Université Sidi Mohammed Ben Abdellah
Faculté des Sciences Dhar El Mahraz- Fès
Centre d'Etudes Doctorales
"Sciences et Technologies"

Formation Doctorale : Sciences et Technologies de l'Information et de Communication
Discipline : Informatique
Spécialité : Informatique
Laboratoire : Laboratoire Informatique, Imagerie et Analyse Numérique

THESE DE DOCTORAT

Présentée par

ABDERAZZAK TAIME

Nouveaux algorithmes et méthodes pour segmenter un maillage 3D en vision par ordinateur

Soutenue le 20/06/2019 devant le jury composé de :

Pr. TAIRI Hamid	**Faculté des Sciences Dhar El Mahraz Fés**	**Président**
Pr. HALLI Akram	**FSJES, Meknès**	**Rapporteur**
Pr. OUANAN Mohammed	**Faculté des Sciences, Meknès**	**Rapporteur**
Pr. SATORI Hassan	**Faculté des Sciences Dhar El Mahraz Fés**	**Rapporteur**
Pr. El ABDERRAHMANI Abdellatif	**Faculté Polydisciplinaire, Larache**	**Examinateur**
Pr. ROUKHE Ahmed	**Faculté des Sciences, Meknès**	**Examinateur**
Pr. SABRI My Abdelouahed	**Faculté des Sciences Dhar El Mahraz Fés**	**Examinateur**
Pr. SATORI Khalid	**Faculté des Sciences Dhar El Mahraz Fés**	**Directeur de thèse**
Pr. SAAIDI Abderrahim	**Faculté polydisciplinaire, Taza**	**Co-directeur de thèse**

Année universitaire : 2018-2019

Résumé

Ce travail de thèse introduit une nouvelle méthode de segmentation sémantique de modèles 3D basée sur la prise en compte de la couleur des pixels des images utilisées pendant la phase de la reconstruction 3D. Cette nouvelle technique permet de segmenter le maillage d'une manière significative et elle est beaucoup plus simple que la plupart des méthodes de segmentation de maillage basées sur les caractéristiques géométriques.

L'approche que nous avons adoptée définie une correspondance entre les couleurs des objets de la scène et le maillage 3D en exploitant le lien entre les points d'intérêt des images analysées et les sommets du maillage 3D reconstruit. Cette méthode comporte quatre étapes fondamentales :

1. La reconstruction 3D à partir des images, avec détermination de la correspondance entre les coordonnées des points 3D de la scène et leurs projections sur les plans de la caméra.
2. L'attribution d'une couleur à chaque facette du maillage.
3. La classification des facettes de couleur.
4. La segmentation du maillage.

Ainsi nous montrons à travers des évaluations qualitatives que notre approche obtient des résultats satisfaisants, reflétant son efficacité et sa performance par rapport aux méthodes de référence avec un temps de calcul nettement inférieur.

Mots-clés : Segmentation sémantique ; Segmentation géométrique ; Primitive couleur ; Points d'intérêt ; Classification.

Abstract

This thesis work introduces a new method of semantic segmentation of 3D models based on considering the pixel color of the images used during the 3D reconstruction phase. This new technique allows to segment the mesh significantly and is much simpler than most mesh segmentation methods based on geometric characteristics.

The approach we have adopted defines a correspondence between the colors of the objects of the scene and the 3D mesh by exploiting the link between the points of interest of the analyzed images and the vertices of the reconstructed 3D mesh. This method has four basic steps:

1. 3D reconstruction from the images, while determining the correspondence between the coordinates of the 3D points of the scene and their projections on the camera's planes.
2. Assigning color to each face of the mesh.
3. Classification of color facets.
4. Mesh segmentation.

Thus, we show through qualitative assessments that our approach gives satisfactory results, reflecting its efficiency and performance compared to the state-of-the-art methods with a calculation time significantly lower.

Keywords: Semantic segmentation ; Geometrical segmentation ; Color primitive ; Points of Interest ; Classification.

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au succès de ma thèse et qui m'ont aidé lors de la rédaction de ce rapport.

Tout d'abord, j'adresse mes remerciements à mon encadrant, M. Abderrahim Saaidi et mon directeur de thèse, M. Khalid Satori, Professeurs de l'Université Sidi Mohamed Ben Abdellah, pour leur apport scientifique considérable qui m'a beaucoup appris. C'est un honneur pour moi d'avoir travailler avec eux durant toutes ces années. Enfin, j'ai extrêmement sensible à leurs qualités humaines d'écoute et de compréhension tout long de ce travail doctoral. Je leur suis également reconnaissant pour avoir cru en moi, pour avoir su m'orienter et pour m'avoir soutenue généreusement à tous les stades de ce travail de thèse.

Je voudrais remercier les rapporteurs de cette thèse M. Akram HALLI et M. Mohammed OUANAN, Professeurs de l'Université Moulay Ismaïl et M. Hassan SATORI, Professeur de l'Université Sidi Mohamed Ben Abdellah pour leurs remarques et commentaires qui m'ont permis d'améliorer ce mémoire et pour l'intérêt qu'ils ont porté à mon travail.

J'associe à ces remerciements M. Abdellatif El ABDERRAHMANI, professeur de l'Université Abdelmalek Essaâdi, M. Ahmed ROUKHE, Professeur de l'Université Moulay Ismaïl et M. My Abdelouahed SABRI, Professeur de l'Université Sidi Mohamed Ben Abdellah pour avoir accepté d'examiner mon travail. Merci également à M. Hamid TAIRI, Professeur de l'Université Sidi Mohamed Ben Abdellah d'avoir présidé mon jury de thèse.

Je voudrais aussi remercier mes parents qui m'ont toujours laissé le choix de mes études et sans qui je n'aurais jamais pu mener à bout ce travail de thèse. Je remercie également ma femme, mes enfants, mes sœurs, mes frères, ainsi que mes amis pour leur soutien au cours de ces années. Leur soutien inconditionnel et leur confiance m'ont donné l'envie d'avancer et ont toujours été ma source d'inspiration.

Table des matières

Introduction Générale **8**

Problématique **8**

Contributions **9**

Organisation de la thèse **10**

Chapitre 1 - Fondements mathématiques et état de l'art **12**

1.1 Introduction **12**

1.2 Fondements mathématiques de la reconstruction 3D et de la segmentation de maillage **12**

1.2.1 La mise en correspondance stéréoscopique 12

1.2.1.1 Méthode SIFT 13

1.2.1.2 Méthodes de rejection des faux appariements 14

1.2.1.2.1 Méthode de rééchantillonnage 14

1.2.1.2.1.1 RANSAC 14

1.2.1.2.1.2 PROSAC 15

1.2.1.2.1.3 MLESAC 16

1.2.1.2.1.4 LO-RANSAC 16

1.2.1.2.2 Méthode de régression statistiques 16

1.2.1.2.2.1 LMedS 16

1.2.1.2.2.2 M-estimators 17

1.2.1.2.2.3 ICF 17

1.2.1.2.2.4 RSC 18

1.2.2 Reconstruction 3D 21

1.2.2.1 Estimation des paramètres de la caméra pour chaque image 22

1.2.2.1.1 Modèle de caméra 22

1.2.2.1.2 Autocalibrage : Géométrie à deux vues 23

1.2.2.1.2.1 Paramètres intrinsèques et extrinsèques 23

1.2.2.1.3 Géométrie à multi vues 25

1.2.2.2 Reconstruction de la géométrie 3D de la scène 26

1.2.2.2.1 Vues multiples via Structure à partir du mouvement (SFM) 27

1.2.2.2.1.1 Cadre commun 27

1.2.2.2.1.2 Ajustement de faisceaux 28

1.2.2.2.2 Stéréovision multi-vues (MVS) 28

1.2.3 Maillage 3D 30

1.2.3.1 Définitions 30

1.2.3.1.1 Maillage 30

1.2.3.1.2 Maillage triangulaire 31

1.2.3.2 Reconstruction de surface 32

1.2.3.2.1 Triangulation de Delaunay 32

1.2.3.2.2 Diagramme de Voronoï 34

1.2.3.2.3 Algorithmes de Crust 36

1.2.3.2.3.1 Définition de Crust 36

1.2.3.2.3.2 Algorithme 36

1.2.4 Segmentation de maillage 37
1.2.4.1 Formulation du problème de segmentation d'un maillage 37
1.2.4.1.1 Définition 37
1.2.4.1.2 Le problème de métrique de similarité 37
1.2.4.2 Fondement mathématique : méthodes géométriques 38
1.2.4.2.1 Ajustement des primitives 38
1.2.4.2.2 Marches aléatoires 40
1.2.4.3 Fondement mathématique : méthodes sémantiques 41
1.2.4.3.1 Fonction diamètre forme 41
1.2.4.3.2 Extraction des parties principales 43
1.2.4.3.3 Randomized cuts 45

1.3 Etat de l'art de la segmentation de maillage 3D 46
1.3.1 Méthodes géométriques 46
1.3.1.1 Ajustement des primitives 47
1.3.1.2 K-Means 47
1.3.1.3 Normalized cuts 48
1.3.1.4 Marches aléatoires 48
1.3.2 Méthodes sémantiques 49
1.3.2.1 Fonction diamètre forme 49
1.3.2.2 Extraction core 50
1.3.2.3 Randomized cuts 51
1.3.2.4 Autres approches 52

1.4 Conclusion 53

Chapitre 2 - Nouvelles méthodes de mise en correspondance 54

2.1 Introduction 54

2.2 Identification des correspondances de points par les cercles correspondants ICC 55
2.2.1 Méthode proposée 55
2.2.1.1 Principe 55
2.2.1.2 Moyens de vérification 56
2.2.1.3 Extension des cercles dans la scène 3D 56
2.2.1.3.1 Répartition virtuelle d'une scène 57
2.2.1.3.2 Critères de localisation des valeurs aberrantes 58
2.2.1.3.3 Hypothèse et preuve 59
2.2.1.4 Prétraitement 61
2.2.1.5 Algorithme de l'approche proposée 63
2.2.1.6 Complexité de calcul 65

2.3 Identification rapide des correspondances de points par les cercles correspondants FICC 65
2.3.1 L'approche adoptée 66
2.3.1.1 Limitation de la qualité ICC 68
2.3.1.2 Limitation d'efficacité de ICC 68
2.3.1.3 Amélioration de la qualité du FICC 69
2.3.1.4 Amélioration de l'efficacité du FICC 70
2.3.1.5 Complexité de calcul 71
2.3.2 Expérimentation et résultats 73
2.3.2.1 Mise en correspondance de seize paires d'images de différentes bases de données 75
2.3.2.2 Correspondance de caractéristiques basée sur la base de données Mikolajczyk 78

2.3.2.3 Résultats concernant le temps d'exécution 83
2.3.3 Discussion 83
2.4 Conclusion 84
Chapitre 3 - Segmentation significative des maillages 3D complexes via la primitive couleur des images stéréoscopiques 85
3.1 Introduction 85
3.2 Etapes de la segmentation significative via la primitive couleur 86
3.2.1 Reconstruction 3D 88
3.2.2 Attribution des couleurs au maillage 88
3.2.3 Classification des couleurs RVB des facettes 89
3.2.4 La segmentation du maillage 90
3.2.5 Algorithme de la méthode proposée 90
3.3 Résultats expérimentaux et étude comparative 91
3.3.1 Critères d'évaluation 93
3.3.2 Séquence 1 94
3.3.3 Séquence 2 95
3.3.4 Séquence 3 98
3.3.5 Séquence 4 99
3.3.6 Discussion 101
3.4 Conclusion 103
Conclusion générale 104
Table des figures 106
Liste des tableaux 110
Liste des symboles 111
Bibliographie 114
Liste des publications 120

Introduction Générale

Au cours des dernières années, les maillages 3D sont les modèles surfaciques les plus utilisés et qui ne cessent d'avoir une place prépondérante au fil des années. En effet, ils offrent une représentation tridimensionnelle des objets. Le maillage 3D est aussi adapté à la conception des formes. Le maillage triangulaire plus particulièrement représente une solution privilégiée pour diverses modélisations grâce à sa simplicité et sa flexibilité. L'analyse de haut niveau des modèles 3D est une étape primordiale pour plusieurs applications, telles que la modélisation des formes par composition, le mappage de texture, la reconnaissance de formes et l'indexation du modèle par parties. Ces applications exigent la compréhension de l'information sémantique associée à ces modèles 3D.

Dans la littérature, la majorité des algorithmes de segmentation de maillage proposés sont basés sur les caractéristiques géométriques des objets. En effet, ils s'appuient uniquement sur l'ensemble des points 3D extraits à partir de la mise en correspondances des images. Mais on peut noter que ces méthodes ignorent les autres informations contenues dans les images (par exemple la primitive couleur) qui ont un rôle important pour l'amélioration de l'efficacité de la segmentation du maillage. C'est dans ce contexte, où la rareté des méthodes reliant la segmentation du maillage directement aux images mises en jeu, que nous avons proposé une nouvelle approche de segmentation. Cette approche permet donc d'établir un lien entre le maillage reconstruit et les images mises en jeu, ainsi la nécessité d'élaborer une méthode qui prenne en considération les différentes étapes de reconstruction 3D des objets en partant de la mise en correspondance des images stéréoscopiques.

Problématique

De nombreux algorithmes ont été développés pour le traitement spécifique des maillages 3D en vue d'extraire automatiquement leurs parties significatives. Ces algorithmes sont basés sur l'analyse globale d'un maillage pour décrire la sémantique de l'objet [1]–[4].

Il faut noter que la segmentation automatique des maillages aux parties significatives est un problème fondamental en infographie. La plupart des méthodes s'appuient essentiellement sur les attributes de surface des objets, incluant les directions normales, la courbure et les distances géodésiques moyennes. Ces attributes de surface dépendent souvent de leur

topologie et de la pose de l'objet, cela peut induire la segmentation du maillage du même objet à des résultats différents. D'autre part, ces méthodes donnent généralement des décompositions plus faciles à interpréter, particulièrement en cas d'utilisation des modèles avec des structures relativement simples. Par exemple, des modèles des animaux ou des humains aux parties sémantiques spécifiques, e.g., la main, la tête, le tronc et les jambes, en se basant sur des techniques telles que la squelettisation, les distances géodésiques, etc. Cependant, la partition des objets aux parties significatives, qui vise à obtenir des résultats similaires à ceux obtenus par les humains, reste un problème, particulièrement, dans le cas du maillage ayant une géométrie complexe.

Les caractéristiques géométriques de l'objet considérés généralement un très bon moyen pour la segmentation, paradoxalement, ils représentent un défi pour la meilleure segmentation aux parties significatives visuelle, par exemple, segmenter un bouquet de fleurs comme objet aux parties spécifiques qui pourraient nous permettre de distinguer clairement les fleurs les unes des autres.

En générale, les méthodes géométriques de segmentation de maillage s'articulent autour des sommets de maillage en les considérant comme les seules données initiales. En revanche, les images stéréoscopiques contiennent encore d'autres informations comme l'information « couleur », et qui peuvent aussi être utilisées pour une segmentation plus pertinente du maillage.

Contributions

Afin de remédier aux limitations citées ci-dessus, nous avons proposé trois méthodes. La première méthode proposée permet de segmenter le maillage en parties significatives. La partie principale de cette méthode consiste en un découpage efficace du maillage en se basant sur une classification des couleurs des images mises en jeu dans l'étape de la reconstruction 3D du modèle. De cette manière, on garantit une segmentation réellement significative, tout en remédiant aux problèmes des autres approches qui concernent la complexité de la géométrie de la scène et la détermination précises des régions en parties significatives. En effet, l'information fournie par la couleur représente le moyen le plus naturel pour segmenter les objets, permettant ainsi une bonne perception visuelle de ces objets. Par exemple, on pourrait facilement distinguer les fleurs dans un bouquet de fleurs par l'information de la couleur plutôt que par les caractéristiques géométriques. Cette première méthode proposée

exige d'établir une liaison entre la scène 3D et les images 2D afin d'utiliser les couleurs des pixels, d'où l'importance de nôtre deuxième et troisième contribution qui concernent l'étape cruciale de la mise en correspondance. Les principales contributions de ce travail sont les suivantes :

- Elaboration d'une méthode de la mise en correspondance stéréoscopique appelé ICC. Cette méthode vise à éliminer les faux appariements à partir d'un ensemble des points de correspondances putatives tout en utilisant des cercles correspondants. Cette contribution a été l'objet d'une publication intitulée : «**Robust point matching via corresponding circles**» dans le journal **Multimedia Tools and Applications (Springer, 2018).**
- Amélioration de la méthode précédente pour une meilleure efficacité, en proposant une nouvelle version plus rapide et plus efficace appelée FICC : «**Fast point matching using corresponding circles**» dans le journal **Multimedia Tools and Applications (Springer, 2018).**
- Elaboration d'une méthode de la segmentation permettant de segmenter le maillage en parties significatives. Cette contribution a été l'objet d'une publication intitulée : «**A new semantic segmentation approach of 3D mesh using the stereoscopic image colors**» dans le journal **Multimedia Tools and Applications (Springer, 2018).**

Organisation de la thèse

Cette thèse s'articule autour de trois chapitres.

Dans le **chapitre 1** nous allons présenter les différentes étapes et les fondements mathématique sur lesquels sont fondés la reconstruction 3D et la segmentation de maillage. Ensuite, nous présenteront un état de l'art concernant les principales méthodes de segmentation de maillage 3D.

Le **chapitre 2** présente nos deux contributions concernant la mise en correspondance. Cette dernière est l'étape préliminaire la plus critique dans tout processus de la stéréovision incluant la segmentation de maillage. Ces deux méthodes contribuent à l'amélioration de la segmentation de maillage. La première méthode vise à améliorer la qualité des appariements en rejetant les faux appariements contaminant un ensemble de points de correspondances putatives, tout en se basant sur les cercles correspondants. La deuxième méthode vise à

améliorer la méthode précédente en utilisant de nouveaux critères permettant de produire une précision plus pertinente tout en minimisant la complexité de l'algorithme.

Le **chapitre 3** sera consacré à la description des étapes de notre méthode de segmentation. Comparativement aux méthodes classiques, une nouvelle stratégie complètement différente a été adoptée. En effet, notre méthode tient compte des différentes étapes de la reconstruction 3D, depuis l'étape de la mise en correspondance des images stéréoscopiques jusqu'à la phase de maillage. En fait, elle est basée sur l'attribution d'une couleur, extraite à partir des images, à chaque triangle du maillage. En exploitant le lien entre les sommets du maillage 3D et les points d'intérêt des images mises en jeu. Ensuite, une classification de l'ensemble des triangles de maillage est réalisée en se basant sur la couleur affectée à chaque triangle. Finalement, cette classification nous permet de segmenter le maillage en parties significatives. L'algorithme élaboré a été implémenté en utilisant le langage java et la bibliothèque JAMA.

Chapitre 1 - Fondements mathématiques et état de l'art

1.1 Introduction

La méthode de segmentation proposée, comme il est indiqué dans l'introduction générale, est basée sur une connaissance à priori sur le lien entre les sommets du maillage 3D et les points d'intérêt des images mises en jeu. Par conséquent on tient compte des différentes étapes de la reconstruction 3D, depuis l'étape de la mise en correspondance des images jusqu'à la phase de maillage. Dans ce chapitre, nous allons présenter les différentes étapes et les fondements mathématiques concernant la reconstruction 3D et la segmentation de maillage. Ensuite, nous allons présenter l'état de l'art concernant les méthodes de segmentation de maillages 3D.

1.2 Fondements mathématiques de la reconstruction 3D et de la segmentation de maillage

Dans cette section, nous allons présenter les différentes étapes et les fondements mathématiques concernant la reconstruction 3D et la segmentation de maillages.

1.2.1 La mise en correspondance stéréoscopique

Le problème de la mise en correspondance, comme celui de la segmentation, est l'un des problèmes les plus épineux, et bien sûr loin d'être résolu, en vision par ordinateur. De manière générale, la mise en correspondance stéréoscopique consiste à retrouver dans deux images d'une même scène les pixels homologues, c'est-à-dire, les pixels qui représentent la projection de la même entité de la scène. La mise en correspondance est appliqué dans de nombreux domaines tels que la construction de carte et reconnaissance aérienne [5], la navigation robotique [6], la reconstruction d'objets [7], la reconnaissance de concept visuel [8], le tatouage pour la vidéo 3D [9], la détection de petites cibles infrarouge (IR) [10], etc.

Une stratégie fréquemment adoptée pour résoudre le problème de mise en correspondance consiste à utiliser un processus en deux étapes. La première étape, un ensemble des correspondances putatives sont calculés en utilisant une contrainte de similarité.

Cet ensemble de correspondances putatives comprenant généralement la plupart des vrais appariements, mais aussi un nombre considérable de faux appariements, ou des aberrants, dues à des ambiguïtés liées aux contraintes de similarité (particulièrement si les images contiennent des motifs répétitifs). La deuxième étape vise à éliminer les points aberrants [11]–[14]. Nous nous intéressons dans ce manuscrit surtout aux méthodes de rejet des faux appariements d'un ensemble de points de correspondances.

En effet, dans la première étape un ensemble de correspondances putatives sont calculées en utilisant une contrainte de similarité entre deux ensembles de points discrets dont chaque point est extrait par un détecteur des caractéristiques locales. Comme l'un des plus populaires descripteurs [15]–[18], le SIFT (Scale Invariant Feature Transform) [18] est robuste à l'éclairage, au changement affine, au changement d'échelle, et aux distorsions locales [18].

1.2.1.1 Méthode SIFT

L'extraction des caractéristiques invariantes distinctives des images représente une approche efficace afin d'établir une correspondance fiable entre des différentes vues d'un objet ou d'une scène. Dans ce stade, l'algorithme SIFT proposée par DAVID G. LOWE figure parmi les plus importants. Cet algorithme rend les images convenables pour réaliser la mise en correspondance, en outre, les caractéristiques extraites sont invariantes par rapport au changement d'échelle des images, à la rotation, partiellement au changement d'éclairage et au changement de point de vue de la caméra.

Voici donc les principales étapes de calcul utilisées pour générer un ensemble des caractéristiques de l'image :

1. Détection d'extrema échelle-espace : cette première étape vise la recherche toutes les échelles. Elle est mise en œuvre en utilisant une fonction de différence gaussienne afin d'identifier les points d'intérêts potentiels, invariants à l'échelle et à la direction.

2. Localisation des points-clés : pour chaque position candidate, un modèle détaillé est adapté permet de déterminer la position et l'échelle. Les points-clés sont sélectionnés en se basant sur les mesures de leur stabilité.

3. Attribution de direction : un ou plusieurs directions sont attribuées à chaque position de point-clé en se basant sur les directions des gradients locaux de l'intensité de l'image.

4. Descripteur de point-clé : les gradients locaux d'intensité de l'image sont évalués selon l'échelle sélectionnée dans la région autour de chaque point-clé. Ils sont

transformés en une représentation qui permet un certain niveau de distorsion de forme et du changement de l'éclairage.

Les points-clés SIFT sont particulièrement utiles en raison de leur caractère distinctif, permettant de sélectionner les vrais appariements parmi une importante base de données d'autres points-clés. Ce caractère distinctif est atteint par assemblage d'un vecteur représentant les gradients d'image à l'intérieur d'une région locale. Les points-clés montrent qu'ils sont invariants à la rotation, à l'échelle d'image, à la distorsion affine et au bruit.

1.2.1.2 Méthodes de rejection des faux appariements

Malgré la robustesse de certaines descripteurs locaux comme SIFT [18], il est quasiment impossible d'avoir un ensemble des correspondances complètement correct ; due par exemple au bruit. Pour remédier à ce problème, une autre étape de rectification semble indispensable, dans ce cadre plusieurs travaux proposent des méthodes visant à éliminer les pairs de points aberrants. Ces méthodes peuvent être classifiées en deux catégories principales : les méthodes de rééchantillonnage et les méthodes de régression statistique.

1.2.1.2.1 Méthode de rééchantillonnage

Les méthodes de rééchantillonnage peuvent éliminer les aberrants par génération répétitive d'hypothèses fondée sur l'estimation d'un modèle paramétrique à partir d'un nombre minimum de points de correspondances choisie au hasard. Dans cette section on présente les méthodes de rééchantillonnage les plus utilisées dans ce stade :

1.2.1.2.1.1 RANSAC

RANSAC [11] est un algorithme à la fois simple et puissant que l'on applique généralement pour l'estimation des paramètres d'un modèle en utilisant une base de données qui est généralement contaminée par des points aberrants. RANSAC estime la relation globale qui décrit mieux les données, tout en regroupant les données en vrais appariements et faux appariements (aberrants).

RANSAC fonctionne dans le cadre d'hypothèse-et-vérification. Etant donnée, un ensemble $\mathcal{U}$ contenant N correspondances putatives que l'on suppose déjà obtenues. De manière itérative, des n uplet $\mathcal{U}'$ sont échantillonnés aléatoirement parmi $\mathcal{U}$, où n est le nombre minimal requis afin de trouver la solution. Lorsque l'hypothèse a été émise à partir de ce sous-ensemble $\mathcal{U}'$, elle sera vérifiée en regard de l'ensemble des points $\mathcal{U}$. On recommence le processus jusqu'à ce qu'un critère d'arrêt soit satisfait. Ce critère repose sur le

nombre minimal d'échantillons requis afin d'assurer un certain niveau de confiance η_0, sous condition qu'aumoins l'un des points sélectionnés de $\mathcal{U}'$ soit un aberrant libre. Etant donné le ratio des vrais appariements ε, la probabilité de sélectionner un échantillon non contaminé de n appariements est ε^n . La probabilité de sélectionner k échantillons dont chacun est contaminé avec au moins un aberrant, est donnée par $(1-\varepsilon^n)^k$. Par conséquence, le nombre minimum des échantillons qu'il faut établir en vue de s'assurer que cette probabilité soit inférieure au seuil $1-\eta_0$ est :

$$k \geqslant \frac{\log(1-\eta_0)}{\log(1-\varepsilon^n)} \tag{1.1}$$

La qualité d'un échantillon est définie par son cardinal. L'algorithme s'arrête au terme de i_{max} itérations. Le meilleur échantillon identifié est validé si son cardinal est plus grand qu'un second seuil.

En revanche, le nombre d'échantillons requis augmente de façon exponentielle, et par conséquence le coût de calcul associé est important [19], ce qui représente un inconvénient majeur.

En fait, l'algorithme de RANSAC génère des hypothèses en échantillonnant uniformément l'ensemble de données d'entrée. Par conséquence, RANSAC suppose implicitement qu'il n'y a pas d'information à priori disponible sur la précision des points de données, ce qui n'est pas forcément vrai et dans plusieurs cas pratiques la réalité a montré l'inverse. En effet, une information à priori est souvent disponible, et peut être utilisée pour générer de meilleures hypothèses.

En vue d'améliorer la Génération d'hypothèses plusieurs méthodes sont apparues.

1.2.1.2.1.2 PROSAC

Les correspondances entre deux images ou plus sont obtenues en utilisant un algorithme de correspondance local. Une fonction de similarité est évaluée sur un certain nombre de points, puis consécutivement seuillée pour obtenir un ensemble de correspondances provisoires. En partant de l'hypothèse que les points présentant une forte similarité soient plus susceptibles d'être vrais appariements que les points faiblement similaires. Il est possible de générer de meilleures hypothèses en échantillonnant à partir d'un ensemble réduit de points ayant des scores de similitude élevés. Ce fait est exploité dans [20], où les correspondances sont ordonnées en fonction de leurs scores de similarité.

1.2.1.2.1.3 MLESAC

RANSAC procède en générant de manière répétée des solutions estimées à partir d'ensembles minimaux de correspondances collectés à partir des données, puis en testant chaque solution à l'aide de l'ensemble complet des correspondances putatives.

Tordoff et Murray proposent un nouvel estimateur qui prend en charge la vraisemblance logarithmique de la solution (en tenant compte de la distribution des valeurs aberrantes) et utilise un échantillonnage aléatoire pour maximiser la vraisemblance. La nouvelle méthode d'échantillonnage Maximum Likelihood Estimation Sample Consensus MLESAC est une généralisation de RANSAC, adopte une nouvelle fonction de coût utilisant une stratégie de vote pondérée, tout en se basant sur M-estimation et en choisissant la solution qui maximise la probabilité plutôt que le nombre d'appariement.

1.2.1.2.1.4 LO-RANSAC

RANSAC standard présume qu'un modèle calculé à partir d'un échantillon non contaminé est cohérent avec tous les vrais appariements. En pratique, ce n'est souvent pas le cas, en particulier lorsque les points de données sont bruyants. Chum et al. [21] définissent une variante RANSAC localement optimisée pour traiter ce problème. En observant qu'un bon modèle a tendance à trouver une fraction significative des vrais appariements, une stratégie interne de RANSAC est conçue en définissant un nombre constant d'hypothèses (~ 20) générées en utilisant seulement l'ensemble des vrais appariements au meilleur modèle actuel.

Bien que ces méthodes aient beaucoup de succès dans de nombreuses situations, elles ont eu un succès limité si les contraintes géométriques sont non-paramétriques et elles ont aussi tendance à se dégrader si la proportion de valeurs aberrantes dans l'ensemble de correspondance putative devient importante [13].

1.2.1.2.2 Méthode de régression statistiques

Les méthodes de régression statistique visent à éliminer les aberrants en se basant sur la somme des erreurs quadratiques. Quatre approches figurent parmi les méthodes de régression statistiques les plus performantes :

1.2.1.2.2.1 LMedS

LMedS est une méthode robuste proposée d'abord en statistiques [22]. LMedS dépend du concept de la minimisation de la médiane quadratique. Dans cette méthode l'hypothèse est

évaluée sur la base de résidu médian des correspondances putatives. Deriche et Zhang et al. [23], [24] ont abordé le problème des faux appariements avec LMedS pour l'estimation de la matrice fondamentale. Torr et Murray [25] ont étudié l'estimation de la matrice fondamentale et la segmentation du mouvement. LMedS peut gérer un grand pourcentage de valeurs aberrantes, mais le principal défaut est sa faible efficacité de calcul.

1.2.1.2.2.2 M-estimators

Un autre type de méthodes robustes de régression consiste aux M-estimators (Maximum-likelihood estimators) [26]–[28]. Elles dépendent du concept de la somme des fonctions symétriques à définition positive de la minimisation des résidus. Reza et Alireza [26] ont étudié l'estimation de la matrice fondamentale et la segmentation du mouvement avec des M-estimateurs. Chen et al [28] proposent une méthode d'élimination des faux appariements basé sur l'estimation de la relation spatiale cohérente d'une paire d'images donnée. Il estime alternativement l'ensemble des vrais appariements et récupère la relation d'image dans le cadre EM Cependant, les M-estimateurs nécessitent une bonne estimation initiale pour le ou les paramètres à estimer.

1.2.1.2.2.3 ICF

Li et al [13] proposent une méthode permettant d'éliminer les aberrants en utilisant une fonction de correspondance de régression à vecteurs de support qui cartographie des points d'une image à leurs correspondantes dans une autre. Ils rejettent les aberrants en vérifiant s'ils sont compatibles avec les estimations de la fonction de correspondance.

L'idée fondamentale réside repose sur deux fonctions qui sont définies pour relier les points correspondants entre deux images en mappant un point dans une image à son point correspondant dans une autre. Etant données, par exemple, deux images I et I' d'une scène, nous pouvons définir une fonction f, par laquelle un point $p = (u, v) \in I$ pourrait être mappé à un point $p' = (u', v') \in I'$ où p et p' sont une paire de points correspondants entre les images I et I'. Les fonctions f et f' peuvent être définies comme suit :

$$\begin{cases} f : I \rightarrow I' \\ f' : I' \rightarrow I \end{cases} \tag{1.2}$$

Les fonctions de correspondance (FC) f et f' sont deux fonctions vectorielles entre une paire d'images I et I'. Elles peuvent être réécrites comme suit :

$$f(u, v) = (u', v') = \big(g_1(u, v), g_2(u, v)\big) \tag{1.3}$$

$$f'(u',v') = (u,v) = \left(g_1'(u',v'), g_2'(u',v')\right) \quad (1.4)$$

Où g_i et g_i' sont des fonctions scalaires avec deux variables, i=1, 2. Elles sont appelées fonction de composant FC.

Pour faire face à l'exception de la contrainte d'unicité, une autre fonction de I à I′ est nécessaire comme complémentarité de f. Ainsi, tous les points correspondants corrects sont cohérents avec au moins l'une des deux fonctions f et f′, où f et f′ sont appelées fonctions de correspondance et (f, f′) sont appelées une paire de fonctions de correspondance (FC). Ainsi, cette paire de fonctions est utilisé pour identifier la validité de chaque correspondant putatif.

1.2.1.2.2.4 RSC

Chen et al proposent une méthode [28] pour éliminer les faux appariements entre les points de correspondances putatives dans les paires d'images en se basant sur «Relations Spatiales Cohérentes». Dans le cadre bayésien, on formule cette approche comme un problème de maximum de vraisemblance et l'on résout une relation spatiale cohérente entre les points de correspondances putatives en utilisant l'algorithme EM (Espérance-Maximisation). Cette approche associe chaque correspondance de points à une variable latente qui évalue cette paire comme vrai appariement ou une paire aberrante, et estime alternativement les vrais appariements et récupère la relation spatiale cohérente. Pour paramétrer la relation spatiale cohérente, nous choisissons la géométrie à deux vues comme modèles pour les cas rigides et non rigides. Les faux appariements pourraient être éliminés avec succès grâce aux relations spatiales cohérentes après la convergence de l'algorithme EM.

Dans la section suivante, on présente une formulation de maximum de vraisemblance pour résoudre le problème de suppression de faux appariements et présenter une solution EM. Ensuite, on décrive la méthode de récupération de la relation d'image et on discute la complexité du calcul.

- Formulation de vraisemblance maximale

Étant donné un ensemble de points de correspondance putatives $S = \{(u_n, v_n)\}_{n=1}^{N}$ en deux vues qui peuvent être perturbées par le bruit et les paires aberrantes, l'objectif est d'éliminer les paires aberrantes contenues dans l'ensemble de points afin d'établir des points de correspondances fiables.

On note M comme la relation spatiale cohérente entre l'ensemble des vrais appariements, et $M(u, v) = 0$ pour un vrai appariement (u, v). Pour cette méthode, on utilise les coordonnées homogènes pour l'ensemble de points d'image, c'est à dire $u = (u^x, u^y, 1)^T$.

Évidemment, si on récupère avec succès la relation spatiale cohérente M, les paires aberrantes peuvent être facilement distinguées et nous obtenons des correspondances précises. Cependant, pour estimer la relation spatiale cohérente, il faut également des points correspondances fiables. Pour résoudre ce dilemme, on formule le problème sous la forme d'un problème de maximum de vraisemblance et, alternativement, on récupère la relation spatiale cohérente en estimant la valeur initiale dans un cadre EM.

Sans perte de généralité, on fait l'hypothèse que le bruit sur les vrais appariements est gaussien sur chaque composante avec une moyenne nulle et un écart-type uniforme σ, et la distribution des paires aberrantes est uniforme 1/a, où a est juste une constante. Soit γ le pourcentage des vrais appariements. Ainsi, la probabilité est une densité mélange comme

$$p(S|\theta) = \prod_{n=1}^{N} \left(\frac{\gamma}{2\pi\sigma^2} e^{\frac{\|M(u_n, v_n)\|^2}{2\sigma^2}} + \frac{1-\gamma}{a} \right) \tag{1.5}$$

Où $\theta = \{M, \sigma^2, \gamma\}$ est un ensemble des paramètres inconnus.

En général, les paramètres exacts de l'ensemble θ maximisent la vraisemblance(1.5). Maintenant, on donne une estimation du maximum de vraisemblance de, c.-à-d., $\theta^* = \text{argmax}_\theta \, p(S|\theta)$. Cela équivaut à rechercher l'énergie minimale

$$E(\theta) = -\sum_{n=1}^{N} \text{In } p(u_n, v_n|\theta) \tag{1.6}$$

La relation spatiale cohérente M sera directement obtenue à partir de la solution optimale θ^*.

Il y a plusieurs manières d'estimer les paramètres de la densité mélange, tels que l'algorithme EM, la descente de gradient et l'inférence variationnelle. L'algorithme EM fournit un cadre naturel pour résoudre ce problème. L'étape Estimation estime essentiellement la responsabilité qui indique dans quelle mesure un échantillon appartient à l'ensemble précédent sous la relation spatiale cohérente, tandis que l'étape Maximisation met à jourles paramétres en fonction de l'estimation actuelle de la responsabilité.

Considérant la fonction de vraisemblance logarithmique négative, c.-à-d. l'équation (1.6), on associe l'échantillon n à une variable latente $z_n \in \{0, 1\}$, où $z_n = 1$ indique une distribution gaussienne et $z_n = 0$ pointe vers une distribution uniforme. Le log-vraisemblance est alors donnée par

$$\mathcal{Q}(\theta, \theta^{old}) = -\frac{1}{2\sigma^2}\sum_{n=1}^{N} P(z_n = 1|u_n, v_n, \theta^{old})\|M(u_n, v_n)\|^2$$

$$-In\,\sigma^2 \sum_{n=1}^{N} P(z_n = 1|u_n, v_n, \theta^{old})$$

$$+In\,\gamma \sum_{n=1}^{N} P(z_n = 1|u_n, v_n, \theta^{old})$$

$$+In\,(1-\gamma) \sum_{n=1}^{N} P(z_n = 0|u_n, v_n, \theta^{old}) \qquad (1.7)$$

Cela peut être maximisé en traitant le z_n comme les données manquantes à partir de densité mélange dans l'équation (1.5).

