HAL
open science

Inférence géométrique discrète

Louis Cuel

To cite this version:

Louis Cuel. Inférence géométrique discrète. Interface homme-machine [cs.HC]. Université de Grenoble, 2014. Français. NNT : 2014GRENM074 . tel-01551790

HAL Id: tel-01551790

https://theses.hal.science/tel-01551790v1

Submitted on 30 Jun 2017

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.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques Appliquées**

Arrêté ministériel : 7 août 2006

Présentée par

Louis Cuel

Thèse dirigée par **Jacques-Olivier Lachaud**
et codirigée par **Boris Thibert**

préparée au sein **Laboratoire de Mathématiques de l'université de Savoie** / **Laboratoire Jean Kuntzmann**
et de **Mathématiques, Sciences et Technologie de l'Information, Informatique**

Inférence Géométrique Discrète

Thèse soutenue publiquement le **18 décembre 2014**,
devant le jury composé de :

M. Simon Masnou
Professeur à l'université Lyon 1, Président
M. Pierre Alliez
Directeur de recherche à l'INRIA Sophia-Antipolis, Rapporteur
M. Rémy Malgouyres
Professeur à l'université d'Auvergne, Rapporteur
M. Jacques-Olivier Lachaud
Professeur à l'université de Savoie, Directeur de thèse
M. Boris Thibert
Maître de conférence à l'université Joseph Fourier, Co-Directeur de thèse

UNIVERSITÉ JOSEPH FOURIER

École Doctorale MSTII
Mathématiques, Sciences et Technologie de
l'Information, Informatique

THÈSE

pour obtenir le grade de **DOCTEUR de l'Université Joseph Fourier**

Spécialité : **Mathématiques Appliquées**

préparée au laboratoire **Laboratoire de Mathématiques de l'université de Savoie / Laboratoire Jean Kuntzmann**

présentée et soutenue publiquement le 18 décembre 2014 par

Louis Cuel

Inférence Géométrique Discrète

Directeur de thèse : **Jacques-Olivier Lachaud**
Co-directeur de thèse : **Boris Thibert**

Jury

M. Pierre Alliez,	Rapporteur
M. Rémy Malgouyres,	Rapporteur
M. Simon Masnou,	Examinateur
M. Jacques-Olivier Lachaud,	Directeur de thèse
M. Boris Thibert,	Co-directeur de thèse

Sommaire

Sommaire iii

INTRODUCTION GÉNÉRALE 1
1 Inférence géométrique 2
2 Tour d'horizon des méthodes d'estimation 3
3 Fonctions distances 5
4 Contributions / Résumé de la thèse 6

I INFÉRENCE GÉOMÉTRIQUE ET FONCTION DISTANCE 11

1 FONCTION DISTANCE 13
1 Fonction distance et ensemble à reach positif 14
2 Formule des tubes 16
2.1 Mesure de Hausdorff 16
2.2 Volume d'un offset 17
3 Stabilité de la projection 20
3.1 Stabilité L_1 20
3.2 Pseudo-continuité de la projection 21

2 ESTIMATION DE NORMALES ET ANALYSE EN COMPOSANTES PRINCIPALES 25
1 Stabilité de l'analyse en composantes principales 26
1.1 Généralités 26
1.2 Stabilité dans $\mathbb{R}^2$ 27
1.3 Démonstration du Théorème de stabilité. 28
2 Application en Géométrie Digitale 33
2.1 Prérequis en Géométrie Digitale 33
2.2 Convergence multigrille de l'estimateur de normale ACP . . 35
3 Expérimentations 38
3.1 Expérimentations sur les nuages de points 38
3.1.1 Validation en 2D 38
3.1.2 Heuristique sur les Surfaces en 3D 39

3.2 Expérimentations en géométrie discrète 40
3.3 Analyse en composantes principales pondérée 43

3 DIGITAL VORONOI COVARIANCE MEASURE 47
1 Voronoi covariance measure 48
1.1 Nuages de points 49
1.2 Surface lisse 49
1.3 Stabilité 49
2 VCM sur des ensembles digitaux 50
2.1 Définition 50
2.2 Convergence multigrille de l'estimateur VCM 50
2.3 Estimation du VCM d'un nuage de points. 51
2.4 Stabilité du VCM. 53
3 Convergence multigrille de l'estimateur de normale 54
4 Expérimentations 57

II DISTANCES ROBUSTES POUR L'INFÉRENCE GÉOMETRIQUE 61

4 FONCTIONS DISTANCE-LIKE 63
1 Généralités 64
1.1 Définition 64
1.2 Stabilité des gradients des fonctions distance-like 65
2 Distance à une mesure 66
2.1 Définition 66
2.2 Stabilité de la distance à une mesure 67
2.3 Distance à une mesure et distance à un compact 67
3 Exemples de fonctions distance-like calculables 68
3.1 Witnessed-k-distance 68
3.2 Median-k-distance 69
3.3 Center-k-distance 70
3.4 Une approximation de la k-distance supportée par P 70

5 VORONOI COVARIANCE MEASURE GÉNÉRALISÉ 71
1 Introduction 71
2 δ-Voronoi covariance measure 72
2.1 Généralités 72
2.2 Stabilité du δ-VCM 74
2.3 démonstration du théorème 5.5 75
3 Exemples de δ-VCM 77

3.1 VCM utilisant la distance à une mesure 77
3.2 VCM utilisant la witnessed k-distance 77
3.3 Autres variantes du δ-VCM 78
4 Implémentation et Expérimentations 78
4.1 Calcul du VCM pour une distance de puissance 78
4.1.1 Intersection 78
4.1.2 Intégration 79
4.2 Évaluation du $d^{\mathrm{w}}_{P,k}$-VCM 79
4.2.1 Estimation des normales 80
4.2.2 Estimation de la courbure et détection des arêtes vives 83
4.3 Comparaison avec le $d^{\mathrm{m}}_{P,k}$-VCM 85
4.4 $d^{\mathrm{w}}_{P,k}$-VCM digital 85

CONCLUSION GÉNÉRALE **89**

NOTATIONS **91**

Bibliographie **93**

Introduction Générale

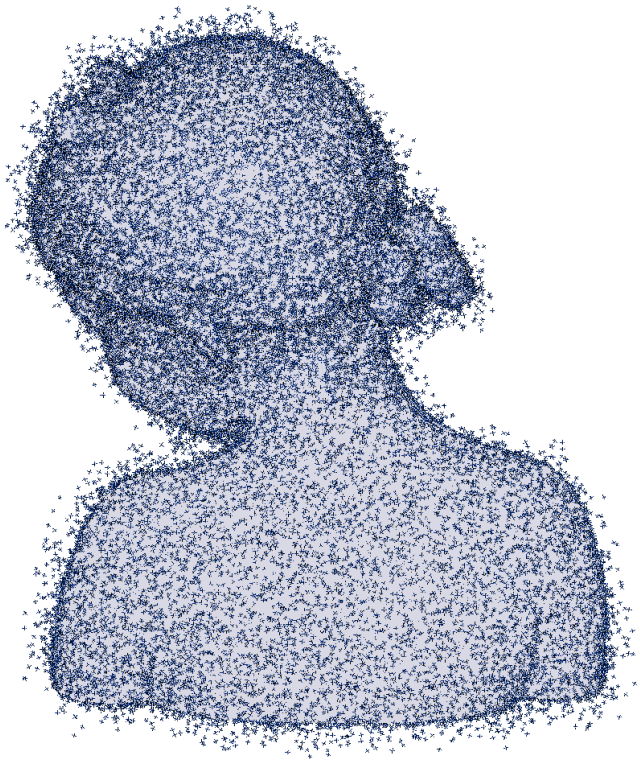


Figure 1 – Un nuage de points approchant une surface

Sommaire

1 **Inférence géométrique** 2

2 **Tour d'horizon des méthodes d'estimation** 3

3 **Fonctions distances** 5

4 **Contributions / Résumé de la thèse** 6

1 INFÉRENCE GÉOMÉTRIQUE

La problématique de l'inférence géométrique consiste à estimer l'information géométrique d'une forme à partir de données discrètes partielles. En pratique, étant donné un objet S comme une surface ou un volume, les méthodes d'acquisition aboutissent à la donnée d'un objet K approchant S. En inférence géométrique, on cherche tout d'abord à définir des quantités géométriques sur les objets K et S comme les courbures ou les mesures de courbure, puis on montre que ces quantités sont proches sous des hypothèses les plus larges possible.

L'estimation de quantités différentielles, la reconstruction de surfaces, et la détection d'arêtes vives sont motivées par un grand nombre d'applications en traitement des images 3D ("geometry processing" et "computer graphics"). De nombreux systèmes d'acquisition 3D donnent en sortie, un nuage de points ou ensemble de voxels [1]. Il est donc crucial de développer des algorithmes d'extraction d'informations géométriques directement sur des nuages de points. Ces algorithmes seront en particulier applicables aux ensembles digitaux et aux triangulations. Il apparaît, depuis un certain temps, que l'inférence géométrique peut être utile dans l'analyse de données telles que celles récoltées dans des forêts [Par06] ou des milieux urbains [CGAY13] et provenant de diverses domaines tels que l'imagerie médicale [LDF04]. Les informations géométriques sont utiles dans divers algorithmes de traitement de ces données dont la visualisation, la segmentation, la détection de symétries ou la reconstruction de surfaces sont les exemples principaux. La reconstruction de surfaces cherche à reconstruire un objet 2-dimensionnel, le plus souvent une triangulation ou une surface implicite, à partir d'un nuage de points (voir [BTS+14] pour un état de l'art récent des méthodes de reconstruction). Plusieurs algorithmes définissent leur reconstruction comme une filtration de la triangulation de Delaunay [CG06]. La plupart de ces méthodes n'ont pas de garanties de retrouver la bonne topologie ou les bonnes normales. Certaines d'entre elles garantissent une bonne reconstruction dans le cadre restreint où le nuage de départ est situé sur la surface sous-jacente [AB99] ou avec un bruit Hausdorff assez faible [DG06]. Les hypothèses mises en jeu dans les précédentes méthodes sont souvent assez fortes et la reconstruction de surfaces à partir de données bruitées est encore peu garantie. Pour palier à ce problème, certaines méthodes de reconstruction de surfaces utilisent en entrée un nuage de points muni d'informations géométriques telles que les normales. La qualité de la reconstruction est alors liée à la qualité des informations fournies. Par exemple, les méthodes de reconstruction de Poisson [KBH06] sont très dépendantes de la qualité des normales données en entrée.

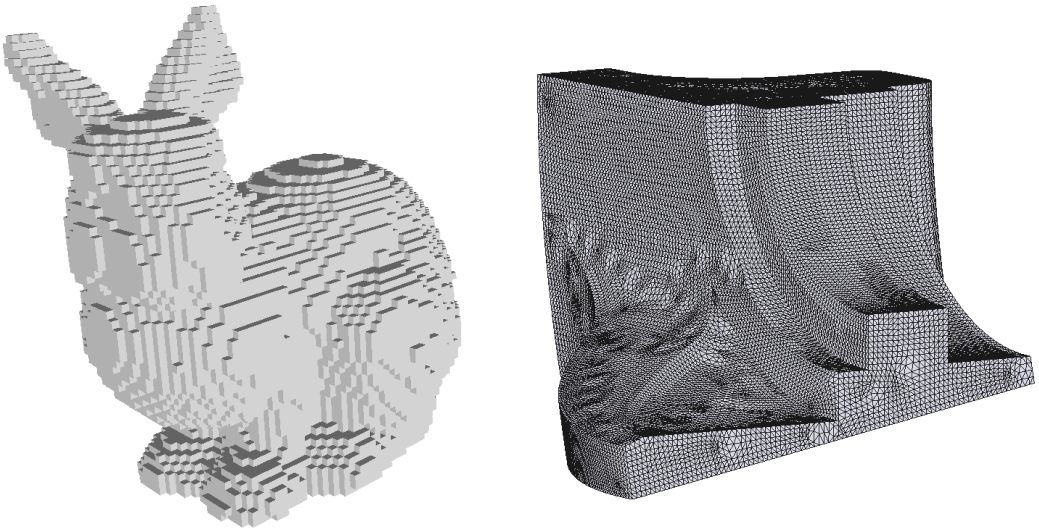
La recherche de garanties en inférence géométrique sur des objets bruités est donc un prérequis à l'obtention de garanties sur de nombreux traitements. Dans notre approche, c'est la stabilité de la fonction distance à un compact qui permet d'obtenir la stabilité d'une méthode d'inférence. La fonction distance à un compact est stable pour la distance de Hausdorff et permet d'obtenir des garanties quand les approximations sont proches au sens de Hausdorff. La distance à une mesure est un outil qui permet d'obtenir une meilleur stabilité, par exemple aux points aberrants (appelés *outliers*). Les travaux relatifs aux garanties obtenues avec la

1. contraction de "volumetric pixels", ce terme désigne les cubes composant une image 3D

fonction distance et la distance à une mesure seront présentés dans la section 3 de l'introduction. Au préalable, je donne un tour d'horizon de diverses méthodes d'estimation n'ayant pas forcément de garanties, mais qui peuvent être intéressantes pour leur précision et leur stabilité expérimentale.

2 Tour d'horizon des méthodes d'estimation

Cette section constitue un survol des méthodes d'estimation géométrique.



a) b)

FIGURE 2 – Différents types de données. a) nuage de points digital. b) triangulation

Estimation des normales et courbures. Un grand nombre de résultats d'inférence géométrique se basent sur des données structurées telles que les contours digitaux [KL12] ou les triangulations [MT04]. Il existe un grand nombre d'algorithmes d'estimation de courbures sur les triangulations [Ham93, KLM98, SW92, Tau95]. On peut se référer à [MSR07] pour une comparaison des méthodes d'estimation des courbures moyennes et Gaussiennes sur les triangulations. En pratique, c'est souvent intéressant d'estimer des quantités directement sur des nuages de points non structurés obtenus par exemple à l'aide d'un scanner. Je présente ici à travers quelques exemples, trois familles de méthodes d'estimations géométriques sur des nuages de points.

La méthode la plus répandue pour estimer les normales et courbures à partir d'un nuage de points consiste à approcher localement la surface sous-jacente par un polynôme. La méthode de l'analyse en composante principale a pour but d'approcher localement la surface par le plan qui minimise l'erreur au sens des moindres carrés. Cela correspond à un polynôme de degré 1. Dans [MNG04], les auteurs montrent la convergence d'un estimateur de normale à une surface basé sur l'analyse en composante principale. Dans [CP05], les auteurs utilisent la

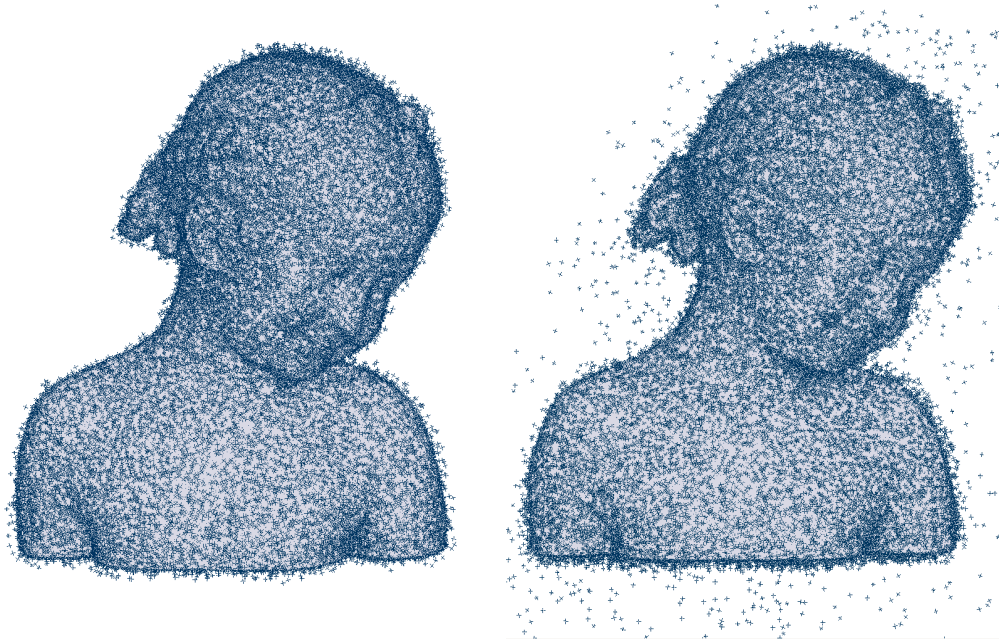
formule de Taylor pour définir des fonctions polynomiales proches du nuage de points. Cette dernière méthode a l'inconvénient de ne pas être garantie théoriquement.

Une deuxième classe de méthodes permet l'estimation de la normale et des courbures à partir d'un nuage de points en reconstruisant une surface lisse proche du nuage à l'aide d'une fonction implicite. La surface implicite est alors définie comme une ligne de niveau d'une fonction réelle. Les méthodes *moving least square* [AK04, PY07] cherchent une fonction qui minimise localement les moindres carrés avec un poids. La méthode étudiée dans [FCOS05] s'intéresse particulièrement à la robustesse aux *outliers* tout en prenant en compte la présence d'arêtes vives. Des garanties topologiques et géométriques d'une méthode *moving least square* sont données dans [Kol08] mais la convergence des normales n'est pas garantie. D'autres méthodes basées sur les surfaces implicites considèrent la surface comme élastique et la déforment à chaque pas de temps pour se rapprocher d'un minimum d'énergie [ZOF01, ZOMK98]. Si la surface donnée en initialisation de l'algorithme est suffisamment proche de la surface sous-jacente, alors cette méthode est assez robuste. Mais on ne peut pas garantir pour ces méthodes d'atteindre un minimum global d'énergie.

Il existe aussi des méthodes statistiques pour estimer des quantités géométriques. Dans [YLL+07], des méthodes statistiques sont utilisées pour améliorer la méthode présentée dans [HDD+92] basée sur l'analyse en composante principale. Plus récemment, Boulch et al. [BM12] a proposé un estimateur statistique des normales basé sur le "Randomized Hough Transform" très robuste au bruit et aux *outliers*. Néanmoins, cette méthode est plus un classifieur de normales qu'un estimateur de normales. En effet, le nuage de points est divisé en un nombre fini de groupes de points sur un critère de proximité des normales. Cette méthode nécessite le choix d'un paramètre de taille qui contrôle le cardinal des groupes.

DÉTECTION DES ARÊTES VIVES. Le problème de détection d'arêtes vives est très proche du problème d'estimation des courbures et comme pour les courbures, la plupart des algorithmes existants pour la détection des arêtes vives prennent une triangulation en entrée. Dans [HPW05], les auteurs utilisent la courbure discrète et dans [WHH10], les auteurs utilisent l'application de Gauss pour detecter les arêtes. Les résultats sont visuellement bons, mais non robustes au bruit Hausdorff et aux *outliers*. Quelques algorithmes utilisent en entrée un nuage de points. Dans [PKG03], les auteurs présentent une méthode basée sur l'ACP qui est intéressante par ses propriétés multi-échelles. Cependant, ces méthodes n'ont pas de garanties puisqu'elles sont définies par une séries de processus souvent complexes. Plus récemment, la méthode d'estimation des normales développée dans [BM12] et présentée dans le précédent paragraphe permet la détection d'arêtes vives. Elle utilise le fait que la distribution statistique des normales dans une zone proche d'une arête est radicalement différente que sur une partie lisse.

INFÉRENCE GÉOMÉTRIQUE EN GÉOMÉTRIE DIGITALE. En géométrie digitale, la façon classique de relier la quantité estimée dans l'espace discret à la quantité euclidienne est la convergence multigrille : quand la forme est discrétisée avec un pas de grille h tendant vers 0, la quantité estimée doit converger vers celle attendue.



a) b)

FIGURE 3 – Différents types de bruits. a) Nuage de points avec un fort bruit Hausdorff. b) Nuage de points avec un fort bruit Hausdorff et des *outliers*

En dimension 2, quelques estimateurs convergents multigrille ont été introduits pour approcher la normale, [dVL08, PG11] et la courbure [PG11, KL09, RL11]. En 3D, des méthodes empiriques pour l'estimation des normales et courbures ont été introduites dans [FM08]. Plus récemment, un estimateur de courbures convergeant basé sur les invariants intégraux a été présenté dans [CLL14].

3 FONCTIONS DISTANCES

DISTANCE À UN COMPACT. Un des principaux problèmes qui se pose en inférence géométrique est la question de l'échelle à laquelle on regarde l'objet. Il apparaît nécessaire de choisir un paramètre qui définit l'échelle à laquelle faire les calculs. L'information à une échelle $R \in \mathbb{R}^+$ est contenue dans la ligne de niveau R de la fonction distance. L'étude du R-voisinage tubulaire de K a ainsi permis de garantir des méthodes d'inférence en topologie et en géométrie. Cette idée a été utilisée pour développer des méthodes d'estimation de la topologie [NSW08], du cône normal [MOG11, CCSL09] et des mesures de courbure [MOG11, CCSLT09]. Dans [NSW08], les auteurs garantissent ainsi qu'il est possible, sous certaines conditions d'échantillonnage, de retrouver l'homologie. La fonction distance et l'offset sont aussi utilisés dans [CCSL09] pour obtenir des garanties sur l'estimation du cône normal dans le cadre des surfaces à μ-*reach* positif. Une condition d'échantillonnage est donnée pour que la reconstruction de l'offset soit isotope à l'objet sous-jacent. Dans [CCSLT09], les auteurs utilisent la fonction distance et la théorie du cycle normal pour prouver la stabilité des mesures de courbure des offsets dans le cas des compacts à μ-*reach* strictement positif. Dans [MOG11], les

auteurs utilisent un estimateur de normale et de courbures défini à l'aide d'une matrice de covariance : le *Voronoi Covariance Measure* (VCM). Le VCM capte l'information géométrique contenue dans les cellules de Voronoi du nuage de points. Les méthodes présentées ci-dessus s'appuient sur la stabilité de la fonction distance. L'une des propriété principale de la fonction distance est sa stabilité au bruit Hausdorff qui peut s'exprimer ainsi : si K' est un compact proche de K, alors les distance $d_{K'}$ et d_K sont proches. La fonction distance a deux autres propriétés importantes de régularité :

- d_K est 1-Lipschitzienne, i.e. $\forall x, y \in \mathbb{R}^d,\ \|d_K(x) - d_K(y)\| \leq \|x - y\|$.
- d_K^2 est 1-semi concave, i.e. $\|.\|^2 - d_K^2$ est convexe.

Ces deux propriétés interviennent directement ou indirectement dans de nombreuses études de stabilité d'estimateurs.

DISTANCE À UNE MESURE. Les méthodes du paragraphe précédent s'appuient sur la stabilité de la fonction distance qui est robuste au bruit Hausdorff. Or la présence d'un seul *outlier* dans le nuage de points fait tomber l'hypothèse Hausdorff et les théorèmes ne s'appliquent plus. Il a été montré dans [CCSM11] que la distance à une mesure possédait, comme la distance à un compact, les deux propriétés de régularité (Lipschitz et 1-semi concave). Les méthodes basées sur la distance à une mesure peuvent en conséquence s'appuyer sur ces propriétés. La distance à une mesure est un outil qui contient de l'information sur la géométrie et qui est stable pour la distance de Wasserstein (en un sens qui sera précisé par la suite), ce qui permet la présence d'*outliers* [CCSM11]. La stabilité de cet outil a récemment été utilisé dans une méthode d'analyse topologique de données [BCOS13].

4 CONTRIBUTIONS / RÉSUMÉ DE LA THÈSE

Nous venons de voir qu'il existait un grand nombre de méthodes d'estimation géométrique sur des approximations de surfaces. Seulement une petite partie d'entre elles sont garanties robustes au bruit Hausdorff et encore moins aux *outliers*. L'objectif principal de cette thèse est d'obtenir des garanties théoriques sur la stabilité de nos estimateurs au bruit Hausdorff (et éventuellement aux *outliers*).

Un autre objectif de ce travail est d'établir les liens entre les méthodes d'estimation de normales et de courbures en géométrie algorithmique et celles utilisées en géométrie digitale. Comme on l'a vu dans la section 2 de l'introduction, de nombreuses méthodes d'inférence en géométrie digitale utilisent la géométrie spécifique des ensembles de pixels et voxels. On peut appliquer les algorithmes définis sur des nuages de points quelconques à des ensembles de voxels. Ainsi, les différentes méthodes présentées dans cette thèse font l'objet d'une application en géométrie digitale à travers des expérimentations mais aussi des résultats de convergence multigrille. La difficulté principale du passage d'un résultat de stabilité Hausdorff sur des nuages de points à un résultat de convergence en géométrie digitale est de déterminer les valeurs des paramètres pour lesquelles on peut établir la convergence multigrille. Enfin, nous attachons une importance particulière au fait que les méthodes développées dans cette thèse soient calculables efficacement. Le travail de programmation lié à l'implémentation des méthodes

étudiées dans cette thèse représente une grande partie du travail fourni pour cette thèse.

Première Partie : Inférence géométrique et fonction distance

Les chapitres 1, 2 et 3 constituent une première partie s'intéressant à des méthodes d'inférence utilisant la fonction distance.

CH. I : FONCTION DISTANCE

Nous rappelons, dans ce premier chapitre, quelques propriétés de la fonction distance à un compact. La première section rappelle des propriétés de la fonction distance et des propriétés des compacts à *reach* positif. Dans la deuxième section, nous rappelons la formule des tubes. La section 3 s'intéresse à la stabilité de la projection à un compact. Dans le cas où K et K' sont deux compacts proches au sens de Hausdorff, nous rappelons un résultat de stabilité L_1 des projections sur K et K' (Théorème 1.12). Nous démontrons enfin un résultat qui permet de contrôler le comportement de la projection sur un compact K quand celui-ci approche une hyper-surface lisse S de $\mathbb{R}^d$ (théorème 1.14).

CH. II : ESTIMATION DE NORMALES ET ACP

L'analyse en composante principale (ACP) est un outil classique d'analyse de données. Nous montrons dans ce chapitre un résultat de stabilité d'estimateur de normale à l'aide de l'ACP (théorème 2.3) qui donne une borne explicite de l'erreur. Ce résultat quantifie un résultat de convergence montré dans [MNG04] et implique directement le corollaire suivant qui établie qu'en choisissant convenablement le paramètre d'échelle r et sous certaines conditions d'échantillonnage, on a la convergence de l'estimateur de normale.

Corollaire 2.4 . *Soit C une courbe de classe $\mathcal{C}^2$ de reach $\rho > 0$ et p un point de C. Soit $\varepsilon \in \mathbb{R}^{+*}$, $a \in \mathbb{R}^{+*}$ et $r = a\varepsilon^{\frac{1}{2}}$. Si $P \subset \mathrm{B}_p(r)$ est un (ε, m)-échantillon de $C \cap B_p(r)$ tel que $m \leq \frac{|P|}{18}$, alors l'angle β entre la normale à C en p et la normale au nuage P au sens des moindres carrés satisfait*

$$\beta = O(m^2\varepsilon^{\frac{1}{2}}),$$

où la constante intervenant dans le O ne dépend que de ρ et de a.

Appliqué à des ensembles digitaux, ce résultat permet d'obtenir une convergence multigrille de l'estimateur (théorème 2.13). Des expérimentations sur des nuages de points et des ensembles digitaux viennent confirmer la stabilité en présence de bruit Hausdorff (section 3). Le dernier paragraphe propose une heuristique en ajoutant des poids aux points pour obtenir plus de robustesse au bruit.

CH. III : DIGITAL VORONOI COVARIANCE MEASURE

Alors que l'analyse en composante principale cherche à estimer un plan tangent approchant la surface, le *Voronoi covariance measure* (VCM) tente d'approcher le cône normal. Dans ce chapitre nous appliquons la méthode du VCM aux ensembles de voxels dans $\mathbb{R}^3$ pour établir un résultat de stabilité d'un estimateur de normale à l'aide du VCM (théorème 3.6) qui donne une borne explicite de l'erreur. Ce résultat implique directement le corollaire suivant qui donne la convergence multigrille de cet estimateur.

Corollaire 3.7 . *Soit X un domaine compact de $\mathbb{R}^3$ dont le bord ∂X est une surface de classe $\mathcal{C}^2$ de reach $\rho > 0$. Soient $a, b \in \mathbb{R}^+$, $r = ah^{\frac{1}{4}}$ et $R = bh^{\frac{1}{4}}$. Alors pour $h > 0$, on a*

$$\langle \widehat{n}_{r,R}(\widehat{p_0}), \mathbf{n}_S(p_0) \rangle = O\Big(h^{\frac{1}{8}}\Big),$$

où $\mathbf{n}_S(p_0)$ est la normale à S en p_0 et $\widehat{n}_{r,R}(\widehat{p_0})$ est son approximation à l'aide de l'ACP.

Les expérimentations permettent de confirmer la précision et la robustesse de cette méthode. Le VCM contient, en outre, des informations sur les courbures. Les tests numériques montrent que ces informations peuvent être utilisées pour détecter les arêtes vives. D'autre part, nous automatisons le choix des paramètres en choisissant ceux qui minimisent l'erreur garantie par notre résultat de stabilité multigrille.

Deuxième Partie : Distances robustes pour l'inférence géométrique

Les chapitres 4 et 5 forment une deuxième partie consacrée à l'inférence géométrique utilisant des fonctions distance robustes aux *outliers*, afin de garantir une estimation robuste aux *outliers*. Cette partie constitue la contribution principale de cette thèse.

CH. IV : FONCTIONS DISTANCE-LIKE

La majorité des méthodes d'inférence géométrique se basent, directement ou indirectement, sur la stabilité de la fonction distance. Celle-ci est stable au bruit Hausdorff mais pas aux *outliers*. Ce chapitre a pour objet de rappeler les propriétés des fonctions *distance-like*. On montre la stabilité des gradients des fonctions *distance-like* (Proposition 4.4) qui est une conséquence directe de la stabilité L_1 des gradients des fonctions convexes. On propose enfin des exemples de fonctions *distance-like* robustes et calculables efficacement.

CH. V : VORONOI COVARIANCE MEASURE GÉNÉRALISÉ

On développe ici un estimateur de quantités géométriques robuste aux *outliers* à l'aide des fonctions *distance-like*. On généralise le *Voronoi covariance measure* aux fonctions *distance-like*. Pour chaque fonction distance like δ, le δ-VCM associe à toute fonction mesurable χ une matrice $\mathcal{V}_{\delta,R}(\chi)$. On montre le résultat suivant qui établie la stabilité de notre estimateur.

Théorème 5.5 . *Soit K un compact et δ une fonction distance-like. Pour toute fonction bornée et Lipschitzienne $\chi : \mathbb{R}^d \longrightarrow \mathbb{R}^+$, on a*

$$\|\mathcal{V}_{\delta,R}(\chi) - \mathcal{V}_{d_K,R}(\chi)\|_{\mathrm{op}} \leq C_1 \ \|\chi\|_{\mathrm{BL}} \, \|\delta - d_K\|_\infty^{\frac{1}{2}} \, ,$$

où C_1 est une constante qui dépend seulement de R, d et $\mathrm{diam}(K)$.

Ce résultat assure que si une fonction distance-like δ est proche de la fonction distance à un compact d_K, alors le δ-VCM $\mathcal{V}_{\delta,R}$ est proche du VCM $\mathcal{V}_{d_K,R}$. Appliqué à la distance à une mesure, ce résultat implique que le VCM généralisé permet une estimation robuste aux *outliers* de la normale à une hyper-surface. Des tests numériques sur des nuages de point et sur des ensembles de voxels confirment en pratique la précision et la robustesse au bruit Hausdorff et aux *outliers* des estimations de normales, courbures, directions de courbures (figure 4) et arêtes vives.

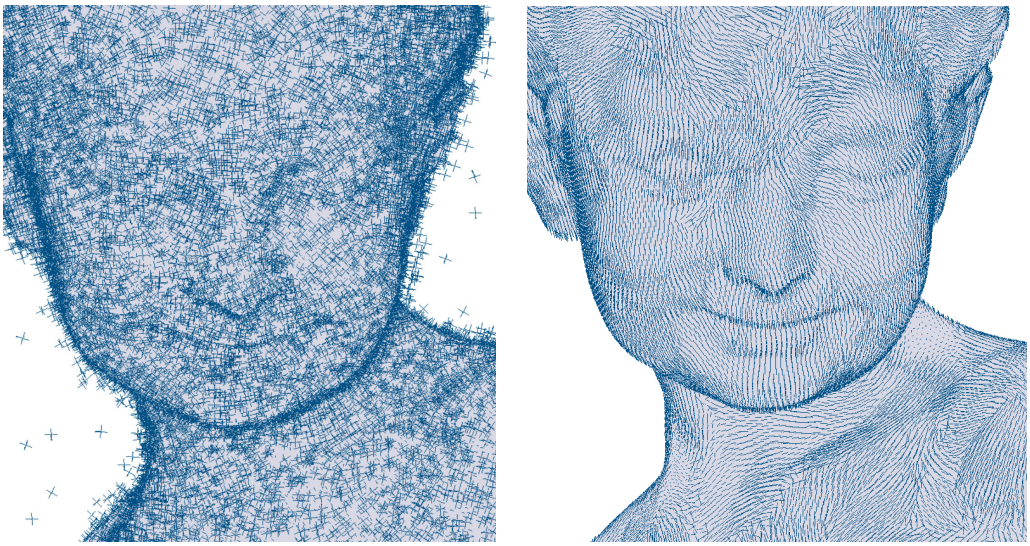


FIGURE 4 – Estimation des directions principales sur “Bimba”. Gauche : Données en entrée avec *outliers* (80% des points sont translatés d'une distance au plus 0.02, 10% sont translatés d'une distance entre 0.02 et 0.1, et 10% sont des *outliers* pris uniformément dans une boîte englobante). Droite : pour chaque point, on projette l'estimation de la direction principale minimale du VCM généralisé sur la triangulation initiale.

Publications :

[1] L. Cuel, J.O. Lachaud, B. Thibert. Voronoi-based geometry estimator for 3D digital surfaces, *Discrete Geometry for Computer Imagery*, 2014.

[2] L. Cuel, J.O. Lachaud, Q. Mérigot, B. Thibert. Robust Geometry Estimation using the Generalized Voronoi Covariance Measure, *SIAM Journal on Imaging Sciences (SIIMS), (soumis)*.

Communications :

[3] L. Cuel, J.O. Lachaud, Q. Mérigot, B. Thibert. Robust Inference with a Generalised Voronoi Covariance measure, *euroCG*, (2014) Israel.

Première partie

Inférence Géométrique et Fonction Distance

Chapitre 1

Fonction Distance

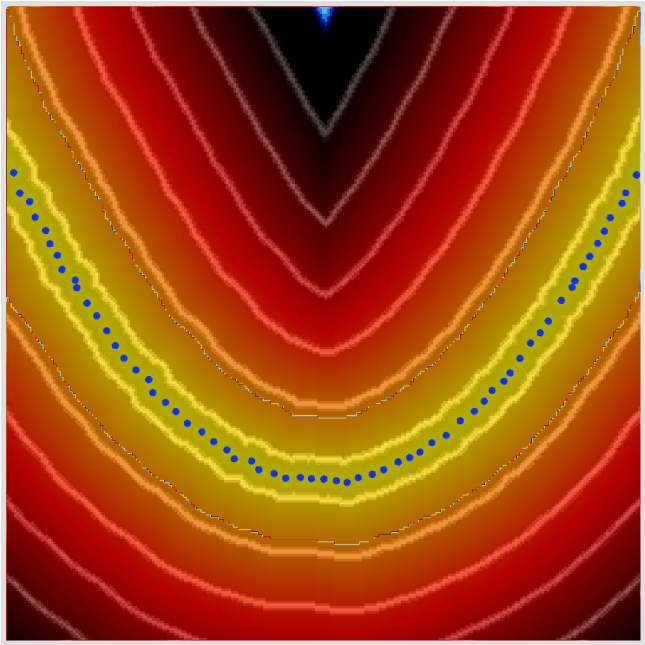


FIGURE 1.1 – Lignes de niveau de la fonction distance à un nuage de point

Sommaire

1 Fonction distance et ensemble à reach positif 14

2 Formule des tubes 16

2.1 Mesure de Hausdorff 16

2.2 Volume d'un offset 17

3 Stabilité de la projection 20

3.1 Stabilité L_1 20

3.2 Pseudo-continuité de la projection 21

INTRODUCTION

Comme nous l'avons rappelé dans le chapitre précédent, la fonction distance à un compact K contient des informations géométriques sur le compact K. Pour établir un résultat de stabilité sur une méthode d'inférence géométrique basée sur la fonction distance, il peut être utile de connaître les propriétés de la fonction distance et en particulier sa stabilité et celle de son gradient. C'est en cela que ce chapitre constitue une introduction aux deux chapitres suivants qui décrivent deux résultats d'inférence géométrique basés sur la fonction distance. Les propriétés de la fonction distance ont largement été étudiées [Fed59, Gro93]. On s'intéresse à des résultats de stabilité de la fonction distance car ils permettent d'obtenir des résultats de stabilité d'inférence géométrique [CCSLT09, MOG11, CCSL09].

La section 1 de ce chapitre rappelle les résultats de régularité de la fonction distance à un compact et les propriétés des compacts à *reach* positif. Nous rappelons ensuite dans la section 2 la formule des tubes dans le cadre des hyper-surfaces à *reach* positifs. Cette formule sera utilisée dans le chapitre 3, pour la preuve de la stabilité d'un estimateur de normale. La section 3 de ce chapitre étudie la stabilité de la projection. Nous rappelons tout d'abord le résultat de stabilité L_1 de la projection [CCSM10]. Mon apport dans cette section est un résultat qui permet de contrôler le comportement de la projection sur un compact K quand celui-ci approche une hyper-surface lisse S (Théorème 1.14). Ces deux résultats seront utiles dans le chapitre 3.

1 FONCTION DISTANCE ET ENSEMBLE À REACH POSITIF

La fonction distance à un compact K associe à tout point $x \in \mathbb{R}^d$ sa distance à K :

$$d_K(x) := \min \{\|x - y\|,\ y \in K\}$$

où $\|.\|$ est la norme euclidienne de $\mathbb{R}^d$. La distance au compact K caractérise entièrement le compact K puisque $K = \{x \in \mathbb{R}^d,\ d_K(x) = 0\}$. Elle contient donc toute l'information sur la géométrie du compact K. L'objectif de cette section est d'étudier les propriétés de la fonction distance, fortement liées à la notion de *reach*. On rappelle la proposition suivante qui établit que la distance est Lipschitzienne et qui est une conséquence directe de l'inégalité triangulaire.

Proposition 1.1. *Pour tout compact K, la distance d_K est 1-Lipschitzienne. i.e.*

$$\forall x, y \in \mathbb{R}^d,\ |d_K(x) - d_K(y)| \leq \|x - y\|.$$

En particulier d_K est continue.

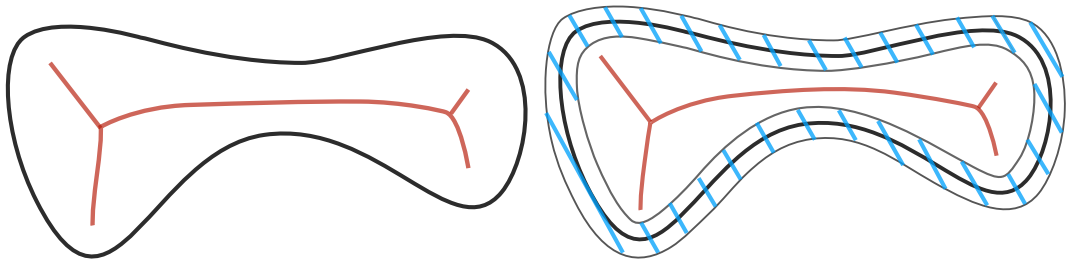


FIGURE 1.2 – Axe médian et offset.

AXE MÉDIAN. L'axe médian est introduit par Blum en 1967 [Blu67] et a été largement étudié car il contient des informations géométriques. Le survey [ABE09] fait le point sur la stabilité et le calcul de l'axe médian. Ce dernier est défini comme l'ensemble des points de $\mathbb{R}^d$ ayant plusieurs projections possibles sur K.

