THÈSE

Pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ DE POITIERS
UFR des sciences fondamentales et appliquées
XLIM-SIC
(Diplôme National - Arrêté du 7 août 2006)

École doctorale : Sciences et ingénierie pour l'information, mathématiques - S2IM (Poitiers)
Secteur de recherche : Traitemement du signal et des images

Présentée par :
Alexandre Bony

Modélisation de l'interaction lumière/matière pour l'analyse de surfaces rugueuses texturées par stéréo photométrie

Directeur(s) de Thèse :
Majdi Khoudeir, Benjamin Bringier

Soutenue le 02 décembre 2013 devant le jury

Jury :

Président	**Fabrice Mériaudeau**	Professeur des Universités, Université de Dijon
Rapporteur	**Abdelaziz Bensrhair**	Professeur des Universités, INSA de Rouen
Rapporteur	**Hassan Zahouani**	Professeur des Universités, École centrale de Lyon
Membre	**Majdi Khoudeir**	Professeur des Universités, Université de Poitiers
Membre	**Benjamin Bringier**	Maître de conférences, Université de Poitiers
Membre	**Jean-Denis Durou**	Maître de conférences, Université Paul Sabatier, Toulouse 3

Pour citer cette thèse :
Alexandre Bony. *Modélisation de l'interaction lumière/matière pour l'analyse de surfaces rugueuses texturées par stéréo photométrie* [En ligne]. Thèse Traitemement du signal et des images. Poitiers : Université de Poitiers, 2013. Disponible sur Internet <http://theses.univ-poitiers.fr>

Thèse

Pour l'obtention du grade de

Docteur de l'université de poitiers

Faculté des sciences Fondamentales et Appliquées
(Diplôme national – arrêté du 7 août 2006)

École Doctorale **: Sciences et Ingénierie pour l'Information**
Secteur de Recherche **: Traitement du Signal et des Images**

Présentée par

Alexandre BONY

****

Modélisation de l'interaction lumière/matière pour l'analyse de surfaces rugueuses texturées par stéréo photométrie

****

Thèse soutenue le 2 decembre 2013 devant le jury composé de :

Abdelaziz BENSRHAIR, Professeur, LITIS – INSA de Rouen ..... *Rapporteur*
Hassan ZAHOUANI, Professeur, LTDS Ecole Centrale de Lyon ..... *Rapporteur*
Fabrice MERIAUDEAU, Professeur, Laboratoire Le2i Le Creusot ..... *Membre du jury*
Jean-Denis DUROU, HDR, Université Paul Sabatier de Toulouse ..... *Membre du jury*
Majdi KHOUDEIR, Professeur, Université de Poitiers ..... *Directeur de Thèse*
Benjamin BRINGIER, Maitre de Conférence, Université de Poitiers ..... *Co-Directeur de Thèse*

Remerciements

Je tiens a exprimer ma plus profonde gratitude à tous ceux qui, de près ou de loin, ont contribué à ce mémoire de thèse.

Je voudrais tout d'abord exprimer mes profonds remerciements au professeur Majdi KHOUDEIR mon directeur de thèse, qui m'a offert l'opportunité de réaliser ces travaux en travaillant à ses cotés.

Le remerciement suivant revient très justement à mon co-directeur de thèse Benjamin BRINGIER pour son aide précieuse, ses conseils avisés et son écoute. Je lui suis également reconnaissant pour sa disponibilité et ses qualités scientifiques et pédagogique grâce auxquelles j'ai beaucoup appris.

Je tiens à remercier les professeurs Abdelaziz BENSRHAIR et Hassan ZAHOUANI pour avoir accepté d'être les rapporteurs de ce manuscrit. Je remercie également Jean-Denis DUROU ainsi que Fabrice MERIAUDEAU d'avoir accepté d'être membre du jury et d'assister à la présentation de ce travail.

Je tiens aussi à remercier l'ensemble du personnel du laboratoire XLIM-SIC de Poitiers pour leur accueil et leurs conseils. Plus particulièrement, j'exprime un vif remerciement à Philippe DUBOIS, pour sa disponibilité ainsi que son aide avisée dans l'élaboration des prototypes.

Merci à tous mes amis pour leur soutien et leur patience, mais aussi pour tous les bons moments passés à leurs côtés le long de ces années.

Un énorme remerciement à toute ma famille. Merci à mes parents qui ont su m'apporter leur amour et leur soutien sans lesquels je n'aurais pu arriver au terme de ces études. Merci à Muriel (ma sœur), Jean François (mon beau-frère) ainsi que Maëva et Jade (mes nièces) qui malgré de dures épreuves m'ont toujours apporté de la joie et un soutien ininterrompus.

Le dernier et non le moindre, je remercie chaleureusement ma compagne, Emma, pour sa patience, son soutien et ses encouragements tout le long de ces années.

Je conclurais ces remerciements par une pensée adressée à Germaine BONY (ma grand-mère) qui n'a pas pu suivre à nos cotés le chemin vers mon accession.

à Poitiers, le 28 octobre 2013.

Résumé

Les techniques de reconstruction 3d sont devenues incontournables pour des applications telles que la caractérisation et l'analyse de surfaces. Les travaux réalisés au cours de cette thèse ont pour objectif d'améliorer la qualité des reconstructions 3d par stéréo-photométrie. Cette méthode repose sur deux principes, l'inversion d'un modèle d'interaction lumière/matière (BRDF) et la configuration d'un système d'éclairage et de prises de vues. Pour des surfaces diffuses, la stéréo-photométrie est réalisée à partir d'un minimum de trois images acquises d'un point de vue fixe pour des directions d'éclairages différentes. Son avantage est d'extraire simultanément les propriétés géométriques et colorimétriques des surfaces analysées même en cas de forte rugosité. Néanmoins, son application exige la formulation de plusieurs hypothèses qui sont difficilement respectables dans un contexte réel. Ceci génère des erreurs significatives dans les reconstructions. Pour les réduire, nous proposons différentes contributions qui s'articulent autour de la prise en compte globale de la chaine d'acquisition. Les apports de nos travaux se situent aux niveaux de la caractérisation et de la modélisation du système d'éclairage, du capteur d'acquisition et de l'amélioration de la qualité des images. Nous nous sommes aussi intéressés à l'optimisation des protocoles de prises de vues dans le cas de spécularité surfacique ou d'ombrage dus à la présence de rugosité. Les résultats obtenus montrent que la prise en compte de ces caractéristiques dans l'inversion d'un modèle de BRDF permet une nette amélioration des reconstructions et offre la possibilité de réduire la taille des systèmes d'acquisition.

Mots-clés : *reconstruction 3d, acquisition d'image, photométrie, stéréo-photométrie, brdf, imagerie.*

Abstract

Tridimensional reconstruction method has become essential for applications such as the characterization and analysis of surfaces. In this thesis, aims are to increase the quality of 3d reconstructions by photometric stereo. This method is based on two principles, reversing light-matter interaction model and configuration of a lighting system. With diffuse surfaces, the photometric stereo use three captured images from a fixed point of view for different illumination directions. Its main advantage is to extract the color and geometric properties for the textured rough surfaces. However, its application requires to make assumptions that are not credible in real cases. This problem generates significant errors in the reconstructions. To reduce them, we offer various solutions around the overall consideration of the acquisition chain. Our contribution focuses on the characterization and modeling of the lighting system, the acquisition sensor and improved image quality. We are also interested to optimize acquisition protocol in the case of specular surface or shading due to the surface geometry. Our results show that the inclusion of these features in the inversion of a BRDF model allows an improvement of 3d reconstructions as well as the possibility of reducing the size of the acquisition systems.

Keywords : *3d reconstruction, image acquisition, photometry, photometric stereo, BRDF, imaging.*

Table des matières

Table des matières ix

Notations xiii

1 Introduction 1
1.1 Contexte et motivation . . . 2
1.2 Structure du document . . . 2

2 Notions de lumière 5
2.1 Définition historique de la lumière . . . 6
2.1.1 Théorie corpusculaire . . . 6
2.1.2 Théorie ondulatoire . . . 7
2.1.3 Complémentarité des deux théories . . . 9
2.1.4 Le spectre lumineux . . . 10
2.2 Les sources lumineuses . . . 11
2.2.1 Classification des sources lumineuses . . . 11
2.2.1.1 Sources à spectre continu . . . 11
2.2.1.2 Sources à spectre discret . . . 12
2.2.1.3 Sources à spectre mixte . . . 12
2.2.2 Métrologie : définitions des valeurs . . . 13
2.2.2.1 La radiométrie . . . 13
2.2.2.2 La photométrie . . . 17
2.2.3 Métrologie : bilan . . . 18
2.3 Interactions physiques lumière matière . . . 18
2.3.1 L'absorption . . . 19
2.3.2 Spectre de réflectance . . . 19
2.3.3 La transmission . . . 20
2.3.4 La réflexion . . . 20
2.3.5 La diffusion . . . 21
2.3.6 Autres phénomènes . . . 21
2.4 Modélisation mathématique . . . 22
2.4.1 Le modèle de Lambert . . . 22
2.4.2 Le modèle de Phong . . . 24
2.4.3 Autres modèles . . . 26
Conclusion . . . 27

3 La capture d'image 29
3.1 Principe et Historique . . . 30
3.2 L'objectif et sa géométrie : la perte dimensionnelle . . . 33
3.2.1 Principe de convergence et modèle de projection . . . 33
3.2.1.1 Le cas de la lentille mince . . . 33

3.2.1.2 Le cas de l'objectif . . . 35
3.2.1.3 Modélisation mathématique de la projection . . . 35
3.2.2 Le diaphragme . . . 39
3.2.3 Les défauts . . . 40
3.2.3.1 Les aberrations . . . 40
3.2.3.2 La profondeur de champ . . . 40
3.2.3.3 L'éclairement énergétique . . . 41
3.3 Le capteur numérique . . . 42
3.3.1 L'effet photoélectrique . . . 43
3.3.2 Le photosite . . . 44
3.3.2.1 La pleine capacité . . . 45
3.3.2.2 Le temps d'intégration . . . 45
3.3.2.3 Rendement quantique . . . 46
3.3.2.4 Efficacité quantique . . . 46
3.3.3 Le circuit de lecture . . . 46
3.3.3.1 Les bruits de conversion . . . 47
3.3.3.2 Bilan : tension de sortie . . . 48
3.3.4 La conversion analogique numérique . . . 48
3.3.4.1 Cas idéal : quantification . . . 49
3.3.4.2 Erreur de quantification . . . 49
3.3.4.3 Cas réel : non linéarité . . . 50
3.3.5 La dynamique . . . 51
3.3.5.1 Dynamique photonique . . . 51
3.3.5.2 Dynamique photoélectrique . . . 51
3.3.5.3 Dynamique numérique . . . 51
3.3.5.4 Dynamique du capteur . . . 52
3.3.6 Les technologies des capteurs d'images . . . 52
3.3.6.1 La structure MOS-XY . . . 53
3.3.6.2 La structure CID . . . 53
3.3.6.3 La structure CCD . . . 54
3.3.6.4 La structure APS-CMOS . . . 55
3.3.6.5 Bilan technologique . . . 56
3.3.7 Le capteur : outil métrologique . . . 57
3.3.7.1 Les dispositifs tri-CCD . . . 59
3.3.7.2 Les dispositifs à filtrage intégré . . . 60
3.3.7.3 Les dispositifs à triples jonctions enterrées . . . 60
3.3.7.4 Bilan . . . 61
3.3.8 La gestion des données numériques . . . 61
Conclusion . . . 63

4 La Reconstruction 3D par vision **65**
4.1 État de l'art des méthodes . . . 66
4.1.1 Analyse du temps de vol . . . 66
4.1.2 Analyse des déformations d'un signal lumineux . . . 66
4.1.3 Analyse des intensités énergétiques mesurées . . . 67
4.1.3.1 Le Shape from Shading . . . 67
4.1.3.2 Le Shape from Stéréo-Photométrie . . . 69
4.1.3.3 Le Shape from Contours . . . 69
4.1.3.4 Le Shape from Texture . . . 69
4.1.3.5 Le Shape from Focus/Defocus . . . 70

4.1.3.6 Le Shape from Stéréo-Vision . . . 70
4.2 La Stéréo-Photométrie . . . 71
4.2.1 Extraction de l'information . . . 72
4.2.2 Intégration des champs de gradients . . . 74
4.2.2.1 Méthodes Locales . . . 75
4.2.2.2 Méthodes basées sur la relaxation . . . 77
4.2.2.3 Méthodes Globales . . . 78
4.2.2.4 Conclusion . . . 80
4.3 Ombres et Spécularités dans la SP . . . 80
4.3.1 Élimination des ombres et points spéculaires (s>3) . . . 82
4.3.2 Notre méthode d'élimination . . . 85
4.3.2.1 Élimination des ombres . . . 86
4.3.2.2 Élimination de la spécularité . . . 87
4.3.2.3 Résultats par simulation . . . 88
4.3.2.4 Acquisitions réelles . . . 91
Conclusion . . . 95

5 Éclairage réaliste en SP **97**
5.1 Influence des paramètres d'éclairage . . . 98
5.1.1 Etude des erreurs de position des sources . . . 98
5.1.1.1 Erreurs sur l'angle zénith . . . 99
5.1.1.2 Erreurs sur l'angle azimut . . . 101
5.1.1.3 Erreurs globales sur les positions des sources . . . 102
5.1.2 Étude de la non-uniformité de l'éclairage . . . 103
5.2 Etat de l'art . . . 106
5.2.1 Calibration des positions des sources . . . 107
5.2.1.1 Méthodes basées sur l'analyse statistique des intensités . . . 107
5.2.1.2 Méthodes basées sur l'analyse des ombres . . . 111
5.2.1.3 Méthodes basées sur l'analyse des points spéculaires . . . 113
5.2.1.4 Bilan de l'étude comparative . . . 116
5.2.2 Méthodes de correction de la non-uniformité des sources . . . 117
5.3 Méthodes proposées . . . 119
5.3.1 Calibration des positions des sources . . . 119
5.3.2 Correction de la non-uniformité de l'éclairage dans la SP . . . 124
Conclusion . . . 128

6 Chaine d'acquisition en SP **131**
6.1 Les limites du capteur d'image dans la SP . . . 132
6.2 Les images hautes dynamiques . . . 133
6.2.1 État de l'art du principe HDR . . . 133
6.2.2 Images HDR, une solution en SP ? . . . 138
6.3 Une nouvelle méthode d'acquisition : images AIQ . . . 141
6.3.1 Le compromis Quantification / Dynamique . . . 142
6.4 Les images AIQ pour la SP . . . 144
6.4.1 Protocole d'acquisition AIQ . . . 146
6.4.2 Les images AIQ et les reconstructions 3D . . . 146
6.4.2.1 Résultats de simulation . . . 146
6.4.2.2 Surfaces réelles . . . 147
Conclusion . . . 151

Conclusion générale **153**

A Annexes **157**
A.1 Tableau de correspondance Radiométrie-Photométrie 158
A.2 Nos systèmes d'acquisition 159
A.2.1 NIKON_1 159
A.2.2 NIKON_2 160
A.3 Étude comparative des méthodes d'intégration numérique 161
A.3.1 Comparaison 1 : qualité des reconstructions 162
A.3.2 Comparaison 2 : robustesse face au bruit 165
A.3.3 Comparaison 3 : temps de calcul 165
A.4 Étude comparative des méthodes de calibration des positions des sources 167
A.4.1 Étude des méthodes basées sur l'analyse statistique des intensités 167
A.4.2 Etude des méthodes basées sur l'analyse des ombres 175
A.4.3 Etude des méthodes basées sur l'analyse des points spéculaires 177

Bibliographie **181**

Notations

$[\mathbf{A}]$	Une matrice (entre crochet et lettre en gras)
$\vec{a}$	Un vecteur (flèche)
$\bar{a}$	Une moyenne (barre)
$\mathrm{d}a$	La dérivée
$var(a)$	La variance
A, B, a, b, C, c	Lettres génériques utiliser pour exprimer des matrices, des vecteurs ou autres
i, j	Lettres génériques réservées à l'indexation
r, R	Rayon d'un cercle ou d'une sphère
$\epsilon, \bar{\epsilon}$	Une erreur et sa moyenne
$\mathcal{R}$	Désigne un repère euclidien
$\mathcal{B}$	Désigne un bruit
$\mathcal{H}$	Désigne les hypothèses applicatives de la stéréo-photométrie
$\mathcal{D}$	Désigne une dynamique
$Ie, \hat{Ie}$	Intensité lumineuse emise par une scène et son estimation
Ie_0	Intensité lumineuse incidente à la surface
$Ie_{0source}$	Intensité lumineuse emise par une source en direction de la surface
$Ie_{ambiente}$	Intensité lumineuse ambiante
Ie_{spec}	Intensité lumineuse générée par la composante spéculaire
$Le, \hat{Le}$	Radiance émise par une scène et son estimation
$Ee, \hat{Ee}$	Irradiance et son estimation
λ	Longueurs d'ondes
$V(\lambda)$	Courbes de sensibilité CIE
M	Valeurs numériques quantifiées issues des capteurs d'images
M'	Valeurs numériques quantifiées jugées sans ombre en un point
crf, crf^{-1}	Réponse de la caméra (camera response function) et son inverse
U, V, u, v	Résolution d'une image suivant les lignes / colonnes et indexation
N, N', n	Nombre de Bits d'une image LDR, HDR et indexation
qt	Quantum d'un quantifieur
$\vec{n}$	Normales à la surface
$\vec{l}$	Direction d'une source lunineuse
$\vec{r}$	Direction de reflexion
$\vec{v}$	Direction d'observation
$\vec{h}$	Direction dans le modèle de BLINN
θ	Angle d'incidence source lumineuse / normale à la surface
τ, σ	Angles azimutaux et zénithaux des sources lumineuses
$\hat{\tau}, \hat{\sigma}$	Angles azimutaux et zénithaux estimés des sources lumineuses
η	Angle entre $\vec{r}$ et $\vec{v}$

δ_{LC}	Distance entre un point de la scène et une source lumineuse
$\vec{\delta}_{LC}$	Vecteur des distances entre un point de la scène et les sources lumineuses
$\zeta, \hat{\zeta}$	Élévation des surface suivant l'axe z et son estimation
$[\mathbf{L}]$	Matrice regroupant les directions des sources lumineuses
l_x, l_y, l_z	Coordonnées cartésiennes d'une source lumineuse
$[\mathbf{M}]$	Matrice regroupant des valeurs numériques quantifiées
ρ	Albédos d'une surface
p, q	Champs de gradients suivant les axes x et y
S, s	Nombre de sources lumineuses et indexation
S', s'	Nombre de sources lumineuses et indexation (ne générant pas d'ombre en un point)
f	Distance focale
d	Diamètre d'ouverture d'un objectif
γ	Grandissement optique
k	Facteur de conversion radiance/irradiance définit par Horn et relatif à l'objectif
g_n	Gain d'amplification relatif à la sensibilité ISO
k_x, k_y	Taille métrique d'un pixel suivant les axes x et y
$\mathrm{d}S$	L'aire d'une surface métrique
c_x, c_y	Indexes du centre optique dans une image
Ψ	Représentation angulaire des champs de gradients
ϵ_Ψ	Représentation de l'erreur angulaire entre des champs de gradients
$\bar{\epsilon}_\Psi$	Erreur angulaire moyenne entre des champs de gradients
α, β	Points (x, y, z) d'une scène ou (u, v) d'une image
Δt	Temps d'obturation (aussi appelé temps d'acquisition, d'intégration ou shutter time)
Δt_G	Temps d'acquisition global nécessaire pour une série multi-expositions
G, g	Nombre d'images dans une série multi-exposition et indexation
w	Fonction de poids pour la fusion d'images
χ	Ratio d'exposition entre les image LDR d'une série multi-exposition

Introduction 1

Sommaire

1.1 Contexte et motivation 2
1.2 Structure du document 2

La perception en trois dimensions est naturelle pour l'être humain. Les yeux et des traitements complèxes par le cerveau permettent d'apprécier des distances, des formes et des couleurs. Dès l'époque de la renaissance, les grands peintres tels que De Vinci ou DellaPorta ont cherché à réaliser des esquisses dont la vision permet une restitution fidèle du relief. Depuis ces premières représentations, l'intéret de l'être humain pour la conservation et l'archivage le plus fidèle possible de scènes réelles n'a fait que croître au fil du temps. Lors de ces dernières décennies, notre société moderne à entrepris de grands travaux pour la conservation et le partage du patrimoine. Ainsi, la numérisation 3d couleur est un outil important pour archiver ou mettre en valeur les plus grandes oeuvres et vestiges de notre société.

Mais l'acquisition 3d ne se cantonne pas uniquement au domaine de la conservation. L'essor des techniques industrielles génèrent d'autres enjeux tout aussi importants. Les nouvelles méthodes d'acquisition 3d facilitent l'application d'algorithmes pour le contrôle qualité. Il est ainsi possible de mesurer des propriétés ou des défauts surfaciques comme la rugosité. Ces méthodes sont nombreuses et diverses, elles sont généralement classées suivant deux catégories : avec ou sans contact. Les travaux de cette thèse se positionnent dans la seconde famille de méthodes. Nous proposons des techniques robustes et précises pour mesurer la colorimétrie et la géométrie d'échantillons sans risque de détérioration.

1.1 Contexte et motivation

Ces travaux de thèse sont réalisés au sein du laboratoire XLIM-SIC à Poitiers. Ils s'inscrivent dans un projet appelé Surface Image Matière (SIM) dont l'objectif principal est la réalisation de mesures métrologiques sans contact et non destructives par le traitement d'images numériques. Fort de plusieurs années d'expériences, ce projet fait l'objet de multiples collaborations pour le contrôle qualité des surfaces autoroutières, des bâtiments durables ou pour l'étude d'objets et de vestiges archéologiques.

Parmi l'ensemble des axes du projet, l'un des objectifs principaux est la reconstruction topographique tridimensionnelle afin de faciliter l'étude des caractéristiques en termes de relief/rugosité ainsi que l'extraction des propriétés colorimétriques et photométriques des surfaces analysées. De plus, la méthode utilisée doit impérativement être rapide et facile à mettre en oeuvre. Pour répondre à ces objectifs, nous proposons d'utiliser la stéréo-photométrie (SP) comme technique de reconstruction tridimensionnelle car elle permet de dissocier les constituants d'une image : géométrie, couleur et condition d'éclairage. Sa mise en oeuvre est simple et elle permet la reconstruction de surfaces de grandes tailles ($> 10\ cm^2$) en un temps faible ($< 1\ min$) pour des résolutions importantes ($50\ \mu m$).

Malgré des performances théoriques élevées, l'application de la stéréo photométrie se révèle imparfaite en conditions réelles. Elle nécessite la formulation d'hypothèses contraignantes et restrictives qui, lorsqu'elles ne sont pas respectées, entraînent des erreurs de reconstruction. Chaque étape de la chaine d'acquisition est impactée par ces hypothèses, la première concerne la nature photométrique des surfaces analysées. La BRDF (Bidirectional Reflectance Distribution Function) des surfaces analysées doit être parfaitement connue et inversible. Deux autres hypothèses concernent la géométrie ainsi que le profil énergétique des sources d'éclairage utilisées.

Enfin l'une des hypothèses les plus réductrices provient du processus d'acquisition des images par les appareils photographiques considérés comme parfaits. Malgré bon nombre de travaux sur l'imperfection des mesures due à la quantification, le bruit d'acquisition, les non linéarités ou encore les distorsions géométriques de l'image, la stéréo-photométrie est encore très largement appliquée à partir des valeurs numériques fournies par les capteurs d'images sans aucun traitement.

L'objectif de nos travaux est d'améliorer les résultats de reconstruction 3d dans un contexte d'application réel. Pour atteindre cet objectif, nous proposons des techniques robustes basées sur les méthodes existantes afin de conserver la rapidité et la faible complexité de mise en œuvre de la stéréo-photométrie. Les méthodes proposées permettent de se soustraire des contraintes fixées par les hypothèses applicatives et offre une amélioration significative des résultats.

1.2 Structure du document

Le chapitre 2 est consacré à la description de la lumière et son interaction. Le cheminement scientifique sur sa nature est rappelé par un bref historique. La suite du chapitre est dédiée à la description d'une source lumineuse et son interaction avec une surface représentable par des modèles mathématiques. Il permet aussi de définir les valeurs métrologiques liées à la radiométrie et la photométrie.

Le chapitre 3 propose la description du principe géométrique et énergétique de capture d'une image par le biais d'un capteur numérique. L'accent est porté sur la modélisation mathématique du capteur ainsi que la caractérisation de ses limitations. Grâce à une description détaillée du principe de conversion photo-électrique-numérique, il fait apparaitre l'imperfection des mesure énergétiques.

Le chapitre 4 permet de répondre à l'objectif initial, il est constitué d'un état de l'art des méthodes existantes de reconstruction tridimensionnelle à partir d'images. Le principe et les hypothèses applicatives de la stéréo-photométrie sont décrites ainsi que les limitations et les problèmes rencontrés. Notre première contribution apparait et permet de lever l'hypothèse sur la nature photométrique des surfaces analysées.

Le chapitre 5 est consacré au système d'éclairage nécessaire à l'application de la stéréo-photométrie. Il comprend une analyse démonstrative des problèmes rencontrés lors de la négligence des contraintes d'application. Cette étude est suivie d'un état de l'art détaillé des techniques existantes permettant d'obtenir et de corriger les paramètres d'éclairage. Enfin en dernière partie sont présentées nos contributions qui visent à lever ces contraintes tout en améliorant la fidélité des reconstructions et en permettant la miniaturisation des systèmes d'acquisition.

Le chapitre 6 est dédié à l'imperfection des messures issues des capteurs optiques. La première partie permet de démontrer l'impact négatif de ces mesures biaisées sur les reconstructions tridimensionnelles. Ensuite sont abordées les méthodes d'acquisitions censées outrepasser les limitations des capteurs et lever l'hypothèse d'un processus d'acquisition parfait. Dans ce chapitre, nous proposons une adaptation de ces méthodes et des nouveaux algorithmes de traitement qui permettent d'adapter les images à la stéréo-photométrie.

Notions de lumière 2

Sommaire

2.1 Définition historique de la lumière . . . 6
2.1.1 Théorie corpusculaire . . . 6
2.1.2 Théorie ondulatoire . . . 7
2.1.3 Complémentarité des deux théories . . . 9
2.1.4 Le spectre lumineux . . . 10
2.2 Les sources lumineuses . . . 11
2.2.1 Classification des sources lumineuses . . . 11
2.2.2 Métrologie : définitions des valeurs . . . 13
2.2.3 Métrologie : bilan . . . 18
2.3 Interactions physiques lumière matière . . . 18
2.3.1 L'absorption . . . 19
2.3.2 Spectre de réflectance . . . 19
2.3.3 La transmission . . . 20
2.3.4 La réflexion . . . 20
2.3.5 La diffusion . . . 21
2.3.6 Autres phénomènes . . . 21
2.4 Modélisation mathématique . . . 22
2.4.1 Le modèle de Lambert . . . 22
2.4.2 Le modèle de Phong . . . 24
2.4.3 Autres modèles . . . 26
Conclusion . . . 27

Dans ce chapitre, nous reprenons la définition de la lumière et les outils métrologiques qui permettent sa mesure. Cette définition est primordiale car pour être visualisées, analysées ou mesurées, les surfaces ont besoin de lumière et de l'interaction produite avec la matière dont elles sont constituées. C'est pourquoi, nous consacrons la seconde partie de ce chapitre à référencer les interactions lumière/matière en fonction de la nature géométrique et des propriétés physiques de surfaces analysées. Enfin nous terminons par la description des modèles mathématiques capables de reproduire ces interactions.

2.1 Définition historique de la lumière

De l'origine de l'être humain jusqu'aux premières études menées par la civilisation grecque, la lumière fut considérée comme divine. En 300 av. JC, le mathématicien Euclide pour qui la lumière est étroitement liée à la vision, la représente sous forme de minces filets rectilignes appelés rayons lumineux. D'après sa théorie, l'œil envoie des rayons frapper les objets ce qui permet de les voir.

Quelques centaines de siècles plus tard, les Arabes récupèrent et traduisent les ouvrages grecs durant leurs conquêtes. Ainsi ils héritent des savoirs antiques, les approfondissent et innovent. Vers l'an 1000, Ibn Al-Haitan comprend le fonctionnement de l'œil et renverse la théorie d'Euclide. Il affirme que la lumière provient des objets. De plus, aux travers de la science basée sur l'observation et la mesure, il ébauche les premières lois de réflexion et réfraction. Durant le moyen âge, l'apogée du peuple arabe permet d'exporter les écrits d'Ibn Al-Haitan vers l'Europe. Traduits en latin, ces livres inspirent la communauté scientifique. Ainsi en 1500, Galilée passe de l'optique géométrique à l'optique instrumentale et construit la première lunette astronomique. Elle permet de découvrir les étoiles de la voie lactée, d'observer les détails de la lune et d'autres astres tel que *Io* le satellite de Jupiter. Pour Galilée, ces découvertes sont synonymes d'interrogations sur la vitesse de la lumière. D'ailleurs, il tente de mesurer sa vitesse de trajet entre deux collines de Toscane avec une lanterne à volets mais cette expérience reste vaine car les instruments de l'époque sont trop rudimentaires. Cet échec ne fait que conforter la plupart des scientifiques qui croient depuis l'antiquité que la lumière se propage instantanément. Parmi eux, Snell et Descartes établissent en 1625-1637 les lois de la réfraction et de la réflexion publiées dans [Des65]. Pour démontrer la loi de la réfraction, Descartes utilise l'exemple d'une balle qui pénètre un milieu aqueux. Cette démonstration est sujette à controverse car le changement de milieu entraîne la perte de vitesse du projectile. Ce phénomène est inconcevable avec sa théorie de transmission instantanée de la lumière. Il émet alors plusieurs hypothèses hasardeuses sur ce phénomène et lance un débat sur la nature de la lumière qui va durer trois siècles.

Après la mort de Descartes, Römer relève des irrégularités dans le déplacement du satellite *Io* utilisé comme horloge par les navigateurs. En effet, sa visibilité diffère en fonction de la distance entre la Terre et Jupiter. Römer attribue ce phénomène au temps supplémentaire que met la lumière à parvenir jusqu'à la Terre. Ainsi ces observations rejoignent la théorie de Galilée, la lumière n'a pas une transmission instantanée et possède une vitesse finie. Cette affirmation conforte Huygens dans son idée que la lumière est une onde. Cette idée ce heurte à celle de Newton pour qui la lumière est un flot de particules. Ce débat sur la nature de la lumière scinde la communauté scientifique en deux groupes, l'un partisan de la théorie ondulatoire initiée par Huygens et l'autre partisan de la théorie corpusculaire énoncée par Newton.

2.1.1 Théorie corpusculaire

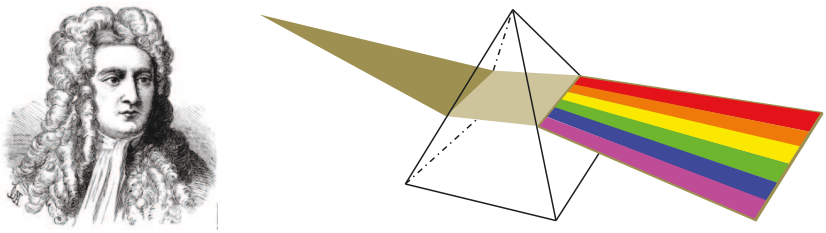
Pour Newton la lumière est composée de particules dont les différentes masses provoquent des sensations colorimétriques distinctes sur la rétine. A partir de ce concept, il établit logiquement que chaque particule se propage de façon rectiligne et peut être réfléchie selon les lois de Snell-Descartes.

La réflexion est assimilée à un rebond des particules sur la surface des objets. Les particules sont considérées comme des projectiles qui rebondissent sur les objets en un seul point.

D'après la mécanique Newtonienne, la déviation d'un rayon incident résulte de l'action d'une force sur les particules. Cette force est perpendiculaire à la surface de réflexion au point d'incidence et le rayon réfléchi est symétrique au rayon incident.

Pour ce qui est de la réfraction, Newton utilise les règles de la balistique des projectiles dans lesquelles la réfraction est régie par la vitesse des corpuscules. Plus la densité du milieu est importante, plus la vitesse des particules lumineuses augmente. Ce concept sert aussi de base à l'explication du phénomène de diffraction. La densité d'une surface définit l'interaction avec les particules lumineuses selon la force exercée perpendiculairement à la surface du milieu. Dans le cas d'un changement de milieu air/eau, il considère que la vitesse des particules lumineuses augmente et leurs trajectoires dévient car l'eau est plus dense que l'air.

Dans sa théorie achevée, Newton affirme que la lumière est composée de corpuscules de masses différentes qui se propagent dans le vide à une vitesse très élevée. Lorsqu'ils atteignent la surface d'un milieu, ces corpuscules subissent l'action d'une force réfringente perpendiculaire à la surface et proportionnelle à la densité du milieu. Cette force dévie la trajectoire des corpuscules et cause la réflexion, la diffraction, la dispersion et la réfraction. Avec cette dernière formulation de sa théorie, il explique le principe de sa célèbre expérience de diffraction par un prisme (figure 2.1). A travers son traité d'optique en 1704 [N+22], il démontre que la lumière blanche est composée de plusieurs corpuscules de masses différentes (les couleurs). Lorsque la lumière traverse le prisme, chaque corpuscule subit la gravité des particules qui composent le prisme, cela entraîne la déviation en fonction des lois de la gravité.



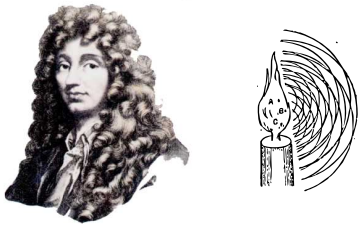
(a) (b)

Figure 2.1 – Newton *(a), diffraction par le prisme (b).*

Bien que la théorie de la gravitation soit incontestable, l'aspect corpusculaire de la lumière souffre d'incohérence et de confusion. En effet, d'après les opposants de la théorie corpusculaire, certaines observations ne sont pas explicables tels que le croisement de deux rayons lumineux qui ne crée pas de déviation ou comment un même milieu peut produire soit la réflexion, soit la diffraction alors que la densité reste identique. Malgré l'ensemble de ces incohérences, Newton par sa suprématie dans le monde scientifique a réussi à imposer ses idées durant un siècle et ce aux dépens de la théorie ondulatoire.

2.1.2 Théorie ondulatoire

Parallèlement à Newton, Huygens et ses partisans défendent l'aspect ondulatoire de la lumière. En 1667 dans [Huy67], la lumière apparait comme une vibration qui se transmet de proche en proche dans un milieu appelé Éther. L'Éther contenu dans tout l'univers serait un milieu fluide constitué de particules invisibles qui oscilleraient et transporteraient la lumière. L'oscillation de ces particules forme des ondes (figure 2.2) qui se propagent dans toutes les directions de façon circulaire.



(a) (b)

FIGURE 2.2 – HUYGENS *(a), schéma (b) de déplacement des ondes lumineuses. (schéma 4, Traité de la lumière p. 19, 1667).*

HUYGENS ébranle fortement la théorie corpusculaire en se basant sur les incohérences. Selon les théories Newtoniennes, le fait que deux faisceaux lumineux se croisent devrait entraîner une déviation car l'inertie des particules croisent une infinité d'autres trajectoires. Or cette expérience est sans impact visuel. Pour expliquer ce phénomène, HUYGENS se base sur sa théorie ondulatoire et l'existence de l'Éther. Afin de tenter d'imposer sa théorie à la communauté scientifique, il effectue une expérience avec le pendule de NEWTON dont il considère les boules métalliques comme des particules d'Éther. D'après les observations, lorsqu'une boule est soulevée puis lâchée d'un côté du pendule la boule opposée se soulève car il y a propagation d'une onde longitudinale et l'énergie cinétique est transmise. Mais si les boules aux deux extrémités sont soulevées puis lâchées, elles rebondissent en deux mouvements contraires et les forces exercés s'annulent. Pour HUYGENS, c'est selon ce principe que deux ondes lumineuses peuvent se croiser sans s'affaiblir ni dévier. Et contrairement à NEWTON, il considère que le changement de milieu air/eau entraîne une diffraction par la diminution de la vitesse des ondes.

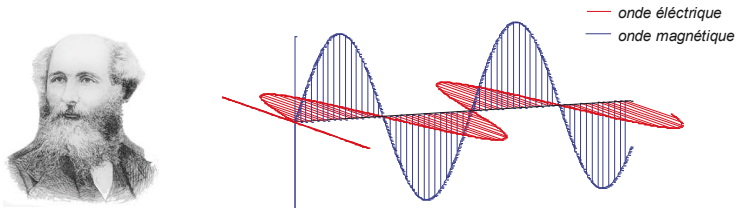
Durant le 18ème siècle, YOUNG, MALUS, FRESNEL et ARAGO renforcent la théorie ondulatoire par la découverte des phénomènes d'interférence et de polarisation. Mais de nombreux savants tel que BIOT, LAPLACE, POISSON restent partisans de la théorie corpusculaire. Pour trancher, ARAGO propose de comparer la vitesse de la lumière lors du changement de milieu air/eau. Il utilise les dispositifs de laboratoire révolutionnaires (figure 2.3) mis au point par FOUCAULT et FIZEAU qui permettent de mesurer la vitesse de la lumière par l'utilisation de miroirs ou de roues dentées.



(a) (b) (c)

FIGURE 2.3 – *Le dispositif de* FOUCAULT *composé d'un système de turbine (a)à vapeur et de miroirs rotatifs (b), le dispositif de* FIZEAU *(c) avec sa roue dentée (« Ecole Polytechnique Inventaire général. M. Lebée. ») .*

Cette expérience cruciale est réalisée en 1850 et le verdict donne raison à HUYGENS car la vitesse de la lumière mesurée dans l'eau est plus faible que dans l'air. Ce triomphe est parachevé en 1865 par la synthèse de l'électricité et du magnétisme réalisée par MAXWELL [Max65] . Il conclut que la lumière est le résultat de l'addition entre une onde électrique et une onde magnétique (figure 2.4).

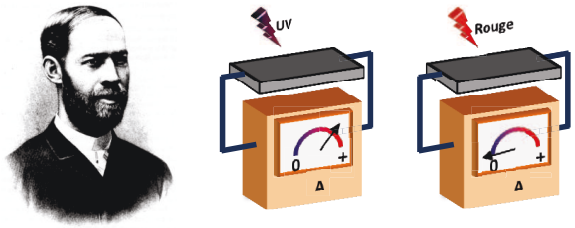


(a) (b)

FIGURE 2.4 – *Maxwell (a), la lumière est une onde électromagnétique (b) « A dynamical theory of the electromagnetic field » 1865.*

2.1.3 Complémentarité des deux théories

Malgré le triomphe de la théorie ondulatoire, le débat n'est pas définitivement achevé. De nouveaux phénomènes (figure 2.5) sont mis en évidence par HERTZ en 1885. Il observe qu'une plaque métallique exposée à une lumière ultraviolette produit de l'électricité, alors que la même plaque exposée à une lumière rouge ne produit absolument rien même après augmentation de l'intensité lumineuse. Ce phénomène n'a aucun sens si la lumière est une onde transportant de l'énergie.



(a) (b)

FIGURE 2.5 – *Hertz (a) réalise en 1885 l'expérience photo-électrique (b).*

Inspiré par les travaux de PLANCK [Plao1], EINSTEIN valorise de nouveau la théorie corpusculaire. En 1905 dans [Eino5], il annonce que la lumière est en réalité composée de Quanta baptisée plus tard Photons. Les Quanta sont des particules sans masse dont l'énergie dépend de la fréquence de l'onde associée.

$$e_{photons} = h\frac{c}{\lambda} \tag{2.1}$$

avec :

h	:	constante de Planck.
c	:	célérité (vitesse de la lumière dans le vide).
λ	:	la longueur d'onde.

Selon son interaction avec la matière, la lumière se manifeste soit comme une onde, soit comme des corpuscules. Cette théorie répond au phénomène mis en évidence par HERTZ.

Cette dualité onde-corpuscule est généralisée en 1924 par DE BROGLIE dans [DB24]. Elle met fin au débat sur la nature de la lumière lancé environs trois siècles auparavant.

2.1.4 Le spectre lumineux

D'après la dualité onde corpuscule, la lumière est caractérisée par sa longueur d'onde et l'énergie des photons qu'elle transporte. Comme toute onde électromagnétique, elle possède un spectre qui lui est propre.

Pendant le débat sur sa nature, le premier à caractériser la lumière est Newton grâce à son prisme. Il observe que la lumière du soleil peut être décomposée par diffraction en sept couleurs. Ces couleurs vont du rouge au violet avec un passage par la couleur indigo. Cette couleur est inexistante car elle a été ajoutée par Newton pour parvenir à sept, le chiffre énigmatique du 18ème siècle.

En 1800, Hershel place des thermomètres dans chaque couleur et par surprise il s'aperçoit que la température est différente. De plus, il observe qu'au-delà du rouge là où l'œil ne voit plus rien, la température est plus élevée. Il vient de découvrir la première lumière invisible. Un an plus tard, le chimiste Riter expose au spectre solaire une plaque photographique recouverte de chlorure d'argent et constate qu'elle réagit particulièrement au-delà du violet. Cette expérience lui permet de découvrir la seconde lumière invisible, les ultra-violets. La même année, Young offre une graduation pour caractériser la lumière. Il interprète les couleurs comme une manifestation de la longueur d'onde grâce à la célèbre expérience des franges d'interférence. Cette expérience consiste à faire interférer deux faisceaux lumineux issus d'une même source par l'intermédiaire de deux petits trous percés dans un plan opaque. Elle permet de mettre en évidence le déphasage entre les faisceaux et de mesurer les longueurs d'onde de chaque couleur qui sont de l'ordre du micron.

Le spectre électromagnétique s'enrichit des ondes radio à la suite des travaux de Hertz en 1885, puis des rayons-X grâce à Conrad en 1895. La découverte de la radioactivité permet d'ajouter les rayons gamma en 1914 et plus récemment la frontière floue entre les ondes radios et l'infrarouge est comblée par la découverte des micro-ondes.

Dans l'ensemble du spectre électromagnétique (figure 2.6), l'œil humain ne perçoit qu'une infime partie dont les longueurs d'ondes sont comprises entre 380 et 780 *nanomètre*. Ces valeurs ont évolué au cours du temps et sont aujourd'hui normalisées par la CIE (Commission Internationale de l'Eclairage).

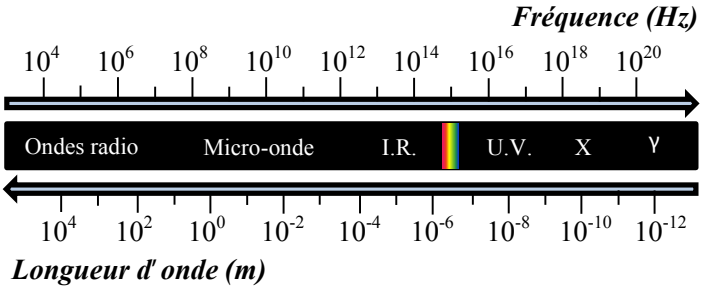


Figure 2.6 – *Spectre électromagnétique.*

La technologie créée par l'homme permet de pallier le manque de sensibilité du SVH (Système Visuel Humain). En effet, des capteurs ultra-sensibles permettent d'acquérir la lumière visible mais aussi invisible car selon la nature des sources lumineuses, le rayonnement produit peut s'étendre dans l'infrarouge ou l'ultra-violet.

2.2 Les sources lumineuses

Le terme source lumineuse désigne tous dispositifs naturels ou artificiels capables de transformer (figure 2.7) une énergie primaire en un flux énergétique rayonnant non nul dans le spectre visible.

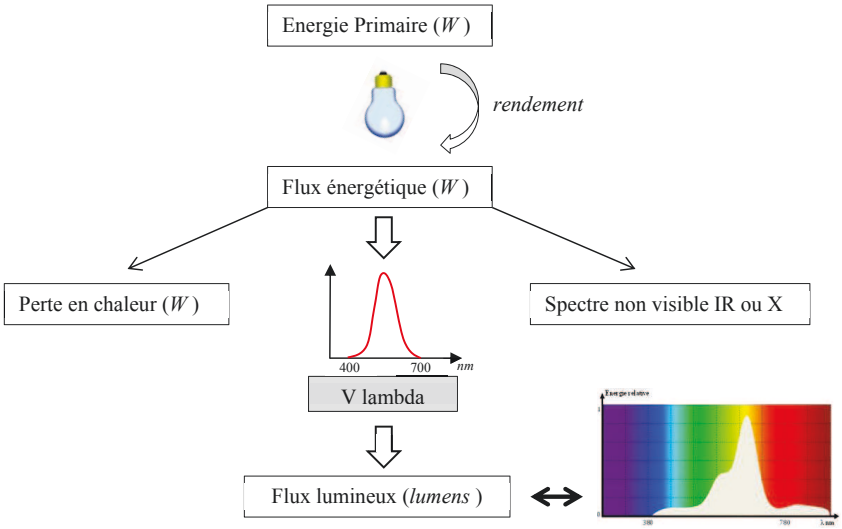


FIGURE 2.7 – *Echange energétique source lumineuse.*

La plus ancienne et principale source lumineuse est le soleil. Le rayonnement perçu étant cyclique, l'être humain a toujours désiré prolonger et contrôler la durée des cycles d'éclairage. Pour cela, il a cherché à compléter le soleil par d'autres sources lumineuses telles que les flammes dans un premier temps. Puis grâce à l'évolution technologique sont apparus l'ensemble des ampoules (incandescentes, tungstène) ou encore les tubes (néons) dont la source d'énergie primaire est l'électricité.

2.2.1 Classification des sources lumineuses

Les diverses sources lumineuses peuvent être classées en fonction de leur technologie, de l'énergie primaire, de la température du rayonnement produit ou encore de l'efficacité lumineuse. Afin de faciliter la classification, l'intérêt est porté uniquement sur la lumière produite. Ainsi, elles sont classables en fonction de leurs distributions spectrales (figure 2.8) suivant trois catégories : spectre continu, spectre discret et spectre mixte.

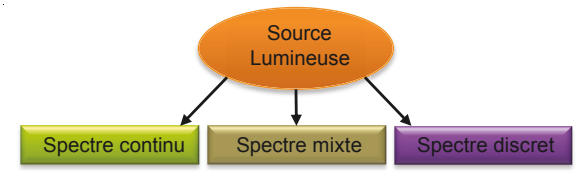
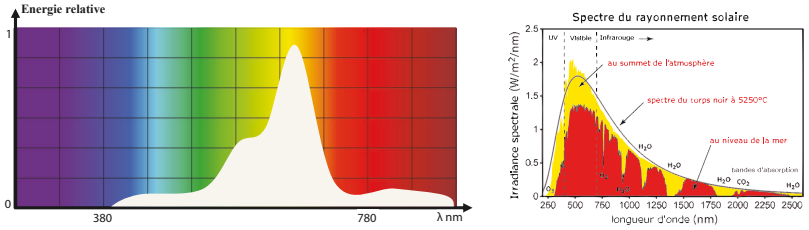


FIGURE 2.8 – *Classification des sources lumineuses.*

2.2.1.1 Sources à spectre continu

Le rayonnement produit par ce type de source est constitué (figure 2.9a) d'un ou plusieurs intervalles continus de longueur d'onde. Cette catégorie regroupe l'ensemble des sources émettant de l'énergie lumineuse sous l'effet de la chaleur (lampe à incandescence classique,

flamme) ou du soleil. Par exemple, le soleil possède une distribution spectrale d'énergie couvrant l'ensemble des longueurs d'ondes du spectre visible mais une grande partie se situe dans le spectre non visible (figure 2.9b), notamment dans les ultra-violets et surtout dans l'infra-rouge. L'amplitude énergétique n'est pas identique pour l'ensemble des longueurs d'onde. Pour le spectre visible, l'intensité maximum se situe autour de 500 *nm*.

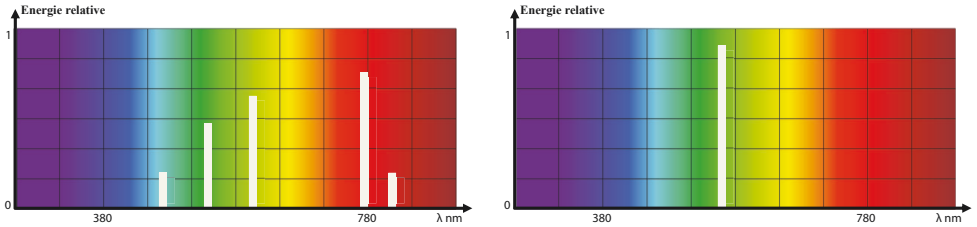


(a) (b)

FIGURE 2.9 – *Exemple de source à spectre continue (a), spectre du rayonnement solaire (b).*

2.2.1.2 Sources à spectre discret

Cette catégorie regroupe les sources constituées (figure 2.10a) d'un ensemble limité de rayonnement dont les longueurs d'onde sont déterminées. Ces caractéristiques spectrales se retrouvent principalement pour la technologie des lampes à décharge. Le cas particulier (figure 2.10b) où le spectre est composé d'une unique longueur d'onde se retrouve pour le cas de la technologie laser. Ces sources sont qualifiées de sources monochromatiques.



(a) (b)

FIGURE 2.10 – *Exemple de source à spectre discret (a), spectre d'une source monochromatique (b) par exemple le laser.*

2.2.1.3 Sources à spectre mixte

Les sources à spectre mixte possèdent une distribution spectrale (figure 2.11) qui regroupe les caractéristiques des sources à spectre continu et à spectre discret. Leur spectre est continu et certaines longueurs d'onde présentent des raies de forte intensité. La technologie représentative est typiquement le tube fluorescent.

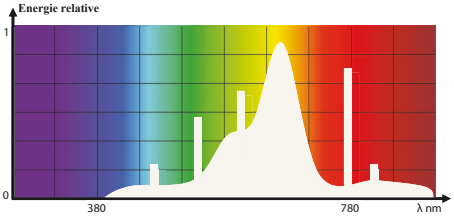


FIGURE 2.11 – *Exemple de source à spectre mixte.*

2.2.2 Métrologie : définitions des valeurs

La classification des sources lumineuses nécessite la mesure et l'analyse énergétique spectrale. Ces mesures sont obtenues par l'intermédiaire des valeurs métrologiques telles que la radiométrie et la photométrie.

2.2.2.1 La radiométrie

La radiométrie consiste à mesurer le rayonnement produit par les ondes électromagnétiques sans se limiter à la lumière visible. Les unités de mesure qui la composent sont unifiées par le système de norme internationale SI (Système International). L'unité de base est le flux énergétique exprimé en Watt. L'ensemble des termes regroupés et décrits dans ce paragraphe proviennent des ouvrages de référence [IoNA87] et [ML99].

Le flux énergétique

Le flux énergétique (Fe ou Φe) correspond au débit d'énergie transporté par un rayonnement à l'instant t suivant la célérité c. Ce rayonnement est considéré comme la superposition d'ondes planes quasi monochromatique (unique longueur d'onde) appelée mode dont la répartition en fréquence de vibration et en directions de propagation dépend de sa complexité. D'après les travaux de Maxwell, la densité d'énergie électromagnétique dans le vide est caractérisée par un vecteur de champs électrique $\vec{E}$ et un vecteur de champs magnétique $\vec{H}$ tel que :

$$\begin{aligned} \vec{E}(\vec{r},t) &= \vec{E}_0 \cos(\vec{k}.\vec{r} - \omega t) \\ \vec{H}(\vec{r},t) &= \vec{H}_0 \cos(\vec{k}.\vec{r} - \omega t) \end{aligned} \tag{2.2}$$

avec :

ω	:	$(2\pi c/\lambda)$ la pulsation.
$\vec{r}$	:	le vecteur position du point considéré.
$\vec{k}$	:	est le vecteur d'onde.

Le flux énergétique d'un mode est proportionnel à la valeur moyenne temporelle de son vecteur de Poynting :

$$< |S| >= \frac{|\vec{E}_0|^2}{2Z} \tag{2.3}$$

avec :

Z	:	l'impédance du milieu de propagation.

La densité du flux

La densité du flux est le flux énergétique par unité d'aire à un point d'une surface réelle ou imaginaire. L'unité de mesure est le Watt par mètre carré ($W.m^{-2}$) et deux cas sont possibles. Le flux peut être reçu par la surface depuis n'importe quelle direction (figure 2.12a), dans ce cas la densité du flux énergétique est appelée Irradiance Ee et peut être définie par :

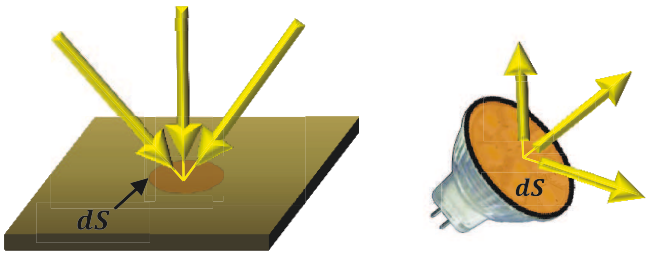
$$Ee = \frac{\mathrm{d}Fe}{\mathrm{d}S} \tag{2.4}$$

où $\mathrm{d}Fe$ est le flux énergétique reçu à un point et $\mathrm{d}S$ est l'aire de la surface qui entoure le point.

Le flux énergétique peut aussi être émis ou réfléchi par la surface. Dans ce cas (figure 2.12b), la densité du flux émise dans n'importe quelle direction est appelée Exitance Me et peut être définie par :

$$Me = \frac{\mathrm{d}Fe}{\mathrm{d}S} \tag{2.5}$$

où $\mathrm{d}Fe$ est le flux énergétique émis à un point et $\mathrm{d}S$ est l'aire de la surface qui entoure le point.



(a) (b)

FIGURE 2.12 – *(a) irradiance Ee, (b) existance Me.*

L'unité stéradian

Pour le moment, seules les notions temporelles et spatiales (la surface) ont été abordées car le flux provient de toutes les directions possibles. Maintenant si le flux provient d'une direction précise, la notion spatiale angulaire doit être considérée. Les unités radiométriques sont alors l'Intensité et la Radiance. Ces mesures font intervenir une unité appelée le stéradian.

Le stéradian au même titre que le radian est une unité SI pour définir un angle solide (couramment noté Ω), cette mesure est utilisée pour le calcul tridimensionnel. Afin de déterminer la valeur d'un angle solide, il est nécessaire d'utiliser une sphère de rayon unitaire : 1 mètre. Prenons une surface (figure 2.13) quelconque, projetons la sur une sphère unitaire nous obtenons une surface, l'angle solide est le rapport entre la surface occupée sur la sphère et le rayon de la sphère au carré.

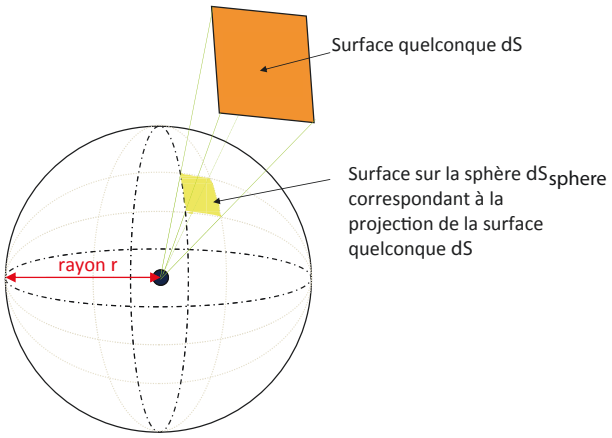


FIGURE 2.13 – *Schéma angle solide.*

$$\Omega = \frac{dS_{sphere}}{r^2} \tag{2.6}$$

avec :

dS_{sphere}	:	la surface projetée. sur la sphère unitaire.
r	:	le rayon de la sphère. unitaire $(= 1)$.

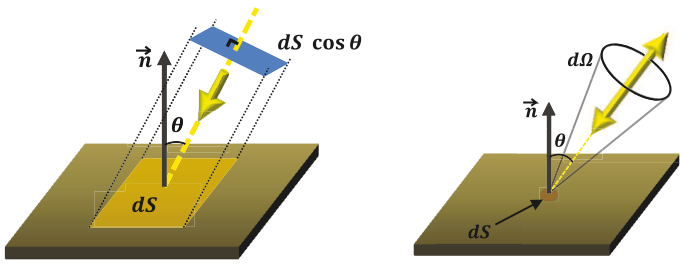
La radiance

Pour définir le terme Radiance, il est nécessaire d'imaginer un rayon émis ou reçu par un point sur une surface dans une direction donnée (figure 2.14a). La Radiance est tout simplement la quantité infinitésimale de flux énergétique contenue dans ce rayon. L'angle d'incidence θ entre le rayon et la normale à la surface doit être pris en compte. La surface dS est projetée suivant cet angle et devient $dS \cos\theta$.

Une définition plus formelle décrit un rayon comme un cône élémentaire infiniment étroit dont le sommet est un point sur une surface réelle ou imaginaire (figure 2.14b). Ce cône a un angle différentiel solide $d\Omega$ dont l'unité est le stéradian. Le calcul de la Radiance Le émise ou reçue, mesurée en Watt par mètre carré par stéradian ($W.m^{-2}.sr^{-1}$) peut être définie par :

$$Le = \frac{d^2Fe}{dS(d\Omega\ \cos\theta)} \tag{2.7}$$

où dFe est le flux énergétique émis à un point, dS est l'aire de la surface qui entoure le point, $d\Omega$ est l'angle solide d'un cône élémentaire et θ est l'angle d'incidence entre le rayon et la normale à la surface.



(a) (b)

FIGURE 2.14 – *La radiance Le.*

L'intensité

L'Intensité Ie est une mesure ponctuelle et quantifie le flux énergétique émis en un point et dans une direction donnée le long du centre d'un cône englobant un angle solide (figure 2.15). Cette mesure énergétique est exprimée en Watt par stéradian ($W.sr^{-1}$) et est exprimée par :

$$Ie = \frac{\mathrm{d}Fe}{\mathrm{d}\Omega} \tag{2.8}$$

où $\mathrm{d}Fe$ est le flux énergétique émis en un point, $\mathrm{d}\Omega$ est l'angle différentiel solide du cône élémentaire dans une direction donnée.

Pour une source isotrope, le flux rayonnant est émis dans toutes directions et l'intensité du rayonnement est égale à $Fe/4\pi$.

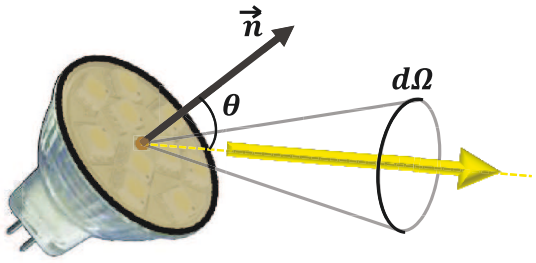


FIGURE 2.15 – *Intensite Ie d'une source.*

Radiométrie et spectre

Dans le cas des sources à spectre continu tel que le soleil, le rayonnement électromagnétique s'étend des ondes radio aux rayons gamma. L'énergie de rayonnement spectrale est la quantité d'énergie rayonnante sur intervalle de longueur d'onde non nulle. Elle se mesure en Joule par nanomètre et peut être définie comme l'intégrale des flux énergétiques. A l'inverse, le flux énergétique peut être exprimé en nombre de photons si le rayonnement produit dans le spectre visible est monochromatique, telles les sources à spectre discret. Par la relation de PLANCK, le flux énergétique d'un rayonnement monochromatique λ correspond à un débit de photons appelé flux photonique (Fp ou Φp) et exprimé par unité de temps (s) :

$$Fp_\lambda = Fp_\lambda / e_{photons} \tag{2.9}$$

Lorsque, le rayonnement électromagnétique est mesuré et étudié uniquement pour le spectre visible, les valeurs radiométriques sont généralement converties en valeurs photométriques.

2.2.2.2 La photométrie

Le SVH n'a pas la même sensibilité selon les différentes longueurs d'onde du spectre visible. La CIE a donc établi des courbes (figure 2.16) appelées $V(\lambda)$ qui permettent de faire le lien entre la sensibilité du SVH et les mesures radiométriques. Par conséquent, la photométrie décrite dans [dl83] reprend les mêmes mesures que la radiométrie mais elles sont limitées au spectre visible et pondérées par une courbe de sensibilité. Pour plus de détails entre les unités radiométriques et photométriques, nous fournissons un tableau de correspondance en annexe A.1.

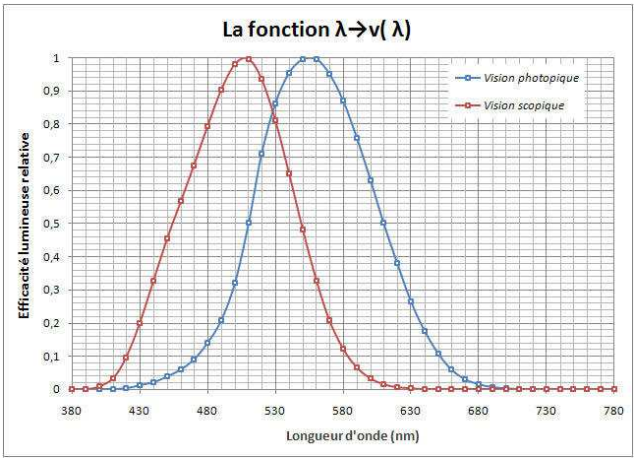


FIGURE 2.16 – *Courbes $V(\lambda)$ definies par le CIE.*

Le flux lumineux

L'ensemble des mesures photométriques est basé sur le flux lumineux (Fv ou Φv), dont l'unité est le lumen (lm). Il est défini à partir de son homologue radiométrique le flux énergétique, puis pondéré par la courbe de sensibilité et intégré sur l'ensemble du spectre visible.

$$Fv = \int_{380}^{780} Km\ V(\lambda)\ Fe(\lambda) \quad \mathrm{d}(\lambda) \tag{2.10}$$

où Km est la constante de conversion des Watt en lumen ($= 683$), $V(\lambda)$ est la courbe de sensibilité pour la valeur λ et $Fe(\lambda)$ est le flux énergétique exprimé en Watt pour la longueur d'onde λ.

Autres mesures photométriques

L'Exitance lumineuse Mv et l'Éclairement lumineux Ev sont respectivement la quantité de flux émise et reçue en fonction d'une surface $\mathrm{d}S$. Ces valeurs sont exprimées en lumen par mètre carré ($lm.m^{-2}$) ou directement en lux (lx). Comme pour leurs homologues radiométriques, il n'y a pas de notion de direction dans ces mesures, c'est-à-dire que le flux lumineux est omnidirectionnel.

$$Mv = \frac{\mathrm{d}Fv}{\mathrm{d}S} \qquad \text{et} \qquad Ev = \frac{\mathrm{d}Fv}{\mathrm{d}S} \tag{2.11}$$

Comme pour l'Intensité radiométrique, le flux lumineux Fv est mesuré sur un angle solide. La mesure obtenue est l'Intensité lumineuse Iv dont l'unité est le lumen par stéradians ($lm.sr^{-1}$), plus couramment appelé candéla (cd).

La dernière mesure à définir est la Luminance Lv, pendant photométrique de la Radiance pondérée par la courbe de sensibilité. C'est la seule mesure qui est perceptible par le SVH. Pour être plus exact la Luminance est la mesure approximative du degré de luminosité sur une surface. Il existe une formule qui lie directement la Radiance à la Luminance :

$$Lv = \int_{380}^{780} Km\ V(\lambda)\ Le(\lambda) \quad \mathrm{d}(\lambda) \tag{2.12}$$

Il existe tout de même une ambiguïté, car pour une même valeur de Luminance Lv plusieurs valeurs de Radiance $Le(\lambda)$ peuvent correspondre du fait de l'utilisation de pondération générée par la courbe $V(\lambda)$. Ce phénomène est appelé le méta-mérisme.

2.2.3 Métrologie : bilan

D'après les définitions de la mesure radiométrique et photométrique, aucun terme n'est attribué à un flux dont le rayonnement est concentré dans le spectre visible ou au-delà et observé par une sensibilité différente du SVH. Par conséquent dans le cadre de ce document consacré à l'acquisition et au traitement des images numériques, nous utilisons par abus de langage les termes SI de la radiométrie. Il est tout de même nécessaire de conserver à l'esprit que ces termes sont intégrés et pondérés suivant la sensibilité des systèmes d'acquisition proche du SVH. De plus, les valeurs mesurées sont dépendantes des sources lumineuses et de l'interaction avec la scène observée.

2.3 Interactions physiques lumière matière

Selon les propriétés de "la matière éclairée", l'interaction globale (figure 2.17) est une succession de phénomènes pouvant être indexés selon leur nature spectrale et géométrique.

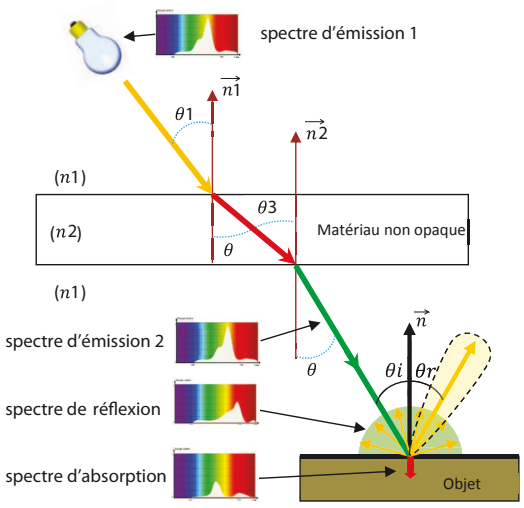


FIGURE 2.17 – *Exemple complexe d'un éclairage artificiel.*

Pour comprendre ces phénomènes, il convient de les décomposer suivant la géométrie optique (émission, transmission, réflexion) et leurs caractéristiques spectrales (spectre d'émission, spectre d'absorption et spectre de réflexion).