E-étape : Notons $P = diag(p_1, \dots, p_2)$, où la responsabilité $p_n = P(z_n = 1|u_n, v_n, \theta^{old})$ peut être calculée en appliquant la règle de Bayes :

$$p_n = \frac{\gamma e^{\frac{\|\epsilon_n\|^2}{2\sigma^2}}}{\gamma e^{\frac{\|\epsilon_n\|^2}{2\sigma^2}} + 2\pi\sigma^2(1-\gamma)/a}, \qquad (1.8)$$

Où $\epsilon_n = M(u_n, v_n)$. La probabilité à postériori p_n est une décision souple indiquant dans quelle mesure l'échantillon n est compatible avec les paramètres estimés.

M-étape : On détermine l'estimation de paramètre révisée $\theta^{nouveau}$ comme suit :

$$\theta^{nouveau} = \text{argmax}_\theta \mathcal{Q}(\theta, \theta^{old}) \qquad (1.9)$$

En prenant la dérivée de $\mathcal{Q}(\theta)$ par rapport à la variance σ^2 et au coefficient de mélange γ et en les mettant à zéro, on obtient

$$\sigma^2 = \frac{\sum_{n=1}^{N} P_n \|\epsilon_n\|^2}{\sum_{n=1}^{N} P_n} \qquad (1.10)$$

$$\gamma = \sum_{n=1}^{N} P_n / N \qquad (1.11)$$

Une fois que l'algorithme EM converge, on obtient la relation spatiale cohérente M. On peut alors éliminer les faux appariements en vérifiant si elles sont cohérentes avec la relation spatiale cohérente.

La mise en correspondance est un problème qui n'est pas facile à résoudre et il n'existe pas jusqu'à présent un algorithme unique permettant de la réaliser. Nous avons présenté les contributions les plus importantes relatives à cette problématique, en distinguant deux catégories de méthodes. La mise en correspondance représente la phase la plus délicate dans la reconstruction 3D qui sera abordée dans la section suivante.

1.2.2 Reconstruction 3D

Récemment, les systèmes modernes de stéréovision multi-vues (MVS) [29], [30] pourraient être utilisé pour calculer une représentation dense des objets complexes par génération des maillages surfaciques de haute précision à partir des images calibrées : en effet, une étude récente [31] a montré que plusieurs algorithmes atteint environ 96% de la surface traitée et une précision de la profondeur d'environ 0.4 mm d'un objet ayant 10 cm de diamètre capturé par 16 caméras à basse résolution (640 x 480). Cette technologie prometteuse est inégalée dans les tâches de modélisation à partir des images, cependant elle impose des exigences énormes relatives à la procédure de calibrage utilisée pour estimer les paramètres intrinsèques et extrinsèques, les coefficients de distorsion, etc.

La structure à partir du mouvement (SFM) est l'une des principales approches qui font face au problème de calibrage. En effet, SFM permet de retrouver le calibrage et la pose de la caméra ainsi que la structure de la scène [32], [33]. Dans ce contexte, les paramètres intrinsèques sont souvent supposés à priori connus [32], ou récupérés a posteriori à travers l'autocalibrage [34]. Une dernière étape d'ajustement de faisceaux (BA) est habituellement utilisée pour affiner les positions des points de la scène et l'ensemble des paramètres de la caméra (y compris les paramètres intrinsèques et, éventuellement, les coefficients de distorsion) [35].

Furakuwa et al [29] propose un algorithme BA simple et efficace qui prend un ensemble d’images, et les paramètres de la caméra qui pourraient être obtenus par le système SFM avec certaines estimations prudentes des erreurs de reprojection comme données d'entrées.

Dans les sections qui suivent, les approches de récupération des paramètres intrinsèques et extrinsèques de la caméra à partir de deux ou plusieurs images et les algorithmes SFM

seront traitées. Par la suite, un processus de reconstruction de la géométrie 3D via MVS sera présenté.

1.2.2.1 Estimation des paramètres de la caméra pour chaque image

Une caméra peut être modélisée à l'aide des paramètres intrinsèques et extrinsèques. Les paramètres intrinsèques décrivent les propriétés optiques et les paramètres extrinsèques décrivent la position de la caméra dans la scène 3D. L'estimation de ces paramètres est appelée calibrage de la caméra.

1.2.2.1.1 Modèle de caméra

Dans ce travail, la projection générale du modèle de caméra sténopé [36] est utilisé pour l'estimation des paramètres. Une caméra sténopé est modélisée suivant sa distance focale f qui représente la distance entre le centre camera et le plan image. C'est le modèle couramment utilisé avec les appareils photo numériques autonomes qui capturent des photographies. Toute matrice de projection P de taille 3×4 peut être décomposée en produit d'une matrice triangulaire supérieure K de taille 3×3 et d'une matrice de pose [R|T] de taille 3×4.

En effet, un point 3D M est relié à un point m associé dans l'image par l'équation (1.12) suivante :

$$m = PM \text{ où } P = \underbrace{\begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}}_{K} \cdot \underbrace{\left(\begin{array}{ccc|c} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{array}\right)}_{R \quad T} \tag{1.12}$$

La matrice K est communément appelée matrice intrinsèque, car elle est composée de grandeurs intrinsèques de la caméra : les distances focales verticales et horizontales (f_x, f_y), le point principal $[c_x, c_y]^T$, et l'inclinaison s. La matrice [R|T] est appelée matrice extrinsèque, où R et T sont respectivement la rotation et la translation de la caméra. Notez que, en raison de la qualité des capteurs numériques, on estime rarement les 11 paramètres de la matrice de projection. En particulier, les pixels sont supposés n'avoir aucune inclinaison $(s = 0)$, et être carrés $(f_x = f_y)$. De même, si une image n'a pas été recadrée, il est prudent de supposer que le point principal est au centre de l'image. En conséquence, un modèle de caméra à sténopé commun est composé de 7 paramètres : la distance focale f, la matrice de rotation R et le vecteur de translation T.

Afin de récupérer les paramètres intrinsèques et extrinsèques, l'autocalibrage est l'une des approches la plus utilisée dans ce stade.

1.2.2.1.2 Autocalibrage : Géométrie à deux vues

Dans la pratique, les paramètres intrinsèques sont inconnus, et les appariements sont les seules informations qu'on peut extraire à partir d'une séquence d'images. L'autocalibrage consiste à calculer les paramètres intrinsèques commençant par les points de correspondances.

1.2.2.1.2.1 Paramètres intrinsèques et extrinsèques

En fait, deux procédures principales sont nécessaires pour réaliser l'autocalibrage soit on se base sur les équations de kruppa soit on utilise la matrice essentielle. La matrice essentielle est équivalente à kruppa d'un point de vue mathématique, mais elle est plus facile à manipuler par rapport aux procédures numériques. Une différence importante entre les deux approches, concernant la nature des contraintes d'autocalibrage, est que les contraintes de l'autocalibrage basées sur la matrice essentielle sont de nature algébrique, par contre celles de kruppa sont de nature géométrique. Dans la partie suivante, grâce à sa simplicité, nous étudions surtout la matrice essentielle.

- Paramètres intrinsèques

La contrainte de Huang et Faugeras [37] peut être utilisée afin de dériver la fonction de coût des distance focales (f_x, f_y) par unité de pixels. Lorsque les paramètres intrinsèques encodés par la matrice de calibrage K sont supposés fixes entre deux vues liées à la matrice fondamentale F, il en résulte que la matrice essentielle E peut s'écrire sous la forme :

$$E = K^T F K. \tag{1.13}$$

Cette relation peut être exploitée pour dériver la fonction de coût des distances focales. On obtient l'équation suivante :

$$\mathrm{tr}(K^T F K K^T F^T K)^2 - 2\mathrm{tr}((K^T F K K^T F^T K)^2) = 0 \tag{1.14}$$

Cette équation comporte tous les paramètres intrinsèques de la caméra [38].

- Paramètres extrinsèques

Cette section aborde le problème du calcul de la position et de l'orientation relative à la caméra dans l'espace de la scène 3D. D'abord, les paramètres extrinsèques dans le cas de deux caméras ont été étudiés, ce qui est la condition minimale pour une reconstruction de scène métrique. La méthode décrite ici est basée sur la propriété ESV de la matrice essentielle présentée au-dessus dans la dernière section.

On suppose que la première position de la caméra se trouve à l'origine du cadre de coordonnées et que sa rotation relative est décrite par $R_1 = I$ où I est la matrice d'identité 3×3. Ainsi, la matrice de projection de cette caméra a la forme $P_1 = K[I\ 0]$. Le but est alors de trouver les paramètres extrinsèques R et t de la deuxième caméra qui est la composante rotationnelle et translationnelle du mouvement respectivement. A partir des paramètres intrinsèques connus et du mouvement estimé, la matrice de projection $P_2 = K\,[R\ t]$ peut être assemblée.

Si la matrice fondamentale d'une image est calculée et les paramètres intrinsèques de la caméra sont estimées en utilisant l'algorithme classique d'autocalibrage, la matrice essentielle E peut être calculée directement. Cette matrice essentielle peut être exprimée en produit d'une matrice antisymétrique $[t]_\times$ et la matrice de rotation R comme $E = [t]_\times R$.

En conséquence de la propriété ESV, E peut être exprimée en SVD :

$$\begin{aligned} E &= U\mathrm{diag}(1,1,0)\, WU^T R \\ &= UDV^T. \end{aligned} \tag{1.15}$$

A partir de cette décomposition, il est possible d'extraire la position et l'orientation relatives [R t] de la deuxième caméra si la première avoir la forme $P = K[I\ 0]$. La matrice antisymétrique $[t]_\times$ peut être directement déterminée à partir de la propriété de ESV, il en résulte que :

$$[t]_\times = U(\pm Z)U^T. \tag{1.16}$$

Ainsi, la translation de la caméra est déterminée jusqu'à un facteur d'échelle inconnu. Notez que jusqu'à présent, le signe et donc la direction du mouvement est également inconnue. Puisque $[t]_\times t = 0$. Il découle de ce qui précède que le vecteur de translation est $t = \pm u_3$ où u_3 est la dernière colonne de U.

La matrice orthogonale V^T encode la matrice rotation R de la deuxième caméra. Par conséquent, la rotation R peut s'écrire :

$$R = UW^TV^T \quad \text{ou} \quad R = UWV^T \tag{1.17}$$

Comme dans le cas de la translation, il existe deux solutions pour la matrice de rotation qui sont cohérentes avec la géométrie épipolaire. Puisque chaque solution pour la translation peut être combinée avec chaque matrice de rotation. En résumé, il y a au total quatre solutions possibles pour la matrice de projection P_2 de la seconde caméra :

$$P_2 = K[UW^TV^T \quad + u_3] \qquad P_2 = K[UWV^T \quad - u_3]$$

$$P_2 = K[UW^TV^T \quad - u_3] \qquad P_2 = K[UWV^T \quad + u_3] \tag{1.18}$$

Il est évident qu'une seule de ces solutions désigne correctement le mouvement physique de la caméra. Pour déterminer quelle solution est correcte, une paire de points correspondants $x \leftrightarrow x'$ est analysée. Pour la solution correcte, le point X correspondant à cette paire de points est situé devant les deux caméras alors que pour toutes les autres solutions, il est situé derrière au plus une des caméras [39]. Pour la reconstruction 3D de X, une méthode linéaire simple pour la triangulation stéréo peut être utilisée, voir par exemple [36] pour plus de détails. La récupération des paramètres extrinsèques en utilisant la matrice essentielle est résumée dans l'algorithme suivant :

Algorithme Récupération des paramètres extrinsèques

Entrée : Les paramètres intrinsèques contenus dans K et la matrice fondamentale F avec une paire associée de points correspondants $x \leftrightarrow x'$.

Sortie : Les paramètres extrinsèques R et t pour la deuxième position de la caméra.

1. Calculer la matrice essentielle $E = K^TFK$.
2. Décomposer E via un SVD en UDV^T.
3. Calculer les vecteurs de translation putatifs et les matrices de rotation et assembler quatre matrices de projection de caméra :
   - Translation: $t = +u_3$ ou $t = -u_3$ (1.19)
   - Rotation: $R = UW^TV^T$ ou $R = UWV^T$ (1.20)
4. Trouver le point X qui correspond aux paires de points $x \leftrightarrow x'$ pour chacune des quatre matrices de projection possibles.
5. Sélectionner les paramètres extrinsèques [R t] pour lesquels X est devant les deux caméras.

1.2.2.1.3 Géométrie à multi vues

Nous avons vu que la matrice essentielle ou fondamentale encapsule la contrainte géométrique relative aux paires de vues. Ensuite, nous porterons notre attention sur la résolution du problème d'un nombre arbitraire de vues. L'étape finale est généralement

l'ajustement du faisceau, qui est utilisé de manière itérative pour affiner les paramètres de structure et de mouvement. (Voir la figure 1.1).

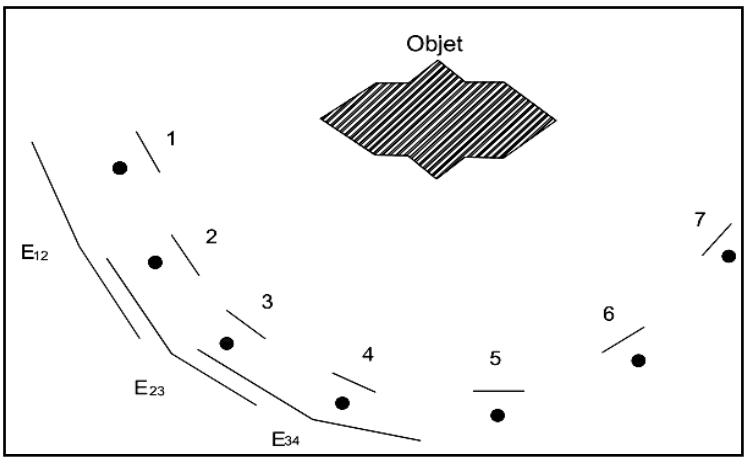


Figure 1.1 − Illustration d'un enregistrement séquentiel. Les vues 1 à 7 sont enregistrées une par une en calculant les matrices essentielles E_{12}, E_{23}, etc.

1.2.2.2 Reconstruction de la géométrie 3D de la scène

Il existe de nombreux indices qui peuvent être utilisés pour extraire la géométrie des photographies : texture, contours et correspondance stéréo. Les deux derniers ont eu beaucoup de succès, la correspondance stéréo étant la plus performante en termes de robustesse et de nombre d'applications. MVS (Multi View Stereo) est le terme général donné à un ensemble de techniques qui utilisent la correspondance stéréo comme source de données principale et utilisent plus de deux images [40], [41].

Un algorithme MVS dépend de la qualité des images d'entrées et des paramètres de la caméra. De plus, une grande partie du récent succès de MVS est due au succès des algorithmes sous-jacents de SFM (Structure From Motion) qui calculent les paramètres de la caméra.

L'approche globale est toujours similaire :

• Calculer les paramètres de la caméra pour chaque image,

• Reconstruire la géométrie 3D de la scène à partir des images mises en jeux et des paramètres de caméra correspondants.

1.2.2.2.1 Vues multiples via Structure à partir du mouvement (SFM)

1.2.2.2.1.1 Cadre commun

Les techniques de structure à partir du mouvement sont utilisées dans une grande variété d'applications, notamment l'étude photogrammétrique [42], la reconstruction automatique de modèles de réalité virtuelle [43] et la détermination du mouvement de la caméra.

La théorie géométrique de la structure du mouvement permet de calculer simultanément les matrices de projection et les points 3D en utilisant uniquement des points correspondants dans chaque vue.

Les algorithmes de SFM prennent en entrée un ensemble d'images et produisent deux choses : les paramètres de caméra de chaque image et un ensemble de points 3D visibles dans les images qui sont souvent encodés en pistes. Une piste est définie comme les coordonnées 3D d'un point 3D reconstruit et la liste des coordonnées 2D correspondantes dans un sous-ensemble des images d'entrée.

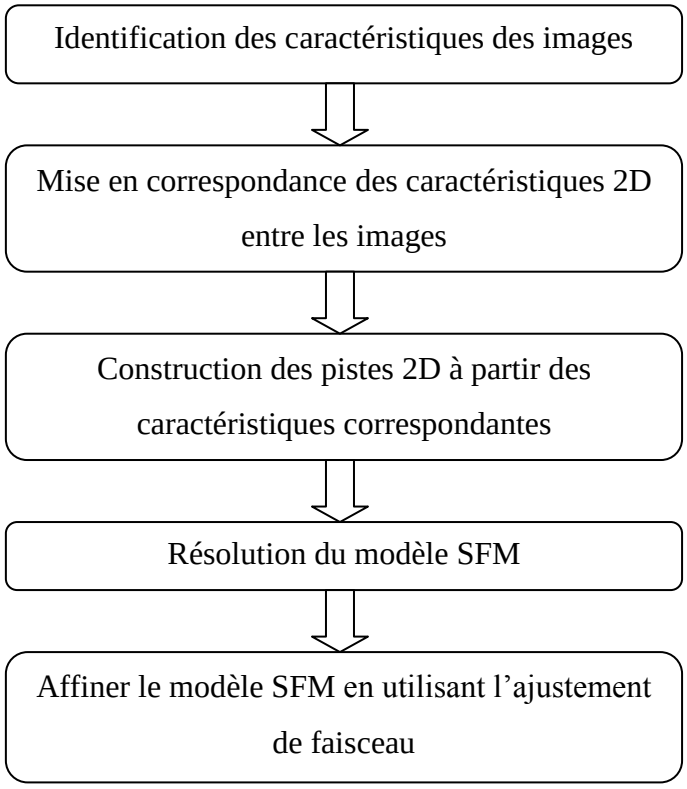


Figure 1.2 − Principales étapes d'un pipeline SFM générique : détection des entités, mise en correspondance des entités, génération des pistes, structure de mouvement et ajustement de faisceau.

La plupart des algorithmes SFM de référence partagent le même pipeline de traitement de base (voir la figure 1.2).

La figure 1.3 illustre des exemples de SFM à grande échelle de [44].



Figure 1.3 − Exemples de SFM: Modèle de la Colisée.

1.2.2.2.1.2 Ajustement de faisceaux

L'ajustement de faisceaux [45] est une étape très courante pour affiner le modèle initial de SFM. Étant donné un ensemble de paramètres de caméra $\{P_i\}$ et un ensemble de pistes $\left\{M^j,\ \{m_i^j\}\right\}$, où M^j désigne la coordonnée 3D d'une piste, et m_i^j la coordonnée d'image 2D de sa projection d'image dans la $i-$ éme caméra, l'ajustement de faisceau minimise l'erreur de moindres carrés non linéaire suivante :

$$E(P,M) = \sum_j \sum_{j \in V(j)} \left| P_i\left(M^j\right) - m_i^j \right|^2 \qquad (1.21)$$

$V(j)$ représente la liste des indices de caméra où le point M^j est visible et $P_i\left(M^j\right)$ représente la coordonnée du point de projection du point 3D M^j sur l'image 2D en relative avec la $i-$ éme camera, tout en utilisant ses paramètres P_i.

1.2.2.2.2 Stéréovision multi-vues (MVS)

Les algorithmes de stéréovision à deux vues se considèrent un domaine de recherche très actif et fructueux [46]. La version multi-vues de stéréo est une amélioration naturelle du cas à deux vues. Au lieu de capturer deux photographies de deux points de vue différents, la stéréo multi-vues capturerait davantage de points de vue entre les deux pour augmenter la robustesse.

Bien que les algorithmes MVS partagent les mêmes principes que les algorithmes stéréo classiques, ils sont conçus pour traiter des images avec des points de vue plus variés, tels qu'un ensemble d'images entourant un objet, et traitant également un très grand nombre d'images.

Notez que dans le cas de MVS, où les paramètres de la caméra sont connus, la résolution de la géométrie 3D de la scène est exactement équivalente à la résolution du problème de correspondance entre les images d'entrée. En effet, considérons un point 3D appartenant à la géométrie de la scène 3D (voir la figure 1.4 à gauche ci-dessous). La projection du point 3D dans l'ensemble des caméras visibles établit une correspondance unique entre les coordonnées projetées sur chaque image.

Étant donné un pixel dans une image, la recherche des pixels correspondants dans d'autres images nécessite deux ingrédients :

• Un moyen efficace de générer des candidats de pixels possibles dans les autres images.

•Une mesure pour déterminer la probabilité qu'un candidat donné soit la bonne correspondance.

Si la géométrie de la caméra n'est pas connue, comme c'est généralement le cas dans le flux optique [47], chaque pixel d'une image peut correspondre à tout autre pixel d'une autre image. C'est-à-dire que pour chaque pixel, il faut effectuer une recherche 2D dans l'autre image. Cependant, lorsque les paramètres de la caméra sont connus (et que la scène est rigide), le problème de correspondance d'image est simplifié d'une recherche 2D à une recherche 1D (voir la figure 1.4 à droite). Un pixel dans une image génère un rayon optique 3D qui traverse le pixel et le centre de la caméra de l'image. Le pixel correspondant sur une autre image ne peut se trouver que sur la projection de ce rayon optique dans la seconde image. Les différentes contraintes géométriques créées lorsque plusieurs caméras visualisent la même scène 3D à partir de différents points de vue sont connues sous le nom de géométrie épipolaire [36].

Nous avons vu dans cette partie, un aperçu des méthodes de reconstruction 3D à partir d'une ou plusieurs images. Deux méthodes de structure à partir du mouvement (SFM) sont qualifiées pour la reconstruction éparse de modèles [42], [43]. En effet, les algorithmes SFM prennent un ensemble d'images comme entrée et générant en sortie les paramètres de caméra de chaque image et un ensemble de points 3D visibles dans les images, souvent codés sous

forme de pistes. En se basant sur SFM, la reconstruction 3D dense est calculée à partir de la MVS [41], [48].

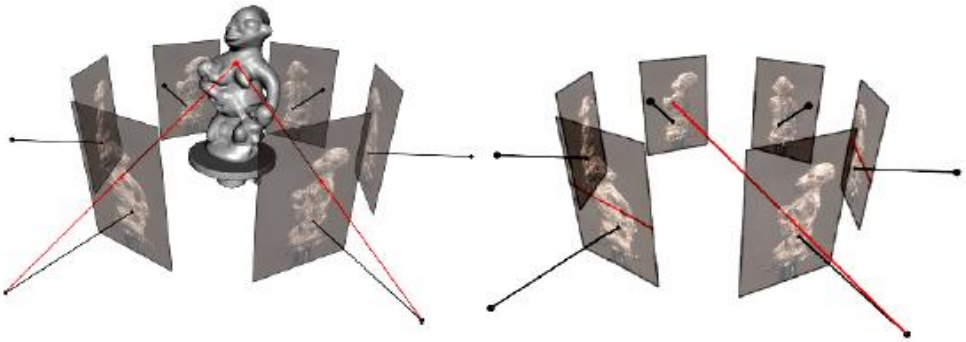


Figure 1.4 − Correspondance des images basée sur une connaissance à priori des paramètres de caméra. A gauche : Recherche 2D entre les pixels de différentes images. A droite :Recherche 1D entre les pixels [49].

L'obtention de la structure finale du modèle exige l'utilisation l'une des méthodes de maillage 3D. Les parties suivantes présentent les algorithmes des maillages triangulaires les plus utilisées dans ce stade.

1.2.3 Maillage 3D

Dans cette section, nous présentons dans un premier temps quelques définitions du maillage et particulièrement celles de maillages triangulaires, les différentes méthodes de reconstruction de surfaces basées sur la triangulation de Delaunay DT(P), le Diagramme de Voronoi Vor(V) et l'approche de Crust**.**

1.2.3.1 Définitions

1.2.3.1.1 Maillage

Shamir [50] définie un maillage M avec m sommets comme un triplet $(V; E; F)$ de la manière suivante :

- V : les sommets $=\{P_i | P_i \in \mathbb{R}^3, 1 \leq i \leq m\}$
- E : les arrêtes $=\{(P_i, P_j) | P_i, P_j \in V\}$
- F : les faces $=\{< P_i, P_j, P_k, \ldots > | P_i, P_j, P_k, \ldots \in V\}$.

Un maillage (figure 1.5) donc est un objet tridimensionnel constitué de sommets. Une arête est une ligne reliant deux sommets. Une face est un cycle d'arêtes.

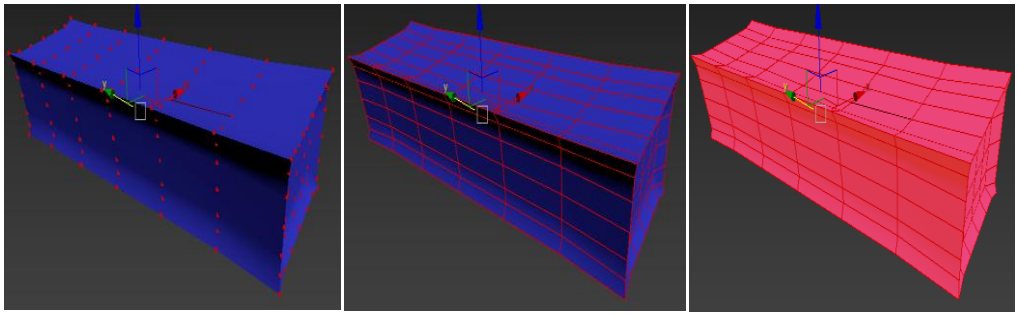


Figure 1.5 − Illustration de maillage quadrangulaire. A gauche :Les points rouges représentent les sommets. Au milieu : Les arêtes sont les traits rouges reliant les sommets. A droite : Le carré constitue une face.

La connectivité d'un maillage est définie à partir de ses arêtes E et de ses faces F. Les maillages triangulaires, dont chaque face connecte trois sommets, sont les plus courants.

1.2.3.1.2 Maillage triangulaire

Définition 1 : Un maillage triangulaire M_T est défini par des données de type topologique et de type géométrique.

Les informations topologiques décrivent la connectivité des faces du maillage et sont données par le graphe $G = \{V, E, F\}$ où :

- $V = \{p_1, \dots, p_v\}$ est l'ensemble des sommets.
- $E = \{e_1, \dots, e_e | \ \forall i \in (1, \dots, e), \exists v, v' \in E \text{ tel que } e_i = (v, v') = (v', v)\}$ L'ensemble des arêtes.
- $F_{i,j,k} = \{(v_i, v_j, v_k) | v_i, v_j, v_k \in V \text{ et } (v_i, v_j), (v_i, v_k), (v_j, v_k) \in E\}$ L'ensemble des faces triangulaires du maillage.

L'information géométrique est donnée par le plongement géométrique de la surface maillée M_T dans $\mathbb{R}^3$. Il est obtenu en assignant une position $p_i = (x_i, y_i, z_i)$ de $\mathbb{R}^3$ à chaque sommet v_i de V.

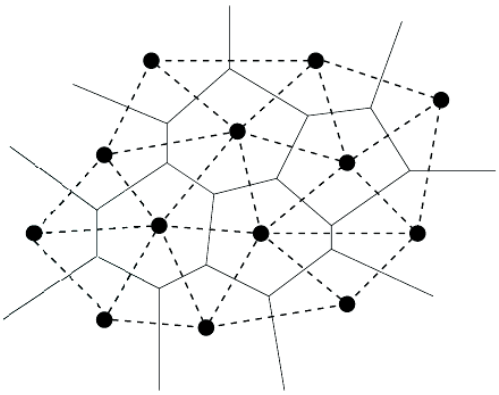


Figure 1.6 − Illustration de maillage (en pointillé) et son graphe dual, [51].

A partir de la description topologique G d'un maillage M_T, on peut repérer les faces adjacentes du maillage et construire le graphe dual de G.

Définition 2 Deux faces de M_T sont dites adjacentes si et seulement si elles ont une arête de E_T en commun.

Définition 3 Le graphe dual $G_d = (V_G, E_G)$ du graphe primal G du maillage M_T est construit de la façon suivante (voir figure 1.6) :

Les nœuds de G_d correspondent aux faces de M_T.

Les arêtes de G_d relient les noeuds de G_d correspondant aux faces adjacentes dans M_T.

1.2.3.2 Reconstruction de surface

La reconstruction de surface consiste à calculer un modèle qui approxime une surface connue seulement par un échantillon de points. L'une des approches la plus utilisée dans ce stade est la triangulation de Delaunay et le Diagramme de Voronoi.

1.2.3.2.1 Triangulation de Delaunay

Un triangle est le terme générique pour un triangle en 2D ou un tétraèdre en 3D. Une triangulation T (S) de S est un ensemble de triangles non chevauchants qui recouvre exactement l'enveloppe convexe Ω (S) avec tous les points de S étant parmi les sommets de la triangulation. La propriété Delaunay s'assure simplement que la sphère circonscrite des tétraèdres ne contient aucun sommet d'autres tétraèdres. En deux dimensions, cette propriété revient à dire qu'aucun cercle entourant un triangle ne contient pas le sommet d'un autre triangle. Une illustration simple de la propriété de Delaunay en deux dimensions est donnée dans la figure 1.7 ci-dessous. Remarquez que le sommet rouge se trouve à l'intérieur du cercle circonscrit d'un triangle voisin et, par conséquent, la propriété Delaunay est violée. Notez qu'il est possible que les points qui forment deux triangles ou tétraèdres adjacents soient co-circulaires ou co-sphériques. Ainsi, le cercle circonscrit / circumsphère contiendra un sommet non défini. Les triangulations de Delaunay sont populaires dans la communauté de maillage car il existe des algorithmes rapides qui permettent de générer la complexité DT (S) en $O\ (nlogn)$.

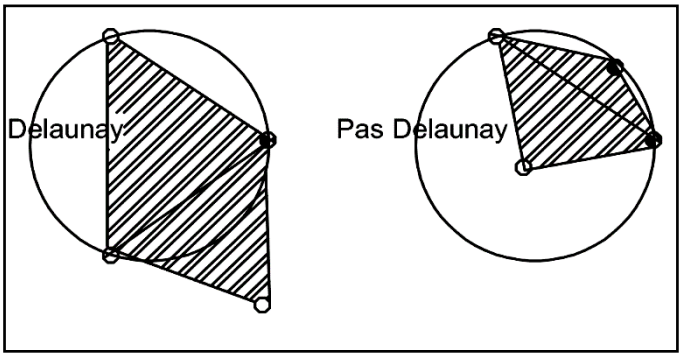


Figure 1.7 − Illustration de la propriété de Delaunay. A gauche : Delaunay. A droite : La propriété de Delaunay est non respectée.

L'algorithme de triangulation le plus couramment utilisé, et celui qui inspire le plus d'algorithmes de tétraèdres tridimensionnels, a été développé indépendamment et publié simultanément dans le même journal par Bowyer [52] et Watson [53]. Leur algorithme est résumé comme suit : Étant donné une triangulation de Delaunay, insérez un nouveau point dans la triangulation en trouvant tous les triangles dont les cercles circonscrits contiennent le nouveau point. Supprimez ces triangles et connectez les points pour former de nouveaux triangles de manière appropriée. Cet algorithme est généralement connu sous le nom d'insertion "Boywer-Watson".

Les algorithmes les plus rapides permettant de construire DT(S) sont basés sur le noyau de Delaunay [54]. Soit DT_k la triangulation de Delaunay d'un ensemble de points $S_k = \{p_1, \dots, p_k\} \subset \mathbb{R}^d$. Delaunay kernel est une procédure qui permet l'insertion incrémentielle d'un point donné $p_{k+1} \in \Omega(S_k)$ dans DT_k et de construire la triangulation de Delaunay DT_{k+1} de $S_{k+1} = \{p_1, \dots, p_k, p_{k+1}\}$. Le noyau de Delaunay peut être écrit de la manière abstraite suivante :

$$DT_{k+1} = DT_k - \mathcal{C}(DT_k, p_{k+1}) + \mathcal{B}(DT_k, p_{k+1}) \qquad (1.22)$$

Où $\mathcal{C}(DT_k, p_{k+1})$ est la cavité de Delaunay. Elle représentel'ensemble de tous les triangles dont les cercles circonscritscontiennent le nouveau point p_{k+1} (voir la figure 1.8 où les triangles de la cavité ne peuvent faire partie de DT_{k+1}) et la boule de Delaunay $\mathcal{B}(DT_k, p_{k+1})$ est un ensemble de triangles qui remplissent le trou polyédrique laissé vide, tout en enlevantla cavité Delaunay $\mathcal{C}(DT_k, p_{k+1})$ de DT_k.

Le bloc de construction le plus important dans toute implémentation du noyau de Delaunay est le calcul de la cavité de Delaunay simplement connectée $\mathcal{C}(DT_k, p_{k+1})$ [54].

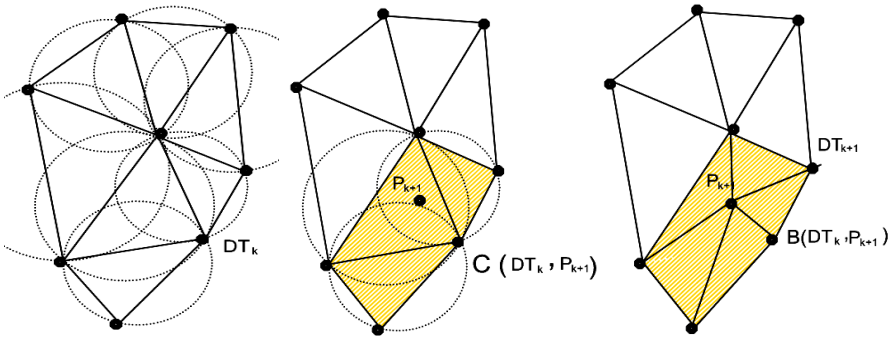


Figure 1.8 − Triangulation de Delaunay $\mathrm{DT_k}$ (gauche), Cavité Delaunay $\mathcal{C}(\mathrm{DT_k}, \mathrm{p_{k+1}})$ (au centre) et $\mathcal{B}(\mathrm{DT_k}, \mathrm{p_{k+1}})$ (droite).

Delaunay a montré aussi que toute triangulation avec la propriété décrite ci-dessus était le maillage double du diagramme de Voronoï [55]. Une illustration de cette double relation entre le diagramme de Voronoï et la triangulation de Delaunay est donnée sur la figure 1.9.

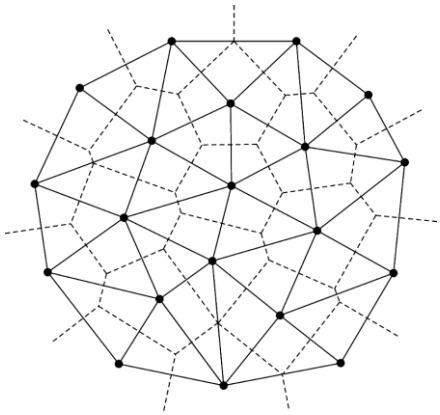


Figure 1.9 − Illustration de relation entre le diagramme de Voronoï (en pointillé) et la triangulation de Delaunay (en noire), [56]

La triangulation / tétraèdre de Delaunay d'un plan / espace est unique pour un ensemble donné de points, à l'exception de la possibilité de points co-circulaires / co-sphériques.

Dans cette optique, Chew [57] a montré qu'avec un ensemble donné de n sommets et un ensemble prescrit d'arêtes, un maillage Delaunay contraint peut être construit avec une complexité $O(nlogn)$.

1.2.3.2.2 Diagramme de Voronoï

Etant donné un ensemble de sites, $\mathrm{V} = \{\mathrm{v_k}\}_{\mathrm{k=1}}^{\mathrm{n}} \subset \mathbb{R}^{\mathrm{d}}$, la cellule de Voronoï du site $\mathrm{v_k} \in \mathrm{V}$ est définie comme suit :

$$\mathrm{Vor}(\mathrm{v_k},\ \mathrm{V}) = \left\{\mathrm{x} \in\ \mathbb{R}^{\mathrm{d}},\ \|\mathrm{v_k} - \mathrm{x}\| \leq \|\mathrm{v_l} - \mathrm{x}\| \text{ pour tout } \mathrm{l} \neq \mathrm{k}\right\}.$$

Les voisins Voronoi du site v_k sont définis comme les sites $v_l \in V$ tel que $\mathrm{Vor}(v_l, V) \cap \mathrm{Vor}(v_k, V) \neq \emptyset$.

Etant donné une paire de sites $(v_k, v_l) \in V^2$, un ensemble de $x \in \mathbb{R}^d$ tel que $\|v_k - x\| = \|v_l - x\|$ est appelé bissectrice de v_k et v_l. La bissectrice délimite deux demi-espaces. Soit $H_{k,l}$ celui qui contient v_k et exclut v_l. Il n'est pas difficile de voir que la cellule Voronoï $\mathrm{Vor}(v_k, V)$ est l'intersection de tous les demi-espaces $H_{k,l}$ définie par des bissectrices entre le site v_k et ses voisins Voronoïv_l. Le diagramme de Voronoï $\mathrm{Vor}(V)$ est la collection de toutes les cellules de Voronoï, $\mathrm{Vor}(V) = \{\mathrm{Vor}(v, V), \forall v \in V\}$, (voir la figure 1.10 ci-dessous).