Définition 1.2. *(axe médian)*
L'axe médian d'un compact $K \subset \mathbb{R}^d$ est défini par

$$\mathcal{AM}(K) = \{x \in \mathbb{R}^d,\ \exists y, z \in K,\ tq\ y \neq z\ et\ \|x-y\| = \|x-z\| = d_K(x)\}.$$

La figure 1.3 donne l'exemple d'un axe médian dans le cas d'une courbe fermée de $\mathbb{R}^2$. On note $\mathfrak{p}_K$ la projection sur le compact K, c'est à dire l'application qui à un point $x \in \mathbb{R}^d$ associe son plus proche voisin dans K. L'axe médian correspond au lieu géométrique où la projection n'est pas définie, mais aussi à l'ensemble sur lequel la fonction distance n'est pas différentiable [Fed59]. Nous rappelons le résultat suivant qui donne l'expression du gradient de la fonction distance (Théorème 4.8 de [Fed59]).

Proposition 1.3. *Pour tout compact K, la distance d_K est de classe $\mathcal{C}^1$ sur $\mathbb{R}^d \backslash (K \cup \mathcal{AM}(K))$ et son gradient vaut*

$$\forall x \in \mathbb{R}^d \backslash (K \cup \mathcal{AM}(K)),\ \nabla d_K(x) = \frac{x - \mathfrak{p}_K(x)}{d_K(x)}.$$

La fonction distance étant lipschitzienne, le théorème de Rademacher implique qu'elle est de classe $\mathcal{C}^1$ presque partout au sens de Lebesgue. Cela implique en particulier que l'axe médian est de mesure de Lebesgue nulle.

ENSEMBLE À REACH POSITIF. La notion de *reach* permet de définir un voisinage dans lequel la projection est bien définie.

Définition 1.4 (*reach*)**.**
Le reach d'un compact K est la distance entre K et son axe médian $\mathcal{AM}(K)$, c'est à dire

$$\inf\{\|x-y\|,\ x \in K,\ y \in \mathcal{AM}(K)\}.$$

Le *reach* correspond à la distance critique à laquelle on peut s'éloigner de K en étant sûr de n'avoir qu'un seul plus proche point dans K. Tout au long de cette thèse nous considèrerons des objets lisses à *reach* strictement positif et nous

nous intéresserons au voisinage tubulaire (voir figure 1.3) d'une taille plus petite que le *reach* dans lequel la projection est bien défini. Formellement, on définit le voisinage tubulaire de K de paramètre t (ou t-offset) par : $K^t = \{x \in \mathbb{R}^d,\ d_K(x) \leq t\}$. La proposition suivante donne une propriété importante de la fonction distance démontrée par Federer [Fed59], dont on se servira tout au long de cette thèse.

Proposition 1.5. *Soient K un compact de reach $\rho > 0$ et r un réel tel que $r < \rho$. Alors $\mathfrak{p}_K$ est $\frac{\rho}{\rho - r}$-lipschitzienne sur K^r :*

$$\|\mathfrak{p}_K(x) - \mathfrak{p}_K(y)\| \leq \frac{\rho}{\rho - r}\|x - y\|.$$

En particulier, si t est un réel tel que $0 < t < r$, d_K est de classe $\mathcal{C}^{1,1}$ sur $K^r \backslash K^t$, i.e. la différentielle de d_K est lipschitzienne sur $K^r \backslash K^t$.

HYPER-SURFACE. Soit $S \subset \mathbb{R}^d$ une variété plongée orientée de classe $\mathcal{C}^k$ et de dimension $d - 1$. On peut alors trouver pour chaque point $p \in S$ un voisinage dans S qui est l'image d'une fonction $\psi_p : \mathcal{O} \subset \mathbb{R}^{d-1} \to \mathbb{R}^d$ de classe $\mathcal{C}^k$ dont la différentielle est de rang maximal et telle que $\psi_p(\mathbf{0}) = p$. On dit qu'une hyper-surface est de classe $\mathcal{C}^{k,1}$ si les ψ_p peuvent être choisis de classe $\mathcal{C}^{k,1}$, c'est à dire que la différentielle d'ordre k de d_K est lipschitzienne. On note T_pS le plan tangent à S en p et $\mathbf{n}_S(p)$ la normale unitaire orientée à S en p. L'application $-D\mathbf{n}_S(p)$ est un endomorphisme symétrique de l'espace tangent T_pS, et est en conséquence diagonalisable. Les courbures principales de S en p sont notées λ_1, λ_2, ..., λ_{d-1}, et sont définies comme les $d - 1$ valeurs propres de $-D\mathbf{n}_S(p)$. On appelle directions principales les vecteurs propres correspondant e_1, e_2, ..., e_{d-1}. Alors que la normale est une information d'ordre 1, les courbures principales sont liées à la différentielle d'ordre 2. Federer donne dans [Fed59] la caractérisation suivante des hypersurfaces à *reach* positif :

Proposition 1.6. *Une hyper-surface de classe $\mathcal{C}^1$ est de classe $\mathcal{C}^{1,1}$ si et seulement si elle est à reach strictement positif.*

Le théorème de Rademacher implique qu'une hyper-surface de classe $\mathcal{C}^1$ à *reach* positif est de classe $\mathcal{C}^2$ presque partout. Les courbures d'une hyper-surface à *reach* $\rho > 0$ sont alors définies presque partout. Notons que, en les points où c'est défini, les valeurs absolues des courbures principales sont majorées par $\frac{1}{\rho}$.

2 FORMULE DES TUBES

Nous rappelons ici la formule des tubes qui donne une expression du volume d'un voisinage tubulaire d'un objet. On déduira ainsi la proposition 1.10 qui sera utile dans le chapitre 3. Nous rappelons avant cela quelques définitions utiles pour la suite.

2.1 MESURE DE HAUSDORFF

Nous définissons ici la mesure de Hausdorff qui est une généralisation des notions d'aire et de volume. Pour tout ensemble $A \subset \mathbb{R}^d$, on note $\mathrm{diam}(A)$ son

diamètre, c'est à dire la quantité

$$\operatorname{diam}(A) := \sup \{\|x-y\|,\ x,y \in A\}.$$

Définition 1.7. *(mesure de Hausdorff) Soit $K \subset \mathbb{R}^d$ et $k \in \mathbb{R}^+$, on définit la k-mesure de Hausdorff de K par*

$$\mathcal{H}^k(K) := \lim_{\varepsilon \to 0} \inf \left\{ \sum_{i=1}^{\infty} \alpha(k) r(A_i)^k,\ K \subset \bigcup_{i \in \mathbb{N}} A_i,\ \operatorname{diam}(A_i) \leq \varepsilon \right\} \tag{1.1}$$

où les A_i sont des parties de $\mathbb{R}^d$, $r(A_i) = \operatorname{diam}(A_i)/2$ et $\alpha(k)$ est le volume de la boule unité de dimension k.

L'existence de la limite est assurée par le fait que $\inf \{ \sum_{i=1}^{\infty} \alpha(k) r(A_i)^k,\ K \subset \bigcup_i A_i,\ \operatorname{diam}(A_i) \leq \varepsilon\}$ est une fonction décroissante de ε [Mor08]. La terminologie "mesure de Hausdorff" est justifiée car on peut montrer que c'est une mesure de Borel. Pour comprendre intuitivement cette mesure, plaçons-nous dans le cas d'une courbe C plongée dans $\mathbb{R}^2$. Sa "longueur", qui correspond à sa 1-mesure de Hausdorff, est définie à l'aide d'une suite de recouvrements de plus en plus fins auquel correspond une quantité fini $\sum_i \alpha(1) r(A_i)$. Au contraire un objet "plein" de $\mathbb{R}^2$ tel que la boule unité, a une 1-mesure de Hausdorff infinie.

ENSEMBLES RECIFIABLES.

Définition 1.8. *Un ensemble $A \subset \mathbb{R}^d$ est dit m rectifiable si $\mathcal{H}^m(A) < \infty$ et s'il existe une famille dénombrable de fonctions lipschitziennes $f_i : \mathbb{R}^m \to \mathbb{R}^d$ telle que $A \subset \bigcup_i f_i(\mathbb{R}^m)$ $\mathcal{H}^m$-presque partout.*

On peut noter que les ensembles rectifiables admettent un espace tangent $\mathcal{H}^m$-presque partout ce qui permet en particulier d'intégrer sur ces ensemble [Mor08].

k-JACOBIEN. Soien $V \subset \mathbb{R}^d$ et $W \subset \mathbb{R}^{d'}$ deux ensembles rectifiables. Pour toute fonction lipschitzienne $f : V \to W$, le k-Jacobien de f peut être défini comme la limite

$$J_k f(p) = \lim_{r \to 0} \frac{\mathcal{H}^k\left(f(\mathrm{B}_p(r) \cap V)\right)}{\mathcal{H}^k\left(\mathrm{B}_p(r) \cap V\right)},$$

où $\mathrm{B}_p(r)$ est la boule de $\mathbb{R}^d$ de centre p et de rayon r. Si V et W sont deux variétés de même dimension m, alors pour $\mathcal{H}^m$-presque tout p de V, la k-Jacobienne n'est rien d'autre que le déterminant de la différentielle de f en p qui est l'application linéaire $D_f(p) : T_p V \to T_{f(p)} W$.

2.2 VOLUME D'UN OFFSET

Nous nous intéressons ici au volume des offsets des hyper-surfaces de $\mathbb{R}^d$ de classe $\mathcal{C}^1$ à *reach* positif. La formule des tubes établit que ce volume est un polynôme en le paramètre d'offset. Elle a été donnée dans le cas des ensembles convexes par Steiner [Ste43], et dans le cas lisse dans [Wey39]. Federer a généralisé ces deux formules dans le cadre des ensembles à *reach* positif [Fed59]. On rappelle ici la formule des tubes dans le cas des hyper-surfaces à *reach* positif.

Lemme 1.9 (Formule des tubes)**.** *Soit $S \subset \mathbb{R}^d$ une hyper-surface de classe $\mathcal{C}^{1,1}$ de reach $\rho > 0$. Soient $R > 0$ et B un borélien de $\mathbb{R}^d$, on a :*

$$\mathcal{H}^d(S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)) = \int_{\mathrm{B}\cap S} \int_{-R}^{R} \Pi_{i=1}^{d-1}(1 - t\lambda_i(x))dtdx,$$

où les $\lambda_i(x)$ sont les courbures principales de S.

Démonstration. La démonstration utilise la formule générale de l'aire-coaire. Soit f la fonction définie par

$$\begin{array}{rccl} f: & S\times]-R,R[& \longrightarrow & S^R \\ & (x,t) & \longmapsto & x + t.\mathbf{n}_S(x) \end{array}$$

Comme S est une hyper-surface de classe $\mathcal{C}^{1,1}$, la fonction f est Lipschitzienne. On peut donc appliquer la formule générale de l'aire-coaire (formule 3.13 de [Mor08]) qui nous donne :

$$\int_{S^R} \mathrm{d}\mathcal{H}^d(x) = \int_{S\times]-R,R[} J_d(f)\mathrm{d}\mathcal{H}^d(x).$$

D'après le théorème de Rademacher, $\mathbf{n}_S$ est différentiable presque partout, et on a pour presque tout x, $J_d(f)(x) = \Pi_{i=1}^{d-1}(1 - t\lambda_i(x))$, ce qui permet de conclure. □

On en déduit la proposition suivante qui sera utile au chapitre 3.

Proposition 1.10. *Soit $S \subset \mathbb{R}^d$ une $d-1$-variété de classe $\mathcal{C}^{1,1}$ et de reach $\rho > 0$. Soit $R > 0$ et $\varepsilon > 0$ tels que $R + \varepsilon < \frac{\rho}{2}$. Alors pour tout borélien B tel que $\partial(B \cap S)$ est une variété de dimension $d-2$, on a :*

1) $\mathcal{H}^d\left(S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)\right) \leq 2\left(\frac{3}{2}\right)^{d-1} R\, \mathcal{H}^{d-1}(\mathrm{B} \cap S)$

2) $\mathcal{H}^d\left((S^{R+\varepsilon} \backslash S^{R-\varepsilon}) \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)\right) \leq \frac{2}{\rho}\left(\frac{3}{2}\right)^{d-1} \mathcal{H}^d(\mathrm{B} \cap S)\varepsilon$

3) $\mathcal{H}^{d-1}\left(\partial[S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)]\right) \leq \mathcal{H}^{d-2}(\mathcal{C})R\left(\frac{3}{2}\right)^{d-2} + 2\left(\frac{3}{2}\right)^{d-1} \mathcal{H}^{d-1}(\mathrm{B} \cap S)$

En particulier, si B est une boule de rayon r, on a :

a) $\mathcal{H}^d\left(S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)\right) = O(Rr^{d-1})$.

b) $\mathcal{H}^d\left((S^{R+\varepsilon} \backslash S^{R-\varepsilon}) \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)\right) = O(\varepsilon r^{d-1})$.

c) $\mathcal{H}^{d-1}\left(\partial[S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)]\right) = O(Rr^{d-2} + r^{d-1})$,

où la notation O fait intervenir une constante qui ne dépend que du reach ρ.

Démonstration. Le lemme 1 donne :

$$\mathcal{H}^d(S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)) = \int_{\mathrm{B}\cap S} \int_{-R}^{R} \Pi_{i=1}^{d-1}(1 - t\lambda_i(x))dtdx.$$

On peut majorer tous les $|\lambda_i(x)|$ par $\frac{1}{\rho}$, on a donc

$$\mathcal{H}^d(S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)) \leq \int_{\mathrm{B}\cap S} dx \times \int_{-R}^{R} \left(1 + \frac{t}{\rho}\right)^{d-1} dt \leq \mathcal{H}^{d-1}(\mathrm{B} \cap S) \times 2R\left(\frac{3}{2}\right)^{d-1}.$$

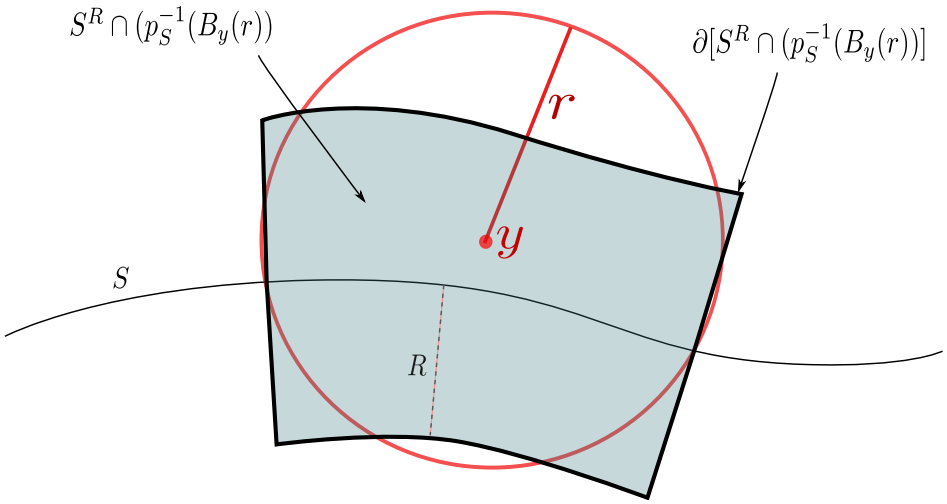


FIGURE 1.3 – Volume d'un offset.

La preuve du deuxième point est similaire. Le volume de $(S^{R+\varepsilon} \backslash S^{R-\varepsilon}) \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)$ est plus petit que

$$\int_{\mathrm{B}\cap S} dx \times \left[\int_{R-\varepsilon}^{R+\varepsilon} \left(1 + \frac{t}{\rho}\right)^{d-1} dt + \int_{-R-\varepsilon}^{-R+\varepsilon} \left(1 - \frac{t}{\rho}\right)^{d-1} dt \right].$$

Il suffit alors d'utiliser l'égalité suivante :

$$\int_{R-\varepsilon}^{R+\varepsilon} \left(1 + \frac{t}{\rho}\right)^{d-1} dt = \frac{1}{d}\left(1 + \frac{R+\varepsilon}{\rho}\right)^d - \frac{1}{d}\left(1 + \frac{R-\varepsilon}{\rho}\right)^d.$$

En factorisant par $\frac{1}{d}\left[\left(1 + \frac{R+\varepsilon}{\rho}\right) - \left(1 + \frac{R-\varepsilon}{\rho}\right)\right]$, on conclut à l'aide de la majoration suivante

$$\frac{1}{d}\left[\left(1 + \frac{R+\varepsilon}{\rho}\right) - \left(1 + \frac{R-\varepsilon}{\rho}\right)\right] \times \sum_{k=0}^{d-1} \left(1 + \frac{R+\varepsilon}{\rho}\right)^k \left(1 + \frac{R-\varepsilon}{\rho}\right)^{d-1-k} \leq \frac{2\varepsilon}{\rho}\left(\frac{3}{2}\right)^d.$$

Pour le point 3), on a

$$\partial[S^R \cap \mathfrak{p}_S^{-1}(\mathrm{B} \cap S)] = A_1 \cup A_2,$$

où

$$\begin{aligned} A_1 &= \{x \in \mathbb{R}^3,\ \|x - \mathfrak{p}_S(x)\| = R \text{ et } \mathfrak{p}_S(x) \in \mathrm{B} \cap S\} \\ A_2 &= S^R \cap \mathfrak{p}_S^{-1}(\partial(\mathrm{B} \cap S)) \end{aligned}$$

Et similairement au premier point, on a

$$\begin{aligned} \mathcal{H}^{d-1}(A_1) &= \textstyle\int_{\mathrm{B}\cap S} \Pi_{i=1}^{d-1}(1 - \lambda_i(x)R)\mathrm{d}x \\ &\leq 2\mathcal{H}^{d-1}(\mathrm{B} \cap S)\left(1 + \tfrac{R}{\rho}\right)^{d-1} \\ &\leq 2\left(\tfrac{3}{2}\right)^{d-1}\mathcal{H}^{d-1}(\mathrm{B} \cap S). \end{aligned}$$

On note $\mathcal{C} = \partial[\mathrm{B} \cap S]$. Comme pour le lemme 1.9, l'ensemble A_2 a pour paramétrisation la fonction

$$\begin{array}{rccl} f: & \mathcal{C} \times (-R, R) & \longrightarrow & A_2 \\ & (x,t) & \longmapsto & x + t.\mathbf{n}_S(x) \end{array}$$

Par hypothèse, $\mathcal{C}$ admet en tout point un plan tangent pour lequel on peut définir une base orthonormée. Le 2-jacobien de f en (x,t) est donnée par $\Pi_{i=1}^{d-2}(1 - t\lambda_{i,C}(x))$, où les $\lambda_{i,C}(x)$ sont les courbures de S en x dans les directions de la base orthonormée de l'espace tangent à $\mathcal{C}$ en x. On a

$$\begin{aligned} \mathcal{H}^{d-1}(A_2) = \int_{\mathcal{C}} \int_{-R}^{R} |\Pi_{i=1}^{d-2}(1 - t\lambda_{i,C}(x))| \mathrm{d}t \mathrm{d}x &\leq \mathcal{H}^{d-2}(\mathcal{C}) \int_{-R}^{R} \left(1 + \frac{1}{2}\right)^{d-2} \mathrm{d}t \\ &\leq \mathcal{H}^{d-2}(\mathcal{C}) R \left(\frac{3}{2}\right)^{d-2}. \end{aligned}$$

La preuve des points a, b et c suivent en remarquant que si B est une boule de rayon r, alors $\mathcal{H}^{d-1}(\mathrm{B} \cap S) = O(r^{d-1})$ et $\mathcal{H}^{d-2}(\mathcal{C}) = O(r^{d-2})$. □

3 STABILITÉ DE LA PROJECTION

Sur les nuages de point ou sur les triangulations, il n'existe pas de définition standard de normales et courbures comme sur les surfaces lisses, mais une grande diversité d'estimateurs. Nous nous intéresserons, dans les chapitres suivants, aux garanties théoriques sur la qualité des estimateurs ainsi qu'à leur résistance aux perturbations. Ces garanties sont tout naturellement liées à la stabilité de la fonction distance et de son gradient.

Nous rappelons d'abord un résultat de stabilité en norme L_1 de la projection sur un nuage de points proche d'une hyper-surface à *reach* positif (section 3.1). Ce résultat sera utilisé dans le chapitre 3 et généralisé dans le chapitre 4. La contribution de cette section est le Théorème 1.14 qui permet de contrôler le comportement de la projection sur un compact approchant une hyper-surface. Ce dernier résultat sera utilisé dans le chapitre 3.

3.1 STABILITÉ L_1

Définition 1.11. *Soient K et K' deux compacts, on appelle distance de Hausdorff entre K et K' le réel positif :*

$$d_H(K, K') = \inf\{\varepsilon \in \mathbb{R}^+,\ K \subset K'^{\varepsilon} \text{ et } K' \subset K^{\varepsilon}\}.$$

Soit D un domaine de $\mathbb{R}^d$ et $E \subset D$ un ensemble rectifiable. Pour toute fonction ϕ définie sur D et à valeur dans $\mathbb{R}^d$, on note $\|\phi\|_{1,E}$ la norme L_1 restreinte à E, c'est-à-dire le réel $\|\phi\|_{1,E} = \int_E \|\phi(x)\| dx$. De même, $\|\phi\|_{\infty,E}$ est la norme uniforme restreinte à E : $\|\phi\|_{\infty,E} = \sup_{x \in E} \|\phi(x)\|$. On rappelle le résultat suivant qui garantit la stabilité de la projection [CCSM10].

Théorème 1.12. *Soient E un ouvert de $\mathbb{R}^d$ de bord rectifiable, K et K' deux compacts de $\mathbb{R}^d$. Soit $R = \max(\|d_K\|_{\infty,E}, d_{K'}\|_{\infty,E})$. On a :*

$$\|\mathfrak{p}_K - \mathfrak{p}_{K'}\|_{1,E} \leq Cst \times [\mathcal{H}^d(E) + \mathrm{diam}(K)\mathcal{H}^{d-1}(\partial E)] \times \sqrt{Rd_H(K, K')}$$

où la constante ne dépend que de la dimension.

La démonstration s'appuie sur un théorème plus général donnant la stabilité du gradient de fonctions convexes qui est un résultat clé de [CCSM10] et [MOG11] :

Théorème 1.13. *Soient E un sous-ensemble de $\mathbb{R}^d$ de bord rectifiable, et $f, g : E \to \mathbb{R}^d$ deux fonctions localement convexes sur E tels que $\mathrm{diam}(\nabla f(E) \cup \nabla g(E)) \leq C_1$, alors,*

$$\|\nabla f - \nabla g\|_{L_1(E)} \leq C_2 \times [\mathcal{H}^d(E) + (C_1 + \|f - g\|_{\infty,E}^{1/2})\mathcal{H}^{d-1}(\partial E)] \times \|f - g\|_{\infty,E}^{1/2}$$

où la constante C_2 ne dépend que de la dimension.

3.2 Pseudo-continuité de la projection

Le théorème suivant établit que la projection $\mathfrak{p}_K$ sur l'ensemble K qui est proche d'une hyper-surface lisse S se comporte "presque comme" la projection $\mathfrak{p}_S$. Ce résultat n'implique pas la continuité de la projection sur le compact K, mais il indique que si K est suffisament proche de S et x de x', alors $\mathfrak{p}_K(x)$ sera proche de $\mathfrak{p}_K(x')$. Ce résultat est lié aux propriétés classiques de la projection sur un ensemble à *reach* positif.

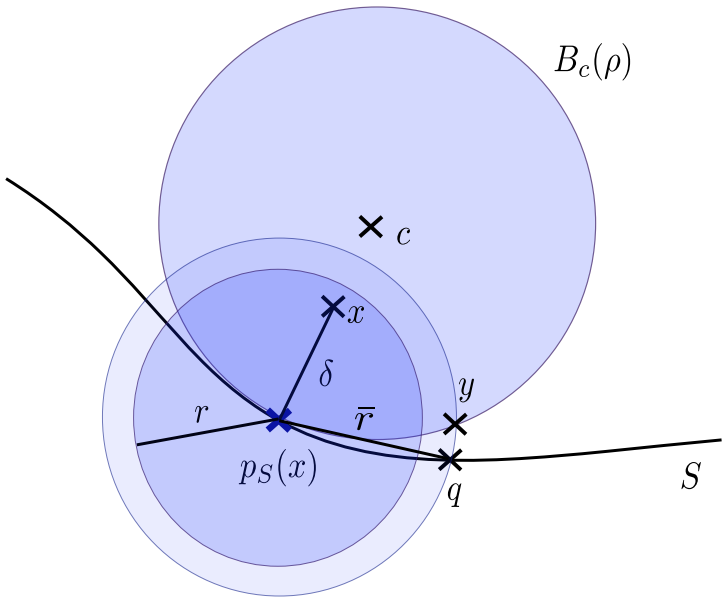


FIGURE 1.4 – Démonstration du lemme 1.15

Théorème 1.14. *Soit S une hypersurface plongée dans $\mathbb{R}^d$ de classe $\mathcal{C}^2$ et de reach $\rho > 0$. Soit K un compact tel que $d_H(S, K) = \varepsilon < 2\rho$ et $R < \rho$ un nombre positif. Si x et x' sont des points de S^R tels que $\|x - x'\| \leq \eta$, alors :*

$$\|\mathfrak{p}_K(x) - \mathfrak{p}_K(x')\| \leq 2\sqrt{8\varepsilon\rho} + 2\varepsilon + \frac{1}{1 - \frac{R}{\rho}}\eta.$$

Ce résultat sera utilisé dans le chapitre 3. La preuve du théorème est une conséquence des lemmes 1.15 et 1.16. On note $\mathrm{B}_p(r)$ la boule de centre p et de rayon r.

Lemme 1.15. *Soit S une hyper-surface plongée dans $\mathbb{R}^d$ de classe $\mathcal{C}^2$ et de reach $\rho > 0$. Soit $r < \rho$ et $x \in \mathbb{R}^d$ tel que $d(x, S) = \delta < r$. Si $q \in S$ satisfait $\|q - \mathfrak{p}_S(x)\| > r$ alors $\|x - q\|^2 > \frac{\delta}{\rho} r^2 + \delta^2$.*

Démonstration. La démonstration est illustrée sur la figure 1.4. Comme S a un *reach* ρ strictement positif, il existe une boule $\mathrm{B}_c(\rho)$ de centre c et de rayon ρ tangente à S au point $\mathfrak{p}_S(x)$. L'intérieur de cette boule n'intersecte pas S et les points c, x, et $\mathfrak{p}_S(x)$ sont alignés. Soit q un point de S tel que $\|q - \mathfrak{p}_S(x)\| = \bar{r} \geq r$. On note $\mathcal{P}$ le plan passant par c, x, $\mathfrak{p}_S(x)$ et p. Soit $y \in \partial\mathrm{B}_c(\rho) \cap \partial\mathrm{B}_{\mathfrak{p}_S(x)}(\bar{r}) \cap \mathcal{P}$. On a clairement que $\mathfrak{p}_S(x)$ est à égale distance de q et y. De plus, comme $q \notin \mathrm{B}_c(\rho)$, c est plus proche de y que de q, Par suite, le segment $[c, \mathfrak{p}_S(x)]$ est plus proche de y que de q, ce qui implique que $\|x - q\| \geq \|x - y\|$. Comme $\|x - y\|$ est la longueur de la diagonale d'une trapèze régulier, on a $\|x - y\|^2 \geq \frac{\delta}{\rho}\bar{r}^2 + \delta^2$, ce qui nous permet de conclure. □

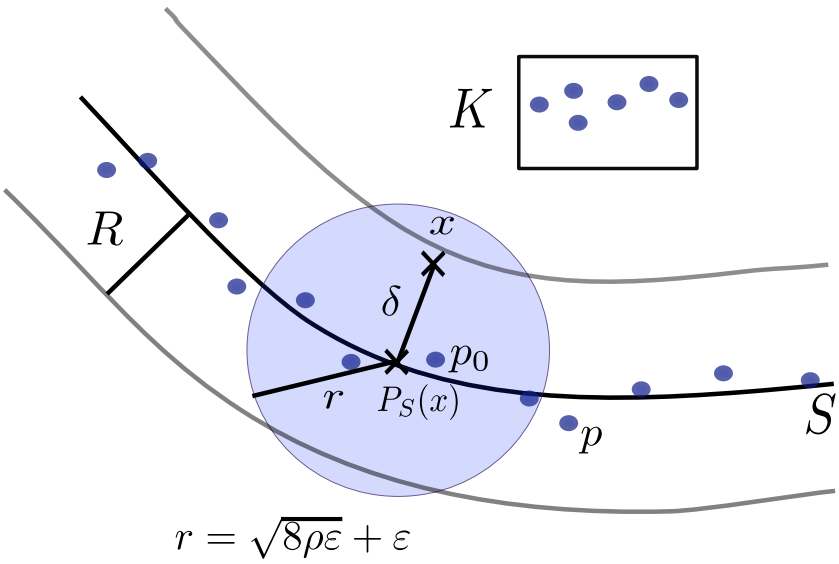


FIGURE 1.5 – Démonstration du lemme 1.16.

Lemme 1.16. *Soit S une hyper-surface plongée dans $\mathbb{R}^d$ de reach $\rho > 0$. Soit K un compact tel que $d_H(S, K) = \varepsilon$ avec $\varepsilon \leq 2\rho$. Soit R un nombre tel que $R < \rho$. Pour tout $x \in S^R$, on a*

$$\mathfrak{p}_K(x) \in B(\mathfrak{p}_S(x), \sqrt{8\varepsilon\rho} + \varepsilon)$$

Démonstration. La démonstration est illustrée sur la figure 1.5. Soit $x \in S^R$. On note $\delta = \|x - \mathfrak{p}_S(x)\|$. Il existe $p_0 \in K$ tel que $\|\mathfrak{p}_S(x) - p_0\| < \varepsilon$. On a donc $\|x - p_0\| \leq \|x - \mathfrak{p}_S(x)\| + \|\mathfrak{p}_S(x) - p_0\| \leq \delta + \varepsilon$, ce qui implique que $\|x - \mathfrak{p}_K(x)\| \leq \delta + \varepsilon$. Soit $p \in K$ tel que $p \notin B(\mathfrak{p}_S(x), \sqrt{8\rho\varepsilon} + \varepsilon)$. Il nous faut maintenant prouver

que $p \neq \mathfrak{p}_K(x)$. Commençons par supposer que $\delta < \varepsilon$. L'inégalité $\varepsilon \leq 2\rho$ implique que $3\varepsilon < \sqrt{8\rho\varepsilon} + \varepsilon$ et donc $\|p - \mathfrak{p}_S(x)\| \geq 3\varepsilon$. On a

$$\|x - p\| > \|p - \mathfrak{p}_S(x)\| - \|\mathfrak{p}_S(x) - x\| > 3\varepsilon - \delta > 2\varepsilon \geq \|x - \mathfrak{p}_K(x)\|,$$

ce qui assure que $p \neq \mathfrak{p}_K(x)$.

On suppose maintenant que $\delta \geq \varepsilon$. Il existe $q \in S$ tel que $\|q - p\| \leq \varepsilon$ et donc $\|q - \mathfrak{p}_S(x)\| \geq \sqrt{8\rho\varepsilon}$. Le lemme 1.15 implique de plus que $\|x - q\|^2 > \frac{\delta}{\rho} \times 8\rho\varepsilon + \delta^2$. Le fait que $\delta \geq \varepsilon$ implique que $\|x - q\| > \delta + 2\varepsilon$. On a finalement

$$\|x - p\| \geq \|x - q\| - \varepsilon > \delta + \varepsilon \geq \|x - \mathfrak{p}_K(x)\|.$$

Ce qui assure que $p \neq \mathfrak{p}_K(x)$.

□

Démonstration du Théorème 1.14.
L'inégalité triangulaire nous donne

$$\|\mathfrak{p}_K(x) - \mathfrak{p}_K(x')\| \leq \|\mathfrak{p}_K(x) - \mathfrak{p}_S(x)\| + \|\mathfrak{p}_S(x) - \mathfrak{p}_S(x')\| + \|\mathfrak{p}_S(x') - \mathfrak{p}_K(x')\|.$$

La proposition 1.5 de la section précédente assure que $\mathfrak{p}_S$ est $\frac{1}{1-\frac{R}{\rho}}$-Lipschitzienne sur S^R. On a donc $\|\mathfrak{p}_S(x) - \mathfrak{p}_S(x')\| \leq \frac{1}{1-\frac{R}{\rho}}\eta$. Les deux autres termes sont bornés avec le lemme 1.16. □

Chapitre 2

Estimation de Normales et Analyse en Composantes Principales

Sommaire

1 Stabilité de l'analyse en composantes principales 26
1.1 Généralités 26
1.2 Stabilité dans $\mathbb{R}^2$ 27
1.3 Démonstration du Théorème de stabilité. 28
2 Application en Géométrie Digitale 33
2.1 Prérequis en Géométrie Digitale 33
2.2 Convergence multigrille de l'estimateur de normale ACP . 35
3 Expérimentations 38
3.1 Expérimentations sur les nuages de points 38
3.2 Expérimentations en géométrie discrète 40
3.3 Analyse en composantes principales pondérée 43

Introduction

L'analyse en composantes principales est une méthode très utilisée en statistiques pour trouver des directions privilégiées dans des jeux de données [Jol86, ZHT06], mais aussi dans des domaines proches de la géométrie comme la vision par ordinateur [BMVS04, KRV06]. Hoppe et. al. [HDD+92] ont proposé un algorithme pour estimer la normale à un nuage de points approximant une courbe de $\mathbb{R}^2$ ou une surface de $\mathbb{R}^3$ utilisant l'analyse en composantes principales (ACP). La normale à un point p peut alors être définie comme la normale au plan minimisant la somme des distances au carré aux points situés dans une boule centrée en p. Mitra et. al. [MNG04] ont montré un résultat de convergence des normales calculées avec l'ACP.

Contribution. Dans ce chapitre, nous nous intéressons à des résultats effectifs en quantifiant l'erreur faite sur l'estimation de la normale. Nous montrons une

version plus géométrique du résultat de convergence de [MNG04] dans le cadre des courbes de reach positif de $\mathbb{R}^2$ (section 1). Dans le théorème de stabilité de [MNG04], les auteurs utilisent une notion d'uniformité de l'échantillonnage se basant sur un modèle statistique. Les points ne sont pas tirés uniformément autour de la courbe sous-jacente, mais ils sont l'image par une fonction de points tirés uniformément sur l'espace tangent. Nous avons choisi de contrôler l'échantillonnage à l'aide d'une notion plus géométrique, la notion de (ε, m)-échantillon. Nous appliquons ensuite ce théorème dans le contexte des contours digitaux pour obtenir la convergence multigrille d'un estimateur de normale (section 2). Nous présentons dans la section 3 des expérimentations sur deux types de données : des nuages de points et des contours digitaux. Les expérimentations permettent de confirmer la précision de la méthode. Nous présentons en complément une heuristique sur les surfaces en 3 dimensions. Nous proposons également dans la section 3.3 une heuristique basée sur l'analyse en composantes principales pondérée.

1 STABILITÉ DE L'ANALYSE EN COMPOSANTES PRINCIPALES

Dans cette section, nous établissons un résultat de stabilité Hausdorff d'un estimateur de normale basé sur l'analyse en composantes principales, sur des nuages de points approchant des courbes planes de classe $\mathcal{C}^2$. Mitra et. al. [MNG04] ont montré un résultat de convergence avec une hypothèse d'uniformité de l'échantillonnage se basant sur un modèle statistique. Nous remplaçons cette condition par une condition géométrique d'(ε, m)-échantillon. Nous donnons également une borne explicite de l'erreur faite sur l'estimation des normales.

1.1 GÉNÉRALITÉS

On rappelle tout d'abord le principe et les outils de l'analyse en composantes principales en dimension quelconque d. On se place dans un repère orthonormé $(p, \mathbf{i_1}, ..., \mathbf{i_d})$ de $\mathbb{R}^d$. Soit $P = (p_i)_{1 \leq i \leq n}$ un nuage de n points de $\mathbb{R}^d$. Pour tout entier $i \in \{1, ..., n\}$, on note $(x_{i,j})_{1 \leq j \leq d}$ les coordonnées de p_i. Soit $\overline{p}$ le barycentre des points $(p_i)_{1 \leq i \leq n}$ dont la j-ème coordonnée vérifie $\overline{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{i,j}$. On note $\mathcal{M}_{n,d}(\mathbb{R})$ l'ensemble des matrices $n \times d$ à coefficients réels. Soit $N \in \mathcal{M}_{n,d}(\mathbb{R})$ la matrice réelle $n \times d$ des coordonnées de P définie par $N_{i,j} = (x_{i,j} - \overline{x_j})$. La matrice $M = \frac{1}{n} N^T N$ est appelée *matrice de covariance* du nuage de point P. Un vecteur propre correspondant à la plus petite valeur propre de la matrice M est une normale au plan le plus proche du nuage P au sens des moindres carrés. Un tel vecteur est appelé *normale au nuage P au sens des moindres carrés* ou bien, en référence à la méthode statistique d'analyse en composantes principales, une *normale ACP*.

Remarque 2.1. *La normale au sens des moindres carrés ne dépend pas du choix du repère orthonormé, mais seulement du nuage de points P. En effet, soit $(p', \mathbf{i'_1}, ..., \mathbf{i'_d})$ un autre repère orthonormé. On remarque tout d'abord que le vecteur N reste invariant par changement d'origine. Soit Q la matrice de passage de changement de base.*

Alors dans la nouvelle base, la matrice des coordonnées centrées s'écrit $N' = Q^T NQ$. La matrice de covariance dans cette nouvelle base vérifie :

$$M' = \frac{1}{n}(Q^T NQ)^T Q^T NQ = \frac{1}{n} Q^T N^T QQ^T NQ = \frac{1}{n} Q^T N^T NQ = Q^T MQ.$$

La matrice M' est semblable à M et a donc les mêmes valeurs propres et les mêmes vecteurs propres.

En plus d'une hypothèse de proximité Hausdorff du nuage de points à la courbe lisse, nous avons besoin d'une hypothèse sur l'uniformité de l'échantillonnage des points pour établir le théorème de stabilité de l'estimateur de normale. Nous utilisons une notion classique de (ε, m)-échantillon inspirée de celle de [Dey06] qui garantit à la fois la proximité et l'uniformité de l'échantillonnage. Pour tout ensemble fini E, on note $|E|$ son cardinal.

Définition 2.2. *Soient S une hyper-surface de $\mathbb{R}^d$ de reach $\rho > 0$, P un nuage de points du plan, $\varepsilon \in \mathbb{R}^+$ et $m \in \mathbb{N}^*$. On dit que P est un (ε, m)-échantillon de S si*

- $d_H(S, P) < \varepsilon\rho$
- $\forall p \in P$, $|\{q \in P, \|p - q\| < 2\varepsilon\rho\}| < m$

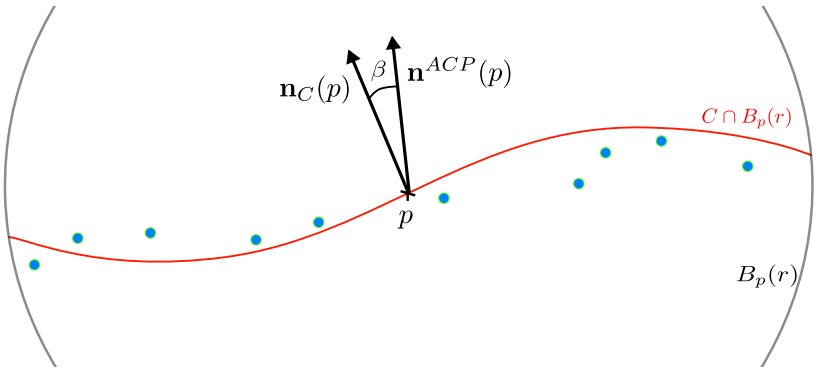


FIGURE 2.1 – Normale au sens des moindres carrés. Le vecteur $\mathbf{n}_C(p)$ est la normale à C en p et $\mathbf{n}^{ACP}(p)$ est la normale ACP du nuage de point P.

1.2 STABILITÉ DANS $\mathbb{R}^2$

Le résultat suivant contrôle l'erreur faite dans l'estimation de la normale à l'aide de l'analyse en composantes principales. Le théorème affirme que si P est un (ε, m)-échantillon d'une partie de la courbe C de *reach* ρ localisée autour d'un point $p \in C$, alors la normale estimé au point p à l'aide de l'analyse en composantes principales appliqué au nuage de points $P \cap \mathrm{B}_p(r)$ est bonne si on prend ε et r assez petits devant m et ρ.

