la topographie de la surface originale pratiquement parfaitement, et qu'il peut donc s'agir d'une voie à

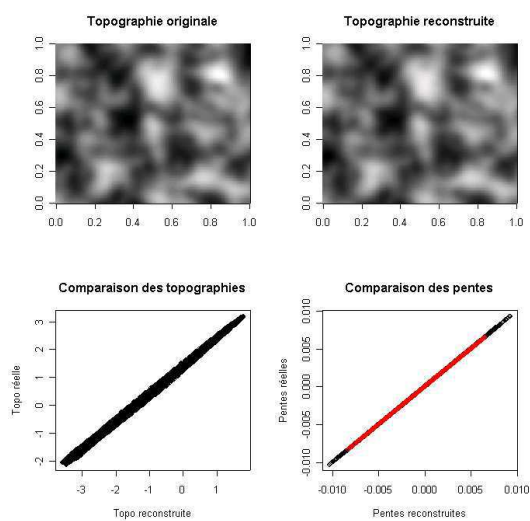
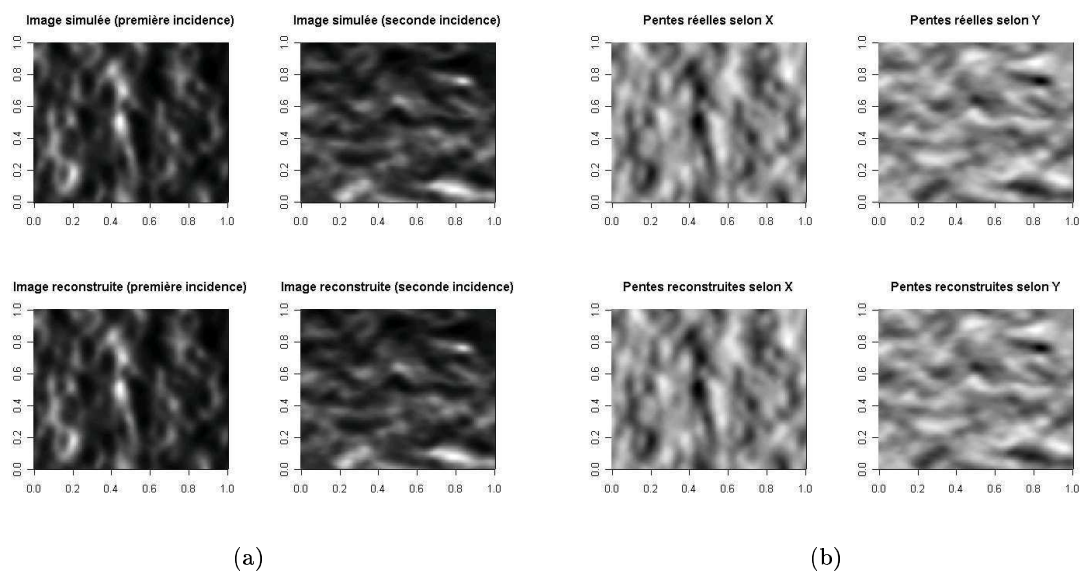


FIG. 7.19.: Reconstruction à partir de deux incidences

explorer si l'utilisation d'une seule caméra ne suffisait pas à caractériser pleinement ou à classer correctement certains produits.

Sur le plan technique, nous avons pu remarquer que les résultats de reconstruction de surfaces possédant une certaine régularité (filtré par une gaussienne) étaient meilleurs que sur des topographies non filtrées comprenant des composantes de haute fréquence : cela suggère peut-être un traitement multi-échelles traitant séparément les différentes composantes fréquentielles de l'image originale, de façon à reconstruire des rugosités de plus en plus fines de la topographie originale.

Chapitre 8.

Conclusion et perspectives

Au début de cette thèse, nous sommes parti d'une question simple : Comment caractériser en ligne la qualité de l'aspect des surfaces ? Nous avons apporté une réponse à cette question par le biais de nouveaux descripteurs de surfaces, de l'utilisation de techniques de bootstrap, d'une méthodologie générale d'analyse et de classification et finalement d'une étude expérimentale ayant permis la mise en ligne d'un premier prototype de contrôle de l'aspect des surfaces.