Le diagramme de Voronoi et sa double triangulation de Delaunay ont été largement étudiés [58] et des algorithmes robustes existent pour leur calcul [59]–[61]. Les algorithmes de référence implémentent généralement une construction incrémentale de la triangulation de Delaunay : les sites sont insérés un par un, vérifiant leurs validité et remaillage de la zone couverte par les sites invalides. Lorsque l'ensemble des sites est connu à l'avance, l'ordre d'insertion est généralement optimisé en utilisant des techniques de tri spatial pour améliorer la localisation des données [62]. Un aspect important de ces méthodes consiste à prouver une performance géométrique dans le cas d'un échantillon de points dense fourni comme données d'entrée, mais généralement la qualité topologique de ces méthodes dans la génération des maillages surfaciques n'est pas toujours garantie. Pour faire face à ces limitations topologiques, plusieurs méthodes basées sur le diagramme de Voronoi et la triangulation de Delaunay ont été proposées [54], [63], [64][65].

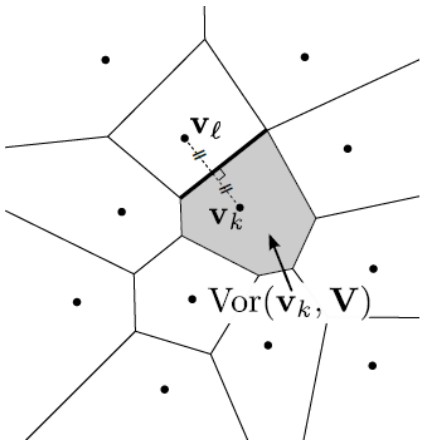


Figure 1.10 − Diagramme de Voronoï pour un ensemble de sites $V = \{v_k\}_{k=1}^n$.

1.2.3.2.3 Algorithmes de Crust

L'algorithme de Crust est basé sur le diagramme de Voronoï tridimensionnel et la triangulation de Delaunay. Il produit un ensemble de triangles que nous appelons crust des points échantillonnés. Tous les sommets des triangles de Crust sont des points échantillonnés. En fait, tous les triangles de Crust apparaissent dans la triangulation de Delaunay des points échantillonnés.

Un bon échantillon est celui dans lequel la densité d'échantillonnage est inversement proportionnelle à la distance par rapport à l'axe médian. En effet, un échantillon S est un r-échantillon d'une surface F lorsque la distance euclidienne entre n'importe quel point $p \in F$ et le point d'échantillonnage le plus proche est au plus r fois la distance de p au point le plus proche de l'axe médial de F. La constante de proportionnalité r est généralement inférieure à un.

1.2.3.2.3.1 Définition de Crust

Une arête e appartient à Crust s'elle représente un cercle circonscrit vide non seulement de tous les autres points d'échantillonnage mais aussi de tous les sommets de Voronoï de S. Crust obéit au théorème suivant :

Théorème. Crust d'un r-échantillon à partir d'une courbe lisse F, pour $r \leq 0.25$, relie uniquement les points d'échantillonnage adjacents en F.

Lorsque S est un r-échantillon, pour r suffisamment petit, les sommets de Voronoï se rapprochent de l'axe médian, et tout cercle circonscrit d'une arête entre des échantillons non adjacents contient un autre échantillon ou un sommet de Voronoï. D'autre part, une arête entre deux échantillons adjacents est circonscrite par un petit cercle loin de l'axe médian et par conséquence de tous les sommets de Voronoï.

1.2.3.2.3.2 Algorithme

La définition de Crust bidimensionnelle conduit à l'algorithme suivant :

Soit S un échantillon de points, et V l'ensemble des sommets de Voronoï.

1- Calculer le diagramme de Voronoï de S, et soit V l'ensemble des sommets de Voronoï

2- Calculer ensuite la triangulation de Delaunay de $S \cup V$.

Crust consiste aux arêtes de Delaunay entre les points de S, puisque ceux-ci sont des arêtes avec des cercles circonscrits vides des points $S \cup V$ (Voir la figure 1.11).

Notez que Crust est également un sous-ensemble de la triangulation de Delaunay. Sachant que les sommets de Voronoi filtrent les arêtes indésirables de la triangulation de Delaunay. Nous appelons cette technique le filtrage de Voronoi.

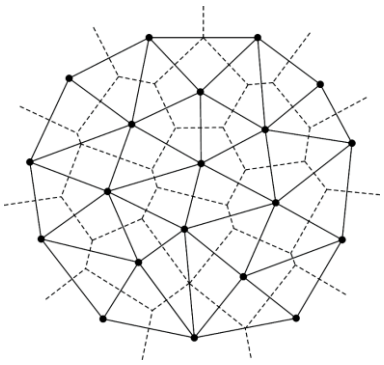


Figure 1.11 − L'algorithme bidimensionnel. Les arêtes de pointillés représentent le diagramme de Voronoï d'un ensemble de points S échantillonné à partir d'une courbe. Les arêtes en noir représentent la triangulation de Delaunay de S ∪ V, [56].

1.2.4 Segmentation de maillage

Dans cette section, nous présentons le problème de la segmentation d'un maillage et le fondement mathématique sur lequel sont basées les approches de segmentation des modèles 3D citées dans l'état de l'art.

1.2.4.1 Formulation du problème de segmentation d'un maillage

1.2.4.1.1 Définition

Considérons un maillage triangulaire 2-variété M_T et $G = \{V, E, F\}$ son graphe associé. Le problème peut être énoncé brièvement en tant que tâche visant à trouver une partition S d'un maillage M_T. Mathématiquement, une partition S d'un ensemble M_T est une collection disjointe de N sous ensembles $S = \{S_1, \dots, S_N\}$ non vide et distinct de nœuds [66]. Les N partitions sont induites par le partitionnement $P_N = \{V_1^{Gd}, \dots, V_N^{Gd}\}$ de son graphe dual G_d.

On peut donc définir une segmentation d'un maillage M_T en N partitions comme suit :

$$\begin{cases} \forall V_i^G \in P_N, \quad V_i^G \neq \emptyset \text{ et } V_i^G \subset V^G \\ \bigcup_{i=1..N} V_i^G = V^G \\ \forall (V_i^G, V_j^G) \in P_N^2, i \neq j \Rightarrow V_i^G \cap V_j^G = \emptyset \end{cases} \quad (1.23)$$

1.2.4.1.2 Le problème de métrique de similarité

La segmentation de maillage vise à décomposer un maillage surfacique aux partitions perceptuelles et cohérentes. Dans cette perspective, une grande diversité des méthodes

computationnelles a été proposée jusqu'à présent. L'un des défis les plus importants que ces méthodes partagent en commun consiste à définir une métrique quantifiant la dissemblance entre les points de la surface.

En fait, une métrique de similarité est une mesure quantifiant le niveau de la similarité visuelle entre les éléments contigus, et par conséquence, leurs probabilités d'appartenance au même segment. Il fournit donc un outil numérique pour comparer les éléments d'une forme géométrique. Plusieurs méthodes de segmentation [67]–[70] utilisent certaines mesures inhérentes qu'on peut calculées directement à partir du maillage, par exemple, les angles dièdres entre les faces, courbures. Ces mesures souvent offre une raisonnable fidélité dans la segmentation puisque la perception visuelle humaine dépend largement des courbures de la surface [71], [72]. Toutefois, elles sont incapables de capturer les caractéristiques géométriques dans un cadre global et volumétrique, par exemple cous étroits (i.e., régions où la surface transversale réduit rapidement) ou lorsque l'objet à segmenter ayant une topologie complexe, un bouquet de fleurs par exemple.

1.2.4.2 Fondement mathématique : méthodes géométriques

Dans cette section nous allons présenter le fondement mathématique sur lequel se base les méthodes géométriques citées dans la section 1.3.1 (état de l'art).

1.2.4.2.1 Ajustement des primitives

Attene et al. [68] proposent un algorithme de classification hiérarchique basé sur l'ajustement des primitives.

Dans cette section, nous allons décrire comment calculer les paramètres des primitives : plan, cylindre et sphère.

- Plans ajustés

Pour calculer le meilleur plan primitif dans un ensemble des triangles on utilise une méthode classique basée sur ACP [69], [70]. Plus précisément, on calcule la matrice de covariance cov_v des sommets de la classe :

$$\mathrm{Cov_v} = \sum_i \mathrm{a(v_i)}\,(\mathrm{v_i} - \bar{\mathrm{v}})(\mathrm{v_i} - \bar{\mathrm{v}})^{\mathrm{T}} \qquad (1.24)$$

Où $\bar{\mathrm{v}} = \frac{\sum_i \mathrm{a(v_i)v_i}}{\sum_i \mathrm{a(v_i)}}$

Sachant que v_i est un sommet et $a(v_i)$ est la zone Voronoi restreinte de v_i.

Le plan le mieux ajusté passe par $\bar{v}$, et son normale n est le vecteur propre correspondant au minimum vecteur propre de cov_v. L'erreur d'ajustement L^2 peut être calculée en utilisant :

$$L^2 = \sum_{i=1}^{n} a(v_i) \left(n(v_i - \bar{v})\right)^2 \tag{1.25}$$

- Sphères ajustées

La détermination des paramètres (centre et rayon) des moindres carrés de la sphère la mieux ajustée pour un ensemble des points 3D peut être écrit comme suit :

Soit P un ensemble de n points$(x_i,\ y_i, z_i)$, et soit $(x - c_x)^2 + \left(y - c_y\right)^2 + (z - c_z)^2 - r^2 = 0$ l'équation de la sphère S du rayon r et de centre $c = \left(c_x, c_y, c_z\right)$. La distance euclidienne quadratique entre un point $p_i = (x_i, y_i, z_i)$ et la sphère S est :

$$d^2(p_i, S) = \left(\sqrt{(x_i - c_x)^2 + \left(y_i - c_y\right)^2 + (z_i - c_z)^2} - r\right)^2 \tag{1.26}$$

Au sens des moindres carrés, calculant la sphère la mieux ajustée d'un ensemble de points P équivaut à déterminer le centre $\left(c_x, c_y, c_z\right)$ et le rayon r tels que la somme de toutes les distances quadratiques est minimisée par :

$$\min\left(\sum_{i=1}^{n} d^2\,(p_i, S)\right) \tag{1.27}$$

Les paramètres c et r de la sphère étant établis, Le L^2 erreur d'ajustement peut être calculée en utilisant :

$$L^2 = \sum_{i=1}^{n} a(v_i)(\|v_i - c\|_2 - r)^2 \tag{1.28}$$

- Cylindres ajustés

On trouve utile de représenter chaque cylindre par un rayon, un vecteur unitaire parallèle à son axe, et un point appartenant à son axe qu'on appelle le centre de cylindre.

En établissant la direction n de cylindre ajusté, Les autres paramètres à déterminer sont le rayon et le centre. Il faut noter que cette opération peut être réduite à un problème d'un

cercle d'ajustement de 2 dimensions en projetant les sommets sur le plan de normale n et passant par le centre de gravité de la classe.

Les paramètres du cercle 2D ajusté peut être calculé que celle antérieurement décrite pour les sphères ajustées en réduisant simplement la dimensionnalité du problème.

Les paramètres n, c et r du cylindre étant établis. L'erreur d'ajustement L^2 peut être calculée en utilisant la formule :

$$L^2 = \sum_{i=1}^{n} a(v_i)(\|(v_i - c) \times n\|_2 - r)^2 \quad (1.29)$$

1.2.4.2.2 Marches aléatoires

Marches aléatoires (**Random walks**) Lai et al. [73] décrivent une procédure de segmentation qui se déroule en deux phases.

1. Sélection automatique des germes :

Cette méthode répartie essentiellement les germes sur la surface du modèle en optimisant un certain retour d'énergie [74]. Après répartition, on sélectionne les faces contenant des germes.

2. Fusionnement des segments :

Pour déterminer le coût relatif de fusion entre deux segments adjacents S_i et S_j, on désigne d'abord par $\partial S_i \cap \partial S_j$ la frontière commune entreles deux segments, et par $\partial S_i \cup \partial S_j$ la combinaison du deux frontières. On intègre la mesure de différence d le long de la frontière commune, et on la désigne par $D_{\partial S_i \cap \partial S_j} = \sum_{e \in \partial S_i \cap \partial S_j} |e|\, d_e$, avec $|e|$ et d_e sont la longueur de l'arête e et la mesure de différence d entre les deux faces adjacentes de e. On définit aussi la longueur totale de la frontière commune comme suit :

$$L_{\partial S_i \cap \partial S_j} = \sum\nolimits_{e \in S_i \cap S_j} |e| \quad (1.30)$$

$D_{\partial S_i \cup \partial S_j}$ et $L_{\partial S_i \cup \partial S_j}$ peut être définie d'une manière similaire. Nous définissons donc le coût relatif de fusion $C_{i,j}$ comme suit :

$$C_{i,j} = \frac{D_{\partial S_i \cap \partial S_j} / L_{\partial S_i \cap \partial S_j}}{D_{\partial S_i \cup \partial S_j} / L_{\partial S_i \cup \partial S_j}} \quad (1.31)$$

Pour chaque paire de segments adjacents S_i et S_j, on calcule le coût de fusionnement $C_{i,j}$ et on remet les paires par ordre de priorité. Le processus de fusionnement procède en choisissant la paire ayant un coût de fusion minimal, en les fusionnant dans un seul segment et mettant à jour l'ordre de priorité en conséquence. Ce processus peut être terminé soit lorsqu'il existe une augmentation importante de $C_{i,j}$ concernant la paire actuelle, soit lorsque on atteint un nombre de régions prédéterminé par l'utilisateur. La figure 1.12 illustre une segmentation automatique avant et après la fusion. Le nombre initial de germes est 60 et le nombre de segments après la fusion est 30.

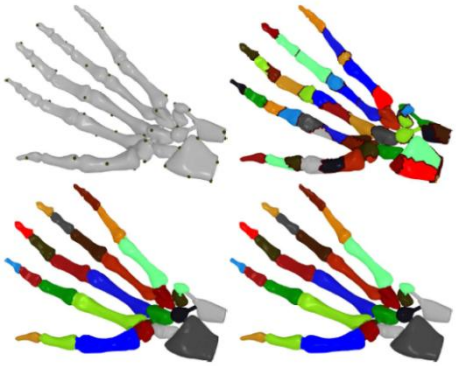


Figure 1.12 − Illustration des étapes principales de la méthode Marches aléatoires (figure extraite de [73]). En haut : Modèle d'entrée avec placement des germes et sur-segmentation. En bas de gauche à droite : Résultat après la fusion et résultat final après lissage des bords.

1.2.4.3 Fondement mathématique : méthodes sémantiques

Dans cette section nous allons présenter le fondement mathématique sur lequel se base les méthodes sémantiques citées dans la section 1.3.2 (état de l'art).

1.2.4.3.1 Fonction diamètre forme

Fonction diamètre forme SDF : Shapira et al. [4] décrivent un algorithme basé sur une fonction nommée "fonction diamètre forme" (SDF).

L'algorithme SDF se déroule en deux étapes :

La première consiste à utiliser GMM k gaussienne d'ajustement à travers un histogramme de l'ensemble des valeurs SDF des faces. Ceci est obtenu en utilisant EM (Expectation Maximization). Pour plus de détail concernant GMM, le lecteur est prié de se reporter à [75], [76]. Le résultat de ce processus de classification de chaque face est un vecteur de longueur k, ce qui signifie sa probabilité d'être affecter à l'une des classes SDF. Noter que chaque classe SDF peut contenir plusieurs parties de maillage tel que les pieds et

les doigts. Les résultats des différentes valeurs de k sont utilisés dans les différentes segmentations hiérarchiques (Voir la figure 1.13).

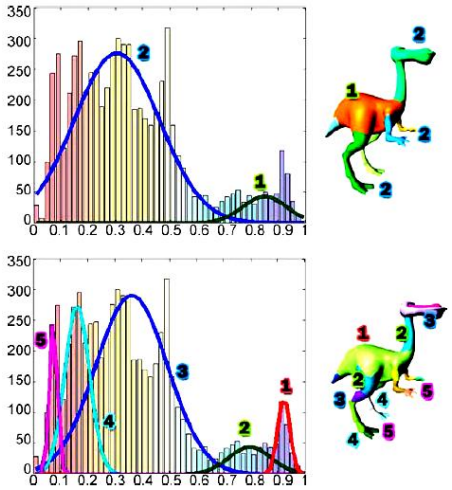


Figure 1.13 − L'ajustement de l'histogramme du SDF permet d'obtenir un partitionnement plus fin de l'objet en parties [4].

La deuxième étape consiste à adoucir les frontières entre les parties. En effet, dans cette étape on utilise un algorithme alpha-expansion avec de graphe-partition [77] pour parvenir à une partition définitive. Le k-way graphe-partition attribue à chaque face une seule partition, compte tenu à la fois du vecteur de probabilité issue de la première étape, et la qualité des frontières. Le graphe-partition sert à optimiser la fonction d'énergie suivante, tout en se basant sur $\mathbf{e_1}$ terme de donnée et $\mathbf{e_2}$ terme de lissage :

$$E(\bar{x}) = \sum_{f \in F} e_1\,(f, x_f) + \lambda \sum_{\{f,g\} \epsilon N} e_2\,(x_f, x_g) \tag{1.32}$$

Où $\quad e_1(f, x_f) = -\log(P(f|x_f) + \varepsilon)$ et $e_2(x_f, x_g) = \begin{cases} -\log(\theta(f,g)/\pi) & x_f \neq x_g \\ 0 & x_f = x_g \end{cases}$

Avec $\mathbf{x_f}$ la classe attribuée à une face $\mathbf{f}$. $\mathbf{P(f|x_p)}$ représente la probabilité d'attribuer une face $\mathbf{f}$ à une classe $\mathbf{p}$. Ces valeurs sont issues à partir de GMM (première étape). $\mathbf{N}$ est un ensemble des couples des faces adjacentes, $\boldsymbol{\lambda}$ est un paramètre défini le degré de lissage, et $\boldsymbol{\theta(f, g)}$ est l'angle dièdre entre les facettes $\mathbf{f}$ et $\mathbf{g}$. Le résultat de l'algorithme graphe partition est une partition lisse du maillage.

La figure 1.14, décrite par Shapira et al. dans [4], illustre les deux étapes.

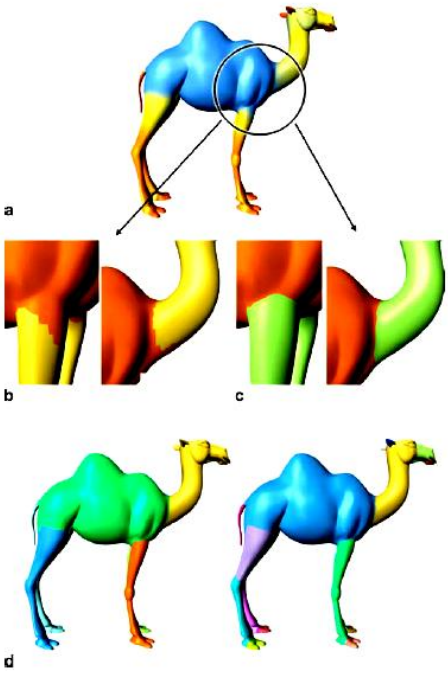


Figure 1.14 − L'algorithme de partitionnement en deux étapes : a. partitionnement souple du maillage. b. seuillage simple. c. limites lisses. d. Le partitionnement hiérarchique est robuste, pose insensible et rapide, [4].

1.2.4.3.2 Extraction des parties principales

Katz et al. [1] proposent un algorithme basé sur la théorie de positionnement multidimensionnel (MDS) [78], [79] calculant une décomposition hiérarchique d'un maillage donné. Pour chaque nœud de l'arbre hiérarchique, l'algorithme se compose des étapes suivantes (voir la figure 1.15) :

1- Déraffinement de maillage : Le déraffinement de maillage est appliqué en tant qu'étape de prétraitement. Il aide non seulement à l'accélération de l'algorithme lorsqu'il est exécuté sur des maillages de grandes tailles, mais aussi à diminuer la sensibilité de l'algorithme en présence de bruit.

2- Représentation insensible à la pose : le positionnement multidimensionnel MDS est utilisée pour transformer le maillage S à un maillage canonique S_{GMD}, comme illustré à la figure 1.15a. Les distances euclidiennes entre les points situés sur le S_{MDS} sont similaires aux distances géodésiques entre leurs points correspondants situés sur S. Cette propriété rend la représentation invariante à la pose, car les organes repliés (i.e. bras) sont redressés par la transformation.

3- Détection des points caractéristiques : les points les plus caractéristiques sont calculés sur S_{MDS}, et mis en correspondance avec ceux appartenant à S, comme il est indiqué dans la figure 1.15b.

4. Extraction des composants essentiels : le composant essentiel est extrait en se basant sur une nouvelle opération appliquée sur le maillage "miroir sphérique" (Figure 1.15c).

5- Raffinement : raffinement des bords en suivant les limites naturelles du maillage. (Figure 1.15d).

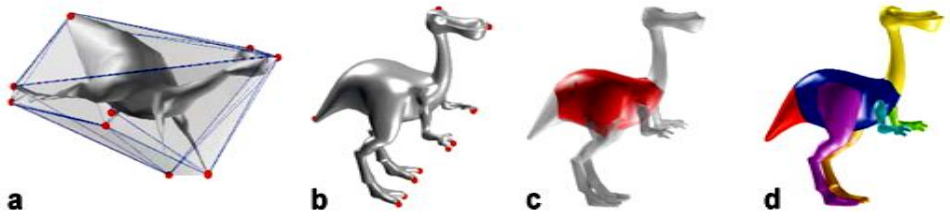


Figure 1.15 − Illustration des principales étapes de l'algorithme : **a** Une transformation MDS et son enveloppe convexe **b** Points caractéristiques **c** Partie essentielle **d** Segmentation et affinement des coupes, [1].

Définitions et équations

- ✓ Le positionnement multidimensionnel (« Multidimensional scaling » ou MDS) est utilisé pour transformer les sommets du maillage à une représentation invariante à la pose. MDS découvre la structure géométrique d'un ensemble de donnés à partir de la (dis)similarité d'information existant entre eux, en représentant la dissimilarité comme une distance dans un espace euclidien de dimensions m. Pour plus de détails concernant MDS, voir [78]–[81].
- ✓ Point caractéristique : un sommet de maillage est un point caractéristique s'il satisfait l'équation (1.33), en plus il doit résider dans l'enveloppe convexe de S_{MDS}.

  $\forall v \in S$, soit N_v un ensemble des sommets aux voisinages du sommet v. Soit $\mathrm{GeodDist}(v_i, v_j)$ la distance géodésique entre les sommets v_i et v_j du maillage S. La condition locale qu'un point caractéristique doit satisfaire est :

$$\sum_{v_i \in S} \mathrm{GeodDist}(v,\ v_i) > \sum_{v_i \in S} \mathrm{GeodDist}(v_n,\ v_i) \qquad (1.33)$$

- ✓ Fonctions Stress : les fonctions stress sont utilisées pour mesurer le degré de correspondance des distances entre les sommets. Soit δ_{ij} la distance géodésique entre v_i et v_j de S et d_{ij} la distance euclidienne entre leurs points correspondants dans l'espace MDS. La fonction stress utilisée est la suivante :

$$F_s = \frac{\sum_{i<j}\left(f(\delta_{ij}) - d_{ij}\right)^2}{\sum_{i<j} {d_{ij}}^2} \qquad (1.34)$$

1.2.4.3.3 Randomized cuts

Golovinskiy et al. [82] proposent une méthode de décomposition hiérarchique permettant la décimation du maillage à un nombre fini des triangles qui composent le segment initial, puis procède à une subdivision binaire des segments afin de guider la segmentation.

Le processus se déroule en quatre étapes :

1. Construction graphique : la première étape produise un graphe G d'arrêt-pondéré représentant le maillage d'entré. De manière similaire de [83], on attribue un poids de concavité à chaque arête du maillage comme suit : soit θ est l'angle dièdre extérieur en face de l'arrêt, on définit le poids de concavité par $w(\theta) = \min((\theta/\pi)^{\alpha}, 1)$ et 1 pour les arrêts convexes. On crée les coûts de coupe par multiplication des longueurs des arêtes du maillage par $w(\theta)$, et les coûts de traversé en divisant la distance entre les centres de gravité des faces par $w(\theta)$. Le coût de coupe d'un segment de maillage est conçu pour favoriser les frontières concaves. Par contre, le coût de traversé représente les distances géodésiques sur un maillage, conçues pour faire trajet à travers les arêtes concaves les plus longues.

2. Coupes aléatoires : Cette étape consiste à générer un ensemble aléatoire de segmentations k-way. Un ensemble de segmentation issues de cette étape est notéS_i donne un ensemble d'arêtes C à travers la segmentation.

3. Fonction de partition : Cette étape consiste à estimer la fonction de partition pour chaque arête. Etant donnée un ensemble de segmentations k-way $\hat{S}$, échantillonnées de manière aléatoire à partir de S. Cette opération peut être effectuée trivialement pour chaque arête e en calculant la somme des poids de toutes les segmentations $\hat{S}$ qui ont une frontière avec e et divisant la somme par le total des poids des segments $\hat{S}$.

4. Coupes cohérentes : Cette dernière étape donne une note d'uniformité pour chaque segmentations S_i de $\hat{S}$. La note d'uniformité est calculée en considérant la moyenne des poids des valeurs de la fonction de partition le long des arêtes sur les frontières des segments S_i. Les M segmentations avec les plus élevés notes constituent un ensemble de coupes les plus cohérentes.

Après avoir présenté les différentes étapes qui composent le processus de la reconstruction 3D et de la segmentation du maillage, il est temps de voir les différentes méthodes proposées dans la littérature. Cette étude nous permettra de comprendre les différentes approches et méthodes existantes et d'identifier les points forts et faibles de chaque méthode. Cela nous conduira par la suite à élaborer nos propres méthodes qui

permettent d'améliorer la qualité de la reconstruction du maillage et de segmenter significativement celui-ci.

1.3 Etat de l'art de la segmentation de maillage 3D

Segmenter un maillage consiste à découper sa surface en un ensemble de régions, appelées segments. Chaque segment est conçu en vue de satisfaire des propriétés géométriques ou sémantiques en reposant sur la perception humaine. Plusieurs travaux comparatifs [1], [4], [82], [84]–[87]; [50], [84], [88], [89], [90] concernant les techniques de segmentation des maillages 3D ont été proposées. La plupart de ces méthodes reposent généralement sur les attributs de surface des objets, y compris les directions normales, la courbure et les distances géodésiques moyennes. Ces méthodes peuvent être divisées en deux grandes familles, les méthodes reposant uniquement sur les critères géométriques et celles basées sur les critères sémantiques. La figure 1.16 montre un exemple de segmentation sémantique et un exemple de segmentation géométrique concernant le même modèle.



Figure 1.16 − A gauche, une segmentation sémantique. A droite, une segmentation géométrique [50].

Dans ce chapitre, nous allons étudier plusieurs méthodes qui visent à segmenter le maillage surfacique en parties significatives.

1.3.1 Méthodes géométriques

Concernant cette première famille, elle inclut les méthodes reposant sur des critères géométriques (courbure [67], approximation par des primitives [68], etc.) pour segmenter le maillage. Dans ce contexte plusieurs méthodes ont été développées. Dans cette section nous allons présenter certaines méthodes parmi les plus utilisées dans ce stade.

1.3.1.1 Ajustement des primitives

Attene et al. [68] proposent un algorithme de classification hiérarchique basée sur l'ajustement des primitives. La meilleure primitive géométrique ajustée (plan, cylindre ou sphère) est calculée afin de fusionner des segments adjacents (voir figure 1.18 ajustement primitive). L'algorithme procède donc de manière ascendante en fusionnant les segments et se termine lorsqu'un nombre de segments déterminé par l'utilisateur a été atteint.

En effet, l'algorithme proposé par **Attene et al.** est une variante de méthode de regroupement hiérarchiques des faces (HFC) décrite en [69]. L'idée de base de l'approche HFC consiste à fusionner les triangles voisins en regroupement représentatives. Ici, un regroupement est un ensemble des triangles connectés qui peuvent être approximés par une simple primitive. En se référant à [69], par exemple, les regroupement sont approximés par des plans ajustés calculés au moyen de l'analyse en composante principale (ACP), et le coût de fusionnement d'un ensemble des triangles à une simple classe représentative est une distance intégrale L^2 de ses sommets à partir d'un plan ajusté. La méthode produit une arborescence qui peut être représenté par un arbre binaire des regroupements.

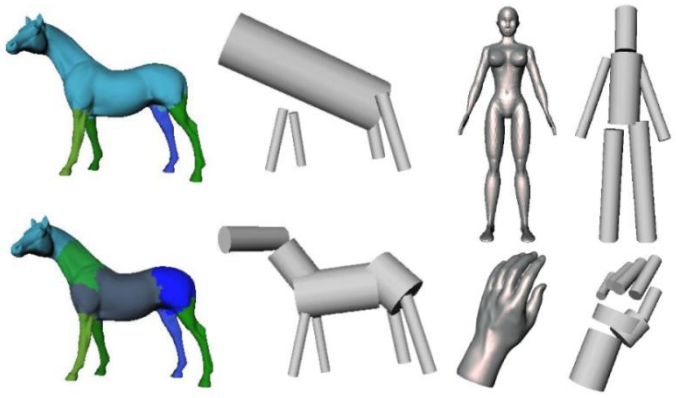


Figure 1.17 − Quelques modèles avec les abstractions correspondantes représentées par les cylindres ajustés. Pour le modèle du cheval, deux niveaux de résolution sont affichés avec le regroupement de visage correspondant, [91].

1.3.1.2 K-Means

Shlafman et al. [92] décrivent un algorithme basé sur la classification k-moyenne des faces. Etant donné un nombre de segments k par un utilisateur, l'algorithme choisit tout d'abord un ensemble de k faces germes pour représenter des regroupements en sélectionnant continuellement de nouvelles faces déjà sélectionnées antérieurement. Par conséquence, l'algorithme alterne entre : 1) L'affectation de toutes les faces à la classe représentant le germe le plus proche, et 2) L'adaptation du germe de chaque classe afin de se situer au centre de la

face associée. Cette itération continue jusqu'à ce que l'affectation des faces aux classes atteigne une convergence.

1.3.1.3 Normalized cuts

Golovinskiy et al. [82] décrivent un algorithme de regroupement hiérarchique dans lequel chaque face du maillage débute par son propre segment, et les segments sont fusionnées hiérarchiquement dans un ordre déterminé selon une fonction de coût d'une zone normalisée réduite : la somme de chaque périmètre d'un segment donné (pondéré par concavité) divisée par sa zone. La classification est terminée lorsqu'un nombre de segments déterminé par l'utilisateur est atteint. Cet algorithme favorise les segments d'avoir des petites frontières de segmentations le long de jointure concave tout en maintenant les segments ayant des zones à peu près similaires.

1.3.1.4 Marches aléatoires

D'autres approches essayent de segmenter le maillage en se basant sur le processus stochastique. Marches aléatoires (**Random walks**) Lai et al. [73] décrivent une procédure de segmentation se déroule en deux phases. Durant la première phase, un ensemble de k faces germes est défini pour le modèle 3D et une sur segmentation est calculée en affectant chaque face f du modèle au germe qui a la probabilité la plus élevée d'atteindre f par une marche aléatoire sur le graphe dual du maillage. Il convient de noter que k est un nombre supérieur au nombre de segments requis. Durant la deuxième phase, les segments sont fusionnés de manière hiérarchique fondée sur les longueurs relatives des intersections et de périmètres totaux des segments adjacents. L'algorithme est terminé lorsqu'un nombre de segments déterminé par l'utilisateur a été atteint. (Voir la figure 1.19).

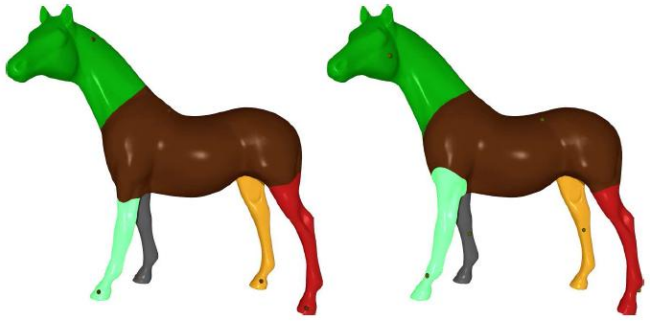


Figure 1.18 − Segmentation du cheval avec des positions de germes différents : les boules représentent les positions des germes (figure extraite de [73]).

En utilisant une telle approche, on s'attend à ce que beaucoup de segments aient plusieurs germes, ce qui naturellement conduit à une sur-segmentation. Par contre, les résultats d'une telle segmentation ne sont pas en générale sensibles à la place précise des germes. On peut simplement fusionner les segments résultants et générer les régions finales convenables. On effectue la fusion comme un processus itératif. Pour déterminer le coût relatif de fusion entre deux segments adjacents.

Ces méthodes donnent généralement des résultats satisfaisants en terme de découpage, cependant, elles souffrent de certains problèmes : premièrement, les algorithmes basés sur la hiérarchie [1], [93] pourraient conduire à des résultats erronés. Deuxièmement, la plupart des algorithmes de classification [82], [94], [95] basés sur la courbure, la distance géodésique et les angles dièdres sont sensibles aux caractéristiques locales de la surface et de changement de pose de modèle. Troisièmement, ces méthodes reposent seulement sur les caractéristiques géométriques de la surface, ce qui est parfois insuffisant pour la segmenter en parties significatives, surtout dans le cas des modèles avec des caractéristiques géométriques complexes.

1.3.2 Méthodes sémantiques

Concernant cette deuxième famille, elle inclut les méthodes reposant sur des critères sémantiques en segmentant un maillage en partie significatives [1], [4], [84]**,** toute en se basant sur les caractéristiques géométriques de l'objet. En effet, le but principal de ces méthodes est de segmenter un objet en entités correspondant à des parties de l'objet (les membres d'un être humain tels que les bras, les jambes...). Ce type de segmentation repose sur la perception humaine.

1.3.2.1 Fonction diamètre forme

Fonction diamètre forme SDF : Shapira et al. [4] décrivent un algorithme basé sur une fonction nommée "fonction diamètre forme" (SDF),qui consiste à une mesure du diamètre du volume de l'objet au voisinage d'un point de la surface. le SDF est calculé à partir du centre de gravité de chaque face, puis la segmentation se déroule en deux étapes. Premièrement, le Modèle de Mélange de Gaussiennes (GMM) est utilisé pour ajuster k gaussiennes à travers un histogramme de l'ensemble des valeurs SDF en vue de générer un vecteur de longueur k pour chaque face en indiquant sa probabilité d'être assigné à chaque classe de SDF. Deuxièmement, l'algorithme alpha-expansion combiné avec de graphe-partition [77] est utilisé afin d'affiner la segmentation tout en minimisant une fonction énergie qui combine les

vecteurs probabilistes calculées dans la première étape avec les bordures lisses et concaves. L'algorithme se déroule hiérarchiquement jusqu' à un nombre de "partitionnement candidats" qui fixe le nombre de segments en sortie. En fait, suivant les auteurs [4], il est recommandé de fixer le nombre de partitionnement candidats à 5, et laisser l'algorithme détermine le nombre de segments automatiquement, La figure 1.19 illustre cette approche.

La SDF est invariante aux transformations rigides de la totalité du maillage. En outre, la SDF est insensible aux déformations qui ne changent pas localement la forme volumétrique. Cela comprend le squelette de base et la fonction affine par morceaux des transformations rigides. En substance, la SDF reste largement insensible à la pose (voir la figure 1.19).

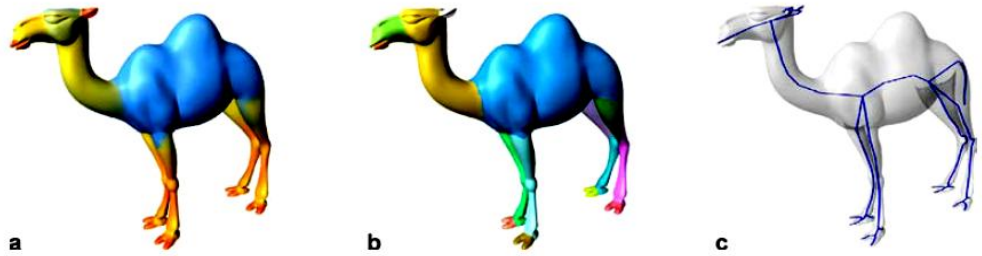


Figure 1.19 − La fonction de diamètre de forme locale (SDF) (a). Partitionnement hiérarchique (b) Squelettes robustes simples (c), [4].