Théorème 2.3. *Soit C une courbe de classe $\mathcal{C}^2$, de reach $\rho > 0$ et p un point de C. Soient $r \in \mathbb{R}$ tel que $0 < r < \frac{\rho}{4}$, $\varepsilon \in \mathbb{R}$ tel que $0 < \varepsilon < 1$. Soit $P \subset \mathrm{B}_p(r)$ un (ε, m)-échantillon de $C \cap B_p(r)$ tel que $m \leq \frac{|P|}{18}$. Si*

$$\gamma := 32m^2 \left(\frac{2r}{\rho} + \frac{2\rho\varepsilon}{r} + \frac{r^2}{\rho^2} + 2\varepsilon + \frac{\rho^2\varepsilon^2}{r^2} \right) < \frac{1}{2},$$

alors l'angle β entre la normale à C en p et une normale au nuage P au sens des moindres carrés satisfait :

$$\beta \leq \frac{\gamma(1+\gamma)}{(1-\gamma)^2}.$$

Le corollaire suivant nous donne le comportement asymptotique quand les paramètres r et ε sont liés par la relation $r = a\varepsilon^{\frac{1}{2}}$ (avec $a > 0$).

Corollaire 2.4. *Soit C une courbe de classe $\mathcal{C}^2$ de reach $\rho > 0$ et p un point de C. Soit $\varepsilon \in \mathbb{R}^{+*}$, $a \in \mathbb{R}^{+*}$ et $r = a\varepsilon^{\frac{1}{2}}$. Si $P \subset \mathrm{B}_p(r)$ est un (ε, m)-échantillon de $C \cap B_p(r)$ tel que $m \leq \frac{|P|}{18}$, alors l'angle β entre la normale à C en p et une normale au nuage P au sens des moindres carrés satisfait*

$$\beta = O(m^2\varepsilon^{\frac{1}{2}}),$$

où la constante intervenant dans le O ne dépend que de ρ et de a.

Ce résultat indique une assez forte dépendance à l'uniformité de l'échantillonnage puisque la borne dépend de manière quadratique en m. Néanmoins, si l'échantillon étudié en pratique est relativement uniforme, alors le terme en m^2 sera peu important.

1.3 Démonstration du Théorème de stabilité.

La démonstration donnée ici suit les étapes de la démonstration du résultat de [MNG04], mais en explicitant la borne en fonction de r, ε et ρ. Sans perdre en généralité, on se place dans le repère orthonormé $(p, \mathbf{i}, \mathbf{j})$ où $\mathbf{i}$ est un vecteur tangent à C en p. L'angle β est l'angle entre la normale au sens des moindres carrés et l'axe (Oy). La matrice de covariance M est une matrice symétrique pouvant s'écrire

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix}$$

où m_{11} et m_{22} sont positifs. Nous rappelons tout d'abord le théorème de Gershgorin [Bel65].

Théorème 2.5 (Gerschgorin)**.** *Soit $A = (A_{ij}) \in \mathcal{M}_{d,d}(\mathbb{R})$. Si les d disques de Gerschgorin $D_i := \{z \in \mathbb{R}, |z - A_{ii}| \leq \sum_{j \neq i} |A_{ij}|\}$ sont 2 à 2 disjoints, alors A admet d valeurs propres réelles distinctes $\lambda_1, ..., \lambda_d$. Plus précisément, pour tout entier $i \in \{1, ..., d\}$, on a $\lambda_i \in D_i$.*

La démonstration du théorème 2.3 résulte directement des lemmes 2.6 et 2.10. Les lemmes 2.7, 2.8 et 2.9 sont des résultats techniques intermédiaires. On introduit $\alpha = \frac{|m_{12}| + m_{22}}{m_{11}}$.

Lemme 2.6. *Sous les hypothèses du Théorème 2.3, si $\alpha < \frac{1}{2}$, alors*

$$\beta < \frac{\alpha(1+\alpha)}{(1-\alpha)^2}.$$

Démonstration. L'hypothèse $\alpha < \frac{1}{2}$, entraîne $m_{22} \leq \frac{1}{2}m_{11} - |m_{12}|$. Comme m_{22} est positif, on a $m_{11} - |m_{12}| > m_{22} + |m_{12}|$, ce qui entraîne que les disques de Gerchgorin sont disjoints. Si on note λ_1 et λ_2 les deux valeurs propres, on a donc d'après le théorème de Gerchgorin : $\lambda_1 \geq m_{11} - |m_{12}| > m_{22} + |m_{12}| \geq \lambda_2$. Soit $\begin{bmatrix} v \\ 1 \end{bmatrix}$ le vecteur propre correspondant à la plus petite valeur propre λ_2. Alors

$$\begin{aligned}
\begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix} \begin{bmatrix} v \\ 1 \end{bmatrix} &= \lambda_2 \begin{bmatrix} v \\ 1 \end{bmatrix} \\
\Rightarrow v \begin{bmatrix} m_{11} - \lambda_2 \\ m_{12} \end{bmatrix} &= -\begin{bmatrix} m_{12} \\ m_{22} - \lambda_2 \end{bmatrix} \\
\Rightarrow v &= -\frac{(m_{11} - \lambda_2)m_{12} + m_{12}(m_{22} - \lambda_2)}{(m_{11} - \lambda_2)^2 + m_{12}^2} \\
\Rightarrow |v| &\leq \frac{|m_{12}|(m_{11} - \lambda_2 + |m_{12}|)}{(m_{11} - \lambda_2)^2} \\
\Rightarrow |v| &\leq \frac{\alpha(1+\alpha)}{(1-\alpha)^2} \\
\Rightarrow \beta = \arctan(|v|) &\leq \arctan(\frac{\alpha(1+\alpha)}{(1-\alpha)^2}) \\
\Rightarrow \beta &\leq \frac{\alpha(1+\alpha)}{(1-\alpha)^2}
\end{aligned}$$

□

Nous avons besoin du résultat classique suivant, qui est une conséquence directe de la proposition 1.5 du chapitre précédent.

Lemme 2.7. *Soit C'' une courbe du plan de longueur finie, de reach $\rho > 0$ et d'extrémités q_1'' et q_2''. Si le segment $[q_1''q_2'']$ est situé dans le ρ-voisinage tubulaire de C'', alors la longueur $l(C'')$ de C'' vérifie*

$$l(C'') \leq \frac{\|q_2'' - q_1''\|}{1 - \frac{\eta}{\rho}},$$

où $\eta = \max\{\|q - \mathfrak{p}_{C''}(q)\|, q \in [q_1''q_2'']\}$.

Démonstration. Comme le *reach* de C'' est plus grand que ρ, la projection $\mathfrak{p}_{C''}$ sur C'' est bien définie en tout point du segment $[q_1''q_2'']$. Par la proposition 1.5, on sait que sa restriction à $[q_1''q_2'']$ est $\frac{1}{1-\frac{\eta}{\rho}}$-Lipschitzienne, ce qui donne

$$l(C'') = l\left(\mathfrak{p}_{C''}([q_1''q_2''])\right) \leq \frac{\|q_2'' - q_1''\|}{1 - \frac{\eta}{\rho}}.$$

□

Lemme 2.8. *Soit C une courbe du plan de reach $\rho > 0$ et p un point de C. Si $r < \rho$ et q un point quelconque de $C \cap \mathrm{B}_p(r)$, alors la pente α_q de la courbe en q par rapport à la tangente à C en p vérifie*

$$\alpha_q \leq \frac{r}{\sqrt{\rho^2 - r^2}}$$

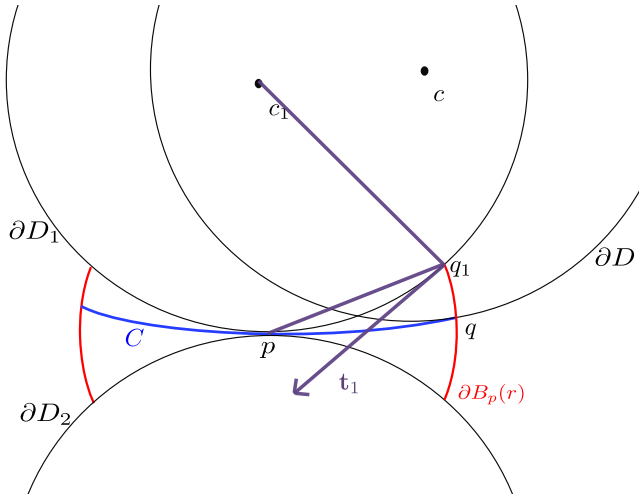


FIGURE 2.2 – Démonstration du lemme 2.8

Démonstration. On peut se ramener au cas où $q \in \partial \mathrm{B}_p(r) \cap C$. En effet, si ce n'est pas le cas, alors $q \in \partial \mathrm{B}_p(\tilde{r}) \cap C$ avec $\tilde{r} < r$. Le fait que la borne $\frac{\tilde{r}}{\sqrt{\rho^2-\tilde{r}^2}}$ soit plus petite que $\frac{r}{\sqrt{\rho^2-r^2}}$ suffit pour se restreindre à $q \in \partial \mathrm{B}_p(r) \cap C$. Le lemme 7 de [AB99] implique que $\mathrm{B}_p(r) \cap C$ n'a qu'une seule composante connexe. Le fait que C soit à *reach* positif ρ implique que la courbe C est contenue dans le complémentaire de l'union de deux disques D_1 et D_2 (cf. figure 2.2). Soit q_1 un plus proche voisin de q dans l'ensemble $(\partial D_1 \cup \partial D_2) \cap \partial \mathrm{B}_p(r)$. Supposons que q_1 appartienne à ∂D_1 de centre c_1. Le cercle ∂D_1 a une pente égale à $\frac{r}{\sqrt{\rho^2-r^2}}$ en q_1. Notons D le disque de centre c et de rayon ρ tangent à C en q. Notons $\mathbf{t}$ et $\mathbf{t}_1$ des vecteurs tangents respectivement à ∂D en q et à ∂D_1 en q_1 tels que ces vecteurs forment un angle aigu avec le vecteur $(p - q_1)$. On a alors

$$\frac{\pi}{2} = |\langle \mathbf{t_1}, p - q_1 \rangle| + |\langle p - q_1, c_1 - q_1 \rangle| = |\langle \mathbf{t}, p - q \rangle| + |\langle p - q, c - q \rangle|$$

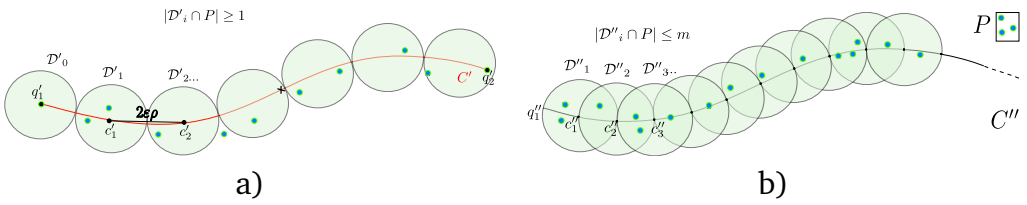
Par l'absurde, si la pente de C en q est plus grande que celle de ∂D_1 en q_1, alors l'angle $\langle \mathbf{t_1}, p - q_1 \rangle$ est plus petit que l'angle $\langle \mathbf{t}, p - q \rangle$ en valeur absolue. Cela entraîne que l'angle $\langle p - q, c - q \rangle$ est plus petit que l'angle $\langle p - q_1, c_1 - q_1 \rangle$ en valeur absolue. On a donc $\|c - p\| < \|c_1 - p\| = \rho$, ce qui est impossible. □

Lemme 2.9. *Soit C' une courbe du plan d'extrémités q'_1 et q'_2, de reach $\rho > 0$. Soient $\varepsilon \in \mathbb{R}^+$ tel que $\varepsilon < 1$, et $m \in \mathbb{N}^*$. Soit P un (ε, m)-échantillon de C'. Alors*

$$a) \qquad |P| \geq \left\lfloor \frac{\|q'_2 - q'_1\|}{2\varepsilon\rho} \right\rfloor.$$

De plus, pour toute courbe connexe $C'' \subset C'$, le nombre de points du nuage P situés dans le voisinage tubulaire $C''^{\varepsilon\rho}$ est majoré :

$$b) \quad |P \cap C''^{\varepsilon\rho}| \leq (\left\lfloor \frac{l(C'')}{2\varepsilon\rho} \right\rfloor + 3)m.$$



a) b)

FIGURE 2.3 – Démonstration du lemme 2.9 a) Définition de l'ensemble $\mathcal{R}'$ b) Définition de l'ensemble $\mathcal{R}''$. L'ensemble $\mathcal{R}''$ est un recouvrement de P.

Démonstration. On note $c'_0 = q'_1$. On définit l'ensemble de points $\{c'_i\}$ de C' en parcourant la courbe C' entre q'_1 et q'_2 de la manière suivante : c'_i est le point de $B_{c'_{i-1}}(2\varepsilon\rho) \cap C'$ qui vient après c'_{i-1} dans le parcours de C'. Le dernier point est noté c'_M (figure 2.3.a. L'ensemble $\mathcal{R}' = \{\mathcal{D}'_i = B_{c'_i}(\varepsilon\rho),\ 0 \leq i \leq M\}$ est constitué de M disques ouverts d'intersection deux à deux vides. La longueur $\mathcal{L}'$ de la ligne polygonale reliant les points c'_0, c'_1,... c'_M, puis q'_2 vérifie

$$\mathcal{L}' = \left(\sum_{i=0}^{M-1} \|c'_{i+1} - c'_i\|\right) + \|q'_2 - c'_M\| = M \times 2\varepsilon\rho + \|q'_2 - c'_M\| \leq (M+1) \times 2\varepsilon\rho$$

En utilisant le fait que $\mathcal{L}' \geq \|q'_2 - q'_1\|$, on a alors $(M+1) \times 2\varepsilon\rho > \|q'_2 - q'_1\|$. Or le fait que P soit un (ε, m)-échantillon implique que $|D'_i| \geq 1$. Comme les disques D'_i sont disjoints, on a $|P| \geq M+1$, ce qui donne la première inégalité.

Pour la seconde inégalité, on note q''_1 et q''_2 les extrémités de C''. On définit de manière itérative les points c''_i en parcourant la courbe C'' de q''_1 à q''_2 de la manière suivante : $c''_0 = q''_1$ et c''_i est le point de $B_{c''_{i-1}}(2\varepsilon\rho) \cap C''$ qui suit c''_{i-1} dans le sens de parcours (figure 2.3.b). Le dernier point avant q''_2 est alors noté c''_{N-1} avec $N > 0$ et on note $c''_N = q''_2$. L'ensemble $\mathcal{R}'' = \{\mathcal{D}''_i = B_{c''_i}(2\varepsilon\rho),\ 0 \leq i \leq N\}$ forme un recouvrement de $C''^{\varepsilon\rho}$. La longueur $\mathcal{L}''$ de la ligne polygonale reliant les points c''_0, c''_1,... c''_N vérifie

$$\mathcal{L}'' = \left(\sum_{i=0}^{N-2} \|c''_{i+1} - c''_i\|\right) + \|q''_2 - c''_{N-1}\| \geq 2(N-1)\varepsilon\rho.$$

Le fait que $\mathcal{L}'' \leq l(C'')$ donne l'inégalité $|\mathcal{R}''| = N+1 \leq \left\lceil \frac{l(C'')}{2\varepsilon\rho} \right\rceil + 2 = \left\lfloor \frac{l(C'')}{2\varepsilon\rho} \right\rfloor + 3$. Or l'hypothèse P est un (ε, m)-échantillon de C' implique que pour tout entier i, $|\mathcal{D}''_i \cap P| \leq m$, ce qui donne la deuxième inégalité.

□

Lemme 2.10. *Sous les hypothèses du théorème 2.3*

$$\alpha \leq 32m^2 \left(\frac{2r}{\rho} + \frac{2\rho\varepsilon}{r} + \frac{r^2}{\rho^2} + 2\varepsilon + \frac{\rho^2\varepsilon^2}{r^2}\right).$$

Démonstration. Le fait que C soit de *reach* ρ implique que C est contenu dans le complémentaire de l'union de deux disques tangents de rayon ρ, ce qui donne que

pour tout point $p' \in C \cap \mathrm{B}_p(r)$ de coordonnées (x', y') dans le repère $(p, \mathbf{i}, \mathbf{j})$, on a $|x'| \leq r$ et

$$|y'| \leq \rho - \sqrt{\rho^2 - x'^2} \leq \frac{x'^2}{\rho + \sqrt{\rho^2 - x'^2}} \leq \frac{r^2}{\rho}.$$

Pour tout point $q \in P \subset \mathrm{B}_p(r)$ de coordonnées (x, y), on a $|x| \leq r$. Comme la distance de Hausdorff entre P et C est inférieur à $\varepsilon\rho$, on a $|y| \leq \frac{r^2}{\rho} + \varepsilon\rho$. On note n le cardinal du nuage de point P. On peut donc majorer les deux termes :

$$\begin{aligned} |m_{12}| &= \left| \frac{1}{n} \sum_{i=1}^{n} x_i y_i - \frac{1}{n^2} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \right| \leq \frac{2r^3}{\rho} + 2\varepsilon\rho r \\ m_{22} &\leq \frac{1}{n} \sum_{i=1}^{n} y_i^2 \leq \frac{r^4}{\rho^2} + 2r^2\varepsilon + \varepsilon^2\rho^2. \end{aligned}$$

Enfin, dans le but de minorer $m_{11} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$, on considère le sous-ensemble du plan $T = \{(x, y) \in \mathbb{R}^2,\ |x - \overline{x}| < \frac{r}{4m}\} \cap \mathrm{B}_p(r)$ (figure 2.4) et on minore $N = \left|\{(x, y) \in P, |x - \overline{x}| \geq \frac{r}{4m}\}\right| = n - |T \cap P|$. On majore pour cela $|T \cap P|$. L'hypothèse $r \leq \frac{\rho}{4}$ permet d'appliquer le lemme 7 de [AB99] qui implique que $C'' := C \cap T$ n'a qu'une composante connexe. On note q_1'' et q_2'' les deux extrémités de C''. Comme $|P \cap T| \leq |P \cap C''^{\varepsilon\rho}|$, le lemme 2.9.b appliqué à C'' entraîne $|T \cap P| < \left(\left\lfloor \frac{l(C'')}{2\epsilon\rho} \right\rfloor + 3\right) m$.

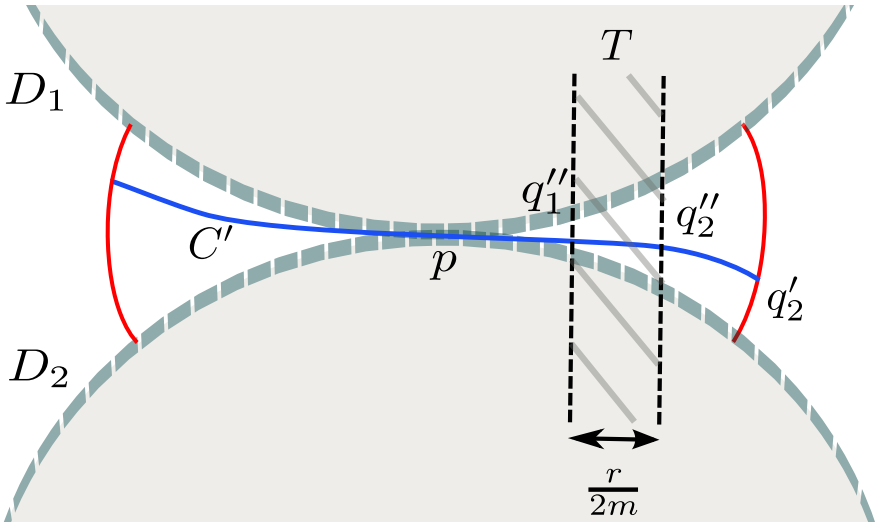


FIGURE 2.4 – Représentation de l'ensemble T.

On applique le lemme 2.7 à $C'' = C \cap T$. Ceci nous donne $l(C'') \leq \frac{\|q_2'' - q_1''\|}{1 - \frac{\eta}{\rho}}$, où $\eta = \max\{\|q - \mathfrak{p}_{C''}(q)\|, q \in [q_1'' q_2'']\}$. Le fait que la courbe C soit de *reach* $\rho > 0$ implique qu'elle est contenue dans le complémentaire de l'union de deux disques D_1 et D_2 tangents de rayon ρ (voir figure 2.4). De plus, la pente de la courbe

est majorée en valeur absolue par le lemme 2.8 par $\frac{r}{\sqrt{\rho^2-r^2}} \leq \frac{1}{\sqrt{15}}$. On a alors $\|q_2'' - q_1''\| \leq \sqrt{1+\left(\frac{1}{\sqrt{15}}\right)^2}\frac{r}{2m} \leq \frac{2r}{\sqrt{15}m}$. De même, on a $\eta = \max\{\|q - \mathfrak{p}_{C''}(q)\|, q \in [q_1''q_2'']\} \leq \frac{r}{2\sqrt{15}m} \leq \frac{\rho}{30}$. Cela implique $l(C \cap T) \leq \frac{60r}{29\sqrt{15}m} = \frac{4\sqrt{15}r}{29m}$. D'après le lemme 2.9.b, on a :

$$|T \cap P| \leq \left\lfloor \frac{4\sqrt{15}r}{29m\varepsilon\rho} \right\rfloor m + 3m \leq \left\lfloor \frac{4\sqrt{15}rm}{29m\varepsilon\rho} \right\rfloor + 3m = \left\lfloor \frac{4\sqrt{15}r}{29\varepsilon\rho} \right\rfloor + 3m.$$

Majorons maintenant chacun des deux termes en fonction de n. L'hypothèse $m \leq \frac{n}{18}$ donne immédiatement que $3m \leq \frac{n}{6}$. Pour le premier terme, on applique le lemme 2.9.a à $C' = C \cap \mathrm{B}_p(r)$ d'extrémité q_1' et q_2'. Cela nous donne $\left\lfloor \frac{\|q_1'-q_2'\|}{\epsilon\rho} \right\rfloor \leq n$. Pour minorer $\|q_1' - q_2'\|$, on remarque que tout point q de coordonnées (x_q, y_q) appartenant à une des deux composantes connexes de $\left(\mathbb{R}^d \backslash (D_1 \cup D_2)\right) \cap \partial \mathrm{B}_p(r)$ vérifie $|y_q| \leq \frac{r^2}{\rho}$. Cela implique que $|x_q| \geq \sqrt{r^2 - \frac{r^4}{\rho^2}} \geq r\sqrt{1-\frac{1}{16}} = r\frac{\sqrt{15}}{4}$. Or le fait que $C' = C \cap \mathrm{B}_p(r)$ soit contenu dans le complémentaire de $D_1 \cup D_2$ implique que q_1' et q_2' appartiennent à deux composantes connexes différentes de $\left(\mathbb{R}^d \backslash (D_1 \cup D_2)\right) \cap \partial \mathrm{B}_p(r)$ (voir figure 2.4). Ceci implique $\|q_1' - q_2'\| \geq r\frac{\sqrt{15}}{2}$. Cela nous donne $\left\lfloor \frac{r\sqrt{15}}{2\epsilon\rho} \right\rfloor \leq n$. Cela entraîne $\left\lfloor \frac{4\sqrt{15}r}{29\epsilon\rho} \right\rfloor \leq \frac{1}{3}n$. On peut donc minorer $N = n - |T \cap P|$ par $N \geq n - \frac{n}{3} - \frac{n}{6} = \frac{n}{2}$. Enfin, cela nous donne la minoration de m_{11} :

$$m_{11} = \frac{1}{n}\sum_{i=1}^{n}(x-\overline{x})^2 \geq \frac{1}{n} \times \frac{n}{2} \times \left(\frac{r}{4m}\right)^2 = \frac{r^2}{32m^2}.$$

Ce qui donne $\alpha < 32m^2\left(\frac{2r}{\rho} + \frac{2\rho\varepsilon}{r} + \frac{r^2}{\rho^2} + 2\varepsilon + \frac{\rho^2\varepsilon^2}{r^2}\right)$. □

2 Application en Géométrie Digitale

Le but de cette partie est d'appliquer la méthode précédente d'estimation de normale à l'aide de l'ACP dans le cas particulier des ensembles digitaux. On rappelle tout d'abord les outils nécessaires en géométrie digitale. On établit ensuite un résultat de convergence multigrille de notre estimateur.

2.1 Prérequis en Géométrie Digitale

Contours digitaux. Comme nous nous intéressons au comportement d'un estimateur de quantité différentielle sur des objets digitaux, nous devons tout d'abord établir les liens entre l'objet euclidien sous-jacent et l'objet digital. On définit pour tout point $x \in \mathbb{R}^d$ avec $x = (x_1, ..., x_d)$, son voxel associé $\mathrm{vox}(x) = [x_1 - \frac{1}{2}h, x_1 + \frac{1}{2}h] \times ... \times [x_d - \frac{1}{2}h, x_d + \frac{1}{2}h]$. Considérons un domaine $X \subset \mathbb{R}^d$. On note $Dig_h(X)$ la discrétisation de Gauss de X dans une grille d-dimensionnelle de pas h :

$$Dig_h(X) := \mathrm{vox}(X \cap (h\mathbb{Z})^d).$$

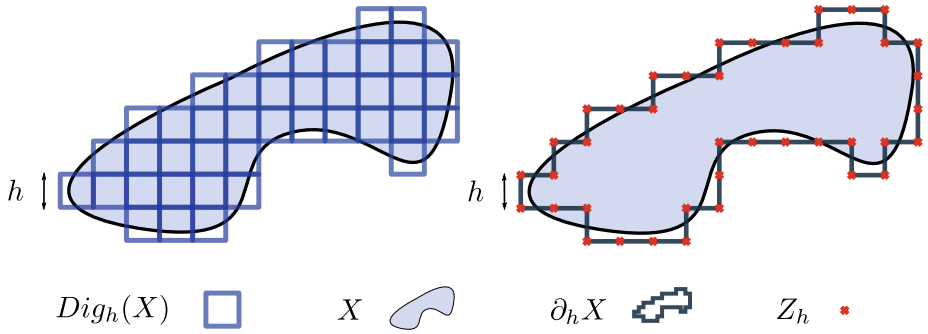


FIGURE 2.5 – Définition et représentation de Z_h

On note ∂X le bord de X et $\partial_h X$ la h-frontière, c'est à dire le bord de $Dig_h(X)$. On définit Z_h comme l'ensemble de points : $Z_h = \partial_h X \cap h(\mathbb{Z} + \frac{1}{2})^d$. L'ensemble Z_h est un nuage de points qui a la propriété d'être proche de ∂X au sens de Hausdorff :

Lemme 2.11. *Soit X un domaine de $\mathbb{R}^d$ dont le bord ∂X est de reach $\rho > 0$, pour tout réel h tel que $0 < h < \frac{2\rho}{\sqrt{d}}$, on a*

$$d_H(\partial X, Z_h) < \left(\frac{\sqrt{d} + \sqrt{d-1}}{2}\right) h$$

Démonstration. On a l'inégalité $d_H(\partial X, Z_h) \leq d_H(\partial X, \partial_h X) + d_H(\partial_h X, Z_h)$. Le théorème 1 de [LT14] établie que, $d_H(\partial X, \partial_h X) \leq \frac{\sqrt{d}}{2} h$. D'autre part, on a immédiatement par construction que $d_H(\partial_h X, Z_h) \leq \frac{\sqrt{d-1}}{2} h$, ce qui permet de conclure. □

CONVERGENCE MULTIGRILLE. La convergence multigrille est une approche naturelle pour garantir la stabilité d'estimateurs appliqués à des discrétisation d'un objet lisse sous-jacent. Étant donné un estimateur défini sur $Dig_h(X)$, l'idée de la convergence multigrille est de garantir que l'estimateur tend bien vers la quantité à estimer sur X quand h tend vers 0. On peut se référer à [CLR12] pour l'étude de la convergence multigrille de quantité géométrique élémentaires telles que la longueur d'une courbe ou l'aire d'une surface. La figure 2.6 représente les discrétisés successifs de plus en plus fins d'un domaine de $\mathbb{R}^2$. Cette figure illustre la convergence multigrille de l'estimateur d'aire puisque l'on visualise intuitivement que l'aire du domaine discrétisé tend vers l'aire du domaine X quand la grille devient fine. La plupart du temps, on ne peut avoir une convergence multigrille que sur une famille restreinte d'objet X. La définition suivante donne la définition plus formelle de la convergence multigrille sur une famille d'objet $\mathcal{X}$.

Définition 2.12. *Soit $\widehat{E}_{X,h}$ un estimateur d'une quantité géométrique E défini sur ∂X. On dit que $\widehat{E}_{X,h}$ est convergent multigrille sur la famille $\mathcal{X}$ si pour tout $X \in \mathcal{X}$,*

$$\exists h_0 > 0,\ \forall h < h_0,\ \forall p \in \partial X \quad \|\widehat{E}_{X,h}(p) - E(p)\| \leq \sigma(h)$$

où $\sigma(h)$ tend vers 0 quand h tend vers 0.

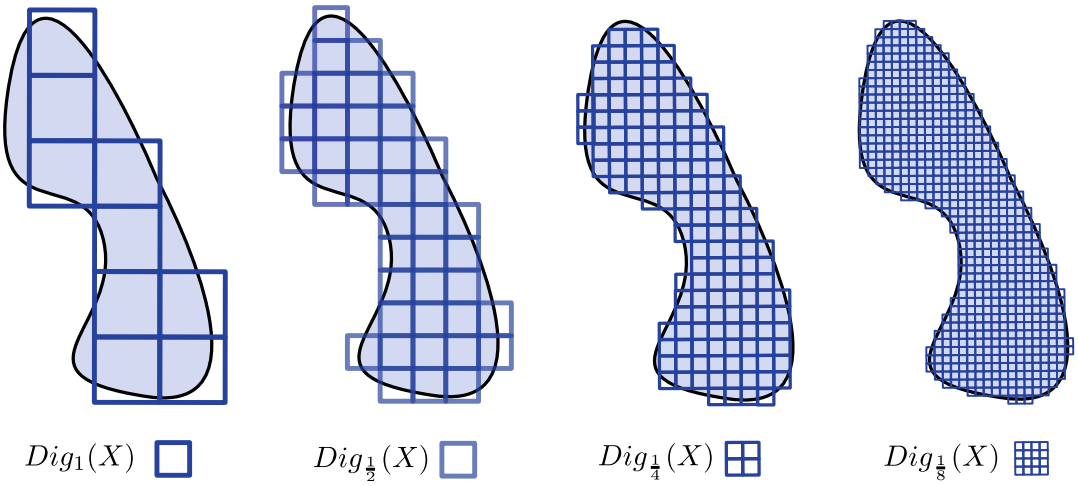


FIGURE 2.6 – Suite de discrétisations de plus en plus fines.

Cette définition de convergence multigrille d'un estimateur est une version relaxée de la convergence usuelle [CLR12]. En effet, nous supposons que nous connaissons le point $p \in \partial X$, ce qui est faux en général. Il serait plus naturel de définir $\widehat{E}_{X,h}$ sur l'ensemble $\partial_h X$ ou sur Z_h.

2.2 CONVERGENCE MULTIGRILLE DE L'ESTIMATEUR DE NORMALE ACP

L'estimateur de normale ACP étudié dans la section 1 est défini sur tout nuage de points. En particulier, on l'applique dans cette section à l'ensemble Z_h. On s'intéresse donc à la convergence multigrille de l'estimateur $\widehat{n}_{X,h} = \widehat{\mathbf{n}}^{ACP}_{Z_h,r}$ qui est pour tout point p la normale ACP associé au nuage de point $Z_h \cap \mathrm{B}_p(r)$. Le théorème de la section 1 permet d'établir la convergence multigrille de cet estimateur en 2D :

Théorème 2.13. *Soit $\mathcal{X}$ la famille des domaines $X \subset \mathbb{R}^2$ dont le bord ∂X est de reach positif. Si a un réel strictement positif et $r = ah^{\frac{1}{2}}$, alors l'estimateur $\widehat{\mathbf{n}}^{ACP}_{Z_h,r}$ est convergent multigrille sur $\mathcal{X}$ et vérifie pour tout $X \in \mathcal{X}$:*

$$\forall p \in \partial X, \quad \|\widehat{\mathbf{n}}^{ACP}_{Z_h,r}(p) - \mathbf{n}_{\partial X}(p)\| = O(h^{\frac{1}{2}})$$

où la constante intervenant dans le $O(.)$ ne dépend que du reach de ∂X.

La démonstration découle directement du corollaire 2.4 et du lemme 2.15. On peut supposer sans perdre en généralité que le *reach* ρ vaut 1. Le lemme suivant est un résultat intermédiaire :

Lemme 2.14. *Soit X un domaine du plan dont le bord ∂X est de reach $\rho = 1$ et p un point de ∂X. Soient $\varepsilon > 0$ et $r = \sqrt{\varepsilon}$. Soient q_1 et q_2 les deux extrémités d'une composante connexe de $\mathrm{B}_p(r+\varepsilon) \backslash \mathrm{B}_p(r) \cap \partial X$. Si ε est assez petit, alors*

$$\|q_1 - q_2\| \leq 2\varepsilon.$$

Démonstration. On se place dans le repère orthonormé $(p, \mathbf{t}_{\partial X}(p), \mathbf{n}_{\partial X}(p))$ où $\mathbf{t}_{\partial X}(p)$ et $\mathbf{n}_{\partial X}(p)$ sont respectivement le vecteur unitaire tangent et normal à ∂X en p. Le fait que ∂X soit de *reach* 1 implique que la courbe est contenu dans le complémentaire de l'union de deux boules tangentes en p de rayon 1. Supposons que $q_1 \in \partial X \cap \partial \mathrm{B}_p(r)$ et $q_2 \in \partial X \cap \partial \mathrm{B}_p(r+\varepsilon)$ ont pour coordonnées (x_{q_1}, y_{q_1}) et (x_{q_2}, y_{q_2}). Si α_{max} est la pente maximale de $\partial X \cap \mathrm{B}_p(r)$, on a alors

$$\|q_1 - q_2\| \leq \sqrt{1+\alpha_{max}^2}|x_{q2} - x_{q1}|.$$

Or on a par définition $|x_{q_2}| \leq r + \varepsilon$. Pour minorer $|x_{q_1}|$, majorons $|y_{q_1}|$. Le fait que q_1 est situé entre les deux cercles ∂D_1 et ∂D_2 implique que

$$|y_{q_1}| \leq 1 - \sqrt{1-r^2} = \frac{1}{2}r^2 + o(r^2).$$

Quand r est assez petit, on a $|y_{q_1}| \leq r^2$ Comme $q_1 \in \mathrm{B}_p(r)$, on a alors

$$|x_{q_1}| \geq \sqrt{r^2 - r^4} = r - \frac{1}{2}r^2 + o(r^2).$$

Pour r assez petit, on a alors $|x_{q_1}| \geq r - r^2$ et en conséquence $|x_{q_2} - x_{q_1}| \leq r^2$. Le lemme 2.8 implique que la pente α_{max} est majorée en valeur absolue par

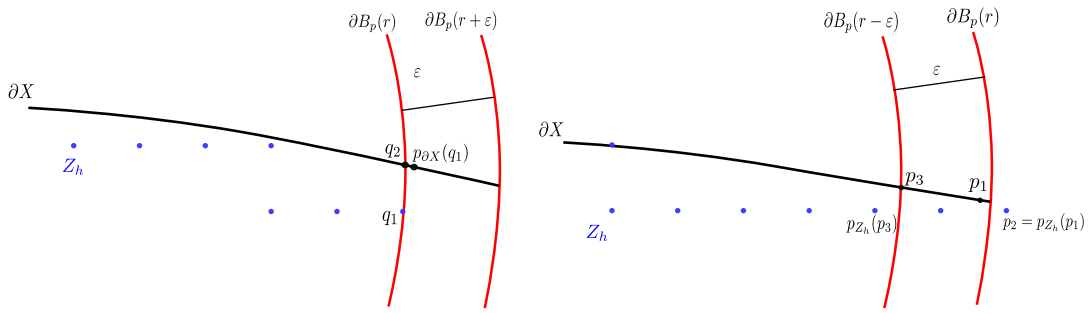
$$\alpha_{max} \leq \frac{r+\varepsilon}{\sqrt{1-(r+\varepsilon)^2}} = (r+\varepsilon)\left(1 + \frac{1}{2}(r+\varepsilon)^2 + o\left((r+\varepsilon)^2\right)\right).$$

Si on prend r et ε assez petit, on a alors $\alpha_{max} \leq \frac{3}{2}(r+\varepsilon)$. On a enfin

$$\begin{aligned} \|q_1 - q_2\| \leq \sqrt{1+\alpha_{max}^2}|x_{q2} - x_{q1}| &\leq \sqrt{1 + \left(\frac{3}{2}\right)^2 (r+\varepsilon)^2 r^2} \\ &= r^2\left(1 + \frac{4}{9}(r+\varepsilon)^2 + o((r+\varepsilon)^2)\right) \end{aligned}$$

Comme $r = \sqrt{\varepsilon}$, si ε est assez petit, $\|q_1 - q_2\| \leq 2\varepsilon$

□



a) $Z_h \subset (\partial X \cap \mathrm{B}_p(r))^{3\varepsilon}$ b) $\partial X \subset (Z_h \cap \mathrm{B}_p(r))^{3\varepsilon}$

FIGURE 2.7 – Démonstration du lemme 2.15.

Lemme 2.15. *Soit X un domaine dont le bord ∂X est de reach $\rho = 1$. Si $h \in \mathbb{R}^{+\star}$ est assez petit et $r = \frac{2}{\sqrt{3}} h^{\frac{1}{2}}$, alors $Z_h \cap \mathrm{B}_p(r)$ est un $(4h, 289)$-échantillon de $\partial X \cap \mathrm{B}_p(r)$ et $289 \leq \frac{|Z_h \cap \mathrm{B}_p(r)|}{16}$.*

Démonstration. Par définition du (ε, m)-échantillon (Définition 2.2), il faut montrer que :

- $d_H(\partial X \cap \mathrm{B}_r(p), Z_h \cap \mathrm{B}_r(p)) \leq 4h$
- $\forall p \in \partial X$, $|\{q \in Z_h, \|p - q\| \leq 8h\}| \leq 289$

Le second point est immédiat car on remarque que toute boule de rayon $8h$ contient moins de $17^2 = 289$ points de $h\mathbb{Z}^2$. Pour le second point, on utilise le lemme 2.11 qui nous donne $d_H(C, Z_h) \leq \frac{\sqrt{3}+\sqrt{2}}{2} h \leq \frac{4}{3} h$. Notons $\varepsilon = \frac{4}{3} h$. On ne peut pas conclure que $d_H(C \cap \mathrm{B}_r(p), Z_h \cap \mathrm{B}_r(p)) \leq \varepsilon$ car un problème de bord peut apparaître.
Soit $q_1 \in Z_h \cap \mathrm{B}_p(r)$, montrons qu'il existe un élément de ∂X à une distance de q_1 plus petite que 3ε (voir figure 2.7.a) . Si $\mathfrak{p}_{\partial X}(q_1) \in \mathrm{B}_p(r)$, alors $\mathfrak{p}_{\partial X}(q_1)$ convient. Sinon, $\mathfrak{p}_{\partial X}(q_1) \in (\mathrm{B}_p(r+\varepsilon) \backslash \mathrm{B}_p(r)) \cap \partial X$. Pour h assez petit, il existe un point $q_2 \in \partial \mathrm{B}_p(r) \cap \partial X$ tel que $\|q_2 - \mathfrak{p}_{\partial X}(q_1)\| \leq 2\varepsilon$ par le lemme précédent. Réciproquement, pour $p_1 \in \partial X \cap \mathrm{B}_p(r)$, montrons qu'il existe un élément de Z_h à une distance de p_1 plus petite que 3ε (voir figure 2.7.b). Soit $p_2 \in Z_h$ un plus proche voisin de p_1 dans Z_h. Si $p_2 \in \mathrm{B}_p(r)$ alors il convient. Sinon, on a $p_1 \in (\mathrm{B}_p(r) \backslash \mathrm{B}_p(r-\varepsilon)) \cap \partial X$. Pour h assez petit, il existe un point $p_3 \in \partial \mathrm{B}_p(r-\varepsilon) \cap \partial X$ tel que $\|p_3 - p_1\| \leq 2\varepsilon$ par le lemme précédent. Comme tous les plus proches voisins de p_3 dans Z_h sont dans $\mathrm{B}_p(r)$ et à une distance plus petite que ε de p_3, on a le résultat.

Les deux derniers points impliquent que $d_H(\partial X \cap \mathrm{B}_r(p), Z_h \cap \mathrm{B}_r(p)) \leq 4h$. Pour montrer que $289 \leq \frac{|Z_h \cap \mathrm{B}_p(r)|}{16}$, le lemme 2.9.a appliqué à $\mathrm{B}_p(r) \cap Z_h$ implique que $|Z_h \cap \mathrm{B}_p(r)|$ est minoré à une constante près par $\frac{1}{\sqrt{h}}$.