Si notre système d'analyse a trouvé une application directe dans le cadre de l'étude des surfaces "Peau d'orange", nous nous sommes efforcés de conserver une certaine généralité dans le développement de nos outils, de même qu'une certaine indépendance entre les organes principaux de notre méthode. Ceci ajouté à une structure informatique adaptée, doit faciliter l'intégration de nouveaux descripteurs aussi bien que de nouvelles méthodes. Nous avons par exemple choisi de diminuer les dimensions de l'espace des descripteurs par analyse factorielle, mais nous avons aussi signalé qu'il existait des méthodes "à noyaux" (*kernel pca* par exemple, ou encore analyse de Fisher par noyaux) dont l'utilisation permettrait de faire intervenir des comportements non linéaires vis à vis des descripteurs originellement choisis. Idéalement, l'implémentation finale de nos outils devrait permettre de substituer facilement une stratégie de réduction des dimensions à une autre, ou une méthode de classification à une autre, c'est une condition sine qua non pour que nos outils restent évolutifs et puissent améliorer leur performances en intégrant de nouveaux développements mathématiques, algorithmiques ou numériques.

8.1. Cuique suum (A chacun le sien)

Une question simple n'admet pas toujours de réponse simple, ni surtout directe, et nous avons été amené à nous intéresser à de nombreux domaines scientifiques, parfois très différents : morphologie mathématique, bien sûr, mais aussi statistique, géostatistique, optique, optimisation ... Réunir dans un tout cohérent des briques issues de chacune de ces disciplines afin de répondre à une attente industrielle constitue à notre avis l'une des principales richesses de notre travail. Malgré tout, il nous paraît important ici de faire un inventaire détaillé des outils que nous avons pu utiliser en précisant explicitement quels ont été nos apports personnels et quelles ont été les méthodes développées par d'autre qui ont simplement été réutilisées dans un contexte particulier.

Dans le second chapitre, la plupart des descripteurs que nous avons mis en oeuvre sont issus de méthodes existantes que nous avons adaptées à l'étude des surfaces rugueuses. L'algorithme de segmentation des surfaces cependant, bien qu'issue d'une approche classique (extraction de marqueur puis ligne de partage des eaux du gradient), est original, de même que l'obtention d'un profil de tortuosités après *normalisation* d'une fonction distance pour des ensembles de départ et d'arrivée particuliers.

Dans le troisième chapitre, notre apport principal a été *l'adjonction systématique d'une distribution d'erreur* aux descripteurs de surface. Les méthodes utilisées pour appliquer une forme particulière de bootstrap adaptée à chaque variable sont originales de même que la vérification systématique de la pertinence des résultats obtenus sur des modèles morphologiques de fonctions aléatoires.

Dans le quatrième chapitre, nous présentons une méthodologie générale d'étude des surfaces dont chaque élément pris séparément est classique. L'analyse factorielle par exemple, avait déjà été utilisée par A. Aubert [1, 3] comme "instrument de mesure de ressemblance texturale", mais les descripteurs étaient tous traités séparément : nous avons ajouté une phase *d'analyse globale* à ce schéma existant. L'utilisation des *échantillons bootstrap* pour évaluer les *densités de probabilités* nécessaires à la classification bayésienne est aussi de notre initiative.

Dans le cinquième chapitre, nous décrivons l'application industrielle de notre méthodologie. L'ensemble des mesures ont été réalisés par le département Mesure-Contrôle d'Arcelor Research, de même que la mise au point du prototype destiné aux premiers essais en ligne. Notre rôle principal réside donc dans *l'interprétation des mesures brutes et des résultats* et dans *les traitements* à effectuer pour les rendre utilisables.

Dans le sixième chapitre, nous proposons une représentation arborescente pour caractériser les surfaces. Si la construction d'un arbre d'inondation est une pratique courante, l'utilisation de celui-ci pour caractériser des textures l'est beaucoup moins, a fortiori lorsque l'on prend en compte la structure de plusieurs sous-arbres, de façon à pouvoir facilement utiliser des méthodes de bootstrap. *L'algorithme de division en sous-arbre* est également original, de même que l'idée d'un *sur-échantillonnage "mixte"* composé d'un modèle d'arbre markovien dont les feuilles seraient obtenues par bootstrap.

Dans le septième chapitre, nous adaptons dans un premier temps des algorithmes classiques de lancer de rayons à une configuration optique particulière tenant compte des effets de flou dûs à l'ouverture de diaphragme de la caméra. Cet algorithme n'est pas en soi véritablement nouveau, mais la possibilité de remplacer avec succès des tests physiques par une simulation réaliste serait une belle performance expérimentale et les applications industrielles seraient nombreuses : nous avons effectué un pas dans ce sens. Dans la section traitant du problème inverse, en revanche, notre apport principal a été *le choix d'un nouveau système de variables* permettant de respecter systématiquement la

condition d'éclairément sans utiliser de phase de projection comme dans l'algorithme de Worthington et Hancock[67]. Nous avons aussi utilisé *une fonction de coût adaptée aux surfaces de type "peau d'orange"* contenant des termes originaux absents des publications concernant les problèmes de "shape from shading".