2.3.1 L'absorption

Pour certains matériaux, les photons reçus sont absorbés et transformés en chaleur. Cette propriété va permettre de percevoir ou de capturer les objets en couleur. Dans le cas où un objet reçoit une lumière à spectre continu dont l'amplitude énergétique est constante pour l'ensemble des longueurs d'ondes, si les propriétés physiques du matériau composant l'objet font en sorte que l'ensemble des longueurs d'ondes soit totalement absorbé, il est alors perçu par le SVH comme noir. Par contre si les propriétés physiques du matériau font en sorte qu'il n'absorbe qu'une certaine plage de longueurs d'ondes, les autres plages sont réfléchies et l'objet est perçu de la couleur de ces plages. Le principe d'absorption est la base de la capture d'image.

2.3.2 Spectre de réflectance

Il est possible de caractériser une surface en fonction des longueurs d'ondes qu'elle absorbe et/ou réfléchie (figure 2.18). Pour cela, il est nécessaire de déterminer son spectre de réflectance SRr (longueur d'onde réfléchie et absorbée). De même, il est possible d'obtenir le spectre de radiance de la source lumineuse SRl afin de calculer celui de la radiance réfléchie par l'objet SRf :

$$SRf(\lambda) = SRl(\lambda)\ SRr(\lambda) \tag{2.13}$$

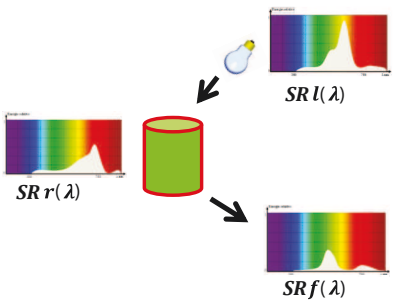
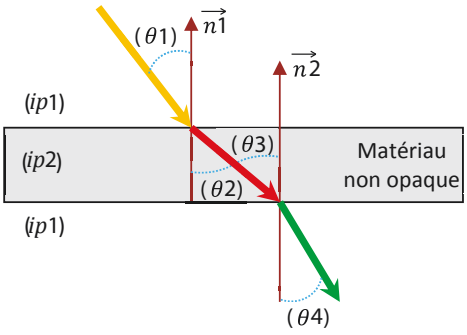


FIGURE 2.18 – *Spectre de réflectance.*

Il faut noter que cette équation présente plusieurs limites car elle ne prend pas en compte l'aspect géométrique (angle d'incidence de la lumière ou d'observation). Par exemple, dans le cas d'une réflexion spéculaire, il est obligatoire de considérer l'aspect géométrique afin de permettre la modélisation de l'interaction entre la lumière et l'objet. Pour tenir compte de ces effets, la fonction de réflectance spectrale bidirectionnelle (BRDF : Bidirectional Reflectance Distribution Function) definie par NICODEME, WYSZECKI, STILES ou encore SOUAMI doit être utilisée. Toutefois le modèle de réflectance spectrale, en tenant compte des limitations, permet une analyse simple dans beaucoup de cas.

2.3.3 La transmission

Lorsqu'un matériau est non opaque (verre et certains polymères), il laisse passer un certain nombre de photons et la lumière est transmise au travers. Lors du changement de milieu, passage de l'air au matériau et inversement, un phénomène de réfraction se produit. La loi de Snell-Descartes permet de modéliser (figure 2.19) ce phénomène en liant les indices de réfraction $ip1$ et $ip2$ de chacun des milieux et les angles : incident $\theta 1$ et réfracté $\theta 2$.



$$ip1 \times \sin\theta 1 = ip2 \times \sin\theta 2$$

$$ip2 \times \sin\theta 3 = ip1 \times \sin\theta 4$$

$$\theta 2 = \theta 3$$

$$ip1 \times \sin\theta 1 = ip2 \times \sin\theta 2 = ip1 \times \sin\theta 4$$

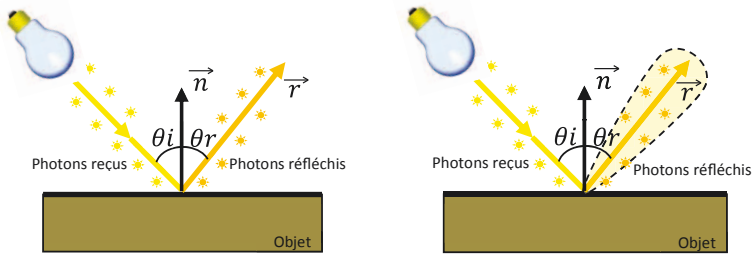
Figure 2.19 – *La transmission.*

2.3.4 La réflexion

La réflexion parfaite (figure 2.20a) appelée aussi réflexion spéculaire répond au principe optique selon les lois édictées par Snell-Descartes. L'angle d'incidence est de même valeur que l'angle réfléchi :

$$\theta i = \theta r \tag{2.14}$$

Mais certains matériaux ont une surface qui présente de légères irrégularités microscopiques. Ainsi, la réflexion se fait dans une unique direction en respectant le principe optique de réflexion, mais elle n'est pas parfaite. La lumière provenant de la source lumineuse est réfléchie dans la direction d'un vecteur de réflexion $\vec{r}$, suivant un coefficient de diminution exponentiel n_s et un coefficient d'absorption K_s. Cela forme un lobe (figure 2.20b) de réflexion spéculaire.

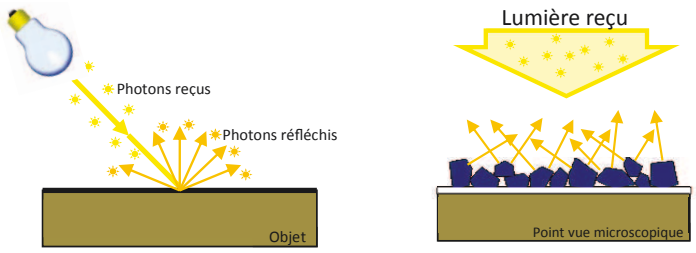


(a) (b)

Figure 2.20 – *Schéma d'une reflexion parfaite (a), schéma d'une reflexion imparfaite (b).*

2.3.5 La diffusion

La diffusion est un cas particulier de réflexion, elle se produit lorsque la surface du matériau qui compose l'objet est de nature très irrégulière. La lumière est réfléchie de manière uniforme (figure 2.21a) dans toutes les directions. Ce principe repose sur le facteur d'échelle d'observation, par exemple un point à l'échelle macroscopique provoque une diffusion si à l'échelle microscopique (figure 2.21b) il présente une forte irrégularité de surface. Dans ce cas, l'intensité du flux énergétique émis par un point de la surface est indépendante de la direction d'observation et est uniquement fonction du flux d'énergie incident à cette surface.



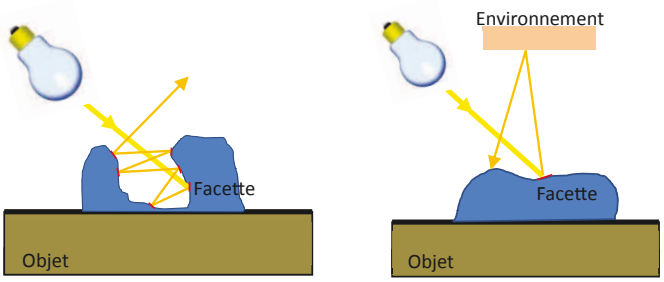
(a) (b)

FIGURE 2.21 – *Schéma d'une diffusion à l'échelle macroscopique (a), schéma d'une diffusion à l'échelle microscopique (b).*

2.3.6 Autres phénomènes

En complément des interactions spectrales et géométriques basiques, il existe d'autres phénomènes d'interaction plus complexes tels que :

– Inter-réflexion : lorsqu'un flux énergétique est réfléchi par une facette d'une surface, les photons réémis interagissent avec d'autres facettes de la surface et se comportent comme des sources d'éclairage secondaires. Ce phénomène est appelé inter-réflexion directe (figure 2.22a) mais il existe aussi le cas indirect (figure 2.22b) où les photons interagissent avec l'environnement proche de la surface éclairée.



(a) (b)

FIGURE 2.22 – *Schéma d'une inter-reflexion directe multiple (a), schéma d'une inter-reflexion indirecte (b).*

– Ombres : lorsqu'un objet opaque est éclairé par une source de lumière, certaines zones (figure 2.23) situées derrière l'objet ne reçoivent pas de lumière et constituent l'ombre de l'objet. Deux type d'ombres se distinguent la première dite propre est la zone de l'objet

opposée à la direction de la source lumineuse. La seconde appelée ombre portée se situe sur une surface occultée par l'objet. Elles possèdent une forme corrélée aux contours de l'objet et à la géométrie de la source lumineuse.

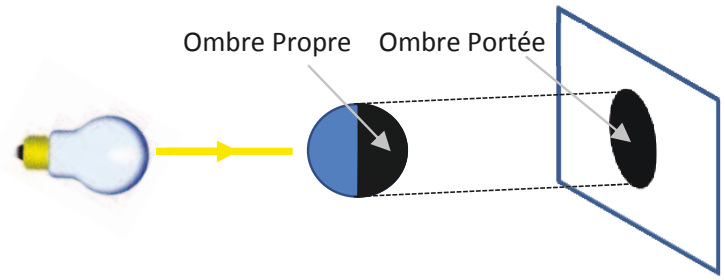


FIGURE 2.23 – *Schéma illustrant le principe de l'ombre propre et de l'ombre portée.*

2.4 Modélisation mathématique

Des modèles mathématiques permettent de modéliser le comportement de la lumière lorsqu'elle interagit avec les surfaces des matériaux. Les modèles de base (LAMBERT et PHONG) sont les plus simples mais offrent une modélisation limitée. Ils permettent de modéliser respectivement la diffusion et la réflexion. Mais il faut garder à l'esprit qu'un grand nombre de modèles existe et que certains plus complexes combinent l'ensemble des interactions possibles.

2.4.1 Le modèle de Lambert

Le modèle Lambertien formulé par LAMBERT en 1760 [Lam60] est le plus simple pour décrire l'interaction de la lumière sur des surfaces parfaitement diffuses. Ce modèle (figure 2.24) suppose que la surface présente une réflectance homogène et considère que l'intensité réfléchie est identique quelle que soit la direction de l'observateur. De plus, les phénomènes d'ombres et d'inter-réflexions sont ignorés.

A partir de ces hypothèses et pour un éclairage directionnel, l'intensité réfléchie par un point de la surface est proportionnelle à l'angle entre la normale à la surface et la direction de la source lumineuse.

$$Ie = Ie_0\ \rho\ \cos\theta = Ie_0\ \rho\ (\vec{n}.\vec{l}) \tag{2.15}$$

avec :

Ie	:	l'intensité réfléchie.
Ie_0	:	l'intensité incidente de la source lumineuse.
$\rho(x,y)$	:	l'albédo correspondant au coefficient de réflexion.
$\vec{n}$	=	la normale à la surface.
$\vec{l}$	=	le vecteur direction de la source lumineuse.
θ	:	l'angle d'incidence de la direction de la source lumineuse.

Cette équation se retrouve couramment sous la forme discrète :

$$ie(x,y) = ie_0\ \rho(x,y)\frac{-p(x,y)\ \cos\tau\ \sin\sigma - q(x,y)\ \sin\tau\ \sin\sigma + \cos\sigma}{\sqrt{p^2(x,y)+q^2(x,y)+1}} \tag{2.16}$$

avec :

$ie(x,y)$	:	l'intensité réfléchie au point (x,y).
ie_0	:	l'intensité incidente de la source lumineuse.
$\rho(x,y)$	:	l'albédo du point (x,y).
τ,σ	:	les angles azimut et zénith de la source lumineuse.
$\vec{n}$	=	$\left[\frac{-p}{\sqrt{p^2+q^2+1}}\ \frac{-q}{\sqrt{p^2+q^2+1}}\ \frac{1}{\sqrt{p^2+q^2+1}}\right]$ la normale.
p	=	$[\mathrm{d}S(x,y)/\mathrm{d}x]$ la dérivée partielle de la surface s suivant x.
q	=	$[\mathrm{d}S(x,y)/\mathrm{d}y]$ la dérivée partielle de la surface s suivant y.
$\vec{l}$	=	$[\cos\tau\ \sin\sigma \quad \sin\tau\ \sin\sigma \quad \cos\sigma]$ la direction source.

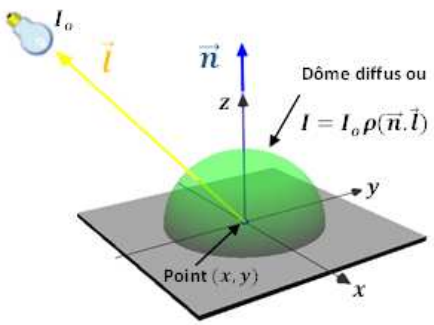


FIGURE 2.24 – *Modèle diffus de* LAMBERT.

L'angle d'incidence est important (figure 2.25), lorsqu'il est en alignement avec la normale de la surface (zénith $\theta = 0^o$) l'intensité reçue est conservée, dans le cas où l'angle a une incidence ($0^o < \theta < 90^o$) l'intensité reçue est dispersée. Par contre si l'angle a une incidence nulle ($\theta > 90^o$), l'intensité reçue est nulle sur la surface au point (x,y).

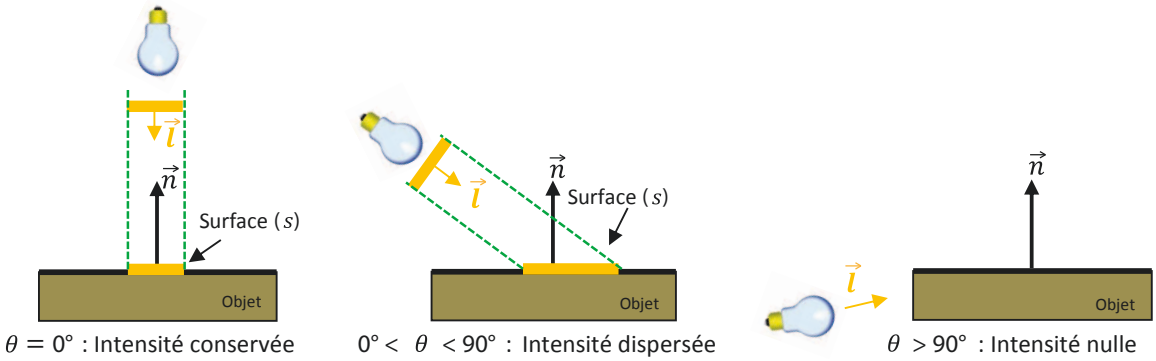


FIGURE 2.25 – *Loi du cosinus de* LAMBERT.

Dans le cas de plusieurs sources lumineuses, il est possible d'obtenir le modèle Lambertien par sommation de toutes les sources en fonction de leurs intensités et leurs angles d'incidence. Cette souplesse permet aussi la prise en compte de l'intensité réfléchie de la lumière ambiante si elle n'est pas nulle :

$$Ie = \sum_{s=1}^{S} Ie_0(s)\ \rho\ \cos\theta(s) + Ie_{ambiante} \tag{2.17}$$

Ce modèle présente tout de même certaines limitations, Wolff a démontré qu'il ne peut être appliqué que si l'angle d'incidence est faible. Dans ces travaux [Wol96], il propose une modification simple qui permet de prendre en compte la position d'observation. D'autres contributions ont tenté d'améliorer le modèle, par exemple Oren et Nayar dans [ON95] ont proposé des modifications plus importantes pour inclure le cas des surfaces imparfaitement diffuses. D'après leurs travaux, la réflexion d'une surface dépend de l'angle d'incidence et de la direction d'observation. Le modèle proposé prend en compte les géométries complexes et exclut les phénomènes d'ombres ou d'inter-réflexions. Mais au détriment des améliorations apportées, ce type de méthode est généralement négligé du fait de l'augmentation de la complexité. Car malgré les contraintes et les limitations du modèle original, sa simplicité et la possibilité de l'inverser font de lui le modèle le plus utilisé dans le domaine de la vision par ordinateur. Son utilisation est omniprésente dans les technique de reconstruction tridimensionnelle comme le Shape from Shading et la Stéréo-Photométrie. Il est également utilisé par les méthodes de stéréo vision, de détection de mouvement ou de correction d'image.

2.4.2 Le modèle de Phong

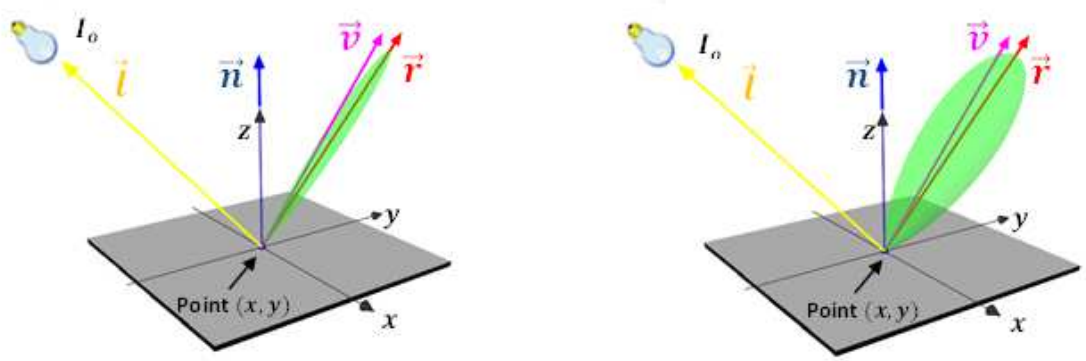
Le modèle [Pho75] proposé par Phong permet d'obtenir l'intensité lumineuse d'un point spéculaire suivant une direction d'observation particulière. L'intensité réfléchie Ie_{spec} dépend de l'angle η entre la direction d'observation $\vec{v}$ et la direction de réflexion $\vec{r}$ définie selon la loi de Snell-Descartes.

$$Ie_{spec} = Ie_0\ K_s\ \cos^{n_s}\ \eta = Ie_0\ K_s\ (\vec{r}.\vec{v})^{n_s} \tag{2.18}$$

avec :

Ie_{spec}	:	l'intensité spéculaire réfléchie d'un point.
Ie_0	:	l'intensité incidente de la source lumineuse.
K_s	:	le coefficient empirique de réflexion spéculaire du point.
$\vec{r}$	:	le vecteur unitaire de la direction de réflexion.
$\vec{v}$	:	le vecteur unitaire de la direction d'observation.
η	:	l'angle entre $\vec{r}$ et $\vec{v}$.
n_s	:	le coefficient de réflexion de la surface.

Le coefficient n_s est proportionnel à l'imperfection microscopique des points spéculaires. Il influe directement sur la taille du lobe de réflexion et par conséquent sur l'intensité observable. Dans le cas d'une surface parfaitement lisse, n_s tend vers l'infini et l'intensité réfléchie est exclusivement concentrée dans la direction de réflexion (figure 2.26a). L'intensité observée est donc nulle si $\vec{v}$ est différent de $\vec{r}$. A l'inverse (figure 2.26b), lorsque le coefficient n_s diminue,



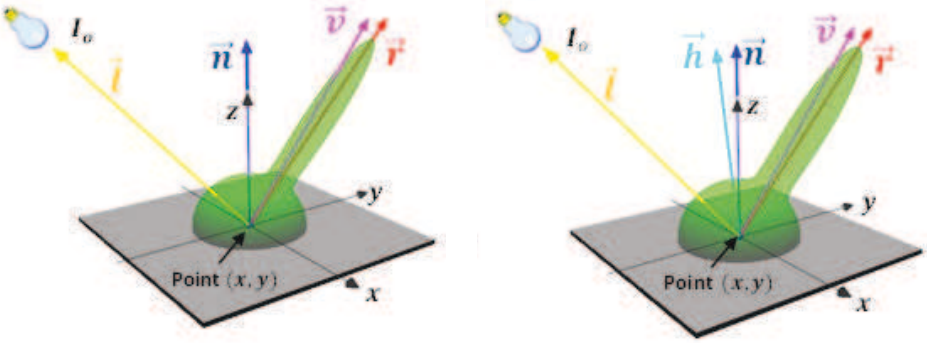
(a) (b)

FIGURE 2.26 – *Modèle de* PHONG *avec (a) n_s élevé, (b) n_s faible.*

la taille du lobe spéculaire ainsi que l'intensité spéculaire augmente et ce quelle que soit la direction d'observation. Mais la diminution du coefficient n_s présente une limite physique car la réflexion doit être considérée comme diffuse lorsque les imperfections microscopiques sont trop importantes. Pour contourner cette limite le modèle de PHONG est généralement associé à celui de Lambert (figure 2.27a).

$$Ie = Ie_{ambiante} + Ie_0 \; K_d \; \cos\theta + Ie_{spec} \tag{2.19}$$

où K_d correspond à l'albédo ρ défini par LAMBERT.



(a) (b)

FIGURE 2.27 – *Modèle de* PHONG *accompagné du modèle de* LAMBERT *(a), modèle de* PHONG *avec la modification de* BLINN*(b).*

Le calcul du vecteur $\vec{r}$ présente une complexité et un coût non négligeables mais BLINN dans [Bli77] propose une simplification. Ce modèle (figure 2.27b) est généralement appelé BLINN-PHONG et consiste à approximer la direction du vecteur $\vec{r}$ par l'ajout d'un vecteur $\vec{h}$. Le vecteur $\vec{h}$ est la normale d'une surface hypothétique entre la direction d'observation et la direction de

la source lumineuse. D'après Blinn le terme spéculaire $\cos^{n_s}\eta$ peut être remplacé par $(\vec{n}.\vec{h})^{n_s}$ et l'intensité spéculaire observée devient :

$$Ie_{spec} = Ie_0\ K_s\ (\vec{n}.\vec{h})^{n_s} \tag{2.20}$$

avec :

$\vec{n}$	:	la normale à la surface.
$\vec{h}$	=	$\frac{\vec{l}+\vec{v}}{2}$ la normale à la surface hypothétique à mi-chemin entre $\vec{l}$ et $\vec{v}$.

Ce modèle est tout de même limité, les ombres sont totalement négligées et le coefficient de réflexion de la surface doit être fixé empiriquement par l'utilisateur. L'inversion est possible mais difficile sur des surfaces à coefficient de réflexion non homogène.

2.4.3 Autres modèles

Les modèles de Lambert et de Phong sont empiriques et ne tiennent pas compte des caractéristiques surfaciques. Leur but est de fournir une modélisation simple et déconnectée des phénomènes mais qui correspondent aux données expérimentales. D'autres modèles existent, ils sont basés sur l'approche géométrique ou physique et s'appuient sur l'étude théorique de la réflexion. L'approche géométrique est basée sur l'optique géométrique et suppose que la dimension des irrégularités surfaciques est grande par rapport à la longueur d'onde du flux énergétique reçu. Cette hypothèse est valable pour des surfaces à fortes rugosités mais incompatible avec des surfaces lisses. Cette incompatibilité peut être éliminée par l'approche physique qui utilise la théorie électromagnétique pour construire un modèle plus complexe mais compatible avec les deux types de surface.

Un modèle de réflexion basé sur l'optique géométrique a été introduit par Torrance et Sparrow dans [TS67]. Le modèle proposé décompose une surface à forte rugosité en série de micro-facettes en forme de "V" dont les orientations sont aléatoires, mais dont les pentes suivent une distribution gaussienne autour de la normale principale à la surface. Il prend en compte les effets d'inter-masquages et d'ombrages entre les micro-facettes mais il exclut la présence de pics spéculaires et ne peut donc être employé pour des surfaces lisses.

Le modèle de Beckmann-Spizzichino [BS87] est basé sur les lois électromagnétiques et sur la modélisation de la surface par des micro-aspérités. La rugosité des surfaces est caractérisée par l'écart type dans une distribution normale de moyenne nulle. Un paramètre de corrélation est ajouté afin de distinguer les surfaces et leur rugosité. La réflexion spéculaire est modélisée soit par un lobe soit par un pic en fonction de la rugosité. Ce modèle est tout de même incomplet car il exclut les phénomènes d'ombres et d'inter-réflexions et nécessite des surfaces conductrices pour être mis en œuvre.

Cook et Torrance dans [CT82] se basent sur le modèle de Blinn mais reprennent l'hypothèse d'une distribution normale faite par Beckmann-Spizzichino. Cette hypothèse permet d'accélérer le processus de calcul mais dépend de la distribution spatiale de la surface.

Nayar propose dans [NIK90] un modèle hybride qui a la simplicité des modèles géométriques, ainsi que le caractère complet d'un modèle physique. Il estime l'intensité observable à partir du modèle de Lambert pour la composante diffuse et la composante spéculaire est

obtenue par mélange du modèle Torrance-Sparrow et Beckmann-Spizzichino.

Ces modelés sont principalement utilisés dans le domaine de l'informatique graphique afin d'obtenir des rendus réalistes sur des surfaces dont les caractéristiques diélectriques sont hétérogènes. Dans le cadre de l'estimation de surfaces tridimensionnelles à partir des valeurs d'intensité, l'inversion de ces modèles n'est pas courante car elle est jugée complexe et coûteuse en temps de calcul. Néanmoins, leur utilisation se révèle efficace dans le cas particulier de la reconstruction de matériaux diélectriques ou à forte réflexion spéculaire comme les métaux, les plastiques, la céramique, les verres opaques ou les peintures brillantes. Dans [HMI10], les auteurs offrent un référencement des différentes méthodes de reconstruction et les associent aux modèles utilisés. De plus, ils proposent une méthode hybride dont le modèle varie en fonction des matériaux des surfaces analysées.

Conclusion du chapitre

Ce chapitre permet de définir la nature de la lumière et l'évolution de sa caractérisation au cours des siècles. A l'heure actuelle, il s'avère que sa nature est une dualité entre onde électromagnétique et corpuscules appelés photons. Cette dualité offre des outils de mesure énergétique appelée Radiométrie et Photométrie en fonction des paramètres d'observation. Elle permet aussi d'appliquer les lois fondamentales de la géométrie optique afin de caractériser les sources lumineuses et les interactions entre la lumière et la matière. Ces interactions sont dépendantes des propriétés intrinsèques des matériaux composant les surfaces. Pour l'observateur, la surface devient une source lumineuse dont les propriétés physiques en chaque point sont mathématiquement transposables grâce aux modèles de réflectances.

Dans le contexte de nos travaux, l'objectif est de reconstruire la géométrie d'une surface à partir de l'énergie qu'elle réfléchit sans connaissances préalables des propriétés de réflectance. Pour réaliser cette tâche, le modèle mathématique utilisé doit être parfaitement inversible. La géométrie des surfaces analysées est inconnue mais présente une forte probabilité de micro rugosité au-delà de l'échelle d'observation. A partir de ces informations, nous pouvons établir que le modèle mathématique le plus adapté pour déterminer la géométrie des surface analysées est celui de Lambert.

La capture d'image 3

Sommaire

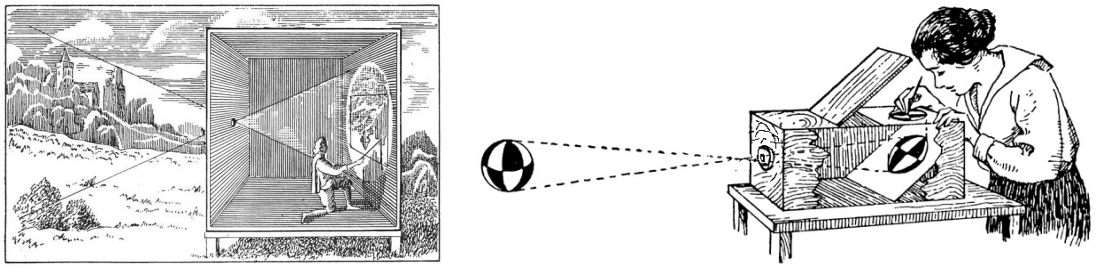
3.1 Principe et Historique . . . 30
3.2 L'objectif et sa géométrie : la perte dimensionnelle . . . 33
3.2.1 Principe de convergence et modèle de projection . . . 33
3.2.2 Le diaphragme . . . 39
3.2.3 Les défauts . . . 40
3.3 Le capteur numérique . . . 42
3.3.1 L'effet photoélectrique . . . 43
3.3.2 Le photosite . . . 44
3.3.3 Le circuit de lecture . . . 46
3.3.4 La conversion analogique numérique . . . 48
3.3.5 La dynamique . . . 51
3.3.6 Les technologies des capteurs d'images . . . 52
3.3.7 Le capteur : outil métrologique . . . 57
3.3.8 La gestion des données numériques . . . 61
Conclusion . . . 63

La vision par ordinateur nécessite l'acquisition des intensités émises par une scène. Pour répondre à cette contrainte, le principe de la capture d'image a grandement évolué depuis ses début en 300 av. JC. Un bref historique est décrit dans la première partie de ce chapitre. Puis, l'intérêt est porté sur les transformations géométriques opérées par les éléments optiques d'un dispositif d'acquisition. Enfin, la dernière partie est consacrée à la description des capteurs numériques. Chacune des étapes du processus de conversion des intensités d'une scène en valeurs numériques est décrite.

3.1 Principe et Historique

Le terme image a pour définition la représentation physique d'une scène réelle. Depuis l'aube des temps, l'être humain a cherché à reproduire son environnement comme en attestent les premières peintures rupestres. Par l'intermédiaire de ses yeux et de sa main, il a reproduit l'interaction observée entre la lumière et les objets qui l'entourent. L'évolution de la connaissance et de la technologie ont permis des reproductions plus réalistes, accessibles à tous et appelées photographies.

Du point de vue technologique, la photographie est un processus qui lie une image projetée sur une surface par un système optique et la technique qui permet la fixation permanente de l'image. Historiquement la photographie trouve ses origines 300 ans av JC grâce à Aristote. Lors d'une éclipse solaire il dessine les contours des deux astres projetés par l'intermédiaire d'un petit trou sur un mur dans une pièce noire. Ce principe fut très longtemps utilisé par les savants avant d'être correctement décrit en 1430 par Brunelleschi grâce à l'évolution des connaissances en optique géométrique. Il est alors question de chambre noire ou camera obscura en latin. Le principe de base consiste à percer un petit trou appelé sténopé dans la paroi d'une boite noire et d'observer la projection inversée des objets extérieurs sur la paroi opposée au trou. Au 16ème siècle, la popularisation du procédé permet d'élaborer des dispositifs portables (figure 3.1) appelés par extension sténopés. Ils servent principalement aux travaux topographiques et permettent de reproduire les contours de l'image projetée sur une feuille de papier ou une plaque de verre.



(a) (b)

Figure 3.1 – *Camera obscura (a) et un dispositif portable (b) appelé par extension sténopé.*

Le principal inconvénient du sténopé est son manque de luminosité. En effet, la finesse des détails des images produites est dépendante de la dimension du trou. Pour obtenir une image suffisamment détaillée, celui-ci doit être le plus petit possible. Dans ce cas, très peu de lumière passe et l'image est peu visible. En 1550, Cardan propose de remplacer le trou par une lentille convergente qui a la propriété de concentrer les rayons lumineux. Les dispositifs se dotent d'objectifs à base de lentilles convergentes (figure 3.2) et permettent d'obtenir des images plus précises et plus lumineuses.



Figure 3.2 – *Une des premières lentilles convergentes.*

Après cette évolution technique qui permet de créer des images projetées, il reste une étape importante à franchir. Les scientifiques cherchent à automatiser la transformation de l'image projetée en une image physique permanente. Afin de supprimer l'interprétation personnelle et les erreurs de reproduction faites par les dessinateurs. Cette automatisation est élaborée par Niépce au début du 19ème siècle, il est le premier à obtenir et à conserver une image due à l'action de la lumière. Il propose d'utiliser du sel d'argent placé au fond d'une chambre noire car ses propriétés photosensibles le font noircir sous l'action de la lumière. Malheureusement, les images obtenues sont instables car le sel d'argent continue de noircir après l'exposition et les images finissent par disparaitre. Niépce se rend compte de l'importance d'interrompre l'action du produit après la période d'exposition. Pour cela il utilise une autre substance photosensible appelée bitume de Judée dont la propriété est de perdre sa solubilité sous l'action de la lumière. Une plaque de métal enduite de bitume est exposée plusieurs heures puis rincée au solvant et rongée par l'acide aux endroits où le bitume est dissous. Cette plaque (figure 3.3) est ensuite utilisée en imprimerie pour reproduire l'image sur un support papier. Ce principe fonctionne mais les résultats sont peu précis car les dégradés sont mal reproduits.



Figure 3.3 – *Plaque de bitume de judée obtenue par la technique de* Niépce.

Après la mort de Niépce en 1833, Daguerre poursuit l'amélioration du procédé et découvre le principe de l'image latente après plusieurs changements de surface photosensible et de révélateur. Il parvient à obtenir une image stable à partir de l'iodure d'argent et révélée par les vapeurs de mercure. Grâce à cette découverte, Daguerre aidé par l'état français commercialise le premier appareil photographique appelé daguerréotype (figure 3.4a). Cet appareil est constitué d'un objectif, d'une chambre noire, d'une surface photosensible et pour la première fois d'un obturateur dont le but est de limiter la durée d'exposition de la surface photosensible car le temps de pose est réduit à quelques minutes.



(a) (b)

Figure 3.4 – *Le daguerréotype (a) et une pellicule (b) de* Eastman.

Parallèlement aux découvertes de Daguerre, Talbot invente le calotype en 1840. Ce procédé négatif-positif consiste à introduire une feuille de papier enduite de chlorure d'argent dans la chambre noire. Les propriétés photosensibles du chlorure d'argent permettent d'obtenir un négatif, le positif est ensuite crée par tirage contact. Les tirages peuvent être reproduits en plusieurs exemplaires. Le procédé de Talbot est la base de la photographie argentique encore connue de nos jours. L'invention de la pellicule est attribuée à Eastman en 1884 et permet la

réduction de la taille des appareils d'acquisition. En 1888, il fonde la célèbre société Kodak dont le rôle est la fabrication industrielle des pellicules (figure 3.4b) et le tirage des négatifs. Le photographe n'a plus besoin d'avoir des connaissances en chimie, l'acquisition d'image se simplifie et se démocratise auprès du grand public.

En 1903 les frères Lumière révolutionnent la capture d'image par l'invention de l'acquisition couleur en une seule prise. En marge des techniques de l'époque, leur méthode (figure 3.5) emploie la synthèse additive et enregistre l'image sur une plaque photographique en verre. La technique consiste à saupoudrer une plaque de verre avec des millions de grains de fécule de pomme de terre teints en rouge, vert et bleu. Ce filtre est scellé par une laque qui le protège pendant les opérations de développement de la surface sensible. Développée en diapositive, la plaque conserve cette trame de micro-filtres et restitue par transparence les couleurs originales. Durant la suite du 20ème siècle, la technique de filtre couleur des frères Lumière est perfectionnée et adaptée aux pellicules par Kodak. Les surfaces photosensibles sont perfectionnées et les temps d'exposition sont très largement réduits grâce à l'évolution des émulsions d'argent. L'ensemble des procédés photographiques argentiques couleurs tel que nous les connaissons sont le fruit du perfectionnement de toutes ces inventions. Mais l'entrée dans l'ère numérique au début du 21ème siècle a condamné la pellicule argentique à un usage marginal au profit du capteur électronique.

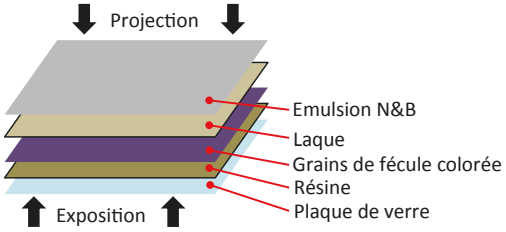


Figure 3.5 – *Technique de filtre couleur des frères* Lumière.

La photographie numérique caractérise l'image obtenue par un capteur électronique utilisé comme surface photosensible mais aussi le traitement qui en découle. La première image numérique fut réalisée en 1975 par Sasson, ingénieur de la société Kodak. A cette époque le véritable défi fut de stocker les fichiers. Mais l'évolution rapide de la technologie électronique et informatique ont permis l'invention d'appareils (figure 3.6) de plus en plus sophistiqués. En 2013, les appareils photographiques numériques (APN) sont capables de capturer, stocker, traiter et diffuser un nombre quasi illimité d'images.

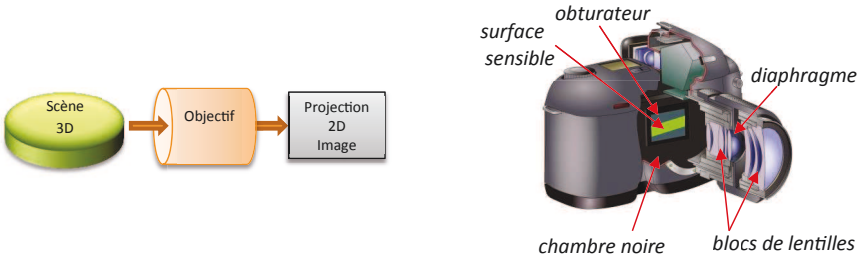


(a) (b)

Figure 3.6 – *Premier APN (a) de* Sasson *et APN moderne (b) le* Nikon *D300S.*

3.2 L'objectif et sa géométrie : la perte dimensionnelle

Un objectif est un dispositif optique qui permet de concentrer la radiance émise par une scène tridimensionnelle dans une chambre noire où les rayons lumineux sont projetés sur une surface plane généralement photosensible. Cette projection entraîne la perte d'une dimension (profondeur) et l'homothétie des deux autres. Elle est réalisée à partir d'un objectif (figure 3.7) qui se compose principalement d'une ou de plusieurs lentilles qui peuvent être mobiles afin d'ajuster les caractéristiques de la projection. L'autre élément de l'objectif est le diaphragme dont le rôle principal est de limiter l'énergie lumineuse à laquelle est exposée la surface sensible. Cette limitation est complétée par le rideau obturateur placé entre l'objectif et la surface sensible.



(a) (b)

FIGURE 3.7 – *Schéma de principe de la perte d'une dimension (a) et éléments optiques de l'appareil photographique (b).*

3.2.1 Principe de convergence et modèle de projection

3.2.1.1 Le cas de la lentille mince

Une lentille est un milieu transparent (verre ou polymère) limité par deux dioptres sphériques de rayon r_1 et r_2. La droite qui joint les centres C_1 et C_2 des deux cercles est l'axe optique. Lorsque la lentille (figure 3.8) est plus mince aux bords qu'au centre, elle est convergente sinon elle est divergente.

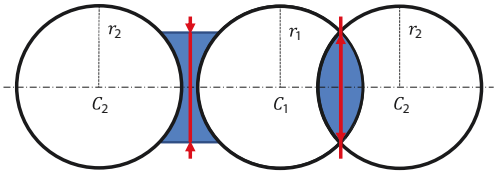


FIGURE 3.8 – *Schéma géométrie des lentilles convergentes et divergentes.*

La densité du milieu est différente de l'air, par conséquent les lois de la transmission optique s'appliquent à l'entrée et à la sortie des rayons suivants trois cas (figure 3.9) :

1. le rayon lumineux a un angle d'incidence nul. Il est transmis suivant l'axe optique.
2. le rayon lumineux a un angle d'incidence oblique mais la première réfraction l'oblige à passer par le centre optique. Il est transmis avec un décalage mais sa direction n'est pas modifiée.

3. le rayon lumineux a un angle d'incidence oblique mais ne passe pas par le centre optique. Il est transmis suivant deux réfractions et sa direction est modifiée.

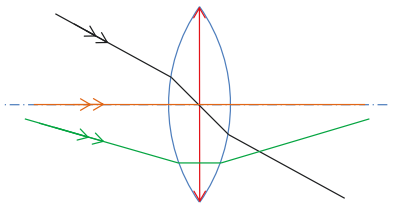


FIGURE 3.9 – *La transmission à travers une lentille mince suivant les trois cas : (1) orange, (2) noir et (3) vert.*

Si les lentilles sont très minces par rapport à l'échelle d'observation, il convient de ne pas tenir compte de leurs épaisseurs et de négliger le déplacement latéral des rayons obliques passant par le centre optique dans l'analyse du principe de convergence.

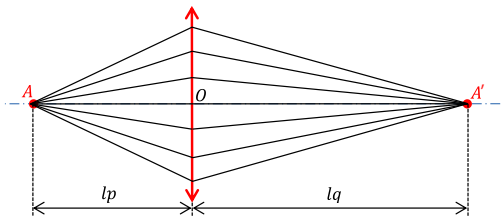
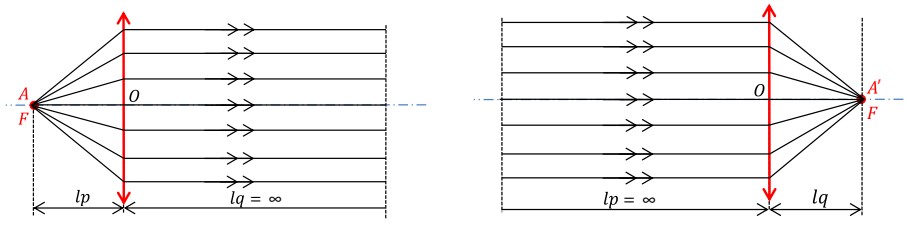


FIGURE 3.10 – *Analyse du principe de convergence et des distances focales.*

Comme illustré figure 3.10, un point A placé à une distance lp sur l'axe optique d'une lentille mince converge vers un point A' placé de l'autre côté de la lentille à une distance lq.

Lorsque lp diminue, la distance lq du point A' augmente. Mais si lp est suffisamment petit, les rayons sortant sont parallèles et la distance lq est infinie. Le principe de convergence n'est plus assuré et lp est appelée distance focale f de la lentille (figure 3.11a). A l'inverse lorsque lp augmente vers l'infini (figure 3.11b), lq se rapproche de la lentille pour atteindre le minimum et devenir aussi la distance focale f.



(a) (b)

FIGURE 3.11 – *Analyse du principe de convergence : (a) lp diminue et atteint le minimum $lp = f$, (b) lp augmente, lq diminue et atteint le minimum $lq = f$.*

Ces deux points minimums de part et d'autre de la lentille sont appelés foyer focaux (F, F') et permettent la construction de l'image d'un objet (figure 3.12).

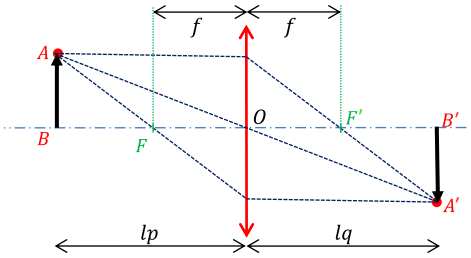


FIGURE 3.12 – *Construction géométrique d'une image à l'aide des foyers focaux.*

L'image A' d'un point lumineux A s'obtient par les vecteurs passant par le centre optique O et les foyers focaux F et F'. D'après les règle de la triangulation, l'objet (AB) forme l'image $(A'B')$ dont l'homothétie peut être définie grâce à la relation du grandissement :

$$\gamma = \frac{A'B'}{AB} = \frac{lq}{lp} \tag{3.1}$$

DESCARTES a par ailleurs démontré la relation entre la distance focale f et les distances lp et lq :

$$\frac{1}{f} = \frac{1}{lp} + \frac{1}{lq} \tag{3.2}$$

3.2.1.2 Le cas de l'objectif

L'objectif d'un APN est un bloc optique composé de plusieurs lentilles divergentes et convergentes, mais l'ensemble des relations sont assemblables sous la forme d'une lentille mince convergente (figure 3.13) équivalente.

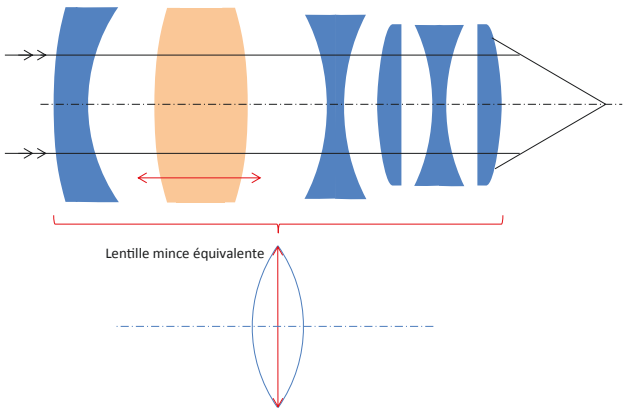


FIGURE 3.13 – *Schéma illustrant l'équivalence entre l'objectif d'un APN et une lentille mince convergente.*

Suivant la technologie, certains objectifs appelés zoom ou multi-focale possèdent une lentille mobile dont le réglage permet de modifier la distance focale et ainsi le grandissement des objets projetés sur la surface sensible.

3.2.1.3 Modélisation mathématique de la projection

La projection réalisée par un objectif ou plus globalement la lentille mince convergente est mathématiquement modélisable (figure 3.14) par l'intermédiaire de trois transformations

élémentaires successives :

1. changement de repère entre la scène réelle et l'objectif.
2. changement de repère entre l'objectif et le plan de l'image projetée.
3. changement de repère entre le plan de l'image projetée et l'indexation des pixels.

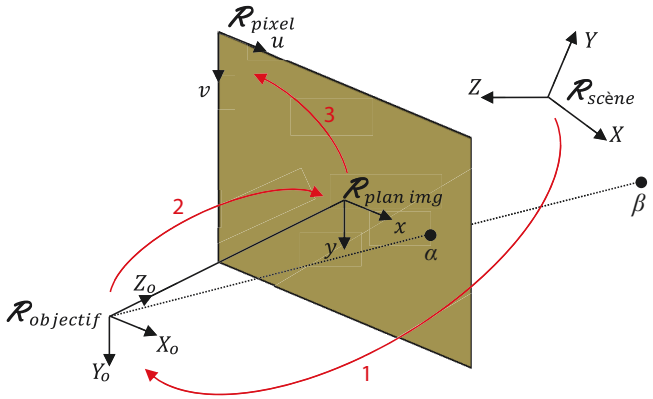


FIGURE 3.14 – *Changement de repères successifs modélisant le principe de la projection perspective.*

La transformation (1) se décompose en une rotation $[\mathbf{R}]$ et une translation $[\mathbf{T}]$. Les valeurs de cette transformation sont des paramètres intrinsèques de l'objectif, ils dépendent de la position du point focal par rapport au repère $\mathcal{R}_{scène}$.

$$\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix} = [\mathbf{R}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + [\mathbf{T}] = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.3}$$

avec :

$$[\mathbf{T}] = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad et \quad [\mathbf{R}] = \begin{bmatrix} r_{11} & \cdots & r_{13} \\ \vdots & \ddots & \vdots \\ r_{31} & \cdots & r_{33} \end{bmatrix} \tag{3.4}$$

Par souci de simplification lorsqu'une unique caméra est utilisée, il convient de définir le repère $\mathcal{R}_{scène}$ en alignement avec le repère $\mathcal{R}_{objectif}$. Ainsi la rotation $[\mathbf{R}]$ est nulle.

La seconde transformation (2) est une projection perspective notée $[\mathbf{P}]$ qui transpose un point tridimensionnel (X_0, Y_0, Z_0) en coordonnées 2d sur le plan image (x, y) en fonction de la distance focale f :

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix} = [\mathbf{P}] \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix} \tag{3.5}$$

par simplification :

$$x = f\frac{X_0}{Z_0} \quad et \quad y = f\frac{Y_0}{Z_0} \tag{3.6}$$

La dernière transformation géométrique (3) correspond à la conversion des coordonnées métriques sur le plan image en indexation (u, v) des pixels :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & k_x \cot\vartheta & c_x + c_y \cot\vartheta \\ 0 & k_y / \sin\vartheta & c_y / \sin\vartheta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = [\mathbf{K}] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3.7}$$

où ϑ est la distorsion éventuelle en orthogonalité, ce paramètre est désigné par le terme anglais "skew factor". Peu fréquent sur les objectifs de qualité, il est très souvent négligé :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & k_x & c_x \\ 0 & k_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = [\mathbf{K}_{simplifié}] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3.8}$$

avec :

k_x	:	taille métrique d'un pixel suivant x.
k_y	:	taille métrique d'un pixel suivant y.
c_x, c_y	:	index du centre optique.

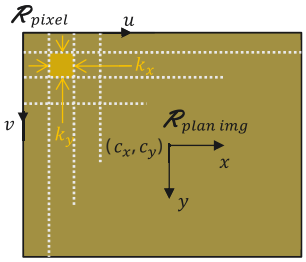


FIGURE 3.15 – *Définition des valeurs lors la conversion des coordonnées métrique sur le plan image en indexation (u, v) des pixels.*

La modélisation géométrique complète d'un index pixel α s'écrit à partir des paramètres extrinsèques $[\mathbf{T}]$ et intrinsèques $[\mathbf{P}]$ et $[\mathbf{K}]$:

$$\alpha = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = [\mathbf{K}][\mathbf{P}][\mathbf{T}]\beta = [\mathbf{K}][\mathbf{P}][\mathbf{T}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.9}$$

Cette modélisation est valable pour un objectif idéal en projection perspective et néglige les éventuelles distorsions géométriques induites par l'imperfection des éléments optiques. Plusieurs auteurs [Bro71, Bey92] ont démontré qu'il s'avère nécessaire d'introduire et de corriger par calibration ces distorsions dans le modèle de projection.

Les principales distorsions de type radial, prismatique ou décentrage [WCH92] sont modélisées par une quatrième transformation [**D**] qui est ajoutée entre les transformations (2) et (3) du modèle complet :

$$(X,Y,Z) \xrightarrow{[\mathbf{T}]} (X_0,Y_0,Z_0) \xrightarrow{[\mathbf{P}]} \boxed{(x,y) \xrightarrow{[\mathbf{D}]} (\tilde{x},\tilde{y})} \xrightarrow{[\mathbf{K}]} (\tilde{u},\tilde{v}) \tag{3.10}$$

La calibration est un problème d'estimation qui consiste à déterminer l'ensemble des paramètres du modèle. Généralement, ce problème est résolu par l'utilisation d'un objet géométrique spécifique qui fournit des points 3d connus dans le repère scène. De nombreuses méthodes [Pen91, WCH92, Zha00, Hem03, EM13] ont été proposées et au fil du temps elles se sont sophistiquées et conduisent vers une estimation de plus en plus juste des paramètres du modèle. L'étape de calibration géométrique est fastidieuse à mettre en œuvre mais permet d'obtenir un modèle très réaliste.

Néanmoins, sous certaines conditions le modèle de projection peut être simplifié en réduisant le nombre de paramètres. Ces simplifications diminuent le réalisme des modèles mais facilitent le traitement mathématique et éliminent une ou plusieurs étapes de la calibration. Les modèles de projection simplifiés les plus rencontrés sont :

Cas simplificateur : projection orthographique

Dans le cas du modèle orthographique, la profondeur des points 3d n'a aucune influence sur la projection. La transformation [**P**] est modifiée et la transformation [**T**] n'a plus lieu d'être car les projections des points (figure 3.16) sont toutes parallèles à l'axe optique :

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = [\mathbf{P}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.11}$$

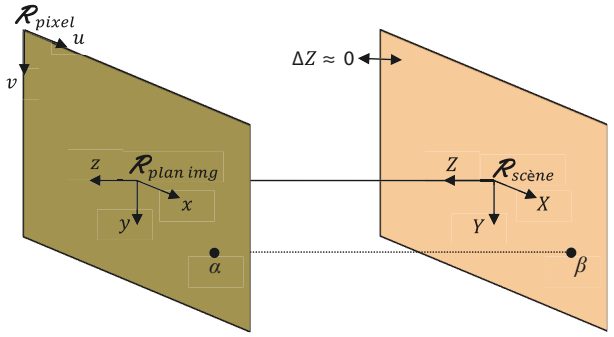


FIGURE 3.16 – *Changement de repère dans le cas simplificateur de la projection orthographique.*

Ce modèle très simpliste nécessite deux conditions élémentaires pour être applicable :

- la distance lp (entre l'objectif et la scène) et la distance focale f sont très élevées par rapport à la profondeur ΔZ entre les points 3d de la scène.
- chaque point de la scène est projeté parrallèlement à l'axe optique.

Cas simplificateur : projection orthographique à facteur d'échelle

La technologie des objectifs permet de réunir les conditions particulières de la projection orthographique. Mais avec un objectif classique la taille de la zone acquise n'est généralement pas identique à celle de l'objectif. Par conséquent, le modèle géométrique inclut un facteur d'échelle et se nomme orthographique à facteur d'échelle ou perspective faible. Comme le modèle orthographique, la transformation $[\mathbf{T}]$ n'est pas considérée. Contrairement au modèle perspectif classique le facteur d'échelle est constant pour l'ensemble des points 3d de la scène, il est défini comme le rapport k_z entre la moyenne des profondeurs $\bar{Z}$ et la distance focale f, ($k_z = \bar{Z}/f$) :

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} k_z & 0 & 0 & 0 \\ 0 & k_z & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = [\mathbf{P}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.12}$$

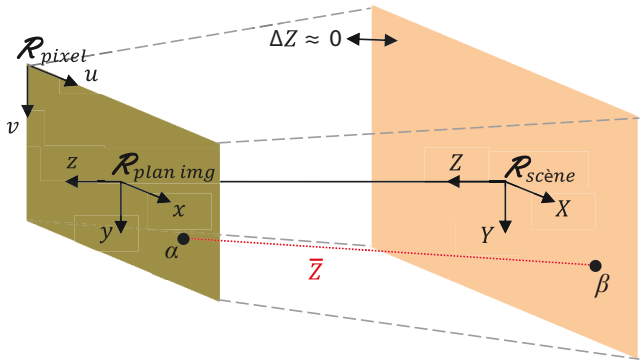


FIGURE 3.17 – *Changement de repère dans le cas simplificateur de la projection orthographique à facteur d'échelle ou perspective faible.*

Ce modèle ne peut être appliqué que si les variations de profondeur ΔZ de la scène sont quasi-inexistantes ou du moins négligeables par rapport à f et $\bar{Z}$.

3.2.2 Le diaphragme

Disposé entre les différentes lentilles de l'objectif, le diaphragme est un dispositif mécanique mince et opaque qui comporte une ouverture centrale réglable. Son rôle est de limiter les rayons lumineux projetés sur la surface sensible par le réglage manuel ou électrique des lamelles mécaniques disposées sous la forme d'un iris. Le diaphragme a trois fonctions :

- limiter les aberrations optiques et chromatiques des lentilles.
- modifier la profondeur de champ.
- réguler l'éclairement de la surface sensible.

En photographie, le diamètre d'ouverture du diaphragme est appelé par simplification ouverture. La valeur associée est relative car elle désigne le rapport entre la distance focale f et le diamètre d du trou dans l'iris. Elle se note $f/\#$:

$$f/\# = \frac{f}{d} \tag{3.13}$$

Les valeurs d'ouverture sont définies par convention comme une suite géométrique de raison $\sqrt{2}$ soit $(f/1, f/1.4, f/2, \ldots)$.

3.2.3 Les défauts

Le passage des rayons lumineux aux travers des éléments de l'objectif n'est pas exempt de défauts. Ils subissent des modifications géométriques et énergétiques qui conduisent vers des aberrations ou des limitations.

3.2.3.1 Les aberrations

Une aberration géométrique (figure 3.18) est un écart de focalisation. Les rayons lumineux reçus par la lentille ne se focalisent pas en un unique point, ce qui conduit à la formation d'une tâche dans la projection. Ce phénomène est dû à l'imperfection des lentilles et leurs épaisseurs négligées lors de l'étude de la convergence. Pour minimiser les aberrations, la courbure des lentilles est usinée de façon très précise afin de réduire la courbure initiale. Mais ce procédé n'assure pas une correction idéale, surtout aux extrémités là où la divergence est la plus importante.

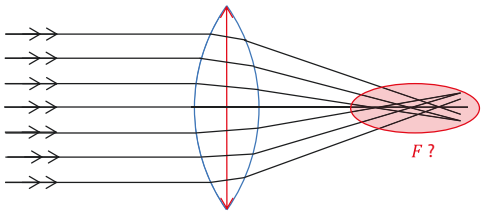


FIGURE 3.18 – *Illustration du principe de l'aberration géométrique, les rayons ne se focalisent pas en un point.*

L'aberration chromatique (figure 3.19) est due à la variation de l'indice de réfraction du matériau composant la lentille. En fonction de la longueur d'onde des rayons lumineux reçus, la lentille se comporte comme le prisme de NEWTON et décompose la lumière. Il en résulte des points focaux variables en fonction de la couleur. Comme l'aberration géométrique, ce phénomène est plus important aux extrémités de la lentille.

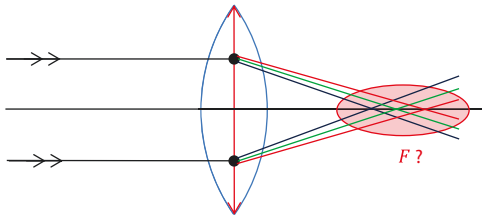


FIGURE 3.19 – *Illustration du principe de l'aberration chromatique, un rayon est divisé en fonction de sa couleur en plusieurs points focaux.*

La fermeture du diaphragme contribue à réduire ces aberrations car seul le centre de la lentille moins assujettie aux aberrations est utilisé.

3.2.3.2 La profondeur de champ

L'ouverture agit directement sur la profondeur de champ (figure 3.20), c'est-à-dire la zone d'espace ($\mathcal{R}_{scène}$) dans laquelle l'image projetée ($\mathcal{R}_{image}$) est nette autour du plan de mise au

point. La distance focale détermine le plan de mise au point pour lequel tout point C de la scène est projeté comme un unique point C'. Néanmoins, la taille des pixels ou la sensibilité spatiale des surfaces sensibles permet d'établir une zone de netteté autour du point C.

Ainsi tout point de la scène situé entre les plans passant par A et B donnent lors de leurs projections une tache de diamètre acceptable pour être considérée comme nette. La distance entre A et B définit la profondeur de champ. Par géométrie, la réduction de l'ouverture du diaphragme permet d'augmenter la profondeur de champ.

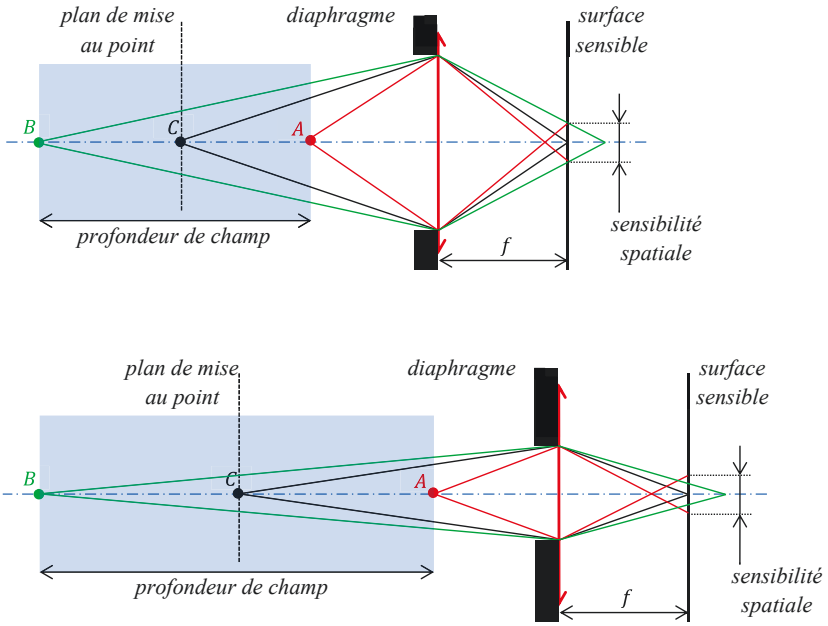


FIGURE 3.20 – *(haut) Schéma de principe de la profondeur de champ, (bas) plus l'ouverture du diaphragme diminue plus la profondeur de champ augmente.*

3.2.3.3 L'éclairement énergétique

Jusque-là seul l'aspect géométrique a été abordé, mais il ne faut pas oublier que la lumière est une onde électromagnétique transportant de l'énergie. Ainsi le principe de projection géométrique opéré par l'objectif permet de convertir l'intensité Ie émise par chaque point d'une scène en valeurs d'éclairement énergétique (irradiance) Ee auxquelles est exposée la surface du capteur pendant un laps de temps Δt. D'après HORN dans [Hor86], l'éclairement énergétique est définie comme :

$$Ee_{\Delta t} = \left(Ie \, \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \theta \right) \Delta t \tag{3.14}$$

avec :

Ie	:	l'intensité rayonnante émise par un point de la scène.
d	:	diamètre d'ouverture du diaphragme.
f	:	distance focale.
θ	:	angle d'incidence.
Δt	:	laps de temps pendant laquelle le capteur est exposé.

Le terme $\cos^4 \theta$ correspond à l'angle du rayon qui atteint la surface du capteur par rapport à

l'axe optique. La modulation du diamètre du diaphragme d module directement les valeurs d'irradiance. En dehors du temps d'exposition Δt, la surface du capteur est plongée dans le noir par obturation totale.

Historiquement, cette fonction était assurée par l'occultation de l'objectif grâce à un bouchon, puis par la fermeture complète du diaphragme qui s'ouvrait au diamètre voulu pendant le temps désiré. Mais ce type d'obturation est très lente et l'évolution des surfaces sensibles a conduit à des temps très faibles. Afin d'outrepasser ce problème, l'obturateur mécanique est apparu. Il est placé entre la sortie de l'objectif et la surface sensible. Son principe consiste à ouvrir un volet mécanique et permet d'atteindre des temps d'exposition courts ($1/1000$ *sec*). A présent, le passage au capteur numérique rend l'obturateur obsolète car le temps d'obturation appelé temps d'intégration est géré électroniquement par le capteur. Tout de même, l'obturateur mécanique reste présent sur un grand nombre d'appareils numériques pour assurer l'initialisation (mise à zéro des valeurs capturées) des capteurs.

3.3 Le capteur numérique

Dans le cas des images numériques, les surfaces sensibles sont des capteurs électroniques. Géométriquement, ces capteurs d'images sont des surfaces photosensibles subdivisées selon la résolution en éléments pixels. Chaque pixel est indexé par des coordonnées discrètes (u, v) deux dimensions (2d) en fonction de sa position dans le repère plan image.

Chaque pixel possède une taille réelle dans le repère plan image qui est définie comme le rapport entre la taille de la surface sensible et la résolution, soit le nombre de pixels suivant les lignes et les colonnes :

$$Taille\ réelle\ pixel_{(u\ ou\ v)} = \frac{Taille\ surface\ sensible_{(u\ ou\ v)}}{Nombre\ de\ pixel_{(u\ ou\ v)}} \tag{3.15}$$

La taille d'un pixel dans le plan scène est définie par projection à partir de la distance focale :

$$Taille\ pixel\ scène_{(u\ ou\ v)} = \frac{Taille\ réelle\ pixel_{(u\ ou\ v)}}{f} \tag{3.16}$$

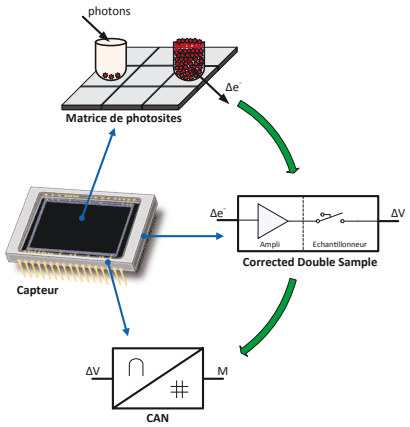


FIGURE 3.21 – *Éléments principaux du capteur numérique.*

L'énergie émise par chaque portion de scène correspondant à un pixel est mesurée par le capteur. Cette mesure s'opère suivant trois étapes schématisées par la figure 3.21, la première étape consiste à collecter l'énergie lumineuse : les photons. Ensuite, ils sont convertis en énergie électrique grâce à un circuit de lecture. Enfin, la dernière étape consiste à obtenir des valeurs numériques quantifiées.

3.3.1 L'effet photoélectrique

Les capteurs d'image numérique sont qualifiés de détecteurs à l'état solide et sont constitués principalement de silicium. Ce matériau est semi-conducteur et suit les lois de la physique quantique appliquées aux électrons des atomes. Elles définissent que les niveaux d'énergies sont regroupés en "bandes". Pour le silicium, il existe deux bandes principales :

- la bande de valence.
- la bande de conduction.

Elles sont séparées par une bande intermédiaire dite d'énergie interdite appelée "gap" dont la largeur est de $e_g = 1.92\ eV$ selon [SN81]. La bande de valence correspond aux électrons dits de valence ou liés qui assurent les liaisons entre les atomes et la bande de conduction correspond aux électrons excités ou libres.

La transmission des électrons liés de la bande de valence vers des électrons libres de la bande de conduction est réalisée sous l'effet d'une absorption d'un rayonnement électromagnétique. Ainsi lorsqu'un photon de longueur d'onde λ est absorbé, si son énergie est supérieure à e_g, elle permet d'arracher un électron de la bande de valence vers la bande de conduction par la libération d'une charge positive ou "trou" et génère une paire d'électron/trou. Les électrons et les trous libérés se déplacent dans la bande de conduction (figure 3.22) et créent un courant. Du fait de leur provenance lumineuse, les électrons sont appelés photo-électrons.

D'après [SN81], pour respecter la condition sinequanon ($e_{photons} > e_g$), la longueur d'onde λ du rayonnement doit être inférieure à 1100 nm. Par conséquent les photons dont la longueur d'onde est supérieure à cette limite soit le proche infrarouge, ne produisent pas d'effet photo-électrique dans le silicium.