□

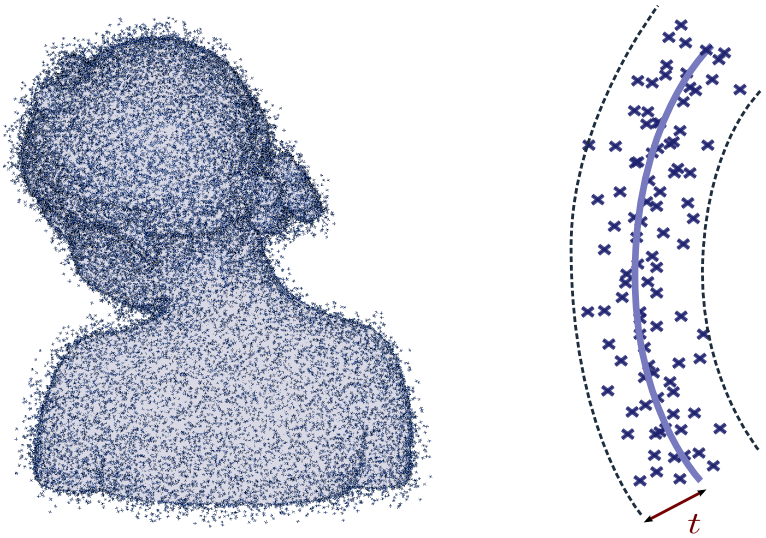


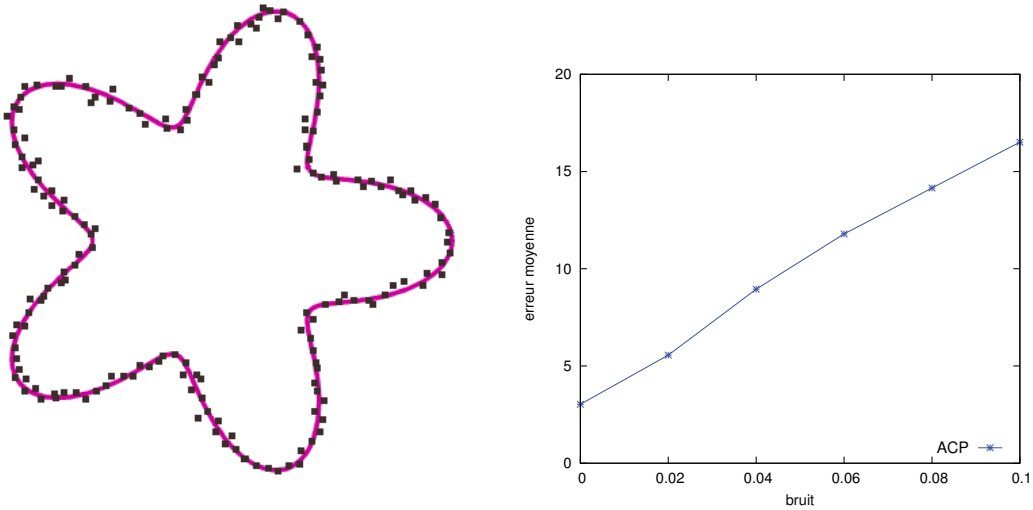
FIGURE 2.8 – Modèle de bruit Hausdorff.

3 EXPÉRIMENTATIONS

3.1 EXPÉRIMENTATIONS SUR LES NUAGES DE POINTS

Dans cette sous-section, nous présentons des résultats expérimentaux de l'estimateur de normale ACP sur des nuages de points. Nous présentons, dans un premier temps, des résultats validant le résultat de stabilité de la section 1 sur les courbes en 2D. L'essentiel de cette thèse s'intéresse à l'estimation des normales et courbures sur des approximations de surfaces en 3D ; nous présentons dans un second temps une heuristique dans ce contexte.

MODÈLE DE BRUIT HAUSDORFF. Dans le but de tester la résistance au bruit Hausdorff des différentes méthodes développées dans cette thèse, je définis le modèle de bruit Hausdorff. Ce modèle de bruit a essentiellement pour but de confirmer expérimentalement les résultats théoriques faisant intervenir une hypothèse Hausdorff. Il est défini par l'algorithme suivant : pour un paramètre de bruit t chaque point du nuage initial non bruité est translaté d'une distance aléatoire uniforme entre 0 et t dans une direction aléatoire de l'espace ambiant. La figure 2.8 montre l'exemple d'une courbe bruitée avec un tel modèle.



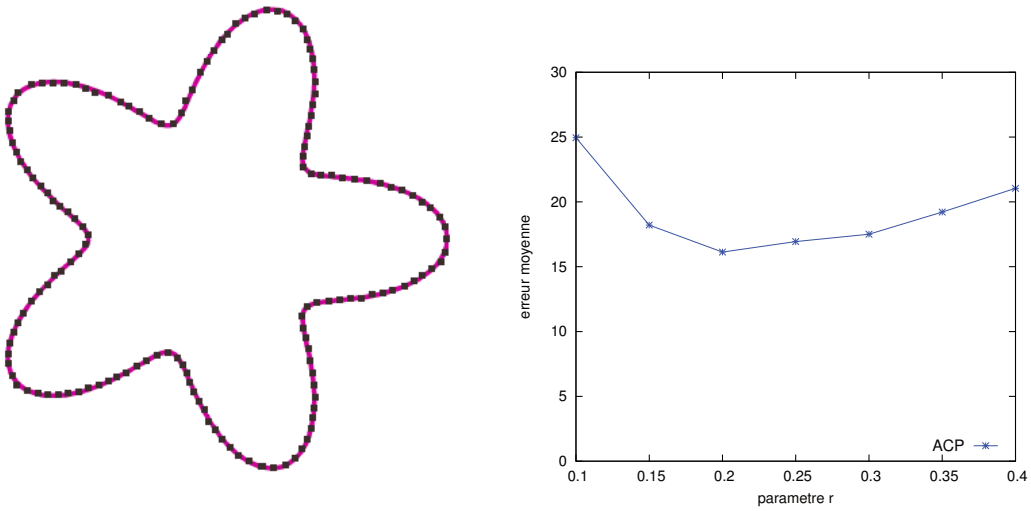
a) b)

FIGURE 2.9 – a) Visualisation du modèle "fleur" d'équation paramétrique polaire $\|p\| = \frac{3}{4} + \frac{1}{4}\cos(5t)$ (diamètre $= 2$) échantillonné avec 160 points avec un bruit Hausdorff 0.1. b) Déviation moyenne (en degrés) des normales ACP en fonction du paramètre de bruit Hausdorff sur la forme précédente.

3.1.1 VALIDATION EN 2D

RÉSISTANCE AU BRUIT. Le Corollaire 2.4 affirme que si P est un nuage de points assez proche d'une courbe C, alors l'estimateur de normale au sens des moindres carrés est proche de la vrai normale. Cette propriété implique une stabilité à

un bruit Hausdorff. Les tests numériques présentés dans la figure 2.9 utilisent le modèle de bruit Hausdorff et permettent en conséquence de confirmer expérimentalement cette stabilité au bruit. La comparaison avec d'autres méthodes est faite dans la dernière section du chapitre 5.



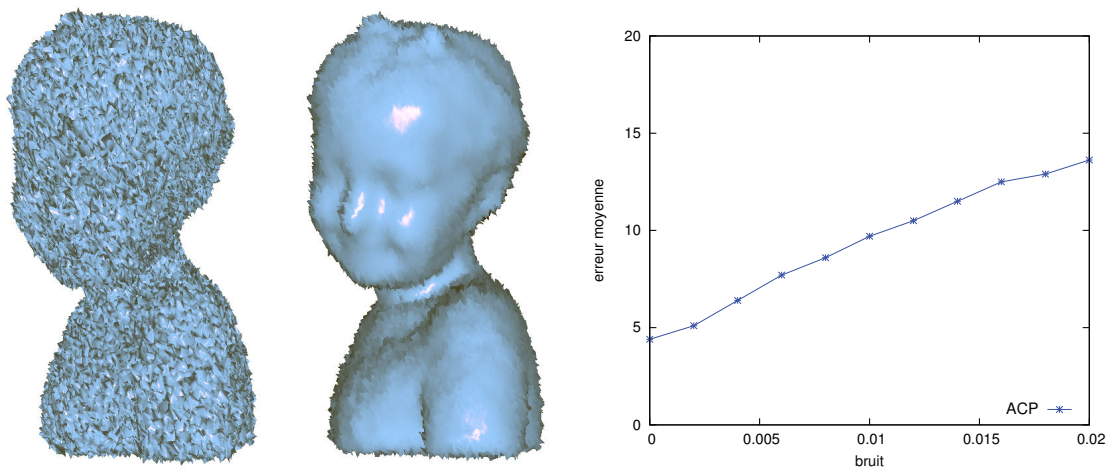
a) b)

FIGURE 2.10 – a) Visualisation du modèle "fleur" d'équation paramétrique polaire $\|p\| = \frac{3}{4} + \frac{1}{4}\cos(5t)$ (diamètre $= 2$) échantillonné avec 160 points non bruité. b) Déviation moyenne (en degrés) des normales ACP en fonction du paramètre r sur la forme précédente.

SENSIBILITÉ AUX PARAMÈTRES. La méthode d'estimation des normales à l'aide de l'analyse en composantes principales nécessite le choix d'un paramètre réel r qui correspond à la taille du voisinage (rayon de la boule) dans lequel la matrice de covariance est calculée. La figure 2.10 présente l'erreur d'estimation moyenne de la normale en fonction du choix du paramètre r. Comme le théorème 2.3 le prévoyait, on observe une certaine dépendance en r. Un paramètre r trop petit entraîne une mauvaise estimation car l'analyse en composantes principales est faite sur un trop petit nombre de points. À ε constant, si r tend vers 0, le calcul de la matrice de covariance est même impossible puisque le nombre de voisins contenu dans la boule de rayon r est alors 0. À l'opposé, un choix de paramètre r trop élevé a tout naturellement tendance à "lisser" les normales et en conséquence à perdre en précision.

3.1.2 HEURISTIQUE SUR LES SURFACES EN 3D

L'essentiel des résultats théoriques obtenus et des expérimentations faites dans cette thèse ont pour cadre général l'approximation de surfaces plongées dans l'espace 3D. Je présente ici des résultats expérimentaux observant la stabilité de l'estimateur de normale ACP dans ce cadre.



a) b) c)

FIGURE 2.11 – a) Visualisation du modèle bimba avec un bruit Hausdorff 0.01. b) Visualisation des normales ACP sur le modèle bimba (diamètre $= 1$) avec un bruit Hausdorff 0.01 (utilisation de l'ombrage de Phong). c) Déviation moyenne (en degrés) des normales en fonction du paramètre de bruit Hausdorff.

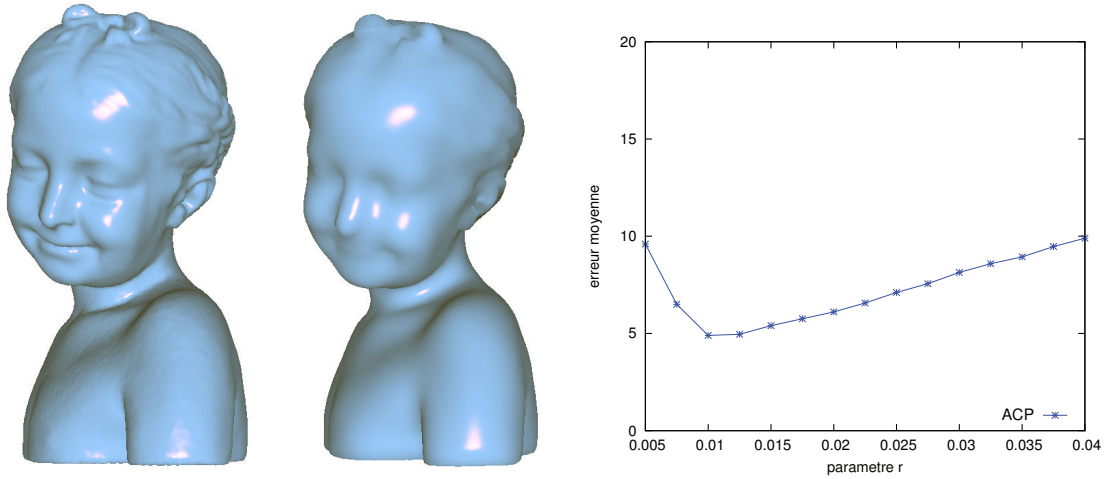
RÉSISTANCE AU BRUIT. La figure 2.11 illustre la résistance au bruit Hausdorff de la normale ACP. Les figures 2.11 a) et b) montrent le résultats de la visualisation du modèle "bimba" avec le même bruit et éclairés respectivement selon les normales à chaque triangle, et selon les normales des moindres carrés. J'utilise un ombrage de Phong sur les normales pour que la visualisation soit plus parlante. La bonne orientation des normales permet ainsi de gommer les imperfections géométriques qui apparaissent alors invisibles à l'œil, excepté sur les contours. Nous remarquons que, malgré la géométrie très bruitée, les normales ACP calculées sont assez bonnes.

SENSIBILITÉ AUX PARAMÈTRES. Les calculs en trois dimensions viennent confirmer la sensibilité de la méthode au choix du paramètre observé dans le cas des courbes du plan. La figure 2.12 illustre cette sensibilité et en particulier le lissage des normales obtenu si le paramètre r est trop grand.

3.2 EXPÉRIMENTATIONS EN GÉOMÉTRIE DISCRÈTE

Dans cette partie, nous validons la convergence multigrille de l'estimateur des normales à travers des expérimentations. Une première série de tests vient confirmer la convergence multigrille de l'estimateur dans le cas d'une courbe en 2 dimension. Une seconde série présente une heuristique montrant la convergence de l'estimateur de normale à une surface dans le cas 3D.

MODÈLE DE BRUIT KANUNGO. Dans le but de confirmer expérimentalement les résultats théoriques faisant intervenir une hypothèse Hausdorff, on définit le modèle de bruit Kanungo sur les contours digitaux. Pour une taille de bruit $t \in \mathbb{N}$ la probabilité qu'un voxel situé à une distance λ de ∂X appartienne au contour bruité est



a) b) c)

FIGURE 2.12 – Visualisation des normales ACP sur le modèle "bimba" non bruité avec $r = 0.01$ (a) et $r = 0.03$ (b). c) Déviation (en degrés) de la normale en fonction du paramètre r sur le modèle bimba non bruité.

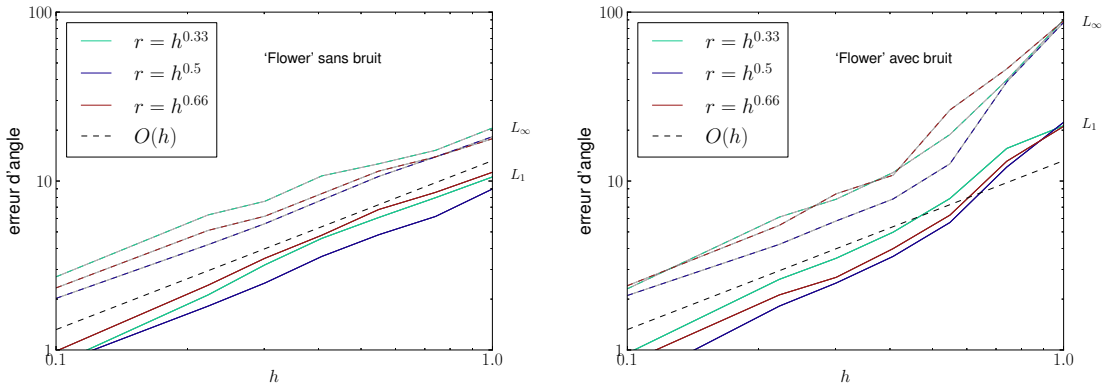


FIGURE 2.13 – **Convergence multigrille en 2D.** Déviation de la normale ACP (en degrés) en fonction du pas h en échelle logarithmique sur le modèle "Flower" non bruité (gauche) et avec un bruit Kanungo de paramètre 0.3 (droite).

t^λ. Ce n'est pas exactement un bruit Hausdorff, mais la plupart des perturbations se situe dans une bande d'épaisseur $2h/(1-t)$. Remarquons que le modèle Kanungo dépend du pas de discrétisation h. Plus précisément, plus le pas de discrétisation est petit, et plus le bruit est faible. Ce modèle est intéressant dans la mesure où il garantit la convergence des contours bruités vers la forme sous-jacente et permet ainsi de faire une étude de convergence multigrille.

OBSERVATIONS DE LA CONVERGENCE MULTIGRILLE. Une première série de données est exposée dans la figure 2.15. Elle confirme expérimentalement la convergence du théorème 2.13. On observe, dans le cas du modèle "Flower" en dimension 2, une convergence plus rapide que $O(h)$. La figure 2.13 b) montre que cette

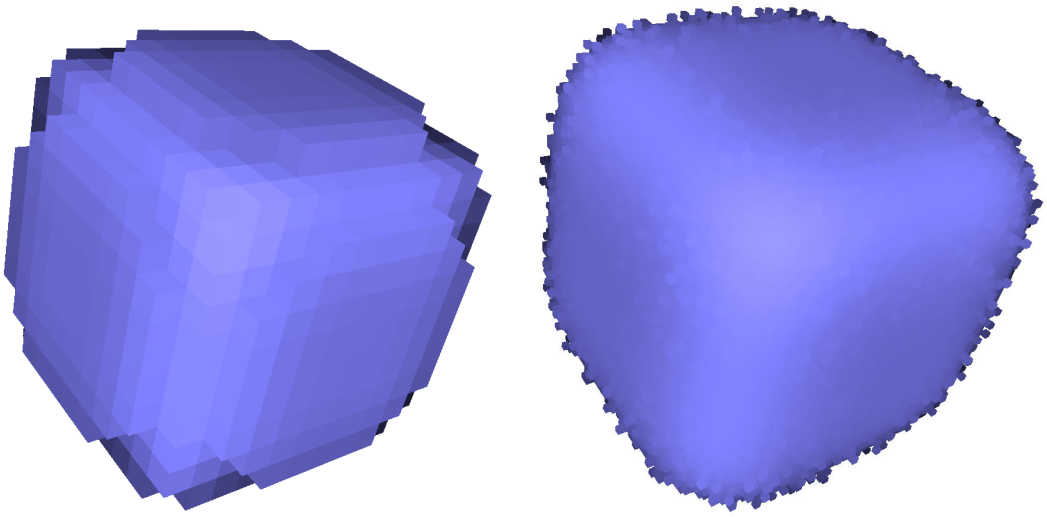


FIGURE 2.14 – **Visualisation des données en 3D.** Visualisation du modèle "Rounded cube" d'équation "$100 - x^4 - y^4 - z^4$" avec les normale théoriques. Avec le paramètres $h = 0.5$ et un bruit Kanungo de taille $t = 0$ (gauche), et $h = 0.1$, $t = 0.3$ (droite).

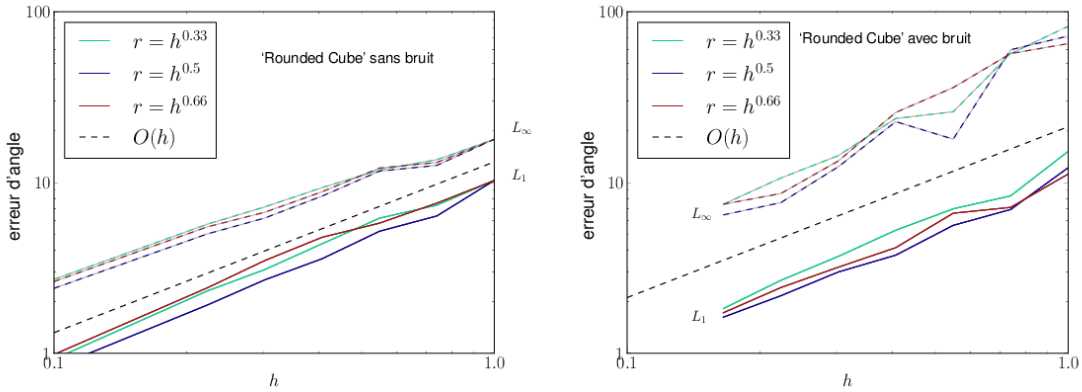


FIGURE 2.15 – **Convergence multigrille en 3D.** Déviation de la normale ACP (en degrés) en fonction du pas h en échelle logarithmique sur le modèle "Rounded Cube" non bruité (gauche) et avec un bruit Kanungo de paramètre 0.3 (droite).

convergence est stable en pratique à une perturbation de type Kanungo de paramètre 0.3. Nous remarquons qu'en présence de bruit, le choix d'un paramètre de rayon r trop faible entraîne une moins bonne précision.

Dans le cadre d'une surface de $\mathbb{R}^3$, nous proposons les résultats des estimateurs de normales au sens des moindres carrés pour le modèle "Rounded Cube". Comme en dimension 2, la présence de bruit ne vient que peu perturber la convergence multigrille.

Cette méthode sera comparée dans le chapitre 5 à la méthode du jet fitting [CP05], et à un algorithme statistique très résistant au bruit et aux *outliers* [BM12].

3.3 Analyse en composantes principales pondérée

Dans cette partie, je présente une heuristique pour améliorer l'estimateur de normale ACP. En ajoutant des poids au nuage de points, le but est de rendre insignifiant les points éloignés du reste du nuage afin de gagner en robustesse au bruit, notamment aux *outliers*.

Description de la méthode. On se place dans un repère orthonormé de $\mathbb{R}^d$. Soit $P = (p_i)_{1\leq i\leq n}$ un nuage de n points de $\mathbb{R}^d$ tels que pour tout entier $i \in \{1, ..., n\}$, p_i a pour coordonnées $(x_{i,j})_{1\leq j\leq d}$. soit $(w_i)_{i\in\{1,...,n\}} \in \mathbb{R}^n$ le vecteur des poids des points de P vérifiant $\sum_{i=1}^n w_i = 1$. Soit $\overline{p}$ le barycentre des points pondérés $(p_i, w_i)_{1\leq i\leq n}$. Le point $\overline{p}$ de coordonnée $(\overline{x}_j)_{1\leq j\leq d}$ vérifie donc : $\overline{x}_j = \sum_{i=1}^n w_i x_{i,j}$.

Définition 2.16. *On appelle normale au sens des moindres carrés pondérés le vecteur* $\mathbf{n}^{ACPP}$ *minimiseur de :*

$$\min_{\|\mathbf{n}\|=1} \sum_i w_i \left((p_i - \overline{p}).\mathbf{n}\right)^2$$

L'estimateur de normale avec analyse en composantes principales classique est facile à calculer puisqu'il correspond à un vecteur propre d'une matrice de covariance. La formulation de la proposition suivante permet l'implémentation de la version pondérée de l'estimateur avec une matrice de covariance pondérée.

Proposition 2.17. *La normale* $\mathbf{n}^{ACPP}$ *est le vecteur propre correspondant à la plus petite valeur propre de la matrice* $M \in \mathcal{M}_{n,n}(\mathbb{R})$ *définie par*

$$M = CC^T$$

où $C \in \mathcal{M}_{n,d}(\mathbb{R})$ *est définie par* $C_{i,j} = \sqrt{w_i}(x_{i,j} - \overline{x})$.

Démonstration. Remarquons que C_i est le vecteur ligne des coordonnées de : $\sqrt{w_i} \times (p_i - \overline{p})$. On a donc

$$\sum_{i=1}^n w_i \left((p_i - \overline{p}).\mathbf{n}\right)^2 = \sum_{i=1}^n (\sqrt{w_i}(p_i - \overline{p}).\mathbf{n})^2 = \sum_{i=1}^n (\sum_{j=1}^n C_{i,j}\mathbf{n}_j)^2 = (C\mathbf{n})^T(C\mathbf{n}).$$

Cette fonction quadratique atteint son minimum quand $\mathbf{n}$ est le vecteur propre de norme 1 correspondant à la plus petite valeur propre.

□

En pratique, j'ai choisi de pondérer les points du nuage à l'aide de la k-distance. La k-distance d'un point x au nuage de points P est la racine carrée de la moyenne des carrés des distance de x aux k plus proches voisins dans P. C'est une notion centrale de la deuxième partie de cette thèse.

Définition 2.18. *Soit* P *un nuage de* n *points de* $\mathbb{R}^d$ *et* $k \in \mathbb{N}$ *tel que* $k \leq n$. *On définit la* k*-distance d'un point* $x \in \mathbb{R}^d$ *au nuage* P *par*

$$d^2_{P,k}(x) = \frac{1}{k} \sum_{y\in \mathrm{N}_{P,k}(x)} d^2(x, y)$$

où $\mathrm{N}_{P,k}(x)$ *est l'ensemble des* k *plus proches voisins de* x *dans* P.

La k-distance est un outil intéressant pour calculer l'éloignement au nuage de points. Cette notion d'éloignement nous est utile dans l'analyse en composantes principales pondérée. Intuitivement, un point qui se trouve loin du reste du nuage doit avoir un poids faible pour qu'il soit moins pris en compte dans le calcul des directions principale. Pour tout point $p_i \in P$, on choisit en conséquence le poids : $\bar{w}_i = \frac{w_i}{\sum_i w_i}$ où $w_i = \frac{1}{d_{P,k}(p_i)}$.

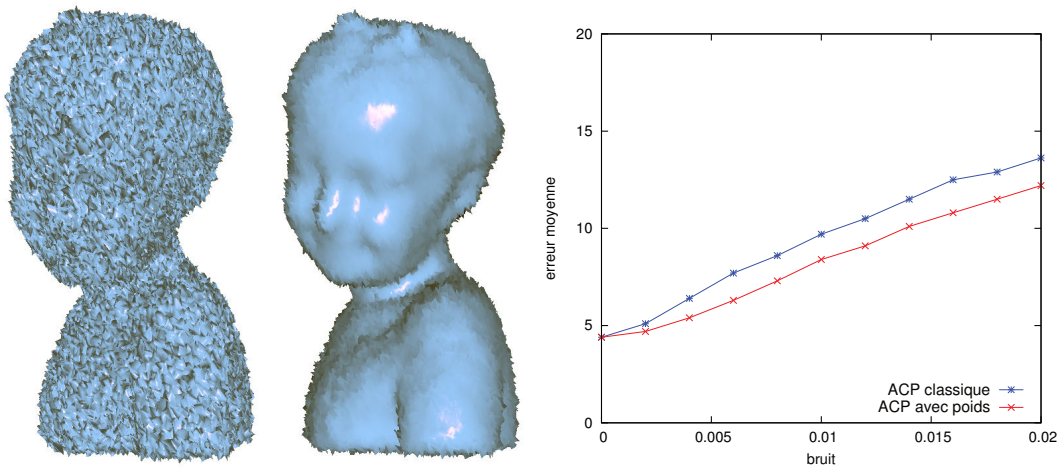


FIGURE 2.16 – Visualisation des normales estimées avec l'ACP pondérée, et moyenne des déviation en degrés de cette normale en fonction du bruit sur le modèle : "bimba".

EXPÉRIMENTATIONS. Cette sous-section a pour but de présenter les résultats expérimentaux de l'estimateur de normale au sens des moindres carrés pondérés. Nous présentons des expériences sur des nuages de points approximant des surfaces plongées dans $\mathbb{R}^3$. Cette méthode a été pensée pour résister davantage au bruit, c'est pourquoi les tests présentés ici se concentrent sur cette stabilité. Nous comparerons cet estimateur avec les autres estimateurs développés dans cette thèse à la fin du chapitre 5. J'ai effectué un grand nombre de tests pour comprendre quel type de poids induisaient de bons résultats. Je ne présente ici que les résultats avec les poids présentés dans le paragraphe précédent définis comme l'inverse de la k-distance car c'est avec ce modèle de poids que les résultats sont les meilleurs.

L'estimateur de normale au sens des moindres carrés pondéré est calculé en suivant le même algorithme que l'estimateur ACP classique, excepté que le nuage de points considéré est pondéré. La première étape est donc de calculer les poids w_i associés aux points p_i à l'aide de la k-distance. Ensuite, on se donne un paramètre de voisinage r et on calcule le vecteur propre associé à la plus petite valeur propre de la matrice définie dans la proposition 2.17.

Les résultats présentés dans la figure 2.16 montrent une amélioration de la déviation des normales et de la visualisation d'autant plus grande que le bruit est grand. Les paramètres choisis pour ces expériences sont constants et valent $k = 10$ et $r = 0.015$.

CONCLUSION

Nous avons étudié la précision et la stabilité de l'estimateur de normale au sens des moindres carrés en dimension 2 et sur deux types de données : les nuages de points et les ensembles digitaux. L'étude théorique a permis d'expliciter la constante dans un résultat de stabilité de l'estimateur ; ce qui nous a permis d'obtenir la convergence multigrille de notre estimateur sur des ensembles de pixels. Les expérimentations sur les deux types de données, en 2 dimensions et en 3 dimensions ont permis de confirmer la précision et la stabilité au bruit Hausdorff.

Chapitre 3

Digital Voronoi Covariance Measure

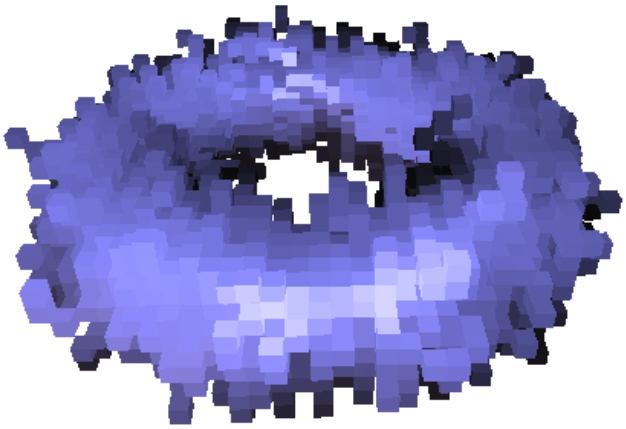


FIGURE 3.1 – Visualisation de la normale estimée avec le VCM sur un tore voxélisé.

Sommaire

1 Voronoi covariance measure 48

1.1 Nuages de points 49

1.2 Surface lisse 49

1.3 Stabilité 49

2 VCM sur des ensembles digitaux 50

2.1 Définition 50

2.2 Convergence multigrille de l'estimateur VCM 50

2.3 Estimation du VCM d'un nuage de points. 51

2.4 Stabilité du VCM. 53

3 Convergence multigrille de l'estimateur de normale 54

4 Expérimentations 57

INTRODUCTION

On propose dans ce chapitre un estimateur de quantités géométriques telles que les normales, les courbures et les arêtes vives pour les surfaces digitales en 3D. Cet estimateur ne dépend que du pas de discrétisation et est une version digitale du *Voronoi Covariance Measure* (VCM) [MOG11].

ESTIMATION GÉOMÉTRIQUE BASÉE VORONOI. Les méthodes d'analyse en composantes principales classiques essaient d'approcher la normale en cherchant localement à approcher la surface par un plan ou un polynôme de degré plus élevé (voir par exemple [MT04]). Les méthodes basées sur le diagramme de Voronoi approchent le cône normal d'une forme sous-jacente, soit géométriquement [AB99] soit en utilisant la covariance des cellules de Voronoi [ACSTD07, MOG11]. Les auteurs de [MOG11] ont défini le *Voronoi Covariance Measure* (VCM) d'un compact de $\mathbb{R}^d$. Ils montrent que le VCM d'une surface lisse contient de l'information géométrique en particulier sur les normales. Ils établissent un résultat de stabilité Hausdorff. Ce résultat de stabilité permet en conséquence d'utiliser le VCM pour estimer les quantités différentielles d'une surface à partir d'une approximation Hausdorff telle qu'un nuage de points ou une triangulation.

CONTRIBUTIONS. Les contributions de ce chapitre sont les suivantes. On définit, dans le cadre de données digitales, un estimateur du VCM qui à chaque surfel du bord d'un objet digital, associe une matrice de covariance. On montre un résultat de convergence multigrille (Section 2, théorème 3.2) de cet estimateur. On en déduit un estimateur de normale (défini comme le vecteur propre associé à la plus grande valeur propre de l'estimateur du VCM) pour lequel on montre la convergence multigrille avec une vitesse de convergence en $O(h^{\frac{1}{8}})$ (Section 3, Corollaire 1). Enfin, on présente des résultats expérimentaux montrant que la vitesse de convergence est proche de $O(h)$ en pratique (Section 4). Les expérimentations indiquent que l'estimateur VCM peut être aussi utilisé pour estimer la courbure et détecter les arêtes vives dans le cas de données digitales perturbées par un bruit Hausdorff.

1 VORONOI COVARIANCE MEASURE

On rappelle dans cette section la définition et les propriétés du *Voronoi covariance measure* (VCM). Le VCM a été introduit dans [MOG11] pour l'estimation des normales et des courbures. Soient K un compact de $\mathbb{R}^d$, d_K la *fonction distance* à K et $\mathfrak{p}_K : \mathbb{R}^d \to K$ la projection sur K. On sait par la proposition 1.3 du chapitre 1 que la projection est définie $\mathcal{H}^d$-presque partout. On rappelle que pour tout réel positif R, l'ensemble K^R est le R-offset de K. Le VCM associe à toute fonction intégrable $\chi : \mathbb{R}^d \to \mathbb{R}^+$, la matrice

$$\mathcal{V}_{K,R}(\chi) := \int_{K^R} (x - \mathfrak{p}_K(x)).(x - \mathfrak{p}_K(x))^t.\chi(\mathfrak{p}_K(x))\mathrm{d}x.$$

1.1 Nuages de points

Dans le cas où le compact K est un nuage de points P, le VCM fait naturellement intervenir le diagramme de Voronoi de P. La cellule de Voronoi d'un point p de P est l'ensemble $\mathrm{Vor}_P(p) := \{x \in \mathbb{R},\ \forall q \in P\ \|x-p\| \leq \|x-q\|\}$, et $\mathfrak{p}_K$ envoie ($\mathcal{H}^d$-presque) tout point x sur le centre de sa cellule de Voronoi. En conséquence,

$$\mathcal{V}_{K,R}(\chi) = \sum_{p\in P} \chi(p) \int_{\mathrm{Vor}_P(p)\cap P^R} (x-p).(x-p)^t \mathrm{d}x.$$

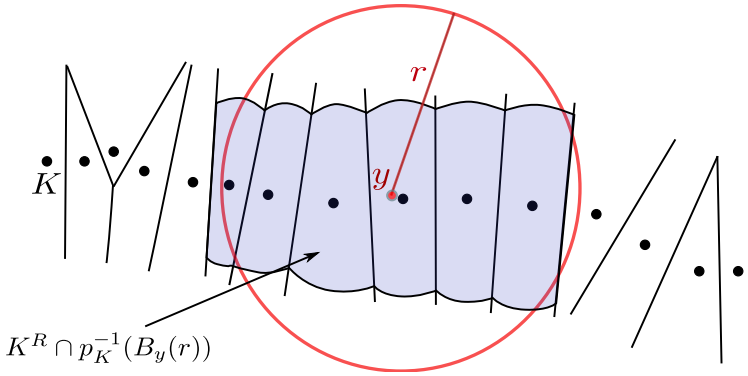


FIGURE 3.2 – Domaine d'intégration du VCM

1.2 Surface lisse

Le VCM d'une surface lisse S de $\mathbb{R}^3$ contient de l'information sur sa normale $\mathbf{n}_S$ [MOG11]. En appliquant le VCM à une fonction indicatrice de la boule $\mathrm{B}_p(r)$ centrée en un point p situé sur la surface donne quand $r \to 0$,

$$\mathcal{V}_{S,R}(\chi) \approx \frac{2\pi}{3} R^3 r^2 \left[\mathbf{n}_S(p).\mathbf{n}_S(p)^t\right]. \tag{3.1}$$

1.3 Stabilité

Le résultat de stabilité suivant [MOG11] implique que le VCM est stable aux perturbations Hausdorff. L'équation (3.1) implique alors que les normales peuvent être estimées à l'aide du VCM d'un nuage de point approchant une surface. Une heuristique est utilisé pour estimer d'autres quantités géométriques telles que les courbures, les directions principales ou pour détecter des arêtes [MOG11]. On note $\|.\|_{\mathrm{op}}$ la norme matricielle induite par la norme euclidienne. Soit une fonction $f : \mathbb{R}^d \to \mathbb{R}$, on note $\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|$ et $\mathrm{Lip} f = \sup_{x\neq y} |f(x)-f(y)|/\|x-y\|$ sa constante de Lipschitz, qui peut être infinie. On définit la norme *bounded-Lipschitz* par $\|f\|_{\mathrm{BL}} = \|f\|_\infty + \mathrm{Lip}(f)$.

Théorème 3.1. *[MOG11] Soit χ une fonction bornée et Lipschitzienne, K un compact de $\mathbb{R}^d$ et $R > 0$. Alors pour tout compact K', on a*

$$\|\mathcal{V}_{K,R}(\chi) - \mathcal{V}_{K',R}(\chi)\|_{\mathrm{op}} \leq C_0 \|\chi\|_{\mathrm{BL}} \mathrm{d}_H(K,K')^{\frac{1}{2}},$$

où la constante C ne dépend que de K, R et de la dimension.

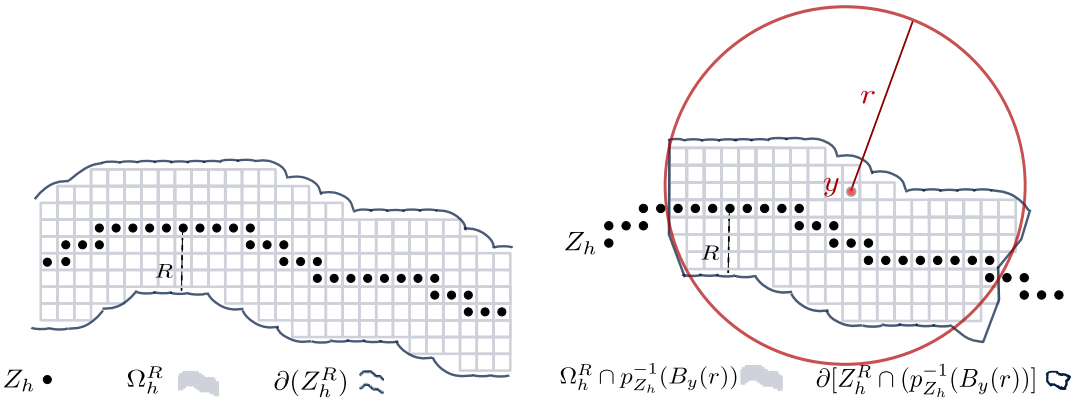


FIGURE 3.3 – La discrétisation d'un offset et sa localisation.

Dans le chapitre 5, nous généraliserons ce résultat à une plus grande classe d'estimateurs faisant intervenir des fonctions distances robustes. En particulier, nous quantifierons la constante C intervenant dans le résultat précédent et montrerons qu'elle ne dépend que de R et du diamètre de K.

2 VCM SUR DES ENSEMBLES DIGITAUX

Dans cette section, on définit un estimateur du VCM dans le cas de données digitales 3D. Le théorème 3.2 explicite les conditions sous lesquelles cet estimateur est convergent multigrille.

2.1 DÉFINITION

Soit X un domaine compact de $\mathbb{R}^3$ dont le bord est une surface de classe $\mathcal{C}^2$. Comme dans le chapitre 2, on note ∂X le bord de X, $Dig_h(X) = \mathrm{vox}(X \cap (h\mathbb{Z})^3)$ la discrétisation de Gauss de X, et $\partial_h X \subset \mathbb{R}^3$ l'ensemble des surfels du bord de $Dig_h(X)$. On définit le nuage de points $Z_h = \partial_h X \cap h(\mathbb{Z} + \frac{1}{2})^3$. On sait d'après le lemme 2.11 que Z_h est proche au sens de Hausdorff de ∂X. On définit enfin l'estimateur digital du VCM par l'application qui à toute fonction mesurable $\chi : \mathbb{R}^3 \to \mathbb{R}$ associe la matrice

$$\widehat{\mathcal{V}}_{Z_h,R}(\chi) := \sum_{x \in \Omega_h^R} h^3 (x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^t \chi(\mathfrak{p}_{Z_h}(x)),$$

où $\Omega_h^R = \{x \in Z_h^R \cap h(\mathbb{Z} + \frac{1}{2})^3, \mathrm{vox}(x) \subset Z_h^R\}$ est l'ensemble des centres des voxels entièrement contenus dans Z_h^R, le R-offset de Z_h (voir Fig. 3.3).