1.3.2.2 Extraction core

Page et al. [96] développent un algorithme basé sur le principe d'une colline grimpant de bassins versants pour segmenter un maillage triangulaire à des parties significatives. Dans leurs implémentations, ils calculent les principaux rayons de courbure et leurs principales directions pour chaque sommet du maillage. Les principales valeurs des rayons de courbure sont utilisées comme métrique pour fixer le seuillage concernant la classification de bassins versants. L'indépendance de pose est aussi une propriété importante dans ce stade (un maillage d'un même objet dans différentes positions est segmenté de manière similaire), en effet **Katz et al.** [1] proposent un algorithme invariant à la pose, tout en se basant sur la théorie de positionnement multidimensionnelle (MDS) [78], [79] calculant une décomposition hiérarchique d'un maillage donné.

Dans leur approche, la probabilité qu'une face appartienne à un certain segment dépend de la distance la séparant des autres faces. Ils se sont aussi appuyés sur l'angle dièdre entre deux faces adjacentes et en outre ils ont utilisé la distance géodésique pour déterminer la probabilité d'appartenance de deux faces au même segment (voir la figure 1.20).

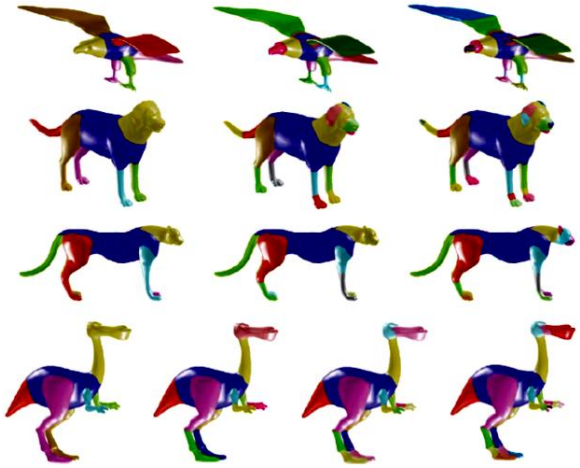


Figure 1.20 − Segmentation hiérarchique d'un aigle, d'un chien, d'un guépard et du dino-pet. Figure extraite de [1].

1.3.2.3 Randomized cuts

Golovinskiy et al. [82] proposent une méthode de décomposition hiérarchique permettant la décimation de maillage à un nombre fini des triangles qui composent le segment initial, puis procède à une subdivision binaire de segments afin de guider la segmentation. En fait, la segmentation du maillage est induite via un ensemble d'échantillons. Cet ensemble est obtenu en exécutant plusieurs fois l'algorithme de partitionnement avec des initialisations différentes sur le maillage 3D. Au cours de cet algorithme proposé, une fonction de partition estime la vraisemblance d'appartenance de chaque arête à une frontière de segmentation, puis algorithme choisit les segments ayant le coût minimal. L'algorithme est terminé lorsqu'un nombre de segments déterminé par l'utilisateur est atteint. La figure 1.21, décrite par Aleksey Golovinskiy et al. [82], illustre cette approche.



(a) Segmentations randomisées (b) Fonction de partition

Figure 1.21 − Échantillon de segmentations randomisé produites avec le regroupement hiérarchique (a) et la fonction de partition résultante (b), [82].

1.3.2.4 Autres approches

Récemment, il existe des méthodes [2], [3], [97] qui reposent sur les réseaux de neurones et l'apprentissage profond donnant aussi des résultats satisfaisants. Shu et al [3] proposent une efficace méthode non supervisée de segmentation / co-segmentation de forme en se basant sur ce que l'on appelle l'apprentissage profond (deep learning). Au départ, chaque modèle 3D se décompose aux patchs primitifs afin de générer une sur-segmentation et calculé plusieurs signatures formes des modèles entrés. Les signatures utilisées sont les suivantes : SIHKS [98], SDF [4] et GC [99]. Cependant, chaque signature séparée peut uniquement caractériser une partie des caractéristiques géométriques. Par conséquence, un réseau d'apprentissage profond non supervisé se reconstruit dans l'étape suivante, un tel haut niveau des caractéristiques essentielles pourrait être extrait à partir des signatures géométriques calculées dans la première étape. Il faut noter que cette étape n'exige pas une opération d'étiquetage. Avec l'appui de l'apprentissage profond non supervisé, nous pouvons finalement segmenter les modèles en se basant sur les caractéristiques essentielles de haut niveau. La figure 1.22 montre un exemple de segmentation et co-segmentation en mettant en œuvre cet algorithme.

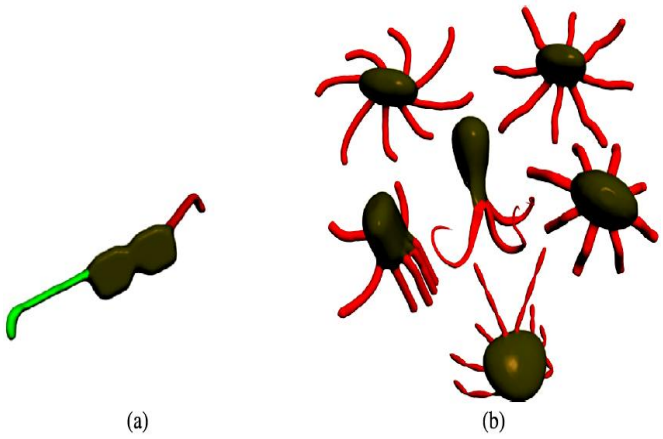


Figure 1.22 − Exemple de segmentation et co-segmentation. (a) Segmentation d'une seule forme. b) Co-segmentation d'une famille de formes similaires [3].

En dépit de ces efforts de recherche, il y a encore de nombreuses questions de recherche inexplorées concernant ce type d'approches. Premièrement, il n'est généralement pas évident de choisir la méthode d'apprentissage profond la plus adéquate pour segmenter un maillage, et quelles caractéristiques essentielles sont les plus efficaces. En outre, ces méthodes comportent des limitations importantes. Premièrement, la qualité des résultats finaux de segmentation dépend fortement de ceux de la sur-segmentation. Deuxièmement, le temps

d'exécution nécessaire pour l'apprentissage du réseau profond et la génération des caractéristiques essentielles de haut niveau est encore trop long.

En général, les méthodes de segmentation de maillage présentées dans cette partie donnent des résultats satisfaisants, mais souvent avec des modèles géométriques assez simple. En effet, la segmentation d'un objet ayant des caractéristiques géométriques assez complexe en parties significatives représente un défi épineux pour ces méthodes, comme il est montré par les expériences que nous avons réalisées dans ce travail.

1.4 Conclusion

Dans ce chapitre nous avons, premièrement, présenté les différentes étapes et les fondements mathématique concernant la reconstruction 3D et la segmentation de maillage. En particulier, nous avons présenté les méthodes avec lesquelles nous avons comparé les méthodes proposées dans cette thèse. Ensuite, nous avons présenté un état de l'art concernant plusieurs méthodes de segmentation de maillage 3D, tout en se focalisant sur celles segmentant le maillage en parties significatives.

Ces méthodes donnent généralement des résultats satisfaisants. Cependant, dans le cas d'un l'objet 3D avec des caractéristiques géométriques complexes, ces méthodes sont en face d'un problème épineux car ces méthodes sont basées sur la squelettisation ou sur la variation d'angle dièdre ou sur la distance géodésique. Ces deux critères restent insuffisants pour réaliser une segmentation significative des objets 3D. L’objectif du prochain chapitre et de proposer une nouvelle approche offrant une segmentation de maillage plus pertinente surtout au cas d'un maillage avec géométrie complexe.

Chapitre 2 - Nouvelles méthodes de mise en correspondance

2.1 Introduction

Depuis que D. Marr [91] a proposé une théorie de la vision stéréoscopique humaine, la communauté scientifique a porté un intérêt croissant aux systèmes binoculaires qui sont devenus très répandus. La réussite de l'établissement d'une correspondance fiable entre deux images est une étape clé dans de nombreux algorithmes de vision par ordinateur [6], [33], [100]. Par exemple : l'autocalibrage, la vision stéréo, la reconstruction 3D, les mosaïquages d'images, etc. La plupart des algorithmes de vision par ordinateur, y compris la stéréovision, reposent sur l'hypothèse de la détermination précise d'un ensemble de points correspondants entre les images mises en jeu. Cet ensemble de correspondances putatives inclut généralement la plupart des vraies correspondances (inliers) mais également un grand nombre de fausses correspondances (outliers), ou de valeurs aberrantes dues aux ambiguïtés des contraintes de similarité (en particulier si les images contiennent des motifs répétitifs).

Parmi les méthodes les plus robustes utilisées pour éliminer ou atténuer l'influence indue des données aberrantes (ou outliers), on trouve la méthode de RANSAC [11] et les méthodes analogues se situent dans le cadre d'hypothèse et de vérification [12], [20], [21]. Bien que ces méthodes aient beaucoup de succès dans de nombreuses situations, elles ont eu un succès limité si les contraintes géométriques sont non paramétriques et elles ont aussi tendance à se dégrader fortement si la proportion des données aberrantes dans l'ensemble de correspondances putatives devient importante [13]. Récemment, plusieurs méthodes ont été proposées en se basant sur les relations spéciales et cohérentes pour éliminer les faux appariements. Parmi ces méthodes, on trouve la méthode proposée par Chen et al. [28] qui résout le problème des relations spéciales et cohérentes entre un ensemble de points de correspondances putatives en s'appuyant sur une approche probabiliste sous un cadre bayésien. Cette méthode donne des résultats satisfaisants, par contre son efficacité diminue si la proportion des données aberrantes devient importante [101].

Afin d'améliorer la segmentation de maillage et pour remédier au problème de la mise en correspondances. Nous avons proposé deux contributions qui permettent de rejeter

efficacement les faux appariements et de choisir le sous ensemble le plus pertinent de points correspondants à la dernière étape de l'algorithme.

La suite de ce chapitre est organisée comme suit : Dans un premier temps, nous allons présenter la première méthode d'identification des correspondances de points par les cercles correspondants ICC [L1] (Identification of point correspondences by Corresponding Circles). Ensuite nous présenterons une amélioration de celle-ci en proposant l'algorithme de l'identification rapide des correspondances de points par les cercles correspondants FICC [L2] (Fast Identification of point correspondences by Corresponding Circles).

2.2 Identification des correspondances de points par les cercles correspondants ICC

Dans ce travail, nous avons introduit un nouveau concept d'algorithme de rejection des données aberrantes (ou faux appariements) à partir d'un ensemble des correspondances putatives. En se basant sur des cercles correspondants et du concept de similarité des images stéréoscopiques. Les faux appariements sont rejetés par le test d'appartenance à ces cercles. L'algorithme proposé améliore non seulement la précision de la détection des vrais appariements mais il offre également une grande souplesse pour sélectionner les vrais appariements en fonction des valeurs calculées qui expriment la chance d'être une valeur aberrante. Dans les sections qui suivent, nous présentons l'approche proposée. Ensuite nous montrons l'efficacité de celle-ci à travers une étude comparative.

2.2.1 Méthode proposée

L'objectif de l'algorithme ICC proposé est le rejet des données aberrantes dans les images stéréoscopiques en vue d'améliorer la segmentation de maillages. Cet algorithme reçoit comme données d'entrée deux ensembles discrets de points où chaque point fourni une information caractéristique et ponctuelle sur l'image. Il est important de rappeler que la stéréovision est basée sur le fait que l'information de profondeur peut être obtenue par triangulation à partir de deux images ayant une partie commune dans leur champ de vue.

2.2.1.1 Principe

La méthode proposée exploite la similitude entre les couples d'images stéréoscopiques en utilisant des couples de cercles basées sur des correspondances putatives. Ces cercles correspondants sont spécifiquement conçus pour décider, par un processus itératif, l'exactitude

de chaque correspondance de points putatifs, selon leur appartenance ou non aux cercles. En se basant sur une valeur cumulée associée à chaque correspondance de points putatifs, nous avons, à la fin du processus, la possibilité de regrouper les correspondances putatives en deux sous-ensembles de vraies et fausses correspondances.

Dans les sections suivantes, nous allons présenter l'algorithme proposé et nous analyserons sa complexité de calcul.

2.2.1.2 Moyens de vérification

La stratégie de notre approche consiste à supposer initialement que tous les points de correspondance donnés ne représentent ni les vrais appariements ni les faux appariements (voir la figure 2.1). Par la suite, nous procédons à la vérification de l'appariement des paires des points mises en jeu en utilisant un processus itératif basé sur l'utilisation de cercles ayant chacun un diamètre égal à la distance euclidienne séparant deux correspondances putatives, tout en exploitant la répartition de la scène3D.

La figure 2.1 montre les deux cas possibles de chaque paire de points parmi les correspondances putatives. La partie gauche illustre les vraies appariements (pr_i , pl_i), sachant que PR_i et PL_i désignent le même point 3D. Cependant, la partie droite illustre les faux appariements (pr_j , pl_j). PR_j et PL_j désignent deux points 3D différents dans la scène.

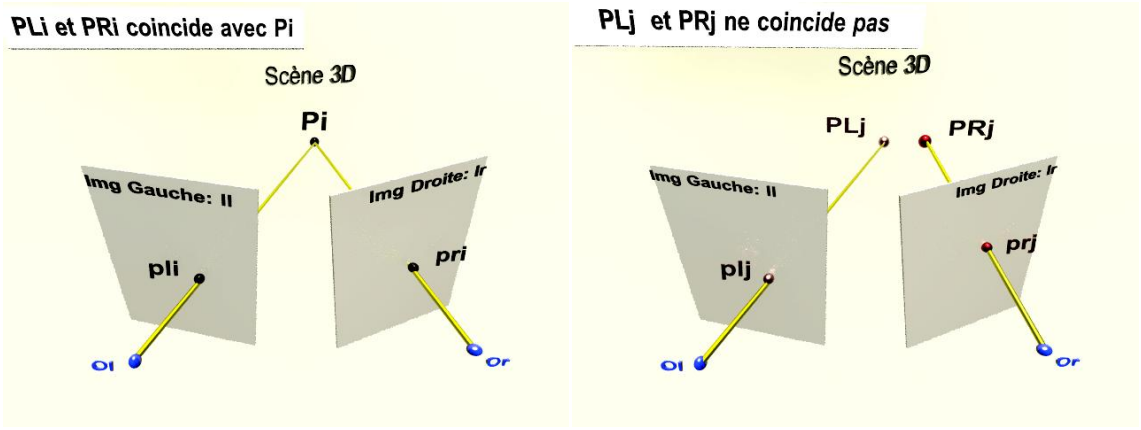


Figure 2.1 − Illustration de deux cas possibles de chaque paire de points parmi les correspondances putatives.

2.2.1.3 Extension des cercles dans la scène 3D

Les sous-sections suivantes établissent comment l'extension des cercles dans la scène 3D et la répartition de la scène 3D en trois zones peuvent nous aider à démontrer la validité et l'exactitude de notre approche.

2.2.1.3.1 Répartition virtuelle d'une scène

Comme le montre la figure 2.2 ci-dessous, nous considérons n correspondances putatives (ensemble de données initial). Parmi eux, nous choisissons deux correspondances (pr_i,pl_i) et (pr_j,pl_j) qui représentent respectivement la projection des points 3D P_i et P_j sur les deux images I_l et I_r. Ces deux paires permettent de créer une paire de cercles (cr_{ij}, cl_{ij}), tel que le centre de ck_{ij} (k=l ou k=r) est le milieu du segment [pk_i, pk_j], et le rayon est défini par la moitié de la distance euclidienne entre les deux points d'intérêt pk_i et pk_j. En prenant en considération la similarité des couples des images stéréoscopiques, ces deux cercles représentent deux régions ayant une similarité importante.

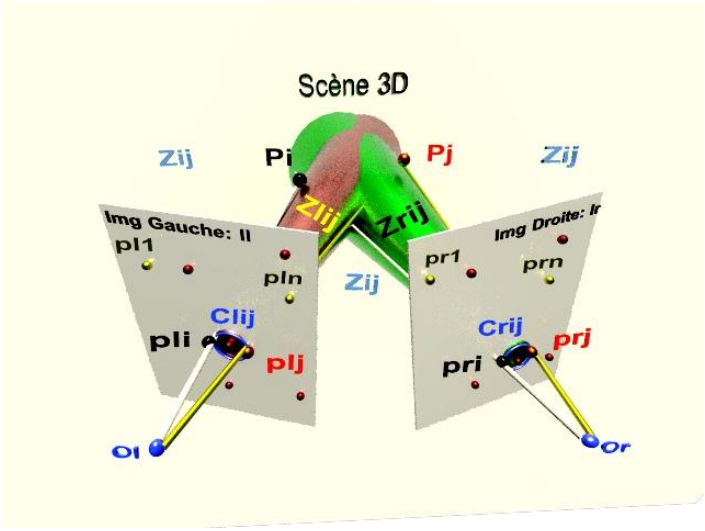


Figure 2.2 − Zones d'une scène 3D.

En se basant sur ces deux cercles, la scène 3D peut être divisée en trois zones : les deux zones Zk_{ij} (avec k = l ou k = r), représentant chacune un cône tronqué oblique (voir la figure 2.3) et la zone restante Z_{ij} .

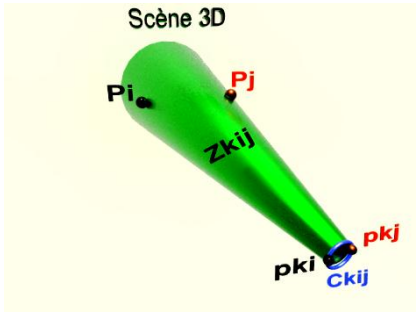


Figure 2.3 − Illustration d'un cône tronqué oblique représentant la zone virtuelle Zk_{ij} (k = l ou r) de la scène 3D. Cette zone est une extension du cercle ck_{ij} dans la scène 3D passant par les deux points 3D P_i et P_j, dont la petite base circulaire est le cercle ck_{ij}, et la hauteur dépend de la scène 3D.

2.2.1.3.2 Critères de localisation des valeurs aberrantes

La répartition de la scène décrite dans la section précédente permet, d'une part, de localiser chaque couple de correspondance et de vérifier, d'autre part, leur appartenance à l'une des zones de subdivision en s'appuyant sur des critères bien définis.

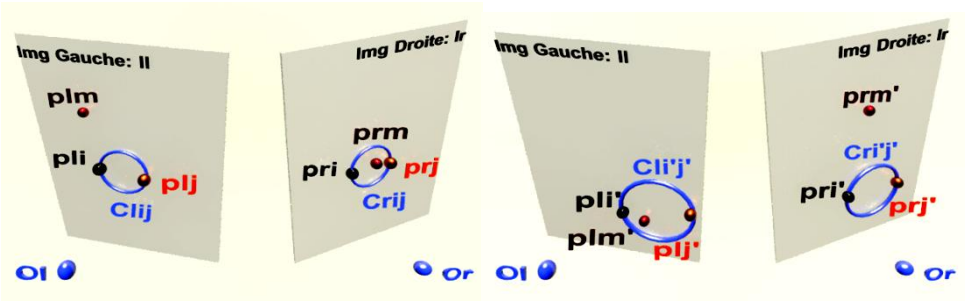


Figure 2.4 − Illustration de la stratégie adoptée. A gauche : $pr_m \in cr_{ij}$ et $pl_m \notin cl_{ij}$. A droite : $pr_{m'} \notin cr_{i'j'}$ et $pl_{m'} \in cl_{i'j'}$.

Pour mieux décrire l'approche proposée, nous considérons n correspondances putatives (pr_k, pl_k) comme ensemble de données initiales, avec k un nombre allant de 1 à n. On considère aussi deux couples de cercles (cr_{ij}, cl_{ij}) et $(cr_{i'j'}, cl_{i'j'})$, dont la création est basée respectivement sur les couples $\left((pr_i, pl_i), (pr_j, pl_j)\right)$ et $\left((pr_{i'}, pl_{i'}), (pr_{j'}, pl_{j'})\right)$. Enfin, on utilise deux correspondances (pr_m, pl_m) et $(pr_{m'}, pl_{m'})$ pour étudier leurs appartenances aux cercles correspondants, avec (i, i', j, j', m, m') ∈ (1..n).

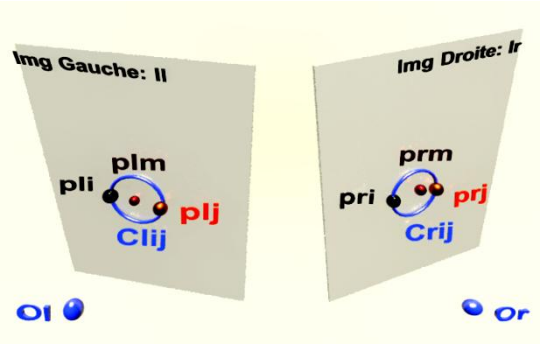


Figure 2.5 − Illustration de la deuxième stratégie non adoptée :$pr_m \in cr_{ij}$ et $pl_m \in cl_{ij}$.

En vue de localiser les vrais appariements, il y a seulement deux stratégies envisageables : soit l'on cherche à localiser les correspondances putatives (pr_m, pl_m) et $(pr_{m'}, pl_{m'})$ qui satisfont le critère suivant, $pr_m \in cr_{ij}$ et $pl_m \notin cl_{ij}$ (ou $pr_{m'} \notin cr_{i'j'}$ et $pl_{m'} \in cl_{i'j'}$ (voir la figure 2.4), ou bien nous essayons de localiser les points correspondants qui appartiennent complètement aux cercles correspondants (voir la figure 2.5). Puisque les deux stratégies représentent les deux cas possibles, alors il est inutile d'appliquer en même

temps les deux stratégies et, pour cette raison, nous avons choisi la première qui est considérée comme le premier critère.

2.2.1.3.3 Hypothèse et preuve

La méthode est basée sur les cercles et sur le critère adopté dans la section précédente, en vue d'identifier les données aberrantes à travers un processus itératif comme suit : pour chaque paire de cercles créés, toutes les correspondances putatives sont soumises au test d'appartenance à ce couple de cercles correspondant selon le critère adopté. Dans le cas où un point correspond à un point qui appartient au premier cercle et à un autre point qui n'appartient pas au second. Cela se traduit par une valeur cumulative qui dépend de la taille des cercles et qui est associée à cette paire de points. Lorsque cette valeur est augmentée, la possibilité que cette paire de points soit une donnée aberrante augmente également. A la fin du processus, les correspondances putatives sont regroupées en fonction de leurs valeurs cumulatives associées.

Afin de pouvoir prouver l'hypothèse selon laquelle les correspondances de points putatifs vérifiant le critère adopté plusieurs fois sont les plus susceptibles d'être des valeurs aberrantes, nous considérons une correspondance putative (pr_m, pl_m) et un couple de cercles (cr_{ij}, cl_{ij}) (où i, j, m sont trois indices qui représentent trois correspondances putatives parmi l'ensemble putatif). Nous essayons par la suite de prouver que quand (pr_m, pl_m) vérifie le critère d'appartenance aux cercles correspondants (cr_{ij}, cl_{ij}), alors sa possibilité d'être une paire aberrante augmente. De plus, afin de lever toute ambiguïté, la vérification est limitée à un seul cas d'appartenance ($pr_m \in cr_{ij}$ et $pl_m \notin cl_{ij}$) en ignorant le second cas. En effet, les deux cas sont très similaires, il suffit de substituer les symboles {R, r} qui se réfèrent à la droite par {L, l} qui se réfèrent à la gauche et vice versa. Sous ces contraintes, deux possibilités pourraient être traitées : la première c'est que la paire de points (pr_m, pl_m) est un vrai appariement, et la seconde, la paire constitue une donnée aberrante :

Première hypothèse : le couple (pr_m, pl_m) représenteles vrais appariements, ce qui implique que les deux points 3D PR_m et PL_m sont confondus, où (PR_m et PL_m) $\in Zr_{ij}$ (voir la figure 2.6).

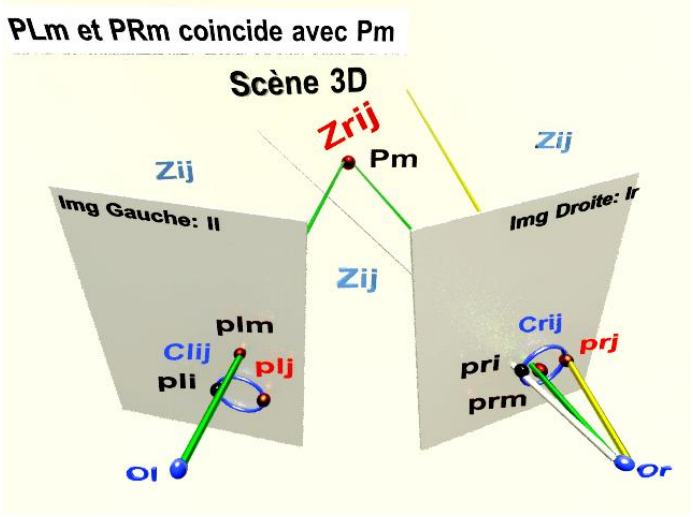


Figure 2.6 − La première hypothèse :(pr_m,pl_m) est un vrai appariement, avec $pr_m \in cr_{ij}$ et $pl_m \notin cl_{ij}$.

Deuxième hypothèse : le couple (pr_m, pl_m) représente les données aberrantes c.-à-d. les faux appariements, ce qui implique que les deux points 3D PR_m et PL_m ne sont pas confondus, où $PR_m \in Zr_{ij}$ et $PL_m \in Zr_{ij} \cup Z_{ij}$ (Figure 2.7).

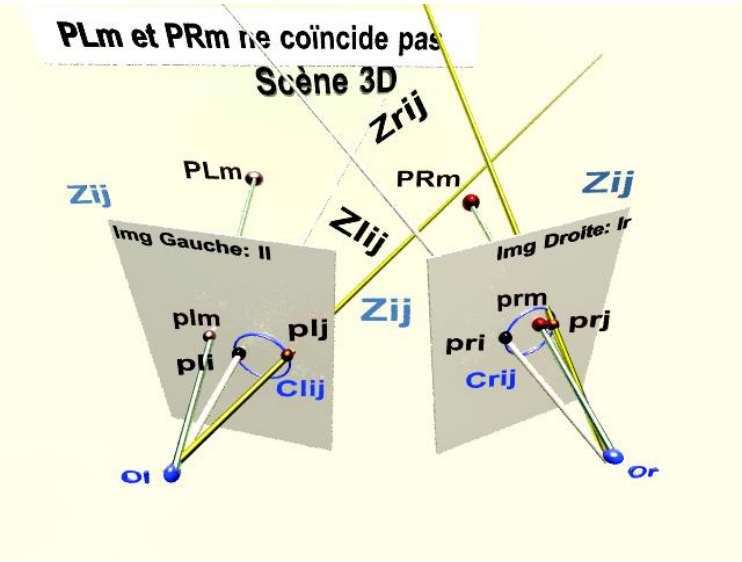


Figure 2.7 − La deuxième hypothèse : (pr_m, pl_m) est une donnée aberrante, avec $pr_m \in cr_{ij}$ et $pl_m \notin cl_{ij}$.

Deux résultats peuvent être tirés de ces deux hypothèses : si la première hypothèse est vérifiée, cela implique que :

$$PL_m \in Zr_{ij} \tag{2.1}$$

Si la deuxième hypothèse est vérifiée, cela implique que :

$$PL_m \in Zr_{ij} \cup Z_{ij} \tag{2.2}$$

On peut déduire de (2.1) et (2.2) que PL_m appartient à une zone plus large lorsqu'il s'agit d'un point aberrant (pr_m, pl_m), donc la possibilité que la paire (pr_m, pl_m) est une donnée aberrante sous ces conditions (c.à.d. la similarité des images et le premier critère adopté) est plus forte qu'un vrai appariement. De plus, cette possibilité dépend également de la taille du cercle (deuxième critère).

En effet, comme illustré sur la droite de la figure 2.8, le cercle $cl_{i'j'}$ tel que $pl_{m'} \notin cl_{i'j'}$ ayant la taille la plus grande, nous aide à classer la paire $(pr_{m'}, pl_{m'})$ avec davantage de précision comme donnée aberrante qu'un cercle cl_{ij} sur la gauche de la figure 2.8, où $pl_m \notin cl_{ij}$ de petite taille, tout en prenant en considération la notion de similarité des images stéréoscopiques pour arriver à cette conclusion .

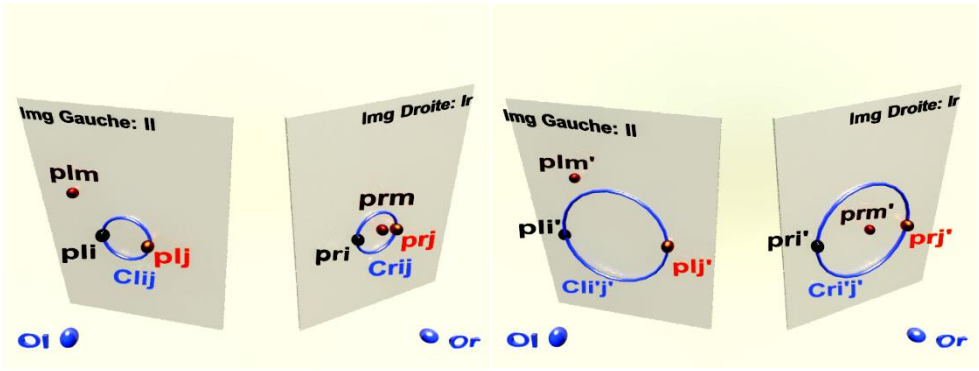


Figure 2.8 − Critère de la taille des cercles sous forme graphique : la taille du cercle $cl_{i'j'}$ est utilisée pour estimer la non-correspondance de la paire $(pr_{m'}, pl_{m'})$ avec davantage de précision que le cercle cl_{ij}.

Un grand nombre de données aberrantes peut provoquer une perturbation de toute la procédure et risque de créer des paires de cercles représentant deux zones non similaires de deux images stéréoscopiques et donc peu fiables. Pour pallier ce problème, nous envisageons une étape de prétraitement ayant comme objectif la sélection d'un sous-ensemble d'appariements fiables afin de créer des paires de cercles (cr, cl) adéquates.

2.2.1.4 Prétraitement

Le principe de cette étape repose sur la constatation que les images stéréoscopiques, grâce à la grande similitude qui caractérise généralement ces images, permettent de classer n correspondances putatives en deux catégories, comme le montre la figure 2.9 ci-dessous. La première catégorie, les segments noirs $[pr_j, pl_j]$ ont presque des angles identiques avec l'axe Ox et ont presque les mêmes distances euclidiennes entre (pr_j, pl_j). Il convient de noter que l'ensemble des correspondances peut contenir plusieurs groupes qui sont classés en fonction

de la paire angle-distance. La deuxième catégorie, les segments rouges [pr_i, pl_i] forment des angles avec l'axe Ox largement différents de ceux de la première catégorie et les segments verts [pr_k, pl_k] ayant des distances euclidiennes sont considérablement différents de ceux de la première catégorie, avec (i, j, k) des nombres compris entre 1 et n.

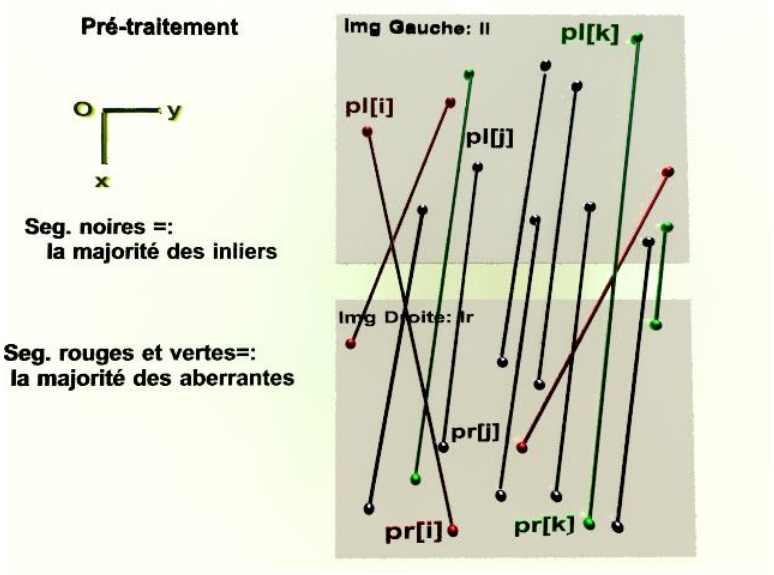


Figure 2.9 − Illustration des deux catégories de points correspondants que l'on peut trouver dans les images stéréoscopiques.

Par conséquent, tout en s'appuyant sur la notion de similarité, on peut dire que les segments noirs représentent la majorité des vrais appariements. La deuxième catégorie représente la majorité des paires aberrantes. Dans cette classification, deux variables sont utilisées : l'angle entre la ligne (pr, pl) et l'axe Ox et la distance euclidienne entre chaque correspondance (pr, pl). En utilisant ensuite la méthode (Expectation Maximization-Gaussian Mixtures) EM-GM [102] pour sélectionner les correspondances putatives (pr, pl) dont les angles et les distances euclidiennes appartiennent au même groupe.

Dans le pire des cas, le sous-ensemble sélectionné comprend un nombre considérable de paires aberrantes. Malgré cela, ces couples sont toujours valables pour créer des cercles correspondants (cr, cl) , grâce à la similarité importante qui caractérise les images stéréoscopiques. En effet, comme le montre la figure 2.10, les cercles jaunes sont des zones incertaines dans lesquelles chaque zone peut contenir, au pire, une paire aberrante. Cependant, ces zones sont relativement petites, en raison des contraintes d'angle et de distance adoptés lors de l'étape de prétraitement. Par conséquent, les paires de cercles créés sont toujours valables.

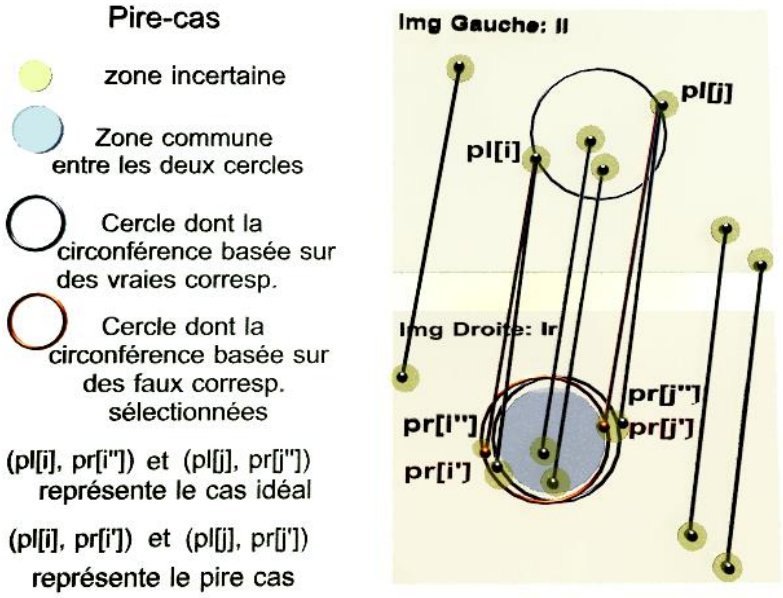


Figure 2.10 – Illustration d'une paire aberrante sélectionnée lors de l'étape de prétraitement.

2.2.1.5 Algorithme de l'approche proposée

Algorithme. 2.1 : Algorithme de sélection des vrais appariements (cas des images stéréoscopiques)

Donnéesn : nombre de correspondances putatives (pr, pl) détectées dans la paire d'images, avec t est un nombre compris entre 1 et n.

N : nombre de paires dans les classes sélectionnées.//Etape 1 de l'algorithme.

Variables CTBO[t]: Variable associée à la paire d'appariement(pr_t, pl_t). Plus cette valeur est grande, plus cela signifie que la paire (pr_t, pl_t) est une donnée aberrante, t représente un nombre compris entre 1 et n.

SC_k[i, j](K = l ou K = r): diamètre du cercle ck_{ij}, avec (i, j) représentent des nombres compris entre 1 et n

// **SC_k** est une matrice qui dépend de la taille du cercle ck_{ij}, dont le point $pk_t \notin ck_{ij}$. La possibilité d'identifier correctement la paire (pr_t, pl_t) dépend également de la taille des cercles, d'où la variable **SC_k**.

Résultat Sous-ensemble des vrais appariements.

Étape 1 Etape de prétraitement

Pour i = 1 à n faire

Calculer la distance euclidienne entre chaque paire (pr_i,pl_i).

Calculer l'angle entre la droite (pr_i,pl_i) et l'axe Ox pour chaque paire.

Fin pour

Classifier toutes les correspondances putatives en fonction de leurs valeurs (angle et distance).// Utiliser EM-GM [102] pour le regroupement.

Étape 2 Initialiser toutes les variables CTBO[t] à zéro

Pour i = 1 à N − 2 faire

Pour j = (i + 1) à N − 1 faire

// Créer deux cercles correspondants ck_{ij} (k = l ou k = r).

Choisissez deux couples (pr_i,pl_i) et (pr_j, pl_j)

$SC_k[i, j]$ = dist((pk_i,pk_j)). (k=l ou k=r).

