N° d'ordre : D.U. ?
EDSPIC : ?

Université Clermont Auvergne

École Doctorale
Sciences Pour l'Ingénieur de Clermont-Ferrand

Thèse présentée par :
Alexandre Morgand

Formation Doctorale :
Électronique et Système

en vue de l'obtention du grade de

Docteur d'Université

Spécialité : Vision pour la Robotique

Un modèle géométrique multi-vues des taches spéculaires basé sur les quadriques avec application en réalité augmentée

---

Mme. Marie-Odile Berger	Examinatrice, Présidente
M. Vincent Lepetit	Rapporteur
M. Eric Marchand	Rapporteur
M. Mohamed Tamaazousti	Co-encadrant
M. Adrien Bartoli	Directeur de thèse

Résumé

La réalité augmentée (RA) consiste en l'insertion d'éléments virtuels dans une scène réelle, observée à travers un écran ou en utilisant un système de projection sur la scène ou l'objet d'intérêt. Les systèmes de réalité augmentée peuvent prendre des différentes formes pour obtenir l'équilibre désiré entre trois critères : précision, latence et robustesse. Il est possible d'identifier trois composants principaux à ces systèmes : *localisation, reconstruction et affichage*. Les contributions de cette thèse se concentrent essentiellement sur l'affichage et plus particulièrement le rendu des applications de réalité augmentée.

À l'opposé des récentes avancées dans le domaine de la localisation et de la reconstruction, l'insertion d'éléments virtuels de façon plausible et esthétique reste une problématique compliquée, mal-posée et peu adaptée à un contexte temps réel. En effet, cette insertion requiert une reconstruction de l'illumination de la scène afin d'appliquer les conditions lumineuses adéquates à l'objet inséré. L'illumination de la scène peut être divisée en plusieurs catégories. Nous pouvons modéliser l'environnement de façon à décrire l'interaction de la lumière incidente et réfléchie pour chaque point 3D d'une surface. Il est également possible d'expliciter l'environnement en calculant la position des sources de lumière, leur type (lampe de bureau, néon, ampoule, ...), leur intensité et leur couleur. Pour insérer un objet de façon cohérente et réaliste, il est primordial d'avoir également une connaissance de la surface recevant l'illumination. Cette interaction lumière/matériaux est dépendante de la géométrie de la surface, de sa composition chimique (matériau) et de sa couleur. Pour tous ces aspects, le problème de reconstruction de l'illumination est difficile, car il est très complexe d'isoler l'illumination sans connaissance a priori de la géométrie, des matériaux de la scène et de la pose de la caméra observant la scène.

De manière générale, sur une surface, une source de lumière laisse plusieurs traces telles que les ombres, qui sont créées par l'occultation de rayons lumineux par un objet, et les réflexions spéculaires ou spécularités qui se manifestent par la réflexion partielle ou totale de la lumière. Bien que ces spécularités soient souvent considérées comme des éléments parasites dans les applications de localisation de caméra, de reconstruction ou encore de segmentation, ces éléments donnent des informations cruciales sur la position et la couleur de la source lumineuse, mais également sur la géométrie de la surface et la réflectance du matériau où elle se manifeste. Face à la difficulté du problème de modélisation de la lumière et plus particulièrement du calcul des l'ensemble des paramètres de la lumière, nous nous sommes focalisés, dans cette thèse, sur l'étude des spécularités et sur toutes les informations qu'elles peuvent fournir pour la compréhension de la scène. Plus particulièrement, nous savons qu'une spécularité est définie comme la réflexion d'une source de lumière sur une surface réfléchissante. Partant de cette remarque, nous avons exploré la

possibilité de considérer la spécularité comme étant une image issue de la projection d'un objet 3D dans l'espace.

Nous sommes partis d'un constat simple, mais peu traité par la littérature qui est que les spécularités présentent une forme elliptique lorsqu'elles apparaissent sur une surface plane. À partir de cette hypothèse, pouvons-nous considérer un objet 3D fixe dans l'espace tel que sa projection perspective dans l'image corresponde à la forme de la spécularité ? Plus particulièrement, nous savons qu'un ellipsoïde projeté perspectivement donne une ellipse. Considérer le phénomène de spécularité comme un phénomène géométrique a de nombreux avantages. Premièrement, la reconstruction d'un objet 3D et plus particulièrement d'une quadrique est une opération qui a fait l'objet de plusieurs publications dans l'état de l'art. Deuxièmement, cette modélisation permet une grande souplesse sur le suivi de l'état de la spécularité et plus particulièrement de la source de lumière. En effet, si une lumière s'éteint, il est facile de visualiser dans l'image que la spécularité disparait si nous en connaissons le contour (et réciproquement si la lumière s'allume à nouveau). En considérant le problème de façon photométrique, il est plus difficile de séparer la composante spécularité du reste de la scène, car les modèles de l'état de l'art n'explicitent pas en général la spécularité. De plus, une spécularité se présente généralement comme un élément identifiable dans l'image en raison de son intensité élevée et de son contour très visible dans l'image. Pour ces raisons, un nombre important d'informations peut être récupéré en étudiant seulement la forme de la spécularité sans qu'intervienne la connaissance du matériau ou de la rugosité de la surface.

Dans un premier temps, nous avons formalisé un modèle géométrique dédié à la prédiction de spécularités appelé *JOint LIght-MAterial Specularity* (JOLIMAS). Le formalisme et la reconstruction de ce modèle ont d'abord été traités pour le cas de surfaces planes. Ce premier modèle correspond à la forme primale de JOLIMAS. Nous avons par la suite étendu ce modèle aux surfaces convexes de courbures constantes afin d'augmenter la généricité de notre modèle par un nouveau modèle appelé JOLIMAS dual. Cette amélioration a été engendrée par une représentation virtuelle d'une caméra qui permet de reconstruire le modèle de façon plus générique et plus flexible. Pour JOLIMAS primal et dual, la prédiction de spécularité est réalisée en projetant la quadrique obtenue en fonction de la pose de caméra (réelle pour la version primale et virtuelle pour la version duale). Enfin, nous avons généralisé le modèle JOLIMAS à toute courbure en étudiant le lien entre changement de courbure et changement de la forme de spécularité. Plus spécifiquement, pour une spécularité observée sur une surface de courbure quelconque, nous modifions les contours associés à cette spécularité afin de simuler la forme qu'aurait la spécularité sur une surface plane (non-courbe). En appliquant cette transformation pour chaque spécularité observée, nous pouvons reconstruire notre modèle JOLIMAS dans une représentation canonique de la quadrique indépendante de la courbure des surfaces sur lesquelles les spécularités se présentent. Pour le processus de prédiction de spécularité, nous faisons le procédé inverse en projetant, sous forme d'ellipse, la quadrique en fonction de la pose de caméra ce qui donne la forme de spécularité pour une surface plane. Par la suite, nous modifions les contours de cette ellipse projetée afin de correspondre à la courbure de la surface actuelle et ainsi correspondre à la forme de la spécularité pour un nouveau point de vue. Notons que pour toutes ces versions de JOLIMAS, la pose de caméra et la géométrie de la surface, où les spécularités se présentent, sont connues. De plus, nous supposons que la réflectance des surfaces utilisées pour reconstruire notre modèle est constante. Afin de montrer l'étendue applicative du modèle, nous avons développé une application de *retexturing* qui permet

de changer la texture d'une surface présentant une spécularité et de simuler la spécularité retirée sur la nouvelle texture en utilisant le modèle JOLIMAS. Le modèle JOLIMAS ne fournissant que le contour externe d'une spécularité, nous avons proposé une nouvelle fonction d'intensité qui simule le comportement photométrique d'une spécularité.

Cette thèse a fait l'objet de trois publications dans des conférences internationales : VISAPP (2014), ISMAR (2016; 2017c), deux journaux internationaux TVCG (2017e; 2017d), quatre conférences nationales : deux articles à ORASIS (2015b; 2015a), RFIA (2015) et ORASIS (2017) et un brevet accepté en juin (2018). Une soumission au journal TVCG est en attente de revue au moment de l'écriture du manuscrit.

Mots clés : Prédiction de spécularité, Réalité augmentée, Retexturing, Reconstruction de quadrique, Conique, Illumination, Multi-vues.

Abstract

Augmented Reality (AR) consists in inserting virtual elements in a real scene, observed through a screen or a projection system on the scene or the object of interest. The augmented reality systems can take different forms to obtain a balance between three criteria : precision, latency and robustness. It is possible to identify three main components to these systems : *localization, reconstruction and display*. The contributions of this thesis focus essentially on the display and more particularly the rendering of augmented reality applications.

Contrary to the recent advances in the field of localization and reconstruction, the insertion of virtual elements in a plausible and aesthetic way remains a complicated problematic, ill-posed and not adapted to a real-time context. Indeed, this insertion requires a good understanding of the lighting conditions of the scene. The lighting conditions of the scene can be divided in several categories. First, we can model the environment to describe the interaction between the incident and reflected light pour each 3D point of a surface. Secondly, it is also possible to explicitly the environment by computing the position of the light sources, their type (desktop lamps, fluorescent lamp, light bulb, ...), their intensities and their colors. Finally, to insert a virtual object in a coherent and realistic way, it is essential to have the knowledge of the surface's geometry, its chemical composition (material) and its color. For all of these aspects, the reconstruction of the illumination is difficult because it is really complex to isolate the illumination without prior knowledge of the geometry, material of the scene and the camera pose observing the scene.

In general, on a surface, a light source leaves several traces such as shadows, created from the occultation of light rays by an object, and the specularities (or specular reflections) which are created by the partial or total reflection of the light. These specularities are often described as very high intensity elements in the image.

Although these specularities are often considered as outliers for applications such as camera localization, reconstruction or segmentation, these elements give crucial information on the position and color of the light source but also on the surface's geometry and the material's reflectance where these specularities appear. To address the light modeling problem, we focused, in this thesis, on the study of specularities and on every information that they can provide for the understanding of the scene. More specifically, we know that a specularity is defined as the reflection of the light source on a shiny surface. From this statement, we have explored the possibility to consider the specularity as the image created from the projection of a 3D object in space.

We started from the simple but little studied in the literature observation that spe-

cularities present an elliptic shape when they appear on a planar surface. From this hypothesis, can we consider the existence of a 3D object fixed in space such as its perspective projection in the image fit the shape of the specularity ? We know that an ellipsoid projected perspectivally gives an ellipse. Considering the specularity as a geometric phenomenon presents various advantages. First, the reconstruction of a 3D object and more specifically of an ellipsoid, has been the subject to many publications in the state of the art. Secondly, this modeling allows a great flexibility on the tracking of the state of the specularity and more specifically the light source. Indeed, if the light is turning off, it is easy to visualize in the image if the specularity disappears if we know the contour (and reciprocally of the light is turning on again). By considering the problem of light modeling in a photometric way, it is way more difficult to separate the specular component from the rest of the scene because the models from the state of the art does not explicit in general the specular part. Moreover, a specularity is often in an explicit and detailed form in a video sequence. An important quantity of information can be recovered by studying the shape of the specularity without prior knowledge of the material or roughness of the surface.

Firstly, we have formalized a geometric model dedicated to the specularity prediction called *JOint LIght-MAterial Specularity* (JOLIMAS). The formalism and the reconstruction of this model have been studied for the case of planar surfaces first. Afterwards, we extended our model to convex surfaces in order to improve the genericity of our model in a new model called dual JOLIMAS. This improvement has been realized by using a virtual representation of a camera which allows to reconstruct the model in a more generic and flexible way. Finally, we have generalized the JOLIMAS model for every curvature by studying the link between the curvature changes and the changes in the specularity shape. More specifically, for a specularity observed on a surface of any curvature, we modify the contours associated to the specularity in order to simulate the shape that the specularity would have specularity on a planar surface (non-curved). By applying this transformation for each specularity observed, we can reconstruct our JOLIMAS model in a canonical representation of the quadric independent of the curvature of the surfaces where the specularities are observed. For the specularity prediction, we do the inverse process by projection, in the form of an ellipse shape, the quadric according to the camera pose which gives the specularity's shape for a planar surface. We then modify the contours of the projected ellipse to match the curvature of the surface and fit the shape of the specularity for a new viewpoint. Note that for every version of JOLIMAS, the camera poses and the geometry of the surface, where the specularities is present, are known. Moreover, we suppose that the reflectance of the surfaces used to reconstruct our model is constant. To show the application scope of the model, we have developed a rexturing application which allows to change the texture of a specular surface and to synthesize the specularity that has been removed from the new texture by using the JOLIMAS model. Since our model only provides the external contour of a specularity, we proposed new intensity functions to synthesize the photometric behavior of a specularity.

This thesis has been the object of several international publications : VISAPP (2014), ISMAR (2016; 2017c), two international TVCG journals (2017e; 2017d), four national conferences : two ORASIS articles (2015b; 2015a), RFIA (2015) and ORASIS (2017) and a patent accepted in June (2018). A submission to a TVCG journal is under review at tje time of the writing of the manuscript.

Key-words : Specularity prediction, Augmented Reality, Retexturing, Quadric reconstruction, Conic, Illumination, Multiple views.

Table des matières

1 Introduction **1**

2 Notions de base **5**

2.1 Notations . . . 5
2.2 Espace projectif et géométrie associée . . . 8
2.2.1 Géométrie projective . . . 8
2.2.2 Le plan projectif . . . 8
2.2.3 Espace projectif . . . 9
2.3 Caméras perspectives . . . 9
2.3.1 Projection perspective . . . 9
2.3.2 Rétroprojection . . . 12
2.4 Image . . . 13
2.4.1 Définition et notations . . . 13
2.4.2 Opérations . . . 13
2.5 Conique et quadrique . . . 14
2.5.1 Conique . . . 14
2.5.1.1 Transformation projective . . . 15
2.5.1.2 Droite tangente à une conique . . . 15
2.5.1.3 Enveloppe de la conique . . . 15
2.5.1.4 Coniques dégénérées . . . 16
2.5.1.5 Enveloppes de coniques dégénérées . . . 16
2.5.2 Quadriques . . . 16
2.5.2.1 Transformation projective . . . 17
2.5.2.2 Enveloppe de quadrique . . . 17
2.5.2.3 Quadriques dégénérées . . . 17
2.5.2.4 Enveloppe de quadriques dégénérées . . . 18
2.6 Géométrie multi-vues et reconstruction 3D . . . 18
2.6.1 Géométrie épipolaire . . . 19
2.6.2 Calcul de la géométrie de l'environnement . . . 21
2.6.2.1 Poses de caméras et déplacement relatif . . . 22
2.6.2.2 Calcul du déplacement relatif par associations 2D/2D . . . 22

2.6.2.3 Calcul de la structure de l'environnement . . . 23
2.6.2.4 Calcul de pose par associations 2D/3D . . . 24
2.7 Notion de courbure d'une surface . . . 24
2.7.1 Première et deuxième formes fondamentales . . . 26
2.7.2 Courbure normale et principale . . . 27
2.7.3 Courbure de Gauss . . . 28
2.8 Optimisation numérique . . . 29
2.8.1 Méthodes linéaires . . . 29
2.8.1.1 Moindres carrés linéaires . . . 29
2.8.1.2 Décomposition en valeurs singulières . . . 29
2.8.2 Méthodes non linéaires . . . 30
2.8.2.1 Descente de gradient . . . 30
2.8.2.2 Newton . . . 31
2.8.2.3 Levenberg-Marquardt . . . 32
2.9 Interaction lumière/matière . . . 32
2.9.1 Phénomène de réflexion spéculaire . . . 33
2.9.2 Réflectivité bidirectionnelle . . . 33
2.10 Illumination globale . . . 34
2.11 Modèles d'illumination locale . . . 35
2.11.1 Modèle dichromatique . . . 35
2.11.2 Modèles de Phong et Blinn-Phong . . . 36
2.11.2.1 Composante ambiante . . . 37
2.11.2.2 Composante diffuse . . . 37
2.11.2.3 Composante spéculaire . . . 37

3 État de l'art 39
3.1 Prédiction de spécularité . . . 39
3.1.1 Rendu en illumination globale . . . 40
3.1.2 Modèles d'illumination locale . . . 41
3.1.3 Modèles géométriques . . . 42
3.2 Positionnement . . . 44

4 Un modèle primal pour les surfaces planes 45
4.1 Introduction . . . 45
4.2 Approche proposée . . . 46
4.3 Motivation théorique . . . 47
4.3.1 Modèles de Phong et Blinn-Phong . . . 47
4.3.2 Terme spéculaire . . . 48
4.3.2.1 Point d'intensité maximale . . . 49
4.3.2.2 Contour externe . . . 49
4.3.2.3 Isocontours internes . . . 50
4.3.3 Validation empirique de l'hypothèse du modèle à partir de données simulées . . . 50
4.4 Estimation du modèle . . . 53
4.4.1 Vue globale . . . 53
4.4.2 Reconstruction de quadrique par la méthode de Cross *et al.* . . . 53
4.5 Résultats expérimentaux . . . 54

4.5.1 Validation du modèle sur des données de synthèse . . . 54
4.5.2 Données réelles . . . 56
4.6 Limites de l'approche . . . 57
4.6.1 Surfaces miroirs et spécularités à forme non-elliptique . . . 57
4.6.2 Influence de la rugosité de la surface . . . 58
4.7 Conclusion . . . 58

5 Un modèle dual pour les surfaces convexes à courbure constante 61
5.1 Introduction . . . 61
5.2 Approche proposée . . . 62
5.3 JOLIMAS dual . . . 62
5.3.1 Notation et formalisation . . . 62
5.3.2 Similarité du comportement d'une surface spéculaire avec un miroir 63
5.3.3 Point d'intensité maximale . . . 63
5.4 Représentation par caméras virtuelles . . . 64
5.5 Triangulation de la position de la source lumineuse . . . 66
5.5.1 Estimation initiale . . . 66
5.5.2 Raffinement de la position de la lumière et du point d'intensité maximale . . . 66
5.5.3 Interpolation de normale . . . 68
5.6 Surfaces convexes par morceaux . . . 68
5.6.1 Principe . . . 68
5.6.2 Décomposition d'un modèle CAO en surfaces convexes approximatives 69
5.7 Résultats expérimentaux . . . 70
5.7.1 Comparaison avec JOLIMAS primal . . . 71
5.7.2 Sensibilité aux changements de courbure . . . 71
5.7.3 Prédiction de spécularité sur des données réelles . . . 71
5.8 Limites de l'approche . . . 74
5.9 Conclusion . . . 75

6 Un modèle général pour les surfaces à courbure variable 77
6.1 Introduction . . . 77
6.2 Lien entre déformation, spécularité, courbure de la surface et champ de vision 78
6.3 Approche proposée - Représentation canonique pour la reconstruction 3D de quadrique . . . 79
6.3.1 Utilisation de la courbure gaussienne . . . 79
6.3.2 Mise en pratique sur un modèle CAO - Notion d'angles limites . . . 82
6.4 Modifications du processus de prédiction de spécularité . . . 83
6.5 Résultats expérimentaux . . . 83
6.5.1 Données de synthèse . . . 83
6.5.2 Données réelles . . . 85
6.6 Limites de l'approche . . . 87
6.7 Conclusion . . . 87

7 Application à la réalité augmentée 91
7.1 Introduction . . . 91
7.2 Suivi de spécularité en contexte multi-lumières . . . 92
7.3 Filtrage multi-vues . . . 93

7.4 Gestion de lumières à état changeant . . . 93
7.5 Retexturing sur des objets réels . . . 94
7.5.1 Première approche : modélisation par fonction Gaussienne pour surface planes . . . 95
7.5.2 Deuxième approche : retexturing pour tous types de surfaces . . . 96
7.6 Limitations . . . 96
7.7 Temps d'exécution . . . 97
7.8 Perspectives . . . 97
7.9 Conclusion . . . 97

Conclusion et perspectives **103**

Annexes **106**

A Détection de spécularité dans les images en temps réel **107**
A.1 Travaux antérieurs . . . 108
A.1.1 Approches hors-ligne . . . 108
A.1.2 Approches en ligne . . . 109
A.2 Méthode proposée . . . 110
A.2.1 Vue d'ensemble . . . 110
A.2.2 Prétraitement . . . 110
A.2.3 Étape de seuillage . . . 111
A.2.4 Post-traitement . . . 112
A.2.5 Complexité algorithmique . . . 116
A.3 Évaluation expérimentale . . . 116
A.3.1 Protocole expérimental . . . 116
A.3.2 Évaluation quantitative . . . 117
A.3.3 Évaluation qualitative . . . 118
A.4 Conclusion et discussion . . . 118

B Reconstruction d'objet sous forme de quadrique **121**
B.1 Introduction . . . 121
B.2 État de l'art . . . 121
B.2.1 Approches par reconstruction multi-vues . . . 122
B.2.2 Autres approches . . . 122
B.3 Solutions proposées . . . 123
B.3.1 Contrainte épipolaire . . . 123
B.3.2 Distance 2D ellipse/point . . . 123
B.3.2.1 Distance de Jaccard . . . 123
B.3.2.2 Paramétrisation d'ellipse et distance point/ellipse . . . 124
B.3.2.3 Paramétrisation d'ellipse pour l'association point/point . . . 124
B.3.2.4 Distance point/point d'ellipses . . . 127
B.3.3 Raffinement non-linéaire . . . 127
B.3.3.1 Fonction de coût avec la distance de Jaccard . . . 127
B.3.3.2 Limitations . . . 127
B.3.3.3 Fonction de coût avec variables latentes . . . 128
B.3.4 Critère d'images clés . . . 128
B.4 Distance 3D de quadrique . . . 129

B.4.1 Distance de centre de quadrique . . . 130
B.4.2 Distance globale 3D . . . 130
B.4.3 Discrétisation de quadrique . . . 130
B.4.4 Recherche de points voisins les plus proches de la deuxième ellipsoïde 130
B.5 Vectorisation du système de projection de la quadrique duale . . . 131

C Étude analytique des modèles de Phong et Blinn-Phong 143
C.1 Modèles étudiés . . . 143
C.1.1 Modèle de Blinn-Phong . . . 143
C.1.2 Analyse des contours externes de la spécularité pour $\tau = 0$ pour le modèle de Blinn-Phong . . . 144
C.1.3 Cas spécial, analyse de $\tau = 1$ pour le modèle de Phong . . . 145
C.2 Point d'intensité maximale dans une spécularité sous le modèle de Phong . 146
C.3 Notation et contexte . . . 146
C.4 Résultat principal . . . 147

Bibliographie 151

CHAPITRE 1

Introduction

Cette thèse a été réalisée entre novembre 2014 et novembre 2017 au Laboratoire Vision et Ingénierie des Contenus (LVIC) du CEA LIST, à Saclay (91) et à l'unité de recherche ISIT (Institut des Sciences de l'Image pour les Techniques interventionnelles, UMR 6284 UdA, CNRS) à Clermont-Ferrand.

Contexte

La réalité augmentée consiste en l'insertion d'éléments virtuels (que nous appellerons augmentations) dans la réalité en temps réel. Au moment de la rédaction de ce mémoire, les applications scientifiques, industrielles et commerciales de la réalité augmentée se sont fortement développées notamment grâce à l'essor des méthodes d'affichage comme le casque Hololens de Microsoft [1] et les kits de développement de réalité augmentée sur *smartphone* d'Apple (ARkit [2]) et Android (ARcore [3]). Les évolutions dans le domaine de la localisation de caméra permettent une industrialisation d'applications plus ambitieuses comme illustrée à la figure 1.1 :

- aide à la vente : la visualisation immédiate des différentes configurations possibles d'un produit (par exemple une voiture, un meuble, ...) sans contraintes de stock, pourrait aider le choix de l'utilisateur et contribuer à une vente réussie tout en réduisant les coûts associés aux points de vente ;
- aide à la maintenance : les systèmes industriels complexes nécessitent des procédures de maintenance complexes ; la réalité augmentée pourrait simplifier la formation des techniciens et augmenter la sécurité en affichant les instructions pas à pas sur le système réel ;
- aide à la navigation : que ce soit à l'intérieur ou à l'extérieur, un «GPS amélioré» pourrait aider à signaler précisément l'itinéraire à suivre et d'autres informations importantes ainsi que les dangers imminents ;

1. https://www.microsoft.com/en-us/hololens
2. https://developer.apple.com/arkit/
3. https://developers.google.com/ar/

– aide à la chirurgie : la plupart des opérations lourdes font l'objet d'une acquisition 3D préliminaire des organes concernés (scanner, . . .), l'affichage de ces informations directement sur le patient en réalité augmentée pourrait grandement aider le praticien en le guidant vers la zone à traiter et en mettant en évidence les zones à éviter, par exemple les artères et les nerfs.

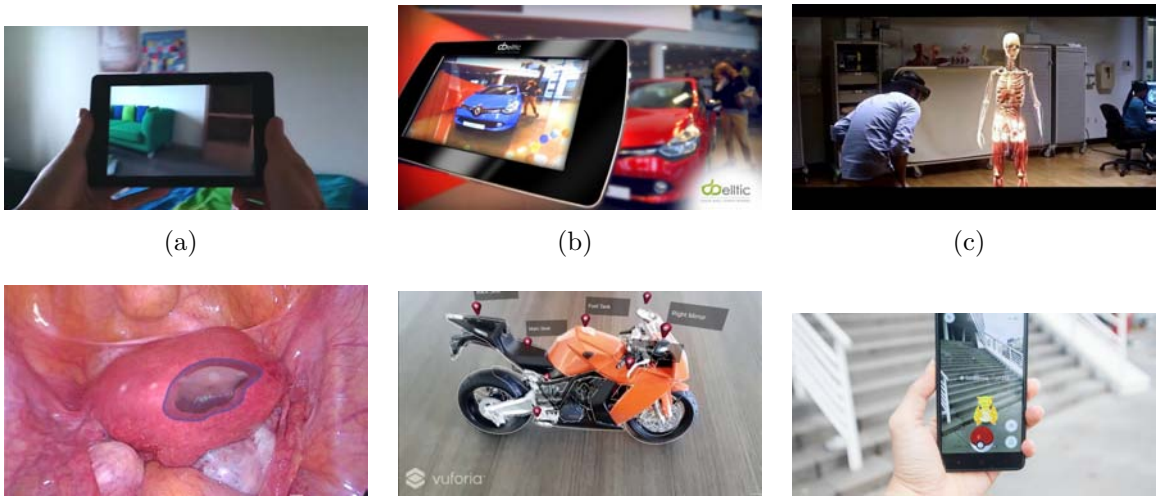


FIGURE 1.1 – Diverses applications de réalité augmentée : aide à la vente immobilière (a) et automobile (b), aide à la formation médicale (c) et pour de l'affichage complémentaire pour de la chirurgie (d) comme la méthode de Ozgür et al. (2017), information complémentaire sur un objet (e) et jeu vidéo (f).

Un système de réalité augmentée est complexe et fait intervenir des composants variés, que l'on peut classer en trois briques principales : reconstruction, localisation et rendu. Dans ce mémoire, nous nous concentrerons sur les problématiques de reconstruction ce qui implique également une meilleure compréhension de la scène. De plus, nous avons contribué à l'amélioration du rendu des applications de réalité augmentée par l'intermédiaire d'une application de *retexturing*. Ces améliorations se font par l'utilisation d'un nouveau modèle géométrique permettant la prédiction de spécularités en temps réel pour un point de vue donné. Ce modèle a fait l'objet de plusieurs extensions afin d'en augmenter les performances et la généricité. Le chapitre 2 introduit les notions de base utiles à la compréhension du mémoire : mathématiques (algèbre, géométrie projective, coniques, quadriques, courbure), vision par ordinateur (localisation, reconstruction, mise en correspondance d'images), modélisation de l'illumination (globale, locale et autres) et optimisation numérique. Nous détaillons le positionnement de notre modèle géométrique de prédiction de spécularité par rapport à l'état de l'art dans le chapitre 3.

Problématique

Le chapitre 4 est dédié à l'introduction de notre modèle géométrique appelé *JOint LIght Material-Specularity* (JOLIMAS) pour la prédiction de spécularités. Ces éléments lumineux sont primordiaux dans le réalisme des applications de réalité augmentée, car ils

participent activement au processus de perception humaine de la courbure d'un objet et du type de matériau de l'objet comme expliqué par Blake (1990). Cependant, la spécularité est un phénomène photométrique complexe, car il est décrit par de nombreux paramètres (couleur, forme, position) et dépend de nombreux facteurs externes (géométrie de la scène, matériau de la surface, paramètres de caméra comme le temps d'exposition ou l'ouverture de l'objectif). Afin de simplifier la problématique, nous nous sommes restreints, dans un premier temps, au cas de surfaces planes avec connaissance de la pose de la caméra par la méthode de Tamaazousti et al. (2011b) et plus récemment de Loesch et al. (2015). Sur ces surfaces, nous avons observé que les spécularités conservaient une forme elliptique. De plus, une source de lumière ne produit qu'une seule spécularité sur une surface plane ce qui permet un meilleur suivi de la spécularité tout le long de la vidéo. De ce constat, une question subsiste : sachant qu'une spécularité représente la réflexion d'une lumière fixe, existe-t-il une forme 3D fixe dans l'espace telle que la projection de cette forme 3D corresponde aux spécularités ? Pour répondre à ce problème, nous introduisons les contributions suivantes :

- Un détecteur de spécularité temps réel à partir d'une image (présenté en annexe A)
- JOLIMAS : un modèle géométrique représentant la spécularité comme étant la projection perspective d'un objet 3D dans l'espace sur des surfaces planes gérant les sources de lumière ponctuelles (ampoules) mais aussi étendues (néons).
- Un processus de prédiction de spécularité pour un nouveau point de vue.

Ces contributions ont fait l'objet de deux publications en conférence internationale : VISAPP (2014), ISMAR (2017a), deux conférences nationales : deux articles à ORASIS (2015a) et (2015b), RFIA (2015) et un brevet accepté en juin (2018).

Par la suite, dans le chapitre 5, nous abordons l'extension du modèle de prédiction de spécularité JOLIMAS dans un contexte plus varié tel que les surfaces courbes et plus précisément les surfaces convexes à courbure constante. Nous avons fait le constat que les spécularités conservaient une forme elliptique sur les surfaces convexes. De plus, pour ces surfaces, une source de lumière ne produit également qu'une seule spécularité. Cependant, le comportement de ces spécularités sur ces surfaces met en échec le modèle JOLIMAS initial que nous appelons primal, car la forme et la position de la spécularité sont grandement affectées par la courbure de la surface. En étudiant la géométrie épipolaire du système, nous observons que l'objet observé (source de lumière) est en mouvement dans le formalisme utilisé par notre modèle JOLIMAS primal alors que cette source de lumière est fixe. De ce fait, il n'est pas possible en l'état de reconstruire une quadrique telle que la projection perspective corresponde au contour de la spécularité pour un nouveau point de vue. Dans JOLIMAS primal, adapté aux surfaces planes, nous utilisions les caméras réelles afin de reconstruire la quadrique utilisée pour prédire la forme de la spécularité. Nous avons constaté que cette quadrique est localisée à l'emplacement du symétrique de la source de lumière. Ce raisonnement est contre-intuitif pour généraliser le modèle, car la source de lumière est fixe dans l'espace. Il est plus intéressant d'utiliser une représentation de caméra virtuelle calculée à partir des poses de caméras réelles qui reconstruit la quadrique autour de la lumière fixe. Pour répondre à ce problème, nous introduisons les contributions suivantes :

- Formalisme d'un JOLIMAS dual en considérant une représentation de caméra virtuelle.
- Reconstruction d'une quadrique 3D pour chaque surface convexe. Généralisation du

modèle pour un maillage plus complexe en décomposant celui-ci en un ensemble de surfaces convexes.
- Prédiction de spécularité adaptée aux surfaces convexes.

Ce chapitre a fait l'objet d'une publication internationale : ISMAR (2017d), d'une publication nationale ORASIS (2017) et d'un journal international TVCG (2017b).

Dans le chapitre 6, nous étendons à nouveau le modèle JOLIMAS dans sa forme duale afin de le généraliser à toute surface de courbure quelconque (convexe, concave et autres). Plus précisément, la courbure de l'objet d'intérêt n'est pas contrainte à rester constante dans le temps comme précédemment dans JOLIMAS dual ce qui permet de gérer des objets plus complexes (ellipsoïde, Tore par exemple ainsi que des objets déformables). Dans cette extension, nous avons étudié le lien entre modification de la courbure d'une surface et modification de la forme d'une spécularité. En effet, si une surface n'a pas une courbure constante (au sens de Gauss), JOLIMAS dual ne pourra pas prédire le changement affectant les contours de la spécularité. Pour répondre à ce problème, nous introduisons les contributions suivantes :
- Formalisation de la notion d'angle limite, qui est une propriété des contours de spécularité pour une pose de caméra donnée et une position de lumière donnée qui reste constante pour tout changement de courbure.
- Reconstruction d'un modèle géométrique canonique invariant aux changements de courbure. Nous reconstruisons notre quadrique à partir des ellipses correspondant aux contours de spécularités transformées afin de revenir à une surface plane (JOLIMAS primal).
- Prédiction de spécularité en déformant la projection perspective de la quadrique afin de correspondre à la courbure de la scène.

Ce chapitre a fait l'objet d'une publication internationale : VRST (2017e) et d'une soumission dans un journal international TVCG 2018.

Dans le chapitre 7, nous résumons l'ensemble des contributions actuelles et potentielles en RA de notre modèle de prédiction. Dans un premier temps, nous détaillons le pipeline complet de notre application de *retexturing* et les différentes techniques utilisées permettant de changer une texture pour une autre. Ce retexturing est réalisé en simulant synthétiquement la spécularité supprimée pour augmenter le réalisme de la texture ajoutée. Dans un second temps, nous détaillons les techniques utilisées afin d'assurer une meilleure flexibilité de nos applications comme un processus de suivi de spécularité permettant une prédiction de spécularité dans un environnement multi lumières. Nous présentons également quelques résultats non publiés de changements synthétiques de rugosité et de source lumineuse.

CHAPITRE 2

Notions de base

Dans ce chapitre, nous introduisons les notions de base et les notations nécessaires à la compréhension du mémoire. Après avoir présenté les caméras perspectives et la géométrie associée, nous rappellerons le formalisme des coniques et quadriques. Par la suite, nous donnerons quelques détails pour décrire l'interaction lumière/matière, l'illumination locale, globale ainsi qu'un rappel de notion de courbure d'une surface. Nous terminerons par un rappel des plus importantes techniques d'optimisation utilisées.

2.1 Notations

Dans ce mémoire, nous utilisons des lettres minuscules en gras pour représenter les vecteurs et des lettres majuscules en gras pour représenter les matrices. Les scalaires sont représentés par des lettres minuscules.

Mathématiques

$\mathtt{M}$	Matrice
$\mathbf{v}$	Vecteur
$\hat{\mathbf{v}}$	Vecteur normalisé
$\mathcal{F}$	Fonction mathématique

Géométrie euclidienne et projective

$\sim$	Égalité à un facteur non-nul près
$\mathbf{p}$	Point 2D
$\tilde{\mathbf{p}}$	Coordonnées homogènes du point 2D
$\mathbf{P}$	Point 3D

$\tilde{\mathbf{P}}$	Coordonnées homogènes du point 3D
$\mathbf{d}$	Droite de l'espace tel que $\mathbf{d} = \begin{pmatrix} a & b & c \end{pmatrix}^\top$ est sa représentation en coordonnées homogènes
$\boldsymbol{\pi}$	Plan de l'espace
$\boldsymbol{\pi}_{\mathbf{P}}$	Plan tangent au point $\mathbf{P}$
$\hat{\mathbf{N}}(\mathbf{P})$	Normale au point $\mathbf{P}$
S	Surface étudiée
$\mathtt{Q}$	Quadrique
$\mathtt{C}$	Conique

Caméras et reconstruction 3D

$\mathcal{C}$	Caméra
I	Image capturée par la caméra
$\mathtt{P}$	Matrice de projection
$\mathtt{K}$	Matrice de paramètre intrinsèque de la caméra
$\mathtt{F}$	Matrice fondamentale
$\mathtt{E}$	Matrice essentielle
$\mathtt{R}$	Rotation de la caméra
$\mathbf{V}$	Position de la caméra
$\mathtt{\Pi}$	Pose de la caméra
$\tilde{\mathtt{\Pi}}$	Pose de la caméra en représentation virtuelle
$\pi(\tilde{\mathbf{p}})$	Fonction de *déhomogénisation* tel que $\pi(\tilde{\mathbf{p}}) = \mathbf{p}$

Acronymes

BRDF	Bidirectionnal Reflectance Distribution Function
CAO	Conception Assistée par Ordinateur
DoF	Degree of Freedom (degré de liberté)
JOLIMAS	JOint LIght-MAterial Specularity
RA	Réalité augmentée
RD	Réalité diminuée
SfM	Structure-from-Motion
SLAM	Simultaneous Localization and Mapping
SVD	Singular value decomposition (décomposition en valeurs singulières)

Algèbre linéaire

x	Scalaire

$\mathbb{R}^n$	Espace vectoriel de dimension n
$\mathbf{x}$	Vecteur
$x_1 \dots$	Coordonnées
$\|.\|$	Norme euclidienne L^2
$\|.\|$	Norme L^1
$\wedge$	Produit vectoriel
$\mathtt{M}$	Matrice
$M_{11} \dots$	Coefficients
$\mathtt{M}^\top$	Transposée
$\mathtt{M}^{-1}$	Inverse
$\mathbf{0}_{m\times n}$	Matrice nulle de dimensions $m \times n$
$\mathtt{I}$	Matrice identité
$\otimes$	Produit de Kronecker
.	Produit scalaire
vec	Opération de vectorisation de matrice
vech	Opération de vectorisation de matrice symétriques
μ	Opération de normalisation tel que $\mu(\mathbf{U}) \stackrel{\text{def}}{=} \frac{\mathbf{U}}{\|\mathbf{U}\|}$
$\mathtt{M}^*$	Matrice adjointe de $\mathtt{M}$
$\mathbf{M}_v$	Matrice vectorisée
$[\mathbf{t}]_\times$	Matrice antisymétrique créée à partir du vecteur $\mathbf{t}$
$\mathtt{M}^+$	Pseudo-inverse de la matrice $\mathtt{M}$

Images et éléments d'illumination

I	Image
$I(\mathbf{p})$	Intensité du pixel $\mathbf{q}$
I_s	Composante spéculaire de l'image
I_d	Composante diffuse de l'image
I_a	Composante ambiante de l'image
i_s	Intensité spéculaire incidente
i_d	Intensité diffuse incidente
i_a	Intensité ambiante incidente
k_s	Ratio de réflexion spéculaire
k_d	Ratio de réflexion diffuse
k_a	Ratio de réflexion ambiante
τ	Seuil d'intensité
$\mathbf{L}$	Position de la source de lumière ponctuelle
$\mathbf{R}$	Position du symétrique de la source de lumière par rapport à la surface S

2.2 Espace projectif et géométrie associée

2.2.1 Géométrie projective

Sur l'espace vectoriel $\mathbb{R}^{n+1}$, il est possible de définir la relation d'équivalence suivante :

$$\mathbf{u} \sim \mathbf{v} \Leftrightarrow \exists \lambda \in \mathbb{R}^* \mid \mathbf{u} = \lambda \mathbf{v}. \tag{2.1}$$

L'ensemble des classes d'équivalence de $\mathbb{R}^{n+1}$ pour cette relation « $\sim$ » définit un espace appelé *espace projectif*. Cet espace, de dimension n, sera noté $\mathbb{P}^n$. Si des études théoriques de ces espaces existent, nous nous intéresserons dans nos travaux à la *géométrie projective* qui leur est associée et qui permet en particulier de formaliser la notion de point à l'infini dans les espaces affines.

Un vecteur de l'espace projectif $\mathbb{P}^n$ aura pour coordonnées :

$$\mathbf{h} = (h_1 \ \dots \ h_{n+1})^\mathsf{T} \tag{2.2}$$

avec les h_i non tous nuls. Si h_{n+1} est non nul, ce vecteur $\mathbf{h}$ représente le vecteur $\mathbf{x}$ de $\mathbb{R}^n$ avec $\mathbf{x} = (h_1/h_{n+1} \ \dots \ h_n/h_{n+1})^\mathsf{T}$. Dans le cas contraire, le vecteur $\mathbf{h}$ décrit un point à l'infini. Les coordonnées de $\mathbf{h}$ sont appelées *coordonnées homogènes* de $\mathbf{x}$. La représentation canonique d'un vecteur de $\mathbb{R}^n$ en coordonnées homogènes est notée avec un point et définie comme suit :

$$\forall\, \mathbf{x} = (x_1 \ \dots \ x_n)^\mathsf{T} \in \mathbb{R}^n : \qquad \tilde{\mathbf{x}} = (x_1 \ \dots \ x_n, 1)^\mathsf{T} \in \mathbb{P}^n. \tag{2.3}$$

Nous appellerons π la fonction de *déhomogénisation* permettant de passer des coordonnées homogènes aux coordonnées euclidiennes, à savoir :

$$\begin{array}{rccc} \pi : & \mathbb{P}^n & \to & \mathbb{R}^n \\ & (x_1 \ \dots \ x_{n+1})^\mathsf{T} & \mapsto & (x_1/x_{n+1} \ \dots \ x_n/x_{n+1})^\mathsf{T}. \end{array} \tag{2.4}$$

En particulier $\forall \mathbf{x} \in \mathbb{R}^n : \pi(\tilde{\mathbf{x}}) = \mathbf{x}$. Nous nous intéressons maintenant aux cas particuliers de cette géométrie en deux puis trois dimensions.

2.2.2 Le plan projectif

L'espace projectif de dimension 2 est appelé *plan projectif*. Un point de $\mathbb{P}^2$ est représenté par un vecteur de dimension 3 : $\mathbf{h} = (x \ y \ w)^\mathsf{T}$. De même, une droite d'équation $ax + by + c = 0$ peut être représentée par le vecteur $\mathbf{l} = (a \ b \ c)^\mathsf{T}$. Cette notation homogène permet de définir simplement la notion d'appartenance du point $\mathbf{h}$ à la droite $\mathbf{l}$, à savoir :

$$\mathbf{l}^\mathsf{T}\mathbf{h} = 0. \tag{2.5}$$

L'équation de la droite $\mathbf{l}$ passant par les points $\mathbf{h}_1$ et $\mathbf{h}_2$ est obtenue en calculant leur produit vectoriel :

$$\mathbf{l} = \mathbf{h}_1 \wedge \mathbf{h}_2. \tag{2.6}$$

Dans l'espace $\mathbb{P}^2$, droites et points jouent un rôle équivalent : c'est ce qu'on appelle le *principe de dualité*. En particulier, à partir de l'équation duale de l'équation précédente,

il est possible de calculer le point d'intersection de deux droites $\mathbf{l}_1$ et $\mathbf{l}_2$:

$$\mathbf{h} = \mathbf{l}_1 \wedge \mathbf{l}_2. \tag{2.7}$$

2.2.3 Espace projectif

Un point $\mathbf{P}$ de $\mathbb{R}^3$ aura pour coordonnées homogènes dans $\mathbb{P}^3$ le vecteur $\mathbf{v} = (X\ Y\ Z\ 1)^\mathsf{T}$. Dans cet espace de dimension 3, le dual du point $\mathbf{P}$ est le plan $\mathbf{\Pi}$ d'équation $aX + bY + cZ + d = 0$ qui est représenté par le vecteur $\mathbf{\Pi} = (a\ b\ c\ d)^\mathsf{T}$. L'appartenance du point $\mathbf{P}$ au plan $\mathbf{\Pi}$ est alors donnée par la relation :

$$\mathbf{\Pi}^\mathsf{T}\mathbf{v} = 0. \tag{2.8}$$

2.3 Caméras perspectives

Dans le cadre de nos travaux, nous utiliserons des caméras perspectives respectant le modèle des *caméras sténopés* idéales. Ce modèle considère que l'ensemble des rayons lumineux passent par un seul et unique point avant d'atteindre le capteur (voir figure 2.1).

Dans la suite, nous présenterons tout d'abord les différents paramètres caractérisant ce type de caméras. Nous présenterons alors la géométrie reliant les images observées par plusieurs caméras. Enfin, nous étudierons les méthodes permettant de retrouver le déplacement de ces caméras ainsi que la structure de l'environnement à partir des images qu'elles observent.

2.3.1 Projection perspective

La projection perspective vise à calculer, pour tout point $\mathbf{P} \in \mathbb{R}^3$, la position 2D $\mathbf{p}$ de sa projection dans l'image. Basée sur la projection centrale, cette transformation consiste à calculer l'intersection du plan de la *rétine* de la caméra (*i.e.* le capteur) avec le *rayon de projection* de $\mathbf{P}$. Ce dernier est défini comme étant la droite reliant le point $\mathbf{P}$ au centre de la caméra (figure 2.1).

Cette projection peut être vue comme un enchaînement de trois transformations géométriques (figure 2.1) :

- La première transformation est un changement de repère qui consiste à exprimer les coordonnées de $\mathbf{P}$ dans le repère lié à la caméra. Ce changement de repère est défini par les *paramètres extrinsèques* de la caméra. Dans la suite du mémoire, en cas d'ambiguïté, les exposants $\mathcal{W}$ ou $\mathcal{C}$ indiqueront le repère dans lequel les coordonnées du point 3D sont exprimées (respectivement monde ou caméra).

- La deuxième transformation est la *projection centrale* du point 3D. Elle revient à passer du point 3D (exprimé dans le repère caméra) au point d'intersection du rayon de projection et du capteur. Les coordonnées 2D du point résultant sont alors exprimées dans le plan de la rétine (en mm).

- La troisième transformation est un changement de repère 2D qui vise à passer du repère rétine (repère lié à la physique du capteur et où les coordonnées sont exprimées

en mm) au repère *image* de la caméra (repère géométrique où les coordonnées sont exprimées en pixels). Cette transformation est définie par les *paramètres intrinsèques* de la caméra.

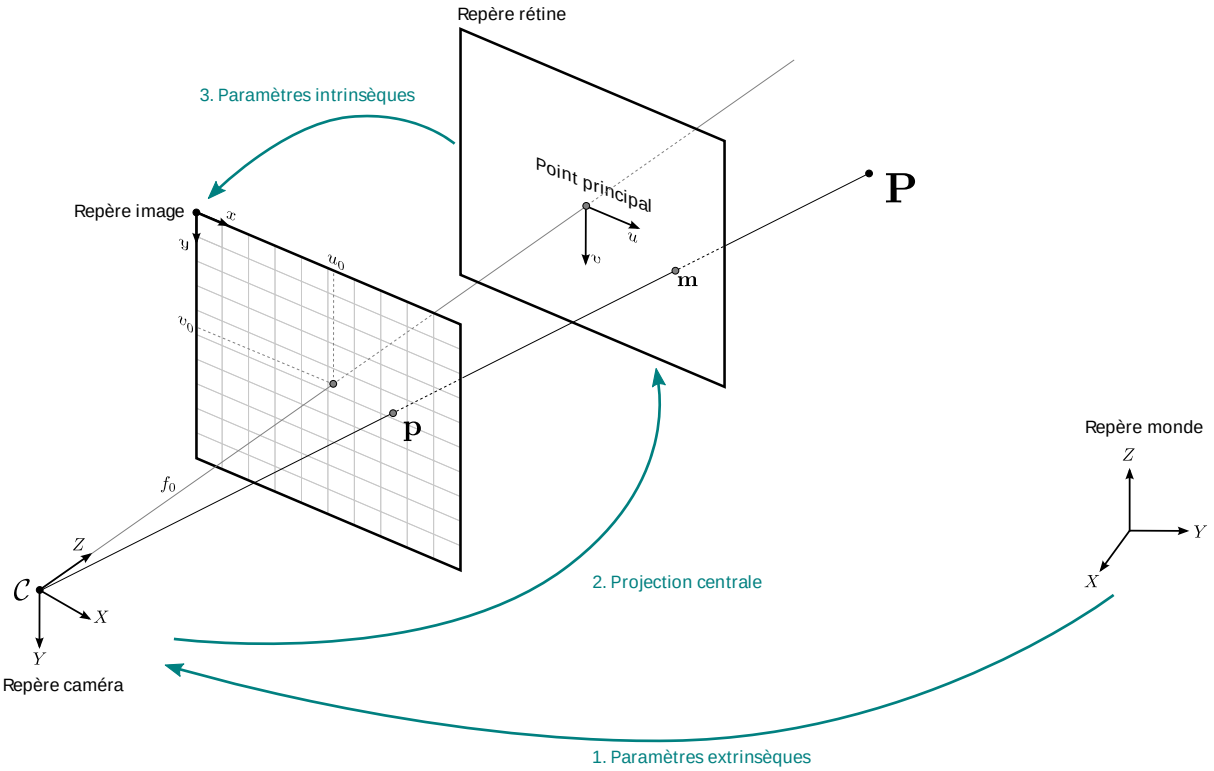


FIGURE 2.1 – **Projection perspective.** La projection perspective peut être vue comme trois transformations géométriques consécutives pour les points 3D.

La projection perspective est donc une transformation projective de $\mathbb{P}^3 \to \mathbb{P}^2$. En pratique, elle sera représentée par une *matrice de projection* $\mathtt{P}$ de dimension (3×4). La projection perspective s'exprime alors par la relation matricielle suivante :

$$\tilde{\mathbf{p}} \sim \mathtt{P}\tilde{\mathbf{P}}_{\mathcal{W}}. \tag{2.9}$$

La matrice $\mathtt{P}$ se décompose selon les trois transformations citées précédemment :

$$\mathtt{P} = \mathsf{K} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathtt{R}^\mathsf{T} & -\mathtt{R}^\mathsf{T}\mathbf{t} \\ 0_{1\times 3} & 1 \end{pmatrix}, \tag{2.10}$$

où K est la *matrice de calibrage* (de taille 3×3) de la caméra, $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ la matrice de *projection centrale* et $\begin{pmatrix} \mathtt{R}^\mathsf{T} & -\mathtt{R}^\mathsf{T}\mathbf{t} \\ 0_{1\times 3} & 1 \end{pmatrix}$ la *matrice de pose*.

Paramètres extrinsèques Les paramètres extrinsèques d'une caméra caractérisent la pose de celle-ci dans le repère monde. La pose d'une caméra possède six degrés de liberté :

- La position 3D du centre optique, décrit par le vecteur $\mathbf{t} = ((\mathbf{t})_x \ (\mathbf{t})_y \ (\mathbf{t})_z)^\mathsf{T}$.
- L'orientation 3D de la caméra. En pratique, cette orientation sera représentée sous la forme d'une matrice de rotation $\mathtt{R}$, cette matrice pouvant être obtenue à partir des trois angles roulis, tangage et lacet $\mathbf{r} = (\alpha \ \beta \ \gamma)^\mathsf{T}$ (voir figure 2.2).

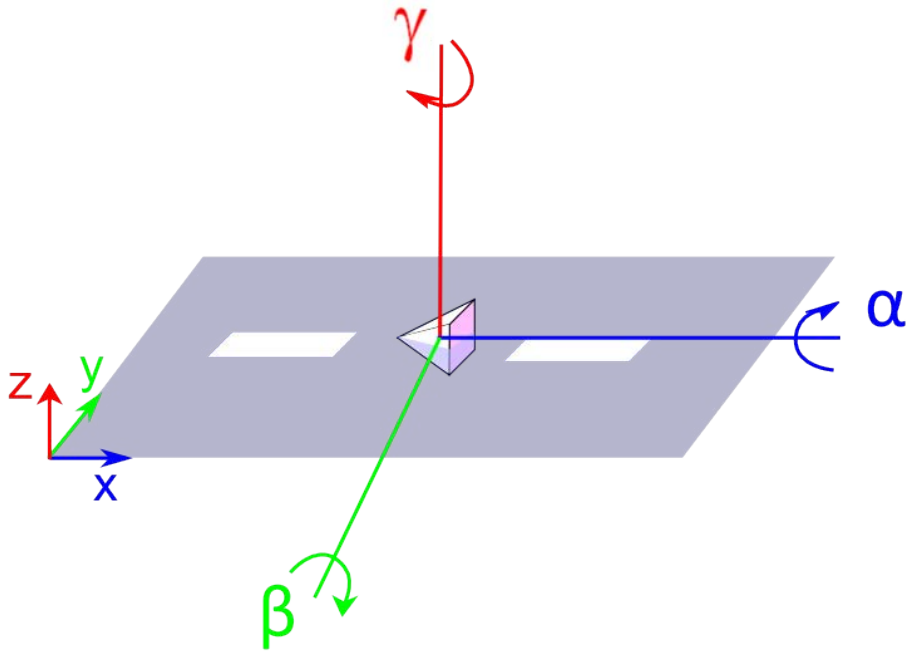


FIGURE 2.2 – **Définition des angles : roulis, tangage et lacet.**

Les paramètres extrinsèques de la caméra permettent d'établir les changements de repère monde/caméra, à savoir :

$$\tilde{\mathbf{P}}_\mathcal{C} \sim \begin{pmatrix} \mathtt{R}^\mathsf{T} & -\mathtt{R}^\mathsf{T}\mathbf{t} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_\mathcal{W} \tag{2.11}$$

$$\tilde{\mathbf{P}}_\mathcal{W} \sim \begin{pmatrix} \mathtt{R} & \mathbf{t} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_\mathcal{C} \tag{2.12}$$

Projection centrale. Lorsqu'on utilise les coordonnées homogènes, la projection centrale d'un point $\mathbf{P}_\mathcal{C}$ est une fonction linéaire de $\mathbb{P}^3 \mapsto \mathbb{P}^2$ caractérisée par la matrice de dimension 3×4 :

$$\tilde{\mathbf{p}} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tilde{\mathbf{P}}_\mathcal{C}. \tag{2.13}$$

Dans de nombreuses publications, le changement de repère 3D et la projection centrale sont vus comme une unique projection centrale à partir d'un point 3D dans le repère

monde. Cela s'écrit sous forme matricielle :

$$\begin{aligned}\tilde{\mathbf{p}} &\sim \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\mathtt{R}^\mathsf{T} & -\mathtt{R}^\mathsf{T}\mathbf{t}\\ 0_{1\times 3} & 1\end{pmatrix}\tilde{\mathbf{P}}_{\mathcal{W}} \\ &\sim \begin{pmatrix}\mathtt{R}^\mathsf{T} & -\mathtt{R}^\mathsf{T}\mathbf{t}\end{pmatrix}\tilde{\mathbf{P}}_{\mathcal{W}}.\end{aligned} \tag{2.14}$$

La projection d'un point exprimé dans le repère monde dans l'image (équation 2.9) s'écrit donc généralement sous cette forme :

$$\tilde{\mathbf{p}} \sim \mathsf{K}\begin{pmatrix}\mathtt{R}^\mathsf{T} & -\mathtt{R}^\mathsf{T}\mathbf{t}\end{pmatrix}\tilde{\mathbf{P}}_{\mathcal{W}}. \tag{2.15}$$

Paramètres intrinsèques et distorsion. Les paramètres intrinsèques définissent les propriétés géométriques de l'image de la caméra. Dans notre étude, nous considérons que les pixels sont carrés. La *matrice de calibrage* K peut alors s'exprimer sous la forme :

$$\mathsf{K} = \begin{pmatrix}f_0 & 0 & u_0\\ 0 & f_0 & v_0\\ 0 & 0 & 1\end{pmatrix}. \tag{2.16}$$

Nous retrouvons dans la matrice de calibrage les différents paramètres intrinsèques, à savoir :

- f_0 la *distance focale*. Exprimée en pixels par unité de mesure, elle décrit la distance orthogonale entre le centre et la rétine de la caméra.
- $(u_0\ v_0)^\mathsf{T}$ le *point principal*. Souvent approximé comme étant le centre du capteur, il est plus précisément l'intersection entre l'axe optique et la rétine de la caméra (figure 2.1).

Il est important de noter que les capteurs à courte focale peuvent présenter un phénomène de distorsion important. Ceci se traduit visuellement par une déformation des lignes droites dans l'image sous forme de courbes. Pour corriger cela, il est possible d'ajouter au calibrage de la caméra des paramètres de distorsion permettant de passer de la position observée d'un point 2D dans l'image à sa position réelle, c'est-à-dire corrigée de toute distorsion.

Dans le cadre de ce mémoire, nous considérerons à la fois que la matrice de calibrage est connue et que les entrées de nos algorithmes ont été préalablement corrigées en distorsion. En pratique, la distorsion radiale est modélisée en utilisant 5 coefficients. La distorsion tangentielle étant beaucoup plus faible, elle sera négligée dans nos travaux. Pour plus de renseignements sur le calibrage des caméras, nous invitons le lecteur à se référer à l'article de Lavest et al. (1998).

2.3.2 Rétroprojection

La *rétroprojection* peut être vue comme l'opération inverse de la projection. Son but est de donner le rayon de vue afin de contraindre la position d'un point 3D $\mathbf{P}$ à partir de son observation $\mathbf{p}$ dans l'image. Néanmoins, à partir d'une seule image, il est impossible d'obtenir la position exacte du point 3D. En effet, l'utilisation d'une seule caméra ne permet pas de retrouver la profondeur à laquelle se situe ce point. La rétroprojection d'un

point de l'image se traduit sous la forme du rayon optique qui passe à la fois par le centre de la caméra $\mathcal{C}$ et par l'observation $\mathbf{p}$. La position du point 3D est donc exprimée à un facteur λ près qui reflète la profondeur du point sur ce rayon :

$$\mathbf{P}(\lambda) = \lambda \mathtt{P}^{+} \tilde{\mathbf{p}} \tag{2.17}$$

où $\mathtt{P}^{+}$ désigne la pseudo-inverse de la matrice $\mathtt{P}$:

$$\mathtt{P}^{+} \;=\; \mathtt{P}^{\mathsf{T}}(\mathtt{P}\mathtt{P}^{\mathsf{T}})^{-1} \tag{2.18}$$

de telle sorte que $\mathtt{P}\mathbf{P}(\lambda) = \lambda(\mathtt{P}\mathtt{P}^{+})\tilde{\mathbf{p}} = \lambda\tilde{\mathbf{p}}$.

2.4 Image

2.4.1 Définition et notations

En matière d'informatique, une image est un tableau à deux dimensions contenant un nombre fini de pixels. Sauf mention contraire, nous considérons des images en niveau de gris à valeurs réelles. Une image I de dimensions $w \times h$ est donc une application bidimensionnelle :

$$\begin{aligned} I : [\![1, w]\!] \times [\![1, h]\!] &\to \mathbb{N} \\ (x, y) &\mapsto I[x, y]. \end{aligned} \tag{2.19}$$

Pour simplifier les expressions mathématiques, il est possible de considérer I définie sur le domaine continu :

$$\Omega_I = [1, w] \times [1, h] \tag{2.20}$$

en interpolant les valeurs entre les pixels voisins pour les coordonnées non entières. On a alors :

$$\begin{aligned} I : \quad \Omega_I &\to \mathbb{N} \\ (x, y) &\mapsto I(x, y). \end{aligned} \tag{2.21}$$

2.4.2 Opérations

Les opérations suivantes sur les images seront utiles par la suite :

Homographie. Une homographie décrit une transformation linéaire dans l'espace projectif. L'homographie $H_I^{\mathtt{H}}$ de l'image I décrite par la matrice $\mathtt{H} \in \mathbb{R}^{3\times3}$ est définie par :

$$\forall \mathbf{p} \in \mathbb{R}^2 \quad \text{t.q.} \quad \pi(\mathtt{H}\tilde{\mathbf{q}}) \in \Omega_I \quad : \quad H_I^{\mathtt{H}}(\mathbf{p}) = I\left(\pi(\mathtt{H}\tilde{\mathbf{p}})\right). \tag{2.22}$$

Deux observations d'une même surface rigide et plane sont toujours liées par une homographie.

Translation. La translation de l'image I de vecteur $\mathbf{t} \in \mathbb{R}^2$ est notée $T_I^{\mathbf{t}}$ et définie telle que :

$$\forall \mathbf{p} \quad \text{t.q.} \quad (\mathbf{p} + \mathbf{t}) \in \Omega_I \quad : \quad T_I^{\mathbf{t}}(\mathbf{p}) = I(\mathbf{p} + \mathbf{t}). \tag{2.23}$$

Convolution. La convolution, notée $*$ est une opération permettant d'appliquer un filtre linéaire à une image. En coordonnées discrètes, une convolution sur l'image I est définie par une image W de domaine Ω_W, appelée *masque* ou *noyau* :

$$\forall W : [\![-m, m]\!] \times [\![-n, n]\!] \to \mathbb{R}$$
$$(I * W)[x, y] = \sum_{i=-m}^{m} \sum_{j=-n}^{n} W[i, j] I[x - i, y - j].$$

En coordonnées réelles, la définition est similaire :

$$\forall W : \Omega_W \to \mathbb{R} \qquad (I * W)(x, y) = \iint_{(x',y') \in \Omega_W} W(x', y') I(x - x', y - y') dx dy. \tag{2.24}$$

Différentiation. Le gradient de l'image I est défini et noté comme suit :

$$\forall (x, y) \in \Omega_I \quad : \quad \nabla I(x, y) = \begin{pmatrix} \frac{\partial I(x,y)}{\partial x} \\ \frac{\partial I(x,y)}{\partial y} \end{pmatrix}. \tag{2.25}$$

Le laplacien est la somme des dérivées partielles au second ordre :

$$\forall (x, y) \in \Omega_I \quad : \quad \nabla^2 I(x, y) = \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2}. \tag{2.26}$$

2.5 Conique et quadrique

2.5.1 Conique

Une conique est une famille de courbes planes générées en prenant toutes les sections possibles d'un cône infini avec un plan. La famille des coniques non dégénérées inclut les paraboles, hyperboles et ellipses. Nous évoquerons les coniques dégénérées par la suite. Chaque point $(x, y)^\top$ de la conique respecte l'équation :

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \tag{2.27}$$

avec a, b, c, d, e et f des paramètres (scalaires) constants. L'équation (2.27) s'exprime en remplaçant $x = \frac{x_1}{x_3}$ et $y = \frac{x_2}{x_3}$ ce qui donne :

$$ax_1^2 + 2bx_1x_2 + cx_2^2 + 2dx_1x_3 + 2ex_2x_3 + fx_3^2 = 0, \tag{2.28}$$

ou sous forme matricielle :

$$\mathbf{x}^\top \mathtt{C} \mathbf{x} = 0, \tag{2.29}$$

avec les coefficients de la matrice $\mathtt{C}$ donnée par :

$$\begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} \tag{2.30}$$

Plusieurs types de coniques sont illustrés à la figure 2.3.

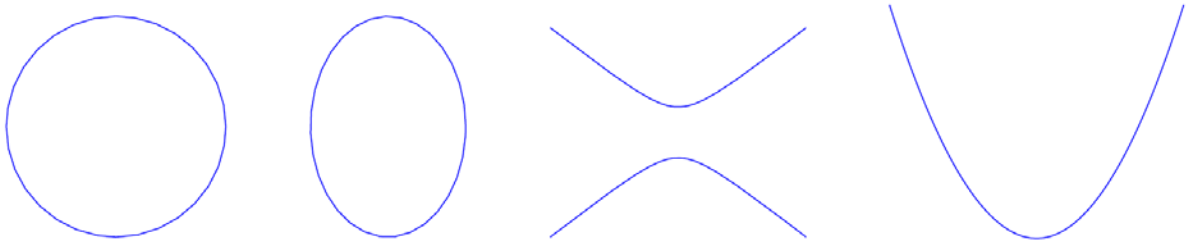


FIGURE 2.3 – Exemple de coniques. De gauche à droite, nous avons un cercle, une ellipse, une hyperbole et une parabole

Nous pouvons noter que $\mathtt{C}$ est une matrice symétrique. De plus, nous pouvons observer qu'une conique a cinq degrés de liberté (six paramètres à une échelle près).

2.5.1.1 Transformation projective

Considérons une transformation projective 2D $\mathtt{H}$ transformant un point $\mathbf{x}$ en $\mathbf{x}'$ tel que $\mathbf{x}' = \mathtt{H}\mathbf{x}$. En assumant que la transformation est non-singulière et en substituant $\mathbf{x} = \mathtt{H}^{-1}\mathbf{x}'$ dans l'équation (2.29), nous avons :

$$\mathbf{x}'^{\top}\mathtt{H}^{-\top}\mathtt{C}\mathtt{H}^{-1}\mathbf{x}' = 0. \tag{2.31}$$

Puisque la matrice $\mathtt{H}^{-\top}\mathtt{C}\mathtt{H}^{-1}$ est symétrique (comme à l'équation (2.29)) l'équation (2.31) définit une nouvelle conique, $\mathtt{C}' = \mathtt{H}^{-\top}\mathtt{C}\mathtt{H}^{-1}$. En général, une transformation projective d'une conique donne également une conique.

2.5.1.2 Droite tangente à une conique

La droite $\mathbf{l}$ tangente à la conique $\mathtt{C}$ au point $\mathbf{x}$ (appartenant à la conique) est donnée par :

$$\mathbf{l} = \mathtt{C}\mathbf{x}. \tag{2.32}$$

Cette formule est justifiée par le fait que la ligne $\mathbf{l}$ passe par $\mathbf{x}$ car $\mathbf{l}^{\top}\mathbf{x} = \mathbf{x}^{\top}\mathtt{C}\mathbf{x} = 0$ si $\mathbf{x}$ appartient à la conique.

2.5.1.3 Enveloppe de la conique

Vu qu'il existe une droite tangente pour chaque point de la conique, il semble logique que la conique soit définie par les droites tangentes. Ce constat est un résultat standard en géométrie projective : *Pour tout théorème de géométrie projective en 2 dimensions, il existe un théorème dual qui est décrit par inter-changer les rôles des points et des lignes du théorème originel.*

Pour une conique non dégénérée, l'équation (2.32) peut être inversée afin de trouver le point où la droite $\mathbf{l}$ est tangente à la conique $\mathtt{C}$: $\mathbf{x} = \mathtt{C}^{-1}\mathbf{l}$. Par injection dans l'équation (2.29), nous avons :

$$(\mathtt{C}^{-1}\mathbf{l})^{\top}\mathtt{C}(\mathtt{C}^{-1}\mathbf{l}) = \mathbf{l}^{\top}\mathtt{C}^{-1}\mathbf{l} = 0, \tag{2.33}$$

avec $\mathtt{C}$ (réciproquement $\mathtt{C}^{-1}$) est symétrique. La conique duale $\mathtt{C}^{-1}$ représente l'enveloppe de la conique et est directement associée à la conique $\mathtt{C}$.

Plus généralement, nous avons la formule :

$$\mathbf{l}^\top \mathtt{C}^* \mathbf{l} = 0, \tag{2.34}$$

où $\mathtt{C}^*$ dénote l'adjoint de $\mathtt{C}$. Pour une matrice symétrique non singulière nous avons $\mathtt{C}^* = \mathtt{C}^{-1}$ (à une échelle près).

2.5.1.4 Coniques dégénérées

Considérons la conique $\mathbf{x}^\top \mathtt{C} \mathbf{x} = 0$ avec $\mathtt{C}$ une matrice 3×3 singulière. Ces coniques sont définies comme *dégénérées* et sont souvent considérées comme des cas spécifiques.

Si le rang de $\mathtt{C}$ est de 2, elle possède un espace nul de dimension 1. Ainsi, il existe un point unique (en coordonnées homogènes), $\mathbf{n}$ pour lequel $\mathtt{C}\mathbf{n} = 0$. En ce point, la tangente de la conique par l'équation (2.32) est non définie. Cette conique particulière est constituée d'une paire de lignes distinctes, $\mathbf{n}$ étant le point d'intersection des lignes ($\mathbf{n}$ n'est pas fini dans le cas de deux droites parallèles). Logiquement, une conique dégénérée de rang 2 a 4 degrés de liberté ce qui correspond à 2 degrés de liberté pour chaque droite.