2.2 CONVERGENCE MULTIGRILLE DE L'ESTIMATEUR VCM

Le principal résultat théorique de ce chapitre est le théorème suivant. Intuitivement, il quantifie l'erreur dans l'approximation du VCM d'une surface lisse par le VCM digital de sa discrétisation de Gauss.

Théorème 3.2. *Soit X un domaine compact de $\mathbb{R}^3$ dont le bord ∂X est une surface de classe $\mathcal{C}^2$ et de reach $\rho > 0$. Soit $R < \frac{\rho}{2}$ et $\chi : \mathbb{R}^3 \to \mathbb{R}^+$ une fonction intégrable*

dont le support est contenu dans une boule de rayon r. Pour tout $h > 0$ tel que $h \leq \min\left(\frac{1}{\sqrt{3}}R, \frac{1}{2\sqrt{3}}r, \frac{1}{32\sqrt{3}\rho}r^2\right)$, on a

$$\begin{aligned}\left\|\mathcal{V}_{\partial X,R}(\chi) - \widehat{\mathcal{V}}_{Z_h,R}(\chi)\right\|_{\mathrm{op}} &= O\Big(\mathrm{Lip}(\chi) \times [(r^3R^{\frac{5}{2}} + r^2R^3 + rR^{\frac{9}{2}})h^{\frac{1}{2}}] \\ &\quad + \|\chi\|_\infty \times [(r^3R^{\frac{3}{2}} + r^2R^2 + rR^{\frac{7}{2}})h^{\frac{1}{2}} + r^2Rh]\Big).\end{aligned}$$

Dans le théorème et dans la suite du chapitre, les constantes intervenant dans la notation $O(.)$ ne dépendent que du *reach* de ∂X et de la dimension (qui est égale à 3 ici). Pour la preuve du théorème 3.2, on utilise le VCM d'un nuage de points Z_h, c'est à dire $\mathcal{V}_{Z_h,R}(\chi)$. Par l'inégalité triangulaire, on a

$$\left\|\mathcal{V}_{\partial X,R}(\chi) - \widehat{\mathcal{V}}_{Z_h,R}(\chi)\right\|_{\mathrm{op}} \leq \|\mathcal{V}_{\partial X,R}(\chi) - \mathcal{V}_{Z_h,R}(\chi)\|_{\mathrm{op}} + \left\|\mathcal{V}_{Z_h,R}(\chi) - \widehat{\mathcal{V}}_{Z_h,R}(\chi)\right\|_{\mathrm{op}}.$$

La preuve du théorème découle alors directement des propositions 3.3 et 3.4 qui bornent respectivement le second et le premier terme du membre de droite. On pose $\varepsilon = \sqrt{3}h$ qui est une borne de la distance de Hausdorff entre Z_h et ∂X (lemme 2.11).

2.3 Estimation du VCM d'un nuage de points.

On rappelle les hypothèses du théorème. L'ensemble X est un domaine compact de $\mathbb{R}^3$ dont le bord ∂X est une surface de classe $\mathcal{C}^2$ de *reach* $\rho > 0$. On a $R < \frac{\rho}{2}$ et $\chi : \mathbb{R}^3 \to \mathbb{R}^+$ est une fonction intégrable dont le support est contenu dans une boule $\mathrm{B}_y(r)$ de centre y et de rayon r.

Proposition 3.3.
Sous les hypothèses du théorème 3.2, on a

$$\left\|\mathcal{V}_{Z_h,R}(\chi) - \widehat{\mathcal{V}}_{Z_h,R}(\chi)\right\|_{\mathrm{op}} = O\left[r^2R^2(\mathrm{Lip}(\chi)R + \|\chi\|_\infty)\, h^{\frac{1}{2}} + r^2R\|\chi\|_\infty h\right].$$

Démonstration.
Étape 1 : Le but de cette étape est de prouver que

$$\mathcal{V}_{Z_h,R}(\chi) = \int_{\mathrm{vox}(\Omega_h^R)} (x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))\mathrm{d}x + R^2\|\chi\|_\infty O(hr^2).$$

Comme $\mathrm{vox}(\Omega_h^R) \subset Z_h^R$, on a

$$\begin{aligned}\mathcal{V}_{Z_h,R}(\chi) &= \textstyle\int_{\mathrm{vox}(\Omega_h^R)}(x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))\mathrm{d}x \\ &+ \textstyle\int_{Z_h^R\setminus\mathrm{vox}(\Omega_h^R)}(x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))\mathrm{d}x\end{aligned}$$

En utilisant le fait que $\|x - \mathfrak{p}_{Z_h}(x)\| \leq R$, que χ est borné par $\|\chi\|_\infty$, et que le support de χ est contenu dans la boule $\mathrm{B}_y(r)$ (voir figure 3.3), le second terme de l'équation est borné par

$$R^2 \times \|\chi\|_\infty \times \mathcal{H}^3\left(\left[Z_h^R\setminus\mathrm{vox}(\Omega_h^R)\right] \cap \mathfrak{p}_{Z_h}^{-1}(\mathrm{B}_y(r))\right).$$

Montrons maintenant que $Z_h^R \cap \mathfrak{p}_{Z_h}^{-1}(\mathrm{B}_y(r)) \subset \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$. En effet, soit $x \in Z_h^R \cap \mathfrak{p}_{Z_h}^{-1}(\mathrm{B}_y(r))$. Le fait que la distance de Hausdorff entre Z_h et ∂X soit bornée par ε implique que $x \in \partial X^{R+\varepsilon}$. Comme $\varepsilon = \sqrt{3}h \leq R$, le lemme 1.16 (page 22) implique que $\|\mathfrak{p}_{\partial X}(x) - \mathfrak{p}_{Z_h}(x)\| \leq \sqrt{8\varepsilon\rho} + \varepsilon$, ce qui donne

$$\|\mathfrak{p}_{\partial X}(x) - y\| \leq \|\mathfrak{p}_{\partial X}(x) - \mathfrak{p}_{Z_h}(x)\| + \|\mathfrak{p}_{Z_h}(x) - y\| \leq \sqrt{8\varepsilon\rho} + \varepsilon + r \leq 2r.$$

Établissons maintenant que $Z_h^R \backslash \mathrm{vox}(\Omega_h^R) \subset \partial X^{R+\varepsilon} \backslash \partial X^{R-(\sqrt{3}+1)\varepsilon}$. En effet, comme nous l'avons dit précédemment, on a $Z_h^R \subset \partial X^{R+\varepsilon}$. De plus, si $x \in \partial X^{R-(\sqrt{3}+1)\varepsilon}$, alors le fait que la distance de Hausdorff entre Z_h et ∂X est inférieur à ε implique que $x \in Z_h^{R-\sqrt{3}\varepsilon}$. Soit $c \in h(\mathbb{Z} + \frac{1}{2})^3$ le centre du voxel contenant x. Le fait que $\mathrm{diam}(\mathrm{vox}(c)) = \sqrt{3}h = \varepsilon$ implique que $\mathrm{vox}(c) \subset Z_h^R$, et donc $x \in \mathrm{vox}(\Omega_h^R)$. On obtient donc $Z_h^R \backslash \mathrm{vox}(\Omega_h^R) \subset \partial X^{R+\varepsilon} \backslash \partial X^{R-(\sqrt{3}+1)\varepsilon}$. On déduit finalement que

$$\left[Z_h^R \backslash \mathrm{vox}(\Omega_h^R)\right] \cap \mathfrak{p}_Z^{-1}(\mathrm{B}_y(r)) \subset \left[\partial X^{R+3\varepsilon} \backslash \partial X^{R-3\varepsilon}\right] \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r)), \tag{3.2}$$

dont le volume est borné par $O(\varepsilon r^2) = O(hr^2)$ par la proposition 1.10.b, ce qui permet de conclure l'étape 1.

Étape 2 : Nous devons maintenant borner le terme restant :

$$\Delta = \int_{\mathrm{vox}(\Omega_h^R)} (x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}} \chi(\mathfrak{p}_{Z_h}(x)) \mathrm{d}x - \widehat{\mathcal{V}}_{Z_h,R}(\chi).$$

En décomposant Δ sur l'ensemble des voxels de $\mathrm{vox}(\Omega_h^R)$, on a

$$\begin{aligned}\Delta = \sum_{c \in \Omega_h^R} \int_{\mathrm{vox}(c)} & \big[(x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}} \chi(\mathfrak{p}_{Z_h}(x)) \\ & -(c - \mathfrak{p}_{Z_h}(c))(c - \mathfrak{p}_{Z_h}(c))^{\mathbf{t}} \chi(\mathfrak{p}_{Z_h}(c))\big] \, \mathrm{d}x\end{aligned}$$

Comme dans l'étape 1, on peut localiser le calcul autour du support de χ. On introduit l'ensemble des centres $D = \Omega_h^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$. On a l'égalité $\Delta = \Delta_1 + \Delta_2$, avec

$$\begin{aligned}\Delta_1 &= \sum_{c \in D} \int_{\mathrm{vox}(c)} (x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}} [\chi(\mathfrak{p}_{Z_h}(x)) - \chi(\mathfrak{p}_{Z_h}(c))] \mathrm{d}x \\ \Delta_2 &= \sum_{c \in D} \int_{\mathrm{vox}(c)} [(x - \mathfrak{p}_{Z_h}(x))(x - \mathfrak{p}_{Z_h}(x))^{\mathbf{t}} - (c - \mathfrak{p}_{Z_h}(c))(c - \mathfrak{p}_{Z_h}(c))^{\mathbf{t}}] \chi(\mathfrak{p}_{Z_h}(c))\end{aligned}$$

Nous allons maintenant borner Δ_1 et Δ_2. On a

$$\|\Delta_1\|_{\mathrm{op}} \leq \sum_{c \in D} \int_{\mathrm{vox}(c)} \|x - \mathfrak{p}_{Z_h}(x)\| \left\|x - \mathfrak{p}_{Z_h}(x)^{\mathbf{t}}\right\| \|\chi(\mathfrak{p}_{Z_h}(x)) - \chi(\mathfrak{p}_{Z_h}(c))\| \, \mathrm{d}x.$$

Pour tout $c \in D$ et $x \in \mathrm{vox}(c)$, on a $\|x - c\| \leq \frac{\sqrt{3}}{2}h$. De plus, par définition de Ω_h^R, on a que x et c appartiennent à $Z_h^R \subset \partial X^{R+\varepsilon}$. Donc, comme $h \leq R \leq \frac{\rho}{2}$, le théorème 1.14 (page 21) implique $\|\mathfrak{p}_{Z_h}(x) - \mathfrak{p}_{Z_h}(c)\| = O(\varepsilon^{\frac{1}{2}} + \varepsilon + h) = O(h^{\frac{1}{2}})$ et donc $\|\chi(\mathfrak{p}_{Z_h}(x)) - \chi(\mathfrak{p}_{Z_h}(c))\| = \mathrm{Lip}(\chi) O(h^{\frac{1}{2}})$. En utilisant le fait que $\|x - \mathfrak{p}_{Z_h}(x)\| \leq R$, on a

$$\|\Delta_1\|_{\mathrm{op}} = \mathcal{H}_3(\mathrm{vox}(D)) \times R^2 \times \mathrm{Lip}(\chi) \times O(h^{\frac{1}{2}}).$$

Comme $\mathrm{vox}(D) \subset Z_h^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r)) \subset \partial X^{R+\varepsilon} \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$ et $\varepsilon \leq R$, la proposition 1.10.a implique que $\mathcal{H}_3(\mathrm{vox}(D)) = O(r^2 R)$. Cela donne $\|\Delta_1\|_{\mathrm{op}} = \mathrm{Lip}(\chi) \times O(r^2 R^3 h^{\frac{1}{2}})$.

Bornons de manière similaire $\|\Delta_2\|_{\mathrm{op}}$. On pose $u = (x - c)$, $v = c - \mathfrak{p}_{Z_h}(c)$ et $w = \mathfrak{p}_{Z_h}(c) - \mathfrak{p}_{Z_h}(x)$. On peut écrire $x - \mathfrak{p}_{Z_h}(x) = u + v + w$, et on obtient

$$\Delta_2 = \sum_{c \in D} \left[\int_{\mathrm{vox}(c)} [(u+v+w)(u+v+w)^t - vv^t] \chi(\mathfrak{p}_{Z_h}(c)) \right].$$

Grâce aux inégalités $\|u\| \leq h$, $\|v\| \leq R$ et $\|w\| = O(h^{\frac{1}{2}})$, on borne le terme dans l'intégrale par $O(\|\chi\|_\infty (R\, h^{\frac{1}{2}} + h))$. Et en utilisant $\mathcal{H}_3(\mathrm{vox}(D)) = O(r^2 R)$, on obtient $\|\Delta_2\|_{\mathrm{op}} = O(\|\chi\|_\infty\, (R^2 r^2 h^{\frac{1}{2}} + r^2 R h))$. □

2.4 Stabilité du VCM.

Il a été montré dans [MOG11] que le VCM d'un nuage de points est Hausdorff stable. Plus précisément, le théorème 5.1 de [MOG11] établit que

$$\|\mathcal{V}_{\partial X,R}(\chi) - \mathcal{V}_{Z_h,R}(\chi)\|_{\mathrm{op}} = O(h^{\frac{1}{2}}).$$

Cependant, la constante intervenant dans $O(h^{\frac{1}{2}})$ dépend de la surface ∂X. Nous voulons ici une constante plus précise faisant intervenir des quantités locales, r et R. La preuve est similaire à celle de [MOG11], excepté que nous localisons les calculs.

Proposition 3.4. *Sous les hypothèses du théorème 3.2, on a*

$$\begin{aligned}
&\|\mathcal{V}_{\partial X,R}(\chi) - \mathcal{V}_{Z_h,R}(\chi)\|_{\mathrm{op}} \\
&= O\Big(Lip(\chi) \times [(r^3 R^{\frac{5}{2}} + r^2 R^{\frac{7}{2}} + r R^{\frac{9}{2}}) h^{\frac{1}{2}}] + \|\chi\|_\infty \times [(r^3 R^{\frac{3}{2}} + r^2 R^{\frac{5}{2}} + r R^{\frac{7}{2}}) h^{\frac{1}{2}}] \Big).
\end{aligned}$$

Démonstration. Comme dans la preuve de la proposition précédente le fait que $h \leq \min\left(\frac{1}{\sqrt{3}} R, \frac{1}{2\sqrt{3}} r, \frac{r^2}{32\sqrt{3}\rho}\right)$ implique que $\mathfrak{p}_{Z_h}^{-1}(\mathrm{supp}(\chi)) \subset \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$. On introduit donc l'ensemble $E = \partial X^{R-\varepsilon} \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$, sur lequel nous allons intégrer. on a

$$\begin{aligned}
\mathcal{V}_{\partial X,R}(\chi) &= \int_{\partial X^R} (x - \mathfrak{p}_{\partial X}(x))(x - \mathfrak{p}_{\partial X}(x))^{\mathbf{t}} \chi(\mathfrak{p}_{\partial X}(x)) \\
&= \int_{\partial X^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))} (x - \mathfrak{p}_{\partial X}(x))(x - \mathfrak{p}_{\partial X}(x))^{\mathbf{t}} \chi(\mathfrak{p}_{\partial X}(x)) \\
&= \int_E (x - \mathfrak{p}_{\partial X}(x))(x - \mathfrak{p}_{\partial X}(x))^{\mathbf{t}} \chi(\mathfrak{p}_{\partial X}(x)) + Err_1,
\end{aligned}$$

où l'erreur Err_1 satisfait

$$\|Err_1\|_{\mathrm{op}} \leq R^2 \times \|\chi\|_\infty \times \mathcal{H}_3(\partial X^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r)) \setminus E).$$

De plus, on a $\partial X^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r)) \setminus E = \left[\partial X^R \setminus \partial X^{R-\varepsilon}\right] \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$, dont le volume est borné par la proposition 1.10.b par $O(r^2 \varepsilon) = O(r^2 h)$. On a donc

$$\|Err_1\|_{\mathrm{op}} = \|\chi\|_\infty \times O(R^2 r^2 h).$$

Similairement, on a

$$\begin{aligned}\mathcal{V}_{Z_h,R}(\chi) &= \int_{Z_h^R}(x-\mathfrak{p}_{Z_h}(x))(x-\mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))\\ &= \int_{Z_h^R\cap\mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))}(x-\mathfrak{p}_{Z_h}(x))(x-\mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))\\ &= \int_E(x-\mathfrak{p}_{Z_h}(x))(x-\mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))+Err_2,\end{aligned}$$

où l'erreur Err_2 satisfait

$$\|Err_2\|_{\mathrm{op}} \leq (R+\varepsilon)^2 \times \|\chi\|_\infty \times \mathcal{H}_3(Z_h^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r)) \setminus E).$$

Comme l'ensemble $Z_h^R \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r)) \setminus E \subset \left[\partial X^{R+\varepsilon}\backslash\partial X^{R-\varepsilon}\right] \cap \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(2r))$, son volume est borné par $O(r^2\varepsilon) = O(r^2h)$, on a

$$\|Err_2\|_{\mathrm{op}} = (R+\varepsilon)^2 \times \|\chi\|_\infty \times O(r^2h).$$

Il nous faut maintenant comparer les deux intégrales sur l'ensemble commun E

$$\Delta = \int_E \left[(x-\mathfrak{p}_{\partial X}(x))(x-\mathfrak{p}_{\partial X}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{\partial X}(x)) - (x-\mathfrak{p}_{Z_h}(x))(x-\mathfrak{p}_{Z_h}(x))^{\mathbf{t}}\chi(\mathfrak{p}_{Z_h}(x))\right].$$

En suivant la preuve du théorème 5.1 de [MOG11], on a

$$\|\Delta\|_{\mathrm{op}} \leq (R^2\mathrm{Lip}(\chi)+2R\|\chi\|_\infty)\times[\mathcal{H}_3(E)+(diam(E)+R+\sqrt{R\varepsilon})\times\mathcal{H}^2(\partial E)]\times\sqrt{R\varepsilon}.$$

Les propriétés b. et c. de la proposition 1.10 donnent respectivement que $\mathcal{H}_3(E)$ est borné par $O(r^2R)$ et que $\mathcal{H}^2(\partial E)$ est borné par $O(rR+r^2)$. Comme $h = \frac{1}{\sqrt{3}}\varepsilon$, on a

$$\|\Delta\|_{\mathrm{op}} = O\Big(\mathrm{Lip}(\chi)\times[(r^3R^{\frac{5}{2}}+r^2R^{\frac{7}{2}}+rR^{\frac{9}{2}})h^{\frac{1}{2}}]+\|\chi\|_\infty\times[(r^3R^{\frac{3}{2}}+r^2R^{\frac{5}{2}}+rR^{\frac{7}{2}})h^{\frac{1}{2}}]\Big).$$

Comme $\|Err_1\|_{\mathrm{op}}$ et $\|Err_2\|_{\mathrm{op}}$ sont négligeable devant $\|\Delta\|_{\mathrm{op}}$, on obtient la borne :

$$\begin{aligned}&\|\mathcal{V}_{\partial X,R}(\chi)-\mathcal{V}_{Z_h,R}(\chi)\|_{\mathrm{op}}\\ &= O\Big(Lip(\chi)\times[(r^3R^{\frac{5}{2}}+r^2R^{\frac{7}{2}}+rR^{\frac{9}{2}})h^{\frac{1}{2}}]+\|\chi\|_\infty\times[(r^3R^{\frac{3}{2}}+r^2R^{\frac{5}{2}}+rR^{\frac{7}{2}})h^{\frac{1}{2}}]\Big).\end{aligned}$$

□

3 CONVERGENCE MULTIGRILLE DE L'ESTIMATEUR DE NORMALE

Soit X un domaine compact de $\mathbb{R}^3$ dont le bord ∂X est une surface de classe $\mathcal{C}^2$ de *reach* ρ. On souhaite ici estimer la normale à ∂X à partir de sa discrétisation de Gauss. Si $h \leq \frac{1}{\sqrt{3}}\rho$, le lemme 2.11 (page 33) implique que la distance de Hausdorff $d_H(Z_h, \partial X)$ est majoré strictement par le *reach* ρ. La projection $\mathfrak{p}_{\partial X}$ est donc défini sur Z_h. Soit $\widehat{p_0}$ un point de Z_h et p_0 la projection de $\widehat{p_0}$ sur ∂X. On définit l'estimateur de normale en appliquant le VCM digital à une fonction Lipschitzienne qui approche la fonction indicatrice de la boule $\mathrm{B}_{\widehat{p_0}}(r)$. Pour tout point p, soit $\chi_{r,p}$ une fonction Lipschitzienne qui vaut 1 sur $\mathrm{B}_p(r)$, vaut $1-(\|x-p\|-r)/r^{\frac{3}{2}}$ sur $\mathrm{B}_p(r+r^{\frac{3}{2}}) \setminus \mathrm{B}_p(r)$, et vaut 0 sur le domaine restant.

Définition 3.5. *L'* estimateur de normale $\widehat{n}_{r,R}(\widehat{p_0})$ *est le vecteur propre associé à la plus grande valeur propre de* $\widehat{\mathcal{V}}_{Z_h,R}(\chi_{r,\widehat{p_0}})$.

Remarquons que l'estimateur de normale est défini au signe près. Le théorème suivant donne une erreur d'estimation entre $\pm\widehat{n}_{r,R}(\widehat{p_0})$ et $\mathbf{n}_S(p_0)$.

Théorème 3.6. *Soit X un domaine compact de $\mathbb{R}^3$ dont le bord ∂X est une surface de classe $\mathcal{C}^2$ et de reach $\rho > 0$. Soit $R < \frac{\rho}{2}$. Donc pour tout $h > 0$ tel que $h \leq \min\left(\frac{1}{\sqrt{3}}R, \frac{1}{2\sqrt{3}}r, \frac{r^2}{32\sqrt{3}\rho}\right)$, l'angle entre les lignes portées par $\widehat{n}_{r,R}(\widehat{p_0})$ et $\mathbf{n}_S(p_0)$ satisfait*

$$\begin{aligned}&\langle \widehat{n}_{r,R}(\widehat{p_0}), \mathbf{n}_S(p_0)\rangle = \\ &O\Big((rR^{-\frac{3}{2}} + R^{-1} + r^{-\frac{1}{2}}R^{-\frac{1}{2}} + r^{-\frac{3}{2}} + r^{-\frac{5}{2}}R^{\frac{3}{2}})h^{\frac{1}{2}} + R^{-2}h + r^{\frac{3}{2}}R^{-1} + R^2 + r^{\frac{1}{2}}\Big).\end{aligned}$$

Le corollaire suivant est une conséquence directe du théorème.

Corollaire 3.7. *Soit X un domaine compact de $\mathbb{R}^3$ dont le bord ∂X est une surface de classe $\mathcal{C}^2$ de reach $\rho > 0$. Soient $a, b \in \mathbb{R}^+$, $r = ah^{\frac{1}{4}}$ et $R = bh^{\frac{1}{4}}$. Alors pour tout $h > 0$ assez petit, on a*

$$\langle \widehat{n}_{r,R}(\widehat{p_0}), \mathbf{n}_S(p_0)\rangle = O\Big(h^{\frac{1}{8}}\Big).$$

Démonstration du théorème 3.6. On introduit le VCM normalisé $\widehat{N}_{r,R}(\widehat{p_0}) = \frac{3}{2\pi r^2 R^3}\widehat{\mathcal{V}}_{Z_h,R}(\chi_{r,\widehat{p_0}})$. D'après le théorème $\sin(\theta)$ de Davis-Kahan [Dav63], au signe près de $\pm\widehat{n}_{r,R}(p_0)$, on a

$$\|\widehat{n}_{r,R}(\widehat{p_0}) - \mathbf{n}_S(p_0)\| \leq 2\|\widehat{N}_{r,R}(\widehat{p_0}) - \mathbf{n}_S(p_0)\mathbf{n}_S(p_0)^t\|_{\mathrm{op}}.$$

Majorons le membre de droite. L'inégalité triangulaire donne

$$\begin{aligned}\|\widehat{N}_{r,R}(\widehat{p_0}) - \mathbf{n}_S(p_0)\mathbf{n}_S(p_0)^t\|_{\mathrm{op}} \leq\ & \frac{3}{2\pi R^3 r^2}\|\widehat{\mathcal{V}}_{Z_h,R}(\chi_{r,\widehat{p_0}}) - \mathcal{V}_{\partial X,R}(\chi_{r,\widehat{p_0}})\|_{\mathrm{op}} \\ &+ \frac{3}{2\pi R^3 r^2}\|\mathcal{V}_{\partial X,R}(\chi_{r,\widehat{p_0}}) - \mathcal{V}_{\partial X,R}(\chi_{r,p_0})\|_{\mathrm{op}} \\ &+ \frac{3}{2\pi R^3 r^2}\|\mathcal{V}_{\partial X,R}(\chi_{r,p_0}) - \mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0(r)}})\|_{\mathrm{op}} \\ &+ \|\frac{3}{2\pi R^3 r^2}\mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0(r)}}) - \mathbf{n}_S(p_0)\mathbf{n}_S(p_0)^t\|_{\mathrm{op}}.\end{aligned}$$

La preuve du théorème est alors une conséquence du théorème 3.2, qui contrôle le premier terme, et des trois lemmes suivants qui contrôlent les autres termes.

Lemme 3.8. *Sous les hypothèses du théorème 3.6, on a*

$$\frac{3}{2\pi R^3 r^2}\|\mathcal{V}_{\partial X,R}(\chi_{r,\widehat{p_0}}) - \mathcal{V}_{\partial X,R}(\chi_{r,p_0})\|_{\mathrm{op}} \leq O(R^{-1}(h + r^{\frac{3}{2}}))$$

Démonstration. On décompose le domaine d'intégration

$$\mathrm{supp}(\chi_{r,\widehat{p_0}}) = (\mathrm{B}_{\widehat{p_0}}(r) \cap \mathrm{B}_{p_0}(r)) \bigcup \mathrm{B}_{\widehat{p_0}}(r + r^{\frac{3}{2}}) \setminus (\mathrm{B}_{\widehat{p_0}}(r) \cap \mathrm{B}_{p_0}(r)).$$

On pose $\varepsilon = \sqrt{3}h$ qui majore la distance de Hausdorff de Z_h et ∂X. Par hypothèse, on a $\varepsilon \leq \frac{r}{2}$. On a donc que les points $\widehat{p_0}$ et p_0 appartiennent à $\mathrm{B}_{\widehat{p_0}}(r) \cap \mathrm{B}_{p_0}(r)$.

De plus les fonctions $\chi_{r,\widehat{p_0}}$ et χ_{r,p_0} sont égales à 1 sur l'intersection $\mathrm{B}_{\widehat{p_0}}(r) \cap \mathrm{B}_{p_0}(r)$ et comme $\|\widehat{p_0} - p_0\| \leq d_H(\partial X, Z_h) \leq \varepsilon$, l'ensemble $\mathrm{B}_{\widehat{p_0}}(r + r^{\frac{3}{2}}) \setminus (\mathrm{B}_{\widehat{p_0}}(r) \cap \mathrm{B}_{p_0}(r))$ est contenu dans l'offset $[\partial(\mathrm{B}_{\widehat{p_0}}(r) \cap \mathrm{B}_{p_0}(r))]^{\varepsilon + r^{\frac{3}{2}}}$ dont le volume est borné par la proposition 1.10.a par $O(r^2(\varepsilon + r^{\frac{3}{2}}))$. On décompose de manière symétrique le support de χ_{r,p_0}, ce qui donne

$$\|\mathcal{V}_{\partial X,R}(\chi_{r,\widehat{p_0}}) - \mathcal{V}_{\partial X,R}(\chi_{r,p_0})\|_{\mathrm{op}} = O(R^2 r^2(\varepsilon + r^{\frac{3}{2}})).$$

Comme $\varepsilon = \sqrt{3}h$ on a le résultat. □

Lemme 3.9. *Sous les hypothèses du théorème 3.6, on a*

$$\frac{3}{2\pi r^2 R^3}\|\mathcal{V}_{\partial X,R}(\chi_{r,p_0}) - \mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0}(r)})\|_{\mathrm{op}} = O(r^{\frac{1}{2}}).$$

Démonstration. Comme $\chi_{r,p_0} = \mathbf{1}_{\mathrm{B}_{p_0}(r)}$ sur la boule $\mathrm{B}_{p_0}(r)$, en utilisant les mêmes arguments que dans les démonstrations de la partie précédente, on a

$$\|\mathcal{V}_{\partial X,R}(\chi_{r,p_0}) - \mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0}(r)})\|_{\mathrm{op}} \leq \mathcal{H}^3\left(\partial X^R \cap \left[\mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(r + r^{\frac{3}{2}})) \setminus \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(r))\right]\right) \times R^2.$$

La proposition 1.10 implique que le volume $\mathcal{H}^3\left(\partial X^R \cap \left[\mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(r + r^{\frac{3}{2}})) \setminus \mathfrak{p}_{\partial X}^{-1}(\mathrm{B}_y(r))\right]\right)$ est plus petit que $4R \times \mathcal{H}^2\left(\partial X \cap \mathrm{B}_y(r + r^{\frac{3}{2}}) \setminus \mathrm{B}_y(r)\right)$. Le fait que cette aire est bornée par $O(r^{\frac{5}{2}})$ nous permet de conclure. □

Lemme 3.10. *Sous les hypothèses du théorème 3.6, on a*

$$\|\frac{3}{2\pi R^3 r^2}\mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0}(r)}) - n(p_0)n(p_0)^t\|_{\mathrm{op}} = O(r + R^2)$$

Démonstration. On a la relation suivante (voir théorème 1 de [Mér09])

$$\mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0}(r)}) = \frac{2}{3}R^3\left[1 + O(R^2)\right]\int_{p \in \mathrm{B}_{p_0}(r) \cap S} n(p)n(p)^t \, \mathrm{d}p. \tag{3.3}$$

Par le théorème des valeurs intermédiaires appliqué à la normale à ∂X, on a

$$\|n(p) - n(p_0)\| \leq \sup_{q \in S} \|Dn(q)\|_{op} \, l_{p,p_0},$$

où l_{p,p_0} est la longueur de la géodésique joignant p et p_0. Comme la corde (pp_0) est contenue dans l'offset ∂X^R, où $R < \rho$, on a $l_{p,p_0} = O(\|p - p_0\|)$ (voir [MT04] par exemple). D'où $\|n(p) - n(p_0)\| = O(r)$ et donc $n(p)n(p)^t - n(p_0)n(p_0)^t = O(r)$. Par conséquent

$$\int_{p \in \mathrm{B}_{p_0}(r) \cap S} n(p)n(p)^t \, \mathrm{d}p = \mathcal{H}^2(\mathrm{B}_{p_0}(r) \cap S)n(p_0)n(p_0)^t + \mathcal{H}^2(\mathrm{B}_{p_0}(r) \cap S)\, O(r).$$

En combinant avec l'équation. (3.3), on a

$$\frac{3}{2R^3\mathcal{H}^2(\mathrm{B}_{p_0}(r) \cap S)}\mathcal{V}_{\partial X,R}(\mathbf{1}_{\mathrm{B}_{p_0}(r)}) = \left[1 + O(R^2)\right] \times \left(n(p_0)n(p_0)^t + O(r)\right).$$

On conclut en utilisant le fait que $\mathcal{H}^2(\mathrm{B}_{p_0}(r) \cap S)$ est equivalent à πr^2. □

4 EXPÉRIMENTATIONS

On évalue expérimentalement la convergence multigrille, la précision et la robustesse au bruit Hausdorff de notre estimateur de normale, mais aussi sa stabilité pour détecter les arêtes vives.

La première série d'expérimentations analyse la convergence de l'estimation de normale avec le VCM par rapport à la normale du bord de la forme ∂X (figures 3.4 et 3.5). Les calculs sont effectués sur un tore de grand rayon 6 et de petit rayon 2, et sur un ellipsoïde de demi axes $\sqrt{90}$, $\sqrt{45}$ et $\sqrt{45}$. On mesure l'erreur d'angle $\epsilon(p) = \cos^{-1}(\hat{n}(p) \cdot n(p))$ pour tout point $p \in Z_h$ de la forme discrétisée avec deux différentes normes :

$$l_1(\epsilon) \stackrel{def}{=} \frac{1}{\mathrm{Card}(Z_h)} \sum_{p \in Z_h} \epsilon(p), \qquad l_\infty(\epsilon) \stackrel{def}{=} \sup_{p \in Z_h} \epsilon(p). \tag{3.4}$$

Dans les expérimentations, nous avons essayé différents noyaux χ_{r,p_0} et nous donnons les résultats pour deux d'entre eux : le noyau χ^0_{r,p_0} défini par $\chi^0_{r,p_0}(x) = 1$ si $\|x - p_0\| \leq r$, 0 sinon ; le noyau χ^1_{r,p_0} défini par $\chi^1_{r,p_0}(x) = 1 - \|x - p_0\|/r$ si $\|x - p_0\| \leq r$, 0 sinon. La figure 3.4 affiche les normes des erreurs d'angles estimées en degré, pour des pas de discrétisation de plus en plus fins. Le corollaire 3.7 prédit la convergence multigrille en $O(h^{\frac{1}{8}})$ de l'estimateur quand $r = ah^{\frac{1}{4}}$ et $R = bh^{\frac{1}{4}}$. On observe en pratique que la vitesse de convergence est proche de $O(h)$ pour les deux normes et pour les différents choix de paramètres $R = r = 3h^{\frac{1}{4}}$, $R = r = 3h^{\frac{1}{3}}$ ou $R = r = 3h^{\frac{1}{2}}$. D'autres expérimentations montrent que les résultats les plus précis sont obtenus pour $\alpha \in [\frac{1}{3}, \frac{1}{2}]$ si $R = r = ah^{\alpha}$. Nous notons que le choix du noyau n'a pas un grand impact sur la qualité de l'estimation, tant que celui-ci a un support comparable au noyau χ^0_{r,p_0}.

On perturbe le modèle du tore avec un bruit Kanungo de paramètre $t = 0.25$ (voir section 3.2 du chapitre 3 pour la définition). Ce n'est pas exactement une perturbation Hausdorff mais la plupart des points se situent dans une bande de taille $2h/(1 - t)$. La figure 3.5 montre que la normale est toujours convergente pour les deux normes. Ici encore, la convergence est proche d'une vitesse $O(h^{\frac{2}{3}})$, ce qui est mieux que la convergence théorique en $O(h^{\frac{1}{8}})$.

On s'intéresse à la qualité visuelle de notre estimateur sur différentes formes, en affichant les surfels éclairés selon la normale estimée. Tout d'abord, la figure 3.6 affiche l'estimation des normales sur le modèle du tore perturbé avec un fort bruit Kanungo de paramètre $p = 0.5$. La figure 3.7 affiche les améliorations visuelles en utilisant les normales calculées avec l'estimateur VCM par rapport à celles calculées avec un autre estimateur trivial. En particulier, la comparaison de la figure 3.7.b avec la figure 3.7 c et d montre que l'estimateur du VCM est plus précis que la convolution sur la géométrie des surfels. De plus, on a testé notre estimateur sur de nombreuses formes standards (voir figure 3.8).

Notre estimateur VCM est une matrice qui contient aussi des informations sur la courbure selon les autres directions. Mérigot *et al.* [MOG11] ont proposé de détecter les arêtes vives en utilisant les trois valeurs propres l_1, l_2, l_3 du VCM comme suit : si $l_1 \geq l_2 \geq l_3$, on calcule $l_2/(l_1 + l_2 + l_3)$ et on marque les points si la valeur dépasse un seuil T. La figure 3.9 montre cette détection sur le modèle "bunny" à différentes échelles , avec $T = 0.1$ pour toutes les données (cela correspond

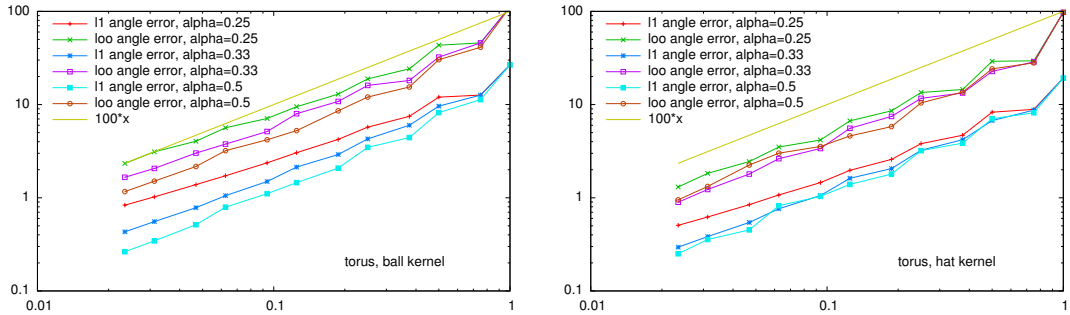


FIGURE 3.4 – Convergence multigrille de l'erreur d'angle de l'estimateur de normale (en degré). Les tests sont faits sur un tore avec les paramètres ($R = r = 3h^\alpha$ avec $\alpha \in \{\frac{1}{4}, \frac{1}{3}, \frac{1}{2}\}$), pour deux normes ($l_1$, l_∞), avec le noyau χ^0_{r,p_0} (gauche), et le noyau χ^1_{r,p_0} (droite).

à un angle de $\approx 25°$). Cela montre que l'information du VCM est géométriquement stable. Pour conclure, on liste ci-dessous des informations sur les temps de calcul. Cet estimateur a été implémenté et est disponible dans la librairie DGtal [DGtrg].

Image	taille	#surfels	(R, r)	calcul χ_r-VCM.	Orient. normales
"Al"	150^3	48017	$(30, 3)$	0.73 s	0.88 s
"rcruiser"	250^3	66543	$(30, 3)$	1.26 s	0.99 s
"bunny"	516^3	933886	$(30, 5)$	30.1 s	15.9 s
"Dig. Snow"	512^3	3035307	$(30, 5)$	82.1 s	53.6 s

CONCLUSION

On a présenté des estimateurs géométriques stables pour les ensembles digitaux, le premier approchant le *Voronoi covariance measure* et le second approchant le champ de normales. On a montré que sous certaines conditions, ils sont convergents multigrille. Ces conditions nous permettent de déterminer les paramètres R and r en fonction du pas de discrétisation h. Les expérimentations ont confirmé à la fois la précision et la stabilité de nos estimateurs. Nous confrontons la précision de notre estimateur avec d'autres techniques dans le dernier paragraphe du chapitre 5.

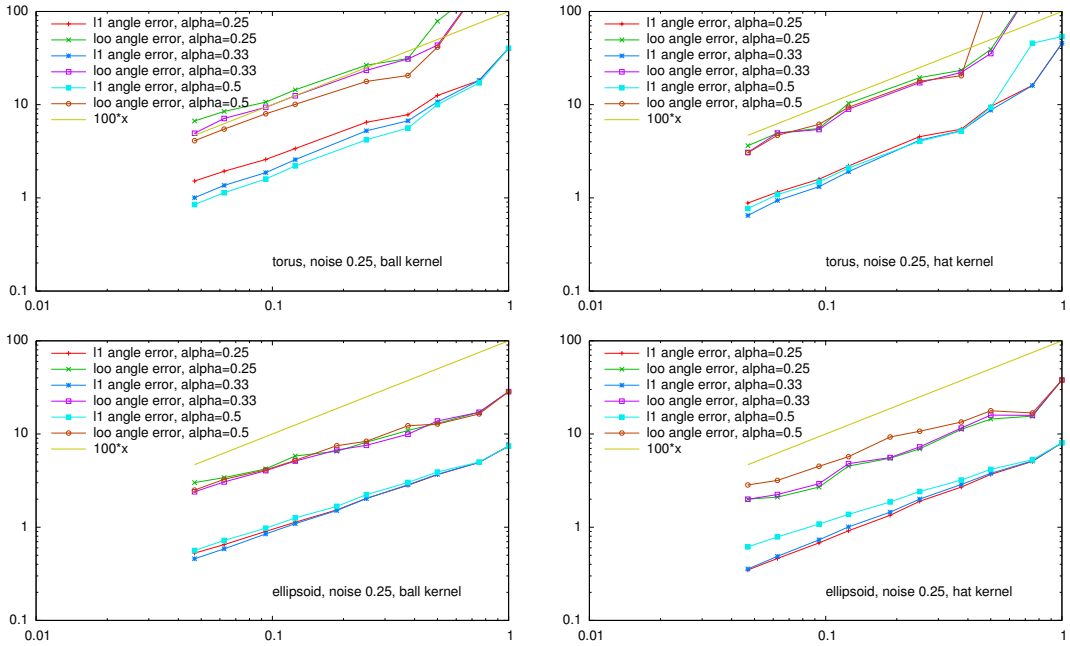


FIGURE 3.5 – Convegence multigrille de l'erreur d'angle de l'estimateur de normale (en degré) sur un modèle bruité. Les tests sont faits sur un tore (première ligne) et sur un ellipsoide (deuxième ligne) perturbé avec un bruit Kanungo 0.25 avec les paramètres ($R = r = 3h^\alpha$ avec $\alpha \in \{\frac{1}{4}, \frac{1}{3}, \frac{1}{2}\}$), pour deux normes ($l_1$, l_∞) : et avec le noyau χ^0_{r,p_0} (gauche), et noyau χ^1_{r,p_0} (droite).