Pour m = 0 à n − 1 faire // identification d'appartenance de chaque paire de points

//(pr_m,pl_m) aux cercles correspondants ck_{ij}. avec m # (i et j)

Si $pk_m \notin ck_{ij}$ et $pk'_m \in ck'_{ij}$ **alors**

CTBO[m] := CTBO[m] + $SC_k[i, j]$

Fin Si // avec ((k=l et k'=r) ou (k=r et k'=l))

Fin pour

Fin pour

Fin pour

Étape 3 Regroupement des paires (pr_t, pl_t) en deux sous-ensembles de points: Vrais appariements et données aberrantes. // Utiliser EM-GM [102] pour le regroupement.

Étape 4 Sélectionner un sous ensemble dont les CTBO sont les plus pertinents, parmi les celles déjà classifiés.

L'algorithme se compose de quatre étapes : La première est celle du prétraitement, visant à sélectionner un sous ensemble de points valides pour créer N (N − 1) / 2 paires de cercles. Ces cercles sont utilisés dans la deuxième étape pour localiser les différentes correspondances putatives et finalement quantifier la possibilité que ces correspondances soient aberrantes via la variable CTBO[t]. La troisième étape est celle de la classification des couples dont les valeurs les plus élevées de CTBO[t] représentent les paires aberrantes. Dans la quatrième étape, on peut sélectionner un sous-ensemble des vrais appariements les plus pertinents qui ont les plus petites valeurs de CTBO[t] sans devoir recalculer ce dernier. Cette caractéristique représente l'un des points forts de cette approche. Il faut préciser que le nombre des vrais appariements sélectionnés à la quatrième étape varie en fonction des besoins individuels.

2.2.1.6 Complexité de calcul

La complexité de notre méthode dépend du nombre de cercles créés lors de la deuxième étape de l'algorithme proposé. La complexité est alors proportionnelle à $N\,(N-1)/2$. Ainsi, l'algorithme proposé pour supprimer les paires aberrantes a une complexité $O(N^2 * n)$. Cela signifie que la complexité temporelle diminue lorsqu'un sous-ensemble N de points est faible. Par contre, l'algorithme proposé n'est pas influencé par l'augmentation de la proportion de paires aberrantes comme les algorithmes RANSAC [11] ou MLESAC [12]. En effet, lorsque la proportion de paires aberrantes augmente, le nombre de cercles créés devrait être réduit.

La méthode ICC se base sur le test d'appartenance aux cercles correspondants (cr, cl) créés pour localiser les points correspondants. Par un processus itératif, ce test permet d'affecter pour chaque paire (pr, pl) de n correspondances putatives, une valeur qui exprime leur chance d'être une paire aberrante basée sur deux critères. Ces valeurs offrent une grande flexibilité qui permet de sélectionner les vrais appariements à la fin du processus.

2.3 Identification rapide des correspondances de points par les cercles correspondants FICC

L'utilisation de l'algorithme ICC présente certains défis. Tout d'abord, chaque cercle correspondant créé lors de l'étape de prétraitement, comme décrit dans l'algorithme [L2], est utilisé pour étudier l'ensemble des correspondances putatives. Cependant, il y a généralement peu de correspondances putatives qui vérifient le premier critère concernant le test d'appartenance aux cercles correspondants. Deuxièmement, à la fin du processus, chaque correspondant putatif a une valeur qui exprime sa chance d'être une paire aberrante selon le deuxième critère lié à la taille des cercles. Malgré sa meilleure efficacité pour éliminer les paires aberrantes, le deuxième critère comme il sera expliqué plus loin, est encore insuffisant pour une sélection efficace des points correspondants fiables.

Dans cette partie, nous proposons un algorithme appelé : Identification rapide des correspondances de points par les cercles correspondants FICC (Fast Identification of point correspondences by Corresponding Circles) qui fournit une meilleure qualité et une meilleure efficacité par rapport à l'algorithme ICC.

FICC améliore la qualité de rejet des paires aberrantes. En particulier, au lieu de se baser sur la taille des cercles comme critère clef d'évaluation des correspondances putatives,

la FICC adopte un nouveau critère prenant en compte la distance entre chaque point (pr ou pl) et les cercles. Comme le montrent les résultats, ce critère permet la rejection des données aberrantes avec une précision plus grande et acceptable.

FICC augmente l'efficacité du calcul. En particulier, au lieu d'étudier l'ensemble des correspondances putatives dans chaque itération à l'instar de la méthode ICC, FICC utilise une approche pour sélectionner un petit sous-ensemble de correspondances putatives concernées par le test d'appartenance aux cercles correspondants. Cette modification entraîne des gains d'efficacité significatifs.

Dans la section suivante, nous présentons les étapes de l'approche proposée, puis nous montrons l'efficacité de celle-ci à travers une étude comparative.

2.3.1 L'approche adoptée

La méthode FICC proposée est une amélioration de l'algorithme ICC [L1], (voir la figure 2.11 ci-dessous). L'approche ICC originale utilise, à chaque itération de la méthode, un cercle correspondant pour localiser et quantifier chaque (pr, pl) selon deux critères. Le premier critère est conçu pour déterminer un sous-ensemble de points correspondants qui a une chance d'être une paire aberrante selon la condition suivante : dans le cas où une paire de points correspondante (pr_m, pl_m) avec un point appartenant au premier cercle $(cr_{ij}$ ou $cl_{ij})$, et l'autre n'appartient pas à la seconde, la probabilité que (pr_m, pl_m) soit une donnée aberrante devient plus forte. Où i, j et m sont trois indices représentant trois correspondances putatives parmi l'ensemble putatif. Ce critère est illustré à la Figure 2.4 ci-dessus. Le deuxième critère (voir la figure 2.8 ci-dessus) vise à estimer, à chaque itération du processus, une valeur CTBO[m] pour chaque (pr_m, pl_m) vérifiant le premier critère. CTBO contient des valeurs qui expriment la chance de chacun (pr_m, pl_m)d'être une paire aberrante en fonction de la taille des cercles $(cr_{ij},\ cl_{ij})$ en appliquant l'équation suivante :

$$\mathrm{CTBO[m]} = \mathrm{CTBO[m]} + \mathrm{SC[i][j]} \tag{2.3}$$

Où SC[i][j] est le diamètre du cercle.

Selon l'équation(2.3), la grande taille du cercle signifie que la possibilité de (pr_m, pl_m) d'être aberrante est élevée. À la fin du processus, une opération du regroupement nous a permis de sélectionner les vrais appariements en fonction de leurs valeurs CTBO. Le cadre de la FICC proposée est illustré sur la figure 2.11.

(a)



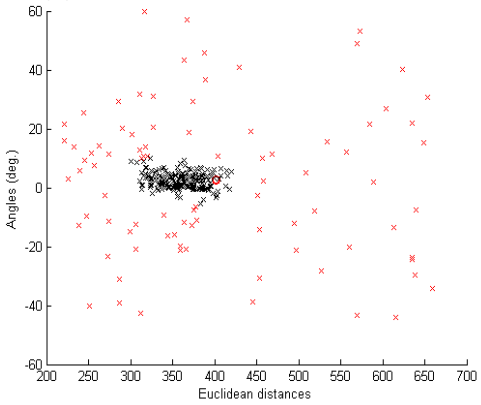
Détection des correspondances putatives initiales

Sélection d'un sous-ensemble des paires les plus appropriées pour créer les cercles correspondants

(b)



(c)



Utilisation des cercles correspondants pour évaluer la possibilité que chaque correspondance putative soit aberrante à travers les valeurs $CTBO$ basées sur leur position relative par rapport aux cercles.

(d)



Sélectionne les correspondances ayant les valeurs CTBO les plus faibles.

Figure 2.11 − Illustration de la méthode proposée. (**a)** une paire d'image d'entrée. (**b)** et (**c)** Une ligne-croix noire représente un sous-ensemble des correspondances putatives sélectionnées pour créer les cercles correspondants. (**d)** Quelques correspondances putatives avec une importante possibilité d'être aberrantes.

2.3.1.1 Limitation de la qualité ICC

Malgré son efficacité, le fait de se baser sur la taille du cercle comme critère d'estimation de la probabilité pour que chaque paire correspondante $(\mathrm{pr}, \mathrm{pl})$ soit aberrante est insuffisant. En effet, si par exemple deux correspondances putatives $(\mathrm{pr}_m, \mathrm{pl}_m), (\mathrm{pr}_n, \mathrm{pl}_n)$ vérifient le premier critère avec le même cercle$(\mathrm{cr}_{ij}, \mathrm{cl}_{ij})$, alors la méthode ICC estime que $(\mathrm{pr}_m, \mathrm{pl}_m)$ et $(\mathrm{pr}_n, \mathrm{pl}_n)$ sont les mêmes chances d'être une paire aberrante en s'appuyant sur la taille des cercles, quelle que soit leur position relative par rapport à ceux-ci, voir la figure 2.12. Ainsi, le deuxième critère est encore insuffisant pour une sélection efficace des points correspondants fiables.

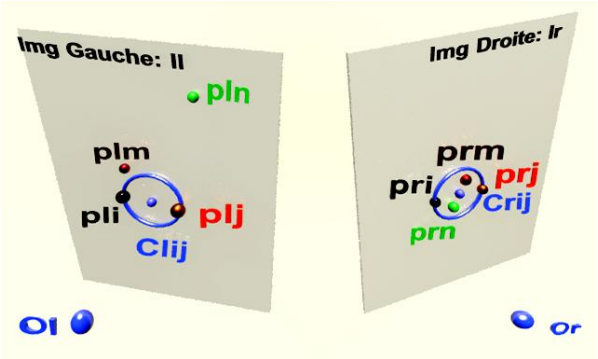


Figure 2.12 − Deux paires $(\mathrm{pr}_m, \mathrm{pl}_m)$ et $(\mathrm{pr}_n, \mathrm{pl}_n)$ qui satisfont au premier critère ICC avec le même cercle correspondant $(\mathrm{cr}_{ij}, \mathrm{cl}_{ij})$: malgré les différences entre les deux paires, la méthode ICC considère que CTBO[m]= CTBO[n].

2.3.1.2 Limitation d'efficacité de ICC

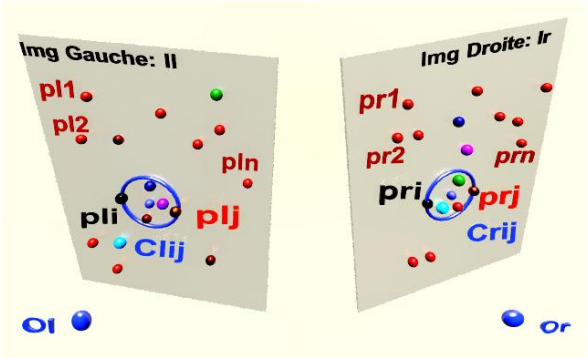


Figure 2.13 − Illustration de l'utilisation d'un cercle correspondant (cr_{ij}, cl_{ij}) pour estimer la correspondance correcte entre n points correspondances putatives. Les paires qui ne vérifient pas le premier critère sont illustrées par les points rouges et que les autres points représentent ceux qui satisfont au premier critère.

Selon la méthode ICC, l'ensemble des points correspondants ont été soumis au test, à chaque itération, sous le premier critère afin de sélectionner un sous-ensemble de points correspondants pouvant être des paires aberrantes. En fait, peu de points vérifient effectivement le premier critère parmi l'ensemble des points de correspondance, comme le montre la figure 2.13. Par conséquent, la complexité est relativement élevée.

2.3.1.3 Amélioration de la qualité du FICC

Pour améliorer la qualité de l'élimination des paires aberrantes, au lieu de se baser sur la taille des cercles (cr, cl) comme critère, nous proposons un nouveau critère qui prend en compte la position relative de chaque (pr, pl) par rapport aux cercles. Ce critère consiste à définir $CTBO[m]$ par la distance euclidienne d (Équation 2.4) entre pk_m et le centre d'un ck_{ij} donné, avec $pk_m \notin ck_{ij}$ ((pr_m, pl_m) doit d'abord vérifier le premier critère). Où m représente une paire correspondante parmi l'ensemble putatif, voir Figure 2.14.

$$CTBO[m] = CTBO[m] + d(pk_m, pk_{ij}) \tag{2.4}$$

Où pk_{ij} est le centre de cercle ck_{ij} et $pk_m \notin ck_{ij}$, avec ($k = r$ ou $k = l$).

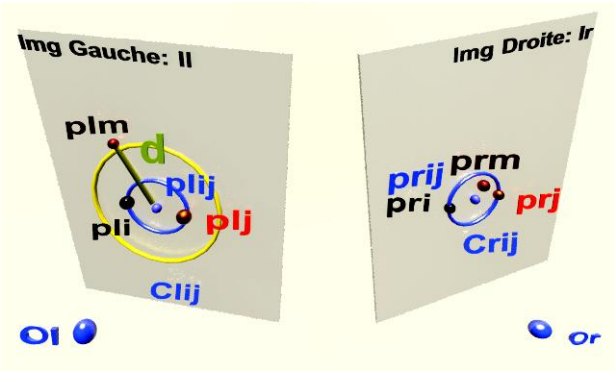


Figure 2.14 − L'illustration représente le deuxième critère FICC : le couple de cercles en bleu (cr_{ij}, cl_{ij}) est utilisé pour estimer la correspondance correcte d'une paire (pr_m, pl_m). Le cercle jaune qui a le même centre que cl_{ij} et de rayon plus long d. d représente la valeur la plus élevée exprime la chance que (pr_m, pl_m) d'être une paire aberrante en utilisant un cercle jaune au lieu de cl_{ij}.

Le cercle jaune vérifie également le premier critère, car cl_{ij} est inclus dans le cercle jaune et pl_m n'appartient pas au cercle jaune. De plus, le cercle jaune représente la valeur la plus élevée de d qui exprime la chance de (pr, pl) d'être une paire aberrante.

Par conséquent, en considérant d comme une valeur exprimant la chance de (pr_m, pl_m) d'être une paire aberrante sur une itération donnée, le FICC est plus précis que la méthode ICC pour la raison que FICC prend en compte la position de (pr_m, pl_m) par rapport au cercle.

2.3.1.4 Amélioration de l'efficacité du FICC

Pour améliorer l'efficacité, nous modifions l'algorithme afin de minimiser, à chaque itération, le nombre de correspondances putatives soumises au test d'appartenance au cercle, sous le premier critère, voir figure 2.15.

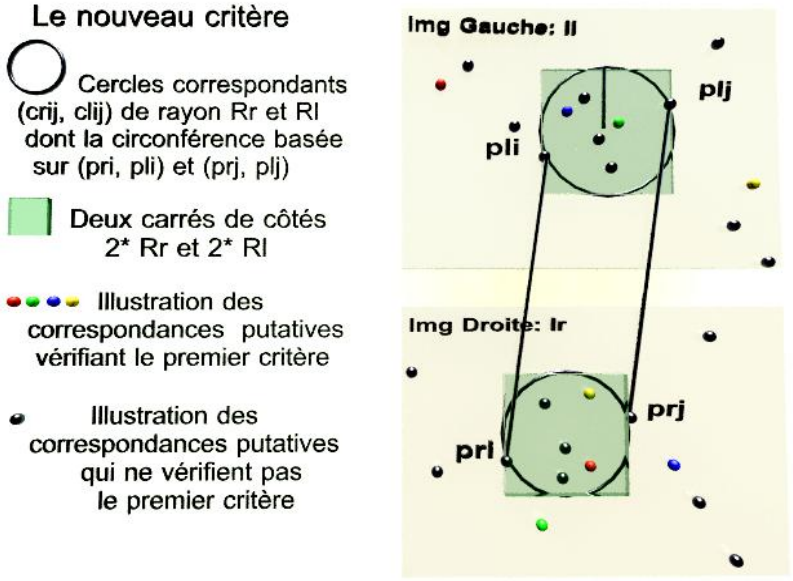


Figure 2.15 − Illustration de l'amélioration obtenue par la nouvelle stratégie visant à réduire le nombre de correspondances putatives testées à chaque itération.

Nous avons d'abord trié les deux ensembles de points pr et pl suivant les coordonnées (x, y), en utilisant la méthode de tri par fusion [103]. Ensuite, à chaque itération, nous étudions uniquement les correspondances de points qui ont eu la possibilité d'appartenir aux cercles. En effet, pour une paire de cercles donnée (cr_{ij}, cl_{ij}), le centre (xk_{ij}, yk_{ij}) de ck_{ij} $(k = l \text{ ou } k = r)$ est le milieu du segment $[pk_i, pk_j]$, et le rayon Rk_{ij} est défini par la moitié de la distance euclidienne entre les deux points d'intérêt pk_i et pk_j. Nous étudions ensuite uniquement les correspondances putatives $(pr(x_r, y_r)\ pl(x_l, y_l))$ où x_r, x_l, y_r, et y_l sont des coordonnées vérifiant l'une des deux conditions suivantes (selon le premier critère) :

$$\begin{cases} \begin{cases} xr_{ij} - Rr_{ij} < x_r < xr_{ij} + Rr_{ij} \text{ et } yr_{ij} - Rr_{ij} < y_r < yr_{ij} + Rr_{ij} \text{ et} \\ x_l \le xl_{ij} - Rl_{ij} \text{ et } x_l \ge xl_{ij} + Rl_{ij} \text{ et } y_l \le yl_{ij} - Rl_{ij} \text{ et } y_l \ge yl_{ij} + Rl_{ij} \end{cases} \\ \begin{cases} xl_{ij} - Rl_{ij} < x_l < xl_{ij} + Rl_{ij} \text{ et } yl_{ij} - Rl_{ij} < y_l < yl_{ij} + Rl_{ij} \text{ et} \\ x_r \le xr_{ij} - Rr_{ij} \text{ et } x_r \ge xr_{ij} + Rr_{ij} \text{ et } y_r \le yr_{ij} - Rr_{ij} \text{ et } y_r \ge yr_{ij} + Rr_{ij} \end{cases} \end{cases} \quad (2.5)$$

Comme le montre la figure 2.15, à travers les équations (2.5) , seules les correspondances putatives colorées vérifiant le nouveau premier critère sont prises en compte. Cette amélioration fait une grande différence d'efficacité entre les algorithmes FICC et ICC.

Le diagramme, présenté dans la Figure 2.16 ci-dessous, décrit l'enchaînement de notre qui se compose de quatre étapes essentielles :

1. Prétraitement : étape vise à atteindre deux objectifs principaux. Premièrement, sélectionner un sous-ensemble de N points, valide pour créer $N\,(N-1)/2$ paires de cercles, tout en utilisant la méthode EM-GM [102]. Deuxièmement, en triant les deux ensembles de points pr et pl suivant les coordonnées (x, y), en utilisant la méthode de tri par fusion [103]. La méthode de tri par fusion [103] est un algorithme de tri efficace dont la plupart des implémentations produisent un tri stable. La pire complexité de la méthode tri par fusion est $O\,(n \log n)$.

2. Evaluation : chaque cercle correspondant créé parmi l'ensemble de (cr, cl) est utilisé pour évaluer l'exactitude de correspondance de chaque correspondance putative et, finalement, quantifier la possibilité que ces correspondances soient des données aberrantes en utilisant la variable $CTBO[t]$ sur la base de l'équation(2.4). Les correspondances putatives k_j sont sélectionnées, à la jéme itération, selon les équations(2.5).

3. Regroupement : la méthode EM-GM [102] est utilisée pour regrouper les correspondances en utilisant. Les valeurs les plus basses de $CTBO[t]$ représentent les vrais appariements.

4. Paires pertinentes : nous pouvons sélectionner un sous-ensemble des vrais appariements les plus pertinents qui ont les plus petites valeurs de $CTBO[t]$ sans les recalculer. Il est aussi spécifié que le nombre des vrais appariements sélectionnés à la quatrième étape varie en fonction des besoins individuels.

2.3.1.5 Complexité de calcul

Le nombre d'échantillons prélevés par ICC dépend des $N\,(N-1)/2$ cercles correspondants et des npaires de points initiaux. Ensuite, utiliser l'algorithme ICC [L1] pour éliminer les paires aberrantes prend du temps $O(N^2 * n)$. Cependant, d'une part, l'algorithme FICC utilise le même nombre $N\,(N-1)/2$ des cercles correspondants. Par contre, à chaque itération j, nous n'utilisons que k_j correspondances putatives. Typiquement $k_j \ll n$.

Ainsi, l'utilisation de l'algorithme FICC pour rejeter les paires aberrantes prend du temps $O(N^2 * k)$. Où $k = \left(\sum_{j=1}^{N(N-1)/2} k_j\right)/(N(N-1)/2)$.

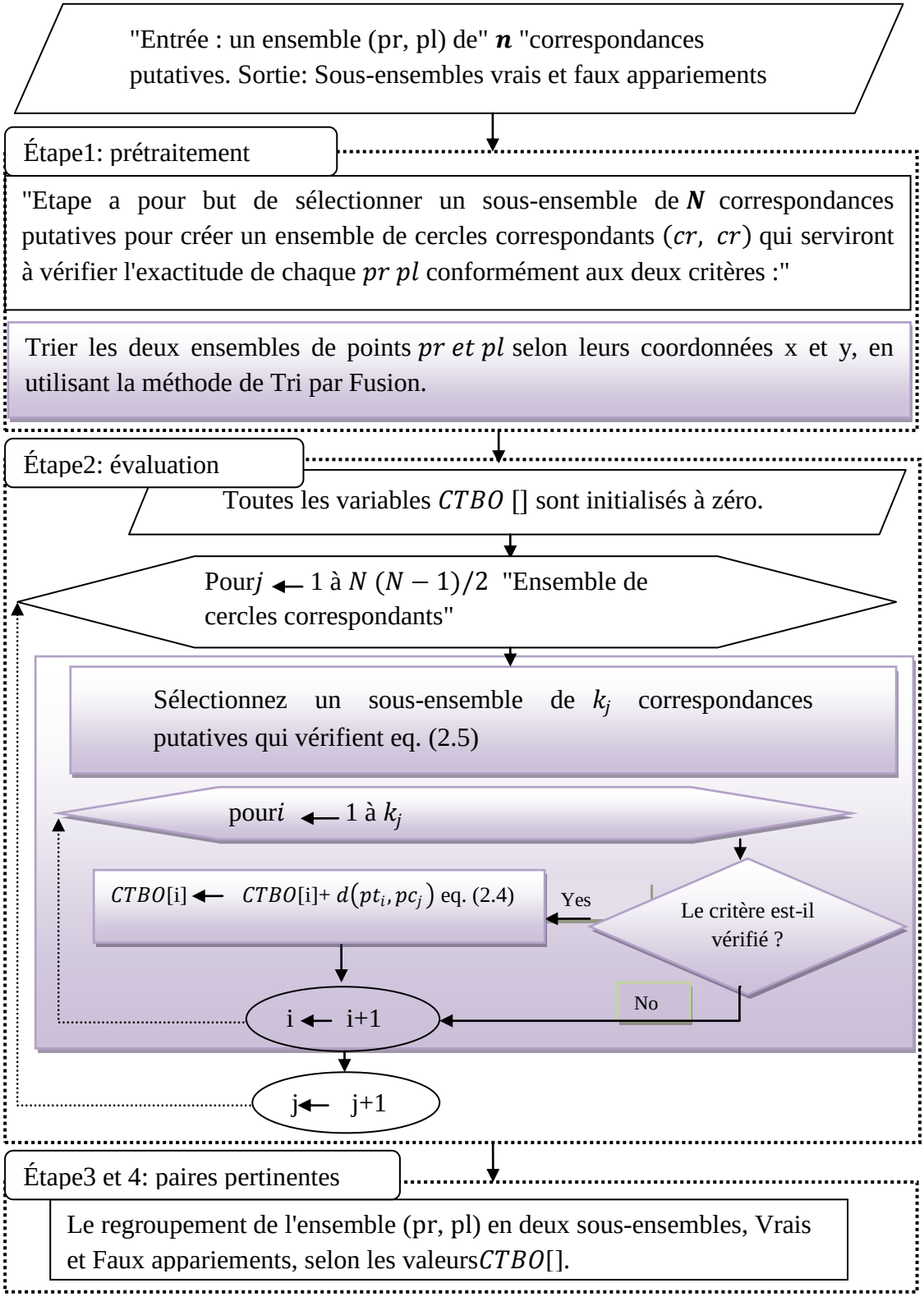


Figure 2.16 − Schéma descriptif de l'approche proposée. L'amélioration de la méthode FICC par rapport à ICC est illustrée par des parties colorées.

2.3.2 Expérimentation et résultats

Pour évaluer la robustesse des méthodes ICC et FICC, une étude comparative a été réalisée en se basant sur des données simulées. Nous avons comparé notre algorithme avec trois autres méthodes qui rejettent les paires aberrantes d'un ensemble de correspondances de points putatifs données. Premièrement, la méthode RSC proposée par Chen et al. [28], qui se base sur une approche probabiliste et qui associe chaque point de correspondance à une variable latente indiquant son exactitude. En conséquence, il existe une similarité entre la méthode de Chen [28] et les méthodes ICC et FICC. Cette similarité consiste à la classification des points correspondants tout en se basant sur une variable associée à chaque point de correspondance. Deuxièmement, RANSAC [11] et MLESAC [12] qui sont basées sur la contrainte géométrique (homographie). Ces méthodes montrent une grande efficacité dans de nombreux travaux [20], [21]. Nous avons implémenté la méthode RSC [28], et ajusté tous les paramètres en conséquence pour trouver les paramètres optimaux. Noté que la méthode RSC a été implémentée avec un modèle rigide. Pour la méthode RANSAC, nous avons utilisé sept paires de correspondances pour estimer la géométrie épipolaire, ainsi que le seuil de distance entre un point de données et le modèle mis à 0,0002. Pour le MLESAC, il a été implémenté en se basant sur des codes accessibles au public. Tout au long des expériences, tous les paramètres ont été fixés. Pour notre algorithme, il a été implémenté en Javafx en raison de la diversité de ses bibliothèques.

En ce qui concerne les méthodes FICC et ICC, deux résultats différents peuvent être obtenus avec et sans la quatrième étape de notre algorithme. Ces expériences ont été conçues afin d'évaluer l'efficacité de la variable CTBO à rejeter les paires aberrantes.

Dans toutes les expériences, les correspondances putatives initiales sont calculées à partir des points-clés SIFT [18] avec les paramètres par défaut, à l'exception du seuil de ratio de distance **t**. D'autre part, les méthodes EM-GM [102] et tri par fusion [103] ont été utilisées respectivement pour sélectionner des sous-ensembles de N correspondances putatives valides et pour l'opération de tri, comme il est illustré dans l'algorithme (étape de prétraitement). L'EM-GM [102] a de nouveau été utilisé à l'étape trois pour sélectionner les vrais appariements parmi les paires aberrantes.

Nous avons utilisé deux données de simulation pour tester notre méthode. Premièrement, 16 paires d'images stéréoscopiques appartenant aux bases de données stéréo Middlebury [46], [104]–[106] et, deuxièmement, l'ensemble de données de Mikolajczyk et al. [40]. Ces bases de données ont été choisies, d'une part, car elles contiennent des images

rigides et c'est le même type d'images demandé pour la reconstruction de maillage 3D dans cette thèse. D'une autre part, ces bases de données constituent un défi considérable pour les méthodes proposées et également pour les autres méthodes, car il y a tellement des transformations géométriques et photométriques et différents types de scènes. Nous avons ensuite testé la robustesse de notre algorithme en utilisant ces bases de données et la comparée aux trois autres algorithmes, tout en utilisant deux méthodes. Premièrement, nous avons généré trois ensembles de correspondances en ajoutant différents nombres de paires aberrantes supplémentaires. Deuxièmement, nous avons généré trois ensembles de correspondance en réduisant les vrais appariements.

La procédure que nous utilisons pour déterminer l'exactitude des correspondances pour les données de simulation décrites ci-dessous se compose de deux étapes. Nous générons d'abord des correspondances en utilisant SIFT en fonction des valeurs par défaut. En effet, la valeur par défaut de t est de 1.5, ce qui générera des correspondances très précises. Ensuite, nous ajustons la matrice fondamentale par RANSAC et l'utilisons pour déterminer la correction de correspondance.

Les résultats expérimentaux sont évalués via la paire précision-rappel de et la mesure F_β, où la précision est définie comme le rapport entre le nombre des vrais appariements préservé et le nombre de correspondances putatives préservé, équation 2.6. Le rappel est défini comme le rapport entre le nombre des vrais appariements préservé et le nombre des vrais appariements effectivement existants parmi les correspondances putatives, équation 2.7. La mesure F_β est une évaluation globale de la précision plus que du rappel, équation 2.8.

$$\text{Précision} = \frac{\text{Nombre des vrais appariements préservé}}{\text{Nombre de correspondances putatives préservé}} \tag{2.6}$$

$$\text{Rappel} = \frac{\text{Nombre des vrais appariements préservé}}{\text{Nombre des vrais appariements effectivement existants}} \tag{2.7}$$

$$F_\beta = \frac{(1+\beta) \times \text{Précision} \times \text{Rappel}}{\beta \times \text{Précision} + \text{Rappel}} \tag{2.8}$$

Où β est une valeur positive réelle et est fixé à $\beta = 0{,}5$.

2.3.2.1 Mise en correspondance de seize paires d'images de différentes bases de données

Dans cette section, nous présentons les résultats de 16 couples d'images stéréoscopiques appartenant aux bases de données stéréo Middlebury suivantes :ensembles de données de scènes planes par morceaux (Vénus, Poster et Tsukuba) [46], ensembles de données sur la réalité de terrain obtenues à l'aide de lumière structurée (Cône, Teddy, Art, Livres, Poupées, Moebius, Renne, Ordinateur et Nains) [104], [105], [107], enfin, ensembles de données stéréo haute résolution avec précision sous-pixel de la réalité de terrain (Flowers, Piano, Classroom1 et Umbrella) [106]. Pour chaque paire d'images, nous avons généré trois ensembles de correspondance selon la procédure suivante :

Pour chaque paire d'images, nous avons défini le seuil t de ratio des distances SIFT à 1.5, 1.2 et 1.0 respectivement, en vue d'obtenir un mélange initial des vrais et faux appariements. En effet, la plus grande valeur du seuil t augmente la proportion des vrais appariements. Pour augmenter davantage la difficulté de cette expérience, nous avons généré trois ensembles de correspondance en ajoutant respectivement 250, 500, 1500 données aberrantes aléatoires. En conséquence, le pourcentage des vrais appariements pour la totalité des paires d'images devient respectivement 45.11%, 31.89% et 14.70%. Ici, chaque paire aberrante est générée en choisissant aléatoirement un sommet de chacune des surfaces. Le taux moyen des vrais appariements initiaux de toutes les paires d'images est de 77,05% et environ 25% des ensembles de correspondance ont un pourcentage inférieur à 55%. En ce qui concerne la quatrième étape de notre méthode, nous sélectionnons 88% des vrais appariements les plus pertinents qui ont les plus petites valeurs de CTBO parmi celles sélectionnées à l'étape trois.

Les résultats des cinq méthodes sont résumés dans le tableau 2.1 ci-dessous qui représente la précision-rappel des paires d'images. Les moyennes de paires précision-rappel sont (95.01%, 93.17%) (92.82%, 97.36%) (94.47%, 97.04%) (97.93%, 93.02%) et (98.71%, 93,10%), respectivement pour RSC, RANSAC, MLESAC, ICC et FICC (avec la quatrième étape) respectivement. Comme le montre dans le tableau 2.1, toutes les méthodes donnent, généralement, des résultats satisfaisants et elles ont réalisés un F_β qui dépasse 60%. Cela est du, en grande partie, à la génération aléatoires des paires aberrantes, qui permet aux méthodes de sélectionner facilement les vrais appariements.

La méthode RSC permet d'obtenir les meilleurs résultats par rapport aux autres méthodes, à l'exception de nos méthodes ICC et FICC avec la quatrième étape. Ces résultats

obtenus par RSC montrent l'efficacité de la variable latente associée à chaque point de correspondance qui permet à la fin du processus de déterminer l'exactitude de chaque paire. Le MLESAC est un peu plus performant que le RANSAC et tous les deux réalisent des résultats tout à fait satisfaisants. D'une part, concernant la performance de notre méthode, avec et sans la quatrième étape, elle est satisfaisante et peut tolérer même 90% des données aberrantes. Ces résultats montrent l'importance et l'efficacité de la classification effectuée au cours de l'étape de prétraitement pour choisir un sous ensemble fiable permettant de créer les cercles correspondants. En effet, la performance de la classification n'est pas beaucoup influencée par les paires aberrantes aléatoires ajoutées, car la correspondance entre deux points aléatoires sera minimale.

Vrais appariements %	77.05% (original pct.)	45.11%	31.89%	14.70%
RSC	(95.01, 93.17) 94.83	(89.63, 85.16) 89.47	(84.43, 82.39) 84.08	(57.91, 80.45) 64.48
RANSAC	(92.82, 97.36) 94.28	(86.27, 90.30) 87.57	(77.52, 85.42) 79.98	(44.76, 78.34) 52.22
MLESAC	(94.47, 97.04) 95.31	(88.23, 84.89) 87.08	(83.05, 81.64) 83.01	(52.44, 71.36) 57.52
ICC Sans 4^{éme} étape	(92.15, 95.44) 93.22	(87.13, 88.35) 87.53	(81.17, 85.93) 82.57	(59.49, 69.78) 62.56
ICC Avec 4^{éme} étape	(97.93, 93.02) 96.23	(92.76, 86.14) 90.44	(87.22, 83.06) 85.78	(64.48, 66.21) 65.04
FICC Sans 4^{éme}étape	(93.25, 95.23) 93.90	(87.91, 89.28) 88.36	(83.47, 84.25) 83.72	(60.37, 68.72) 62.71
FICC Avec 4^{éme}étape	(98.71, 93.10) 96.76	(94.16, 87.03) 91.65	89.94, 81.90) 87.09	(64.14, 67.03) 65.07

Tableau 2.1 − Pourcentage des vrais appariements est modifié en ajoutant des paires aberrantes.

Au fur et à mesure que le pourcentage des vrais appariements initial diminue, la précision et le rappel de notre méthode et également des autres méthodes diminuent progressivement surtout sans la quatrième étape. Cependant, les résultats sont acceptables par rapport aux trois autres méthodes. D'autre part, nous observons que la performance de nos méthodes avec application de la quatrième étape est bien meilleure que sans elle, surtout en ce qui concerne la précision. En revanche, l'application de la quatrième étape diminue le rappel

de ICC et FICC, et pourtant, la mesure F_β a toujours augmenté. Cela montre que la majorité des paires de points supprimées à la quatrième étape sont des paires aberrantes, ce qui démontre l'efficacité de la variable CTBOdans l'élimination des paires aberrantes.

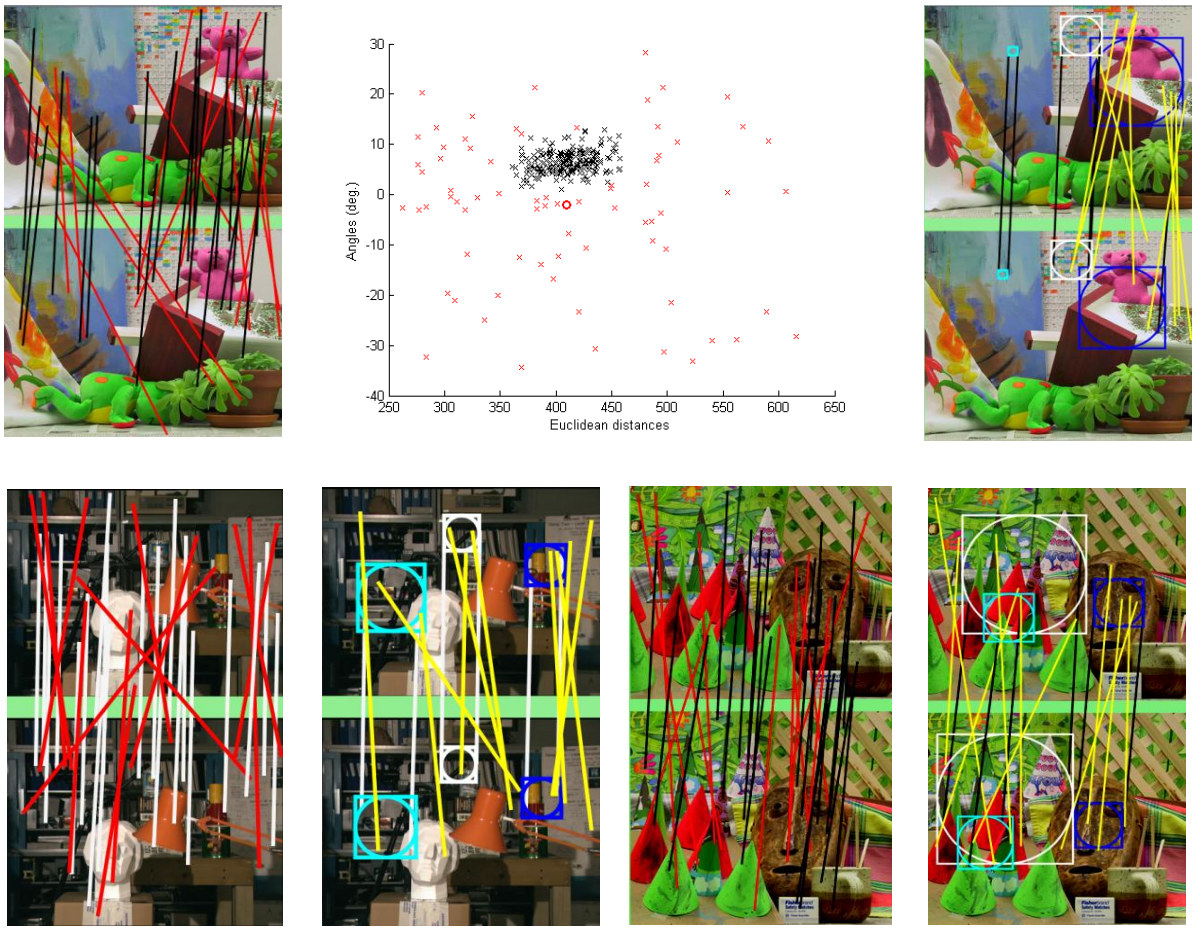


Figure 2.17 − Résultats expérimentaux sur trois paires d'images : Teddy, Tsukuba et Cônes. En haut : les résultats sur Teddy. En bas : les résultats sur Tsukuba et Cônes. Les carrés illustrent le critère établi par la méthode FICC concernant les paires suspectes d'être aberrantes. Pour la visibilité, un sous ensemble des correspondances est choisi au hasard pour l'illustration.