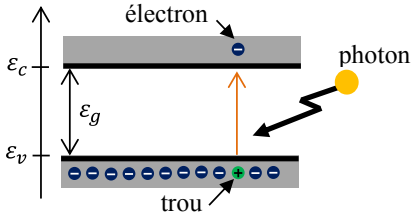


FIGURE 3.22 – *L'effet photoélectrique avec le silicium.*

D'autres matériaux tel que le Germanium possède une bande gap plus faible $\epsilon_g = 0,7\ eV$, ils induisent un effet photoélectrique pour des longueurs d'onde jusqu'à 1800 nm. Mais il existe peu de capteurs utilisant ces matériaux car le silicium est très bien maitrisé au niveau de la technologie de fabrication et permet d'obtenir des dispositifs performants à faible coût.

3.3.2 Le photosite

Tous les pixels à la surface d'un capteur sont des puits de lumière appelés aussi photosites dont le but est de réaliser la transconductance lumière/électricité en exploitant l'effet photoélectrique. Cette transconductance est effectuée soit par des photodiodes, soit par une structure photoMOS.

Les photodiodes (figure 3.23a) sont polarisées en inverse afin de générer un courant lorsqu'elles sont éclairées, la technologie CMOS (Complenetary Metal Oxide Semiconductor) permet de concevoir des photodiodes basées sur une structure N^+/P. La région P correspond au substrat silicium dopé positif tandis que la région N^+ est réalisée par diffusion d'un dopant négatif directement sur le substrat.

Dans le cas des photoMOS, la structure (figure 3.23b) est légèrement différente. Elle est basée sur une capacité MOS (Métal Oxide Semicondutor) et se compose d'un substrat de silicium dopé P, d'une fine couche isolante d'oxyde de silicium et d'une électrode métallique (ou grille) en polysilicium dont la résistance est nulle.

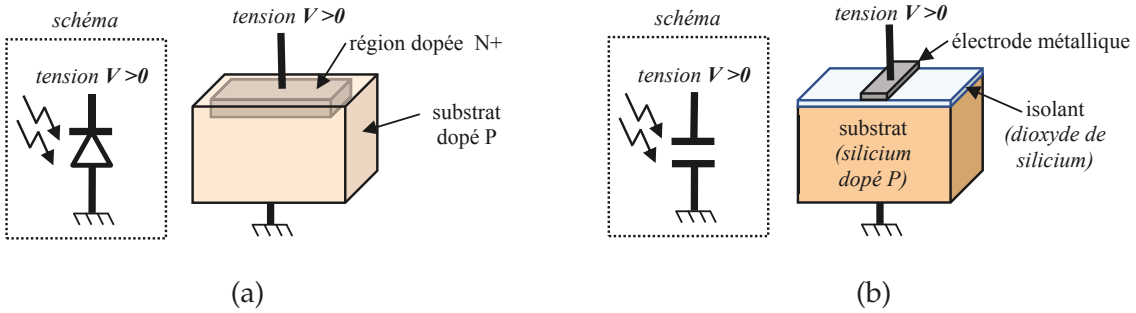


FIGURE 3.23 – *Le photosite : (a) technologie de la photodiode, (b) technologie photoMOS.*

Dans les deux cas, le substrat dopé P contient des porteurs positifs (trous). Par application d'une tension sur l'électrode métallique, les trous sont repoussés et une zone de déperdition dont la taille est proportionnelle à l'amplitude de la tension est créée (figure 3.24a).

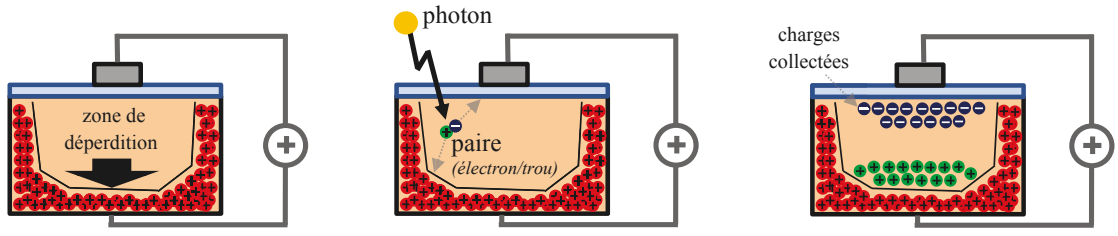


FIGURE 3.24 – *Principe photoélectrique dans le photosite.*

Selon le principe photoélectrique, à chaque fois qu'un photon atteint le substrat, il crée une paire d'électron/trou (figure 3.24b). Durant le temps d'intégration (ou d'exposition) les électrons sont piégés dans la zone de déperdition, la charge totale recueillie est proportionnelle à l'intensité lumineuse reçue (figure 3.24c). Par analogie, le principe du photosite est comparable à un condensateur de faible valeur qui se charge sous l'effet de la lumière. Mais la conversion

photons-électrons n'est pas parfaite et illimitée, elle dépend de plusieurs paramètres tels que :

- la pleine capacité du photosite.
- le temps d'intégration.
- le rendement quantique.
- l'efficacité de détection ou quantique.

3.3.2.1 La pleine capacité

La pleine capacité Δe^-_{max} ou "full capacity" correspond au nombre maximum d'électrons collectables (figure 3.25) par photosite. Cette valeur dépend de la taille des photosites et de la tension appliquée sur l'électrode métallique car elle est proportionnelle à la zone de déperdition.

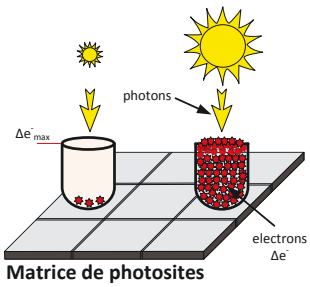


FIGURE 3.25 – *Pleine capacité d'un photosite.*

A tort, beaucoup attribuent la taille des photosites à la résolution du capteur. Il est logique qu'une même surface subdivisée en plus de pixel entraîne des photosites plus petits donc une pleine capacité plus faible. Mais cette affirmation est incomplète car la taille des photosites dépend aussi de la technologie employée lors de la construction du substrat. Par conséquent, un capteur 24×36 mm subdivisé en $5M$ $pixels$ peut avoir des valeurs pleine capacité inférieures à un capteur de la même taille subdivisé en $12M$ $pixels$ si la technologie est différente.

3.3.2.2 Le temps d'intégration

La collecte des photons par le photosite est effective pendant l'exposition à la lumière. Le nombre de photons reçus est proportionnel à l'éclairement énergétique du flux et au temps d'exposition fixé par l'obturateur. Les contraintes mécaniques de l'obturateur empêchent la limitation temporelle de l'exposition. Par conséquent, si l'éclairement énergétique du flux est trop importante, le photosite atteint sa valeur de pleine capacité (figure 3.26) et devient inexploitable car il est saturé.

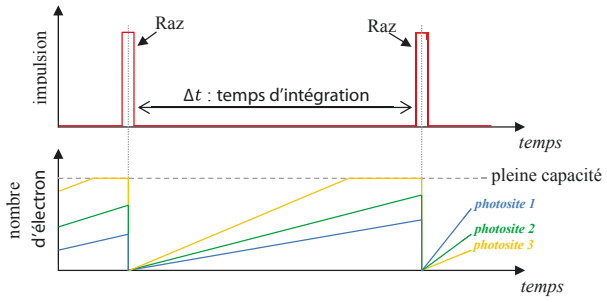


FIGURE 3.26 – *Pleine capacité d'un photosite.*

Pour pallier ce problème, les photosites possèdent un temps d'intégration Δt dont la limite permet d'atteindre un temps plus court que celui de l'obturateur. Il correspond à la période entre deux impulsions de remise à zéro des photosites.

3.3.2.3 Rendement quantique

Le rendement quantique qualifie le nombre d'électrons collectés par rapport au nombre de photons reçus en fonction de la longueur d'onde. Ce paramètre subit l'influence de la technologie du capteur car en fonction de la nature et de l'épaisseur des couches de substrat traversées, les photons incidents sont assujettis aux phénomènes de réflexion (figure 3.27) qui conduisent à leurs pertes. Le calcul du rendement quantique est défini par :

$$N(\lambda) = \frac{\Delta e^-}{\Delta photons} = \frac{électron\ collectés}{photons\ reçus} \tag{3.17}$$

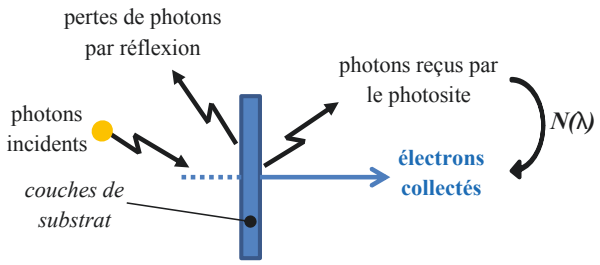


FIGURE 3.27 – *Le rendement quantique d'un photosite.*

3.3.2.4 Efficacité quantique

L'efficacité QE définit la capacité des photosites à convertir les photons incidents en électrons (figure 3.28) en fonction de la longueur d'onde :

$$QE(\lambda) = N(\lambda) \times Facteur\ de\ remplissage \tag{3.18}$$

où le *Facteur de remplissage* est le rapport entre la taille de la zone de détection et la taille totale du photosite. La technologie actuelle permet d'obtenir des facteurs de remplissage supérieurs à 70%.

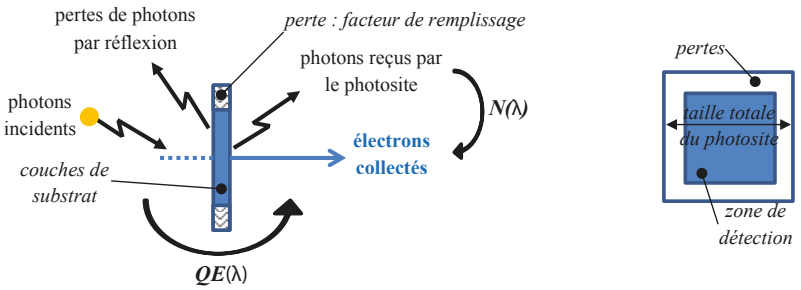


FIGURE 3.28 – *L'efficacité d'un photosite et l'illustration du Facteur de remplissage.*

3.3.3 Le circuit de lecture

La conversion des charges (électrons) en tension est assurée par le circuit de lecture. Ce circuit est généralement composé d'un amplificateur et d'un échantillonneur bloqueur. Le

facteur de conversion (Q/V) exprimé en électrons/volt définit la capacité du circuit à convertir les électrons en tension. La tension de sortie est définie par :

$$\Delta V = \Delta e^{-} \times (Q/V) \times \frac{1}{g_n} \tag{3.19}$$

avec g_n la valeur d'amplification réglable définie par la sensibilité ISO. La valeur maximale de la tension de sortie ΔV_{max} est indépendante de l'amplification et fixe en fonction de la technologie :

$$\Delta V_{max} \neq \Delta e^{-}_{max} \times (Q/V) \times \frac{1}{g_n} \tag{3.20}$$

De plus, l'intégralité de la tension de sortie n'est pas exploitable à cause des diverses perturbations ou bruits générés lors de la conversion. La partie dite utile du signal correspond à la tension maximale de sortie à laquelle sont soustraits les bruits :

$$\Delta V_{utile} = \Delta V_{max} - bruits \tag{3.21}$$

3.3.3.1 Les bruits de conversion

Les perturbations aléatoires qui se superposent au signal utile et limitent la sensibilité des photosites sont d'origines diverses. Parmi elles, il est possible d'isoler deux catégories.

Le bruit temporel

Ce type de bruit est dépendant de la sensibilité ISO et se compose de trois sources distinctes :

- le bruit de grenaille appelé aussi courant de fuite $\mathcal{B}_{DC}$ (Dark Current) est la source principale et intervient tout le long du temps d'intégration. Il provient de la fuite du courant dans les composants et engendre la fluctuation du nombre de charges collectées. Sa distribution est couramment associée à une loi de Poisson dont la variance dépend du temps d'intégration. Ainsi même à éclairement nul, les photosites se remplissent de charges si le temps d'intégration est long. Sa valeur est généralement fournie par les constructeurs (exemple : 5 *électrons/seconde*) et peut être convertie en tension.
- le bruit thermique $\mathcal{B}_{TH}$ est produit par l'agitation thermique des électrons aux bornes de tout semi-conducteur [CMB96, TFG01]. Il en résulte une tension dont le biais est dépendant de la température absolue et de la résistance du composant. Ce bruit intervient principalement lors de la remise à zéro des photosites, la tension de sortie ΔV n'est pas nulle et présente un biais.
- le bruit de vacillement $\mathcal{B}_{FN}$ (Flicker Noise) est associé à la fréquence du rayonnement magnétique reçu. Il est dû aux transistors qui composent les photosites et fut très longtemps soumis à controverse [Hoo94, VLR94]. Des travaux plus récents tels que [DLMFoo] ont prouvé son existence et fournissent un modèle complet du phénomène physique. Mais la prise en compte reste marginale car les valeurs sont négligeables par rapport aux bruits $\mathcal{B}_{DC}$ et $\mathcal{B}_{TH}$

Le bruit spatial

Le bruit spatial fixe $\mathcal{B}_{FPN}$ (Fixed Pattern Noise) est le niveau de bruit invariant dans le temps associé aux imperfections entre les photosites placés à la surface du capteur. Ces

imperfections apparaissent lors de la gravure des capteurs et entraînent une variance reproductible d'une acquisition à l'autre dans la capacité et la réponse des photosites. Pour limiter l'influence des imperfections, les constructeurs fixent une marge de sécurité ΔV_{marge} sur la valeur ΔV_{max} de tous les photosites.

3.3.3.2 Bilan : tension de sortie

La description des bruits de conversion permet d'établir la tension utile de sortie (figure 3.29) à partir de l'équation 3.21 où $\mathcal{B}_{FN}$ est négligé :

$$\Delta V_{utile} = (\Delta V_{max} - \Delta V_{marge}) - \frac{1}{g_n}(\mathcal{B}_{DC} + \mathcal{B}_{TH}) \tag{3.22}$$

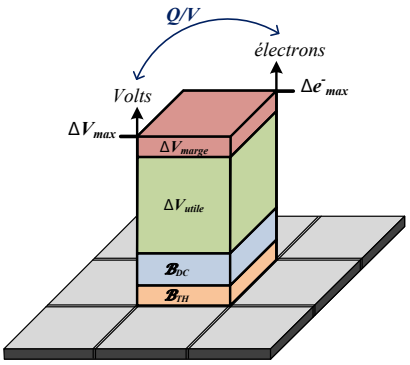


FIGURE 3.29 – *Bilan de la conversion électrons/tension à la sortie du photosite.*

3.3.4 La conversion analogique numérique

La dernière étape du capteur numérique consiste à convertir les tensions de sortie des photosites en valeurs numériques codées sur un nombre de Bits N défini et appelé résolution ou profondeur numérique. Cette conversion fait intervenir un composant électronique appelé CAN (Convertisseur Analogique Numérique) et nécessite un temps de traitement pendant lequel la tension de sortie est fixe car bloquée par l'échantillonneur bloqueur du circuit de conversion. Le temps de traitement est dépendant de la technologie employée pour la conversion :

- CAN à approximations successives, le principe repose sur un unique comparateur. En interne, un séquenceur fournit une tension analogique qui est comparée au signal à convertir. La comparaison est réalisée par dichotomie du bit de poids fort (MSB) jusqu'au bit de poids faible (LSB). Le temps de conversion est directement lié à la résolution du CAN, il est d'environ une dizaine de microsecondes par photosite pour une résolution de 12 Bits.
- CAN flash, le principe repose sur l'utilisation de plusieurs comparateurs dont l'une des entrées est reliée à des tensions de référence et l'autre à la tension de sortie du photosite. pour obtenir N Bits en sortie, N comparateurs sont reliés à une chaine potentiométrique constituée de N résistances de valeurs différentes. La comparaison s'effectue en parallèle du MSB au LSB. Le temps de conversion est indépendant de la résolution du CAN et avoisine la dizaine de nanosecondes par photosite. Mais la précision est directement liée à la chaine potentiométrique : les courants de fuite, la tolérance des résistances et les biais des comparateurs influent directement sur les valeurs numériques de sorties.

3.3.4.1 Cas idéal : quantification

Pour un CAN à N Bits, le nombre de valeurs numériques en sortie est 2^N soit des valeurs binaires entre 0 et $2^N - 1$. Le CAN est également caractérisé par sa plage acceptable de tension analogique d'entrée appelée tension Pleine Echelle V_{PE} (Full Scale).

Dans le cas du capteur d'image, la tension V_{PE} doit être adaptée à la tension de sortie des photosites :

$$V_{PE} = (\Delta V_{max} - \Delta V_{marge}) \tag{3.23}$$

Pour une quantification uniforme idéale (figure 3.30), V_{PE} est divisée en autant de plages d'égale dimension que de valeurs numériques possibles. Le quantum qt définit la dimension des plages :

$$qt = \frac{V_{PE}}{2^N} = \frac{(\Delta V_{max} - \Delta V_{marge})}{2^N} \tag{3.24}$$

Les tensions de seuil V_{Sn} correspondent aux transitions entre les valeurs numériques de sortie :

$$V_{Sn} = n \times qt \qquad n \in \{1, 2, \ldots, N-1\} \tag{3.25}$$

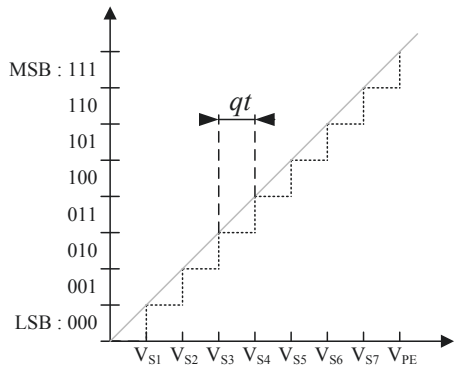


FIGURE 3.30 – *Exemple d'une quantification pour un CAN 3 bits.*

3.3.4.2 Erreur de quantification

L'erreur de quantification est la différence entre la valeur du signal à quantifier et la valeur analogique qui correspond à la valeur numérique de sortie. Elle s'exprime en fonction du quantum ou par analogie en LSB. Sa valeur est comprise entre 0 et 1 LSB dans le cas idéal car toutes les valeurs d'entrée comprises entre V_{Sn} et V_{Sn+1} sont représentées par la valeur numérique 2^n. Cette erreur est intrinsèque au CAN et dépend uniquement de sa résolution. Plus le nombre de bits est important plus l'erreur de quantification est faible.

L'erreur absolue de quantification peut être divisée par deux par changement des tensions de seuil V_{Sn}. Physiquement ce changement est réalisé (figure 3.31) par l'ajout d'un biais équivalent à $qt/2$ sur la tension ΔV à quantifier. La valeur absolue de l'erreur est alors de LSB/2.

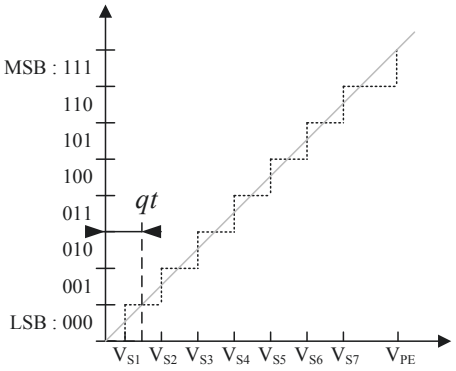


FIGURE 3.31 – *Exemple, décalage des valeurs de seuil pour limiter l'erreur absolue de quantification à LSB/2.*

3.3.4.3 Cas réel : non linéarité

Comme tous les composants électroniques, le CAN est assujetti aux bruits thermiques et grenailles. En plus de l'erreur de quantification, ces bruits sont générateurs d'erreurs de gain et de non linéarités. L'erreur de gain est l'écart entre la pente de la caractéristique idéale de conversion et la caractéristique réelle aux centres des paliers. Cette erreur a pour effet de réduire la plage des tensions convertibles dans le cas d'une pente plus forte ou dans le cas inverse, la plage des valeurs numériques possibles. Les erreurs de non linéarités (figure 3.32) caractérisent les variations locales des tensions de seuil. Elles se distinguent en deux catégories :

– les non linéarités différentielles DNL (Differential Non Linearity) qui correspondent à l'écart entre la dimension d'un palier de la caractéristique de transfert réelle et la dimension idéale 1LSB.

$$DNL(n) = \frac{V_{Sn+1} - V_{Sn} - qt}{qt} \tag{3.26}$$

– les non linéarités intégrales INL (Integral Non Linearity) sont la représentation cumulative des DNL et matérialisent l'écart entre le centre d'un palier et la fonction de transfert idéale.

$$INL(n) = \frac{V_{Sn} - V_{Sn}idéale}{qt} = \sum_{n \leq N} DNL(n) \tag{3.27}$$

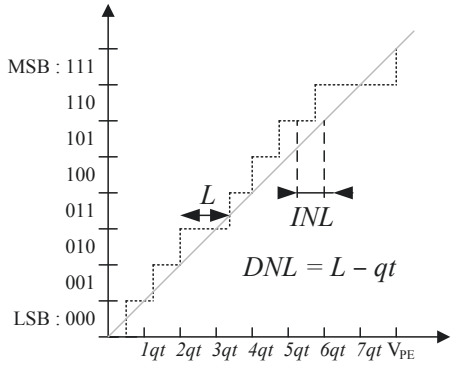


FIGURE 3.32 – *Exemple, d'un CAN présentant des non linéarité dans la conversion.*

L'ensemble des erreurs rencontrées lors de la conversion sont dues aux propriétés intrinsèques des CAN. Bien qu'elles soient décelables et mesurables, les corrections aveugles (sans connaissance préalable du signal d'entrée) sont impossibles à réaliser.

3.3.5 La dynamique

La dynamique se mesure généralement en décibels (dB) et correspond à l'écart entre deux valeurs :

$$\mathcal{D} = 20 \times \log_{10}\left(\frac{Valeur\ maximum}{Valeur\ minimum}\right) \tag{3.28}$$

Dans le cas du capteur d'image, les diverses conversions de l'énergie conduisent à définir trois types de dynamique : la dynamique photonique, la dynamique photoélectrique, la dynamique numérique, dont l'interaction permet d'obtenir la dynamique globale du capteur.

3.3.5.1 Dynamique photonique

La dynamique photonique correspond au rapport entre le nombre de photons nécessaire pour remplir d'électrons un photosite et le nombre minimum de photons pour générer un électron. Ces valeurs sont dépendantes de l'efficacité $QE(\lambda)$ à la longueur d'onde des photons incidents :

$$\mathcal{D}_{photonique} = 20 \times \log_{10}\left(\frac{\Delta e^-_{max} \times 1/QE(\lambda)}{1 \times 1/QE(\lambda)}\right) \tag{3.29}$$

Par simplification la dynamique photonique est calculée à partir du nombre d'électrons collectables par photosite :

$$\mathcal{D}_{photonique} = 20 \times \log_{10}\left(\Delta e^-_{max}\right) \tag{3.30}$$

3.3.5.2 Dynamique photoélectrique

La dynamique photoélectrique intervient après la conversion des charges collectées en tension. Elle est définie comme le rapport entre le maximum et le minimum du signal utile, soit :

$$\mathcal{D}_{photoélectrique} = 20 \times \log_{10}\left(\frac{\Delta V_{max} - \Delta V_{marge}}{\frac{1}{g_n}(\mathcal{B}_{DC} + \mathcal{B}_{TH}}\right) \tag{3.31}$$

Par logique, la dynamique photoélectrique est inférieure à la dynamique photonique :

$$\mathcal{D}_{photoélectrique} < \mathcal{D}_{photonique} \tag{3.32}$$

3.3.5.3 Dynamique numérique

La dynamique numérique est calculée à partir du nombre de bits utiles du CAN, soit le rapport entre la plus haute valeur numérique MSB et la plus faible valeur numérique non nulle :

$$\mathcal{D}_{numérique} = 20 \times \log_{10}\left(\frac{2^N - 1}{1}\right) \tag{3.33}$$

Par convention, elle se calcule aussi en fonction du quantum et de la valeur pleine échelle du CAN :

$$\mathcal{D}_{numérique} = 20 \times \log_{10}\left(\frac{V_{PE}}{qt}\right) \tag{3.34}$$

En toute logique la dynamique numérique dépend des paramètres intrinsèques du CAN et doit être adaptée à la dynamique photoélectrique des photosites :

$$\begin{aligned} V_{PE} &= (\Delta V_{max} - \Delta V_{marge}) \\ qt &\geq \frac{1}{g_n}(\mathcal{B}_{DC} + \mathcal{B}_{TH}) \\ qt &\doteq (Q/V) \end{aligned} \tag{3.35}$$

Au même titre que la résolution spatiale des capteurs, une dynamique numérique élevée est très souvent utilisée comme argument commercial. Mais un quantum inférieur au facteur de conversion charge/tension du circuit de lecture ($qt < (Q/V)$) est totalement inutile car les valeurs numériques de poids faibles sont le fruit d'une interpolation.

3.3.5.4 Dynamique du capteur

La combinaison de l'ensemble des dynamiques (figure 3.33) et le respect des valeurs des plages permet d'établir la dynamique globale du capteur comme étant le minimum entre la dynamique photonique, photoélectrique et numérique.

$$\mathcal{D}_{capteur} = \min\left[\mathcal{D}_{photonique}, \mathcal{D}_{photoélectrique}, \mathcal{D}_{numérique}\right] \tag{3.36}$$

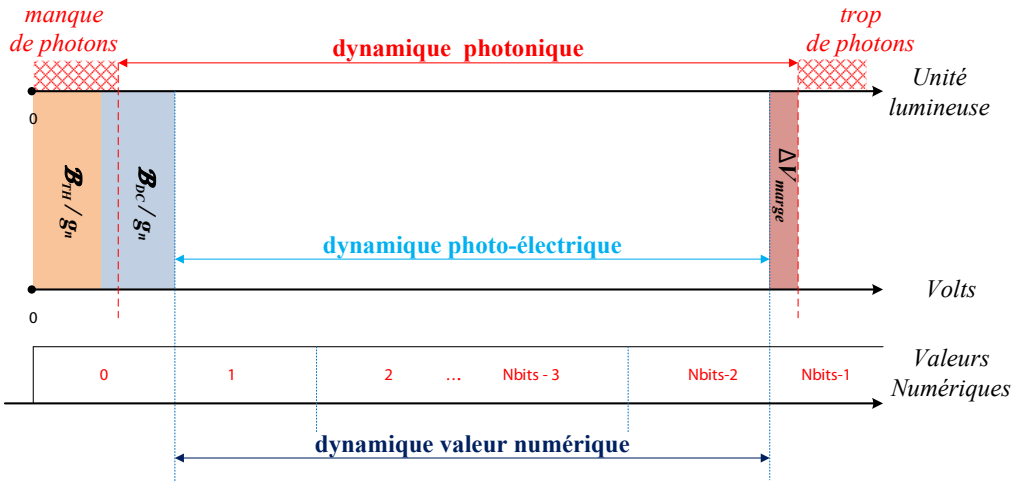


FIGURE 3.33 – *La dynamique du capteur.*

3.3.6 Les technologies des capteurs d'images

L'objectif de cette section est de recenser, décrire et comparer les technologies de capteurs d'images développées en micro-électronique depuis 1975.

3.3.6.1 La structure MOS-XY

La surface sensible de ce capteur est composée d'une matrice de photodiodes. Chaque photosite est muni d'un interrupteur MOS et de deux décodeurs d'adresses X et Y pour lire les valeurs recueillies comme des éléments mémoires. Les décodeurs sont disposés à la périphérie de la matrice de photosites [Aa85, Aa85, Ka80] (figure 3.34).

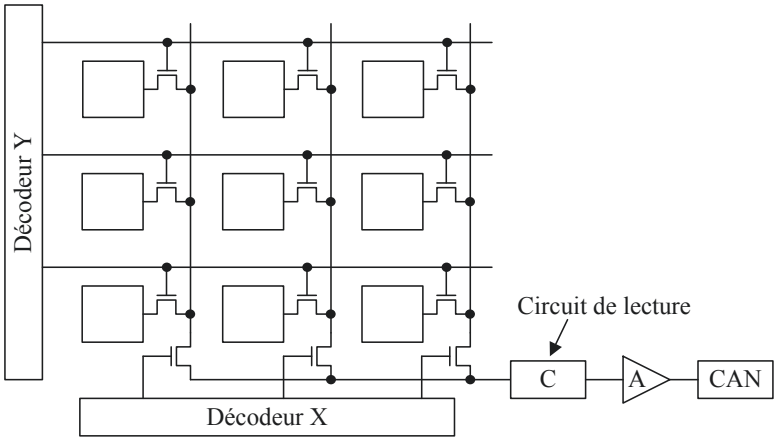


FIGURE 3.34 – *Structure d'un capteur MOS-XY.*

Chaque photodiode est directement reliée à un bus analogique vertical. Le flux d'information série est obtenu par l'adressage séquentiel de la matrice deux dimensions. Après le temps d'intégration, tous les photosites sont lus et remis à zéro les uns après les autres à l'aide des décodeurs X et Y.

En termes de performance, ce type de capteur souffre d'un niveau de bruit spatial fixe $\mathcal{B}_{FPN}$ important. Ce défaut provient du fait que les photodiodes ne sont pas directement gravées sur la surface mais collées et présentent donc de fortes disparités. De plus, elles sont de faible capacité et directement reliées aux bus analogiques de forte capacité. Par conséquent, ces capteurs appelés PPS (Passive Pixel Sensor) produisent un haut niveau de bruit et une faible dynamique.

3.3.6.2 La structure CID

Les capteurs CID (Charge Injectin Device) possèdent une architecture similaire aux capteurs MOS-XY [ZWPa91]. Seuls les photosites sont différents car les photodiodes sont remplacées par deux capacités MOS photosensibles.

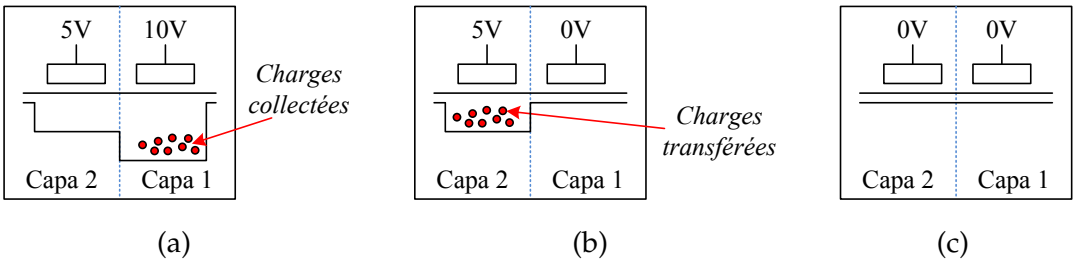


FIGURE 3.35 – *Structure des photosites sur un capteur CID.*

La collecte des charges est réalisée par la capacité ayant la plus haute tension durant le temps d'intégration (figure 3.35a). Après le temps d'intégration, la tension de la capacité 1 est ramenée à $0V$ et les charges sont transférées vers la capacité 2 (figure 3.35b). Ce transfert génère une tension proportionnelle à la quantité des charges échangées entre les deux capacités, donc proportionnelle à l'éclairement énergétique reçue par le photosite. Enfin, la remise à zéro est réalisée par l'application d'une tension nulle aux deux capacités (figure 3.35c).

Ce type de capteur souffre des mêmes défauts que les capteurs MOS-XY. Les photosites présentent de fortes disparités et sont fortement perturbées lors de la lecture [ZEAa93]. Les performances médiocres de ces capteurs ont marginalisé leur utilisation. A partir des années 80, l'apparition des capteurs CCD plus performants a rendu les capteurs MOS-XY et CID complètement obsolètes et a conduit à leur disparition.

3.3.6.3 La structure CCD

Pour les capteurs CCD (Charge Coupled Device) chaque photosite est constitué d'une photoMOS qui accumule les charges lors du temps d'intégration. Après ce temps, les charges collectées sont transférées de photoMOS en photoMOS par un registre de transfert. Ce procéssus permet d'évacuer l'information lumineuse sous forme de charges qui sont ensuite converties en tension par un unique circuit de lecture situé en dehors de la matrice photosensible [The95] (figure 3.36).

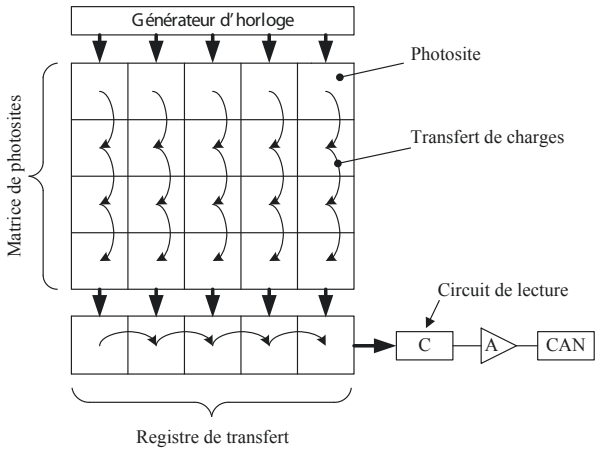
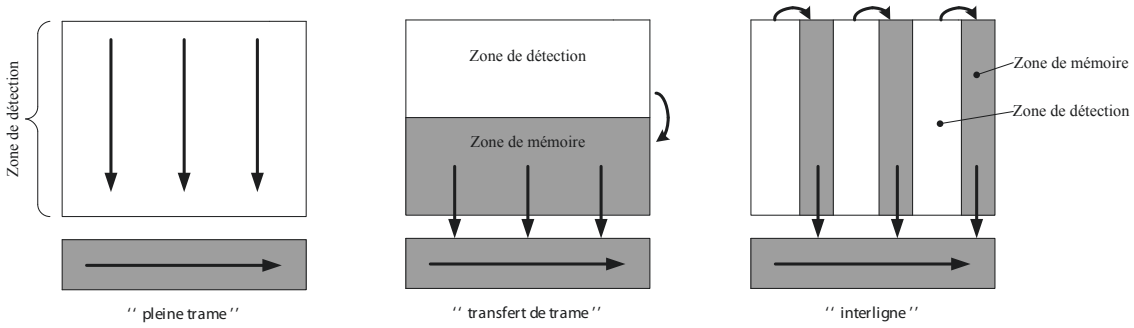


FIGURE 3.36 – *Structure d'un capteur CCD.*

Dans le cas standard appelé pleine trame (figure 3.37a), chaque photosite est sollicité lors des deux phases. Durant la phase d'exposition, les photoMOS collectent les charges puis pendant la phase de lecture elles constituent des éléments mémoires qui permettent le transport vers le registre de transfert. Ce transport de proche en proche est réalisé grâce à la propriété qu'ont les électrons à s'accumuler dans le photosite le plus profond. Pour augmenter la profondeur des photosites, la tension appliquée à chaque photoMOS est augmentée et contrôlée par le générateur d'horloge. Pendant la phase de lecture, les photosites doivent impérativement être isolés de la lumière, d'où la conservation de l'obturateur mécanique sur les APN. Dans le cas contraire, le temps d'intégration diverge entre le premier et le dernier photosite de la matrice, ce qui fausse les mesures d'éclairement énergétique. Car en effet, si le temps de lecture n'est pas négligeable les photosites lus en dernier sont exposés plus longtemps.

Pour pallier à l'obligation d'un obturateur mécanique, d'autres structures séparant les

photo-détecteurs et le transfert des charges ont été élaborées. La première consiste à rajouter un registre de stockage de la même taille que la matrice photosensible. Cette structure appelée transfert de trame (figure 3.37b) double la surface de silicium utilisée. L'autre structure appelée interligne (figure 3.37c) consiste à ajouter après chaque ligne verticale une interligne servant au transfert et à isoler de la lumière. Malgré les solutions apportées, ces structures ne sont pas fréquentes car le principal défaut est de diminuer considérablement le facteur de remplissage des photosites et d'augmenter les coûts de fabrication.



(a) (b) (c)

FIGURE 3.37 – *Principales structures de transfert des charges dans un capteur CCD.*

En termes de performance, les capteurs CCD ne sont malheureusement pas parfaits. Le principal inconvénient provient du fait que chaque photosite n'est pas isolé électriquement des autres. Ainsi, lorsque la pleine capacité est atteinte par l'un, les charges accumulées débordent facilement vers les voisins ce qui fausse les mesures d'éclairement énergétique ponctuelles. Ce phénomène est appelé débordement (ou blooming) et se produit lorsque le capteur est exposé à un point très lumineux. Il se traduit par une tâche sur l'image résultante.

La technologie CCD présente un autre inconvénient lié à l'architecture qui oblige le transfert des charges. En effet, les signaux du générateur sont très complexes et nécessitent de très hauts niveaux de tension ($15V$) pour assurer un transfert efficace. Ces niveaux de tension sont incompatibles avec le standard actuel ($3V$), ce qui rend le CCD coûteux en terme de fabrication et de consommation énergétique malgré une qualité d'acquisition performante.

3.3.6.4 La structure APS-CMOS

Le CMOS (Complementary Metal Oxide Semi-conductor) n'est pas à proprement parler une structure de capteur photosensible mais une technique de lithographie servant à la conception des composants électroniques. Les premières structures MOS-XY utilisent des transistors directement lithographiés sur le substrat silicium, par conséquent ils peuvent être qualifiés de capteurs CMOS. Mais l'évolution technique du CMOS durant les années 80-90 a permis d'améliorer considérablement la finesse de gravure (0.35 $\mu m/transistor$) ce qui a entrainé l'apparition d'un nouveau concept baptisé APS (Active Pixel Sensor).

L'APS repose sur l'idée d'isoler électriquement chaque photosite afin d'éliminer les inconvénients liés aux transferts des charges (CCD) ou aux impédances du bus analogique (MOS-XY, CID). Pour cela, dans un premier temps chaque photosite intègre un détecteur photosensible et son propre circuit de lecture (figure 3.38) sans pour autant nuire au facteur de remplissage grâce à la finesse de gravure. Les détecteurs photosensibles peuvent aussi

bien être des photodiodes que des photoMOS indépendantes. Ils sont lus par deux décodeurs d'adresses X et Y comme la structure MOS-XY, mais sans altération du signal due à l'impédance du bus analogique. De plus, les photo-détecteurs présentent peu de disparité car ils sont directement fondus sur le même substrat, ce qui réduit considérablement le bruit spatial fixe $\mathcal{B}_{FPN}$.

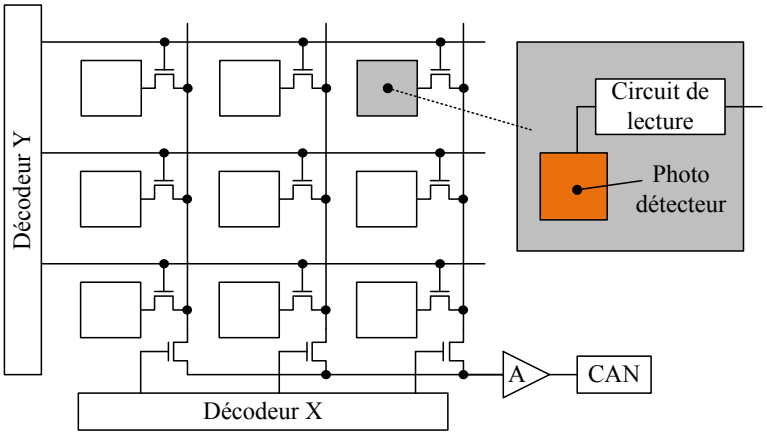


FIGURE 3.38 – *Structure première génération de capteur APS – CMOS.*

Plus récemment, la lithographie CMOS a évolué et la finesse de gravure s'est encore améliorée. Les photosites sont désormais capables d'intégrer indépendamment l'amplificateur et le CAN (figure 3.39) pour une efficacité quasi-égale au CCD et une consommation d'énergie considérablement réduite.

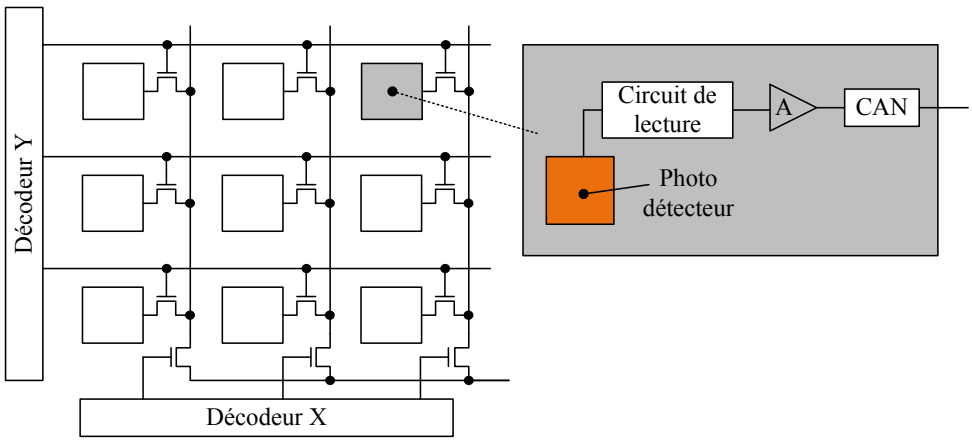


FIGURE 3.39 – *Structure récente de capteur APS – CMOS.*

3.3.6.5 Bilan technologique

Bien que révolutionnaire, la structure APS-CMOS fut tout d'abord jugée moins performante que son ainée le CCD et resta longtemps cantonnée aux applications bas de gamme ne nécessitant pas une qualité d'image élevée. Cependant depuis ces dernières années, la donne a changé. La maturité des techniques de lithographie CMOS et les progrès réalisés dans le domaine électronique ont permis d'atteindre des performances équivalentes au CCD (tableau 3.1) pour une consommation très faible et un coût de fabrication toujours plus réduit.

A présent, les capteurs APS-CMOS sont devenus la référence que ce soit en vision industrielle ou grand public et le CCD est réservé à des usages marginaux ou des applications très particulières telles que l'astronomie.

Caractéristiques	CCD	APS-CMOS
Signal en sortie du photosite	charge	tension ou bits
Signal en sortie du capteur	tension	bits
Signal de sortie	bits	bits
Facteur de remplissage	proche 100%	$< 100\%$
Bruit : blooming	élevé	faible
Bruit : $\mathcal{B}_{DC}$ et $\mathcal{B}_{TH}$	faible	faible
Bruit : $\mathcal{B}_{FPN}$	faible	faible
Dynamique	élevée	modérée mais en augmentation
Temps de lecture	élevé	faible
Obturateur mécanique	obligatoire pour le pleine trame	pas obligatoire mais conseillé
Complexité photosite	faible	élevée
Complexité capteur	très élevée (horloge)	faible
Consommation d'énergie	très élevée	faible
Coût	faible	modéré mais en diminution

TABLE 3.1 – *Comparaison des caractéristiques et performances des structures de capteur CCD et APS-CMOS.*

3.3.7 Le capteur : outil métrologique

Les capteurs d'images peuvent être vus comme des outils métrologiques car ils permettent d'effectuer des mesures ponctuelles du nombre de photons reçus par une surface pendant un laps de temps défini. Les mesures rendues au format numérique peuvent être transcrites en valeurs radiométriques en fonction du temps suivant la formule :

$$M = crf(Ee.\Delta t) \tag{3.37}$$

avec :

M	:	la valeur numérique quantifiée.
Δt	:	le temps d'intégration.
Ee	:	l'éclairement énergétique (irradiance) reçue par un photosite.

La transformation crf (Camera Response Function) correspond à l'ensemble des étapes de quantification du capteur et des perturbations liées aux bruits, aux non linéarités et à l'échantillonnage numérique.

L'éclairement énergétique ponctuelle Ee reçu par le photosite est directement liée à l'intensité ponctuelle émise par un point de la scène suivant la relation établie par l'équation 3.14. Par conséquent les valeurs numériques M peuvent être définies comme :

$$M = crf(k.Ie.\Delta t) \tag{3.38}$$

avec :

M	:	la valeur numérique quantifiée.
Δt	:	le temps d'intégration.
Ie	:	intensité ponctuelle de la scène.
k	=	$\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4\theta$

De plus, il s'avère nécessaire de rappeler que l'éclairement énergétique quantifiée par le photosite n'est pas monochromatique. Il correspond proportionnellement à l'intégration du flux énergétique polychromatique incident en fonction de l'efficacité quantique du silicium, soit :

$$Ee \propto \int_{\lambda=200\ nm}^{1100\ nm} QE(\lambda) \times \frac{Fe(\lambda)}{\mathrm{d}S} \quad \mathrm{d}\lambda \tag{3.39}$$

avec :

Ee	:	l'éclairement énergétique reçu par un photosite.
$QE(\lambda)$	:	l'efficacité quantique pour λ.
$Fe(\lambda)$	:	le flux reçus pour λ.
$\mathrm{d}S$	=	la surface du photosite en (m^2).

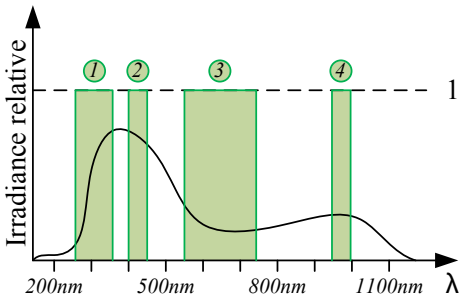
Malheureusement, cette intégration entraîne la perte totale de l'information chromatique (λ). Néanmoins, le principe d'acquisition multi-spectrale permet d'échantillonner les longueurs d'onde et de conserver une partie de l'information spectrale. La mise en oeuvre de ce principe repose sur l'utilisation de filtres passe-bandes interchangeables et l'acquisition multiple d'une même scène. Chaque acquisition permet de mesurer l'éclairement énergétique perçu par le photosite pour une bande spectrale ($\Delta\lambda$) plus ou moins large.

$$M_{\Delta\lambda} = crf(Ee_{\Delta\lambda}.\Delta t) \tag{3.40}$$

$$Ee \propto \int_{\lambda_{min}}^{\lambda_{max}} QE(\lambda) \times \frac{Fe(\lambda)}{\mathrm{d}S} \quad \mathrm{d}\lambda$$

Ce principe (figure 3.40) est notamment utilisé par des instruments métrologiques tels que le spectromètre à capteur CCD où de multiples filtres très étroits sont mis en œuvre pour couvrir une plage spectrale la plus large possible. Un tel dispositif est basé sur une partie mécanique qui déplace successivement les filtres face à l'objectif. Ce type de système est très encombrant et nécessite une scène statique durant les acquisitions. Pour pallier au problème

d'encombrement, des filtres à cristaux liquides ont été réalisés. Ces filtres nommés LCTF (Liquid Cristal Tunable Filter) permettent une sélection de la longueur d'onde avec une bande étroite et réglable (50 $nm < \Delta\lambda < 500\ nm$) pour une large plage (100 $nm < \lambda < 1100\ nm$). De plus, ils admettent deux avantages qui sont l'élimination de la partie mécanique et la possibilité de contrôle par logiciel. Mais le temps d'acquisition reste trop élevé pour autoriser la capture de scènes dynamiques.



filtre	λ_{min}	λ_{max}	$\Delta\lambda$	$Ee_{\Delta\lambda}$
1	*250nm*	*350nm*	*100nm*	*0.55*
2	*400nm*	*450nm*	*50nm*	*0.72*
3	*550nm*	*750nm*	*200nm*	*0.25*
4	*950nm*	*1000nm*	*50nm*	*0.24*

FIGURE 3.40 – *Échantillonnage spectrale par plusieurs filtres.*

Pour ce qui est de l'APN, le but est d'effectuer une acquisition adaptée au SVH. La synthèse additive des couleurs mise en évidence par YOUNG puis HELMHOLTZ démontre qu'à partir de trois couleurs primaires RVB (Rouge, Vert et Bleu) l'ensemble des couleurs perceptibles sont reproductibles pour le SVH. C'est pourquoi, les APN n'ont besoin que de trois filtres centrés sur les longueurs d'onde correspondant aux couleurs primaires. Ainsi l'énergie quantifiée dans ces trois bandes permet par addition sur support papier ou virtuel de stimuler le SVH en retranscrivant l'ensemble des couleurs de la scène. En effet, même si une différence physique existe entre le jaune dont la longueur d'onde est 580 nm et le mélange rouge/vert, les cellules du SVH sont stimulées de façon semblable ce qui permet à l'observateur de ne pas percevoir la différence.

La limitation du nombre de filtres dans l'APN permet d'accélérer considérablement le temps d'acquisition. Mais de tels dispositifs restent tout de même encombrants et coûteux. Par conséquent, des techniques simples ont été développées telles que le tri-CCD, le filtrage intégré au capteur ou des structures semi-conductrices à triples jonctions enterrées.

3.3.7.1 Les dispositifs tri-CCD

Ces dispositifs sont composés de trois capteurs distincts pour chaque couleur primaire. Les rayons pénètrent dans une série de prismes et sont filtrés par des miroirs dichroïques. Chacun des trois filtres laisse passer une couleur vers les capteurs associés (figure 3.41).

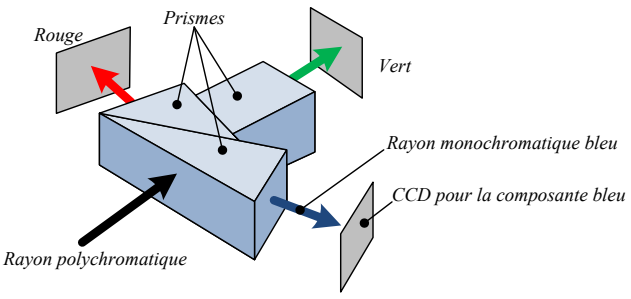


FIGURE 3.41 – *Dispositif tri-CCD.*

La principale difficulté dans ce type de dispositif réside au niveau de la réalisation car les trois images doivent être parfaitement alignées. Le fait d'utiliser trois capteurs et des processus de fabrication exigeants augmentent considérablement le coût de production. Ces dispositifs sont donc réservés à un usage professionnel et surtout à la vidéo car le temps d'acquisition est réduit au temps d'intégration des capteurs.

3.3.7.2 Les dispositifs à filtrage intégré

Même si les dispositifs tri-CCD sont très performants, ils restent trop onéreux. Dans ce contexte, des dispositifs mono-capteur ont été réalisés. Le système de filtrage (figure 3.42) est directement disposé sur le capteur grâce à l'application de filtres colorés CFA (Color Filter Array) sur chaque photosite.

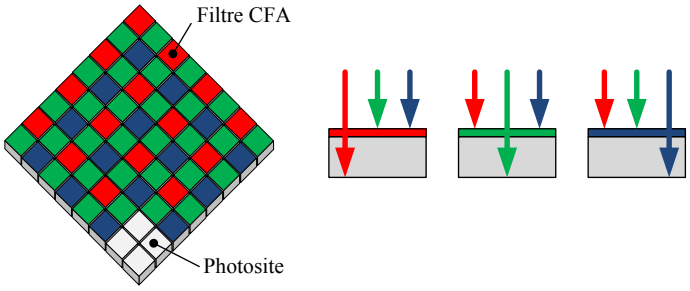


FIGURE 3.42 – *Dispositif à filtre intégré, CFA Bayer.*

Ainsi chaque photosite ne reçoit qu'une plage de longueurs d'onde donnée, ce qui implique un sous-échantillonnage spatial. Il existe beaucoup de combinaisons pour les filtres CFA, le plus fréquent est celui de Bayer [Bay76] où un pixel RVB est composé de quatre photosites filtrés (2 verts, 1 bleu et 1 rouge). Il y a plus de vert pour tenir compte du fait que le SVH est plus sensible à cette longueur d'onde.

Ces dispositifs sont aussi rapides que le tri-CCD et les limitations sont liées au fait que quatre photosites sont nécessaires pour obtenir l'information couleur. Afin de pallier ce problème, des constructeurs ont proposé de modifier la géométrie des photosites. Par exemple, les capteurs FIDJI appelés SuperCCD ont des photosites octogonaux. Selon les constructeurs, contrairement aux photosites carrés, l'organisation de l'espace permet d'obtenir des résolutions proches des caractéristiques perceptuelles du SVH et une sensibilité accrue. Mais scientifiquement, le gain n'est pas prouvé.

3.3.7.3 Les dispositifs à triples jonctions enterrées

Ces derniers dispositifs sont issus d'une innovation technologique proposée par la société FOVEON. Elle propose d'utiliser la profondeur du silicium comme un filtre car la profondeur de pénétration d'un rayonnement est directement liée à la sa longueur d'onde. La capture de chaque couleur primaire se fait par empilement de trois photodiodes et par comparaison des courants générés par chacune.

Cette technologie baptisée X3 (figure 3.43) permet de conserver la résolution originale des capteurs, mais il semble que la séparation des longueurs d'onde ne soit pas parfaite. Les

images résultantes doivent subir un lourd traitement électronique et numérique, ce qui exclut l'utilisation de ces dispositifs dans le cadre de la métrologie optique.

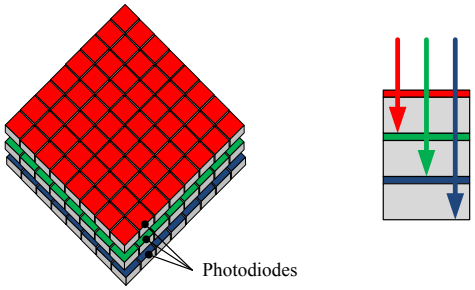


FIGURE 3.43 – *Dispositif à triples jonctions enterrées, FOVEON X3.*

3.3.7.4 Bilan

Dispositifs	Avantages	Inconvénients
Triple filtrage séquentiel	– unique capteur – filtres réglables si LCTF – résolution d'origine	– temps d'acquisitions élevé – réservé aux scènes statiques – encombrement si mécanique
Tri-CCD	– résolution d'origine – temps d'acquisition faible	– fabrication délicate – coût
Filtres intégrés aux capteurs (Bayer)	– temps d'acquisition faible – coût faible – intégration	– résolution d'origine réduite – création de fausses couleurs
Triples jonctions enterrées (FOVEON X3)	– temps d'acquisition faible – intégration – résolution d'origine	– coût de fabrication – mauvaise séparation des couleurs

TABLE 3.2 – *Comparatif des dispositifs de capture couleurs.*

D'après ce comparatif Avantages/Inconvénients (tableau 3.2), les dispositifs à filtres intégrés offrent le meilleur compromis. La technologie actuelle permet l'augmentation de la résolution, ce qui réduit l'effet du sous-échantillonnage et offre une qualité d'image exploitable pour la vision industrielle ainsi que la métrologie optique.

3.3.8 La gestion des données numériques

En sortie des capteurs à filtres intégrés CFA, les données numériques sont fournies en série pour chaque photosite suivant les trois composantes couleurs. Elles sont ensuite stockées sur des supports numériques sous forme de matrices. Selon le filtre utilisé, les matrices

images sont constituées d'une alternance de mesures quantifiées d'intensités en fonction des longueurs d'onde correspondant au rouge, vert et bleu.

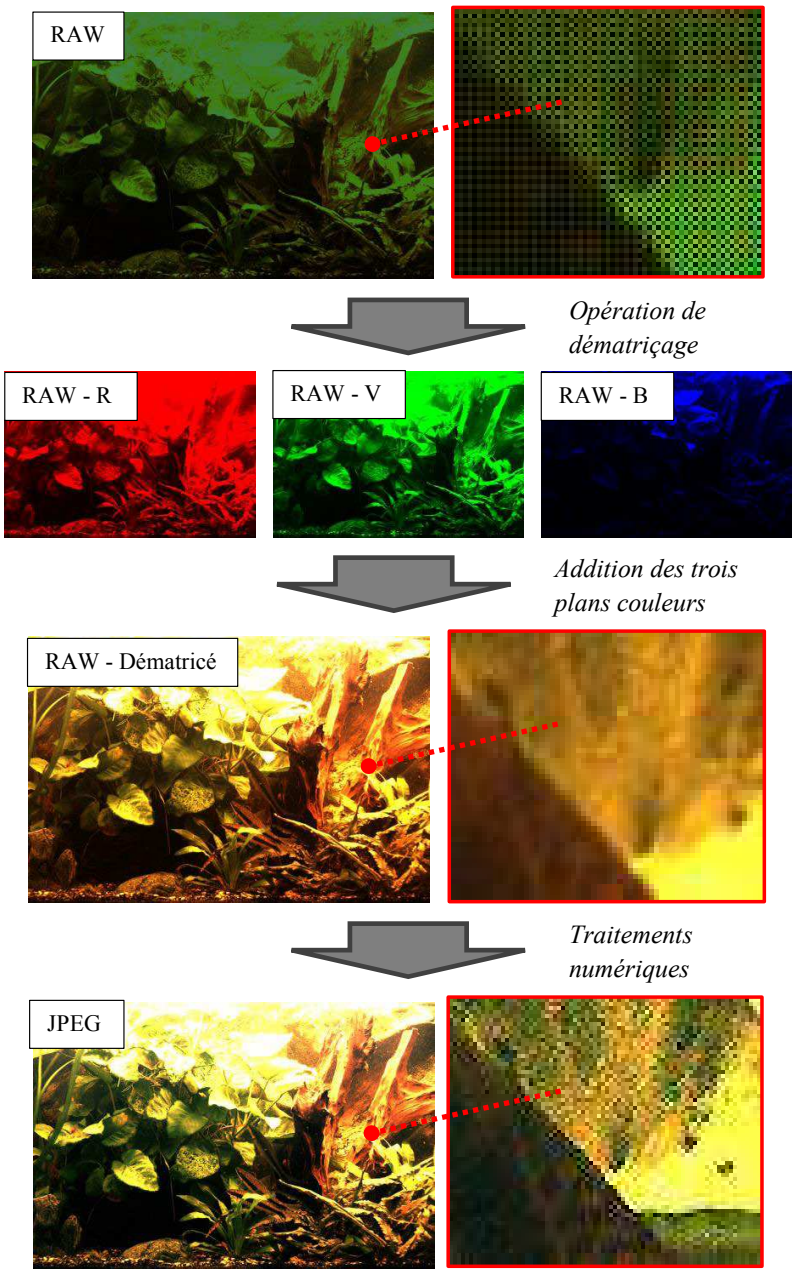


FIGURE 3.44 – *Dématriçage et traitement des fichiers images.*

Ces matrices sont enregistrées sous la forme de fichiers dont les formats sont normalisés. Parmi les formats existants certains conservent les données brutes issues du capteur (figure 3.44) et se nomment RAW. D'autres formats dont le plus connu est le JPEG traitent numériquement les données et les compressent avant de les stocker.

La perception et l'exploitation des images nécessitent une opération de dématriçage qui permet de convertir les matrices images issues du capteur CFA en trois matrices (1 par plan couleur) de taille identique (figure 3.44). De nombreux algorithmes existent, ALLYSON dans [Allo4] propose une revue de trente années de développement des méthodes de dématriçage utilisées pour la reconstruction des images couleurs. Il en ressort que deux principes sont

utilisés :

- soit le nombre de pixels par rapport au nombre de photosites est réduit (sous-échantillonnage spatial).
- soit la résolution des photosites est conservée et les composantes manquantes sont interpolées.

Contrairement au format RAW, les autres formats de fichiers inclut cette opération de dématriçage ainsi que divers traitements numériques visant à améliorer le rendu visuel des images. Mais ces traitements induisent un grand nombre de non linéarités aléatoires et faussent les mesures. Par conséquent, le format RAW beaucoup plus souple d'utilisation est préféré pour les applications dans le domaine de la métrologie.

Conclusion du chapitre

Le processus d'acquisition d'image consiste à projeter une scène tridimensionnelle sur un plan. Géométriquement, cette projection entraîne la perte d'une dimension. Dans le cas des dispositifs d'acquisition, elle est réalisée par l'objectif qui se compose d'une ou plusieurs lentilles convergentes ou divergentes. Le plan de projection est alors une surface photosensible subdivisée en éléments appelés pixels dont le but est de capturer et quantifier l'énergie lumineuse émise par la scène. Ainsi, l'appareil photographique numérique permet de réaliser des images où chaque pixel est une mesure quantifiée correspondant à l'intensité émise par une zone tridimensionnelle de la scène. Néanmoins, la projection, l'acquisition et la quantification sont des processus imparfaits. Ils entraînent des perturbations et des pertes, et sont limités par la technologie des capteurs. Mais grâce à la caractérisation de chaque élément, nous sommes en mesure de comprendre et de modéliser mathématiquement ces divers phénomènes.

La Reconstruction 3D par vision 4

Sommaire

4.1 État de l'art des méthodes . . . 66
4.1.1 Analyse du temps de vol . . . 66
4.1.2 Analyse des déformations d'un signal lumineux . . . 66
4.1.3 Analyse des intensités énergétiques mesurées . . . 67
4.2 La Stéréo-Photométrie . . . 71
4.2.1 Extraction de l'information . . . 72
4.2.2 Intégration des champs de gradients . . . 74
4.3 Ombres et Spécularités dans la SP . . . 80
4.3.1 Élimination des ombres et points spéculaires (s>3) . . . 82
4.3.2 Notre méthode d'élimination . . . 85
Conclusion . . . 95

D'après les chapitres précédents, l'interaction entre la lumière et la matière est mathématiquement modélisable. Les intensités émises par la surface peuvent être déterminées en fonction du système d'éclairage ainsi que des propriétés photométriques et géométriques de la surface analysée. Ces intensités sont mesurables par l'intermédiaire d'un dispositif photographique. Néanmoins, le processus d'acquisition des images entraîne la perte d'une dimension et n'est pas parfait.

Ce chapitre est consacré à l'estimation de la dimension perdue afin d'obtenir des reconstructions tridimensionnelles. Elles doivent être le plus fidèle possible par rapport aux surfaces réelles pour permettre l'application d'algorithmes de mesures métrologiques. Ces derniers ont pour but d'extraire des caractéristiques surfaciques telles que la rugosité ou d'éventuels défauts (fissures, trous, etc).

4.1 État de l'art des méthodes

Les méthodes de reconstruction 3d sans contacts sont variées. Dans la littérature, elles sont généralement classées suivant deux familles. La première nommée "active" comporte des systèmes émettant un signal connu pour réaliser des mesures dans l'environnement 3d. L'énergie employée pour le signal peut être lumineuse, sonore ou tout autre onde électromagnétique. L'autre famille se nomme "passive" et ne fait que recevoir des informations de l'environnement. Cette technologie est principalement basée sur les capteurs optiques pour réaliser les mesures.

Dans le cadre de nos travaux, l'intérêt est porté uniquement sur les méthodes qui utilisent des capteurs optiques. Dans la littérature, ces dispositifs sont utilisés dans les deux familles. Par conséquent, nous différencions ces méthodes par rapport à l'analyse de l'information mesurée par le capteur optique. Elles sont classées suivant trois catégories : analyse du temps de vol, analyse des déformations d'un signal lumineux émis et analyse des intensités énergétiques mesurées.

4.1.1 Analyse du temps de vol

Le principe consiste à émettre un signal lumineux ponctuel en direction de la surface à analyser puis d'attendre que ce signal soit réfléchi et mesuré par le capteur. La différence temporelle entre l'émission et le retour du signal est appelée temps de vol. Connaissant précisément la position de l'émetteur et du récepteur et la vitesse du signal, la distance relative du point mesuré peut être calculée. Pour obtenir toute une surface, ce type d'analyse nécessite un balayage suivant au minimum deux axes dans le cas de la topographie.

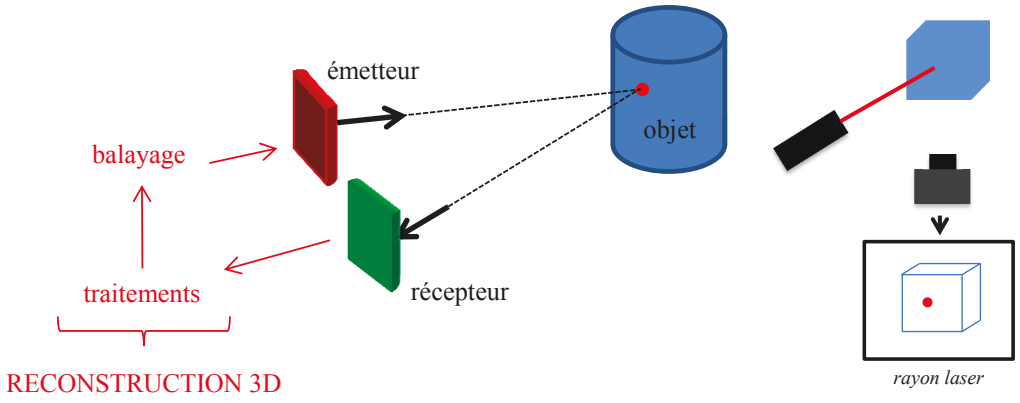


FIGURE 4.1 – *Schéma de principe du temps de vol entre émetteur et récepteur.*

En dehors des capteurs optiques, ce principe est identique au RADAR ou au télémètre à ultrasons. Son principal défaut est le balayage. En effet, il est nécessaire d'effectuer un grand nombre de mesures pour avoir une bonne résolution des reconstructions 3d. Hormis ce point négatif, les mesures par temps de vol sont d'une haute qualité.

4.1.2 Analyse des déformations d'un signal lumineux

Cette technique implique l'émission d'un motif lumineux dont la géométrie est parfaitement connue. Lors de son interaction avec la surface à analyser, la géométrie du motif subit des déformations qui sont enregistrées par le capteur. L'analyse des déformations permet une

estimation des surfaces en 3d. Parmi les méthodes existantes, les plus communément rencontrées sont la projection d'un ou plusieurs plans laser (figure 4.2a) et la projection d'une lumière structurée. Avec la projection d'un plan laser, le balayage de la surface à analyser doit être fait suivant un axe. Dans le cas de plusieurs plans laser ou de la lumière structurée 2d le balayage n'est pas nécessaire. Comparé à la mesure ponctuelle, ces méthodes réduisent le nombre de prise de mesures car elles permettent d'obtenir un reconstruction avec une unique acquisition.