Une conique dégénérée est définie par :

$$\mathtt{C} = \mathbf{l}\mathbf{m}^\top + \mathbf{m}\mathbf{l}^\top, \tag{2.35}$$

avec $\mathbf{l}$ et $\mathbf{m}$ deux droites. On peut retrouver cette relation en observant que tous les points sur $\mathbf{l}^\top \mathbf{x} = 0$ sont sur la conique $\mathbf{x}^\top \mathtt{C} \mathbf{x} = (\mathbf{x}^\top \mathbf{l})(\mathbf{m}^\top \mathbf{x}) + (\mathbf{x}^\top \mathbf{m})(\mathbf{l}^\top \mathbf{x}) = 0$. Tous les points sur $\mathbf{m}^\top \mathbf{x} = 0$ sont aussi sur la conique. La matrice $\mathtt{C}$ est symétrique et de rang 2 (dans le cas où $\mathbf{l}$ et $\mathbf{m}$ ne seraient pas confondus). Dans le cas ultérieur, $\mathtt{C}$ sera de rang 1. Notons qu'il existe d'autres coniques dégénérées, mais leurs connaissances ne sont pas nécessaires pour la bonne compréhension de nos travaux.

2.5.1.5 Enveloppes de coniques dégénérées

L'équation (2.34) définit l'enveloppe de la conique $\mathtt{C}^*$ pour une conique $\mathtt{C}$ qui n'est pas forcément de rang plein. Une démonstration complète peut être trouvée dans Cross (2000) que l'adjoint de la matrice 3×3 est non nul si le rang de la matrice est plus grand que 1. La conséquence de ce résultat est que l'enveloppe d'une conique de rang 2 (paire de droites distinctes) est définie, à l'opposé de l'enveloppe d'une conique de rang 1 (paire de droites confondues) n'est pas définie.

2.5.2 Quadriques

Un point 3D $(x, y, z)^\top$ sur la surface d'une quadrique suit l'équation :

$$ax^2 + 2bxy + cy^2 + 2dxz + 2eyz + fz^2 + 2gx + 2hy + 2iz + j = 0, \tag{2.36}$$

ou en coordonnées homogènes $(x, y, z) \to (x_1/x_4, x_2/x_4, x_3/x_4)$

$$ax_1^2 = 2bx_1x_2 + cx_2^2 + 2dx_1x_3 + 2ex_2x_3 + fx_3^2 + 2gx_1x_4 + 2hx_2x_4 + 2ix_3x_4 + jx_4^2 = 0. \tag{2.37}$$

La définition équivalente en notation matricielle pour une quadrique 3D est décrite par :

$$\mathbf{X}^\top \mathtt{Q} \mathbf{X} = 0, \tag{2.38}$$

avec la matrice de la quadrique :

$$\mathtt{Q} = \begin{bmatrix} a & b & d & g \\ b & c & e & h \\ d & e & f & i \\ g & h & i & j \end{bmatrix} \tag{2.39}$$

Une quadrique possède 9 degrés de liberté et est décrite par une matrice 4×4 de 10 éléments distincts.

Les quadriques incluent des formes comme les sphères, les ellipsoïdes, les hyperboloïdes et paraboloïdes. Plusieurs types de quadriques sont illustrées à la figure 2.4.

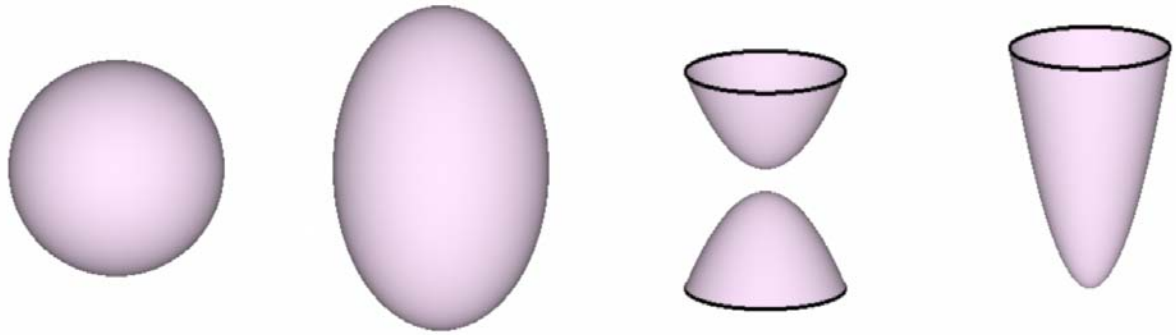


FIGURE 2.4 – Exemples de quadriques non-dégénérées. De gauche à droite, nous avons une sphère, une ellipsoïde, une hyperboloïde et une paraboloïde. Les illustrations sont extraitent de Cross (2000).

2.5.2.1 Transformation projective

Sous la contrainte d'une transformation projective 3D $\mathtt{H}$, la quadrique $\mathtt{Q}$ devient $\mathtt{Q}' = \mathtt{H}^{-\top} \mathtt{Q} \mathtt{H}^{-1}$ (la preuve est quasi identique à celle d'une conique subissant une transformation projective 2D).

2.5.2.2 Enveloppe de quadrique

Le plan tangent $\pi_{\mathbf{X}}$ à la quadrique $\mathtt{Q}$ au point $\mathbf{X}$ est donné par :

$$\Pi = \mathtt{Q} \mathbf{X}. \tag{2.40}$$

L'enveloppe de la quadrique ou "quadrique duale" est donnée de manière similaire à l'équation (2.34) par :

$$\Pi^\top \mathtt{Q}^* \Pi = 0. \tag{2.41}$$

2.5.2.3 Quadriques dégénérées

Comme pour une conique, une quadrique dégénérée est caractérisée par la singularité de la matrice $\mathtt{Q}$. Un cylindre, un cône ou une paire de plans (distincts ou confondus) sont

des exemples de quadriques dégénérées. La dégénérescence est invariante à la transformation projective (ce qu'on peut voir dans les signatures d'une quadrique). Une quadrique dégénérée a $9 - 1 = 8$ degrés de liberté au maximum (la contrainte de singularité fournit une contrainte unique sur les éléments de $\mathtt{Q}$). Plusieurs types de quadriques dégénérées sont illustrées à la figure 2.5.

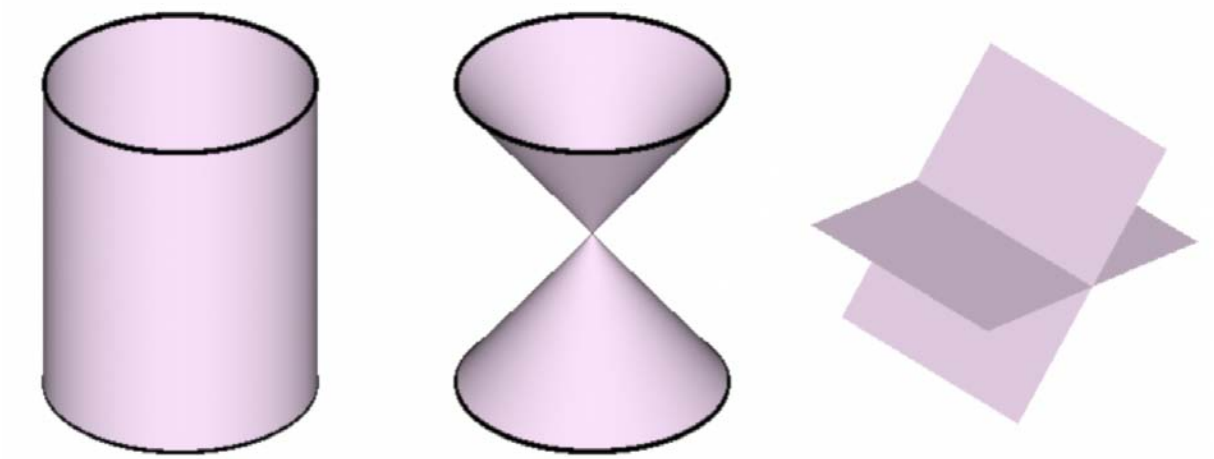


FIGURE 2.5 – Exemples de quadriques dégénérées. De gauche à droite nous avons un cylindre, un cône et une paire de plans (de rang 2). Les illustrations sont extraitent de Cross (2000).

2.5.2.4 Enveloppe de quadriques dégénérées

On peut noter qu'une enveloppe d'une quadrique de rang 1 ou 2 est non définie car $\mathtt{Q}^*$ atteint 0. L'enveloppe d'une quadrique de rang 3 (*i.e.* un cylindre ou un cône) est définie par l'équation (2.41). Ces propriétés sont évoquées plus en détail dans Semple et Kneebone (1998).

2.6 Géométrie multi-vues et reconstruction 3D

Lorsqu'une même scène est observée par plusieurs vues, il est possible d'estimer le déplacement relatif entre les différentes caméras et de calculer la géométrie 3D de l'environnement observé. Ce cas de figure peut apparaître dans différentes configurations :

- **Configuration spatiale.** Cette configuration correspond au cas où plusieurs caméras observent simultanément une même scène à partir de différents points de vue.
- **Configuration temporelle.** Dans ce cas, une seule caméra se déplace dans l'environnement. L'ensemble des vues correspond alors aux points de vue de la caméra capturée à des instants différents.

Dans le cadre de ce mémoire, nous nous intéresserons à la configuration temporelle. Néanmoins, il est important de noter que ces deux configurations, dans le cas d'une scène rigide, sont équivalentes et peuvent être traitées de façon identique.

Cette partie se consacrera à l'étude de la géométrie entre deux vues. Des méthodes complémentaires sur 3 et N vues peuvent être trouvées dans le livre de Hartley et Zisserman (2004).

2.6.1 Géométrie épipolaire

La *géométrie épipolaire* décrit les contraintes reliant les observations d'une même scène observée par deux caméras, notées $\mathcal{C}_1$ et $\mathcal{C}_2$ (figure 2.6). Ces contraintes sont directement liées au déplacement relatif (également appelé positionnement relatif) entre les deux caméras, mais sont totalement indépendantes de la structure de la scène. Toutefois, il est important de rappeler que dans le cas du déplacement d'une caméra (*i.e.* dans le cas de la configuration temporelle), la géométrie épipolaire est uniquement vérifiée si la scène observée est rigide.

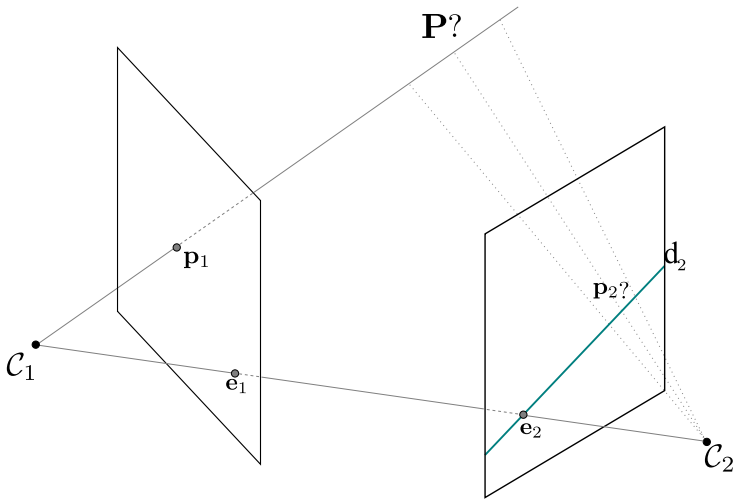


FIGURE 2.6 – **Géométrie épipolaire.** La géométrie épipolaire définit des contraintes géométriques entre les différentes observations d'un même point de l'espace.

Matrice fondamentale. La *matrice fondamentale* exprime la relation épipolaire dans le cas où les paramètres internes des caméras seraient inconnus. Ainsi, pour un point $\mathbf{p}_1$ de l'image de la première caméra, il est possible de calculer la droite $\mathbf{d}_2$ sur laquelle se situe l'observation correspondante dans la deuxième caméra (figure 2.6) :

$$\mathbf{d}_2 \sim \mathtt{F}\tilde{\mathbf{p}}_1. \tag{2.42}$$

La droite $\mathbf{d}_2$ est appelée *droite épipolaire* associée à $\mathbf{p}_1$. De plus, si deux observations $\mathbf{p}_1$ et $\mathbf{p}_2$ correspondent au même point de l'espace, elles vérifient :

$$\tilde{\mathbf{p}}_2{}^{\mathsf{T}}\mathtt{F}\tilde{\mathbf{p}}_1 = 0. \tag{2.43}$$

Cette relation permet d'estimer la matrice fondamentale à partir d'associations 2D entre deux images. En pratique, $\mathtt{F}$ peut se calculer à l'aide de 8 points (Hartley (1997)) ou à partir de 7 points (Torr et Murray (1997)). Notons que pour le cas des 7 points, nous obtenons une ou trois solutions.

Dans chacune des images, un point joue un rôle particulier. Il s'agit des deux *épipôles* $\mathbf{e}_1$ et $\mathbf{e}_2$. Ils correspondent à la projection dans l'image du centre optique de l'autre caméra. Les épipôles présentent deux caractéristiques intéressantes. Tout d'abord, elles définissent le noyau de $\mathtt{F}$: $\mathtt{F}\tilde{\mathbf{e}}_i = 0, \forall i \in \{1, 2\}$. De plus, les épipôles correspondent aux points d'intersection de toutes les droites épipolaires de chacune des images.

Matrice essentielle. La *matrice essentielle* $\mathtt{E}$ peut être vue comme le cas particulier de la matrice fondamentale dans le cas où le calibrage des caméras (K_1 et K_2) est connu, ce qui est le cas qui nous intéresse en particulier. La relation entre matrice essentielle et matrice fondamentale est la suivante :

$$\mathtt{E} \sim {\mathsf{K}_2}^{\mathsf{T}} \mathtt{F} \mathsf{K}_1. \tag{2.44}$$

L'équation (2.43) devient dans ce cas :

$$\tilde{\mathbf{p}}_2{}^{\mathsf{T}} \left({\mathsf{K}_2}^{-\mathsf{T}} \mathtt{E} {\mathsf{K}_1}^{-1} \right) \tilde{\mathbf{p}}_1 = 0, \tag{2.45}$$

où ${\mathsf{K}_2}^{-\mathsf{T}}$ est la transposée inverse de K_2. Pour estimer la matrice essentielle, Nistér (2004) a proposé un algorithme efficace appelé *algorithme des 5 points*.

Relation entre matrice essentielle et déplacement relatif. En fonction des cas d'étude, la matrice essentielle peut avoir différentes utilisations. En effet, il existe une relation qui lie la matrice essentielle du couple de caméras ($\mathcal{C}_1$, $\mathcal{C}_2$) au déplacement relatif entre ces caméras. Le déplacement relatif est défini par le couple $(\mathtt{R}_{1\to2}, \mathbf{t}_{1\to2})$. Une formalisation en sera faite à la section 2.6.2.1. La relation entre $\mathtt{E}$, $\mathtt{R}_{1\to2}$ et $\mathbf{t}_{1\to2}$ s'écrit :

$$\mathtt{E} = [\mathbf{t}_{1\to2}]_\times \mathtt{R}_{1\to2}, \tag{2.46}$$

où $[\mathbf{t}]_\times$ est la matrice antisymétrique construite à partir du vecteur $\mathbf{t}$, à savoir :

$$[\mathbf{t}]_\times = \begin{pmatrix} 0 & -\mathbf{t}_z & \mathbf{t}_y \\ \mathbf{t}_z & 0 & -\mathbf{t}_x \\ -\mathbf{t}_y & \mathbf{t}_x & 0 \end{pmatrix}. \tag{2.47}$$

Dès lors, deux cas de figure sont possibles. Si le déplacement entre les caméras est connu, la matrice essentielle permet de réduire la recherche de point d'intérêt correspondant à 1 dimension (le long de la droite épipolaire). Dans le cas contraire, une estimation de la matrice essentielle (grâce à l'appariement d'au moins 5 points) permet de retrouver le déplacement relatif entre les caméras. Cette notion sera développée dans la section 2.6.2.2.

Homographies 2D Dans le cas où la scène observée est plane, la relation qui existe entre deux caméras est définie par une *homographie* $\mathtt{H}$ (ou *transformation projective*) 2D qui représente une transformation linéaire inversible de $\mathbb{P}^2$ dans $\mathbb{P}^2$ qui conserve l'alignement. Le théorème suivant permet de caractériser de façon matricielle les homographies 2D : Une fonction $\mathtt{H} : \mathbb{P}^2 \to \mathbb{P}^2$ est une homographie si et seulement s'il existe une matrice $\mathtt{H}$ de taille 3×3 telle que pour tout point $\tilde{\mathbf{p}}$ de $\mathbb{P}^2$, $\mathtt{H}(\tilde{\mathbf{p}}) = \mathtt{H}\tilde{\mathbf{p}}$. La matrice $\mathtt{H}$ est homogène : elle est définie à un facteur près et possède donc 8 degrés de liberté.

Un des cas courants d'utilisation des homographies 2D est celui décrit dans la figure 2.7. Nous nous plaçons ici dans le cas de deux caméras observant un plan $\mathbf{\Pi}$ de l'espace. La projection centrale du plan de l'espace au plan image (et réciproquement) définit une homographie 2D (les coordonnées des points 2D étant exprimées dans le repère 2D relatif à chacun des plans). Un résultat intéressant est alors que, pour tout point 3D $\mathbf{P}$ appartenant au plan $\mathbf{\Pi}$, la fonction passant des coordonnées de son observation $\mathbf{p}_1$ dans l'image 1 aux coordonnées de son observation $\mathbf{p}_2$ dans l'image 2 est également une homographie. En effet, la composition de 2 homographies est une homographie. Le lien entre les observations peut donc s'écrire :

$$\begin{aligned} \tilde{\mathbf{p}}_2 &\sim \mathtt{H}_{1\to 2}\tilde{\mathbf{p}}_1 \\ &\sim \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \tilde{\mathbf{p}}_1. \end{aligned} \tag{2.48}$$

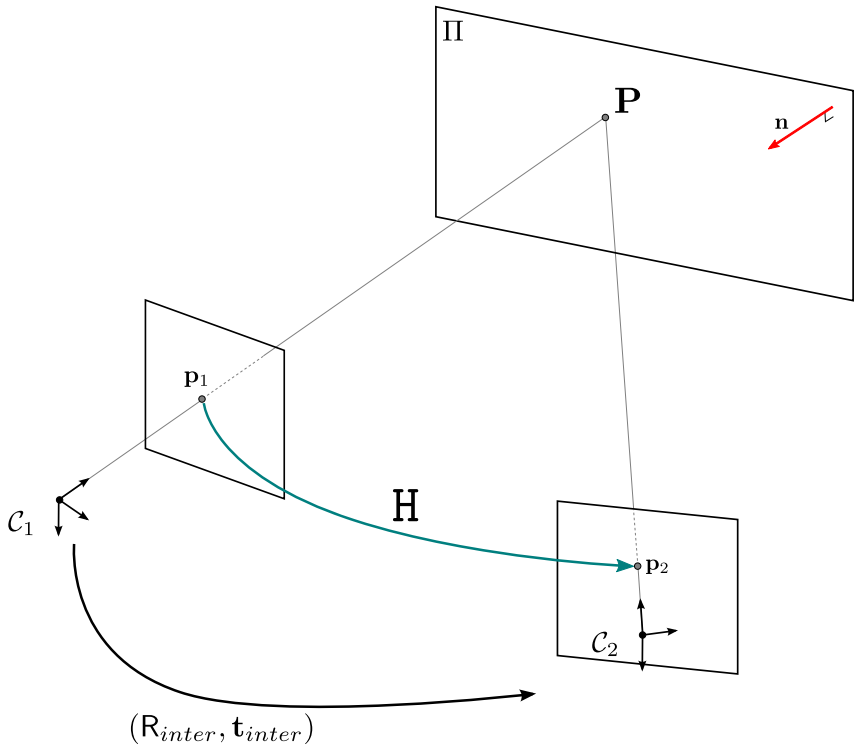


FIGURE 2.7 – Les coordonnées des observations correspondantes de points 3D situés sur un même plan de l'espace sont reliées par une homographie 2D.

2.6.2 Calcul de la géométrie de l'environnement

Dans cette section, nous allons présenter l'ensemble des outils mathématiques élémentaires qui permettent de calculer la géométrie d'une scène 3D, à savoir la pose des différentes caméras ainsi que le nuage de points 3D associés aux points d'intérêt observés.

2.6.2.1 Poses de caméras et déplacement relatif

Le but de cette section est de formaliser la notion de *déplacement relatif* entre deux caméras ainsi que les notations associées. Comme nous l'avons vu précédemment, la pose des caméras peut être vue comme un changement de repère entre le repère monde et les repères attachés aux caméras :

$$\tilde{\mathbf{P}}_{\mathcal{C}_1} \sim \begin{pmatrix} \mathtt{R}_1{}^\top & -\mathtt{R}_1{}^\top \mathbf{t_1} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_{\mathcal{W}} \tag{2.49}$$

$$\tilde{\mathbf{P}}_{\mathcal{C}_2} \sim \begin{pmatrix} \mathtt{R}_2{}^\top & -\mathtt{R}_2{}^\top \mathbf{t_2} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_{\mathcal{W}} \tag{2.50}$$

avec $\tilde{\mathbf{P}}_{\mathcal{C}_1}$ et $\tilde{\mathbf{P}}_{\mathcal{C}_2}$ les coordonnées de $\mathbf{P}$ respectivement dans les repères liés aux caméras $\mathcal{C}_1$ et $\mathcal{C}_2$ et $\tilde{\mathbf{P}}_{\mathcal{W}}$ ce même point exprimé dans le repère monde. Afin de fixer les notations, nous appellerons $(\mathtt{R}_{1\to 2}, \mathbf{t}_{1\to 2})$ le déplacement relatif entre les caméras, c'est-à-dire la transformation permettant de passer du repère lié à $\mathcal{C}_1$ à celui lié à $\mathcal{C}_2$:

$$\tilde{\mathbf{P}}_{\mathcal{C}_2} \sim \begin{pmatrix} \mathtt{R}_{1\to 2} & \mathbf{t}_{1\to 2} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_{\mathcal{C}_1}. \tag{2.51}$$

Des équations (2.49) et (2.50), on peut obtenir le système d'équations suivant :

$$\begin{cases} \tilde{\mathbf{P}}_{\mathcal{W}} \sim \begin{pmatrix} \mathtt{R}_1 & \mathbf{t_1} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_{\mathcal{C}_1} \\ \tilde{\mathbf{P}}_{\mathcal{C}_2} \sim \begin{pmatrix} \mathtt{R}_2{}^\top & -\mathtt{R}_2{}^\top \mathbf{t_2} \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_{\mathcal{W}} \end{cases} \tag{2.52}$$

d'où

$$\tilde{\mathbf{P}}_{\mathcal{C}_2} \sim \begin{pmatrix} \mathtt{R}_2{}^\top \mathtt{R}_1 & \mathtt{R}_2{}^\top (\mathbf{t}_1 - \mathbf{t}_2) \\ 0_{1\times 3} & 1 \end{pmatrix} \tilde{\mathbf{P}}_{\mathcal{C}_1}. \tag{2.53}$$

Le déplacement relatif entre les caméras est donc défini comme suit :

$$\begin{cases} \mathtt{R}_{1\to 2} = \mathtt{R}_2{}^\top \mathtt{R}_1 \\ \mathbf{t}_{1\to 2} = \mathtt{R}_2{}^\top (\mathbf{t}_1 - \mathbf{t}_2) \end{cases} \tag{2.54}$$

2.6.2.2 Calcul du déplacement relatif par associations 2D/2D

Lorsque la structure de l'environnement est inconnue, il est tout de même possible de calculer le déplacement relatif entre deux caméras. Cela nécessite d'associer les observations des 2 caméras qui correspondent aux mêmes points 3D de l'environnement. Comme nous l'avons vu précédemment (section 2.6.1), ceci permet de calculer la matrice fondamentale (algorithme des 8 points, Hartley (1997)) ou essentielle (algorithme des 5 points, Nistér (2004)). Il est alors possible d'extraire d'une de ces matrices le déplacement inter-caméra $(\mathtt{R}_{1\to 2}, \mathbf{t}_{1\to 2})$.

Dans le cas de caméras non calibrées, $\mathtt{R}_{1\to 2}$ et $\mathbf{t}_{1\to 2}$ sont calculés à partir de la matrice fondamentale (Hartley et Zisserman (2004)). Dans ce cas, le déplacement inter-caméras

ne peut être retrouvé qu'à une transformation projective près. En particulier, ceci induit qu'il est impossible de retrouver les rapports de distance et les angles.

Le calibrage des caméras étant connu dans notre étude, il est préférable d'utiliser la matrice essentielle. La décomposition en valeurs singulières SVD (Faugeras (1993)) de celle-ci permet en effet d'en extraire quatre couples de solutions possibles pour $\mathtt{R}_{1\to2}$ et $\mathbf{t}_{1\to2}$. Parmi ces 4 couples, on retient le couple permettant de reconstruire les 5 points ayant servi au calcul de $\mathtt{E}$ devant les 2 caméras . Le détail de cette décomposition peut être trouvé dans l'article de Nistér (2004).

Dans le cas calibré, le déplacement relatif entre les 2 caméras (et donc toute la structure 3D sous-jacente) est défini à un facteur près. En effet, dans le cas du calcul du déplacement par associations 2D/2D, le facteur d'échelle de la scène (c'est-à-dire sa métrique) n'est pas observable. En pratique, cette échelle est donc fixée arbitrairement.

Notons également que seul le déplacement relatif est défini mais pas la pose des caméras dans le repère monde. En effet, aucune information de localisation absolue n'est fournie de sorte que les deux caméras obtenues sont positionnées à une rotation et une translation près dans le monde. Ainsi, si le déplacement relatif est défini à un facteur près, la pose absolue des caméras est définie à 7 degrés près. Une transformation 3D possédant ces 7 degrés de liberté est appelée *similitude* et peut être représentée par la matrice homogène suivante :

$$\mathtt{S} \sim \begin{pmatrix} s\mathtt{R} & \mathbf{t} \\ 0_{1\times3} & 1 \end{pmatrix}, \tag{2.55}$$

avec s le facteur d'échelle, $\mathtt{R}$ la rotation et $\mathbf{t}$ la translation.

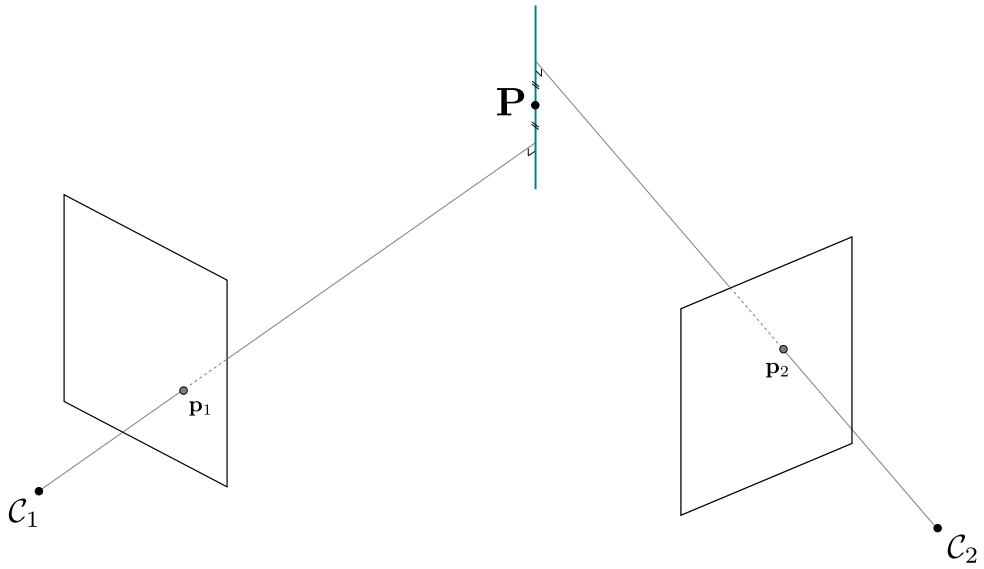


FIGURE 2.8 – **Triangulation de points 3D.** La structure de l'environnement peut être obtenue par triangulation des observations dans les images.

2.6.2.3 Calcul de la structure de l'environnement

Nous avons vu que la rétroprojection d'une observation 2D d'un point de l'espace permet d'obtenir sa position 3D à la profondeur près (section 2.3.2). Dès lors qu'au moins

2 caméras dont la pose et le calibrage sont connus observent ce point, la profondeur du point peut être estimée. On parle alors de *triangulation* du point. L'idée de la triangulation est de calculer l'intersection des rayons optiques issus des 2 observations. En pratique, à cause des bruits sur les différentes données (calibrage, pose des caméras, position des observations, *etc.*), les rayons ne s'intersectent pas. Dans le cas de 2 caméras, le résultat de la triangulation est le point équidistant des deux rayons (figure 2.8).

Dans un but de robustesse et de précision des calculs numériques, la notion de triangulation peut être généralisée à plus de 2 caméras. Par exemple, dans le cas de 3 caméras, il est possible de calculer 3 triangulations différentes à partir des couples de caméras (1,2), (2,3) et (1,3). Le résultat final de la triangulation est alors le barycentre de ces 3 points. Il existe également une approche linéaire permettant de trianguler un point observé par N-vues en utilisant la méthode DLT (Hartley et Zisserman (2004)).

2.6.2.4 Calcul de pose par associations 2D/3D

Une fois la structure de l'environnement partiellement connue, il est possible de calculer la pose d'une caméra tierce à partir d'associations réalisées entre les observations 2D de son image et la position 3D de 3 points de l'environnement. De nombreuses méthodes ont été proposées pour résoudre ce problème. Une comparaison de certaines de ces méthodes peut être trouvée dans l'article de Haralick et al. (1994). Plus récemment, Lepetit et al. (2009) ont proposé une nouvelle approche plus performante (en temps de calcul et en précision) du calcul de pose.

L'utilisation d'associations 2D/3D plutôt que 2D/2D présente plusieurs avantages. Nous avons vu précédemment que l'extraction des paramètres à partir de la matrice essentielle ne permet pas d'estimer le facteur d'échelle et donc en particulier la norme de la translation entre les différentes caméras. Avec l'approche 2D/3D, le facteur d'échelle peut être estimé à partir de l'observation de la distance entre les différents points de l'espace. Enfin, Tardif et al. (2008) ont montré que l'utilisation de l'approche 2D/3D offre un calcul plus précis de la position de la caméra.

2.7 Notion de courbure d'une surface

Dans l'étude métrique des courbes du plan et de l'espace, la courbure mesure la manière dont une courbe, ou un arc géométrique s'éloigne localement d'une ligne droite. Elle évalue le rapport entre la variation de la direction de la tangente à la courbe et un déplacement d'une longueur infinitésimale sur celle-ci : plus ce rapport est important, plus la courbure est importante. En langage imagé, la courbure indique de combien il faut tourner le volant d'une voiture pour aborder un virage (volant tourné modérément pour une courbure faible et fortement pour une courbure forte). Plus précisément, si f est une courbe régulière de classe $\mathtt{C}^k$ avec $k \geq 2$, *e.g* une courbe paramétrée par une fonction dérivable au moins deux fois, dont la dérivée première n'est jamais nulle et dont la dérivée seconde est continue. On sait que f possède localement une paramétrisation normale, c'est-à-dire qu'au voisinage d'un point $\mathbf{M}$, il existe une fonction g paramétrant f et telle que $\|g'\| = 1$. Si $g(s) = \mathbf{M}$, la courbure de f au point $\mathbf{M}$ est :

$$\gamma = \|g''(s)\| \tag{2.56}$$

Si la courbure au point **M** est non nulle, son inverse donne le rayon du cercle osculateur, c'est-à-dire le rayon du cercle s'approchant au plus près de la courbe au point **M**. Dans le cas d'une courbe plane orientée, dans un plan orienté, on peut définir une courbure algébrique, qui indique non seulement l'intensité de l'incurvation, mais aussi sa direction. Pour reprendre l'image de la route, dans un plan orienté selon le sens trigonométrique, une courbure algébrique positive indique qu'il faut tourner le volant à gauche pour aborder le tournant. La courbure algébrique est liée à l'orientation de la courbe, c'est-à-dire son sens de parcours : si pour un conducteur, il faut tourner à gauche pour aborder un tournant, pour les voitures roulant en sens inverse, il faut tourner à droite pour aborder ce même virage. Si g est une paramétrisation normale de f, pour tout s, il existe un vecteur unitaire $\mathbf{n}(s)$ tel que $(g'(s), n(s))$ soit une base orthonormée directe du plan, et il existe une fonction réelle γ telle que, pour tout s, $g''(s) = \gamma(s)n(s)$. La valeur $\gamma(s)$ est la courbure algébrique de l'arc orienté au point $\mathbf{M} = g(s)$. La valeur absolue de la courbure algébrique donne la courbure géométrique. Dans le cas d'une courbe gauche (c'est-à-dire non plane), il n'est pas possible de définir de courbure algébrique.

Cas d'un cercle Quand on se déplace sur un cercle de rayon R d'une longueur ΔS, la tangente au cercle change de direction. Les tangentes au point de départ et au point d'arrivée font un angle de $\Delta\alpha$. Cet angle correspond à l'angle au centre entre le point d'arrivée et le point de départ. On a ainsi la relation :

$$\Delta\alpha = \frac{\Delta S}{R} \tag{2.57}$$

Le rapport $\frac{\Delta\alpha}{\Delta S}$ est constant et égal à $\frac{1}{R}$. C'est ce qu'on appelle la courbure du cercle. Plus le rayon est grand, plus cette courbure est faible. Cette propriété est illustrée à la figure 2.9(a). L'idée est de considérer une courbe localement comme un cercle et de chercher la limite du rapport précédent pour un déplacement infiniment petit sur la courbe.

Cas du graphe d'une fonction Considérons un plan muni d'un repère orthonormé (0, x, y), et un arc défini comme le graphe d'une fonction f, c'est-à-dire défini par l'équation $y = f(x)$. Si cet arc est "suffisamment lisse", alors il admet en tout point x une tangente, et la pente de la tangente est la dérivée en ce point, $f'(x)$. Si cet arc est une droite, la tangente est la même partout, f' est constante. Si cet arc présente une courbure non nulle, c'est qu'il s'écarte de la notion de droite : sa dérivée varie. Intuitivement, on voit que plus l'arc est courbé, plus la dérivée "varie vite". On peut ainsi lier la courbure à la variation de la dérivée, et donc à la dérivée seconde. Si l'on cherche à chiffrer cette courbure, il faut s'intéresser aux variations infinitésimales d'angle des tangentes et de longueur d'arc. La tangente à la courbe fait avec l'axe des abscisses un angle $\alpha(x) = arctan(f'(x))$. Pour une variation infinitésimale dx, on obtient

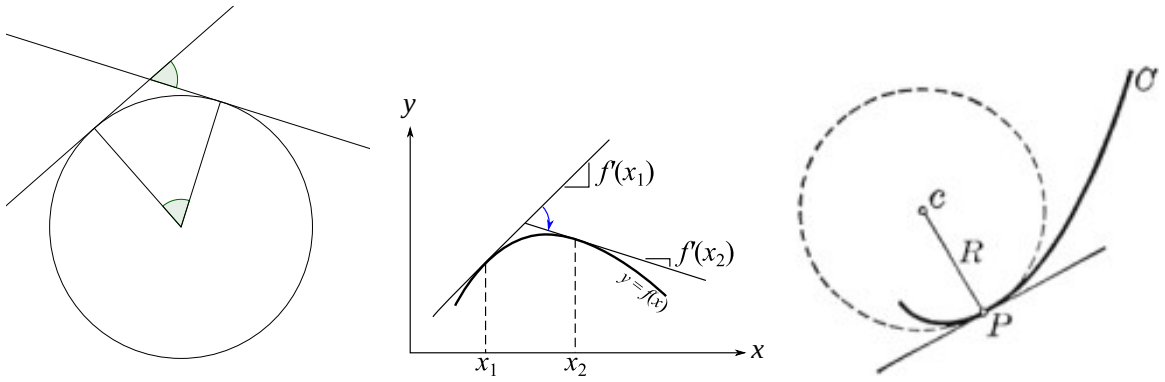
$$d\alpha(x) = \frac{f''(x)}{1 + f'^2(x)}dx$$

$$ds(x) = \sqrt{1 + f'^2(x)}dx$$

La courbure algébrique est alors donnée par le rapport de ces deux variations

$$\gamma(x) = \frac{f''(x)}{(1+f'^2(x))^{3/2}} \tag{2.58}$$

Nous illustrons le cas du graphe d'une fonction à la figure 2.9(b).



(a) (b) (c)

FIGURE 2.9 – Dans (a), nous montrons que l'angle entre les deux tangentes est égal à l'angle au centre. Dans (b), nous montrons que plus un arc est courbé, plus la pente de la tangente varie vite. Dans (c), un exemple de courbure est illustré. Au point **P**, le rayon de courbure de la courbe C est égal à R.

Introduction géométrique L'analyse dimensionnelle montre que pour un problème de cinématique, γ est homogène à l'inverse d'une longueur. On introduit donc fréquemment le rayon de courbure (algébrique) :

$$R = \frac{1}{\gamma} = \frac{ds}{d\alpha}$$

Pour comprendre la signification de cette grandeur, il est intéressant d'examiner le cas particulier d'un cercle :

$$\begin{cases} x(t) = rcos(t) \\ y(t) = rsin(t) \end{cases}. \tag{2.59}$$

L'application de la formule de calcul de la courbure donne $R = r$. Plus généralement, pour tout arc C^2 birégulier au point **P** d'abscisse curviligne s, on montre qu'il existe un unique cercle qui épouse cette courbe le mieux possible dans un voisinage de **P** : le cercle osculateur. Il est tangent à la courbe en **P** et son rayon est égal à la valeur absolue du rayon de courbure. La sinuosité décrit la courbure de plusieurs arcs reliés avec des points d'inflexion. Un exemple est illustré à la figure 2.9(c).

2.7.1 Première et deuxième formes fondamentales

Première forme fondamentale. En géométrie différentielle, la première forme fondamentale est le produit interne sur l'espace tangent d'une surface dans un espace euclidien de dimension trois. Elle permet le calcul de propriétés de courbure et de métrique d'une

surface telles que la longueur et l'aire. La première forme fondamentale est notée :

$$\mathrm{I}(x, y) = \langle x, y \rangle.$$

Soit $X(u, v)$ une surface paramétrique. Ainsi, le produit interne de deux vecteurs tangents est :

$$\begin{aligned}
&\mathrm{I}(aX_u + bX_v, cX_u + dX_v) \\
&= ac\langle X_u, X_u \rangle + (ad + bc)\langle X_u, X_v \rangle + bd\langle X_v, X_v \rangle \\
&= Eac + F(ad + bc) + Gbd,
\end{aligned}$$

où E, F et G sont les coefficients de la première forme fondamentale. La première forme fondamentale peut être également représentée en tant que matrice symétrique :

$$\mathrm{I}(x, y) = x^\top \begin{pmatrix} E & F \\ F & G \end{pmatrix} y$$

Deuxième forme fondamentale. En géométrie différentielle, la deuxième forme fondamentale est une forme quadratique sur le plan tangent d'une surface lisse dans l'espace euclidien de dimension trois et noté généralement II. La deuxième forme fondamentale permet de définir les courbures principales. Soit $X(u, v)$ une surface paramétrique dans $\mathbb{R}^3$. Cette paramétrisation permet de définir un ensemble de vecteurs normaux $\mathbf{n}$:

$$\mathbf{n} = \frac{X_u \times X_v}{|X_u \times X_v|}.$$

La deuxième forme fondamentale est écrite généralement telle que :

$$\mathrm{II} = Ldu^2 + 2Mdudv + Ndv^2,$$

sa représentation matrice dans la base $\{X_u, X_v\}$ sur le plan tangent est :

$$\begin{bmatrix} L & M \\ M & N \end{bmatrix}.$$

Les coefficients L, M et N pour un point donné dans le plan sont donnés par les projections des dérivées partielles de second ordre de $\mathbf{X}$ en ce point sur la droite normale de S et peuvent être calculés à l'aide du produit scalaire tel que :

$$L = X_{uu} \cdot \mathbf{n},\ M = X_{uv} \cdot \mathbf{n},\ X_{vv} \cdot \mathbf{n}. \tag{2.60}$$

2.7.2 Courbure normale et principale

Soit $\mathbf{u_p}$ le vecteur tangent d'une surface régulière $M \in \mathbb{R}^3$. La courbure normale de M dans la direction de $\mathbf{u_p}$ est :

$$H = S(\mathbf{u_p}) \cdot \mathbf{u_p}, \tag{2.61}$$

avec S l'opérateur de forme correspond à la dérivée négative telle que :

$$S(\mathbf{v}) = -D_{\mathbf{v}}\mathbf{N}, \tag{2.62}$$

avec $\mathbf{N}$ la normale au point de surface donné. Prenons x un patch régulier du M avec $p = x(u_x, v_y)$ et :

$$\mathbf{v_p} = a\mathbf{x_u}(u_0, v_0) + b\mathbf{x_v}(u_0, v_0), \tag{2.63}$$

où $\mathbf{v_p} \in M_{\mathbf{p}}$. La courbure normale dans la direction $\mathbf{v_p}$ est :

$$H(\mathbf{vp}) = \frac{ea^2 + 2fab + gb^2}{Ea^2 + 2Fab + Gb^2}, \tag{2.64}$$

où E, F et G sont les coefficients de la première forme fondamentale et e, f et g sont les coefficients de la deuxième forme fondamentale. Les valeurs maximale et minimale de la courbure normale d'un point sur une surface régulière sont appelées les courbures principales κ_1 et κ_2.

2.7.3 Courbure de Gauss

Calculer la courbure de Gauss est complexe. Dans cette section, nous allons évoquer le calcul de cette courbure sur une surface paramétrable et en utilisant les formes fondamentales.

Paramétrage Supposons que la surface soit donnée par une équation $z = f(x, y)$ où f est une fonction de classe C^2. Notons en indice les variables par rapport auxquelles les dérivées sont calculées. Alors, la courbure de Gauss au point de paramètre (x, y) vaut :

$$\mathcal{K} = \frac{f''_{xx} f''_{yy} - f''^2_{xy}}{(1 + f'^2_x + f'^2_y)^2} \tag{2.65}$$

Formes fondamentales Soit une surface paramétrée au moyen de deux paramètres u et v, et soit $\mathtt{I} = Edu^2 + 2Fdudv + Gdv^2$ la première forme fondamentale, $\mathbb{II} = Ldu^2 + 2Mdudv + Ndv^2$ la deuxième forme fondamentale. Alors la courbure de Gauss vaut :

$$\mathcal{K} = \frac{\det \text{II}}{\det \text{I}} = \frac{LN - M^2}{EG - F^2} \tag{2.66}$$

La courbure de Gauss $\mathcal{K}$ et la courbure moyenne H sont liées à κ_1 et κ_2 par :

$$\mathcal{K} = \kappa_1 \kappa_2$$
$$H = \frac{1}{2}(\kappa_1 + \kappa_2).$$

Ce qui peut être écrit par l'équation quadratique :

$$\kappa^2 - 2H\kappa + \mathcal{K} = 0; \tag{2.67}$$

qui a pour solutions :

$$\kappa_1 = H + \sqrt{H^2 - \mathcal{K}}$$
$$\kappa_2 = H - \sqrt{H^2 - \mathcal{K}}.$$

Toutes les notions évoquées ainsi que les démonstrations associées sont décrites dans Kobayashi et Nomizu (1963); Spivak (1970).

2.8 Optimisation numérique

La plupart des problèmes rencontrés au cours de cette thèse peuvent se ramener à la minimisation d'une fonction de coût.

2.8.1 Méthodes linéaires

Lorsque le problème peut se formuler sous la forme d'un système d'équations linéaires, il existe des méthodes matures, robustes et rapides pour trouver la solution optimale.

2.8.1.1 Moindres carrés linéaires

Le problème le plus classique est appelé *moindres carrés linéaires*. Il consiste à résoudre une équation de la forme :

$$\mathtt{A}\mathbf{x} = \mathbf{b} \tag{2.68}$$

avec $\mathbf{b} \in \mathbb{R}^m$ et $\mathtt{A} \in \mathbb{R}^{m \times n}$ connus et $\mathbf{x} \in \mathbb{R}^n$ inconnu en minimisant :

$$\mathcal{C}_{\mathrm{LS}}(\mathbf{x}, \mathtt{A}, \mathbf{b}) = \|\mathtt{A}\mathbf{x} - \mathbf{b}\|^2. \tag{2.69}$$

Il est possible de dériver une solution directe même si les méthodes numériques actuelles obtiennent des résultats plus stables et plus rapides en évitant l'inversion de matrice :

$$\mathbf{x} = argmin\ \mathcal{C}_{\mathrm{LS}}(\mathbf{x}, \mathtt{A}, \mathbf{b}) = \left(\mathtt{A}^\mathsf{T}\mathtt{A}\right)^{-1}\mathtt{A}^\mathsf{T}\mathbf{b} = \mathtt{A}^{+}\mathbf{b}. \tag{2.70}$$

$\mathtt{A}^{+}$ est la pseudo-inverse de $\mathtt{A}$.

2.8.1.2 Décomposition en valeurs singulières

Une autre approche très utile pour la résolution de systèmes linéaires est la décomposition en valeurs singulières. Toute matrice $\mathtt{A} \in \mathbb{R}^{m \times n}$ admet une décomposition de la forme suivante :

$$\mathtt{A} = \mathtt{U}\mathtt{S}\mathtt{V}^\mathsf{T} \tag{2.71}$$

où $\mathtt{U} = \mathtt{U}^{-\mathsf{T}} \in \mathbb{R}^{m \times m}$, $\mathtt{S} \in \mathbb{R}^{m \times n}$ de coefficients diagonaux $\mathbf{s} \in \mathbb{R}^n$ et $\mathtt{V} = \mathtt{V}^{-\mathsf{T}} \in \mathbb{R}^{n \times n}$. Les composantes de $\mathbf{s}$ sont appelées *valeurs singulières*. Les colonnes de $\mathtt{U}$ et $\mathtt{V}$ sont les vecteurs singuliers respectivement à gauche et à droite. Le nombre de valeurs singulières non nulles est égal au rang de la matrice $\mathtt{A}$. La décomposition en valeurs singulières est un des algorithmes numériques les plus utilisés et bénéficie d'implémentations très efficaces.

Pseudo-inverse. La pseudo-inverse de la matrice $\mathtt{A} \in \mathbb{R}^{m \times n}$ s'obtient facilement à partir de la décomposition en valeurs singulières dans le cas où $m = n$:

$$\mathtt{A}^{+} = \mathtt{V}\mathtt{S}^{-1}\mathtt{U}^{\mathsf{T}}. \tag{2.72}$$

Ce qui peut être récrit :

$$\mathtt{A}^{+} = \mathtt{V}(\mathtt{S}^{\top}\mathtt{S})^{-1}\mathtt{S}^{\top}\mathtt{U}^{\mathsf{T}}. \tag{2.73}$$

Système d'équations linéaires homogène. On appelle système d'équations linéaires homogène un problème sous la forme :

$$\mathtt{A}\mathbf{x} = 0 \tag{2.74}$$

où $\mathtt{A}$ est connue. Résoudre le système revient à estimer le noyau de $\mathtt{A}$. La décomposition en valeurs singulières permet une résolution simple. Soit r le rang de la matrice $\mathtt{A}$, la matrice $\mathtt{S}$ possède $n - r$ colonnes nulles. Les colonnes de $\mathtt{V}$ correspondantes forment une base du noyau de $\mathtt{A}$.

2.8.2 Méthodes non linéaires

Lorsque la fonction de coût est non linéaire, il est possible de résoudre le problème posé en utilisant une méthode itérative. Soit une fonction $\mathcal{C}(\mathbf{x})$ à minimiser, en partant d'un jeu de paramètres $\mathbf{x}_n$, chaque itération produit un incrément $\Delta\mathbf{x}_n$ tel que $\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta\mathbf{x}_n$ soit plus proche de la solution recherchée. Ces méthodes font toutes l'hypothèse que la fonction $\mathcal{C}$ est convexe. Dans le cas contraire, il est toujours possible d'utiliser ces méthodes avec une bonne initialisation, mais sans garanties de convergence vers le minimum global. Un problème de minimisation aux moindres carrés minimise une fonction de la forme :

$$\mathcal{C}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})^2. \tag{2.75}$$

Voici les méthodes principalement utilisées en vision par ordinateur.

2.8.2.1 Descente de gradient

La descente de gradient est une méthode de résolution du premier ordre. La direction de déplacement choisie est directement liée au gradient de la fonction étudiée :

$$\mathbf{x}_{n+1}^{(\text{desc. grad.})} = \mathbf{x}_n + \alpha\nabla\mathcal{C}(\mathbf{x}_n). \tag{2.76}$$

La longueur du pas α est cruciale pour garantir la convergence en un nombre d'itérations le plus petit possible. Comme illustré dans la figure 2.10, un pas trop important engendre une divergence du procédé de minimisation et un pas trop petit permet de converger (de façon sûre dans le cas d'un problème convexe) mais pour un nombre d'itérations trop élevé. L'avantage de cette approche est qu'elle converge efficacement même si le jeu de paramètres initial est éloigné du minimum recherché.

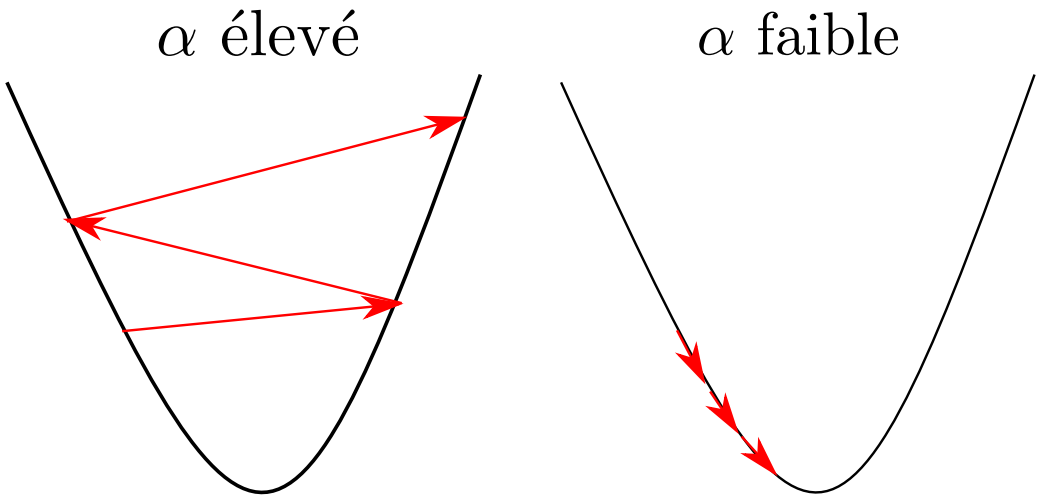


FIGURE 2.10 – Influence de la longueur du pas α sur la convergence de la méthode de descente de gradient. Dans (a), nous montrons qu'un pas trop élevé provoque une divergence et dans (b), qu'un pas trop petit converge, mais pour un nombre d'itérations élevé.

2.8.2.2 Newton

La méthode de Newton est une méthode du second ordre, basée sur une approximation quadratique de la fonction à minimiser. Le développement de Taylor de la fonction de coût $\mathcal{C}$ s'écrit :

$$\mathcal{C}(\mathbf{x}+\Delta\mathbf{x}) \approx \mathcal{C}(\mathbf{x}) + \nabla\mathcal{C}(\mathbf{x})\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathsf{T}}\mathtt{H}(\mathbf{x})\Delta\mathbf{x} \tag{2.77}$$

où la matrice hessienne $\mathtt{H}$ est définie par :

$$\mathtt{H}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2\mathcal{C}(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2\mathcal{C}(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2\mathcal{C}(\mathbf{x})}{\partial x_n x_1} & \cdots & \frac{\partial^2\mathcal{C}(\mathbf{x})}{\partial x_n^2} \end{pmatrix}. \tag{2.78}$$

Un extremum est atteint si et seulement si le gradient de l'équation (2.77) par rapport à $\Delta\mathbf{x}$ est nul, *e.g* :

$$\nabla\mathcal{C}(\mathbf{x}) + \mathtt{H}(\mathbf{x})\Delta\mathbf{x} = 0 \quad \Leftrightarrow \quad \Delta\mathbf{x} = -\mathtt{H}(\mathbf{x})^{-1}\nabla\mathcal{C}(\mathbf{x}) \tag{2.79}$$

L'incrément de Newton est donc :

$$\mathbf{x}_{n+1}^{(\text{Newton})} = \mathbf{x}_n - \mathtt{H}(\mathbf{x})^{-1}\nabla\mathcal{C}(\mathbf{x}). \tag{2.80}$$

La méthode de Newton assure une convergence plus efficace lorsque l'approximation quadratique est valide, ce qui est habituellement le cas lorsque les paramètres sont proches de la solution.

La variante dite de Gauss-Newton permet d'éviter le calcul coûteux de la hessienne.

Elle s'applique uniquement aux problèmes de moindres carrés avec un coût de la forme :

$$\mathcal{C}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})^2 \tag{2.81}$$

avec m le nombre de résidus à minimiser et n le nombre de paramètres tel que $\mathbf{x} \in \mathbb{R}^m$. La hessienne peut alors être approximée par : $\mathtt{H} \approx \mathtt{J}^\mathsf{T}\mathtt{J}$. L'incrément de Gauss-Newton est donc :

$$\mathbf{x}_{n+1}^{(\text{Gauss-Newton})} = \mathbf{x}_n - \left(\mathtt{J}(\mathbf{x})^\mathsf{T}\mathtt{J}(\mathbf{x})\right)^{-1}\mathtt{J}^\mathsf{T}(\mathbf{x}). \tag{2.82}$$

2.8.2.3 Levenberg-Marquardt

La méthode d'optimisation non linéaire de Levenberg-Marquardt (Marquardt, 1963) repose sur les deux approches précédemment citées afin de profiter de leur avantage respectif. Plus stable que celui de Gauss-Newton, il trouve une solution même s'il est démarré très loin d'un minimum. Cependant, pour des fonctions très régulières, il peut converger légèrement moins vite.

2.9 Interaction lumière/matière

La lumière correspond à un type particulier de radiation électromagnétique. Nous définissons le spectre visible comme étant la fréquence tel que la lumière peut être détectée par l'œil humain (spectre visible). En effet, les infrarouges et ultraviolets font partie de la lumière, mais ne sont pas visibles pour l'humain. La lumière peut être décrite d'une manière physique (énergie) mais également d'une façon perceptuelle (couleur, intensité). La lumière est généralement observée à deux échelles. À un niveau macroscopique, la lumière est vue comme une énergie qui flotte de façon ininterrompue dans l'espace en ligne droite et qui est, par la suite, absorbée ou réfléchie par une surface qu'elle rencontre. À une échelle microscopique, la lumière s'avère quantifiée par un ensemble de paquets appelés photons. La lumière a la particularité d'avoir un comportement proche d'une onde, car c'est une radiation électromagnétique caractérisée par une fréquence f. Le lien entre l'énergie E d'une photo et la fréquence f est réalisé par :

$$E = hf = \frac{hc}{\lambda}, \tag{2.83}$$

avec λ la longueur d'onde de la lumière en mètres, c la vitesse de la lumière dans le vide et h la constante de Plank.

En infographie, nous nous intéressons particulièrement au phénomène macroscopique et ignorons l'indivisibilité des photons.

L'interaction d'une source de lumière (feu, lampe, soleil, ...) avec une surface peut être modélisée de différentes façons selon le besoin. Pour les besoins d'un jeu vidéo ou d'une application de réalité augmentée où la notion d'images par seconde est importante, il est préférable d'avoir une représentation de lumière simplifiée et une interaction avec l'environnement, et plus particulièrement les matériaux, rapide et relativement réaliste. Dans ces cas-là, l'ajout de réflexions spéculaires et d'ombres simples suffit pour ajouter du réalisme. Nous utilisons en général des modèles d'illumination locale pour ces cas-là.

Pour des applications photo-réalistes où le temps de calcul n'est pas primordial, des modèles plus complexes peuvent être utilisés pour modéliser des phénomènes de lumière comme les caustiques, les inter-réflexions ou encore la diffraction de la lumière. Les méthodes d'illumination globale plus réalistes mais aussi plus importantes en complexité sont plus adaptées à cette situation. Ces méthodes mesurent de façon méthodique la fonction bidirectionnelle de distribution de la réflectance ou en anglais *Bidirectional réflectance Distribution Function* (BRDF) sur l'ensemble des points de vue et des points de surface de la scène.

Ces deux catégories sont illustrées à la figure 2.11 en utilisant un moteur de rendu photo-réaliste V-RAY [1] et un moteur de jeux vidéo Unreal-engine [2].

(a) Rendu photo-réaliste (b) Rendu semi-réaliste

FIGURE 2.11 – Rendus d'une scène 3D possibles dans domaine de la synthèse d'image (a) et du jeu vidéo (b).

2.9.1 Phénomène de réflexion spéculaire

Un des phénomènes lumineux les plus visibles par l'œil humain ou une caméra est la réflexion spéculaire. Elle se caractérise par une réflexion quasi totale de la lumière dans une direction principale (comportement proche de celui d'un miroir) sur une surface réfléchissante comme une plaque de métal, du plastique ou encore du bois ciré. Une **réflexion spéculaire (appelée également spécularité)** est dépendante du point de vue d'observation. Ce type de réflexion est souvent associé au phénomène de **réflexion diffuse** où la réflexion se produit dans toutes les directions et qui est ainsi indépendant du point de vue d'observation. Effectuer une séparation efficace de ces deux composantes représente un des problèmes photométriques les plus complexes. Un détail plus complet d'une méthode de détection de spécularité est développé en annexe A.

2.9.2 Réflectivité bidirectionnelle

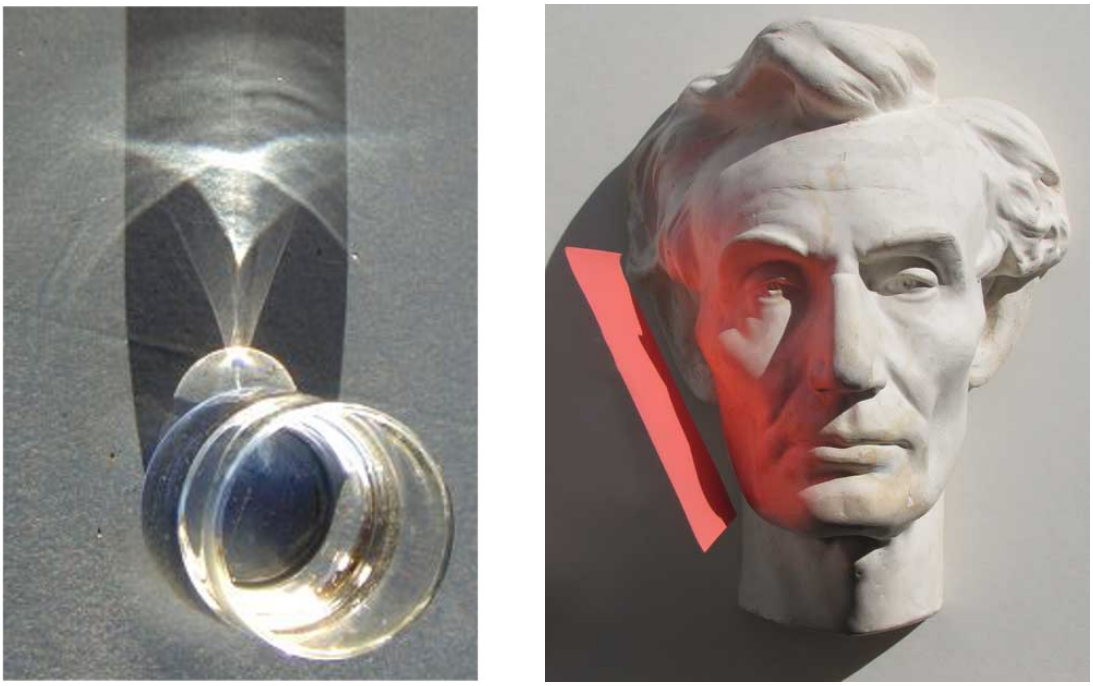
Cette fonction f évoquée dans les équations (2.84) et (2.86) caractérise la réflexion d'une surface. Elle prend en entrée une direction de la lumière incidente ω_i et une direction réfléchie ω_r et indique le ratio de la radiance sortant suivant ω_r en fonction de l'irradiance

1. https://www.chaosgroup.com/
2. https://www.unrealengine.com/en-US/what-is-unreal-engine-4

incidente de la surface suivant la direction ω_i. Cette fonction prend généralement en entrée 4 paramètres représentant ω_i et ω_r en coordonnées sphériques.

2.10 Illumination globale

En ne considérant que l'illumination locale, de nombreux effets lumineux, jouant un rôle important dans le photo-réalisme, sont ignorés. L'illumination globale est caractérisée par la modélisation des rebonds de la lumière entre les sources de lumière et les objets, ce qui permet de synthétiser des phénomènes d'ombres plus avancées et nuancées, d'inter réflexion et de réfraction. L'illumination globale modélise l'illumination indirecte *e.g* la manière dont la lumière rebondit d'une surface sur une autre à l'opposé des méthodes d'illumination directe comme les modèles d'illumination locale que nous évoquerons à la section 2.11. Parmi les effets plus avancés que l'illumination globale peut rendre, nous pouvons citer les caustiques ou les inter-réflexions diffuses (*color bleeding*) comme illustrées aux figures 2.12(a) et 2.12(b).



(a) Caustique crée par un verre d'eau

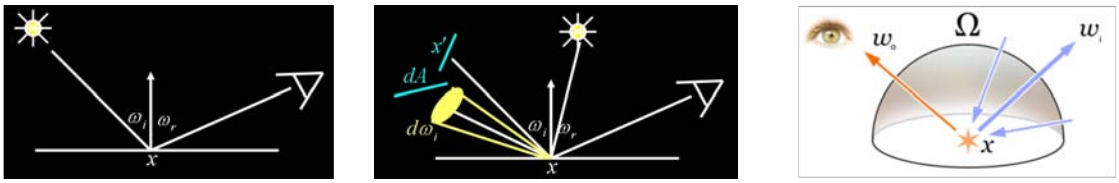
(b) *Color bleeding* d'une surface rouge sur une statue

FIGURE 2.12 – Phénomènes lumineux modélisables par l'illumination globale comme les caustiques (a) et l'inter-réflexion diffuse (b).

L'équation de rendu Une façon précise de modéliser le trajet de la lumière est d'étudier une fonction appelée l'équation de réflectance et plus généralement l'équation de rendu et la manière dont elle peut être résolue. L'équation de réflectance est définie par :

$$L_r(\mathbf{P}, \omega_r) = L_e(\mathbf{P}, \omega_r) + \sum L_i(\mathbf{P}, \omega_i) f(\mathbf{P}, \omega_i, \omega_r)(\omega_i \cdot \hat{\mathbf{N}}(\mathbf{P})), \tag{2.84}$$

avec L_r la radiance réfléchie à partir d'un point de la surface $\mathbf{P}$ dans la direction ω_r causée par la radiance incidence L_i des sources de lumière dans la direction ω_i, f la fonction bidirectionnelle de distribution de la réflectance (BRDF) qui donne la proportion de lumière réfléchie de ω_i à ω_r au point $\mathbf{P}$ et L_e la radiance émise. Une version simplifiée de cette équation est illustrée à la figure 2.13(a).



(a) (b) (c)

FIGURE 2.13 – Illustration du calcul de l'équation de réflectance et de rendu.

Cette formulation est trop simpliste, car elle ne prend en compte que les sources de lumière directe et non les portions de lumière réfléchie des autres surfaces proches. Au lieu de prendre en compte la somme des sources de lumière, il est plus naturel de prendre en compte un panel d'angles et d'intégrer la fonction de réflectance sur ce panel d'angles tel que :

$$L_r(\mathbf{P}, \omega_r) = L_e(\mathbf{P}, \omega_r) + \int_\Omega L_i(\mathbf{P}, \omega_i) f(\mathbf{P}, \omega_i, \omega_r) \cos\theta_i d\omega_i \tag{2.85}$$

$$= L_e(\mathbf{P}, \omega_r) + \int_\Omega L_r(\mathbf{P}', -\omega_i) f(\mathbf{P}, \omega_i, \omega_r) \cos\theta_i d\omega_i, \tag{2.86}$$

avec $\mathbf{P}'$ un point d'une autre surface produisant de la radiance sur la surface contenant $\mathbf{P}$. Nous considérons ainsi les surfaces diffuses comme pouvant potentiellement émettre de la radiance de la même manière qu'une source de lumière, ce qui permet de généraliser l'équation et de décrire le déplacement de la lumière de toute la scène. Une illustration de l'équation de rendu est visible à la figure 2.13(b) et 2.13(c).

Nous n'évoquerons pas ici la résolution de l'équation de rendu. Ce sujet est traité de façon plus longue dans le livre de Hughes et Foley (2014).

2.11 Modèles d'illumination locale

2.11.1 Modèle dichromatique

Le modèle de réflexion dichromatique de Shafer (1985) est un modèle physique décrit par :

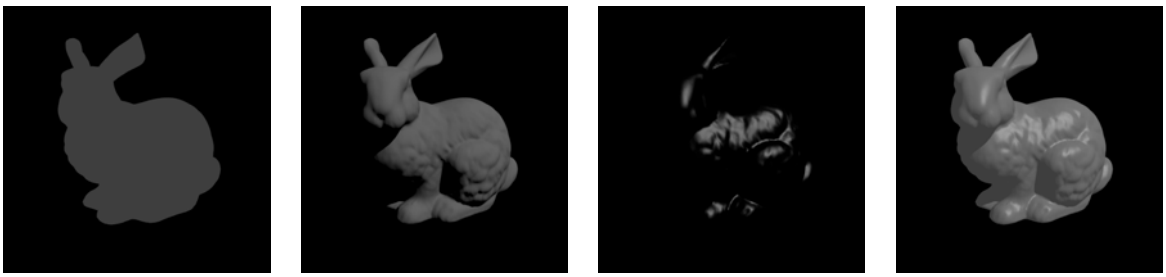
$$\begin{aligned} L(\lambda, i, e, g) &= L_i(\lambda, i, e, g) + L_b(\lambda, i, e, g) \\ &= m_i(i, e, g) c_i(\lambda) + m_b(i, e, g) c_b(\lambda) \end{aligned}$$

Ce modèle présente deux hypothèses sur le phénomène de réflexion de lumière exprimée par les deux parties de l'équation précédente. La première partie énonce que la radiance totale L de la lumière réfléchie est la somme de deux parties indépendantes : la radiance L_i de la lumière réfléchie par le matériau (de façon diffuse) et L_b la radiance de la lumière

directement réfléchie de la surface (composante spéculaire). La deuxième partie détaille le fait que chaque composante de la lumière peut être décomposée en deux parties : une distribution de la puissance spectrale relative c_i et c_b qui dépend de la longueur d'onde, mais indépendante de la géométrie et de la magnitude et un facteur d'échelle géométrique m_i ou m_b qui dépend de la géométrie et indépendante de la longueur d'onde. Le modèle de réflectance dichromatique énonce qu'il existe deux processus de réflexion indépendants et qui ont chacun une caractéristique de couleur pour laquelle sa magnitude (et non sa distribution spectrale) varie avec la direction de l'illumination et de l'observateur. Ce modèle est souvent utilisé dans les applications de séparation de composante diffuse et spéculaire à partir d'une image RGB comme dans l'approche de Hyeongwoo Kim (2013). Pour des applications de vision par ordinateur, dans un souci d'optimisation et de simplicité, il est préférable d'abstraire certains paramètres comme la puissance spectrale relative, le facteur d'échelle et la longueur d'onde. Par la suite, nous présentons les modèles de réflectance de Phong et Blinn-Phong, qui sont des cas spéciaux du modèle proposé par Shafer (1985). Notons que pour des problèmes d'estimation de l'illumination, il a été montré qu'il existe des limites à ce modèle dû à la saturation dans Lin et al. (2017).

2.11.2 Modèles de Phong et Blinn-Phong

La physique expliquant la réflexion de la lumière sur une surface est extrêmement complexe. En conséquence, le pipeline classique des moteurs graphiques repose sur une stratégie d'approximation classique appelée modèle de réflexion (ou illumination) de Phong (1975) qui propose une simulation efficace de la réflexion à un coût très faible. Dans le modèle de Phong, un matériau est décrit par trois composantes distinctes décrivant la réflexion : la composante ambiante I_a (une quantité de lumière constante permettant de décrire de façon grossière l'inter-réflexion par les autres objets) ; la composante diffuse I_d représentant la lumière réfléchie de façon omnidirectionnelle et ne dépendant pas de la position de l'observateur ; la composante spéculaire I_s qui est maximale quand le point de vue de l'utilisateur coïncide avec le rayon réfléchi de la lumière. L'image finale I est obtenue en additionnant ces trois composantes comme illustrées à la figure 2.14.