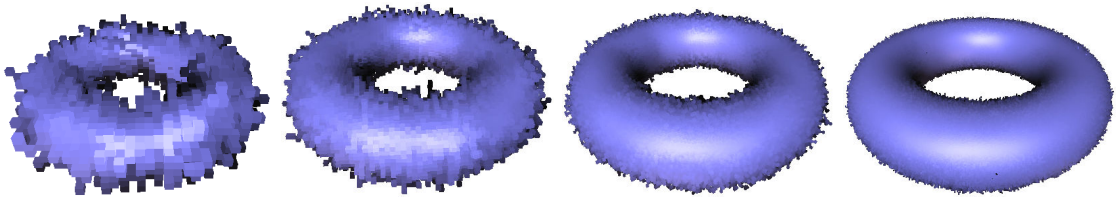
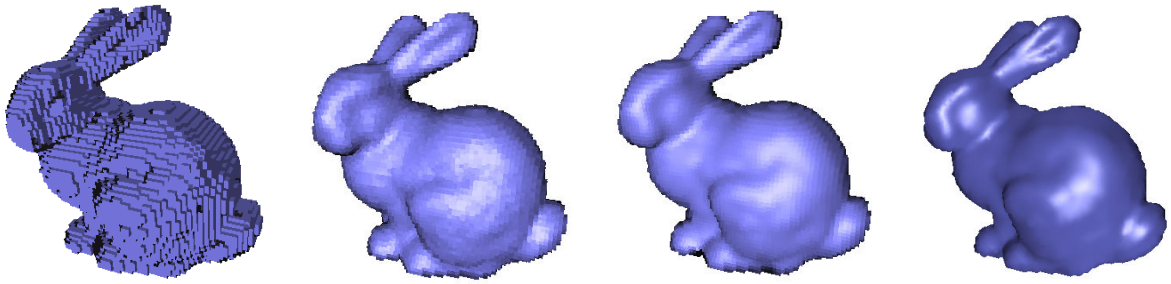


FIGURE 3.6 – Résultat visuel de l'estimation de normale sur le modèle du tore perturbé avec un fort bruit Kanungo ($t = 0.5$) pour différents pas de discrétisation, de gauche à droite $h = 0.5, 0.25, 0.125, 0.0625$.



(a) (b) (c) (d)

FIGURE 3.7 – Aspect visuel de l'estimation de normale sur le modèle "bunny66" pour $r = 3$: (a) normales triviales, (b) normales par convolution χ_r^1 des normales triviales avec un *flat shading*, (c) normales estimées avec le χ_r^1-VCM avec un *flat shading*, (d) normales estimées avec le χ_r^1-VCM avec un *Gouraud shading*.

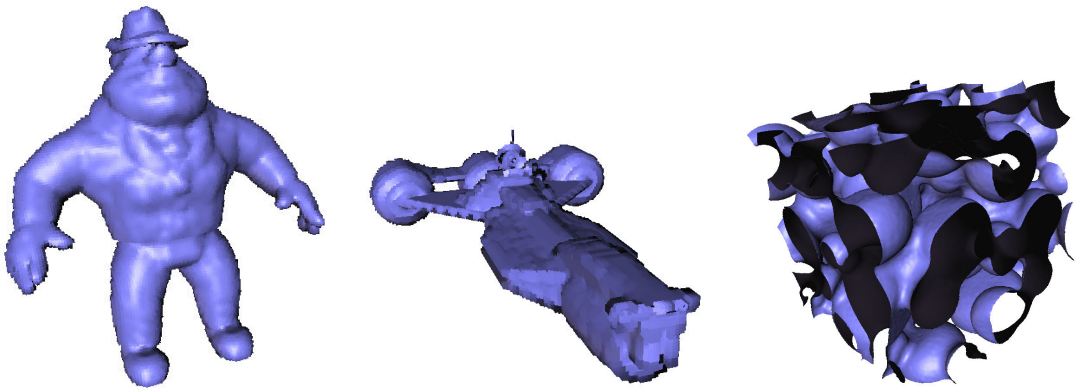


FIGURE 3.8 – Aspect visuel de l'estimation de normale sur les modèles classiques : "Al" 150^3, "Republic cruiser" 250^3, "Digital snow" 512^3.

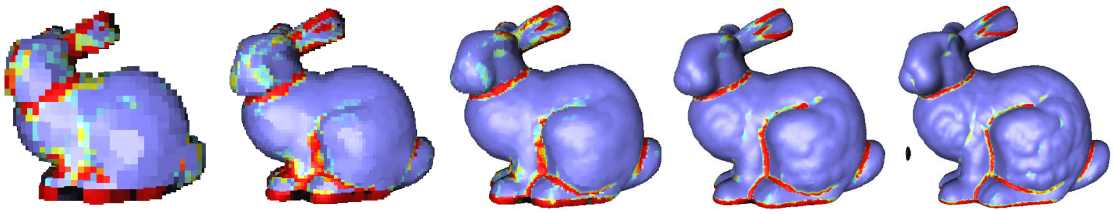


FIGURE 3.9 – Détection d'arêtes vives sur le modèle "bunny" pour différentes résolutions ($R = 30h^{\frac{1}{2}}$,$r = 3h^{\frac{1}{2}}$, $T = 0.1$) : la couleur est bleu métallique quand la valeur appartient à $[0, \frac{2}{3}T]$, puis devient cyan et jaune dans $]\frac{2}{3}T, T[$, jusqu'à rouge dans $[T, +\infty[$.

Deuxième partie

Distances Robustes pour l'Inférence Géometrique

Chapitre 4

FONCTIONS DISTANCE-LIKE

Sommaire

1 Généralités 64
1.1 Définition 64
1.2 Stabilité des gradients des fonctions distance-like 65
2 Distance à une mesure 66
2.1 Définition 66
2.2 Stabilité de la distance à une mesure 67
2.3 Distance à une mesure et distance à un compact 67
3 Exemples de fonctions distance-like calculables 68
3.1 Witnessed-k-distance 68
3.2 Median-k-distance 69
3.3 Center-k-distance 70
3.4 Une approximation de la k-distance supportée par P . . . 70

INTRODUCTION

Dans ce chapitre, nous nous intéressons à la question suivante : avec quelles fonctions distances peut-on faire de l'inférence géométrique robuste à un bruit fort et aux points aberrants (*outliers*) ? Il ne s'agit pas ici de développer une méthode d'inférence géométrique particulière, mais de rappeler les propriétés d'une famille de fonctions distances : les fonctions *distance-like*. A ce titre, ce chapitre constitue une introduction au chapitre suivant qui développe une méthode d'inférence à l'aide de ces fonctions.

FONCTIONS DISTANCES ROBUSTES. De nombreuses méthodes d'estimation géométrique sur des compacts sont liées, directement ou indirectement, à la fonction distance à un compact. La propriété principale de la fonction distance pour ce type d'application est sa stabilité au bruit Hausdorff. Autrement dit, si K et K' sont proches au sens de la distance de Hausdorff, alors $\|d_K - d_{K'}\|_\infty$ est petit. Malheureusement, les données géométriques sont souvent corrompues par des *outliers*, ce qui implique que l'hypothèse de bruit Hausdorff n'est pas réaliste. En effet, la présence d'un seul *outlier* dans un nuage de points peut perturber drastiquement

la fonction distance à ce nuage de points. Pour résoudre cette difficulté, une variante robuste de la fonction distance, qui se comporte bien en présence d'*outliers*, a été introduite dans [CCSM11] : la fonction distance à une mesure (voir Section 2). La classe des fonctions *distance-like* englobe à la fois la fonction distance à un compact et la fonction distance à une mesure.

CONTRIBUTION. On rappelle, dans ce chapitre, la définition et les propriétés de la distance à une mesure et de l'une de ses approximation : la *witnessed k-distance*. On montre un résultat de stabilité des gradients des fonctions *distance-like* en norme L_1 (proposition 4.4) qui est une conséquence du théorème 1.13 ([CCSM10]). Ce résultat généralise le théorème 1.12, qui dans le cadre de la fonction distance à un compact, établit la stabilité de la projection. On propose aussi deux fonctions *distance-like* utilisant le concept de médiane. En pratique, ces fonctions distances sont davantage résistantes aux *outliers*.

1 GÉNÉRALITÉS

1.1 DÉFINITION

Définition 4.1 (Fonctions *distance-like*)**.** *Une fonction $\delta : \mathbb{R}^d \to \mathbb{R}^+$ est dite distance-like si*

- *δ est propre, i.e. $\lim_{\|x\|\to\infty} \delta(x) = \infty$.*
- *δ^2 est 1-semi-concave, i.e. $\|.\|^2 - \delta^2(.)$ est convexe.*

La proposition 3.1 de [CCSM11] implique qu'une fonction *distance-like* est 1-lipschitzienne. Le fait qu'une fonction *distance-like* δ est 1-lipschitzienne implique que δ est différentiable $\mathcal{H}^d$-presque partout (théorème de Rademacher). D'autre part, la 1-semi-concavité implique que la fonction $\|.\|^2 - \delta^2(.)$ est différentiable $\mathcal{H}^d$-presque partout. La proposition suivante établie que la distance usuelle d_K à un compact K est une fonction *distance-like*.

Proposition 4.2. *Si K est un ensemble compact, alors la distance d_K est distance-like.*

Démonstration. Pour la première propriété, il suffit d'utiliser qu'un compact est borné dans $\mathbb{R}^d$. Soit $x \in \mathbb{R}^d$. On pose :

$$\psi_{d_K}(x) := \sup_{y\in K} ||x||^2 - ||x-y||^2 = \sup_{y\in K} 2x.y - ||y||^2.$$

La quantité ψ_{d_K} s'écrit comme un suprémum de fonctions affines et est donc convexe. □

Exemple 4.3. *L'exemple typique de fonction distance-like que nous considérons est la distance de puissance. Étant donné un nuage de points P et une famille de poids $(\omega_p)_{p\in P}$ positifs ou nuls, on appelle distance de puissance à P la fonction distance-like δ_P définie par*

$$\delta_P(x) := \left(\min_{p\in P} \left(\|x-p\|^2 + \omega_p\right)\right)^{1/2}. \tag{4.1}$$

La fonction distance-like δ_P est en conséquence entièrement définie par le diagramme de puissance associé à l'ensemble P. Les cellules de puissance sont définies par

$$\mathrm{Pow}_P^{\omega}(b) = \{x \in \mathbb{R}^n; \forall a \in P, \|x-b\|^2 - \omega_b \leq \|x-a\|^2 - \omega_a\}. \tag{4.2}$$

On remarque que, quand les poids sont nuls, la distance de puissance n'est rien d'autre que la distance au nuage de points P.

1.2 Stabilité des gradients des fonctions distance-like

Comme nous l'avons vu dans les premiers chapitres, la stabilité des estimateurs géométriques est souvent liée à la stabilité du gradient de la fonction distance. Nous généralisons ici le théorème 1.12 [CCSM10] qui établit la stabilité du gradient de la fonction distance à un compact, aux fonctions *distance-like*. La proposition 4.4 est une conséquence du théorème 1.13 rappelé au chapitre 1 établissant la stabilité des gradients des fonctions convexes. Pour toute fonction *distance-like* δ, on définit le champ de vecteurs $\mathbf{n}_\delta(x) = \frac{1}{2}\nabla\delta^2(x)$, qui est défini pour $\mathcal{H}^d$-presque tout point x de $\mathbb{R}^d$ par le théorème de Rademacher. On choisit cette notation car si δ est une fonction distance d_S à une surface lisse S, alors $\mathbf{n}_{d_S}(x)$ est le vecteur de norme $d_S(x)$ et de direction normale à S en x. On rappelle les notations suivantes. Pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}^k$ et tout ensemble E de $\mathbb{R}^d$, on note

$$\|f\|_{1,E} = \int_E \|f(x)\| \mathrm{d}x \qquad \text{et} \qquad \|f\|_{\infty,E} = \sup_{x\in E} \|f(x)\|.$$

Proposition 4.4. *Soient E un ensemble de $\mathbb{R}^d$ à bord rectifiable et $\varepsilon > 0$. Pour toutes fonctions distance-like δ et δ' telles que $\|\delta - \delta'\|_{\infty,E} \leq \varepsilon$, on a :*

$$\|\mathbf{n}_{\delta'} - \mathbf{n}_\delta\|_{1,E} \leq C \times [\mathcal{H}^d(E) + (2\mathrm{diam}(E) + 4R + \sqrt{2R\varepsilon}) \times \mathcal{H}^{d-1}(\partial E)] \times \sqrt{2R\varepsilon},$$

où $R = \max(\|\delta\|_{\infty,E}, \|\delta'\|_{\infty,E})$ et C est une constante ne dépendant que de d.

On introduit $\psi_\delta(x) := \|x\|^2 - \delta^2(x)$ et $\psi_{\delta'}(x) := \|x\|^2 - \delta'^2(x)$. Pour $\mathcal{H}^d$-presque tout x, on a

$$\mathbf{n}_{\delta'}(x) - \mathbf{n}_\delta(x) = \frac{1}{2}(\nabla\delta'^2(x) - \nabla\delta^2(x)) = \frac{1}{2}(\nabla\psi_\delta(x) - \nabla\psi_{\delta'}(x)).$$

En utilisant la convexité de ψ_δ et $\psi_{\delta'}$ et le théorème 1.13, la proposition 4.4 est une conséquence du lemme suivant.

Lemme 4.5. *Sous les hypothèses de la proposition 4.4, on a*
(i) $\forall x \in E,\ |\psi_\delta(x) - \psi_{\delta'}(x)| \leq 2R\varepsilon$
(ii) $\mathrm{diam}(\nabla\psi_\delta(E) \cup \nabla\psi_{\delta'}(E)) \leq 2\mathrm{diam}(E) + 4R.$

Démonstration. Pour tout point x de E, on a

$$\begin{aligned} |\psi_\delta(x) - \psi_{\delta'}(x)| &= |\delta^2(x) - \delta'^2(x)| \\ &= |\delta(x) - \delta'(x)| \times |\delta(x) + \delta'(x)| \\ &\leq 2R\varepsilon. \end{aligned}$$

Soit X et X' deux points de $\nabla\psi_\delta(E)$ et $\nabla\psi_{\delta'}(E)$ respectivement. Il existe x, x' de E tels que $X = 2x - 2\delta(x)\nabla\delta(x)$ and $X' = 2x' - 2\delta'(x')\nabla\delta'(x')$. Alors

$$\begin{aligned}\|X - X'\| &\leq \|2x - 2x'\| + \|2\delta'(x').\nabla\delta'(x') - 2\delta(x).\nabla\delta(x)\| \\ &\leq 2\,\mathrm{diam}(E) + 2R \times \|\nabla\delta\|_{\infty,E} + 2R \times \|\nabla\delta'\|_{\infty,E} \\ &\leq 2\,\mathrm{diam}(E) + 4R.\end{aligned}$$

Dans la dernière inégalité, on utilise le fait qu'une fonction *distance-like* est 1-Lipschitzienne. Le résultat reste vrai si X, X' appartiennent tous les deux à $\nabla\psi_\delta(E)$ ou à $\nabla\psi_{\delta'}(E)$. □

2 Distance à une mesure

2.1 Définition

La distance à une mesure est un exemple de fonction *distance-like* qui a été introduit dans [CCSM11] et est définie pour toute mesure de probabilité sur $\mathbb{R}^d$.

Définition 4.6. *Soit μ une mesure de probabilité sur $\mathbb{R}^d$ et m_0 un paramètre de régularisation dans $]0,1[$. La fonction distance à la mesure μ est définie pour tout point x de $\mathbb{R}^d$ par*

$$\mathrm{d}_{\mu,m_0}(x) := \left(\frac{1}{m_0}\int_0^{m_0} \delta^2_{\mu,m}(x)\mathrm{d}m\right)^{1/2}, \tag{4.3}$$

où $\delta_{\mu,m}(x) = \inf\{r \geq 0,\ \mu(\mathrm{B}_x(r)) \geq m\}$.

La formule définissant la fonction $\delta_{\mu,m}$ s'inspire d'une formule similaire pour la fonction distance à un ensemble K : $\mathrm{d}_K(x) = \inf\{r \geq 0,\ K \cap \mathrm{B}_x(r) \neq \emptyset\}$, comme on le voit sur la figure 4.1. La fonction $\delta_{\mu,m}$ n'étant pas *distance-like* [CCSM11], on la régularise à l'aide de l'eq. (4.3) afin de la rendre *distance-like*. Le corollaire IV.6. de [Mér09] assure que la distance à une mesure est *distance-like*.

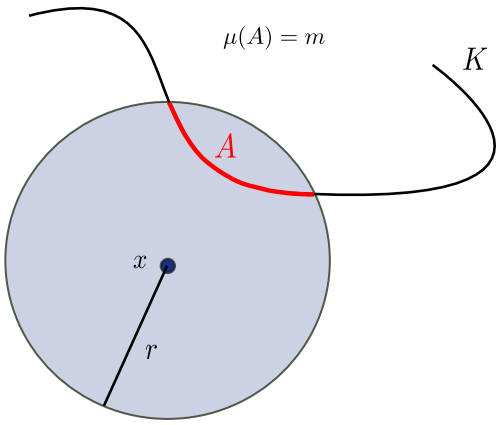


FIGURE 4.1 – Distance à une mesure. On suppose ici que la mesure μ est supportée par un compact K. La quantité $\delta_{\mu,m}(x)$ est le rayon minimum r tel que $\mathrm{B}_x(r) \cap K$ a une masse m.

Le cas du nuage de points Soit $P \subset \mathbb{R}^d$ un ensemble fini de n points, $k \in]0, n[$ un réel et $m_0 = k/n$. Comme dans [GMM13], on appelle *k-distance à P* et on note $d_{P,k}$ la distance à la mesure pour une mesure uniforme sur P avec le paramètre m_0. Dans le cas particulier où k est un entier, un calcul simple [CCSM11] montre que cette définition coïncide bien avec la définition 2.18 de la k-distance donnée dans le chapitre 2,

$$\forall x \in \mathbb{R}^d, \quad d_{P,k}^2(x) = \frac{1}{k} \sum_{p_i \in \mathrm{N}_{P,k}(x)} \|x - p_i\|^2 ,$$

où $\mathrm{N}_{P,k}(x)$ sont les k plus proches voisins de x dans P. De plus, la k-distance $d_{P,k}$ est une distance de puissance [GMM13, proposition 3.1]. Plus précisément, si on note $\mathrm{Bary}_{P,k}$ l'ensemble des isobarycentres de k points distincts de P, on a

$$\forall x \in \mathbb{R}^d \quad d_{P,k}^2(x) = \min_{b \in \mathrm{Bary}_{P,k}} \left(\|x - b\|^2 + \omega_b \right) , \tag{4.4}$$

où le poids $\omega_b = \frac{1}{k} \sum_{p_i \in \mathrm{N}_{P,k}(b)} \|b - p_i\|^2$.

2.2 Stabilité de la distance à une mesure

Il a été montré dans le théorème 3.5 de [CCSM11] que la distance à une mesure est robuste aux *outliers*. Plus précisément, on a

$$\|d_{\mu,m_0} - d_{\mu',m_0}\|_\infty \leq \frac{1}{\sqrt{m_0}} \mathrm{W}_2(\mu, \mu'),$$

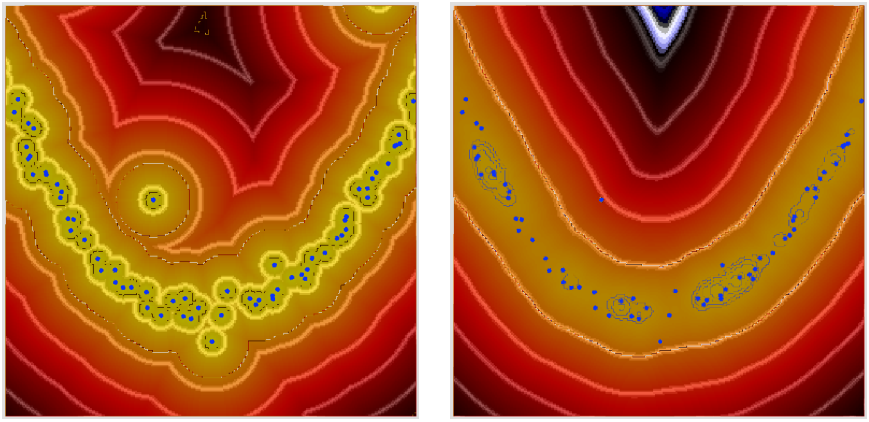
où W_2 est la 2-distance de Wasserstein entre les mesures. Pour plus de détails sur les distances de Wasserstein, qui sont aussi connues sous le nom de *distances de Earthmover* en informatique, on peut se référer à [Vil08]. Pour donner une intuition, quand μ_K et $\mu_{K'}$ sont des mesures de probabilité uniformes sur deux nuages de points K et K' avec le même nombre de points, la distance $\mathrm{W}_2(\mu_K, \mu_{K'})$ est la racine carrée du coût d'un assignement entre K et K'.

La figure 4.2 illustre la stabilité de la k-distance à un nuage de points en présence d'*outliers*. Remarquons que les lignes de niveau de la k-distance (seconde image), sont plus lisses et significatives que les lignes de niveau de la fonction distance classique (première image).

2.3 Distance à une mesure et distance à un compact

On rappelle ici un résultat de stabilité qui indique que l'on peut approcher la distance à un compact par la distance à une mesure. La borne de ce résultat indique qu'en choisissant convenablement le paramètre m_0, la distance à une mesure sera proche de la distance au compact, si la 2-distance de Wasserstein entre les deux mesures associées est petite. Ce résultat implique en particulier une certaine résistance aux *outliers*. Nous avons besoin d'introduire la définition suivante :

Définition 4.7. *On dit qu'une mesure μ sur $\mathbb{R}^d$ a une dimension au plus l avec une constante α_μ si la quantité de masse contenue dans la boule $\mathrm{B}_p(r)$ est au moins $\alpha_\mu r^l$, pour tout point p dans le support de μ et tout rayon r plus petit que le diamètre du support. Si μ a une dimension au plus l, alors il existe une telle constante α_μ.*



a) b)

FIGURE 4.2 – Lignes de niveau de la fonction distance à un compact (a) et de la fonction distance à une mesure (b)

Remarquons que si S est une variété n-dimensionnelle, alors la mesure uniforme μ_S sur S a une dimension au plus n.

Proposition 4.8 (théorème 4.2 de [GMM13]). *Soient μ une mesure probabilité de dimension au plus l, supportée par un compact K et ν une autre mesure, alors pour tout paramètre de masse $m_0 \in]0,1[$,*

$$||d_{\nu,m_0} - d_K||_\infty \leq m_0^{-\frac{1}{2}} W_2(\nu,\mu) + \alpha_\mu^{-\frac{1}{l}} m_0^{\frac{1}{l}}$$

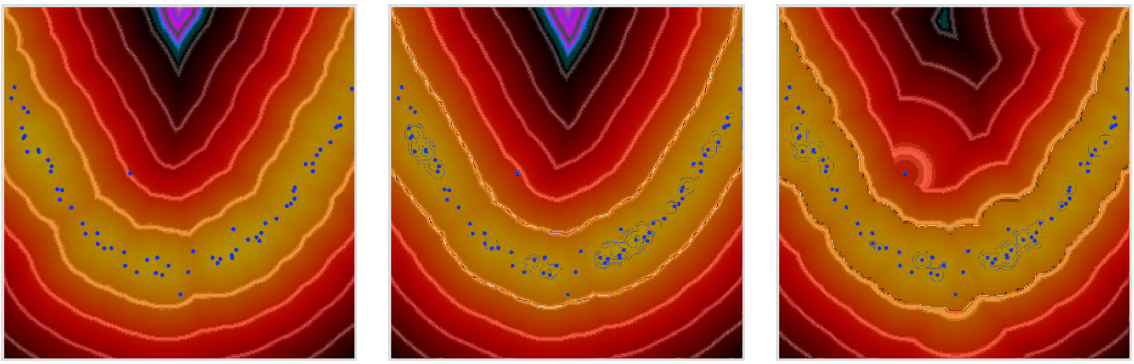
3 EXEMPLES DE FONCTIONS DISTANCE-LIKE CALCULABLES

La k-distance est calculable efficacement en un point, mais il est très coûteux de calculer le diagramme de puissance correspondant. La *witnessed-k*-distance est une relaxation de la k-distance introduite dans [GMM13] dont la robustesse est aussi garantie, et dont le diagramme de puissance est calculable en pratique. Dans cette section, je rappelle tout d'abord la définition et les propriétés de la *witnessed-k*-distance, puis je décris quatre autres exemples de fonctions distances-like inspirées de cette dernière.

3.1 WITNESSED-k-DISTANCE

Soit $P \subset \mathbb{R}^d$ un nuage de points, k un entier. On a vu dans la section précédente (eq. (4.4)) que la k-distance est un diagramme de puissance sur l'ensemble $\mathrm{Bary}_{P,k}$ des barycentres de k points de P. Malheureusement, ce diagramme n'est pas calculable en pratique. L'ensemble $\mathrm{Bary}_{P,k}$ a en effet un cardinal très grand puisqu'il est de l'ordre de $\binom{n}{k}$, si n est le cardinal de P. Pour palier à ce problème, on utilise la *witnessed-k*-distance qui a été introduite dans [GMM13].

On appelle k *witness* un point de $\mathbb{R}^d$ si c'est le barycentre de k points $p_1, \ldots, p_k$ de P tels que $p_2, \ldots, p_k$ sont les $(k-1)$ plus proches voisins de p_1 dans $P \setminus \{p_1\}$ et on appelle *witnessed-k*-uplet un tel k-uplet. On note $\mathrm{Bary}^{\mathrm{w}}_{P,k}$ cet ensemble de



a) b) c)

FIGURE 4.3 – Lignes de niveau de la *witnessed-k*-distance (a), de la median-k-distance (b), et de l'approximation de la k-distance supportée par le nuage P de la section 3.4 (c).

points. La *witnessed-k-distance* est alors définie comme la distance de puissance suivante :

$$d^{\mathrm{w}}_{P,k}(x) := \left(\min_{b \in \mathrm{Bary}^{\mathrm{w}}_{P,k}} \|x-b\|^2 + \omega_b \right)^{1/2},$$

où $\omega_b = \frac{1}{k} \sum_{p_i \in \mathrm{N}_{P,k}(b)} \|b - p_i\|^2$. Notons que le nombre de cellules dans le diagramme de puissance est borné par le nombre de points de P, et peut en conséquence être calculé efficacement avec CGAL [cga].

La proposition suivante [GMM13] assure que la *witnessed-k*-distance est une bonne approximation de la distance au compact sous certaines conditions d'échantillonnage et de choix du paramètre m_0. La borne de ce résultat est la même borne que dans la proposition 4.8 à une constante multiplicative près.

Proposition 4.9. *[GMM13] Soit μ une mesure de probabilité de dimension au plus l, supportée par un compact K. On considère la mesure uniforme $\mathbf{1}_P$ d'un nuage de points P, alors pour tout paramètre de masse $m_0 \in]0,1[$, on a :*

$$||d^{w}_{\mathbf{1}_P,m_0} - d_K||_\infty \leq 3m_0^{-\frac{1}{2}} W_2(\mathbf{1}_P, \mu) + 12\alpha_\mu^{-\frac{1}{l}} m_0^{\frac{1}{l}}$$

Intuitivement, ce résultat indique que la distance à une mesure peut remplacer la distance à un compact puisqu'elle en est une bonne estimation. Nous utiliserons cette proposition dans le chapitre suivant.

3.2 MEDIAN-k-DISTANCE

On introduit ici la *median-k-distance* qui est une fonction *distance-like* dérivée de la *witnessed-k*-distance. Nous n'avons pas de résultat de stabilité pour cette fonction, mais elle est très robuste en pratique (voir figure 4.3). Soient P un nuage de points et $k > 1$ un entier. Un point est appelé *k-median* de P si c'est la médiane géométrique de k points $p_1, \ldots, p_k$ de P tels que $p_2, \ldots, p_k$ soient les $(k-1)$ plus proches voisins de p_1 dans $P \setminus \{p_1\}$. Autrement dit, un *k-median* est la médiane géométrique d'un *witness-k*-uplet de P. On note $\mathrm{Med}_{P,k}$ l'ensemble des *k-median*

de P. La définition de la *median-k*-distance s'inspire de celle de la *witnessed-k*-distance en remplaçant les *k-witness* par les *k-median*.

Définition 4.10. *La median-k-distance est la distance de puissance définie par*

$$d^{\mathrm{m}}_{P,k}(x) := \left(\min_{b \in \mathrm{Med}_{P,k}} \|x - b\|^2 + \omega_b \right)^{1/2},$$

où $\omega_b = \frac{1}{k} \sum_{p_i \in \mathrm{N}_{P,k}(b)} \|b - p_i\|^2$.

Le fait d'avoir remplacé les *k-witness* par les *k-median* revient à remplacer la norme L^2 par la norme L^1. En effet, le barycentre de $p_1, \ldots, p_k$ est le point b qui minimise $\sum_{i=1}^{d} \|b - p_i\|^2$. Similairement, la médiane géométrique est le point b qui minimise $\sum_{i=1}^{d} \|b - p_i\|$. Notons que la médiane géométrique est unique quand les k points ne sont pas colinéaires. La median-k-distance est *distance-like* comme toutes fonctions distance de puissance (cf. exemple 4.3).

3.3 Center-k-distance

On peut tester beaucoup d'exemples de fonctions *distance-like*. Je donne ici un deuxième exemple défini selon le même principe que pour la fonction *distance-like* précédente, en remplaçant le *k-witness* par le *centre géométrique*. Un point c est appelé le centre géométrique d'un nuage P, si pour tout $j \in \{1, ..., k\}$ la j-ème coordonnée de c est la médiane 1-dimensionnel de toutes les j-èmes de chacun des points du nuage P. La *center*-k-distance est alors définie comme la distance de puissance :

$$d^{\mathrm{c}}_{P,k}(x) = \left(\min_{b_c} \|x - b_c\|^2 + \omega_{b_c} \right)^{1/2},$$

où les points pondérés b_c sont les centres géométriques des *witness* k-uplets.

3.4 Une approximation de la k-distance supportée par P

Dans le contexte plus général des espaces métriques, les auteurs de [BCOS13] définissent une approximation de la k-distance à un nuage de points. Dans un espace métrique, il n'est pas possible en général de définir le barycentre de k points et en conséquence on ne peut utiliser la *witnessed-k*-distance. Ils définissent une approximation de la k-distance à l'aide d'une distance de puissance supportée par le nuage de points initial P pondéré par la k-distance. Cette approximation est calculable efficacement puisqu'elle nécessite le calcul de $|P|$ cellules de puissance au maximum. Le théorème 18 de [BCOS13] assure la stabilité de cette approximation. Néanmoins, la figure 4.3.c montre que la régularité des lignes de niveau est moins bonne que les méthodes précédentes. Cela peut s'expliquer par le fait que cette fonction *distance-like* est définie comme une distance de puissance sur le nuage de point lui-même alors que les autres fonctions *distance-like* présentées dans cette section sont des distances de puissance définies sur un nuage de points choisi pour sa représentativité.

Chapitre 5

VORONOI COVARIANCE MEASURE GÉNÉRALISÉ

Sommaire

1 Introduction 71

2 δ-Voronoi covariance measure 72

2.1 Généralités 72

2.2 Stabilité du δ-VCM 74

2.3 démonstration du théorème 5.5 75

3 Exemples de δ-VCM 77

3.1 VCM utilisant la distance à une mesure 77

3.2 VCM utilisant la witnessed k-distance 77

3.3 Autres variantes du δ-VCM 78

4 Implémentation et Expérimentations 78

4.1 Calcul du VCM pour une distance de puissance 78

4.2 Évaluation du $d^{\mathrm{w}}_{P,k}$-VCM 79

4.3 Comparaison avec le $d^{\mathrm{m}}_{P,k}$-VCM 85

4.4 $d^{\mathrm{w}}_{P,k}$-VCM digital 85

1 INTRODUCTION

Alors que les méthodes classiques d'analyse en composante principale tentent d'approcher l'espace tangent, les méthodes basées sur le diagramme de Voronoi essaient d'estimer la normale en approchant le cône normal. On a vu dans le chapitre 3 que ces dernières méthodes permettaient une estimation précise de quantités géométriques telles que la normale ou les courbures. Je propose dans ce chapitre un estimateur de normales, courbures, et d'arêtes vives généralisant le *Voronoi Covariance Measure* (VCM) aux fonctions *distance-like*. Appliqué à la distance à une mesure, cet estimateur se révèle être résistant au bruit Hausdorff et aux *outliers*.

CONTRIBUTIONS

- On étend la définition du *Voronoi covariance measure* d'un compact. Plus précisément, on associe à toute fonction *distance-like* δ, une famille de tenseurs appelés le δ-*Voronoi covariance measure* (δ-VCM).
- On montre la stabilité du δ-VCM. Notre principal théorème (théorème 5.5) établit que si une fonction *distance-like* δ approche convenablement la fonction distance à un compact K, alors le δ-VCM est proche du VCM de K. Quand on l'applique à un nuage de points P approchant une surface S de $\mathbb{R}^3$, ceci implique que l'on retrouve les normales de S très précisément (proposition 5.8). Cette estimation est Hausdorff stable et robuste aux *outliers*.
- La distance à une mesure à un nuage de points n'étant pas calculable en pratique, on la remplace par la witnessed k-distance [GMM13]. On montre que le VCM associé approche toujours bien le VCM d'une surface sous-jacente (proposition 5.9), ce qui nous permet l'implémentation du δ-VCM.
- On montre à travers divers exemples que le δ-VCM induit un estimateur robuste de normales résistant au bruit Hausdorff et aux *outliers*. On l'utilise aussi pour estimer les courbures et détecter les arêtes vives. Nos estimateurs améliorent les résultats du VCM[MOG11] ou du *Jet Fitting* [CP05], même en l'absence d'*outliers*. Nos estimations sont aussi comparées favorablement à l'estimateur de normales de Boulch *et al.* [BM12].
- Dans le cadre de données digitales, on compare des résultats d'estimation des normales de notre méthode avec les résultats obtenus avec l'ACP, le VCM et la méthode *integral invariant* [CLL14]. Ces tests suggèrent expérimentalement la convergence multigrille de l'estimateur de normale basé sur le δ-VCM.

2 δ-VORONOI COVARIANCE MEASURE

La notion de *Voronoi covariance measure* a été introduite dans [MOG11] dans le cadre de l'inférence géométrique. Comme nous l'avons rappelé dans la section 1 du chapitre 3, le VCM d'une surface lisse contient de l'information sur les normales. Dans cette section, on étend cette notion selon le principe suivant. Pour toute fonction *distance-like* δ (voir chapitre 4 pour la définition), on construit le δ-*Voronoi covariance measure* (δ-VCM). Appliqué à la fonction distance classique, on retrouve le VCM original défini dans [MOG11]. Le principal résultat théorique de ce chapitre est le théorème 5.5. Ce théorème établit de manière quantitative que si une fonction *distance-like* δ est uniformément proche de la fonction distance à un compact, alors le δ-VCM est proche du VCM de ce compact. Grossièrement, cela montre que l'on peut retrouver l'information géométrique sur le compact en utilisant seulement une approximation de sa distance par une fonction *distance-like*.

2.1 GÉNÉRALITÉS

Afin de définir le δ-*Voronoi covariance measure* d'une fonction *distance-like* δ nous aurons besoin de deux constructions géométriques :

- Le champ de vecteurs $\mathbf{n}_\delta(x) = \frac{1}{2}\nabla\delta^2(x)$, qui est défini pour $\mathcal{H}^d$-presque tout point x de $\mathbb{R}^d$.
- L'ensemble $\delta^R := \delta^{-1}\left(]-\infty, R]\right)$, où $R > 0$.

On identifie le produit tensoriel $v \otimes w$ de deux vecteurs v et w à la $d \times d$ matrice $v.w^t$ dont les composantes sont définies par $((v \otimes w)_{i,j} = v_i w_j)$

Définition 5.1 (δ-VCM)**.** *Le δ-Voronoi covariance measure est le tenseur qui à toute fonction intégrable $\chi : \mathbb{R}^d \to \mathbb{R}$, associe la matrice suivante :*

$$\mathcal{V}_{\delta,R}(\chi) := \int_{\delta^R} \mathbf{n}_\delta(x) \otimes \mathbf{n}_\delta(x).\chi\left(x - \mathbf{n}_\delta(x)\right) \mathrm{d}x, \tag{5.1}$$

Exemple 5.2 (δ-VCM d'une fonction distance)**.** *Si on considère la fonction distance à un compact K, l'ensemble d_K^R coïncide avec l'offset de K de rayon R, ce qui explique notre choix de notation. D'autre part, le champ de vecteurs $\mathbf{n}_{d_K}$ a une expression explicite faisant intervenir la projection $\mathfrak{p}_K$ sur K :*

$$\mathbf{n}_{d_K}(x) = x - \mathfrak{p}_K(x). \tag{5.2}$$

En comparant l'eq. (5.1) *avec l'eq. (4.2) de [MOG11] et en utilisant cette remarque, on en déduit que le d_K-VCM défini ici coïncide avec la définition originale du VCM de K.*

Considérons une surface lisse orientée S de $\mathbb{R}^3$, de normale unitaire $\mathbf{n}_S$. Si on choisit R assez petit, on rappelle l'expression suivante quand $r \to 0$ [MOG11] :

$$\left\| \mathcal{V}_{d_S,R}(\mathbf{1}_{\mathrm{B}_p(r)}) - \frac{2\pi}{3} R^3 r^2 \left[\mathbf{n}_S(p) \otimes \mathbf{n}_S(p)\right] \right\|_{\mathrm{op}} = O(r^3), \tag{5.3}$$

où $\|.\|_{\mathrm{op}}$ est la norme opérateur classique. Cette équation montre que l'on peut retrouver l'information locale sur les quantités différentielles.

Exemple 5.3 (δ-VCM d'une distance de puissance)**.** *Nous donnons une expression du VCM d'une distance de puissance, que nous utiliserons en pratique. Chaque distance de puissance définit une décomposition de l'espace en une famille de polyèdres convexes (voir eq. 4.2), appelés cellules de puissance, et sur lesquelles la fonction δ_P^2 est quadratique. Quand les poids sont nuls, on retrouve la notion de cellule de Voronoi. Le lemme suivant généralise l'eq. (2.1) de [MOG11], et montre que calculer le VCM d'une distance de puissance revient à calculer la matrice de covariance de l'intersection de chaque cellule de puissance avec une boule (voir aussi l'algorithme 1).*

Lemme 5.4. *Soit (P, ω) un nuage de points pondérés. Soit χ une fonction intégrable de $\mathbb{R}^d$ dans $\mathbb{R}$, alors*

$$\mathcal{V}_{\delta_P,R}(\chi) = \sum_{p \in P} \chi(p)\ M_p, \tag{5.4}$$

où M_p est la matrice de covariance de $C_p := \mathrm{Pow}_P(p) \cap \mathrm{B}_p((R^2 - \omega_p)^{1/2})$,

$$M_p := \int_{C_p} (x - p) \otimes (x - p) \mathrm{d}x. \tag{5.5}$$

Démonstration. Pour tout point x dans l'intérieur d'une cellule de puissance $\mathrm{Pow}_P(p)$, on a l'égalité $\mathbf{n}_{\delta_P}(x) = x - p$, par conséquent, la fonction $\chi(x - \mathbf{n}_{\delta_P}(x)) = \chi(p)$ est constante sur chaque cellule et on peut décomposer l'intégrale sur les cellules de puissance :

$$\mathcal{V}_{\delta_P,R}(\chi) = \sum_{p \in P} \chi(p) \int_{\mathrm{Pow}_P(p) \cap \delta_P^R} (x - p) \otimes (x - p) \mathrm{d}x.$$

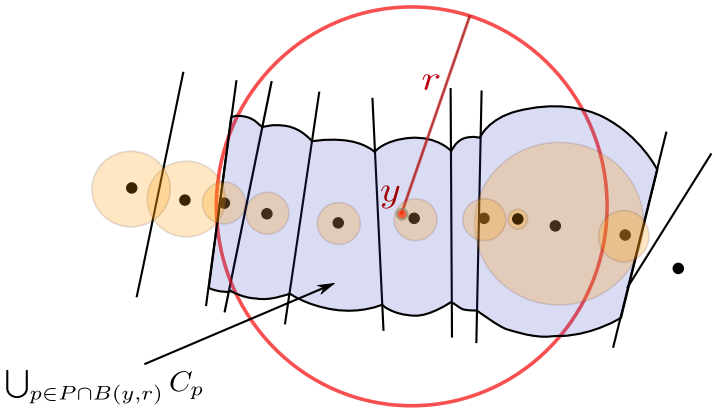


FIGURE 5.1 – Domaine d'intégration du δ_P-VCM. Le nuage de points pondérés P est représenté par une union de disques dont les rayons sont les poids. On suppose que χ est la fonction indicatrice d'une boule $\mathrm{B}_y(r)$ de centre y et de rayon r. Le nombre de cellules de puissance est inférieur au cardinal du nuage de points puisque certaines cellules peuvent être vides.