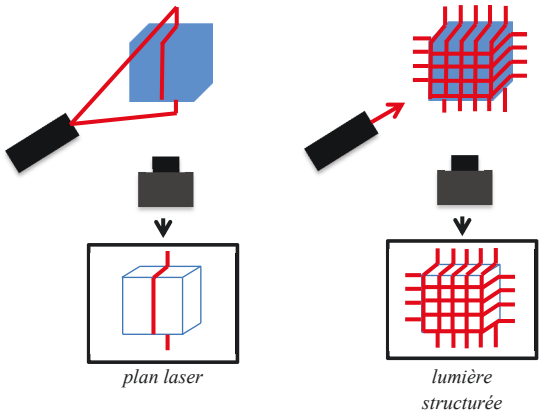


FIGURE 4.2 – *Schéma de principe de la profilométrie (a) et de la lumière structurée 2d (b).*

4.1.3 Analyse des intensités énergétiques mesurées

Ce type d'analyse répond à nos contraintes d'application car il ne nécessite pas de balayage ni d'un système d'éclairage complexe et coûteux. L'estimation des surfaces est directement calculée à partir du traitement des valeurs numériques fournies par le capteur d'images. Les méthodes existantes forment un ensemble appelé "Shape From X" où "X" désigne une caractéristique particulière tel que : multi-vues, contours, ombres (shading), focus etc.

4.1.3.1 Le Shape from Shading

Le Shape From Shading (SFS) consiste à déterminer la géométrie 3d d'une surface par l'analyse des valeurs d'une unique image. Cette technique repose sur l'utilisation d'un modèle de réflectance généralement Lambertien et d'une source lumineuse dont la position est connue ou estimée. Dans ce contexte, chaque valeur de l'image en niveau de gris dépend de la normale à la surface en un point donné et de la position de la source lumineuse.

Initialement formulé par HORN [Hor75], le SFS se décompose en deux étapes. La principale consiste à déterminer deux champs de gradients suivant les axes x et y. Ces champs de gradients notés p et q représentent les dérivées partielles de la surface ζ suivant les axes x et y tel que :

$$p = \frac{\mathrm{d}\zeta}{\mathrm{d}x} \quad et \quad q = \frac{\mathrm{d}\zeta}{\mathrm{d}y} \tag{4.1}$$

Puis la seconde étape consiste à intégrer numériquement les champs de gradients pour obtenir une estimation de la profondeur.

Dans le cas du modèle de réflectance Lambertien, les hypothèses simplificatrices sont :

- la valeur de l'albédo et l'intensité de la source lumineuse incidente sont considérées comme constantes en tous points de la surface et de valeur connue.
- le modèle de projection est supposé orthographique et le centre optique coïncide avec l'axe z de la scène.
- les valeurs numériques M de l'image sont proportionnelles aux intensités Ie des points de la scène.

Dans ces conditions l'équation principale du SFS est :

$$M \propto (\vec{n}.\vec{l}) \tag{4.2}$$

avec :

$\vec{n} = \frac{1}{\sqrt{1+p^2+q^2}} \begin{bmatrix} -p \\ -q \\ 1 \end{bmatrix}$	:	la normale à la surface en un point donné.
$\vec{l} = (lx, ly, lz)^T$	:	la direction connue de la source lumineuse.

Tel quel, le SFS est qualifié de problème mal posé car il consiste à résoudre deux inconnues (p et q) à partir d'une unique équation. Pour pallier à ce problème, des contraintes supplémentaires sont ajoutées afin de reformuler l'équation 4.2. La thèse de Daniel [Danoo] met en évidence certaines des contraintes nécessaires à faire du SFS un problème bien posé et résolvable à partir d'une équation aux dérivées partielles (EDP). Néanmoins, le SFS n'est pas applicable sur des surfaces à fortes variations géométriques ou à albédo variable. De plus, il est impossible de discerner si une surface est concave ou convexe (figure 4.3) comme l'a démontré judicieusement Pentland [Pen84].

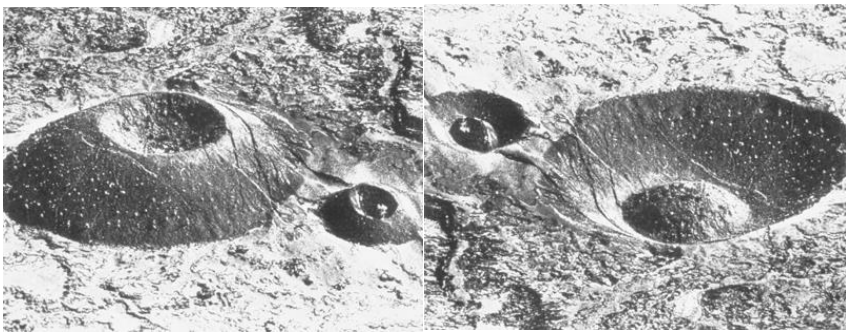


Figure 4.3 – *Démonstration de l'ambigüité concave/convexe faite par* Pentland, *photo de* Whitman Richards, *à droite deux monticules de terres sont visibles, après retournement (à gauche) il est possible de distinguer deux cratères.*

Récemment, Prados et Fougeras [PFao4, PFao6] ont reformulé le problème et apporté des résultats plus concluants. Durou dans [DCD+11] apporte de nouvelles solutions pour résoudre le problème de l'ambigüité concave/convexe.

4.1.3.2 Le Shape from Stéréo-Photométrie

Le Shape From Stéréo-Photométrie ou couramment appelé Stéréo-Photométrie (SP) consiste à reconstruire une surface 3d à partir d'une série d'images acquises d'un point de vue fixe mais avec des conditions différentes d'éclairage pour chaque acquisition. Introduite en 1980 par Woodham [Woo80], elle se base sur le principe du SFS tout en cherchant à éliminer les contraintes du problème. Dans le cas Lambertien, Woodham part du principe que trois équations doivent être utilisées pour déterminer les trois inconnues qui sont les champs de gradients p et q ainsi que l'albédo. Par conséquent, un minimum de trois images pour trois directions d'éclairage non coplanaires est nécessaire pour poser et résoudre correctement le problème. Tout comme le SFS, la SP nécessite l'intégration numérique des champs de gradients afin d'obtenir une estimation $\hat{\zeta}$ du relief de la surface analysée.

4.1.3.3 Le Shape from Contours

Le Shape From Contours (SFC) est une variante du SFS et consiste à contraindre le problème par l'extraction des points correspondant aux contours d'un objet. En effet, grâce à la connaissance du modèle de projection, la géométrie des points de contours est connue ce qui permet de contraindre l'estimation des champs de gradients à une unique estimation possible. Initialement formulé pour des surfaces cylindriques [RDLR91], Durou et Courteille [CCDG07] l'utilisent pour la remise à plat de documents scannés (figure 4.4). Néanmoins, le SFC souffre des mêmes défauts que le SFS et nécessite des surfaces simples avec des contours marqués.

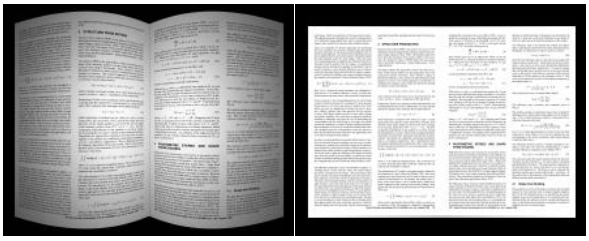


Figure 4.4 – *Exemple d'application du SFC pour la remise à plat de document, ces images sont tirées de la publication de* Durou *et* Courteille *[CCDG07].*

4.1.3.4 Le Shape from Texture

Le Shape From Texture (SFT) utilise aussi une unique image et a pour objectif l'exploitation des informations contenues dans la texture des surfaces analysées. En effet, lorsqu'un observateur humain visualise une image texturée, il est capable d'interpréter virtuellement la forme de la surface par analyse des déformations (figure 4.5). Le SFT s'inspire de cette idée et par de lourds algorithmes, il permet d'interpoler une surface [CM02]. Cette méthode très particulière est applicable uniquement pour des surfaces fortement et régulièrement texturées.

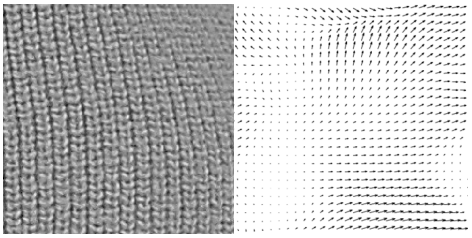


Figure 4.5 – *Exemple de d'une surface texture et de son champs de normales permettant l'interpolation du relief, images tirées de [CM02].*

4.1.3.5 Le Shape from Focus/Defocus

Le Shape From Focus (SFF) regroupe l'ensemble des méthodes de reconstruction basées sur l'analyse de la profondeur de champ ou du flou optique. Ces méthodes utilisent la modification de la distance focale pour mesurer les distances entre la surface et le point de référence correspondant au centre optique de l'appareil d'acquisition. En effet, lors du processus de formation de l'image, seuls les points sur le plan de mise au point apparaissent nets sur l'image. Ainsi il est possible de déterminer la hauteur de chacun par une série d'images dont le nombre dépend de la résolution en z désirée et de l'élévation totale de la surface.

La différence entre Focus et Defocus (figure 4.6) provient de la technique de modification de la distance focale entre les images acquises. Dans le premier cas, la mise au point est réalisée dynamiquement au cours de l'acquisition afin de s'adapter à la surface analysée [NN94]. Dans le second cas, les plans de mise au point sont fixés à intervalles réguliers [FS05].

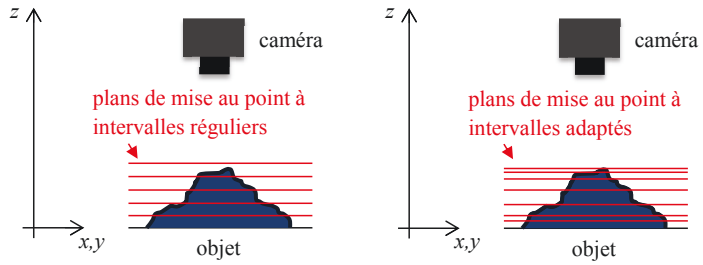


FIGURE 4.6 – *Schémas du principe de SFF, (à droite) méthode de Focus les pas de mise au point s'adapte à la surface analysée, (à gauche) méthode de Defocus les pas sont fixés à intervalles réguliers.*

Cette méthode nécessite une extrême précision sur le contrôle de la distance focale f et elle est généralement réservée à l'analyse des surfaces de très petites tailles.

4.1.3.6 Le Shape from Stéréo-Vision

Le Shape From Stereo-Vision couramment appelé Stéréo-Vision (SV) repose sur une triangulation passive. Le principe consiste à utiliser au moins deux caméras pour obtenir différents points de vue d'une même surface afin de permettre le calcul des coordonnées 3d. Cette méthode implique la connaissance préalable des modèles de projection, de tous les paramètres intrinsèques et extrinsèques ainsi que les positions des deux caméras. En revanche, elle ne nécessite pas de conditions d'éclairage particulières contrairement aux autres méthodes de Shape from X. La SV nécessite la mise en correspondance des points ($\alpha 1$ et $\alpha 2$) sur les deux images d'un même point β de la surface (figure 4.7). Les coordonnées 3d du point β sont obtenues par triangulation aux centres optiques $C1$ et $C2$ des caméras.

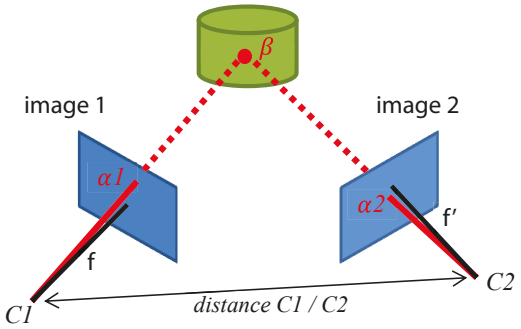


FIGURE 4.7 – *Schéma du principe de triangulation passive de la SV.*

Ici, la principale difficulté est la mise en concordance des points entre l'image de droite et celle de gauche. Ce problème peut toutefois être réduit grâce à l'utilisation de la Cross-Corrélation ou l'emploi de la géométrie épipolaire. Une variante, la SV N-images utilise plus de deux caméras afin de maximiser et faciliter la mise en correspondance des points. Mais il est fréquent qu'elle ne puisse pas être déterminée à cause d'une occultation ou d'une irrégularité géométrique trop importante. Cependant, cette méthode est très performante lorsque les surfaces possèdent des contours ou des points facilement identifiables. Mais elle nécessite un appariement très minutieux réalisé par la mise en œuvre des techniques de calibration. La SV N-images étant très coûteuse d'un point de vue matériel, une méthode dérivée appelée SFM (pour Shape From Motion) s'est développée. Elle consiste à utiliser une unique caméra qui effectue des prises de vue sous différents angles. Cette méthode impose de connaitre les déplacements entre les prises de vues ou de les estimer [WB92].

4.2 La Stéréo-Photométrie

Dans l'ensemble des méthodes de reconstruction 3d par vision, la stéréo-photométrie est particulièrement adaptée à notre contexte applicatif. En effet, elle ne nécessite pas de système d'éclairage complexe et coûteux contrairement aux méthodes basées sur l'analyse des déformations d'un motif lumineux. Parmi les autres méthodes de Shape from X, elle présente l'avantage d'extraire les propriétés colorimétriques et l'appariement est peu complexe par rapport aux méthodes de stéréo-vision. Enfin, à temps et taille de surface égale, elle permet d'obtenir une qualité de reconstruction équivalente aux méthodes basées sur l'analyse du temps de vol.

La stéréo-photométrie est donc utilisée comme méthode de référence dans la suite de cette étude. Comme l'illustre la figure 4.8, le but est d'obtenir une reconstruction 3d estimée à partir d'une série d'images acquises d'un point de vue fixe avec des configurations variables d'éclairage. L'estimation 3d repose sur l'inversion d'un modèle de BRDF (Bidirectional Reflectance Distribution Function). Deux étapes sont nécessaires, la première consiste à dissocier les constituants d'une image : géométrie (champs de gradients) et couleur d'une surface (albédo) ainsi que les phénomènes d'éclairage. La seconde étape consiste à intégrer numériquement les champs de gradients afin d'estimer une carte des hauteurs. Enfin la superposition de la carte des hauteurs et de l'albédo permet l'obtention d'une modélisation tridimensionnelle.

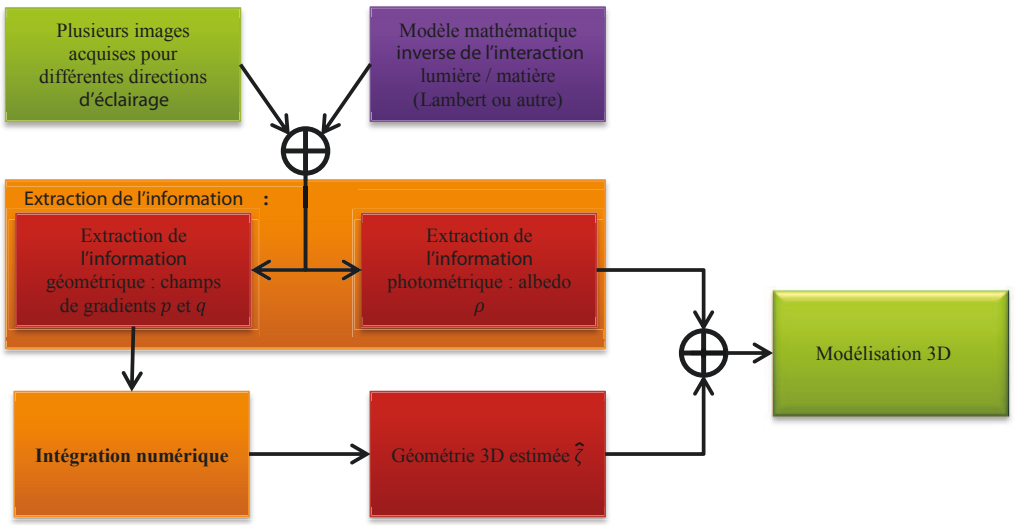


FIGURE 4.8 – *Organigramme décomposant les différentes étapes de la SP.*

Du point de vue des performance, cette méthode nécessite un temps d'acquisition plus élevé que le Shape from Shading mais il est largement compensé par la facilité de résolution du système. Comparée à la stéréo-vision, elle ne nécessite pas d'appariement car le point de vue est fixe et elle peut être appliquée à des surfaces fortement et aléatoirement texturées. De plus, elle permet l'analyse d'une grande zone comparée aux techniques de Shape from Focus ou aux méthodes actives telles que le temps de vol basé laser. Toutefois, l'application de la SP implique l'émission d'hypothèses restrictives afin de simplifier les calculs :

- $\mathcal{H}1$: le modèle de BRDF doit être connu et inversible.
- $\mathcal{H}2$: les directions des sources lumineuses doivent être connues par l'algorithme avant les calculs.
- $\mathcal{H}3$: les sources lumineuses doivent être parfaitement directionnelles et produire un éclairage uniforme (géométrie et énergie) sur l'ensemble de la surface analysée.
- $\mathcal{H}4$: les valeurs numériques sont des mesures exactes de l'intensité émise par les surfaces.

Ces hypothèses restreignent le domaine et les conditions d'application de la SP car dans le cas où elles ne sont pas respectées les reconstructions estimées souffrent de distorsions très importantes. Dans la suite de ce document, nous proposons des techniques permettant de lever ces hypothèses pour étendre le domaine d'application tout en conservant la rapidité et la simplicité de la SP. La suite de cette section est consacrée à l'étude et la comparaison des différentes méthodes d'extraction des champs de gradients/albédo et d'intégration.

4.2.1 Extraction de l'information

Extraire l'information géométrique et photométrique consiste à inverser un modèle de BRDF. Sous l'hypothèse $\mathcal{H}4$, les valeurs M fournies par la série d'images correspondent aux intensités Ie en chaque point de la surface pour une direction d'éclairage particulière (équation 4.3). La fonction crf est considérée comme linéaire, k et Δt ne varient pas entre les acquisitions et sont donc négligés.

$$M_s = crf(Ie_s\ k\ \Delta t) \tag{4.3}$$

Dans la formulation originale de la SP, Woodham se base sur l'inversion du modèle de Lambert. Ce modèle comporte trois inconnues qui sont les champs de gradient p et q ainsi que l'albédo ρ. Pour extraire l'information, le système doit comporter au minimum trois équations pour être résolvable. Par conséquent, il utilise trois images en niveaux de gris acquises suivant trois directions d'éclairage non coplanaires.

$$M_s \propto Ie_s = Ie_{0_s}\rho\left(\vec{n}.\vec{l}\right) \tag{4.4}$$

avec $\vec{n} = [-p\ -q\ 1]^T / \sqrt{(p^2+q^2+1)}$ la normale des champs de gradients en un point de la surface et $s = (1 \cdots S)$ le nombre de directions d'éclairage. Ie_{0_s} correspond à l'intensité incidente à la surface de la source lumineuse s mais l'hypothèse $\mathcal{H}3$ implique que les sources sont énergétiquement identiques. Par conséquent leurs intensités sont identiques et seules leurs directions varient (figure 4.9) et forment une matrice d'éclairage $[\mathbf{L}]$.

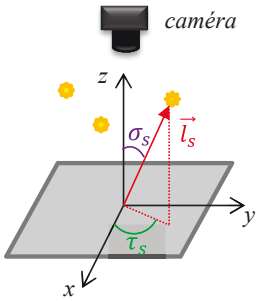


FIGURE 4.9 – *Schéma d'un système d'acquisition SP à trois sources lumineuses.*

$$[\mathbf{L}] = \begin{bmatrix} \vec{l}_1 \\ \vdots \\ \vec{l}_S \end{bmatrix}$$

avec :

$$\vec{l}_s = \begin{bmatrix} \cos\tau_s & \sin\sigma_s \\ \sin\tau_s & \sin\sigma_s \\ \cos\sigma_s \end{bmatrix}^T$$

où : τ_s est l'angle azimut et σ_s est l'angle zénith de la source lumineuse s.

Ainsi le modèle Lambertien peut s'écrire matriciellement :

$$M = \begin{bmatrix} M_1 \\ \vdots \\ M_S \end{bmatrix} = \rho \left([\mathbf{L}] . \begin{bmatrix} -p/\sqrt{(p^2+q^2+1)} \\ -q/\sqrt{(p^2+q^2+1)} \\ 1/\sqrt{(p^2+q^2+1)} \end{bmatrix} \right) \tag{4.5}$$

L'inversion et la résolution par moindre carrés permet d'obtenir une estimation des normales ainsi que de l'albédo :

$$\rho = |([\mathbf{L}]^T.[\mathbf{L}])^{-1}.[\mathbf{L}]^T.M|$$

$$\vec{n} = \frac{(([\mathbf{L}]^T.[\mathbf{L}])^{-1}.[\mathbf{L}]^T.M)}{\rho} \tag{4.6}$$

L'inversion du modèle Lambertien est simple et permet d'extraire rapidement l'information à partir d'un nombre restreint d'acquisitions. En revanche, il présente certaines limitations et n'est applicable qu'aux surfaces parfaitement diffuses et dont la BRDF est constante. Ces limitations peuvent être outre-passées grâce à des modifications discutées à la section 4.3 de ce chapitre.
La formulation de la SP est basée sur l'exploitation d'au moins trois intensités Ie pour chaque point de la surface et correspondant à une direction particulière d'éclairage. Dans le cas des images en niveaux de gris et de l'hypothèse $\mathcal{H}4$, chacune des trois images capturées fournie une valeur M proportionnelle à Ie. Mais les dispositifs d'acquisition ont évolué et fournissent actuellement des images multi-spectrales généralement sur trois bandes (rouge, verte et bleu). Ainsi pour trois directions d'éclairage chaque point de la surface est donné par une matrice $[\mathbf{M}]$ dont les lignes correspondent à la source d'éclairage s et les colonnes aux bandes capturées :

$$[\mathbf{M}] = \begin{bmatrix} M_{1_{rouge}} & M_{1_{vert}} & M_{1_{bleu}} \\ \vdots & \vdots & \vdots \\ M_{S_{rouge}} & M_{S_{vert}} & M_{S_{bleu}} \end{bmatrix} \tag{4.7}$$

Cette évolution technologique a fait évoluer le processus d'extraction SP suivant deux approches. La première consiste à conserver l'information colorimétrique dans l'albédo qui est

déterminé comme un vecteur $\vec{\rho}$ et les champs de gradients sont estimés à partir de la moyenne suivant les colonnes de la matrice $[\mathbf{M}]$. La seconde approche [HVB+07, BHK08] consiste à acquérir une unique image pour trois sources lumineuses dont la longueur d'onde d'émission diffère. Chaque source émet soit dans le spectre rouge, vert ou bleu. Ainsi les trois plans de l'image acquise fournissent des valeurs numériques pour une direction donnée (figure 4.10). La résolution est identique à la méthode originale mais une seule image est nécessaire pour l'inversion, ce qui permet l'analyse des surfaces dynamiques.

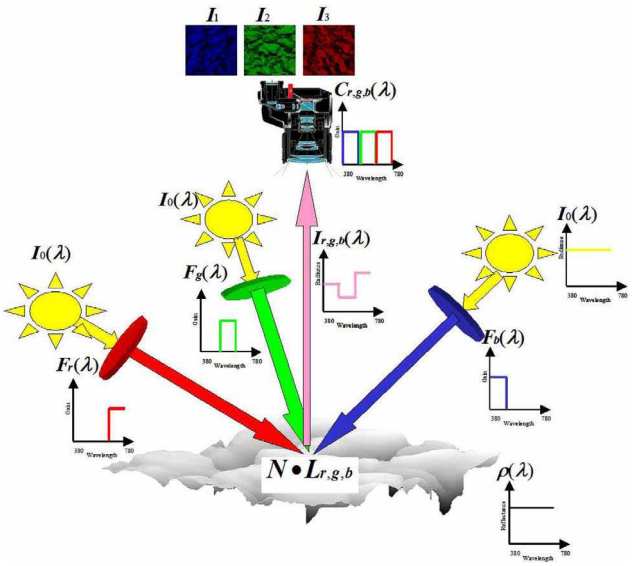


FIGURE 4.10 – *Schéma du principe proposé par [BHK08] de stéréo-photométrie couleur. Trois lumières blanches avec trois filtres spectraux différents éclairent la surface. Ensuite, le filtre de Bayer de la caméra produit trois images spectrales isolées. La surface 3D est estimée avec ces images selon la méthode de la stéréo photométrique classique.*

4.2.2 Intégration des champs de gradients

Au cours de la section précédente, nous avons abordé l'extraction de l'information pour obtenir les champs de gradients p et q de la géométrie des surfaces analysées. Pour obtenir une reconstruction 3d, cette première étape doit être complété par une intégration numérique pour obtenir une estimation de la carte des hauteurs $\hat{\zeta}$. La suite de cette section est consacrée à la présentation et la description de différentes méthodes d'intégration numériques existantes. Une comparaison est donnée en annexe A.3. Ces méthodes sont généralement répertoriées suivant trois catégories :

- locale : COLEMAN et JAIN [CJ82], RODEHORST [Rod93], WU et LI [WL88], ...
- relaxation : IKEUCHI [INI83], PRES [Pa02], ...
- globale : FRANKOT et CHELLAPPA [FC88], WEI et KLETTE [WK02], ...

Pour rappel, la phase d'intégration consiste à déterminer une estimation de l'élévation $\hat{\zeta}$ des surfaces suivant l'axe z du repère $\mathcal{R}_{scène}$ aligné avec l'axe optique de l'appareil d'acquisition. Ce calcul utilise les champs de gradients $p = \mathrm{d}\zeta/\mathrm{d}x$ et $q = \mathrm{d}\zeta/\mathrm{d}y$ de la surface réelle ζ.

L'ensemble des valeurs estimées est représenté par une matrice de réels et de taille $U \times V$ *pixels*. Ainsi chaque valeur est indexable suivant la convention présentée figure 4.11 et le résultat de l'estimation est appelé carte des hauteurs estimées.

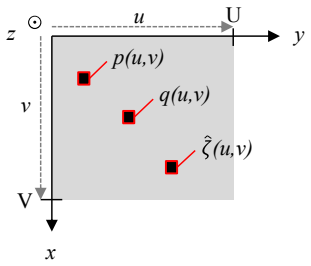


FIGURE 4.11 – *Convention d'indexation des champs de gradient et de la carte des hauteurs estimées.*

Pour que les champs de gradients soient intégrables, il est nécessaire et suffisant que p et q vérifient l'égalité de SHWARTZ :

$$\frac{\mathrm{d}p}{\mathrm{d}y} = \frac{\mathrm{d}q}{\mathrm{d}x} \tag{4.8}$$

Néanmoins les champs de gradients ne sont jamais parfaitement intégrables à cause des erreurs numériques dues aux calculs et surtout des imperfections liées aux hypothèses applicatives $\mathcal{H}1$ à $\mathcal{H}4$ de la SP. En fonction des méthodes d'intégration, il existe deux solutions. La première consiste à utiliser des chemins d'intégration différents et d'utiliser la moyenne pour les méthodes locales. La deuxième consiste à approcher la solution du problème au sens des moindres carrés, c'est le cas des méthodes basées sur la relaxation et les méthodes globales.

4.2.2.1 Méthodes Locales

Les méthodes locales ne reposent pas sur traitement de l'ensemble de la surface mais sur un traitement ponctuel des champs de gradients. L'ensemble de la carte des hauteurs est obtenue suivant le parcours des matrices, appelé chemin d'intégration. Ces méthodes reposent sur le développement en série de TAYLOR où la surface ζ est assimilée à une fonction $ft(z)$ telle que :

$$ft(z+\mathrm{d}z) = ft(z) + ft'(z)\ \mathrm{d}z + \frac{1}{2}ft''(z)\ \mathrm{d}z^2 + \cdots \tag{4.9}$$

La méthode des carrés repose sur le développement au premier ordre de cette série de TAYLOR suivant les axes x et y. La méthode des trapèzes s'obtient quant à elle de la même manière mais en conservant le développement au second ordre :

$$\begin{aligned} \textit{carré} \quad &: \quad \hat{\zeta}(u,v) = \frac{\hat{\zeta}(u-1,v) + \hat{\zeta}(u,v-1)}{2} - \frac{p(u-1,v) + q(u,v-1)}{2} \\ \textit{trapèze} \quad &: \quad \hat{\zeta}(u,v) = \frac{\hat{\zeta}(u-1,v) + \hat{\zeta}(u,v-1)}{2} - \frac{p(u,v) + p(u-1,v) + q(u,v) + q(u,v-1)}{2} \end{aligned} \tag{4.10}$$

D'après ces équations, il est nécessaire d'initialiser la première ligne et la première colonne de la carte des hauteurs estimées. Afin de satisfaire cette contrainte, $\hat{\zeta}(0,0)$ est fixé à 0 et la première ligne ainsi que la première colonne sont calculées dans le cas monodirectionnel. Ensuite l'ensemble de la carte est obtenue par un balayage suivant les colonnes et ligne après ligne. Ce chemin d'intégration a pour effet de propager et de cumuler les erreurs de reconstruction d'un bout à l'autre de la carte (figure 4.12).

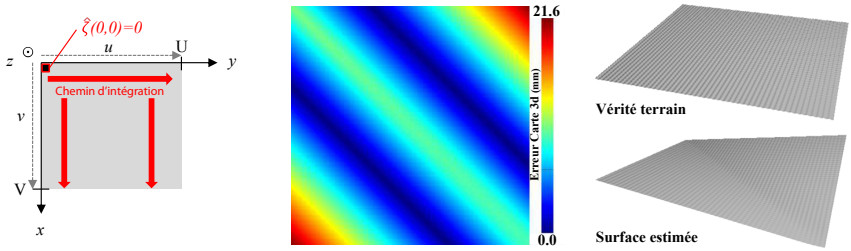


FIGURE 4.12 – *Chemin d'intégration des méthodes carré et trapèze ainsi qu'un exemple de propagation de l'erreur d'intégration pour une simulation de surface micro-rugueuse.*

Pour limiter ce problème, la méthode de COLEMAN et JAIN [CJ82] utilise l'intégration locale carré mais modifie le chemin d'intégration. Les auteurs préconisent une initialisation du centre de la carte et un balayage ligne/colonne suivant quatre directions. Ces différents chemins d'intégration ont pour effet de réduire l'erreur due à la propagation (figure 4.13).

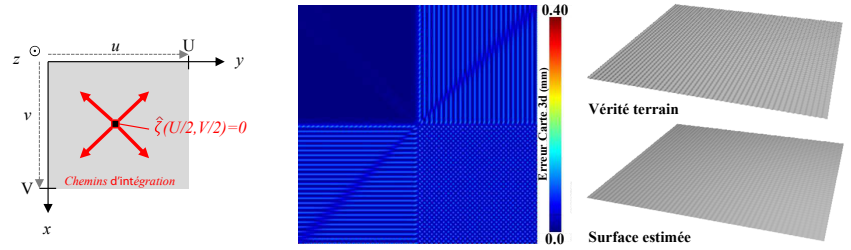


FIGURE 4.13 – *Chemin d'intégration proposé par* COLEMAN *et* JAIN *ainsi qu'un exemple de propagation de l'erreur d'intégration pour une simulation de surface micro-rugueuse.*

RODEHORST a proposé une autre solution [Rod93] pour réduire la propagation de l'erreur d'intégration. Sa méthode initiale préconise de suivre quatre chemins d'intégration partant des quatre coins de la carte et de calculer la moyenne. Cette technique oblige l'initialisation des quatre coins et exclut la reconstruction des surfaces dont l'élévation des coins est inconnue. Néanmoins une version modifiée permet la reconstruction grâce à un recalage des coins après intégration des quatre chemins. L'erreur d'intégration n'est pas nulle mais sa propagation est unifiée sur l'ensemble de la carte estimée (figure 4.14).

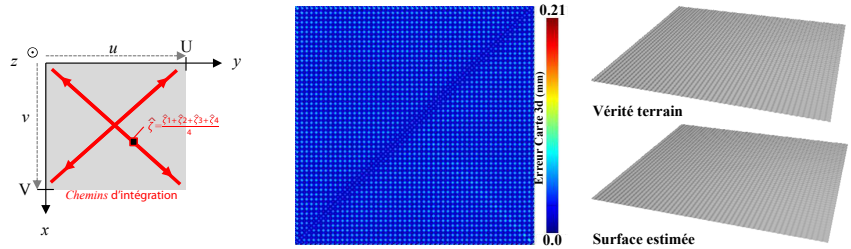


FIGURE 4.14 – *Chemin d'intégration proposé par* RODEHORST *ainsi qu'un exemple de propagation de l'erreur d'intégration pour une simulation de surface micro-rugueuse.*

WU et LI ont proposé une méthode d'intégration [WL88] proche des trapèzes mélangée à l'idée des chemins multiples d'intégration. Les valeurs $\hat{\zeta}$ sont calculées suivant les diagonales à partir des deux chemins moyennés présentés figure 4.15.

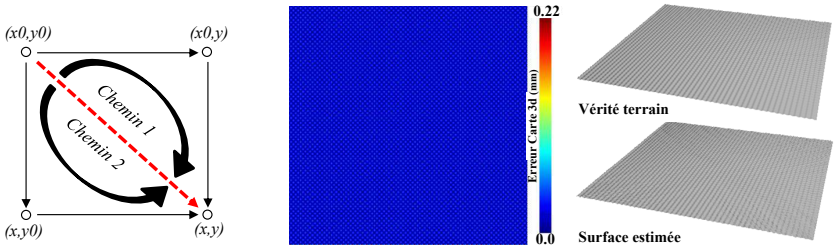


FIGURE 4.15 – *Chemin d'intégration proposé par* WU *et* LI *ainsi qu'un exemple de propagation de l'erreur d'intégration pour une simulation de surface micro-rugueuse.*

Ce schéma permet d'obtenir la formule approchée suivante :

$$\begin{aligned}\hat{\zeta}(u,v) = \hat{\zeta}(u-1,v-1) \quad &+ \quad \frac{p(u-1,v-1)+p(u,v-1)+p(u-1,v)+p(u,v)}{4} \\ &+ \quad \frac{q(u-1,v-1)+q(u,v-1)+q(u-1,v)+q(u,v)}{4}\end{aligned} \tag{4.11}$$

Cette méthode n'élimine pas totalement les erreurs d'intégration mais les réduit considérablement pour des géométries simples (figure 4.15).

4.2.2.2 Méthodes basées sur la relaxation

HORN et al. [HB86, INI83] ont proposé des méthodes d'intégration itératives basées sur la minimisation de l'écart entre la dérivée de la surface estimée et les dérivées partielles données par les champs de gradients :

$$\mathcal{W} = \int\int \left[\left(\frac{\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}x}\right) - p\right]^2 + \left[\left(\frac{\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}y}\right) - q\right]^2 \mathrm{d}x\mathrm{d}y \tag{4.12}$$

En passant par l'équation différentielle d'EULER-LAGRANGE, le minimum de $\mathcal{W}$ est donnée par un schéma itératif de la forme :

$$\hat{\zeta}(x,y)^{i+1} = \bar{\hat{\zeta}}(x,y)^i - \left(\frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\mathrm{d}q}{\mathrm{d}y}\right) \tag{4.13}$$

avec :

i	:	itérations.
$\bar{\hat{\zeta}}$	:	la moyenne.

La discrétisation conduit à un algorithme itératif baptisé schéma itératif de JACOBI :

$$\begin{aligned}\hat{\zeta}(u,v)^{i+1} \quad = \quad &\frac{1}{4}\left[\hat{\zeta}(u-1,v)^i + \hat{\zeta}(u+1,v)^i + \hat{\zeta}(u,v-1)^i + \hat{\zeta}(u,v+1)^i\right] \\ &-\frac{1}{4}\left[p(u+1,v) - p(u,v) + q(u,v+1) - q(u,v)\right]\end{aligned} \tag{4.14}$$

Au fur et à mesure des itérations i, la surface estimée $\hat{\zeta}$ s'affine pour tendre vers la surface réelle ζ. Mais cette méthode nécessite une initialisation totale (mise à zéro de toutes les

valeurs) de la surface pour débuter. La plus simple façon d'initialiser la carte des hauteurs est de fixer $\hat{\zeta}(u,v) = 0 \ \forall (u,v)$. Néanmoins, la convergence du schéma itératif de Jacobi est très lente. Or pour réduire le nombre d'itérations, il est possible d'initialiser la carte $\hat{\zeta}^0$ en utilisant une méthode locale telle que Wu-Li [Danoo, DCo7].

Afin d'accélérer la convergence, Terzopoulos puis Horovitz et Kiryati ont proposé des solutions [Ter84, Ter88, HKo4] basées sur la résolution d'équations aux dérivées partielles. Ces solutions sont nommées méthodes de Gauss-Seidel et définies par :

$$\begin{aligned}\hat{\zeta}(u,v)^{i+1} &= \frac{1}{4}\left[\hat{\zeta}(u-1,v)^{i+1} + \hat{\zeta}(u+1,v)^{i} + \hat{\zeta}(u,v-1)^{i+1} + \hat{\zeta}(u,v+1)^{i}\right] \\ &\quad -\frac{1}{4}\left[p(u+1,v) - p(u,v) + q(u,v+1) - q(u,v)\right] \end{aligned} \tag{4.15}$$

Cette solution proche du schéma itératif de Jacobi permet d'augmenter théoriquement par deux la vitesse de convergence en fonction de la complexité géométrique des surfaces.

La vitesse de convergence peut encore être augmentée par la méthode SOR (Successive Over Relaxation) [Pao2] proposée par Pres et Al. Cette algorithme consiste à introduire un coefficient ($1 < a < 2$) dans l'équation itérative de Gauss-Seidel :

$$\begin{aligned}\hat{\zeta}(u,v)^{i+1} &= (1-a)\hat{\zeta}(u,v)^{i} \\ &\quad +\frac{a}{4}\left[\hat{\zeta}(u-1,v)^{i+1} + \hat{\zeta}(u+1,v)^{i} + \hat{\zeta}(u,v-1)^{i+1} + \hat{\zeta}(u,v+1)^{i}\right] \\ &\quad -\frac{a}{4}\left[p(u+1,v) - p(u,v) + q(u,v+1) - q(u,v)\right] \end{aligned} \tag{4.16}$$

La vitesse de convergence et l'erreur d'intégration sont proportionnelles à la valeur a. Pour des cas géométriquement simples, plus a est proche de la valeur optimale 2, plus le nombre d'itérations nécessaires est bas. A l'inverse pour des cas géométriquement difficiles, a doit être proche de 1 pour assurer la robustesse de la méthode.

4.2.2.3 Méthodes Globales

Comme les méthodes basées sur la relaxation, les méthodes globales cherchent à minimiser la fonction définie par $\mathcal{W}$ dans l'équation 4.12. Parmi les méthodes globales, la plus connue et utilisée est celle proposée par Frankot et Chellappa [FC88]. Elle utilise la transformée de Fourier ainsi que son inverse pour estimer les surfaces dans le domaine fréquentiel. La transformée de Fourier 2d est définie par :

$$\zeta^{TF}(u,v) = \int\int \zeta(u,v) e^{-j(ux+vy)} \, \mathrm{d}x\mathrm{d}y \tag{4.17}$$

et son inverse :

$$\zeta^{TF-1}(u,v) = \frac{1}{2\pi}\int\int \zeta(u,v)^{TF} e^{\,j(ux+vy)} \, \mathrm{d}u\mathrm{d}v \tag{4.18}$$

A partir de ces définitions et en appliquant le théorème de PARSEVAL, la fonction $\mathcal{W}$ permet d'obtenir une relation entre la transformée de FOURIER de la surface à estimer et les transformées de FOURIER des champs de gradients pour $(u,v) \neq (0,0)$:

$$\hat{\zeta}^{TF}(u,v) = \frac{-jup^{TF}(u,v) - jvq^{TF}(u,v)}{u^2+v^2} \tag{4.19}$$

Grâce à cette relation simple, la surface estimée $\hat{\zeta}$ est directement obtenue par la transformée de FOURIER inverse de $\hat{\zeta}^{TF}$. Néanmoins, cette méthode est limitée à la reconstruction de surface dont la valeur moyenne des gradients est nulle $p^{TF}(0,0) = 0$ et $q^{TF}(0,0) = 0$ car les cas $(u,v) = (0,0)$ sont exclus. Ainsi le plan moyen des surfaces estimables doit être parfaitement perpendiculaire à l'axe optique ou z.

Dans le cas de surfaces à plan moyen non nul, la méthode peut être appliquée à condition de séparer la valeur moyenne des champs de gradients :

$$\begin{aligned} p0(x,y) &= p(x,y) - \bar{p} \\ q0(x,y) &= q(x,y) - \bar{q} \end{aligned} \tag{4.20}$$

Pour tous (x,y), $\bar{p}$ et $\bar{q}$ sont les champs de gradients moyens et $p0, q0$ sont les champs de gradients dont la moyenne est nulle. Le fait que l'opération dérivée soit linéaire permet de la décomposer :

$$\hat{\zeta}(x,y) = \hat{\zeta}0(x,y) + \bar{\hat{\zeta}}(x,y) \tag{4.21}$$

avec $\hat{\zeta}0$ la carte des hauteurs estimée par la méthode [FC88] et $\bar{\hat{\zeta}}$ le plan moyen estimable par :

$$\bar{\hat{\zeta}}(x,y) = x\bar{p} + y\bar{q} + a = xp^{TF}(0,0) + yp^{TF}(0,0) + a \tag{4.22}$$

avec a une constante pouvant être nulle car la surface est toujours estimée à une constante additive près.

Cette méthode passe par le domaine fréquentiel ce qui permet d'effectuer un filtrage simple et efficace des champs de gradients. L'élimination des plus hautes fréquences permet de réduire sensiblement les effets du bruit et augmente la robustesse fasse aux erreurs d'intégration. A l'inverse un filtrage des basses fréquences permet d'éliminer le plan moyen. Enfin un filtrage passe bande permet d'extraire des phénomènes à étudier pour certaines applications spécifiques telles que l'analyse d'un type particulier de micro-rugosités.

Toujours dans l'idée d'améliorer la méthode, WEI et KLETTE ont proposé une variante [WK01, WK02] en introduisant des paramètres supplémentaires dans la fonction $\mathcal{W}$ à minimiser.

$$\begin{aligned}\mathcal{W} &= \int\int \left[\left(\frac{-\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}x}\right) - p\right]^2 + \left[\left(\frac{-\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}y}\right) - q\right]^2 \mathrm{d}x\mathrm{d}y \\ &+aa \int\int \left[\left(\frac{\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}x}\right)\right]^2 + \left[\left(\frac{\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}y}\right)\right]^2 \mathrm{d}x\mathrm{d}y \\ &+bb \int\int \left[\left(\frac{-\mathrm{d}^2\hat{\zeta}(x,y)}{\mathrm{d}x^2}\right)\right]^2 + 2\left[\left(\frac{-\mathrm{d}\hat{\zeta}(x,y)}{\mathrm{d}x\mathrm{d}y}\right)\right]^2 + \left[\left(\frac{\mathrm{d}^2\hat{\zeta}(x,y)}{\mathrm{d}y^2}\right)\right]^2 \mathrm{d}x\mathrm{d}y\end{aligned} \tag{4.23}$$

où le second terme représente les variations d'aire de la surface et le troisième représente les variations de courbures de la surface. Ces termes sont gérés par les paramètres *aa* et *bb* positifs. Cette nouvelle écriture permet d'obtenir :

$$\hat{\zeta}^{TF}(u,v) = \frac{-jup^{TF}(u,v) - jvq^{TF}(u,v)}{(1+aa)(u^2+v^2) + bb(u^2+v^2)^2} \tag{4.24}$$

D'après les auteurs, cette amélioration augmente la robustesse au bruit mais malheureusement les résultats obtenus sont plus difficiles à exploiter car en plus de la constante additive près, les surfaces sont estimées à une constante multiplicative près. Or cette constante est difficile à estimer ce qui fausse la mesure ou la comparaison.

4.2.2.4 Conclusion

D'après l'ensemble des comparaisons effectuées (Annexe A.3), nous pouvons constater que la méthode globale Frankot et Chellappa offre la meilleur qualité de reconstruction. De plus, elle est robuste face au bruit et fait partie des méthodes les plus rapides. Par conséquent au vu de ses nombreux avantages elle nous sert de référence pour l'ensemble de ce document.

4.3 Ombres et Spécularités dans la SP

Hormis le cas particulier d'une surface parfaitement diffuse et plane, l'interaction lumière matière produit généralement un grand nombre de phénomènes tel que des ombres, des inter-réflexions ou des zones de spécularité. Les ombres propres ou portées sont fonction de la géométrie des surfaces et sont proportionnelles au niveau de rugosité, d'aspérité et d'amplitude des variations géométriques. Dans le cas d'une lumière ambiante nulle et d'une limitation des inter-réflexions, les ombres apparaissent comme des zones sans information dans les images acquises. Avec la stéréo-photométrie à trois sources lumineuses, les trois valeurs d'intensité *Ie* fournies par les images en chaque point de la surface sont nécessaires pour inverser le modèle de BRDF. Ainsi tout manque d'information pour l'une de ces trois valeurs conduit vers un système non résolvable où l'extraction de l'information géométrique et photométrique est erronée. Il en va de même pour les phénomènes spéculaires, ces zones sont saturées dans les images acquises. Elles apparaissent comme des points ou des tâches dont l'intensité lumineuses est trop importante et l'information ne peut être conservée. Pour illustrer ces propos, la figure 4.16 démontre qu'une demi-sphère conçue dans un matériau ayant une BRDF non Lambertienne génère des ombres propres et portées ainsi qu'un lobe spéculaire lorsqu'elle est éclairée par une source lumineuse dans une direction donnée.

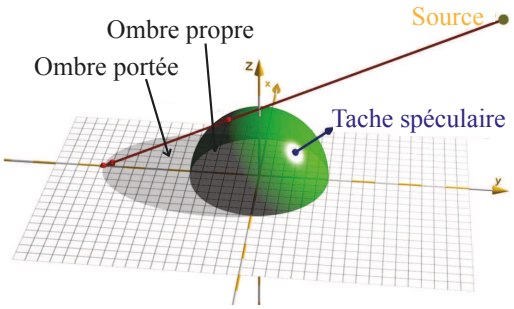


FIGURE 4.16 – *Schéma d'une demi-sphère démontrant la présence des ombres ainsi que celle d'une tache spéculaire pour une direction d'éclairage donnée.*

Ces phénomènes se retrouvent dans les acquisitions. En effet, quelle que soit la direction d'éclairage, les trois images (figure 4.17) obtenues par lancer de rayons et le modèle de PHONG sont affectées par des ombres et des tâches spéculaires.

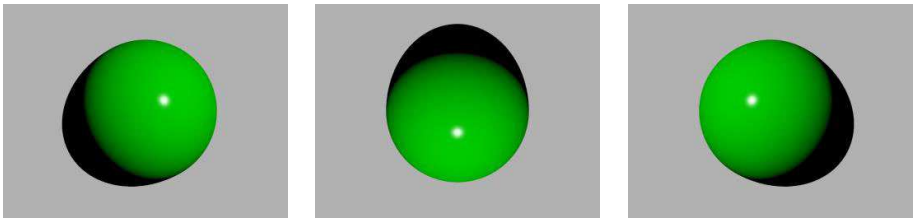
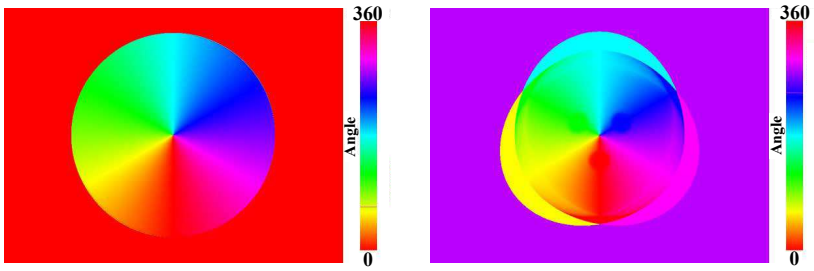


FIGURE 4.17 – *Images acquises de d'une demi-sphère pour trois directions d'éclairage différentes.*

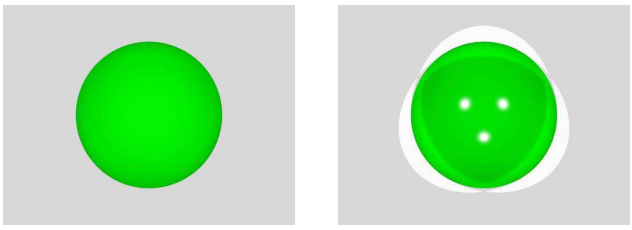
L'extraction de l'information géométrique et photométrique à partir de ce triplet d'images pour le modèle Lambertien est erronée dans les zones d'ombre ou spéculaires. Les champs de gradients p et q (figure 4.18) et l'albédo (figure 4.19) extraits subissent des distorsions dues à la perte d'information.



(a) (b)

FIGURE 4.18 – *Représentation angulaire des champs de gradients p et q, (a) géométrie réelle, (b) extraction suivant le processus classique de la SP.*

Pour remédier à ces problèmes et éviter les distorsions importantes lors de l'extraction de l'information, plusieurs méthodes sont utilisées. Pour certaines, le nombre d'acquisitions reste restreint au minimum de trois images et des modèles de réflectances plus sophistiqués sont



(a) (b)

FIGURE 4.19 – *Albédo, (a) géométrie réelle, (b) extraction par SP.*

utilisés [Nay91, Geo03, SI96, HMI10] ou une régularisation des valeurs est réalisée [HVC08]. Ces méthodes alourdissent considérablement la complexité de résolution du système et ne peuvent corriger l'ensemble des problèmes lorsqu'ils sont multiples. C'est pourquoi, d'autres méthodes proposent d'augmenter le nombre d'acquisitions et sélectionnent les valeurs non affectées par les phénomènes d'ombres ou de spécularités [CJ82, SSS+07, APB10, EGSSF11]. Ainsi au minimum trois valeurs respectant le modèle Lambertien sont conservées pour respecter les contraintes fixées par l'hypothèse $\mathcal{H}1$. Le nombre d'acquisitions nécessaire à cette condition dépend de la complexité géométrique des surfaces analysées ainsi que la configuration des sources lumineuses.

4.3.1 Élimination des ombres et points spéculaires (s>3)

COLEMAN et JAIN dans [CJ82] sont les premiers à s'intéresser à l'élimination des points spéculaires pour pallier aux distorsions. L'algorithme qu'ils proposent utilise le modèle Lambertien pour des surfaces dont la réflexion peut être spéculaire et traitée comme une déviation du modèle. La méthode n'utilise plus trois mais quatre images de la même surface afin de détecter les points spéculaires en l'absence d'ombres propres ou portées. Sous l'hypothèse que les régions spéculaires ne s'intersectent pas dans les quatre images, la détection est réalisée par comparaison des quatre triplets de valeurs possibles :

Triplet 1	M_1, M_2, M_3
Triplet 2	M_2, M_3, M_4
Triplet 3	M_1, M_3, M_4
Triplet 4	M_1, M_2, M_4

Dans ces triplets, si l'albédo extrait diffère de façon significative, la présence d'un point spéculaire est supposée. Ainsi les trois triplets fournissant les plus fortes valeurs d'albédo sont susceptibles d'être affectés et par conséquent seul le triplet qui fournit la plus faible valeur est utilisée pour extraire les champs de gradients. Cette méthode alourdit la résolution de la SP mais surtout elle ne fournit pas de solution en cas d'ombre.

SOLOMON et IKEUCHI ont proposé une variante [SI96] qui permet de prendre en compte les ombres propres grâce à une segmentation par région des images. Mais leur méthode est incapable de détecter les ombres portées ce qui la rend peu robuste et très ambiguë sur l'élimination des valeurs supposées affectées.

Toujours à partir de quatre sources lumineuses, la méthode [BP03] proposée par BARSKY et PETROU vise à réduire les distorsions causées par les ombres et les tâches spéculaires. Les quatre images acquises fournissent un quadruplet de valeurs en chaque point de la surface

et les quatre directions des sources $\vec{l}_{s=1...4}$ sont linéairement dépendantes ce qui permet de définir un rapport linéaire entre elles :

$$a_1\vec{l}_1 + a_2\vec{l}_2 + a_3\vec{l}_3 + a_4\vec{l}_4 = 0 \tag{4.25}$$

En multipliant par l'albédo et la normale à la surface cette équation devient :

$$a_1\rho(\vec{l}_1.\vec{n}) + a_2\rho(\vec{l}_2.\vec{n}) + a_3\rho(\vec{l}_3.\vec{n}) + a_4\rho(\vec{l}_4.\vec{n}) = 0 \tag{4.26}$$

Ce qui correspond par simplification en suivant $\mathcal{H}4$ à :

$$a_1M_1 + a_2M_2 + a_2M_2 + a_4M_4 = 0 \tag{4.27}$$

Ou sous la forme vectorielle :

$$\vec{a}.\vec{b} = 0 \tag{4.28}$$

avec :

$$\begin{array}{rcl} \vec{a} & = & (a_1\ a_2\ a_3\ a_4)^T \\ \vec{b} & = & (M_1\ M_2\ M_3\ M_4)^T \text{ pour des images NVG.} \end{array}$$

Les valeurs de $\vec{a}$ sont estimées par la méthode des moindres carrés. Dans le cas des images couleur, le vecteur $\vec{b}$ est obtenu par une analyse en composantes principales (ACP) :

$$\begin{array}{rcl} [\mathbf{M}] & = & \begin{bmatrix} M_{1_r} & M_{1_v} & M_{1_b} \\ M_{2_r} & M_{2_v} & M_{2_b} \\ M_{3_r} & M_{3_v} & M_{3_b} \\ M_{4_r} & M_{4_v} & M_{4_b} \end{bmatrix} \\ [\mathbf{C}] & = & [\mathbf{M}]^T[\mathbf{M}] \rightarrow \vec{c}\ (\textit{le vecteur propre}) \\ \vec{b} & = & [\mathbf{M}]\vec{c} \end{array} \tag{4.29}$$

Les quadruplets affectés par l'ombre sont détectés par seuillage de l'erreur quadratique générée par l'équation 4.27. Pour les quadruplets affectés, la valeur la plus faible est écartée et l'extraction de l'information est réalisée à partir des trois valeurs les plus élevées. Dans le cas contraire, l'étape suivante consiste à calculer la disparité chromatique. Si cette disparité dépasse un seuil cela signifie que la valeur M la plus élevée est spéculaire et est écartée. L'extraction de l'information est réalisée à partir des trois valeurs les plus faibles. Les seuils sont fixés empiriquement par l'utilisateur. Si les quadruplets sont parfaitement Lambertiens et sans ombre, les quatre valeurs sont utilisées pour calculer p, q et ρ. Le problème majeur

de cette méthode est qu'elle ne peut être utilisée pour plus de quatre sources lumineuses car avec plus de sources lumineuses l'égalité (équation 4.27) peut être vraie malgré une erreur. Elle n'est pas en mesure d'identifier les cas compliqués où deux valeurs sont affectées par des ombres ou des phénomènes spéculaires, ce qui aboutit à des distorsions.

La méthode proposée par Rushmeir et al. [RTG97] utilise cinq sources lumineuses. L'extraction de l'information est effectuée à partir de trois valeurs M car la plus élevée et la plus faible sont automatiquement rejetées pour éviter l'ombre et la spécularité. Cette méthode est basée sur l'indéterminisme mais peut offrir des résultats satisfaisants pour des surfaces dont les variations géométriques sont faibles. De plus, elle permet de tirer des conclusions statistiques car les chances d'extraire trois valeurs parfaitement Lambertiennes et non ombrées augmentent proportionnellement avec le nombre de sources utilisées. De plus la probabilité d'avoir deux points spéculaires pour deux directions d'éclairage différentes est très faible.

A partir de ces conclusions statistiques, Sun et al. [SSS+07] ont proposé un algorithme (figure 4.20) de détection basé sur le seuillage et l'utilisation de six sources lumineuses. Comme Barsky et Petrou, les valeurs seuillages des ombres Sb et des points spéculaires Sh sont fixés empiriquement. L'algorithme fonctionne avec des images couleur dont les valeurs sont extraites par une ACP ou un moyennage des trois plans. Les résultats obtenus sont plus performants du fait de l'utilisation de six sources lumineuses mais le choix empirique des seuils et la simplicité de l'algorithme conduisent à un grand nombre de fausses et de non détections.

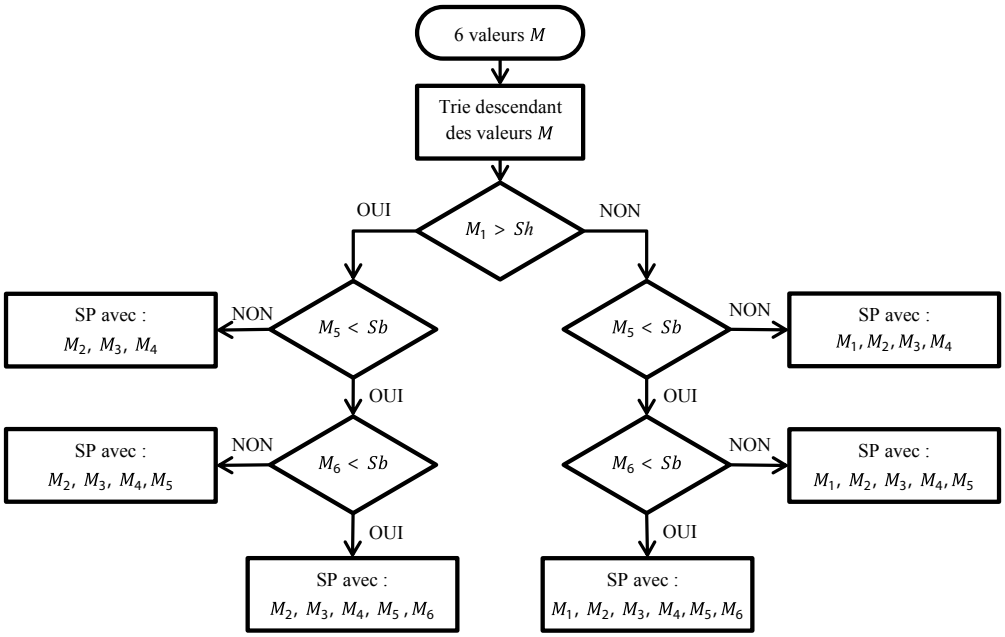


Figure 4.20 – *Algorithme de* Sun *et al. [SSS+07].*

Dans [APB10], Argyriou propose une généralisation à S images de la méthode [BP03]. Les résultats qu'il obtient montre une amélioration de la détection par rapport aux méthodes précédentes mais le seuillage des ombres reste empirique et doit être adapté à la complexité géométrique des surfaces analysées.

4.3.2 Notre méthode d'élimination

Partant du constat que les méthodes à seuillages empiriques sur $S \geq 6$ images offrent des résultats intéressants, nous proposons un nouveau schéma d'élimination des phénomènes spéculaires et d'ombre inspiré par Sun et Argyriou. L'atout principal de notre méthode [BBK12] est le rejet des valeurs M affectées par l'ombre grâce à un seuillage adaptatif automatisé. Le rejet de la spécularité est effectué à partir du calcul de l'erreur Lambertienne proposée par [BP03] que nous avons généralisé. De plus, comme la probabilité est forte que les valeurs soient affectées par l'ombre plutôt que par des phénomènes spéculaires, nous avons proposé une élimination hiérarchique afin d'améliorer la robustesse.

Afin d'évaluer l'efficacité de notre méthode, nous proposons de simuler des effets spéculaires et d'ombres à partir d'images de synthèses produites par lancer de rayons et le modèle de Phong :

$$M \propto Ie = Ie_{ambiante} + Ie_0\ Kd\ (\vec{l}.\vec{n}) + Ie_0\ Ks\ (\vec{r}.\vec{v})^{n_s} \tag{4.30}$$

Les images générées (figure 4.21) représentent deux demi-sphères de diamètre 640 *pixels* dont la réflexion Ks est non nulle et placées sur une surface plane parfaitement Lambertienne. Pour obtenir ces données nous utilisons le logiciel de lancer de rayons POV-Ray et nous ajoutons un bruit blanc Gaussien de variance 0.05 pour plus de réalisme. Chacune des six images de la scène est associée à l'une des six sources lumineuses disposées autour de l'appareil d'acquisition et distribuées tous les 60^o. Cette configuration respecte les hypothèses $\mathcal{H}2$ et $\mathcal{H}3$. Toutes les ombres sont dues à la géométrie de la scène, toutes les facettes sont visibles et les zones spéculaires ne présentent aucun recouvrement parmi les six images.

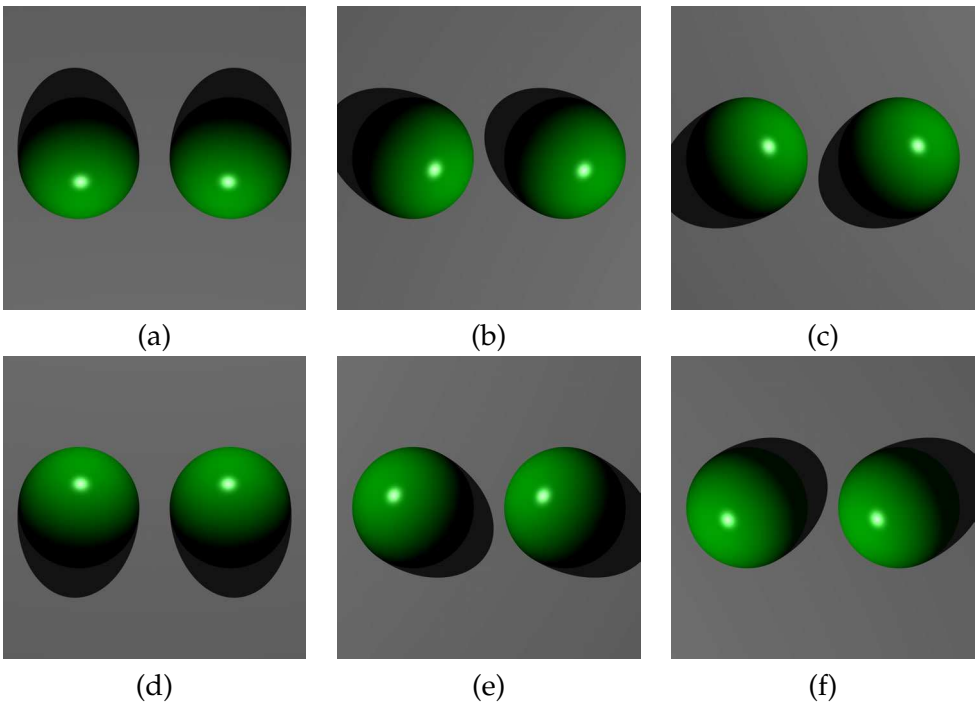


Figure 4.21 – *Images de synthèse de deux demi-sphères obtenues suivant le modèle de* Phong *par le logiciel POV-RAY.*

4.3.2.1 Élimination des ombres

Pour inclure les ombres dans le modèle de LAMBERT, nous redéfinissons une nouvelle équation de formation des images :

$$M \propto Ie = \begin{cases} Ie_{ambiante} + \rho(\vec{l}.\vec{n}) & : \textit{si la source } \vec{l} \textit{ est visible pour la facette} \\ Ie_{ambiante} & : \textit{dans le cas contraire} \end{cases} \tag{4.31}$$

En absence de lumière ambiante les valeurs M pour les zones d'ombre tendent vers zéro.

La première étape de notre méthode consiste à effectuer un tri descendant des valeurs numériques M_s et de rejeter pour le moment la plus forte. Car statistiquement, si les sources lumineuses sont fixes et correctement distribuées, il est raisonnable de supposer qu'il ne peut y avoir qu'une valeur affectée par un phénomène spéculaire. De plus, les valeurs M_2 et M_3 sont obligatoirement supposées sans ombre et sans spécularité car la résolution de la SP nécessite au minimum trois valeurs. La troisième valeur est soit M_1 si elle n'est pas spéculaire ou M_4 si elle n'est pas dans une zone d'ombre.

A partir de ces contraintes, notre critère de rejet des valeurs est défini par :

$$\frac{\frac{1}{2}(M_2 + M_3) - M_s}{\frac{1}{2}(M_2 + M_3) + M_s} \geq Ts \tag{4.32}$$

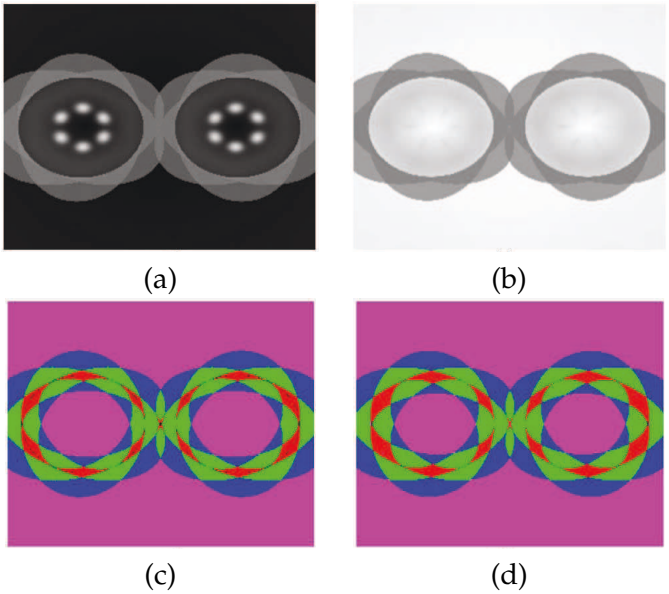
où M_s est jugée affectée par l'ombre si la partie gauche de l'équation dépasse le seuil adaptatif Ts. Ce seuil unique pour chaque source s est défini par une fonction fs dépendante de la variance des valeurs M_s tel que :

$$Ts = fs(var(M_s)) \tag{4.33}$$

Si la variance des valeurs M_s est forte, la probabilité qu'une valeur M_s soit affectée par l'ombre est élevée. A l'inverse, si la variance des valeurs M_s est faible, la probabilité qu'une valeur M_s soit affectée par l'ombre est faible. La figure 4.22a représente la variance de M dans le cas de nos images de synthèse. Dans notre algorithme les valeurs M sont normalisées entre 0 et 1. Pour les images couleurs, les valeurs Ms sont les moyennes des trois plans RVB. Les valeurs de Ts sont alors définies par :

$$Ts = 1 - var(M_2 \dots M_s) \tag{4.34}$$

La figure 4.22b-d illustre le seuillage adaptatif et l'étiquetage des zones d'ombres détectées. Par rapport à la vérité terrain notre détection est quasi-identique et nous pouvons constater que la fonction Ts est plus faible dans le cas des ombres propres que dans le cas des ombres portées.