	RSC	MLESAC	RANSAC	ICC: Sans $4^{éme}$étape	FICC: Sans $4^{éme}$étape
Teddy	(96.63, 92.47)	(95.29, 93.58)	(93.78, 97.31	(93.72, 96.24)	(94.65, 95.16)
Tsukuba	(97.58, 91.45)	(95.67, 94.19)	(94.65, 100.0)	(93.89, 95.74)	(96.78, 96.58)
Cônes	(94.45, 87.45)	(93.36, 96.01)	(92.78, 95.14)	(93.08, 96.31)	(93.86, 96.24)

Tableau 2.2 − Comparaison des performances sur les paires d'images de Teddy Tsukuba et Cônes. Les paires dans le tableau sont Précision-Rappel(%).

Pour étudier davantage les deux méthodes sans application de la quatrième étape, nous avons choisi d'étudier trois paires d'images parmi les 16 couples, chacune contenant des

motifs répétés, Teddy, Tsukuba et Cônes. La paire Tsukuba et Cônes sont des scènes structurées et la paire Teddy est une scène non structurée, comme il est montré sur la Figure 2.17. La Figure 2.17 en haut montre les résultats sur Teddy, il y a 253 correspondances putatives avec 67 paires aberrantes, le pourcentage initial des vrais appariements est de 73,51%. Le graphique du milieu illustre l'opération de clustering pour sélectionner un sous-ensemble de correspondances putatives afin de créer des cercles correspondants. La Figure 2.17 en bas montre les résultats sur Tsukuba et Cônes, le pourcentage initial des vrais appariements est 77,09% et 70,07%, respectivement. Les lignes-croix noires et les lignes-croix rouges sur certaines paires d'images représentent, respectivement, le seul groupe de correspondances putatives choisi dans la première étape en se basant sur la méthode EM-GM [102] et les correspondances putatives non sélectionnées. D'autre part, les autres paires d'images représentent des correspondances putatives (lignes jaunes) qui vérifient les équations (2.6 et 2.7) et vérifie ensuite le premier critère (appartenant ou non aux cercles correspondants). Les carrés représentent les limites définies par la méthode FICC, (équations 2.6 et 2.7).

Les résultats de cette expérience sont présentés dans le tableau 2.2 ci-dessus. Les cinq méthodes réalisent des résultats satisfaisants. La RSC montre le meilleur score de précision. En effet, l'approche probabiliste adoptée par la RSC permet de sélectionner les paires pertinentes mais également de rejeter certains vrais appariements ce qui réduit le score de rappel. Le MLESAC est légèrement meilleur que RANSAC. La FICC et l'ICC ont un bon compromis entre précision et rappel. Cependant, FICC a un meilleur score de précision que la ICC. Ces résultats suggèrent que même avec la contrainte des motifs répétitifs, nos méthodes sont capables de rejeter avec efficacité les paires aberrantes. En fait, la classification des paires initiales pour créer des cercles correspondants en vue d'évaluer chaque correspondance et l'utilisation d'une variable associée à chaque paire permettent aux méthodes ICC et FICC, sans application de la quatrième étape, de réaliser des résultats satisfaisants.

2.3.2.2 Correspondance de caractéristiques basée sur la base de données Mikolajczyk

Dans cette section, nous avons testé notre méthode sur la base de données de Mikolajczyk et al. [40]. Elle contient huit sous-ensembles, chacun contenant six images prises par caméra dans une position fixe pendant l'acquisition. Pour chaque sous-ensemble, nous construisons cinq paires d'images en associant la première image à chacune des cinq images restantes. En conséquence, il y a au total 40 paires d'images pour l'expérience. La précision

moyenne de toutes les paires d'images est de 67.19% et environ 34% des ensembles de correspondance ont un pourcentage inférieur à 50%. Ici, nous avons testé la robustesse de notre méthode en réduisant le pourcentage des vrais appariements. Pour chaque paire d'images, nous avons généré trois ensembles de correspondance selon la procédure suivante : Le seuil de ratio des distances SIFT est défini à 1.4, 1.1 et 1.0 respectivement, puis on élimine de manière aléatoire le même pourcentage des vrais appariements, de sorte que le pourcentage des vrais appariements pour la totalité des paires d'images devient respectivement 46.39%, 30.90% et 12.95%. En ce qui concerne la quatrième étape de notre méthode, nous sélectionnons 15 % de la correspondance préservée.

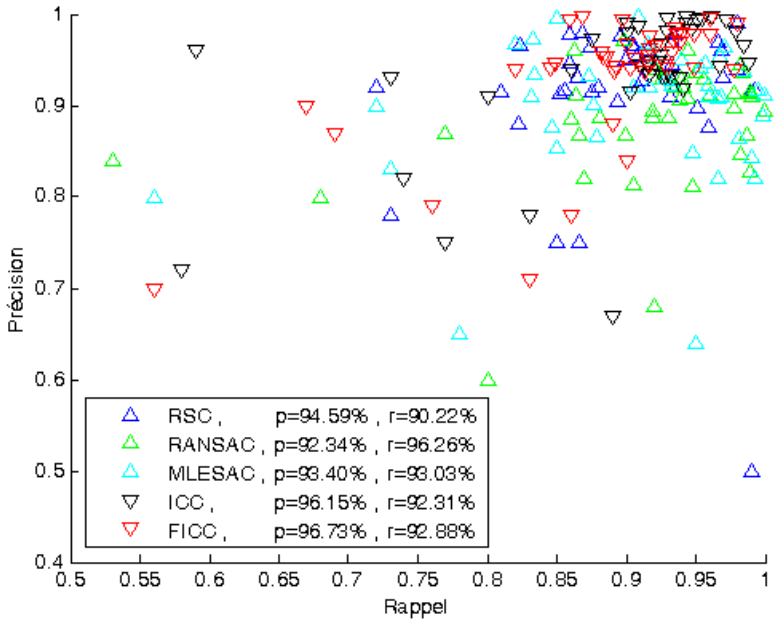


Figure 2.18 − Résultats expérimentaux du base de données Mikolajczyk [40]. Statistiques de Précision-rappel sans réduire les vrais appariements pour RSC, RANSAC, MLESAC, ICC avec 4ème étape et notre FICC avec 4ème étape. ICC et FICC ont les meilleures précisions globales.

Les résultats des cinq méthodes sont résumés sur la figure 2.18. Dans ces résultats, chaque point dispersé représente une paire précision-rappel pour une paire d'images. Les précision-rappel moyennes sont (94,59%, 90,22%) (92,34%, 96,26%) (93,40%, 93,03%) (96,15%, 92,31%) et (96,73%, 93,88%) pour RSC, RANSAC, MLESAC, ICC avec la quatrième étape et FICC avec la quatrième étape respectivement.

Le tableau 2.3 présente les résultats complets de cinq méthodes avec des taux réduits des vrais appariements. Pour la mesure globale F_β, toutes les méthodes sont supérieures à 45%. En revanche, ces résultats sont inférieurs par rapport à ceux réalisés avec l'expérience

illustrée dans le tableau 2.1. En effet, l'expérience du tableau 2.1 se base sur la génération des paires aberrantes, à l'inverse de cette expérience.

Inlier pct.	67.19% (originale pct.)	46.39%	30.90%	12.95%
RSC	(94.59, 90.22) 93.08	(86.43, 87.08) 86.64	(80.52, 80.70) 80.57	(44.46, 72.91) 51.10
RANSAC	(92.34, 96.26) 93.61	(81.58, 89.64) 84.10	(72.01, 90.15) 77.18	(35.94, 77.42) 43.75
MLESAC	(93.40, 93.03) 93.27	(83.13, 88.46) 84.83	(75.24, 88.30) 79.14	(42.71, 71.87) 49.38
ICC Sans 4^{éme}étape	(90.08, 95.56) 91.83	(81.23, 90.07) 83.97	(75.63, 84.38) 78.33	(46.82, 66.03) 51.84
ICC Avec 4^{éme}étape	(96.15, 92.31) 94.83	(88.15, 86.42) 87.56	(80.45, 80.74) 80.54	(52.37, 62.05) 55.24
FICC Sans 4^{éme}étape	(91.13, 95.67) 92.59	(82.94, 89.04) 84.87	(75.93, 85.78) 78.95	(48.92, 60.41) 52.23
FICC Avec 4^{éme}étape	(96.73, 92.88) 95.41	(89.26, 85.69) 88.03	(81.02, 82.14) 81.38	(53.10, 56.80) 54.27

Tableau 2.3 – Le pourcentage des vrais appariements est modifié en réduisant le nombre des vrais appariements.

Comme indiqué dans le tableau 2.3, RSC, d'une part, a un compromis satisfaisant entre la précision et le rappel, et d'autre part, MLESAC et RANSAC produisent des résultats assez similaires, et tous deux obtiennent des performances assez satisfaisantes. En revanche, au fur et à mesure que le ratio de paires aberrantes augmente, la performance de MLESAC et RANSAC devient assez faible, et c'est le point faible de la majorité des méthodes de rééchantillonnages. L'expérience montre également que l'ICC et la FICC peuvent, à la fois, identifier des correspondances correctes et rejeter les paires aberrantes, par rapport aux trois autres méthodes alternatives, avec une probabilité élevée, en particulier avec la quatrième étape.

La FICC est plus performante que la ICC, quel que soit le ratio des paires aberrantes. Cependant, au fur et à mesure que le pourcentage des vrais appariements diminue, la précision et le rappel de la ICC et FICC diminuent. En effet, les méthodes proposées se basent sur le regroupement des paires à l'étape de prétraitement, ce regroupement devient moins efficace

avec l'augmentation de la proportion des paires aberrantes. En revanche, nous observons que les performances du FICC à la quatrième étape sont bien meilleures que sans elle, surtout en ce qui concerne la précision. Cela montre que la majorité des correspondances rejetées lors de la quatrième étape sont des paires aberrantes, ce qui démontre l'efficacité du nouveau critère, à travers la variable CTBO, pour éliminer les paires aberrantes.

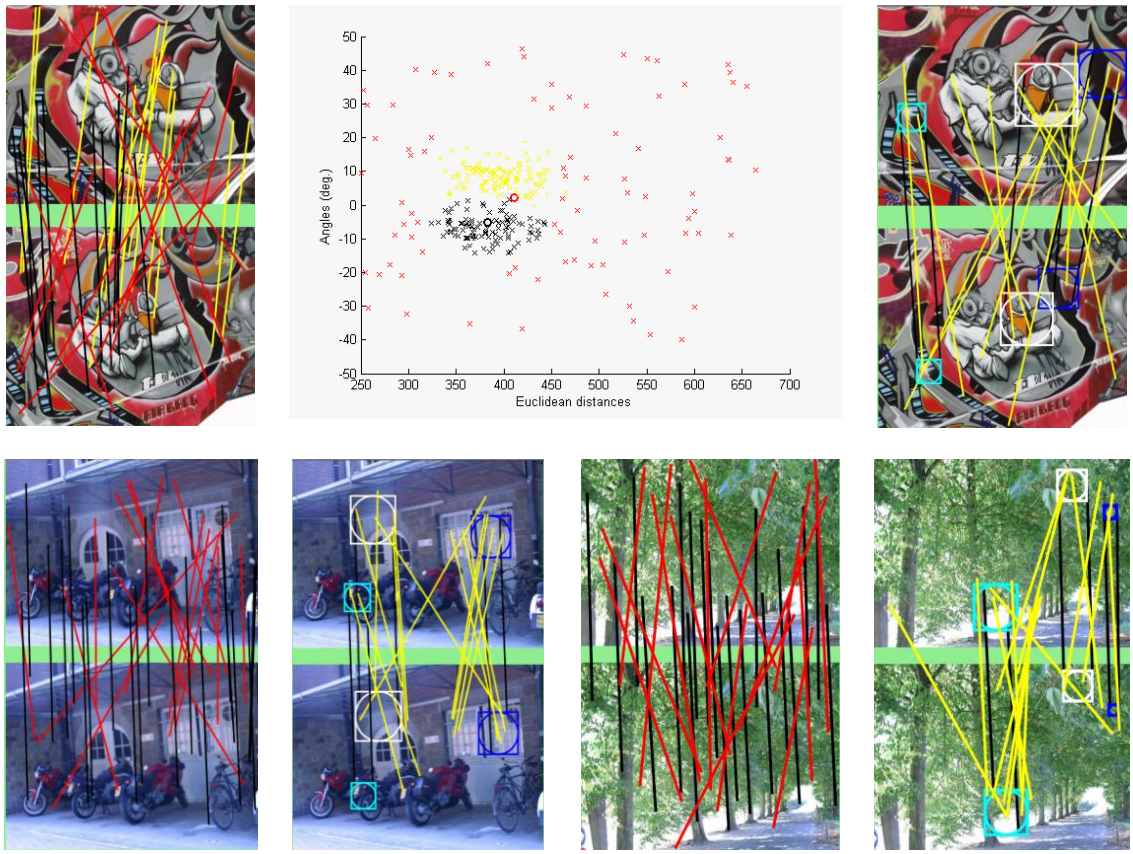


Figure 2.19 − Résultats expérimentaux sur trois paires d'images : *Graffiti*, *Bikes* et *Trees*. En haut : résultats sur *Graffiti*. En bas : résultats sur *Bikes* et *Trees*. Les carrés illustrent le critère établir par la méthode FICC concernant les paires suspectes d'être aberrantes. Pour la visibilité, un sous ensemble des correspondances est choisi au hasard pour l'illustration.

Pour étudier plus avant les méthodes proposées avec application cette fois de la quatrième étape (rejet de 10%), nous avons choisi d'étudier trois paires d'images (*Graffiti*, *Bikes* et *Trees*) de la base de données Mikolajczyk. La première paire d'images (*Graffiti*) se compose de scènes de grands changements de points de vue. La deuxième paire d'images (*Bikes*) comprend des scènes qui ont subis des changements importants d'éclairage. La troisième paire d'images (*Trees*) sont constituées de scènes de motifs répétitifs, comme illustré sur la Figure 2.19 ci-dessus. Pour la paire *Graffiti* comme indiqué sur la ligne supérieure de la figure 2.19, le pourcentage des vrais appariements initiaux donnés par la méthode SIFT [18] est de 50,38%.

Le graphique du milieu illustre l'opération de clustering effectuée à l'étape de prétraitement. Pour les paires *Bikes* et *Trees*, comme indiqué sur la ligne inférieure de la figure 2.19, le pourcentage initial des vrais appariements est 71,92 et 49,60%, respectivement. Les lignes-croix noires et les lignes-croix rouges sur certaines paires d'images représentent, respectivement, le seul groupe de correspondances putatives choisis dans la première étape en utilisant la méthode EM-GM [102] et de correspondances putatives non sélectionnées. Les lignes noires et jaunes tracées sur les autres paires d'images indiquent, respectivement, certaines des correspondances putatives utilisées pour créer les cercles correspondants, et les correspondances putatives les plus susceptibles d'être des paires aberrantes basées sur ces cercles. Les carrés représentent les limites définies par la méthode FICC, (équations 2.5).

	RSC	MLESAC	RANSAC	ICC: Avec 4éme étape	FICC: Avec 4éme étape
Graffiti	(91.47, 89.39)	(89.78, 93.18)	(88.73, 96.45)	(91.73, 92.42)	(92.37,91.67)
Bikes	(97.82, 92.12)	(94.53, 100.0)	(95.14, 100.0)	(96.00, 94.21)	(98.00,94.03)
Trees	(90.10, 86.47)	(87.05, 88.33)	(85.53, 92.65)	(91.61, 90.28)	(92.12, 89.37)

Tableau 2.4 −Tableau montrant la comparaison des performances sur les paires d'images de *Graffiti*, *Bikes* et *Trees*. Les paires dans le tableau sont Précision-Rappel(%).

Les performances des cinq méthodes sont présentées dans le tableau 2.4. La RSC montre un score de précision-rappel satisfaisant et sa performance n'est pas beaucoup influencée par le changement de points de vue entre les deux images. Cela est due à l'efficacité de l'approche probabiliste adoptée par la RSC. MLESAC et RANSAC produisent un score précision-rappel satisfaisant. Par contre, la performance des deux méthodes est un peu diminuée avec l'augmentation de la proportion des paires aberrantes. Après avoir utilisé les méthodes ICC et FICC avec la quatrième étape, les résultats montrent que même avec la présence de la contrainte des motifs répétitifs, de flou et de grand changement de point de vue, les méthodes proposées sont toujours efficaces dans le rejet des paires aberrantes. Par exemple pour la paire *Graffiti*, 122 et 121 correspondances sont préservées, respectivement, par la ICC et la FICC parmi 132 vrais appariements initialement existants.

Ces résultats suggèrent que la ICC et la FICC sont robustes au changement de point de vue. En effet, comme il est montré sur la figure 2.19, les deux classes des paires initiales obtenues au cours de l'étape de prétraitement, permettent aux méthodes ICC et FICC de créer des cercles correspondants plus adaptatifs et plus flexibles avec le changement de point de vue entre les images Graffiti. Cette flexibilité permet d'augmenter la performance des deux méthodes.

2.3.2.3 Résultats concernant le temps d'exécution

Afin de comparer la complexité temporelle de notre nouvelle méthode avec la version ICC, le même nombre de correspondances putatives pour créer le même nombre de cercles correspondants est utilisé sur les deux méthodes. De plus, nous avons ajouté le temps nécessaire à l'opération de tri au temps obtenu par FICC.

Pour mener cette étude, nous évaluons le temps de calcul des deux méthodes sur les bases de données Middlebury Stereo et Mikolajczyk, et nous définissons le seuil t de ratio des distances SIFT à 1, 1.1 et 1.3 respectivement. Les résultats du calcul sont présentés dans le tableau 2.5. Le FICC est environ douze et onze fois plus rapide que la méthode ICC avec les bases de données Middlebury Stereo et Mikolajczyk respectivement. Cette grande amélioration est principalement due à la nouvelle stratégie visant à réduire le nombre de correspondances putatives testées lors de la deuxième étape.

La caractéristique la plus notable de FICC est son efficacité. Le FICC est beaucoup plus efficace que la ICC, quel que soit le ratio des données aberrantes.

Base dedonnée	FICC	ICC
Middlebury Stereo (16 paires images)	12.65	153.46
Mikolajczyk (40 paires images)	33.06	371.20

Tableau 2.5 − Tableau montrant le temps de calcul des méthodes FICC et ICC (en Sec).

2.3.3 Discussion

En analysant les résultats obtenus à partir des expériences, nous avons constaté que la FICC se caractérise par sa précision par rapport aux trois autres méthodes tout en utilisant deux bases de données différentes, comme le montrent les tableaux 2.1 et 2.3 ci-dessus. En effet, RANSAC, comme la plupart des méthodes de rééchantillonnage, manque de robustesse quand il y a beaucoup de paires aberrantes. RSC, en général, a une grande précision. D'autre part, la FICC, comme le montrent les résultats, est un outil efficace pour rejeter les paires aberrantes. Tout d'abord, grâce à une stratégie basée sur la sélection, dans un premier temps, d'un sous-ensemble de correspondances putatives fiables entre les images afin d'évaluer la justesse de la correspondance par un processus itératif. Deuxièmement, grâce à l'utilisation de l'algorithme de clustering pour sélectionner les éléments les plus pertinents dans la troisième étape de l'algorithme FICC, sur la base du nouveau premier critère qui a montré son efficacité à travers plusieurs expériences. De plus, le FICC est une grande amélioration en termes de

temps de calcule de ICC, comme montre dans le tableau 2.5 ci-dessus, en raison du nouveau deuxième critère.

Dans cette partie, deux nouveaux algorithmes pour rejeter les faux appariements appelés ICC et FICC sont proposés. Ce dernier réduit le nombre de correspondances putatives testées à chaque itération, sans entraîner de diminution des performances. D'autre part, elle apporte une amélioration apparente en prenant en compte la position relative des correspondances putatives qui vérifient le premier critère avec les cercles correspondants. Nos résultats expérimentaux montrent plusieurs avantages par rapport à la ICC d'origine et se comparent favorablement aux différentes bases de données.

2.4 Conclusion

Dans ce chapitre, nous avons proposé deux algorithmes robustes et flexibles d'élimination des paires aberrantes. Pour atteindre un tel objectif, les correspondances des caractéristiques entre les paires d'images sont déterminées en se basant sur des cercles correspondants, tout en exploitant la similarité des paires d'images stéréoscopiques. À travers le test d'appartenance à ces cercles, nous pouvons, par un processus itératif, attribuer à chaque paire de correspondances putatives une valeur qui exprime leur chance d'être une paire aberrante. Ces valeurs offrent une grande flexibilité qui permet de sélectionner les vrais appariements à la fin du processus. L'application des algorithmes proposés sur deux bases de données réelles démontre leur efficacité d'atteindre des performances globalement satisfaisantes, et elle montre une efficacité envers la contrainte de changement d'éclairement extrême, le changement de point de vue et les transformations affines.

Un nombre prépondérant d'application de vision par ordinateur telles que la création de modèles 3D commence par supposer que les points de correspondances ont été récupérés avec succès. Par conséquence, ICC et FICC représentent deux solutions convenables qui peuvent être utilisées dans le processus de la reconstruction du maillage 3D.

Chapitre 3 - Segmentation significative des maillages 3D complexes via la primitive couleur des images stéréoscopiques

3.1 Introduction

La segmentation du maillage 3D représente une solution préférée pour de nombreuses tâches en raison de sa simplicité et de sa flexibilité, par exemple, le placage de texture [108], la reconnaissance de formes [109], la paramétrisation de maillage [110], le remaillage [111], la simplification [112], la détection de collision [113], l'ingénierie inverse et l'extraction de squelettes [4], [85]–[87], [94], [114].

La plupart des méthodes de segmentation de maillages 3D reposent uniquement sur des techniques géométriques et négligent l'information de la couleur. En fait, l'information fournie par la couleur constitue le moyen le plus naturel pour segmenter les objets. Elle permet une bonne perception visuelle des objets, par exemple on peut facilement distinguer les fleurs dans un bouquet de fleurs à travers les informations de couleur que par leurs caractéristiques géométriques.

Dans ce contexte, nous proposons une méthode qui permet à la fois une grande précision et une meilleure efficacité. La méthode proposée permet de segmenter un maillage 3D en parties significatives. Le but de celle-ci est de partitionner le maillage de manière efficace en se basant sur une classification des couleurs des images mises en jeu dans l'étape de reconstruction 3D du modèle.

Comparée aux méthodes existantes, la méthode proposée présente deux principaux avantages : 1) Au lieu de s'appuyer sur les caractéristiques géométriques des objets, on peut segmenter efficacement les maillages 3D en parties significatives en se basant sur les couleurs des images mises en jeu dans l'étape de reconstruction 3D avec des erreurs acceptables. En effet, cette méthode permet de tirer profit de l'intuition et de la simplicité relative du traitement de la couleur par rapport à celle de la géométrie, notamment pour les objets à topologie complexe. 2) Il s'agit d'une méthode efficace sur le plan computationnel par rapport aux méthodes de l'état de l'art, dans la mesure où la principale étape de la méthode consistant

à classifier les couleurs extraites des images n'est pas coûteuse en termes de temps de calcul, (Voir la figure 3.1).

La méthode proposée prend en entrée un ensemble d'images, observant un objet rigide. En effet, elle repose sur l'approche de reconstruction stéréo multi-vues (MVS) qui permet la reconstruction quasi-dense de la géométrie 3D à partir d'un ensemble d'images prises à partir de différents points de vue et l'estimation des paramètres correspondants de la caméra tout en se basant sur l'approche SFM. En fait, de grands efforts ont été déployés pour améliorer MVS, et des algorithmes efficaces ont été proposés [115]–[117]. Sur la base des méthodes de Goesele et Lhuillier [115], [116], la méthode MVS qui est assez précise (PMVS) a été développée par Furukawa et al en se basant sur les patches [29]. Cette méthode part d'un ensemble fragmenté de points-clés appariés et les développe de manière répétée jusqu'à ce que des contraintes de visibilité soient invoquées pour filtrer les faux appariements.

Dans la section suivante, nous allons présenter les étapes de l'approche proposée. Ensuite, nous allons discuter de l'efficacité de la méthode proposée à travers une étude comparative. Enfin, une conclusion sera donnée.

3.2 Etapes de la segmentation significative via la primitive couleur

Dans ce chapitre, nous présentons une méthode dont l'objectif est la segmentation du maillage en régions en adoptant une nouvelle stratégie complètement différente des approches existantes. En effet, notre méthode prend en compte les différentes étapes de la reconstruction 3D, en partant de l'étape d'autocalibrage de la caméra, jusqu'à la phase du maillage. A l'inverse des approches basées sur les caractéristiques géométriques du maillage, la méthode proposée repose sur les informations fournies par les images mises en jeu lors de la reconstruction 3D de l'objet.

Le principe de la méthode proposée est d'attribuer une couleur, à partir des images, à chaque triangle du maillage tout en exploitant le lien entre les sommets du maillage 3D et les points d'intérêt des images mises en jeu. Ensuite, on effectue une classification de l'ensemble de ces couleurs. Enfin, une segmentation du maillage est effectuée conformément à cette classification. La figure 3.1 décrit les principales étapes de notre méthode qui sont :

- ✓ Reconstruction 3D à partir d'images.
- ✓ Attribuer une couleur au maillage.
- ✓ Classification des facettes de couleur.

✓ Segmentation du maillage.

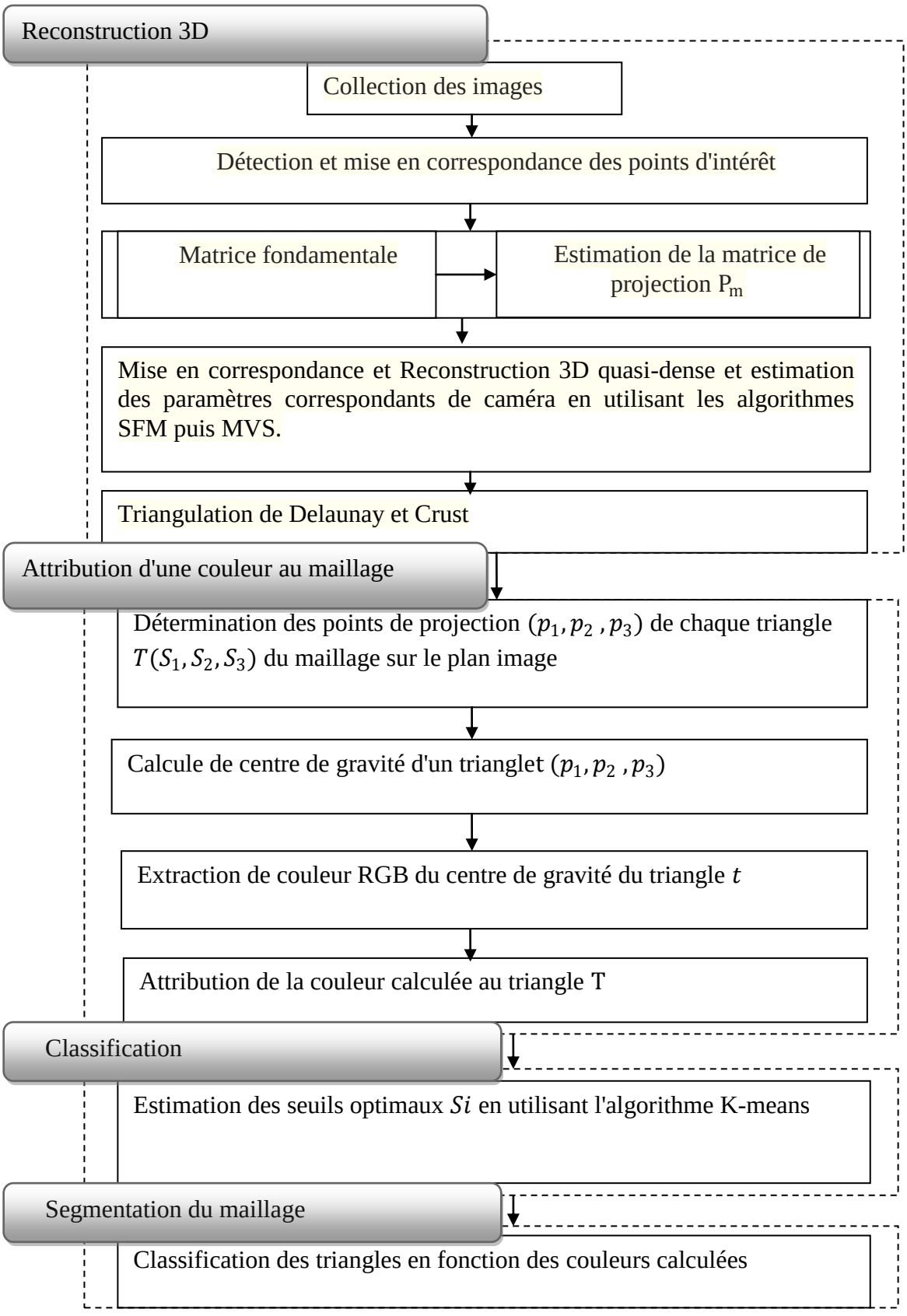


Figure 3.1 − Schéma descriptif de l'approche proposée.

3.2.1 Reconstruction 3D

La mise en correspondance des images constitue la première étape du processus de reconstruction dans lequel les points d'intérêt sont extraits en utilisant un descripteur local (par exemple, SIFT [18]). A partir de cette première étape nous pouvons reconstruire un nuage de points de la scène en trois dimensions et trouver la géométrie de celle-ci, tout en déterminant la correspondance entre les coordonnées des points 3D de la scène et leur projection sur les plans de la caméra. En effet, Considérant le problème général de la reconstruction de scène 3D avec W matrices de projection $P_1, P_2, \dots, P_W$ et W caméras $C_1, C_2, \dots, C_W$. Dans un premier temps, la matrice fondamentale de chaque image est estimée à partir de huit correspondances à l'aide de l'algorithme RANSAC [11]. L'estimation de la matrice fondamentale des images permet de récupérer les paramètres intrinsèques et extrinsèques de la caméra utilisée [29], [42], [118]. Enfin, la relation entre chaque point S_i de la scène et les plans de la caméra est définie par la géométrie épipolaire, cette relation entre les coordonnées spatiales et les coordonnées du plan de l'image est donnée par l'équation(3.1).

$$p_{mi} = P_m S_i \text{ avec } P_m = K_m [R|t] \tag{3.1}$$

Où p_{mi} est le point projeté dans la plan caméra C_m. P_m est la matrice de projection de la caméra C_m. t, R et K_m indiquent, respectivement, la translation et la rotation du centre de coordonnées de la caméra par rapport au système de coordonnées cartésiennes et la matrice intrinsèque de la caméra, sachant que la matrice intrinsèque est calculée une fois pour toutes dans le cas d'une seul caméra.

Enfin, après avoir obtenu les coordonnées des points 3D de la scène, la méthode Delaunay [119] et la méthode Crust [120] permettent de mieux interpoler le maillage.

3.2.2 Attribution des couleurs au maillage

Notre connaissance approfondie des points 3D de la scène et leur projection sur les plans de caméra fournis au cours de l'étape de reconstruction 3D, nous permet d'attribuer une couleur correspondante à partir des images mises en jeu à chaque triangle T_i du maillage. En effet, nous procédons d'abord par déterminer les points $(p_{mi1}, p_{mi2}, p_{mi3})$ représentant la projection des trois sommets (S_{i1}, S_{i2}, S_{i3}) du triangle T_i sur l'image I_m contenant à la fois ces points (Figure 3.2) sur la base de l'équation(3.1), puis par récupérer le centre de gravité de la couleur $Cg(t_i)$ du triangle $t_i(p_{mi1}, p_{mi2}, p_{mi3})$, qui sera attribué au triangle $T_i(S_{i1}, S_{i2}, S_{i3})$. Ce

processus est répété pour chaque triangle. Il convient de noter que, dans le cas où aucune image ne contient à la fois les trois points de projection, on considère alors que la couleur moyenne associée aux sommets du triangle t_i est la couleur du triangle T_i.

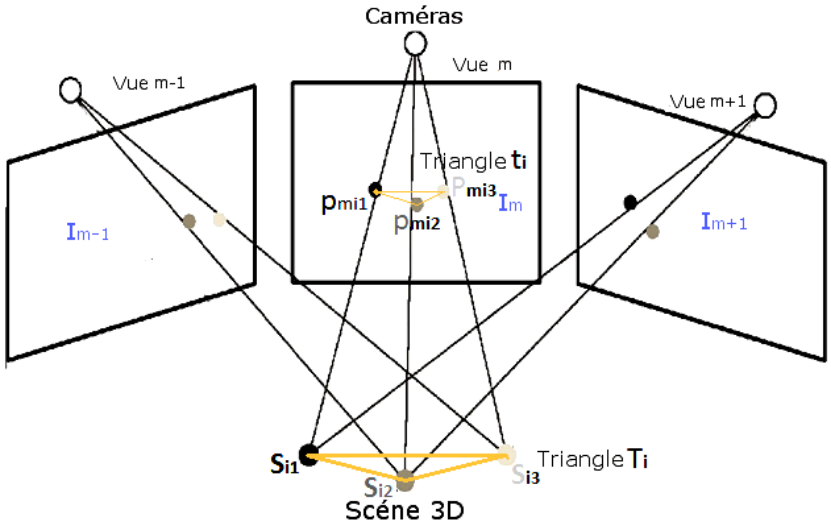


Figure 3.2 − Illustration du triangle Ti du maillage et de la projection de leurs sommets (S_{i1}, S_{i2}, S_{i3}) sur différents plans d'images.

3.2.3 Classification des couleurs RVB des facettes

Dans cette étape, nous procédons à la classification des couleurs déterminées à l'étape précédente, en construisant d'abord un histogramme contenant tous les couleurs RGB des triangles T_i, puis on effectue le seuillage d'histogramme des K classes déterminées à partir de $K-1$ seuils s_i. Celles-ci peuvent être définies par le maximum de l'histogramme h (j) sur chaque intervalle, où j représente la valeur de la couleur.

$$h(s_i) = \max h(j), \qquad \text{for } j \in k_i \tag{3.2}$$

Où k_i indique la $i-$*éme* classe.

Le problème du seuil consiste à déterminer les seuils optimaux qui séparent les différentes classes de l'histogramme. Pour cette raison, nous avons utilisé k-means [121] qui est considéré comme l'une des méthodes de clustering de haute performance. De plus, la version utilisée de k-means [121] présentait un algorithme permettant d'effectuer un regroupement en K-means de nature adaptative et qui ne dépend pas de la sélection de ksites pour initialiser les classes.

3.2.4 La segmentation du maillage

Après avoir calculé l'histogramme et estimé les seuils si, les triangles sont classés selon leur couleur déjà déterminée en fonction des informations fournies par l'histogramme. Enfin, nous regroupons les triangles adjacents appartenant à la même classe, formant ainsi l'un des segments.

3.2.5 Algorithme de la méthode proposée

L'algorithme, ci-dessous, décrit les étapes de la méthode proposée :

Prétraitement : la reconstruction 3D se compose de deux parties. Le but de la première partie est la mise en correspondance des images à partir desquelles les points d'intérêt sont extraits en utilisant un descripteur local (par exemple, SIFT [18]). Ensuite, la méthode FICC est utilisée pour rejeter les paires aberrantes. La deuxième partie consiste à une reconstruction quasi-dense de la géométrie 3D à partir d'un ensemble d'images prises à partir de différents points de vue en se basant sur l'approche MVS [29].

Attribution des couleurs : l'étape de prétraitement permet de déterminer pour chaque point 3D de la scène ses projections sur un ensemble des plans de caméra. Cette connaissance approfondie permet d'attribuer une couleur correspondante à chaque triangle du maillage à partir des images mises en jeu.

Classification : classification des couleurs RVB des facettes en construisant un histogramme. La méthode k-means [121] est utilisée pour la classification.

Grouper les triangles adjacents : les triangles appartenant à la même classe constituent un segment du maillage.