8.2. Perspectives

Au sein d'Arcelor Research, cette thèse s'inscrit dans une dynamique générale de caractérisation, mais aussi de prédiction de l'aspect des surfaces après peinture. Il s'agit essentiellement d'une étape vers le problème inverse, c'est à dire vers la capacité à calculer puis à produire physiquement une topographie dont l'aspect peint remplira les desiderata d'un client. Un tel projet requiert de nombreux efforts de recherche complémentaires. Tout d'abord, il faut comprendre et simuler l'effet d'une peinture sur une topographie : certaines rugosités sont nivelées par la revêtement, mais de nouveaux motifs de plus grande longueur d'onde peuvent éventuellement apparaître... Actuellement, une prédiction de la topographie après peinture est possible lorsque l'opération se fait sur une tôle horizontale. Lorsqu'elle se fait sur une tôle verticale, des termes non linéaires apparaissent (en raison de la gravitation) et viennent perturber les équations de l'écoulement. On doit également pouvoir fournir au client un moyen de définir l'aspect recherché, ce qui peut s'obtenir par le biais d'un jeu de descripteurs adaptés ou par celui d'exemples de rendu optiques simulés. Une fois la cible caractérisée en terme d'aspect, il "suffit" de résoudre deux problèmes inverses : le problème inverse optique permettant de caractériser la topographie après peinture requise pour obtenir la l'aspect visé, puis le problème inverse de l'opération de revêtement permettant de caractériser la topographie avant peinture. Enfin, il faut maîtriser les paramètres du processus de production pour obtenir physiquement une surface remplissant les conditions ainsi définies.

Qu'il s'agisse de caractérisation des surfaces au travers de nouveaux descripteurs, de bootstrap pour assurer la stabilité et la pertinence des variables utilisées, de simulation de rendu ou de "shape from shading", nous espérons que les outils que nous avons développés pourront être enrichis et utilisés utilement dans cette entreprise.

Bibliographie

- [1] A. AUBERT. « *Propriétés optiques des surfaces rugueuses aléatoires* ». PhD thesis, Ecole des Mines de Paris, 1999.
- [2] A. AUBERT et D. JEULIN. « L'optique géométrique comme outil de classification de surfaces rugueuses aléatoires. Problème direct et inverse ». Dans *SF 1999*, Palaiseau, 28 Juin - 2 Juillet, 1999. Ecole Polytechnique.
- [3] A. AUBERT et D. JEULIN. « Classification morphologique de surfaces rugueuses ». *Revue de Métallurgie - CIT/Science et Génie des Matériaux*, pages 247–262, feb 2000.
- [4] A. AUBERT, D. JEULIN, et R. HASHIMOTO. « Surface texture classification from morphological transformations ». Dans L. Vincent J. GOUTSIAS et D.S. BLOOMBERG, éditeurs, *Proc. ISMM'2000, Palo Alto, USA (16-28 June 2000)*, in *Mathematical morphology and its applications to Image and Signal Processing*, pages 253–252. Kluwer Academic Publishers, 2000.
- [5] A. AUBERT, D. JEULIN, et P.J. KRAUTH. « Propriétés optiques des surfaces rugueuses aléatoires ». Dans *Colloque Nouveaux Moyens Optiques pour l'Industrie II*, Mittelwihr, 17-19 Novembre 1999. Club SFO "Contrôles et mesures optiques pour l'industrie".
- [6] N.T.J. BAILEY. *The elements of stochastic processes, with applications to the natural sciences*. J. Wiley & Sons, 1964.
- [7] S. BALBI. « On stability in non symmetrical correspondence analysis using bootstrap ». *Statistica Applicata*, 4(4), 1992.
- [8] R. BELLMAN et T.E. HARRIS. « On age-dependent binary branching processes ». Dans *Ann. Math.*, volume 55, pages 280–295, 1952.
- [9] J.P. BENZECRI. *L'analyse de données*. Dunod, 1973.
- [10] S. BEUCHER. « *Segmentation d'images et morphologie mathématiques* ». PhD thesis, Ecole des Mines de Paris, 1990.
- [11] S. BEUCHER. « The watershed transformation applied to image segmentation ». Dans *Scanning Microscopy International, suppl. 6*, pages 299–314, Cambridge, UK, 1992. 10th Pfefferkorn Conf. on Signal and Image Processing in Microscopy and Microanalysis, 1991.
- [12] S. BEUCHER. « Watershed, hierarchical segmentation and waterfall algorithm ». Dans Jean SERRA et Pierre SOILLE, éditeurs, *Proc. Mathematical morphology and it's applications to image processing, Fontainebleau, Sept. 1994*, pages 69–76, Nld, 1994. Kluwer Ac. Publ.