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

FIGURE 4.22 – *Détection des ombres suivant notre méthode, (a) $Var(M)$, (b) seuil adaptatif Ts, (c) classification des ombres pour la vérité terrain, (d) classification des ombres selon notre méthode.*

4.3.2.2 Élimination de la spécularité

La seconde étape de notre méthode consiste à extraire les valeurs spéculaires. Nous redéfinissons un vecteur $M'_{s'}$ correspondant à M_s avec $s' > 3$ sans les valeurs potentiellement affectées par l'ombre. La plus forte valeur M'_1 est détectée spéculaire ou non en fonction de l'erreur Lambertienne selon une généralisation du détecteur proposé par [BP03], équations 4.25 à 4.28. Nous modifions ces équations pour obtenir une généralisation de l'erreur pour un nombre arbitraire $s' > 4$ de sources lumineuses :

$$\epsilon_{Lambert} = \sum_{i=0}^{C_{s'}^4} \sum_{j=1}^{S'} \left| a_{ij} M'_j \right| \tag{4.35}$$

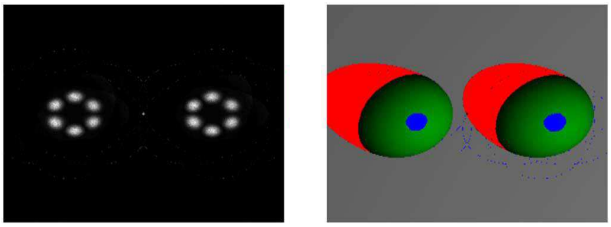
avec :

$\epsilon_{Lambert}$	:	erreur Lambertienne.
a_{ij}	:	les coefficients obtenus par l'équation 4.25.
M'_j	:	les valeurs des pixels pour la source lumineuse j.

Contrairement à la méthode proposée par ARGYRIOU dans [APB10], la détection de la spécularité n'est pas effectuée sur l'ensemble (S) des sources lumineuses afin de limiter les fausses détections dues au fait que plusieurs valeurs M sont potentiellement affectées par l'un des deux phénomènes. En effet les fausses détections lors de l'étape de détection des ombres peuvent affecter l'extraction de la spécularité. Dans notre méthode ces valeurs supposées ombrées ne sont pas prise en compte lors des combinaisons permettant le calcul de l'erreur Lambertienne, par conséquent le nombre de combinaisons est :

$$C_{s'}^4 = \frac{s'!}{4!(s'-4)!} \tag{4.36}$$

Si $\epsilon_{Lambert}$ est supérieur à un seuil défini, la valeur M'_1 est automatiquement rejetée et les champs de gradients p et q ainsi que l'Albédo sont déterminés à partir de $(S'-1)$valeurs de M' restantes. Dans le cas contraire les S' valeurs de M' sont utilisées. La figure 4.23 montre l'erreur $\epsilon_{Lambert}$ pour nos images de synthèse ainsi qu'une détection ombre et spécularité pour une source lumineuse. Comme nous pouvons le constater, la détection des deux phénomènes est correctement réalisée.



(a) (b)

FIGURE 4.23 – *Détection de la spécularité suivant notre méthode, (a) Erreur Lambertienne $\epsilon_{Lambert}$, (b) classification des ombres et des spécularités (en bleu) selon notre méthode.*

4.3.2.3 Résultats par simulation

Afin d'évaluer les performances de notre méthode de détection ($\mathcal{OUR}$), nous appliquons la SP à partir de six sources lumineuses et nous la comparons à : la méthode SP classique sans détection ($\mathcal{M}1$), la méthode avec détection proposée dans [SSS+07] ($\mathcal{M}2$) et la méthode avec détection proposée dans [APB10] ($\mathcal{M}3$). Pour cette première phase d'analyse, les surfaces utilisées sont connues pour être texturées ou à fortes variations géométriques (bitume, crépi, etc). Les vérités terrains proviennent d'une base de données acquises par un système temps de vol laser et sont utilisées pour obtenir des images de synthèse. La taille des surfaces est de $40,96 \times 40,96$ mm avec un pas de 80 μm suivant les axes x et y. Les variations en profondeur z sont au maximum de 10 mm pour un pas d'acquisition de 1 μm. Des exemples de la base sont donnés figure 4.24. Pour obtenir les images de synthèse (figure 4.25) nous utilisons le modèle BRDF de PHONG et six acquisitions simulées pour des angles d'éclairage zénith $\sigma = 30^o$ et azimuts $\tau = [0^o, 60^o, 120^o, 180^o, 240^o, 300^o]$.

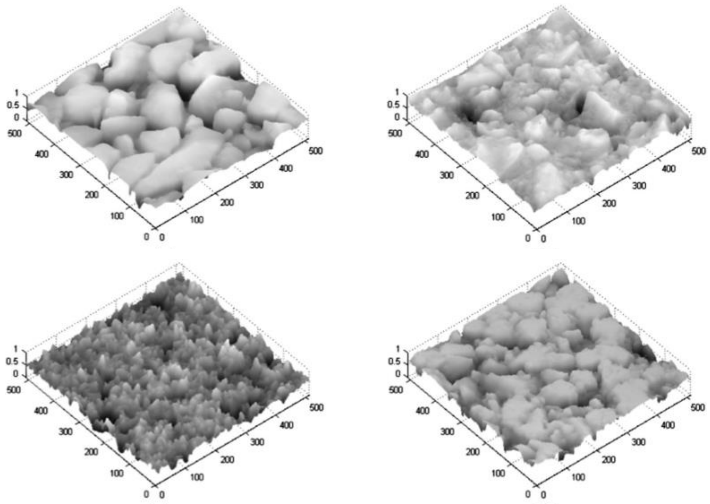
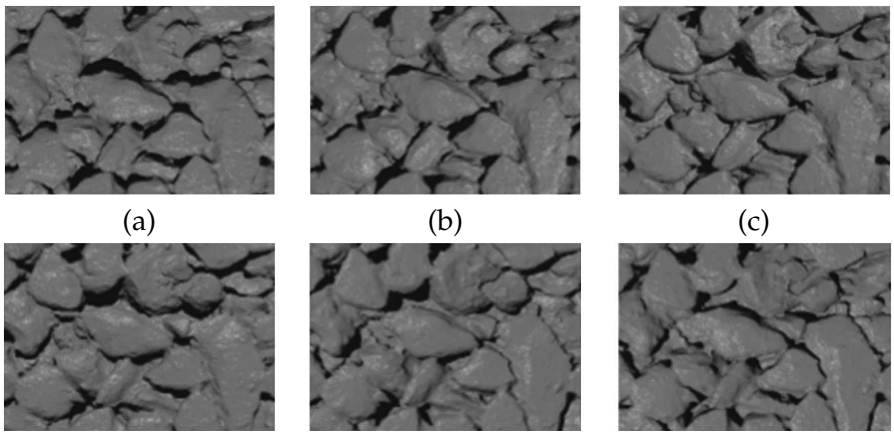


FIGURE 4.24 – *Exemples de vérités terrains acquises par laser.*



(a) (b) (c)

(d) (e) (f)

FIGURE 4.25 – *Six images de synthèses pour un exemple de surface présentant un fort taux d'ombre et de spécularités. L'angle d'acquisition zénith est de $\sigma = 30^o$ et les angles azimuts sont (a) $\tau = 0^o$, (b) $\tau = 60^o$, (c) $\tau = 120^o$, (d) $\tau = 180^o$, (e) $\tau = 240^o$ et (f) $\tau = 300^o$.*

Pour chaque point d'une surface, nous calculons la proportion des valeurs problématiques non détectées VN (Vrais Négatif) et la proportion des valeurs non problématiques détectées comme problématiques FP (Faux Positif) pour l'ombre, la spécularité et les deux phénomènes réunis. Les tableaux 4.1 et 4.2 présentent les résultats pour deux surfaces. Les colonnes $\mathcal{M}1$-VN donnent le pourcentage de valeurs problématiques présentent pour la surface analysée car il n'y a aucune détection effectuée par la méthode $\mathcal{M}1$. Ces valeurs nous permettent d'effectuer une normalisation des résultats pour l'ensemble de la base de données des vérités terrains présentée tableau 4.3.

Méthodes	$\mathcal{M}1$		$\mathcal{M}2$		$\mathcal{M}3$		$\mathcal{OUR}$	
Mal classées	VN	FP	VN	FP	VN	FP	VN	FP
Ombres %	21,76	0,00	11,35	10,98	10,18	13,62	6,82	0,09
Spéculaire %	30,69	0,00	16,08	20,41	5,48	25,70	5,67	19,35
Les deux %	50,01	0,00	26,81	30,71	15,28	31,51	12,06	19,43
SNR *dB*	2,18		7,22		10,88		24,61	

TABLE 4.1 – *Résultats de détection des ombres et des spécularités ainsi que le SNR entre la vérité terrain sans phénomènes et les autres méthodes de détection pour la surface présentée figure 4.24a.*

Méthodes	$\mathcal{M}1$		$\mathcal{M}2$		$\mathcal{M}3$		$\mathcal{OUR}$	
Mal classées	VN	FP	VN	FP	VN	FP	VN	FP
Ombres %	30,66	0,00	8,88	20,28	7,13	25,00	5,77	0,35
Spéculaire %	33,47	0,00	18,59	19,06	8,75	36,61	8,62	23,08
Les deux %	59,16	0,00	26,44	37,43	15,17	45,10	12,89	23,41
SNR *dB*	7,79		18,85		23,84		38,85	

TABLE 4.2 – *Résultats de détection des ombres et des spécularités ainsi que le SNR entre la vérité terrain sans phénomènes et les autres méthodes de détection pour la surface présentée figure 4.24b.*

Méthodes	$\mathcal{M}1$		$\mathcal{M}2$		$\mathcal{M}3$		$\mathcal{OUR}$	
Mal classées	VN	FP	VN	FP	VN	FP	VN	FP
Ombres %	100,0	0,00	40,80	57,74	35,17	69,94	23,02	0,58
Spéculaire %	100,0	0,00	54,17	76,28	22,57	123,1	23,64	73,62
Les deux %	100,0	0,00	49,62	61,11	28,32	67,13	22,31	36,32
SNR dB	3,21		12,36		16,39		33,61	

TABLE 4.3 – *Résultats de détection des ombres et des spécularités ainsi que le SNR entre la vérité terrain sans phénomènes et les autres méthodes de détection pour l'ensemble des surfaces de notre base de données vérité terrain.*

Comme nous pouvons le constater, la détection des ombres est plus faible par notre méthode. Cela est dû au fait qu'elle utilise seulement deux valeurs sur six alors que les autres méthodes utilisent au minimum quatre valeurs pour réaliser la détection. En effet si 3/6 valeurs sont affectées par l'ombre dans $\mathcal{M}3$ et pas de spéculaire dans $\mathcal{M}2$, seules les deux dernières valeurs sont détectables ce qui crée une valeur VN. Or, la conservation de cette valeur ombrée a tendance dans les méthodes $\mathcal{M}2$ et $\mathcal{M}3$ à augmenter la détection des fausses spécularités ($\mathcal{M}2$-FP ou $\mathcal{M}3$-FP lignes spéculaires élevées). De plus, les valeurs FP sont élevées car l'ajustement des seuils pour ces méthodes est difficile à définir [APB10]. Pour ce qui est de la détection de la spécularité, notre méthode est équivalente à $\mathcal{M}3$ pour les valeurs VN par contre elle est supérieure sur les valeurs FP. Dans tous les cas, notre méthode est plus efficace que les trois autres comme en attestent les résultats.

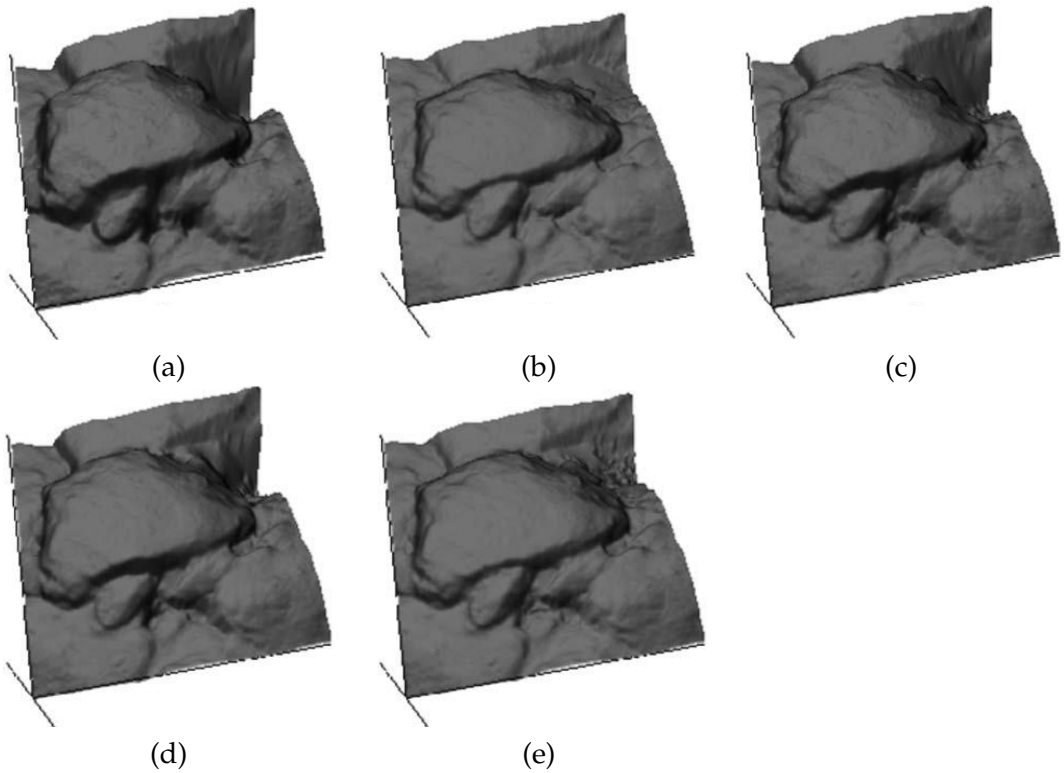


FIGURE 4.26 – *Comparaison des reconstructions 3d pour un échantillon d'une surface obtenu suivant (a) $\mathcal{M}1$, (b) vérité terrain, c) $\mathcal{M}2$, (d) $\mathcal{M}3$ et e) $\mathcal{OUR}$.*

Afin de démontrer l'importance de l'élimination des ombres et des spécularités, nous proposons l'analyse du rapport signal sur bruit SNR entre les surfaces de vérités terrains ζ et les surfaces estimées $\hat{\zeta}$ par la SP à partir des champs de gradients obtenus suivant l'une des quatre méthodes. La méthode d'intégration utilisée pour obtenir $\hat{\zeta}$ est [FC88] (voir 4.2.2.3), les SNR sont calculés par l'équation A.1. Les résultats obtenus sont présentés en dernière ligne des tableaux (4.1, 4.2, et 4.3). La figure 4.26 montre les résultats de reconstruction pour une petite zone d'une surface. Comme nous pouvons le constater, les cavités ne sont pas reconstruites de la même façon selon la méthode d'élimination utilisée car ces zones sont généralement très atteintes par les ombres. Cependant, les résultats entre notre méthode et la vérité terrain sont visuellement très proches. Ces résultats se corrèlent aux valeurs VN, VP et le SNR. La méthode sans détection $\mathcal{M}1$ offre un résultat très médiocre (SNR$< 10dB$) alors que les méthodes $\mathcal{M}2$ et $\mathcal{M}3$ sont légèrement meilleures. Ces résultats démontrent l'importance d'une élimination correcte des phénomènes d'ombres et de spécularités ainsi que les améliorations apportées par notre principe de détection.

4.3.2.4 Acquisitions réelles

Notre système d'acquisition (NIKON_1 A.2.1) permet d'obtenir six images qui couvrent une zone de $141,6 \times 94,8$ mm pour une résolution de 4288×2848 $pixels$, soit un échantillonnage spatial de pas 33×33 μm. Les images présentées figure 4.27 montrent le premier échantillon utilisé. Ce sont deux demi-sphères dont le matériau peut être décrit par le modèle BRDF de Phong. Avec ce cas réel, nous ne disposons pas de la vérité terrain. Par conséquent, nous effectuons une classification manuelle des phénomènes d'ombres et de spécularités pour les six images. Après suppression des valeurs problématiques, nous appliquons la SP. La reconstruction 3d issue de ce procédé est considérée comme la vérité terrain.

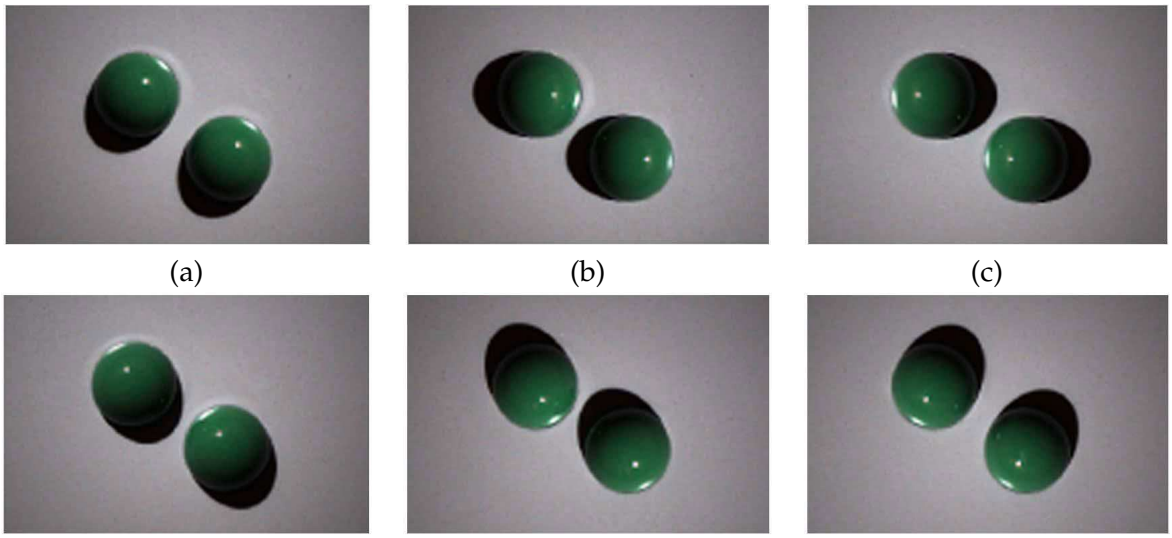


Figure 4.27 – *Six images réelles représentant deux demi-sphères et acquises par notre système NIKON_1. L'angle d'éclairage zénith est de* $\sigma = \frac{\pi}{6}$ *et les angles azimuts sont (a)* $\tau = 0$*, (b)* $\tau = \frac{\pi}{3}$*, (c)* $\tau = \frac{2\pi}{3}$*, (d)* $\tau = \pi$*, (e)* $\tau = \frac{4\pi}{3}$ *et (f)* $\tau = \frac{5\pi}{3}$*.*

Comme pour les résultats par simulation, nous calculons les pourcentages des valeurs problématiques VN et FP (présenté tableau 4.4) issues de la détection automatique réalisée par $\mathcal{M}1$, $\mathcal{M}2$, $\mathcal{M}3$ et $\mathcal{OUR}$ par rapport à la classification manuelle.

Méthodes	$\mathcal{M}1$		$\mathcal{M}2$		$\mathcal{M}3$		$\mathcal{OUR}$	
Mal classées	VN	FP	VN	FP	VN	FP	VN	FP
Ombres %	100,0	0,00	20,91	24,74	17,28	49,56	7,33	7,32
Spéculaire %	100,0	0,00	49,78	402,9	45,14	577,2	65,53	8,01
Les deux %	100,0	0,00	24,13	72,15	20,48	84,77	15,53	7,92
SNR dB	7,33		33,35		37,31		67,53	

TABLE 4.4 – *Résultats de détection des ombres et des spécularités ainsi que le SNR entre la vérité terrain sans phénomènes et les autres méthodes de détection pour nos acquisitions réelles.*

Notre méthode offre une meilleure détection des valeurs affectées par l'ombre. Pour la détection des valeurs spéculaires, le nombre de VN a augmenté par rapport aux résultats de simulation mais il est certainement possible de le diminuer car le taux de FP est très bas comparé aux autres méthodes. Enfin le SNR démontre que notre méthode offre des résultats très proches de la vérité terrain. La figure 4.28 montre les reconstructions 3d obtenues pour l'une des demi-sphères pour l'ensemble des méthodes. Nous pouvons remarquer que les méthodes $\mathcal{M}2$ et $\mathcal{M}3$ fournissent une reconstruction erronée des zones où plus de deux ombres s'entrecroisent. De plus, la comparaison entre $\mathcal{M}1$ sans élimination des phénomènes et les autres méthodes démontrent l'importance de la détection des valeurs problématiques.

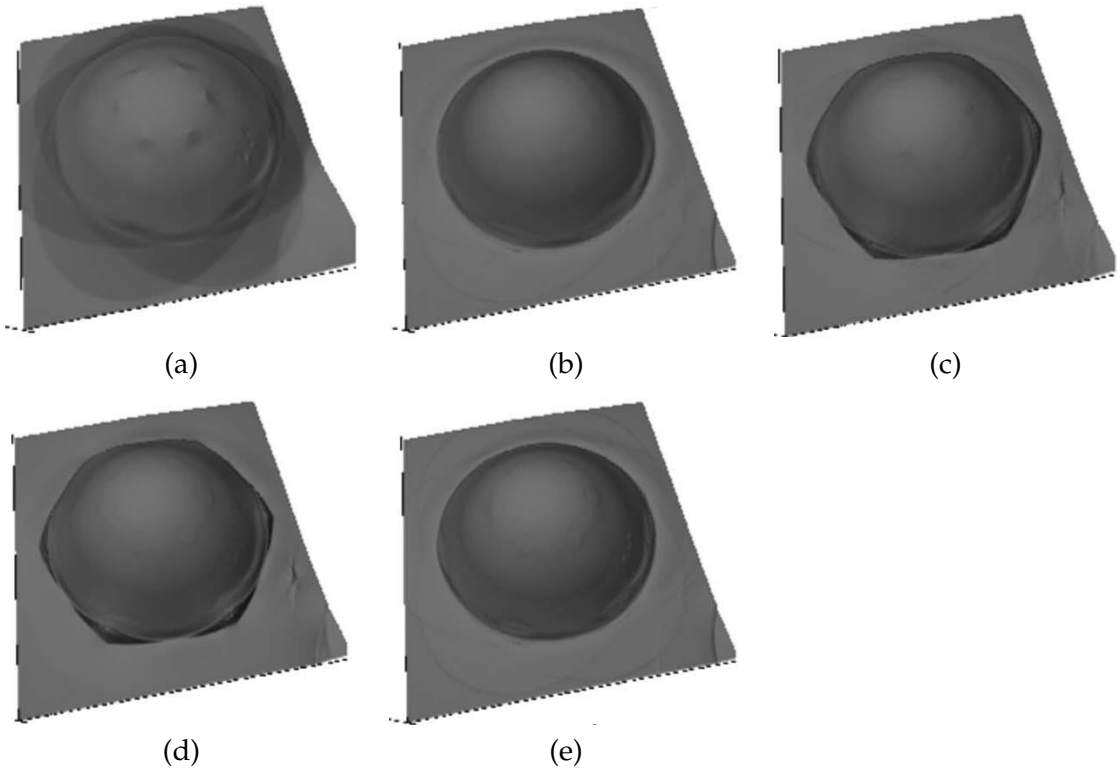


FIGURE 4.28 – *Comparaison des reconstructions 3d pour l'une des demi-sphères réelles obtenues suivant (a) $\mathcal{M}1$, (b) vérité terrain, c) $\mathcal{M}2$, (d) $\mathcal{M}3$ et e) $\mathcal{OUR}$.*

Afin de poursuivre l'analyse sur des cas réels, nous appliquons les différentes méthodes de détection à plusieurs surfaces rugueuses dont la géométrie est très variée (exemple figure 4.29). Nous proposons une analyse visuelle afin d'outrepasser le manque de vérité terrain. La figure 4.31 montre les résultats de détection des valeurs d'ombres ou spéculaires et la figure 4.30 montre des résultats de reconstruction 3d pour deux surfaces. Ces comparaisons permettent de constater que notre méthode offre une meilleure détection des ombres et produit moins de

FP sur les spécularités. Seul point négatif, lorsque la surface présente des cavités très importantes où l'ombre est très présente, aucune méthode ne permet une détection correcte (figure 4.29, dernière surface). Néanmoins, tout comme les résultats obtenus par simulation, notre méthode fournit des reconstructions 3d plus lisses en adéquation avec les surfaces réelles ce qui prouve son efficacité et le gain apporté dans l'extraction de l'information géométrique et photométrique pour la SP.

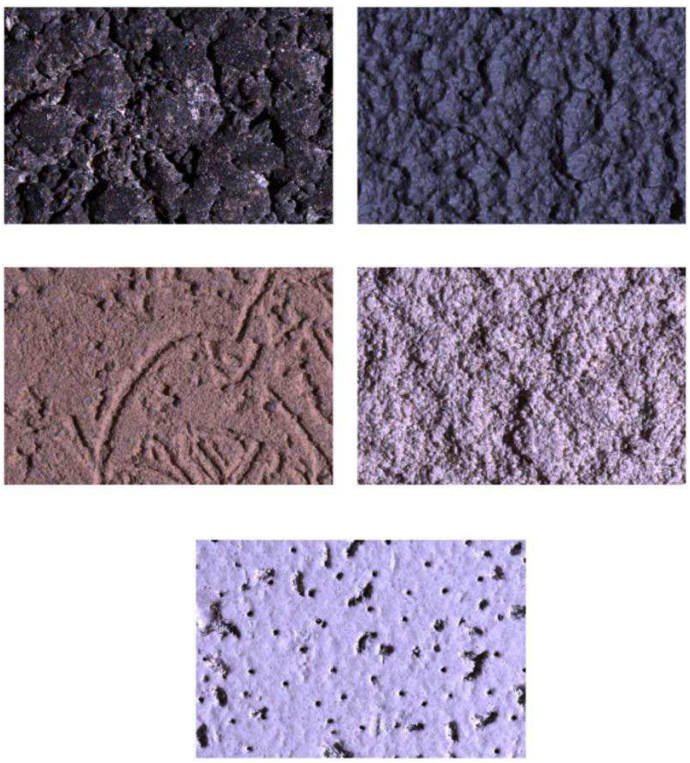


FIGURE 4.29 – *Exemples de surfaces réelles acquises et analysée.*

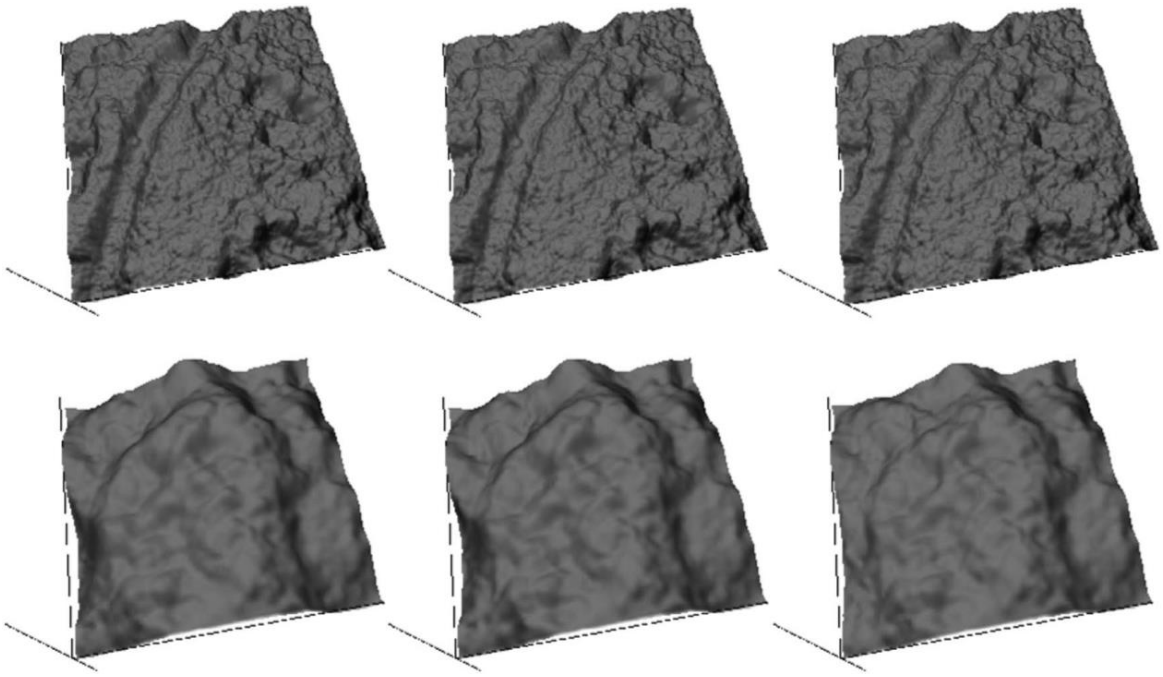


FIGURE 4.30 – *Comparaison des reconstructions 3d pour de gauche à droite la méthode $\mathcal{M}2$, $\mathcal{M}3$ et $\mathcal{OUR}$.*

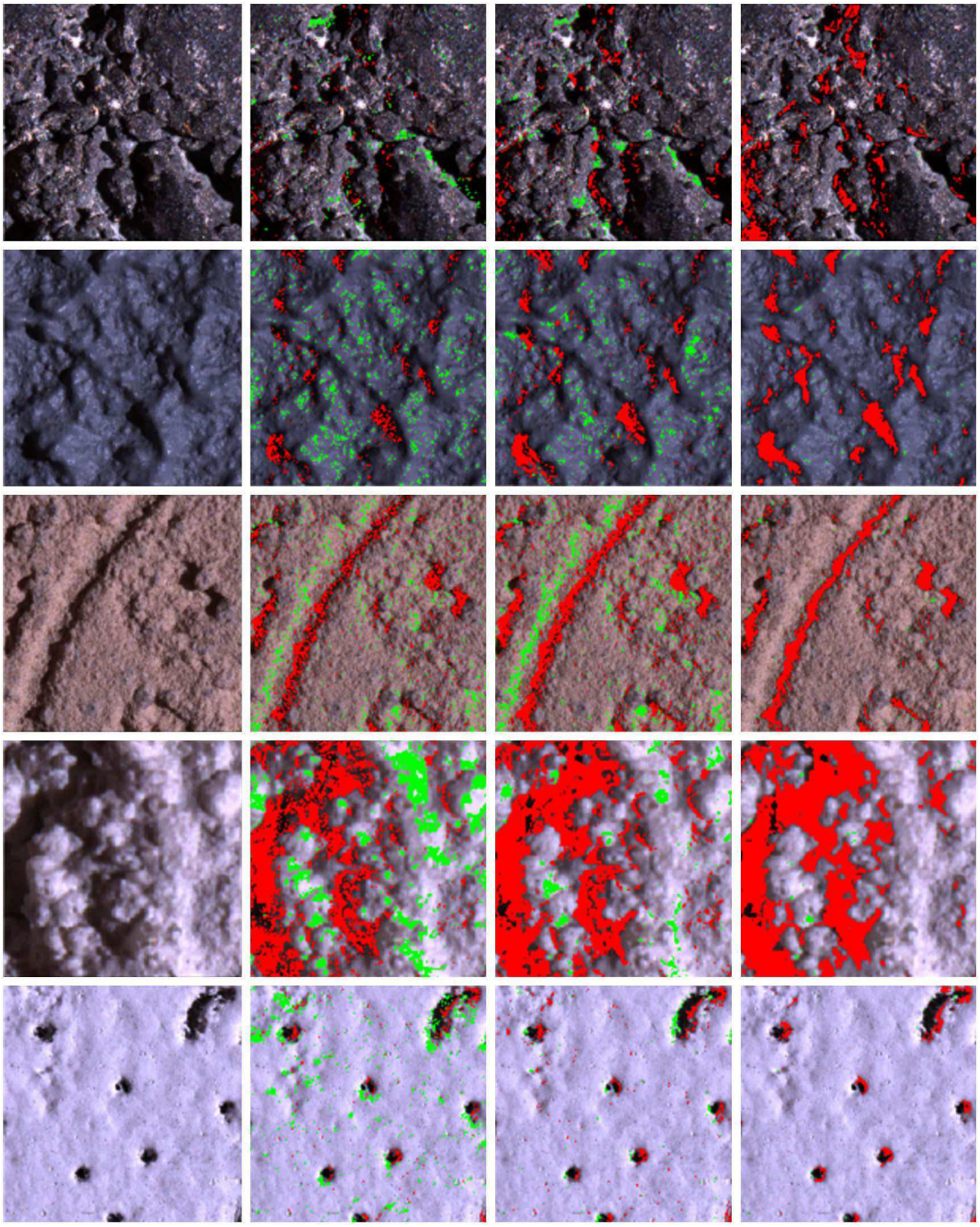


FIGURE 4.31 – *Comparaison de la détection des ombres (rouge) et de la spécularité (vert), pour de gauche à droite l'image de référence, méthode $\mathcal{M}2$, $\mathcal{M}3$ et $\mathcal{OUR}$.*

Conclusion du chapitre

Parmi les méthodes de reconstruction 3d par vision, notre intérêt est porté sur la stéréo-photométrie car elle présente l'avantage d'extraire les propriétés géométriques et colorimétriques des surfaces analysées même en cas de forte rugosité. Néanmoins, cette méthode n'est pas parfaite et son application nécessite la formulation d'hypothèses ($\mathcal{H}1$ à $\mathcal{H}4$) pour éviter des erreurs de reconstruction. Parmi ces hypothèses, la première $\mathcal{H}1$ est en rapport avec la nature des surfaces analysées. En effet, la formulation classique de la stéréo-photométrie, impose des surfaces parfaitement diffuses ou du moins la BRDF doit être connue et inversible en chaque point. Dans notre contexte d'application, les surfaces analysées sont majoritairement diffuses mais, selon des conditions d'éclairage particulières elles peuvent présenter certaines zones spéculaires ou ombrées. Pour outrepasser ces problèmes, la solution proposée consiste à combler le manque d'information par la mise en œuvre d'éclairages supplémentaires et les prises de vue correspondantes. En chaque point de la surface, nos traitements numériques permettent de conserver uniquement les pixels comportant une information utile pour l'inversion du modèle. Cette méthode ne nécessite pas de paramétrage manuel et s'adapte quel que soit le nombre de sources lumineuses utilisées.

Éclairage réaliste en SP 5

Sommaire

5.1 Influence des paramètres d'éclairage . . . 98
5.1.1 Etude des erreurs de position des sources . . . 98
5.1.2 Étude de la non-uniformité de l'éclairage . . . 103
5.2 Etat de l'art . . . 106
5.2.1 Calibration des positions des sources . . . 107
5.2.2 Méthodes de correction de la non-uniformité des sources . . . 117
5.3 Méthodes proposées . . . 119
5.3.1 Calibration des positions des sources . . . 119
5.3.2 Correction de la non-uniformité de l'éclairage dans la SP . . . 124
Conclusion . . . 128

Au cours du chapitre précédent, nous avons défini les contraintes d'application de la SP sous forme d'hypothèses $\mathcal{H}1$ à $\mathcal{H}4$. Notre première contribution permet de lever l'hypothèse $\mathcal{H}1$ liée aux propriétés photométriques des surfaces analysées. Ce nouveau chapitre est consacré à l'analyse des contraintes du système d'éclairage imposées par $\mathcal{H}2$ et $\mathcal{H}3$.

Selon l'hypothèse $\mathcal{H}2$, l'ensemble des méthodes de SP imposent la connaissance préalable des directions des sources lumineuses pour être applicables. Traditionnellement ces mesures sont réalisées manuellement par l'utilisateur, puis fournies aux algorithmes de calculs. Cette tâche est obligatoire, fastidieuse et sujette aux erreurs. La seconde hypothèse $\mathcal{H}3$ oblige l'utilisation de sources lumineuses directionnelles placées à une distance infinie pour produire un éclairage uniforme sur l'ensemble de la surface analysée. Cette contrainte n'est jamais respectée lors de la mise en œuvre, mais la plupart des auteurs émettent l'hypothèse stipulant que les sources sont situées à une distance suffisamment élevée par rapport à la taille et au relief de la surface à reconstruire pour être considérées comme placées à l'infini. Ainsi tout point de la surface est considéré comme éclairé par un rayon lumineux de même intensité et de même direction d'incidence. De plus, toutes les sources sont supposées parfaitement identiques et seule leur position géométrique par rapport à la surface diffère. En pratique et pour un grand nombre d'applications, cette hypothèse s'avère difficile voire impossible à respecter pour des raisons d'encombrement. En effet, par souci de miniaturisation et de confinement afin d'éliminer les phénomènes d'inter-réflexion ou de lumière ambiante, les sources sont généralement ponctuelles et situées à une faible distance de la surface. Nous présentons dans ce chapitre un état de l'art pour calibrer l'éclairage dans le cadre de la SP. Nous apportons une contribution significative pour tenir compte des conditions réelles d'éclairage.

5.1 Influence des paramètres d'éclairage

La plupart des applications SP suivent un schéma d'éclairage où seul l'angle azimut varie (figure 5.1). Mais il s'avère difficile de placer mécaniquement les sources dans des directions exactes. Par conséquent, nous proposons une première étude consacrée à l'influence des erreurs de position entre le système d'éclairage réel et la matrice d'éclairage [**L**] utilisée par la SP.

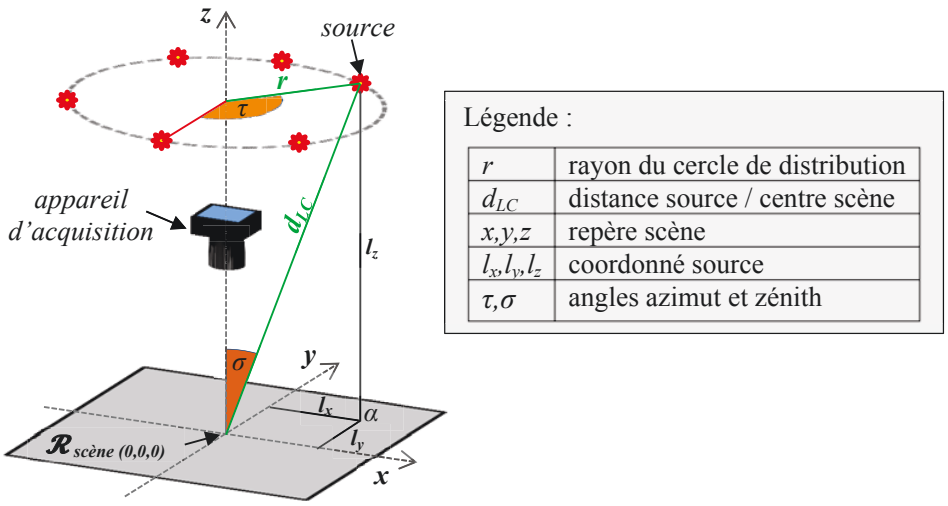


FIGURE 5.1 – *Schéma d'un système d'acquisition 6 sources lumineuses utilisé pour la stéréo-photométrie.*

Pour cette étude et dans un premier temps, nous conservons l'hypothèse $\mathcal{H}3$: les sources sont considérées comme directionnelles et uniformes. Elles sont orientées en direction du centre du repère $\mathcal{R}_{scène}$ correspondant à l'intersection entre l'axe optique de la caméra et le plan support des axes x et y. Seuls les angles azimuts et zéniths peuvent être changés. Dans un second temps, nous proposons d'étudier les distorsions des reconstructions 3d générées par la non-uniformité de l'éclairage due au non respect des contraintes fixées par l'hypothèse $\mathcal{H}3$.

5.1.1 Etude des erreurs de position des sources

Au cours de cette étude, nous utilisons des images de synthèse obtenues suivant le modèle Lambertien et par le logiciel de lancer de rayons POV-RAY. L'objet modélisé est une demi-sphère de rayon 70.8 mm placée au centre d'une image dont la taille réelle couvre une zone de 141.6×94.8 mm. La caméra est située à 90 cm du centre de $\mathcal{R}_{scène}$ et le modèle de projection est orthographique à facteur d'échelle. Le système d'éclairage initial est modélisé suivant le schéma classique présenté figure 5.1 où six sources sont placées sur un plan parallèle à la surface analysée et sont distribuées tous les 60^o (angles azimuts) suivant un cercle dont le centre est la caméra. Cette configuration correspond à notre système d'acquisition NIKON_1 A.2.1.

Afin d'évaluer uniquement l'influence du système d'éclairage, nous reprenons les contraintes imposées par $\mathcal{H}1$ en simulant une surface parfaitement diffuse, sans phénomènes d'ombres portées, d'inter-réflexions ou de lumière ambiante. Pour ne pas subir les ombres portées seule une portion de la demi-sphère est modélisée et les sources lumineuses sont placées dans un cône (figure 5.2) formé par les tangentes à la surface aux points de contact entre la demi-sphère et le plan.

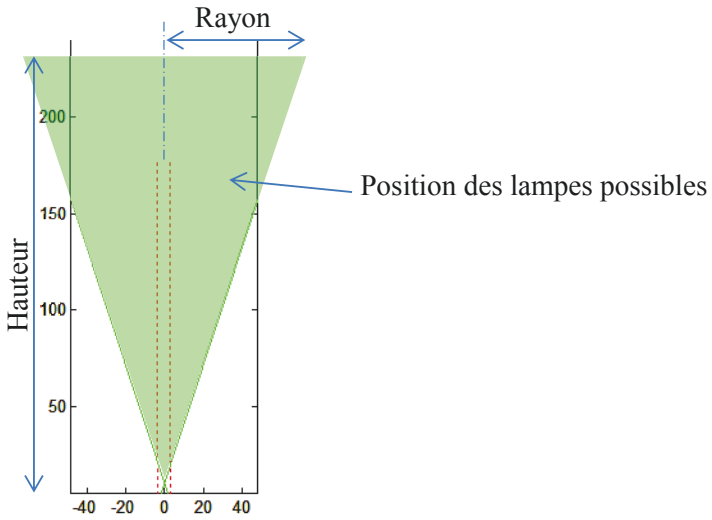


FIGURE 5.2 – *Schéma illustrant le principe de placement des sources lumineuses afin d'éviter les phénomènes d'ombre portées.*

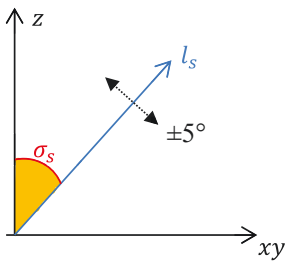
5.1.1.1 Erreurs sur l'angle zénith

Pour une matrice **L** dont tous les angles zénithaux sont identiques et égaux, nous appliquons une erreur de position sur cet angle pour une des sources lumineuses (tableau 5.1) lors de la simulation d'éclairage. Pour déterminer si des distorsions apparaissent sur les cartes des hauteurs et étudier leurs tendances en fonction de l'angle zénith initial, nous simulons deux cas : un angle zénith initial à 7^o puis à 15^o. Afin de visualiser les distorsions, les cartes des hauteurs obtenues à partir d'une matrice d'éclairage erronée sont directement comparées à la carte sans erreurs. Cette analyse est complétée par une mesure de SNR sur les cartes des hauteurs et une représentation de l'erreur angulaire entre les champs de gradients définie par :

$$\epsilon_\Psi = |\Psi_1 - \Psi_2| \tag{5.1}$$

où Ψ est la représentation angulaire des champs de gradients p et q tel que :

$$\Psi = \arctan(\frac{p}{q}) \pm \pi \tag{5.2}$$



(a)

s	τ_s en ^{o}	σ_s en ^{o}
1	0	7 ± 5
2	60	7
3	120	7
4	180	7
5	240	7
6	300	7

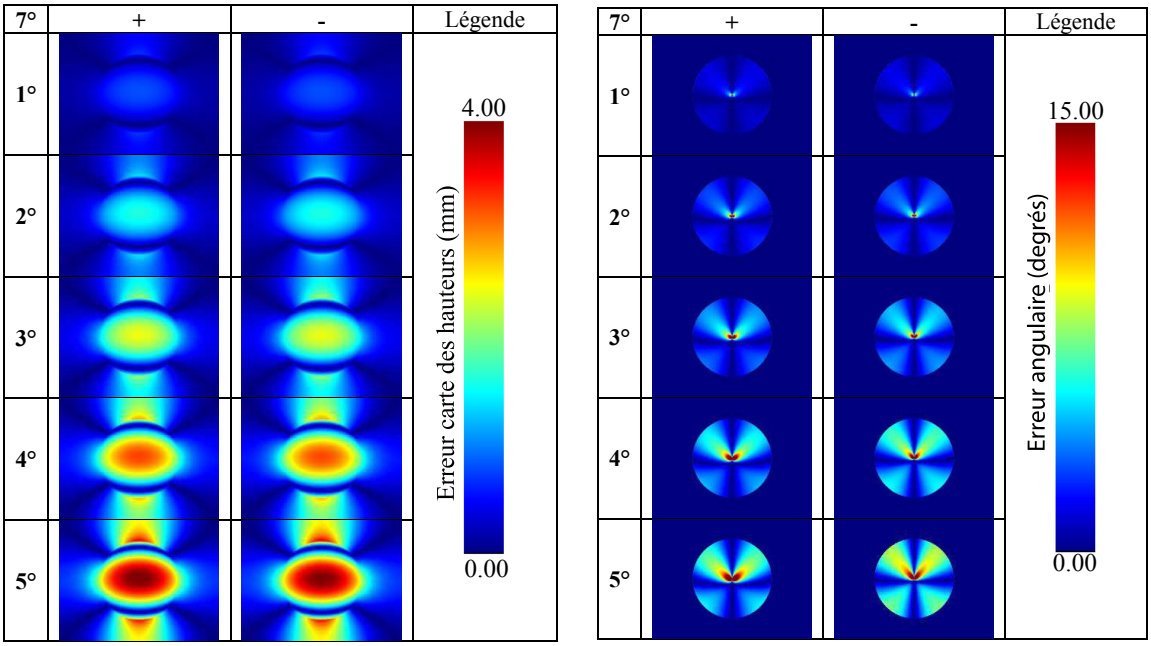
(b)

s	τ_s en ^{o}	σ_s en ^{o}
1	0	15 ± 5
2	60	15
3	120	15
4	180	15
5	240	15
6	300	15

(c)

TABLE 5.1 – *Schéma illustrant la variation angulaire zénith d'une source lumineuse (a), matrice du système d'éclairage réel (b-c).*

Pour un angle zénith initial de 7^o, les déformations dans les cartes des hauteurs estimées apparaissent (figure 5.3a) dès 1^o d'erreur. Elles proviennent des distorsions dans les champs de gradients (figure 5.3b). Ces distorsions évoluent proportionnellement à l'erreur angulaire (figure 5.5) et entraînent une chute significative du SNR (figure 5.4). Avec une erreur angulaire élevée mais potentiellement réaliste (5^o), les distorsions dans la reconstruction atteignent des valeurs significatives ($SNR < 28dB$). L'erreur moyenne en élévation est alors supérieure à 1.15 mm pour un maximum de 3.66 mm, soit 10.66% de l'élévation totale de la surface. Les champs de gradients présentent alors une distorsion angulaire moyenne supérieure à 1.14^o.



(a) (b)

FIGURE 5.3 – *Comparaison (a) des erreurs de reconstruction 3d, (b) des distorsions angulaires dans les champs de gradients en fonction de l'erreur zénith pour une plage allant de -5^o à 5^o et un angle initial de 7^o.*

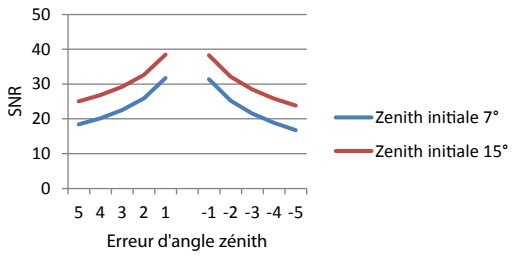


FIGURE 5.4 – *Evolution des SNR en fonction des erreurs d'angles zéniths.*

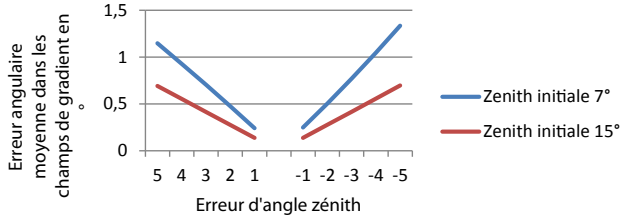
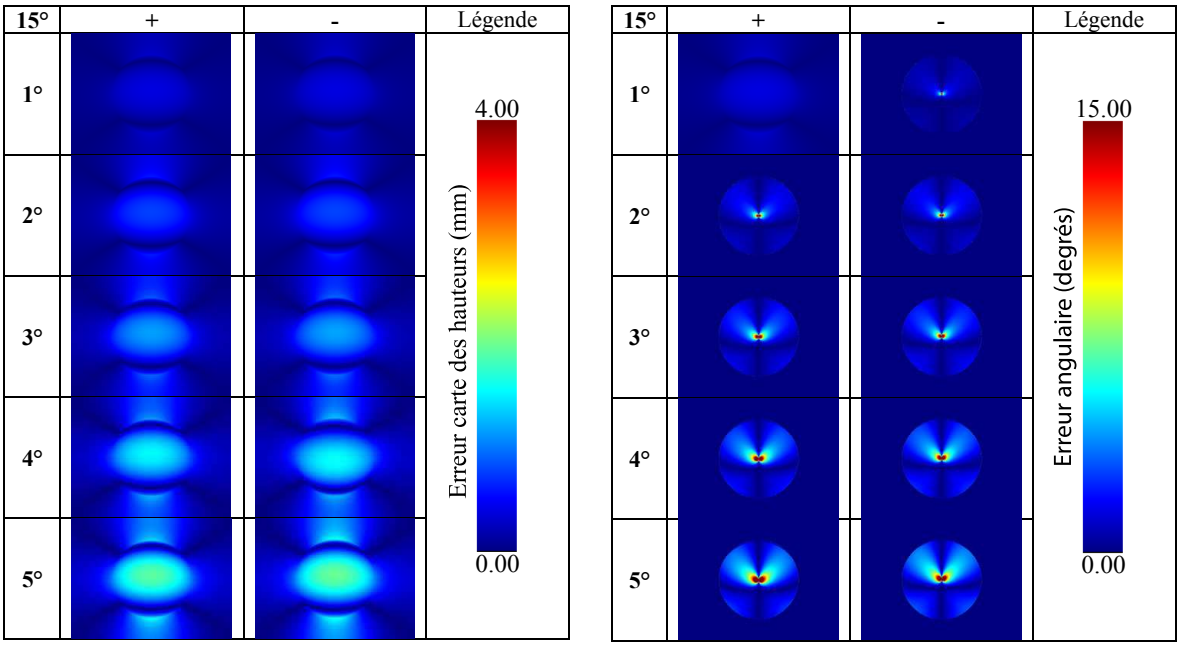


FIGURE 5.5 – *Evolution de l'erreur angulaire moyenne des champs de gradient en fonction des erreurs d'angles zéniths.*

Avec un angle zénith initial de 15^o, les distorsions des reconstructions et des champs de gradients suivent les mêmes tendances (figure 5.6) que le cas précédent. Ainsi quel que soit l'angle zénith initial, une erreur de position entraîne une distorsion des reconstructions 3d.



(a) (b)

FIGURE 5.6 – *Comparaison (a) des erreurs de reconstruction 3d, (b) des distorsions angulaires dans les champs de gradients en fonction de l'erreur zénith pour une plage allant de -5^o à 5^o et un angle initial de 15^o.*

5.1.1.2 Erreurs sur l'angle azimut

Seul l'angle azimut d'une des six sources est soumis à des variations. Le protocole de comparaison reste identique à l'étude de l'erreur sur l'angle zénith. La source choisie est celle placée initialement à 0^o et les variations s'étendent sur une plage de 10^o. Nous analysons uniquement le cas positif car l'inverse est identique. Les déformations des reconstructions et les distorsions des champs de gradients sont présentées figure 5.8. Les valeurs de SNR et les erreurs angulaires relevées (figure 5.7) en fonction de l'erreur azimut montrent des tendances similaires à celles obtenues pour les erreurs d'angle zénith. Les distorsions des champs de gradients évoluent proportionnellement à l'erreur d'angle azimut ce qui entraîne une chute du SNR.

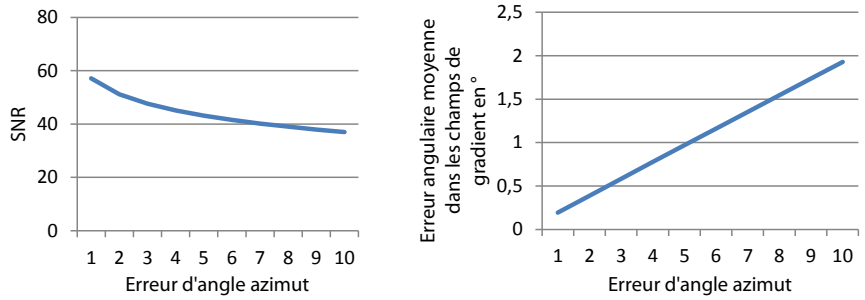
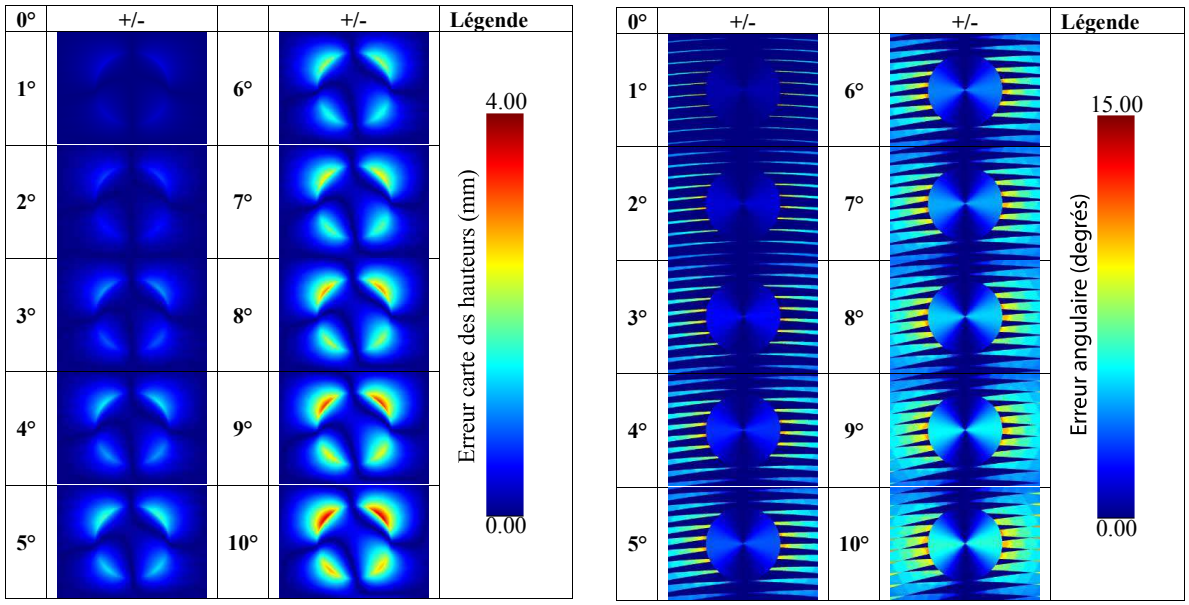


FIGURE 5.7 – *Evolution du SNR et de l'erreur angulaire moyenne des champs de gradient en fonction de l'erreur d'angle azimut.*



(a) (b)

FIGURE 5.8 – *Comparaison (a) des erreurs de reconstruction 3d, (b) des distorsions angulaires dans les champs de gradients en fonction de l'erreur azimut pour une plage allant de 1^o à 10^o et un angle initial de 0^o.*

5.1.1.3 Erreurs globales sur les positions des sources

A titre individuel, les erreurs sur les directions des sources lumineuses sont génératrices de distorsions dans le calcul des champs de gradients et entraînent des distorsions non négligeables dans l'estimation des cartes 3d. Pour démontrer l'impact cumulé de ces erreurs, nous proposons une dernière analyse où plusieurs erreurs minimes ($< 2^o$) sont aléatoirement distribuées (tableau 5.2) entre les positions réelles des sources et les angles azimuts et zéniths utilisés par la SP.

s	τ_s en ^{o}	σ_s en ^{o}
1	2.0	15.5
2	59.5	14.9
3	121.8	16.8
4	178.1	15.2
5	240.3	13.1
6	301.0	14.2

(a)

s	τ_s en ^{o}	σ_s en ^{o}
1	0.0	15.0
2	60.0	15.0
3	120.0	15.0
4	180.0	15.0
5	240.0	15.0
6	300.0	15.0

(b)

TABLE 5.2 – *Matrice du système d'éclairage réel (a) et matrice du système d'éclairage utilisée pour la résolution SP (b).*

Ces multiples erreurs sont représentatives des applications réelles. Malgré leurs faibles valeurs, elles entraînent des distorsions non négligeables dans le calcul des gradients ainsi que des déformations 3d (figure 5.9). L'erreur angulaire moyenne des champs de gradients est supérieure à 28^o, elle s'accompagne d'un SNR de $32,47dB$ qui correspond à une erreur de 0.20 mm en élévation pour un maximum de 0.871 mm, soit 2.53% de l'élévation totale de la surface analysée.

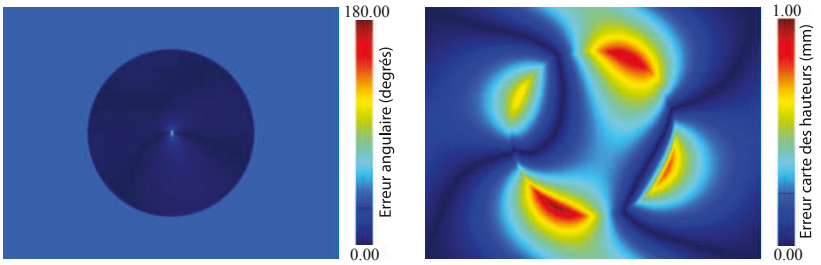


FIGURE 5.9 – *Comparaison des distorsions angulaires dans les champs de gradients et des erreurs de reconstruction 3d en fonction de minimes erreurs angulaires.*

Cette étude sur l'influence des positions des sources lumineuses démontre l'importance de la connaissance des angles azimuts et zéniths lors de la reconstruction 3d par SP.

5.1.2 Étude de la non-uniformité de l'éclairage

Pour cette étude nous avons besoin de définir la différence entre une source lumineuse directionnelle et une source lumineuse ponctuelle. Une source de lumière directionnelle est une source dont les rayons ont la même direction en tout point de l'espace. A l'inverse, la lumière ponctuelle provient d'un point dans l'espace qui éclaire uniformément dans toutes les directions. Une source ponctuelle peut être considérée comme directionnelle lorsqu'elle est très éloignée de la scène car les rayons reçus sont estimés parallèles. Un exemple parfait de lumière directionnelle à l'échelle de la vision humaine est le soleil car tous ses rayons vont dans la même direction en tout point du globe du fait de la taille et de la distance entre les deux astres.

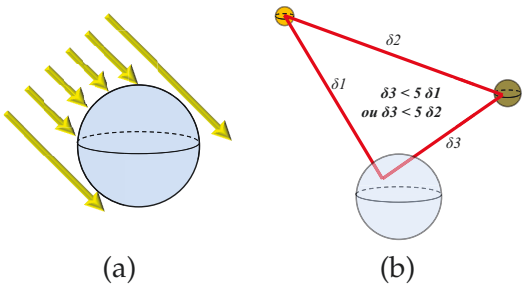
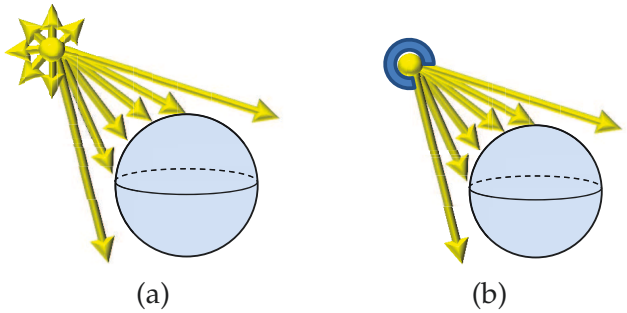


FIGURE 5.10 – *Schémas, (a) exemple d'une source directionnelle, (b) limite de l'hypothèse directionnelle.*

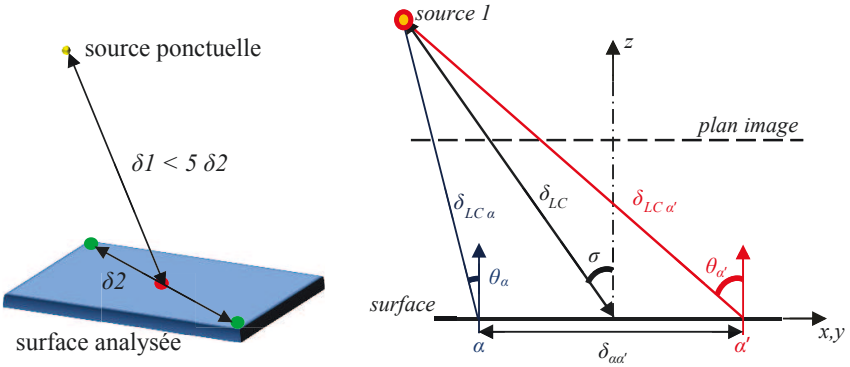
Cette assimilation du soleil comme source lumineuse directionnelle n'est valable que lorsque les points d'observation sont peu distants par rapport à la distance de la source. Dans la littérature, ce rapport inter-distances est généralement fixé à cinq. Par conséquent, si la distance entre deux points d'observation est cinq fois plus grande qu'une des distances source/points, le soleil doit être considéré comme une source lumineuse ponctuelle. La lumière ponctuelle provient d'un point dans l'espace qui éclaire uniformément dans toutes les directions.

Suivant la technologie employée, les sources ponctuelles peuvent être non uniformes (appelée aussi spot), la lumière est diffusée dans une direction précise selon un angle d'ouverture. Par souci d'encombrement, la mise en œuvre de la SP se fait généralement à partir de ce type de source lumineuse dont la distance $\delta 1$ par rapport au centre de la surface analysée est inférieure à cinq fois la distance $\delta 2$ entre les points les plus éloignés à la surface de la scène (figure 5.12a).



(a) (b)

FIGURE 5.11 – *Schémas, (a) exemple d'une source ponctuelle uniforme, (b) limite de l'hypothèse ponctuelle uniforme (spot).*



(a) (b)

FIGURE 5.12 – *Schémas, (a) distance entre deux points d'une surface par rapport à la distance entre la source lumineuse et le centre de la surface, (b) coupe 2d du modèle Lambertien.*

A partir d'une coupe 2d du modèle de Lambert (figure 5.12b), il est facile de constater qu'une source ponctuelle positionnée à une distance relativement faible du centre d'une surface ne produit pas les mêmes conditions d'éclairement en tout point. Pour étudier l'influence de ces variations, nous analysons la différence Δ de l'angle d'incidence θ, de l'intensité Ie_0 incidente à la surface et l'intensité Ie émise par la surface en fonction du rapport de distance $\delta_{LC}/\delta_{\alpha\alpha'}$ pour deux points distincts α et α' positionnés à équidistance du centre de la surface.

Ces différences (figure 5.13) dépendent de l'angle zénith σ et se répercutent sur les intensités émises par la surface Ie_α et $Ie_{\alpha'}$ proportionnellement au rapport entre la distance δ_{LC} et $\delta_{\alpha\alpha'}$. En effet, d'après le modèle de réflectance Lambertien, Ie dépend de l'angle d'incidence et surtout de l'intensité Ie_0 reçue. Ie_0 varie en fonction du carré de la distance entre le point d'émission (la source) et le point de réception (surface).

$$Ie_{0\alpha} = \frac{Ie_{0source}}{\delta_{LC\alpha}^2} \tag{5.3}$$

avec :

$Ie_{0source}$	:	intensité constante au point d'émission quel que soit α.
$\delta_{LC\alpha}$	:	la distance entre le point source et le point de la surface.