(a) Terme ambiant (b) Terme diffuse (c) Terme spéculaire (d) Image finale

FIGURE 2.14 – Modèle de Phong produisant l'image (d) par la somme de 3 composantes : ambiante (a), diffuse (b) et spéculaire (c).

2.11.2.1 Composante ambiante

La lumière ambiante est constante sur l'ensemble de la scène, ce qui rend le calcul de la composante ambiante triviale et rapide. Elle est définie par :

$$I_a = k_a i_a, \tag{2.87}$$

avec k_a un scalaire de réflexion ambiante, ce qui correspond au ratio de réflexion que la composante ambiante présente sur chaque point de la scène et i_a l'intensité de la composante ambiante (niveau de gris ou RGB).

2.11.2.2 Composante diffuse

Dans la composante diffuse, les sources de lumière interagissent avec la géométrie de la surface telle que pour un point $\mathbf{P}$ de la surface :

$$I_d(\mathbf{P}) = k_d i_d \sum_{l=1}^{n} (\hat{\mathbf{L}}_l(\mathbf{P}) \cdot \hat{\mathbf{N}}(\mathbf{P})), \tag{2.88}$$

avec k_d un scalaire de réflexion diffuse qui correspond au ratio de réflexion de la composante diffuse de la lumière incidente (appelée aussi réflectance Lambertienne Lambert (1760)), $\hat{\vec{L}}_l$ la direction normalisée de la lumière d'indice l parmi n lumières en direction d'un point de la surface tel que $\hat{\mathbf{L}}_l(\mathbf{P}) = \mu(\mathbf{L}_l - \mathbf{P})$ et $\hat{\mathbf{N}}(\mathbf{P})$ la normale de la surface au point $\mathbf{P}$ et i_d l'intensité de la composante diffuse (niveau de gris ou RGB).

Nous pouvons observer que ce terme n'a besoin d'être calculé qu'une seule fois pour chaque point 3D de la surface, car le terme diffus n'est pas dépendant du point de vue de l'observateur. Plus spécifiquement, la normale $\hat{\mathbf{N}}(\mathbf{P})$ ne varie pas au point $\mathbf{P}$ et la source de lumière reste fixe ce qui fait que $\mu(\mathbf{L}_l - \mathbf{P})$ conserve sa valeur constante quand l'observateur se déplace.

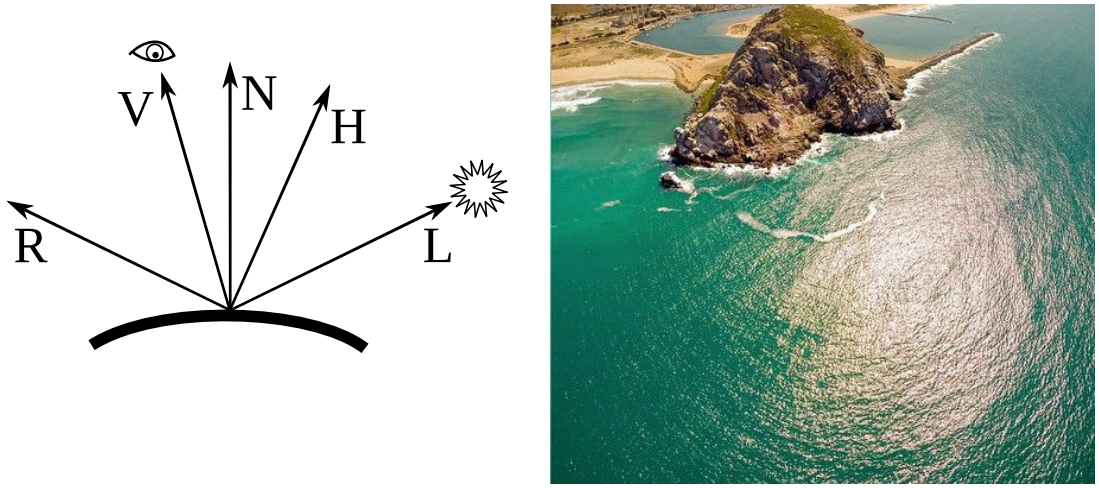
2.11.2.3 Composante spéculaire

La composante spéculaire est le terme le plus coûteux et correspond à la somme des intensités calculées pour chaque source de lumière telle que pour un point $\mathbf{P}$ de la surface :

$$I_s(\mathbf{P}) = k_s i_s \sum_{l=1}^{n} (\hat{\mathbf{R}}_l(\mathbf{P}) \cdot \hat{\mathbf{V}}(\mathbf{P}))^s, \tag{2.89}$$

avec k_s un scalaire de réflexion spéculaire qui correspond au ratio de réflexion de la composante spéculaire de la lumière incidente et i_s l'intensité de la composante spéculaire (niveau de gris ou RGB), $\hat{\mathbf{R}}_l = 2(\hat{\mathbf{L}}_l(\mathbf{P}) \cdot \hat{\mathbf{N}}(\mathbf{P})\hat{\mathbf{N}}(\mathbf{P}) - \hat{\mathbf{L}}_l(\mathbf{P})$ le rayon réfléchi du rayon lumineux incident, $\hat{\mathbf{V}}(\mathbf{P}) = \mu(\mathbf{V} - \mathbf{P})$ le point de vue de l'observateur et s l'exposant spéculaire.

Nous pouvons remarquer que l'intensité spéculaire diminue quand l'observateur s'éloigne du rayon réfléchi (l'angle entre $\hat{\mathbf{V}}(\mathbf{P})$ et $\hat{\mathbf{R}}_l(\mathbf{P})$ augmente). Afin de contrôler la vitesse d'atténuation, un exposant spéculaire s est utilisé qui est en général associé à un type de matériau. Pour une réflexion légère comme sur une pomme par exemple, ce coefficient aura une valeur autour de 10 alors que pour une surface métallique, cette valeur tournera autour de 100 à 1000.



(a) Calcul de la composante spéculaire (b) Spécularité de forme elliptique sur l'océan

FIGURE 2.15 – Calcul de la composante spéculaire en utilisant le rayon réfléchi $\mathbf{R}$ dans le modèle de Phong ou $\mathbf{H}$ le vecteur mi-chemin entre $\mathbf{V}$ et $\mathbf{L}$ (a) et (b) spécularité visible sur l'océan de forme elliptique.

Ce modèle de Phong a été pendant des décennies une méthode rapide et efficace pour donner une impression de réalisme à des applications comme le jeu vidéo. Même si ce modèle n'a pas de fondements physiques précis, la séparation de l'image I en trois composantes et la simplicité de la fonction d'intensité spéculaire font de ce modèle une fondation des modèles modernes et un modèle assez adapté aux applications de Réalité Augmentée.

Une variation du calcul de la composante spéculaire a été proposée par Blinn-Phong Blinn (1977) en remplaçant le calcul de $\hat{\mathbf{R}}_l(\mathbf{P}) \cdot \hat{\mathbf{V}}(\mathbf{P})$ par $\hat{\mathbf{H}}_l(\mathbf{P}) \cdot \hat{\mathbf{N}}(\mathbf{P})$ avec $\hat{\mathbf{H}}_l(\mathbf{P})$ le vecteur de mi-chemin entre le rayon de lumière incidente et le vecteur associé à l'observateur tel que $\hat{\mathbf{H}}_l(\mathbf{P}) = \mu(\hat{\mathbf{L}}_l(\mathbf{P}) + \hat{\mathbf{V}}(\mathbf{P}))$. Les vecteurs réfléchis $\hat{\mathbf{R}}$ et $\hat{\mathbf{H}}$ sont illustrés à la figure 2.15(a). En plus d'optimiser le calcul du terme spéculaire en évitant le calcul du rayon réfléchi, les spécularités prennent une forme plus elliptique contrairement à celles de Phong qui sont majoritairement rondes. Dans le cas d'une spécularité observée dans un point de vue assez incliné, sa forme décrit naturellement une ellipse (reflet du soleil observé à l'horizon par exemple sur la figure 2.15(b)). De plus, Blinn-Phong produit, en général, des spécularités suffisamment précises comme évoqué dans Ngan et al. (2004).

D'autres modèles ayant un fondement physique plus poussé existent, comme celui de Torrance et Sparrow (1967) et Cook et Torrance (1982) mais nous n'entrerons pas dans le détail de ces modèles dans ce mémoire.

CHAPITRE 3

État de l'art

Comme évoqué en introduction, reconstruire de façon rapide et précise les phénomènes lumineux comme les ombres et spécularités est un problème mal-posé et complexe. Dans nos travaux de thèse, nous nous sommes concentrés essentiellement sur une tâche plus réduite, mais également plus simple : la prédiction de spécularité. Il est intéressant de considérer les spécularités comme des éléments à part entière du fait de leurs importances en vision par ordinateur. En provoquant une variation importante de l'intensité dans l'image, de nombreux algorithmes de vision par ordinateur peuvent être perturbés comme les méthodes de segmentation, de reconstruction ou encore de localisation de caméra. Cependant, ces spécularités fournissent également des informations pertinentes sur les sources de lumière, la géométrie et les matériaux de la scène. En raison de la nouveauté de notre étude sur le phénomène des spécularités et la possibilité de les prédire pour de nouveaux points de vue, l'état de l'art des méthodes dédiées uniquement à la prédiction est peu étoffé. Dans ce chapitre, nous décrivons les deux approches de l'état de l'art pouvant réaliser de la prédiction de spécularité, mais généralement incluses dans un processus de reconstruction d'illumination : l'illumination globale et l'estimation de sources lumineuses (incluant l'illumination locale).

3.1 Prédiction de spécularité

La prédiction de spécularité est un problème difficile. En effet, les spécularités représentent un phénomène photométrique complexe décrit par une taille, une position, une forme et une intensité associée. De plus, ces éléments sont fortement influencés par la caméra (sa pose et sa fonction de réponse), la scène comprenant sa géométrie et les matériaux associés (réflectance et rugosité) et les sources de lumière (position, forme, intensité, isotropie ou anisotropie). À notre connaissance, aucune méthode de l'état de l'art ne permet la prédiction explicite de spécularité. En effet, l'approche naturelle serait d'estimer un modèle physique de la scène avec les paramètres associés : matériaux (réflectance et rugosité) et lumières (forme, pose, intensité et atténuation). Contrairement à la localisation de caméra et à la reconstruction géométrique de la scène, l'estimation de ces paramètres

est possible, mais difficile en pratique.

La plupart des approches de prédiction de spécularité modélisent explicitement les sources de lumière ou l'illumination de la scène afin de créer un rendu des spécularités par la suite. Ces méthodes sont divisées en deux catégories : le rendu par illumination globale et l'estimation des sources de lumière.

3.1.1 Rendu en illumination globale

Les méthodes dans cette catégorie privilégient la qualité de rendu des applications en approximant une solution à l'équation de rendu comme expliqué à la sous-section 2.10 comme les méthodes de Immel et al. (1986); Kajiya (1986). Pour rappel, cette équation décrit la quantité de lumière émise à un point **P** pour une direction d'observation donnée et à partir d'une BRDF de la lumière incidente donnée. Ces approches, en général, ne calculent pas les attributs physiques des sources de lumière et des matériaux de la scène. Par exemple, Jachnik et al. (2012) reconstruisent un *light field* 4D sur une surface spéculaire plane de petite dimension. En reconstruisant les composantes diffuses et spéculaires, cette méthode présente des résultats de rendu photo-réaliste convaincant pour des applications de réalité augmentée et en temps-réel par une implémentation GPU. Cependant, elle est incapable de prédire la composante spéculaire pour des points de vue inconnus de la phase d'initialisation. En effet, cette estimation de l'illumination globale ne permet pas de distinguer les sources primaires et secondaires, ce qui est essentiel pour la prédiction de spécularité sur des points de vue inconnus. De plus, les sources de lumière à état changeant (allumées ou éteintes) ne peuvent pas être gérées avec ce type de méthode. Des approches plus récentes comme Thomas et al. (2016) ont étendu Jachnik et al. (2012) en rajoutant une segmentation de matériau accompagnée d'une reconstruction géométrique réalisée avec un capteur RGB-D. Cette méthode souffre cependant des mêmes limitations que Jachnik et al. (2012). Une autre approche de Meilland et al. (2013) présente également une méthode d'estimation d'illumination globale en reconstruisant les sources primaires en tant que sources ponctuelles. Ces sources de lumière sont observées directement à partir d'une caméra RGB-D. Malgré la qualité des résultats obtenus, la méthode manque de flexibilité. En effet, elle représente une source étendue comme un néon par un ensemble de sources ponctuelles. Ainsi, cette méthode doit calculer l'intensité lumineuse des différentes sources ponctuelles alors qu'elles devraient être considérées comme une unique source lumineuse. De plus, aucune lumière dynamique (qui peut s'allumer ou s'éteindre dans la séquence) ne peut être utilisée et les matériaux pour la prédiction de spécularité ne sont pas calculés. En conséquence, cette méthode ne peut pas prédire les spécularités pour des points de vue inconnus.

La tendance actuelle des méthodes d'illumination globale est d'utiliser des méthodes d'apprentissage profond comme Mandl et al. (2017) en donnant en entrée un objet 3D de géométrie et albédo connus. Le réseau de neurones convolutif utilise les coefficients de la fonction d'harmoniques sphérique calculée afin de généraliser l'illumination de façon immédiate après apprentissage. Cette méthode fournit un résultat à partir d'une seule image, mais nécessite une période d'apprentissage de 4 heures par pose et l'illumination apprise se limite à l'objet appris, ce qui limite le panel d'application de la méthode. Quelques méthodes d'illumination globale avec une contrainte temps réel sont illustrées à la figure 3.2 et sans contraintes temporelles à la figure 3.1.

En conséquence, les méthodes citées ci-dessus ne peuvent pas prédire les spécularités

(a) Illumination globale avec des lumières polygonales

(b) Illumination globale avec aprioris géométriques

FIGURE 3.1 – Illustration de deux méthodes sans contraintes temporelles de (a) Heitz et al. (2016) et (b) Karsch et al. (2011). Dans (a), le rendu visé est synthétique, ce qui permet d'avoir un environnement contrôlé en termes de bruit du capteur, des paramètres des sources lumineuses et de la géométrie. Une application de rendu réaliste à partir de sources de lumière polygonales est illustrée. Dans (b), à partir d'une image réelle et de contraintes géométriques par indication manuelle d'éléments de la scène (murs, angles d'une pièce, localisation de l'illumination), un rendu photo-réaliste est réalisé incluant de nombreux phénomènes d'illumination comme les ombres, les caustiques, inter-réflexion et spécularités.

pour de nouveaux points de vue, car elles dépendent souvent d'une phase d'initialisation/apprentissage méticuleuse qui requiert beaucoup de puissance de calcul et beaucoup d'images. Ces méthodes ne nous semblent pas adaptées à un contexte temps réel en réalité augmentée.

3.1.2 Modèles d'illumination locale

En parallèle, d'autres travaux ont été proposés pour la reconstruction d'une source primaire. Idéalement, chaque modèle physique de lumière doit être associé à une géométrie (position et forme), une couleur (Lindsay et Agu (2014)) et une valeur d'intensité pour correspondre de façon réaliste à la scène. Bien qu'il existe de nombreux modèles de lumière en infographie, les modèles utilisés pour la vision par ordinateur se divisent généralement en deux catégories : sources de lumière directionnelles et sources de lumière ponctuelles. Dans un contexte extérieur, une source directionnelle semble plus naturelle pour représenter l'illumination comme le soleil, mais peut aussi être pertinente dans un contexte intérieur (lumières au plafond par exemple). Par exemple, Lagger et Fua (2006) décrivent une méthode pour calculer des sources de lumière directionnelles à partir d'un objet mobile sur un point de vue fixe. Cette application est limitée, car les spécularités sont fortement dépendantes du point de vue et les sources lumineuses doivent être estimées pour chaque position. Ni la forme ou la position de la source lumineuse ni le matériau de l'objet ne sont pris en compte, ce qui rend cette méthode incapable de prédire les

(a) Light field (b) Reconstruction avec capteur RGB-D (c) Apprentissage profond

FIGURE 3.2 – Trois illustrations de méthodes d'illumination globale. Dans (a), la méthode de Jachnik et al. (2012) permet d'ajouter un objet en réalité augmentée (cube en brique) dans des conditions d'illumination réalistes en reconstruisant un *light field*. Dans (b), à l'aide d'une caméra RGB-D, une reconstruction de la géométrie de la scène et de l'illumination sous forme d'un *light field* est réalisée. Dans (c), à partir d'une image RGB et d'un processus d'apprentissage profond des harmoniques sphériques de l'objet, un rendu réaliste est réalisé sur la figurine. Cependant, ces trois méthodes sont limitées à la zone d'initialisation/apprentissage et ne sont pas assez flexibles dans un environnement plus dynamique (grands déplacements de caméra, lumières qui s'allument ou s'éteignent).

spécularités de façon précise et explicite. De plus, les sources étendues comme les néons ne peuvent pas être modélisées correctement par des sources de lumière ponctuelles, ce qui peut limiter l'applicabilité de la méthode.

Des approches similaires comme Boom et al. (2013); Buteau et Saito (2015); Einabadi et Grau (2015); Kanbara et Yokoya (2004); Wong et al. (2008) souffrent des mêmes problèmes avec les mêmes limitations. Pour l'estimation de sources lumineuses ponctuelles, l'approche de Boom et al. (2013) estime une source de lumière primaire dans un contexte Lambertien et en utilisant une caméra RGB-D. À partir du terme diffus et par comparaison avec un rendu synthétique de la scène, une unique source ponctuelle est estimée. Cependant, cette méthode n'est pas adaptée pour des applications temps réel comportant plusieurs sources de lumière pouvant être ponctuelles comme une lampe de bureau ou une ampoule ou étendues comme un néon. En calculant une seule source lumineuse sans expliciter la forme et les matériaux des surfaces de la scène, il est difficile de réaliser une prédiction de spécularité réaliste et cohérente.

Prédire les spécularités à partir d'une reconstruction paramétrique d'une source de lumière requiert de calculer de nombreux paramètres de lumière et de matériau complexes et parfois ambigus. Par exemple, estimer conjointement la réflectance du matériau et l'intensité de la source de lumière est un problème mal posé comme évoqué dans Lombardi et Nishino (2012). En pratique, une approche physique comme celles citées précédemment ne permet pas la prédiction d'un phénomène photométrique aussi complexe que la spécularité. Des exemples de rendu utilisant des modèles d'illumination locale sont illustrés à la figure 3.3.

3.1.3 Modèles géométriques

Considérer la spécularité comme un élément géométrique est une piste qui a été peu étudiée. Dans les travaux de Blake et Brelstaff (1988), le lien entre courbure de la surface

(a) (b)

FIGURE 3.3 – Deux méthodes de reconstruction de modèle d'illumination locale de Boom et al. (2013) (a) et Kanbara et Yokoya (2004) (b). Nous observons que dans ces exemples la qualité n'est pas aussi avancée que dans les méthodes d'illuminations globales, mais elles permettent un rendu très rapide et aussi plus flexible.

et mouvement de la spécularité a été étudié et montre des résultats intéressants sur la compréhension du comportement d'une spécularité de façon géométrique. Dans un contexte où les poses de caméra et la géométrie sont connues, le mouvement de la spécularité peut être prédit. Cependant, aucune information n'est donnée dans la littérature expliquant la transformation de la forme d'une spécularité due aux changements de courbure.

Il est possible de considérer la source de lumière et plus particulièrement la spécularité comme étant un objet 3D dans l'espace au même titre que tout autre objet 3D dans la scène. En effet, une spécularité correspond au reflet d'une source de lumière réfléchie par une surface spéculaire. Le comportement de ces surfaces est très proche de celui d'une surface miroir. Ainsi, nous pouvons rapprocher ce type de méthode des méthodes de SfM (Structure-from-Motion) faisant une reconstruction 3D à partir de réflexions miroir comme Kuthirummal et Nayar (2007); Mitsumoto et al. (1992); Savarese et al. (2005). En géométrie multi-vues, nous distinguons différents cas pour reconstruire un objet statique par projection perspective. Le cas le plus simple est celui d'un objet directement observé. Les points, lignes ou autres caractéristiques visuelles 3D sont obtenus dans l'image par simple projection perspective en fonction de la pose de la caméra. Dans le cas d'un objet, observé à travers un miroir parfait, les caractéristiques visuelles 3D sont obtenues dans l'image en calculant leur symétrique par rapport à la normale du miroir. Pour une source de lumière ponctuelle observée directement dans une image, l'intensité combinée à la sensibilité du capteur fait que l'image de la source de lumière ne correspond pas exactement à la projection des caractéristiques visuelles 3D d'une source de lumière. En effet, il est très difficile d'observer une source de lumière de façon consistante et précise, car cette source de lumière augmente fortement l'intensité générale de l'image. En présence d'une surface spéculaire, en complément de l'intensité de la source de lumière et de la sensibilité du capteur, l'image de la source de lumière ponctuelle est observée à travers une surface brillante (de comportement proche de celui d'un miroir) et affectée par le matériau de la surface (réflectance et rugosité). Ainsi, cette image ne correspond pas exactement à la projection perspective des caractéristiques visuelles 3D. Ce décalage entre la forme réelle de la source de lumière et celle observée dans une image semble être représenté par un

changement d'échelle de la source de lumière. Dans ce mémoire, nous aborderons particulièrement ce problème de modélisation de spécularité par une forme 3D ; comment appliquer cette modélisation dans un processus de prédiction de spécularité et quels sont les avantages d'un tel type de modèle par rapport aux autres approches de l'état de l'art.

3.2 Positionnement

Afin de traiter ce problème complexe et fortement non-linéaire, nous proposons un modèle virtuel empirique basé sur la reconstruction d'une quadrique à partir de coniques calculées sur les spécularités dans chaque point de vue. Malgré sa simplicité, ce modèle empirique abstrait les paramètres des matériaux et des sources lumineuses ce qui permet la prédiction de spécularité pour des points de vue inconnus en présence d'une ou de plusieurs sources de lumière.

Dans le chapitre 4, nous présentons notre modèle géométrique appelé *JOint LIght-MAterial Specularity* (JOLIMAS) permettant d'abstraire l'interaction lumière/matériau. Ce modèle est né de l'observation de la forme des spécularités sur des surfaces planes. Or, sur ces surfaces, ces éléments présentent une forme elliptique. Les chapitres suivants 5 et 6 traiteront de la généralisation de ce modèle pour des surfaces plus complexes (courbes, concaves, convexes). Notre modèle, en plus de prédire la position des spécularités pour chaque source de lumière et chaque surface de la scène, peut être utilisé pour diverses applications, notamment la réalité augmentée. Dans le chapitre 7, nous mettrons en lumière l'efficacité et la pertinence de cette approche en application à la réalité augmentée.

CHAPITRE 4

Un modèle primal pour les surfaces planes

Dans ce chapitre, nous posons les bases de notre modèle géométrique que nous utilisons pour la prédiction des spécularités. À l'opposé des méthodes d'illumination locale et d'illumination globale, il est intéressant de considérer la lumière comme un élément géométrique, car le comportement d'une spécularité est très similaire à celui d'un objet réfléchi dans un miroir. Nous explorons cette piste dans ce chapitre par un exemple relativement simple (une surface plane), mais permettant de mettre en évidence nos premières hypothèses et le principe général de notre modèle géométrique.

4.1 Introduction

Le phénomène photométrique des spécularités est souvent présent dans les images. En effet, ces éléments se manifestent sur une surface lorsque la taille des imperfections de celle-ci est inférieure ou de l'ordre de grandeur de la longueur d'onde incidente. C'est pourquoi une surface de métal brut qui diffuse fortement devient parfaitement réfléchissante quand on la polit (on l'abrase jusqu'à ce que la taille des défauts soit comparable à la longueur d'onde de la lumière). Dans ces conditions, la lumière est complètement réfléchie sous une forme spéculaire (l'angle du rayon de lumière incident avec la normale de la surface à un point de surface donné est égal à l'angle de rayon réfléchi avec la normale). Si cette lumière réfléchie rencontre l'objectif de la caméra, une spécularité est créée dans l'image. De ce fait, ces éléments sont importants dans plusieurs domaines. En vision par ordinateur et traitement d'image, elles provoquent une augmentation de l'intensité de l'image ce qui a pour effet de perturber les algorithmes de détection ou de suivi 2D/3D, de localisation de caméra ou encore de reconstruction 3D.

Cependant, si on les considère comme des primitives de localisation, leur présence apporte des informations pertinentes permettant une meilleure estimation de la profondeur d'une scène et donc de sa géométrie permettant d'améliorer la localisation de caméra Lagger et al. (2008); Meilland et al. (2011); Silveira et Malis (2007), la reconstruction 3D Fleming et al. (2004); Savarese et al. (2004) et l'analyse de matériaux d'une scène Chang et al. (2009); Netz et Osadchy (2013a,b); Shroff et al. (2011). En réalité augmentée (RA)

ou en infographie, ces primitives permettent d'augmenter significativement le réalisme du rendu des applications Csongei et al. (2014); Gruber et al. (2014); Hachisuka et al. (2008); Jachnik et al. (2012); Jarosz et al. (2008); Knorr et Kurz (2014); Rohmer et al. (2014); Sato et al. (1999). Il a été montré dans les études de Blake (1990) que les spécularités jouent un rôle important dans la perception de la profondeur et du relief d'un objet. Afin d'améliorer le plus possible ces applications, il est intéressant de pouvoir prédire les spécularités dans une scène pour un point de vue donné.

Nous avons observé empiriquement que la forme de la spécularité sur une surface plane peut être modélisée par une ellipse. De plus, nous savons qu'une ellipse est issue de la projection d'une quadrique sur un plan. À partir de ces éléments, il est naturel de se demander s'il est possible de reconstruire une quadrique fixe dont la projection correspond aux spécularités pour chaque point de vue. Dans ce cas, ce modèle représenterait un lien entre le phénomène photométrique des spécularités (lumières et matériaux) et la géométrie multi-vues Hartley et Zisserman (2004).

Nous proposons un modèle empirique que nous avons nommé *JOint LIght-MAterial Specularity* (JOLIMAS) permettant de prédire facilement la position et la forme des spécularités pour de nouveaux points de vue comme illustré à la figure 4.5. Ce modèle est composé d'une quadrique qui inclut les propriétés des lumières et des matériaux de la scène. Cette quadrique est reconstruite à partir de chaque ellipse associée à une spécularité observée sur une surface plane, pour un point de vue donné. Dans ce chapitre, nous présentons les différents modèles physiques utilisés en général en vision par ordinateur et analysons leur capacité à prédire le phénomène complexe des spécularités. Nous travaillons essentiellement sur des surfaces planes car il est facile de retrouver ce type de surface dans une scène, comme exploité dans Buteau et Saito (2015); Jachnik et al. (2012). L'extension de ce modèle à des surfaces plus complexes requiert une étude approfondie qui fera l'objet des chapitres 5 et 6.

Notre modèle est présenté dans la section 4.2. Son estimation est développée dans la section 4.4 en comparant différentes approches de l'état de l'art. Notre modèle et sa capacité à prédire les spécularités ont été testés sur des séquences de synthèse et réelles comprenant des sources de lumière de type ampoule et néon dans la section 4.5. Nous illustrons également la capacité de notre modèle à gérer plusieurs sources de lumière et des lumières à état changeant (allumées et éteintes) dans le chapitre 7.

4.2 Approche proposée

Nous partons de l'observation qu'une spécularité présente une forme elliptique sur une surface plane. Si une ellipse correspond à la projection perspective d'une quadrique, existe-t-il une quadrique fixe dans l'espace dont la projection perspective correspond aux spécularités ? Dans ce cas, un modèle faisant abstraction des paramètres physiques liés à l'apparition de la spécularité peut être obtenu.

Nous proposons un modèle empirique incluant les propriétés des sources de lumière et des matériaux en reconstruisant une quadrique issue des ellipses appliquées aux spécularités pour chaque point de vue comme illustré sur la figure 4.1. Nous assumons des surfaces planes non Lambertiennes où chaque spécularité est associée à une source de lumière. Notre étude se concentre sur ce modèle permettant de prédire les spécularités provenant de n'importe quel type de lumière.

Une démonstration de notre approximation de la spécularité, observée sur une surface

plane, par une ellipse est détaillée dans les sections suivantes. Nous évaluerons empiriquement cette hypothèse dans la section 4.3.3.

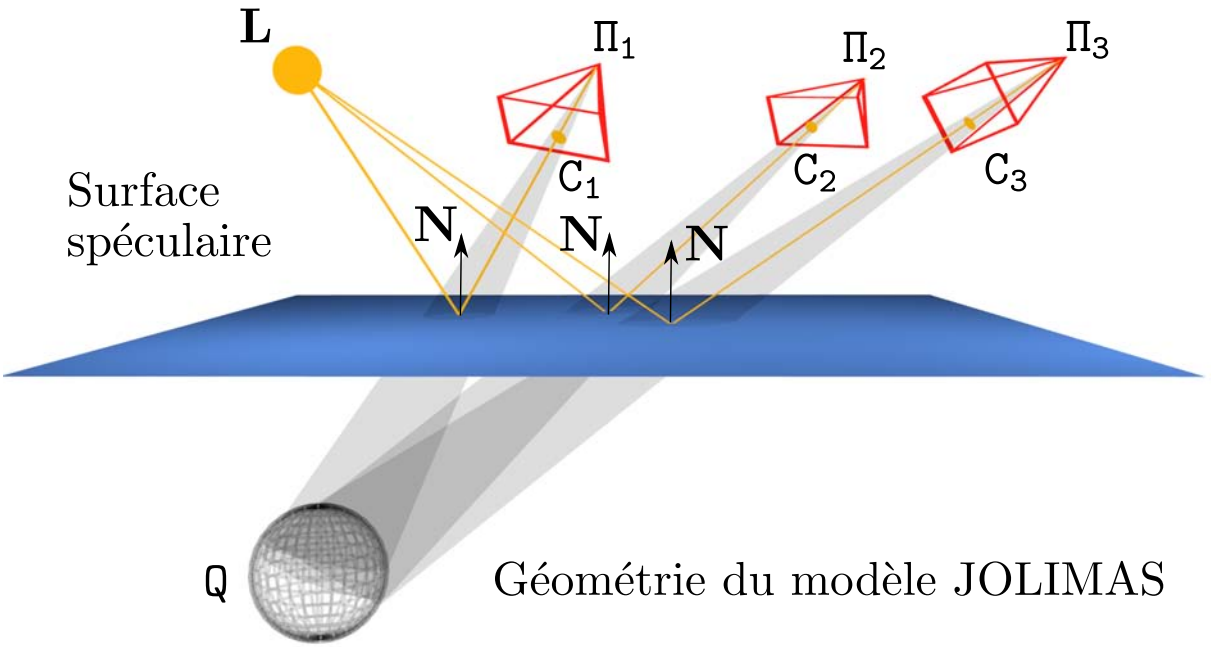


FIGURE 4.1 – Lien entre le modèle physique (lumière et matériaux) et la géométrie multi-vues (notre modèle empirique). À partir de plusieurs points de vue, une quadrique est reconstruite à partir de spécularités de forme elliptique. Cette quadrique virtuelle projetée sur chaque point de vue correspond aux spécularités.

4.3 Motivation théorique

Dans cette section, nous mettons en évidence l'utilisation des modèles classiques d'illumination locale comme Phong et Blinn-Phong comme base théorique de notre modèle. En effet, dans ces modèles, afin d'assurer un rendu rapide et cohérent, de nombreuses approximations sont réalisées sur l'interaction lumière/matériau. Ces approximations sont judicieuses à étudier, car elles fournissent une fondation intéressante pour l'abstraction des paramètres physiques liés à l'apparition de la spécularité de notre modèle. L'objectif de cette section est de montrer que notre hypothèse de spécularité de forme conique (et plus spécifiquement elliptique) est cohérente.

4.3.1 Modèles de Phong et Blinn-Phong

Nous utilisons le modèle d'illumination de Phong (1975) qui est à la base de nombreux modèle de BRDF Hughes et Chraibi (2012). Nous n'utilisons pas le modèle de Phong pour reconstruire notre modèle JOLIMAS mais plutôt comme un moyen théorique de développer certaines de ses propriétés fondamentales. Ce modèle d'illumination divise l'image en 3 composantes : diffuse, ambiante et spéculaire. Des modèles plus récents comme ceux de Blinn (1977), Ward (1994), Cook et Torrance (1982) diffèrent majoritairement dans

le calcul du terme spéculaire. Ces modèles proposent également une estimation de la rugosité qui correspond à la micro géométrie d'une surface et est un élément important dans la perception et le réalisme d'un matériau. Même si le modèle de Phong est considéré comme obsolète dans les méthodes de rendu photo-réaliste, il convient bien pour expliciter l'approximation effectuée dans notre étude de par sa simplicité analytique.

La fonction d'intensité du modèle de Phong (1975) d'un point 3D sur une surface est donnée par la formule :

$$I(\mathbf{p}) = i_a k_a + i_d k_d(\hat{\mathbf{L}}(\mathbf{p}) \cdot \hat{\mathbf{N}}) + i_s k_s(\hat{\mathbf{R}}(\mathbf{p}) \cdot \hat{\mathbf{V}}(\mathbf{p}))^n, \tag{4.1}$$

avec $\hat{\mathbf{R}}(\mathbf{P}) = -\mu(\mathbf{R} - \mathbf{P})$ la direction normalisée d'un rayon réfléchi de lumière $\mathbf{L} = \begin{bmatrix} L_X & L_Y & L_Z \end{bmatrix}^\top$ tel que $\mathbf{R} = \begin{bmatrix} L_X & L_Y & -L_Z \end{bmatrix}^\top$, $\hat{\mathbf{L}}$ et $\hat{\mathbf{V}}$ la direction normalisée du point de vue orientée vers la source de lumière $\mathbf{L}$ et le capteur $\mathbf{V} = \begin{bmatrix} V_X & V_Y & V_Z \end{bmatrix}^\top$, n la brillance de la surface, k_s, k_a et k_d le ratio de réflexion des termes spéculaire, ambiant et diffus de la lumière incidente, $\hat{\mathbf{N}}$ la normale de la surface S et i_s, i_a et i_d l'intensité entrante sur la surface pour les termes spéculaire, ambiant et diffus.

Notre repère monde est choisi tel que la surface $S \subset \mathbb{R}^3$ correspond au plan $(\mathbf{XY})$ et $S = \{\mathbf{P} \in \mathbb{R}^3 | P_Z = 0\}$. Nous pouvons paramétrer S par le point $\mathbf{P} = \begin{bmatrix} P_X & P_Y & P_Z \end{bmatrix}^\top$.

4.3.2 Terme spéculaire

Au point $\mathbf{P}$, la composante spéculaire I_s est donnée par l'équation :

$$I_s(\mathbf{P}) = \max(0, (\hat{\mathbf{R}}(\mathbf{P}) \cdot \hat{\mathbf{V}}(\mathbf{P}))^n). \tag{4.2}$$

Nous voulons analyser les isocontours d'une spécularité sur une surface S et pour un point de vue $\mathbf{V}$, une source de lumière $\mathbf{L}$ et $\mathbf{R}$. Nous partons du modèle de réflexion exprimé par :

$$\tilde{I}_s(\mathbf{P}) = -\mu(\mathbf{R} - \mathbf{P})^\top \mu(\mathbf{V} - \mathbf{P}). \tag{4.3}$$

Dans un premier temps, nous développons Eq.4.3 pour un τ général en la résolvant pour $\tau = 1$ et $\tau = 0$.

$$\tilde{I}_s(\mathbf{p}) = \tau. \tag{4.4}$$

$\tilde{I}_s$ est un produit scalaire entre deux vecteurs normalisés, nous avons $-1 \leq \tilde{I}_s \leq 1$. De plus, puisque $P_Z = 0$, $L_Z > 0$ et $V_Z > 0$, $\tilde{I}_s > 0$, de manière générale, $0 < \tilde{I}_s \leq 1$. En multipliant l'équation (4.4) $\|\mathbf{R} - \mathbf{P}\|\|\mathbf{V} - \mathbf{P}\|$, en mettant l'expression au carré et en soustrayant la partie de droite, nous obtenons une quartique à deux variables (en P_X et P_Y) :

$$\Big((\mathbf{R} - \mathbf{P})^\top (\mathbf{V} - \mathbf{P})\Big)^2 - \tau^2 \|\mathbf{R} - \mathbf{P}\|^2 \|\mathbf{V} - \mathbf{P}\|^2 = 0.$$

En développant le terme spéculaire à l'équation (4.1) et en rassemblant les monômes

de même degré, on obtient :

$$\begin{aligned}
&(d^\circ 4)\ (1-\tau^2)||\mathbf{P}||^4 \\
&(d^\circ 3)\ 2(\tau^2-1)(\mathbf{R}+\mathbf{V})^\top \mathbf{P}||\mathbf{P}||^2 \\
&(d^\circ 2)\ \mathbf{P}^\top(\mathbf{R}\mathbf{R}^\top + 2\tau(1-2\tau^2)\mathbf{R}\mathbf{V}^\top + \\
&\qquad (2\mathbf{R}^\top\mathbf{V} - \tau^2||\mathbf{R}||^2 - \tau^2||\mathbf{V}||^2)\mathtt{I})\mathbf{P} \\
&(d^\circ 1)\ 2(-\mathbf{R}^\top\mathbf{V}\mathbf{R}^\top - \mathbf{R}^\top\mathbf{V}\mathbf{V}^\top + \\
&\qquad \tau^2||\mathbf{R}||^2\mathbf{V}^\top + \tau^2||\mathbf{V}||^2\mathbf{R}^\top)\mathbf{P} \\
&(d^\circ 0)\ (\mathbf{R}^\top\mathbf{V})^2 - \tau^2||\mathbf{R}||^2||\mathbf{V}||^2.
\end{aligned}$$

Pour $\tau = 1$, les monômes de degré 3 et 4 disparaissent tels que les termes restants forment une équation quadratique tels que :

$$\tilde{\mathbf{P}}^\top \mathtt{Q} \tilde{\mathbf{P}} = 0, \tag{4.5}$$

où $\tilde{\mathbf{P}} \overset{\text{def}}{=} \begin{bmatrix}\mathbf{P} & 1\end{bmatrix}^\top$ sont les coordonnées homogènes de $\mathbf{P}$. La matrice $\mathtt{Q} \in \mathtt{R}^{4\times 4}$ est symétrique et définie par :

$$\mathtt{Q} = \begin{bmatrix} [\mathbf{R}-\mathbf{V}]_\times^2 & [\mathbf{V}\times\mathbf{R}]_\times(\mathbf{V}-\mathbf{R}) \\ ([\mathbf{V}\times\mathbf{R}]_\times(\mathbf{V}-\mathbf{R}))^\top & (\mathbf{R}^\top\mathbf{V})^2 - ||\mathbf{R}||^2||\mathbf{V}||^2 \end{bmatrix}$$

Nous observons que la somme des termes de degrés 3 et 4 donne :

$$(1-\tau^2)||\mathbf{P}||^2\mathbf{P}^\top(\mathbf{P}-2\mathbf{R}-2\mathbf{V}). \tag{4.6}$$

4.3.2.1 Point d'intensité maximale

À la valeur d'intensité maximale $\tau = 1$, les monômes de degrés 3 et 4 disparaissent. Les termes restants forment une équation quadratique qui peut être réécrite de la manière suivante :

$$\tilde{\mathbf{P}}^\top \mathtt{J} \tilde{\mathbf{P}} = 0, \tag{4.7}$$

où $\tilde{\mathbf{P}} \overset{\text{def}}{=} \begin{bmatrix}\mathbf{P} & 1\end{bmatrix}^\top$ sont les coordonnées homogènes de $\mathbf{P}$. La matrice $\mathtt{J} \in \mathbb{R}^{4\times 4}$ est symétrique et définie par :

$$\mathtt{J} = \begin{bmatrix} [\mathbf{R}-\mathbf{V}]_\times^2 & [\mathbf{V}\times\mathbf{R}]_\times(\mathbf{V}-\mathbf{R}) \\ ([\mathbf{V}\times\mathbf{R}]_\times(\mathbf{V}-\mathbf{R}))^\top & (\mathbf{R}^\top\mathbf{V})^2 - ||\mathbf{R}||^2||\mathbf{V}||^2 \end{bmatrix}$$

Nous avons montré dans l'appendice C que $\text{rang}(\mathtt{J}) = 2$ et $\mathtt{J}$ est non négative. Ce qui signifie que $\mathtt{J}$ est une quadrique de classe point ce qui représente une ligne contenant $\mathbf{R}$ et $\mathbf{V}$. Son intersection avec S est définie comme étant le point d'intensité maximale. En conséquence, la spécularité a un point d'intensité maximale bien défini et unique.

4.3.2.2 Contour externe

Pour comprendre en détail la structure du problème, l'Eq.4.4 est résolue pour $\tau = 0$ ce qui révèle la nature du contour externe. En développant le terme spéculaire de Eq.4.1,

on obtient :

$$\|\mathbf{P}\|^2 - (\mathbf{R}+\mathbf{V})^\top \mathbf{P} + \mathbf{R}^\top \mathbf{V} = 0, \tag{4.8}$$

ce qui correspond à une quadrique et plus particulièrement une sphère décrite par la matrice :

$$\mathtt{J} = \begin{bmatrix} \mathtt{I}_3 & -\frac{1}{2}(\mathbf{R}+\mathbf{V}) \\ -\frac{1}{2}(\mathbf{R}+\mathbf{V})^\top & \mathbf{R}^\top \mathbf{V} \end{bmatrix}. \tag{4.9}$$

En définissant la projection orthographique sur S :

$$\mathtt{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.10}$$

le contour de la spécularité est donné par la conique :

$$\mathtt{C} = \mathtt{A}^\top \mathtt{J} \mathtt{A} = \begin{bmatrix} \mathtt{I}_2 & -\frac{1}{2}(\bar{\mathbf{R}}+\bar{\mathbf{V}}) \\ -\frac{1}{2}(\bar{\mathbf{R}}+\bar{\mathbf{V}})^\top & \mathbf{R}^\top \mathbf{V} \end{bmatrix},$$

avec $\bar{\mathbf{R}} \overset{\text{def}}{=} \begin{bmatrix} R_X, R_Y \end{bmatrix}^\top$ et $\bar{\mathbf{V}} \overset{\text{def}}{=} \begin{bmatrix} V_X, V_Y \end{bmatrix}^\top$ correspondent aux projections orthographiques de $\mathbf{R}$ et $\mathbf{V}$ sur S respectivement. $\mathtt{C}$ représente un cercle pour $\tau = 0$.

En observant le cercle $\mathtt{C}$ sur le plan image de la caméra de centre optique $\mathbf{V}$, une conique est obtenue et plus particulièrement une ellipse si la spécularité est contenue entièrement dans le plan image.

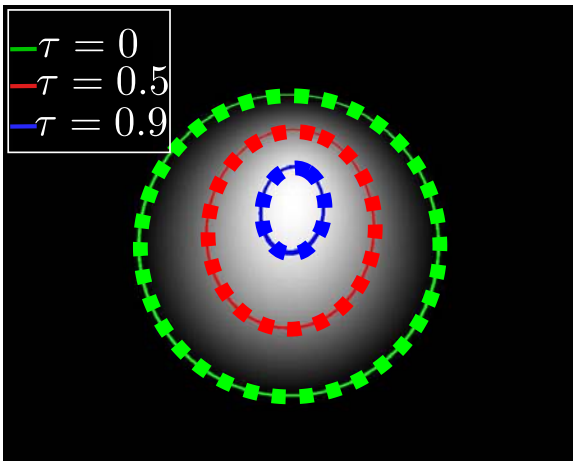
4.3.2.3 Isocontours internes

Afin d'étudier les contours internes correspondant à une valeur de $\tau \in [0, 1]$ arbitraire, les réflexions spéculaires peuvent être approximées par une ellipse dans le modèle de Phong comme illustré dans la figure 4.2(a). Cette approximation a également été testée et validée sur le modèle de Blinn (1977) comme illustré à la figure 4.2(b). Initialement, ce modèle a été défini en tant qu'approximation du modèle de Phong et régulièrement présenté en infographie comme un modèle plus précis en comparaison avec Phong. Ce modèle est aussi considéré comme physiquement plus cohérent, car il vérifie l'équation de Helmholtz Haines et Akenine-Moller (2002).

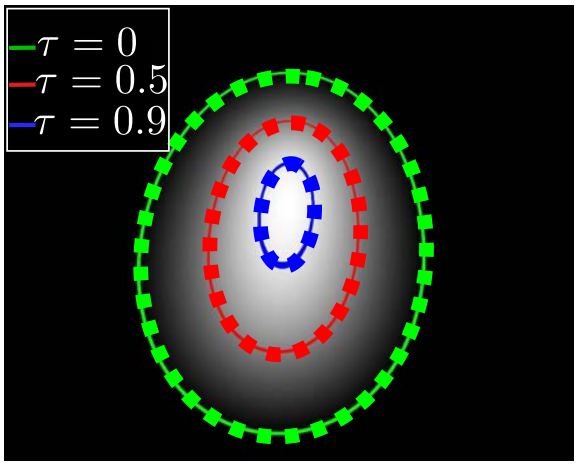
Cette démonstration illustre la pertinence et la précision d'une modélisation d'une spécularité par une ellipse. Afin de démontrer davantage cette propriété, une validation empirique sur des données de synthèse est réalisée dans la prochaine section.

4.3.3 Validation empirique de l'hypothèse du modèle à partir de données simulées

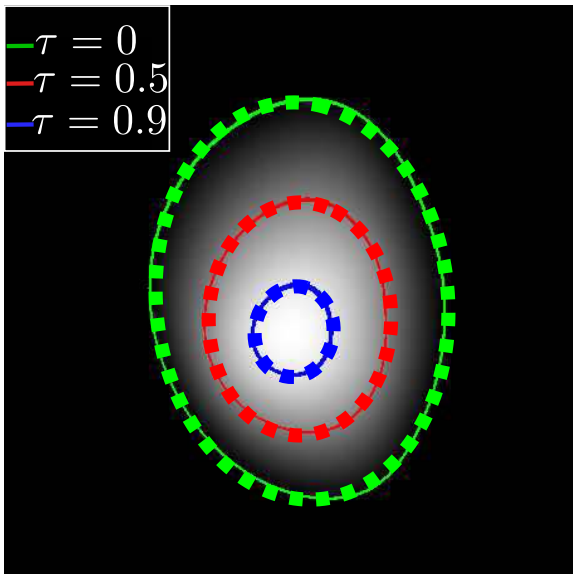
Afin de valider notre approximation de spécularités, nous utilisons la méthode de Fitzgibbon et al. (1996) d'estimation d'ellipse pour différentes valeurs de $\tau \in [0, 1]$ sur la partie spéculaire uniquement et ensuite sur la partie diffuse combinée à la partie spéculaire. Dans un contexte de RA, les composantes diffuse et spéculaire sont corrélées. Afin de calculer nos données synthétiques, nous générons des images 2000×2000 en utilisant les modèles d'illumination locale de Phong (1975) et Blinn (1977) à partir d'une position



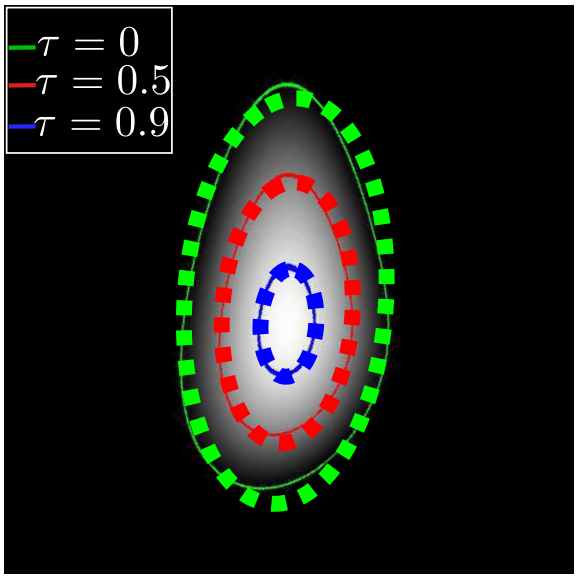
(a) Phong, source ponctuelle



(b) Blinn-Phong, source ponctuelle



(c) Phong, source étendue

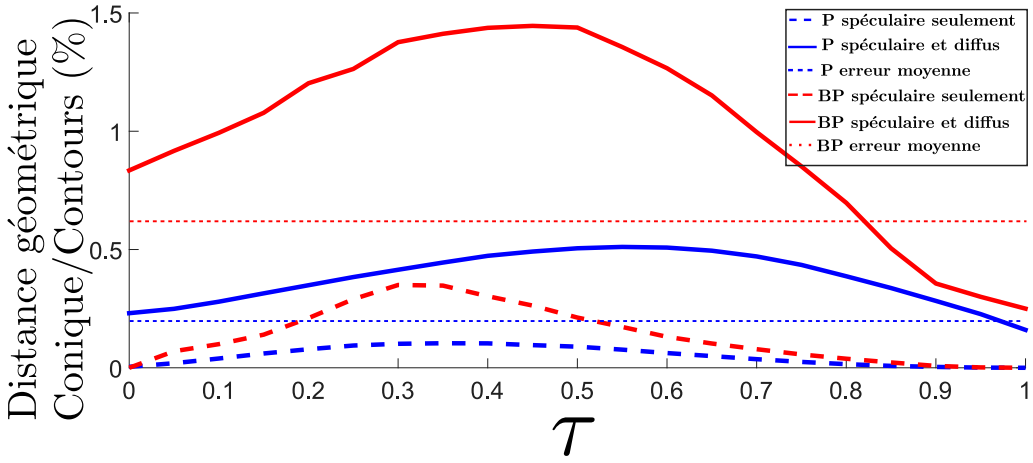


(d) Blinn-Phong, source étendue

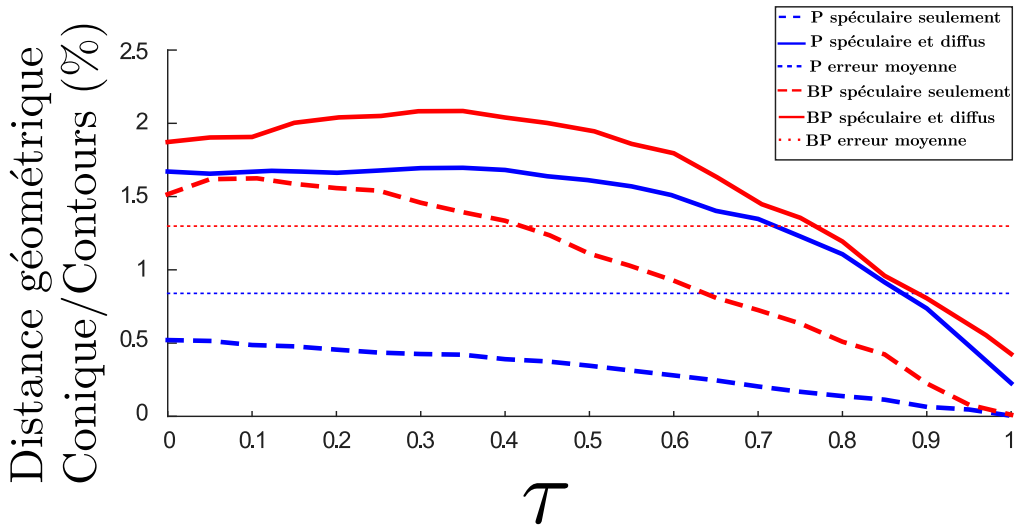
FIGURE 4.2 – Illustration de notre hypothèse de spécularités de forme elliptique. Une spécularité est créée à partir du terme spéculaire du modèle de réflexion de Phong dans (a) pour une simple lumière ponctuelle, et dans (c) pour une source de lumière étendue (une ligne composée de plusieurs sources ponctuelles) et à partir du terme spécularité du modèle de réflexion de Blinn-Phong dans (b) pour une simple lumière ponctuelle et dans (d) pour une source de lumière étendue. Les spécularités sont clairement de forme elliptique dans tous ces cas. En observant ces formes dans le plan images, nous obtenons des ellipses. Nous montrons dans ces exemples les différents isocontours associés à une intensité spécifique $\tau \in \{0, 0.5, 0.9\}$ (contours en vert, rouge et bleu) ainsi que les ellipses correspondantes (en pointillés verts, rouges et bleus). Notre hypothèse de spécularité elliptique est distinctement vérifiée dans ces exemples.

aléatoire de source lumineuse et d'une caméra localisées au-dessus de la surface plane de taille fixe. L'objectif est d'analyser la forme de la spécularité pour une intensité fixe $\tau \in [0, 1]$. Pour chaque valeur de τ, avec un pas de 0.05 entre chaque valeur, nous lançons 1000 scénarios de position de lumière et de pose de caméra. L'erreur d'estimation de l'ellipse

sous forme d'une distance point/point entre les isocontours et l'ellipse correspondante est calculée par la méthode de Sturm et Gargallo (2007). Cette distance est calculée pour chaque scénario. La validation empirique de l'estimation d'ellipse est donnée par la figure 4.3(a) pour des sources de lumière ponctuelles et par la figure 4.3(b) pour des sources étendues. En infographie, les sources étendues sont souvent représentées comme un ensemble de sources de lumière ponctuelles *e.g* une ligne pour représenter un néon.



(a) Source ponctuelle



(b) Source étendue

FIGURE 4.3 – Validation empirique à partir de données synthétiques pour (a) des sources de lumière ponctuelles et (b) des sources de lumière étendues. Notre hypothèse de spécularités de forme elliptique est testée sur les modèles d'illumination de Phong (P) et de Blinn-Phong (BP) pour différentes valeurs de $\tau \in [0, 1]$. Cette hypothèse est testée pour le terme spéculaire dans un premier temps et en combinant le terme spéculaire et le terme diffus combinés.

Cette erreur est calculée sous forme d'un ratio entre la distance géométrique ellipse/contours avec le diamètre de l'ellipse en pixels pour $\tau = 0$. Les résultats montrent que notre approximation correspond aux données de synthèse avec moins de 0.2% d'erreur pour le modèle de Phong et 0.6% pour le modèle de Blinn-Phong pour une source de lu-

mière ponctuelle et une erreur moyenne de 0.7% et 1.3% respectivement pour les modèles de Phong et Blinn-Phong pour une source étendue. Ces résultats confirment la fiabilité et la précision de notre représentation elliptique de spécularité pour les sources de lumière ponctuelles et étendues. Pour les sources étendues comme les néons, notre approximation d'ellipses est toujours pertinente comme illustré à la figure 4.3(b). Notre montrons empiriquement qu'une quadrique fixe et unique existe à partir de plusieurs points de vue et qui justifie les formes elliptiques des spécularités dans la section 4.5.1.

4.4 Estimation du modèle

Nous avons montré dans la section 4.3.3 qu'une spécularité observée sur une surface plane peut être approximée par une ellipse. La contribution principale de notre méthode est de pouvoir représenter chaque lumière par une quadrique $\mathtt{Q}$ associée par les projections (spécularités) sur un plan S pour une matrice de projection donné $\Pi = \mathtt{K}\begin{bmatrix}\mathtt{R}_{rot} & \mathbf{V}\end{bmatrix}$.

4.4.1 Vue globale

Le pipeline de la méthode proposée est divisé en plusieurs étapes :

- Detection de spécularité dans les images par notre approche expliquée dans l'annexe A
- Calcul de l'ellipse enveloppant les spécularités Fitzgibbon et al. (1996)
- Reconstruction de quadrique Cross et Zisserman (1998)
- Prédiction de spécularité par projection de quadrique

Notre processus de détection de spécularité est divisé en 3 étapes. Dans un premier temps, afin de réduire le bruit dans l'image due au contexte lumineux dynamique, nous appliquons une balance de la luminosité ; par la suite, plusieurs candidats de spécularités sont sélectionnés en réalisant un seuillage automatique par l'intermédiaire d'une valeur moyenne de luminance. Enfin, afin de réduire davantage les mauvaises détections, notre méthode utilise un critère sur le gradient de l'intensité de l'image de gradient afin de discriminer les surfaces spéculaires des textures blanches détectées à tort. Cette méthode suppose que l'intensité d'une spécularité est isotropique et décroissante à partir du point d'intensité maximale. Le processus de calcul d'ellipse englobante utilise une image en niveaux de gris en sortie du processus de détection de spécularité afin de fournir les ellipses d'entrée pour le processus de reconstruction de quadrique. Le détail de notre méthode de détection est fourni dans l'annexe A.

4.4.2 Reconstruction de quadrique par la méthode de Cross*et al.*

Nous utilisons l'approche de Cross et Zisserman (1998) qui reconstruit une quadrique duale $\mathtt{Q}^*$ à partir de plusieurs coniques duales $\mathtt{C}^*$ en reformulant la relation :

$$\Pi \mathtt{Q}^* \Pi^\top = \mathtt{C}^*. \tag{4.11}$$

La vectorisation de l'équation (4.11) est donnée dans l'appendice B.5. Par vectorisation de $\mathtt{Q}^*$ et $\mathtt{C}^*$ en $\mathbf{Q}_v^*$ et $\mathbf{C}_v^*$, nous posons l'équation (4.12) avec $\mathtt{C}$ nos spécularités modélisées

en ellipses et $\mathtt{Q}$ notre modèle de prédiction représenté en quadrique pour n points de vue tels que $n \geq 3$. Nous construisons le système :

$$\mathtt{M}\mathtt{w} = 0,$$

avec

$$\underbrace{\begin{pmatrix} \mathtt{B}_1 & -\mathbf{C}^*_{1,v} & 0 & \dots & 0 \\ \mathtt{B}_2 & 0 & -\mathbf{C}^*_{2,v} & & 0 \\ \vdots & 0 & 0 & \ddots & \vdots \\ \mathtt{B}_n & 0 & 0 & \dots & -\mathbf{C}^*_{n,v} \end{pmatrix}}_{\mathtt{M}} \underbrace{\begin{pmatrix} \mathbf{Q}^*_v \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}}_{\mathtt{w}} = 0, \tag{4.12}$$

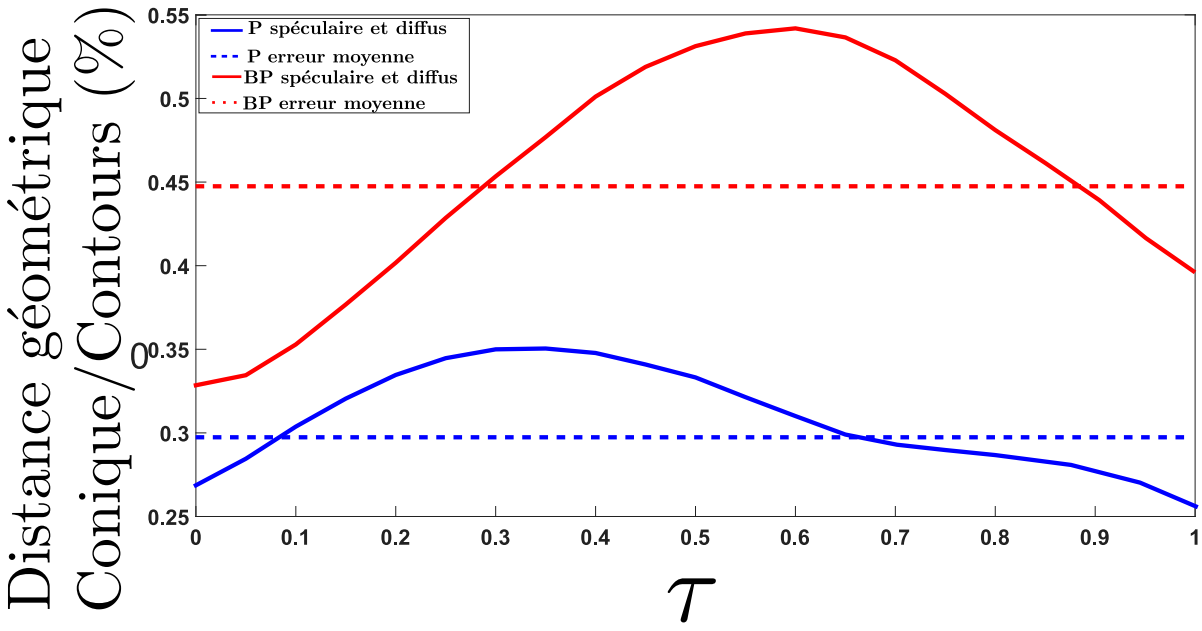
avec la matrice $\mathtt{B}_i \in \mathbb{R}^{6\times 10}$. La solution de ce système linéaire est illustrée par l'équation 4.12 qui est calculée par décomposition en valeurs singulières (SVD) de $\mathtt{M}$. α_i correspond à l'échelle pour l'image d'indice i telle que : $\alpha_i \mathbf{C}^*_{i,v} = \mathtt{B}_i \mathbf{Q}^*_{i,\mathtt{v}}$. Le processus de reconstruction de la quadrique est illustré à la figure 4.1. Cependant, le système (4.12) est relativement sensible aux erreurs du calcul des ellipses en entrée. Afin d'assurer la qualité des ellipses en entrée utilisées pour la reconstruction, nous utilisons un procédé de correction épipolaire des ellipses en entrée. Cette correction assure que les ellipses respectent la contrainte épipolaire avant le procédé initial de reconstruction de quadrique. Afin d'améliorer la reconstruction de quadrique, un raffinement non-linéaire est proposé.

4.5 Résultats expérimentaux

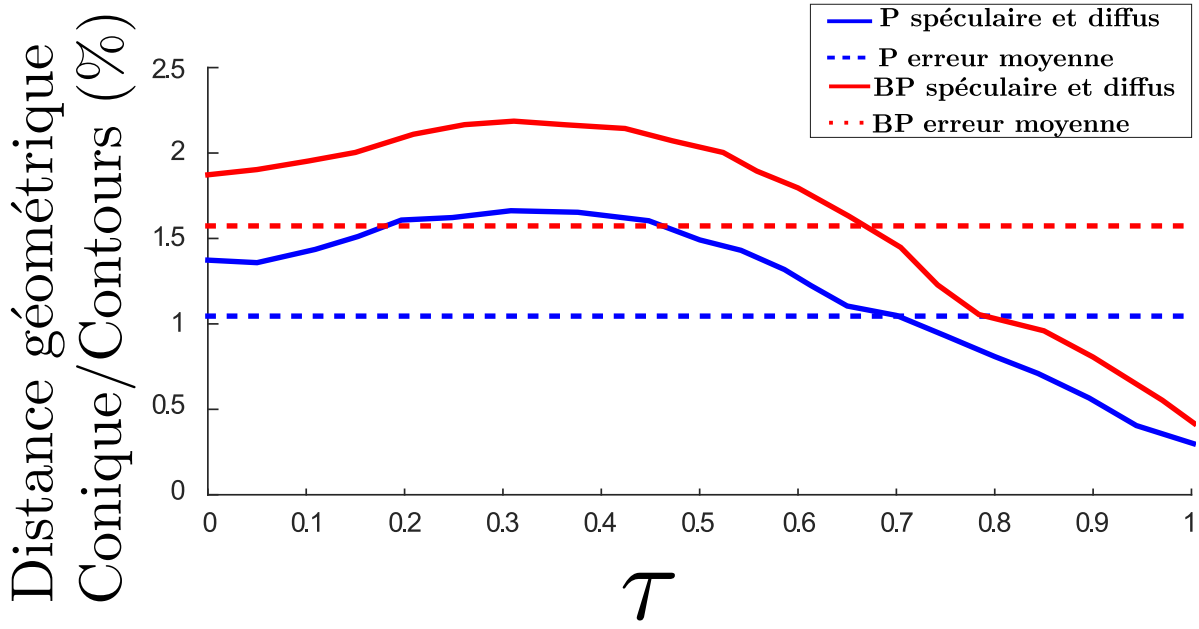
Afin d'évaluer la pertinence et la précision de notre modèle, deux expériences sont réalisées. Nous analysons, dans un premier temps, la performance de notre modèle pour la prédiction de spécularité sur des données de synthèse générées à partir du modèle de Phong et Blinn-Phong. Cette capacité de prédiction est ensuite évaluée sur cinq séquences réelles incluant des sources de lumière de différents types comme les ampoules ou les néons.

4.5.1 Validation du modèle sur des données de synthèse

Dans cette section, nous détaillons la validation de notre approximation de quadrique sur des données de synthèse. La caméra utilisée a une distance focale de $50mm$ avec le point principal centré dans l'image. La taille des pixels est de $0.45 \times 0.45mm$ pour une résolution de 1000×1000 pixels. Notre scène est constituée d'un plan réfléchissant de $20\times 20cm$ où l'on observe une spécularité en mouvement. Nous sélectionnons aléatoirement les points de vue et la position de la source de lumière au-dessus du plan. À partir des modèles d'illumination de Phong et Blinn-Phong pour tout $\tau \in [0, 1]$, 100 quadriques sont estimées par le biais de 100 points de vue. Nous calculons la distance géométrique entre le terme diffus et le terme spéculaire combinés avec la projection de quadrique associée au point de vue au sens de notre distance Ellipse/Point. Comme illustré à la figure 4.4, notre approximation par quadrique est efficace pour la prédiction de spécularité pour tout $\tau \in [0, 1]$ pour les modèles de Phong et Blinn-Phong avec respectivement 0.3% et 0.45% d'erreur en moyenne en fonction de la taille de la scène. Cette expérience a été également



(a) Source ponctuelle



(b) Source étendue

FIGURE 4.4 – Validation de notre modèle en mesurant l'erreur de prédiction sur des données synthétiques pour une source de lumière ponctuelle (a) et une source de lumière étendue (b). À partir de 100 points de vue, une quadrique fixe est reconstruite et sa projection est comparée avec les contours de la spécularité pour chaque point de vue en utilisant notre distance Ellipse/Point. Cette erreur est normalisée en fonction de la dimension de la scène.

réalisée pour des sources de lumière étendues composées de plusieurs sources ponctuelles pour les modèles d'illumination de Phong et Blinn-Phong pour une erreur moyenne de 1% et 1.6% respectivement. On peut noter que le processus de raffinement ne joue pas un

rôle significatif sur des données de synthèse où les conditions sont idéales.