- [13] S. BEUCHER, M. BLANC, et T. HERSANT. « Analyse quantitative de clichés stéréoscopiques ». *J. of Microsc. Spectrosc. Electron.*, 7 :105–106, 1982.
- [14] S. BEUCHER et F. MEYER. « *Mathematical Morphology in Image Processing* », Chapitre The morphological approach of segmentation : the watershed transformation. Dougherty E.(Editor), New York, 1992.
- [15] A.T. BHARUCHA-REID. *Elements of the theory of Markov processes and their applications*. McGraw-Hill, 1960.
- [16] M. BIGERELLE, D. NAJJAR, et A. IOST. « Description d'une nouvelle méthode de corrélation entre la rugosité et une propriété de surface. Application à la brillance de tôles skin-pasées ». *La revue de Métallurgie*, 2002.
- [17] L. BIRGE et Y. ROZENHOLC. « How many bins should be put in a regular histogram ». Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599, Université Paris VI et Paris VII, Prépublication 721, 2002.
- [18] J. CARR, W. R. FRIGHT, et R. K. BEATSON. « Surface interpolation with radial basis functions for medical imaging ». Dans *IEEE transactions on medical imaging*, volume XX, février 1997.
- [19] D. COMANICIU et P. MEER. « Mean shift : a robust approach toward feature space analysis ». *IEEE transactions on pattern analysis and machine intelligence*, 24(5), 2002.
- [20] D. COMANICIU et P. MEER. « An algorithm for data-driven bandwidth selection ». *IEEE transactions on pattern analysis and machine intelligence*, 25(5), 2003.
- [21] D. COMANICIU, V. RAMESH, et P. MEER. « The variable bandwidth mean shift and data-driven scale selection ». volume 1, pages 438–445, Vancouver, Canada, 2001. IEEE Int. Conf. Computer Vision (ICCV'01).
- [22] B. EFRON et R. TIBSHIRANI. *An introduction to the bootstrap*. Chapman & Hall, 1993.
- [23] W. FELLER. *An Introduction to Probability Theory and its Applications*, volume 1,2. J. Wiley & Sons, 1957.
- [24] G. FRICOUT et D. JEULIN. A model of random tree to simulate the growth of kidneys. Ecole des Mines de Paris, oct 2001. note interne N-51/01/MM.
- [25] G. FRICOUT et D. JEULIN. « On line characterization of painted surfaces aspect by morphological and image processing tools ». Dans *Proceedings of the european coil coating association autumn congress*, nov 2003.
- [26] G. FRICOUT et D. JEULIN. « Propriétés morphologiques et classification de surfaces rugueuses ». Dans *Quatrième colloque francophone, Méthodes et techniques optiques pour l'industrie*, nov 2003.
- [27] G. FRICOUT et D. JEULIN. « Texture classification from their morphological properties ». Dans B.-G. Rosén T.R. Thomas L. BLUNT, éditeur, *Transactions of the 9th Int. Conf. on Metrology and Properties of Engineering Surfaces*, Halmstad, Sweden, sep 2003.