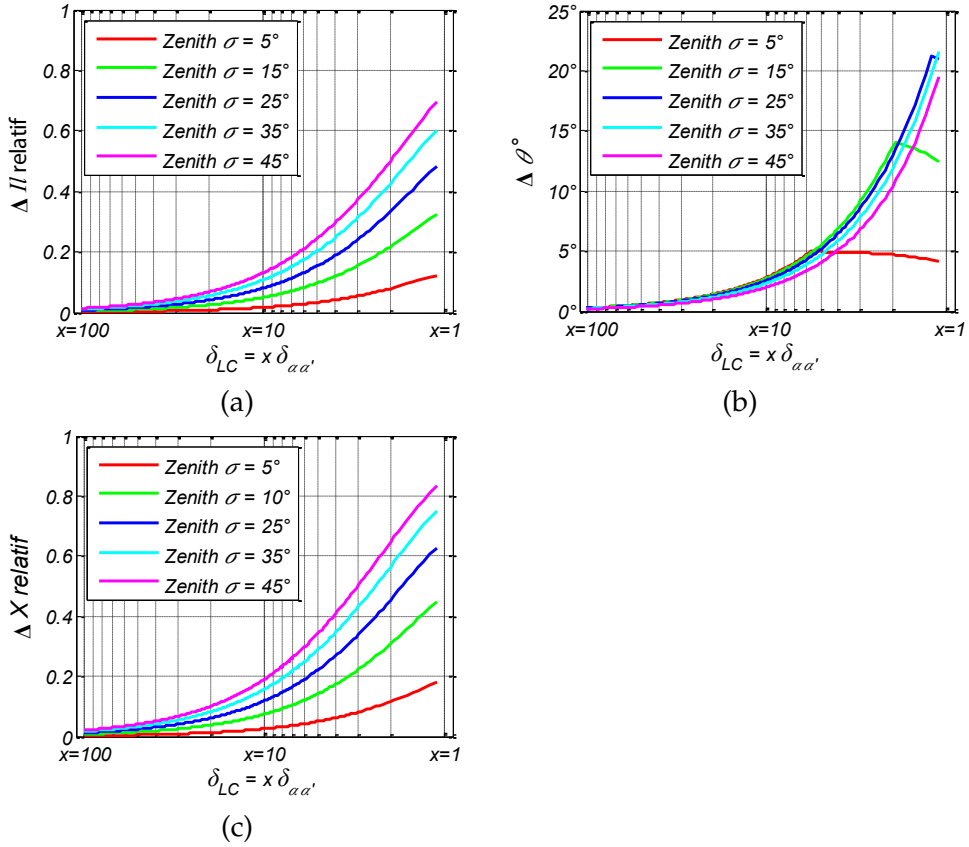


FIGURE 5.13 – *Évolution Δ de (a) l'intensité Ie_0, (b) l'angle θ et (c) de l'intensité Ie émise par la surface en fonction du rapport de distance $\delta_{LC}/\delta_{\alpha\alpha'}$ et l'angle zénith σ de la source lumineuse.*

La distance $\delta_{LC\alpha}$ est un calcul de distance en trois dimensions, effectué à partir des coordonnées de la source lumineuse et les coordonnées des points de la surface :

$$\delta_{LC\alpha} = \sqrt{(\alpha_x - l_{s_x})^2 + (\alpha_y - l_{s_y})^2 + (\alpha_z - l_{s_z})^2} \tag{5.4}$$

avec :

$\alpha_x, \alpha_y, \alpha_z$	:	les coordonnées 3d de la surface éclairée au point α.
$l_{s_x}, l_{s_y}, l_{s_z}$	:	les coordonnées 3d de la source s.

Or l'hypothèse $\mathcal{H}3$, utilisée pour simplifier la résolution de la SP, considère que les valeurs Ie_0 sont identiques pour l'ensemble des sources et l'ensemble des points de la surface. Cette hypothèse génère des erreurs de reconstruction car chaque point de la surface est éclairé avec une intensité et une direction d'incidence différente. Les champs de gradients et l'estimation de la carte des hauteurs sont biaisés, les surfaces planes sont reconstruites sous forme de dôme (figures 5.14 a,d et g) dont le sommet correspond au point qui bénéficie du maximum d'éclairement des trois sources lumineuses. Les surfaces plus complexes (figures 5.14 b,e,h et c,f,i) présentent des déformations de la macro et micro géométrie.

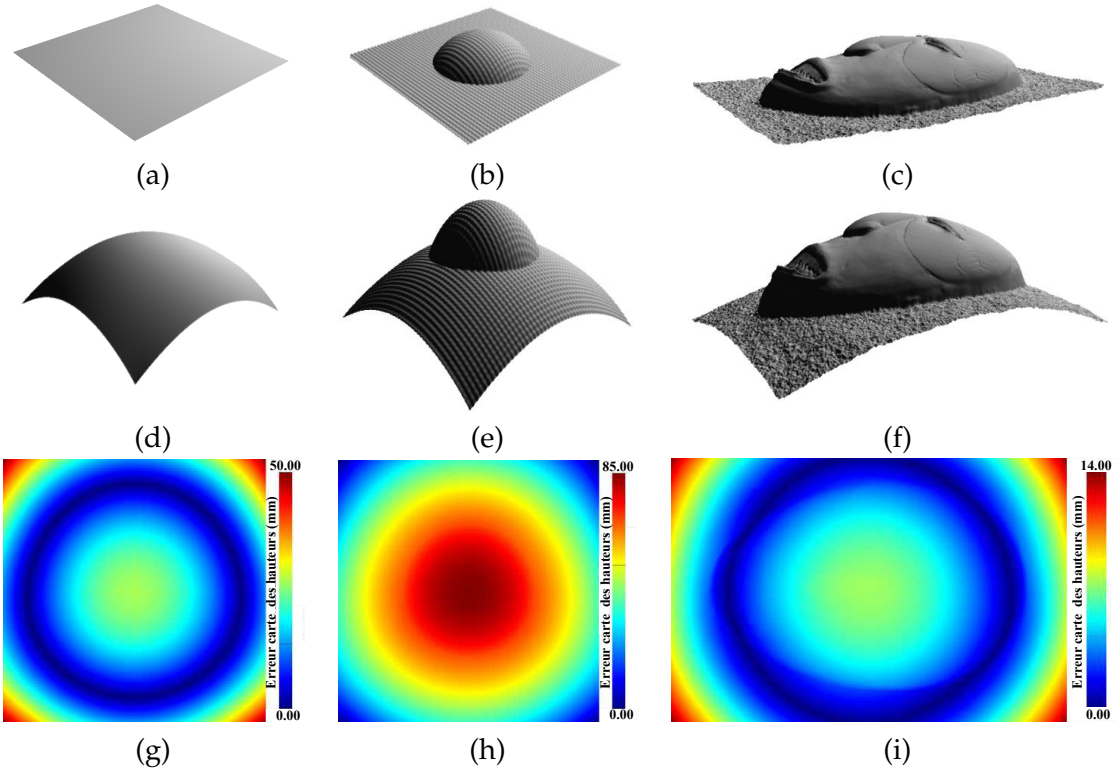


FIGURE 5.14 – *(a-c) surfaces réelles vérités terrains, (d-f) surfaces estimées par la méthode SP classique, (g-i) différences de hauteurs entre les carte estimée et la vérité terrain ou son approximation.*

5.2 Etat de l'art

Les premières méthodes de calibration des sources lumineuses [BH85, CBG94, LR85, Pen82] sont apparues dans les années 1980-90 avec le Shape From Shading et reposent sur l'analyse statistique des intensités produites par une surface Lambertienne dont les normales sont isotropiquement distribuées. Elles sont capables d'estimer une unique direction ainsi que l'albédo ou l'intensité de la source. Au début des années 1990, des auteurs tels que [ZC91, Wei90, YY91] délaissent l'analyse des intensités au profit de l'analyse des contours produits par les ombres. A la fin des années 1990 et début 2000, les méthodes basées sur l'analyse des contours évoluent [ZY01]. Le domaine du ré-éclairage de scène s'intéresse à l'estimation de la position des sources lumineuses afin d'incruster des objets virtuels dans des scènes réelles. De nouvelles méthodes émergentent, par exemple DEBEVEC [Deb08] ou POWELL [PSG01] proposent d'utiliser les points spéculaires sur des sphères miroir pour trianguler la position des sources. Sur le même principe de triangulation, [ZK02] estime la position des sources lumineuses à partir d'une paire d'images et le principe de la Stéréo-Vision. La triangulation permet l'estimation des sources ponctuelles (angles zénith, azimut, distance source objet et intensité) car auparavant seule les sources directionnelles (angles zénith, azimut et intensité) pouvaient être estimées. Plus récemment, l'estimation des sources lumineuses est utilisée pour améliorer la classification de surfaces texturées. Les travaux [LLLS03] utilisent l'analyse statistique des intensités, des ombres et des points spéculaires sur des surfaces texturées pour estimer la position des sources. Les travaux les plus récents combinent l'analyse statistique et géométrique pour fournir une estimation quelque soit la géométrie et la réflectance de la scène. A l'inverse, les travaux sur la non-uniformité de l'éclairage en SP sont peux nombreux. Néanmoins quelques auteurs [PCF05, SSSF10] se sont intéressés à ce problème et ont apporté des solutions basées sur le prétraitement des images.

5.2.1 Calibration des positions des sources

Nous proposons un état de l'art des méthodes d'estimation de la position de sources lumineuses non visibles dans une image. Les trois principes de bases de l'estimation de la position des sources sont : l'analyse statistique des intensités, l'analyse des ombres et contours et l'analyse des points spéculaires. Pour chacun de ces principes, l'intérêt est porté sur la méthode originale et les extensions les plus importantes. Afin de pouvoir mener une étude comparative, nous avons implémenté et testé diverses méthodes en utilisant un maximum de paramètres communs. Cette étude est présentée en Annexe A.4. Les méthodes trop éloignées du domaine de la SP ou utilisant plusieurs principes de base ne sont pas abordées.

5.2.1.1 Méthodes basées sur l'analyse statistique des intensités

Pentland est le premier à s'intéresser aux problèmes d'estimation d'une source de lumière non directement visible dans une image. Ces travaux [Pen82] utilisent l'analyse statique des intensités émises par une surface connue pour récupérer la direction d'éclairage d'une scène sous conditions particulières :

- la surface doit être parfaitement Lambertienne.
- une unique source lumineuse éclaire la surface.
- les normales de la surface doivent être distribuées de façon isotrope.

Les deux premières conditions se retrouvent dans beaucoup de travaux, elles permettent d'appliquer une analyse statistique. La troisième condition limite le type d'objet analysable car les normales doivent impérativement être distribuées de façon isotrope. Ainsi, des surfaces planes ou cylindriques sont inappropriées car elles présentent un fort biais dans une direction et ne répondent pas aux conditions d'application.

$$\mathbb{E}(\mathrm{d}\vec{n}) = \frac{\sum \mathrm{d}\vec{n}}{\mathrm{card}(\vec{n})} = 0 \tag{5.5}$$

L'approche de Pentland repose sur l'exploitation du modèle Lambertien et l'exclusion du cas $(\vec{n}.\vec{l}) \leq 0$. Par conséquent, les zones d'ombre doivent être exclues du domaine d'analyse. Avec un albédo et une direction constante de la source lumineuse, il est possible de réécrire le modèle Lambertien comme :

$$\mathrm{d}Ie = \rho\ Ie_0\ (\mathrm{d}\vec{n}.\vec{l}) \tag{5.6}$$

Cette équation sert de base pour construire un système de régression linéaire et obtenir une estimation des directions $\hat{l}_x$ et $\hat{l}_y$ de la source lumineuse. Le système de régression utilise la variation moyenne de l'intensité suivant une direction spécifique de dérivation :

$$\begin{bmatrix} \mathrm{d}\bar{I}e_{\theta 1} \\ \vdots \\ \mathrm{d}\bar{I}e_{\theta i} \end{bmatrix} = \begin{bmatrix} \cos\theta 1 & \sin\theta 1 \\ \vdots & \vdots \\ \cos\theta i & \sin\theta i \end{bmatrix} \begin{bmatrix} \hat{l}_x \\ \hat{l}_y \end{bmatrix} = [\mathbf{A}] \begin{bmatrix} \hat{l}_x \\ \hat{l}_y \end{bmatrix} \tag{5.7}$$

où $\mathrm{d}\bar{I}e_{\theta i}$ est la moyenne de la dérivée de l'intensité suivant la direction $(\cos\theta i \quad \sin\theta i)$. Généralement, $i = 1 \ldots 8$ directions sont utilisées : (2 verticales, 2 horizontales et 4 diagonales).

Pentland préconise de résoudre le système par la méthode des moindres carrés approchés :

$$\begin{bmatrix} \hat{l}_x \\ \hat{l}_y \end{bmatrix} = ([\mathbf{A}]^T[\mathbf{A}])^{-1}[\mathbf{A}]^T \begin{bmatrix} \mathrm{d}\bar{I}e_{\theta 1} \\ \vdots \\ \mathrm{d}\bar{I}e_{\theta i} \end{bmatrix} \tag{5.8}$$

Après résolution, l'angle azimutal est directement estimé à partir des valeurs $\hat{l}_x$ et $\hat{l}_y$:

$$\hat{\tau} = \arctan(\frac{\hat{l}_y}{\hat{l}_x}) \pm \pi \tag{5.9}$$

L'angle zénithal est estimé grâce à la dérivée de l'image suivant l'angle $\hat{\tau}$:

$$\hat{\sigma} = \arccos\left[1 - \frac{\hat{l}_x^{\,2} + \hat{l}_y^{\,2}}{a^2}\right]^{\frac{1}{2}} \tag{5.10}$$

Avec a l'écart-type de la dérivée de l'intensité suivant la direction $(\cos\theta i \quad \sin\theta i)$:

$$a = \sqrt{\mathbb{E}\left[\mathrm{d}Ie_{\hat{\tau}}^2\right] - \mathbb{E}\left[\mathrm{d}Ie_{\hat{\tau}}\right]^2} \tag{5.11}$$

Les travaux de Lee et Rosenfeld [LR85] reprennent l'estimateur de Pentland [Pen82]. Les auteurs tirent la conclusion qu'il n'est pas nécessaire de dériver les intensités d'une image suivant toutes les directions. L'estimation de l'angle azimut est calculée à partir des dérivées premières de l'intensité suivant la direction x et y. Il supprime le système de régression linéaire car son utilisation offre peu d'avantages, sauf peut-être la réduction des effets du bruit. La formule d'estimation de l'angle azimut devient :

$$\hat{\tau} = \arctan(\frac{\mathbb{E}(\mathrm{d}Ie_y)}{\mathbb{E}(\mathrm{d}Ie_x)}) \pm \pi \tag{5.12}$$

L'angle $\hat{\sigma}$ doit satisfaire :

$$\frac{\mathbb{E}(Ie)}{\sqrt{\mathbb{E}(Ie^2)}} = 8\frac{\pi\cos\hat{\sigma} - \hat{\sigma}\cos\hat{\sigma} + \sin\hat{\sigma}}{3\pi(1+\cos\hat{\sigma})^{3/2}} \tag{5.13}$$

où $\mathbb{E}(Ie^2)$ est la moyenne des intensités au carré et la partie droite de l'égalité correspond à une fonction monolithique décroissante (figure 5.15). Il existe une unique solution pour $\hat{\sigma}$ à condition que :

$$0.8488 < \frac{\mathbb{E}(Ie)}{\sqrt{\mathbb{E}(Ie^2)}} < \frac{2\sqrt{2}}{3} \tag{5.14}$$

Dans le cas ou $\frac{\mathbb{E}(Ie)}{\sqrt{\mathbb{E}(Ie^2)}}$ n'est pas compris dans ces limites, la valeur de l'angle $\hat{\sigma}$ est 0^o ou 90^o.

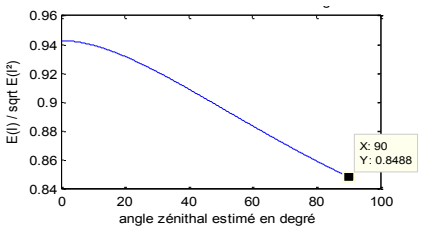


Figure 5.15 – Lee *et* Rosenfeld, *fonction monolithique décroissante pour l'estimation de l'angle zénith.*

Comme Pentland ou Lee et Rosenfeld, Zheng et Chellappa [ZC91] basent leurs travaux sur le modèle Lambertien. Ils proposent deux méthodes pour l'estimation de l'angle azimut, la première appelée *"voting"* utilise les changements d'intensité locale. La seconde appelée *"contour"* est basée sur l'étude des contours des objets.

Avec la méthode *"voting"*, l'estimation repose sur la connaissance préalable des normales de la surface utilisée. Cette surface doit présenter une distribution isotrope des normales. En partant d'une surface sphérique, les auteurs reformulent un système de régression semblable à l'estimateur azimut de Pentland. Le système est construit à partir de l'évolution des intensités $\mathrm{d}Ie_u$ entre le centre (x_0, y_0) et plusieurs points (x_u, y_u) de la surface sphérique. Sa résolution permet d'estimer la position de la source lumineuse suivant les directions $\hat{l}_x$ et $\hat{l}_y$:

$$\mathrm{d}\vec{Ie} = \begin{bmatrix} \mathrm{d}Ie_{11} \\ \mathrm{d}Ie_{22} \\ \vdots \\ \mathrm{d}Ie_{UV} \end{bmatrix} = \begin{bmatrix} \mathrm{d}x_1 & \mathrm{d}y_1 \\ \mathrm{d}x_2 & \mathrm{d}y_2 \\ \vdots & \vdots \\ \mathrm{d}x_U & \mathrm{d}y_V \end{bmatrix} \begin{bmatrix} \hat{l}_x \\ \hat{l}_y \end{bmatrix} = [\mathbf{A}]\,\hat{\vec{l}}$$
$$\hat{\vec{l}} = ([\mathbf{A}]^T[\mathbf{A}])^{-1}[\mathbf{A}]^T\,\mathrm{d}\vec{Ie} \tag{5.15}$$

avec :

$$\mathrm{d}Ie(x_u, y_v) = Ie(x_0, y_0) - Ie(x_u, y_v)$$

et

$$\mathrm{d}x_u = x_0 - x_u$$
$$\mathrm{d}y_v = y_0 - y_v$$

avec :

x_0, y_0	:	position métrique du centre optique dans $\mathcal{R}_{scène}$.
x_u, y_u	:	position métrique d'un pixel (u, v) dans $\mathcal{R}_{scène}$.

Les auteurs sont évasifs sur les nombres U, V et les directions des points à choisir pour résoudre le système. Après résolution, l'estimation de l'angle azimut se fait suivant :

$$\hat{\tau} = \arctan\left(\frac{\mathbb{E}\left(\frac{\hat{l}_y}{\sqrt{\hat{l}_x{}^2+\hat{l}_y{}^2}}\right)}{\mathbb{E}\left(\frac{\hat{l}_x}{\sqrt{\hat{l}_x{}^2+\hat{l}_y{}^2}}\right)}\right) \pm \pi \tag{5.16}$$

La méthode *"contour"* utilise les contours des objets. En chaque point de contour, l'angle zénith v de la normale à la surface est constant et l'angle azimut φ_α est perpendiculaire au contour dans le plan image. Toujours en partant du modèle Lambertien, l'équation de l'intensité du α^{th} point de contour de l'objet peut s'écrire :

$$\begin{aligned} Ie_\alpha &= \rho(\cos\tau\sin\sigma\ \cos\varphi_\alpha\sin v + \sin\tau\sin\sigma\ \sin\varphi_\alpha\sin v + \cos\sigma\cos v) \\ &= [\cos\varphi_\alpha\ \ \sin\varphi_\alpha]\begin{bmatrix} \rho\sin\sigma\sin v\cos\tau \\ \rho\sin\sigma\sin v\sin\tau \end{bmatrix} + \rho\cos\sigma\cos v \end{aligned} \tag{5.17}$$

où la normale $\vec{n} = [\cos\varphi \sin v,\ \sin\varphi \sin v,\ \cos v]$ est donnée pour φ et v l'angle azimutal et zénithal entre la normale à la surface et le plan image. Et la direction de la source lumineuse $\vec{l} = [\cos\tau \sin\sigma,\ \sin\tau \sin\sigma,\ \cos\sigma]$ est donnée pour τ et σ l'angle azimutal et zénithal.

Pour l'ensemble des valeurs $\vec{Ie}$ correspondantes aux points α de contour Λ ($\alpha \in \Lambda$), un système linéaire est défini :

$$\vec{Ie} = [\mathbf{A}]\vec{C} + \vec{B} \ \ avec\ : \vec{C} = ([\mathbf{A}]^T[\mathbf{A}])^{-1}[\mathbf{A}]^T\vec{Ie} - ([\mathbf{A}]^T[\mathbf{A}])^{-1}[\mathbf{A}]^T\vec{B} \tag{5.18}$$

Ce système est contraint car dans le cas d'un contour fermé $[\mathbf{A}]^T\vec{B} = 0$. La résolution du système $\vec{C} = ([\mathbf{A}]^T[\mathbf{A}])^{-1}[\mathbf{A}]^T\vec{Ie}$ permet d'obtenir une estimation de l'angle azimut :

$$\hat{\tau} = \arctan\left(\frac{C(2)}{C(1)}\right) \pm \pi \tag{5.19}$$

Les auteurs préconisent l'utilisation d'un algorithme de segmentation tel que le détecteur de passage à zéro du Laplacien *"zero-crossing edge detector"* pour extraire les contours de l'objet dans l'image. Pour l'estimation de l'angle zénithal, ils utilisent la même méthode que Lee et Rosenfeld et changent uniquement la fonction monolithique décroissante d'estimation (figure 5.16) par un polynôme :

$$\begin{aligned} \frac{\mathbb{E}(Ie)}{\sqrt{\mathbb{E}(Ie^2)}} &= f_3(\sigma) \\ &= 0.5577 + 0.6240\cos\sigma + 0.1882\cos^2\sigma - 0.6514\cos^3\sigma \\ &\quad -0.5350\cos^4\sigma + 0.9282\cos^5\sigma + 0.3476\cos^6\sigma - 0.4984\cos^7\sigma \end{aligned} \tag{5.20}$$

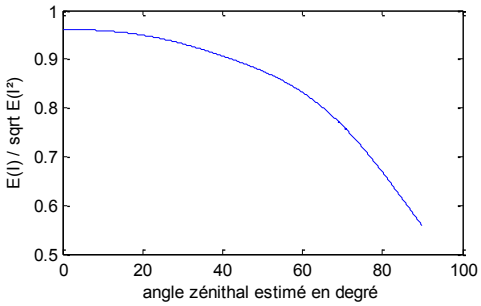


Figure 5.16 – Zheng *et* Chellappa, *fonction monolithique décroissante pour l'estimation de l'angle zénith.*

Ce polynôme est issue de l'intégration des intensités pour l'ensemble des angles zéniths et azimuts sur une surface sphérique suivant le modèle de Lambert. Son utilisation permet de prendre en compte l'ombre propre des sphères surtout lorsque l'angle zénith est proche de 90^o.

La méthode proposée par Chojnacki et Brooks [CBG94] est appelée *"disk method"*. Elle remet en question l'estimateur de Pentland [Pen82]. Les auteurs démontrent que l'ensemble de l'image d'une demi-sphère ne permet pas l'estimation correcte des angles de la source lumineuse à cause des ombres. Ils proposent de masquer une portion de la demi-sphère par un disque afin d'utiliser uniquement les points non ombrés de la surface. Les travaux de Lee

et Rosenfeld [LR85] sont réutilisés, l'estimation de l'angle azimut est obtenue à partir de deux directions de dérivation des intensités de l'image.

$$\hat{\tau} = \arctan\left(\frac{\mathbb{E}^{Dar}(\mathrm{d}Ie_y)}{\mathbb{E}^{Dar}(\mathrm{d}Ie_x)}\right) \pm \pi \tag{5.21}$$

où $\mathbb{E}^{Dar}(\mathrm{d}Ie_x)$ et $\mathbb{E}^{Dar}(\mathrm{d}Ie_y)$ sont les moyennes des dérivées premières des intensités de l'image suivant les directions respectives x et y pour les points Dar appartenant à un disque de rayon ar et de centre identique au centre de la demi-sphère.

La formulation de l'estimateur d'angle zénith est légèrement modifiée :

$$\begin{aligned} \hat{\sigma} &= \arccos\left(1 - \frac{fct(a)\left[\mathbb{E}^{Dar}(\mathrm{d}Ie_x)^2 + \mathbb{E}^{Dar}(\mathrm{d}Ie_y)^2\right]}{bb^2}\right)^{\frac{1}{2}} \pm \pi \\ bb &= \sqrt{\mathbb{E}^{Dar}(\mathrm{d}Ie_{\hat{\tau}}^2) - \mathbb{E}^{Dar}(\mathrm{d}Ie_{\hat{\tau}})^2} \\ fct(a) &= -\frac{1}{2} - \frac{1}{2a^2}\ln(1-a^2) \end{aligned} \tag{5.22}$$

5.2.1.2 Méthodes basées sur l'analyse des ombres

Les ombres propres à la surface d'une sphère (figure 5.17) répondent à certaines propriétés géométriques et photométriques. Les normales à la surface aux points limites entre l'éclairage direct et l'ombre propre sont perpendiculaires à la direction de la source lumineuse. Dans le cas d'un albédo identique sur l'ensemble de la surface, les points limites possèdent tous la même intensité.

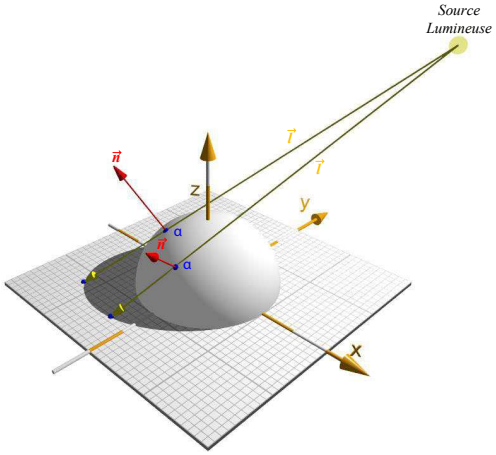


Figure 5.17 – *Schéma des propriétés géométriques aux points α entre les directions de la source lumineuse $\vec{l}$ et les normales $\vec{n}$ à la surface.*

Pour tous points α_j $(\alpha_{xj}, \alpha_{yj}, \alpha_{zj})$ appartenant à la limite Λ entre la surface directement éclairée par une unique source lumineuse et l'ombre propre , l'intensité Ie_{α_j} est nulle ou égale à l'intensité de la lumière ambiante $Ie_{ambient}$:

$$Ie_{\alpha_j} = [0||Ie_{ambient}]\,\forall(\alpha_j \in \Lambda) \tag{5.23}$$

car le modèle d'éclairage Lambertien est défini par :

$$Ie_{\alpha_j} = \rho Ie_0(\vec{n}_{\alpha_j}.\vec{l}) + [0||Ie_{ambient}] \tag{5.24}$$

Or le produit scalaire entre la direction $\vec{l}$ de la source et la normale à la surface $\vec{n}$ est nul :

$$(\vec{n}_{\alpha_j}.\vec{l}) = \vec{n}_{\alpha_j} \perp \vec{l} = 0 \ \ \forall \ (\alpha_j \in \Lambda) \tag{5.25}$$

Utiliser une demi-sphère de géométrie parfaitement connue permet de construire un système linéaire de la forme $[\mathbf{A}]\vec{b} = 0$. La résolution de ce système par la méthode des moindres carrés offre une estimation de la direction $\vec{l}$ de la source lumineuse :

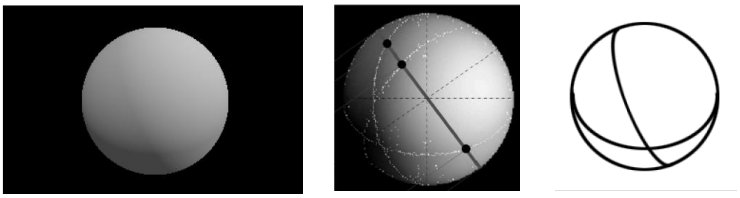
$$\begin{bmatrix} n_{x\alpha_1} & n_{y\alpha_1} & n_{z\alpha_1} \\ \vdots & \vdots & \vdots \\ n_{x\alpha_j} & n_{y\alpha_j} & n_{z\alpha_j} \end{bmatrix} \times \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix} = 0 \tag{5.26}$$

La méthode originale est donnée par Zhang et Yang dans [ZY01], elle permet d'estimer la direction de plusieurs sources lumineuses ainsi que leurs intensités à partir des points limites entre les zones directement éclairées et les ombres propres. Le principe reprend les propriétés géométriques et photométriques des ombres propres. Ces propriétés sont données pour une unique source lumineuse mais dans le cas de plusieurs sources, l'intensité au point α_j n'est pas nulle. Elle dépend des différentes sources :

$$Ie_{\alpha_{js}} = \sum_{s}^{S} \rho Ie_0 s(\vec{n}_{\alpha_{js}}.\vec{l}_s) + [0||Ie_{ambient}] \tag{5.27}$$

Chaque source engendre une limite différente :

$$(\vec{n}_{\alpha_{js}}.\vec{l}_s) = \vec{n}_{\alpha_{js}} \perp \vec{l}_s = 0 \forall (\alpha_{js} \in \Omega_s) \tag{5.28}$$



(a) (b) (c)

Figure 5.18 – *Processus de détection des ombres propres par Zhang et Yang, (a) exemple d'une sphère éclairée par deux sources lumineuses, (b) détection des "critical points" à la surface de la sphère, (c) "cutoff curves".*

La méthode consiste à déterminer tous les points α_{js} appelés *"critical points"* à la surface de la demi-sphère (figure 5.18a) susceptibles d'être des limites. La deuxième phase (figure 5.18b) consiste à regrouper dans un ensemble *"cutoff curves"* les points jugés comme critique en fonction de leur possible appartenance à l'une ou l'autre des sources lumineuses. La dernière étape consiste à résoudre un système linéaire par la méthode des moindres carrés pour estimer la direction et l'intensité de chacune des sources lumineuses.

L'approche proposée par Wang et Samaras [WS02] est similaire à [ZY01]. La surface utilisée est non plus une sphère mais une forme dont la géométrie est arbitraire et inconnue. Afin

de pouvoir utiliser la même approche pour déterminer la direction des sources lumineuses, la forme inconnue est plaquée sur une sphère. Toujours sur le même principe, Wei propose une amélioration [Wei03]. La principale différence réside dans la méthode utilisée pour déterminer les points critiques entre la partie directement éclairée de la sphère et l'ombre propre. D'après l'auteur, la méthode proposée est plus robuste mais plus coûteuse en temps de calcul.

5.2.1.3 Méthodes basées sur l'analyse des points spéculaires

Ces méthodes délaissent complètement le modèle de Lambert au profit des lois de la réflexion optique de Smell-Descartes. Debevec dans [Debo8] utilise les points spéculaires comme indicateurs de la position des sources lumineuses. Le principe consiste à déterminer l'emplacement des points spéculaires sur l'image d'une sphère miroir, puis d'appliquer les principes de la réflexion optique avec pour connaissances la topologie exacte de la scène (position camera et sphère, rayon de la sphère).

D'après les lois de Smell-Descartes pour la réflexion (Chapitre 2 : 2.3.4 La réflexion), l'angle θ_i formé entre la normale $\vec{n}$ à la surface et la direction d'incidence $\vec{l}$ est égale en valeur absolue à l'angle θ_r formé entre $\vec{n}$ et la direction de réflexion $\vec{r}$:

$$\begin{aligned} \|\theta_i\| &= \|\theta_r\| \\ \vec{l} &= -\vec{r} + 2\cos(\theta_r)\vec{r} \\ \vec{l} &= 2(\vec{n}.\vec{r})\vec{n} - \vec{r} \end{aligned} \tag{5.29}$$

L'utilisation du modèle de Phong permet de détecter les points spéculaires dans l'image car la valeur numérique M de chaque pixel (u, v) est définie comme la somme de la composante ambiante $Mamb(u, v)$, la composante diffuse $Mdiff(u, v)$ et la composante spéculaire $Mspec(u, v)$ soit :

$$M(u, v) = Mamb(u, v) + Mdiff(u, v) + Mspec(u, v) \tag{5.30}$$

Le ou les pixels spéculaires sont ceux ayant une valeur maximale, car la composante spéculaire est maximale lorsque la direction de l'observateur $\vec{v}$ est identique à la direction de réflexion $\vec{r}$. Pour estimer la direction $\vec{l}$ de la source lumineuse, Debevec utilise une sphère. Cette géométrie particulière permet la connaissance des normales en chaque point de la surface. Cette estimation repose sur deux hypothèses qui permettent de réduire la complexité des calculs. La première hypothèse concerne le modèle de projection de la caméra, il est considéré comme orthographique. Le vecteur $\vec{v}$ est défini en tout point par $[0\ 0\ -1]$. La seconde hypothèse considère que la source lumineuse est placée dans la direction $\vec{l}$ mais à une distance infinie. Ainsi la direction d'éclairage est approximativement identique en tous points de la scène.

L'estimation (figure 5.19) de $\vec{l}$ revient à :

- déterminer le rayon a_r et le centre $\alpha' = [\alpha'_u, \alpha'_v]$ de la sphère dans l'image.
- déterminer la position du point spéculaire $\alpha = [\alpha_u, \alpha_v]$ dans l'image.
- transposer les indexes en coordonnées réelles grâce à la matrice de projection $[\mathbf{K}]$.
- calculer la normale $\vec{n}$ au point spéculaire.
- estimer $\vec{l}$ puis calculer l'angle azimut et zénith.

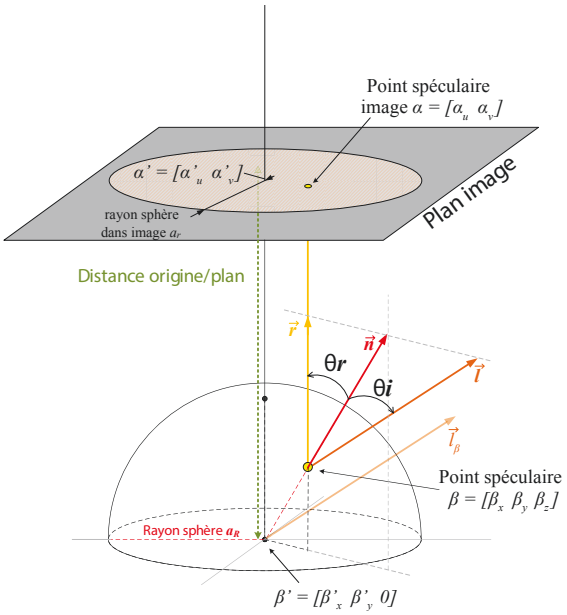


FIGURE 5.19 – *Projection point spéculaire avec un modèle orthographique.*

Hyphotèse 1 : $\vec{r} = [0\ 0\ -1]$

Hyphotèse 2 : $\vec{l}_\beta = \vec{l}$

Position centre hémisphère :

$\beta' = [\beta'_x\ \beta'_y\ 0] = \alpha'[\mathbf{K}]^{-1} = [\alpha'_u\ \alpha'_v][\mathbf{K}]^{-1}$

Position point spéculaire :

$\beta = [\beta_x\ \beta_y\ 0] = \alpha[\mathbf{K}]^{-1} = [\alpha_u\ \alpha_v][\mathbf{K}]^{-1}$

avec : $[\mathbf{K}]$ *la matrice de projection.*

Normale :

$$\vec{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} (\beta_x - \beta'_x)/a_R \\ (\beta_y - \beta'_y)/a_R \\ \sqrt{1 - n_x^2 - n_y^2} \end{bmatrix}$$

Donc la direction de la source :

$\vec{l}_\beta = \vec{l} = 2(\vec{n}.\vec{r})\vec{n} - \vec{r}$

Précédemment Powell dans [PSG01] a jugé ces hypothèses trop réductrices et a proposé une méthode similaire (figure 5.20) basée sur le modèle de projection perspective capable d'estimer la position d'une source ponctuelle.

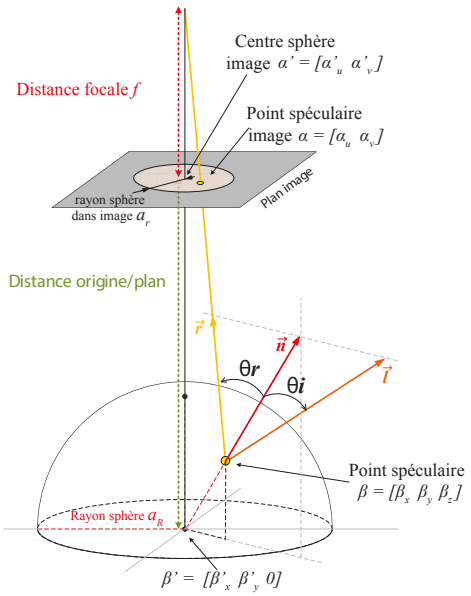


FIGURE 5.20 – *Projection point spéculaire avec un modèle perspectif.*

Projection centre hémisphère :

$$\beta' = \left[\beta'_x\ \beta'_y\ \beta'_z\right] = \left[\frac{\alpha'_u a_R}{a_r}\ \frac{\alpha'_v a_R}{a_r}\ \frac{f a_R}{a_r}\right]$$

Projection point spéculaire :

$$\alpha = [\alpha_u\ \alpha_v] = \left[\frac{f\beta_x}{\beta_z}\ \frac{f\beta_y}{\beta_z}\right]$$

$$(\beta_x - \beta'_x)^2 + (\beta_y - \beta'_y)^2 + (\beta_z - \beta'_x)^2 = {a_R}^2$$

$$donc : {\beta_z}^2 A + \beta_z B + C = 0$$

$$avec : A = \left(\frac{\alpha_u}{f}\right)^2 + \left(\frac{\alpha_v}{f}\right)^2 + 1$$

$$B = -2\left(\frac{\alpha_u - \beta'_x}{f} + \frac{\alpha_v - \beta'_y}{f} + \beta'_z\right)$$

$$C = {\beta'_x}^2 + {\beta'_y}^2 + {\beta'_z}^2 - {a_R}^2$$

Conserver la solution : $\beta_z < 0$

$\beta_x = \alpha_u \beta_z / f$ *et* $\beta_y = \alpha_v \beta_z / f$

Cette méthode est la première à déterminer la distance entre les sources lumineuses et la scène grâce à une triangulation des directions lumineuses estimées sur deux sphères miroirs. Dans le cas d'une projection perspective, ces estimations nécessitent la connaissance de la distance focale f, du rayon réel des sphères a_R et des matrices $[\mathbf{T}], [\mathbf{P}], [\mathbf{D}], [\mathbf{K}]$ de projection de la caméra. Les rayons, les centres des sphères et les positions des points spéculaires dans le repère $\mathcal{R}_{image}$ sont obtenus par segmentation. Puis le principe de projection permet de déterminer leurs positions dans le repère $\mathcal{R}_{scène}$.

Les directions $\vec{r}$, $\vec{n}$ et $\vec{l}$ sont ensuite déterminées par :

$$\vec{n} = \begin{bmatrix} (\beta'_x - \beta_x)/a_R \\ (\beta'_y - \beta_y)/a_R \\ (\beta'_z - \beta_z)/a_R \end{bmatrix} \qquad \vec{r} = \begin{bmatrix} (\beta_x - \alpha_u) \\ (\beta_y - \alpha_v) \\ -\beta_z \end{bmatrix} / \|\vec{r}\|$$

$$\vec{l} = 2(\vec{n}.\vec{r})\vec{n} - \vec{r} \tag{5.31}$$

Pour deux sphères, les directions $\vec{l}_1$ et $\vec{l}_2$ sont estimées (figure 5.21). Dans le cas idéal, leur intersection détermine les coordonnées $[l_x, l_y, l_z]$ de la source ponctuelle.

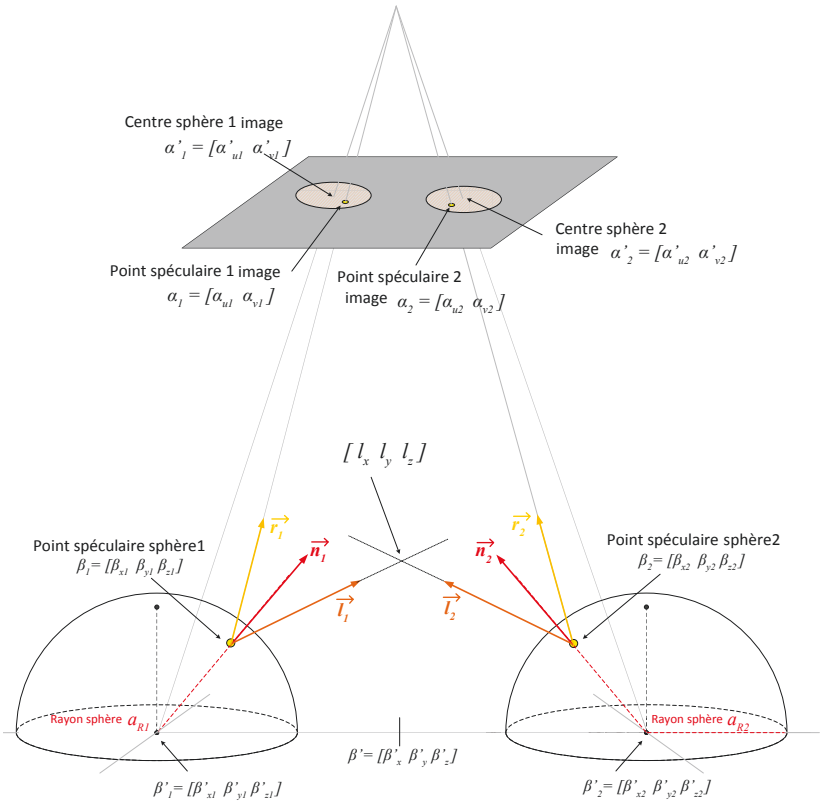


FIGURE 5.21 – *Principe de la triangulation selon la méthode de* POWELL.

En pratique, les deux directions $\vec{l}_1$ et $\vec{l}_2$ peuvent être gauches. Cette singularité est due principalement à l'échantillonnage spatial des images mais aussi aux multiples approximations faites lors du calcul numérique. Pour résoudre ce problème il suffit de déterminer le segment

où les deux directions sont les plus proches. Le milieu du segment est considéré comme la position de la source. Il est alors possible de déterminer l'angle azimut, l'angle zénith et la distance source/scène à partir de l'estimation de la position de la source et du centre de la scène correspondant à l'axe optique :

$$\hat{\delta}_{LC} = \sqrt{(\beta'_x - l_x)^2 + (\beta'_y - l_y)^2 + (\beta'_z - l_z)^2} \tag{5.32}$$

$$\hat{\sigma} = \arctan\left(\frac{l_x}{l_y}\right) \pm \pi \tag{5.33}$$

$$\hat{\tau} = \arctan\left(\frac{\sqrt{(\beta'_x - l_x)^2 + (\beta'_y - l_y)^2}}{l_z}\right) \tag{5.34}$$

En supplément, Powell propose d'estimer la distance focale (figure 5.22). Il utilise une troisième sphère de rayon différent et placée à une profondeur β'_{z2} différente de β'_{z1}. Les équations de projection permettent une estimation de la distance focale à partir des rayons réels des sphères (a_{R1}, a_{R2}), de la distance entre les centres réels des sphères δ_{CS}, des rayons des sphères dans l'image (a_{r1}, a_{r2}) et des centres des sphères dans l'image ($\alpha'_1 = [\alpha'_{u1}\ \alpha'_{v1}]$, $\alpha'_2 = [\alpha'_{u2}\ \alpha'_{v2}]$) :

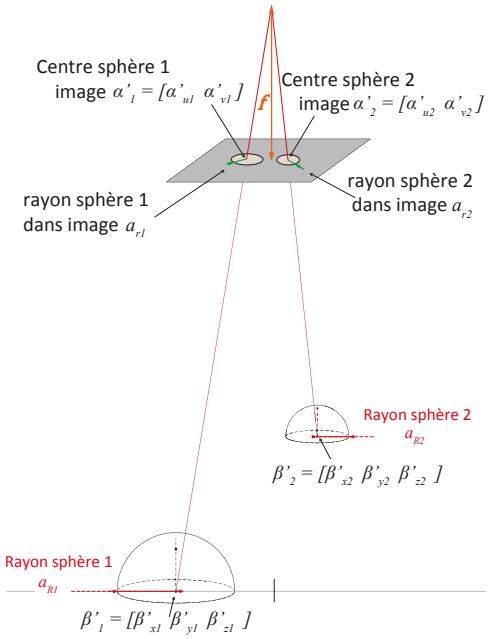


Figure 5.22 – *Estimation de la distance focale selon la méthode de* Powell.

$$\alpha'_1 = [\alpha'_{u1}\ \alpha'_{v1}] = \left[\frac{f\ \beta'_{x1}}{\beta'_{z1}}\ \frac{f\ \beta'_{y1}}{\beta'_{z1}}\right]$$

$$\alpha'_2 = [\alpha'_{u2}\ \alpha'_{v2}] = \left[\frac{f\ \beta'_{x2}}{\beta'_{z2}}\ \frac{f\ \beta'_{y2}}{\beta'_{z2}}\right]$$

$$a_{r1} = \frac{f\ a_{R1}}{\beta'_{z1}} \quad et \quad a_{r2} = \frac{f\ a_{R2}}{\beta'_{z2}}$$

Donc :

$$\beta'_1 = [\beta'_{x1}\ \beta'_{y1}\ \beta'_{z1}] = \left[\frac{\alpha'_{u1}\ a_{R1}}{a_{r1}}\ \frac{\alpha'_{v1}\ a_{R1}}{a_{r1}}\ \beta'_{z1}\right]$$

$$\beta'_2 = [\beta'_{x2}\ \beta'_{y2}\ \beta'_{z2}] = \left[\frac{\alpha'_{u2}\ a_{R2}}{a_{r2}}\ \frac{\alpha'_{v2}\ a_{R2}}{a_{r2}}\ \beta'_{z2}\right]$$

$$\beta'_{z1} = \frac{f\ a_{R1}}{a_{r1}} \quad et \quad \beta'_{z2} = \frac{f\ a_{R2}}{a_{r2}}$$

De plus :

$$(\beta'_{x1} - \beta'_{x2})^2 + (\beta'_{y1} - \beta'_{y2})^2 + (\beta'_{z1} - \beta'_{z2})^2 = {\delta_{CS}}^2$$

Après remplacement :

$$f = \sqrt{\frac{{\delta_{CS}}^2 - (\beta'_{x1} - \beta'_{x2})^2 - (\beta'_{y1} - \beta'_{y2})^2}{\left(\frac{a_{R1}}{a_{r1}}\right)^2 - \left(\frac{2\ a_{R1}\ a_{R2}}{a_{r1}\ a_{r2}}\right) + \left(\frac{a_{R2}}{a_{r2}}\right)^2}}$$

5.2.1.4 Bilan de l'étude comparative

D'après les résultats de l'étude comparative (Annexe A.4), la méthode proposée par Lee et Rosenfeld offre les meilleures estimations parmi les méthodes basées sur l'analyse des

intensités. Néanmoins, pour certaines configurations d'éclairage, le biais d'estimation est important surtout pour l'estimation des angles zéniths. A l'inverse, la méthode basée sur l'analyse des ombres proposée par Zhang et Yang offre de très bonnes performances d'estimation des angles zéniths, mais un fort biais d'estimation des angles azimuts. Ces deux méthodes pourraient être complémentaires, néanmoins elles souffrent des mêmes défauts que la SP. En effet, leur application requiert les valeurs d'intensité émises par la scène. Cette contrainte est approchée grâce aux valeurs numériques capturées et l'hypothèse $\mathcal{H}4$. Or comme nous l'abordons Chapitre 6, cette approximation est imparfaite et peut générer des erreurs importantes qui se répercutent sur l'estimation des angles d'éclairage.

Les méthodes basées sur l'analyse des points spéculaires sont nettement moins dépendantes des valeurs d'intensité émises par la scène. Car l'estimation des angles est uniquement basée sur la géométrie optique. De plus, la détection des points spéculaires et l'extraction des sphères sont facilement réalisables à partir des images. Comme le démontrent les résultats de simulation obtenus avec la méthode proposée par Powel, l'erreur d'estimation angulaire moyenne est faible et régulière pour les deux angles (azimut et zénith). Cette méthode possède aussi l'avantage d'estimer la position des sources ponctuelles, contrairement aux autres méthodes qui ne peuvent estimer que des sources directionnelles. Même si la distance estimée entre la source lumineuse et la scène est fortement biaisée, le principe présente un grand intérêt pour la correction de la non-uniformité de l'éclairage.

5.2.2 Méthodes de correction de la non-uniformité des sources

Les solutions généralement proposées [PCF05, SSSF10] pour résoudre le problème de non-uniformité de l'éclairage consistent à effectuer une compensation préalable des images à partir d'une calibration de la distribution de l'éclairement. Cette calibration (figure 5.23) est obtenue par l'analyse de l'atténuation des valeurs M acquises à partir d'une mire blanche et plane pour chaque source lumineuse. Elle permet d'extraire des cartes d'éclairement (figure 5.24) utilisées ensuite pour effectuer une compensation des images après le processus d'acquisition.

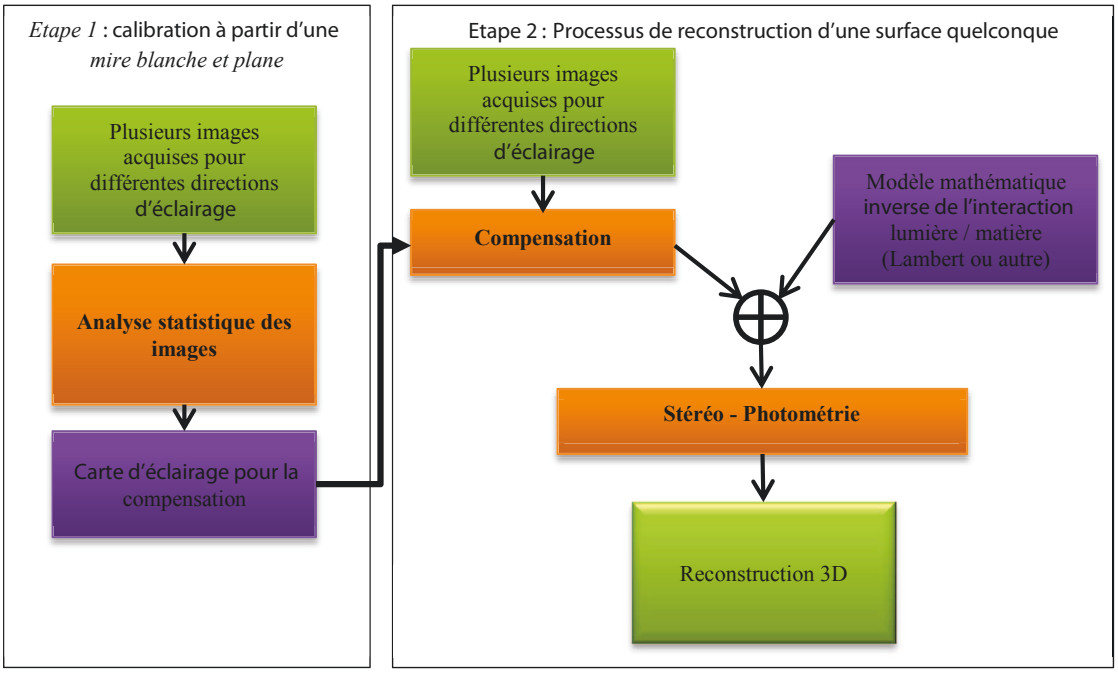


Figure 5.23 – *Diagramme du principe de prétraitement de la SP par l'analyse statistique d'une surface dont la géométrie et l'albédo sont uniformes, aussi appelé méthode de compensation.*

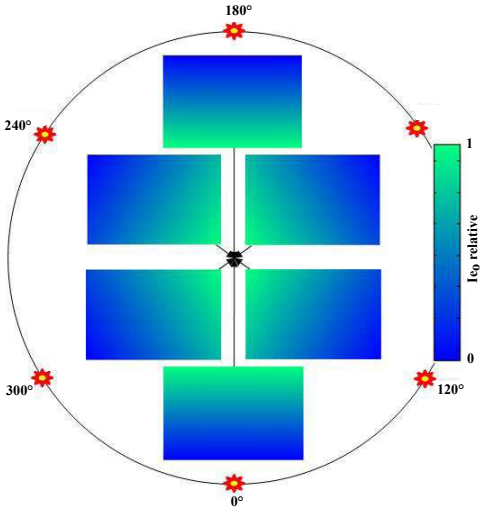


FIGURE 5.24 – *Exemple de carte d'éclairement obtenue à partir d'une surface dont la géométrie et l'albédo sont uniformes.*

Mais ces solutions ne permettent pas une correction complète et active (figure 5.25) car elles ne prennent pas en compte les modifications des directions d'incidence de l'éclairage en chaque point de la surface lors de l'inversion du modèle de réflectance. De plus, l'application des méthodes de compensation n'est pas valable lorsque les surfaces analysées ont une élévation non négligeable par rapport à la distance entre la source lumineuse et la surface analysée car les cartes de compensation sont acquises à partir de surfaces planes.

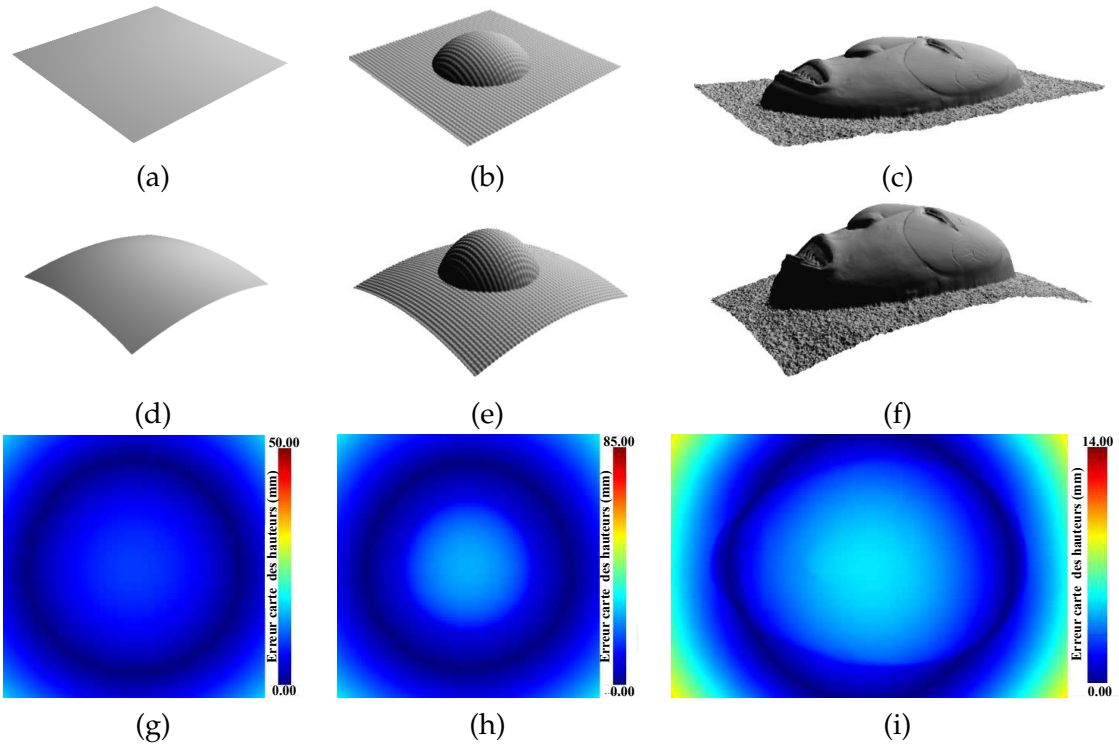


FIGURE 5.25 – *(a-c) surfaces réelles vérités terrains, (d-f) surfaces estimées par la méthode de compensation de stéréo-photométrie, (g-i) différences de hauteurs entre la carte estimée et la vérité terrain ou son approximation.*

5.3 Méthodes proposées

5.3.1 Calibration des positions des sources

Notre méthode de calibration de position des sources lumineuses est une variante de celle proposée par Powell. Elle utilise plusieurs demi-sphères identiques dont le matériau présente de fortes propriétés spéculaires afin d'utiliser le modèle de Phong pour déterminer la position des sources lumineuses par triangulation. Elle néglige l'estimation de la distance focale afin de simplifier les calculs car celle-ci est généralement fournie avec précision par les métadonnées des fichiers images ou par l'appareil d'acquisition. Par ailleurs, nous proposons d'utiliser plus de deux demi-sphères afin d'accroître les performances de la triangulation et notamment l'estimation des distances entre les sources et le centre du plan acquis. De plus, l'algorithme implémenté laisse le choix entre un modèle de projection perspective ou orthographique à facteur d'échelle afin de respecter les caractéristiques du système d'acquisition.

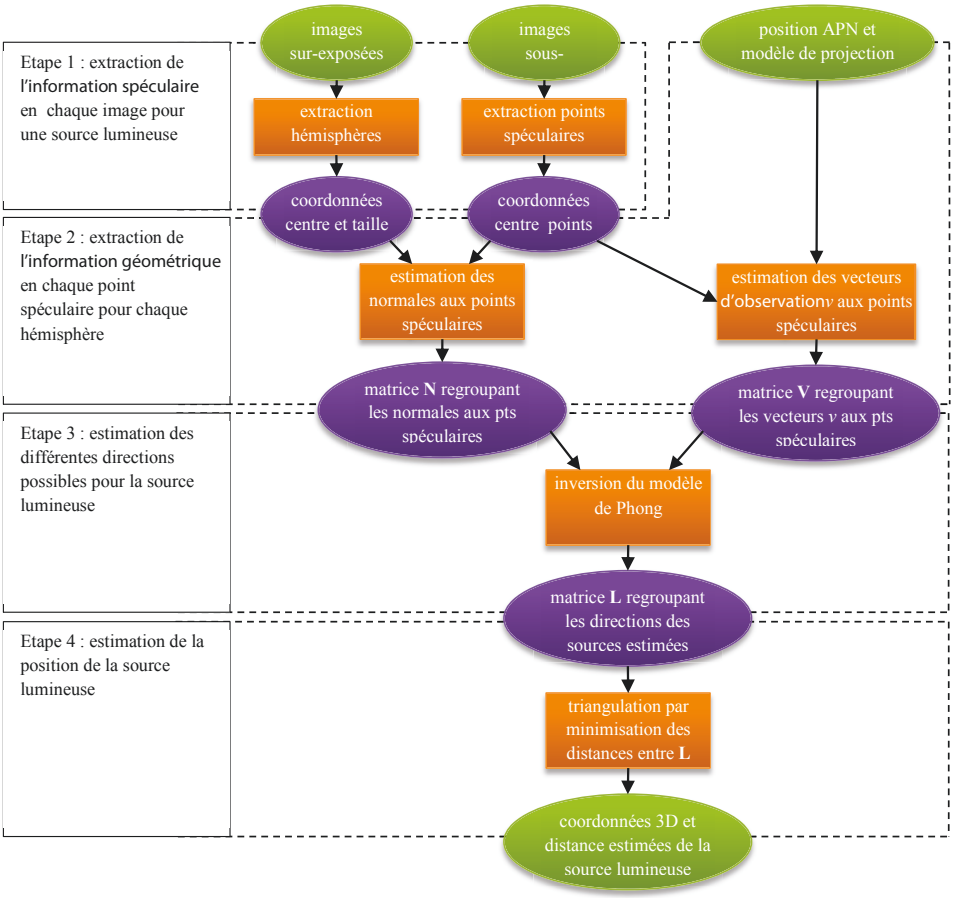
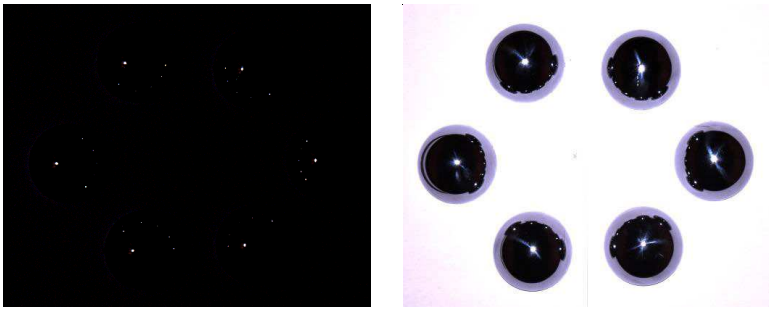


Figure 5.26 – *Algorithme proposé pour déterminer la position d'une source lumineuse par triangulation à partir d'une image de demi-sphères spéculaires.*

La première étape (figure 5.26) consiste à extraire les caractéristiques des demi-sphères (positions et tailles) ainsi que les coordonnées des points spéculaires sur chacune d'entre elles pour une unique source lumineuse. Cette extraction est effectuée à partir d'une segmentation obtenue par filtrage, puis l'utilisation d'un détecteur circulaire de Hough [Hou62]. Nous préconisons d'utiliser deux images d'un même point de vue pour chaque source lumineuse. En effet, l'utilisation d'une image sous exposée (figure 5.27a) limite l'étendue des tâches spéculaires et facilite la détermination du centre de celle-ci. En revanche l'extraction de la taille

et du centre des demi-sphères est impossible d'où l'utilisation d'une seconde image (figure 5.27b) sur exposée.



(a) (b)

FIGURE 5.27 – *Images des demi-sphères spéculaires capturée d'un point de vue fixe pour une unique source lumineuses, (a) sous exposée afin de facilité l'extraction des point spéculaires, (b) sur exposée afin de faciliter l'extraction des caractéristique géométriques des demi-sphères.*

Après extraction des positions, la seconde étape consiste à extraire l'information géométrique en chaque point spéculaire détecté. Cela consiste à estimer les normales $\vec{n}$ au centre des tâches spéculaire et les regrouper dans une matrice. Cette estimation est facilement obtenue à partir de l'équation de la sphère, de la connaissance des tailles réelles des demi-sphères utilisées et des caractéristiques extraites à l'étape précédente. L'autre partie de cette étape consiste à déterminer les vecteurs d'observation $\vec{v}$ en chaque point spéculaire et de les regrouper dans une matrice. Ces vecteurs sont obtenus par le principe de projection, leur détermination nécessite la connaissance des caractéristiques du système d'acquisition ainsi que le modèle de projection à utiliser.

Au cours de la troisième étape, les matrices de normales et de vecteurs d'observation permettent l'inversion du modèle de PHONG et l'obtention d'une matrice **L** regroupant l'ensemble des directions estimées de la source lumineuse pour chaque point spéculaire.

Enfin la dernière étape consiste à effectuer une triangulation entre les directions d'éclairage estimées. Pour cela, nous cherchons un point dans l'espace tridimensionnel qui minimise les distances entre les différents vecteurs $\vec{l}$. Cette quatrième étape permet de valider le nombre de demi-sphères à employer. Ce nombre doit être impair afin de permettre l'élimination d'un des vecteurs $\vec{l}$ si la direction estimée est aberrante ou ne peut converger vers une distance inter vecteur minimale. De plus, le nombre de demi-sphères employées doit être supérieur à trois sinon aucune amélioration n'est possible par rapport à la méthode de POWELL. A l'inverse, le nombre de demi-sphères ne doit pas être excessif car d'un point de vue statistique, il est fort probable que les vecteurs $\vec{l}$ se croissent en plusieurs positions, ce qui rend la triangulation impossible. Dans notre étude nous préconisons l'utilisation de cinq demi-sphères, ce qui permet l'élimination de deux directions $\vec{l}$ aberrantes au maximum, tout en conservant plus de deux directions possibles pour la triangulation.

Sur cas réel, nous utilisons six demi-sphères mais l'une d'entre elle est directement éliminée lors de la première étape. Cette élimination lors de l'extraction des points spéculaires permet de limiter les problèmes d'ombres portées car lorsque la position de la source est rasante, il arrive qu'une des tâches spéculaires soit cachée par une ombre (figure 5.28).

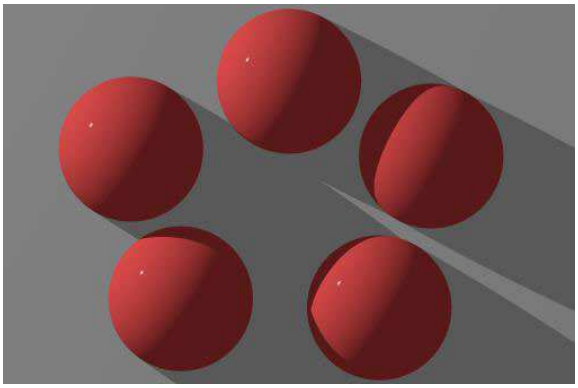


FIGURE 5.28 – *Pour une direction d'éclairage rasante, la tâche spéculaire de l'une des demi-sphères est occultée par l'ombre.*

Afin d'évaluer les performances de notre algorithme de calibration, nous simulons une scène par le logiciel POV-Ray. Cette scène (figure 5.28) représente cinq demi-sphères spéculaires de 75 *mm* de diamètre disposées sur un plan et éclairées suivant les schémas présentés en Annexe A.4.

La synthèse des résultats d'estimation angulaire (angles azimuts et zéniths) est donnée tableau 5.3 et figures 5.29, 5.30, 5.31, 5.32. L'estimation des distances entre la source et le centre du plan acquis est donné tableau 5.4 et figure 5.33. Après analyse, nous pouvons constater que les performances de calibration angulaire sont quasi identiques à la méthode de POWELL. Néanmoins, la triangulation multiple apporte de légères améliorations. Mais globalement, notre algorithme souffre des mêmes défauts. Certaines configurations telles que les angles zéniths extrêmes (proches 0^o et de 90^o) entraînent de fortes erreurs d'estimation. De plus, l'estimation reste dépendante de la résolution des images ainsi que de la capacité des algorithmes de segmentation et de détection. Par contre, l'avantage d'une triangulation multiple se trouve dans l'estimation des distances car nous pouvons constater de nettes améliorations des résultats. Les erreurs d'estimation des distances ont considérablement diminué par rapport à la méthode de POWELL et ce quelque soit la configuration d'éclairage.