4.5.2 Données réelles

Nous évaluons notre modèle JOLIMAS sur des séquences réelles avec trois sources de lumière différentes et cinq matériaux différents. La quadrique est reconstruite à partir des ellipses calculées par la méthode de Fitzgibbon et al. (1996). Les spécularités sont détectées en utilisant notre méthode de détection de spécularité Morgand et Tamaazousti (2014). Les poses de caméra sont calculées en utilisant une méthode de SLAM Tamaazousti et al. (2011b). Le processus de reconstruction est basé sur une méthode temps-réel et la prédiction est également réalisée en temps réel en projetant perspectivement la quadrique reconstruite pour une pose de caméra donnée. La projection de cette quadrique donne une ellipse correspondant à la prédiction pour la pose de caméra. Elle est obtenue par la formule :

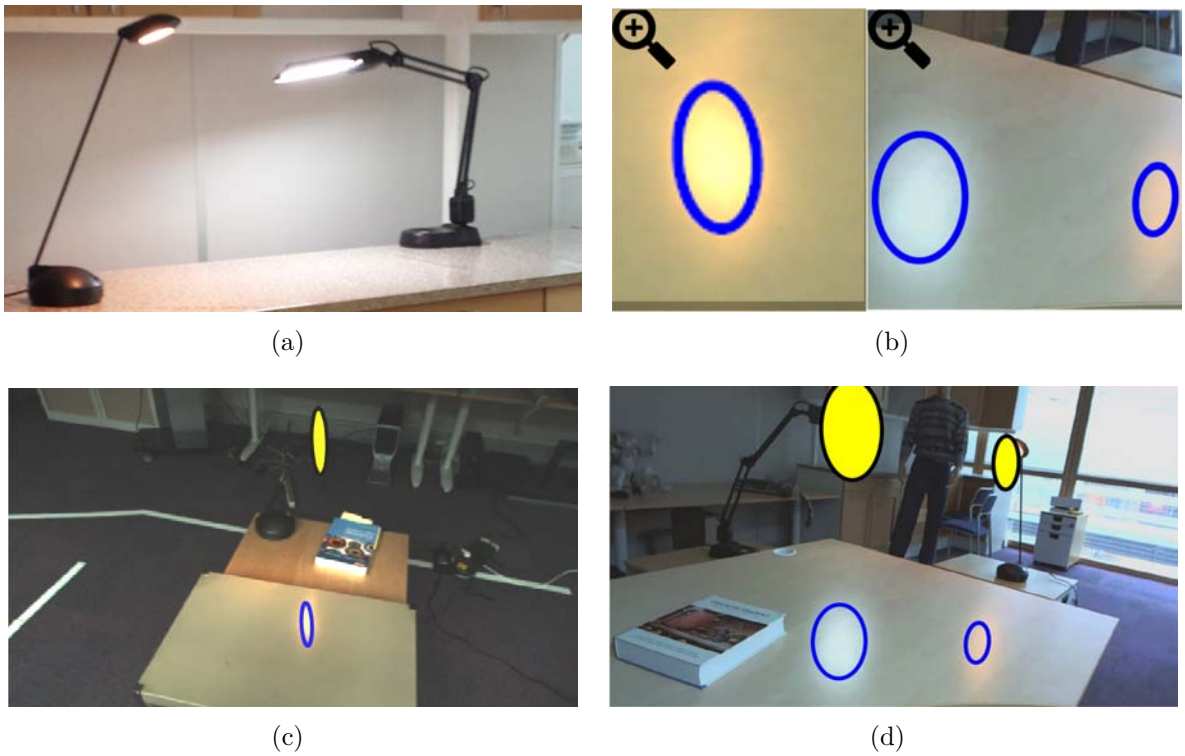
$$\mathtt{C} = (\Pi \mathtt{Q}^* \Pi^\top)^*. \tag{4.13}$$

La précision de notre modèle JOLIMAS est évaluée sur des données réelles afin de quantifier la capacité de prédiction de spécularités de notre approche. Nous utilisons notre distance 2D ellipse/point pour évaluer notre résultat de prédiction avec la vérité terrain. Afin de calculer la vérité terrain, nous avons manuellement annoté les spécularités sur les séquences. Nos résultats sont illustrés sur la table 4.1, la figure 4.5 et la figure 4.6.

TABLE 4.1 – Validation empirique de notre modèle et évaluation de sa capacité à prédire la forme des réflexions spéculaires dans les images pour 5 séquences réelles présentées dans les figures 4.5 et 4.6. L'erreur moyenne de notre prédiction par spécularité est calculée par notre distance Ellipse/Point entre les ellipses prédites et les contours de spécularités détectées dans l'image. Cette prédiction est évaluée en trois étapes : à l'étape d'initialisation, de la correction épipolaire et du raffinement non-linéaire. Nous observons une baisse significative de l'erreur pour chaque étape.

Distance 2D (en pixels) / Sequence	Initialisation	Correction épipolaire	Raffinement non-linéaire
Ampoule 1 (table en fer)	110.3	82.2	62.8
Néon (Tableau blanc)	210.6	80.9	31.6
Ampoule 2 Livre	85.2	66.7	28.9
Ampoule 1 + Néon (table en bois)	162.1	96.9	40.1
Ampoule 1, 2 + néon (plan de travail de cuisine)	240.3	144.9	60.3

Dans notre contexte, l'initialisation de quadrique n'est pas suffisante à elle seule pour prédire les spécularités de façon précise. Estimer des contours de spécularité précis est difficile en pratique ce qui impacte fortement la qualité de la reconstruction de la quadrique. En ajoutant un processus de correction épipolaire aux contours d'ellipses en entrée produit de meilleurs résultats. En combinant la correction épipolaire avec le raffinement non-linéaire, nous atteignons une erreur maximale de 62.8 pixels. Pour des séquences



(a) (b)

(c) (d)

FIGURE 4.5 – Nous observons qu'une spécularité présente une forme elliptique. Nous la prédisons en projetant une quadrique virtuelle fixe incluant les propriétés de lumière et de matériau. Cette figure présente des exemples de reconstruction de quadrique (ellipsoïde jaune) pour deux types de lumière (ampoule (c) et néon (d)) et de matériaux (table en fer (c) et plan de travail d'une cuisine (d)). Une fois notre modèle reconstruit, nous l'utilisons pour prédire les spécularités pour des points de vue inconnus en la projetant sous forme d'ellipse (en bleu). (b) illustre un gros plan de la prédiction dans les séquences respectives.

réelles, JOLIMAS est précis et fournit une prédiction de spécularité à partir de nouveaux points de vue pour différents types de sources de lumière et en temps réel.

4.6 Limites de l'approche

4.6.1 Surfaces miroirs et spécularités à forme non-elliptique

Dans le cas d'une surface parfaitement spéculaire telle qu'un miroir, nous ne pouvons pas garantir la forme elliptique de la spécularité puisque les sources de lumière peuvent avoir des formes diverses et variées (comme les lumières de plafond polygonales). Nous considérons le cas général des lumières de type ampoule et néons, ce qui inclut la plupart des sources de lumière que l'on retrouve dans un contexte intérieur. Même pour une surface de faible rugosité, notre méthode est toujours capable de reconstruire un modèle JOLIMAS. Une extension de notre méthode serait d'utiliser quelques idées de Heitz et al. (2016) qui modèle de façon efficace les sources de lumière polygonales.

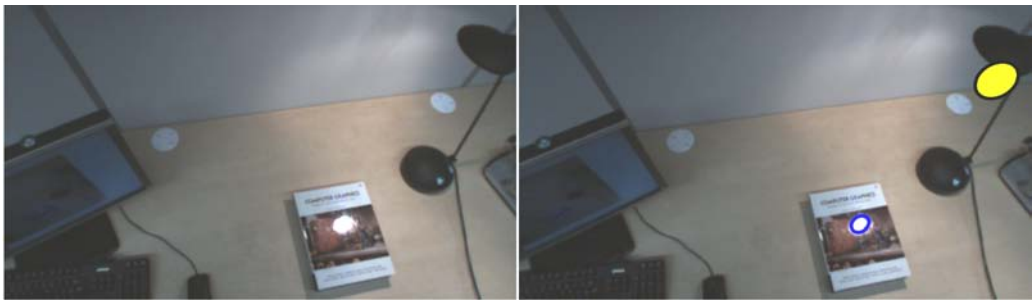
4.6.2 Influence de la rugosité de la surface

Dans le cas des surfaces à forte rugosité, puisque les normales de la surface varient fortement à un niveau microscopique (la surface paraît plane dans l'ensemble), la forme d'une spécularité peut s'éloigner de la forme elliptique initiale. Ces variations ne sont pas incluses dans notre modélisation de spécularité et peuvent affecter la qualité de la reconstruction de quadrique de JOLIMAS. Nous verrons dans le chapitre 6 la manière dont nous avons étudié le lien entre variation de normales et forme de spécularité sur la surface de façon similaire à l'approche de Blake et Brelstaff (1988).

4.7 Conclusion

Dans ce chapitre, nous avons présenté notre modèle de prédiction de spécularité appelé *JOint LIght-MAterial Specularity* (JOLIMAS). Notre modèle repose sur l'observation que les spécularités présentent une forme elliptique sur une surface plane. Nous avons démontré en détail que les modèles de réflexion de Phong et Blinn-Phong pouvaient être utilisés en bases théoriques pour confirmer la pertinence de notre approximation. Nous avons empiriquement validé cette approximation pour des sources de lumière ponctuelles et étendues. En géométrique projective, une projection de quadrique génère une ellipse. En démontrant l'existence d'une quadrique fixe dont la projection est associée au contour d'une spécularité sur un point de vue, nous avons montré un lien entre photométrie (lumières et matériaux) et géométrie multi-vue (quadrique). La quadrique est reconstruite à partir d'ellipses qui correspondent aux contours des spécularités. Ce modèle a été testé sur des données synthétiques et réelles pour des sources de lumière variées en forme et type telles que des lampes de bureau, ampoules et néons. Les contributions évoquées dans ce chapitre ont fait l'objet d'une publication dans une conférence internationale : ISMAR (2017a), 2 conférences nationales : deux articles à ORASIS (2015b), RFIA (2015) et un brevet accepté en juin 2016. Le détecteur de spécularité évoqué dans les annexes a fait l'objet d'une publication dans une conférence internationale Morgand et Tamaazousti (2014). Le chapitre suivant 5 évoque l'extension de notre modèle JOLIMAS à des surfaces courbes.

(a)



(b)

(c)

(d)

FIGURE 4.6 – Résultats de prédiction de spécularité sur quatre séquences réelles. Dans (a), un tableau blanc est illuminé par une lampe fluorescente, dans (b) un livre est illuminé par une lampe de bureau, dans (c), une table en bois est illuminée à la fois par une lampe et un néon et dans (d), un comptoir de cuisine est illuminé par deux lampes de bureau et un néon. Notre quadrique est projetée perspectivement sur les différents plans afin de correspondre aux spécularités dans l'image (ellipses bleues). Nous calculons également le symétrique de la quadrique à partir des normales des plans (tableau blanc, table en bois, comptoir de cuisine et couverture de livre) afin de montrer le rapport entre la quadrique de JOLIMAS (en jaune) et les sources de lumière. On remarque que le symétrique de la quadrique est très proche de la position des sources de lumière.

CHAPITRE 5

Un modèle dual pour les surfaces convexes à courbure constante

Prédire les spécularités dans une image, à partir d'une pose de caméra et de la géométrie de la scène connues, est un problème complexe et ouvert. Cette prédiction est néanmoins essentielle pour de nombreuses applications en réalité augmentée. L'approche JOLIMAS, présentée au chapitre 4, répond partiellement à ce problème sous l'hypothèse que les spécularités ont une forme elliptique et que la scène est composée d'un objet plan. JOLIMAS modélise une spécularité comme étant l'image d'une quadrique fixe dans l'espace. Dans ce chapitre, nous présentons JOLIMAS dual, un modèle géométrique qui répond à la limitation aux plans. Ce modèle repose sur le fait que les spécularités conservent leurs formes elliptiques sur des surfaces convexes et que chaque surface peut être divisée en plusieurs parties convexes. La géométrie de JOLIMAS dual utilise une quadrique pour chaque surface convexe et chaque source lumineuse et prédit les spécularités en utilisant des caméras virtuelles permettant de gérer la courbure de la surface. Nous évaluons l'efficacité et la précision du JOLIMAS dual sur de multiples séquences synthétiques et réelles composées de divers types d'objets et sous différents éclairages.

5.1 Introduction

Le SLAM (*Simultaneous Localization And Mapping*) a récemment gagné en maturité. Son intégration dans l'industrie a fortement augmenté avec l'aide notamment de nombreuses méthodes de localisation de caméra disponibles publiquement telles que Engel et al. (2016); Tan et al. (2013) et de reconstruction 3D telle que Newcombe et al. (2011); Whelan et al. (2015). La plupart des approches SLAM considèrent que la scène observée est Lambertienne (surface mate), une condition qui n'est pas vérifiée dans une scène contenant des objets à matériaux spéculaires tels que du métal, du plastique ou de la porcelaine. En raison d'une interaction proche de celle d'un miroir entre les sources de lumière et les matériaux spéculaires, des spécularités apparaissent et provoquent des changements brusques d'intensité dans l'image.

Dans le chapitre précédent 4, nous avons montré qu'une spécularité sur une surface plane peut être convenablement approchée par une ellipse sous un éclairage provenant

d'une ampoule ou d'un néon. Nous avons présenté notre modèle empirique appelé *JOint LIght-MAterial Specularity* (JOLIMAS) pour abstraire l'interaction lumière-matière et traiter le problème de prédiction de spécularité géométriquement. Ce modèle utilise une quadrique fixe dont la projection prédit la forme des spécularités sur des points de vue existants et nouveaux. La reconstruction de quadrique est obtenue avec au moins trois points de vue. Nous nous référerons à ce modèle comme étant *JOLIMAS primal*.

5.2 Approche proposée

Nous proposons un nouveau modèle, *JOLIMAS dual*, qui est une généralisation de la version primale aux surfaces convexes par morceaux tout en conservant sa fonctionnalité sur des surfaces planes. L'idée principale est que les spécularités conservent leur forme elliptique sur des surfaces convexes. Notre modèle se base sur une représentation virtuelle des caméras qui associe une caméra virtuelle à chaque spécularité avec la pose de la caméra réelle et la géométrie de la surface connues. Du fait que chaque surface peut être divisée en parties convexes, nous pouvons modéliser et prédire les spécularités sur des objets plus complexes comme illustrés à la figure 5.7. En pratique, nous représentons une surface par un maillage et utilisons un procédé de décomposition convexe approximative comme expliqué dans la section 5.6. Nous reconstruisons une quadrique pour chaque partie convexe et chaque source de lumière. Nous expliquons le détail de notre modèle de prédiction de spécularité et le calcul de notre représentation pour de nouveaux points de vue.Cette prédiction n'est pas triviale, car les caméras virtuelles utilisées pour la prédiction requièrent le calcul du point d'intensité maximale d'une spécularité inconnue pour un nouveau point de vue. Nous détaillerons le calcul de ce point particulier par utilisation de notre quadrique reconstruite. L'efficacité et la précision de notre méthode sont validées par de nombreux exemples synthétiques et réels incluant divers types d'objets et diverses conditions lumineuses dans la section 5.7.

Notre modèle étend notre approche précédente présentée dans le chapitre 4 aux surfaces non-planes tout en reprenant l'approximation de spécularités de forme elliptique même pour des sources étendues. Nous proposons JOLIMAS dual qui utilise des caméras virtuelles obtenues par symétrie des caméras réelles selon le plan tangent à un point particulier de la surface à l'opposé de JOLIMAS primal, qui utilise des caméras réelles. Notre approche améliore la reconstruction en matière de généricité avec un nombre de points de vue nécessaires inférieur à la version primale grâce à une meilleure répartition des caméras virtuelles. De nouvelles caméras virtuelles peuvent être ajoutées par l'utilisation de spécularités sur des objets similaires (courbure et matériaux) à partir de la même source de lumière.

5.3 JOLIMAS dual

5.3.1 Notation et formalisation

Une scène 3D typique contient une caméra de pose $\Pi = \begin{bmatrix} \mathtt{R} & \mathbf{V} \end{bmatrix}$ et une source de lumière $\mathbf{L}$. Un point de surface est noté $\mathbf{P}$ et de normale $\hat{\mathbf{N}}(\mathbf{P})$. Nous utilisons $\hat{\mathbf{L}}$ en tant que vecteur normalisé désignant le rayon de lumière incident de la source de lumière $\mathbf{L}$ tel que $\hat{\mathbf{L}}(\mathbf{P}) = \mu(\mathbf{L} - \mathbf{P})$ et $\hat{\mathbf{V}}$ le vecteur normalisé désignant la caméra $\mathbf{V}$ tel que

$\hat{\mathbf{V}}(\mathbf{P}) = \mu(\mathbf{V} - \mathbf{P})$. Le plan tangent $\mathbf{P}$ est défini par $\pi_{\mathbf{P}} = \begin{bmatrix} \hat{\mathbf{N}}(\mathbf{P}) \\ -d \end{bmatrix}$ avec d la distance du plan à l'origine. Chaque entité virtuelle calculée par symétrie est notée avec $\sim$ (tilde en gras contrairement au tilde utilisé pour les coordonnées homogènes), tel que la position de la caméra virtuelle est notée $\tilde{\mathbf{V}}$ et la pose virtuelle $\tilde{\Pi}$.

5.3.2 Similarité du comportement d'une surface spéculaire avec un miroir

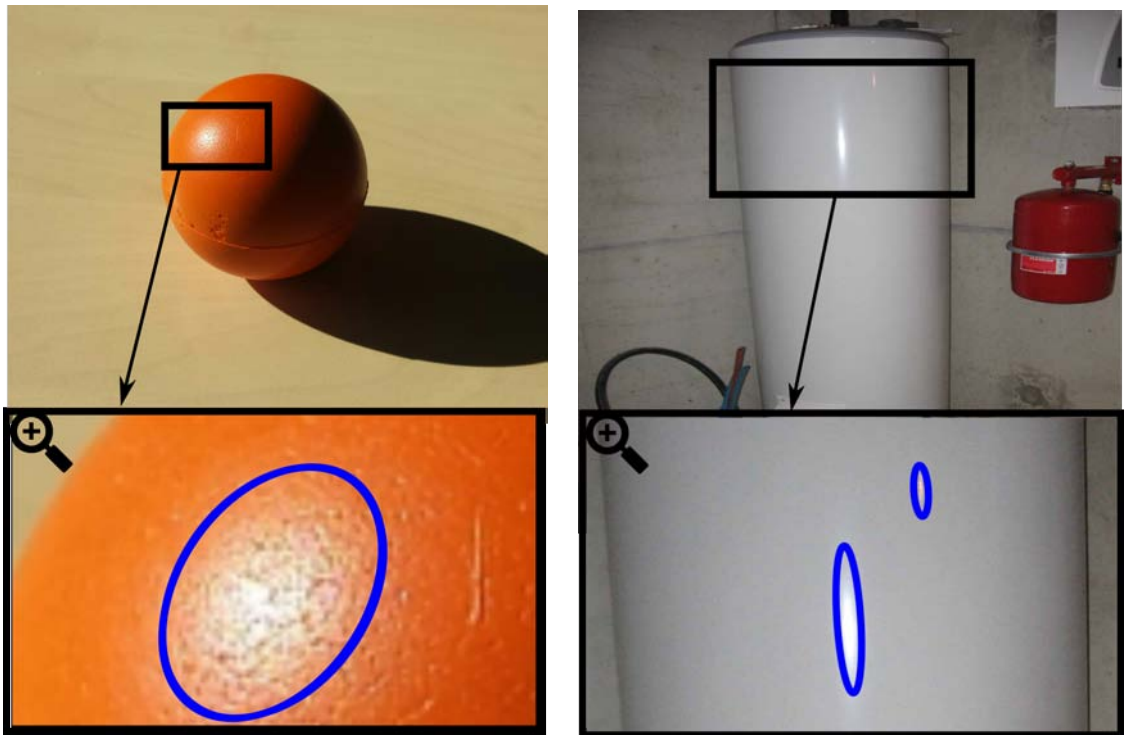
Sur une surface spéculaire plane, nous pouvons interpréter une spécularité comme étant une image miroir de la source de lumière (comportement semi-miroir de la surface). Cependant, pour des surfaces courbes, cette image miroir est distordue et incohérente avec la vraie position et forme de la source de lumière. À partir des spécularités, nous ne pouvons pas reconstruire un objet 3D cohérent localisé autour de la source de lumière dans l'état. L'idée principale de JOLIMAS dual est d'utiliser des caméras virtuelles pour reconstruire la quadrique de façon cohérente avec la source de lumière. Ces caméras virtuelles sont calculées par symétrie de la vraie caméra avec la surface. Contrairement aux surfaces planes, la normale varie le long de la surface courbe, ce qui rend le choix de la caméra virtuelle ambigu. Pour une pose de caméra et une spécularité données, quel point de la surface devons-nous choisir pour calculer cette caméra virtuelle ?

5.3.3 Point d'intensité maximale

Sur des surfaces convexes, plusieurs propriétés peuvent être mises en évidence. Premièrement, une spécularité conserve une forme elliptique de la même manière que sur les surfaces planes comme illustrée à la figure 5.1. Deuxièmement, une source de lumière est associée à une unique spécularité. Enfin, si nous transformions une surface plane en une surface convexe, la forme de la spécularité ne serait affectée que par son échelle. Plusieurs études comme celles du chapitre 4 et Hyeongwoo Kim (2013); Tian et Clark (2013) ont souligné l'importance d'un point spécifique appelé le point d'intensité maximale qui est unique pour chaque spécularité pour des surfaces planes, mais aussi convexes. Pour chaque direction autour de ce point, l'intensité de la spécularité décroit. Ce point est très important en optique, car pour une pose de caméra et une position de source lumineuse données, il suit la loi de la réflexion :

$$\hat{\mathbf{L}}(\mathbf{P}) = 2(\hat{\mathbf{N}}(\mathbf{P})^\top \hat{\mathbf{V}}(\mathbf{P}))\hat{\mathbf{N}}(\mathbf{P}) - \hat{\mathbf{V}}(\mathbf{P}). \tag{5.1}$$

Ceci implique que le point d'intensité maximale n'est pas affecté par la distorsion induite par la courbure de la surface. Si nous calculons le symétrique de la caméra par rapport au plan tangent à ce point, nous obtenons une caméra virtuelle qui pointe directement à la source de lumière en passant à l'intérieur de la spécularité comme illustrée à la figure 5.2. En calculant la caméra virtuelle pour chaque spécularité, nous pouvons ainsi reconstruire une quadrique localisée à une position proche de la source de lumière et dont la projection correspond à la forme des spécularités. Dans l'annexe C, nous démontrons que le point d'intensité maximale du modèle d'illumination de Phong est unique et localisé à l'intersection de la ligne contenant la position de la source de lumière $\mathbf{L}$ et le symétrique de la position de la caméra $\tilde{\mathbf{V}}$ dans le cas de surfaces planes. Pour des surfaces convexes,



(a) (b)

FIGURE 5.1 – Spécularités sur des surfaces convexes telles qu'une balle en mousse (a) et un ballon d'eau chaude (b). L'hypothèse de spécularités elliptiques est toujours valide pour ces surfaces comme le montre les ellipses sur les gros plans (a) et (b). Pour chaque surface convexe, il y a une source de lumière pour chaque spécularité (une source de lumière dans (a) et deux sources de lumière dans (b)).

cette propriété est conservée si le point de la surface utilisé pour calculer $\tilde{\mathbf{V}}$ coïncide avec le point d'intersection comme illustré à la figure 5.2.

Notons qu'en pratique, le calcul du point le plus intense se doit d'être le plus précis possible, car les caméras virtuelles sont sensibles aux erreurs de normale. Nous présentons une méthode d'estimation de ce point d'intensité maximale à partir d'aprioris géométriques sur le maillage dans la section 5.5.2.

5.4 Représentation par caméras virtuelles

Nous calculons JOLIMAS dual à partir de la caméra virtuelle estimée à partir du point le plus intense pour chaque image sur la surface. Nous utilisons une estimation initiale du point du plus intense $\mathbf{P_B}$, à partir d'une détection de spécularité initiale, en prenant le centre de gravité des contours de la spécularité. En pratique, ce point le plus intense n'est pas unique dans l'image, car la spécularité provoque souvent une saturation de l'image. Une pose de caméra virtuelle $\tilde{\Pi}$ est calculée en suivant le formalisme de Long et al. (2015) sur une pose de caméra réelle Π tel que :

$$\tilde{\Pi} = \Pi \mathtt{S} \text{ avec } \mathtt{S} = \begin{bmatrix} \mathtt{I}_3 - 2\hat{\mathbf{N}}(\mathbf{P})\hat{\mathbf{N}}(\mathbf{P})^\top & \mathbf{0} \\ -2d\hat{\mathbf{N}}(\mathbf{P}) & 1 \end{bmatrix}^\top . \tag{5.2}$$

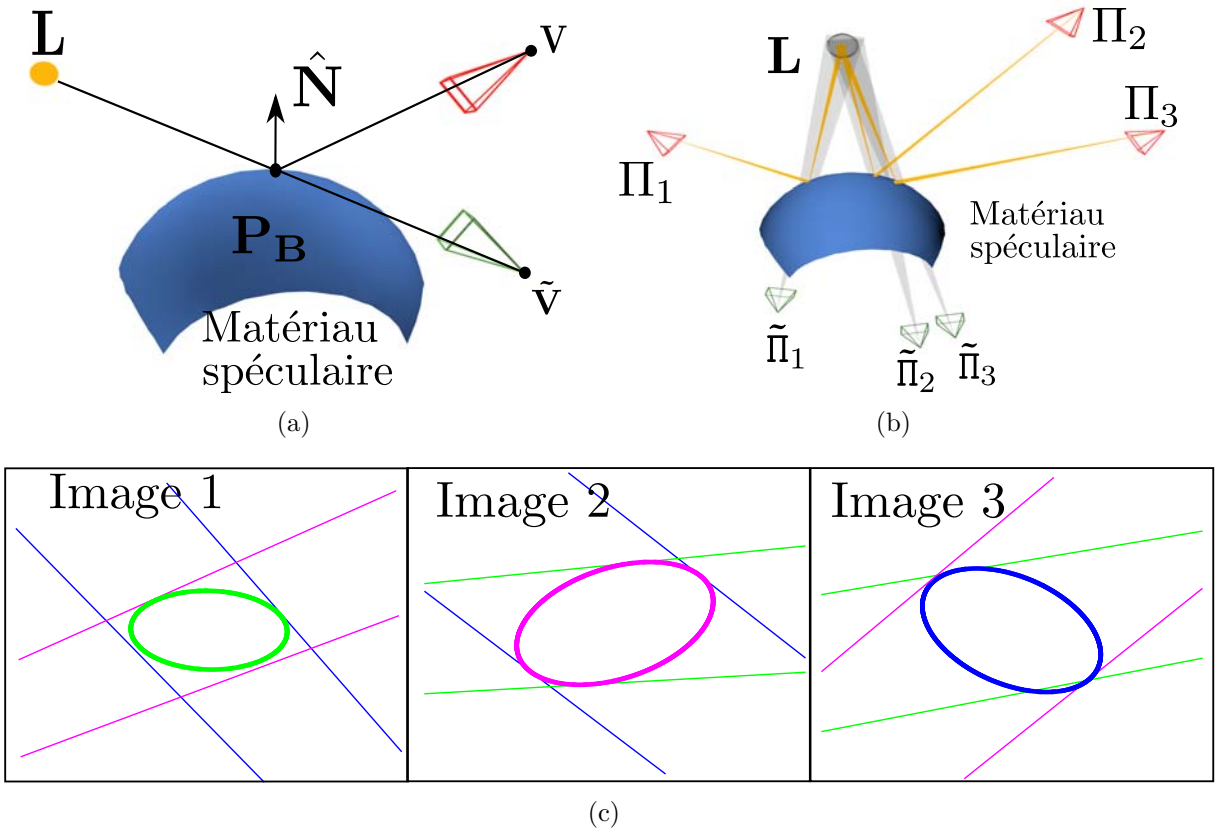


FIGURE 5.2 – JOLIMAS dual utilise des caméras virtuelles produisant des ellipses épipolairement consistantes. Dans (a), nous montrons la loi de la réflexion sur une surface convexe. Le point le plus intense $\mathbf{P_B}$ correspond au point de la surface où l'intersection de la ligne contenant la position de la caméra virtuelle $\tilde{\mathbf{V}}$ (en vert) et la source de lumière $\mathbf{L}$ est confondue avec le point de la surface utilisé pour calculer la caméra virtuelle. Dans (b), nous présentons JOLIMAS dual sur une surface convexe pour 3 poses de caméra données Π_1, Π_2 et Π_3. Les poses de caméra (en vert) $\tilde{\Pi}_1$, $\tilde{\Pi}_2$ et $\tilde{\Pi}_3$ permettent de reconstruire une quadrique proche de la source de lumière réelle $\mathbf{L}$. Dans (c), nous montrons que notre représentation virtuelle de caméra permet d'obtenir des ellipses respectant la géométrie épipolaire.

Pour une source de lumière unique, le pipeline de l'estimation de notre modèle est :

- Détection de spécularité calculant simultanément le point le plus intense pour chaque spécularité.
- Estimation des caméras virtuelles initiales par symétrie des caméras réelles avec la surface pour chaque point d'intensité maximale. Chaque caméra est symétrisée par rapport au point d'intensité maximale que cette caméra observe.
- Correction épipolaire à partir de Reyes et Bayro-Corrochano (2005) et reconstruction de la quadrique à partir de Cross et Zisserman (1998) pour au moins 3 points de vue.
- Correction des points d'intensité maximale de chaque spécularité à partir de la reconstruction de quadrique (décrite dans 5.5.2).
- Raffinement non-linéaire à partir de tous les points de vue disponibles.

Un exemple de notre représentation de caméra virtuelle est présenté à la figure 5.2(b). Ce pipeline est répété pour chaque source de lumière. Un récapitulatif du pipeline est illustré à la figure 5.3.

5.5 Triangulation de la position de la source lumineuse

5.5.1 Estimation initiale

À partir d'un minimum de deux points d'intensité maximale $\mathbf{P_{B,1}}$ et $\mathbf{P_{B,2}}$ provenant de deux spécularités (dans une même image ou une paire d'images clés), il est possible de trianguler la position de source de lumière $\mathbf{L}$. À ces deux spécularités, nous associons deux poses de caméra virtuelles $\tilde{\Pi}_1 = \begin{bmatrix} \mathtt{R}_1 & \mathbf{V_1} \end{bmatrix}$ et $\tilde{\Pi}_2 = \begin{bmatrix} \mathtt{R}_2 & \mathbf{V_2} \end{bmatrix}$. Les vecteurs $\mathbf{d_1}$ et $\mathbf{d_2}$ représentent les directions des rayons partant des centres optiques des caméras virtuelles par les points d'intensité maximale dans les images. Pour la triangulation, nous utilisons la méthode du point médian ou *midpoint* qui est une méthode rapide pour trianguler un point à partir de plusieurs rayons. De meilleures méthodes de triangulation peuvent être utilisées comme évoqué dans Hartley et Sturm (1997). Pour la méthode du point médian, on utilise les matrices de projection $\mathtt{P} = (\mathtt{M}| - \mathtt{M}\mathbf{t})$ et $\mathtt{P}' = (\mathtt{M}'| - \mathtt{M}'\mathbf{t}')$. Pour un point $\mathbf{u}$, le rayon associé est décrit sous la forme $\mathbf{t} + \alpha\mathtt{M}^{-1}\mathbf{u}$. A partir de deux images, les deux rayons sont sécants dans l'espace (si non parallèles). Nous obtenons l'équation suivante $\alpha\mathtt{M}^{-1}\mathbf{u} - \alpha'\mathtt{M}'^{-1}\mathbf{u}' = \mathbf{t}' - \mathbf{t}$. En utilisant une méthode de moindres carrés linéaire, nous minimisons la distance carrée entre les deux rayons. Le point médian est donné par la formule $(\mathbf{t} + \alpha\mathtt{M}^{-1}\mathbf{u} + \mathbf{t}' + \alpha'\mathtt{M}'^{-1}\mathbf{u})/2$. Cette reconstruction de la position de la source de lumière $\mathbf{L}$ peut être réalisée à partir d'au moins trois observations en prenant les différentes paires de caméras virtuelles et en faisant la moyenne des points triangulés obtenus par chaque paire. En pratique, $\mathbf{P_{B,1}}$ et $\mathbf{P_{B,2}}$ sont initialisés en prenant le centre de gravité de la zone d'intensité maximale (lorsque le point d'intensité maximale n'est pas unique) de la spécularité associée.

5.5.2 Raffinement de la position de la lumière et du point d'intensité maximale

Selon l'équation (5.1), un point $\mathbf{P}$ de la surface est considéré comme étant le point d'intensité maximale si $\mu(\tilde{\mathbf{V}} - \mathbf{P})$ et $\mu(\mathbf{L} - \mathbf{P})$ sont colinéaires. Cela signifie que le point de la surface $\mathbf{P}$ utilisé pour calculer la caméra virtuelle doit être identique à l'intersection de la ligne contenant la caméra virtuelle $\tilde{\mathbf{V}}$ et la source de lumière $\mathbf{L}$ avec le plan tangent $\pi_{\mathbf{P}}$ d'une surface S au point $\mathbf{P}$. Cette propriété est illustrée à la figure 5.4(a). Ce point d'intersection $\mathbf{P}_I$ est défini par :

$$\mathbf{P}_I = \tilde{\mathbf{V}} + \frac{\tilde{\mathbf{V}}^\top \hat{\mathbf{N}}(\mathbf{P}) - d}{(\mathbf{L} - \tilde{\mathbf{V}})^\top \hat{\mathbf{N}}(\mathbf{P})}(\mathbf{L} - \tilde{\mathbf{V}}) \tag{5.3}$$

avec

$$\tilde{\mathbf{V}} = \begin{bmatrix} \mathtt{I}_3 - 2\hat{\mathbf{N}}(\mathbf{P})\hat{\mathbf{N}}(\mathbf{P})^\top & -2d\hat{\mathbf{N}}(\mathbf{P}) \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ 1 \end{bmatrix}.$$

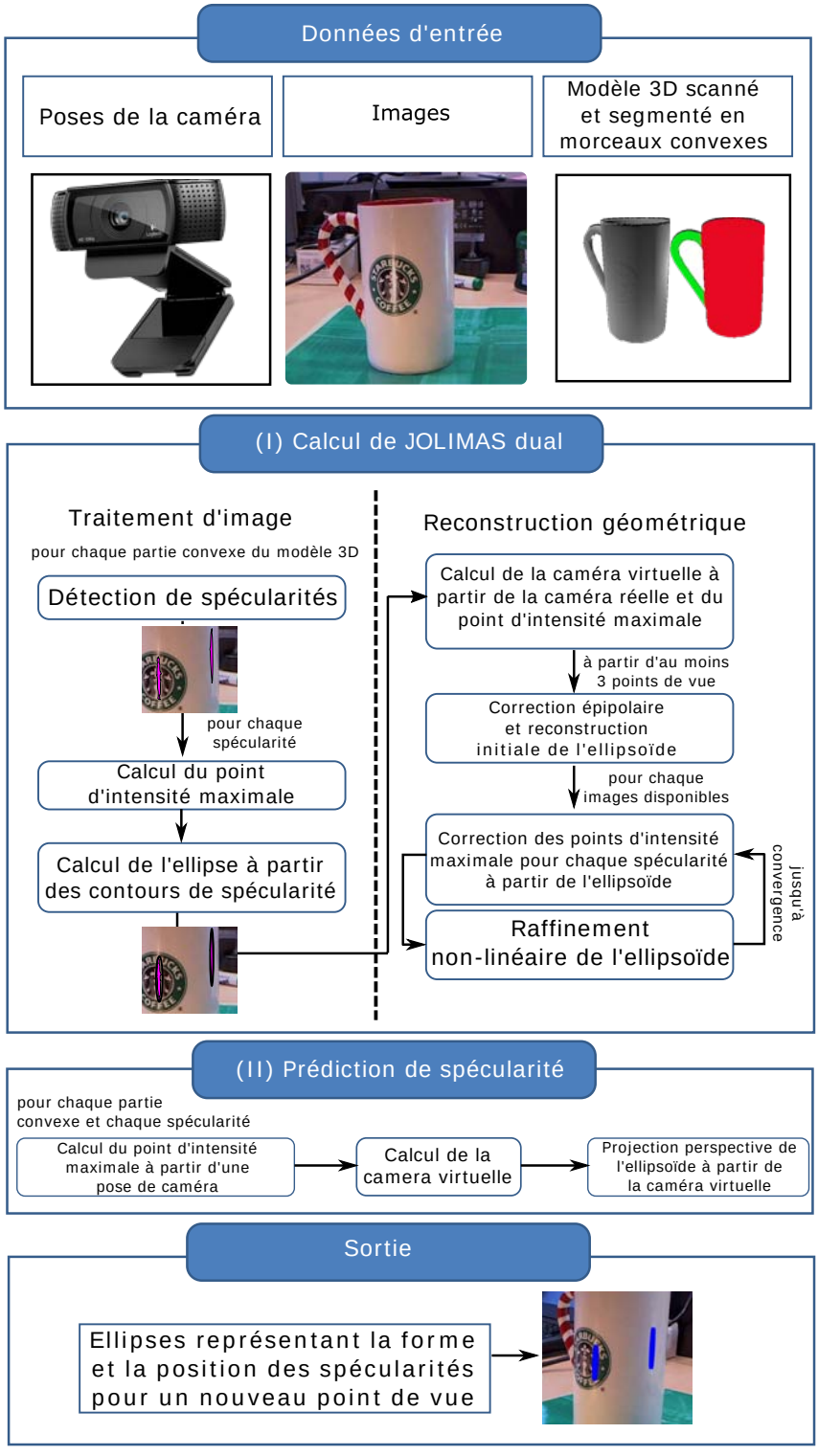


FIGURE 5.3 – Pipeline de notre méthode décrivant les données d'entrée, le calcul de notre modèle dual JOLIMAS et le processus de prédiction de spécularité. Notre méthode produit une ellipse décrivant la forme et la position de la spécularité dans l'image. Cette représentation géométrique abstrait l'interaction de la source de lumière (intensité, position, forme, couleur) avec le matériau (réflectance et rugosité) de la surface.

Le calcul du point le plus intense doit être réalisé pour chaque surface convexe et chaque source de lumière. Notre estimation du point d'intensité maximale est décrite dans l'algorithme 1.

Algorithm 1 Version simplifiée du calcul du point d'intensité maximale à partir d'une pose de caméra, le modèle CAO et la position de la source de lumière obtenue par triangulation la première fois puis en prenant le centre de la quadrique par la suite.

procedure CALCULPOINTINTENSITÉMAX(*Pose, Modèle, L*)
  $BP \leftarrow initial_BP$
  do
    $VirtualPose \leftarrow compute_virtual_pose(Pose, BP)$
    $P \leftarrow model_intersection(L, VirtualPose, Model)$
    $search_direction \leftarrow \frac{P-BP}{\|P-BP\|}$
    $BP \leftarrow BP + search_direction$
  while $BP \neq P$
return BP

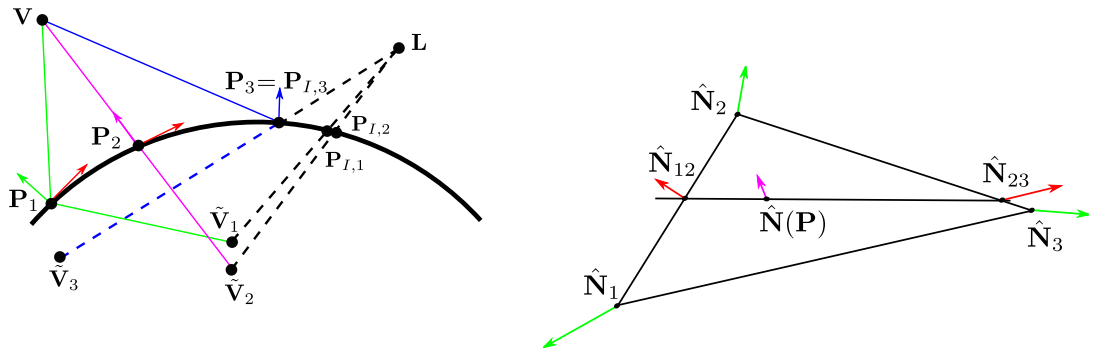
5.5.3 Interpolation de normale

Afin de calculer des caméras virtuelles cohérentes, une information de normale de surface précise est importante. Sur un maillage, cette information est fournie uniquement par les arêtes et les faces du modèle CAO. Afin de trouver le point d'intensité maximale le plus juste, nous devons retrouver la variation de normale précisément sur les différents polygones du maillage, pour répondre à ce problème, nous utilisons une interpolation de normale entre les différents polygones du maillage ce qui est une approche semblable à l'interpolation de normale utilisée par Phong (1975). Nous commençons à partir d'un point $\mathbf{P}$ de normale inconnue $\hat{\mathbf{N}}(\mathbf{P})$ sur un polygone du maillage. Après avoir calculé les normales $\hat{\mathbf{N}}_i$ pour chaque sommet i du triangle, nous traçons une ligne incluant $\mathbf{P}$ dans une direction donnée (l'axe $\mathbf{y}$ par exemple) qui intersecte plusieurs arêtes du polygone. Pour ces points d'intersection, nous calculons leurs normales par interpolation linéaire. La normale finale $\hat{\mathbf{N}}(\mathbf{P})$ est calculée à partir de l'interpolation linéaire des normales des points d'intersection le long de la ligne, comme illustrée à la figure 5.4(b).

5.6 Surfaces convexes par morceaux

5.6.1 Principe

Pour chaque source de lumière, nous avons un unique point d'intensité maximale et plus généralement une spécularité unique pour chaque surface convexe. Cette hypothèse est semblable au comportement d'un miroir de forme convexe. Dans ce miroir, la réflexion d'un objet donné sera distordue, mais l'image de celui-ci sera unique. Nous pouvons représenter un modèle 3D d'un objet comme étant la combinaison de surfaces convexes. En pratique, nous reconstruisons une quadrique pour chaque surface convexe et chaque source de lumière.



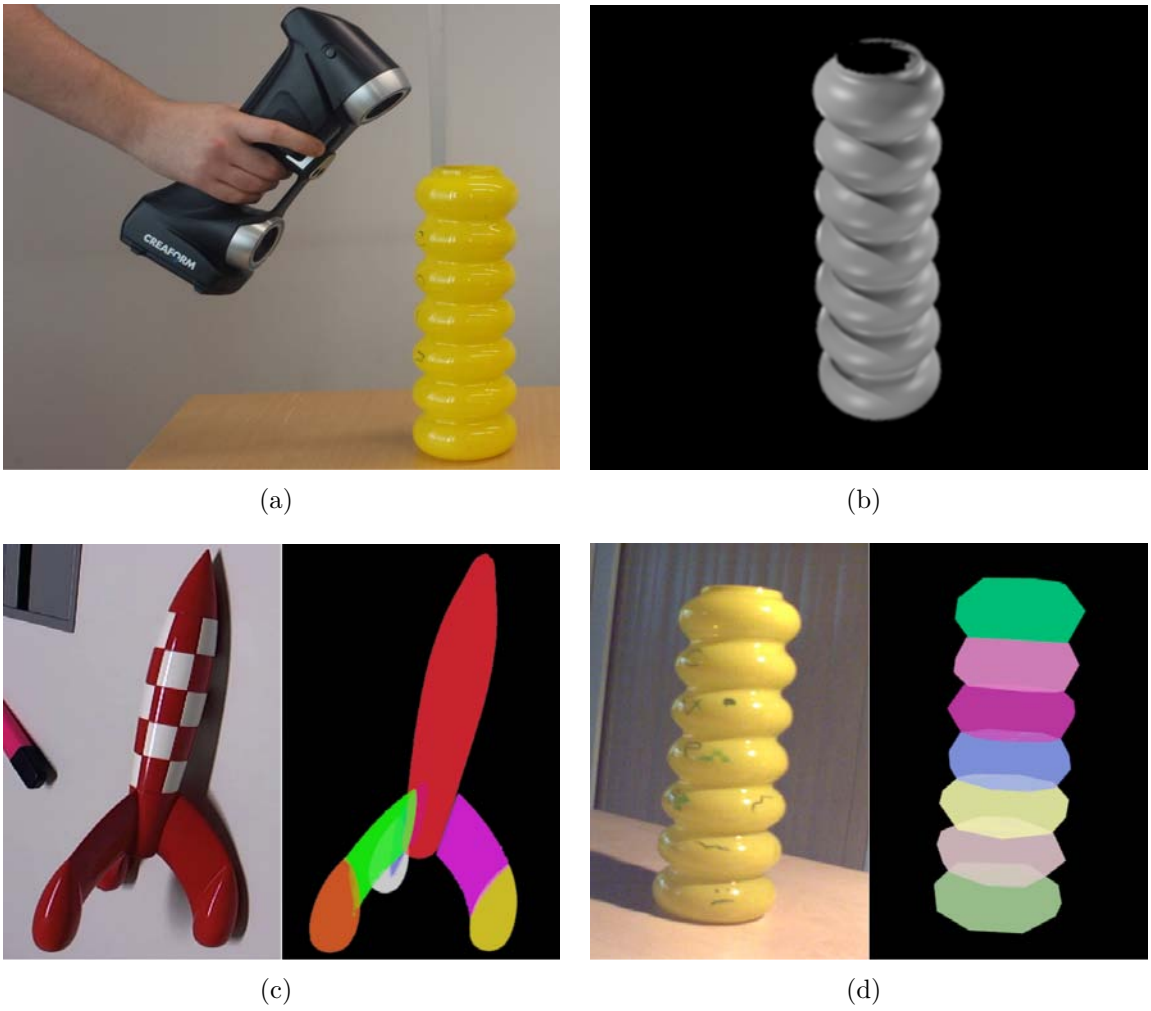
(a) (b)

FIGURE 5.4 – Calcul du point le plus intense et interpolation de normale. Dans (a), notre algorithme pour calculer le point le plus intense est illustré. À partir d'une initialisation $\mathbf{P}_1$, nous calculons la position de la caméra virtuelle $\tilde{\mathbf{V}}_1$ (symétrique de $\mathbf{V}$) par rapport au plan tangent à la surface en $\mathbf{P}_1$ et la ligne contenant $\tilde{\mathbf{V}}_1$ et $\mathbf{L}$. Puisque l'intersection $\mathbf{P}_{I,1}$ entre la ligne et la surface S n'est pas égale à $\mathbf{P}_1$, nous itérons à la face adjacente dans la direction du point d'intersection. Ce procédé s'arrête quand la distance entre $\mathbf{P}$ et $\mathbf{P}_I$ est minimale. Pour améliorer l'estimation sur la face restante, nous calculons une normale plus précise en utilisant une interpolation de normale dans (b). Dans cet exemple, à partir d'un polygone comportant trois sommets avec normales associées, la normale $\hat{\mathbf{N}}(\mathbf{P})$ en $\mathbf{P}$ est calculée en interpolant les normales sur la ligne horizontale contenant $\mathbf{P}$ avec les arrêtes du polygone donnant $\hat{\mathbf{N}}_{12}$ et $\hat{\mathbf{N}}_{23}$. A partir d'une interpolation linéaire de ces deux normales, $\hat{\mathbf{N}}(\mathbf{P})$ est calculé.

5.6.2 Décomposition d'un modèle CAO en surfaces convexes approximatives

Selon Lien et Amato (2007), chaque maillage 3D peut être approximé par la combinaison de plusieurs surfaces convexes. La décomposition convexe approximative est souvent utilisée dans les moteurs physiques pour accélérer la détection de collisions. Notre objectif est de déterminer une décomposition convexe d'une surface S, qui consiste à la partitionner en un ensemble minimal de sous-surfaces convexes. Ce procédé a pour but de déterminer la partition des triangles du maillage avec un nombre d'amas minimal tout en assurant que chaque amas a une concavité plus petite qu'un seuil prédéfini. Une décomposition convexe exacte est un problème *NP-complet* et ne serait pas utilisable du fait d'une fragmentation trop importante des amas. Nous utilisons ainsi une bibliothèque de décomposition convexe approximative appelée V-HACD [1] pour segmenter notre maillage en plusieurs surfaces convexes. En utilisant cette décomposition, notre modèle peut être généralisé aux objets non-plans en reconstruisant une quadrique pour chaque surface convexe. Nous illustrons cette décomposition sur la figure 5.5(c) et 5.5(d). Notre maillage est reconstruit à partir du scanneur HandySCAN 3D de l'entreprise Creaform [2] comme illustré aux figures 5.5(a) et 5.5(b).

1. github.com/kmammou/v-hacd
2. www.creaform3d.com/en



(a) (b)

(c) (d)

FIGURE 5.5 – Numérisation d'un objet et sa décomposition convexe approximative. Dans (a), nous utilisons le capteur HandySCAN 3D sur un vase et nous montrons la reconstruction du maillage en (b). Pour deux maillages scannés à partir d'objets réels présentés dans la figure 5.7, nous montrons les résultats de l'algorithme de décomposition convexe approximative. Dans (a), le maillage de la réplique de fusée est décomposé en 7 surfaces convexes, ainsi que le vase jaune. Nous avons une spécularité par surface convexe et par source de lumière et nous reconstruisons une quadrique virtuelle à chaque fois.

5.7 Résultats expérimentaux

Dans cette section, la prédiction de spécularité est testée sur des données synthétiques et réelles pour évaluer les améliorations par rapport au modèle JOLIMAS primal présenté au chapitre 4, la sensibilité du modèle aux changements de courbure et la robustesse de notre méthode à un point d'intensité maximale bruitée. La prédiction de spécularité est évaluée en utilisant la distance 2D entre les contours de la spécularité et l'ellipse prédite comme définie dans le chapitre précédent 4. Cette erreur de prédiction est décrite en tant que pourcentage en calculant l'erreur moyenne par point de contour comparé à la taille de l'image.

5.7.1 Comparaison avec JOLIMAS primal

Nous utilisons une séquence comportant 300 images illustrées à la figure 5.6(a). Dans cette séquence, une surface plane est courbée progressivement afin d'évaluer la fonctionnalité de notre modèle dual sur des surfaces planes et convexes. Pour chaque image, nous reconstruisons une quadrique à partir de 6 points de vue et nous calculons la distance 2D entre la spécularité prédite et la spécularité détectée. JOLIMAS dual a une erreur moyenne de 1% des pixels en comparaison avec JOLIMAS primal qui accumule 33% d'erreur par ellipse.

5.7.2 Sensibilité aux changements de courbure

La forme d'une spécularité varie en fonction des changements géométriques et plus particulièrement les changements de courbure. Il est difficile de prédire de façon précise la spécularité pour un nouveau point de vue avec un changement de courbure de la scène. En utilisant notre séquence précédente, nous avons testé la sensibilité de notre méthode aux changements de courbure en reconstruisant une quadrique à partir de la première image (surface plane) et en prédisant la spécularité pour toutes les images. Même si l'erreur augmente, l'information fournie par la prédiction reste cohérente avec une erreur moyenne de 7% comme illustrée à la figure 5.6.

5.7.3 Prédiction de spécularité sur des données réelles

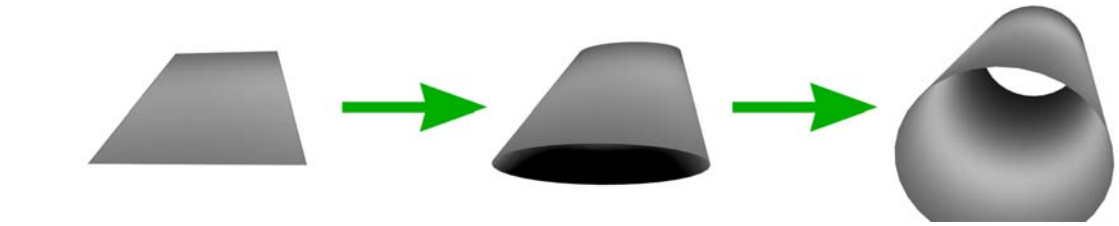
Pour les séquences réelles, des expériences qualitatives sont réalisées sur la prédiction de spécularité en comparant l'ellipse prédite avec les contours de la spécularité associée pour chaque surface présentant des spécularités et avec la prédiction de spécularité de JOLIMAS primal.

Contexte. Nous choisissons différents objets avec des courbures et matériaux variés sous différentes conditions de lumière. Pour les séquences de la réplique de fusée, du mug et du vase jaune, les images ont une résolution de 1280×720. Les poses de caméra sont calculées à partir d'une des variantes Tamaazousti et al. (2011b,a); Gay-Bellile et al. (2012); Besbes et al. (2012) du framework SLAM contraint Tamaazousti et al. (2016) proposé dans la thèse de Tamaazousti (2013). Plus spécifiquement, nous utilisons une amélioration de celui-ci adapté aux objets courbes de Loesch et al. (2015). Cette méthode de SLAM utilise le maillage reconstruit préalablement par le scanner HandySCAN 3D. Pour la séquence de billard, les images ont une résolution de 640×480. Pour cette séquence, les poses de caméra sont restituées par le logiciel Agisoft PhotoScan [3]. La détection de spécularité est fournie par notre méthode de détection présentée dans l'annexe A en tant qu'image en niveau de gris. Nous n'utilisons que 6 images pour reconstruire la quadrique pour chaque séquence.

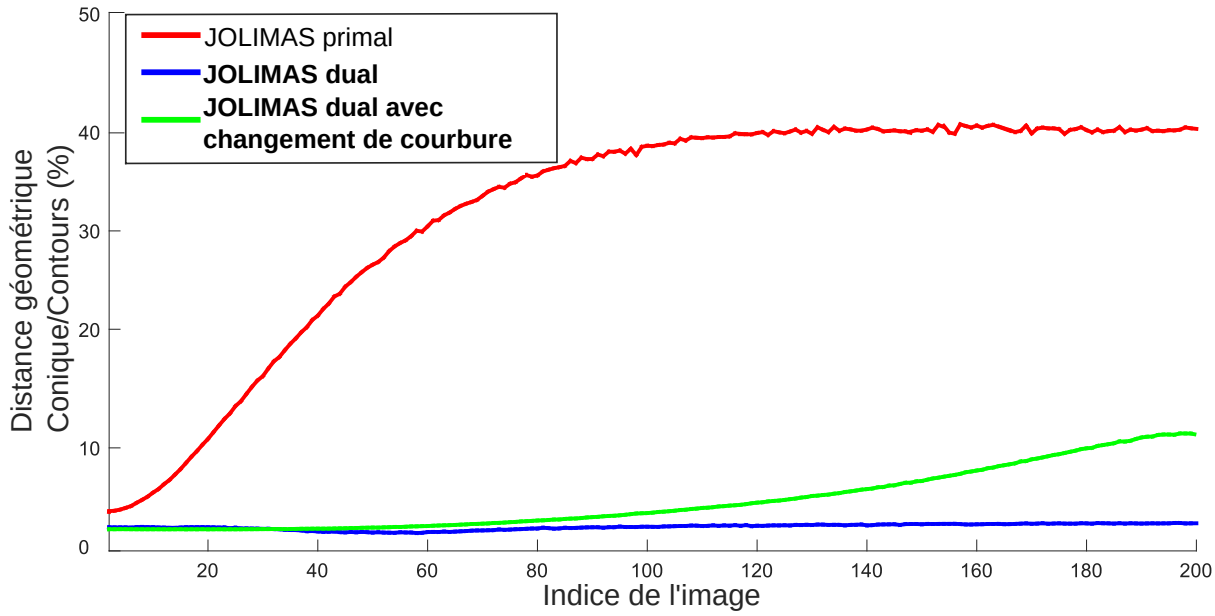
L'ensemble des résultats est illustré sur les figures 5.7 et 5.8. Un résumé des résultats de notre modèle dual comparé au modèle primal est donné par la table 5.1.

Réplique de fusée. Dans cette séquence comportant 1410 images, une réplique de fusée en métal est illuminée par un néon. Comme illustré à la figure 5.5(c), notre maillage de

3. www.agisoft.com



(a)

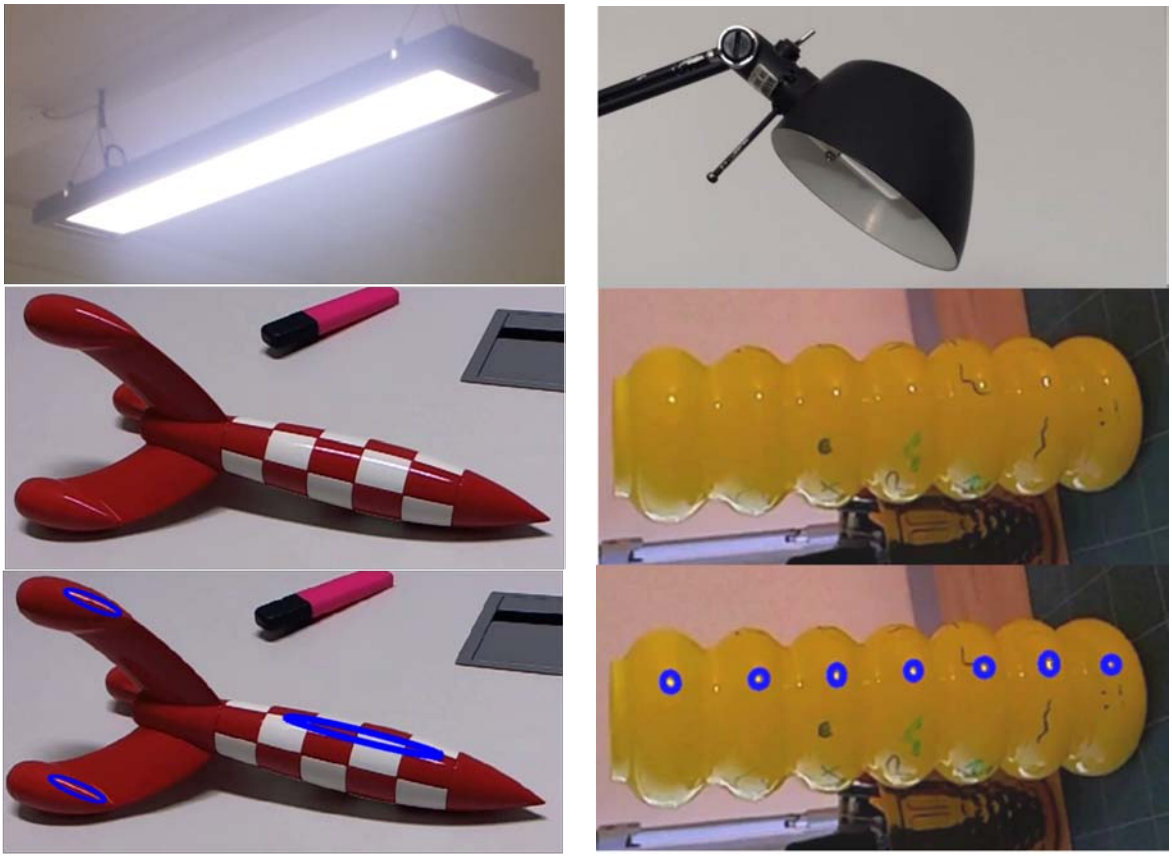


(b)

FIGURE 5.6 – Comparaison entre JOLIMAS primal et notre modèle JOLIMAS dual. Dans cette séquence de synthèse, la courbure augmente progressivement à chaque image comme illustré dans (a). En utilisant 6 poses de caméra, les deux modèles primal et dual sont reconstruits et comparés en utilisant la distance 2D entre la prédiction de spécularité et les contours de la spécularité détectée. Notre approche duale réalise de meilleurs résultats que JOLIMAS primal avec une erreur moyenne de 1% par ellipse comparée à une erreur de 33% par ellipse en moyenne pour JOLIMAS primal. Nous testons également la prédiction de spécularité de notre modèle dual en reconstruisant une quadrique sur une surface plane (première image) et en l'utilisant pour les changements de courbure restants. Même si l'erreur croît pour chaque image, la prédiction de spécularité reste fiable avec une erreur relativement faible (7% en moyenne).

la fusée est divisé en 7 morceaux convexes. Cependant, seulement 3 morceaux convexes sont utilisés, car les pièces restantes ne présentent pas de spécularité. Nous atteignons une précision de 2.1% au niveau de l'erreur de prédiction comparée à 32.8% pour le modèle primal.

Mug. Pour cette séquence comportant 300 images, un mug en porcelaine est illuminé par 2 sources de lumière : une lampe de bureau et un néon différent de celui utilisé pour la séquence de la fusée. Notre méthode est capable de prédire les deux spécularités précisément avec une erreur de prédiction de 1.3% puisque la courbure de l'objet d'inté-



(a) (b)

FIGURE 5.7 – Prédiction de spécularité sur deux surfaces non-planes. Sous l'hypothèse que les spécularités ont une forme elliptique sur les surfaces convexes, nous reconstruisons une quadrique dont la projection perspective correspond aux spécularités pour de nouveaux points de vue. Une quadrique est créée pour chaque surface convexe et chaque source de lumière. (a) présente une séquence réelle d'une réplique de fusée (au milieu) éclairée par un néon (en haut). Dans (b), nous utilisons une ampoule (en haut) pour illuminer un vase en verre (au milieu). Notre modèle a la capacité de prédire les spécularités (en bas) sur les différentes parties de la fusée et du vase pour de nouveaux points de vue (ellipses bleues).

rêt est constante. JOLIMAS primal est capable de prédire les spécularités avec la taille appropriée, mais avec une mauvaise position présentant une erreur moyenne de 18.3%.

Vase jaune. Dans cette séquence de 1900 images, un vase jaune en verre est illuminé par une lampe de bureau. Cette séquence est particulièrement difficile du fait de la symétrie de l'objet rendant l'estimation de la pose de caméra propice aux erreurs. De plus, les spécularités sur le vase ont une petite taille rendant le raffinement de la quadrique difficile. Cependant, nous arrivons à obtenir de bons résultats avec une erreur de prédiction de l'ordre de 1.9% comparés à 35.2% pour le modèle primal. Cet objet est divisé en 7 morceaux convexes comme illustrés sur la figure 5.5(d).

Billard. Dans cette séquence de 415 images, trois boules de billard sont éclairées par deux sources de lumière : une lampe de bureau et une lampe de type ampoule au plafond.

Cette séquence est également difficile à cause de la petite taille des spécularités. JOLIMAS dual obtient une précision de 1.1% comparée à 51.4% pour le modèle primal.

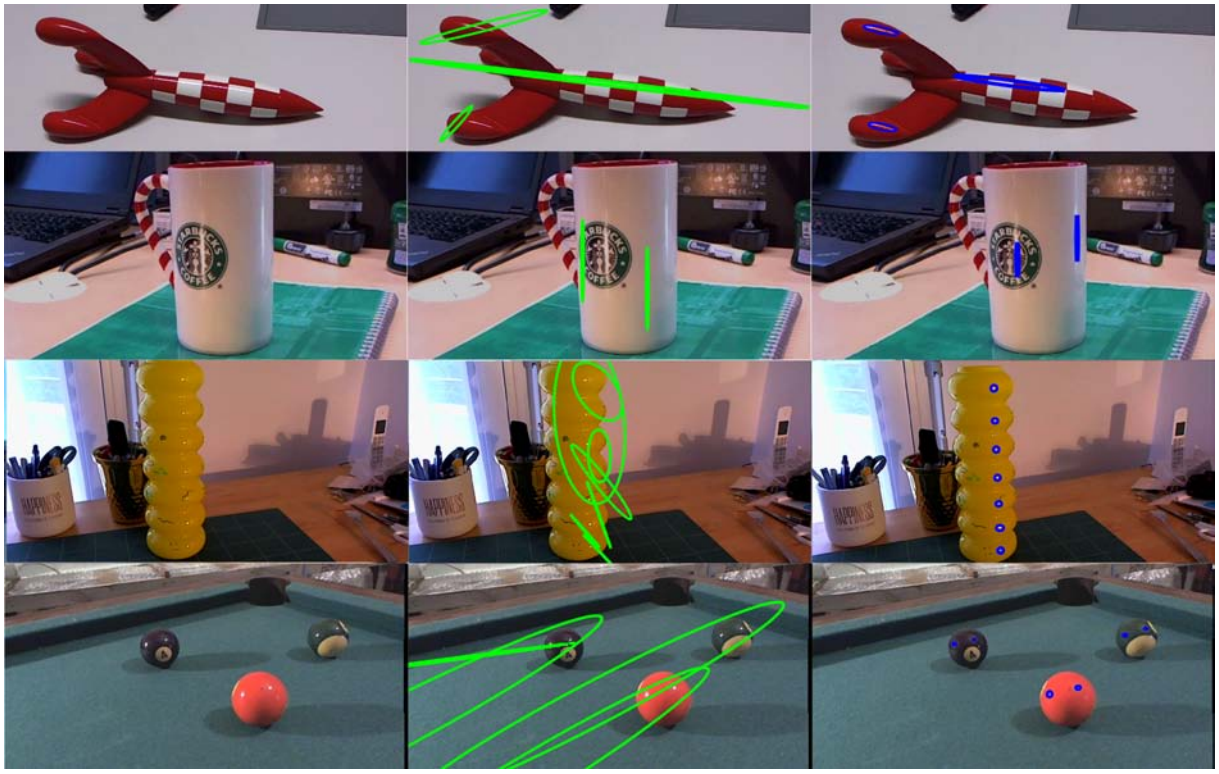


FIGURE 5.8 – Prédiction de spécularité sur des séquences réelles avec différentes conditions de lumière et pour différentes courbures. Dans les images de la première ligne, une réplique de fusée est illuminée par un néon. Nous prédisons 3 spécularités précisément en utilisant JOLIMAS dual à droite contrairement à JOLIMAS primal (au centre) qui échoue en matière de position et de forme de spécularité prédite. Dans la deuxième ligne, un mug, illuminé par une lampe de bureau et une lampe néon, est utilisé pour tester la capacité de notre modèle à prédire plusieurs sources de lumière de forme et d'intensité différentes. Pour cet objet, notre résultat est particulièrement précis grâce à la constance de la courbure de l'objet. JOLIMAS primal parvient à prédire correctement l'échelle des spécularités mais avec des mauvaises positions. Dans la troisième ligne, nous utilisons un vase illuminé par une lampe de bureau. L'objet comporte 7 parties convexes, ce qui donne 7 spécularités. Dans la dernière ligne, nous utilisons trois boules de billard illuminées par une lampe de bureau et une lampe de type ampoule au plafond. Ces deux objets sont particulièrement difficiles à traiter à cause des erreurs de pose de la caméra. En effet, notre méthode de localisation rencontre des difficultés à gérer des objets symétriques et notre détecteur de spécularité est limité par la petite taille des spécularités qui apparaissent les surfaces des objets.

5.8 Limites de l'approche

Notre modèle dual de JOLIMAS apporte une amélioration significative par rapport à JOLIMAS primal mais rencontre quelques limites et présente quelques cas difficiles. La

TABLE 5.1 – Validation empirique de notre modèle et sa capacité à prédire les spécularités dans les images pour 4 séquences réelles présentées à la figure 5.7 et 5.8. Pour des courbures et sources de lumière variées, notre méthode fournit des résultats précis en termes de prédiction de spécularité en comparaison de JOLIMAS primal.