De plus, un calcul donne $\mathrm{Pow}_P(p) \cap \delta_P^R = \mathrm{Pow}_P(p) \cap \mathbb{B}\left(p, (R^2 - \omega_p)^{1/2}\right)$, si on considère qu'une boule de rayon négatif est vide. $\square$

2.2 STABILITÉ DU δ-VCM

Nous pouvons maintenant établir le théorème principal de ce chapitre. Notre théorème assure que si la fonction distance à un compact K est bien approchée par une fonction *distance-like* δ, alors le VCM de K est aussi bien approché par le δ-VCM.

La norme uniforme d'une fonction f de $\mathbb{R}^d$ est notée $\|f\|_\infty = \sup_{\mathbb{R}^d} |f|$. Étant donnée une fonction Lipschitzienne f de constante de Lipschitz $\mathrm{Lip}(f)$, on introduit sa norme *bounded-Lipschitz* $\|f\|_{\mathrm{BL}} = \|f\|_\infty + \mathrm{Lip}(f)$.

Théorème 5.5. *Soit K un compact et δ une fonction distance-like. Pour toute fonction bornée et Lipschitzienne $\chi : \mathbb{R}^d \longrightarrow \mathbb{R}^+$, on a*

$$\|\mathcal{V}_{\delta,R}(\chi) - \mathcal{V}_{d_K,R}(\chi)\|_{\mathrm{op}} \leq C_1 \ \|\chi\|_{\mathrm{BL}} \, \|\delta - d_K\|_\infty^{\frac{1}{2}} ,$$

où C_1 est une constante qui dépend seulement de R, d et $\mathrm{diam}(K)$.

En pratique, en choisissant une fonction mesurable χ dont le support est contenu dans une petite boule, on peut retrouver l'information locale à l'aide du δ-VCM.

Remarque 5.6. *Une caractéristique importante de ce théorème est que la constante de la borne ne dépend que du diamètre de K et non de sa géométrie locale ou de sa régularité.*

Remarque 5.7. *Si le compact K est une surface lisse S et χ_p^r est la fonction indicatrice d'une boule centrée en un point p de S, l'équation (5.3) indique que le vecteur propre associé à la plus grande valeur propre de $\mathcal{V}_{d_K,R}(\chi_p^r)$ donne une estimation de la normale à S en p. Similairement à la démonstration du théorème 3.6, on peut*

montrer que le théorème 5.5 implique que les vecteurs propres du VCM et du δ-VCM sont deux à deux proches. Par conséquent, la diagonalisation de $\mathcal{V}_{\delta,R}(\chi_p^r)$ définit un estimateur de la direction normale à S en p.

2.3 DÉMONSTRATION DU THÉORÈME 5.5

Cette preuve s'inspire de la démonstration du théorème de stabilité de [MOG11]. L'idée est de comparer les deux intégrales sur l'ensemble $E = K^{R-\varepsilon}$ avec $\varepsilon = \|\delta - d_K\|_{\infty,E}$ et de montrer que la partie restante est négligeable. Par la proposition 4.2 de [CCSM10], l'ensemble ∂E est $(d-1)$-rectifiable. Pour tout $R \in \mathbb{R}^+$, on note $\mathcal{N}(\partial K, R)$ le nombre de recouvrement de ∂K avec un paramètre de rayon R, i.e. le nombre minimal de boules de rayon R nécessaire pour recouvrir ∂K. Supposons tout d'abord que $\varepsilon < \frac{R}{2}$. On a

$$\begin{aligned}\mathcal{V}_{d_K,R}(\chi) = &\int_E \mathbf{n}_{d_K}(x) \otimes \mathbf{n}_{d_K}(x)\chi(x - \mathbf{n}_{d_K}(x))\mathrm{d}x \\ &+ \int_{K^R \setminus E} \mathbf{n}_{d_K}(x) \otimes \mathbf{n}_{d_K}(x)\chi(x - \mathbf{n}_{d_K}(x))\mathrm{d}x.\end{aligned}$$

Pour tout $x \in K^R$, on a $\|\mathbf{n}_{d_K}(x)\| = \|x - \mathfrak{p}_K(x)\| \leq R$. Le fait que $\|\chi\|_\infty \leq \|\chi\|_{\mathrm{BL}}$ implique

$$\left\|\mathcal{V}_{d_K,R}(\chi) - \int_E \mathbf{n}_{d_K}(x) \otimes \mathbf{n}_{d_K}(x)\chi(x - \mathbf{n}_{d_K}(x))\mathrm{d}x\right\|_{op} \leq R^2 \cdot \|\chi\|_{\mathrm{BL}} \cdot \mathcal{H}^d(K^R \setminus K^{R-\varepsilon}). \tag{5.6}$$

On procède similairement pour le δ-VCM. Par définition, on a

$$\|\mathbf{n}_\delta(x)\| = \|\delta(x)\nabla\delta(x)\| \leq |\delta(x)| \leq R.$$

De plus, on a $K^{R-\varepsilon} \subset \delta^R \subset K^{R+\varepsilon}$. Par conséquent, le volume de $\delta^R \setminus K^{R-\varepsilon}$ est inférieur au volume de $K^{R+\varepsilon} \setminus K^{R-\varepsilon}$ et

$$\left\|\mathcal{V}_{\delta,R}(\chi) - \int_E \mathbf{n}_\delta(x) \otimes \mathbf{n}_\delta(x)\chi(x - \mathbf{n}_\delta(x))\mathrm{d}x\right\|_{op} \leq R^2 \cdot \|\chi\|_{\mathrm{BL}} \cdot \mathcal{H}^d(K^{R+\varepsilon} \setminus K^{R-\varepsilon}). \tag{5.7}$$

Bornons maintenant le volume de $K^{R+\varepsilon} \setminus K^{R-\varepsilon}$. La formule générale de l'aire-coaire (§ 3.13 de [Mor08]) nous donne

$$\mathcal{H}^d(K^{R+\varepsilon} \setminus K^{R-\varepsilon}) = \int_{R-\varepsilon}^{R+\varepsilon} \mathcal{H}^{d-1}(\partial K^t)\,\mathrm{d}t$$

On utilise la proposition 4.2 de [CCSM10] qui établit que pour tout réel $t > 0$, on a $\mathcal{H}^{d-1}(\partial K^t) \leq \mathcal{N}(\partial K, t) \cdot \omega_{d-1}(2t)$ où $\omega_n(t)$ est le volume de la boule unité n-dimensionnelle de rayon t.

$$\begin{aligned}\mathcal{H}^d(K^{R+\varepsilon} \setminus K^{R-\varepsilon}) &\leq \int_{R-\varepsilon}^{R+\varepsilon} \mathcal{N}(\partial K, t) \cdot \omega_{d-1}(2t)\,\mathrm{d}t \\ &\leq \mathcal{N}(\partial K, R-\varepsilon) \cdot \omega_{d-1}(2(R+\varepsilon)) \cdot 2\varepsilon \\ &\leq 2\,\mathcal{N}(\partial K, \frac{R}{2}) \cdot \omega_{d-1}(3R)\,\varepsilon.\end{aligned} \tag{5.8}$$

Nous devons maintenant borner la norme opérateur de l'intégrale $\int_E [A_{d_K}(x) - A_\delta(x)]dx$, où

$$A_{d_K}(x) := \mathbf{n}_{d_K}(x) \otimes \mathbf{n}_{d_K}(x)\chi(x - \mathbf{n}_{d_K}(x)),$$

et A_δ est défini de manière similaire. La différence entre ces deux termes se décompose selon l'égalité suivante :

$$\begin{aligned} A_{d_K}(x) - A_\delta(x) = \chi(x - \mathbf{n}_{d_K}(x))(\mathbf{n}_{d_K}(x) \otimes \mathbf{n}_{d_K}(x) - \mathbf{n}_\delta(x) \otimes \mathbf{n}_\delta(x)) \\ + \ [\chi(x - \mathbf{n}_{d_K}(x)) - \chi(x - \mathbf{n}_\delta(x))]\mathbf{n}_\delta(x) \otimes \mathbf{n}_\delta(x) \end{aligned}$$

En utilisant les inégalités suivantes :

$$\begin{aligned} &\|\mathbf{n}_{d_K}(x) \otimes \mathbf{n}_{d_K}(x) - \mathbf{n}_\delta(x) \otimes \mathbf{n}_\delta(x)\| \\ &\leq \|\mathbf{n}_{d_K}(x) \otimes (\mathbf{n}_{d_K}(x) - \mathbf{n}_\delta(x))\| + \|(\mathbf{n}_\delta(x) - \mathbf{n}_{d_K}(x)) \otimes \mathbf{n}_\delta(x)\| \\ &\leq 2R \, \|\mathbf{n}_{d_K}(x) - \mathbf{n}_\delta(x)\|, \end{aligned}$$

et le fait que

$$\|[\chi(\mathbf{n}_{d_K}(x) - x) - \chi(x - \mathbf{n}_\delta(x))]\mathbf{n}_\delta(x) \otimes \mathbf{n}_\delta(x)\| \leq \mathrm{Lip}(\chi) \, \|\mathbf{n}_{d_K}(x) - \mathbf{n}_\delta(x)\| \, R^2$$

donne l'inégalité

$$\|A_{d_K}(x) - A_\delta(x)\|_{\mathrm{op}} \leq \|\chi\|_{\mathrm{BL}} \, (R^2 + 2R)\|\mathbf{n}_{d_K}(x) - \mathbf{n}_\delta(x)\|.$$

Ce qui donne en intégrant

$$\left\| \int_E [A_{d_K}(x) - A_\delta(x)]dx \right\|_{\mathrm{op}} \leq \|\chi\|_{\mathrm{BL}} \, (R^2 + 2R)\|\mathbf{n}_{d_K} - \mathbf{n}_\delta\|_{1,E}.$$

La proposition 4.4 implique que la norme $\|\mathbf{n}_{d_K} - \mathbf{n}_\delta\|_{1,E}$ est bornée par

$$C \times [\mathcal{H}^d(E) + (2\mathrm{diam}(E) + 4R + \sqrt{2R\varepsilon}) \times \mathcal{H}^{d-1}(\partial E)] \times \sqrt{2R\varepsilon}.$$

Comme $\mathrm{diam}(E) \leq \mathrm{diam}(K) + 2R$ et $\varepsilon \leq \frac{R}{2}$, les deux dernières équations impliquent

$$\begin{aligned} &\left\| \int_E [A_{d_K}(x) - A_\delta(x)]dx \right\|_{\mathrm{op}} \\ &\leq C \, \|\chi\|_{\mathrm{BL}} \, (R^2 + 2R)[\mathcal{H}^d(K^R) + (2\mathrm{diam}(K) + 9R)\mathcal{H}^{d-1}(\partial K^{R-\varepsilon})] \, \sqrt{2R} \, \sqrt{\varepsilon}. \end{aligned} \tag{5.9}$$

À nouveau par la proposition 4.2 de [CCSM10], on borne $\mathcal{H}^{d-1}(\partial K^{R-\varepsilon})$ par le réel $\mathcal{N}(\partial K, \frac{R}{2}) \cdot \omega_{d-1}(3R)$. De plus, comme le diamètre de K^R est inférieur à $\mathrm{diam}(K) + 2R$, son volume $\mathcal{H}^d(K^R)$ est borné par une constante faisant intervenir $\mathrm{diam}(K)$, R et d. Remarquons que $\mathcal{N}(\partial K, \frac{R}{2})$ est aussi borné par une constante dépendant des mêmes quantités. Par conséquent, la constante intervenant dans les équations (5.9) et (5.8) dépendent seulement de $\mathrm{diam}(K)$, R et d. On conclut la preuve du théorème 5.5 en combinant (5.9) avec (5.6) et (5.7).

Le cas où $\varepsilon \geq \frac{R}{2}$ est trivial (et pas intéressant en pratique) puisque les deux quantités $\mathcal{V}_{\delta,R}(\chi)$ et $\mathcal{V}_{d_K,R}(\chi)$ peuvent être bornées supérieurement par un réel ne dépendant que de $\mathrm{diam}(K)$, R et d.

3 EXEMPLES DE δ-VCM

3.1 VCM UTILISANT LA DISTANCE À UNE MESURE

On a vu dans la section 2 que la *distance à une mesure* est une fonction *distance-like* résistante aux *outliers*. C'est donc naturel de considérer le δ-VCM dans le cas particulier où δ est une distance à une mesure. On étudie ici la stabilité du δ-VCM dans ce contexte.

Soit S une surface de $\mathbb{R}^3$ et μ_S sa mesure uniforme de probabilité. On sait d'après le théorème de Günther-Bishop (page 749 de [CCSM11]) qu'il existe une constante $\alpha_S > 0$ telle que

$$\forall p \in S,\ \forall r \leq \operatorname{diam}(S), \qquad \mu_S(\mathrm{B}_p(r)) \geq \alpha_S r^2. \tag{5.10}$$

Proposition 5.8. *Soient P un nuage de points de cardinal n, S une surface de $\mathbb{R}^3$, μ_P et μ_S les mesures de probabilité uniforme sur P et S respectivement. Si on prend $k = W_2(\mu_S, \mu_P)\sqrt{\alpha_S}n$, alors on a*

$$\|\mathcal{V}_{d_{P,k},R}(\chi) - \mathcal{V}_{d_S,R}(\chi)\|_{\mathrm{op}} \leq C_2\ \alpha_S^{-1/8}\ \|\chi\|_{\mathrm{BL}}\ W_2(\mu_S, \mu_P)^{\frac{1}{4}},$$

où la constante C_2 dépend de $\operatorname{diam}(S)$ et R.

Démonstration. Soit $k \in]0, n[$ et $m_0 = k/n$. Par la proposition 4.8, on a

$$\|d_{P,k} - d_K\|_\infty \leq m_0^{-\frac{1}{2}} W_2(\mu_S, \mu_P) + \alpha_S^{-\frac{1}{2}} m_0^{\frac{1}{2}}.$$

le membre de droite est minimal quand m_0 satisfait $m_0^{-\frac{1}{2}} W_2(\mu_S, \mu_P) = \alpha_\mu^{-\frac{1}{2}} m_0^{\frac{1}{2}}$, ce qui donne $m_0 = W_2(\mu_S, \mu_P)\sqrt{\alpha_S}$. Si $k = W_2(\mu_S, \mu_P)\sqrt{\alpha_\mu}n$, la borne dans l'équation précédente devient $W_2(\mu_S, \mu_P)^{\frac{1}{2}} \alpha_S^{-1/4}$. On conclut à l'aide du théorème 5.5. □

Ce résultat peut être étendu à toute dimension d et dans le cas où S est un compact de dimension au plus l, puisque pour une telle surface, il existe une constante α_S telle que $\mu_S(\mathrm{B}_p(r)) \geq \alpha_S r^l$).

3.2 VCM UTILISANT LA WITNESSED k-DISTANCE

On a vu dans le paragraphe précédent que le $d_{P,k}$-VCM est résistant aux *outliers*. Malheureusement, il n'est pas calculable en pratique puisque la k-distance ne l'est pas. On remplace donc la k-distance par sa version relaxée et calculable dont la définition et les propriétés sont données dans la section 3.1 du chapitre 4.

La démonstration de la proposition suivante est similaire à celle de la proposition 5.8, et utilise la proposition 4.9.

Proposition 5.9. *Soient μ_S la mesure uniforme sur une surface $S \subset \mathbb{R}^3$ et μ_P la mesure de probabilité uniforme sur un nuage de points P. Soit $k = W_2(\mu_P, \mu_S)\sqrt{\alpha_S}n/4$, où n est le nombre de points de P. Alors pour toute fonction $\chi : \mathbb{R}^d \to \mathbb{R}^+$, on a*

$$\|\mathcal{V}_{d^{\mathrm{w}}_{P,k},R}(\chi) - \mathcal{V}_{S,R}(\chi)\|_{\mathrm{op}} \leq C_3\ \alpha_S^{-1/8}\ \|\chi\|_{\mathrm{BL}}\ W_2(\mu_P, \mu_S)^{\frac{1}{4}},$$

où la constante C_3 dépend de $\operatorname{diam}(S)$ et R.

Remarquons que l'on peut calculer le $d^{\mathrm{m}}_{P,k}$-VCM en utilisant le lemme 5.4.

3.3 AUTRES VARIANTES DU δ-VCM

On peut définir un δ-VCM pour chacune des quatre autres distances de puissance définies dans la section 3 du chapitre 4. L'intérêt de la median-k-distance et de la center-k-distance est l'intervention du concept de médiane. Comme pour la witnessed-k-distance, on peut calculer les VCM correspondant à ces deux *distance-like* en utilisant le lemme 5.4. Au vu des résultats expérimentaux, il semblerait que le VCM associé à la median-k-distance soit le plus robuste aux *outliers*. Nous verrons dans la section suivante que la median-k-distance améliore les résultats d'estimations de normales et courbures. Il semble que la médiane de k points soit en conséquence plus représentative de la géométrie du compact sous-jacent (avec nos modèles de bruit au moins).

4 IMPLÉMENTATION ET EXPÉRIMENTATIONS

4.1 CALCUL DU VCM POUR UNE DISTANCE DE PUISSANCE

On décrit ici notre algorithme pour calculer une approximation du VCM d'une distance de puissance. Soit (P, ω) un nuage de points pondérés qui définit une fonction *distance-like* δ_P. Afin d'estimer la normale et la courbure, nous devons calculer pour chaque point q de P la matrice de covariance $\mathcal{V}_{\delta_P,R}(\chi_q^r)$, où χ_q^r est la fonction indicatrice de la boule $\mathrm{B}_q(r)$. En utilisant le lemme 5.4, on a

$$\mathcal{V}_{\delta_P,R}(\chi_q^r) = \sum_{p \in P \cap \mathrm{B}_q(r)} M_p,$$

où la matrice M_p est définie par

$$\begin{aligned} C_p &:= \mathrm{Pow}_P(p) \cap \mathbb{B}\left(p, (R^2 - \omega_p)^{1/2}\right) \\ M_p &:= \int_{C_p} (x-p) \otimes (x-p) \mathrm{d}x \end{aligned}$$

La difficulté principale est de calculer ces matrices de covariance. En pratique, on les approche en remplaçant la boule dans la définition de C_p par un polyèdre convexe. En entrée de notre algorithme (résumé dans l'algorithme 1) on prend un nuage de points pondérés, un rayon R et une approximation de la boule unité par un polyèdre convexe B. Dans tous les exemples, le polyèdre B est un dodécaèdre circonscrit à la boule unité. Pour calculer la matrice de covariance approchée $(M_p^B)_{p \in P}$, on construit tout d'abord le diagramme de puissance du nuage de points pondérés. On procède ensuite en deux étapes pour chaque point :

4.1.1 INTERSECTION

On calcule une approximation de la cellule C_p

$$C_p^B := \mathrm{Pow}_P(p) \cap (p + (R^2 - \omega_p)^{1/2} B) \tag{5.11}$$

Pour effectuer ce calcul efficacement, on rassemble les demi-espaces définissant les deux polyèdres et on calcule leur intersection. Dans [PM79], l'auteur donne

une méthode pour calculer efficacement l'intersection de demi espace. À chaque demi-espace d'équation $ax + by + cz + d \leq 0$ on associe le point de coordonnées $(\frac{-a}{d}, \frac{-b}{d}, \frac{-c}{d})$. Les auteurs montrent que calculer l'intersection des demi-espaces est équivalent à calculer l'enveloppe convexe du nuage de points associé.

4.1.2 Intégration

Le bord du polyèdre C_p^B est décomposé en une union de triangles, et on considère les tétraèdres $\Delta_p^1, \ldots, \Delta_p^{k_p}$ joignant ces triangles au centroid de C_p. On calcule

$$M_p^B = \int_{C_p^B} (x-p) \otimes (x-p) \mathrm{d}x \tag{5.12}$$

en sommant la contribution signée de chaque tétraèdre, qui peut être évaluée exactement en utilisant la même formule que dans [ACSTD07]. Nous avons implémenté cet algorithme avec CGAL [cga] pour le calcul du diagramme de puissance et l'étape d'intersection. Nous avons reporté dans la Table 5.1 des temps d'exécution pour la witnessed-k-distance et la median-k-distance.

Algorithm 1 Calcul de $\mathcal{V}_{\delta_P,R}$.

Require: $P \subseteq \mathbb{R}^d$ nuage de points, $R > 0$, $B =$ approximation de $\mathrm{B}_0(1)$
  Calcul du diagramme de puissance $(\mathrm{Pow}_P(p))_{p \in P}$
  for all $p \in P$ **do**
    $C_p^B \leftarrow \mathrm{Pow}_P(p) \cap (p + (R^2 - \omega_p)^{1/2} B)$ {§4.1.1}
    $\Delta_p^1, \ldots, \Delta_p^{k_p} \leftarrow$ décomposition de C_p^B en tétraèdres {§4.1.2}
    $M_p^B \leftarrow \sum_{i=1}^{k_p} \int_{\Delta_p^i} (x-p) \otimes (x-p) \mathrm{d}x$ {§4.1.2}
  end for
  return $(M_p^B)_{p \in P}$.

formes	nombre de points	$d_{P,k}^{\mathrm{w}}$-VCM	$d_{P,k}^{\mathrm{m}}$-VCM
ellipsoid	10K	3.31 s	3.44 s
hand	36K	12.31s	13.86s
bimba	74K	25.98 s	29.08 s
ceasar	387K	175.16 s	201.20 s

Table 5.1 – Temps de calcul de $(M_p^B)_{p \in P}$ pour le $d_{P,k}^{\mathrm{w}}$-VCM et le $d_{P,k}^{\mathrm{m}}$-VCM (avec un $4 \times 2.5Ghz$ CPU).

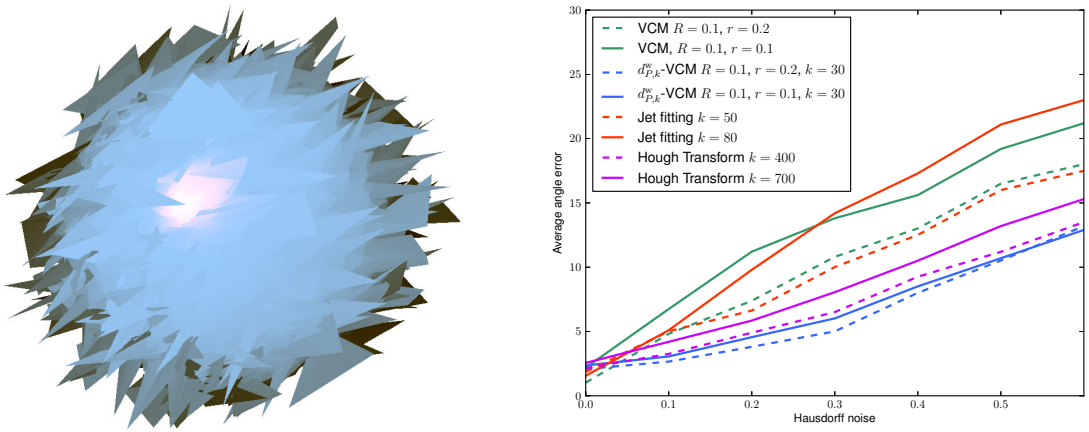
4.2 Évaluation du $d_{P,k}^{\mathrm{w}}$-VCM

On montre dans cette section que le $d_{P,k}^{\mathrm{w}}$-VCM induit en pratique une estimation robuste des normales, courbures et arêtes vives, qui est résistante à la fois au bruit Hausdorff et aux *outliers*. Notons que dans nos expériences, le nuage de points P est l'ensemble des sommets d'une triangulation. La triangulation en elle-même est

seulement utilisée pour la visualisation, notre algorithme étant basé sur le nuage de points P. Le diamètre du nuage de points dans tous les exemples est 2. On considère qu'un nuage de points est non bruité quand l'échantillonnage est fait sur la surface-sous-jacente. On dit que ce nuage de points est bruité avec un bruit Hausdorff de taille ou de paramètre ε si tout point p est translaté d'une distance au plus ε dans une direction quelconque uniforme (voir paragraphe 3.1 du chapitre 2 pour plus d'informations). Nous parlerons d'*outliers* si un certain pourcentage de points est translaté davantage.

4.2.1 Estimation des normales

Comme nous l'avons suggéré dans la remarque 5.7, on définit une normale en chaque point p d'un nuage de points P comme un vecteur propre associé à la plus grande valeur propre de $\mathcal{V}_{\delta_P,R}(\chi_p^r)$.

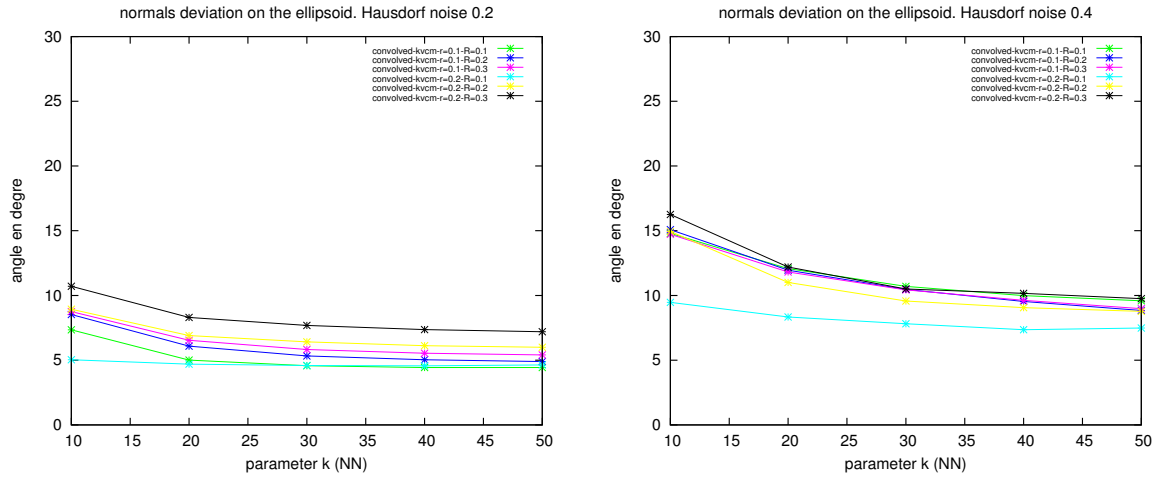


a) b)

FIGURE 5.2 – Estimation de normale sur l'ellipsoïde. Gauche : le rendu illustre la qualité de l'estimation des normales (paramètres : $k = 30$, $R = 0.2$, $r = 0.2$, bruit Hausdorff $= 0.4$). Droite : Comparaison de la déviation moyenne des normales (en degrés) estimées avec le *Jet Fitting*, VCM, $d^{\mathrm{w}}_{P,k}$-VCM et le *Randomized Hough Transform* [BM12] pour différentes valeurs de bruit Hausdorff.

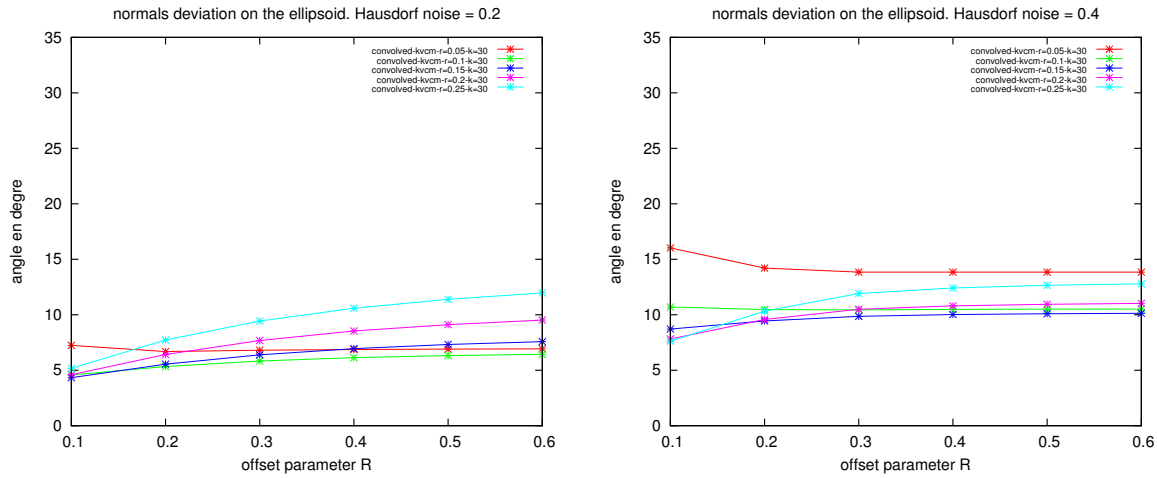
Comparaison avec d'autres méthodes On compare la précision de notre méthode au VCM [MOG11], au *Jet Fitting* [CP05] et au *Randomized Hough Transform* [BM12]. La figure 5.2 présente la déviation angulaire moyenne entre la normale estimée et la vrai normale dans le cas d'un ellipsoïde bruité avec des *outliers*. Notre méthode donne de meilleurs résultats.

Sensibilité aux paramètres Pour mesurer la dépendance de notre méthode aux différents paramètres, on calcule la déviation moyenne des normales sur un ellipsoïde bruité pour différents choix de paramètres k et R. Comme on peut le voir sur les figures 5.3 et 5.4, les résultats sont assez stables. La figure 5.3 montre qu'il faut choisir k plus grand quand le bruit est large. On observe que la valeur



a) b)

FIGURE 5.3 – Dépendance au paramètre k : on peut voir l'influence du paramètre k (en abscisse) sur la déviation moyenne (en degrés) des normales estimées utilisant le $d^{\mathrm{w}}_{P,k}$-VCM. L'expérimentation est faite pour différentes valeurs de r et R et pour deux différents ellipsoïdes bruités. Gauche : bruit Hausdorff $= 0.2$. Droite : bruit Hausdorff $= 0.4$.



a) b)

FIGURE 5.4 – Dépendance au paramètre R : on peut voir l'influence du paramètre R (en abscisse) sur la déviation moyenne (en degrés) des normales estimées utilisant le $d^{\mathrm{w}}_{P,k}$-VCM. L'expérimentation est faite pour différentes valeurs de r et pour deux différents ellipsoïdes bruités. Gauche : bruit Hausdorff $= 0.2$. Droite : bruit Hausdorff $= 0.4$.

$k = 30$ donne de bons résultats dans le cas non bruité ainsi qu'en présence de bruit. Par conséquent toutes nos expérimentations (excepté bien sûr celles de la figure 5.3) sont faites avec la valeur $k = 30$. Notons que le VCM classique coïncide exactement avec le $d^{\mathrm{w}}_{P,k}$-VCM quand $k = 1$.

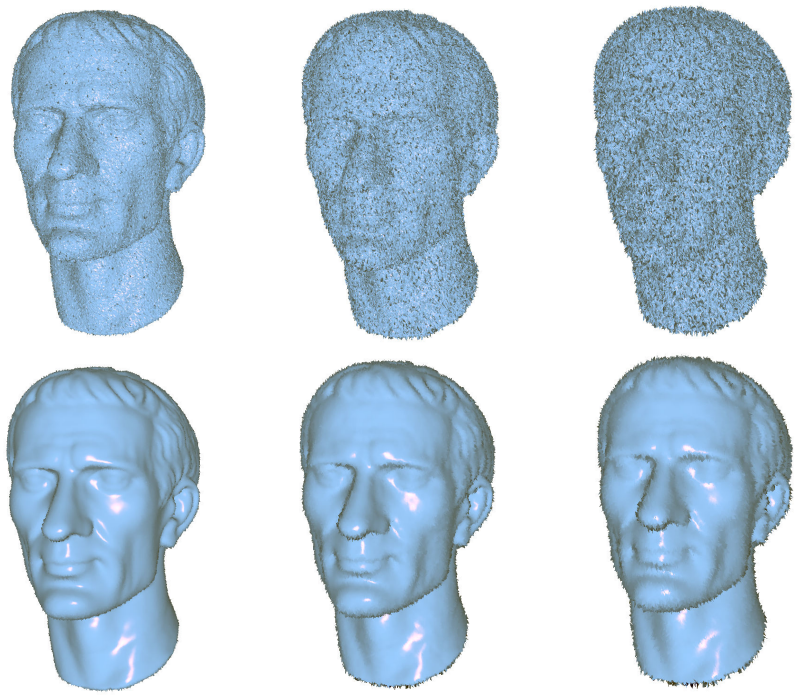


FIGURE 5.5 – Visualisation des données “caesar” utilisant la normale aux triangles (ligne du haut) et la normale estimée avec le $d^{\mathrm{w}}_{P,k}$-VCM (ligne du bas) avec un ombrage de Phong (paramètres $R = 0.04$, $r = 0.04$, $k = 30$). De gauche à droite, le bruit Hausdorff est 0.02, 0.04 et 0.06.

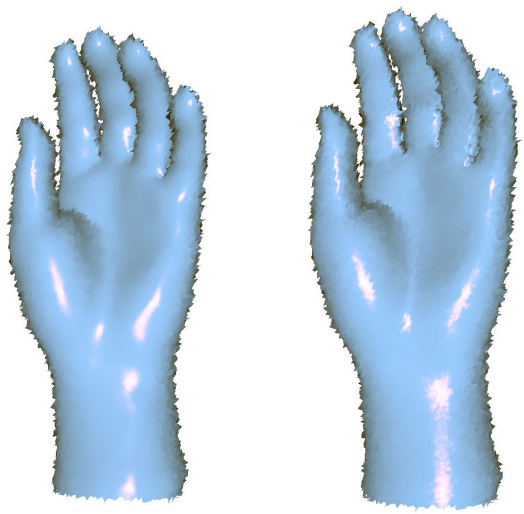


FIGURE 5.6 – Comparaison du rendu sur les données “hand” : (gauche) utilisant notre normale estimée avec le $d^{\mathrm{w}}_{P,k}$-VCM et les paramètres $R = 0.04$, $r = 0.02$, $k = 30$, et (droite) la méthode du *Randomized Hough Transform* [BM12] avec des paramètres standards : nombre de voisins $= 500$ et les autres égaux aux valeurs par défaut. Les triangles sont affichés avec un ombrage de Phong “hand” est perturbé avec un bruit Hausdorff égal à 0.04.

VISUALISATION DES NORMALES ESTIMÉES On teste la qualité de notre estimateur de normale sur des formes standards, en affichant la triangulation éclairée

selon les normales estimées. Dans la figure 5.5, on note que le rendu fait avec les normales calculées avec notre méthode est meilleur que le rendu des normales induites par la géométrie de la triangulation sous-jacente. De plus, le rendu a une bonne résistance au bruit et garde intacte les caractéristiques principales de la forme. Dans la figure 5.6, on compare ce rendu avec le rendu obtenu avec les normales calculées avec la méthode du *Randomized Hough Transform* [BM12], qui est une méthode statistique résistante au bruit et aux *outliers*. Notre méthode présente un rendu plus lisse.

4.2.2 Estimation de la courbure et détection des arêtes vives

La matrice de covariance contient aussi des informations sur la courbure le long des autres directions propres [MOG11]. Si on note $\lambda_0 \geq \lambda_1 \geq \lambda_2$ les trois valeurs propres de $\mathcal{V}_{\delta_P,R}(\chi_p^r)$ en un point p. À une constante multiplicative près, λ_1 et λ_2 correspondent à la valeur absolue de la courbure respectivement minimale et maximale [MOG11]. Les vecteurs propres correspondants sont appelés respectivement les directions principales de courbures minimale et maximale.

On compare notre méthode avec la méthode du *Jet Fitting* [CP05] dans la figure 5.7. Les expérimentations ont été faites pour un large choix de paramètres et notre méthode est plus stable. En présence de beaucoup d'*outliers*, on affiche dans la figure 5.8 la direction principale minimale de courbure estimée avec le $d^{\mathrm{w}}_{P,k}$-VCM et projetée sur la triangulation initiale.

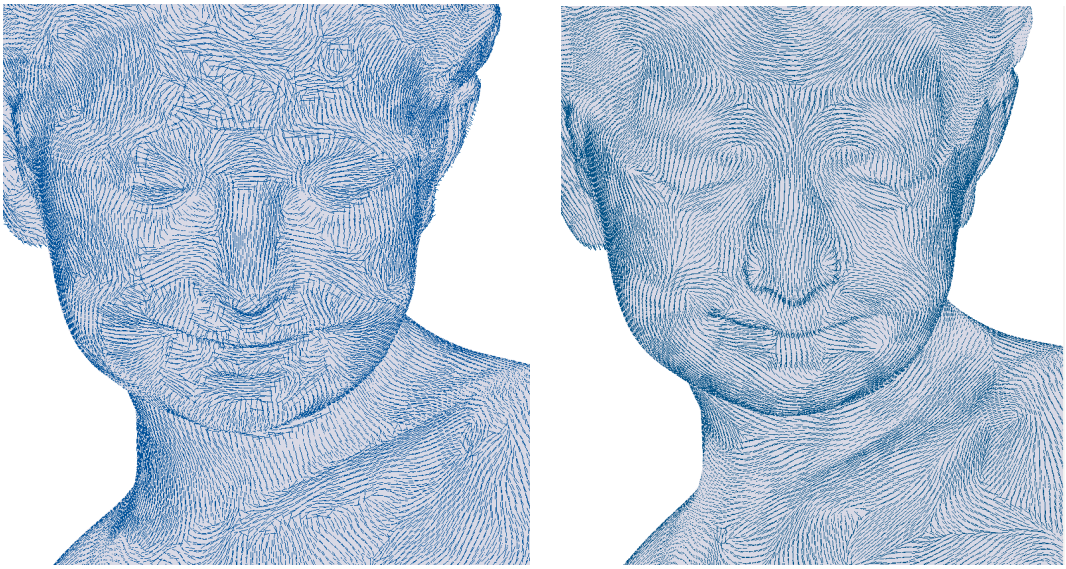


FIGURE 5.7 – Estimation des directions principales sur "Bimba" (sans bruit). On affiche un segment aligné en chaque point avec la direction principale minimale. Gauche : méthode du Jet-fitting (avec un degré de polynôme 4×4 et $k = 100$ voisins pour chaque point). Droite : $d^{\mathrm{w}}_{P,k}$-VCM (paramètres $R = 0.04$, $r = 0.08$, $k = 30$).

Pour illustrer la performance de l'estimation de courbure avec le $d^{\mathrm{w}}_{P,k}$-VCM, la moyenne des valeurs absolue des courbures principales $\lambda_1 + \lambda_2$ est affichée pour les données "caesar" et "hand". Dans la figure 5.9, on affiche aussi le rendu des triangulations avant tout traitement pour illustrer l'amplitude du bruit.