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2	Schema3
ε azimut °	-0.078	0.175	0.176	0.028	-0.006	-0.272	0.124	-0.971
etype	0.662	1.178	0.529	1.644	1.059	0.485	0.541	0.826
ε zenith °	-0.493	-0.0002	0.073	-0.011	-0.093	-0.278	0.211	-0.297
etype	0.187	0.646	0.517	0.814	1.089	0.532	0.693	0.013

TABLE 5.3 – *Algorithme proposé, erreur moyenne d'estimation des angles azimuts, zéniths et l'écart-type pour chaque série d'image test.*

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2	Schema3
ε distance cm	2.224	0.796	0.239	4.357	4.494	-4.066	-3.59	-4.947
etype	1.748	4.968	5.478	2.774	4.186	3.621	2.93	14.166

TABLE 5.4 – *Algorithme proposé, erreur moyenne d'estimation des distances et l'écart-type pour chaque série d'images test.*

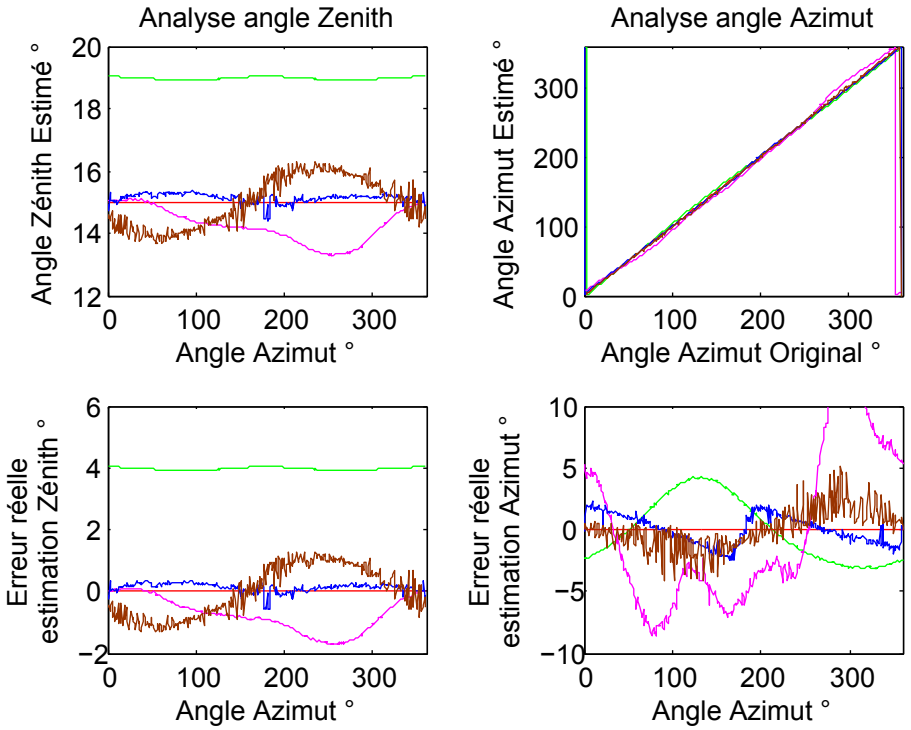


FIGURE 5.29 – *Comparaison entre les estimations idéales (rouge), Powell (bleu),* ZHANG *et* YANG *(magenta),* LEE *et* ROSENFELD *(vert) et notre algorithme (marron) pour azimut* (0 – 360°) *et zénith* 15°.

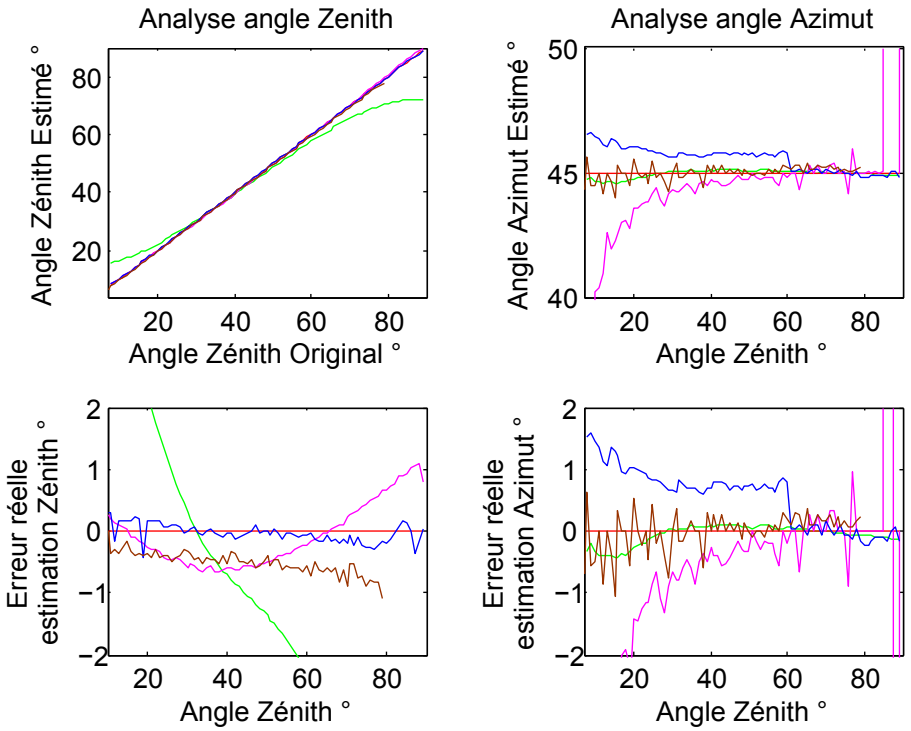


FIGURE 5.30 – *Comparaison entre les estimations idéales (rouge), Powell (bleu),* ZHANG *et* YANG *(magenta),* LEE *et* ROSENFELD *(vert) et notre algorithme (marron) pour azimut* 45° *et zénith* (4 – 89°).

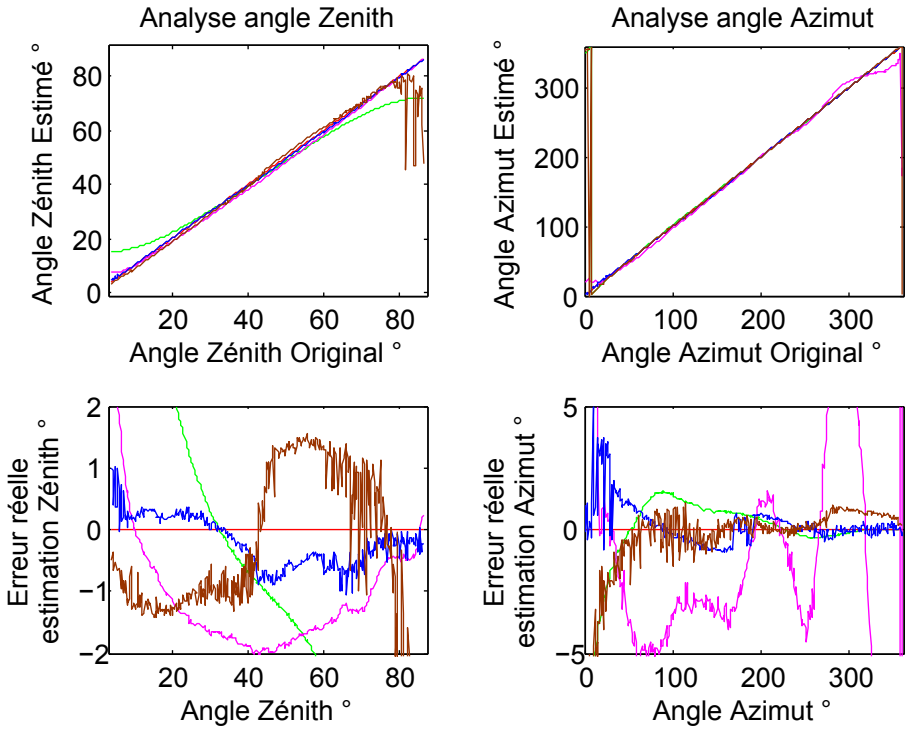


FIGURE 5.31 – *Comparaison entre les estimations idéales (rouge), Powell (bleu),* ZHANG *et* YANG *(magenta),* LEE *et* ROSENFELD *(vert) et notre algorithme (marron) pour éclairage schéma 1 (figure A.12).*

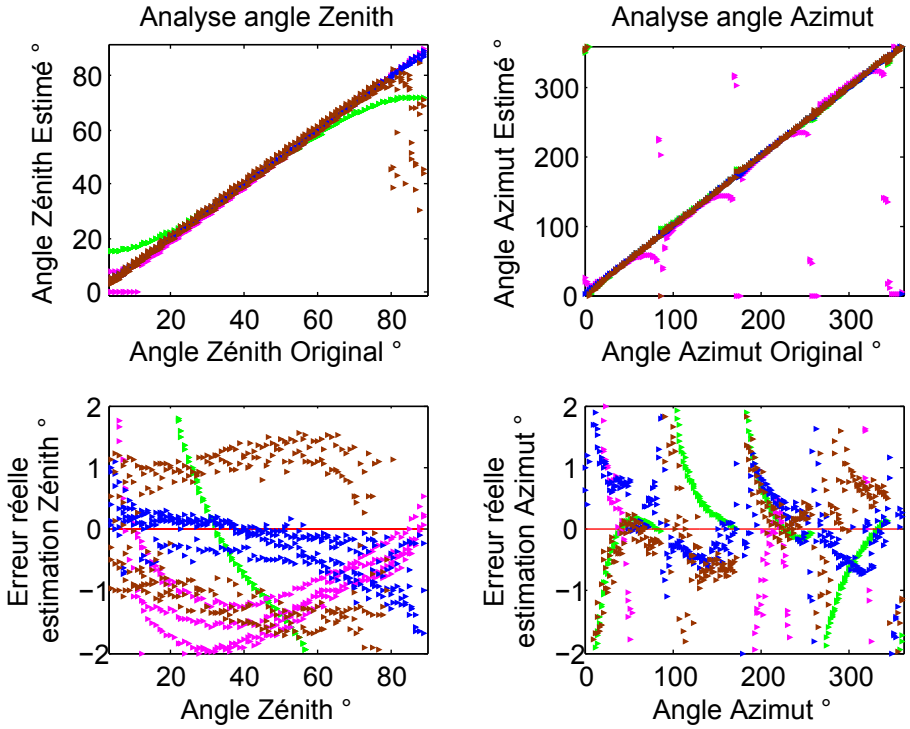


FIGURE 5.32 – *Comparaison entre les estimations idéales (rouge), Powell (bleu),* ZHANG *et* YANG *(magenta),* LEE *et* ROSENFELD *(vert) et notre algorithme (marron) pour éclairage schéma 2 (figure A.12).*

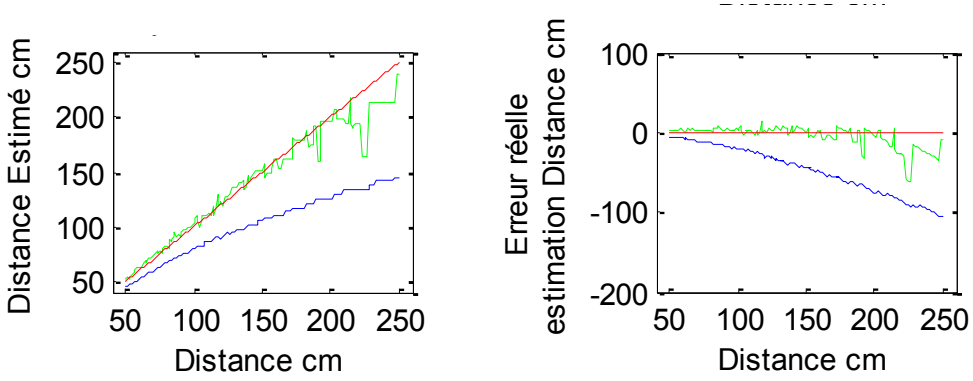


FIGURE 5.33 – *Comparaison entre l'idéal (rouge), Powell (bleu) et notre algorithme (vert) erreur d'estimation des distances pour une série d'image, la distance de la source varie et les angles restent fixes (azimut* 45^o*, zénith* 15^o*).*

En cas réel, nous utilisons notre méthode pour calibrer le système d'acquisition (NIKON_1 A.2.1) dont la configuration d'éclairage peut varier. Nous présentons (tableau 5.5) les résultats entre les estimations et des mesures manuelles pour la configuration d'éclairage initiale. Comme nous pouvons le constater, l'erreur d'estimation angulaire est très faible et l'erreur d'estimation de la distance est similaire aux simulations, mais n'excède pas 5 *cm*.

	Notre méthode			Mesures manuelles			Différence		
Lampe	azimut	zénith	distance	azimut	zénith	distance	azimut	zénith	distance
1	0,89	30,05	172,15	1,2	29,89	174,4	0,31	-0,16	2,25
2	60,02	30,25	178,92	59,9	30,2	175,3	-0,12	-0,05	-3,62
3	120,26	30,01	178,02	120,2	30,09	173,8	-0,06	0,08	-4,22
4	183,59	28,61	169,01	183,6	28,62	172,9	0,01	0,01	3,89
5	239,52	29,55	172,3	238,9	29,65	174,1	-0,62	0,1	1,8
6	300,59	31,46	178,23	300,2	31,54	173,8	-0,39	0,08	-4,43

TABLE 5.5 – *Comparaison entre une calibration par notre méthode d'estimation des sources des mesures manuelles pour le système d'acquisition NIKON_1.*

5.3.2 Correction de la non-uniformité de l'éclairage dans la SP

Dans le cas d'un système d'acquisition SP à sources lumineuses ponctuelles (exemple NIKON_2 A.2.2), l'éclairage des surfaces analysées est non uniforme et génère d'importantes distorsions dans les reconstructions 3d. Par conséquent, pour être correct le modèle Lambertien doit être réécrit en chaque point α de la surface en fonction de l'intensité Ie_{0s} émise par la source lumineuse s par rapport à la distance $\delta_{LC\alpha}$ et la direction $\vec{l}_{s\alpha}$:

$$Ie_\alpha = \frac{Ie_{0s}}{(\delta_{LC\alpha})^2}\rho\cos\theta_\alpha = \frac{Ie_{0s}}{(\delta_{LC\alpha})^2}\rho\left(\vec{n}_\alpha.\vec{l}_{s\alpha}\right) \tag{5.35}$$

La méthode proposée [BBK13b, BBK13c], de correction de la non uniformité de l'éclairage, est basée sur cette réécriture du modèle Lambertien. L'inversion se fait toujours à partir des valeurs M des pixels considérées selon $\mathcal{H}4$ comme parfaitement proportionnelles aux valeurs

Ie. De plus, les sources utilisées sont identiques ainsi l'intensité Ie_{0s} émise est constante quelle que soit la source s :

$$\vec{n}_\alpha = \begin{bmatrix} -p_\alpha / \sqrt{(p^2+q^2+1)} \\ -q_\alpha / \sqrt{(p^2+q^2+1)} \\ -1 / \sqrt{(p^2+q^2+1)} \end{bmatrix} = \frac{\left(([\mathbf{L}_\alpha]^T [\mathbf{L}_\alpha])^{-1} . [\mathbf{L}_\alpha]^T . \vec{M'}_\alpha \right)}{\left| ([\mathbf{L}_\alpha]^T \mathbf{L}_\alpha)^{-1} . \mathbf{L}_\alpha{}^T . \vec{M'}_\alpha \right|} \tag{5.36}$$

où $[\mathbf{L}_\alpha]$ est la matrice d'éclairage au point α :

$$[\mathbf{L}_\alpha] = \begin{bmatrix} \vec{l}_\alpha(1) \\ \vdots \\ \vec{l}_\alpha(S) \end{bmatrix} \; avec \; \vec{l}_\alpha(s) = \begin{bmatrix} l_x(s) - \alpha_x \\ l_y(s) - \alpha_y \\ l_z(s) - \alpha_z \end{bmatrix} / \left| \vec{l}_\alpha(s) \right| \tag{5.37}$$

et $\vec{M'}_\alpha$ un vecteur regroupant les valeurs des pixels pondérés par les carrés des distances entre le point α et les S sources lumineuses :

$$\vec{M'}_\alpha = \begin{bmatrix} M_\alpha(1) / \delta^2_{LC\alpha}(1) \\ \vdots \\ M_\alpha(S) / \delta^2_{LC\alpha}(S) \end{bmatrix} \tag{5.38}$$

avec :

$$\vec{\delta}_{LC\alpha}{}^2(s) = [l_x(s) - \alpha_x]^2 + [l_y(s) - \alpha_y]^2 + [l_z(s) - \alpha_z]^2 \tag{5.39}$$

Les positions $(l_x(s), l_y(s), l_z(s))$ des sources lumineuses $(s = 1 \dots S)$ sont parfaitement connues selon l'hypothèse $\mathcal{H}2$ ou obtenues par calibration selon la méthode proposée section 5.3.1 de ce chapitre. Les coordonnées réelles (α_x, α_y) représentent la position d'un point α de la surface suivant le plan (x, y). Ils sont déterminées à partir de l'indexation des pixels et des paramètres intra et extrinsèques de la caméra (équation 3.16 - chapitre 3) :

$$\begin{aligned} \alpha_x &= \left(\alpha_u - \tfrac{U}{2}\right) \times Taille\ pixel\ scène_{(u)} \\ \alpha_y &= \left(\alpha_v - \tfrac{V}{2}\right) \times Taille\ pixel\ scène_{(v)} \end{aligned} \tag{5.40}$$

avec :

α_u, α_v	:	l'indexation du pixel au point α.
U, V	:	la résolution de l'image.

Malheureusement, α_z est l'inconnue et correspond à l'élévation de chaque point suivant l'axe z, soit $\zeta\alpha$. Afin de résoudre cette ambiguïté, notre méthode est itérative (figure 5.34). Lors de la première itération, les matrices $\mathbf{L}_\alpha$ et les distances $\vec{\delta}_{LC\alpha}$ sont déterminées pour une surface initiale nulle $\alpha_z \ \forall \alpha = 0$. Cette première itération permet d'extraire l'information et d'obtenir une première estimation de la carte des hauteurs $\hat{\zeta}\alpha$. Ces élévations sont donc réutilisées lors de l'itération suivante afin de recalculer au plus juste les matrices $\mathbf{L}_\alpha$ et les distances $\vec{\delta}_{LC\alpha}$. Les itérations suivantes permettent une convergence de la carte des hauteurs $\hat{\zeta}$ vers les valeurs réelles ζ.

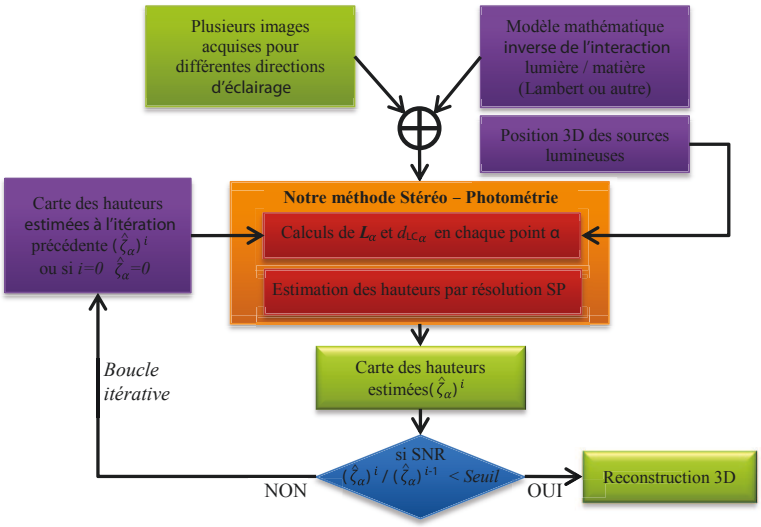


FIGURE 5.34 – *Diagramme du principe itératif de notre méthode de résolution de la SP pour des sources lumineuses ponctuelles non uniformes.*

Afin de prouver l'efficacité de notre méthode, nous la comparons à la méthode classique de SP sur des images de synthèse et sur cas réel. Les images de synthèse sont obtenues à partir d'un lancer de rayon (POV-Ray), la résolution est 800×800 *pixels* pour une surface de 100×100 *mm* représentant une demi-sphère à laquelle nous ajoutons une rugosité sinusoïdale de période 16 *pixels*. Notre système d'acquisition réel NIKON_2 (A.2.2) possède une configuration d'éclairage identique à la simulation et permet d'obtenir trois images réelles qui couvrent une zone de 68.9×103 *mm* pour une résolution de 1434×2160 *pixels*.

Les figures 5.35a et 5.36a représentent une coupe 2D des surfaces selon l'axe x. Elles permettent de comparer les déformations subies par les reconstructions lors de l'application de la méthode classique de SP ainsi que l'évolution l'affinage entre la première et la dernière itération de notre méthode. Que ce soit sur cas réel ou simulé, la méthode classique entraîne une déformation globale sous forme de dôme du profil. Cette déformation affecte principalement la macro mais aussi la micro géométrie de reconstruction. Elle est liée aux variations de distance et d'angle d'incidence en chaque point de la surface. Nous pouvons remarquer que notre méthode permet de corriger l'ensemble de ces défauts. Dès la première itération, la macro géométrie est corrigée. Comme le prouve l'évolution du SNR (tableau 5.6 et figures 5.35b-d, 5.36b), les itérations suivantes permettent un affinage pour obtenir la micro géométrie avec précision. Le nombre d'itérations dépend de la complexité de la surface (figure 5.35c et 5.36c) pour un seuil fixé arbitrairement à $110dB$. Au-delà, l'amélioration apportée par l'affinage n'est pas significative. La surface de synthèse nécessite 6 itérations et le cas réel 5 itérations.

Itération	1	2	3	4	5	6	Classic
Cas simulé: hémisphère rugueuse							
SNR $\hat{\zeta}/\zeta$	53.7	79.3	80.5	80.6	82.6	82.7	-2.19
SNR $\hat{\zeta}^i/\hat{\zeta}^{i-1}$	23.7	51.7	78.3	94.9	114.1		
Cas réel : masque archéologique africain							
SNR $\hat{\zeta}/\zeta$	27.7	30.9	31.0	31.2	31.6	31.6	3.51
SNR $\hat{\zeta}^i/\hat{\zeta}^{i-1}$	37.7	70.9	106.6	139.4	186.7		

TABLE 5.6 – *Evolution du SNR entre la vérité terrain et la reconstruction en fonction des itérations.*

Les résultats des reconstructions 3D selon notre méthode (figure 5.37d-f) sont visuellement

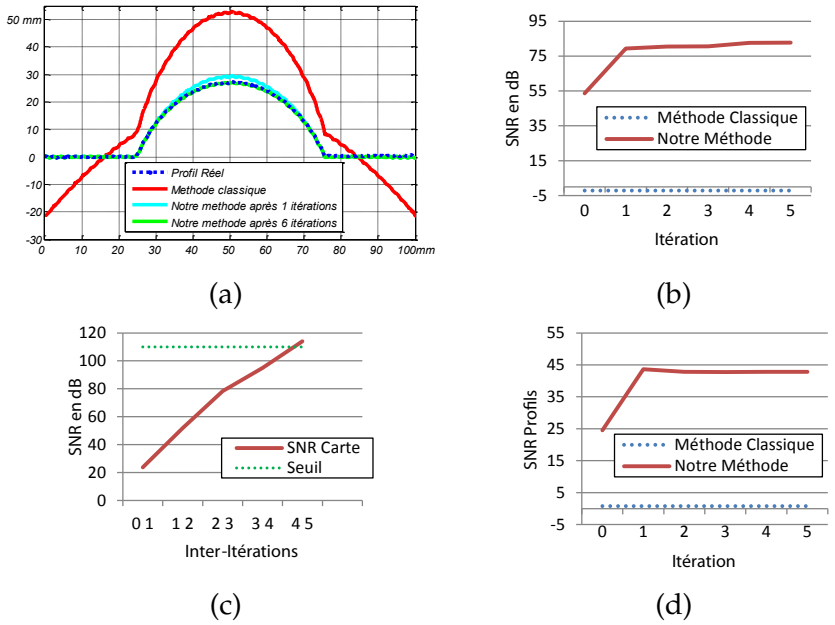


FIGURE 5.35 – *Résultats sur une surface de synthèse, (a) coupes de profil, (b) évolution du SNR pour le profil, (c) évolution du SNR de l'ensemble de la carte $\hat{\zeta}$ entre les itération de notre méthode, (d) évolution du SNR de l'ensemble de la carte $\hat{\zeta}$.*

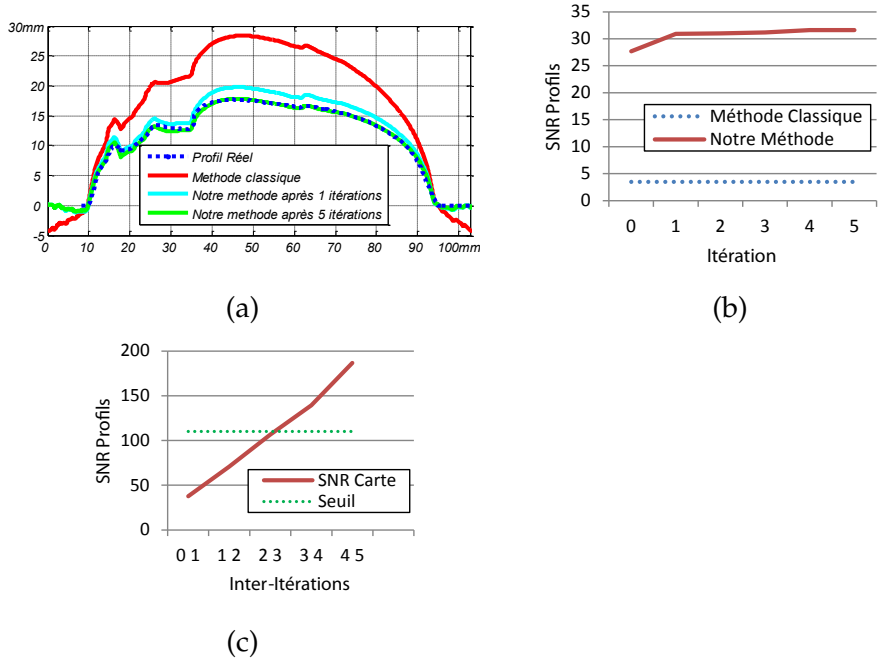


FIGURE 5.36 – *Résultats sur une surface réelle, (a) coupes de profil, (b) évolution du SNR pour le profil, (c) évolution du SNR de l'ensemble de la carte $\hat{\zeta}$ entre les itérations de notre méthode.*

identiques à la vérité terrain (figure 5.37a-c) comme le révèle les cartes d'erreurs sur les hauteurs estimées (figure 5.37g-i), calculées et affichées selon les mêmes échelles que pour la méthode SP classique (figure 5.14) ou la méthode SP par compensation (figure 5.25). Comparés aux méthodes existantes, nos résultats offrent une nette amélioration des reconstructions 3d grâce à la prise en compte de la géométrie des sources lumineuses, la géométrie de la surface estimée ainsi que le processus itératif.

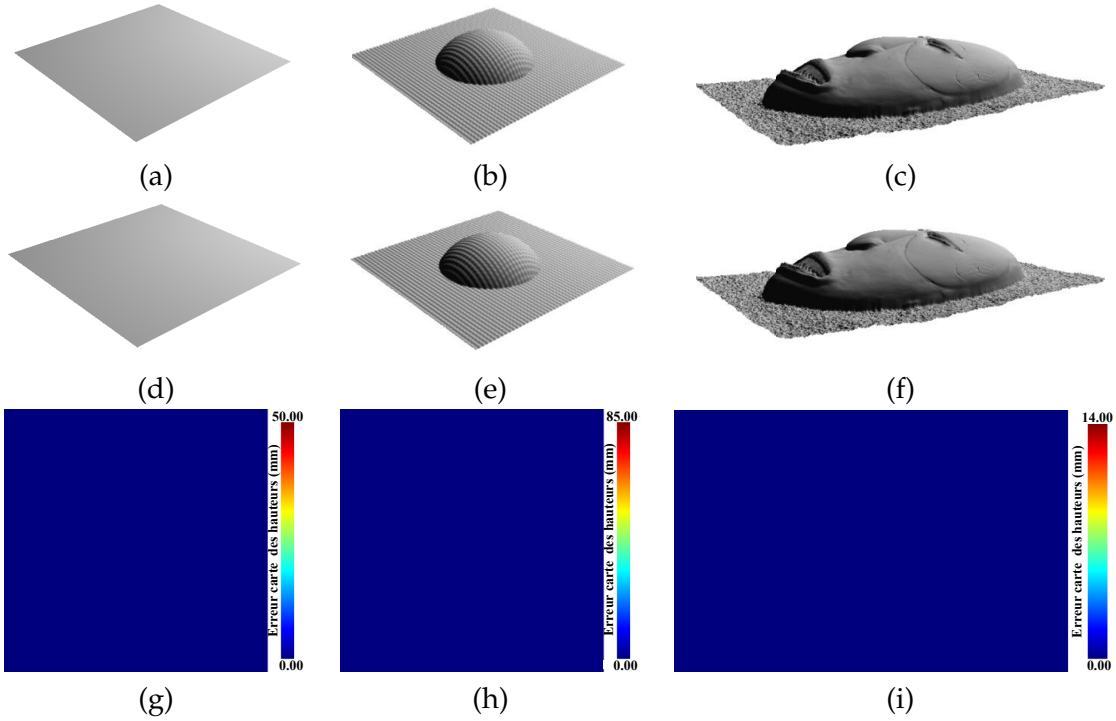


FIGURE 5.37 – *(a-c) surfaces réelles et vérités terrains, (d-f) surfaces estimées avec notre méthode de stéréo-photométrie, (g-i) différences de hauteurs entre les cartes estimées et la vérité terrain ou son approximation.*

Conclusion du chapitre

Ce chapitre met en évidence la nature et la configuration du système d'éclairage sur les résultats de reconstruction par stéréo-photométrie. Comme l'exprime l'hypothèse applicative $\mathcal{H}2$, la connaissance préalable des directions des sources lumineuses est primordiale dans le processus d'inversion d'un modèle de BRDF. De plus, selon $\mathcal{H}3$, l'utilisation de sources lumineuses directionnelles et uniformes est obligatoire. Néanmoins, dans la majeure partie des cas applicatifs réels, les sources lumineuses utilisées sont généralement ponctuelles et non uniformes. En effet, ce type de sources est plus facile à mettre en œuvre et permet de réduire la taille des systèmes d'acquisition. Mais la nature de l'éclairage produit ne respecte pas $\mathcal{H}3$ et entraîne des erreurs significatives dans les géométries reconstruites. L'étude menée sur la calibration des sources lumineuses montre que plusieurs méthodes ont été proposées. Mais dans le cas de sources ponctuelles, la plupart d'entre elles s'avèrent incomplètes ou pas suffisamment précises. Seules les méthodes basées sur l'analyse des points spéculaires offrent la possibilité d'estimer la position des sources lumineuses alors que les autres méthodes sont uniquement capables d'estimer les directions. Dans ce contexte, nous proposons une méthode de calibration basée sur l'analyse des points spéculaires à la surface de plusieurs

demi-sphères. A partir d'une triangulation, elle permet d'obtenir les positions précises et le respect de l'hypothèse $\mathcal{H}2$ dans le cas de sources lumineuses ponctuelles. Néanmoins, la nature de l'éclairage ne peut être en accord avec l'hypothèse $\mathcal{H}3$. Pour pallier à ce problème, nous proposons une seconde méthode qui intègre la non-uniformité de l'éclairage ainsi que la géométrie des surfaces dans un processus itératif d'inversion du modèle de BRDF. Chaque point de la surface est traité indépendamment grâce à la détermination des positions et des intensités reçues des sources lumineuses. Le processus itératif permet d'affiner l'estimation de la géométrie. Les résultats obtenus montrent l'efficacité et les apports de nos méthodes, les reconstructions 3d sont plus fidèles à la réalité et ce malgré une géométrie d'éclairages initialement incompatibles avec la stéréo-photométrie.

Chaine d'acquisition en SP 6

Sommaire

6.1 Les limites du capteur d'image dans la SP . . . 132
6.2 Les images hautes dynamiques . . . 133
6.2.1 État de l'art du principe HDR . . . 133
6.2.2 Images HDR, une solution en SP ? . . . 138
6.3 Une nouvelle méthode d'acquisition : images AIQ . . . 141
6.3.1 Le compromis Quantification / Dynamique . . . 142
6.4 Les images AIQ pour la SP . . . 144
6.4.1 Protocole d'acquisition AIQ . . . 146
6.4.2 Les images AIQ et les reconstructions 3D . . . 146
Conclusion . . . 151

Les méthodes de stéréo-photométrie supposent la connaissance des valeurs d'intensité Ie émises par les surfaces. Or, seules les valeurs M issues des appareils d'acquisition sont connues et ne sont qu'une estimation des valeurs Ie plus ou moins bruitées. Généralement, la relation entre M et Ie est supposée parfaite selon l'hypothèse $\mathcal{H}4$. Malheureusement, cette supposition est éloignée de la réalité car les valeurs numériques issues du capteur sont assujetties à divers traitements et erreurs dus au principe d'acquisition. Parmi ces phénomènes, nous pouvons citer les limitations des plages dynamiques d'acquisition, les pertes d'informations géométriques et radiométriques lors de l'étape de quantification, les non linéarités ou encore les perturbations dues aux bruits.

Au cours de ce chapitre, nous démontrons que l'hypothèse $\mathcal{H}4$ est trop réductrice si elle est appliquée directement. Et, dans le but d'inverser les modèles de réflectance par SP, nous proposons des solutions qui permettent d'estimer au plus juste les valeurs d'intensité Ie émises par une surface à partir des valeurs numériques M issues des appareils d'acquisition.

6.1 Les limites du capteur d'image dans la SP

Dans le contexte applicatif de la stéréo photométrie, les capteurs d'image sont supposés parfaits sous l'hypothèse $\mathcal{H}4$. En réalité, ces outils de mesures sont imparfaits et présentent des limitations géométriques et radiométriques. En effet, un capteur est un échantillonneur spatial d'une image géométriquement continue. Le découpage de l'image projetée en éléments matriciels (pixels) génère une perte d'information. Ce problème a été largement traité ces dernières décades [CKK+96, PPK03, vO06]. De plus, l'évolution technologique permet aujourd'hui un échantillonnage spatial à grande résolution [TN13]. En revanche, les limitations radiométriques sont peu abordées lors de l'application de la stéréo photométrie. Or, les imperfections des mesures sont généralement les principales causes d'erreurs de reconstruction 3d.

Dans l'exemple présenté figure 6.1, la quantification et la non linéarité des valeurs mesurées génèrent des défauts dans la reconstruction 3d d'une demi-sphère. En effet, la géométrie est reconstruite avec un effet "marches d'escalier" en fonction des pas de quantification. Plus ceux-ci sont important, plus l'effet est visible. De même, les limitations des plages dynamiques capturées entraînent une perte d'information dans les zones les plus sombres (faibles valeurs radiométriques) et les plus claires (fortes valeurs radiométriques) de l'image. Ce défaut se traduit par des pertes totales de la géométrie estimée. Enfin, l'exemple présenté figure 6.2 démontre qu'un fort niveau de bruit dans les images entraîne l'apparition de micro-rugosités dans la reconstruction 3d d'un plan initialement plat.

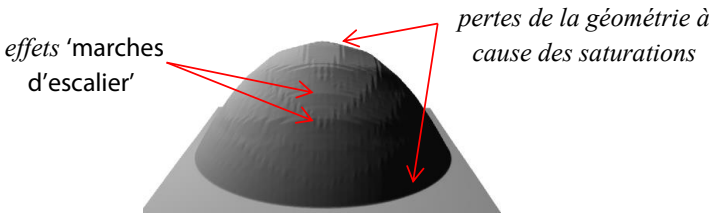


FIGURE 6.1 – *Reconstruction 3d d'une demi-sphère, affectée par le manque de quantification et les saturations du capteur.*

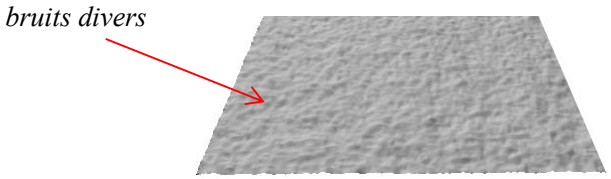


FIGURE 6.2 – *Reconstruction 3d d'un plan initialement lisse, affectée par le bruit lors de l'étape d'acquisition.*

Ces défauts se cumulent lors de l'acquisition d'images, l'hypothèse $\mathcal{H}4$ n'est plus respectée et des biais sont introduits dans les reconstructions 3d. Pour outrepasser les limitations du capteur, certains auteurs [AZK08, GCHS10] proposent d'utiliser des images dites hautes dynamiques ou High Dynamic Range (HDR). Cette parade permet d'écarter les défauts de sur et sous exposition, néanmoins les défauts dus à la quantification sont toujours présents. La suite de ce chapitre a pour but d'inclure la prise en compte de l'ensemble des imperfections du capteur dans le processus d'acquisition et de proposer des solutions qui améliorent la qualité des reconstructions 3d par stéréo photométrie.

6.2 Les images hautes dynamiques

Les images hautes dynamiques sont censées outrepasser les limitations technologiques des capteurs (chapitre 3) et ainsi améliorer les performances des méthodes de reconstruction 3d. Dans le cadre de la stéréo photométrie, l'apport de l'utilisation des images HDR n'a jamais été démontré. Par conséquent, nous proposons un état de l'art du principe HDR ainsi que l'analyse de son utilisation dans les processus de reconstruction 3d par SP.

6.2.1 État de l'art du principe HDR

A l'origine, les images HDR ont été conçues pour outrepasser les limites de la dynamique d'acquisition des capteurs. En effet, une scène intérieur/extérieur (figure 6.3) peut facilement dépasser les $120dB$ alors que les capteurs ont généralement une plage dynamique d'acquisition limitée à $60dB$.

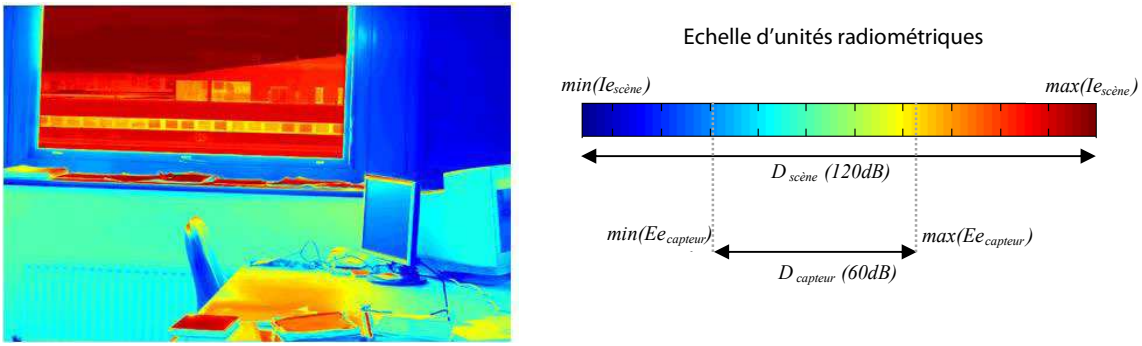


FIGURE 6.3 – *Scène intérieurs/extérieurs représentée en échelle de couleurs proportionnelles aux valeurs radiométriques (à gauche) et représentation des plages dynamiques (à droite).*

Dans ce contexte, les images issues des capteurs sont incapables de conserver l'ensemble des informations de la scène, elles sont dites de faible dynamique ou Low Dynamic Range (LDR). En effet, l'acquisition d'une scène haute dynamique entraîne des phénomènes de sur ou sous exposition (figure 6.4). Les détails des zones dont l'intensité lumineuse dépasse les capacités du capteur sont perdus et apparaissent comme des taches uniformes blanches quantifiées par la plus haute valeur de M. A l'inverse, les zones de trop faible intensité apparaissent uniformément noires et sont quantifiées par la valeur nulle $M = 0$.



A : zone sur exposée, intensité lumineuse trop importante pour les capacités du capteur *M=Mmax*

B : zone sous exposée, intensité lumineuse trop faible pour le capteur donc *M=0*

FIGURE 6.4 – *Exemple d'une image saturée acquise avec un capteur de* $60dB$ *pour une scène de* $120dB$.

Pour acquérir les détails dans l'une des zones sur ou sous exposées, la plage dynamique capturée peut être ajustée grâce aux paramètres du système d'acquisition : ouverture, temps

d'exposition ou sensibilité ISO. Cepandant, la dynamique du capteur n'est pas modifiable, l'acquisition des détails dans les zones à forte intensité se fait donc au détriment des zones à faible intensité (figure 6.5a) et inversement (figure 6.5b).



(a) (b)

FIGURE 6.5 – *Modification de la plage dynamique capturée vers (a) les zones à fortes intensité, (b) les zones à faible intensité.*

Le principe de la prise de vue multi-expositions consiste à effectuer plusieurs acquisitions d'une scène statique avec un point de vue fixe mais des paramètres différents. La modification des paramètres d'acquisition permet la modification de la plage dynamique capturée à chaque image. Ainsi, en fonction du nombre d'acquisitions, la dynamique de la scène peut être entièrement acquise (figure 6.6).



FIGURE 6.6 – *Série d'images multi-expositions capturant toutes la dynamique d'une scène intérieur/extérieur.*

Selon le processus d'acquisition des images (équation 3.38), la fonction crf lie les valeurs numériques quantifiées M aux valeurs d'irradiance réelles Ee en fonction du temps d'obturation Δt. L'irradiance à la surface du capteur est proportionnelle à un facteur multiplicatif k, à la radiance Le ou dans le cas d'une surface diffuse à l'intensité Ie émise par la scène en direction de l'appareil d'acquisition.

$$M = crf(Ee.\Delta t) = crf(k.Le.\Delta t) = crf(k.Ie.\Delta t) \tag{6.1}$$

Les valeurs minimales et maximales d'irradiance capturées par le capteur sont fixes, la solution la plus simple pour modifier la plage dynamique des valeurs capturées Ie consiste à modifier l'une des pondérations k ou Δt. Cependant, la modification de la pondération k [Hor86] est réalisée par la variation des paramètres d'ouverture d ou de la distance focale f ce qui entraîne une modification géométrique des images acquises. Par conséquent, la variation de k est à proscrire lors d'une prise de vue multi-expositions car la scène doit être capturée

selon la même configuration géométrique. A l'inverse, la modification du temps d'obturation Δt permet la capture de plusieurs plages d'intensité émises par la scène sans altération de la configuration géométrique d'acquisition. Il en résulte une série d'images géométriquement identiques qui selon le nombre de temps d'obturation différents sont capables de capturer l'ensemble de la dynamique radiométrique de la scène.

Une alternative est basée sur la modification du paramètre variable d'amplification interne au capteur (voir ISO, chapitre 3). Pour étudier ce principe, il convient de décomposer la fonction crf. En effet, cette fonction lie les valeurs numériques M aux intensités Ie et peut être définie comme :

$$M = crf(k.Ie.\Delta t) = f_{ADC}\left(\min\left(\frac{k.Ie.\Delta t}{g_n} + S_0 + s, S_{max}\right)\right) \tag{6.2}$$

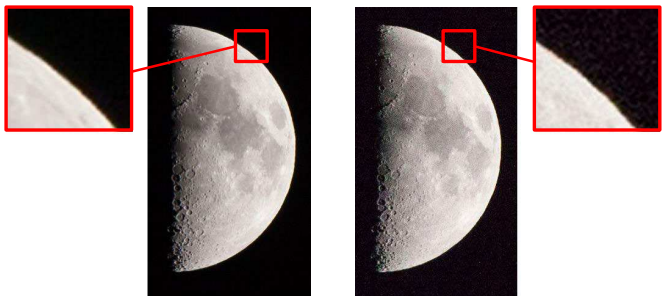
avec :

$S_0 = \frac{S_{DC}\Delta t}{g_n}$	:	signal généré par le bruit $\mathcal{B}_{DC}$ provenant du courant d'obscurité et exprimé par S_{DC} en électron/sec.
S_{max}	:	maximum du signal quantifiable par le CAN.
f_{ADC}	:	fonction de transfert du CAN.
g_n	:	gain d'amplification définie par la sensibilité ISO.
s	:	signal issue des différents bruits.

Le signal s est un bruit aléatoire à moyenne nulle modélisable par :

$$var(s) = \frac{k.Ie.\Delta t}{g_n^2} + \frac{\sigma_{read}^2}{g_n^2} + \sigma_{ADC}^2 \tag{6.3}$$

où le premier terme correspond à une distribution de Poisson des photons collectés. Les autres termes sont indépendants de la scène et correspondent respectivement au bruit de lecture et la conversion du CAN.



(a) (b)

FIGURE 6.7 – *Démonstration de l'augmentation du bruit en fonction des paramètres d'acquisition, (a) 200 ISO $g_n = 1.37$ et Δt, (b) 3200 ISO $g_n = 0.082$ et $0.06\Delta t$, ces images sont acquises avec des conditions identiques seule la sensibilité ISO et le temps d'obturation varient, [Cla12].*

Cette reformulation permet de constater que la variation de la sensibilité ISO entraîne la modification de la plage des valeurs Ie capturables, soit la plage dynamique d'acquisition. Tout comme la modification du temps d'exposition, modifier la sensibilité n'entraîne pas de variation géométrique de l'image entre les acquisitions. En revanche, ces deux solutions entraînent

des variations différentes du bruit d'acquisition. Le rapport signal sur bruit dans l'image diminue en fonction de l'augmentation du temps d'obturation Δt ou du gain d'amplification g_n. Néanmoins pour la prise de vue multi-expositions, la variation du temps d'obturation reste la meilleure solution. L'expérience présentée figure 6.7 démontre ce phénomène, l'impact du bruit est moins important pour une variation du temps d'obturation. En effet, d'après l'équation 6.3, les variations de g_n ont beaucoup plus d'influence sur le bruit qui perturbe les photons collectés.

A partir d'une prise de vue multi-expositions où chacune des G images de la série apporte de nouvelles informations et de la redondance par rapport aux autres. Il est nécessaire de fusionner les images en une estimation des valeurs réelles $\hat{I}e$, $\hat{L}e$ ou $\hat{E}e$ de la scène capturée pour obtenir une image HDR. Les premiers travaux de fusion d'images sont attribués à Mann et Picard dans [MP95]. Leur idée consiste à combiner trois images dont l'une est correctement exposée et les deux autres sont sous et sur exposées afin de couvrir au maximum la dynamique de la scène. Cette fusion (figure 6.8) est réalisée en deux étapes, la première consiste à filtrer les images afin d'éliminer les saturations. Lors de la seconde étape, les trois images sont combinées.

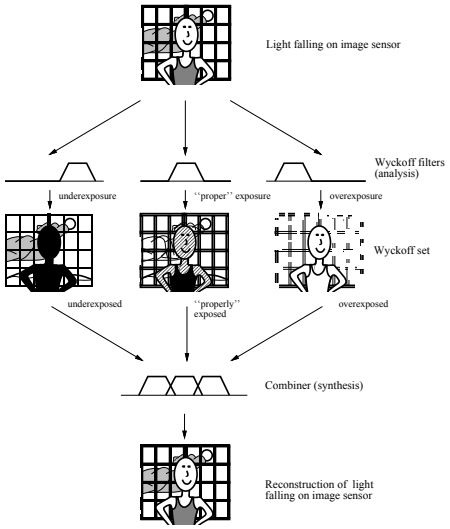


Figure 6.8 – *Principe de fusion proposé par* Mann *et* Picard *issu du document [MP95].*

Peu de temps après, Debevec et Malik établissent une méthode [DM97] plus aboutie basée sur le même principe. Elle utilise une série d'images multi-expositions obtenues par variations des temps d'obturation. Pour réaliser la fusion, les auteurs proposent de déterminer la fonction crf de la caméra à partir de la série d'images. La crf étant définie monotone, l'inversion est possible grâce aux valeurs M tel que :

$$crf^{-1}(M) = Ee.\Delta t \tag{6.4}$$

Si les temps d'obturation sont parfaitement connus, cette définition comporte deux inconnues : la fonction crf^{-1} et les valeurs Ee. Par conséquent un système linéaire utilisant plusieurs valeurs M doit être défini et résolu au sens des moindres carrés afin d'estimer la fonction de réponse inverse. La seconde étape consiste à estimer l'irradiance $\hat{E}e$ à la surface du capteur. Ces valeurs correspondent à un facteur près aux intensités émises par la scène dans chaque

image g :

$$\hat{E}e_u = \frac{crf^{-1}(M_g)}{\Delta t} \tag{6.5}$$

puis elles sont fusionnées :

$$\hat{E}e = \frac{\sum_g w(M_g) \times \hat{E}e_g}{\sum_g w(M_g)} \tag{6.6}$$

avec $w(M_g)$ une fonction de poids utilisée pour réduire le bruit et éliminer les valeurs M saturées.

A partir de cette formulation de la fusion, plusieurs méthodes ont été élaborées [MN99, GN03, RBS03]. Le principe de fusion reste identique mais les fonctions $w(M_g)$ sont différentes (figure 6.9) afin d'éliminer au maximum les effets du bruit ou des non linéarités des courbes de réponse crf [KA06, GAW+10]. De plus, le processus de fusion exige la connaissance des temps d'obturation. Dans le but de supprimer cette contrainte, la méthode [MN99] proposée par Mitsunaga et Nayar propose de les estimer par analyse des rapports entre les valeurs M inter-exposition.

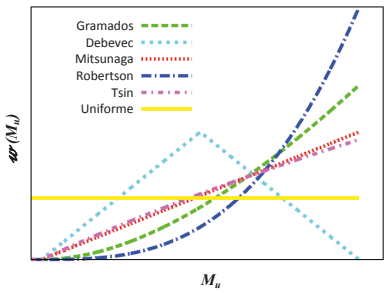


Figure 6.9 – *Différentes fonctions de poids en fonction des méthodes proposées, graphique issu de* Gramados *dans [GAW+10].*

Les valeurs d'intensité estimées numériquement sont des flottants ou du moins elles sont codées sur un nombre de Bits N' (généralement N' =32 Bits/plan) supérieurs aux N Bits d'une image LDR. Elles forment une carte de radiance ou dans le cas des images couleur chaque plan est utilisé de façon marginale. La visualisation des résultats (figure 6.10) exige des opérations de Tome Mapping [RSSF02] afin de les rendre affichables sur un écran. Mais dans le contexte de la SP, les opérations de Tome Mapping sont inutiles car les valeurs d'intensité estimées sont directement utilisées.

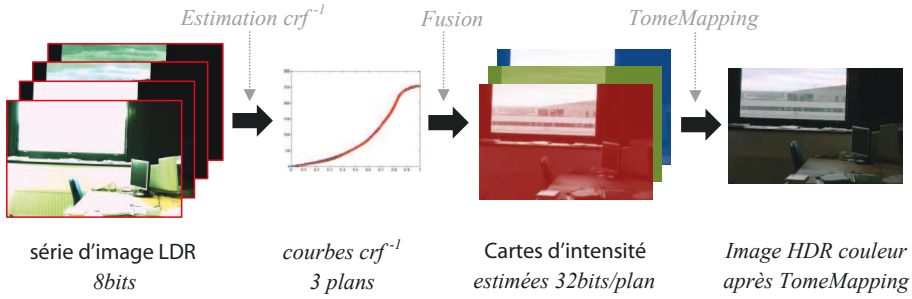


Figure 6.10 – *Schéma de principe de l'algorithme de création d'image HDR affichable sur un écran.*

Afin d'acquérir au maximum la dynamique des scènes, différents auteurs proposent d'utiliser

des ratios χ linéaires et proches de 0.5 entre les G images LDR de la série. Le ratio χ d'une série multi-expositions par variation linéaire des temps d'obturation est défini par :

$$\chi = \frac{\Delta t_g}{\Delta t_{g+1}} \forall g \tag{6.7}$$

Il permet d'influencer la plage dynamique capturée ainsi que la redondance d'information entre les images LDR. En effet, plus la valeur de χ est proche de 1, plus la redondance d'information est importante. A l'inverse, plus sa valeur est faible, plus la plage dynamique capturée est grande :

$$\mathcal{D}_{capturée} = \mathcal{D}_{capteur} + 20 \log \left(\frac{1}{\chi^{G-1}} \right) \tag{6.8}$$

Les limites de χ sont définies par la capacité de variation des temps d'obturation ou de la sensibilité ISO des systèmes d'acquisition. En règle générale, un appareil grand public permet d'atteindre des valeurs comprises entre $0.5 < \chi < 0.8$. Ainsi, une série de $G = 15$ images LDR acquises avec un ratio $\chi = 0.5$ et une dynamique capteur $\mathcal{D}_{capteur} = 60dB$, il est possible de capturer une scène dont la dynamique dépasse les $140dB$. Cette méthode d'acquisition permet de capturer des scènes extérieures avec des dynamiques très élevées et repousse les limites des capteurs d'images numériques (figure 6.11).



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

FIGURE 6.11 – *Exemple d'images HDR obtenues pour des scènes extérieures à hautes dynamique, (a) meilleure image LDR prise dans la série Memorial Church de* DEBEVEC *et (b) image HDR résultante de la fusion [DM97] et du TomeMapping proposé par* WARD*, (c) meilleure image LDR prise dans la série BigFog de* TUMBLIN *et (d) image HDR résultante de la fusion [DM97] et du TomeMapping proposé par* WARD*.*

6.2.2 Images HDR, une solution en SP ?

En théorie, l'inversion d'un modèle de BRDF exige la connaissance des intensités Ie émises par les points de la scène. Or dans le cas applicatif, ces valeurs sont inconnues et remplacées par les valeurs quantifiées M issues de l'appareil d'acquisition grâce à l'hypothèse $\mathcal{H}4$. Mais les capteurs sont imparfaits et des erreurs sont générées lors de la détermination des champs de gradients. Nous proposons d'évaluer l'impact des images LDR et HDR sur les reconstructions 3d à partir d'images de synthèse obtenues par lancer de rayon et le modèle Lambertien. La surface simulée est une demi-sphère de 780 *pixels* de diamètre placée sur un plan de taille 800×800 *pixels*. Des micro-rugosités sont ajoutées grâce à une fonction sinusoïdale bi-directionnelle de période 16 *pixels* pour une amplitude de 10 *pixels*. Les valeurs

de radiance de cette scène sont proportionnelles aux intensités Ie émises et comprises entre $min(Le_{scène}) = 0.015W.m^{-2}.sr^{-1}$ et $max(Le_{scène}) = 30W.m^{-2}.sr^{-1}$. Cette plage dynamique correspond aux configurations d'éclairage rencontrées en laboratoire lors de l'application de la SP. La vérité terrain est obtenue à partir des données de calcul du logiciel POV-Ray. Elle correspond à des images 32 Bits de la scène dont l'éclairage varie suivant trois directions distinctes comme dans le cas de notre système d'acquisition (annexe NIKON_2 A.2.2). La figure 6.12b-c montre une des images 32 Bits ainsi qu'une représentation angulaire Ψ des champs de gradients pour la vérité terrain.

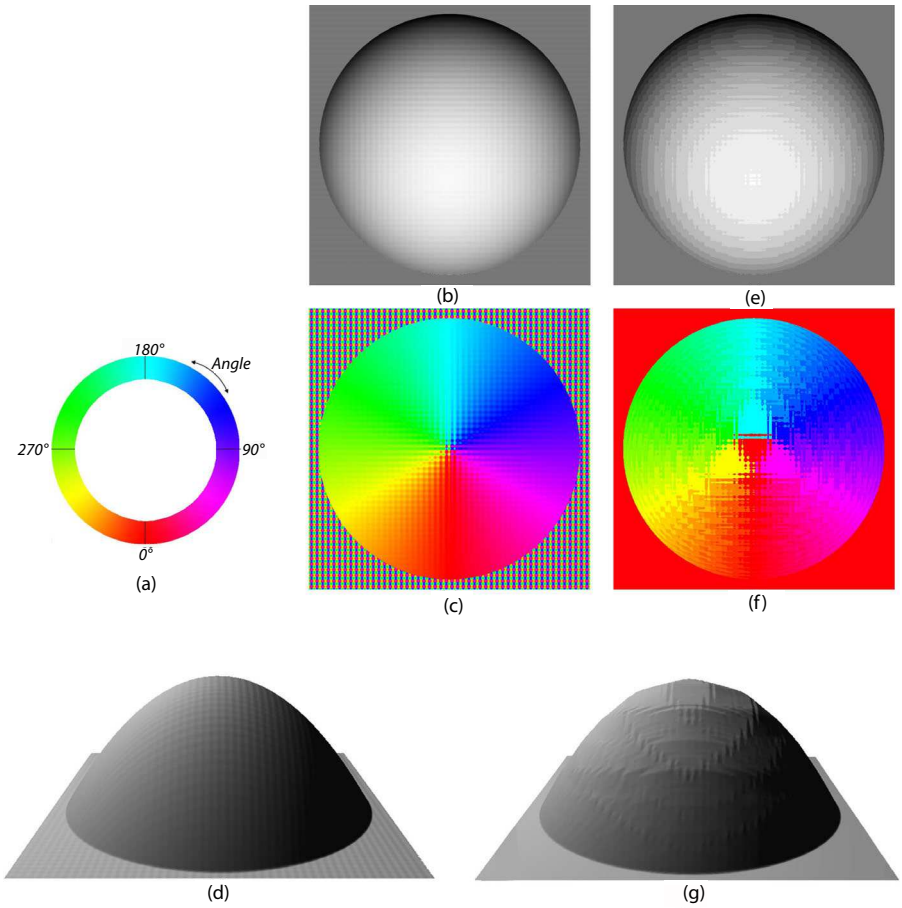


FIGURE 6.12 – *Légende de la représentation angulaire des champs de gradients (a), chaque couleur correspond à une direction, (b) images vérité terrain et (e) images LRD 4 Bits pour une direction d'éclairage, représentation angulaire des champs de gradients et reconstruction 3d par SP pour 3 directions d'éclairage avec (c-d) les images vérités terrain et (f-g) les images LDR 4 Bits.*

Cette scène est aussi acquise avec un capteur d'image dont la dynamique est limitée à $60dB$ pour être plus réaliste. De plus, les valeurs numériques sont volontairement quantifiées sur 4 Bits afin de proposer une analyse visuelle de l'influence de la quantification sur les résultats. Comme précédemment, la figure 6.12e-f montre l'une des images LDR 4 Bits ainsi que la représentation angulaire des champs de gradients obtenus. Toute la dynamique de la scène n'est pas acquise, la plage de radiance capturée est comprise entre $min(Le_{capturée}) = 0.030W.m^{-2}.sr^{-1}$ pour éviter la surexposition et $max(Le_{capturée}) = 30W.m^{-2}.sr^{-1}$. Les pas de quantification sont clairement visibles à la fois sur l'image LDR, la représentation des champs de gradient et la reconstruction 3d (figure 6.12g). Cet effet marche d'escalier est dû au manque d'information dans les images 4 Bits par rapport à la vérité terrain. Dans le cas d'images LDR quantifiées par un nombre de Bits plus important (8-12 Bits), il est moins visible mais il reste présent et influe

sur la qualité des reconstructions. Les premières zones affectées sont les micro-rugosités, elles disparaissent presque entièrement dans notre cas.

La suite de cette étude porte sur l'impact de l'utilisation des images HDR dans le processus de SP. En effet, ces images permettent une capture totale de la plage dynamique de la scène et une redondance d'information susceptible d'augmenter la quantification. Pour cela, nous utilisons la méthode [DM97]. Conformément aux préconisations des auteurs, le nombre G d'images LDR est déterminé par la plage dynamique de la scène et le ratio fixe $\chi = 0.5$. Les temps d'obturations sont décroissants suivant χ. La première valeur $\Delta t(1)$ est choisie arbitrairement par l'utilisateur afin de capturer les détails des zones les plus sombres. Le processus d'acquisition est terminé lorsque les zones les plus claires sont acquises sans sur exposition. Pour notre scène simulée, au maximum 15 images LDR peuvent être acquises pour chaque direction d'éclairage. Elles sont quantifiées sur 4 Bits puis fusionnées afin d'obtenir une carte d'intensité (image HDR) estimée sur 32 Bits. Les trois cartes d'intensité estimées sont ensuite utilisées pour résoudre la SP et obtenir une reconstruction 3d. Dans ce contexte, la dynamique capturable est très largement supérieure à celle de la scène et est comprise entre $min(Le_{capturée}) = 3.75e-5 W.m^{-2}.sr^{-1}$ et $max(Le_{capturée}) = 6144 W.m^{-2}.sr^{-1}$. Beaucoup d'images LDR sont sous ou surexposées, mais nous respectons $G = 15$ images afin de comparer équitablement les plages dynamiques capturées et les distributions des pas de quantification entre les méthodes.

La figure 6.13 compare la distribution des pas de quantification entre la meilleure image LDR et l'image HDR 32 Bits. Comme nous pouvons le constater, l'image HDR permet de capturer l'ensemble de la plage dynamique de la scène. De plus, les zones sombres sont densément quantifiées alors que les zones claires sont quantifiées de façon équivalente à l'image LDR 4 Bits malgré les 32 Bits de l'image HDR.

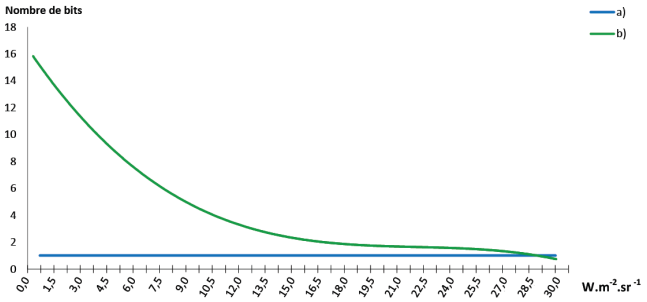


FIGURE 6.13 – *Graphique comparant le nombre de pas de quantification entre (a) une image LDR 4 Bits et (b) une image HDR 32 Bits pour la même plage dynamique radiométrique d'une scène.*

La figure 6.14d-e montre un exemple d'une image HDR 32 Bits pour l'une des trois directions d'éclairage utilisées et la représentation angulaire obtenue après application de la SP. Nous proposons une comparaison des champs de gradients entre les résultats issus des images LDR et ceux issus des images HDR. Cette comparaison est réalisée grâce au calcul de l'erreur moyenne angulaire définie comme :

$$\bar{\epsilon}_\Psi = |\Psi_1 - \Psi_2| \tag{6.9}$$

où Ψ_1 correspond à la représentation angulaire des champs de gradients issus de la vérité terrain et Ψ_2 correspond à la représentation angulaire des champs de gradients issus des images LDR ou des images HDR. Les valeurs d'erreurs angulaires moyennes obtenues (images

LDR : $\bar{\epsilon}_\Psi = 33.736^o$ et images HDR : $\bar{\epsilon}_\Psi = 33.545^o$) sont relativement proches et permettent d'affirmer que l'utilisation des images HDR n'apporte pas de contribution significative dans la résolution de la SP. En effet, les défauts de reconstruction 3d sont légèrement atténués mais restent similaires (figure 6.14f).

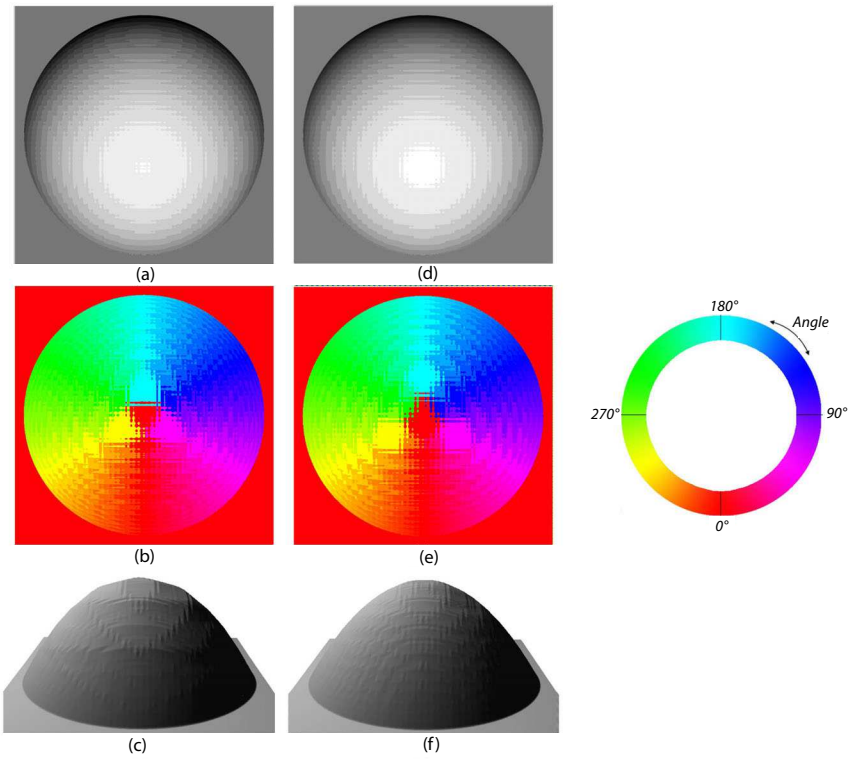


FIGURE 6.14 – *Images LDR 4 Bits (a) et images HDR 32 Bits (b) pour une direction d'éclairage, représentation angulaire des champs de gradients et reconstruction 3d par SP pour 3 directions d'éclairage avec (c-e) les images LDR 4 Bits et (d-f) les images HDR 32 Bits.*

Cette étude permet de conclure que les méthodes d'acquisition et de fusion HDR sont particulièrement adaptées aux hautes dynamiques des scènes extérieures. Mais dans le cadre applicatif de la stéréo-photométrie, le contexte est différent. En effet, les contraintes d'acquisition obligent l'utilisation d'un éclairage contrôlé dont l'intensité est beaucoup plus limitée que celle d'une scène extérieure. De plus, les surfaces analysées sont obligatoirement diffuses et les phénomènes spéculaires sont éliminés (voir chapitre 4). Par conséquent, les dynamiques à acquérir sont faibles et généralement proches de celles des capteurs. De plus lors de la fusion, la redondance d'information n'apporte pas une distribution équitable des pas de quantification sur l'ensemble de la plage dynamique capturée.