	Erreur de prédiction (distance 2D en %)	
Séquence	Primal	Dual (notre méthode)
Fusée	32.8	2.1
Mug	18.3	1.3
Vase	35.2	1.9
Boule de billard	51.4	1.1

limitation principale de JOLIMAS dual est le fait qu'il faut reconstruire une quadrique par surface convexe et par source de lumière. Ceci pourrait être vu comme une incohérence dans le modèle et peut être inefficace en matière de temps d'exécution pour des objets plus complexes comportant de nombreuses parties convexes. Il est plus cohérent et efficace d'avoir une quadrique par source de lumière pour toutes les surfaces convexes à la fois. Dans l'exemple du vase illustré aux figures 5.7, nous utilisons les sept spécularités pour reconstruire la quadrique associée à la source de lumière et pour une seule image. À proprement parler, JOLIMAS dual ne permet pas cependant ce type de reconstruction, car notre modèle ne gère pas les changements de courbure. En termes d'images, il faut que les spécularités observées respectent la géométrie épipolaire. En pratique, nous avons vu que des variations légères de la courbure de la surface peuvent être gérées par JOLIMAS dual pour des surfaces lisses. Notons que pour des changements plus importants transformant une surface convexe en concave, nous ne pouvons pas gérer le changement topologique de la forme de la spécularité car il est difficile de modéliser clairement ce phénomène par une projection de quadrique. Un exemple de ce cas est illustré à la figure 5.9(b). Notre modèle est mis en échec lorsque la surface en question est constituée de différents matériaux. Ce cas va provoquer un changement d'échelle de la spécularité lorsqu'elle se déplacera d'un matériau à un autre et ainsi impactera négativement la reconstruction de quadrique de notre modèle. Un exemple de ce cas est illustré à la figure 5.9(a)

5.9 Conclusion

Nous avons traité le besoin de prédiction de spécularité pour les applications de réalité augmentée et de vision par ordinateur en présentant un nouveau modèle géométrique de forme de spécularité appelé JOLIMAS dual. Ce modèle est capable de prédire la forme d'une spécularité sur des surfaces planes, mais également sur des surfaces convexes. Il étend JOLIMAS primal sous l'hypothèse que les spécularités conservent une forme elliptique sur une surface convexe. L'idée générale du modèle est qu'il représente une spécularité en tant qu'image réfléchie de la source de lumière sur une surface spéculaire ayant un comportement proche de celui d'un miroir. Cette source de lumière virtuelle peut être modélisée par une quadrique dont la projection perspective correspond à la forme de la spécularité pour les points de vue existants et nouveaux en fonction de la pose de la caméra. Notre approche précédente, JOLIMAS primal, présentée dans le chapitre 4 était limitée aux surfaces planes. Pour des surfaces non-planes, les ellipses correspondant aux

(a) (b)

Figure 5.9 – Cas difficiles pour JOLIMAS dual. Dans (a), nous montrons que la spécularité est affectée par le matériau de la surface en termes d'échelle même si la courbure est similaire sur les deux objets (boules de Noël). Dans ce contexte, les ellipses associées aux spécularités ne respecteront pas la géométrie épipolaire. Dans (b), nous illustrons le cas d'un extincteur à incendie présentant des changements de courbure conséquents, ce qui va affecter la forme de la spécularité. Dans ce cas précis, nous ne sommes pas en mesure de prédire correctement la spécularité en matière d'échelle et de position.

spécularités ne sont pas épipolairement consistantes, ce qui met en échec JOLIMAS primal. En utilisant une représentation de caméra virtuelle à partir d'un point spécifique de la spécularité (le point d'intensité maximale), qui suit la loi de réflexion de Snell-Descartes en optique et qui n'est pas affecté par la distorsion induite par la courbure de la surface. Cependant, le calcul de ce point n'est pas trivial en raison de la présence de bruit dans les poses de caméra et dans les normales des maillages. Pour toutes ces raisons, nous avons implémenté une méthode d'estimation du point le plus intense respectant les règles de l'optique. Pour nous assurer de la généricité de notre modèle, nous utilisons une décomposition convexe approximative pour diviser le maillage en morceaux convexes. Afin de prédire chaque spécularité, nous reconstruisons une quadrique pour chaque morceau convexe et chaque source de lumière. La spécularité prédite (en tant qu'ellipse) est obtenue en projetant la quadrique reconstruite en utilisant la pose de caméra virtuelle du nouveau point de vue. Notre prédiction de spécularité a été évaluée en termes de robustesse et de précision sur des séquences synthétiques et réelles. Nous avons testé notre modèle dans des contextes variés avec plusieurs sources de lumière et plusieurs objets. Ce chapitre a fait l'objet de deux publications internationales ISMAR (2017d), d'une édition spéciale dans TVCG et d'une publication nationale ORASIS (2017). Cependant, dans le cas d'objets ayant une courbure variable (ellipsoïde par exemple), le modèle JOLIMAS dual présente des limites. En effet, dans l'état actuel, les informations de normales sont utilisées uniquement pour calculer les caméras virtuelles. Ceci ne pose pas de problèmes tant que la courbure de la surface reste uniforme sur les différents points de vue, car la projection de la quadrique reconstruite produira l'ellipse attendue correspondant à la spécularité. Dans le cas d'une courbure variable, la spécularité voit ses dimensions varier drastiquement. Nous avons constaté ce phénomène dans notre expérience synthétique illustrée à la figure 5.6 où nous transformons un plan en cylindre. Nous proposons une solution à cette problématique dans le chapitre suivant 6 sous forme d'un nouveau modèle : JOLIMAS canonique.

CHAPITRE 6

Un modèle général pour les surfaces à courbure variable

Dans les chapitres précédents 4 et 5, nous avons présenté le modèle JOLIMAS qui propose une solution au problème de prédiction de spécularité en utilisant un modèle géométrique sous l'hypothèse que les spécularités aient une forme elliptique sur des surfaces planes (version primale) et convexes (version duale). Nous partons, dans ce chapitre, de la version duale de JOLIMAS qui est limitée aux surfaces planes et convexes et courbure constante. L'intuition de cette approche vient du domaine de reconstruction multi-vue d'objets à partir d'images miroirs. En effet, une surface spéculaire a un comportement similaire à celui d'un miroir. Cette propriété nous permet de considérer la spécularité comme étant la réflexion d'un objet 3D qui est fixe dans l'espace (en position et en angle). Cette réflexion est naturellement affectée par la courbure de la surface locale. Afin de reconstruire notre modèle JOLIMAS par l'intermédiaire de notre quadrique, nous utilisons au moins trois ellipses calculées à partir des contours de spécularité détectés et nous appliquons une transformation spécifique aux contours afin de simuler une surface plane et ainsi simuler un miroir plan qui ne déforme pas l'image de l'objet réfléchi. Après reconstruction, nous projetons la quadrique en perspective et nous transformons l'ellipse pour correspondre à la courbure locale de la surface pour un point de vue nouveau. Nous évaluons notre nouvelle approche à la fois sur des séquences synthétiques et réelles et comparons ces résultats avec l'itération de notre modèle précédent : JOLIMAS dual.

6.1 Introduction

Dans le chapitre 5, nous présentions une méthode qui, à partir des ellipses calculées sur des contours de spécularités, reconstruit une quadrique dont la projection perspective correspond aux contours de la spécularité pour des nouveaux points de vue. Cette approche est limitée aux surfaces planes et convexes de courbure constante ce qui ne permet pas de gérer de changements de courbure locale sur la surface de la spécularité courante. Dans ce chapitre, nous détaillons deux contributions majeures à notre modèle de prédiction de spécularités. Premièrement, nous présentons une représentation canonique dans la section 6.3 de notre modèle qui reconstruit une quadrique à partir des ellipses correspondant à la

forme des spécularités dans l'image après transformation des contours de cette spécularité en fonction de la courbure locale de la surface. L'idée derrière cette transformation est de revenir à une surface de courbure nulle ce qui correspond à une surface plane. Si nous parvenons à transformer les contours des spécularités afin de simuler la forme de la spécularité sur une surface plane, la quadrique reconstruite à partir des ellipses calculées sur ces contours transformés sera notre modèle canonique pouvant être adapté à n'importe quel type de courbure. L'hypothèse utilisée est que l'image réfléchie (ellipse) de notre quadrique sur une surface spéculaire plane n'est pas distordue de la même façon que sur un miroir plan. Deuxièmement, pour un nouveau point de vue, par projection perspective de la quadrique, nous obtenons une ellipse qui correspond à la forme de la spécularité si la surface était plane. À partir de cette ellipse, nous réalisons le procédé inverse de la reconstruction et transformons les contours de l'ellipse en fonction de la courbure locale. Ces ellipses transformées correspondent à la forme des spécularités dans l'image pour n'importe quelle surface. Nous évaluons l'efficacité de notre méthode pour des données synthétiques et réelles comparées à notre précédente itération : JOLIMAS dual.

6.2 Lien entre déformation, spécularité, courbure de la surface et champ de vision

Lors d'un processus de reconstruction 3D, il est toujours difficile de reconstruire une surface miroir ou semi-réfléchissante comme du métal ou du plastique. En effet, lors de la reconstruction d'une surface mate, l'information de couleur d'un point 3D sur la surface reste fixe dans l'espace en fonction de la pose de la caméra. En observant la surface et l'évolution des informations de couleurs des différents points de la surface, la géométrie de la surface peut être reconstruite. Dans le cas d'un miroir, la couleur de l'objet observé à un point 3D de l'objet est reflétée par rapport à la normale au point 3D sur le miroir. De ce fait, afin de reconstruire un objet reflété dans un miroir, il est nécessaire de connaitre la géométrie du miroir, ce qui demeure un problème complexe.

Comme évoqué brièvement dans le chapitre 5, nous pouvons assimiler le comportement d'une surface réfléchissante à celui d'une surface miroir. En effet, dans le cas d'une surface spéculaire, la spécularité correspond au reflet de source de lumière observé à travers la surface. Toute la base théorique de notre modèle JOLIMAS repose sur l'existence d'un objet 3D fixe dans l'espace qui est réfléchi par une surface spéculaire. Si nous trouvons une relation entre la courbure de la surface (miroir ou spéculaire) et l'objet 3D à reconstruire (quadrique dans le cas de JOLIMAS), alors il nous sera possible de généraliser entièrement notre modèle de prédiction de spécularité à n'importe quelle courbure.

Nous constatons dans la figure 6.1 que les objets observés dans un miroir plan sont déformés dans le cas d'un miroir courbé. Il en est de même pour la spécularité observé sur le livre rouge plastifié. Un autre constat remarquable dans le cas du miroir est que le champ de vision semble beaucoup plus large alors que l'aire de la surface n'a pas augmenté. Une première hypothèse serait que le changement de courbure induit un changement dans le champ de vision et ainsi, vu que l'aire de la surface reste constante lors de la déformation, la surface représentant l'objet diminue en devant concave et augmente en devenant convexe. Ce lien est souligné davantage dans nos données synthétiques à la figure 6.2.

Le champ de vision d'une caméra est défini par l'ensemble de ce qu'elle peut observer à tout instant. Ce champ de vision est paramétré par deux composantes : un angle de

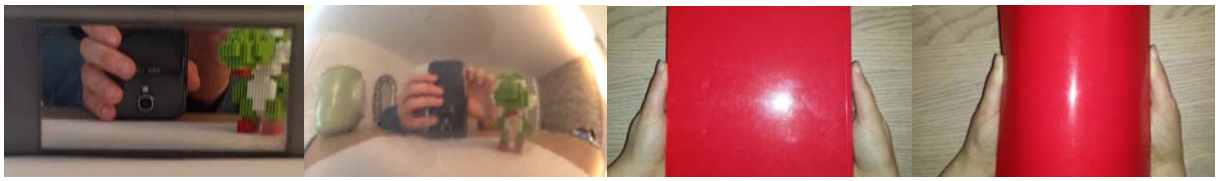


FIGURE 6.1 – Images miroirs d'une scène 3D dans deux cas : une surface miroir (paire de gauche) et une surface spéculaire (paire de droite). Les objets 3D (une statue en plastique et un *smartphone* pour la surface miroir et une source de lumière pour la surface spéculaire) sont déformés en fonction de la courbure locale de la surface. Dans un processus de prédiction de spécularité, ces déformations de surface doivent être prises en compte.

champ de vision horizontal et un angle pour le champ de vision vertical. En présence d'une surface miroir, ce champ de vision peut être augmenté comme on peut le voir avec le rétroviseur d'une voiture ou encore dans une caméra catadioptrique. En effet, les rayons de lumière sont réfléchis sur la surface du miroir ce qui permet d'obtenir un point de vue plus large. Il est possible de faire varier le champ de vision réfléchi par ce miroir en changeant sa courbure. Quand on passe d'un miroir plan à un miroir convexe, le champ de vision augmente et inversement pour un miroir concave. Nous illustrons cette propriété à la figure 6.2.

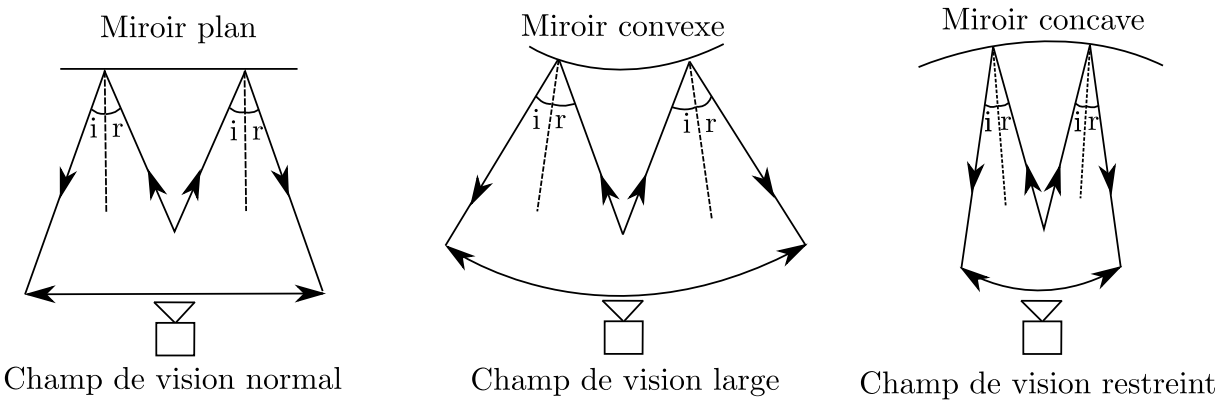


FIGURE 6.2 – Relation entre courbure et champ de vision dans le cas d'un miroir plan, convexe et concave (de gauche à droite). Nous observons que le champ de vision augmente lorsque le miroir devient convexe et diminue lorsque le miroir devient concave.

6.3 Approche proposée - Représentation canonique pour la reconstruction 3D de quadrique

6.3.1 Utilisation de la courbure gaussienne

À partir de la méthode de Healey et Binford (1988) qui est une application de *Shape from Specularity* qui consiste à récupérer des informations sur la géométrie d'un objet à partir du mouvement d'une spécularité sur une surface, plusieurs idées/hypothèses sont intéressantes à explorer afin de trouver un lien encore courbure locale et forme de spécularité. En effet, de façon semblable à notre approche du modèle JOLIMAS, Healey et

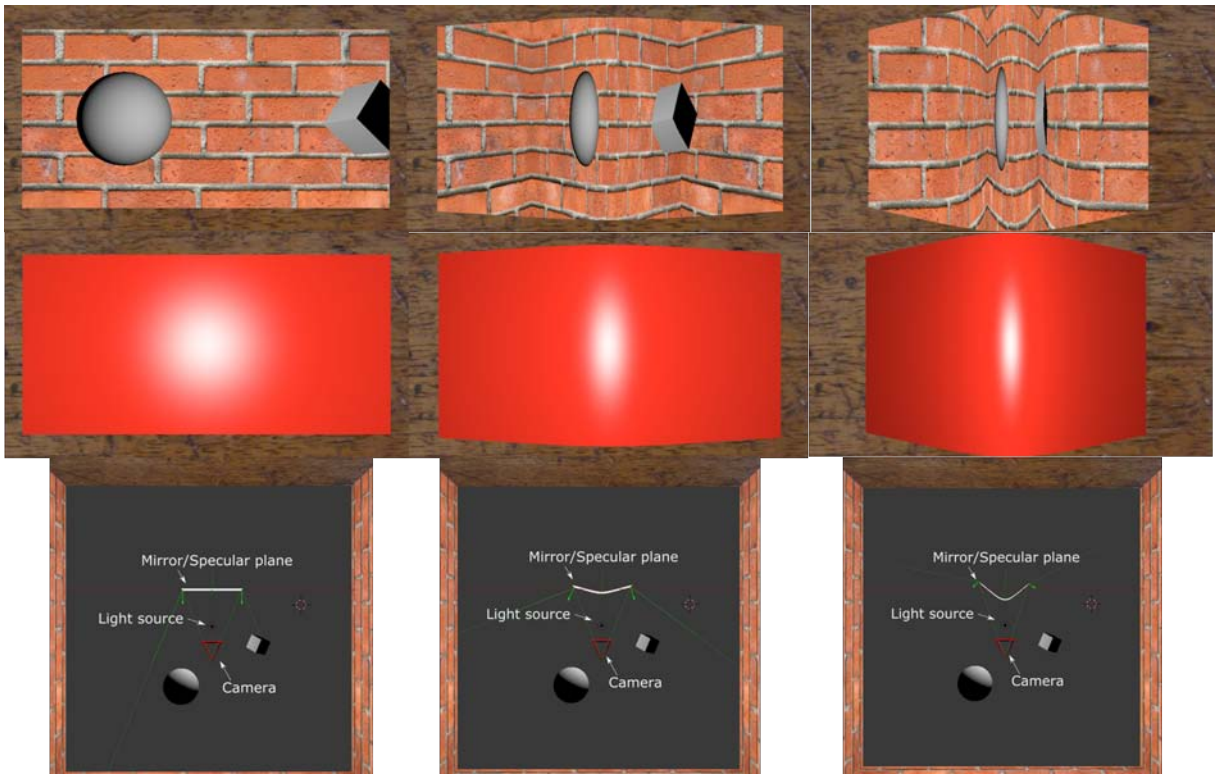
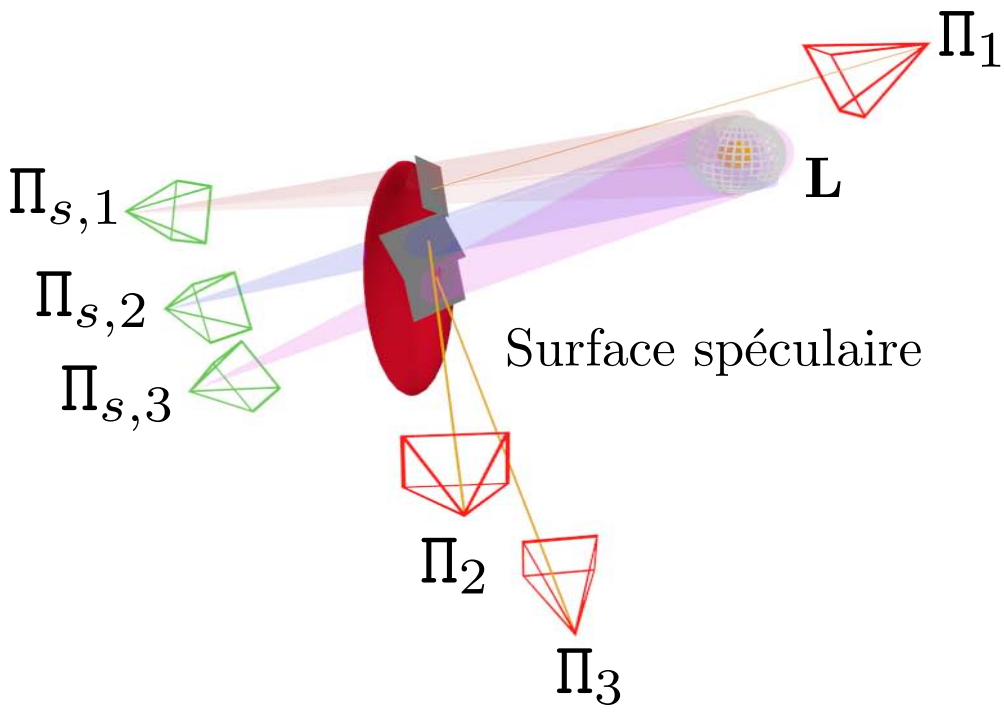


FIGURE 6.3 – Illustration de la relation entre courbure et champ de vision par l'utilisation d'une scène synthétique 3D contenant un cube, une sphère, une source de lumière et une surface réfléchissante (miroir et surface plastique). Afin d'illustrer le changement de courbure et le champ de vision, nous courbons la surface de la scène progressivement et nous observons les changements sur les reflets du miroir (première ligne) et de la spécularité (deuxième ligne). Dans les images de la troisième ligne, nous illustrons une vue aérienne de la scène 3D et nous avons tracé les limites du champ de vision en lignes pointillées vertes. Nous observons, de façon similaire dans la figure 6.1, que la scène 3D reflétée dans le miroir est déformée (première ligne) et que la spécularité est également déformée (deuxième ligne). Nous constatons que le champ de vision augmente lorsque la surface passe de plane à convexe. Contrairement à la déformation qui est appliquée lorsqu'on déforme une texture mate, une image réfléchie est affectée par les rayons réfléchis de la surface spéculaire.

Binford (1988) utilisent un modèle d'illumination locale afin de faire le parallèle entre la courbure de la surface et forme de la spécularité. Malgré la qualité des résultats que fournit le modèle de Phong, Healey et Binford (1988) précisent que les paramètres de ce modèle n'ont pas un sens physique. Afin d'être plus précis, Healey et Binford (1988) utilisent le modèle de Torrance et Sparrow (1967) qui est un modèle développé par des physiciens permettant de donner une formulation plus détaillée de la composante spéculaire. Ce modèle suppose qu'une surface est composée de petites facettes (micro-facettes) orientées de façon aléatoire et ayant un comportement proche de celui d'un miroir. Ce modèle quantifie également l'occultation des facettes par d'autres facettes adjacentes en utilisant un facteur d'atténuation géométrique. Le modèle spéculaire résultant est décrit par la formule :

$$I_S = FDA, \tag{6.1}$$



(a)

FIGURE 6.4 – Illustration de notre modèle canonique JOLIMAS. Pour chaque pose de caméra réelle (Π_1, Π_2 et Π_3), nous créons des poses de caméra virtuelles ($\Pi_{s,1}$, $\Pi_{s,2}$ et $\Pi_{s,3}$) calculées à partir des points d'intensité maximale de chaque spécularité. En synthétisant la forme des spécularités sur le plan tangent aux points d'intensité maximale, nous sommes capables de reconstruire une quadrique localisée près de la source de lumière $\mathbf{L}$ invariante à la courbure de la surface.

avec F le coefficient de Fresnel, D la fonction décrivant la distribution de l'orientation d'une facette et A le facteur géométrique d'atténuation. Le coefficient F modélise la quantité de lumière qui est réfléchie par chaque facette. En général, F dépend de l'angle du rayon incident de lumière et d'un indice de réfraction du matériau réfléchissant. Selon Cook et Torrance (1982), F doit caractériser la couleur de la spécularité afin de synthétiser des images réalistes. La fonction de distribution D décrit l'orientation des micro-facettes par rapport la normale moyenne de la surface $\mathbf{N}$. Dans le cas de Torrance-Sparrow, une distribution gaussienne est utilisée afin de décrire D tel que :

$$D = Ke^{-(\alpha/m)^2}, \tag{6.2}$$

où K est une constance de normalisation, m un indice de rugosité et α l'angle d'incidence. Ainsi pour un α donné, D est proportionnel au nombre de facettes orientées dans la direction du vecteur $\hat{\mathbf{H}}$ décrivant le vecteur de mi-chemin de Blinn (1977) que nous détaillerons par la suite.

Pour analyser la forme de la spécularité, il n'est pas nécessaire de disposer de tous les paramètres de l'équation (6.1). F peut être considéré comme constant par rapport au point de vue de l'observateur. De plus, Healey et Binford (1988) montrent que le facteur

exponentiel présenté à l'équation (6.2) varie de façon beaucoup plus rapide que les autres termes. ce qui permet de simplifier l'équation (6.1) en :

$$I_S = K' e^{-(\alpha/m)^2}, \tag{6.3}$$

où K' est une constante. Healey et Binford (1988) évoquent également l'existence d'un point d'intensité maximale tel que l'intensité varie à partir de ce point. En étudiant cette variation d'intensité, il est possible d'en déduire la courbure de la surface. Le lien entre intensité I_S à un point $\mathbf{P}$ et l'angle d'incidence est donné par la formule :

$$|\alpha| = m\sqrt{-ln\frac{I_S}{K'}}. \tag{6.4}$$

Le lien entre courbure de la surface et angle d'incidence est explicité par la formule :

$$\kappa_n = \frac{\mathrm{d}\alpha}{\mathrm{d}s}\bigg|_{\mathbf{P_B}}, \tag{6.5}$$

avec κ_n la courbure locale dans une direction donnée, dα un léger changement d'angle de la source de lumière réfléchie et ds la longueur d'arc sur la surface au point d'intensité maximale $\mathbf{P_B}$. L'angle α est défini tel que $\alpha = \cos^{-1}(\hat{\mathbf{N}} \cdot \hat{\mathbf{H}})$ avec $\hat{\mathbf{N}}$ la normale et $\hat{\mathbf{H}}$ le vecteur de mi-chemin $\hat{\mathbf{H}} = \frac{\hat{\mathbf{L}}+\hat{\mathbf{V}}}{\|\hat{\mathbf{L}}+\hat{\mathbf{V}}\|}$ entre le rayon de lumière $\hat{\mathbf{L}}$ et le rayon d'observation $\hat{\mathbf{V}}$. En pratique, l'équation (6.5) n'est pas triviale à calculer pour des surfaces non-paramétriques telles que des maillages. Calculer la longueur d'arc implique de calculer une distance géodésique, ce qui peut être coûteux en temps de calcul et imprécis pour des arêtes franches ou des surfaces rugueuses, ce qui arrive souvent dans le cas d'un maillage.

6.3.2 Mise en pratique sur un modèle CAO - Notion d'angles limites

Afin de réduire le temps de calcul de la courbure locale, nous définissons une notion d'angles limites. Celles-ci correspondent aux valeurs d'angles α^i_{max} telles que pour tout angle supérieur à cette limite, l'intensité de la spécularité est considérée comme nulle. Nous calculons ces angles limites α^i_{max} à partir des contours de la spécularité ainsi que des normales de la surface associées à ces points de contours $\hat{\mathbf{N}}$ et les vecteurs de mi-chemin $\hat{\mathbf{H}}$ avec $i \in [\![0, n]\!]$. D'après l'équation (6.3), l'intensité de la spécularité n'est influencée que par l'angle α. L'intensité associée aux points de contours de la spécularité reste constante par rapport à la courbure appliquée à la surface, ce qui implique que ces angles limites conservent des valeurs identiques pour tout changement de courbure locale aux points de contour de la spécularité.

Dans notre représentation canonique, afin de transformer l'ellipse courante associée à la spécularité, nous calculons le point d'intensité maximale $\mathbf{P_B}$ comme expliqué au chapitre précédent 5, et son plan tangent $T_{\mathbf{P_B}}(S)$ associé. Nous échantillonnons n vecteurs $v_i \in T_{\mathbf{P_B}}(S)$ en commençant au point $\mathbf{P_B}$ dans un intervalle de $[0, 2\pi[$. Le choix de n dépend de la taille de la spécularité et de sa forme. En pratique, nous avons fixé une valeur de 36 pour les séquences présentées à la figure 6.8. Cette valeur nous a donné empiriquement les meilleurs résultats. Lorsque nous atteignons les contours externes de la spécularité, nous calculons un angle limite α^i_{max} en utilisant $\hat{\mathbf{H}}$ et $\hat{\mathbf{N}}$ au point projeté

Algorithm 2 Pseudo-algorithme décrivant la transformation de forme de la spécularité sur la surface S en une représentation canonique sur le plan tangent $T_{\mathbf{P_B}}(S)$ au point d'intensité maximale $\mathbf{P_B}$.

procedure CURVEDTOPLANAR($\mathbf{P_B}$, S, *pose*, *Specu*, $\mathbf{L}$)
  $T_{\mathbf{P_B}}(S) \leftarrow ComputeTangentPlane(\mathbf{P_B}, S)$
  $v \leftarrow SamplingVectorInPlane(T_{\mathbf{P_B}}(S), \mathbf{P_B}, [0, 2\pi[)$
  for $i \in length(v)$ **do**
    $\mathbf{P} \leftarrow ReachOutline(Specu, S, v_i)$
    $\alpha_{max_i} \leftarrow AngleWithSurface(S, \mathbf{P}, pose, \mathbf{L})$
    $\mathbf{P_{res_i}} \leftarrow ReachOutlineAngle(v_i, T_{\mathbf{P_B}}(S), pose, \alpha_{max_i}, \mathbf{L})$
  return $CalculEllipse(\mathbf{P_{res}})$

orthogonalement sur la surface S. Par la suite, nous parcourons le même vecteur v_i sur $T_{\mathbf{P_B}}(S)$ jusqu'à ce que l'on trouve un point où la valeur de l'angle atteint α^i_{max} qui est la limite associée au vecteur. L'ellipse transformée est calculée à partir des nouveaux points de contours obtenus. Ce procédé est illustré à la figure 6.5 et la procédure est décrite dans l'algorithme 2.

À partir de ces ellipses transformées, nous sommes en mesure d'utiliser le même formalisme de reconstruction de quadrique utilisant des caméras virtuelles comme dans le modèle JOLIMAS dual au chapitre 5.

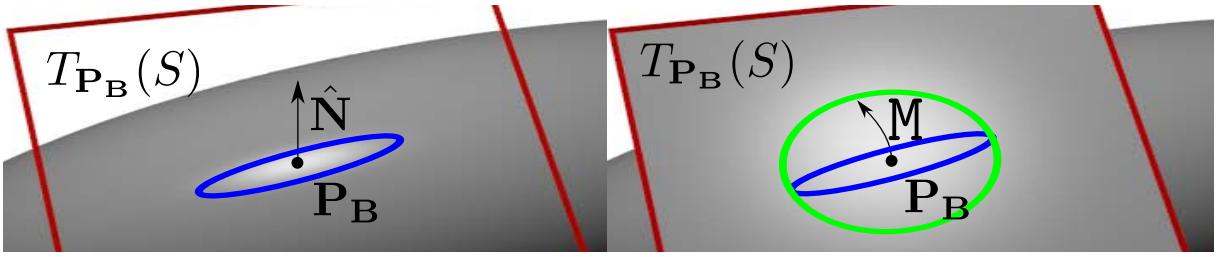
6.4 Modifications du processus de prédiction de spécularité

Dans l'étape de prédiction de spécularité, un processus semblable à la section 6.3 est utilisé afin de prédire la forme des spécularités pour un nouveau point de vue. Nous calculons le point d'intensité maximale pour le nouveau point de vue et nous projetons perspectivement la quadrique reconstruite, ce qui nous donne la forme de la spécularité sur le plan tangent $T_{\mathbf{P_B}}(S)$. Nous calculons les angles limites sur $T_{\mathbf{P_B}}(S)$ et nous déplaçons les contours de la spécularité sur S jusqu'à atteindre l'angle limite sur $T_{\mathbf{P_B}}(S)$.

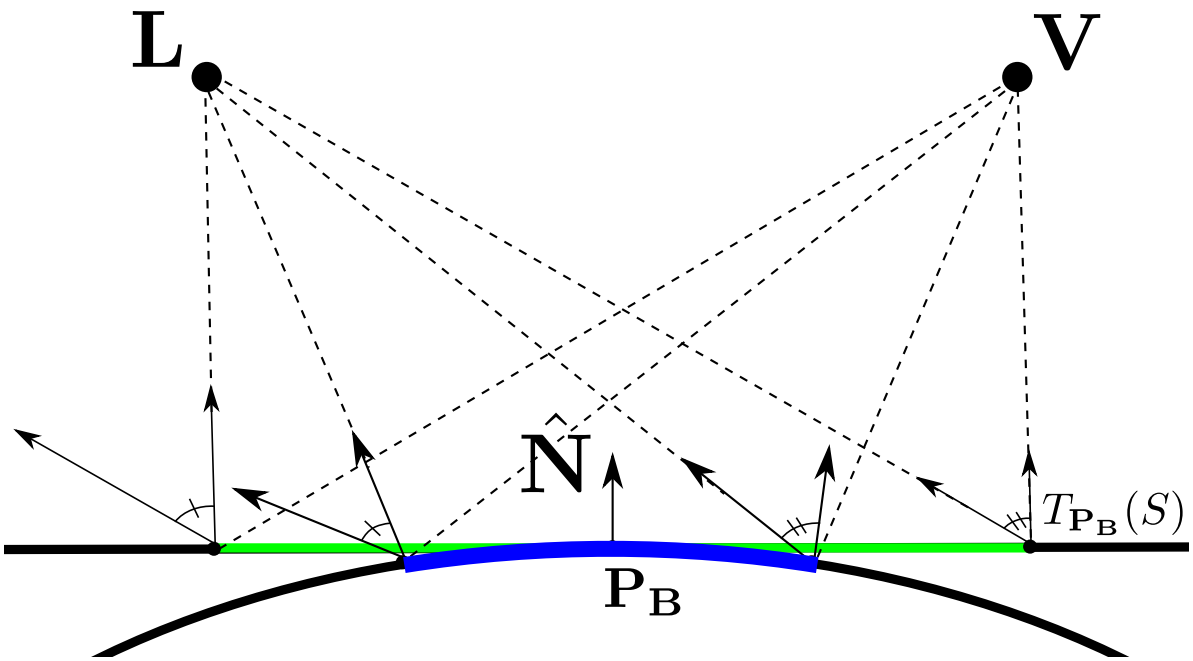
6.5 Résultats expérimentaux

6.5.1 Données de synthèse

Séquence plan/cylindre. Afin de tester l'amélioration de notre modèle canonique par rapport au modèle dual, nous avons repris l'expérience testée au chapitre précédent 5 qui consistait à faire varier le rayon d'un cylindre pour simuler la transformation d'un plan en un cylindre. Nous présentons les résultats à la figure 6.6. Nous observons que notre modèle canonique reconstruit seulement à partir des 6 premières images de la surface plane peut suivre les changements de courbure de la surface tout le long de la séquence en conservant une bonne précision avec une erreur avoisinant 3%. Notre approche précédente duale présente une erreur allant jusqu'à 12%.

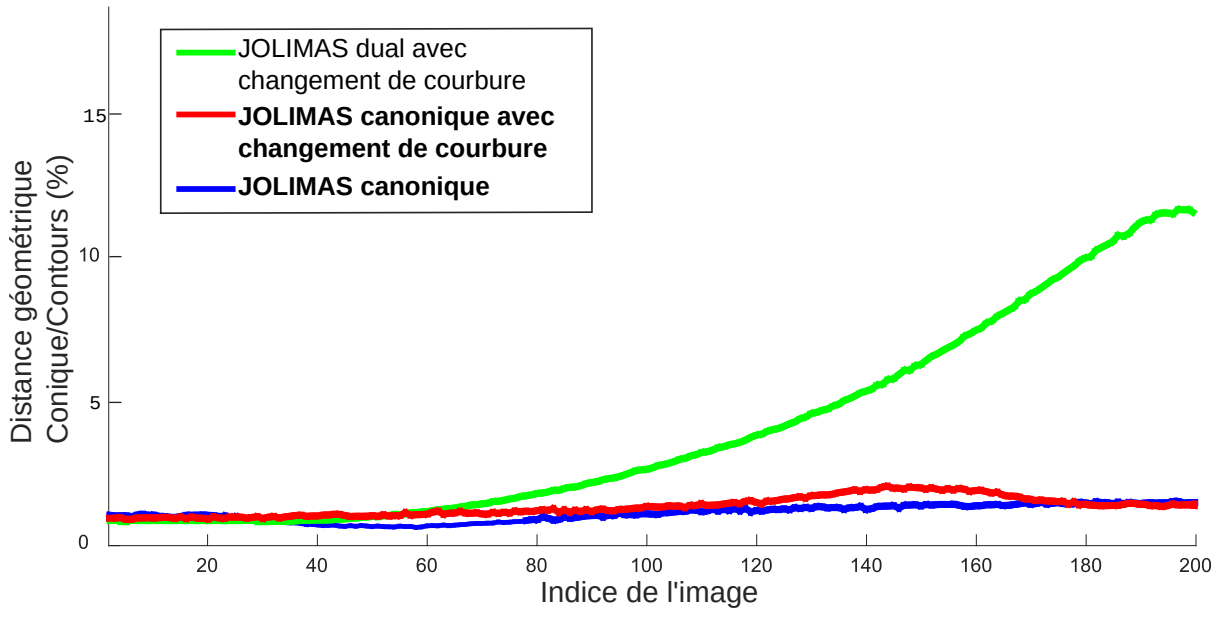


(a)

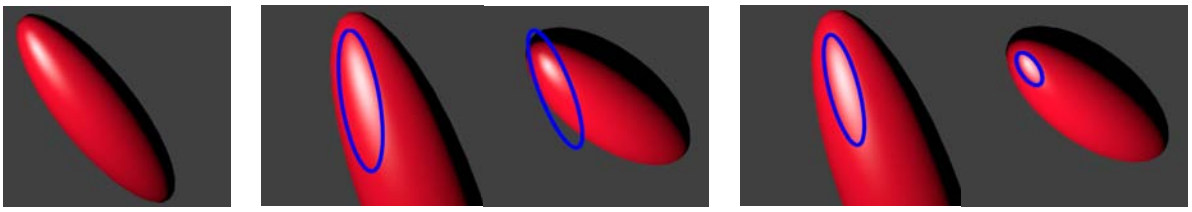


(b)

FIGURE 6.5 – Transformation d'ellipse de sa forme sur la surface courbe S en représentation canonique sur le plan tangent $T_{\mathbf{P_B}}(S)$ au point d'intensité maximale $\mathbf{P_B}$. Dans (a), nous montrons le processus qui, à partir de l'ellipse correspondant à la spécularité (en bleu), applique la transformation M afin d'obtenir l'ellipse corrigée (en vert) dans le plan tangent $T_{\mathbf{P_B}}(S)$. Dans (b), nous montrons que les angles limites conservent leurs valeurs aux contours de la spécularité sans être affectés par la courbure de la surface (bleu pour le courbe et vert pour le plan).



Séquence ellipsoïde. Afin de tester la validité de notre modèle JOLIMAS canonique sur un exemple simple, nous avons créé une séquence de synthèse comportant un ellipsoïde qui est une forme dont la courbure n'est pas uniforme sur toute la surface. Dans la figure 6.7, nous montrons que la prédiction de spécularité de JOLIMAS canonique est bien plus précise que notre approche précédente JOLIMAS dual que ce soit pour la position et la forme prédite de l'ellipse.



(a) (b) (c)

FIGURE 6.7 – Résultats de prédiction de spécularité pour notre séquence de synthèse d'ellipsoïde. Dans (a), nous montrons que l'ellipsoïde n'est pas uniforme en courbe. Dans (b) et (c), nous montrons une paire d'images de la séquence et le résultat de la prédiction de spécularité pour JOLIMAS dual (b) et JOLIMAS canonique (c). L'image de gauche correspond à la courbure utilisée pour reconstruire notre modèle et celle de droite correspond à une courbure plus importante. Nous constatons que notre modèle JOLIMAS canonique donne de meilleurs résultats pour les deux courbures contrairement à JOLIMAS dual qui reste correct, mais améliorable pour la première image, mais échoue complètement lorsque la courbure varie.

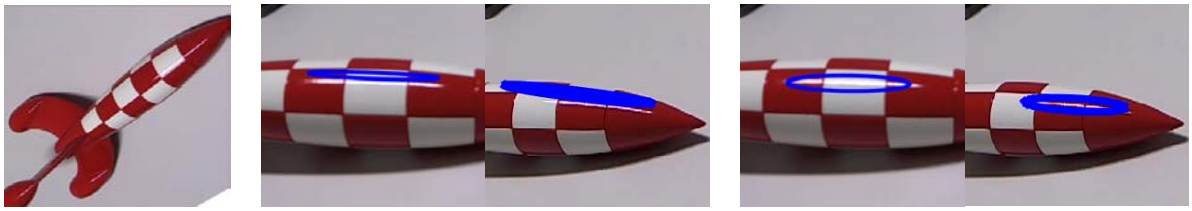
6.5.2 Données réelles

Dans cette section, nous présentons deux résultats de prédiction de spécularité sur des séquences réelles contenant une réplique de fusée en métal et un livre à couverture brillante. La reconstruction du JOLIMAS en forme canonique est réalisée à partir de six images. La prédiction de spécularité est réalisée pour des images ultérieures à celles utilisées pour la reconstruction.

Séquence Fusée. Cette séquence a été présentée précédemment dans le chapitre 5. Même si les résultats étaient déjà satisfaisants, la partie centrale de la réplique de fusée n'a pas une courbure constante, ce qui a provoqué des erreurs dans la prédiction de spécularité du modèle JOLIMAS dual. Nous pouvons observer dans la figure 6.8 que notre modèle JOLIMAS canonique présente de bien meilleurs résultats en termes de prédiction des spécularités.

Séquence Kinect. Afin de montrer le potentiel de notre méthode dans un contexte plus complexe, nous avons réalisé une séquence réelle de 146 images à partir d'une caméra RGB-D (Kinect V2 [1]) donnant la profondeur associée à chaque pixel en plus de la couleur. Le support de la Kinect V2 étant essentiellement assuré sur Windows 8, nous avons dû utiliser la bibliothèque Libfreenect2 du projet OpenKinect [2]. Cette bibliothèque permet

1. https ://developer.microsoft.com/en-us/windows/kinect
2. https ://github.com/OpenKinect/libfreenect2



(a) (b) (c)

FIGURE 6.8 – Prédictions de spécularité sur la séquence réelle présentée dans le chapitre précédent 5. (a) présente l'objet d'intérêt : une réplique de fusée réelle. Nous présentons une paire d'images pour chaque séquence montrant les résultats de prédiction (ellipses bleues) de JOLIMAS dual, dans (b) et JOLIMAS canonique, dans (c). La nouvelle approche estime la forme et la position des spécularités de façon beaucoup plus précise comparée à JOLIMAS dual.

d'accéder aux informations d'infrarouges sous forme d'une image 512×424 et de couleurs sous forme d'une image 1920×1080 du capteur. Cette bibliothèque propose également des outils pour synchroniser ces données et fournir une image de profondeur de dimension 512×424.

Afin de reconstruire notre modèle JOLIMAS canonique, il est nécessaire de récupérer les informations de normales des surfaces sur lesquelles les spécularités se présentent. Dans le cas de la Kinect V2, une carte de normale doit être déduite du nuage de points fourni par la carte de profondeur ; cependant la Kinect V2 ne fournit pas de profondeur pour les cas suivants :

- Objets réfléchissant les rayons infrarouge comme un miroir ou certaines zones d'une surface spéculaire
- Zones occultées dans l'image infrarouge par rapport à l'image couleur en raison de la *baseline* entre les deux caméras. Ces zones correspondent en général aux contours occultant des objets

Afin de pallier ces problèmes, nous avons dans un premier temps appliqué un algorithme de restauration d'image (*inpainting*) de Telea (2004) sur l'image de profondeur afin de combler les trous potentiels. Pour des zones plus larges, nous pouvons également utiliser les méthodes de Chican et Tamaazousti (2014); Bousefsaf et al. (2018). Dans notre séquence, ces trous sont en général nombreux, mais petits ce qui permet de les interpoler avec les informations avoisinantes dans l'image. Nous illustrons ce principe dans la figure 6.9.

Une fois la carte de profondeur complète, nous utilisons la bibliothèque Point Cloud Library (PCL) [3] afin de calculer la carte de normale. Celle-ci est obtenue en utilisant les points 3D présents dans une zone de recherche (définie en fonction du nuage de points). Notre rayon de recherche pour la séquence est de 3 cm. Nous illustrons cette étape dans la figure 6.10.

Nous présentons dans la figure 6.11 nos résultats de prédiction de spécularité sur la séquence Kinect de JOLIMAS canonique comparé à JOLIMAS dual. Comme attendu, JOLIMAS dual n'est pas capable de gérer le changement de courbure du livre contrairement à JOLIMAS canonique.

3. http ://pointclouds.org/

(a) (b) (c)

FIGURE 6.9 – Image couleur de notre séquence Kinect (a), sa carte de profondeur brute (b) et corrigée par *inpainting* (c).

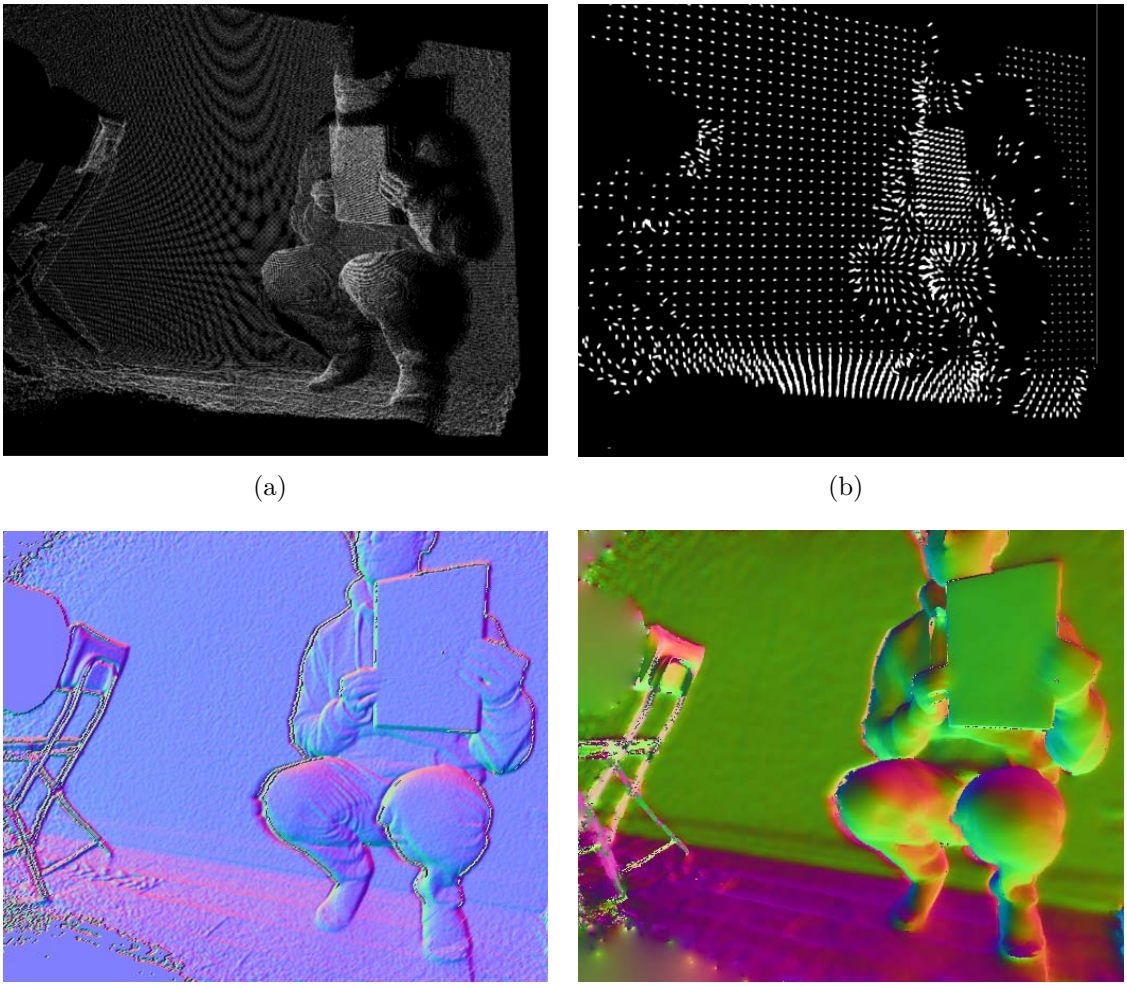
6.6 Limites de l'approche

Notre nouveau modèle JOLIMAS canonique apporte une solution au problème de prédiction de spécularité pour des objets à courbure variable (ellipsoïde, réplique de fusée) ainsi que des objets dont la courbure peut être modifiée en direct (livre que l'on plie au fur et à mesure de la séquence). Les expériences ont pu mettre en avant la relation critique entre qualité de la géométrie (à partir d'un modèle CAO et d'une Kinect V2) et stabilité/précision de la prédiction. Dans le cas de la Kinect, il est très difficile d'obtenir une prédiction sans sautillement provoqué par les approximations de la carte de normale dues au bruit des données brutes du capteur profondeur. Il serait intéressant de proposer un algorithme de cohérence temporelle et spatiale permettant de fournir une stabilité optimale même dans les cas difficiles. Une interpolation de la courbure de la surface et du changement de forme de la spécularité entre deux images pourrait être une approche intéressante.

6.7 Conclusion

Nous avons exposé une représentation canonique de l'approche précédente JOLIMAS dual afin de prédire les spécularités pour n'importe quelle courbure d'une surface. Nous pourrions utiliser cette version améliorée du modèle JOLIMAS afin d'avoir une initialisation précise pour calculer des modèles d'illumination locale comme Blinn-Phong et Cook-Torrance afin de permettre un rendu de meilleure qualité des spécularités et d'autres éléments de lumière comme les ombres. Il serait également intéressant d'étudier l'impact du type de matériau (rugosité et réflectance) sur la forme de la spécularité, ce qui pourrait grandement améliorer la généricité du modèle.

Ce chapitre a fait l'objet d'une publication internationale VRST (2017e) et une soumission dans un journal international TVCG 2018.

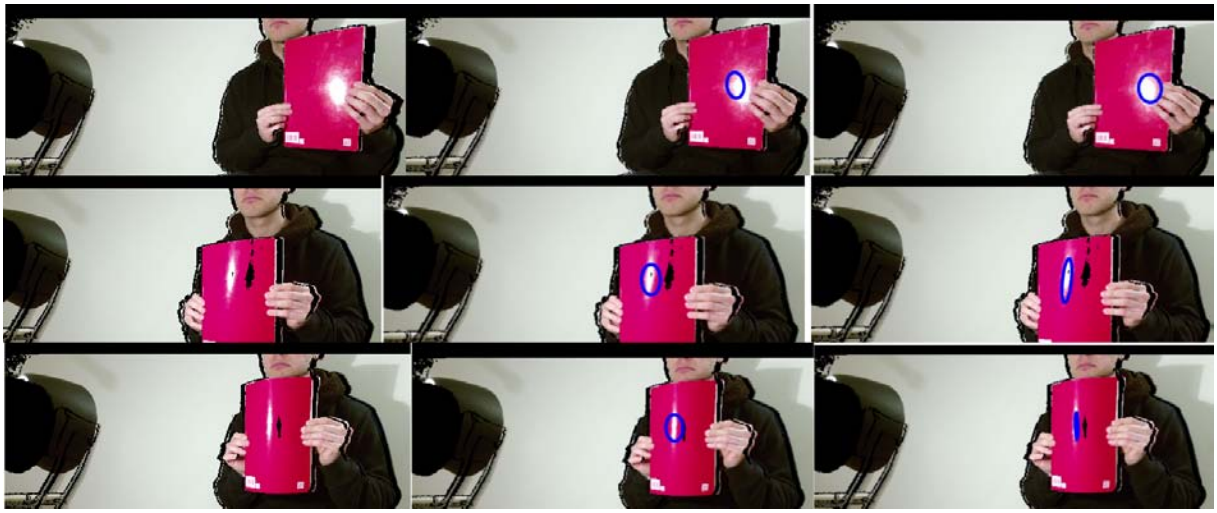


(a) (b)

(c) (d)

FIGURE 6.10 – Calcul de la carte de normale. À partir du nuage de points de la carte de profondeur (a), nous utilisons la bibliothèque PCL afin de calculer les normales à partir des points 3D voisins (b). Dans (c), nous illustrons une approche naïve du calcul de la carte de normale à partir de l'image de profondeur (2D) directement au lieu d'utiliser le nuage de points (3D). Nous constatons un bruit assez important sur des surfaces qui sont censées être planes (mur et livre). La carte de normale (d) obtenue à partir de la PCL est suffisamment précise pour notre application.

(a)



(b)

FIGURE 6.11 – Images utilisées pour la reconstruction de notre modèle (a) et résultats de prédiction de spécularités sur notre séquence Kinect pour trois images présentant des courbures différentes (pour chaque ligne) (b). Les images de gauche correspondent aux images d'origine. Celles du milieu correspondent aux résultats de prédiction de spécularité de JOLIMAS dual et celles de droites de JOLIMAS canonique.

CHAPITRE 7

Application à la réalité augmentée

Après avoir détaillé dans les chapitres précédents les différentes itérations de notre modèle JOLIMAS primal 4, dual 5 et canonique 6, nous détaillons ici comment nous avons utilisé les résultats de prédiction de spécularité afin d'obtenir un rendu en réalité augmentée. Nous présentons plusieurs itérations d'une application de retexturing qui consiste à substituer la texture d'une surface à une autre tout en conservant son contexte lumineux comme les variations d'intensité, les spécularités et les ombres. De plus, nous détaillons la gestion de contextes complexes comme le multi-lumières, les lumières à état changeant pouvant s'éteindre et s'allumer, le choix des images à utiliser pour la reconstruction et une analyse du coût du pipeline de retexturing.

7.1 Introduction

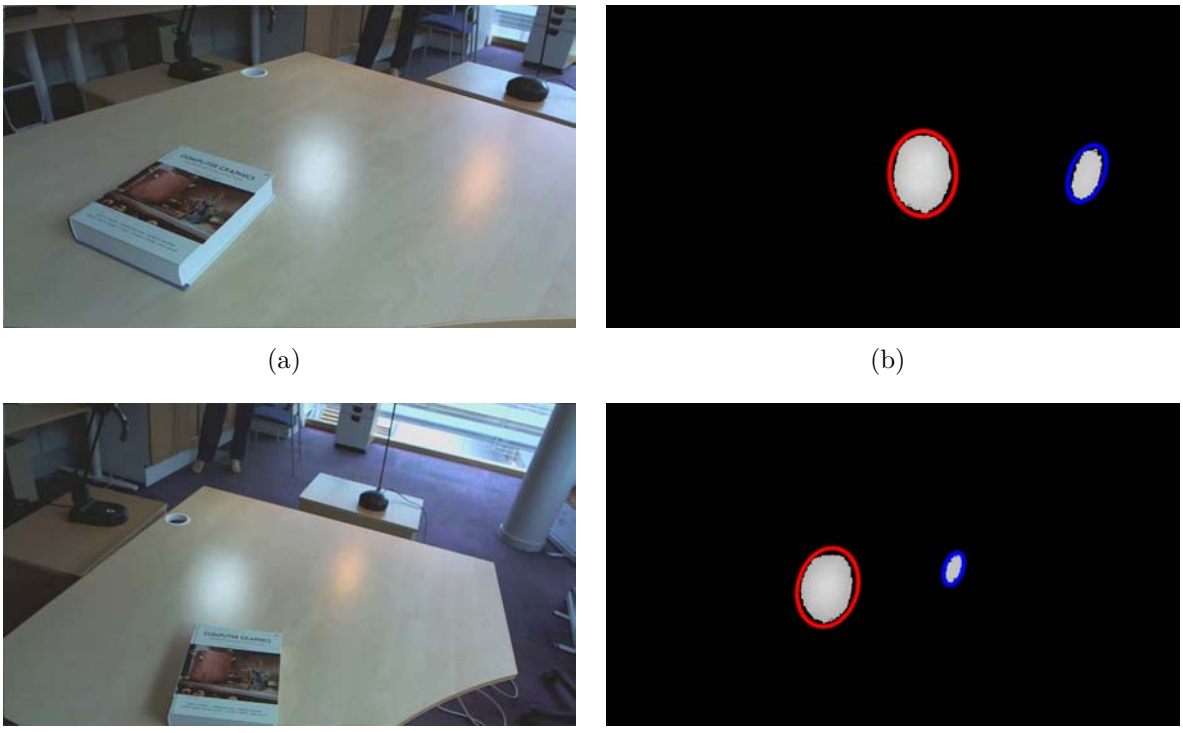
Dans Blake (1990), il a été démontré que la perception humaine prend en compte les spécularités pour mieux comprendre une scène (géométrie, illumination). Pour ces raisons, la modélisation des sources de lumière et plus particulièrement la prédiction des spécularités peut potentiellement améliorer les applications suivantes :

- Le retexturing en réalité augmentée
- La réalité diminuée et plus particulièrement lorsqu'une spécularité traverse une zone de l'image à restaurer Kawai et al. (2015); Herling et Broll (2014); Said et al. (2017)
- Synthétiser des points de vue inconnus à partir d'un flux vidéo (où la caméra n'est jamais allée physiquement)

Dans ce chapitre, nous détaillons l'ensemble des composants implémentés pour assurer le bon fonctionnement de notre modèle dans un contexte de réalité augmentée large, ce qui inclut la gestion de plusieurs sources de lumière simultanément pouvant s'allumer et s'éteindre, la capacité à synthétiser des spécularités ainsi qu'une nouvelle texture pour une application de retexturing et une analyse du temps d'exécution de notre méthode.

7.2 Suivi de spécularité en contexte multi-lumières

Une composante additionnelle de notre modèle est sa capacité à prédire plusieurs spécularités provenant de différentes sources de lumière. En effet, sur des surfaces planes, une spécularité est associée à une seule source de lumière. En annotant et en suivant le déplacement de ces éléments, plusieurs quadriques peuvent être reconstruites à la fois. Après le processus de détection de spécularité, nous suivons le mouvement de chaque spécularité pour chaque image. Le suivi de spécularité est essentiel pour compter et distinguer les sources de lumière. Notre suivi est réalisé en calculant l'évolution des contours de spécularité d'une image à l'autre en détectant les intersections entre les contours de l'image courante et la précédente comme illustré dans la figure 7.1. Même si le mouvement de caméra est important, les spécularités suivent naturellement le mouvement de la caméra (plus particulièrement pour les surfaces planes), permettant à notre algorithme d'être robuste à des cas complexes. Ce suivi est réalisé pour chaque image en utilisant l'algorithme de détection de contour de Suzuki et Satoshi (1985) afin d'avoir des contours précis extraits de l'image binaire produite par notre méthode de détection de spécularité en temps réel. Cependant, suivre des spécularités se chevauchant (lorsque deux sources de lumière ont la même direction d'observation par rapport au point de vue) n'est pas encore géré par notre méthode, bien que ce cas reste relativement rare.



(a) (b)

(c) (d)

FIGURE 7.1 – Annotation multi-lumière et suivi de spécularités (en bleu et rouge). (a) et (c) représentent deux points de vue différents d'une table en bois présentant deux spécularités. (b) et (d) représentent les ellipses associées à chaque source de lumière (en bleu et rouge).

7.3 Filtrage multi-vues

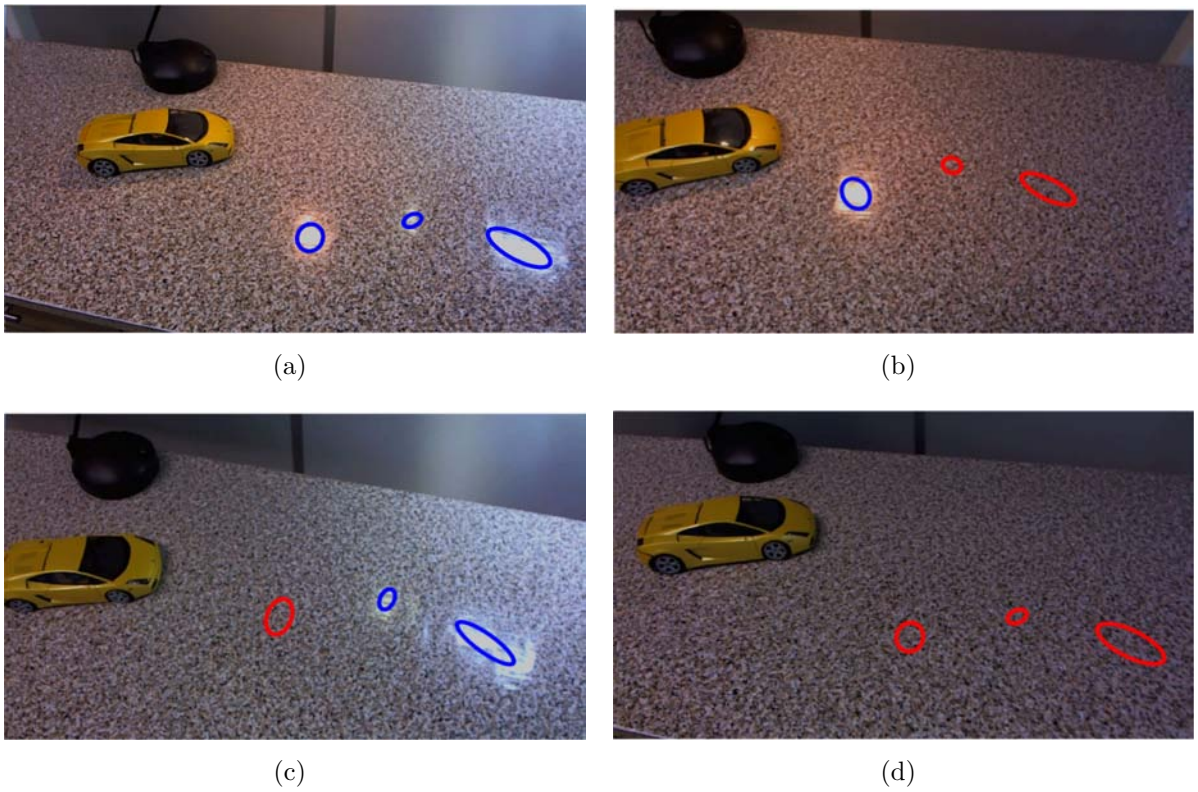
Pour des images surexposées, notre méthode de détection de spécularité peut confondre les textures blanches avec des spécularités. Afin d'assurer la robustesse de cette détection, un processus multi-vues sur les spécularités candidates s'impose. En détail, à l'opposé des contours 3D d'une texture, les contours 3D d'une spécularité se déplacent sur la surface lorsque la caméra se déplace. Ce phénomène est justifié du fait qu'une spécularité correspond à la réflexion d'une source de lumière qui, elle, est fixe. La spécularité que nous observons correspond à un objet situé en dessous de la surface et non sur la surface, d'où le déplacement. En calculant l'homographie associée au plan où se présente la spécularité entre deux poses relativement éloignées, nous transformons (warping) les contours des spécularités candidates . Si le contour est suffisamment proche, le candidat est une texture. Dans le cas contraire, c'est une spécularité.

7.4 Gestion de lumières à état changeant

Nous abordons ici l'analyse de l'état des spécularités en comparant la détection de spécularité dans une image et la prédiction produite par notre quadrique reconstruite. Cette analyse présente de nombreux intérêts :

- Si les ellipses associées respectivement avec la prédiction et la détection ont la même forme, mais une position différente, cela signifie probablement que la pose de caméra n'est pas assez précise.
- Si les ellipses associées respectivement avec la prédiction et la détection ont la même position, mais une forme différente, cela signifie probablement que l'intensité de la lumière a changé ou que le matériau est différent.
- Si l'ellipse associée à la détection n'est plus présente, cela signifie que la source de lumière a été éteinte ou occultée.

Ce dernier point sera abordé dans cette section. Il semble intuitif qu'une source de lumière à état changeant (allumée ou éteinte) provoque des variations d'intensité importantes dans l'image observée, ce qui impacte fortement les algorithmes de vision par ordinateur. De plus, une caméra doit également s'adapter au nouveau contexte lumineux, ce qui affectera le résultat des différentes applications de vision par ordinateur comme la segmentation, la reconstruction et la localisation. Actuellement, aucune solution n'explicite ce problème. En utilisant la prédiction de spécularité de notre modèle JOLIMAS, nous pouvons suivre la spécularité créée par une source de lumière qu'elle soit éteinte ou allumée en comparant la prédiction et la détection de spécularité. Nous illustrons la capacité de notre modèle à prédire les spécularités des sources de lumière à état changeant (allumées et éteintes) dans la figure 7.2. Dans cette séquence, trois sources de lumière sont utilisées (deux lampes de bureau et une lampe néon) sur un comptoir de cuisine. Durant cette séquence, nous allumons et éteignons chaque source de lumière à tour de rôle avant d'éteindre toutes les sources de lumière en une seule fois à la fin de la séquence. Les résultats de cette séquence montrent la capacité de notre modèle à gérer les changements rapides d'état de lumière tels que les clignotements de la lampe néon lors de l'allumage dans un contexte multi-lumières.



(a) (b)

(c) (d)

FIGURE 7.2 – Gestion des changements d'état de plusieurs sources de lumière. Les spécularités prédites sont représentées par des ellipses en bleu quand la lumière est allumée et en ellipses rouges dans la lumière est éteinte. La première lampe de bureau est éteinte en (a) et le néon en (b). Dans (c) la première lampe de bureau et le néon sont allumés alors que la deuxième lampe de bureau est éteinte avant d'être rallumée. Dans (d), toutes les sources de lumière sont éteintes jusqu'à la fin de la séquence.

7.5 Retexturing sur des objets réels

Pour toutes ces applications, afin de synthétiser une spécularité dans une image, on peut envisager, dans un premier temps, de détecter/segmenter les spécularités dans une image et de les appliquer directement sur une nouvelle texture. Cependant, détecter les spécularités de manière uniforme n'est pas une tâche triviale, ce qui rend cette solution non optimale. En effet, les variations au niveau de la détection peuvent provoquer des incohérences temporelles et des vibrations non voulues comme illustrées à la figure 7.3. De plus, la qualité de la détection de spécularité ne peut pas être garantie pour chaque point de vue à cause des changements de conditions lumineuses, des limitations propres au détecteur (voir annexe A), des imperfections de la surface en matière de rugosité et d'occultations potentielles d'objets de la scène. Notons que la bonne gestion de la cohérence temporelle du rendu est particulièrement importante pour les applications comme la réalité diminuée Kawai et al. (2015); Herling et Broll (2014); Said et al. (2017).

7.5.1 Première approche : modélisation par fonction Gaussienne pour surface planes

Afin de répondre au problème de cohérence temporelle pour la synthèse de spécularité lors du *retexturing*, nous utilisons notre modèle JOLIMAS. En effet, celui-ci est une solution naturelle, car notre prédiction de spécularité est majoritairement dépendante de la stabilité de la pose de la caméra vu que cette prédiction est obtenue par projection de la quadrique à partir de la pose de caméra. Afin de synthétiser une spécularité à partir de JOLIMAS, il est nécessaire d'avoir une fonction d'intensité décrivant la variation d'intensité de la spécularité. Nous proposons une approximation basée sur une fonction gaussienne 2D. En effet, une spécularité peut être décrite comme une zone de forte intensité par rapport au reste de l'image. La fonction d'intensité gaussienne suit deux propriétés essentielles de la spécularité : une variation progressive de l'intensité et des isocontours de forme elliptique comme détaillé dans le chapitre 4. La propriété de variation progressive a déjà été exploitée dans la méthode de Hyeongwoo Kim (2013) afin de séparer les composantes spéculaires et diffuses dans une image. Notre méthode de retexturing est divisée en trois étapes.

Premièrement, à partir des spécularités détectées dans la séquence, une couleur moyenne de la spécularité est calculée afin de correspondre aux conditions d'illumination. La fonction gaussienne appropriée est calculée sur les canaux rouges, verts et bleus afin de correspondre aux spécularités détectées dans les séquences.

Deuxièmement, afin de synthétiser les spécularités sur la texture associée à la surface, deux homographies sont calculées : $\mathtt{H}_2$ qui correspond à la transformation des points de la texture jusqu'aux points de la surface et $\mathtt{H}_1$ qui transforme le cercle unitaire en notre conique prédite (la conique est d'abord transformée par $\mathtt{H}_2^{-1}$).

Troisièmement, notre texture synthétique remplace la surface en fusionnant la texture avec notre texture gaussienne en utilisant $\mathtt{H}_1$ et nous transformons cette fusion sur la surface en utilisant $\mathtt{H}_2$. Quatre résultats de *retexturing* sont illustrés dans la figure 7.5. À cause du manque de méthode pour la modélisation géométrique de spécularité comme JOLIMAS, nous avons implémenté l'approche de Buteau et Saito (2015) comme méthode de comparaison pour notre *retexturing*. Cette approche relativement récente reconstruit des sources de lumière ponctuelle à partir des spécularités pour des scènes contenant des surfaces planes, ce qui la rend assez proche de notre première itération de JOLIMAS (primal). Afin de réaliser l'application de *retexturing* pour des sources de lumière ponctuelles, nous calculons manuellement chaque paramètre du modèle d'illumination de Phong (rugosité, intensité et couleurs des sources de lumière) et appliquons ces paramètres seulement pour le terme spéculaire pour générer la nouvelle texture. Comme illustré à la figure 7.5, notre approche synthétise de façon précise l'intensité, la forme et la position des spécularités contrairement à Buteau et Saito (2015) qui créent des spécularités de meilleure qualité, mais avec une mauvaise position et une mauvaise taille. De plus, pour les exemples présentés dans les figures 7.4(a) et 7.4(b), les modèles d'illumination locale comme celui de Phong ne sont pas capables de prédire correctement la forme et la variation d'intensité de la spécularité. En utilisant uniquement des sources de lumière ponctuelles, il est difficile de modéliser les sources de lumière étendue. La simplicité de notre méthode nous permet de fournir de bons résultats pour les surfaces planes. Dans la section suivante, nous montrons comment nous avons généralisé ce processus pour tous types de surfaces et amélioré le rendu en calculant le terme diffus afin de rajouter des effets d'ombre.

7.5.2 Deuxième approche : retexturing pour tous types de surfaces

Nous améliorons notre première approche de synthèse de spécularité en ajoutant un terme diffus qui est essentiel pour représenter les ombres et les variations d'intensité qui apparaissent sur des surfaces non-planes. Afin d'améliorer davantage le réalisme, nous calculons aussi l'intensité et la couleur de la spécularité empiriquement à partir de chaque image utilisée lors du processus de reconstruction de quadrique.

Calcul du terme diffus Nous considérons la texture d'entrée comme la composante ambiante utilisée dans le modèle de Phong. Afin d'ajouter le terme diffus, nous utilisons les centres des quadriques reconstruites comme étant la position des sources de lumière $\mathbf{L}$ et nous calculons la composante diffuse pour chaque point de la surface $\mathbf{P}$ avec :

$$I_d = \sum_{i=1}^{k} \left(\hat{\mathbf{L}}_i(\mathbf{P}) \cdot \hat{\mathbf{N}}(\mathbf{P}) \right), \tag{7.1}$$

avec I_d l'image diffuse, k le nombre de sources de lumière et i l'index de la source de lumière utilisée.