Entrée : Un modèle M et son ensemble S de sommets 3D et ses N triangles T_j.
Les ensembles Aij de points d'intérêt entre les images avec une connaissance préalable des différentes matrices de projections P_m. Les K centres initiaux.
Sortie : Parties sémantiques $\{Mi\}$ de la décomposition de M
Début

pour $j =$ de 1 à N **faire**

- Déterminer $\mathrm{t_j}(\mathrm{p_{mj\,1}}, \mathrm{p_{mj\,2}}, \mathrm{p_{mj\,3}})$ la projection du $\mathrm{T_j}(\mathrm{S_{j1}},\ \mathrm{S_{j2}},\ \mathrm{S_{j3}})$// en utilisant $eq(1)$.

si $\exists t_j$ **alors**

- $Col(T_j) \leftarrow Col\left(Cg(t_j)\right)$. //$Col\left(Cg(t_j)\right)$: couleur du centre de gravité du triangle t_j

sinon

- $Col(T_j) \leftarrow \begin{pmatrix} \frac{\sum_{i=1}^{3} R(p_{ji})}{3} \\ \frac{\sum_{i=1}^{3} G(p_{ji})}{3} \\ \frac{\sum_{i=1}^{3} B(p_{ji})}{3} \end{pmatrix}$ // couleur RGB moyenne de $(\mathrm{p_{j1}}, \mathrm{p_{j2}}, \mathrm{p_{j3}})$.

fin

fin pour

- Construire l'histogramme 3D h contenant l'ensemble ($Col(T_j)$) des couleurs récupérées
- Seuillage de l'histogramme h en K classes Eq. (2) // utilisant la méthode $\mathrm{k-means}$ [121]

Classifier les N triangles T_j en fonction de leur couleur.

- Grouper les triangles adjacents appartenant à la même classe pour déterminer l'un des segments.

Fin

3.3 Résultats expérimentaux et étude comparative

Les images stéréoscopiques utilisées dans cette section sont prises par une caméra qui tourne autour de l'objet ou de la scène avec des angles égaux afin de reconstruire l'objet.

Tout d'abord, on présente une étude comparative entre l'approche proposée et cinq autres approches, à savoir (Shape Diameter Function) SDF [4], Core Extraction [1], Fitting Primitives [68], Random Walks [73] et Randomized Cuts [82]. Les approches sont choisies en raison de leur haute performance. Pour la méthode SDF, nous suivons les conseils des auteurs, en fixant le nombre de partitions candidates à 5.

Notre approche est implémentée en Java sous Eclipse (processeur Core i5, 2,2 GHz, 8 Go de RAM). La bibliothèque JAMA a été utilisée pour les calculs et l'API Java 3D pour la visualisation 3D. Pour les codes source des cinq algorithmes, ils sont disponibles dans «Princeton Mesh Segmentation Benchmark» [84]. Il est important de noter que le nombre de segments est estimé automatiquement par les deux algorithmes Core Extraction et SDF, tandis que l'utilisateur doit définir le nombre de segments comme paramètre d'entrée pour les trois autres algorithmes. Cela exige donc une détermination du nombre de segments pour ces trois méthodes. Pour remédier ce problème, les algorithmes SDF et Core Extraction sont exécutés, puis sur la base des deux nombres de segments automatiquement déterminés par les deux

algorithmes, nous affectons leur moyenne aux trois autres algorithmes comme paramètre d'entrée. Ce choix met les cinq algorithmes sur un pied d'égalité.

Pour tester et valider la robustesse de la méthode proposée, nous utilisons quatre séquences d'images représentant des scènes complexes de différentes natures. La première séquence d'images concernant « Espadrille » de résolution 3264 × 2448. La seconde est une séquence d'images de résolution 2048 × 2048 d'un objet complexe« Mario ». La troisième séquence représente« Bouquet de Fleurs »de résolution 4096 × 4096. La quatrième séquence d'images de résolution 2560 × 1536 représente une « feuille ».

Les trois premières séquences contiennent des objets 3D avec des caractéristiques géométriques complexes ce qui représente un défi épineux pour les différentes méthodes de segmentation de maillage 3D. Par contre, la dernière séquence, illustre un objet « Feuille » d'une forme presque plate, ce qui représente lui aussi un défi difficile dû à sa grande simplicité géométrique. Les images sont prises à l'aide d'une caméra qui effectue des déplacements autour de l'objet ou de la scène de douze degrés. Une analyse du résultat de chaque séquence est d'abord présentée, puis nous interprétons les résultats. La figure 3.3 illustre quatre images représentant les quatre séquences.

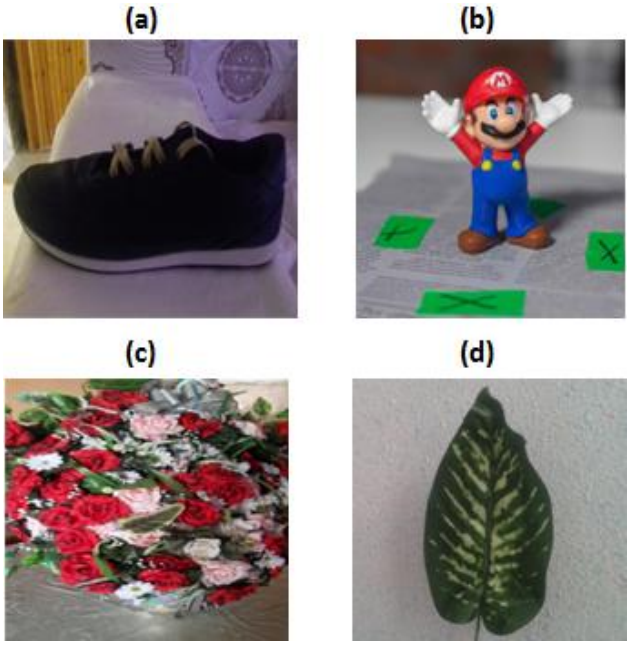


Figure 3.3 − (a) Image de la séquence «Espadrille ». (b) Une image de la séquence « Mario ». (c) Une image de la séquence « Bouquet de Fleurs ». (d) Une image de la séquence « Feuille ».

L'étape de la reconstruction 3D consiste à trouver les différents points de correspondances entre les images consécutives mises en jeu. Pour éliminer les paires aberrantes nous appliquons la méthode FICC (avec rejection des 20% des paires les moins

pertinentes à la quatrième étape). Les points sélectionnés constituent les données d'entrées de la méthode SFM [42] utilisée par la suite pour l'estimation des paramètres correspondants de la caméra. Après l'estimation des différentes matrices de projection et la récupération des points d'intérêt entre les images consécutives, la méthode (PMVS) de furakuwa et al. [29] a été appliquée pour augmenter le nombre de points reconstruits et obtenir une reconstruction 3D plus dense. La reconstruction 3D a été implémentée en utilisant JAMA (Java) et Open-CV.

Les résultats d'exécution de cet algorithme concernant les différentes séquences d'images sont présentées dans le tableau 3.1.

Le nombre de points 3D pour chaque séquence est illustré dans le tableau 3.1. Ce dernier est principalement lié à la résolution et au nombre d'images mises en jeu.

Les matrices de projection calculées dans la reconstruction 3D permettent de déterminer, pour chaque point 3D, les différents points de projection sur les différentes images, ce qui permet de déterminer une couleur pour chaque triangle T du maillage.

Résolution	Nombre de points reconstruits (3D Dense)
Séquence1 3264 x 2448	16072
Séquence2 2048 x 2048	26204
Séquence 3 4096 x 4096	166239
Séquence 4 2560 x 1536	14122

Tableau 3.1 − Résultats de la reconstruction 3D dense de différentes séquences.

3.3.1 Critères d'évaluation

Les critères d'évaluation des méthodes de segmentation de maillage sont spécialement conçus pour les méthodes basées sur la géométrie, et comme nous adoptons une approche non géométrique basée sur la classification de la couleur RVB des images mises en jeu en segmentant le maillage 3D en sous-maillages visuellement significatifs. On se base donc sur l'observation visuelle du maillage pour arriver à une évaluation précise et significative. Pour que ce critère de comparaison soit efficace, les objets les plus significatifs sont déterminés

pour chaque scène et ces méthodes sont évaluées en fonction de leur capacité à isoler ces objets. Ce critère représente donc le taux de réussite en pourcentage de chaque méthode.

Ce critère d'évaluation peut être exprimé comme suit :

$$\mathrm{Cr} = \frac{\text{Nombredesobjets correctementsegmentés}}{\text{Nombretotale desobjets}} * 100\% \qquad (3.3)$$

3.3.2 Séquence 1

Cette première expérience représente une scène contenant trois objets significatifs, ce qui pose un défi aux six méthodes pour segmenter un objet «Espadrille ». Une séquence de vingt images est utilisée pour représenter cette scène. La figure 3.4 montre une image de cette séquence avec les objets les plus significatifs, le résultat de la reconstruction 3D et l'histogramme correspondant.

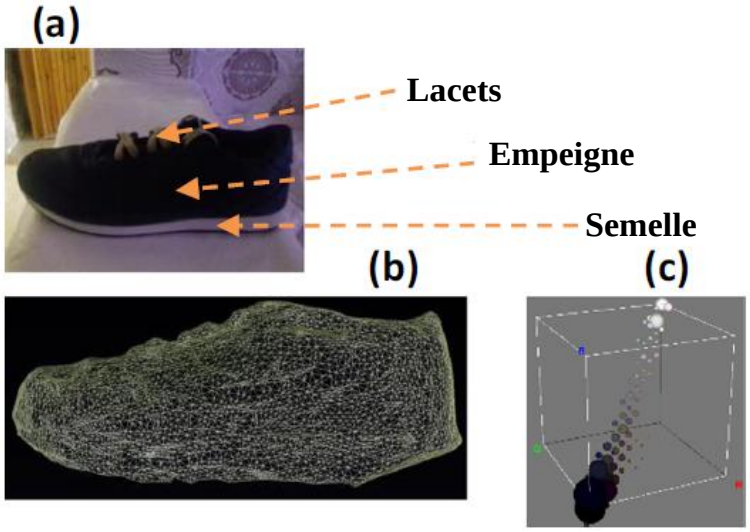


Figure 3.4 − (a) Image de la séquence avec des objets significatifs. (b) Vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 78 couleurs représentatives

Après la reconstruction 3D de l'objet et après le calcul et la classification d'histogramme, en appliquant la méthode k-means à la première séquence, nous avons obtenu deux seuils de couleur RVB ((51, 25, 26), (230, 239, 241)) qui ont été utilisés pour segmenter le maillage. La figure 3.5 montre le résultat de la segmentation des six algorithmes et le tableau 3.2 donne la comparaison entre les six méthodes.

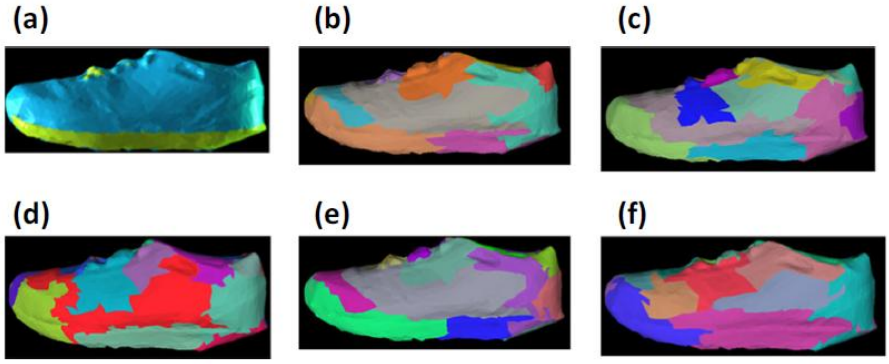


Figure 3.5 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 13 composants. (c) Le résultat obtenu par la méthode SDF avec 21 composants. (d) Le résultat obtenu par Fitting Primitives avec 17 composants. (e) Le résultat obtenu par Random Walks avec 17 composants. f) Le résultat obtenu par Randomized Cuts à 17 composants.

Objet	Lacets	Semelle	Empeigne	Cr
Méthode proposée	0	1	1	**66.6%**
Core Extra	0	0	0	**0%**
SDF	0	0	0	**0%**
Fit Prim	0	0	0	**0%**
Rand Walk	0	0	0	**0%**
Rand Cuts	0	0	0	**0%**

Tableau 3.2 − Comparaison entre les six méthodes selon leur capacité à détecter les objets « Espadrille ».

Comme il est montré dans le tableau 3.2, notre méthode permet de distinguer deux objets avec un taux de précision de 66,6%. Cependant, l'utilisation des outils mathématiques comme la squelettisation, l'angle dièdre et la courbure ne permet pas aux autres méthodes de segmenter l'objet de manière significative. En effet, la segmentation de l'objet s'avère très difficile pour les méthodes géométriques en raison de la nature topologique de «Espadrille» dans laquelle la séparation géométrique de la «Semelle» et de la «Empeigne» représente un problème épineux. Cependant, l'utilisation des informations de couleur résout largement le problème.

3.3.3 Séquence 2

Dans cette expérience, une séquence de dix-sept images d'une scène complexe contenant plusieurs objets significatifs est utilisée. La figure 3.6 montre deux images de cette

séquence, les objets les plus significatifs, le résultat de la reconstruction 3D et l'histogramme correspondant.

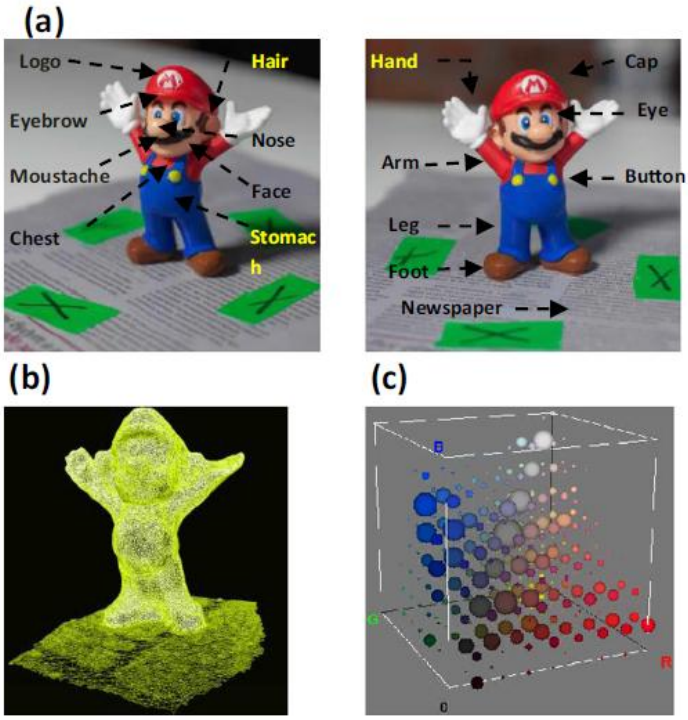


Figure 3.6 − (a) Deux images de la séquence comportant les parties les plus sémantiques (la couleur jaune permet une bonne visualisation). (b) vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 312 couleurs représentatives.

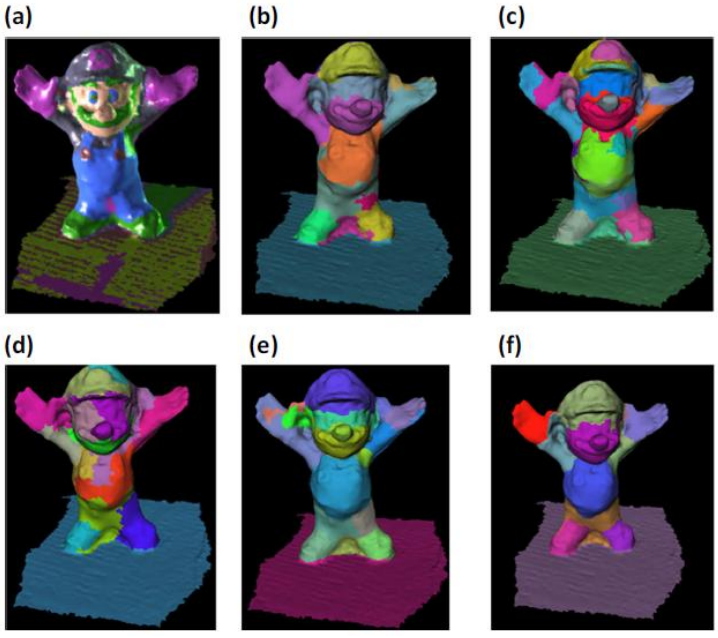


Figure 3.7 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 14 composants. (c) Le résultat obtenu par la méthode SDF avec 20 composants. (d) Le meilleur résultat obtenu par Fitting Primitives avec 17 composants. (e) Le meilleur résultat obtenu par Random Walks avec 17 composants. (f) Le meilleur résultat obtenu par Randomized Cuts à 17 composants.

Après la reconstruction 3D de l'objet et après le calcul et la classification de l'histogramme, en appliquant la méthode k-means, nous avons obtenu six seuils de couleur RVB qui ont été utilisés pour segmenter le maillage. La figure 3.7, ci-dessus, présente le résultat de la segmentation obtenue. Le résultat de la comparaison entre les six méthodes est détaillé dans le tableau 3.3.

Objet	Méthode Proposée	Core Extra	SDF	Fit Prim	Rand Walks	Rand Cuts
Logo	1	0	1	0	0	0
Sourcil	1	0	0	0	0	0
Moustache	1	0	0	0	0	0
Poitrine	1	0	1	0	0	0
Cheveux	1	0	0	0	0	0
Nez	0	0	1	0	1	0
Visage	1	0	1	1	1	1
Estomac	0	0	0	0	1	1
Main	1	0	0	0	0	1
Bras	0	0	1	0	0	1
Jambe	0	1	1	0	0	1
Pied	1	1	1	0	1	1
Journal	0	1	1	1	1	1
Casquette	1	1	0	0	0	0
Œil	1	0	0	0	0	0
Bouton	1	0	0	0	0	0
Total	**11**	**4**	**8**	**2**	**5**	**7**
Cr	**68.7%**	**25%**	**50%**	**12.5%**	**31.2%**	**43.7%**

Tableau 3.3 − Comparaison entre les six méthodes selon leur capacité à détecter 16 objets.

D’après le tableau 3.3, nous observons que la méthode proposée est plus efficace et plus précise que les cinq autres méthodes. En effet, notre méthode est capable de distinguer treize objets avec un taux de précision de 68,7%. En fait, plusieurs composants de l'objet «Mario» peuvent être clairement distinguées, en raison de l’uniformité de couleur des pixels associés sur les images mises en jeu. Cependant, le meilleur résultat est obtenu par la méthode SDF avec une précision de 50%. En effet, celle-ci détermine un squelette médian adéquat pour une extraction pertinente des composants de l'objet, en exploitant la ressemblance relative entre la topologie de l'objet et un corps humain. De plus, la méthode de Fitting Primitives a le score le plus faible avec une précision de 12,5%. Ce résultat insatisfaisant est dû à la nature topologique de l'objet qui ne permet pas à cette méthode de déterminer des composants pertinents en se basant sur les primitives ajustées.

3.3.4 Séquence 3

Cette expérience représente un véritable défi pour les six méthodes en essayant de segmenter un bouquet de fleurs dont la géométrie est très complexe. Une séquence de dix-neuf images est utilisée pour représenter cette scène. La figure 3.8 montre deux images de cette séquence avec les parties les plus sémantiques, le résultat de la reconstruction 3D et l'histogramme correspondant.

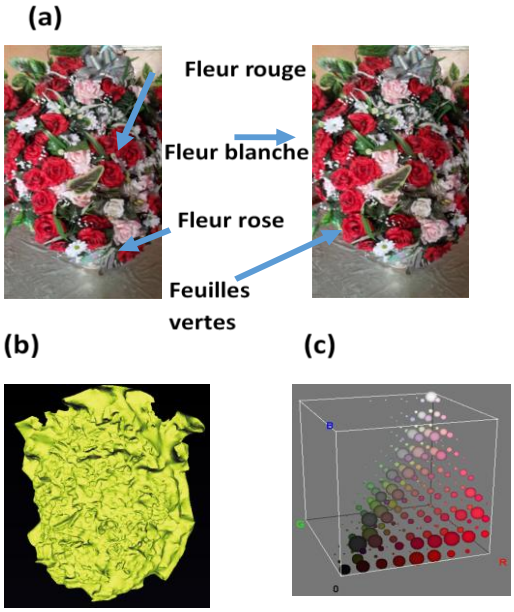


Figure 3.8 − (a) Deux images de la séquence avec des objets significatifs. (b) vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 289 couleurs représentatives.

Après la reconstruction 3D de l'objet et après le calcul et la classification de l'histogramme, nous avons obtenu trois seuils de couleur RVB en appliquant la méthode k-means ((183, 18, 43), (73, 101, 41), (210, 186, 178)). Ces trois seuils sont utilisés pour segmenter le maillage. La figure 3.9 et le tableau 3.4,ci-dessous, montrent, respectivement, le résultat de la segmentation des six algorithmes et le résultat de la comparaison entre les six méthodes.

Selon le tableau 3.4 ci-dessous, nous notons à nouveau que les cinq méthodes sont incapables de surmonter ce test, cet échec est dû en grande partie à la grande complexité géométrique de l'objet. En effet, la méthode SDF ne peut pas déterminer un squelette médian adéquat pour extraire des composants significatifs de l'objet. La méthode Core extraction est incapable d'extraire les composants essentiels depuis un objet avec courbures et angles dièdres qui varient brusquement et continuellement. La méthode Fitting primitive est également incapable de déterminer des composants pertinents en se basant sur les primitives ajustées, en

raison de la complexité géométrique de l'objet. En revanche, bien que les fleurs blanches et les fleurs roses soient considérées comme étant la même en raison du résultat du regroupement, la méthode proposée présente une précision considérable qui vaut 50%. En fait, la méthode proposée par une opération de classification peut distinguer les fleurs vertes et les fleurs rouges et elle est également incapable de différencier les fleurs blanches de celles roses.

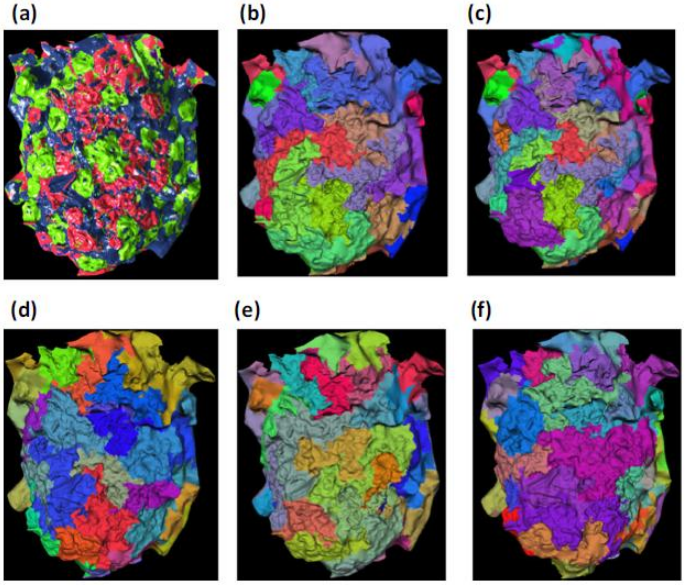


Figure 3.9 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 23 composants. (c) Le résultat obtenu par la méthode SDF avec 48 composants. (d) Le résultat obtenu par fitting primitives avec 36 composants. (e) Le résultat obtenu par Random Walks avec 36 composants. (f) Le résultat obtenu par Randomized Cuts avec 36 composants.

Objet	**Méthode Proposée**	**Core Extra**	**SDF**	**Fit Prim**	**Rand Walks**	**Rand Cuts**
Fleur rouge	1	0	0	0	0	0
Fleur blanche	0	0	0	0	0	0
Fleur rose	0	0	0	0	0	0
Feuilles vertes	1	0	0	0	0	0
Total	**2**	**0**	**0**	**0**	**0**	**0**
Cr	**50%**	**0%**	**0%**	**0%**	**0%**	**0%**

Tableau 3.4 − Comparaison entre les six méthodes selon leur capacité à détecter 4 objets.

3.3.5 Séquence 4

Au cours de cette dernière expérience, nous essayons de tester les performances des six méthodes en utilisant un objet qui présente un défi en raison de sa surface (presque) plate.

Une séquence de dix-sept images est utilisée pour représenter cette scène. La figure 3.10 montre une image de cette séquence avec les objets les plus significatifs, le résultat de la reconstruction 3D et l'histogramme correspondant.

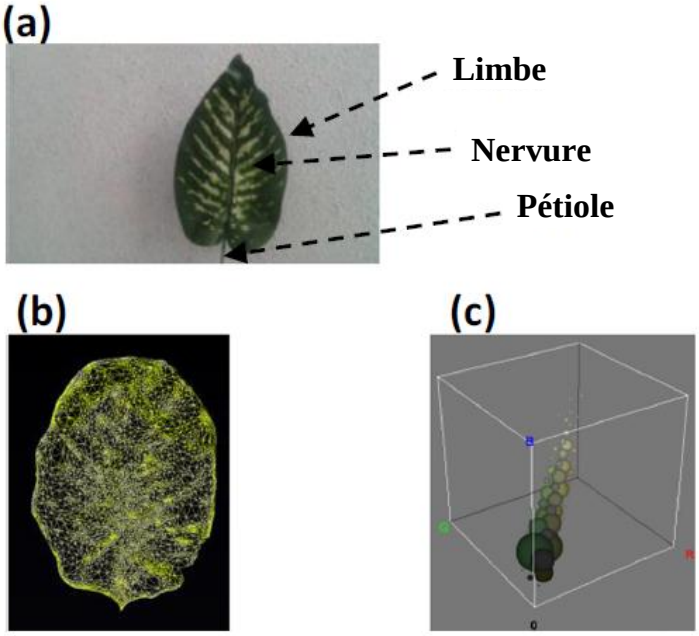


Figure 3.10 − (a) Image de la séquence avec des parties sémantiques. (b) Vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 90 couleurs représentatives.

Après la reconstruction 3D de l'objet et après le calcul et la classification de l'histogramme, la méthode k-means nous a permis d'obtenir deux seuils de couleur RVB ((40,62,38). (172,168,128)). Ces seuils ont servi à segmenter le maillage.

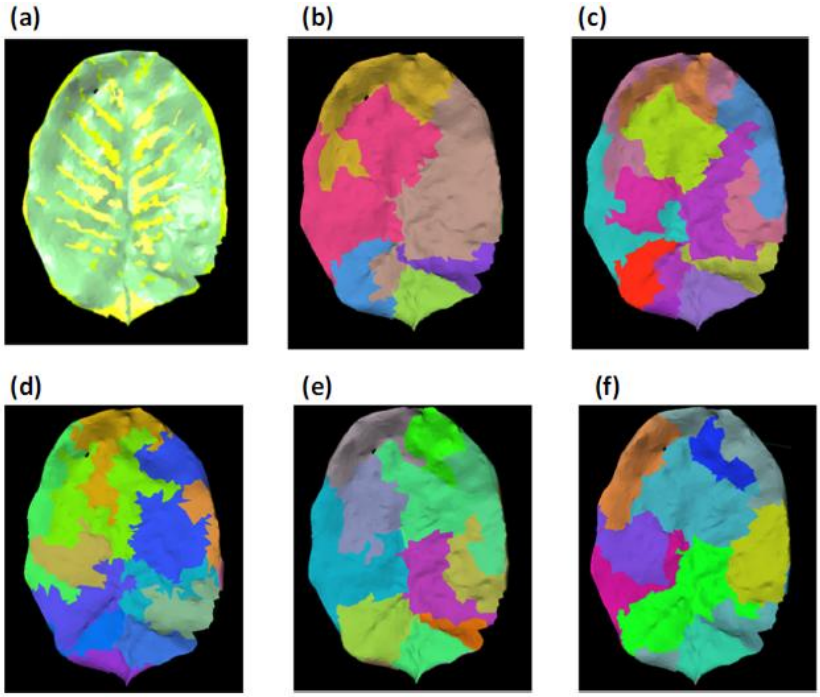


Figure 3.11 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 8 composants. (c) Le résultat obtenu par la méthode SDF avec 16 composants. (d) Le résultat obtenu par fitting primitives avec 12 composants. (e) Le résultat obtenu par Random Walks avec 12 composants. f) Le résultat obtenu par les découpes aléatoires à 12 composants.

La figure 3.11 et le tableau 3.5 ci-dessous montrent, respectivement, le résultat de la segmentation obtenue et le résultat de la comparaison entre les six méthodes.

Le tableau 3.5 ci-dessous montre que la méthode proposée est plus efficace et plus précise que les cinq autres méthodes. En effet, elle permet de distinguer deux objets avec un taux de précision de 66,6%,en se basant sur le résultat de la classification. Cependant, la nature topologique de l'objet qui consiste à une surface presque plane, a diminué l'efficacité des techniques basées sur des critères géométriques, comme la squelettisation, l'angle dièdre, la courbure et la distance géodésique, pour guider la segmentation de maillage. En conséquence, ces méthodes ne peuvent pas segmenter ces objets de manière significative.

Objet	Méthode Proposée	Core Extra	SDF	Fit Prim	Rand Walk	Rand Cuts
Limbe	1	0	0	0	0	0
Nervure	1	0	0	0	0	0
Pétiole	0	0	0	0	0	0
Total	2	0	0	0	0	0
Cr	66.6%	0%	0%	0%	0%	0%

Tableau 3.5 − Comparaison entre les six méthodes selon leur capacité à détecter 3 objets.

3.3.6 Discussion

En analysant les résultats des trois séquences, on constate que la méthode proposée est caractérisée par sa précision par rapport aux cinq autres méthodes. Cette précision est évidente dans le cas de scènes présentant des caractéristiques géométriques très complexes (Figure 3.6 et 3.8) ou très simples (Figure 3.10). En effet, l'obtention des segments significatifs d'une scène est plutôt difficile sur la base de la géométrie, ce qui explique les résultats obtenus par les autres méthodes.

L'efficacité et la performance des méthodes de segmentation dépendent non seulement des qualités de segmentation de maillage, mais également de la vitesse d'exécution. Pour cela,

le temps d'exécution de chaque algorithme dans chaque séquence est calculé. Le tableau 3.6 illustre les résultats obtenus.

Séquence	Méthode Proposée	Core Extra	SDF	Fit Prim	Rand Walk	Rand Cuts
Espadrille	0.32	1.29	0.72	0.38	0.25	4.86
Mario	0.74	3.28	1.60	0.86	0.67	12.98
Bunch of flowers	2.06	14.23	7.62	3.82	2.47	43.60
Leaf	0.26	0.93	0.55	0.37	0.28	4.22

Tableau 3.6 − Comparaison du temps d'exécution moyen (plus de cinquante exécutions) des six algorithmes de segmentation pour chaque séquence (le temps de calcul est exprimé en secondes sur un PC de 8,8 Ghz).

En regardant les résultats présentés dans le tableau 3.6, nous trouvons que la méthode proposée nécessite un temps de calcul nettement inférieur à celui des autres méthodes à l'exception de la méthode Random Walks. Ce résultat n'est pas surprenant puisqu'elle guide la segmentation, d'une part, par la détermination d'une couleur pour chaque triangle du maillage et, d'autre part, par la classification de l'histogramme. Pour ce dernier, le temps d'exécution est compté en quelques millisecondes et cela n'a donc que peu d'influence sur le temps d'exécution.

En ce qui concerne la détermination des couleurs, cela dépend des points de projection des différents sommets du maillage sur les plans d'images concernés, ces points sont déjà calculés lors de la reconstruction 3D, ce qui réduit le temps de calcul à un simple calcul des centres de gravité. Par contre, les autres méthodes nécessitent un temps relativement important pour le calcul. En effet, l'algorithme Core-Extraction [1] demande beaucoup de temps pour exécuter trois étapes de l'algorithme à savoir: la représentation invariante de la pose, la détection des points caractéristiques et l'extraction de l'élément essentiel. De plus, il est nécessaire d'ajouter le temps nécessaire pour compléter la segmentation et le raffinement du maillage, comme décrit dans l'algorithme [1].

Concernant la méthode SDF, la détermination de l'axe médian et de la boule médiane d'une part, et le calcul de SDF pour chaque face d'une autre part, nécessite un temps considérable particulièrement lorsque l'on tient compte du fait qu'il s'agit d'une méthode hiérarchique. La méthode fitting primitives est plus rapide que les deux dernières méthodes grâce à l'utilisation d'un algorithme de clustering simple basé sur le choix de la paire la mieux ajustée pour la fusion. La méthode de Randomized Cuts nécessite plus de temps de calcul, en

raison de la procédure hiérarchique appliquée et du calcul d'un ensemble de coupes aléatoires pour chaque segment et chaque division. Par contre, la méthode Random Walks donne un bon résultat grâce à une première étape de sur-segmentation guidée par un double graphique du maillage.

Enfin, aucun algorithme de segmentation n'est meilleur que les autres pour tous les types d'objets. Cependant, notre méthode, comparée aux méthodes de référence, permet de segmenter un grand nombre de maillages 3D en parties sémantiques basées sur les couleurs. En effet, à l'instar de la segmentation humaine, il est souvent plus facile de segmenter des objets 3D en fonction de la couleur que par les caractéristiques géométriques de ces objets.

De plus, l'extraction des couleurs des images et leur classification ne nécessitent pas un temps de calcul plus élevé, même pour des objets assez complexes. Ces avantages expliquent en grande partie la performance et l'efficacité de la méthode proposée.

3.4 Conclusion

Dans ce chapitre, nous avons décrit une approche dont l'objectif est de résoudre le problème de segmentation de maillage en parties significatives (sémantiques), tout en exploitant la couleur des images mises en jeu. En fait, nous avons exploité les matrices de projections pour reconstruire l'objet 3D en affectant une couleur à chaque triangle du maillage. La classification de ces couleurs nous permet de segmenter le maillage. Les résultats expérimentaux obtenus, en se basant sur des séquences des images réelles, suggèrent l'efficacité de la méthode proposée à la segmentation significative des objets 3D. Cette efficacité est due à l'utilisation de l'information de la primitive couleur par rapport à la géométrie, notamment pour les objets à topologie complexe.

Il serait également intéressant de tester cette méthode de segmentation sur d'autres modèles 3D pour évaluer davantage sa performance et sa pertinence, tout en comparant avec celles d'autres méthodes de segmentation existantes

Conclusion générale

L’objectif visé par le présent travail, consiste à proposer une nouvelle approche de segmentation d’objets 3D en parties significatives qui sera utilisée dans plusieurs applications particulièrement l'étude de la similarité entre les formes dans le contexte de l’indexation partielle et de recherche par le contenu d’objets 3D.

Nous avons conçu et implémenté une méthode efficace et rapide qui effectue la segmentation d’un maillage en se basant sur la couleur des pixels des images stéréoscopiques contrairement à la majorité des méthodes de référence en segmentation de maillages qui s'appuient sur les caractéristiques géométriques des objets 3D. Cette technique est basée sur l'étape de la mise en correspondance stéréoscopique dont la qualité des appariements identifiés est d’une importance primordiale, d'où l'intérêt de nos deux contributions à ce stade de l’algorithme.

En effet, nous avons proposé deux algorithmes pour rejeter les paires aberrantes contenues dans un ensemble de paires extraites à partir des images stéréoscopiques. Pour atteindre un tel objectif, nous avons quantifié la chance de chaque paire d'être aberrante en lui affectant une valeur qui dépend du test d'appartenance aux cercles. La sélection des vrais appariements est ainsi possible grâce à une opération de classification qui se base sur les valeurs associées aux paires de points. Les deux algorithmes proposés démontrent une efficacité importante envers la contrainte de changement de point de vue et les transformations affines. A l'instar des autres approches de mise en correspondance, les méthodes proposées ont une grande importance pour divers applications dans le domaine de la vision par ordinateur, y compris celui de la segmentation de maillage qui a fait l'objet de la troisième contribution.

Le principe de l’algorithme de segmentation proposé consiste à établir un lien entre les couleurs des objets de la scène et le maillage en attribuant une couleur à chaque face du maillage. En fait, sur la base des matrices de projection, calculées lors de l'étape de la reconstruction 3D, nous pouvons construire et classifier l'histogramme. Cela nous permet de segmenter le maillage significativement à la fin du processus. La méthode proposée en se basant sur la primitive couleur, contrairement aux approches géométriques, démontre à travers une étude comparative son efficacité à segmenter les objets 3D, en particulier ceux ayant une grande complexité géométrique.

Notre méthode, comme les autres techniques, est pénalisée par quelques limitations. En effet, si la lumière est trop faible ou si la reconstruction cible est constituée de quasiment la même couleur sur sa surface, la qualité des résultats de la segmentation diminue. Une stratégie bien meilleure doit donc être élaborée à l'avenir :

- Il pourrait être intéressant d'intégrer un prétraitement permettant d'effectuer la mise en correspondance stéréoscopique active à images multiples. Ceci permettrait d'améliorer considérablement la qualité des appariements et par la suite le résultat de la segmentation du maillage.
- Améliorer les résultats dans les zones où la reconstruction cible a presque la même couleur sur sa surface en appliquant une méthode hybride permettant de combiner l'approche géométrique et l'approche chromatique.