- [28] G. FRICOUT et D. JEULIN. « Análisis de imagen y morfología matemática para la caracterización en línea del aspecto de las superficies pintadas ». *pinturas y acabados industriales*, (292), jun 2004.
- [29] C. GOMILA. « *Mise en correspondance de partitions en vue du suivi d'objets* ». PhD thesis, Ecole des Mines de Paris, 2001.
- [30] B.K. HORNE et M.J. BROOKS, éditeurs. *Shape From Shading*. The MIT Press, 1989.
- [31] D. JEULIN. « Morphological modeling of surfaces ». *Surface Engineering*, 14(3) :199–204, 1998.
- [32] D. JEULIN. « Introduction to model of random trees ». Rapport Technique, Paris School of Mines publication, jun 2001.
- [33] D. JEULIN et P. JEULIN. « Synthesis of Rough Surfaces by Random Morphological Models ». Dans *Proc. 3rd European Symposium of Stereology*, volume 3, pages 239–246. Stereol. Iugosl. 1981, 1981.
- [34] D. JEULIN et P. LAURENCE. « Morphological simulation of the roughness transfer on steel sheets ». *Microscopy, Microanalysis, Microstructure*, 7 :541–547, 1996.
- [35] D. JEULIN et P. LAURENCE. « Simulation of rough surfaces by morphological random functions ». *Journal of Electronic Imaging*, 6 :16–30, 1997.
- [36] X.Q. JIANG et P.J. SCOTT. « Review of surface texture work in ISO ». Dans B.-G. Rosén T.R. Thomas L. BLUNT, éditeur, *Transactions of the 9th Int. Conf. on Metrology and Properties of Engineering Surfaces*, Halmstad, Sweden, sep 2003.
- [37] D.G. KENDALL. « On the role of variable generation time in the development of a stochastic birth process ». Dans *Biometrika*, volume 35, pages 316–330, 1948.
- [38] I.S. KWEON et T. KANADE. « Extracting topographic features from elevation maps ». *CVGIP : image understanding*, 59(2) :171–182, 1994.
- [39] S. LAHIRI. *Resampling method for dependant data*. Springer, 2003.
- [40] S. LEE. « *Application de l'analyse continue par ondelettes 1D et 2D a la caractérisation multi-échelle de la rugosité* ». PhD thesis, Ecole Centrale de Lyon, septembre 1997.
- [41] M. LIVNY, R. RAMAKRISHNAN, et T. ZHANG. « Fast density and probability estimation using CF-kernel method for very large databases ». Rapport Technique, University of Wsconsin-Madison, Computer Sciences Department, 1996.
- [42] S. MALLAT. *A wavelet tour of signal processing*. Academic Press, 1999.
- [43] G. MATHERON. *Les variables régionalisées et leur estimation*. Masson, 1965.
- [44] G. MATHERON. *Estimating and choosing : an essay on probability in practice*. Springer, 1989. Translated from the French "Estimer et choisir" by A.M. Hasofer.
- [45] F. MEYER. « Un algorithme optimal de ligne de partage des eaux ». Dans *Proceedings 8ème congrès AFCET*, pages 847–457, Lyon-Villeurbanne France, 1991.
- [46] G. W. MEYER, P. A. WALKER, H. B. WESTLUND, E. MARX, T. V. VORBURGER, et F. Y. HUNT. « A first stop towards photorealistic rendering of coated surfaces and computer based standards of appearance ». *Service Life Prediction Methodology and Metrologies*, 1999.

- [47] S. MEZGHANI et H. ZAHOUANI. « Characterization of the 3d waviness and roughness motifs ». Dans B.-G. Rosén T.R. Thomas L. BLUNT, éditeur, *Transactions of the 9th Int. Conf. on Metrology and Properties of Engineering Surfaces*, Halmstad, Sweden, sep 2003.
- [48] C. A. MICCHELLI. « Interpolation of scattered data : distance matrices and conditionally positive functions ». *Constructive approximation*, 2 :11–22, 1986.
- [49] P. MOERLAND. « Mixtures of latent variable models for density estimation and classification ». Rapport Technique, Dalle Molle Institute for Perceptual Artificial Intelligence, IDIAP, 2000.
- [50] P. MOUTTOU. Analyse des données (Cours). Ecole Des Mines de Paris, 1996.
- [51] R.M. NEAL et G.E. HINTON. « A view of the EM algorithm that justifies incremental, sparse, and other variants ». Department of Computer Science, University of Toronto, Ontario, Canada.
- [52] F. PERNKOPF. « *Automatic visual inspection of metallic surfaces* ». PhD thesis, Institute of automation University of Leoben, 2002.
- [53] D. N. POLITIS et H. WHITE. « Automatic block-length selection for the dependent bootstrap ». *Econometric Reviews*, 23(1) :53–70, 2004.
- [54] E. PRADOS et O. FAUGERAS. « A mathematical and algorithmic study of the lambertian sfs problem for orthographic and pinhole cameras ». Rapport Technique, INRIA, nov 2003.
- [55] W.H. PRESS, S.A. TEUKOLSKY, W.T. VETTERLING, et B.P. FLANNERY. *Numerical recipes in C : The art of scientific computing*. Cambridge university Press, 1992.
- [56] A. ROBLES-KELLY, A.G. BORS, et E.R. HANCOCK. « Surface acquisition from single gray-scale images ». *IEEE*, 2003.
- [57] C. SCOTT et R. NOWAK. « Complexity Regularized Dyadic Classification Trees : Efficient Pruning and Rates of convergence ». Rapport Technique, Rice University, Department of Electrical and Computer Engineering, apr 2002. Technical Report TREE0201.
- [58] J. SERRA. « Les fonctions aléatoires de dilution ». Rapport Technique, Paris School of Mine publication, 1968. N 37, CMM.
- [59] J. SERRA. Cours de Morphologie Mathématique, Première Partie. Ecole des Mines de Paris, 1999. Centre de Morphologie Mathématique.
- [60] J. SERRA. *Image Analysis and Mathematical Morphology, vol. I*. Ac. Press, London, 982.
- [61] J. SERRA. *Image Analysis and Mathematical Morphology, vol. II, theoretical advances*. Ac. Press, London, 988.
- [62] J. TILAHY BEANDRAZANA. « Rapport de stage ingénieur (février-juillet 2003) ». Rapport Technique, ENSMM, Arcelor Research, 2003.
- [63] C. VACHIER. « *Extraction de caractéristiques, segmentation d'image et morphologie mathématique* ». PhD thesis, Ecole des Mines de Paris, 1995.