Reconstruction de la spécularité et fusion. Pour chaque image utilisée pour la reconstruction de la quadrique, nous utilisons les spécularités prédites et les spécularités détectées dans les images afin de reconstruire un motif de spécularité. En mettant en correspondance les formes des différentes spécularités, nous calculons l'évolution de l'intensité et de la couleur de la spécularité. Afin d'assurer une fusion entre la spécularité et la nouvelle texture, nous interpolons la couleur aux bordures de la spécularité avec le terme diffus autour de celle-ci. Nous détaillons le pipeline complet à la figure 7.5.

Comme illustré dans la figure 7.6, notre application de *retexturing* permet de synthétiser l'intensité et la couleur de la spécularité et de créer un rendu cohérent en temps réel.

7.6 Limitations

Croisements de spécularités Si la caméra observe une scène dans une direction telle que deux sources de lumière ont un alignement similaire, les deux spécularités résultantes peuvent se croiser dans l'image. Ce phénomène met en échec le suivi et donc la labellisation de spécularités. Actuellement, ce problème n'est pas géré dans notre application, mais pourrait être traité si les sources de lumière sont de couleurs différentes. Cependant, pour des sources de lumière similaires (forme et couleur), la seule façon de distinguer clairement les spécularités est d'utiliser une information 3D (reconstruire une source de lumière ponctuelle à partir de deux points de vue par triangulation par exemple) afin de s'assurer de la cohérence du mouvement de la spécularité. Dans tous les cas, les images où les spécularités se croisent doivent être exclues de notre reconstruction.

7.7 Temps d'exécution

Dans la table 7.1, nous présentons le temps d'exécution de chaque étape de notre méthode incluant la détection de spécularité, le calcul du point d'intensité maximale, la reconstruction de quadrique et le *retexturing*. Nous calculons les résultats sur un processeur Intel i7 de fréquence 2.70 GHz sur des séquences réelles présentées dans la figure 5.8. La reconstruction de quadrique prend 76.2ms en moyenne et est réalisée quand un minimum de 6 images clés est atteint. La quadrique est raffinée pour chaque nouvelle image clé mais prend cette fois-ci 15ms en moyenne, car seules quelques itérations sont nécessaires contrairement au raffinement initial. Notre analyse a été réalisée sans utilisation de GPU ou de parallélisation. Notre méthode tourne à approximativement 25 images par secondes ce qui la rend adaptée à un contexte temps-réel.

TABLE 7.1 – Calcul du temps d'exécution de la prédiction de spécularité de JOLIMAS dual. Notre méthode tourne en temps réel.

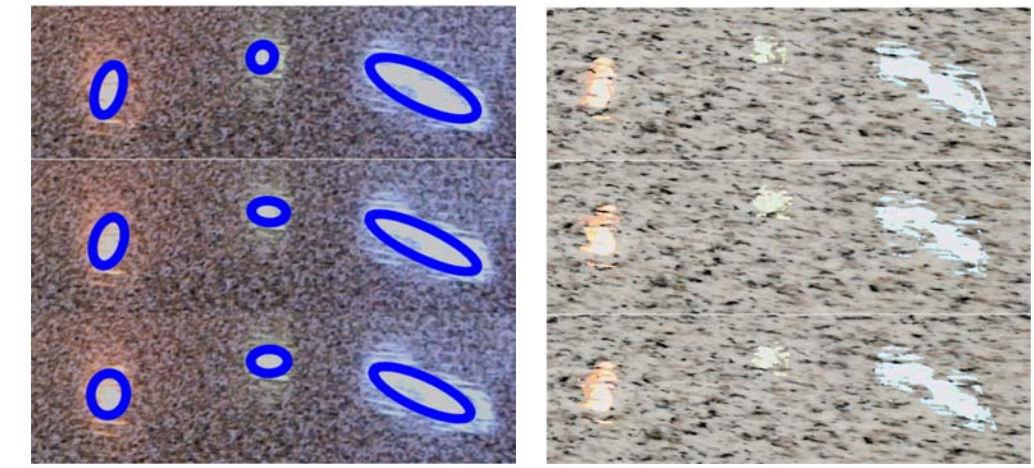
Étape de reconstruction de JOLIMAS dual	Temps d'exécution (ms)
Détection de spécularité	12.1
Correction du point d'intensité maximale	8.3
Reconstruction de quadrique	76.2
Prédiction de spécularité	2.4

7.8 Perspectives

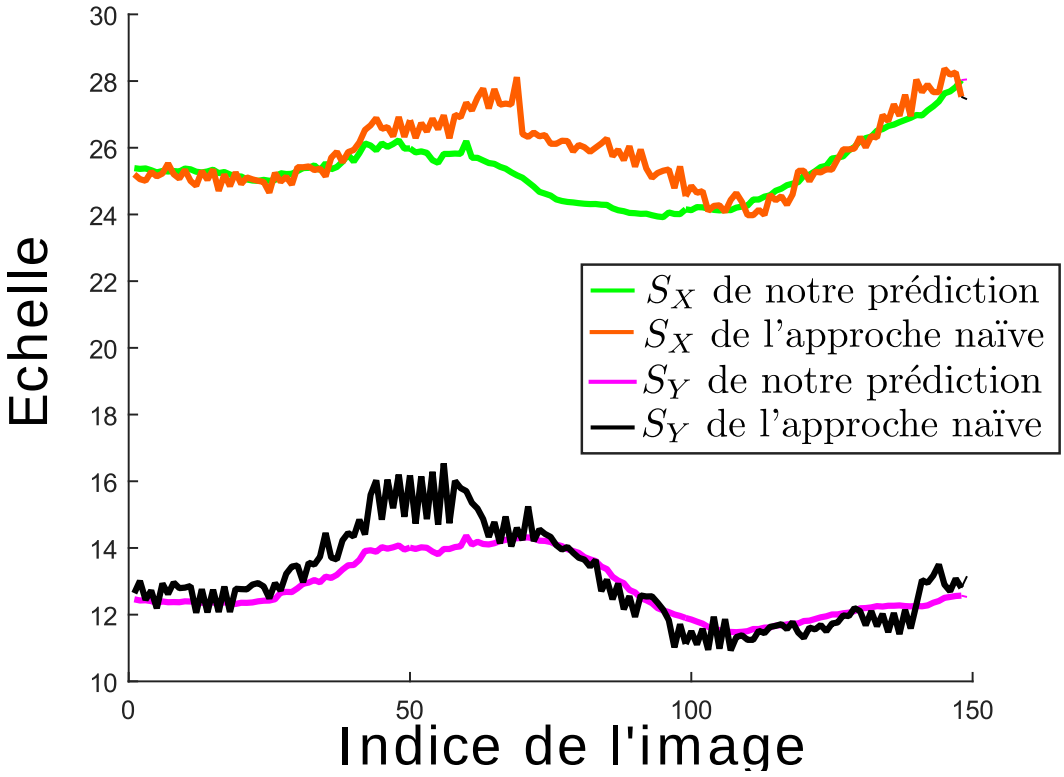
À partir de notre modèle JOLIMAS, il est possible d'utiliser la quadrique comme bonne initialisation pour n'importe quel modèle d'illumination locale (comme Torrance et Sparrow (1967); Cook et Torrance (1982) de façon similaire à ce qui a été fait dans la section 7.5) et d'appliquer une étape de raffinement non-linéaire pour calculer/raffiner les autres paramètres comme la couleur, l'intensité voire la position de la source de lumière. Une autre extension serait d'utiliser les spécularités pour calculer une carte de normale afin de mesurer la rugosité. En effet, comme observé à la figure 7.7, les spécularités mettent en valeur la rugosité sur des surfaces qui semblent majoritairement planes.

7.9 Conclusion

Afin d'assurer le bon fonctionnement de notre modèle dans un contexte de réalité augmentée, plusieurs choses doivent être prises en compte. D'abord, la possibilité d'avoir plusieurs sources de lumière et le fait que ces lumières peuvent potentiellement s'allumer ou s'éteindre. Nous avons également proposé une application de *retexturing* qui consiste à transformer la texture initiale de la surface en une autre. La difficulté de cette application est de synthétiser la nouvelle texture ainsi que les spécularités cohérentes qui la parcourent. Nous avons illustré nos résultats sur de nombreuses séquences réelles et montré que le modèle JOLIMAS est efficace et adapté à un contexte de réalité augmentée.

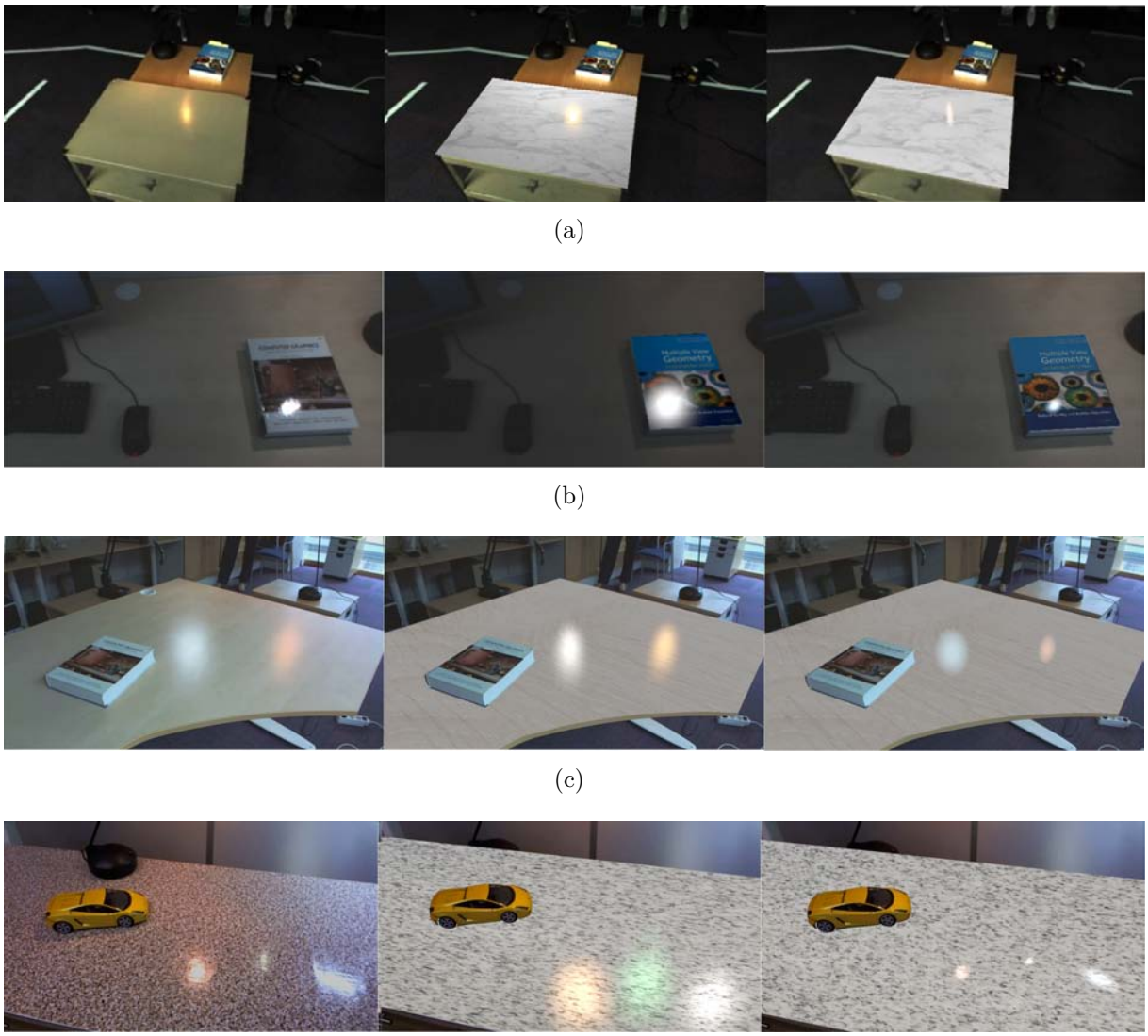


(a) (b)



(c)

FIGURE 7.3 – Illustration d'incohérences temporelles en utilisant un calcul d'ellipse pour une application de *retexturing* dans une séquence multi-lumière présentée dans la figure 4.6(d) pour trois images consécutives (une image par ligne de (a) et (b)). Dans (a), un tremblement sur le processus de calcul d'ellipse est visible pour chaque spécularité où nous calculons l'ellipse directement à partir de la détection de spécularité. Ce tremblement est illustré dans (c) en mesurant l'évolution des demi-axes (S_x et S_Y) de l'ellipse calculée (approche basique) comparée à l'ellipse prédite par notre modèle (JOLIMAS) pour chaque image de la séquence du tableau présenté dans la figure 4.6(a). Dans (b), nous copions directement le résultat de la détection de spécularité sur la nouvelle texture, ce qui provoque un décalage en termes d'intensité qui est difficile à estomper avec la texture. De plus, cette méthode est sensible à la qualité de la détection de spécularité et ne peut pas prédire de nouveaux points de vue.



(a)

(b)

(c)

(d)

FIGURE 7.4 – Comparaison des méthodes de *retexturing* de notre implémentation de Buteau et Saito (2015) (milieu) et de notre approche (droite). La méthode de Buteau et Saito (2015) étant une méthode pour reconstruction des sources de lumière ponctuelles, nous avons manuellement ajouté la rugosité, la couleur et l'intensité de la source de lumière en utilisant le modèle d'illumination locale de Phong. Nous utilisons uniquement le terme spéculaire. Le *retexturing* est illustré sur la séquence ampoule/table de fer (a) en utilisant une texture de marbre, (b) la séquence ampoule/libre en changeant la couverture du livre par une autre, (c) la séquence néon/ampoule/table en bois en changeant la texture de bois d'une table en une autre et (d) la séquence multi-lumière/comptoir de cuisine en utilisant une texture de roche. Sans prendre en compte le terme diffus sur la texture, la texture de la surface peut être changée de façon réaliste en utilisant uniquement le terme spéculaire prédit par JOLIMAS. Contrairement à l'approche de Buteau et Saito (2015), nous calculons de façon précise la forme, la position et l'intensité de la spécularité avec un temps d'exécution faible pour un résultat réaliste. De plus, calculer l'illumination à partir de modèles comme celui de Phong ne peut pas être fait sans données sur les matériaux et le nombre de sources lumineuses.

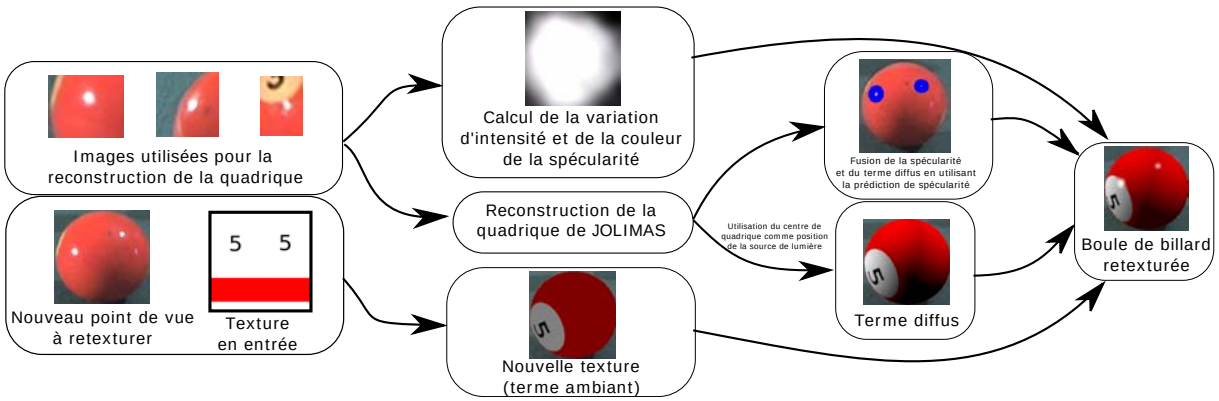


FIGURE 7.5 – Pipeline du processus de *retexturing* généralisé. La méthode est divisée en trois phases : la reconstruction de spécularité à partir des images utilisées pour la reconstruction de la quadrique 3D (partie du haut), calcul du terme diffus (partie du bas) et fusion de la spécularité synthétisée, la spécularité prédite et le terme diffus. La spécularité est reconstruite en termes de variation d'intensité et de couleur de façon empirique à partir des coniques prédites et des spécularités détectées dans les images. Afin de calculer le terme diffus, nous considérons le centre des quadriques reconstruites comme étant des sources ponctuelles. Nous calculons ensuite le terme diffus de Phong (1975). La spécularité et le terme diffus sont ensuite fusionnés afin de créer la nouvelle texture associée à la surface. L'ajout du terme diffus combiné aux spécularités améliore drastiquement le rendu.

FIGURE 7.6 – Exemples de *retexturing* sur trois séquences : la table de billard, le mug et le vase jaune. Nous pouvons voir que les spécularités synthétisées ainsi que les ombres sont cohérentes avec le contexte lumineux des images d'entrée.

(a) (b)

FIGURE 7.7 – Illustration de la rugosité mise en évidence par la spécularité. Une surface paraissant plane et lisse se révèle rugueuse.

Conclusion et perspectives

Travaux réalisés

L'objectif principal de cette thèse était de proposer une solution de modélisation des sources de lumière d'une scène à l'aide d'une méthode de localisation de type SLAM contraint. Cette modélisation vise à améliorer le rendu des applications de réalité augmentée en ajoutant le contexte lumineux lorsqu'un objet virtuel est inséré dans un flux vidéo. Cependant, le problème de modélisation de l'illumination est en l'état mal posé et extrêmement complexe. Afin de proposer une solution à ce problème, nous avons, dans un premier temps, analysé les propriétés des différentes traces de lumière comme les ombres et les spécularités. Une ombre est produite quand un autre objet obstrue la lumière émise par une source lumineuse, créant une diminution de l'intensité reçue par cette surface. Cette variation d'intensité est généralement associée à la notion de *Shading* en anglais. Une spécularité correspond à la réflexion d'une source de lumière sur une surface réfléchissante donnant une information cruciale sur la position de la source de lumière. De plus, nous avons constaté que pour des surfaces réfléchissantes planes et convexes, ces spécularités prennent une forme elliptique. De ce constat, nous avons émis l'hypothèse que ces spécularités peuvent être décrites par la projection perspective d'une quadrique. Cette quadrique permet de fournir la forme des spécularités pour n'importe quel point de vue, mais également de donner de nombreuses informations sur les sources de lumière (forme, position). Nous proposons un modèle géométrique novateur par rapport aux approches photométriques. Notre modèle reconstruit une quadrique à partir des ellipses calculées sur les contours des spécularités détectées dans l'image.

Dans un premier temps, dans une version appelée JOLIMAS primal, nous avons posé les bases de notre modèle dans le contexte plus restreint des surfaces planes. Dans cette version, nous avons implémenté la reconstruction de quadriques à partir d'ellipses calculées à partir des spécularités. Après cette première reconstruction, nous utilisons un processus de raffinement non-linéaire de la quadrique qui minimise la distance entre les ellipses issues de la projection de la quadrique et des contours de spécularités observés dans les images. Ce modèle a été utilisé pour une application de *retexturing* en utilisant une fonction d'intensité gaussienne qui a donné des résultats prometteurs en matière de synthèse de spécularités.

En second lieu, nous avons mis en œuvre une nouvelle itération de JOLIMAS sous

une forme duale pour lever la contrainte de géométrie plane et l'étendre aux surfaces courbes. En effet, lorsqu'une surface est courbée, la reconstruction de quadrique échoue, car les contours des spécularités ne suivent pas la géométrie épipolaire. Afin de résoudre ce problème, nous utilisons un système de caméras virtuelles qui sont calculées à partir des normales aux points d'intensité maximale des spécularités. Dans ce formalisme de caméras virtuelles, les contours de spécularités suivent la géométrie épipolaire. Ces caméras virtuelles permettent de réutiliser le pipeline de reconstruction de quadrique de JOLIMAS primal et d'augmenter la généricité du modèle. Nous avons également amélioré l'application de *retexturing* en reconstruisant un terme diffus qui nous permet de synthétiser des variations d'intensité et des ombres, ce qui augmente le réalisme de notre rendu.

Dans un troisième temps, nous avons proposé un modèle général appelé JOLIMAS canonique permettant de prédire les spécularités sur une surface pour n'importe quelle courbure. Dans cette itération de JOLIMAS, nous avons fait le lien entre courbure et spécularité en utilisant une formule décrivant le terme spéculaire du modèle d'illumination de Cook et Torrance (1982). En ajoutant un prétraitement des contours de la spécularité, nous sommes capables de transformer les contours de la spécularité sur la surface courbe en spécularité sur une surface plane (dans ce cas, le plan étant tangent au point d'intensité maximale). En utilisant le formalisme de caméra virtuelle de JOLIMAS dual et le pipeline de reconstruction de quadrique de JOLIMAS primal, nous fournissons un modèle de prédiction de spécularité générique, en temps réel et qui requiert un nombre d'images assez faible (au moins trois images et en moyenne six images).

Nous avons testé ces différentes itérations sur de nombreuses séquences synthétiques et réelles afin d'évaluer la qualité de la prédiction de spécularité ainsi que la qualité du rendu pour le *retexturing* en réalité augmentée.

Perspectives

Dans cette thèse, nous avons mis l'accent sur une application de *retexturing* utilisant la prédiction de spécularité de JOLIMAS. Cependant, n'utiliser JOLIMAS que pour une amélioration du rendu est réducteur, car ce modèle fournit des informations très complètes en termes de géométrie qui peuvent profiter à de nombreuses applications comme évoquées ci-après.

Amélioration de la pose de caméra. De façon similaire à Lagger et al. (2008) (illustré à la figure 7.8), nous pouvons utiliser les spécularités et plus particulièrement notre prédiction de spécularité pour raffiner davantage la pose de la caméra. En effet, dans un processus d'ajustement de faisceaux qui est une méthode classique de raffinement non-linéaire des points 3D reconstruits et de la pose de caméra, l'erreur de reprojection entre les points 3D projetés et leur correspondance 2D dans l'image doit être minimisée. Il serait envisageable de rajouter à la fonction de coût une minimisation de la distance point/point entre l'ellipse prédite par JOLIMAS et les contours de la spécularité détectée dans les images.

Géométrie à partir de la spécularité. Quelques travaux ont tenté de faire le lien entre courbure et spécularité comme Healey et Binford (1988); Blake et Brelstaff (1988) en prenant les informations d'intensité et de position de la spécularité comme primitives

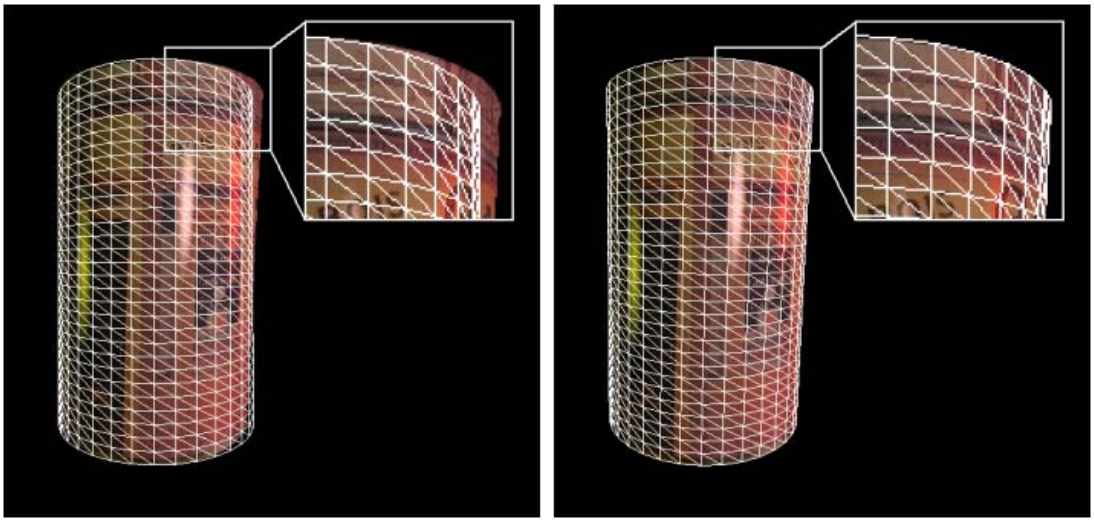


FIGURE 7.8 – Utilisation des spécularités pour raffiner la position de la pose de caméra et respectivement l'alignement du maillage du modèle de boite en métal. Images extraites de l'article de Lagger et al. (2008).

de reconstruction de surface. Nous avons eu l'occasion pendant nos travaux d'étudier ce lien entre forme/position de la spécularité et courbure. Une extension possible serait de reconstruire le modèle JOLIMAS à partir d'une surface dont la géométrie est connue et de calculer la courbure d'une autre surface en comparant la forme de la spécularité détectée avec la spécularité prédite par la quadrique.

***Retexturing* plus avancé.** De façon plus évidente qu'avec le *shading*, les spécularités mettent davantage en évidence la micro-géométrie/rugosité d'une surface. Il serait intéressant de reconstruire une carte de normales associée à une surface afin d'améliorer le rendu de notre *retexturing*. En effet, la micro-géométrie d'une surface a un impact conséquent sur la forme de la spécularité et sur le réalisme.

Initialisation des modèles d'illumination locale. De façon semblable à notre application de *retexturing*, il est pertinent d'utiliser la quadrique permettant la prédiction de spécularité comme position initiale de la source de lumière dans un modèle d'illumination locale de type Phong, Blinn-Phong ou encore Cook-Torrance. En général, il est difficile de calculer simultanément tous les paramètres des modèles, car certains sont de nature photométrique (intensité, couleur) alors que d'autres sont de nature géométrique (position de la lumière). De plus, la forme de la quadrique reconstruite par JOLIMAS donne des indications intéressantes sur le type de lumières présentent dans la scène comme les néons ou les spots. Il serait intéressant de proposer une fonction d'intensité qui prend en compte la forme de la quadrique.

Utilisation en réalité diminuée. Récemment, des approches comme Said et al. (2017) ont mis en évidence l'importance de gérer les spécularités dans le domaine de la réalité diminuée comme illustré à la figure 7.9. En effet, pour les applications de réalité diminuée, l'objectif est de supprimer un objet d'intérêt dans une séquence vidéo en temps réel et de façon naturelle. En présence de spécularités dans une scène, il peut être difficile

(a) (b)

FIGURE 7.9 – Exemple de prise en compte des spécularités dans une application de réalité diminuée. Dans cette figure, une clé USB est supprimée d'une table en bois en présence d'une spécularité. Il est important ici de synthétiser la spécularité lorsqu'elle se déplace sur la zone où la clé USB a été supprimée pour assurer une cohérence visuelle. Images extraites de Said et al. (2017).

de supprimer un objet d'une surface réfléchissante de façon cohérente si les spécularités retirées par le processus d'*inpainting* ne sont pas synthétisées à nouveau. De plus, si une spécularité se déplace sur la zone d'*inpainting*, il ne faut pas que celle-ci disparaisse pour des raisons de cohérence visuelle. Pour pallier ce problème de façon générale, il pourrait être utile d'utiliser le modèle JOLIMAS et sa prédiction de spécularité pour synthétiser les spécularités de la même manière que dans notre application de *retexturing*.

Identification des matériaux. Dans notre modèle JOLIMAS, de nombreux paramètres sont abstraits par la forme de spécularité (intensité, ouverture du capteur, rugosité, réflectance). Après avoir reconstruit un modèle JOLIMAS pour une surface donnée, nous n'avons pas d'indications sur la réflectance ou la rugosité. Lors de nos expériences, nous avons remarqué que la rugosité et la réflectance avaient un impact sur l'échelle de la spécularité. Ainsi, lorsque nous prédisons la spécularité sur une surface de matériau différent de celui de la surface où la reconstruction de la quadrique a été réalisée, nous pouvons calculer un coefficient qui inclut la rugosité et la réflectance correspondant à une différence d'échelle entre la forme des deux spécularités. Nous pourrions également utiliser des méthodes d'apprentissage à partir d'une base de données de réflectance associée à différents matériaux comme celle de MERL Matusik et al. (2003). A noter qu'un des problèmes compliqués en apprentissage profond est d'apprendre des représentations universelles (capable de gérer un maximum de modalité, de domaine visuel et de domaine sémantique) Tamaazousti et al. (2017a,b).

ANNEXE A

Détection de spécularité dans les images en temps réel

Dans cette annexe, nous détaillerons le processus de détection de spécularité utilisé dans ce mémoire. Plus particulièrement, nous proposons une méthode générique pour la détection en temps réel des réflexions spéculaires dans une image. Cette méthode utilise une nouvelle technique de seuillage appliquée dans l'espace colorimétrique Teinte-Saturation-Valeur (TSV). Les méthodes de l'état de l'art ne prennent souvent pas en compte les différents changements du contexte lumineux (sauts de lumières). De plus, ces méthodes sont généralement algorithmiquement complexes et non adaptées à un contexte temps réel. Notre méthode répond à cette problématique de détection en trois étapes : adaptation du contraste de l'image pour gérer les variations d'intensité de la lumière, un seuillage automatique afin d'isoler les réflexions spéculaires et un post-traitement afin de limiter davantage toute mauvaise détection. Cette méthode a été comparée avec l'état de l'art par deux protocoles expérimentaux basés sur la précision des contours et du centre de gravité ; elle propose, en temps réel, des résultats précis et sans a priori sur les images. ou pour des applications de réalité augmentée Jachnik et al. (2012)). Les éléments de lumières comme les ombres ou les reflets peuvent fortement améliorer la qualité de rendu.

Dans le domaine du traitement d'image, pour certaines opérations comme la segmentation, la détection ou encore la correspondance d'objets, la présence de spécularités peuvent perturber les résultats. Par exemple, pour une application de segmentation colorimétrique comme dans Deng et al. (1999), les spécularités sont souvent considérées comme perturbantes. Dans le domaine médical, elles peuvent être à l'origine de faux positifs lors d'une détection d'éléments suspects. Pour cette raison, la restauration d'image est utilisée afin de les retirer (Lee et al. (2010); Oh et al. (2007); Saint-Pierre et al. (2011); Stehle (2006)).

Néanmoins, la détection de spécularités est un processus complexe. Gérer efficacement les variations de luminosité comme les forts contrastes, les images sombres ou légèrement surexposées sous différentes sources lumineuses d'intensités variables, est une problématique difficile. Ainsi, les méthodes de l'état de l'art se limitent à leur domaine d'application et sont donc très spécifiques. Cependant, les spécularités plusieurs caractéristiques communes à de nombreux domaines d'applications peuvent être exploitées. Notre méthode vise

un large panel d'applications sans a priori sur le contexte lumineux pour une détection en temps réel.

Dans la section A.1, nous présentons les méthodes associées en soulignant les différentes applications visées, leurs résultats et leurs limitations. Ces méthodes sont divisées en deux approches : détection hors-ligne A.1.1 et en ligne A.1.2. Dans la section A.2, notre approche est décrite en soulignant sa fiabilité, son efficacité et sa généricité. Nous avons positionné notre méthode par rapport à l'état de l'art sur 3 étapes : prétraitement (section A.2.2), seuillage (section A.2.3) et post-traitement (section A.2.4). Nos résultats sont présentés et comparés, dans la section A.3 pour mettre en avant l'efficacité et la généricité de notre méthode.

A.1 Travaux antérieurs

La détection de spécularités est un processus qui se divise, en général, en 3 étapes. Dans un premier temps, un prétraitement afin de gérer le bruit et le contexte lumineux potentiellement variable. Par la suite, un seuillage est réalisé afin d'isoler les spécularités. Le seuil utilisé peut être prédéfini ou calculé automatiquement. Pour finir, un post-traitement est réalisé afin d'éliminer les faux positifs ou récupérer les faux négatifs manquants.

Ces étapes sont présentées dans le tableau A.1 avec d'autres critères comme le choix de l'espace colorimétrique utilisé et le type de donnée d'entrée (image ou vidéo) afin de donner une vue globale de l'état de l'art et de notre positionnement.

A.1.1 Approches hors-ligne

Dans un contexte où la qualité du résultat prime sur le temps d'exécution, plusieurs méthodes hors ligne sont disponibles pour la détection de spécularité. Stehle (2006) ont utilisé un seuillage sur l'espace colorimétrique YUV, Y étant le canal de luminance et U, V les canaux de chrominance. Le principe de cette méthode est de seuiller le canal Y à la valeur d'intensité correspondant au dernier pic de son histogramme. Selon Stehle (2006) ce dernier pic de petite taille, correspond aux pixels spéculaires. Cette méthode est utilisée dans l'imagerie endoscopique où les images d'entrées ont une meilleure résolution, un histogramme bien égalisé où les problèmes lumineux sont fortement atténués. Néanmoins, dans des images dépourvues de spécularités, cette méthode de seuillage pourrait engendrer des faux positifs comme les objets de couleur blanche. De plus, une image comportant des spécularités n'a pas nécessairement de pic à la fin de son histogramme dans son canal de luminance. Oh et al. (2007) appliquent deux seuillages : un seuillage global donnant une première estimation des réflexions spéculaires et un deuxième seuillage moins restrictif afin de récupérer les derniers éléments spéculaires restants. Ces éléments apparaissent comme ayant un contraste élevé et blanc (illustré à la figure A.4(a) et A.4(c)). Dans leur méthode, l'image est convertie dans l'espace TSV afin d'appliquer un seuillage sur les canaux Valeur et Saturation et d'obtenir une première estimation des candidats spéculaires. Par la suite, une segmentation spatiale est appliquée afin de séparer l'image source en k régions pour convertir celles-ci en de nouvelles zones qui seront fusionnées avec les candidats spéculaires. Cette méthode donne des résultats convenables, mais souffrants des mêmes limitations que Stehle (2006).

Les méthodes de Torres et al. (2003) et Ortiz et Torres (2006) utilisent un histogramme bidimensionnel appelé diagramme MS afin de seuiller l'image pour un seuil estimé à partir

Propriétés / Méthodes	Espace colorimétrique	Seuil	Pré traitement	Post traitement	Vidéo Image
Ortiz et Torres (2006)	MS	Constant	EH	OM	Image
Stehle (2006)	YUV	Auto	Aucun	OM	Image
Oh et al. (2007)	TSV	Constant	OHF	Segmentation et OM	Image
Arnold et al. (2010)	RVB + niveau de gris	Auto	Aucun	OM	Vidéo
Notre méthode	**TSV**	**Auto**	**Contraste**	**Dégradé**	**Vidéo**

TABLE A.1 – Classification de l'état de l'art et positionnement de notre méthode.

de ce diagramme. Une égalisation d'histogramme (EH) est utilisée pour garder un seuil constant pour chaque image. Ces méthodes donnent des résultats précis et rapides, mais l'EH peut entrainer de mauvaises détections (augmentation du bruit). Torres et al. (2003); Ortiz et Torres (2006) ont souligné l'importance du choix de l'espace colorimétrique à utiliser comme le TSV. En effet, cet espace colorimétrique est proche de la manière donc la perception humaine analyse une image. En effet, celle-ci décompose une image en caractéristiques basiques comme l'intensité de la couleur, la teinte et la brillance.

A.1.2 Approches en ligne

Plusieurs domaines requièrent une méthode temps réel pour la détection de spécularités tel que les applications médicales Arnold et al. (2010) ou la réalité augmentée Jachnik et al. (2012). Arnold et al. (2010) utilisent un seuillage à partir d'informations dans l'espace RVB et dans une image en niveau de gris. En RVB, les valeurs de fortes intensités de bleu, rouge ou vert sont considérées comme des spécularités.

Contrairement à l'approche de Oh et al. (2007) utilisant des seuils constants, Arnold et al. (2010) utilisent un seuil adaptatif calculé à partir des canaux verts et bleus. Ainsi un premier seuillage est réalisé afin de sélectionner les premiers candidats. Un deuxième seuillage est utilisé par la suite pour trouver les derniers candidats plus incertains en utilisant un filtre médian et un seuil moins restrictif. Ces derniers candidats vont être traités dans une étape de post-traitement utilisant quelques opérations morphologiques (OM), calculant les contours des spécularités et supprimant les zones de tailles conséquentes. En effet, ces zones sont susceptibles d'être trop lumineuses et non spéculaires. Dans le domaine de l'imagerie endoscopique, cette méthode est fiable et rapide, mais manque de généricité tout en étant sensible aux surfaces blanches. En effet, les systèmes endoscopiques permettent d'avoir un contrôle précis de la source lumineuse et fournissent une correction automatique de l'exposition ce qui permet d'éviter des cas de sur/sous exposition.

(a) (b)

FIGURE A.1 – Résultat de notre égalisation de contraste : (a) présente le cas surexposé, difficilement exploitable en l'état. (b) montre l'image (a) après notre égalisation de contraste

A.2 Méthode proposée

A.2.1 Vue d'ensemble

Notre méthode a pour but d'éviter les restrictions dans la détection de spécularités et de gérer plusieurs spécularités en temps réel. Nous ne considérons pas l'histogramme de l'image comme égalisé. Au contraire, chaque problème lumineux pouvant se produire dans un flux vidéo doit être pris en compte (illustré à la figure A.1(a)).

De plus, nous ne faisons aucun *a priori* sur la taille d'une spécularité et notre détecteur doit être suffisamment rapide pour gérer des applications temps réel.

Nous avons choisi de travailler dans l'espace TSV, car les spécularités y ressortent naturellement (voir figure A.2) et deux critères peuvent être utilisés dans l'espace TSV au lieu d'un seul dans l'espace RVB (voir figure A.2). En effet, dans l'espace TSV, les spécularités sont caractérisées par une intensité basse dans le canal Saturation basse et une haute intensité dans le canal Valeur contrairement à l'espace RVB, où ces spécularités sont représentées aux pixels de haute intensité dans les trois canaux (surfaces blanches).

A.2.2 Prétraitement

Les images surexposées sont, en général (figure A.1(a)) difficilement exploitables. En effet, celles-ci ont un contraste élevé pouvant impacter des méthodes de segmentation ou de détection. Pour limiter ce problème, une égalisation du contraste est appliquée sur l'image en calculant la luminosité de l'image définie dans l'équation (A.1). Si cette valeur est au-dessus d'un certain seuil T_b (seuil de luminosité), l'image (I) est considérée comme surexposée et le contraste est abaissé. Le résultat de cet algorithme est illustré à la figure A.1(b).

$$I_{\text{Luminosité}} = \sum_{x \in I} \frac{\sqrt{0.241x_R^2 + 0.691x_V^2 + 0.068x_B^2}}{Largeur * Hauteur} \tag{A.1}$$

La luminosité est un facteur pertinent pour un seuillage automatique, car une augmentation de la luminosité provoque également une augmentation proportionnelle dans le canal valeur.

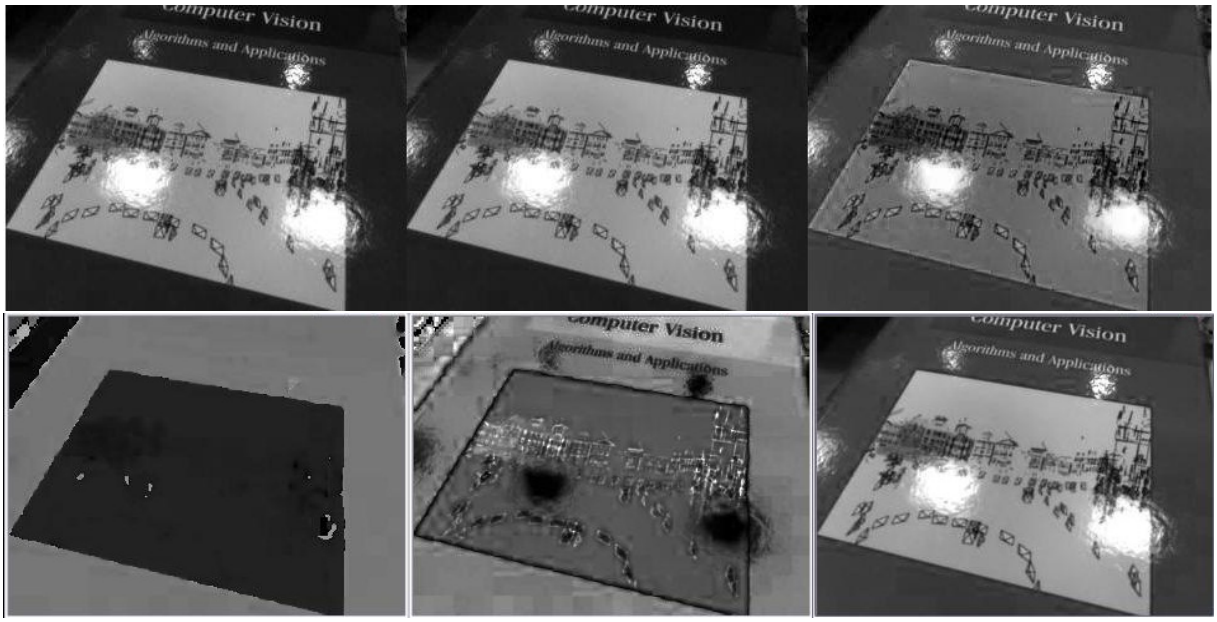


FIGURE A.2 – Représentation RVB (haut) et TSV (bas) d'une image contenant des spécularités. Le canal de Teinte peut être ignoré

A.2.3 Étape de seuillage

Afin de gérer différentes conditions lumineuses et différents contrastes, nous pouvons appliquer une égalisation d'histogramme sur le canal valeur. Cependant, selon Ortiz et Torres (2006) cette égalisation augmente le bruit et l'intensité de certaines zones ce qui peut donner de mauvais résultats et amener de mauvaises détections. Ainsi, au lieu d'utiliser des seuils constants (Oh et al. (2007) et Ortiz et Torres (2005)), un seuil dynamique est utilisé sur le canal valeur afin d'éviter toute opération morphologique pouvant bruiter davantage l'image. Ce seuil est estimé à partir de la relation entre la luminosité et le canal valeur. En effet, nous avons remarqué empiriquement une proportionnalité entre notre seuil T_v, fixé sur le canal valeur à une valeur donnant les meilleurs résultats, et la luminosité d'une image.

Afin d'illustrer cette observation, nous avons réalisé une régression linéaire sur une base de 50 images variées en taille, contraste, contexte lumineux et en intensité lumineuse. Pour chaque image, le seuil T_v, correspondant à la valeur du seuil du canal valeur donnant les meilleurs résultats de détection de spécularités, a été manuellement estimé (voir figure A.3). Le résultat de cette régression est ainsi utilisé pour calculer le seuil k_v qui est le rapport entre T_v, seuil du canal valeur optimal, et la luminosité $I_{Luminosité}$ (A.2).

Afin d'utiliser toutes les informations disponibles, nous avons également seuiller le canal saturation avec le seuil T_s. Ce seuil a été fixé à une valeur constante, car le comportement de la saturation représente l'intensité de la couleur ce qui est plus difficile à exploiter pour la détection de spécularité.

$$T_v = I_{\text{Luminosité}} k_v. \tag{A.2}$$

Nos conditions de seuillage sont présentées dans l'algorithme (A.3).

$$\text{si } S(x) < T_s \text{ et } V(x) > T_v, \text{le pixel est spéculaire} \tag{A.3}$$

avec $S(x)$ et $V(x)$ la valeur du canal saturation et valeur du pixel x.

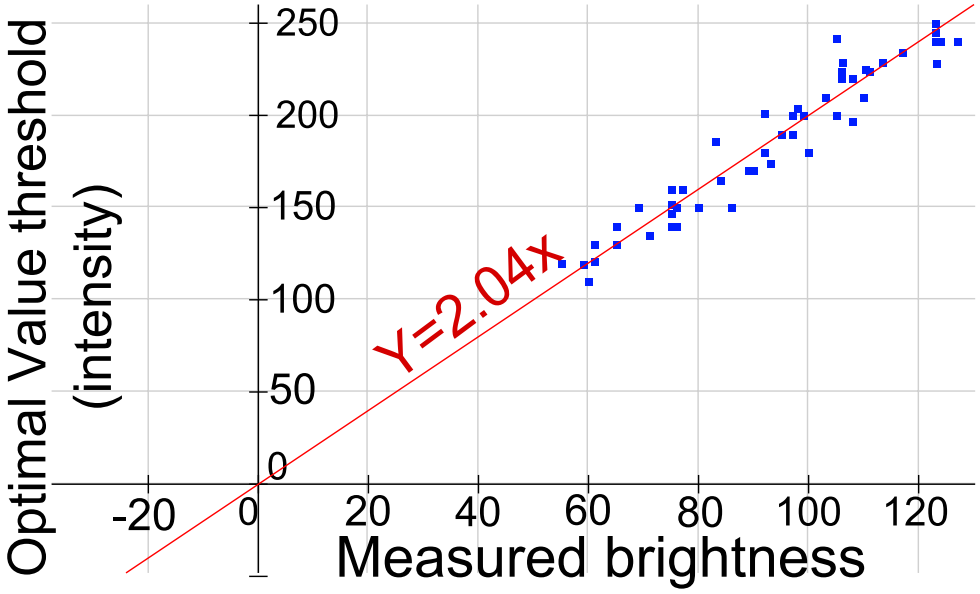


FIGURE A.3 – Résultats de la régression linéaire donnant une relation entre T_v et la luminosité globale de l'image suivant une fonction linéaire $y = k_v x$ avec $k_v = 2$ (ligne rouge). Les carrés bleus représentent la valeur optimale de T_v en fonction de la luminosité pour chaque image.

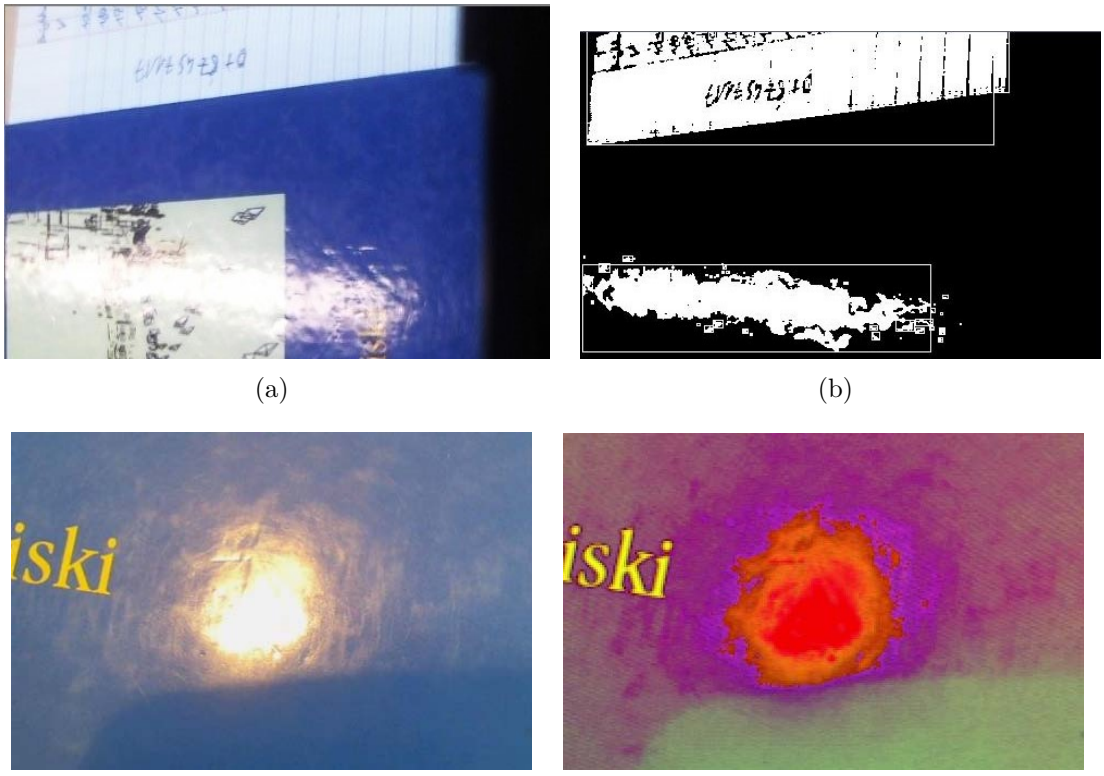
Cette étape de seuillage produit des résultats robustes pour des conditions difficiles et permet un meilleur contrôle du contexte lumineux. En effet, nous pouvons désactiver la détection de spécularités quand l'image est trop exposée (haute luminosité) pendant des sauts de lumières pour éviter des détections incorrectes et fournir plus de fiabilité pour une application temps-réel.

A.2.4 Post-traitement

Après une égalisation du contraste, certaines zones dans l'image restent faussement détectées pour une image surexposée (illustré à la figure A.4(a)). En effet, une texture blanche renvoie une grande quantité de lumière ce qui peut fausser la détection, car son intensité lumineuse est très élevée. Pour répondre à ce problème, nous avons implémenté une méthode pour séparer les taches spéculaires des textures blanches.

En observant l'image dans le canal Valeur du TSV, il est facile de constater une décroissance progressive de l'intensité de celles-ci à partir de leurs centres. Cette propriété est visible à la figure A.4(d).

Pour comprendre ce constat, il est pertinent de rappeler le comportement d'un rayon incident sur une surface quelconque. En effet, plusieurs éléments sont à prendre en considération : la composante diffuse, la composante directionnelle et la composante spéculaire. L'impact de ces différentes composantes est déterminé par la BRDF, qui est une fonction bidirectionnelle permettant de décrire la radiance selon le rayon incident, le matériau et l'angle de réflexion. Sur la figure A.5, on constate que pour un point de vue donné, la composante spéculaire décrit un pic correspondant à la composante spéculaire idéale (réflexion stricte du rayon incident) et d'une intensité décroissante correspondant à la composante directionnelle. On peut modéliser cette zone spéculaire sous forme d'une gaussienne en



(a) (b)

(c) (d)

FIGURE A.4 – Dans l'image (a), le papier a un aspect semblable à une spécularité et pourrait être détecté à tort comme étant une spécularité. (b) montre l'étape de séparation des candidats spéculaires de la figure (a). (c) représente une spécularité vue dans l'espace RVB. (d) illustre la décroissance progressive de l'intensité d'une spécularité à partir de son centre dans l'espace TSV.

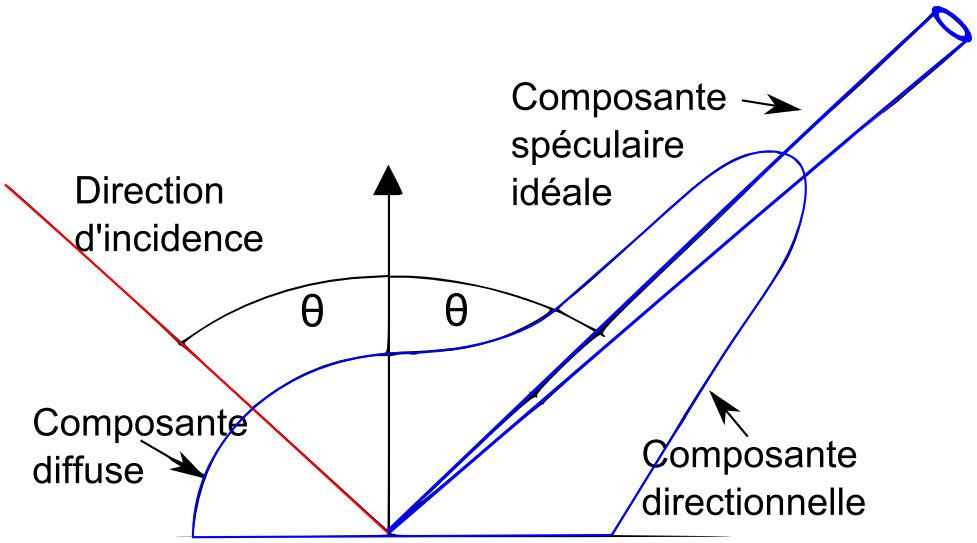


FIGURE A.5 – Représentation de la répartition lumineuse provoquée par la réflexion d'un rayon incident avec un objet. Cette répartition est déterminée en utilisant la BRDF du matériau. Nous constatons également que l'intensité de la lumière pour une réflexion spéculaire décroit progressivement autour du centre de gravité de la spécularité (composante spéculaire idéale).

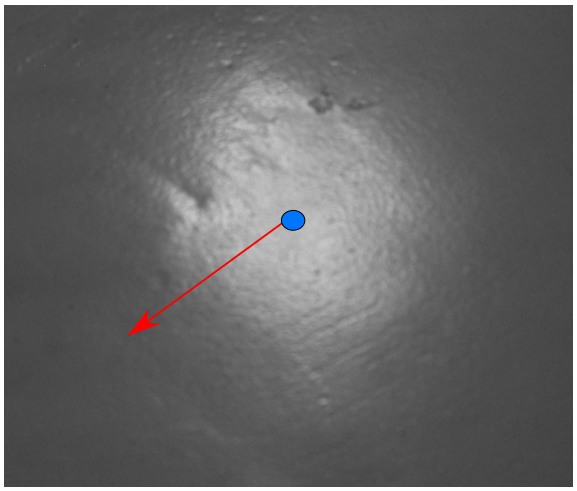
trois dimensions centrées au centre de gravité de la réflexion spéculaire.

Une première idée serait de partir du centre de gravité de chaque spécularité et d'analyser l'évolution de l'intensité sur la droite reliant le centre de gravité à un point de contours. Théoriquement, l'allure de cette courbe d'intensité devrait représenter une fonction sigmoïde. Ainsi, en présence d'une texture, cette courbe décrit une fonction échelon ce qui nous permettrait de discriminer les textures faussement détectées des réflexions spéculaires.

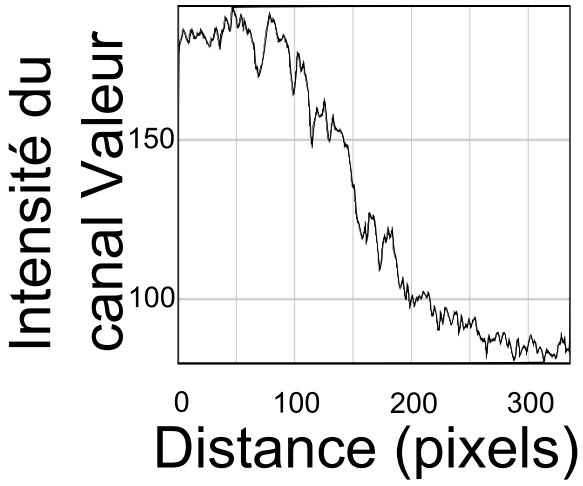
En pratique, ce modèle présente quelques limites, car cette décroissance d'intensité bien que visible n'est pas stricte (illustré à la figure A.6(b)). Il est possible de déduire une sigmoïde caractéristique pour chaque spécularité, mais réaliser une correspondance entre cette variation progressive de l'intensité avec une fonction adaptée comme la sigmoïde est sujet à de nombreuses erreurs.

Afin d'exploiter cette décroissance progressive de l'intensité pour séparer textures des spécularités, nous divisions l'image après notre seuillage en k-régions de candidats spéculaires. Cette segmentation est réalisée en utilisant un algorithme classique de segmentation d'images binaires Suzuki et Satoshi (1985). Nous prenons la boîte englobante la plus grande pour chaque contour. En effet, une spécularité n'est pas uniforme et est généralement très fragmentée. En utilisant une boîte englobante, nous incluons ces fragments dans le calcul. Dans un second temps, en modifiant le seuil T_v d'un pas unitaire pour chaque itération, nous allons observer l'évolution de l'aire de ces boîtes englobantes. Si cette évolution est constante (décroissance légère et régulière (voir figure A.7) ou si la variation est nulle, cette aire est considérée comme une réflexion spéculaire. Si l'aire décroit soudainement, cette zone a été détectée à tort comme étant spéculaire.

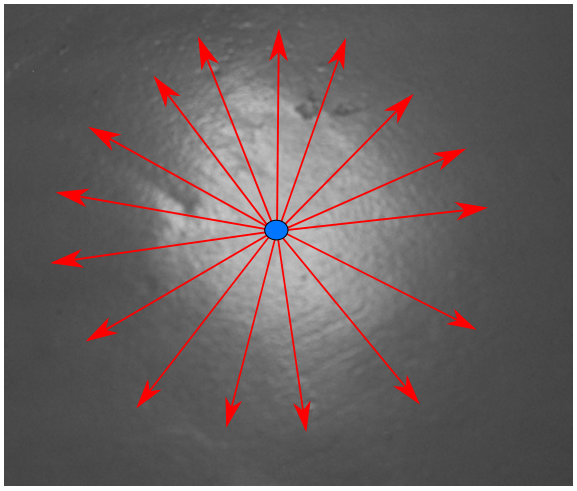
Ce critère de variation d'intensité à partir du centre est visible pour chaque spécularité, mais est difficilement applicable pour de petites surfaces car le dégradé existe, mais de façon infime. On peut noter que dans certains cas, les textures blanches ont un aspect



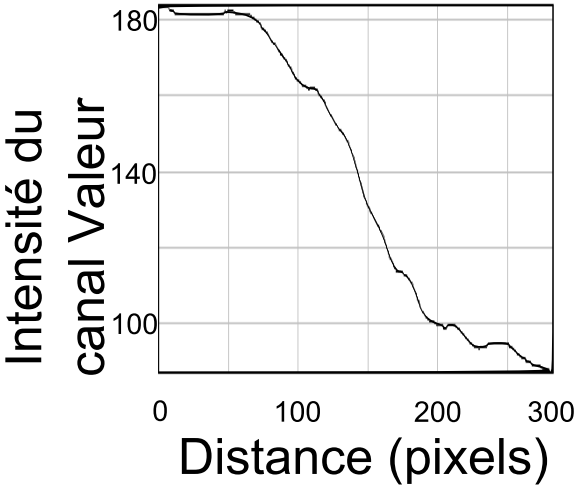
(a)



(b)

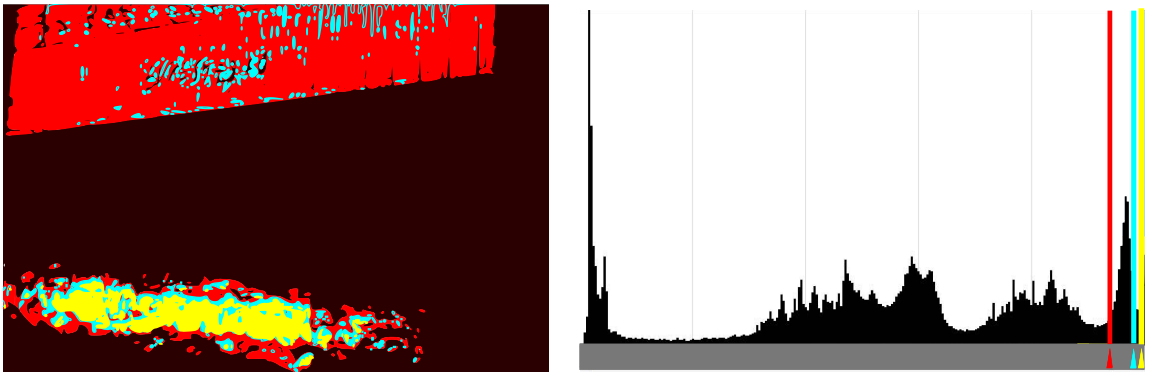


(c)



(d)

FIGURE A.6 – Analyse de la variation d'intensité du canal Valeur suivant. (a) représente un contour allant du centre de gravité et passant par un des contours de la spécularité. On observe une forme de sigmoïde très bruitée au niveau de son intensité dans (b). (c) représente un ensemble de contours sur le même principe que (a). La moyenne de l'intensité de ces contours nous donne un sigmoïde beaucoup plus exploitable.



(a) (b)

FIGURE A.7 – Évolution de l'aire d'une spécularité comparée à une texture blanche pour une image surexposée. (a) représente l'évolution de l'aire de la spécularité de l'image à la figure A.4(b) pour 3 seuils donnés (marqués par les couleurs rouges, bleus et jaunes). Nous pouvons observer que la spécularité a une aire qui décroit progressivement contrairement à la feuille de papier qui s'estompe dès le deuxième seuil. (b) représente notre histogramme du canal Valeur de notre image. Nous pouvons remarquer un pic à la fin de celui qui représente la feuille de papier et la spécularité. Une augmentation progressivement du notre seuil T_v fait disparaitre plus rapidement la texture blanche de la réflexion spéculaire.

dégradé, mais ce cas reste isolé et rare.

A.2.5 Complexité algorithmique

Une étude sur la complexité algorithmique a été menée sur chaque étape de notre méthode (illustré au tableau A.2). La complexité globale calculée est de $O(n+kc)$ avec n représentant le nombre de pixels de l'image, k le nombre de régions trouvées dans notre segmentation et c le nombre d'itérations durant notre étude de la variation d'intensité à partir du centre.

Étapes	Complexité
Égalisation du contraste	$O(1)$
Seuillage	$O(n)$
Segmentation + étude du dégradé	$O(nk) + O(kc)$
Total	$O(n+kc) \sim O(n)$

TABLE A.2 – Complexité globale de notre méthode.

A.3 Évaluation expérimentale

A.3.1 Protocole expérimental

Nous comparons nos résultats avec notre implémentation de la méthode de Arnold et al. (2010) qui est une approche rapide de l'état de l'art donnant de bons résultats en temps réel. De plus, Arnold et al. (2010) se sont positionnés par rapport à Oh et al. (2007)

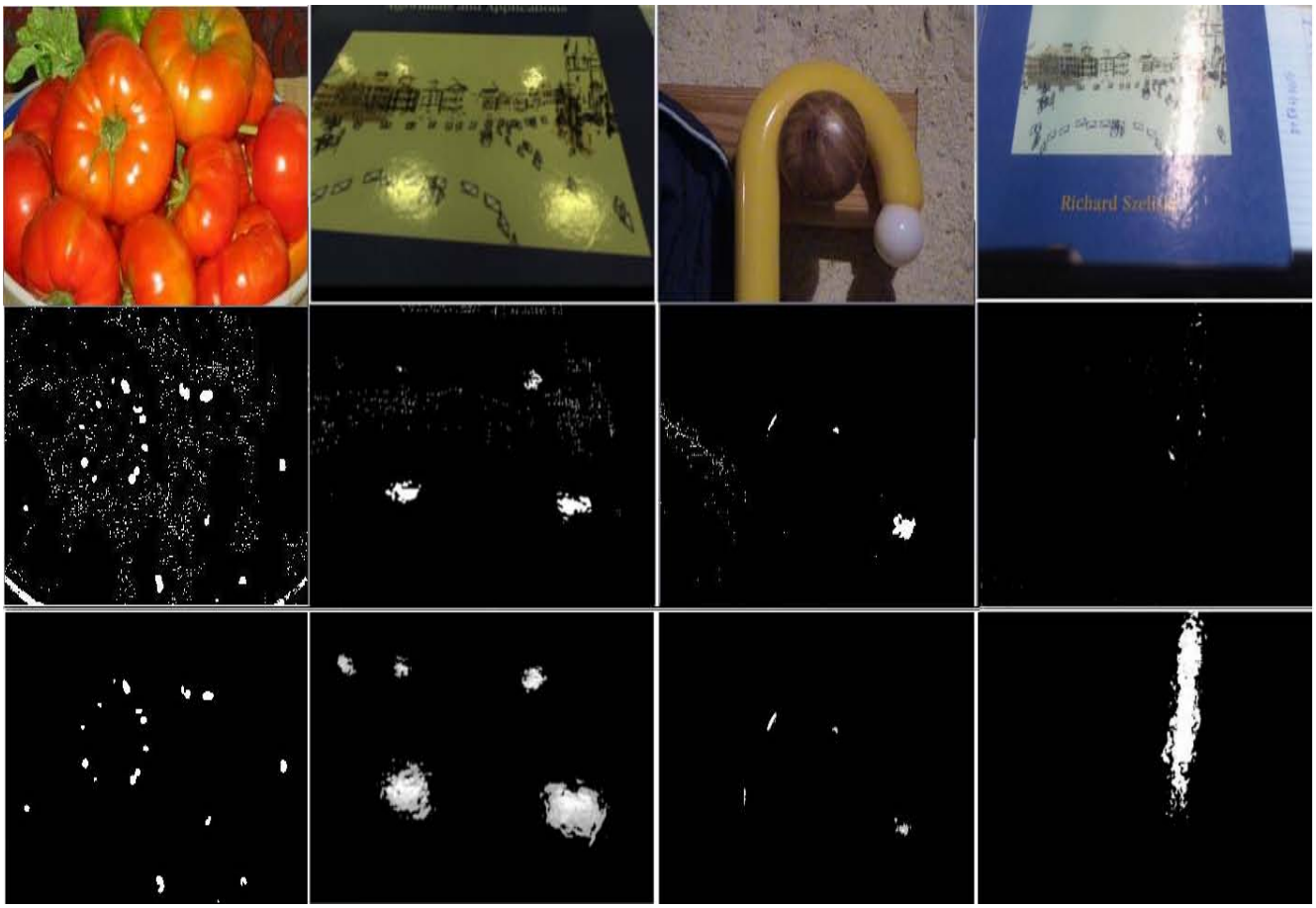


FIGURE A.8 – Résultats de la détection de spécularités de l'approche de Arnold et al. (2010) (au milieu) en comparaison avec notre méthode (en bas).

en mettant en évidence l'efficacité et la vitesse de leur approche. L'implémentation de la méthode a été réalisée en utilisant les paramètres conseillés par Arnold et al. (2010).

Ces approches ont été testées sur la base de données de 100 images tests (différentes de notre base d'entrainement utilisé pour fixer les différents seuils) pour mettre en évidence la généricité et l'efficacité de la méthode proposée.

A.3.2 Évaluation quantitative

Notre évaluation quantitative se fera sur deux critères. Le premier critère à mettre en évidence correspond à la précision des contours des spécularités. L'objectif de ce critère est de comparer le contour d'une spécularité évaluée par l'œil humain avec les résultats de l'algorithme. Ainsi, un contour manuel a été réalisé sur la base de données de 100 images et comparé avec les contours des spécularités détectées. Les contours ont été calculés en utilisant l'opérateur de Sobel avec un noyau 3×3. Un contour est considéré comme précis si pour chaque point, nous trouvons un pixel de notre vérité terrain appartenant au voisinage. Celui-ci est représenté par un patch de taille 9×9 pour être suffisamment strict sans pour autant accepter les erreurs critiques (la figure A.9(a) présente notre évaluation du contour).

Un deuxième critère à mettre en évidence et l'importance d'avoir un centre de gravité

Méthode / Évaluation	Arnold et al. (2010)	**Notre méthode**
Temps (en secondes)	0.0584	**0.0371**
Contour (en %)	70.3	**80.29**
Centre de gravité en %	67.2	**78.13**

TABLE A.3 – Évaluation quantitative. Pour les trois critères, notre méthode est plus pertinente et offre de meilleurs résultats.

précis (voir A.9(b)). En effet, c'est à partir de ce centre de gravité que la décroissance d'intensité dans le canal Valeur se produit. L'intérêt de ce critère est qu'il existe plusieurs cas où une spécularité peut être partiellement obstruée (rebord d'une surface, ombre coupant la spécularité ou diverses occultations comme illustrées aux images A.4(c) et A.4(d)). Dans ces cas, le calcul du centre de gravité simple estimé avec le contour de la spécularité ne correspond pas au centre réel représenté. Pour estimer ce centre, à partir du contour basique, nous réduisons ce contour jusqu'à obtenir la valeur maximale (intensité stable ou à une valeur de 255). Ce nouveau contour réalisé est utilisé pour calculer un centre de gravité plus pertinent qu'une estimation avec le contour initial. La vérité terrain a également été réalisée manuellement en affinant le contour initial de la spécularité jusqu'à avoir le contour du noyau de la spécularité et utiliser celui-ci afin calculer un centre de gravité plus pertinent. Ces deux critères ont été évalués sur une base de données incluant les images de l'article de Arnold et al. (2010) et sur la base de données de 100 images test.