Table des figures

Figure 1.1 − Illustration d'un enregistrement séquentiel. Les vues 1 à 7 sont enregistrées une par une en calculant les matrices essentielles E12, E23, etc. 26

Figure 1.2 − Principales étapes d'un pipeline SFM générique : détection des entités, mise en correspondance des entités, génération des pistes, structure de mouvement et ajustement de faisceau. 27

Figure 1.3 − Exemples de SFM: Modèle de la Colisée. 28

Figure 1.4 − Correspondance des images basée sur une connaissance à priori des paramètres de caméra. A gauche : Recherche 2D entre les pixels de différentes images. A droite :Recherche 1D entre les pixels [49]. 30

Figure 1.5 − Illustration de maillage quadrangulaire. A gauche :Les points rouges représentent les sommets. Au milieu : Les arêtes sont les traits rouges reliant les sommets. A droite : Le carré constitue une face. 31

Figure 1.6 − Illustration de maillage (en pointillé) et son graphe dual, [51]. 32

Figure 1.7 − Illustration de la propriété de Delaunay. A gauche : Delaunay. A droite : La propriété de Delaunay est non respectée. 33

Figure 1.8 − Triangulation de Delaunay DTk (gauche), Cavité Delaunay $\mathcal{C}$DTk, pk + 1 (au centre) et $\mathcal{B}$DTk, pk + 1 (droite). 34

Figure 1.9 − Illustration de relation entre le diagramme de Voronoï (en pointillé) et la triangulation de Delaunay (en noire), [56] 34

Figure 1.10 − Diagramme de Voronoï pour un ensemble de sites V = vkk = 1n. 35

Figure 1.11 − L'algorithme bidimensionnel. Les arêtes de pointillés représentent le diagramme de Voronoï d'un ensemble de points S échantillonné à partir d'une courbe. Les arêtes en noir représentent la triangulation de Delaunay de S ∪ V, [56]. 37

Figure 1.12 − Illustration des étapes principales de la méthode Marches aléatoires (figure extraite de [73]). En haut : Modèle d'entrée avec placement des germes et sur-segmentation. En bas de gauche à droite : Résultat après la fusion et résultat final après lissage des bords. 41

Figure 1.13 − L'ajustement de l'histogramme du SDF permet d'obtenir un partitionnement plus fin de l'objet en parties [4]. 42

Figure 1.14 − L'algorithme de partitionnement en deux étapes : a. partitionnement souple du maillage. b. seuillage simple. c. limites lisses. d. Le partitionnement hiérarchique est robuste, pose insensible et rapide, [4]. 43

Figure 1.15 − Illustration des principales étapes de l'algorithme : **a** Une transformation MDS et son enveloppe convexe **b** Points caractéristiques **c** Partie essentielle **d** Segmentation et affinement des coupes, [1]. 44

Figure 1.16 − A gauche, une segmentation sémantique. A droite, une segmentation géométrique [50]. 46

Figure 1.17 − Quelques modèles avec les abstractions correspondantes représentées par les cylindres ajustés. Pour le modèle du cheval, deux niveaux de résolution sont affichés avec le regroupement de visage correspondant, [91]. 47

Figure 1.18 − Segmentation du cheval avec des positions de germes différents : les boules représentent les positions des germes (figure extraite de [73]). 48

Figure 1.19 − La fonction de diamètre de forme locale (SDF) (a). Partitionnement hiérarchique (b) Squelettes robustes simples (c), [4]. 50

Figure 1.20 − Segmentation hiérarchique d'un aigle, d'un chien, d'un guépard et du dino-pet. Figure extraite de [1]. 51

Figure 1.21 − Échantillon de segmentations randomisé produites avec le regroupement hiérarchique (a) et la fonction de partition résultante (b), [82]. 51

Figure 1.22 − Exemple de segmentation et co-segmentation. (a) Segmentation d'une seule forme. b) Co-segmentation d'une famille de formes similaires [3]. 52

Figure 2.1 − Illustration de deux cas possibles de chaque paire de points parmi les correspondances putatives. 56

Figure 2.2 − Zones d'une scène 3D. 57

Figure 2.3 − Illustration d'un cône tronqué oblique représentant la zone virtuelle Zkij (k = l ou r) de la scène 3D. Cette zone est une extension du cercle ckij dans la scène 3D passant par les deux points 3D Pi et Pj, dont la petite base circulaire est le cercle ckij, et la hauteur dépend de la scène 3D. 57

Figure 2.4 − Illustration de la stratégie adoptée. A gauche : prm ∈ crij et plm ∉ clij. A droite : prm′ ∉ cri′j′ et plm′ ∈ cli′j′ 58

Figure 2.5 − Illustration de la deuxième stratégie non adoptée : prm ∈ crij et plm ∈ clij. 58

Figure 2.6 − La première hypothèse : prm, plm est un vrai appariement, avec prm ∈ crij et plm ∉ clij. 60

Figure 2.7 − La deuxième hypothèse : prm, plm est une donnée aberrante, avec prm ∈ crij et plm ∉ clij. 60

Figure 2.8 − Critère de la taille des cercles sous forme graphique : la taille du cercle cli′j′ est utilisée pour estimer la non-correspondance de la paire prm′, plm′ avec davantage de précision que le cercle clij. 61

Figure 2.9 − Illustration des deux catégories de points correspondants que l'on peut trouver dans les images stéréoscopiques. 62

Figure 2.10 − Illustration d'une paire aberrante sélectionnée lors de l'étape de prétraitement. 63

Figure 2.11 − Illustration de la méthode proposée. (**a)** une paire d'image d'entrée. (**b)** et (**c)** Une ligne-croix noire représente un sous-ensemble des correspondances putatives sélectionnées pour créer les cercles correspondants. (**d)** Quelques correspondances putatives avec une importante possibilité d'être aberrantes. 68

Figure 2.12 − Deux paires (prm, plm) et (prn , pln) qui satisfont au premier critère ICC avec le même cercle correspondant crij, clij : malgré les différences entre les deux paires, la méthode ICC considère que CTBO[m]= CTBO[n]. 68

Figure 2.13 − Illustration de l'utilisation d'un cercle correspondant crij, clij pour estimer la correspondance correcte entre n points correspondances putatives. Les paires qui ne vérifient pas le premier critère sont illustrées par les points rouges et que les autres points représentent ceux qui satisfont au premier critère. 69

Figure 2.14 − L'illustration représente le deuxième critère FICC : le couple de cercles en bleu crij, clij est utilisé pour estimer la correspondance correcte d'une paire prm, plm. Le cercle jaune qui a le même centre que clij et de rayon plus long d. d représente la valeur la plus élevée exprime la chance que prm, plm d'être une paire aberrante en utilisant un cercle jaune au lieu de clij. 69

Figure 2.15 − Illustration de l'amélioration obtenue par la nouvelle stratégie visant à réduire le nombre de correspondances putatives testées à chaque itération. 70

Figure 2.16 − Schéma descriptif de l'approche proposée. L'amélioration de la méthode FICC par rapport à ICC est illustrée par des parties colorées. 72

Figure 2.17 − Résultats expérimentaux sur trois paires d'images : Teddy, Tsukuba et Cônes. En haut : les résultats sur Teddy. En bas : les résultats sur Tsukuba et Cônes. Les carrés illustrent le critère établi par la méthode FICC concernant les paires suspectes d'être aberrantes. Pour la visibilité, un sous ensemble des correspondances est choisi au hasard pour l'illustration. 77

Figure 2.18 − Résultats expérimentaux du base de données Mikolajczyk [40]. Statistiques de Précision-rappel sans réduire les vrais appariements pour RSC, RANSAC, MLESAC, ICC avec 4ème étape et notre FICC avec 4ème étape. ICC et FICC ont les meilleures précisions globales. 79

Figure 2.19 − Résultats expérimentaux sur trois paires d'images : *Graffiti*, *Bikes* et *Trees*. En haut : résultats sur *Graffiti*. En bas : résultats sur *Bikes* et *Trees*. Les carrés illustrent le critère établir par la méthode FICC concernant les paires suspectes d'être aberrantes. Pour la visibilité, un sous ensemble des correspondances est choisi au hasard pour l'illustration. 81

Figure 3.1 − Schéma descriptif de l'approche proposée. 87

Figure 3.2 − Illustration du triangle Ti du maillage et de la projection de leurs sommets (Si1, Si2, Si3) sur différents plans d'images. 89

Figure 3.3 − (a) Image de la séquence «Espadrille ». (b) Une image de la séquence « Mario ». (c) Une image de la séquence « Bouquet de Fleurs ». (d) Une image de la séquence « Feuille ». 92

Figure 3.4 − (a) Image de la séquence avec des objets significatifs. (b) Vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 78 couleurs représentatives 94

Figure 3.5 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 13 composants. (c) Le résultat obtenu par la méthode SDF avec 21 composants. (d) Le résultat obtenu par Fitting Primitives avec 17 composants. (e) Le résultat obtenu par Random Walks avec 17 composants. f) Le résultat obtenu par Randomized Cuts à 17 composants. 95

Figure 3.6 − (a) Deux images de la séquence comportant les parties les plus sémantiques (la couleur jaune permet une bonne visualisation). (b) vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 312 couleurs représentatives. 96

Figure 3.7 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 14 composants. (c) Le résultat obtenu par la méthode SDF avec 20 composants. (d) Le meilleur résultat obtenu par Fitting Primitives avec 17 composants. (e) Le meilleur résultat obtenu par Random Walks avec 17 composants. (f) Le meilleur résultat obtenu par Randomized Cuts à 17 composants. 96

Figure 3.8 − (a) Deux images de la séquence avec des objets significatifs. (b) vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 289 couleurs représentatives..... 98

Figure 3.9 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 23 composants. (c) Le résultat obtenu par la méthode SDF avec 48 composants. (d) Le résultat obtenu par fitting primitives avec 36 composants. (e) Le résultat obtenu par Random Walks avec 36 composants. (f) Le résultat obtenu par Randomized Cuts avec 36 composants. 99

Figure 3.10 − (a) Image de la séquence avec des parties sémantiques. (b) Vue du modèle 3D obtenu. (c) Histogramme 3D du modèle 3D contenant 90 couleurs représentatives. 100

Figure 3.11 − (a) Résultats obtenus par notre algorithme. (b) Les résultats obtenus par la méthode Core Extraction avec 8 composants. (c) Le résultat obtenu par la méthode SDF avec 16 composants. (d) Le résultat obtenu par fitting primitives avec 12 composants. (e) Le résultat obtenu par Random Walks avec 12 composants. f) Le résultat obtenu par les découpes aléatoires à 12 composants. 101

Liste des tableaux

Tableau 2.1 − Pourcentage des vrais appariements est modifié en ajoutant des paires aberrantes. 76

Tableau 2.2 − Comparaison des performances sur les paires d'images de Teddy Tsukuba et Cônes. Les paires dans le tableau sont Précision-Rappel(%)..... 77

Tableau 2.3 – Le pourcentage des vrais appariements est modifié en réduisant le nombre des vrais appariements..... 80

Tableau 2.4 −Tableau montrant la comparaison des performances sur les paires d'images de *Graffiti*, *Bikes* et *Trees*. Les paires dans le tableau sont Précision-Rappel(%). 82

Tableau 2.5 − Tableau montrant le temps de calcul des méthodes FICC et ICC (en Sec). 83

Tableau 3.1 − Résultats de la reconstruction 3D dense de différentes séquences. 93

Tableau 3.2 − Comparaison entre les six méthodes selon leur capacité à détecter les objets « Espadrille ». 95

Tableau 3.3 − Comparaison entre les six méthodes selon leur capacité à détecter 16 objets. 97

Tableau 3.4 − Comparaison entre les six méthodes selon leur capacité à détecter 4 objets. 99

Tableau 3.5 − Comparaison entre les six méthodes selon leur capacité à détecter 3 objets. 101

Tableau 3.6 − Comparaison du temps d'exécution moyen (plus de cinquante exécutions) des six algorithmes de segmentation pour chaque séquence (le temps de calcul est exprimé en secondes sur un PC de 8,8 Ghz). 102

Liste des symboles

η_0	Niveau de confiance
ε	Ratio des vrais appariements
ε^n	Probabilité de sélectionner un échantillon non contaminé de n appariements
C	Fonction de coût
γ	Paramètre de mélange des distributions gaussiennes et uniformes
υ	Constante
σ	Ecart-type de l'erreur sur chaque coordonnée
L	Vraisemblance logarithmique négative
X	Variable aléatoire suivant la distribution normale
$\hat{\sigma}_i$	Estimation de l'écart-type σ_i.
S	Ensemble de correspondances putatives
ξ	Seuillage établi pour rejeter les faux appariements
Y	Variable aléatoire suivant la distribution de $\chi^2(2)$,
$M(u, v)$	Relation spatiale cohérente entre l'ensemble des vrais appariements
γ	Pourcentage des vrais appariements
θ	Ensemble des paramètres
z_n	Variable latente
$\mathcal{Q}(\theta, \theta^{old})$	Log-vraisemblance
F	Matrice fondamentale
P	Matrice de projection 3×4
K	Matrice intrinsèque 3×3
f	Distance focale
(f_x, f_y)	Distances focales verticales et horizontales
$[R \mid T]$	Matrice extrinsèque
$[x, y, z, 1]^T$	Coordonnées homogène
E	Matrice essentielle
$[t]_\times$	Matrice antisymétrique

V^T	Matrice orthogonale
$\det(E)$	Déterminant de la matrice E
$\mathrm{tr}(E)$	Trace de la matrice E
$\left({}^1\sigma ij, {}^2\sigma ij\right)$	Valeurs singulières non nulle
X^T	Transposée de la matrice X
V^T	Matrice orthogonale
$\left\{M^j, \{m_i^j\}\right\}$	Ensemble de pistes, où M^j est un point 3D désigne chaque piste, et m_i^j représente sa projection sur le plan image de la $i-$éme caméra
$\rho(a,b)$	Mesure de similarité qui compare deux vecteurs a et b
$\pi_i(p)$	Indique la projection de p sur l'image i
$\Omega(x)$	Définit un domaine de support autour du point x
$I_i(x)$	Désigne les intensités de l'image i dans le domaine de support autour du point x
M	Maillage
$DT(P)$	Triangulation de Delaunay
V	Nombre de sommets du maillage
E	Nombre d'arêtes du maillage
F	Nombre de faces du maillage
M_T	Maillage triangulaire 2-variété
G	Graphe représentant la structure globale d'un maillage triangulaire
G_d	Graphe dual
DT_k	Triangulation de Delaunay d'un ensemble de points S_k
$\mathcal{C}$	Cavité de Delaunay
$\mathcal{B}$	Boule de Delaunay
$Vor(V)$	Ensemble de cellules de Voronoi
$\{v_k\}_{k=1}^n$	Ensemble de n sites de Voronoi
$Vor(v_k, V)$	Cellule de Voronoi autour de v_k
S_i	$i^{ème}$ segment d'un maillage segmenté
$\left(cr_{ij}, cl_{ij}\right)$	Paire de cercles correspondants
I_l	Image gauche
I_r	Image droite
(pr_i, pl_i)	$i-$éme Correspondance putative

Z_{ij}	Zone de la scène
{R, r}	Symboles qui se réfèrent à la droite
{L, l}	Symboles qui se réfèrent à la gauche
pk_{ij}	Centre de cercle ck_{ij}
SC[i][j]	Diamètre du cercle ck_{ij}
$d(pk_m, pk_{ij})$	Distance euclidienne entre pk_m et le centre d'un cercle ck_{ij}
CTBO[m]	Valeurs qui expriment la chance qu'une paire (pr_m, pl_m) d'être faux appariements
M	Modèle est défini par un ensemble S de points 3D et N triangles $T_i(S_{i1}, S_{i2}, S_{i3})$, chacun avec trois sommets de scène S_{i1}, S_{i2}, S_{i3}.
$S_i = (X_i, Y_i, Z_i, 1)^T$	Point de la scène projeté sur le plan image m au point $p_{mi} = (u_{mi}, v_{mi}, 1)^T$
$t_i(p_{mi1}, p_{mi2}, p_{mi3})$	Projection du triangle $T_i(S_{i1}, S_{i2}, S_{i3})$ sur le plan image m
I_m	La $m^{ème}$ image
$A_i^j = \{p_{ik}/k = 1, \dots, n_{ij}\}$	Ensemble des points d'intérêt de l'image I_i correspond-il aux points d'intérêt de l'image I_j, avec n_{ij} le nombre de correspondances.
$P_m = (M_{uv}^m)_{u=1..3}^{v=1..4}$	Matrice de projection de la caméra correspondant à l'image I_m.
Col(p)	Couleur du point d'intérêt p.
$(R(p), G(p), B(p))$	Valeurs chromatiques (rouge, vert, bleu) du point d'intérêt p.
$Col(T_i)$	Couleur assignée à un triangle T_i
$Cg(T_i)$	Centre de gravité du triangle T_i

Bibliographie

[1] S. Katz, G. Leifman, and A. Tal, “Mesh segmentation using feature point and core extraction,” *Vis. Comput.*, vol. 21, no. 8–10, pp. 649–658, 2005.

[2] D. George, X. Xie, and G. K. L. Tam, “3D mesh segmentation via multi-branch 1D convolutional neural networks,” *Graph. Models*, vol. 96, pp. 1–10, 2018.

[3] Z. Shu *et al.*, “Unsupervised 3D shape segmentation and co-segmentation via deep learning,” *Comput. Aided Geom. Des.*, vol. 43, pp. 39–52, 2016.

[4] L. Shapira, A. Shamir, and D. Cohen-Or, “Consistent mesh partitioning and skeletonisation using the shape diameter function,” *Vis. Comput.*, vol. 24, no. 4, p. 249, 2008.

[5] M. Chevrel, M. Courtois, and G. Weill, “The SPOT satellite remote sensing mission,” *Photogramm. Eng. Remote Sensing*, vol. 47, pp. 1163–1171, 1981.

[6] M. Kumano, A. Ohya, and S. Yuta, “Obstacle avoidance of autonomous mobile robot using stereo vision sensor,” in *Intl. Symp. Robot. Automat*, 2000, pp. 497–502.

[7] P. Fua and Y. G. Leclerc, “Object-centered surface reconstruction: Combining multi-image stereo and shading,” *Int. J. Comput. Vis.*, vol. 16, no. 1, pp. 35–56, 1995.

[8] Y. Yang, Z. Ma, F. Nie, X. Chang, and A. G. Hauptmann, “Multi-class active learning by uncertainty sampling with diversity maximization,” *Int. J. Comput. Vis.*, vol. 113, no. 2, pp. 113–127, 2015.

[9] Y.-H. Chen and H.-C. Huang, “A wavelet-based image watermarking scheme for stereoscopic video frames,” in *Intelligent Information Hiding and Multimedia Signal Processing, 2013 Ninth International Conference on*, 2013, pp. 25–28.

[10] J. Yang, F. Li, Z. Sun, and S. Jiang, “A small target detection method based on human visual system and confidence measurement,” *J Inf. Hiding Multimed. Signal Process*, vol. 7, no. 2, pp. 448–459, 2016.

[11] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[12] P. H. S. Torr and A. Zisserman, “MLESAC: A new robust estimator with application to estimating image geometry,” *Comput. Vis. image Underst.*, vol. 78, no. 1, pp. 138–156, 2000.

[13] X. Li and Z. Hu, “Rejecting mismatches by correspondence function,” *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 1–17, 2010.

[14] J. Zhao, J. Ma, J. Tian, J. Ma, and D. Zhang, “A robust method for vector field learning with application to mismatch removing,” 2011.

[15] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (SURF),” *Comput. Vis. image Underst.*, vol. 110, no. 3, pp. 346–359, 2008.

[16] M. Jian, K.-M. Lam, J. Dong, and L. Shen, “Visual-patch-attention-aware saliency detection,” *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1575–1586, 2015.

[17] R. Sukthankar and Y. K. PCA-SIFT, “A More Distinctive Representation for Local Image Descriptors,” in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattem Recognition. USA: 1EEE*, 2004, pp. 1063–6919.

[18] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[19] B. Tordoff and D. W. Murray, “Guided sampling and consensus for motion estimation,” in *European conference on computer vision*, 2002, pp. 82–96.
[20] O. Chum and J. Matas, “Matching with PROSAC-progressive sample consensus,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1, pp. 220–226.
[21] O. Chum, J. Matas, and J. Kittler, “Locally optimized RANSAC,” in *Joint Pattern Recognition Symposium*, 2003, pp. 236–243.
[22] L. Mili, V. Phaniraj, and P. J. Rousseeuw, “Least median of squares estimation in power systems,” *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 511–523, 1991.
[23] R. Deriche, Z. Zhang, Q.-T. Luong, and O. Faugeras, “Robust recovery of the epipolar geometry for an uncalibrated stereo rig,” in *European Conference on Computer Vision*, 1994, pp. 567–576.
[24] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong, “A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry,” *Artif. Intell.*, vol. 78, no. 1–2, pp. 87–119, 1995.
[25] P. H. S. Torr and D. W. Murray, “Stochastic motion clustering,” in *European Conference on Computer Vision*, 1994, pp. 328–337.
[26] R. Hoseinnezhad and A. Bab-Hadiashar, “Fast estimation of epipolar geometry using high breakdown M-estimators,” in *Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society on*, 2007, pp. 159–166.
[27] P. J. Huber, “Robust statistics,” in *International Encyclopedia of Statistical Science*, Springer, 2011, pp. 1248–1251.
[28] J. Chen, J. Ma, C. Yang, and J. Tian, “Mismatch removal via coherent spatial relations,” *J. Electron. Imaging*, vol. 23, no. 4, p. 43012, 2014.
[29] Y. Furukawa and J. Ponce, “Accurate, dense, and robust multiview stereopsis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, 2010.
[30] O. Faugeras and R. Keriven, *Variational principles, surface evolution, PDE’s, level set methods and the stereo problem*. IEEE, 2002.
[31] F. Simões *et al.*, “Challenges in 3d reconstruction from images for difficult large-scale objects: A study on the modeling of electrical substations,” in *Virtual and Augmented Reality (SVR), 2012 14th Symposium on*, 2012, pp. 74–83.
[32] M. Pollefeys *et al.*, “Visual modeling with a hand-held camera,” *Int. J. Comput. Vis.*, vol. 59, no. 3, pp. 207–232, 2004.
[33] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
[34] D. Nistér, O. Naroditsky, and J. Bergen, “Visual odometry,” in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, vol. 1, pp. I–I.
[35] M. I. A. Lourakis and A. A. Argyros, “SBA: A software package for generic sparse bundle adjustment,” *ACM Trans. Math. Softw.*, vol. 36, no. 1, p. 2, 2009.
[36] R. Hartley and A. Zisserman, “Multiple view geometry in computer vision,” *Robotica*, vol. 23, no. 2, p. 271, 2005.
[37] T. S. Huang and O. D. Faugeras, “Some properties of the E matrix in two-view motion estimation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 12, pp. 1310–1312, 1989.
[38] D.-I. F. H. T. Köhler and R. Westermann, “Comparison of Self-Calibration Algorithms for Moving Cameras,” *Dep. Comput. Sci. Math. Univ. Appl. Sci.*, 2010.
[39] R. I. Hartley, “Estimation of relative camera positions for uncalibrated cameras,” in *European conference on computer vision*, 1992, pp. 579–587.
[40] K. Mikolajczyk *et al.*, “A comparison of affine region detectors,” *Int. J. Comput. Vis.*,

vol. 65, no. 1–2, pp. 43–72, 2005.
[41] C. Strecha, W. Von Hansen, L. Van Gool, P. Fua, and U. Thoennessen, "On benchmarking camera calibration and multi-view stereo for high resolution imagery," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1–8.
[42] K. Kraus, *Photogrammetry: geometry from images and laser scans*. Walter de Gruyter, 2011.
[43] A. Zisserman, A. Fitzgibbon, and G. Cross, "VHS to VRML: 3D graphical models from video sequences," in *Multimedia Computing and Systems, 1999. IEEE International Conference on*, 1999, vol. 1, pp. 51–57.
[44] S. Agarwal *et al.*, "Building rome in a day," *Commun. ACM*, vol. 54, no. 10, pp. 105–112, 2011.
[45] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—a modern synthesis," in *International workshop on vision algorithms*, 1999, pp. 298–372.
[46] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1–3, pp. 7–42, 2002.
[47] E. Hayman and D. W. Murray, "The effects of translational misalignment when self-calibrating rotating and zooming cameras," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 8, pp. 1015–1020, 2003.
[48] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *null*, 2006, pp. 519–528.
[49] Y. Furukawa and C. Hernández, "Multi-view stereo: A tutorial," *Found. Trends® Comput. Graph. Vis.*, vol. 9, no. 1–2, pp. 1–148, 2015.
[50] A. Shamir, "A survey on mesh segmentation techniques," in *Computer graphics forum*, 2008, vol. 27, no. 6, pp. 1539–1556.
[51] B. Krishnamoorthy and A. Tropsha, "Development of a four-body statistical pseudo-potential to discriminate native from non-native protein conformations," *Bioinformatics*, vol. 19, no. 12, pp. 1540–1548, 2003.
[52] A. Bowyer, "Computing dirichlet tessellations," *Comput. J.*, vol. 24, no. 2, pp. 162–166, 1981.
[53] D. F. Watson, "Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes," *Comput. J.*, vol. 24, no. 2, pp. 167–172, 1981.
[54] J.-F. Remacle, V. Bertrand, and C. Geuzaine, "A two-level multithreaded delaunay kernel," *Procedia Eng.*, vol. 124, pp. 6–17, 2015.
[55] G. Voronoi, "Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites.," *J. für die reine und Angew. Math.*, vol. 133, pp. 97–178, 1908.
[56] A. Uwitonze, J. Huang, Y. Ye, W. Cheng, and Z. Li, "Exact and heuristic algorithms for Space Information Flow," *PLoS One*, vol. 13, no. 3, p. e0193350, 2018.
[57] L. P. Chew, "Constrained delaunay triangulations," *Algorithmica*, vol. 4, no. 1–4, pp. 97–108, 1989.
[58] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*, vol. 501. John Wiley & Sons, 2009.
[59] A. Fabri, G. Giezeman, L. Kettner, S. Schirra, and S. Schönherr, "On the design of CGAL a computational geometry algorithms library," *Softw. Pract. Exp.*, vol. 30, no. 11, pp. 1167–1202, 2000.
[60] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, 1996.
[61] J. R. Shewchuk, "Triangle: Engineering a 2D quality mesh generator and Delaunay triangulator," in *Applied computational geometry towards geometric engineering*,

Springer, 1996, pp. 203–222.
[62] J.-D. Boissonnat, O. Devillers, and S. Hornus, “Incremental construction of the Delaunay triangulation and the Delaunay graph in medium dimension,” in *Proceedings of the twenty-fifth annual symposium on Computational geometry*, 2009, pp. 208–216.
[63] P. J. Frey and P.-L. George, *Mesh generation: application to finite elements*. ISTE, 2007.
[64] N. Amenta, S. Choi, and R. K. Kolluri, “The power crust,” in *Proceedings of the sixth ACM symposium on Solid modeling and applications*, 2001, pp. 249–266.
[65] N. Amenta, M. Bern, and M. Kamvysselis, “A new Voronoi-based surface reconstruction algorithm,” in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998, pp. 415–421.
[66] R. R. Stoll, *Set theory and logic*. Courier Corporation, 1979.
[67] G. Lavoué, F. Dupont, and A. Baskurt, “A new CAD mesh segmentation method, based on curvature tensor analysis,” *Comput. Des.*, vol. 37, no. 10, pp. 975–987, 2005.
[68] M. Attene, B. Falcidieno, and M. Spagnuolo, “Hierarchical mesh segmentation based on fitting primitives,” *Vis. Comput.*, vol. 22, no. 3, pp. 181–193, 2006.
[69] M. Garland, A. Willmott, and P. S. Heckbert, “Hierarchical face clustering on polygonal surfaces,” in *Proceedings of the 2001 symposium on Interactive 3D graphics*, 2001, pp. 49–58.
[70] D. Cohen-Steiner, P. Alliez, and M. Desbrun, “Variational shape approximation,” in *ACM Transactions on Graphics (ToG)*, 2004, vol. 23, no. 3, pp. 905–914.
[71] E. H. Cohen and M. Singh, “Geometric determinants of shape segmentation: Tests using segment identification,” *Vision Res.*, vol. 47, no. 22, pp. 2825–2840, 2007.
[72] I. S. Lim and E. C. Leek, “Curvature and the visual perception of shape: Theory on information along object boundaries and the minima rule revisited.,” *Psychol. Rev.*, vol. 119, no. 3, p. 668, 2012.
[73] Y.-K. Lai, S.-M. Hu, R. R. Martin, and P. L. Rosin, “Fast mesh segmentation using random walks,” in *Proceedings of the 2008 ACM symposium on Solid and physical modeling*, 2008, pp. 183–191.
[74] A. P. Witkin and P. S. Heckbert, “Using particles to sample and control implicit surfaces,” in *ACM SIGGRAPH 2005 Courses*, 2005, p. 260.
[75] S. Dasgupta, “Learning mixtures of Gaussians,” in *Foundations of computer science, 1999. 40th annual symposium on*, 1999, pp. 634–644.
[76] N. Vlassis and A. Likas, “A greedy EM algorithm for Gaussian mixture learning,” *Neural Process. Lett.*, vol. 15, no. 1, pp. 77–87, 2002.
[77] R. Zabih and V. Kolmogorov, “Spatially coherent clustering using graph cuts,” in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, vol. 2, pp. II–II.
[78] M. A. A. Cox and T. F. Cox, “Multidimensional scaling,” in *Handbook of data visualization*, Springer, 2008, pp. 315–347.
[79] J. B. Kruskal and M. Wish, “Multidimensional scaling. Number 07–011 in Sage University Paper series on quantitative applications in the social sciences.” Sage Publications, Beverly Hills, 1978.
[80] J. B. Kruskal, “Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis,” *Psychometrika*, vol. 29, no. 1, pp. 1–27, 1964.
[81] T. F. Cox, “Multidimensional scaling,” 1994.
[82] A. Golovinskiy and T. Funkhouser, “Randomized cuts for 3D mesh analysis,” in *ACM transactions on graphics (TOG)*, 2008, vol. 27, no. 5, p. 145.
[83] T. Funkhouser *et al.*, “Modeling by example,” in *ACM Transactions on Graphics (TOG)*, 2004, vol. 23, no. 3, pp. 652–663.
[84] X. Chen, A. Golovinskiy, and T. Funkhouser, “A benchmark for 3D mesh

segmentation," in *Acm transactions on graphics (tog)*, 2009, vol. 28, no. 3, p. 73.
[85] M. Ghosh, N. M. Amato, Y. Lu, and J.-M. Lien, "Fast approximate convex decomposition using relative concavity," *Comput. Des.*, vol. 45, no. 2, pp. 494–504, 2013.
[86] J.-M. Lien and N. M. Amato, "Approximate convex decomposition of polyhedra," in *Proceedings of the 2007 ACM symposium on Solid and physical modeling*, 2007, pp. 121–131.
[87] J.-M. Lien and N. M. Amato, "Approximate convex decomposition of polyhedra and its applications," *Comput. Aided Geom. Des.*, vol. 25, no. 7, pp. 503–522, 2008.
[88] M. Attene, S. Katz, M. Mortara, G. Patané, M. Spagnuolo, and A. Tal, "Mesh segmentation-a comparative study," in *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*, 2006, p. 7.
[89] E. Elghoul, "Segmentation de maillages 3D par l'exemple." Télécom ParisTech, 2014.
[90] R. Arcila, "Séquences de maillages: classification et méthodes de segmentation." Université Claude Bernard-Lyon I, 2011.
[91] D. Marr, T. Poggio, E. C. Hildreth, and W. E. L. Grimson, "A computational theory of human stereo vision," in *From the Retina to the Neocortex*, Springer, 1991, pp. 263–295.
[92] S. Shlafman, A. Tal, and S. Katz, "Metamorphosis of polyhedral surfaces using decomposition," in *Computer graphics forum*, 2002, vol. 21, no. 3, pp. 219–228.
[93] M. Lhuillier and L. Quan, "Match propagation for image-based modeling and rendering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 8, pp. 1140–1146, 2002.
[94] A. Kovacs, "Hierarchical Mesh Decomposition Using Fuzzy Clustering and Cuts Katz and Tal, Siggraph 2003," 2009.
[95] V. Krayevoy and A. Sheffer, "Variational, meaningful shape decomposition," in *ACM SIGGRAPH 2006 Sketches*, 2006, p. 50.
[96] A. F. Koschan, "Perception-based 3D triangle mesh segmentation using fast marching watersheds," in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, 2003, vol. 2, p. II-27-II-32.
[97] Z. Xie, K. Xu, L. Liu, and Y. Xiong, "3d shape segmentation and labeling via extreme learning machine," in *Computer graphics forum*, 2014, vol. 33, no. 5, pp. 85–95.
[98] M. M. Bronstein and I. Kokkinos, "Scale-invariant heat kernel signatures for non-rigid shape recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010, pp. 1704–1711.
[99] R. Gal and D. Cohen-Or, "Salient geometric features for partial shape matching and similarity," *ACM Trans. Graph.*, vol. 25, no. 1, pp. 130–150, 2006.
[100] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*. Cengage Learning, 2014.
[101] R. Liu, C. Zhang, and J. Tian, "Point matching based on non-parametric model," in *MIPPR 2015: Pattern Recognition and Computer Vision*, 2015, vol. 9813, p. 98130X.
[102] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, 2002.
[103] D. E. Knuth, *The Art of Computer Programming, Volume 4A: Combinatorial Algorithms, Part 1*. Pearson Education India, 2011.
[104] H. Hirschmuller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 2007, pp. 1–8.
[105] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, 2003, vol. 1, pp. I–I.

[106] D. Scharstein *et al.*, “High-resolution stereo datasets with subpixel-accurate ground truth,” in *German Conference on Pattern Recognition*, 2014, pp. 31–42.
[107] D. Scharstein and C. Pal, “Learning conditional random fields for stereo,” in *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*, 2007, pp. 1–8.
[108] J. Jeon, Y. Jung, H. Kim, and S. Lee, “Texture map generation for 3D reconstructed scenes,” *Vis. Comput.*, vol. 32, no. 6–8, pp. 955–965, 2016.
[109] K. Fu, *VLSI for pattern recognition and image processing*, vol. 13. Springer Science & Business Media, 2013.
[110] P. Lai and C. Samson, “Applications of mesh parameterization and deformation for unwrapping 3D images of rock tunnels,” *Tunn. Undergr. Sp. Technol.*, vol. 58, pp. 109–119, 2016.
[111] B. Zeramdini, C. Robert, G. Germain, and T. Pottier, “Simulation of metal forming processes with a 3D adaptive remeshing procedure,” in *AIP Conference Proceedings*, 2016, vol. 1769, no. 1, p. 200018.
[112] L. Yao, S. Huang, H. Xu, and P. Li, “Quadratic error metric mesh simplification algorithm based on discrete curvature,” *Math. Probl. Eng.*, vol. 2015, 2015.
[113] R. Liu, H. Zhang, and J. Busby, “Convex hull covering of polygonal scenes for accurate collision detection in games,” in *Proceedings of graphics interface 2008*, 2008, pp. 203–210.
[114] L. Serino, G. S. di Baja, and C. Arcelli, “Object decomposition via curvilinear skeleton partition,” in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 4081–4084.
[115] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, “Multi-view stereo for community photo collections,” 2007.
[116] M. Lhuillier and L. Quan, “A quasi-dense approach to surface reconstruction from uncalibrated images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 418–433, 2005.
[117] A. O. Ulusoy, M. J. Black, and A. Geiger, “Semantic Multi-view Stereo: Jointly Estimating Objects and Voxels.,” in *CVPR*, 2017, pp. 4531–4540.
[118] N. El Akkad, M. Merras, A. Saaidi, and K. Satori, “Camera self-calibration with varying intrinsic parameters by an unknown three-dimensional scene,” *Vis. Comput.*, vol. 30, no. 5, pp. 519–530, 2014.
[119] F. Cazals and J. Giesen, “Delaunay triangulation based surface reconstruction,” in *Effective computational geometry for curves and surfaces*, Springer, 2006, pp. 231–276.
[120] C. M. Gold and M. Dakowicz, “The crust and skeleton–applications in GIS,” in *Proceedings, 2nd. International Symposium on Voronoi Diagrams in Science and Engineering*, 2005, pp. 33–42.
[121] S. K. Bhatia, “Adaptive K-Means Clustering.,” in *FLAIRS conference*, 2004, pp. 695–699.

Liste des publications

[L1] A. Taime, J. Riffi, A. Saaidi, and K. Satori, “Robust point matching via corresponding circles”,Multimedia Tools and Applications, vol. 77, no. 12, pp. 15027–15046, 2018.

[L2] A. Taime, A. Saaidi, and K. Satori, “Fast point matching using corresponding circles”,Multimedia Tools and Applications, vol. 77, no. 23, pp. 31139–31157, 2018.

[L3] A. Taime, A. Saaidi, and K. Satori, “A new semantic segmentation approach of 3D mesh using the stereoscopic image colors”,Multimedia Tools and Applications, vol. 77, no. 20, pp. 27143–27162, 2018.

[L4] A. Taime, A. Saaidi, and K. Satori, “Comparative study of mesh simplification algorithms’’. Int. Proceedings of the Mediterranean Conference on Information & Communication Technologies 2015, pp. 287-295. Springer, Cham.