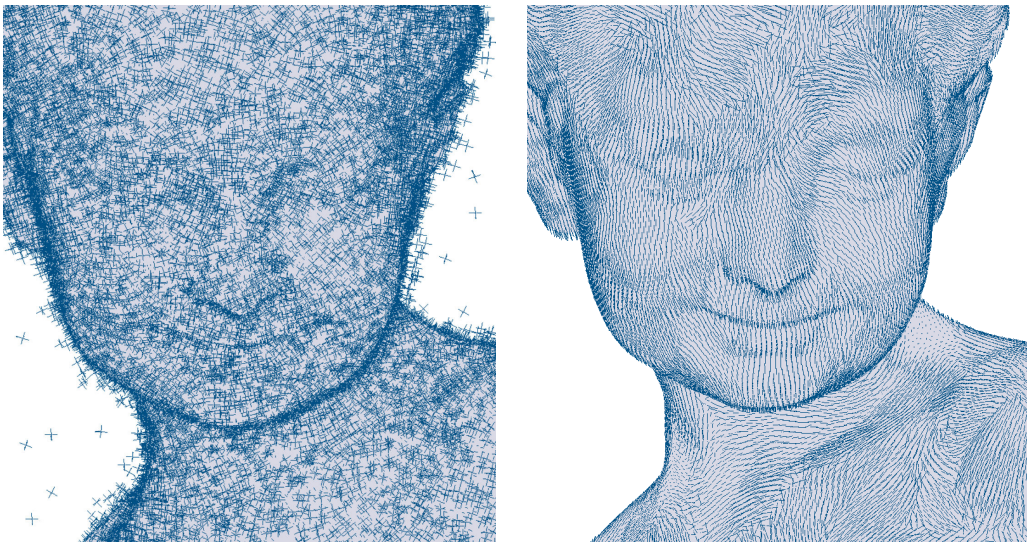


FIGURE 5.8 – Estimation des directions principales sur “Bimba”. Gauche : Données en entrée avec *outliers* (80% des points sont translatés d'une distance au plus 0.02, 10% sont translatés d'une distance entre 0.02 et 0.1, et 10% sont des *outliers* pris uniformément dans une boîte englobante). Droite : pour chaque point, on projette l'estimation de la direction principale minimale du $d^{\mathrm{w}}_{P,k}$-VCM sur la triangulation initiale.

a) b)

FIGURE 5.9 – moyenne des valeurs absolues des courbures sur les données “ceasar” estimées avec le $d^{\mathrm{w}}_{P,k}$-VCM (le jaune représente une courbure faible et le rouge une courbure forte). Paramètres $R = 0.04$, $r = 0.04$, $k = 30$. De gauche à droite : donnée en entrée avec un bruit Hausdorff $= 0.02$; estimation des courbures correspondantes ; donnée en entrée avec un bruit Hausdorff $= 0.06$; estimation des courbures correspondantes.

Notre méthode permet aussi de détecter les arêtes vives, en utilisant le même critère que dans [MOG11] : on dit qu'un point p appartient à une arête vive si $\lambda_1/(\lambda_0 + \lambda_1 + \lambda_2) \geq T$, pour un paramètre de seuil T. La figure 5.11 montre des résultats sur les données ‘fandisk’ pour différents niveaux de bruit (Hausdorff et *outliers*). On remarque que les deux δ-VCM sont moins sensibles aux *outliers* que le VCM classique. Plus précisément, nos méthodes ne détectent que les *outliers*

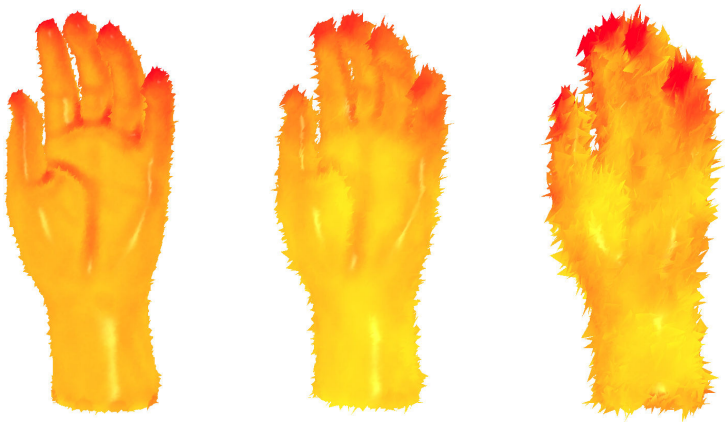


FIGURE 5.10 – moyenne des valeurs absolues des courbures sur les données “ceasar” estimées avec $d^{\mathrm{w}}_{P,k}$-VCM (le jaune représente une courbure faible et le rouge une courbure forte). Paramètres : $R = 0.04$, $r = 0.02$, $k = 30$. De gauche à droite, le bruit Hausdorff est 0.04, 0.08 and 0.16.

proches des arêtes.

4.3 COMPARAISON AVEC LE $d^{\mathrm{m}}_{P,k}$-VCM

J'ai testé la qualité des estimations du δ-VCM pour de nombreuses fonctions *distance-like*, par exemple pour les fonctions *distance-like* présentées dans la dernière section du chapitre 4. Les résultats ont été particulièrement intéressants en utilisant la *median*-k-distance. Les figures 5.12 et 5.13 indiquent que la $d^{\mathrm{m}}_{P,k}$-VCM donne une meilleur estimation de la moyenne des valeurs absolues des courbures que le $d^{\mathrm{m}}_{P,k}$-VCM en présence d'*outliers*. La figure 5.11 montre que la witnessed-k-distance et la median-k-distance donnent des résultats très proches pour la détection d'arêtes. On peut néanmoins remarquer que sur seulement quelques points, la median-k-distance donne de meilleurs résultats.

4.4 $d^{\mathrm{w}}_{P,k}$-VCM DIGITAL

Dans ce paragraphe, je décris les résultats expérimentaux de la version digitale de l'estimateur de normale basée sur le $d^{\mathrm{w}}_{P,k}$-VCM. Comme dans le chapitre 3, on définit le $d^{\mathrm{w}}_{P,k}$-VCM digital d'un contour digital en faisant la somme des contributions à l'intégrale des voxels contenus dans les cellules de puissances. Nous n'avons pas de convergence multigrille théorique de cet estimateur mais avons effectué des tests numériques confirmant cette convergence en pratique. La figure 5.14 compare la déviation moyenne des normales estimées avec les méthodes suivantes : ACP (chapitre 2), VCM (chapitre 3) et *integral invariants* [CLL14]. Les tests numériques sont effectués avec un bruit Kanungo (voir section 3.2) de paramètre 0.4. Les résultats sur le modèle “Leopold” d'équation $100-(x^2.y^2.z^2+4x^2+4y^2+3z^2)$ montrent que l'estimateur du $d^{\mathrm{w}}_{P,k}$-VCM donne des résultats meilleurs que les autres méthodes, même si la vitesse de convergence semble être la même pour toutes les méthodes ($O(h^{0.9})$). Les tests effectués sur le modèle “Blobby” d'équation $1 - 0.1x^4 - 0.1y^4 - 0.1z^4 + 4x^2 + 4y^2 + 4z^2$ contenant moins de partie courbées

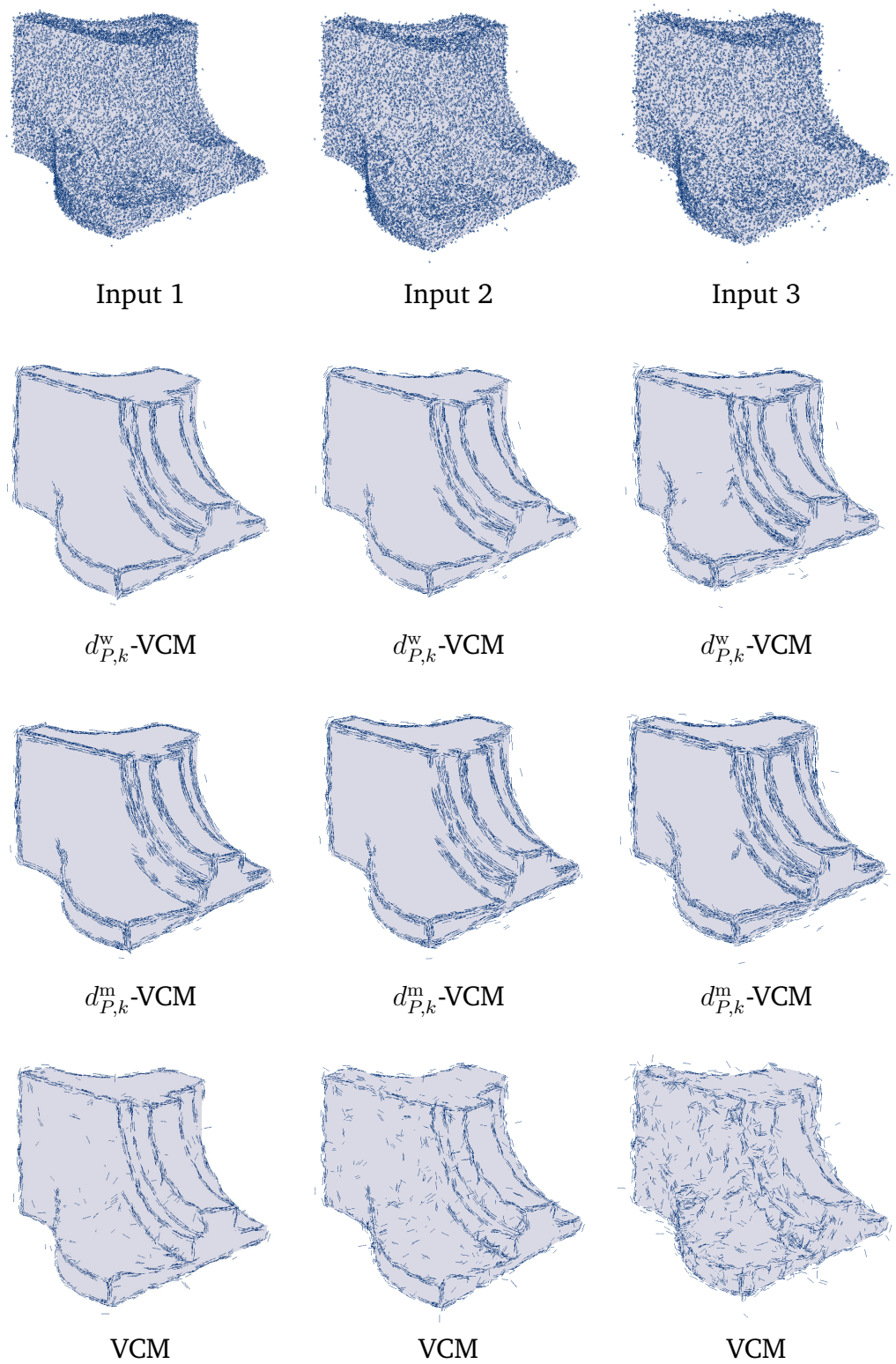
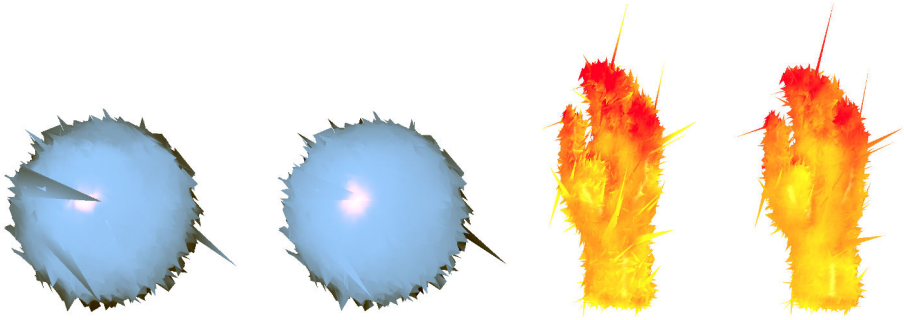
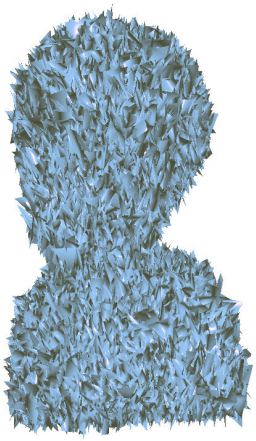


FIGURE 5.11 – Détection d'arêtes sur le modèle 'fandisk' bruité : 98% des points sont translatés d'au plus 0.02 (Input 1), 0.035 (Input 2) ou 0.05 (Input 3), 2% sont translatés d'une distance au plus 0.1 (Input 1), 0.175 (Input 2) ou 0.25 (Input 3). Tous ces résultats sont calculés avec les paramètres $R = 0.06$ et $r = 0.02$. Le $d^{\mathrm{w}}_{P,k}$-VCM et le $d^{\mathrm{m}}_{P,k}$-VCM sont calculés avec $k = 30$.



$d^{\mathrm{w}}_{P,k}$-VCM $d^{\mathrm{m}}_{P,k}$-VCM $d^{\mathrm{w}}_{P,k}$-VCM $d^{\mathrm{m}}_{P,k}$-VCM

FIGURE 5.12 – Comparaison entre le $d^{\mathrm{w}}_{P,k}$-VCM et le $d^{\mathrm{m}}_{P,k}$-VCM (pour les mêmes paramètres $R = 0.04$, $r = 0.02$, $k = 30$). Les données en entrée contiennent des *outliers* (99.9% des points sont translatés d'une distance au plus 0.1, et 0.1% sont translatés d'une distance entre 0.1 et 1). Le rendu illustre la qualité de l'estimation de normale sur l'ellipsoïde. La moyenne des valeurs absolues des courbures est affichée sur le modèle "hand".



Triangulation avec 50 % d'*outliers*

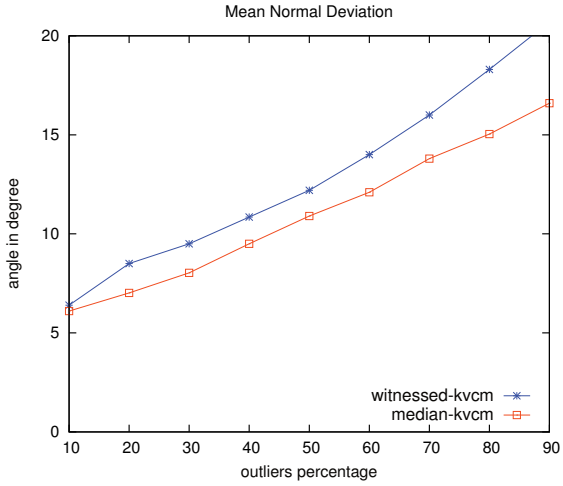


FIGURE 5.13 – Comparaison entre le $d^{\mathrm{w}}_{P,k}$-VCM et le $d^{\mathrm{m}}_{P,k}$-VCM (pour les mêmes paramètres $R = 0.1$, $r = 0.1$, $k = 30$). le bruit Hausdorff est 0.2. Le pourcentage d'*outliers* correspond au pourcentage de points qui sont translatés d'une distance entre 0.05 et 0.1. Gauche : données en entrée avec 50% d'*outliers*. Droite : évolution de la déviation moyenne (en degrés) des normales estimées quand la proportion d'*outliers* grandie.

révèlent que la méthode *integral invariant* obtient de meilleurs résultats mais est plus sensible aux choix des paramètres.

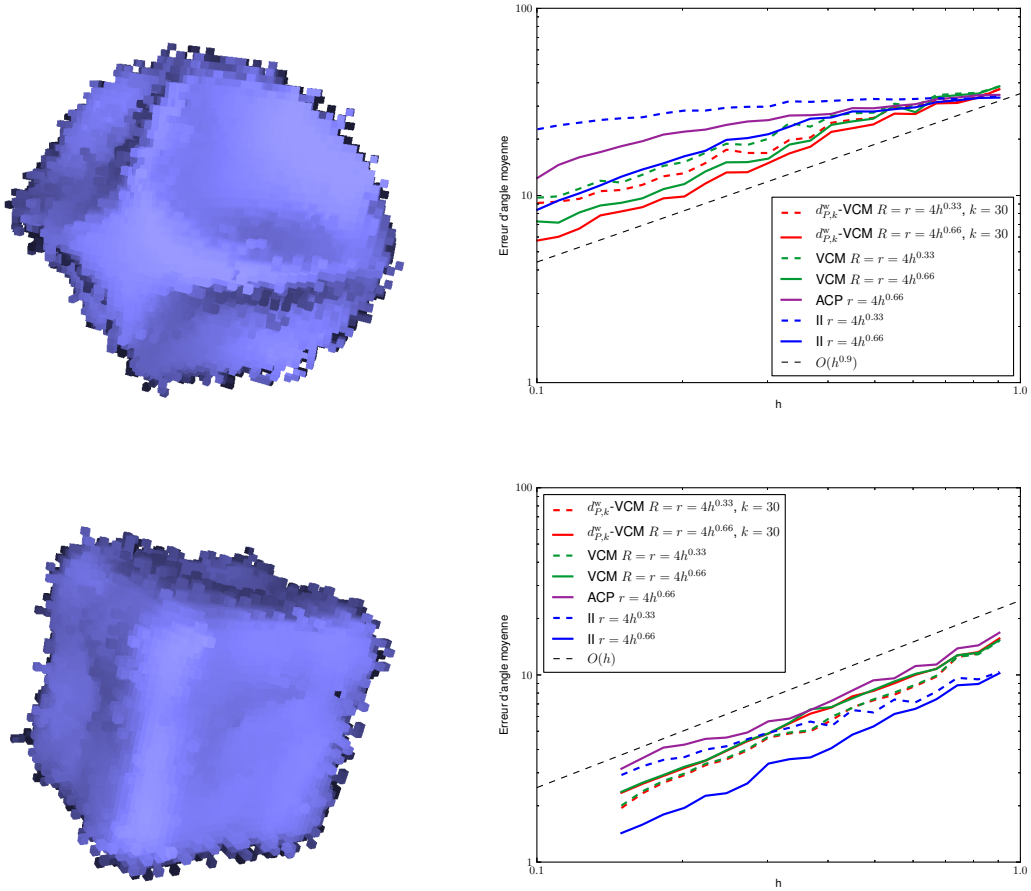


FIGURE 5.14 – À gauche : visualisation des normales estimées avec le $d^{\mathrm{w}}_{P,k}$-VCM sur le modèle “Leopold” (en haut) d'équation $100 - (x^2.y^2.z^2 + 4x^2 + 4y^2 + 3z^2)$ et sur le modèle “Blobby” (en bas) d'équation $1 - 0.1x^4 - 0.1y^4 - 0.1z^4 + 4x^2 + 4y^2 + 4z^2$ bruité avec un bruit Kanungo de paramètre 0.4 et avec un pas de discrétisation $h = 0.3$. À droite : Comparaison des déviations moyennes (en degrés) des normales estimées avec les méthodes VCM, $d^{\mathrm{w}}_{P,k}$-VCM, ACP et Integral Invariant sur le modèle “Leopold” (en haut) et sur le modèle “Blobby” (en bas), bruités avec un bruit Kanungo de paramètre 0.4 pour un pas de discrétisation de plus en plus fin.

Conclusion Générale

Synthèse Dans cette thèse, nous avons présenté des contributions en inférence géométrique. Notre but était de développer des méthodes d'estimation de quantités géométriques et de garantir ces estimations. La stabilité Hausdorff de la fonction distance à un compact nous a permis d'obtenir la stabilité Hausdorff de l'estimateur de normale ACP ainsi que la convergence multigrille des estimateurs digitaux de normale utilisant l'ACP et le VCM.

Nous nous sommes aussi intéressés au développement de méthodes d'inférence géométrique robustes aux *outliers*. La fonction distance n'étant pas robuste aux *outliers*, elle n'était pas l'outil naturel pour aborder ce problème. Les fonctions *distance-like* se sont révélées être un outil très intéressant pour cet objectif. Il n'y avait encore que peu d'exemples de fonctions *distance-like* étudiés dans la littérature. On a proposé différentes fonctions *distance-like* définies à partir d'un nuage de points et on a testé leur stabilité. On a généralisé le *Voronoi covariance measure* aux fonctions *distance-like*. La stabilité théorique de la distance à une mesure et de la *witnessed-*k-distance ont permis de montrer la stabilité aux *outliers* du δ-VCM appliqué à ces fonctions. Les expérimentations ont montré que notre approche permet une estimation très robuste des normales, des courbures, des directions de courbure et des arêtes vives.

Le tableau 5.2 présente un récapitulatif des principales caractéristiques des méthodes étudiées dans cette thèse ainsi que trois méthodes auxquelles nous nous sommes comparés. Ces informations montrent globalement que nos méthodes permettent l'estimation de nombreuses grandeurs avec des garanties théoriques.

Perspectives Parmi les perspectives à court terme, il serait intéressant de généraliser en dimension quelconque la convergence multigrille de l'estimateur digital du VCM (théorème 3.2) et la convergence de l'estimateur de normale associé (corollaire 3.7). On a vu que l'application en géométrie digitale n'est pas immédiate, la difficulté principale étant de trouver pour quelles valeurs des paramètres la convergence multigrille est avérée. Néanmoins, ces résultats de convergence indiquent qu'il est possible d'utiliser les résultats de stabilité d'estimateurs sur des nuages de points pour les appliquer en géométrie digitale. Il serait ainsi intéressant de montrer la convergence multigrille de la version digitale du δ-VCM. Il faudrait pour cela adapter la preuve du théorème 3.2 du chapitre 3 aux fonctions *distance-like*.

Dans la continuité du chapitre 4, un travail théorique reste encore à faire pour montrer la stabilité de certaines fonctions *distance-like*. Il faut adapter la démonstration de la stabilité de la *witnessed-*k-distance [GMM13] pour obtenir, par exemple, la stabilité de la median-k-distance.

Un problème de la méthode du VCM et de sa version généralisée, est sa rela-

Méthodes	Quantités estimées	Détection d'arêtes	Résultat de stabilité	Convergence multigrille théorique	Temps de calcul 20K points
ACP	$\mathbf{n}$	non	Hausdorff (2d) (Conv. [MNG04])	oui (2d)	0.86
VCM	$\mathbf{n}$, λ_1, λ_2 e_1, e_2	oui	Hausdorff(pour $\mathbf{n}$) [MOG11]	oui (pour $\mathbf{n}$)	2.96s
$d^w_{P,k}$-VCM	$\mathbf{n}$, λ_1, λ_2 e_1, e_2	oui	Robuste aux outliers (pour $\mathbf{n}$)	non	4.19s
Jet fitting	$\mathbf{n}$, λ_1, λ_2	non	non	non	1.05s
Hough Transform	$\mathbf{n}$	oui	non	non	1.83s
Integral Invariant	$\mathbf{n}$, λ_1, λ_2 e_1, e_2	non	oui [PWHY09]	oui (λ_1,λ_2) [CLL14]	1.38s

TABLE 5.2 – Résumé des caractéristiques des méthodes d'estimation géométrique sur des nuages de points. Les quatre contributions théoriques principales de cette thèse sont indiquées en vert.

tive lenteur. La *witnessed-k*-distance a été définie comme une version relaxée de la k-distance présentant un nombre réduit de cellules de puissance. Le calcul des cellules de puissance de la *witnessed-k*-distance est en conséquence nettement plus rapide que pour la k-distance. On pourrait aller plus loin en définissant une fonction *distance-like* robuste aux *outliers* qui s'appuient sur un support comportant un nombre considérablement plus réduit de points.

À plus long terme, on peut imaginer généraliser d'autre méthodes d'inférence aux fonctions *distances-like*. En se basant sur l'idée conductrice du chapitre 5, d'autres théorèmes de stabilité Hausdorff d'estimateurs pourraient être généralisés pour obtenir une stabilité aux *outliers*. L'utilisation, par exemple, de la distance à une mesure dans des méthodes de reconstruction de surfaces basées sur des fonctions implicites pourrait améliorer la qualité de la surface reconstruite.

Le VCM généralisé permet des estimations de quantités géométriques très robustes. Il serait très intéressant d'utiliser cette estimation robuste des normales dans des méthodes de reconstruction de surfaces utilisant en entrée un nuage de points munis de normales. On pourrait alors obtenir de bonnes reconstructions en présence d'*outliers* avec les méthodes de reconstruction de Poisson ou certaines méthodes *moving least square*.

NOTATIONS

K	un compact quelconque	14
d_K	fonction distance à K	14
$\|v\|$	norme euclidienne du vecteur v	14
$\mathcal{AM}(K)$	axe médian de K	15
$\mathfrak{p}_K$	projection sur K	15
K^R	voisinage tubulaire de paramètre R (R-offset) de K	16
T_pS	espace tangent à la surface S au point p	16
$\mathbf{n}_S$	normale unitaire à S	16
$\mathrm{diam}(K)$	diamètre de K	16
$\mathcal{H}^k(K)$	k-mesure de Hausdorff de K	17
J_kf	k-jacobien de la fonction f	17
D_f	Différentielle de la fonction f	17
$d_H(K, K')$	distance de Hausdorff entre K et K'	20
$B_x(r)$	boule ouverte de centre x et de rayon r	22
$\mathcal{M}_{k,k'}(\mathbb{R})$	ensemble des matrices réelles $k \times k'$	26
$\|E\|$	cardinal de l'ensemble E	27
$d_{P,k}$	k-distance au nuage de points P	43
$\mathrm{Vor}_P(p)$	Cellule de Voronoi du point p dans le nuage de point P	49
$\|.\|_{op}$	norme matricielle induite par la norme euclidienne	49
$\|f\|_\infty$	norme uniforme de $\mathbb{R}^d$	49
$\mathrm{Lip}f$	Constante de Lipschitz de f	49
$\|f\|_{BL}$	norme bounded-Lipschitz de f	49
$\|f\|_{1,E}$	norme 1 de f restreint à l'ensemble E	65
$\|f\|_{\infty,E}$	norme infinie de f restreint à l'ensemble E	65
d_{μ,m_0}	distance à la mesure μ de paramètre m_0	66
$\mathrm{W}_2(\mu, \mu')$	2-distance de Wasserstein	67

Bibliographie

[AB99] Nina Amenta and Marshall Bern. Surface reconstruction by voronoi filtering. *Discrete & Computational Geometry*, 22 :481–504, 1999.

[ABE09] Dominique Attali, Jean-Daniel Boissonnat, and Herbert Edelsbrunner. Stability and computation of medial axes - a state-of-the-art report. In *Mathematical Foundations of Scientific Visualization*, pages 109–125. Springer Berlin Heidelberg, January 2009.

[ACSTD07] Pierre Alliez, David Cohen-Steiner, Yiying Tong, and Mathieu Desbrun. Voronoi-based variational reconstruction of unoriented point sets. In *Symposium on Geometry processing*, volume 7, pages 39–48, 2007.

[AK04] Nina Amenta and Yong Joo Kil. Defining point-set surfaces. SIGGRAPH, pages 264–270, New York, 2004.

[BCOS13] Mickael Buchet, Frederic Chazal, Steve Y. Oudot, and Donald R. Sheehy. Efficient and robust topological data analysis on metric spaces. In *arXiv :1306.0039*, May 2013.

[Bel65] Howard E. Bell. Gershgorin's theorem and the zeros of polynomials. In *The American Mathematical Monthly*, volume 72, March 1965.

[Blu67] Harry Blum. A transformation for extracting new descriptors of shape. In Weiant Wathen-Dunn, editor, *Models for the Perception of Speech and Visual Form*, pages 362–380, 1967.

[BM12] Alexandre Boulch and Renaud Marlet. Fast and robust normal estimation for point clouds with sharp features. In *Computer Graphics Forum*, volume 31, pages 1765–1774, 2012.

[BMVS04] Volker Blanz, Albert Mehl, Thomas Vetter, and Hans-Peter Seidel. A statistical method for robust 3D surface reconstruction from sparse data. In *2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings*, pages 293–300, September 2004.

[BTS+14] Matthew Berger, Andrea Tagliasacchi, Lee Seversky, Pierre Alliez, Joshua Levine, Andrei Sharf, and Claudio Silva. State of the Art in Surface Reconstruction from Point Clouds. In *EUROGRAPHICS star reports*, volume 1, pages 161–185, April 2014.

[CCSL09] Frédéric Chazal, David Cohen-Steiner, and André Lieutier. Normal cone approximation and offset shape isotopy. In *Computational Geometry*, volume 42, pages 566–581, August 2009.

[CCSLT09] Frédéric Chazal, David Cohen-Steiner, André Lieutier, and Boris Thibert. Stability of curvature measures. In *Computer Graphics Forum*, volume 28, pages 1485–1496, 2009.

[CCSM10] Frédéric Chazal, David Cohen-Steiner, and Quentin Mérigot. Boundary measures for geometric inference. In *Foundations of Computational Mathematics*, volume 10, pages 221–240, April 2010.

[CCSM11] Frédéric Chazal, David Cohen-Steiner, and Quentin Mérigot. Geometric inference for probability measures. In *Foundations of Computational Mathematics*, volume 11, pages 733–751, December 2011.

[CG06] Frédéric Cazals and Joachim Giesen. Delaunay triangulation based surface reconstruction : Ideas and algorithms. In *Effective computational geometry for curves and surfaces*, pages 231–273. Springer, 2006.

[cga] CGAL, Computational Geometry Algorithms Library. http ://www.cgal.org.

[CGAY13] Ricard Campos, Rafael Garcia, Pierre Alliez, and Mariette Yvinec. Splat-based Surface Reconstruction from Defect-Laden Point Sets. In *Graphical Models*, volume 75, pages 346–361. Elsevier, November 2013.

[CLL14] David Coeurjolly, Jacques-Olivier Lachaud, and Jérémy Levallois. Multigrid convergent principal curvature estimators in digital geometry. In *Computer Vision and Image Understanding*, 2014.

[CLR12] David Coeurjolly, Jacques-Olivier Lachaud, and Tristan Roussillon. Multigrid convergence of discrete geometric estimators. In *Digital Geometry Algorithms*, number 2 in Lecture Notes in Computational Vision and Biomechanics, pages 395–424. Springer Netherlands, January 2012.

[CP05] Frédéric Cazals and Marc Pouget. Estimating differential quantities using polynomial fitting of osculating jets. In *Computer Aided Geometric Design*, volume 22, pages 121–146, February 2005.

[Dav63] Chandler Davis. The rotation of eigenvectors by a perturbation. In *Journal of Mathematical Analysis and Applications*, 1963.

[Dey06] Tamal K. Dey. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis*. Cambridge University Press, October 2006.

[DG06] Tamal K. Dey and Samrat Goswami. Provable surface reconstruction from noisy samples. In *Symposium on Computational Geometry*, volume 35, pages 124–141, August 2006.

[DGtrg] DGTAL : Digital geometry tools and algorithms library, `http://libdgtal.org`.

[dVL08] François de Vieilleville and Jacques-Olivier Lachaud. Comparison and improvement of tangent estimators on digital curves. In *Pattern Recognition*, 2008.

[FCOS05] Shachar Fleishman, Daniel Cohen-Or, and Cláudio T. Silva. Robust moving least-squares fitting with sharp features. In *ACM Trans. Graph.*, volume 24, pages 544–552, July 2005.

[Fed59] Herbert Federer. Curvature measures. In *Trans. Amer. Math. Soc*, volume 93, pages 418–491, 1959.

[FM08] Sébastien Fourey and Rémy Malgouyres. Normals and curvature estimation for digital surfaces based on convolutions. In *Discrete Geometry for Computer Imagery*, number 4992 in LNCS, pages 287–298. Springer Berlin Heidelberg, 2008.

[GMM13] Leonidas Guibas, Dmitriy Morozov, and Quentin Mérigot. Witnessed k-distance. In *Discrete & Computational Geometry*, volume 49, pages 22–45, January 2013.

[Gro93] Karsten Grove. Critical point theory for distance functions. In *Differential Geometry, Part 3, Proceedings of Symposia in Pure Mathematics*, volume 54 American Mathematical Society, Providence, RI, pages 357–385, 1993.

[Ham93] Bernd Hamann. Geometric modelling. pages 139–153. Springer-Verlag, London, 1993.

[HDD+92] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface reconstruction from unorganized points. In *SIGGRAPH Comput. Graph.*, volume 26, pages 71–78, July 1992.

[HPW05] Klaus Hildebrandt, Konrad Polthier, and Max Wardetzky. Smooth feature lines on surface meshes. In *Symposium on geometry processing*, pages 85–90, 2005.

[Jol86] Ian T. Jolliffe. *Principal Component Analysis*. Springer, New York, 1986.

[KBH06] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson surface reconstruction. In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, SGP '06, pages 61–70, Aire-la-Ville, Switzerland, 2006. Eurographics Association.

[KL09] B. Kerautret and J. O. Lachaud. Curvature estimation along noisy digital contours by approximate global optimization. volume 42, pages 2265–2278, October 2009.

[KL12] Bertrand Kerautret and Jacques-Olivier Lachaud. Meaningful scales detection along digital contours for unsupervised local noise estimation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 34, pages 2379–2392, 2012.

[KLM98] Pavel Krsek, Gabor Lukács, and Ralph Martin. Algorithms for computing curvatures from range data. In *In The Mathematics of Surfaces VIII, Information Geometers*, pages 1–16, 1998.

[Kol08] Ravikrishna Kolluri. Provably good moving least squares. In *ACM Trans. Algorithms*, volume 4, pages 18 :1–18 :25, May 2008.

[KRV06] Reinhard Knothe, Sami Romdhani, and Thomas Vetter. Combining PCA and LFA for surface reconstruction from a sparse set of control points. In *7th International Conference on Automatic Face and Gesture Recognition, 2006. FGR*, pages 637–644, April 2006.

[LDF04] Christophe Lenglet, Rachid Deriche, and Olivier Faugeras. Inferring white matter geometry from diffusion tensor MRI : Application to connectivity mapping. In *Computer Vision-ECCV*, pages 127–140. Springer, 2004.

[LT14] Jacques-Olivier Lachaud and Boris Thibert. Properties of gauss digitized sets and digital surface integration. Technical report, hal-01070289, July 2014.

[Mér09] Quentin Mérigot. *Détection de structure géométrique dans les nuages de points*. PhD thesis, Université Nice Sophia Antipolis, December 2009.

[MNG04] Niloy J. Mitra, An Nguyen, and Leonidas Guibas. Estimating surface normals in noisy point cloud data. In *International Journal of Computational Geometry & Applications*, volume 14, pages 261–276, October 2004.

[MOG11] Quentin Mérigot, Maks Ovsjanikov, and Leonidas Guibas. Voronoi-based curvature and feature estimation from point clouds. In *IEEE Transactions on Visualization and Computer Graphics*, volume 17, pages 743–756, 2011.

[Mor08] Frank Morgan. *Geometric Measure Theory : A Beginner's Guide*. Academic Press, September 2008.

[MSR07] Evgeni Magid, Octavian Soldea, and Ehud Rivlin. A comparison of gaussian and mean curvature estimation methods on triangular meshes of range image data. In *Computer Vision and Image Understanding*, volume 107, pages 139–159, September 2007.

[MT04] Jean-Marie Morvan and Boris Thibert. Approximation of the normal vector field and the area of a smooth surface. In *Discrete & Computational Geometry*, volume 32, pages 383–400, September 2004.

[NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. In *Discrete & Computational Geometry*, volume 39, pages 419–441, March 2008.

[Par06] Claude-Eric Parveaud. *Propriétés radiatives des couronnes de Noyers (Juglans nigra x J. regia) et croissance des pousses annuelles - Influence de la géométrie du feuillage, de la position des pousses et de leur climat radiatif*. PhD thesis, Université Montpellier II - Sciences et Techniques du Languedoc, March 2006.

[PG11] Laurent Provot and Yan Gérard. Estimation of the derivatives of a digital function with a convergent bounded error. In *Discrete Geometry for Computer Imagery*, number 6607 in LNCS, pages 284–295. 2011.

[PKG03] Mark Pauly, Richard Keiser, and Markus Gross. Multi-scale feature extraction on point-sampled surfaces. In *Computer Graphics Forum*, volume 22, pages 281–289, 2003.

[PM79] F. P. Preparata and D. E. Muller. Finding the intersection of n half-spaces in time o(n log n). In *Theoretical Computer Science*, volume 8, pages 45–55, 1979.

[PWHY09] Helmut Pottmann, Johannes Wallner, Qi-Xing Huang, and Yong-Liang Yang. Integral invariants for robust geometry processing. In *Computer Aided Geometric Design*, volume 26, pages 37–60, January 2009.

[PY07] Xiaoping Qian Pinghai Yang. Direct computing of surface curvatures for point-set surfaces. In *Point-based Graphics, IEEE/Eurographics Symposium*, pages 29–36, 2007.

[RL11] Tristan Roussillon and Jacques-Olivier Lachaud. Accurate curvature estimation along digital contours with maximal digital circular arcs. In *Combinatorial Image Analysis*, volume 6636 of *LNCS*, pages 43–55. 2011.

[Ste43] Jakob Steinter. Über parallele flächen. In *Über die Steinersche Formel für Parallelflächen, Abh Math. Sem. , Hansischen Univ.*, volume 15, 1943.

[SW92] Ernest Stokely and Shang You Wu. Surface parametrization and curvature measurement of arbitrary 3-d objects : five practical methods. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 14, pages 833–840, August 1992.

[Tau95] Gabriel Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. In *Computer Vision*, 1995.

[Vil08] Cédric Villani. *Optimal Transport : Old and New*. Springer, October 2008.

[Wey39] Hermann Weyl. On the volume of tubes. In *American Journal of Mathematics*, pages 461–472. JSTOR, 1939.

[WHH10] Christopher Weber, Stéphanie Hahmann, and Hans Hagen. Sharp feature detection in point clouds. In *Shape Modeling International Conference (SMI)*, pages 175–186, 2010.

[YLL+07] Mincheol Yoon, Yunjin Lee, Seungyong Lee, Ioannis Ivrissimtzis, and Hans-Peter Seidel. Surface and normal ensembles for surface reconstruction. In *Computer-Aided Design*, volume 39, pages 408–420, 2007.

[ZHT06] Hui Zou, Trevor Hastie, and Robert Tibshirani. Sparse principal component analysis. In *Journal of Computational and Graphical Statistics*, volume 15, pages 265–286, June 2006.

[ZOF01] Hong-Kai Zhao, S. Osher, and R. Fedkiw. Fast surface reconstruction using the level set method. In *IEEE Workshop on Variational and Level Set Methods in Computer Vision, 2001. Proceedings*, pages 194–201, 2001.

[ZOMK98] Hong-Kai Zhao, Stanley Osher, Barry Merriman, and Myungjoo Kang. Implicit, non parametric shape reconstruction from unorganized points using a variational level set method. In *Cam report 98-7, UCLA*, 1998.

Discrete Geometric Inference

LOUIS CUEL

Résumé

Ces travaux s'inscrivent dans la thématique de l'inférence géométrique dont le but est de répondre au problème suivant : étant donné un objet géométrique dont on ne connaît qu'une approximation, peut-on estimer de manière robuste ses propriétés ? On se place dans cette thèse dans le cas où l'approximation est un nuage de points ou un ensemble digital dans un espace euclidien de dimension finie. On montre tout d'abord un résultat de stabilité d'un estimateur de normale basé sur l'analyse en composante principale, ainsi qu'un résultat de convergence multigrille d'un estimateur du *Voronoi Covariance Measure* qui utilise des matrices de covariance de cellules de Voronoi.

Ces deux résultats, comme la plupart des résultats en inférence géométrique, utilisent la stabilité de la fonction distance à un compact. Cependant, la présence d'un seul point aberrant suffit pour que les hypothèses des résultats de stabilité ne soient pas satisfaites. La distance à une mesure est une fonction distance généralisée introduite récemment qui est robuste aux points aberrants. Dans ce travail, on généralise le *Voronoi Covariance Measure* à des fonctions distances généralisées et on montre que cet estimateur appliqué à la distance à une mesure est robuste aux points aberrants. On en déduit en particulier un estimateur de normale très robuste. On présente également des résultats expérimentaux qui montrent une forte robustesse des estimations de normales, courbures, directions de courbure et arêtes vives. Ces résultats sont comparés favorablement à l'état de l'art.

Abstract

The purpose of geometric inference is to answer the following problem : Given a geometric object that is only known through an approximation, can we get a robust estimation of its properties ? We consider in this thesis the case where the approximation is a point cloud or a digital set in a finite dimensional Euclidean space. We first show a stability result for a normal estimator based on the principal component analysis, as well as a result of multigrid convergence of an estimator of the Voronoi covariance measure, which uses covariance matrices of Voronoi cells.

As most of geometric inference results, these two last results use the robustness of the distance function to a compact set. However, the presence of a single outlier is sufficient to make the assumptions of these results not satisfied. The distance to a measure is a generalized distance function introduced recently, that is robust to outliers. In this work, we generalize the Voronoi Covariance Measure to generalized distance functions and we show that this estimator applied to the distance to a measure is robust to outliers. We deduce a very robust normal estimator. We present experiments showing the robustness of our approach for normals, curvatures, curvature directions and sharp features estimation. These results are favorably compared to the state of the art.