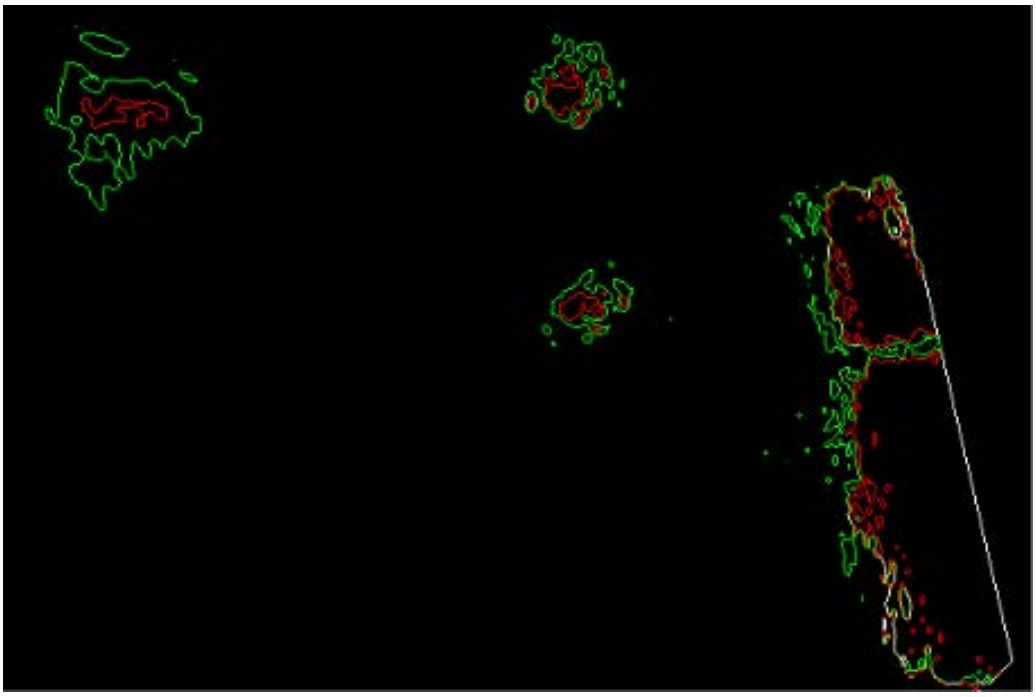
Notre approche s'avère 1.5 fois plus rapide que Arnold et al. (2010), plus précise avec une meilleure estimation du centre de gravité.

A.3.3 Évaluation qualitative

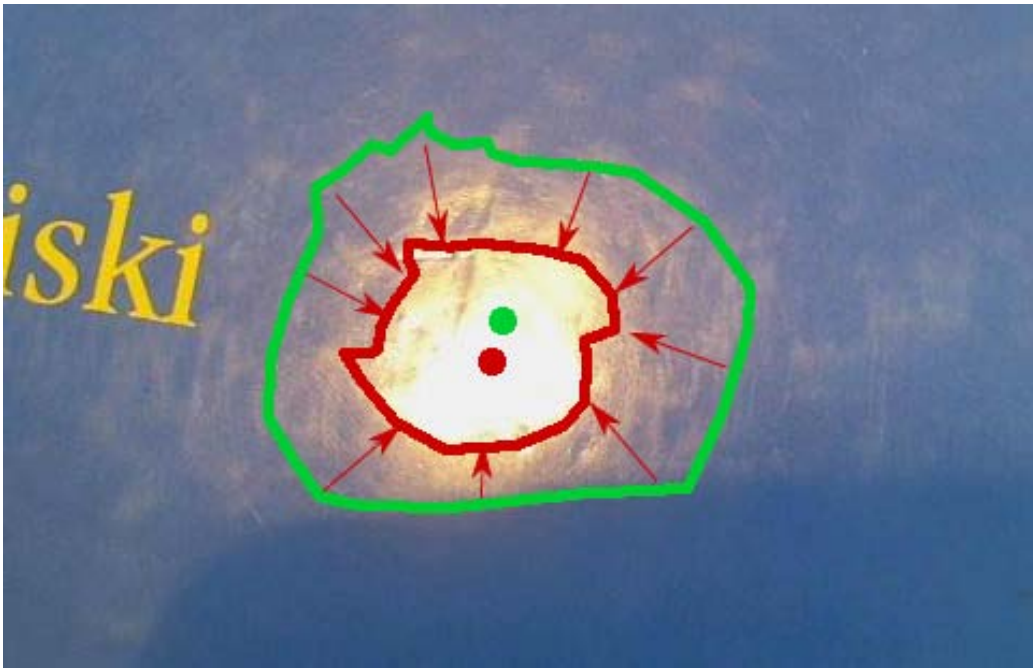
En comparaison avec Arnold et al. (2010), le bruit est grandement réduit pour toutes les images. La figure A.8 met en évidence qu'une spécularité peut être de taille conséquente. Utiliser un critère de taille comme Arnold et al. (2010) peut amener à ignorer certaines spécularités de grande taille ce qui n'est pas souhaitable pour une approche générique de la détection de spécularités.

A.4 Conclusion et discussion

Dans cet article, nous avons présenté une nouvelle approche pour la détection de spécularités en utilisant des propriétés simples et efficaces des spécularités. Nous avons utilisé les canaux de valeur et de saturation de l'espace TSV afin d'estimer les différents seuils calculés automatiquement en fonction de la luminosité globale de l'image. De plus, nous avons proposé un prétraitement d'ajustement automatique du contraste afin de gérer les variations d'illumination et un post-traitement qui observe la décroissance de l'intensité du canal valeur à partir de son centre de gravité afin gérer les régions détectées à tort dans notre seuillage. Notre méthode est générique et fiable pour différentes applications temps-réel. Cette approche a été comparée avec l'état de l'art en utilisant un nouveau protocole expérimental fondé sur deux propriétés : la précision du contour et du centre



(a)



(b)

FIGURE A.9 – Évaluation proposée pour la détection de spécularités : contours et centre de gravité. (a) différences entre notre vérité terrain (en vert) et les résultats de notre détection (en rouge). Les pixels communs sont affichés en gris. (b) Estimation du centre de gravité. Les lignes rouges représentent le vecteur direction du dégradé et l'avancement de notre contour initial (vert) pour s'arrêter au contour du noyau de la spécularité (rouge).

de gravité d'une réflexion spéculaire.

Les approches de l'état de l'art pour la détection de spécularités en temps réel sont

basées sur du seuillage. Néanmoins, l'utilisation de modèle de réflexion comme le modèle Lambertien Brelstaff et Blake (1988) afin de prendre en compte les composantes diffuses de l'image ou séparer les composantes diffuses et spéculaires comme Tan et al. (2004) peut être pertinente pour améliorer la précision et la réduction du bruit dans les résultats.

ANNEXE B

Reconstruction d'objet sous forme de quadrique

La reconstruction d'objet sous forme de quadrique est primordiale dans ce travail de thèse. Cependant, nos contributions vont au-delà du cadre de la thèse puisque la reconstruction de quadrique peut être utilisée dans de nombreux domaines comme la localisation d'objets, l'estimation d'enveloppe englobante d'objet, mais également la localisation de caméra.

B.1 Introduction

Dans le domaine de la vision par ordinateur, les quadriques sont très largement utilisées notamment dans la reconstruction d'objet d'une forme proche d'une quadrique (sphère, cylindre, ellipsoïde, ...), mais également pour approximer la forme d'objets plus complexes. Cette approximation forme une boite de délimitation (*bounding box*) ce qui peut donner une première pose approximative d'un objet ou pour calculer une collision approximative entre deux objets ce qui est moins couteux que de calculer la collision entre deux maillages. Dans ce chapitre, nous évoquons les différentes méthodes de reconstruction de quadrique utilisées ainsi que les métriques et techniques utilisées pour évaluer la qualité de la reconstruction.

B.2 État de l'art

Notre modèle géométrique repose sur l'existence d'une quadrique reconstruite par l'observation d'ellipses calculées à partir des spécularités. Il est important d'évoquer l'état de l'art des méthodes de reconstruction de quadrique, car le phénomène de spécularité étant hautement variable, nous avons besoin d'une méthode de reconstruction particulièrement résistante au bruit des données d'entrée.

B.2.1 Approches par reconstruction multi-vues

Une des premières approches de reconstruction de quadriques correspond à celle de Karl et al. (1994) et Ma et Chen (1994). Ces deux articles ont exprimé l'existence d'une relation linéaire entre l'enveloppe d'une quadrique et la silhouette (ou contour occultant) dans le plan image de la caméra, mais non généralisé à l'ensemble des classes de quadrique. Par la suite, Ma et Li (1996) ont mis en évidence que la reconstruction d'une quadrique n'est pas possible de façon immédiate avec seulement deux vues mais ont omis de préciser les différentes raisons autres que la dégénérescence de la solution algébrique. De plus, aucune précision n'est donnée sur l'ambiguïté de la reconstruction. L'inclusion de la reconstruction de quadriques dégénérées n'est également pas évoquée.

Dans l'approche de Shashua et Toelg (1997), la reconstruction de quadrique est réalisée par une approche hybride mélangeant des correspondances de points d'intérêts et de contours occultant de la quadrique. Cependant, cette approche est difficilement généralisable au-delà de l'algorithme proposé qui permet de reconstruire une quadrique à partir d'une seule vue et 4 correspondances de points.

Par la suite, Kahl et Heyden (1997) ont également remarqué la dégénérescence de la reconstruction de quadrique en utilisant uniquement deux points de vue, mais ont émis qu'une solution était possible en fournissant une preuve algébrique.

Un détail très complet sur les méthodes de reconstruction de quadrique ont été réalisés par Cross et Zisserman (1998); Cross (2000) en considérant la reconstruction dans l'espace dual ce qui permet de donner une solution plus générale et géométriquement complète. Un soin particulier a été porté sur l'ambiguïté de chaque reconstruction (à partir de conique dégénérée ou non) et a été exprimé géométriquement. Plus particulièrement, le concept de contrainte de dégénérescence qui permet de gérer le cas difficile de reconstruction de quadrique dégénérée a été méticuleusement traité géométriquement. La contribution principale de Cross et Zisserman (1998) repose sur une reconstruction générale de quadrique à partir de 2 vues (problème sous contraint), 3 vues (résolution parfaite) et 3+ vues (sur-contraint). Cette reconstruction se fait par l'utilisation des contours occultant et de correspondances de points. Cette reconstruction utilise également en prétraitement une contrainte épipolaire de coniques.

Reyes et Bayro-Corrochano (2005) ont insisté sur l'aspect contrainte épipolaire en mettant en évidence que le critère épipolaire de Cross et Zisserman (1998) requiert des données d'entrée (estimation de conique) très précise. Ils proposent une contrainte épipolaire plus robuste en utilisant les points de contours utilisés pour estimer les coniques en entrée et une reconstruction de quadrique optimisée par ajustement de faisceaux.

B.2.2 Autres approches

Récemment Crocco et al. (2016) ont présenté une méthode permettant à la fois d'estimer la calibration et de reconstruire une quadrique (majoritairement un ellipsoïde) englobant un objet d'intérêt à partir de coniques (majoritairement ellipses) détectées à partir d'une boite englobante d'un objet. L'originalité de la méthode réside dans le traitement de données d'entrées fortement bruitées (ellipses à partir de boite englobante potentiellement ambiguë et paramètres intrinsèques initialement inconnus). Les quadriques reconstruites sont utiles pour calculer la surface d'occupation 3D d'un objet et de permettre d'imposer une contrainte forte sur les objets et peut être amélioré en ajoutant des correspondances de points ou une segmentation plus fine de l'objet pour une meilleure précision de la pose

3D de l'objet.

B.3 Solutions proposées

B.3.1 Contrainte épipolaire

Reconstruire une quadrique à partir de réflexions spéculaires n'est pas une tâche facile. En effet, contrairement à l'approche de Cross et Zisserman (1998), notre reconstruction est réalisée à partir d'observations de réflexions spéculaires dans plusieurs images au lieu d'observations d'un objet directement.

Dans l'approche de Cross et Zisserman (1998), le calcul d'ellipses est moins propice aux erreurs. En effet, les spécularités représentent des images semi-miroir des sources de lumière réfléchie sur des surfaces spéculaires. Ces spécularités ont des contours qui ne sont pas consistants à cause de la variation d'intensité entre les images ce qui cause que les ellipses calculées ne sont pas épipolairement consistantes comme illustré dans la figure B.1.

Ce problème pourrait affecter la qualité de la reconstruction de quadrique à partir des spécularités. En conséquence, la consistance épipolaire devrait être vérifiée pour chaque vue utilisée pour l'initialisation afin d'obtenir un résultat optimal. Sans correspondance entre les contours d'ellipses, le système linéaire (4.12) peut retourner des résultats incohérents. L'espace dual permet une formulation élégante du problème, mais à cause de la nature algébrique des équations, un bruit même léger sur l'estimation des ellipses peut causer de grandes erreurs dans la quadrique reconstruite.

Une ellipse est définie de façon unique à partir de 5 points non-alignés et 5 droites non-colinéaires dans l'espace dual. Pour trois points de vue, chaque ellipse est contrainte par deux paires de droites provenant des autres points de vues ce qui rend le problème d'estimation de la conique duale sous-contrainte. Une des solutions proposées par Reyes et Bayro-Corrochano (2005) est d'estimer les lignes épipolaires de façon plus précises à partir des ellipses provenant des autres points de vue et une droite supplémentaire provenant de la tangente d'un point de contour de l'ellipse de l'image actuelle. En utilisant ces 5 lignes, nous pouvons calculer une conique duale unique. Ce procédé est répété pour chaque point de contour et chaque point de vue. L'ellipse corrigée est celle minimisant notre distance ellipse/Point discutée dans la section B.3.2.2. Nous pouvons noter que l'efficacité du calcul d'ellipse dépend fortement de la qualité des points de contour de la spécularité. Les ellipses obtenues pour chaque point de vues seront utilisées dans le processus de reconstruction, car elles respectent les contraintes épipolaires. Ce procédé est itératif, fiable et efficace. Nous illustrons une étape du procédé de correction à la figure B.2. Dans notre contexte, nous ne pouvons pas nous limiter à seulement trois points de vue. Ainsi, nous incluons un raffinement non-linéaire à partir d'au moins trois points de vue afin d'améliorer la reconstruction de quadrique.

B.3.2 Distance 2D ellipse/point

B.3.2.1 Distance de Jaccard

La distance de Jaccard calcule la similarité entre deux ensembles A et B en comparant la ratio entre entre l'intersection et l'union des deux ensembles. Cette distance est illustrée

à la figure B.3(b). En pratique, cette distance peut facilement être utilisée pour estimer une distance entre deux ellipses dans un processus de maximisation. Cette distance est définie dans l'équation (B.1).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \tag{B.1}$$

B.3.2.2 Paramétrisation d'ellipse et distance point/ellipse

Calculer une distance point/ellipse entre les contours d'une spécularité et son ellipse associée n'est pas optimal en pratique. En effet, calculer la distance entre un point (élément discret) et une ellipse (élément paramétrique) requiert de calculer la racine d'ordre 4 d'un polynôme ce qui est couteux en temps de calcul et ne permet pas une dérivée analytique comme évoquée dans Sturm et Gargallo (2007). Pour ces raisons, une distance point/point est préférable par souci de simplicité et d'optimisation. Pour discrétiser l'ellipse en différents points, une paramétrisation d'ellipse est nécessaire.

B.3.2.3 Paramétrisation d'ellipse pour l'association point/point

Afin de résoudre ce problème, Sturm et Gargallo (2007) proposent une paramétrisation de l'ellipse afin de l'échantillonner en fonction des points de contours les plus proches de l'ellipse. Cette paramétrisation permet ainsi d'obtenir une distance point/point plus rapide, plus cohérente et plus adaptée à notre contexte. Intuitivement, le point sur l'ellipse le plus proche d'un point de contour correspond à l'intersection de la droite passant par le centre de l'ellipse et le point de contour avec l'ellipse. Cependant, il est plus simple de calculer l'homographie $\mathtt{H}$ appliquée à notre ellipse $\mathtt{C}$ afin de transformer celle-ci en un cercle unitaire $\mathtt{C}_{unit}$. En effet, sans cette transformation, cet échantillonnage d'ellipse impliquerait le calcul d'une intersection entre une droite et une ellipse ce qui n'est pas trivial et coûteux en temps de calcul. Cependant, avec la transformation $\mathtt{H}$, le problème se résume à un calcul d'angle déduit par relation trigonométrique. Naïvement, nous pouvons représenter chaque point $\mathbf{q}'_j \in \mathtt{C}_{unit}$ par :

$$\mathbf{q}'_j = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{B.2}$$

Cependant, une façon plus simple de paramétrer une ellipse et de choisir le cercle unitaire $\mathtt{C}_{unit}$ comme support ce qui nous permet de paramétrer chaque $q'_j \in \mathtt{C}_{unit}$ par un angle α_j tel que :

$$\mathbf{q}'_j = \begin{pmatrix} \cos\alpha_j \\ \sin\alpha_j \\ 1 \end{pmatrix}, \tag{B.3}$$

ce qui implique de calculer un seul paramètre d'angle au lieu de deux paramètres de position.

À première vue, cette paramétrisation a l'inconvénient de comporter un nombre important de paramètres à estimer : $n+8$ (pour n points de contours de spécularités à faire correspondre à l'ellipse $\mathtt{C}$ et 8 paramètres pour $\mathtt{H}$) au lieu de 5 paramètres suffisants pour paramétrer une ellipse (position x_o, y_o, échelle s_x, s_y et angle α).

Jusqu'à maintenant, nous avons considéré $\mathtt{H} \in \mathbb{R}^{3\times3}$ tel que notre ellipse $\mathtt{C}$ est obtenue par la relation :

$$\mathtt{C} \sim \mathtt{H}^{-\top}\mathtt{C}_{unit}\mathtt{H}^{\top}, \tag{B.4}$$

avec $\mathtt{C}_{unit}$ le cercle unitaire, centrée à l'origine. Nous avons considéré $\mathtt{C}$ comme étant une générale homographie 2D ce qui est une sur-paramétrisation. Nous pouvons paramétrer $\mathtt{H}$ de façon minimale tels que :

$$\mathtt{C} \sim \mathtt{R}\Sigma = \mathtt{R}\ \mathrm{diag}(a, b, c), \tag{B.5}$$

avec $\mathtt{R}^{3\times3}$ une matrice orthogonale et a, b et c des scalaires. Ces paramètres sont obtenus par décomposition en valeurs singulières (SVD). Contrairement à cette approche de Sturm et Gargallo (2007), nous estimons les 5 paramètres de $\mathtt{C}$ utilisant l'algorithme suivant :

Algorithm 3

procedure TransformationEllipseEnParametrique($\mathtt{C}$)
  $\mathtt{K} \leftarrow \mathtt{C}(1:2, 1:2)$
  $\mathbf{t} \leftarrow \mathtt{C}(1:2, 3)$
  $e \leftarrow \mathtt{C}(3, 3)$
  $\mathbf{p} \leftarrow -\mathtt{K}^{-1}\mathbf{t}$
  $[\mathtt{V}, \mathtt{D}] \leftarrow eig(\frac{1}{(\mathbf{t}^{\top}\mathtt{K}^{-1}\mathbf{t}-e)\mathtt{K}})$
  if $\det(\mathtt{V}) < 0$ **then**
    $\mathtt{V}(:, 1) = -\mathtt{V}(:, 1)$
  $s_x \leftarrow \frac{1}{\sqrt{\mathtt{D}(1,1)}}$
  $s_y \leftarrow \frac{1}{\sqrt{\mathtt{D}(2,2)}}$
  $\alpha \leftarrow \tan(\frac{\mathtt{V}(2,1)}{\mathtt{V}(1,1)})^{-1}$
return $\begin{pmatrix}\mathbf{p} & s_x & s_y & \alpha\end{pmatrix}$

Ainsi $\mathtt{H}$ est obtenu en convertissant $\mathtt{C}$ en représentation naturelle explicitant les coordonnées du centre $\mathbf{p_o} = (x_o, y_o)^{\top}$, les scalaires s_x, s_y et l'angle de rotation de l'ellipse α décrite par l'algorithme 3. $\mathtt{H}$ se construit de la manière évoquée par Hartley et Zisserman (2004) :

$$\mathtt{H} = \begin{pmatrix} s_x cos(\alpha) & -s_y sin(\alpha) & x_o \\ s_x sin(\alpha) & s_y cos(\alpha) & y_o \\ 0 & 0 & 1 \end{pmatrix} \tag{B.6}$$

Cette paramétrisation a pour objectif d'associer les points de contours de la spécularité avec les points associés les plus proches de l'ellipse. Cette étape est réalisée une seule fois afin de fournir un état initial au processus de raffinement non-linéaire. Ces points seront utilisés comme paramètres dans le raffinement pour contraindre davantage la quadrique à raffiner. Ainsi, nous appliquerons également $\mathtt{H}$ pour les points de contours de la spécularité. Ce principe est illustré dans la figure B.4. Les différentes étapes de cet algorithme représentées dans la figure B.4 sont :

1. Calcul de l'homographie $\mathtt{H}$ transformant l'ellipse $\mathtt{C}$ en cercle unitaire $\mathtt{C}_{unit}$ et transformation des points de contours (en noir à gauche) par $\mathtt{H}$ en points de contours à droite autour du cercle unitaire $\mathtt{C}_{unit}$

2. Calcul de l'intersection entre les points de contours transformés et le cercle unitaire $\mathtt{C}_{unit}$. Nous isolons un point de contour (en bleu à droite) comme exemple
3. Transformation des points d'intersection par la transformation $\mathtt{H}^{-1}$. Le point résultant (en bleu à gauche) est cohérent par rapport au point de contour de spécularité initial

Limitations A l'opposé des approches de calcul d'ellipse à partir d'un ensemble de points de contours comme dans Sturm et Gargallo (2007), notre application présente quelques cas spécifiques qui mettent en échec notre paramétrisation d'ellipse $\mathtt{C}$ par rapport aux points de contours de la spécularité. Lors de la reconstruction initiale de notre quadrique, nous n'avons pas de garantie que les ellipses obtenues par projection soient considérablement décalées par rapport à leurs ellipses associées correspondant aux contours de spécularités. Ce phénomène s'explique si la quadrique reconstruite est considérablement décalée. Nous illustrons ce phénomène à la figure B.5(a). Il est également possible que les ellipses obtenues par projection aient une échelle considérablement différente de leurs ellipses associées. Cette problématique s'explique quand la position de la quadrique a bien été estimée, mais que le calcul de l'échelle de cette quadrique soit faux. Nous illustrons ce phénomène à la figure B.5(b).

Ces deux phénomènes ne permettent pas une estimation de la distance entre une ellipse et des points de contours cohérents, car pour une ellipse et des points de contours visuels éloignés et d'échelles différentes, la distance entre les deux entités est trop faible et non représentative.

Amélioration de la paramétrisation Pour notre application, nous recherchons une répartition des points sur $\mathtt{C}$ plus uniforme tout en conservant une correspondance entre un point de contour et un point appartement à l'ellipse la plus petite possible. Nous proposons une amélioration de la paramétrisation permettant un bon compromis entre ces deux objectifs.

Les différentes étapes de notre algorithme sont :

1. Nous partons de l'ellipse initiale $\mathtt{C}_1$ estimée par la projection de la quadrique reconstruite et de l'ellipse $\mathtt{C}_2$ estimée à partir des points de contours de la spécularité et utilisée par le processus de reconstruction de quadrique
2. Nous calculons l'homographie $\mathtt{H}_1$ transformant $\mathtt{C}_1$ en cercle unitaire $\mathtt{C}_{unit}$. Nous appliquons cette homographie à $\mathtt{C}_2$ afin d'obtenir une ellipse $\mathtt{C}'_2$ et aux points de contours de spécularité associés.
3. Par la suite, nous calculons l'homographie $\mathtt{H}_2$ transformant $\mathtt{C}'_2$ en cercle unitaire $\mathtt{C}_{unit}$ et l'appliquons au point de contours transformés par $\mathtt{H}_1$.
4. L'intersection entre les nouveaux points de contours obtenus et le cercle unitaire nous donne une discrétisation que nous appliquons à $\mathtt{C}_1$ en utilisant la transformation $\mathtt{H}_1^{-1}$

Notre processus de paramétrisation est illustré à la figure B.6(b).

En reprenant les exemples de la figure B.5, nous observons dans la figure B.7 les améliorations de notre nouvelle paramétrisation. Nous observons une meilleure répartition des points sur l'ellipse associée aux contours et une distance ellipse/point cohérente pour chaque point de contours.

B.3.2.4 Distance point/point d'ellipses

Notre distance correspond à la somme des distances euclidiennes pour chaque paire de points $\mathbf{q}'_{i,j}$ (points de contours de la spécularité) et $\mathbf{q}_{i,j} \in \mathtt{C}$ tel que :

$$\sum_{j=1}^{m} d(\mathbf{q}'_j, \mathbf{q}_j), \tag{B.7}$$

avec m le nombre de points de contours pour l'ellipse $\mathtt{C}$. Cette distance est illustrée à la figure B.3(a).

B.3.3 Raffinement non-linéaire

Le raffinement non-linéaire est crucial pour gérer des données d'entrée bruitées et pour corriger l'initialisation de quadrique sur un panel de points de vue plus large. Pour des données réelles, le processus d'initialisation n'est pas suffisant. Notre contexte diffère fortement de Cross et Zisserman (1998) due à l'aspect bruité des spécularités ce qui correspond à un problème de modélisation. En détail, un contour de spécularité est difficilement consistant d'une image à l'autre à cause d'une dépendance forte de la forme de la spécularité aux paramètres de caméra, des propriétés des matériaux et des conditions d'illumination. Afin d'améliorer notre initialisation de quadrique, nous utilisons un raffinement non-linéaire qui minimise notre distance Ellipse/Point évoquée à la section B.3.2.2 entre les projections de la quadrique pour chaque point de vue et les contours associés. Notre fonction de coût pour la minimisation est proche de celle de Reyes et Bayro-Corrochano (2005), mais nous n'utilisons pas un ajustement de faisceaux, car nous assumons que les poses de caméra sont suffisamment robustes. De plus, en dehors de la fonction de coût, Reyes et Bayro-Corrochano (2005) ne donne pas énormément de détail sur l'implémentation du raffinement non-linéaire et notamment sur le calcul de la distance Ellipse/Point.

B.3.3.1 Fonction de coût avec la distance de Jaccard

Cette distance a pour objectif de maximiser le recouvrement de la spécularité détectée et l'ellipse estimée par projection perspective de la quadrique sur différents points de vue. Une fonction de coût utilisant la distance de Jaccard est décrite dans l'équation (B.8) :

$$\max_{\mathtt{Q}} \quad \sum_{i=1}^{n} J(\mathtt{C}'_i, E_i), \tag{B.8}$$

avec E l'ensemble des pixels spéculaires dans l'image fournie par notre détecteur de spécularité pour le point de vue d'indice i.

B.3.3.2 Limitations

La distance de Jaccard fournit des résultats précis seulement quand les ensembles comparés sont sécants. Dans notre situation, il n'est pas rare d'avoir l'ellipse projetée de la quadrique et les contours de la spécularité séparés par une distance conséquente après une mauvaise initialisation de la quadrique. Dans ce cas, la distance de Jaccard est nulle et ne fournit aucune information localement ce qui pose un problème lors de l'utilisation

d'algorithmes de raffinement non-linéaire. De plus, le calcul de l'aire d'intersection entre la surface couverte par les points de contours et l'ellipse projetée est relativement coûteux en temps de calcul et relativement complexe comme évoqué dans les travaux de Hughes et Chraibi (2012).

B.3.3.3 Fonction de coût avec variables latentes

Afin de contraindre davantage le raffinement, nous utilisons des variables latentes sous forme de points 2D $\mathbf{q}'_{i,j}$ contraints à appartenir à l'ellipse $\mathtt{C}$ et comparés avec les contours de spécularité $\mathbf{q}_{i,j}$. Cette formulation nous permet de poser les contraintes suivantes durant la minimisation :

$$\mathbf{q}'^{\top}_{i,j}\mathtt{C}_j\mathbf{q}'_{i,j} = 0. \tag{B.9}$$

Les paramètres de notre minimisation sont composés de la représentation matricielle de la quadrique $\mathtt{Q} \in \mathbb{R}^{4\times4}$ accompagnée des différentes variables latentes appartenant à l'ellipse. Avant le processus de raffinement, ces variables latentes $\mathbf{q}'_{i,j}$ sont estimées par notre paramétrisation évoquée dans la section B.3.2.2 pour chaque ellipse $\mathtt{C}_j$ d'indice j. La fonction de coût avec les variables latentes est décrite par :

$$\begin{cases} \min\limits_{\mathtt{Q},\mathbf{q}'_{1,1},\dots,\mathbf{q}'_{n,m_n}} & \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m_i} d(\mathbf{q}'_{i,j}, \mathbf{q}_{i,j})^2, \\ \text{sous} & \mathbf{q}'^{\top}_{i,j}\mathtt{C}_j\mathbf{q}'_{i,j} = 0 \end{cases} \tag{B.10}$$

Une illustration du raffinement d'une quadrique à partir de trois points de vue est illustrée à la figure B.8. Bien que cette minimisation soit fonctionnelle, la forme matricielle de $\mathtt{Q}$ est plus difficile à optimiser dans l'état que la forme paramétrique de la quadrique (9 degrés de liberté : position 3D, échelle en 3 dimensions et 3 angles). Cela est dû au fait que les 10 éléments distincts de la matrice symétrique $\mathtt{Q}$ (9 en fixant la dernière case correspondant à l'échelle à 1) ne sont pas indépendants. Nous transformons la forme matricielle de $\mathtt{Q}$ en représentation paramétrique par l'algorithme suivant :

B.3.4 Critère d'images clés

Lorsque les points de vue de caméra s'étendent sur un ensemble d'images où la pose se déplace faiblement, le système (4.12) utilisé pour reconstruire la quadrique duale $\mathtt{Q}^*$ peut être mal conditionné. Ce contexte peut engendrer une reconstruction de quadrique dégénérée, à la mauvaise échelle (un degré de liberté complètement disproportionné) ou l'obtention d'une mauvaise quadrique comme un hyperboloïde. Pour la phase initiale et le raffinement non-linéaire, il est important de choisir un ensemble d'images pertinentes appelé images clés (notion similaire à Mouragnon et al. (2006)). Une approche naïve serait de sélection toutes les images de la séquence ce qui engendrerait une accumulation d'erreur importante avec une augmentation du temps de calcul de l'application. En pratique, nous limitons la reconstruction initiale de quadrique par 6 images clés.

Une image est considérée comme image clé lorsqu'elle répond à deux critères : estimation correcte de l'ellipse correspondante due à une bonne détection de la spécularité et bonne contrainte épipolaire.

Algorithm 4 Conversion de la forme matricielle de $\mathtt{Q}$ en forme paramétrique correspondant à neuf degrés de liberté.

procedure TransformationQuadriqueEnParametrique($\mathtt{Q}$)
  $\mathtt{K} \leftarrow \mathtt{Q}(1:3, 1:3)$
  $\mathbf{t} \leftarrow \mathtt{Q}(1:3, 4)$
  $e \leftarrow \mathtt{Q}(4,4)$
  $\mathbf{p} \leftarrow -\mathtt{K}^{-1}\mathbf{t}$
  $[\mathtt{V}, \mathtt{D}] \leftarrow eig(\frac{1}{(\mathbf{t}^\top \mathtt{K}^{-1}\mathbf{t} - e)\mathtt{K}})$
  $\mathtt{V} = \mathtt{V}^\top$
  if $\det(\mathtt{V}) < 0$ **then**
    $\mathtt{V} = -\mathtt{V}$
  $\mathtt{V}(1,:) = -\mathtt{V}(1,:)$
  $\mathtt{V}(2,:) = -\mathtt{V}(2,:)$
  $s_x \leftarrow \frac{1}{\sqrt{\mathtt{D}(1,1)}}$
  $s_y \leftarrow \frac{1}{\sqrt{\mathtt{D}(2,2)}}$
  $s_z \leftarrow \frac{1}{\sqrt{\mathtt{D}(3,3)}}$
  $\alpha \leftarrow -\tan(\frac{\mathtt{V}(3,2)}{\mathtt{V}(3,3)})^{-1}$

  $\beta \leftarrow -\tan(-\frac{\mathtt{V}(3,1)}{\sqrt{\mathtt{V}(3,2)^2 + \mathtt{V}(3,3)^2}})^{-1}$

  $\gamma \leftarrow -\tan(\frac{\mathtt{V}(2,1)}{\mathtt{V}(1,1)})^{-1}$
return $\begin{pmatrix} \mathbf{p} & s_x & s_y & s_z & \alpha & \beta & \gamma \end{pmatrix}$

Qualité de détection et correspondance d'ellipse de la spécularité A l'aide de la distance de B.3.2.1, nous considérons l'image courante comme image clé si l'ellipse estimée recouvre une aire proche de l'aire de la spécularité détectée. En pratique, 90% de recouvrement est suffisant pour obtenir une bonne correspondance.

Répartition des tangentes épipolaire. Afin de s'assurer que notre contrainte épipolaire est efficace, il est nécessaire d'avoir les tangentes épipolaires pour chaque ellipse réparties autour de celle-ci. Notre critère est de maximiser l'angle entre les droites contenant l'épipole et le centre de l'ellipse pour chaque épipole. Ce critère est appliqué pour chaque image clé et illustré dans la figure B.9. Nous devons avoir des angles (α, β et γ) les plus proches possible de 120° afin d'assurer une répartition optimale des points de vue permettant une bonne initialisation de la quadrique.

B.4 Distance 3D de quadrique

La distance entre deux quadriques peut être utilisée pour évaluer la qualité de reconstruction d'une quadrique par rapport à une vérité terrain ou pour un calcul de collision rapide entre deux quadriques.

B.4.1 Distance de centre de quadrique

Les quadriques sont des structures paramétriques qui peuvent être décrites en représentation naturelle comme détaillée dans 4 qui explicite le centre, l'échelle et la rotation de la quadrique. Pour cette distance, une distance euclidienne est utilisée entre les centres des deux quadriques.

B.4.2 Distance globale 3D

Dans le cas de deux quadriques très proches, nous voulons une distance 3D très précise. Une première approche serait d'utiliser une distance 3D similaire à la distance de Jaccard. Cependant, une telle distance impliquerait de calculer des intersections de quadriques ce qui peut s'avérer couteux en temps de calcul. De plus, cette distance serait également limitée comme la distance de Jaccard 2D car elle serait nulle lorsque les deux quadriques sont non-sécantes. Pour répondre à ces problèmes, nous proposons une distance point/point 3D entre deux quadriques inspirées de notre distance 2D point/point entre deux ellipses.

B.4.3 Discrétisation de quadrique

Afin de discrétiser la quadrique en k points répartis uniformément sur la surface, nous utilisons l'algorithme de la spirale de Fibonacci qui est décrite dans Hannay et Nye (2004). Prenons par exemple, la quadrique $\mathtt{Q}_1$ obtenue par un processus de reconstruction et $\mathtt{Q}_2$, la vérité terrain. Ces deux quadriques sont converties dans une représentation paramétrique (appelée aussi naturelle) décrite à l'algorithme 4 afin de calculer la matrice de transformation $\mathtt{M}_\mathtt{Q}$ telle que :

$$\mathtt{M}_\mathtt{Q}^{-T}\mathtt{Q}_{sphere}\mathtt{M}_\mathtt{Q}^{-1} = \mathtt{Q}, \tag{B.11}$$

avec $\mathtt{Q}_{sphere}$ la sphère unitaire centrée à l'origine. Ainsi, nous calculons les transformations $\mathtt{M}_{\mathtt{Q}_1}$ et $\mathtt{M}_{\mathtt{Q}_2}$ pour $\mathtt{Q}_1$ et $\mathtt{Q}_2$ respectivement. Il ne reste plus qu'à transformer $\mathtt{Q}_1$ en sphère unitaire en utilisant la transformation $\mathtt{M}_{\mathtt{Q}_1}^{-1}$ et en appliquant notre discrétisation sur la surface de $\mathtt{Q}_{sphere}$. Ensuite, chaque point est réparti sur la surface $\mathtt{Q}_1$ en utilisant $\mathtt{M}_{\mathtt{Q}_1}$ comme illustré dans la figure B.11(a).

B.4.4 Recherche de points voisins les plus proches de la deuxième ellipsoïde

La distance globale 3D correspond à la somme des distances euclidiennes pour chaque paire de points de l'ellipsoïde. Pour un point $\mathbf{p}_{\mathtt{Q}_1} \in \mathtt{Q}_1$, il existe un point $\mathbf{p}_{\mathtt{Q}_2} \in \mathtt{Q}_2$ formant une paire unique $(\mathbf{p}_{\mathtt{Q}_1}, \mathbf{p}_{\mathtt{Q}_2})$ tel que $\mathbf{p}_{\mathtt{Q}_2}$ est le point le plus proche de $\mathbf{p}_{\mathtt{Q}_1}$. Un moyen efficace de trouver le point le plus proche est d'appliquer la transformation $\mathtt{M}_{\mathtt{Q}_2}^{-1}$ aux points appartenant à la surface de $\mathtt{Q}_1$. Ensuite, nous calculons les intersections entre les droites passant par les points transformés et le centre de la sphère unitaire. Ces intersections sont transformées à nouveau avec $\mathtt{M}_{\mathtt{Q}_2}$. Le résultat de notre distance est illustré à la figure B.11(b). Nous distance globale 3D notée GD par l'équation (B.12) :

$$GD(\mathtt{Q}_1, \mathtt{Q}_2) = \int_{\partial\mathtt{Q}_1} \min_{q \in \partial\mathtt{Q}_2} d(\mathbf{p}, \mathbf{q}) \mathrm{d}\mathbf{p}, \tag{B.12}$$

avec $\mathbf{p} \in \mathtt{Q}_1$, $\mathbf{q} \in \mathtt{Q}_2$ et $\mathbf{p} \neq \mathbf{q}$.

Le calcul des points sur la seconde ellipsoïde ne garantit pas une répartition uniforme des points sur la surface. Afin d'assurer la symétrie de notre distance 3D (notée QD), nous la définissons par :

$$QD(\mathtt{Q}_1, \mathtt{Q}_2) = \max(GD(\mathtt{Q}_1, \mathtt{Q}_2), GD(\mathtt{Q}_2, \mathtt{Q}_1)), \tag{B.13}$$

B.5 Vectorisation du système de projection de la quadrique duale

Dans cette section, nous détaillons le processus de vectorisation détaillée dans le chapitre 4. Nous définissons deux opérateurs de vectorisation évoqués dans Henderson et Searle (1979), vec pour toute matrice et vech pour toute matrice symétriques matrices tel que la fore matricielle d'une l'ellipse (et plus généralement d'une conique) $\mathtt{C}$, et d'un ellipsoïde (plus générale une quadrique) $\mathtt{Q}$:

$$\mathtt{X} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}, \ \mathrm{vec}(\mathtt{X}) = \begin{bmatrix} a & b & b & d \end{bmatrix}^\top, \ \mathrm{vech}(\mathtt{X}) = \begin{bmatrix} a & b & d \end{bmatrix}^\top.$$

Nous appliquons vech à l'équation :

$$\mathrm{vech}(\mathtt{\Pi Q}^*\mathtt{\Pi}^\top) = \mathrm{vech}(\mathtt{C}^*)$$

qui peut être réécrit :

$$\mathtt{H}\,\mathrm{vec}(\mathtt{\Pi Q}^*\mathtt{\Pi}^\top) = \mathbf{C}_v^*, \text{ with } \mathtt{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

avec $\mathtt{H} \in \mathbb{R}^{6\times 9}$ tel que $\mathrm{vech}(\mathtt{X}) = \mathtt{H}\,\mathrm{vec}(\mathtt{X})$. De Henderson et Searle (1979), avec trois matrices $\mathtt{A}$, $\mathtt{B}$ et $\mathtt{C}$, $\mathrm{vec}(\mathtt{ABC}) = (\mathtt{C}^\top \otimes \mathtt{A})\,\mathrm{vec}(\mathtt{B})$ avec $\otimes$ le produit de Kronecker, nous développons :

$$\mathtt{H}(\mathtt{\Pi} \otimes \mathtt{\Pi})\,\mathrm{vec}(\mathtt{Q}^*) = \mathbf{C}_v^*.$$

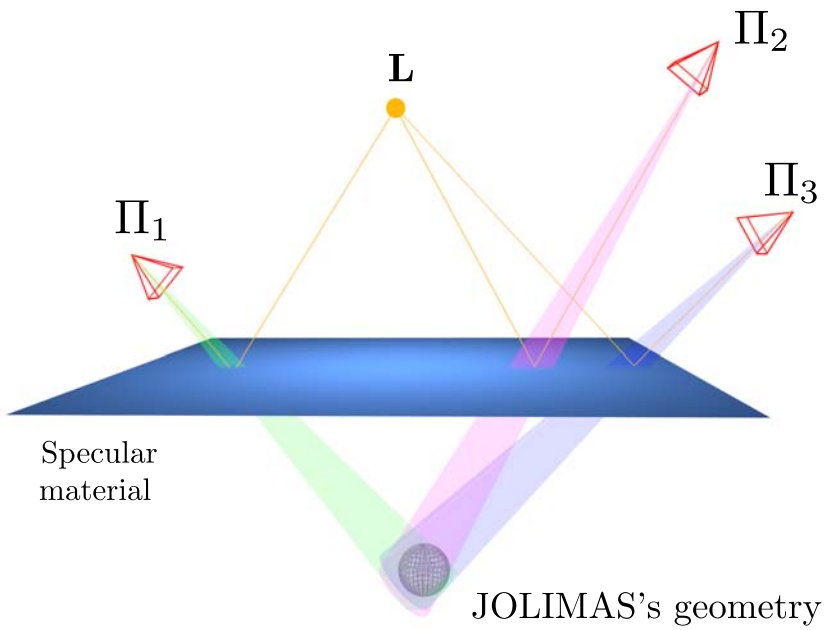
vec et vech sont aussi liés de telle façon que $\mathrm{vec}(\mathtt{X}) = \mathtt{G}\,\mathrm{vec}(\mathtt{X})$ avec $\mathtt{G} \in \mathbb{R}^{16\times 10}$. Nous obtenons la linéarisation suivante :

$$\mathtt{H}(\mathtt{\Pi} \otimes \mathtt{\Pi})\mathtt{G}\,\mathrm{vech}(\mathtt{Q}^*) = \mathbf{C}_v^* \tag{B.14}$$

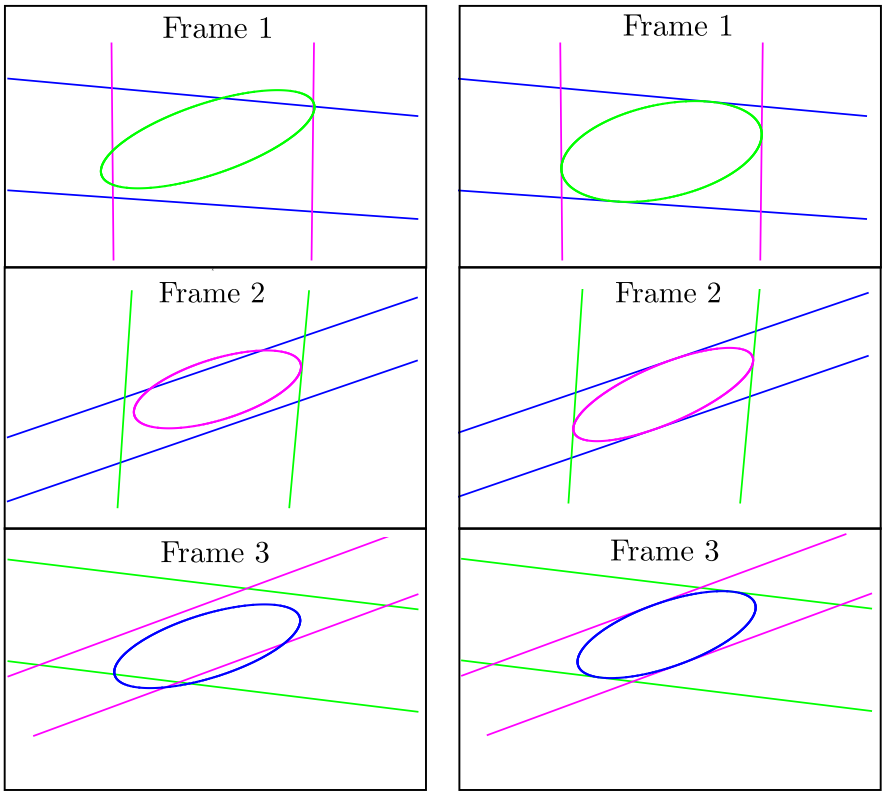
avec

$$\mathtt{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^{\top},$$

ce qui donne la forme finale $\mathtt{B}\mathbf{Q}_v^* = \mathbf{C}_v^*$ avec $\mathtt{B} = \mathtt{H}(\Pi \otimes \Pi)\mathtt{G}$.



(a)



(b) (c)

FIGURE B.1 – Lignes épipolaires des ellipses pour 3 poses de caméra Π_1, Π_2 et Π_3 et une source de lumière **L**. Cette scène est illustrée en (a) et les ellipses associées et les lignes épipolaires le sont dans (b). Dans la forme naïve, la géométrie épipolaire des ellipses n'est pas respectée. Ceci résulte en une reconstruction de quadrique incorrecte et donc une prédiction de spécularité incorrecte. La correction épipolaire corrige ce problème en corrigeant les ellipses afin de répondre aux contraintes épipolaires sans changer drastiquement la forme des ellipses.

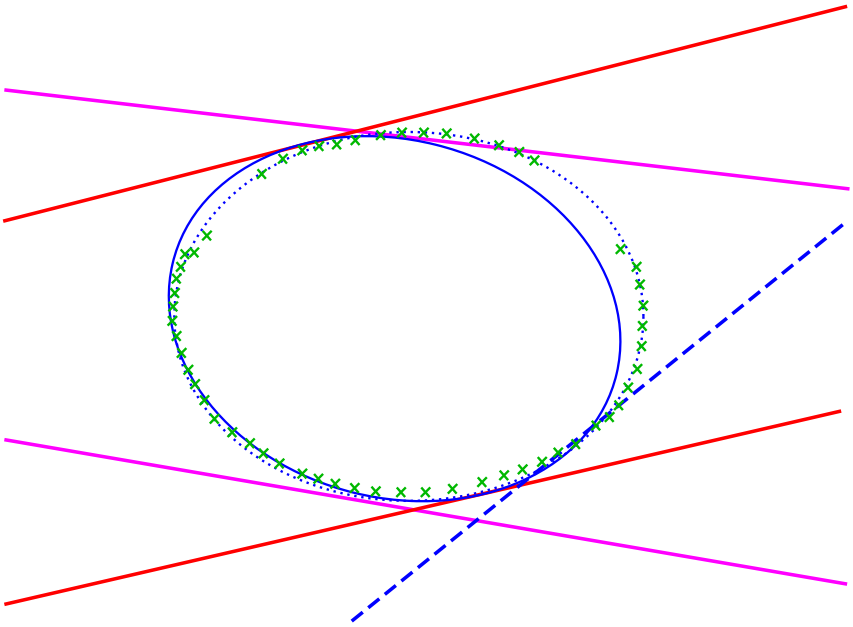
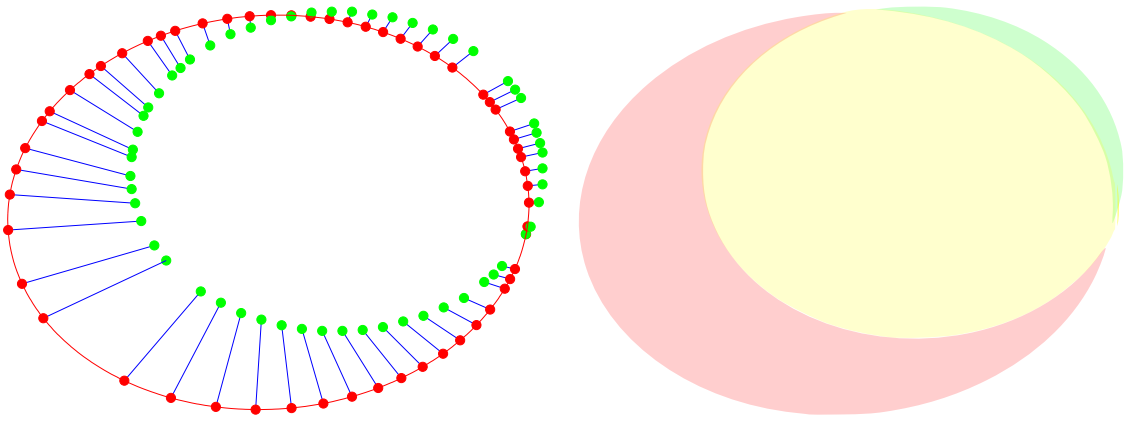


FIGURE B.2 – Correction épipolaire de Reyes et Bayro-Corrochano (2005). L'ellipse observée (points bleus) est calculée à partir des points de contour (croix noires) de la spécularité. Les quatre droites épipolaires créées à partir des autres points de vue sont représentées en rouge et magenta. Afin d'estimer la conique duale respectant la géométrie épipolaire, une cinquième droite calculée à partir de la tangente d'un point de contours (droite en pointillée en bleu). Nous calculons une ellipse dans l'espace dual pour chaque point de contour et nous choisissons celle (représentée ici) minimisant la distance ellipse/point évoquée à section B.3.2.2 pour chacun des trois points de vue. Les ellipses obtenues seront utilisées dans le processus de reconstruction d'une quadrique précise à partir d'ellipses épipolairement consistantes.



(a) (b)

FIGURE B.3 – Dans (a) le distance point/point est représenté comme étant la somme des distances Euclidiennes entre les contours de spécularités (en vert) et l'ellipse discrétisée par notre méthode (points rouges). Dans (b), nous illustrons la distance de Jaccard en utilisant l'intersection (en jaune) de l'aire des deux ellipses (en rouge et vert). L'union correspond à la somme de ces 3 couleurs.

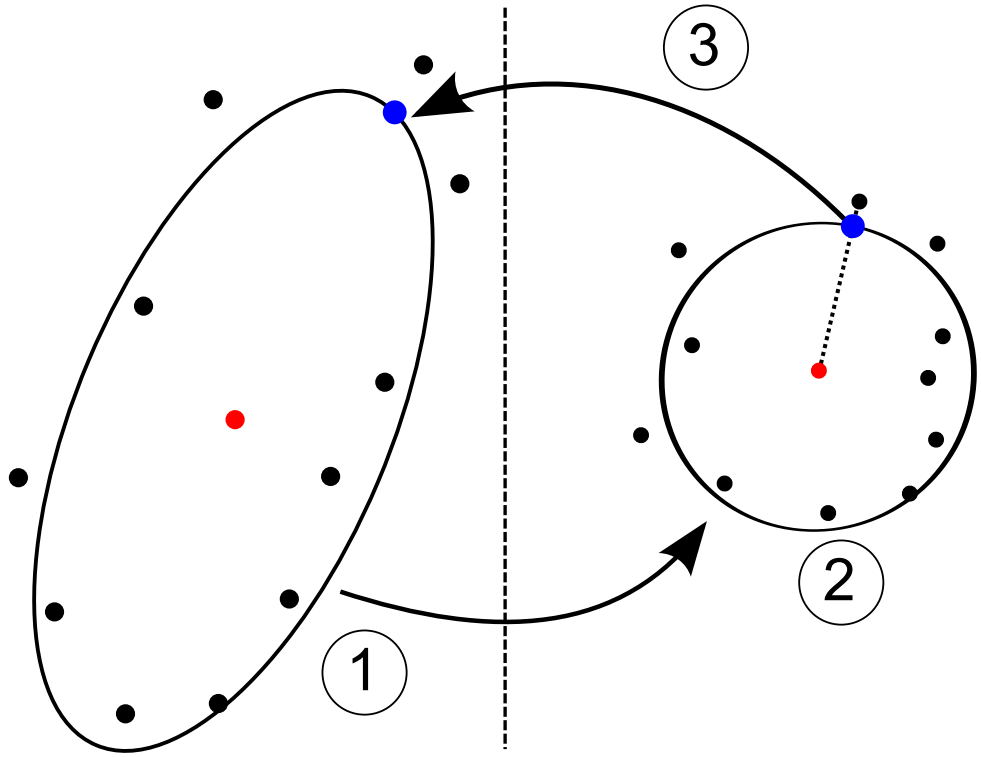
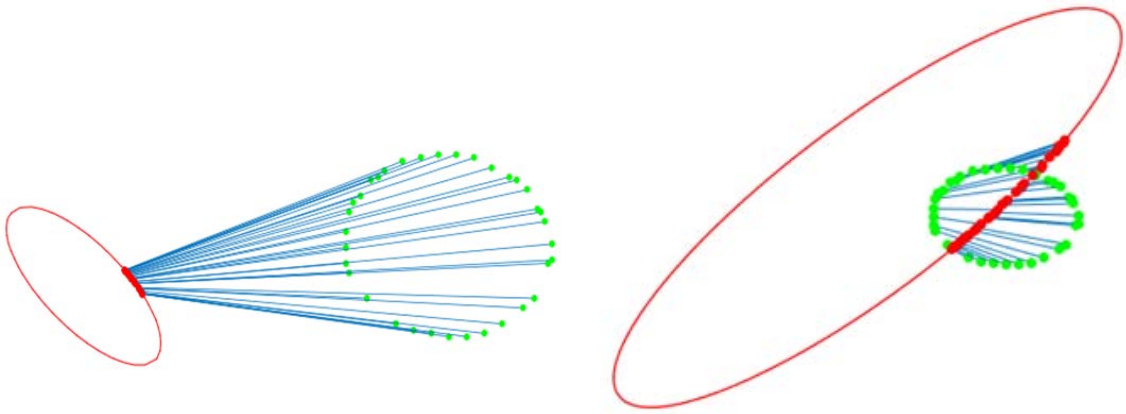


FIGURE B.4 – Paramétrisation d'ellipse. À partir de l'état initial (gauche) avec les contours de la spécularité (points noirs), nous associons à chaque point de contours un point sur l'ellipse. Ces points de l'ellipse sont calculés en transformant l'ellipse (en cercle unitaire (droite)) et les points de contours par l'homographie $\mathtt{H}^{-1}$ (étape (1)). Dans l'étape (2), nous calculons les intersections des lignes passant par les points de contours et le centre du cercle unitaire avec ce même cercle. dans l'étape (3), les intersections (points bleus) sont transformées à nouveau sur l'ellipse avec l'homographie $\mathtt{H}$.



(a) (b)

FIGURE B.5 – Limitation de la paramétrisation classique d'ellipse à partir de points de contours. Dans (a), nous illustrons le problème de la paramétrisation pour des contours éloignés considérablement de l'ellipse associée. Notre paramétrisation recherche les points de l'ellipse les plus proches des points de contours ce qui provoque une mauvaise répartition des points sur l'ellipse. Dans (b), nous illustrons le problème d'une paramétrisation entre des points contours répartis sur une surface beaucoup plus importante que l'ellipse associée. Ces deux problèmes ont pour impact d'avoir une distance ellipse/point trop faible et incohérent.

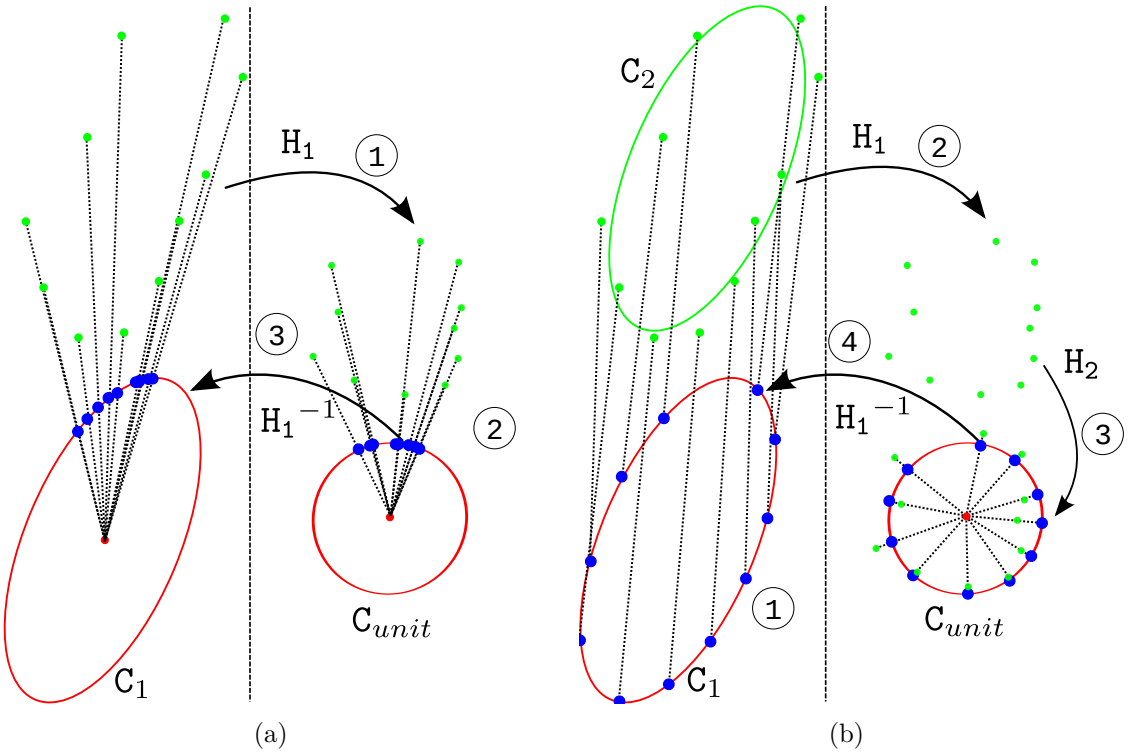
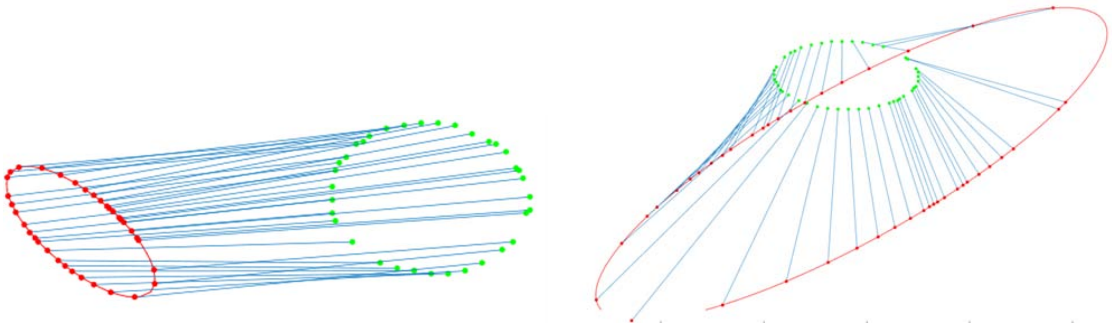
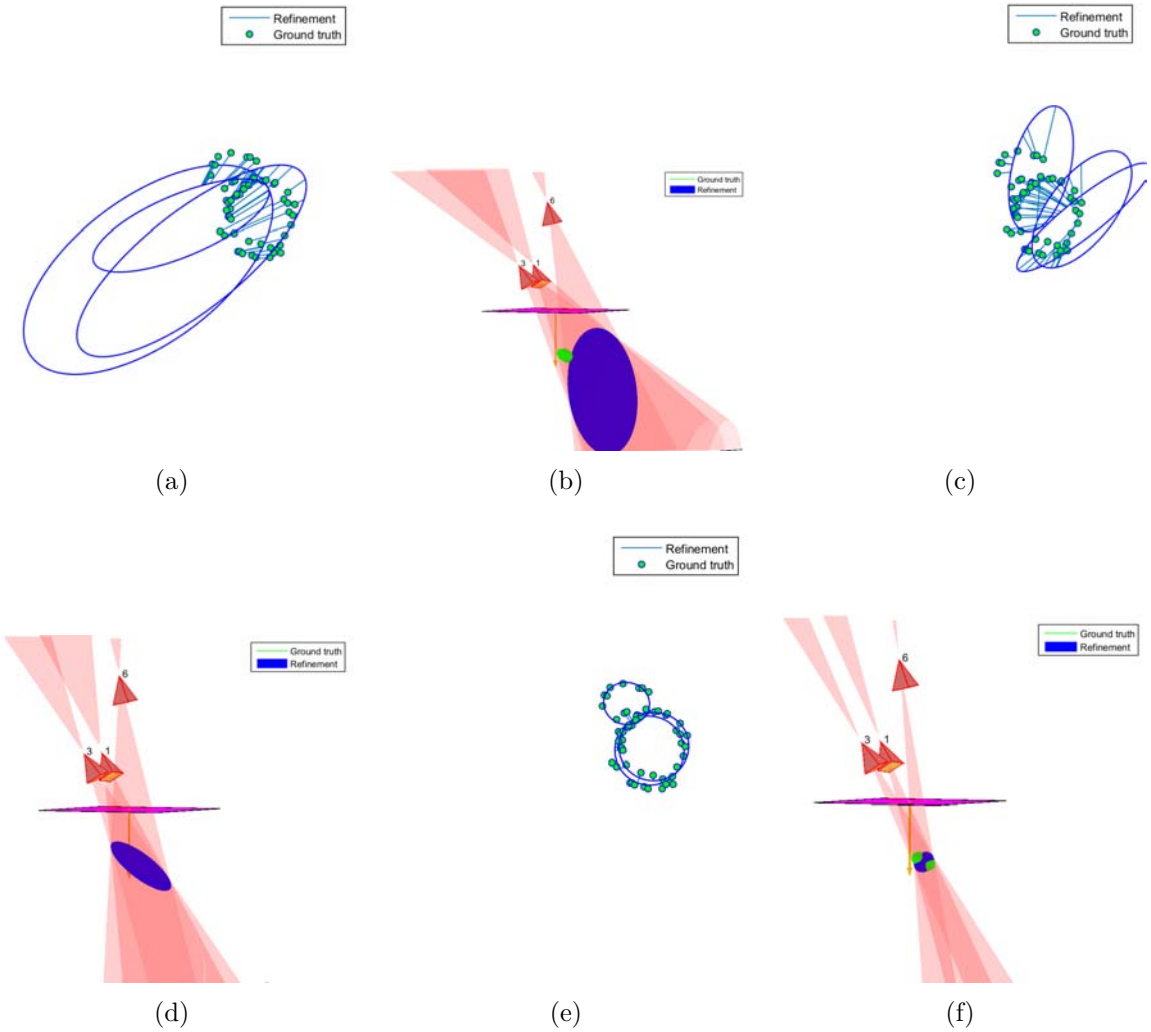


FIGURE B.6 – Amélioration des limitations de la paramétrisation d'ellipse classique. Dans (a), nous illustrons le problème de notre méthode précédente de paramétrisation pour des contours fortement éloignés de l'ellipse associée en appliquant les mêmes étapes que dans la figure B.4. Dans (b), nous présentons le même exemple avec notre nouvelle paramétrisation qui fournit un meilleur compromis entre répartitions des points sur l'ellipse paramétrée et distance entre points de contours et l'ellipse minimale. Nous partons de la partie gauche de la figure avec l'ellipse $\mathtt{C}_1$ estimée par la projection de l'ellipsoïde. Les points de contours (vert) correspondent à nos données d'entrées avec l'ellipse $\mathtt{C}_2$ (en verte) associée aux points. Nous estimons l'homographie $\mathtt{H}_1$ qui transforme $\mathtt{C}_1$ en cercle unitaire et centré (étape 1) et l'appliquons à $\mathtt{C}_1$, $\mathtt{C}_2$ et aux points de contours (étape 2). Pour recentrer les contours, nous estimons $\mathtt{H}_2$ l'homographie qui transforme $\mathtt{C}_2$ en cercle unitaire centré et l'appliquons aux points de contours transformés par $\mathtt{H}_1$ (étape 3). L'intersection entre le cercle unitaire et ces points (bleus à droite) transformés est calculée. Nous utilisons $\mathtt{H}_1^{-1}$ pour transformer ces points d'intersection sur $\mathtt{C}_1$ (étape 4).



(a) (b)

FIGURE B.7 – Amélioration des limitations de la paramétrisation d'ellipse classique. Nous observons une meilleure répartition des points sur l'ellipse associée aux contours et une distance ellipse/point cohérente pour chaque point de contours.



(a) (b) (c)

(d) (e) (f)

FIGURE B.8 – Illustration de l'évolution du raffinement non-linéaire de quadrique (à droite) à partir des contours de spécularité en vert. Nous pouvons observer l'évolution des ellipses correspondant aux projections de la quadrique en fonction des poses de caméra.

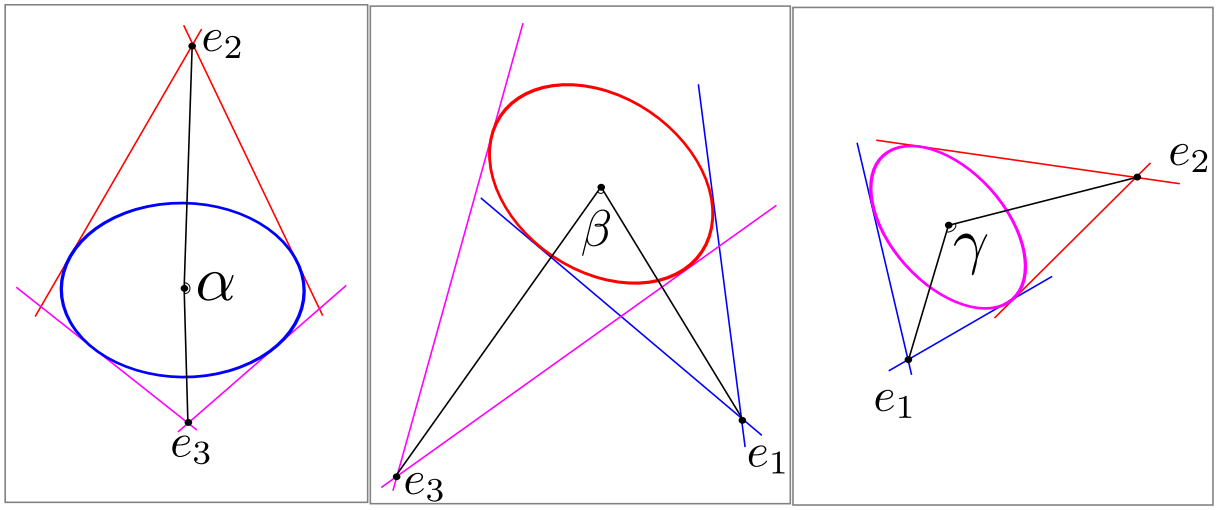


FIGURE B.9 – Critère utilisant la répartition des tangentes épipolaires afin de sélectionner les images clés. L'objectif est de maximiser l'angle entre chaque droite contenant le centre de l'ellipse et chaque épipole à partir de trois images.