- [64] T. WALTER. « *Application de la Morphologie mathématique au diagnostic de la Rétinopathie Diabétique à partir d'images couleurs* ». PhD thesis, Ecole des Mines de Paris, 2003.
- [65] G. WARD. « Measuring and modeling anisotropic reflection ». Dans James J. THOMAS, éditeur, *Proceedings of the 19st annual conference on computer graphics and interactive techniques, SIGGRAPH 1992*.
- [66] G. WARD, F. RUBINSTEIN, et Clear R.. « A ray tracing solution to diffuse inter-reflection ». Dans *Proceedings of the 15th annual conference on Computer graphics and interactive techniques, SIGGRAPH 1988*, 1988.
- [67] P.L. WORTHINGTON et E.R. HANCOCK. « New constraints on data-closeness and needle map consistency for shape from shading ». *IEEE transactions on pattern analysis and machine intelligence*, 21(12), 1999.
- [68] H. ZAHOUANI, S. E. POKOSSI, S. MEZGHANI, R. VARGIOLU, H. JACOBS, et J. PIEZANOWSKI. « Characterization of the painted surfaces appearance by continuous wavelet transform ». Dans B.-G. Rosén T.R. Thomas L. BLUNT, éditeur, *Transactions of the 9th Int. Conf. on Metrology and Properties of Engineering Surfaces*, Halmstad, Sweden, sep 2003.
- [69] R. ZHANG, P.S. TSAI, J.E. CRYER, et M. SHAH. « Shape from shading : A survey ». *IEEE Transactions on pattern analysis and machine intelligence*, 21(8) :690–706, 1999.

Résumé

Les propriétés topographiques et physico-chimiques d'une surface conditionnent en grande partie la qualité de son aspect. La maîtrise des phénomènes déterminant l'apparence d'un produit représente aujourd'hui un fort enjeu industriel et scientifique.

Dans cette thèse nous introduisons des descripteurs morphologiques et statistiques permettant de caractériser une texture aussi bien qualitativement que quantitativement.

Ces informations sont exploitées par l'intermédiaire de techniques d'analyse factorielle, de sur-échantillonnage systématique de type "bootstrap" et de classification bayésienne ayant permis la conception d'un algorithme automatique d'estimation de la qualité de l'apparence d'une surface.

Ces développements théoriques ont été appliqués au contrôle qualité des surfaces sur ligne de production grâce à un système de vision permettant d'acquérir des images en continu.

Abstract

Topographical, physical and chemical properties of rough surfaces are strongly involved in their visual aspect. The understanding of the phenomena determining the appearance of a product represents a growing scientific as well as economic challenge.

In this thesis, we introduce morphological and statistical features designed for characterising textures in a qualitative as well as quantitative way.

These data are processed by factorial analysis methods, systematic bootstrap over-sampling and bayesian classification techniques used together to produce an automatic algorithm of surface appearance quality estimation.

These theoretical advances have been applied to an industrial context. A vision system enables us to take pictures continuously during the production process of orange peel surfaces. These images taken on line are then processed to analyse the visual quality of the product.