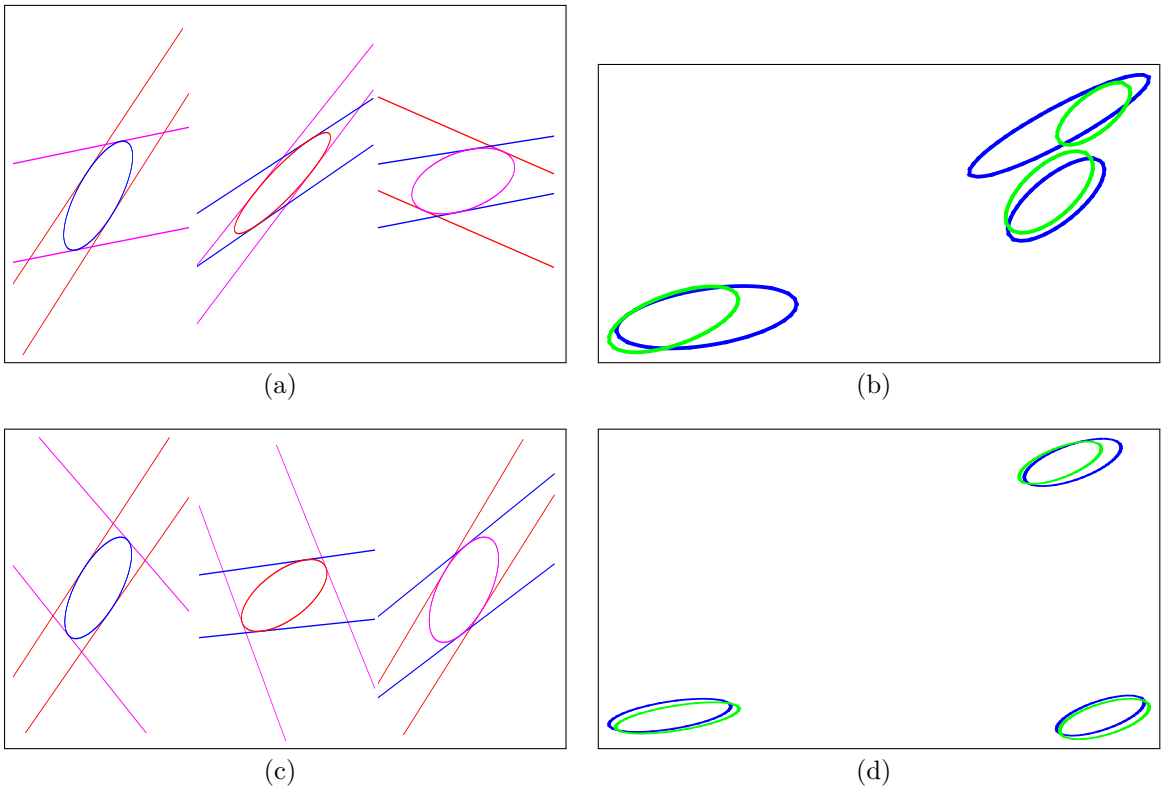
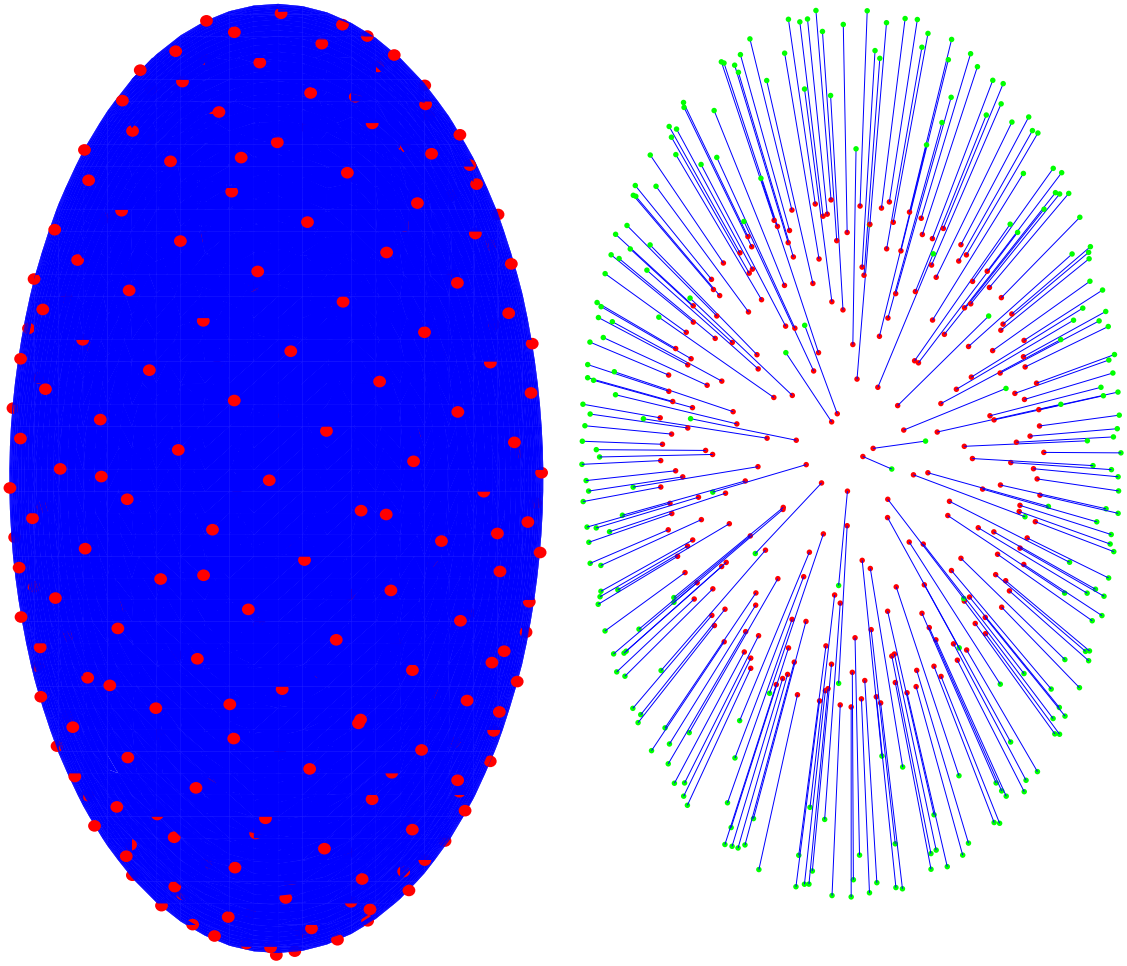


FIGURE B.10 – Comparaison de la qualité de reconstruction de quadrique pour des images choisies manuellement (première ligne) et par notre sélection d'images clés (deuxième ligne). Dans (a, c), nous illustrons les trois ellipses (bleu, rouge et magenta) utilisées pour la reconstruction ainsi que les contraintes épipolaires associées à chacune des ellipses provenant des deux autres points de vue. Nous pouvons observer que dans le cas des images clés, les contraintes épipolaires sont beaucoup plus réparties par rapport aux images sélectionnées manuellement ce qui permet d'améliorer le processus de reconstruction de la quadrique. Les ellipses issues de la projection de la quadrique (en bleu) pour les points de vue issus des images sélectionnées manuellement (b) et des images clé (d) sont comparées avec leurs vérités terrain. Dans le cas des images clé, nous pouvons observer que les ellipses projetées correspondent de manière presque identique à la vérité terrain.



(a) (b)

FIGURE B.11 – (a) détails de la discrétisation d'ellipsoïde. Chaque point est distribué uniformément. (b) illustre notre représentation de la distance globale. Les points de l'ellipsoïde 1 et 2 sont représentés en rouge et vert respectivement. Les paires sont représentées par des droites bleues.

ANNEXE C

Étude analytique des modèles de Phong et Blinn-Phong

Dans ce chapitre, nous donnons des explications complémentaires sur l'analyse des isocontours pour le modèle de Phong lorsque $\tau = 0$ et une analyse des isocontours de Blinn-Phong pour $\tau = 0$. Nous prouvons aussi l'unicité du point d'intensité maximale dans le modèle de Phong.

C.1 Modèles étudiés

C.1.1 Modèle de Blinn-Phong

La fonction d'intensité du modèle de Blinn-Phong d'un point 3D d'une surface $\mathbf{P}$ est donnée par :

$$I(\mathbf{P}) = i_a k_a + i_d k_d(\hat{\mathbf{L}}(\mathbf{P}) \cdot \hat{\mathbf{N}}) + i_s k_s(\hat{\mathbf{N}} \cdot \hat{\mathbf{H}}(\mathbf{P}))^n, \tag{C.1}$$

avec $\hat{\mathbf{H}}$ le vecteur de mi-chemin :

$$\hat{\mathbf{H}} = \frac{\hat{\mathbf{L}}(\mathbf{P}) + \hat{\mathbf{V}}(\mathbf{P})}{||\hat{\mathbf{L}}(\mathbf{P}) + \hat{\mathbf{V}}(\mathbf{P})||}$$

Le terme spéculaire est défini par :

$$I_s(\mathbf{P}) \propto \hat{\mathbf{N}}(\mathbf{P})^\top \frac{\hat{\mathbf{L}}(\mathbf{P}) + \hat{\mathbf{V}}(\mathbf{P})}{||\hat{\mathbf{L}}(\mathbf{P}) + \hat{\mathbf{V}}(\mathbf{P})||}, \tag{C.2}$$

Nous développons l'équation (C.2) par :

$$I_s(\mathbf{P}) \propto \hat{\mathbf{N}}(\mathbf{P})^\top \frac{\mu(\mathbf{L} - \mathbf{P}) + \mu(\mathbf{V} - \mathbf{P})}{||\mu(\mathbf{L} - \mathbf{P}) + \mu(\mathbf{V} - \mathbf{P})||}, \tag{C.3}$$

D'une façon semblable au modèle de Phong, nous voulons analyser les isocontours d'une spécularité à partir de l'équation (C.3).

C.1.2 Analyse des contours externes de la spécularité pour $\tau = 0$ pour le modèle de Blinn-Phong

Nous résolvons l'équation pour $\tau = 0$:

$$\hat{\mathbf{N}}(\mathbf{P})^\top \frac{\mu(\mathbf{L}-\mathbf{P}) + \mu(\mathbf{V}-\mathbf{P})}{||\mu(\mathbf{L}-\mathbf{P}) + \mu(\mathbf{V}-\mathbf{P})||} = 0.$$

En développant et en rassemblant les monômes de même degré, nous avons :

$$\begin{aligned}
&(d^\circ 4)\ P_Z||\mathbf{P}||^2 \\
&(d^\circ 3)\ -2(L_Z P_Z||\mathbf{P}||^2 - V_Z P_Z||\mathbf{P}||^2 - \mathbf{P}^\top\mathbf{L}P_Z^2 + \mathbf{P}^\top\mathbf{V}P_Z^2) \\
&(d^\circ 2)\ 4(L_Z P_Z\mathbf{P}^\top\mathbf{V} - V_Z P_Z\mathbf{P}^\top\mathbf{L}) + \\
&\qquad P_Z^2(||\mathbf{V}||^2 - ||\mathbf{L}||^2) + (L_Z^2 - V_Z^2)||\mathbf{P}||^2 \\
&(d^\circ 1)\ 2(V_Z P_Z||\mathbf{L}||^2 - L_Z P_Z||\mathbf{V}||^2 + V_Z^2\mathbf{P}^\top\mathbf{L} - L_Z^2\mathbf{P}^\top\mathbf{V}) \\
&(d^\circ 0)\ L_Z^2||\mathbf{V}||^2 - V_Z^2||\mathbf{L}||^2.
\end{aligned}$$

Nous notons que, $\mathbf{P} = \begin{bmatrix} P_X & P_Y & P_Z \end{bmatrix}^\top$. Avec $P_Z = 0$, les degrés quatre et trois disparaissent et :

$$\begin{aligned}
&(d^\circ 2)\ (L_Z^2 - V_Z^2)||\mathbf{P}||^2 \\
&(d^\circ 1)\ 2(V_Z^2\mathbf{L}^\top - L_Z^2\mathbf{V}^\top)\mathbf{P} \\
&(d^\circ 0)\ L_Z^2||\mathbf{V}||^2 - V_Z^2||\mathbf{L}||^2.
\end{aligned}$$

Les monômes peuvent être factorisés sous la forme suivante :

$$\tilde{\mathbf{P}}^\top \mathtt{J}\tilde{\mathbf{P}} = 0, \tag{C.4}$$

où $\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{P} & 1 \end{bmatrix}^\top$ sont les coordonnées homogènes de $\mathbf{P}$. La matrice $\mathtt{J} \in \mathbb{R}^{4\times 4}$ est symétrique et définie par :

$$\mathtt{J} = \begin{bmatrix} L_Z^2 - V_Z^2 & (V_Z^2\mathbf{L} - L_Z^2\mathbf{V})^\top \\ V_Z^2\mathbf{L} - L_Z^2\mathbf{V} & L_Z^2||\mathbf{V}||^2 - V_Z^2||\mathbf{L}||^2. \end{bmatrix} \tag{C.5}$$

Nous définissons la projection orthographique de S tel que :

$$\mathtt{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Le contour externe de la spécularité est ainsi donné par l'ellipse :

$$\mathtt{C} = \mathtt{A}^\top\mathtt{J}\mathtt{A} = \begin{bmatrix} L_Z^2 - V_Z^2 & (V_Z^2\bar{\mathbf{L}} - L_Z^2\bar{\mathbf{V}})^\top \\ V_Z^2\bar{\mathbf{L}} - L_Z^2\bar{\mathbf{V}} & L_Z^2||\mathbf{V}||^2 - V_Z^2||\mathbf{L}||^2 \end{bmatrix},$$

avec $\bar{\mathbf{V}} = \begin{bmatrix} V_X & V_Y \end{bmatrix}^\top$ et $\bar{\mathbf{L}} = \begin{bmatrix} L_X & L_Y \end{bmatrix}^\top$ sont les projections orthographiques de $\mathbf{V}$ et $\mathbf{L}$ sur S respectivement.

C.1.3 Cas spécial, analyse de $\tau = 1$ pour le modèle de Phong

L'étude du cas $\tau = 1$ inclut les points ayant une intensité maximale sur la surface S. Pour le modèle de Phong, cette étude est intéressante, car pour $\tau = 1$, les termes de degré trois et quatre disparaissent. Les termes restants sont :

$$\begin{aligned}
&(d^\circ 2)\ \mathbf{P}^\top [\mathbf{R} - \mathbf{V}]_\times^2 \mathbf{P} \\
&(d^\circ 1)\ 2\mathbf{P}^\top [\mathbf{V} \times \mathbf{R}]_\times (\mathbf{V} - \mathbf{R}) \\
&(d^\circ 0)\ (\mathbf{R}^\top \mathbf{V})^2 - \tau^2 ||\mathbf{R}||^2 ||\mathbf{V}||^2.
\end{aligned}$$

Ces termes peuvent être factorisés sous la forme suivante :

$$\tilde{\mathbf{P}}^\top \mathtt{J} \tilde{\mathbf{P}} = 0, \tag{C.6}$$

où $\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{P} & 1 \end{bmatrix}^\top$ sont les coordonnées homogènes de $\mathbf{P}$. La matrice $\mathtt{J} \in \mathbb{R}^{4\times 4}$ est symétrique et définie par :

$$\mathtt{J} = \begin{bmatrix} [\mathbf{R} - \mathbf{V}]_\times^2 & [\mathbf{V} \times \mathbf{R}]_\times (\mathbf{V} - \mathbf{R}) \\ ([\mathbf{V} \times \mathbf{R}]_\times (\mathbf{V} - \mathbf{R}))^\top & (\mathbf{R}^\top \mathbf{V})^2 - ||\mathbf{R}||^2 ||\mathbf{V}||^2 \end{bmatrix}$$

Elle représente ainsi une quadrique et $\hat{\mathbf{P}}$ se situe à l'intersection de cette quadrique avec S. Le bloc principal (3×3) de $\mathtt{J}$ est $[\mathbf{R} - \mathbf{V}]_\times^2$ et est de rang 2. Ainsi, $\mathrm{rank}(\mathtt{J}) \geq 2$. Nous montrons ci-dessous que $\mathbf{Q}\tilde{\mathbf{R}} = \mathbf{Q}\tilde{\mathbf{V}} = 0$. Ce qui signifie que $\mathrm{rank}(\mathtt{J}) \leq 2$ et que $\mathrm{rank}(\mathtt{J}) = 2$. Ceci signifie également que $\mathtt{J}$ est semi-définie, non-positive ou non-négative. La façon dont le polynôme est construit, à partir de la fraction $I_s = \frac{a}{b} = \tau = 1$, avec $a \leq b$ ce qui implique que $a - b \leq 0$ et ainsi que $\mathtt{J}$ est une quadrique point représentant la ligne contenant $\mathbf{R}$ et $\mathbf{V}$. L'intersection de cette ligne avec S nous donne la solution attendue pour $\hat{\mathbf{P}}$. En montrant que $\tilde{\mathbf{R}} \in \mathtt{J}^\perp$ nous avons :

$$[\mathbf{R} - \mathbf{V}]_\times^2 \mathbf{R} + [\mathbf{V} \times \mathbf{R}]_\times (\mathbf{V} - \mathbf{R}).$$

Le premier terme est développé en :

$$\begin{aligned}
&[\mathbf{R} - \mathbf{V}]_\times [\mathbf{R} - \mathbf{V}]_\times \mathbf{R} \\
&= [\mathbf{R} - \mathbf{V}]_\times [\mathbf{V}]_\times \mathbf{R} \\
&= [\mathbf{V}]_\times^2 \mathbf{R} - [\mathbf{R}]_\times [\mathbf{V}]_\times \mathbf{R} \\
&= \mathbf{V} \times (\mathbf{V} \times \mathbf{R}) - \mathbf{R} \times (\mathbf{V} \times \mathbf{R}).
\end{aligned}$$

Le second terme est développé en :

$$(\mathbf{V} \times \mathbf{R}) \times \mathbf{R} - (\mathbf{V} \times \mathbf{R}) \times \mathbf{R} = -\mathbf{V} \times (\mathbf{V} \times \mathbf{R}) + \mathbf{R} \times (\mathbf{V} \times \mathbf{R}),$$

ce qui donne un zéro. Le dernier élément est :

$$(\mathbf{R} - \mathbf{V})^\top [\mathbf{V} \times \mathbf{R}]_\times \mathbf{R} + (\mathbf{R}^\top \mathbf{V})^2 - ||\mathbf{R}||^2 ||\mathbf{V}||^2.$$

Le premier terme est développé en :

$$(\mathbf{R}-\mathbf{V})^\top(\mathbf{R}\times(\mathbf{R}\times\mathbf{V}))=-\mathbf{V}^\top(\mathbf{R}\times(\mathbf{R}\times\mathbf{V}))=-\mathbf{V}^\top[\mathbf{R}]_\times^2\mathbf{V}.$$

Nous avons vu que le deuxième et le troisième terme (les coefficients associés au degré 0) sont aussi égaux à $\mathbf{V}^\top[\mathbf{R}]_\times^2\mathbf{V}$, ce qui conclue.

C.2 Point d'intensité maximale dans une spécularité sous le modèle de Phong

Nous souhaitons trouver le point d'intensité maximale d'une spécularité créé par une source de lumière sur une surface plane dont l'intensité est donnée par le modèle de Phong. Sans perte de généralité, nous travaillons dans le repère monde pour la surface plane de la scène $S \subset \mathbb{R}^3$ soit le plan $(\mathbf{XY})$. En détail, $S = \{\mathbf{P} \mid P_Z = 0\}$. Nous pouvons ainsi paramétrer S par un point $\mathbf{p} \in \mathbb{R}^2$ et définir $\mathbf{P} = \text{stk}(\mathbf{p}, 0)$. Au point $\mathbf{p}$ la composante spéculaire $I_s(\mathbf{p})$ est donnée par :

$$I_s(\mathbf{p}) = \max(0, \cos(\beta(\mathbf{p})))^n, \tag{C.7}$$

où β correspond à l'angle entre la direction d'observation et la direction de réflexion. Cette dernière est obtenue par réflexion de la direction de la lumière incidente en fonction de la normale de la surface. Le paramètre $n > 0$ caractérise la réflectance de la surface. Nous pouvons formellement décrire le problème de la recherche du point d'intensité maximale $\hat{\mathbf{p}} \in S$ en :

$$\hat{\mathbf{p}} = \arg\max_{\mathbf{p}\in\mathbb{R}^2} I_s(\mathbf{p}). \tag{C.8}$$

Nous voulons étudier l'existence, la localisation et l'unicité de $\hat{\mathbf{p}}$. À proprement parler, nous réalisons une étude du point d'intensité maximale de terme spéculaire de Phong. En pratique, ce terme domine le termes diffus et ambiant, et $\hat{\mathbf{p}}$ sera ainsi une bonne approximation du point d'intensité maximale dans les cas incluant la lumière ambiante et la réflexion diffuse.

C.3 Notation et contexte

Nous définissons à présent les vecteurs de direction plus précisément. L'ensemble de ces vecteurs sont unitaires et nous les définissons en utilisant la fonction de normalisation $\mu(\mathbf{U}) \stackrel{\text{def}}{=} \frac{\mathbf{U}}{\|\mathbf{U}\|}$. Le point de vue est à $\mathbf{V} \in \mathbb{R}^3$ et la direction d'observation est $\mu(\mathbf{V}-\mathbf{P})$. La source de lumière est à $\mathbf{L} \in \mathbb{R}^3$ et la direction de la lumière est décrite par $\mu(\mathbf{L}-\mathbf{P})$. Le point de vue et la source de lumière sont localisés sur la même partie de S et nous définissons ce côté comme étant la 'partie positif' tel que $V_Z > 0$ et $L_Z > 0$. La direction de la réflexion parfaite est définie par $-\mu(\mathbf{R}-\mathbf{P})$, où $\mathbf{R} \in \mathbb{R}^3$ est un point 'virtuel' localisé sur la partie négative de S, avec $\mathbf{R} \stackrel{\text{def}}{=} \text{stk}(L_X, L_Y, -L_Z)$. Avec cette notation, nous avons :

$$\cos(\beta(\mathbf{p})) = -\mu(\mathbf{R}-\mathbf{P})^\top\mu(\mathbf{V}-\mathbf{P}). \tag{C.9}$$

C.4 Résultat principal

Théorème 1. Pour un point de vue et une source de lumière donnés aux positions $\mathbf{V}, \mathbf{L} \in \mathbb{R}^3$ respectivement, localisés du même côté du plan S coïncidant avec le plan $\mathbf{XY}$ du plan de référence dans le repère monde, la spécularité créée par le terme spéculaire de Phong a un point d'intensité maximale localisée à $\frac{1}{V_Z+L_Z}\begin{bmatrix} V_Z L_X + V_X L_Z & V_Z L_Y + V_Y L_Z & 0 \end{bmatrix}^\top$ avec une intensité de réflexion de un.

La preuve du théorème 1 requiert la proposition suivante.

Proposition 1. Pour deux points $\mathbf{R}, \mathbf{V} \in \mathbb{R}^3$ et un plan S avec $\mathbf{R}, \mathbf{V} \notin S$, l'ensemble des points $\mathcal{P} \subset S$ tels que le vecteur reliant $\mathbf{R}$ à la projection de la ligne contenant $\mathbf{P}$ et $\mathbf{V}$ est l'opposée du vecteur reliant $\mathbf{V}$ à la projection de la droite contenant $\mathbf{P}$ et $\mathbf{R}$ quand projetée sur S est un point unique $\mathcal{P} = \{\mathbf{P}\}$ situé à l'intersection de la droite contenant $\mathbf{R}$ et $\mathbf{V}$ et S. Deux cas spéciaux se produisent : si $\mathbf{R} = \mathbf{V}$ alors $\mathcal{P} = \{\mathbf{P} \mid \mathbf{P} \in S\}$ et si la droite contenant $\mathbf{R}$ et $\mathbf{V}$ est parallèle à S alors $\mathcal{P} = \emptyset$.

Preuve du théorème 1. Nous reformulons le modèle d'illumination tel que :

$$\tilde{I}_s(\mathbf{p}) = \cos(\beta(\mathbf{p}))^n = \Big(- \mu(\mathbf{R} - \mathbf{P})^\top \mu(\mathbf{V} - \mathbf{P})\Big)^n.$$

L'extremum de $\tilde{I}_s$ est donné en résolvant :

$$\frac{\partial \tilde{I}_s}{\partial \mathbf{p}} = 0.$$

Nous avons $\frac{\partial \mu}{\partial \mathbf{U}}(\mathbf{U}) = \frac{\|\mathbf{U}\|^2 \mathtt{I} - \mathbf{U}\mathbf{U}^\top}{\|\mathbf{U}\|^3}$, nous réécrivons l'équation tel que :

$$\begin{aligned} -n\Big(- \mu(\mathbf{R} - \mathbf{P})^\top \mu(\mathbf{V} - \mathbf{P})\Big)^{n-1} \\ \mathtt{K}\left(\frac{\|\mathbf{R} - \mathbf{P}\|^2 \mathtt{I} - (\mathbf{R} - \mathbf{P})(\mathbf{R} - \mathbf{P})^\top}{\|\mathbf{R} - \mathbf{P}\|^3} \mu(\mathbf{V} - \mathbf{P}) \right. \\ \left. + \frac{\|\mathbf{V} - \mathbf{P}\|^2 \mathtt{I} - (\mathbf{V} - \mathbf{P})(\mathbf{V} - \mathbf{P})^\top}{\|\mathbf{V} - \mathbf{P}\|^3} \mu(\mathbf{R} - \mathbf{P}) \right) = 0, \end{aligned}$$

avec $\mathtt{K} \overset{\text{def}}{=} \frac{\partial \mathbf{p}}{\partial p} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Nous observons que $\mathtt{K}\mathbf{Q}$ se projette simplement en un point $\mathbf{Q} \in \mathbb{R}^3$ sur S. Puisque $n > 0$, nous pouvons supprimer le premier terme. Le deuxième terme disparait pour $\mathbf{R} - \mathbf{P}$ et $\mathbf{V} - \mathbf{P}$ orthogonaux. Ceci peut être facilement vérifié, en analysant la fonction de coût, que les extremums sont des minimums puisqu'ils annulent $\tilde{I}_s$ et I_s. De plus, le deuxième terme peut être supprimé de l'équation. Nous pouvons

multiplier par $\|\mathbf{R}-\mathbf{P}\|\|\mathbf{V}-\mathbf{P}\|$ pour obtenir :

$$\mathtt{K}\left(\left(\mathtt{I}-\frac{(\mathbf{R}-\mathbf{P})(\mathbf{R}-\mathbf{P})^\top}{\|\mathbf{R}-\mathbf{P}\|^2}\right)(\mathbf{V}-\mathbf{P}) + \left(\mathtt{I}-\frac{(\mathbf{V}-\mathbf{P})(\mathbf{V}-\mathbf{P})^\top}{\|\mathbf{V}-\mathbf{P}\|^2}\right)(\mathbf{R}-\mathbf{P})\right) = 0.$$

Cette équation peut être reformulée en une forme plus compacte en utilisant des projecteurs tels que :

$$\Pi_S\left(\overrightarrow{\mathbf{PV}} - \Pi_{\overrightarrow{\mathbf{PR}}}(\overrightarrow{\mathbf{PV}}) + \overrightarrow{\mathbf{PR}} - \Pi_{\overrightarrow{\mathbf{PV}}}(\overrightarrow{\mathbf{PR}})\right) = 0, \tag{C.10}$$

et nous obtenons :

$$\Pi_S\left(\overrightarrow{\mathbf{PV}}\right) + \Pi_S\left(\overrightarrow{\mathbf{PR}}\right) = \Pi_S\left(\Pi_{\overrightarrow{\mathbf{PR}}}(\overrightarrow{\mathbf{PV}})\right) + \Pi_S\left(\Pi_{\overrightarrow{\mathbf{PV}}}(\overrightarrow{\mathbf{PR}})\right). \tag{C.11}$$

Nous pouvons enfin appliquer la proposition 1. Notons que $\mathbf{V}$ est situé sur la partie positive de S et $\mathbf{R}$ est un point 'virtuel' qui est situé sur la partie négative de S. Puisque aucun des vecteurs $\mathbf{V}$ et $\mathbf{R}$ ne sont contenus dans S, nous avons nécessairement $\mathbf{R} \neq \mathbf{V}$ et la droite contenant $\mathbf{R}$ et $\mathbf{V}$ qui ne sont jamais parallèles à S. Aucun des cas spéciaux de la proposition 1 ne s'applique, et nous concluons que $\mathbf{P}$ est définie uniquement comme étant l'intersection de S et la droite contenant $\mathbf{R}$ et $\mathbf{V}$. Un raisonnement géométrique simple montre que $\mathbf{P}$ est le point où $\mathbf{N}$ coïncide avec la bissectrice de l'angle formé par $\mathbf{R}$, $\mathbf{P}$ et $\mathbf{V}$. Nous trouvons que $\mathbf{P} = \mathbf{R} + \lambda(\mathbf{V}-\mathbf{R})$ en utilisant $P_Z = 0$ afin de trouver $\lambda = \frac{R_Z}{R_Z - V_Z}$, et nous obtenons finalement :

$$\mathbf{P} = \begin{bmatrix}\mathbf{p}\\ 0\end{bmatrix} \text{ with } \mathbf{p} = \frac{1}{V_Z + L_Z}\begin{bmatrix}V_Z L_X + V_X L_Z\\ V_Z L_Y + V_Y L_Z\end{bmatrix}. \tag{C.12}$$

La valeur du coût $\tilde{I}_s(\mathbf{p})$ est donnée après quelques manipulations algébriques par :

$$\tilde{I}_s(\mathbf{p}) = -\mu(\mathtt{M}\mathbf{L}-\mathbf{V})^\top \mu(\mathtt{M}\mathbf{V}-\mathbf{V}) = 1.$$

Puisque c'est une valeur positive, cela signifie que le point $\mathbf{P}$ est un maximum de I_s. □

La preuve de la proposition 1 requiert le lemme suivant.

Lemma 1. *Pour* $\mathbf{A}, \mathbf{B} \in \mathbb{R}^3$*, nous avons* $(\|\mathbf{A}\|^2\mathtt{I} - \mathbf{A}\mathbf{A}^\top)\mathbf{B} = \mathbf{A}\times\mathbf{B}\times\mathbf{A}$.

Preuve du lemme 1. En développant le terme de gauche nous pouvons l'écrire tel que $\mathbf{A}^\top\mathbf{A}\mathbf{B} - \mathbf{A}\mathbf{A}^\top\mathbf{B}$. En changeant l'ordre des produits scalaire et vectoriel pour chaque terme, nous obtenons $\mathbf{B}\mathbf{A}^\top\mathbf{A} - \mathbf{A}^\top\mathbf{B}\mathbf{A}$ que nous factorisons en $\mathbf{B}\mathbf{A}^\top - \mathbf{A}^\top\mathbf{B})\mathbf{A}$. Le premier facteur peut être écrit en utilisant la matrice antisymétrique et nous obtenons $[\mathbf{A}\times\mathbf{B}]_\times\mathbf{A}$ ce qui correspond au terme à droite. □

Preuve de la proposition 1. Cette équation peut être réécrite tel que :

$$\mathtt{K}\Bigg(\Bigg(\mathtt{I}-\frac{(\mathbf{R}-\mathbf{P})(\mathbf{R}-\mathbf{P})^\top}{\|\mathbf{R}-\mathbf{P}\|^2}\Bigg)(\mathbf{V}-\mathbf{P}) + \Bigg(\mathtt{I}-\frac{(\mathbf{V}-\mathbf{P})(\mathbf{V}-\mathbf{P})^\top}{\|\mathbf{V}-\mathbf{P}\|^2}\Bigg)(\mathbf{R}-\mathbf{P})\Bigg) = 0.$$

Nous multiplions par $\|\mathbf{R}-\mathbf{P}\|^2\|\mathbf{V}-\mathbf{P}\|^2$ et obtenons :

$$\mathtt{K}\Bigg(\|\mathbf{V}-\mathbf{P}\|^2\Big(\|\mathbf{R}-\mathbf{P}\|^2\mathtt{I}-(\mathbf{R}-\mathbf{P})(\mathbf{R}-\mathbf{P})^\top\Big)(\mathbf{V}-\mathbf{P}) + \|\mathbf{R}-\mathbf{P}\|^2\Big(\|\mathbf{V}-\mathbf{P}\|^2\mathtt{I}-(\mathbf{V}-\mathbf{P})(\mathbf{V}-\mathbf{P})^\top\Big)(\mathbf{R}-\mathbf{P})\Bigg) = 0.$$

En utilisant le lemme 1 deux fois sur le premier terme avec $\mathbf{A}=\mathbf{R}-\mathbf{P}$ et $\mathbf{B}=\mathbf{V}-\mathbf{P}$ et sur le second terme avec $\mathbf{A}=\mathbf{V}-\mathbf{P}$ et $\mathbf{B}=\mathbf{R}-\mathbf{P}$, nous obtenons :

$$\mathtt{K}\Bigg(\|\mathbf{V}-\mathbf{P}\|^2[(\mathbf{R}-\mathbf{P})\times(\mathbf{V}-\mathbf{P})]_\times(\mathbf{R}-\mathbf{P}) + \|\mathbf{R}-\mathbf{P}\|^2[(\mathbf{V}-\mathbf{P})\times(\mathbf{R}-\mathbf{P})]_\times(\mathbf{V}-\mathbf{P})\Bigg) = 0.$$

où $\mathbf{A},\mathbf{B}\in\mathbb{R}^3$, $[\mathbf{A}]_\times\in\mathbb{R}^{3\times3}$ une matrice antisymétrique telle que $[\mathbf{A}]_\times\mathbf{B}=\mathbf{A}\times\mathbf{B}$. Puisque $[\mathbf{A}\times\mathbf{B}]_\times=-[\mathbf{B}\times\mathbf{A}]_\times$ nous avons :

$$\mathtt{K}[(\mathbf{R}-\mathbf{P})\times(\mathbf{V}-\mathbf{P})]_\times(\|\mathbf{V}-\mathbf{P}\|^2(\mathbf{R}-\mathbf{P})-\|\mathbf{R}-\mathbf{P}\|^2(\mathbf{V}-\mathbf{P}))=0. \tag{C.13}$$

Nous obtenons un produit de trois facteurs , f_1, f_2 et f_3, deux matrices et un vecteur, multiplié à un vecteur. Puisque $f_1\neq0$: il ne reste plus que trois possibilités pour résoudre l'équation (i) $f_2=0$, (ii) $f_3=0$ et (iii) $f_3\in(f_1f_2)^\perp$, où $f_1^\perp$ représente l'espace nul de f_i. La condition (iii) est équivalente à $f_2f_3\in f_1^\perp$ et inclut $f_2\in f_1^\perp$ et $f_3\in f_2^\perp$. Puisque $\mathbf{R}\neq\mathbf{P}$ et $\mathbf{V}\neq\mathbf{P}$, (i) reste valable seulement si $\mathbf{R}-\mathbf{V}\propto\mathbf{V}-\mathbf{P}$. Cette condition signifie que les trois points $\mathbf{R}$, $\mathbf{V}$ et $\mathbf{P}$ doivent être alignés. Nous avons ainsi trois cas : (i) $-a$ $\mathbf{R}=\mathbf{V}$, ainsi (i) est vérifié pour n'importe quel $\mathbf{P}\in S$; (i) $-b$ $\mathbf{R}\neq\mathbf{V}$ et la droite contenant $\mathbf{R}$ et $\mathbf{V}$ est parallèle à S, (i) n'est jamais vérifié et (i) $-c$ $\mathbf{R}\neq\mathbf{V}$ et la droite contenant $\mathbf{R}$ et $\mathbf{V}$ n'est pas parallèle à S, ainsi (i) est vérifié pour $\mathbf{P}$ à l'intersection de S et la droite contenant $\mathbf{R}$ et $\mathbf{V}$ n'est pas parallèle à S. Ainsi (i) est vérifié pour $\mathbf{P}$ à l'intersection de S et la droite contenant $\mathbf{R}$ et $\mathbf{V}$. Puisque $\mathbf{R}\neq\mathbf{P}$ et $\mathbf{V}\neq\mathbf{P}$, (ii) est vérifié si et seulement si $\mathbf{R}-\mathbf{P}\propto\mathbf{V}-\mathbf{P}$ et $\|\mathbf{V}-\mathbf{P}\|=\|\mathbf{R}-\mathbf{P}\|$ ce qui signifie que si et seulement si $\mathbf{R}=\mathbf{V}$ et pour tout $\mathbf{P}\in S$. Dans (iii), nous considérons f_1f_2, qui est la matrice $\mathtt{K}$ supprimant la troisième ligne de la matrice antisymétrique dans f_2. Pour $\mathbf{C}\overset{\text{def}}{=}(\mathbf{R}-\mathbf{P})\times(\mathbf{V}-\mathbf{P})$, nous avons $\mathtt{K}[\mathbf{C}]_\times=\begin{bmatrix}0&-c_3&c_2\\c_3&0&-c_1\end{bmatrix}$. Nous avons ainsi trois cas dépendants de $\operatorname{rank}(\mathtt{K}[\mathbf{C}]_\times)$. Le cas (iii) $-a$ se produit lorsque $\operatorname{rank}(\mathtt{K}[\mathbf{C}]_\times)=2$, ce qui est vérifié si et seulement si

$c_3 \neq 0$. (iii) est vérifié si et seulement si :

$$(\mathbf{R}-\mathbf{P}) \times (\mathbf{V}-\mathbf{P}) \propto \|\mathbf{V}-\mathbf{P}\|^2(\mathbf{R}-\mathbf{P}) - \|\mathbf{R}-\mathbf{P}\|^2(\mathbf{V}-\mathbf{P}).$$

La partie de gauche définit un vecteur qui est orthogonal au plan défini par $\mathbf{R}$, $\mathbf{V}$ et l'origine du repère monde est localisé sur S alors que la partie de droite définit un vecteur qui est inclus dans ce plan. Ainsi, (iii) est vérifié si et seulement si les deux parties de l'équation disparaissent ce qui implique que (i) et (ii) sont vérifié en même temps ce qui se produit seulement pour $\mathbf{R} = \mathbf{V}$ et pour tout $\mathbf{P} \in S$ ce qui vérifie dans ce cas (ii). Le cas (iii) $-b$ se produit lorsque $\text{rank}(\mathtt{K}[\mathbf{C}]_\times) = 1$, ce qui est vérifié si et seulement si $c_3 = 0$ et $c_1 \neq 0$, $c_2 \neq 0$ ou $c_1 c_2 \neq 0$, $c_3 = 0$ ce qui signifie que les deux composantes de $\mathbf{R}-\mathbf{P}$ et $\mathbf{V}-\mathbf{P}$ sont linéairement dépendants et ainsi que le plan contenant $\mathbf{R}$, $\mathbf{V}$ et $\mathbf{P}$ est orthogonal à S et ainsi contient $\mathbf{N}$. (iii) est vérifié si :

$$\mathbf{N}^\top\Big(\|\mathbf{V}-\mathbf{P}\|^2(\mathbf{R}-\mathbf{P}) - \|\mathbf{R}-\mathbf{P}\|^2(\mathbf{V}-\mathbf{P})\Big) = 0. \tag{C.14}$$

Ce qui signifie que $\mathbf{N}$ est orthogonal au plan contenant $\mathbf{R}$, $\mathbf{V}$ et $\mathbf{P}$, (iii) ce qui nécessite $\|\mathbf{V}-\mathbf{P}\|^2(\mathbf{R}-\mathbf{P}) - \|\mathbf{R}-\mathbf{P}\|^2(\mathbf{V}-\mathbf{P}) = 0$, ce qui est (ii). Le cas (iii)-c est obtenu quand $\text{rank}(\mathtt{K}[\mathbf{C}]_\times) = 0$ ce qui est valable si et seulement si $\mathbf{C} = 0$ et suit ainsi (i). □

Bibliographie

[Arnold et al. 2010] Arnold, Mirko ; Ghosh, Anarta ; Ameling, Stefan ; Lacey, Gerard : Automatic segmentation and inpainting of specular highlights for endoscopic imaging. In : *Journal on Image and Video Processing* 2010 (2010), S. 9

[Besbes et al. 2012] Besbes, Bassem ; Collette, Sylvie N. ; Tamaazousti, Mohamed ; Bourgeois, Steve ; Gay-Bellile, Vincent : An interactive augmented reality system : a prototype for industrial maintenance training applications. In : *ISMAR*, 2012 (ISMAR)

[Blake 1990] Blake, Andrew : Does the brain know the physics of specular reflection ? In : *Nature* 343 (1990), Nr. 6254, S. 165–168

[Blake et Brelstaff 1988] Blake, Andrew ; Brelstaff, Gavin : Geometry from specularities. In : *International Conference on Computer Vision*, 1988 (ICCV)

[Blinn 1977] Blinn, James F. : Models of light reflection for computer synthesized pictures. In : *Special Interest Group on Computer Graphics and Interactive Techniques*, 1977 (SIGGRAPH)

[Boom et al. 2013] Boom, Bastiaan J. ; Orts-Escolano, Sergio ; Ning, Xin X. ; McDonagh, Steven ; Sandilands, Peter ; Fisher, Robert B. : Point Light Source Estimation based on Scenes Recorded by a RGB-D camera. In : *British Machine Vision Conference*, 2013 (BMVC)

[Bousefsaf et al. 2018] Bousefsaf, Frédéric ; Tamaazousti, Mohamed ; Said, Souheil H. ; Michel, Rémi : Image completion using multispectral imaging. In : *Image Processing* (2018)

[Brelstaff et Blake 1988] Brelstaff, Gavin ; Blake, Andrew : Detecting specular reflections using Lambertian constraints. In : *International Conference on Computer Vision*, 1988 (ICCV)

[Buteau et Saito 2015] Buteau, Paul-Emile ; Saito, Hideo : Retrieving Lights Positions Using Plane Segmentation with Diffuse Illumination Reinforced with Specular Component. In : *International Symposium on Mixed and Augmented Reality*, 2015 (ISMAR)

[Chang et al. 2009] Chang, Ju Y. ; Raskar, Ramesh ; Agrawal, Amit : 3D pose estimation and segmentation using specular cues. In : *Computer Vision and Pattern Recognition*, 2009 (CVPR)

[Chican et Tamaazousti 2014] CHICAN, Guillaume ; TAMAAZOUSTI, Mohamed : Constrained PatchMatch for Image Completion. In : *International Symposium on Visual Computing*, 2014 (ISVC)

[Cook et Torrance 1982] COOK, Robert L. ; TORRANCE, Kenneth E. : A reflectance model for computer graphics. In : *Transactions on Graphics* 1 (1982), Nr. 1, S. 7–24

[Crocco et al. 2016] CROCCO, Marco ; RUBINO, Cosimo ; DEL BUE, Alessio : Structure from Motion with Objects. In : *Computer Vision and Pattern Recognition*, 2016 (CVPR)

[Cross et Zisserman 1998] CROSS, Geoffrey ; ZISSERMAN, Andrew : Quadric reconstruction from dual-space geometry. In : *International Conference on Computer Vision*, 1998 (ICCV)

[Cross 2000] CROSS, Geoffrey M. : *Surface reconstruction from image sequences*, University of Oxford, Dissertation, 2000

[Csongei et al. 2014] CSONGEI, Michael ; HOANG, Liem ; SANDOR, Christian : Interactive near-field illumination for photorealistic augmented reality on mobile devices. In : *Virtual Reality*, 2014 (VR)

[Deng et al. 1999] DENG, Yining ; MANJUNATH, BS ; SHIN, Hyundoo : Color image segmentation. In : *Computer Vision and Pattern Recognition*, 1999 (CVPR)

[Einabadi et Grau 2015] EINABADI, Farshad ; GRAU, Oliver : Discrete light source estimation from light probes for photorealistic rendering. In : *British Machine Vision Conference*, 2015 (BMVC)

[Engel et al. 2016] ENGEL, Jakob ; KOLTUN, Vladlen ; CREMERS, Daniel : Direct sparse odometry. In : *arXiv preprint arXiv :1607.02565* (2016)

[Faugeras 1993] FAUGERAS, Olivier : *Three-dimensional computer vision : a geometric viewpoint*. MIT press, 1993

[Fitzgibbon et al. 1996] FITZGIBBON, Andrew W. ; FISHER, Robert B. et al. : A buyer's guide to conic fitting. In : *DAI Research paper* (1996)

[Fleming et al. 2004] FLEMING, Roland W. ; TORRALBA, Antonio ; ADELSON, Edward H. : Specular reflections and the perception of shape. In : *Journal of Vision* 4 (2004), Nr. 9, S. 10

[Gay-Bellile et al. 2012] GAY-BELLILE, G ; BOURGEOIS, Steve ; TAMAAZOUSTI, Mohamed ; NAUDET-COLLETTE, S ; KNODEL, Sebastian : A mobile markerless augmented reality system for the automotive field. In : *ISMAR Workshop*, 2012 (ISMARW)

[Gruber et al. 2014] GRUBER, Lukas ; LANGLOTZ, Tobias ; SEN, Pradeep ; HOHERER, Tobias ; SCHMALSTIEG, Dieter : Efficient and robust radiance transfer for probeless photorealistic augmented reality. In : *Virtual Reality*, 2014 (VR)

[Hachisuka et al. 2008] HACHISUKA, Toshiya ; OGAKI, Shinji ; JENSEN, Henrik W. : Progressive photon mapping. In : *ACM Transactions on Graphics (TOG)* Bd. 27, 2008, S. 130

[Haines et Akenine-Moller 2002] HAINES, Eric ; AKENINE-MOLLER, Tomas : Real-time rendering. (2002)

[Hannay et Nye 2004] HANNAY, JH ; NYE, JF : Fibonacci numerical integration on a sphere. In : *Journal of Physics A : Mathematical and General* 37 (2004), Nr. 48, S. 11591

[Haralick et al. 1994] Haralick, Bert M. ; Lee, Chung-Nan ; Ottenberg, Karsten ; Nölle, Michael : Review and analysis of solutions of the three point perspective pose estimation problem. In : *International Journal of Computer Vision* 13 (1994), Nr. 3, S. 331–356

[Hartley et Zisserman 2004] Hartley, R. I. ; Zisserman, A. : *Multiple View Geometry in Computer Vision*. Second. Cambridge University Press, ISBN : 0521540518, 2004

[Hartley 1997] Hartley, Richard I. : In defense of the eight-point algorithm. 19 (1997), Nr. 6, S. 580–593

[Hartley et Sturm 1997] Hartley, Richard I. ; Sturm, Peter : Triangulation. In : *Computer vision and image understanding* 68 (1997), Nr. 2, S. 146–157

[Healey et Binford 1988] Healey, Gleen ; Binford, Thomas O. : Local shape from specularity, 1988, S. 62–86

[Heitz et al. 2016] Heitz, Eric ; Dupuy, Jonathan ; Hill, Stephen ; Neubelt, David : Real-time polygonal-light shading with linearly transformed cosines. In : *Transactions on Graphics* 35 (2016), Nr. 4, S. 41

[Henderson et Searle 1979] Henderson, Harold V. ; Searle, SR : Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics. In : *Canadian Journal of Statistics* 7 (1979), Nr. 1, S. 65–81

[Herling et Broll 2014] Herling, Jan ; Broll, Wolfgang : High-quality real-time video inpaintingwith PixMix. In : *Transactions on Visualization and Computer Graphics* 20 (2014), Nr. 6, S. 866–879

[Hughes et Chraibi 2012] Hughes, Gary B. ; Chraibi, Mohcine : Calculating ellipse overlap areas. In : *Computing and Visualization in Science* 15 (2012), Nr. 5, S. 291–301

[Hughes et Foley 2014] Hughes, John F. ; Foley, James D. : *Computer graphics : principles and practice*. 2014

[Hyeongwoo Kim 2013] Hyeongwoo Kim, Sunil Hadap Inso K. : Specular Reflection Separation using Dark Channel Prior, 2013 (CVPR)

[Immel et al. 1986] Immel, David S. ; Cohen, Michael F. ; Greenberg, Donald P. : A radiosity method for non-diffuse environments. In : *Special Interest Group on Computer Graphics and Interactive Techniques*, 1986 (SIGGRAPH)

[Jachnik et al. 2012] Jachnik, Jan ; Newcombe, Richard A. ; Davison, Andrew J. : Real-time surface light-field capture for augmentation of planar specular. In : *International Symposium on Mixed and Augmented Reality*, 2012 (ISMAR)

[Jarosz et al. 2008] Jarosz, Wojciech ; Zwicker, Matthias ; Jensen, Henrik W. : The beam radiance estimate for volumetric photon mapping. In : *Special Interest Group on Computer Graphics and Interactive Techniques*, 2008 (SIGGRAPH)

[Kahl et Heyden 1997] Kahl, F ; Heyden, A : The quadric reference surface : Theory and applications. In : *International Conference on Computer Vision*, 1997 (ICCV)

[Kajiya 1986] Kajiya, James T. : The rendering equation. In : *Special Interest Group on Computer Graphics and Interactive Techniques*, 1986 (SIGGRAPH)

[Kanbara et Yokoya 2004] Kanbara, Masayuki ; Yokoya, Naokazu : Real-time Estimation of Light Source Environment for Photorealistic Augmented Reality. In : *International Conference on Pattern Recognition*, 2004 (ICPR)

[Karl et al. 1994] KARL, William C. ; VERGHESE, George C. ; WILLSKY, Alan S. : Reconstructing ellipsoids from projections. In : *Graphical Models and Image Processing* 56 (1994), Nr. 2, S. 124–139

[Karsch et al. 2011] KARSCH, Kevin ; HEDAU, Varsha ; FORSYTH, David ; HOIEM, Derek : Rendering synthetic objects into legacy photographs. In : *ACM Transactions on Graphics (TOG)* 30 (2011), Nr. 6, S. 157

[Kawai et al. 2015] KAWAI, Norihiko ; SATO, Tomokazu ; YOKOYA, Naokazu : Diminished Reality Based on Image Inpainting Considering Background Geometry. In : *Transactions on Visualization and Computer Graphics* (2015)

[Knorr et Kurz 2014] KNORR, Sebastian B. ; KURZ, Daniel : Real-time illumination estimation from faces for coherent rendering. In : *International Symposium on Mixed and Augmented Reality*, 2014 (ISMAR)

[Kobayashi et Nomizu 1963] KOBAYASHI, Shoshichi ; NOMIZU, Katsumi : *Foundations of differential geometry*. Interscience publishers New York, 1963

[Kuthirummal et Nayar 2007] KUTHIRUMMAL, Sujit ; NAYAR, Shree K. : Flexible mirror imaging. In : *International Conference on Computer Vision*, 2007 (ICCV)

[Lagger et Fua 2006] LAGGER, Pascal ; FUA, Pascal : Using Specularities to Recover Multiple Light Sources in the Presence of Texture. In : *International Conference on Pattern Recognition*, 2006 (ICPR)

[Lagger et al. 2008] LAGGER, Pascal ; SALZMANN, Mathieu ; LEPETIT, Vincent ; FUA, Pascal : 3d pose refinement from reflections. In : *Computer Vision and Pattern Recognition*, 2008 (CVPR)

[Lambert 1760] LAMBERT, Johann H. : *Photometria sive de mensura et gradibus luminis, colorum et umbrae*. 1760

[Lavest et al. 1998] LAVEST, Jean-Marc ; VIALA, Marc ; DHOME, Michel : Do we really need an accurate calibration pattern to achieve a reliable camera calibration ? In : *European Conference on Computer Vision*. Springer, 1998, S. 158–174

[Lee et al. 2010] LEE, Seong-Taek ; YOON, Tae-Ho ; KIM, Kyeong-Seop ; KIM, Kee-Deog ; PARK, Wonse : Removal of specular reflections in tooth color image by perceptron neural nets. In : *International Conference on Signal Processing Systems*, 2010 (ICSPS)

[Lepetit et al. 2009] LEPETIT, Vincent ; MORENO-NOGUER, Francesc ; FUA, Pascal : Epnp : An accurate o (n) solution to the PnP problem. In : *International Journal of Computer Vision* 81 (2009), Nr. 2, S. 155–166

[Lien et Amato 2007] LIEN, Jyh-Ming ; AMATO, Nancy M. : Approximate convex decomposition of polyhedra. In : *Solid and Physical Modeling*, 2007 (SPM)

[Lin et al. 2017] LIN, John ; TAMAAZOUSTI, Mohamed ; SAID, Souheil H. ; MORGAND, Alexandre : Color consistency of specular highlights in consumer cameras. In : *Virtual Reality Software and Technology*, 2017 (VRST)

[Lindsay et Agu 2014] LINDSAY, Clifford ; AGU, Emmanuel : Automatic Multi-light White Balance Using Illumination Gradients and Color Space Projection. In : *International Symposium on Visual Computing*, 2014 (ISVC)

[Loesch et al. 2015] LOESCH, Angelique ; BOURGEOIS, Steve ; GAY-BELLILE, Vincent ; DHOME, Michel : Generic edgelet-based tracking of 3D objects in real-time. In : *Intelligent Robots and Systems*, 2015 (IROS)

[Lombardi et Nishino 2012] Lombardi, Stephen ; Nishino, Ko : Reflectance and natural illumination from a single image. In : *European Conference on Computer Vision*, 2012 (ECCV)

[Long et al. 2015] Long, Gucan ; Kneip, Laurent ; Li, Xin ; Zhang, Xiaohu ; Yu, Qifeng : Simplified Mirror-Based Camera Pose Computation via Rotation Averaging. In : *Computer Vision and Pattern Recognition*, 2015 (CVPR)

[Ma et Chen 1994] Ma, SD ; Chen, X : Quadric surface reconstruction from its occluding contours. In : *International Conference on Pattern Recognition*, 1994 (ICPR)

[Ma et Li 1996] Ma, SD ; Li, Li : Ellipsoid reconstruction from three perspective views. In : *International Conference on Pattern Recognition*, 1996 (ICPR)

[Mandl et al. 2017] Mandl, David ; Yi, Kwang M. ; Mohr, Peter ; Roth, Peter ; Fua, Pascal ; Lepetit, Vincent ; Schmalstieg, Dieter ; Kalkofen, Denis : Learning Lightprobes for Mixed Reality Illumination. In : *International Symposium of Mixed and Augmented Reality*, 2017 (ISMAR)

[Marquardt 1963] Marquardt, Donald W. : An algorithm for least-squares estimation of nonlinear parameters. In : *Journal of the Society for Industrial & Applied Mathematics* 11 (1963), Nr. 2, S. 431–441

[Matusik et al. 2003] Matusik, Wojciech ; Pfister, Hanspeter ; Brand, Matt ; McMillan, Leonard : A Data-Driven Reflectance Model. In : *Transactions on Graphics* 22 (2003), July, Nr. 3, S. 759–769

[Meilland et al. 2013] Meilland, Maxime ; Barat, Christian ; Comport, Andrew : 3D high dynamic range dense visual slam and its application to real-time object relighting. In : *International Symposium on Mixed and Augmented Reality*, 2013 (ISMAR)

[Meilland et al. 2011] Meilland, Maxime ; Comport, A ; Rives, Patrick ; Méditerranée, INRIA Sophia A. : Real-time dense visual tracking under large lighting variations. In : *British Machine Vision Conference*, 2011 (BMVC)

[Mitsumoto et al. 1992] Mitsumoto, Hiroshi ; Tamura, Shinichi ; Okazaki, Kozo ; Kajimi, Naoki ; Fukui, Yutaka : 3-D reconstruction using mirror images based on a plane symmetry recovering method. In : *Pattern Analysis and Machine Intelligence* 14 (1992), Nr. 9, S. 941–946

[Morgand et Tamaazousti 2014] Morgand, Alexandre ; Tamaazousti, Mohamed : Generic and Real-time Detection of Specular Reflections in Images. In : *8th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2014 (VISAPP)

[Morgand et Tamaazousti 2015a] Morgand, Alexandre ; Tamaazousti, Mohamed : Détection générique et temps réel des spécularitÃ©s. In : *Journées francophones des jeunes chercheurs en vision par ordinateur*, 2015 (ORASIS)

[Morgand et Tamaazousti 2015b] Morgand, Alexandre ; Tamaazousti, Mohamed : Reconstruction d'une source lumineuse modélisée par une quadrique à partir d'images multiples. In : *Journées francophones des jeunes chercheurs en vision par ordinateur*, 2015 (ORASIS)

[Morgand et Tamaazousti 2017] Morgand, Alexandre ; Tamaazousti, Mohamed : Modèle empirique pour la prédiction de spécularité : application au retexturing. In : *Journées francophones des jeunes chercheurs en vision par ordinateur*, 2017 (ORASIS)

[Morgand et al. 2015] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : Modèle empirique pour la prédiction de spécularité : application au retexturing. (2015)

[Morgand et al. 2016] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : An Empirical Model for Specularity Prediction with Application to Dynamic Retexturing. In : *International Symposium of Mixed and Augmented Reality*, 2016 (ISMAR)

[Morgand et al. 2017a] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : A Geometric Model for Specularity Prediction on Planar Surfaces with Multiple Light Sources. In : *Transactions on Visualization and Computer Graphics*, 2017 (TVCG)

[Morgand et al. 2017b] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : A Geometric Model for Specularity Prediction on Planar Surfaces with Multiple Light Sources. (2017), S. 1

[Morgand et al. 2017c] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : A Multiple-View Geometric Model of Specularities on Non-Planar Shapes with Application to Dynamic Retexturing. In : *International Symposium of Mixed and Augmented Reality*, 2017 (ISMAR)

[Morgand et al. 2017d] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : A Multiple-View Geometric Model of Specularities on Non-Planar Shapes with Application to Dynamic Retexturing. In : *Transactions on Visualization and Computer Graphics*, 2017 (TVCG)

[Morgand et al. 2017e] MORGAND, Alexandre ; TAMAAZOUSTI, Mohamed ; BARTOLI, Adrien : A multiple-view geometric model of specularities on non-uniformly curved surfaces. In : *Virtual Reality Software and Technology*, 2017 (VRST)

[Mouragnon et al. 2006] MOURAGNON, Etienne ; LHUILLIER, Maxime ; DHOME, Michel ; DEKEYSER, Fabien ; SAYD, Patrick : Real time localization and 3d reconstruction. In : *Computer Vision and Pattern Recognition*, 2006 (CVPR)

[Netz et Osadchy 2013a] NETZ, Aaron ; OSADCHY, Margarita : Recognition Using Specular Highlights. In : *Computer Vision and Pattern Recognition*, 2013 (CVPR)

[Netz et Osadchy 2013b] NETZ, Aaron ; OSADCHY, Margarita : Recognition Using Specular Highlights. In : *Pattern Analysis and Machine Intelligence* 35 (2013), Nr. 3, S. 639–652

[Newcombe et al. 2011] NEWCOMBE, Richard A. ; DAVISON, Andrew J. ; IZADI, Shahram ; KOHLI, Pushmeet ; HILLIGES, Otmar ; SHOTTON, Jamie ; MOLYNEAUX, David ; HODGES, Steve ; KIM, David ; FITZGIBBON, Andrew : KinectFusion : Real-time dense surface mapping and tracking. In : *ACM symposium on User interface software and technology,*, 2011 (ISMAR)

[Ngan et al. 2004] NGAN, Addy ; DURAND, Frédo ; MATUSIK, Wojciech : Experimental validation of analytical BRDF models. In : *Special Interest Group on Computer Graphics and Interactive Techniques*, 2004 (SIGGRAPH)

[Nistér 2004] NISTÉR, David : An efficient solution to the five-point relative pose problem. 26 (2004), Nr. 6, S. 756–770

[Oh et al. 2007] OH, JungHwan ; HWANG, Sae ; LEE, JeongKyu ; TAVANAPONG, Wallapak ; WONG, Johnny ; GROEN, Piet C. de : Informative frame classification for endoscopy video. In : *Medical Image Analysis* 11 (2007), Nr. 2, S. 110–127

[Ortiz et Torres 2005] Ortiz, Francisco ; Torres, Fernando : A new inpainting method for highlights elimination by colour morphology. In : *Pattern Recognition and Image Analysis*. 2005, S. 368–376

[Ortiz et Torres 2006] Ortiz, Francisco ; Torres, Fernando : Automatic detection and elimination of specular reflectance in color images by means of MS diagram and vector connected filters. In : *Systems, Man, and Cybernetics, Part C : Applications and Reviews* 36 (2006), Nr. 5, S. 681–687

[Ozgür et al. 2017] Ozgür, Erol ; Lafont, Alexis ; Bartoli, Adrien : Visualizing In-Organ Tumors in Augmented Monocular Laparoscopy. (2017)

[Phong 1975] Phong, Bui T. : Illumination for computer generated pictures. In : *Communications of the ACM* 18 (1975), Nr. 6, S. 311–317

[Reyes et Bayro-Corrochano 2005] Reyes, Leo ; Bayro-Corrochano, Eduardo : The projective reconstruction of points, lines, quadrics, plane conics and degenerate quadrics using uncalibrated cameras. In : *Image and Vision Computing* 23 (2005), Nr. 8, S. 693–706

[Rohmer et al. 2014] Rohmer, Kai ; Buschel, Wolfgang ; Dachselt, Raimund ; Grosch, Thorsten : Interactive near-field illumination for photorealistic augmented reality on mobile devices. In : *International Symposium on Mixed and Augmented Reality*, 2014 (ISMAR)

[Said et al. 2017] Said, Souheil H. ; Tamaazousti, Mohamed ; Bartoli, Adrien : Image-Based Models for Specularity Propagation in Diminished Reality. In : *Transactions on Visualization and Computer Graphics*, 2017 (TVCG)

[Saint-Pierre et al. 2011] Saint-Pierre, Charles-Auguste ; Boisvert, Jonathan ; Grimard, Guy ; Cheriet, Farida : Detection and correction of specular reflections for automatic surgical tool segmentation in thoracoscopic images. In : *Machine Vision and Applications* 22 (2011), Nr. 1, S. 171–180

[Sato et al. 1999] Sato, Imari ; Sato, Yoichi ; Ikeuchi, Katsushi : Acquiring a radiance distribution to superimpose virtual objects onto a real scene. In : *Visualization and Computer Graphics, IEEE Transactions on* 5 (1999), Nr. 1, S. 1–12

[Savarese et al. 2005] Savarese, Silvio ; Chen, Min ; Perona, Pietro : Local shape from mirror reflections. In : *International Journal of Computer Vision* 64 (2005), Nr. 1, S. 31–67

[Savarese et al. 2004] Savarese, Silvio ; Fei-Fei, Li ; Perona, Pietro : What do reflections tell us about the shape of a mirror ? In : *Proceedings of the 1st Symposium on Applied perception in graphics and visualization* ACM (Veranst.), 2004, S. 115–118

[Semple et Kneebone 1998] Semple, John G. ; Kneebone, Geoffrey T. : *Algebraic projective geometry*. Oxford University Press, 1998

[Shafer 1985] Shafer, Steven A. : Using color to separate reflection components. In : *Color Research & Application* 10 (1985), Nr. 4, S. 210–218

[Shashua et Toelg 1997] Shashua, A ; Toelg, S : The quadric reference surface : Theory and applications. In : *International Journal of Computer Vision*, 1997 (IJCV)

[Shroff et al. 2011] Shroff, Nitesh ; Taguchi, Yuichi ; Tuzel, Oncel ; Veeraraghavan, Ashok ; Ramalingam, Srikumar ; Okuda, Haruhisa : Finding a needle in a specular haystack. In : *International Conference on Robotics and Automation*, 2011 (ICRA)

[Silveira et Malis 2007] SILVEIRA, Geraldo ; MALIS, Ezio : Real-time visual tracking under arbitrary illumination changes. In : *Computer Vision and Pattern Recognition*, 2007 (CVPR)

[Spivak 1970] SPIVAK, Michael D. : *A comprehensive introduction to differential geometry*. Publish or perish, 1970

[Stehle 2006] STEHLE, Thomas : Removal of specular reflections in endoscopic images. In : *Acta Polytechnica : Journal of Advanced Engineering* 46 (2006), Nr. 4, S. 32–36

[Sturm et Gargallo 2007] STURM, Peter ; GARGALLO, Pau : Conic fitting using the geometric distance. In : *Asian Conference on Computer Vision*. 2007 (ACCV)

[Suzuki et Satoshi 1985] SUZUKI ; SATOSHI : Topological structural analysis of digitized binary images by border following. In : *Computer Vision, Graphics, and Image Processing* 30 (1985), Nr. 1, S. 32–46

[Tamaazousti 2013] TAMAAZOUSTI, Mohamed : *L'ajustement de faisceaux contraint comme cadre d'unification des méthodes de localisation : application à la réalité augmentée sur des objets 3D*, Université Blaise Pascal-Clermont-Ferrand II, Dissertation, 2013

[Tamaazousti et al. 2011a] TAMAAZOUSTI, Mohamed ; GAY-BELLILE, Vincent ; COLLETTE, Sylvie N. ; BOURGEOIS, Steve ; DHOME, Michel : Real-time accurate localization in a partially known environment : Application to augmented reality on textureless 3d objects. In : *ISMAR workshop*, 2011 (ISMARW)

[Tamaazousti et al. 2011b] TAMAAZOUSTI, Mohamed ; GAY-BELLILE, Vincent ; NAUDET COLLETTE, S ; BOURGEOIS, Steve ; DHOME, Michel : Nonlinear refinement of structure from motion reconstruction by taking advantage of a partial knowledge of the environment. In : *Computer Vision and Pattern Recognition*, 2011 (CVPR)

[Tamaazousti et al. 2018] TAMAAZOUSTI, Mohamed ; MORGAND, Alexandre ; BARTOLI, Adrien : *Method for processing images having specularities and corresponding computer program product*. Mai 17 2018. – US Patent App. 15/574,475

[Tamaazousti et al. 2016] TAMAAZOUSTI, Mohamed ; NAUDET-COLLETTE, Sylvie ; GAY-BELLILE, Vincent ; BOURGEOIS, Steve ; BESBES, Bassem ; DHOME, Michel : The constrained SLAM framework for non-instrumented augmented reality. In : *Multimedia Tools and Applications* 75 (2016), Nr. 16, S. 9511–9547

[Tamaazousti et al. 2017a] TAMAAZOUSTI, Youssef ; BORGNE, Hervé Le ; HUDELOT, Céline ; SEDDIK, Mohamed El A. ; TAMAAZOUSTI, Mohamed : Learning More Universal Representations for Transfer-Learning. In : *arXiv preprint arXiv :1712.09708* (2017)

[Tamaazousti et al. 2017b] TAMAAZOUSTI, Youssef ; LE BORGNE, Hervé ; HUDELOT, Céline : MuCaLe-Net : Multi Categorical-Level Networks to Generate More Discriminating Features. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, S. 6711–6720

[Tan et al. 2004] TAN, Robby T. ; NISHINO, Ko ; IKEUCHI, Katsushi : Separating reflection components based on chromaticity and noise analysis. In : *Pattern Analysis and Machine Intelligence* 26 (2004), Nr. 10, S. 1373–1379

[Tan et al. 2013] TAN, Wei ; LIU, Haomin ; DONG, Zilong ; ZHANG, Guofeng ; BAO, Hujun : Robust monocular SLAM in dynamic environments. In : *International Symposium on Mixed and Augmented Reality*, 2013 (ISMAR)

[Tardif et al. 2008] Tardif, J-P ; Pavlidis, Yanis ; Daniilidis, Kostas : Monocular visual odometry in urban environments using an omnidirectional camera. In : *International Conference on Intelligent Robots and Systems*, 2008 (IROS)

[Telea 2004] Telea, Alexandru : An image inpainting technique based on the fast marching method. In : *Journal of graphics tools* 9 (2004), Nr. 1, S. 23–34

[Thomas et al. 2016] Thomas, Richter-Trummer ; Denis, Kalkofen ; Jinwoo, Park ; Dieter, Schmalstieg : Instant Mixed Reality Lighting from Casual Scanning. In : *International Symposium of Mixed and Augmented Reality*, 2016 (ISMAR)

[Tian et Clark 2013] Tian, Qing ; Clark, James J. : Real-time Specularity Detection Using Unnormalized Wiener Entropy. In : *Computer and Robot Vision*, 2013 (CRV)

[Torr et Murray 1997] Torr, Philip H. ; Murray, David W. : The development and comparison of robust methods for estimating the fundamental matrix. In : *International Journal of Computer Vision* 24 (1997), Nr. 3, S. 271–300

[Torrance et Sparrow 1967] Torrance, Kenneth E. ; Sparrow, Ephraim M. : Theory for off-specular reflection from roughened surfaces. In : *Journal of the Optical Society of America* 57 (1967), Nr. 9, S. 1105–1114

[Torres et al. 2003] Torres, Fernando ; Angulo, Jesús ; Ortiz, Francisco : Automatic detection of specular reflectance in colour images using the MS diagram. In : *Computer Analysis of Images and Patterns*, 2003 (CAIP)

[Ward 1994] Ward, Gregory J. : The RADIANCE lighting simulation and rendering system. In : *Special Interest Group on Computer Graphics and Interactive Techniques*, 1994 (SIGGRAPH)

[Whelan et al. 2015] Whelan, Thomas ; Leutenegger, Stefan ; Salas-Moreno, Renato F. ; Glocker, Ben ; Davison, Andrew J. : ElasticFusion : Dense SLAM Without A Pose Graph. In : *Robotics : science and systems* Bd. 11, 2015

[Wong et al. 2008] Wong, Kwan-Yee K. ; Schnieders, Dirk ; Li, Shuda : Recovering light directions and camera poses from a single sphere. In : *European Conference on Computer Vision*. 2008 (ECCV)