6.3 Une nouvelle méthode d'acquisition : images AIQ

Toujours sur le principe de la fusion multi-expositions, nous proposons une méthode d'acquisition [BBK11, BBK13a] adaptée au contexte de la stéréo photométrie. Notre but est d'estimer au plus juste les valeurs Ie de toute la dynamique de la scène à partir des valeurs quantifiées imparfaites M fournies par une série d'image LDR. Pour traiter correctement cette problématique et limiter les erreurs de reconstruction 3d, il est nécessaire d'étudier le principe de fusion (équations 6.5 et 6.6) afin d'obtenir une image dont les valeurs $\hat{Ie}$ estiment les valeurs Ie réelles. La méthode d'acquisition proposée permet d'augmenter la précision de la

quantification. Elle est inspirée par les travaux de Grosberg et Nayar [GN03] qui consistent à choisir les images LDR à fusionner en fonction des temps d'exposition pour obtenir une réponse particulière des capteurs. Dans notre cas, nous avons adapté cette caractéristique pour augmenter la qualité des reconstructions 3d dans le cas des contraintes d'acquisition fixées par la stéréo photométrie.

6.3.1 Le compromis Quantification / Dynamique

L'objectif est d'obtenir une quantification des valeurs d'intensité Ie plus précise que celle fournie par une image LDR ou celle obtenue dans une image HDR après fusion. Comme il existe une relation de proportionnalité linéaire (équation 6.1) entre les valeurs d'irradiance Ee, de radiance Le et d'intensité Ie, nous utilisons dans cette étude le terme radiométrique radiance Le plus facilement mesurable. Nous étudions la quantification des valeurs de radiance pour une plage définie par :

$$\hat{Le}_M(1,n) \leq Le \leq \hat{Le}_M(1,n+2) \tag{6.10}$$

où $\hat{Le}_M(1,n)$ est l'estimation de la radiance scène pour le pas de quantification n de l'acquisition LDR avec le temps d'obturation $\Delta t(1)$.

Pour cette plage de radiance, la meilleure image LDR permet d'obtenir deux pas de quantification. Dans le cas d'une image HDR classique avec un ratio de $\chi = 0.5$, l'image LDR suivante à fusionner (temps d'obturation $\Delta t(2)$) fournit deux pas de quantification supplémentaires :

$$\begin{aligned} \hat{Le}_M(2,n) &= \hat{Le}_M(1,n) \\ \hat{Le}_M(2,n+1) &= \hat{Le}_M(1,n+2) \end{aligned} \tag{6.11}$$

La figure 6.15 illustre ce principe. Selon l'équation 6.6 pour la plage de radiance les valeurs estimable de Le sont :

$$\begin{aligned} \hat{Le}(1-2,n) &= \frac{\hat{Le}_M(1,n)\ w(M(1,n)) + \hat{Le}_M(2,n)\ w(M(2,n))}{w(M(1,n)) + w(M(2,n))} \\ \hat{Le}(1-2,n+1) &= \frac{\hat{Le}_M(1,n+1)\ w(M(1,n+1)) + \hat{Le}_M(2,n)\ w(M(2,n))}{w(M(1,n+1)) + w(M(2,n))} \end{aligned} \tag{6.12}$$

Ces valeurs ne sont pas nécessairement identiques à $\hat{Le}_M(1,n)$ ou $\hat{Le}_M(1,n+1)$, mais le nombre de pas de quantification reste à deux comme pour une image LDR.

Par contre lorsque le ratio χ est différent de $1/2a$ avec $a \in \mathbb{N}$, (figure 6.16, où $\chi = 0.6$), les valeurs estimables de Le sont :

$$\begin{aligned} \hat{Le}(1-3,n) &= \frac{\hat{Le}_M(1,n)\ w(M(1,n)) + \hat{Le}_M(3,n)\ w(M(3,n))}{w(M(1,n)) + w(M(3,n))} \\ \hat{Le}(1-3,n+1) &= \frac{\hat{Le}_M(1,n+1)\ w(M(1,n+1)) + \hat{Le}_M(3,n)\ w(M(3,n))}{w(M(1,n+1)) + w(M(3,n))} \\ \hat{Le}(1-3,n+2) &= \frac{\hat{Le}_M(1,n+1)\ w(M(1,n+1)) + \hat{Le}_M(3,n+1)\ w(M(3,n+1))}{w(M(1,n+1)) + w(M(3,n+1))} \end{aligned} \tag{6.13}$$

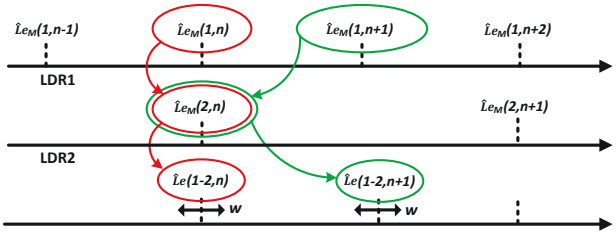


FIGURE 6.15 – *Fusion de deux images LDR avec un ratio* $\chi = 0.5$.

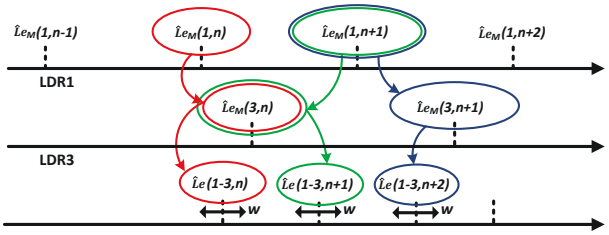


FIGURE 6.16 – *Fusion de deux images LDR avec un ratio* $\chi = 0.6$.

Avec un ratio χ adapté, la quantification est alors plus précise et ce cas idéal permet de constater que chaque image multi-expositions peut ajouter des pas de quantification supplémentaires dans l'estimation d'une plage dynamique. De plus, cette plage dynamique capturée doit être adaptée à celle de la scène pour éliminer les phénomènes de saturations. Pour cela, la première acquisition doit avoir un temps d'obturation qui permet la capture des détails dans les zones les plus sombres :

$$\Delta t(1) = \frac{min(Le_{capteur})}{min(Le_{scène})} \tag{6.14}$$

Les temps d'obturation des $G-1$ images suivantes sont croissants suivant χ défini à partir des plages dynamiques estimées ou mesurées du capteur $\mathcal{D}_{capteur}$ et de la scène $\mathcal{D}_{scène}$:

$$\chi = \left[\frac{1}{10^{\frac{\mathcal{D}_{scène} - \mathcal{D}_{capteur}}{20}}} \right]^{\frac{1}{G-1}} \tag{6.15}$$

L'objectif est d'atteindre un ratio le plus proche possible de $\chi = 1$ afin de maximiser le nombre de pas de quantification supplémentaires. Dans le contexte de notre scène simulée et d'un nombre fixe $G = 15$ images, le ratio est égal à $\chi = 0.95$. La figure 6.17c montre la distribution des pas de quantification pour cette configuration. La distribution est améliorée par rapport au cas de l'image HDR classique où $\chi = 0.5$, mais elle n'est pas optimale. En effet, elle n'est pas constante pour l'ensemble de la plage dynamique capturée. Ce défaut provient du nombre d'image LDR fusionnées qui varie selon les plages de radiance. Cependant nous pouvons définir une plage où la quantification est maximale :

$$min(Le_{AQ}) \leq Le_{AQ} \leq max(Le_{AQ}) \tag{6.16}$$

où $min(Le_{AQ})$ correspond au début de la fusion des G images LDR, soit le deuxième Bits de l'image acquise avec $\Delta t(G)$. $max(Le_{AQ})$ est la fin de la fusion des G images, soit le dernier Bits de l'image acquise avec $\Delta t(1)$.

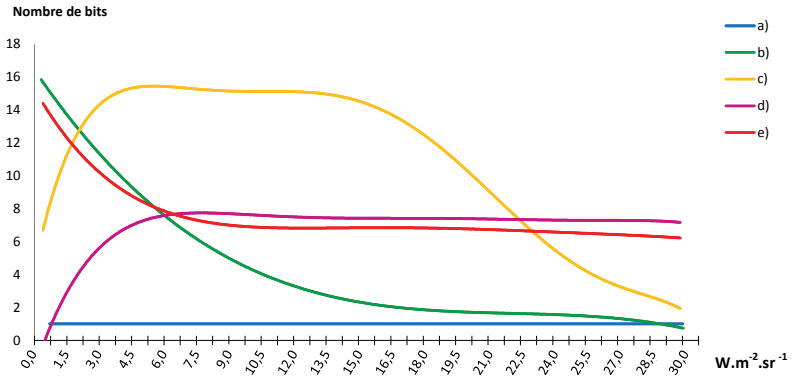


FIGURE 6.17 – *Graphique comparant pour la même plage dynamique d'une scène le nombre de pas de quantification entre une image LDR 4 Bits (a) et une image fusionnée 32 Bits avec $\chi = 0.5$ constant (b), $\chi = 0.95$ constant (c), $\chi = 0.95$ constant et une quantification constante après $min(Le_{AQ})$ (d) et notre méthode AIQ (e).*

Afin d'obtenir une distribution optimale des pas de quantification pour toute la plage dynamique capturée, il est nécessaire que $min(Le_{AQ}) = min(Le_{scène})$ et $max(Le_{AQ}) = max(Le_{scène})$. Or ces deux conditions ne peuvent être atteintes pour un ratio χ constant. Nous proposons de satisfaire la seconde condition car elle est la plus simple à appliquer. Dès lors, il existe deux solutions équivalentes pour capturer toute la dynamique de la scène : la première consiste à réduire χ et conserver $\Delta t(1)$ ou la seconde consiste à conserver χ et diminuer $\Delta t(1)$ de manière à décaler la plage dynamique capturée vers les fortes valeurs de radiance. Nous préconisons la seconde solution car $\Delta t(1)$ peut être facilement déterminé par :

$$\Delta t(1) = \frac{max(Le_{capteur})}{max(Le_{scène})} \tag{6.17}$$

Toujours pour $U = 15$, dans le cas de notre scène simulée $\Delta t(1)$ est fixé à $0.5s$. La figure 6.17d montre la distribution constante des pas de quantification entre $min(Le_{AQ})$ et $max(Le_{scène})$. En revanche la plage des faibles valeurs entre $min(Le_{scène})$ et $min(Le_{AQ})$ est peu quantifiée. Pour résoudre ce problème nous proposons d'utiliser un ratio χ non constant.

Ainsi à partir de la valeur actuelle de $min(Le_{AQ})$, nous augmentons un ou plusieurs temps d'exposition tel que :

$$\Delta t(U) = \frac{max(Le_{capteur})}{min(Le_{AQ})} \tag{6.18}$$

Typiquement $min(Le_{AQ})$ est très faible et seul le temps d'obturation $\Delta t(G)$ doit être modifié. Quand $min(Le_{scène})$ est nul, il est nécessaire de fixer arbitrairement une valeur minimale à capturer. Dans notre cas, $min(Le_{scène}) = 0.015W.m^{-2}.sr^{-1}$ et $\Delta t(G) = 6.7s$. Cette configuration permet à notre méthode d'acquérir complètement la dynamique de la scène avec une meilleure distribution des pas de quantification (figure 6.17e).

6.4 Les images AIQ pour la SP

Pour obtenir le meilleur compromis entre la dynamique et la distribution optimale des pas de quantification, la série doit comprendre un maximum d'images LDR à fusionner. Cette

contrainte a pour conséquence d'augmenter considérablement le temps global d'acquisition Δt_G nécessaire à l'obtention d'une série par variation des temps d'obturation et défini comme :

$$\Delta t_G = \sum_{g}^{G} \Delta t(g) \tag{6.19}$$

De plus, chaque $\Delta t(g)$ est différent et peut atteindre des valeurs élevées (plusieurs secondes) en fonction de l'intensité émise par la scène et de l'intensité initiale des sources lumineuses. Afin de réduire Δt_G, la solution consiste à fixer les temps d'obturation à une faible valeur qui reste identique pour toutes les images. Ainsi, le temps global d'acquisition est uniquement dépendant du nombre G d'images. Mais pour obtenir une série multi-expositions, un des paramètres d'acquisition doit inévitablement varier. Comme nous l'avons démontré précédemment la modification de k entraîne des distorsions géométriques et celle de la sensibilité ISO peut générer une forte amplification du bruit.

Néanmoins, dans le contexte de la SP, il existe une autre solution car les contraintes applicatives exigent le contrôle des sources lumineuses. C'est pourquoi notre méthode propose de remplacer la variation d'un paramètre d'acquisition par la variation de l'intensité lumineuse incidente à la surface analysée. Cette solution apparait en combinant le modèle Lambertien (équation 2.15) au processus d'acquisition des images (équation 3.38) :

$$M = crf(Ie_0 \, \rho \, \cos\theta \, k \, \Delta t) \tag{6.20}$$

avec Ie_0 l'intensité initiale de la source lumineuse éclairant la surface. Cette combinaison permet de redéfinir l'équation 6.5 telle que :

$$\hat{Ie}_g = \frac{crf^{-1}(M_g)}{Ie_0(g)} \tag{6.21}$$

où $\hat{Ie}_g$ correspond à l'intensité émise par la scène dans l'image g au facteur multiplicatif $(k \, \Delta t)$ près.

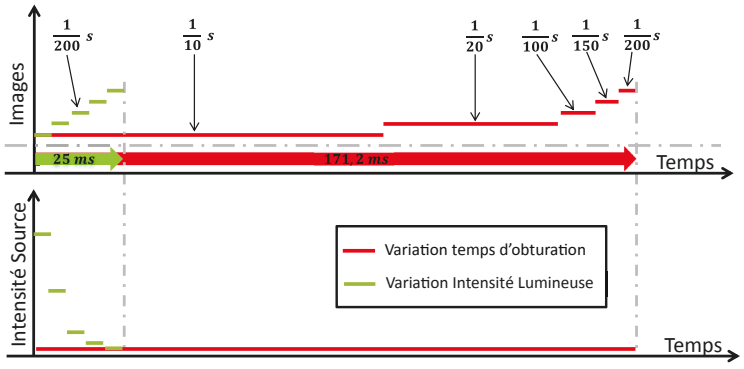


FIGURE 6.18 – *Graphique de comparaison entre notre méthode de variation d'intensité lumineuse et la méthode classique de variation des temps d'obturation, pour* 5 *images le temps d'acquisition global passe de* 171.2 *ms à* 25 *ms lorsque le temps d'obturation est fixé à* 1/200 *s.*

Les résultats après fusion sont identiques mais le temps global d'acquisition Δt_G est constant en fonction du nombre G d'images LDR acquises. De plus, si le temps Δt choisi est faible alors

Δt_G est considérablement réduit. La figure 6.18 compare cette nouvelle méthode de variation d'intensité lumineuse à la méthode classique de variation des temps d'obturation.

6.4.1 Protocole d'acquisition AIQ

Le protocole d'acquisition des images AIQ se compose des étapes suivantes :

1. calibration des valeurs de radiance capturables par le capteur ainsi que celle de la scène.
2. sélection du nombre G d'images LDR à acquérir conformément aux valeurs définies (étape 1) et calcul du ratio optimal χ avec l'équation 6.15.
3. détermination du premier temps d'obturation $\Delta t(1)$ ou de la première intensité source $Ie_0(1)$ avec l'équation 6.17 et calcul de la plage de valeurs Le_{AQ}.
4. détermination du dernier temps d'obturation $\Delta t(G)$ ou de la dernière intensité source $Ie_0(G)$ avec l'équation 6.18.
5. acquisition de la série d'images LDR pour chaque direction d'éclairage.
6. estimation des intensités $\hat{I}e$.
7. reconstruction 3d par stéréo photométrie.

Les valeurs $\hat{I}e$ des images AIQ sont estimées de façon marginale, la fusion est donc effectuée indépendamment sur les trois plans Rouge, Vert et Bleu. Pour éviter les erreurs colorimétriques dues à la réponse spectrale du capteur, un recalage des trois valeurs estimées grâce à des coefficients de balance des blancs. Ces coefficients sont déterminés à partir d'une mire de calibration acquise dans les mêmes conditions que les surfaces analysées. De plus, les valeurs $\hat{I}e$ sont estimées au coefficient $(k\ \Delta t)$ près. Ce coefficient multiplicatif est obtenu par la comparaison entre les valeurs estimées et celles mesurées à l'aide d'un spectrophotomètre sur la mire. La plage dynamique du capteur et celle capturable par variation lumineuse sont estimées à partir de la série multi-exposition. Celle de la scène est déterminée par l'analyse des valeurs $\hat{I}e_g$ d'une image sur puis sous exposée.

6.4.2 Les images AIQ et les reconstructions 3D

6.4.2.1 Résultats de simulation

Pour évaluer l'apport de la méthode d'acquisition AIQ, nous l'appliquons à la stéréo photométrie pour la même scène simulée. Les résultats de reconstruction 3d (micro et macro géométrie) obtenus avec nos images AIQ 32 Bits (figure 6.19) sont similaires à la vérité terrain (figure 6.12).

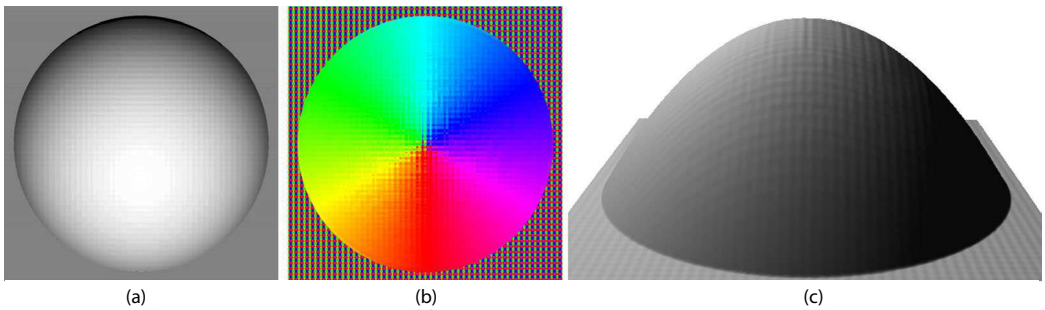


FIGURE 6.19 – *Images AIQ 32 Bits (a) pour une direction d'éclairage, représentation angulaire des champs de gradients et reconstruction 3d par SP pour 3 directions d'éclairage avec (b-c) les images AIQ 32 Bits.*

L'erreur angulaire moyenne des champs de gradients est de 4.573^o, soit une réduction de 13.5% par rapport aux autres méthodes. Ces résultats confirment l'importance de la quantification des images utilisées par la SP et particulièrement pour les surfaces à forte rugosité.

Dans le but d'approfondir cette étude, nous comparons notre méthode AIQ à d'autres méthodes de stéréo-photométrie qui utilisent la redondance d'information par augmentation des directions d'éclairage [BP03, SSS+07, APB10, BBK12]. Pour la scène simulée, 15 directions d'éclairage permettent de produire des images 4 Bits. Comme le démontrent les résultats présentés tableau 6.1, la redondance d'information apportée par l'ajout de sources lumineuses ne permet pas une réelle amélioration tant que la quantification des images reste faible. En effet, l'erreur angulaire moyenne obtenue à partir d'images 4 Bits pour 15 directions d'éclairage reste toujours très élevée par rapport à celle obtenue pour nos images AIQ et seulement 3 directions d'éclairage.

Méthode d'acquisition	**LDR-3**	**HDR**	**LDR-15**	**AIQ-3**	**AIQ-15**
Nbre de Bits de codage	4	32	4	32	32
Nbre d'images fusionnées	1	15	1	15	15
Nbre de directions d'éclairage	3	3	15	3	15
Type de ratio χ	aucun	constant	aucun	non linéaire	non linéaire
Erreur $\bar{\epsilon}_\Psi$ en (^{o})	33.736	33.545	31.111	4.573	1.116

TABLE 6.1 – *Comparaison des erreurs angulaires moyennes des champs de gradients obtenus par SP pour des images LDR, HDR et AIQ.*

Nous pouvons donc affirmer que notre méthode AIQ offre le meilleur compromis entre la distribution des pas de quantification et la plage dynamique capturée. Elle est particulièrement adaptée aux conditions spécifiques d'acquisition exigées par la stéréo photométrie.

6.4.2.2 Surfaces réelles

A présent, nous procédons à l'évaluation des performances de la méthode AIQ dans le cas de mesures métrologiques de rugosité sur des surfaces réelles. Notre système d'acquisition (annexe NIKON_2 A.2.2) est similaire à celui mis en œuvre lors des simulations. Le temps d'obturation est fixé à $1/200$ s et l'intensité des trois sources lumineuses varie suivant 15 valeurs par ratio $\chi = 0.8$. Les images AIQ codées sur 32Bits sont obtenues par la fusion d'images LDR 8 Bits.

Trois surfaces réelles et texturées sont utilisées. La première est un masque Africain avec une forte géométrie facile à visualiser et une fine micro-rugosité. Cette surface présente l'avantage d'être diffuse et possède une plage dynamique proche de celle du capteur. Les deux autres surfaces utilisées sont des échantillons de murs peints, ils ont une très forte rugosité et produisent de grandes variations d'intensité.

Surface	$Sq(\mu m)$	Ssk	Sku
Masque Africain	49	0.05119	2.33349
Murs médiéval peint	200	-0.4901	2.33968 ‘
Murs moderne peint	870	0.26878	2.73753

TABLE 6.2 – *Mesure de rugosité (Norme ISO 25178) sur les surfaces réelles analysées.*

Pour évaluer la rugosité des surfaces analysées, le tableau 6.2 propose plusieurs mesures sta-

tistiques issues de la norme [ISO09]. Cette norme est communément utilisée dans le domaine industriel pour mesurer les caractéristiques d'une surface. Les principales valeurs pour la rugosité sont : la rugosité moyenne quadratique (Sq), l'asymétrie surfacique (Ssk) et le kurtosis (Sku) définis par :

$$Sq = \sqrt{\frac{1}{UV}\sum_{u=1}^{U}\sum_{v=1}^{V}\hat{\zeta}^2(x_u, y_v)} \tag{6.22}$$

$$Ssk = \sqrt{\frac{1}{UVSq^3}\sum_{u=1}^{U}\sum_{v=1}^{V}\hat{\zeta}^3(x_u, y_v)} \tag{6.23}$$

$$Sku = \sqrt{\frac{1}{UVSq^4}\sum_{u=1}^{U}\sum_{v=1}^{V}\hat{\zeta}^4(x_u, y_v)} \tag{6.24}$$

où U et V sont la résolution horizontale et verticale de la carte des hauteurs $\hat{\zeta}$.

Dans un premier temps, nous comparons nos images AIQ 32 Bits aux images HDR 32 Bits issues de la fusion classique avec $\chi = 0.5$ ainsi que les meilleures images LDR 8 Bits pour le masque africain.

La meilleure image LDR est présentée figure 6.20a. Elle a plusieurs zones sur et sous exposées car les limitations du capteur empêchent l'acquisition de toute la dynamique de la scène. En effet, la dynamique du capteur est de $60dB$ alors que celle de la scène est de $87dB$. De plus, la distribution des pas de quantification est dépendante des performances du capteur. Dans le cas de l'image HDR, la plage dynamique capturée ($144dB$) est injustifiée comparée à celle de la scène. La distribution des pas de quantification (figure 6.20c) est principalement concentrée dans les zones les plus sombres et décroît exponentiellement lorsque les zones présentent une plus forte intensité.

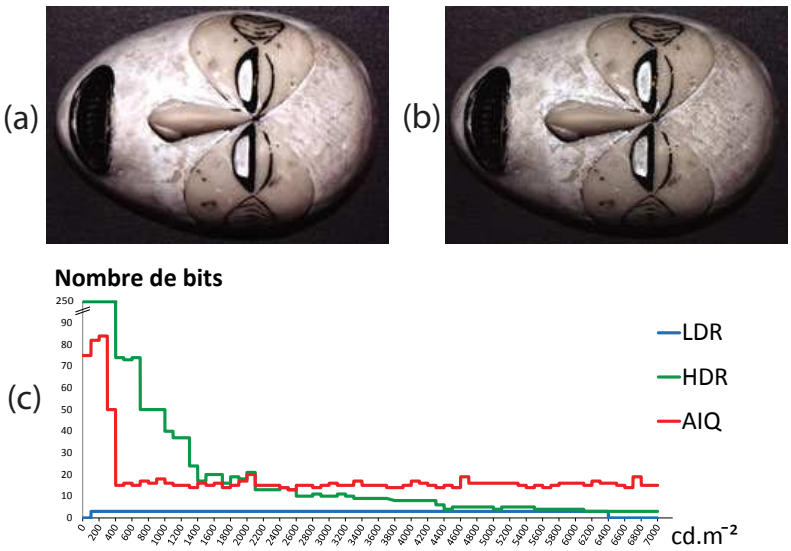


FIGURE 6.20 – *Meilleur compromis entre la sur-ou sous-exposition parmi les images LDR 8 Bits (a) et une image AIQ 32 Bits ayant subi un TomeMapping pour l'affichage (b), la comparaison de la distribution des pas de quantification (c) pour l'image LDR 8 Bits, une image HDR 32 Bits et notre image AIQ 32 Bits. La plage dynamique analysée correspond à la plage dynamique de scène* $[0.2 - 7000]$ cd/m^2. *Le nombre de pas de quantification est donné tous les* 100 cd/m^2.

Les images AIQ sont le meilleur compromis entre l'acquisition de toute la plage dynamique de la scène et une distribution régulière des pas de quantification. Nous pouvons constater que le nombre de pas est au minimum cinq fois supérieur à celui d'une image LDR et reste constant pour les intensités élevées contrairement à celui des images HDR.

Nos images AIQ 32 Bits, les images HDR 32 Bits ainsi que les meilleurs compromis entre sur et sous expositions parmi les images LDR 8 Bits sont utilisées pour la reconstruction 3d par SP à trois sources lumineuses. Avec les images LDR, les formes macro géométriques sont mal reconstruites à cause des saturations, beaucoup d'inversions concaves/convexes ainsi que des pics injustifiés sont visibles (figure 6.21a). L'acquisition complète de la plage dynamique de la scène solutionne ce problème, avec les images HDR ou AIQ la macro géométrie des surfaces reconstruites est conforme à la réalité (figure 6.21b-c).



(a) (b) (c)

FIGURE 6.21 – *Reconstruction 3d du masque africain par SP avec (a) les images LDR 8 Bits, (b) les images HDR 32 Bits, (c) nos images AIQ 32 Bits.*

Pour évaluer la qualité de reconstruction de la microrugosité, il est nécessaire d'observer une représentation 2d des champs de gradients où les directions sont représentées par une variation de la teinte et les normes sont corrélées à la saturation. Pour les images LDR et HDR, ces représentations des champs de gradients (figure 6.22c-d) ont une apparence granuleuse et aléatoire car les directions ainsi que les normes subissent de violents changements dans les zones micro-rugueuses. Cette apparence anormale génère des erreurs de reconstruction et est principalement due au manque de quantification des moyennes et fortes intensités. Pour compléter cette analyse visuelle, nous proposons des mesures d'erreurs angulaires moyennes sur les champs de gradients (équation 6.9). N'ayant pas accès à la vérité terrain, nous prenons comme référence les champs de gradients obtenus par la méthode AIQ avec $G = 15$ images, car d'après les résultats de simulation (tableau 6.1) elle permet d'obtenir une erreur très faible comparée aux autres méthodes.

Les résultats obtenus (tableau 6.3) confirment l'analyse visuelle, l'erreur angulaire moyenne est importante : environ 19.75^o pour les images HDR et 26.07^o pour les images LDR. La différence provient de la non saturation des images HDR mais elle est insuffisante car elle reste trop importante pour l'analyse métrologique de la rugosité. A l'inverse l'usage de nos images AIQ permet d'obtenir des champs de gradients réguliers et les variations sont douces comme sur les surfaces réelles (figure 6.22a). L'erreur angulaire est réduite, sa moyenne pour toutes les zones des surfaces analysées est de 7.4^o et conformément aux résultats de simulation, le minimum atteint est compris entre 4 et 5^o.

Ces résultats démontrent une fois encore l'importance de la quantification des valeurs d'intensité émises par la scène dans le processus de reconstruction 3d par stéréo photométrie. La phase d'acquisition est d'autant plus importante lorsque les surfaces analysées présentent

une forte micro-rugosité. Notre méthode AIQ offre une amélioration significative des résultats et les reconstructions 3d obtenues peuvent être employées pour réaliser des mesures métrologiques de rugosité.

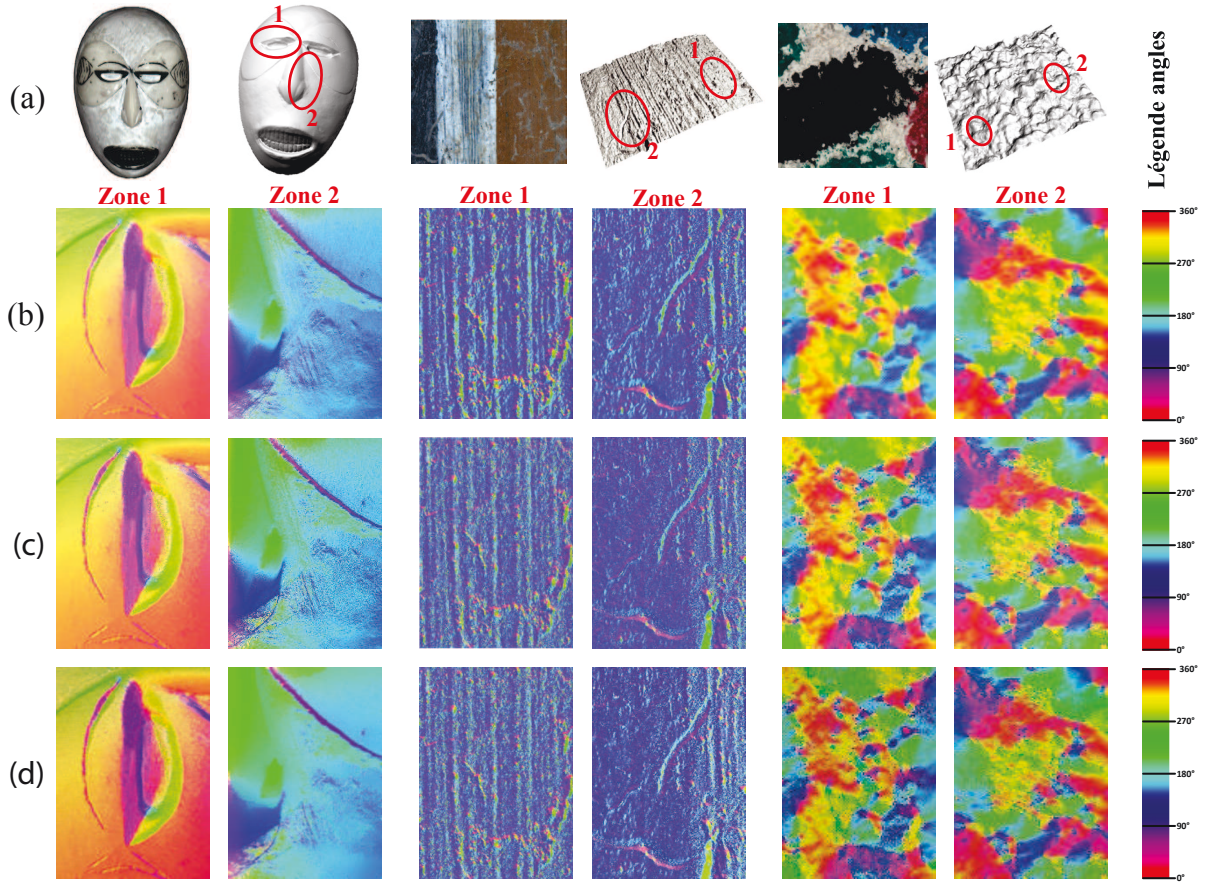


FIGURE 6.22 – *Résultats pour trois reconstructions 3d de surfaces réelles avec la méthode SP à 3 directions d'éclairage : l'albédo et des exemples de reconstruction 3d (a), représentation 2d des champs de gradients pour les images AIQ 32 Bits (b), les images HDR 32 Bits (c) et les meilleurs compromis des images LDR 8 Bits (d).*

Méthode d'acquisition	**LDR**	**HDR**	**AIQ**
Nbre de Bits de codage	4	32	32
Nbre d'images fusionnées	1	15	15
Nbre de directions d'éclairage	3	3	3
Type de ratio χ	aucun	constant	non linéaire
Erreur $\bar{\varepsilon}_\Psi$ Masque Africain : complet	28.747^o	21.319^o	9.265^o
Erreur $\bar{\varepsilon}_\Psi$ Masque Africain : zone 1	42.288^o	15.458^o	9.327^o
Erreur $\bar{\varepsilon}_\Psi$ Masque Africain : zone 2	42.687^o	37.395^o	12.192^o
Erreur $\bar{\varepsilon}_\Psi$ Murs médiéval peint : complet	19.892^o	14.160^o	4.892^o
Erreur $\bar{\varepsilon}_\Psi$ Murs médiéval peint : zone 1	21.871^o	13.309^o	4.133^o
Erreur $\bar{\varepsilon}_\Psi$ Murs médiéval peint : zone 2	9.371^o	14.590^o	5.325^o
Erreur $\bar{\varepsilon}_\Psi$ Murs moderne peint : complet	23.629^o	19.565^o	8.169^o
Erreur $\bar{\varepsilon}_\Psi$ Murs moderne peint : zone 1	23.700^o	23.697^o	6.326^o
Erreur $\bar{\varepsilon}_\Psi$ Murs moderne peint : zone 2	22.452^o	18.186^o	7.060^o

TABLE 6.3 – *Comparaison des erreurs angulaires moyennes des champs de gradients obtenus par SP pour des images LDR, HDR et AIQ sur des surfaces réelles.*

Conclusion du chapitre

Ce chapitre est consacré à l'impact des limitations des capteurs de vision sur les reconstructions 3d par stéréo-photométrie. En effet, l'inversion des modèles de BRDF est faite à partir des intensités lumineuses émises ou réfléchies par les surfaces analysées. Mais dans le cas applicatif, les valeurs utilisées sont celles acquises et quantifiées par les capteurs. Ce processus d'acquisition d'images numériques est un processus imparfait car il présente des limitations en termes de quantification et de dynamique. L'étude menée au cours de la première partie de ce chapitre, montre que l'hypothèse applicative $\mathcal{H}4$ n'est pas respectée et l'impact sur les reconstructions est significatif car l'estimation des surfaces est fortement biaisée. Les erreurs de reconstruction induites sont principalement dues au manque d'informations lorsque la dynamique de la scène n'est pas en adéquation avec celle du capteur ou lorsque l'étape de conversion analogique/numérique produit des pas de quantification trop importants.

Pour pallier à ces problèmes plusieurs auteurs [AZK08, GCHS10] préconisent l'utilisation des images hautes dynamiques (HDR). Cette technique consiste à acquérir une série d'images d'un point de vue fixe mais avec un paramètre d'acquisition variable (temps d'obturation, ouverture, sensibilité ISO) afin de couvrir une plus grande plage dynamique d'acquisition. Après fusion, cette série d'image permet d'obtenir une carte des intensités estimées de la scène. Mais comme le montre notre étude, l'utilisation des images HDR est inadaptée au processus de reconstruction 3d par stéréo-photométrie car l'augmentation de la dynamique capturée est trop importante pour les conditions spécifiques d'éclairage utilisée. De plus, la distribution des pas de quantification est irrégulière.

Pour résoudre ce problème, nous proposons une nouvelle méthode d'acquisition nommée AIQ et particulièrement adaptée au contexte d'application de la stéréo-photométrie. Cette nouvelle méthode adapte la dynamique capturée à celle de la scène. De plus, elle permet d'obtenir une meilleure distribution des pas de quantification afin d'approcher l'hypothèse applicative $\mathcal{H}4$. Comme le montrent les résultats obtenus, la création de nouveaux pas et l'affinage de la quantification offerts par les images AIQ réduisent considérablement l'erreur angulaire moyenne des champs de gradients estimés. Ces améliorations permettent d'éliminer les principaux défauts de reconstruction dus à l'imperfection des capteurs numériques de vision.

Conclusion générale

Grâce aux principes de l'optique géométrique et de l'optique ondulatoire, les phénomènes traduisant l'interaction lumière/matière peuvent être mathématiquement modélisables. Ces modèles permettent l'extraction des propriétés géométriques et colorimétriques d'une surface à partir des valeurs d'intensité lumineuse émises ou réfléchies par la surface étudiée. Dans notre contexte, cette extraction puis la reconstruction 3d a pour objectif de faciliter l'application d'algorithmes de mesures métrologiques afin d'analyser les caractéristiques des surfaces, telles que la rugosité, le relief ou les éventuels défauts de surfaces comme des fissures. Le choix de la méthode d'extraction et de reconstruction 3d est régi par les contraintes applicatives du projet Surface Image Matière (SIM) mené au laboratoire. En effet, la méthode doit être basée sur l'utilisation des capteurs photographiques et d'éclairages adaptés, et doit être opérationnelle en environnement de type laboratoire et en environnement de type industriel avec une forte contrainte de déployabilité sur le terrain. De plus, il est impératif qu'elle permette l'analyse de surfaces de tailles importantes ($\geq$ 10 cm^2) en un minimum de temps ($<$ 1 min).

La stéréo-photométrie présente l'avantage de répondre à ces contraintes et permet aussi l'extraction simultanée des propriétés colorimétriques même en cas de forte rugosité. Cet avantage supplémentaire n'est pas négligeable car il offre des perspectives en termes d'analyse (facilite la segmentation) et permet l'acquisition ainsi que la conservation complète de toutes les caractéristiques des surfaces analysées. Néanmoins, comme nous attachons à le mettre en évidence au cours de ce document, l'application de la stéréo-photométrie reste imparfaite. En effet, des contraintes applicatives formulées au chapitre 4 régissent l'ensemble de la chaine d'acquisition 3d. Mais elles ne peuvent pas toutes être respectées dans le cas d'une application réelle. Ce problème entraîne des erreurs de reconstruction significatives. Pour le résoudre, nous proposons différentes méthodes dont le but est d'éliminer ou à l'inverse d'approcher ces contraintes afin d'obtenir des reconstructions 3d fidèles à la réalité terrain.

Au cours du chapitre 4, nous proposons une méthode multi-sources ($\geq$ 6 sources lumineuses) qui permet de reconstruire des surfaces dont la BRDF n'est pas parfaitement Lambertienne. Cette méthode se base sur la probabilité quasi-nulle qu'un même point d'une surface éclairée suivant plusieurs directions différentes présente plus d'une réflexion non diffuse. Ainsi en augmentant le nombre de direction d'éclairage et le nombre d'images acquises, nous augmentons la probabilité qu'en chaque point de la surface un minimum de trois pixels comporte de l'information utile pour l'inversion du modèle de BRDF. Le choix des pixels à conserver est réalisé par traitement numérique et un processus de seuillage automatisé qui ne nécessite pas de paramétrage manuel. Ainsi, la méthode proposée permet de s'affranchir de la connaissance préalable du type de surface analysée et étend l'emploi de la stéréo-photométrie à toutes les surfaces qui présentent une réflexion diffuse pour au moins trois directions d'éclairages différentes.

Au cours de l'étude du système d'éclairage présenté au chapitre 5, nous justifions que les directions des sources lumineuses doivent être parfaitement connues pour inverser un

modèle de BRDF. Cette information primordiale pour appliquer la stéréo-photométrie peut être obtenue par l'une des méthodes de calibration présentées lors de l'état de l'art ou par la méthode que nous proposons si les sources lumineuses sont ponctuelles. Notre méthode se base sur le principe de la réflexion totale, le modèle de PHONG et la géométrie optique. Elle a l'avantage d'estimer la distance séparant une source lumineuse ponctuelle par rapport au centre du plan d'analyse. Même si ce type de source lumineuse est à proscrire lors de l'application de la stéréo-photométrie, il est néanmoins couramment utilisé afin de réduire la taille des systèmes d'acquisition. En effet, les sources lumineuses ponctuelles placées à faible distance des surfaces produisent un éclairage non uniforme et des directions incidentes variables en chaque point. Pour pallier à ce problème, nous proposons une reformulation de l'inversion du modèle Lambertien adaptée à la stéréo-photométrie. Cette nouvelle méthode prend en compte ponctuellement la géométrie du système d'éclairage et de la surface analysée grâce à un schéma itératif. Les reconstructions obtenues sont identiques à celles obtenues avec l'emploi de sources lumineuses directionnelles uniformes. Les défauts dus à la non uniformité d'éclairage sont corrigés.

Ces contributions permettent d'obtenir des systèmes d'acquisition 3d par stéréo-photométrie transportables (taille réduite) et facilement adaptables pour l'analyse de différents types de surfaces. Néanmoins, comme le démontre le chapitre 6, la stéréo-photométrie est limité par l'imperfection des capteurs photographiques car le processus d'acquisition est tributaire de la technologie adoptée. Les limitations de la dynamique d'acquisition entraînent des pertes d'information et perturbent la reconstruction. Ce problème à déjà été soulevé par différents travaux. Mais la solution générale qui consiste à employer des images haute dynamique n'est pas viable dans le contexte d'un éclairage contrôlé. En effet, les méthodes d'acquisition d'images haute dynamique sont uniquement adaptées à des scènes extérieures avec des variations importantes de l'intensité d'éclairage. De plus, cette solution ne permet pas de résoudre les problèmes de reconstruction liés au manque de quantification des intensités de la scène lors de leur conversion vers les valeurs numériques de l'image. Néanmoins, le principe de fusion d'images multi-expositions peut être adapté à la stéréo-photométrie. Pour cela, nous proposons une nouvelle méthode d'acquisition multi-expositions qui permet d'obtenir le meilleur compromis entre la dynamique capturée et la quantification grâce à un choix adapté des ratios entre les images fusionnées. Le processus de fusion reste identique aux méthodes HDR mais nos images AIQ résultantes estiment au plus juste les valeurs réelles d'intensité émises par la scène. Elles permettent de s'approcher des contraintes fixées par l'hypothèse applicative $\mathcal{H}4$ qui stipule que l'inversion d'un modèle de BRDF doit être réalisée à partir des valeurs d'intensité de la scène. L'utilisation des images AIQ dans le processus de reconstruction 3d par stéréo-photométrie réduit considérablement les erreurs liées au manque d'information. Comme l'illustrent les résultats obtenus, la géométrie et la rugosité des surfaces reconstruites sont proches de la vérité terrain.

Les contributions présentées dans ce document permettent une amélioration significative des résultats de reconstruction 3d par stéréo-photométrie. L'apport de chaque méthode proposée est illustré et quantifié de manière indépendante au cours de ce document. Mais les différentes approches proposées peuvent être combinées pour offrir une méthode complète de stéréo-photométrie adaptée à différents contextes applicatifs. Dans nos travaux, le contexte applicatif exposé est principalement axé sur la reconstruction de surfaces rugueuses texturées dont l'interaction lumière/matière répond en majeure partie au modèle diffus. Néanmoins, la réflexion diffuse n'est pas exclusive et comme nous y avons été confrontés, les phénomènes spéculaires ou d'ombres peuvent être corrélés à la rugosité des surfaces analysées.

Les différents résultats obtenus nous amènent à envisager en prolongement de ce travail une généralisation de la prise en compte de la chaine d'acquisition dans l'inversion de modèles d'interaction lumière/matière plus complexes. Comme le montrent nos résultats dans le cas du modèle Lambertien, les améliorations apportées par la caractérisation du système d'éclairage et d'acquisition sont significatifs. Par conséquent, il est envisageable de fournir une méthode basée sur le principe de la stéréo-photométrie dans laquelle les améliorations apportées sont incluses et où le modèle photométrique à inverser est générique. Ainsi, elle est adaptable à la nature des surfaces analysées. Cette perspective permet d'étendre l'usage de la stéréo-photométrie à différents types de surfaces jusqu'à présent non envisagées du fait de l'insuffisance du modèle Lambertien pour traduire les phénomènes d'interaction lumière/matière dans ce cas. Parmi les types de surfaces envisagés, nous pouvons citer l'analyse des surfaces à fort coefficients de réflexion spéculaire.

Cependant les perspectives de ces travaux ne se limitent pas seulement à la reconstruction 3d. En effet, l'acquisition d'image est primordiale pour un grand nombre de traitement. Mais les caractéristiques du système d'acquisition sont rarement prises en compte. Ainsi, il est envisageable d'introduire dans les algorithmes de traitement, des modèles d'interactions lumière/matière ainsi que des modèles physiques et réalistes du système d'acquisition. Cette voie, introduite par notre principe d'acquisition d'image AIQ, est à développer. En effet, notre étude du principe d'acquisition des images HDR révèle que l'estimation des valeurs d'intensité réelles est inexacte et ne prend pas en compte le bruit ou les éventuelles distorsions liées aux capteurs d'images. Inclure une modélisation physique du capteur et développer un traitement d'exclusion du bruit dans le principe de fusion est une perspective d'amélioration majeure du principe d'acquisition HDR. Cette voie, n'est qu'un des multiples exemples d'amélioration du traitement d'image par la prise en compte de la chaine globale d'acquisition.

Annexes

A

Sommaire

A.1 Tableau de correspondance Radiométrie-Photométrie . . . 158
A.2 Nos systèmes d'acquisition . . . 159
A.2.1 NIKON_1 . . . 159
A.2.2 NIKON_2 . . . 160
A.3 Étude comparative des méthodes d'intégration numérique . . . 161
A.3.1 Comparaison 1 : qualité des reconstructions . . . 162
A.3.2 Comparaison 2 : robustesse face au bruit . . . 165
A.3.3 Comparaison 3 : temps de calcul . . . 165
A.4 Étude comparative des méthodes de calibration des positions des sources 167
A.4.1 Étude des méthodes basées sur l'analyse statistique des intensités . . . 167
A.4.2 Etude des méthodes basées sur l'analyse des ombres . . . 175
A.4.3 Etude des méthodes basées sur l'analyse des points spéculaires . . . 177

A.1 Tableau de correspondance Radiométrie-Photométrie

Radiométrie					Photométrie				
Terme FR	***Terme SI***	***Unité SI***		***Informations***	***Terme FR***	***Terme SI***	***Unité SI***		***Informations***
Énergie	Radiant	Qe	$Joules(J)$	énergie au sens classique du terme	**Quantité de lumière**	Luminous energy	Qv	$lm.s$	produit du flux lumineux par sa durée
Flux énergétique *Puissance (radiante)*	Radiant flux	Fe ou Φe	$Watt(W)$	une puissance au sens commun du terme	**Flux lumineux** *Puissance lumineuse*	Luminous flux	Fv ou Φv	lm	unité historique dont découle toutes les autres
Éclairement énergétique	Irradiance	Ee	$W.m^{-2}$	la puissance rapportée à l'unité de surface réceptrice	**Éclairement lumineux**	Illuminance	Ev	lux ou $lm.m^{-2}$	flux lumineux reçu par une surface
Exitance *Émittance (énergétique)*	Radiant exitance	Me	$W.m^{-2}$	la puissance rapportée à l'unité de surface émettrice	**Exitance lumineuse** *Émittance (lumineuse)*	Luminous emittance	Mv	lux ou $lm.m^{-2}$	flux lumineux émis par une surface
Intensité rayonnante *(énergétique)*	Radiant intensity	Ie	$W.sr^{-1}$	la puissance rapportée à l'unité d'angle solide	**Intensité lumineuse**	Luminous intensity	Iv	cd ou $lm.sr^{-1}$	flux lumineux émis par une source et vu par une surface unitaire
Radiance *(Rayonnance)*	Radiance	Le	$W.sr^{-1}.m^{-2}$	la puissance rapportée à l'unité d'angle solide et à l'unité de surface apparente	**Luminance** *Brillance*	Luminance	Lv	$cd.m^{-2}$	perception visuelle du flux lumineux
Exposition	Radiant exposure	He	$J.m^{-2}$	correspond au produit de l'éclairement et du temps d'exposition	**Exposition lumineuse** *Lumination*	Luminous exposure	Hv	$lux.s$	terme déprécié
Exitance énergétique spectrale	Radiant energy density	ωe	$W.m^{-3}$	l'exitance énergétique rapportée à l'unité de largeur de bande spectrale	***	Luminous energy density	ωv	$lm.s.m^{-3}$	****
Radiosité	Radiosity	Je	$W.m^{-2}$	terme utilisé en simulation d'éclairage					

TABLE A.1 – *Tableau de correspondance des unitées Radiométrie-Photométrie*

A.2 Nos systèmes d'acquisition

A.2.1 NIKON_1

Le système NIKON_1 est un prototype à paramètres variables, il se compose d'un appareil photo reflex NIKON D300 et de deux objectifs interchangeables de 105 et 150 *mm*. La distance entre l'objectif et la surface varie de 90 à 130 *cm* en fonction des besoins. Le système d'éclairage est composé initialement de 6 lampes à décharge identiques et d'intensité constante, elles sont placées le long d'un cercle dont le centre est l'appareil photo. Dans la configuration initiale les angles azimuts sont (0^o, 60^o, 120^o, 180^o, 240^o, 300^o) pour des angles zéniths de 30^o et des distances lampes/surface de 150 *cm*.

Systèmes	NIKON_1		
	initiale	min	max
Nombre de source	6	1	15
Type d'éclairage	Semi-Ponctuelle	-	-
Type de source	Lampe à décharge	-	-
Distance Sources/Sol (cm)	150	50	200
Distance Sources/Centre (cm)	173.20	50	250
Azimut (deg)	0, 60, 120, 180, 240, 300	0	360
Zenith (deg)	30	5	75
APN (marque/ref)	Nikon D300	-	-
Taille capteur (mm)	15.8 x 23.6	-	-
Résolution (pixel)	4320 x 2868	2160x1434	4320 x 2868
Pas acquis (µm)	33 x 33	33 x 33	135 x 135
Zone acquise (mm)	141.6 x 94.8	141.6 x 94.8	292.2 x 195.65
Distance APN/Sol (cm)	90	90	130
Objectif (marque)	Sigma EX DG F2.8	Nikon AF-S VR F2.8	Sigma EX DG F2.8
Objectif (focal)	150mm	105mm	150mm
Rapport Distance :	1732.0 / 141.6 = 12.23	500/292.2= 1.71	2500/141.6= 17.65
Hypothèse d'Uniformité	Valide car r>10	Non valider<10	Valide car r>10

Table A.2 – *Informations système NIKON_1*

A.2.2 NIKON_2

Le système d'acquisition NIKON_2 est un prototype à paramètre fixe, le système d'acquisition et d'éclairage est inclus dans une chambre noire de 0.25 m^3 afin de permettre des acquisitions en extérieur sans subir les phénomènes de lumière ambiante. L'appareil photo est un NIKON D300s équipé d'un objectif NIKON AF-S VR F2.8 de 105 *mm*. Il est situé à une distance de 45.8 *cm* par rapport à la surface analysée. Le système d'éclairage est composé de 3 Flashs NIKON SB 900 disposés en cercle autour de l'appareil photo. Ils sont distribués selon des angles azimuts de 0^o, 120^o et 240^o pour un angle zénith de 20.55^o et une distance source/surface de 42.72 *cm*. Ces flashs ont la capacité de faire varier leurs intensités par pas de 0.8 suivant 15 valeurs différentes.

Systèmes	NIKON_2
Nombre de source	3
Type d'éclairage	Ponctuelle
Type de source	Flashs NIKON SB 900
Distance Sources/Sol (cm)	40
Distance Sources/Centre (cm)	42.72
Rayon Source/APN (cm)	15
Azimut (deg)	0, 120, 240
Zenith (deg)	20.55
APN (marque/ref)	Nikon D300s
Taille capteur (mm)	15.8 x 23.6
Résolution (pixel)	1434x2160
Résolution acquise (µm)	42 x 42
Zone acquise (mm)	103 x 68.9
Distance APN/Sol (cm)	45.8
Objectif (marque)	Nikon AF-S VR F2.8
Objectif (focal)	105mm
Rapport Distance :	427.2/89.9 = 4.75
Hypothèse d'Uniformité	NON car r<10

Table A.3 – *Informations système NIKON_2*

A.3 Étude comparative des méthodes d'intégration numérique

Plusieurs comparatifs des méthodes d'intégration ont déjà été proposées [Danoo, Moro5] mais nous proposons notre propre étude afin de définir les avantages et les inconvénients de chacune. Pour cette étude sont implémentées les méthodes locales trapèze, rectangle, Coleman et Jain, Rodehorst et Wu-Li. Pour les méthodes basées sur la relaxation, sont implémentées Horn, Wu-Li avec Horn et SOR. Les méthodes globales sont représentées par Frankot et Chellappa. Ces implémentations sont réalisées en C^{++} sur un PC de type Pentium i7 64 Bits et 8 Go de mémoire vive.

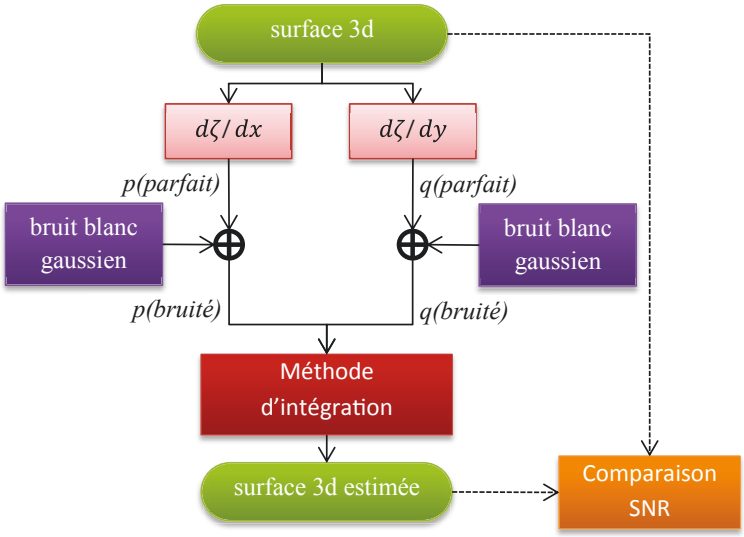
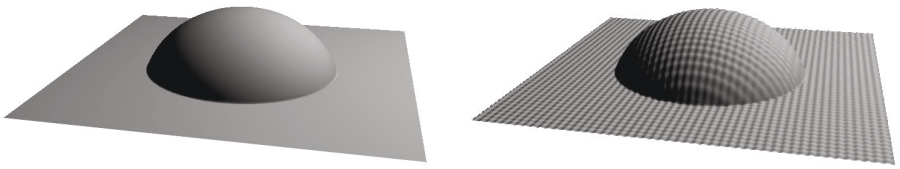


Figure A.1 – *Schéma du protocole de comparaison des méthodes d'intégration.*

Le protocole de comparaison (figure A.1) est similaire à celui de [Moro5]. Nous étudions le temps de calcul nécessaire à l'intégration pour des champs de gradient de petite taille (800×800 *pixels*), la qualité des reconstructions et la robustesse au bruit.

Les surfaces utilisées sont des demi-sphères de synthèse, de diamètre 480 *pixels* et posées sur un plan plat de résolution 800×800 *pixels*. Les champs de gradients initiaux sont directement calculés à partir des surfaces de synthèses et considérés comme la vérité terrain afin de comparer uniquement le processus d'intégration. La première demi-sphère (figure A.2a) est lisse, la seconde (figure A.2b) présente une surface micro-rugueuse obtenue par l'ajout d'une fonction sinusoïdale de fréquence 16 *pixels* suivant les axes x et y.



(a) (b)

Figure A.2 – *Surfaces utilisées pour la comparaison des méthodes d'intégration, (a) demi-sphère de diamètre* 480 *pixels posée sur un plan de résolution* 800×800 *pixels, (b) même demi-sphère avec ajout d'une fonction sinusoïdale de fréquence* 16 *pixels suivant les lignes et les colonnes afin de simuler une micro-rugosité.*

Pour évaluer la robustesse le même protocole est utilisé mais les champs de gradients initiaux de la demi-sphère lisse sont dégradés (figure A.3) par un bruit blanc gaussien de variance 0.05 puis 0.1 soit respectivement 12% et 25% de la valeur maximale des champs de gradient qui s'élève à 0.4 *mm*.

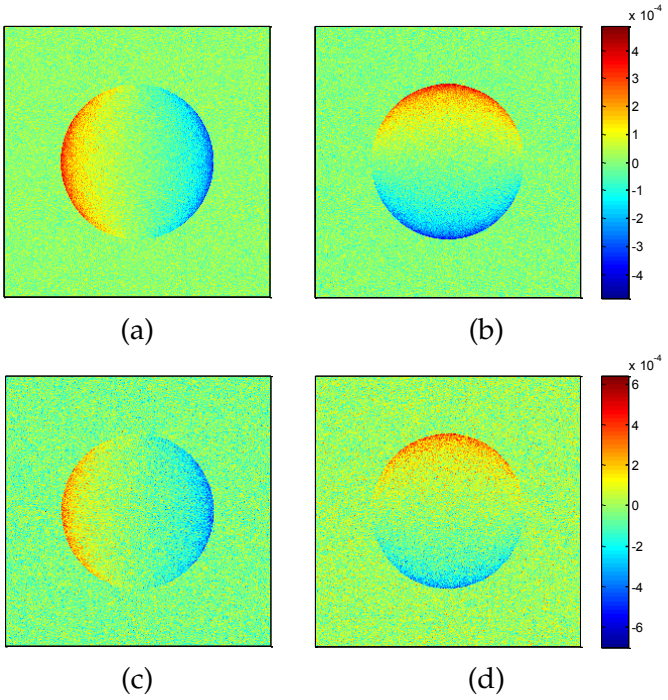


FIGURE A.3 – *Champs de gradient p et q perturbés par une bruit blanc gaussien de variance (a-b)* 12% *et (c-d)* 25% *de la valeur maximale des champs de gradient non bruités.*

La fidélité des reconstructions est analysée par le calcul du rapport signal sur bruit SNR (Signal to Noise Ratio) entre les surfaces initiales ζ et les surfaces estimées $\hat{\zeta}$ après intégration, suivant la formule :

$$SNR = 10 \log_{10} \left(\frac{\frac{1}{U \times V} \sum\sum \zeta(u,v)^2}{MSE} \right) \tag{A.1}$$

où MSE correspond à l'erreur quadratique moyenne entre les surfaces $\hat{\zeta}$ et ζ de résolution $U \times V$ et où $\hat{\zeta}$ est considérée comme la vérité bruitée de ζ :

$$MSE = \frac{1}{U \times V} \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} \left\| \zeta(u,v) - \hat{\zeta}(u,v) \right\|^2 \tag{A.2}$$

A.3.1 Comparaison 1 : qualité des reconstructions

Les résultats SNR sont présentés figure A.4 pour la demi-sphère lisse et figure A.8 pour la demi-sphère micro-rugueuse.

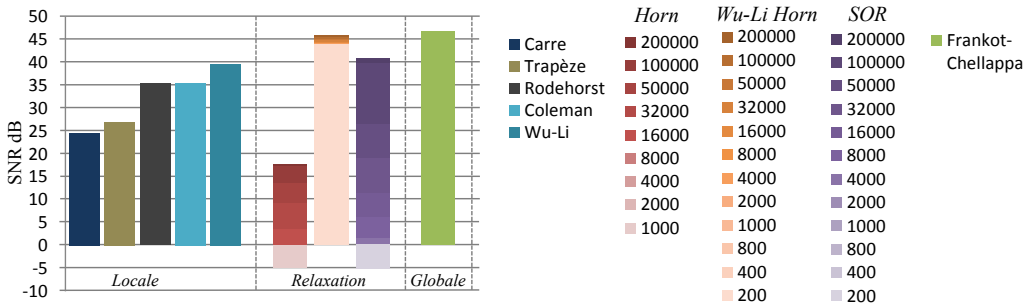


FIGURE A.4 – *Comparaison des reconstructions 3d (SNR) en fonction des méthodes d'intégration pour la demi-sphère lisse.*

Dans le cas de la surface lisse, l'ensemble des méthodes locales permettent d'obtenir des SNR supérieurs à $25dB$, les méthodes carré et trapèze offrent les moins bons résultats à cause de la propagation de l'erreur d'intégration. Pour ce qui est des méthodes basées sur la relaxation, le nombre d'itérations définit la qualité des reconstructions (figure A.5).

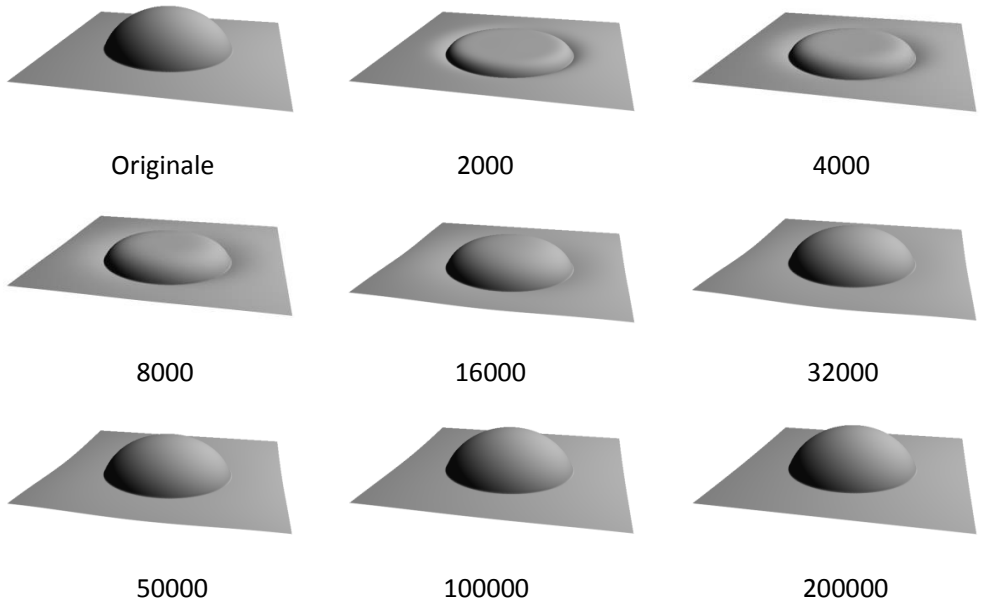


FIGURE A.5 – *Évolution de l'estimation de la surface 3d par intégration numérique suivant la méthode de* HORN *basée sur la relaxation.*

Lors des tests, nous avons fixé ce nombre à 200000 itérations maximum. Nous pouvons constater (figure A.6) que la convergence de la méthode SOR est deux fois plus rapide que HORN. En effet, en partant d'une surface initiale nulle et après 200000 itérations la méthode SOR permet d'obtenir un SNR supérieur à $40dB$ et la méthode de HORN un SNR inférieur à $20dB$. Néanmoins la convergence de ces méthodes est très lente, mais une initialisation par des valeurs non nulles (la méthode WU-LI HORN par exemple) permet d'atteindre très rapidement un niveau de SNR élevé. Dans l'exemple présenté figure A.6, il atteint $40dB$ dès l'initialisation, puis après 200 itérations il est supérieur à $43dB$.

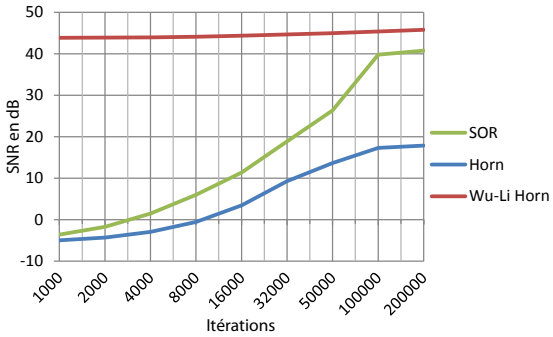
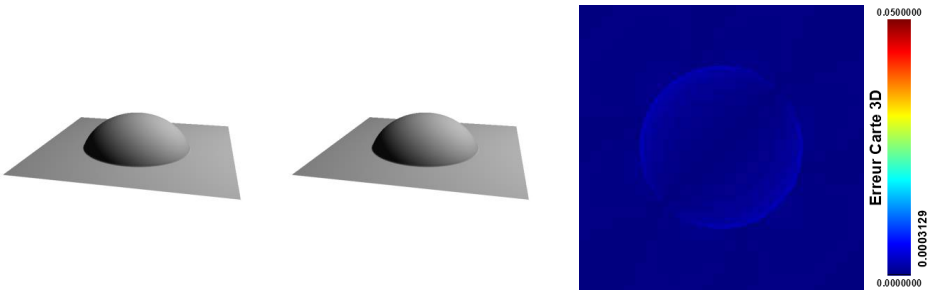


FIGURE A.6 – *Comparaison du nombre d'itérations par rapport au SNR pour les méthodes basées sur la relaxation.*

Enfin la méthode globale offre le meilleur résultat, en effet elle permet d'atteindre un SNR de $46,5dB$ et la surface reconstruite est visuellement identique à la surface initiale (figure A.7).



(a) (b) (c)

FIGURE A.7 – *Erreur d'intégration par la méthode globale proposée par* FRANKOT *et* CHELLAPPA, *(a) surface originale en 3d, (b) estimation 3d par intégration numérique, (c) erreur entre la surface originale et l'estimation.*

Cette analyse est confirmée par les résultats (figure A.8) obtenus pour la surface micro-rugueuse. Malgré une légère chute ($< 2dB$), les méthodes WU-LI HORN ainsi que FRANKOT et CHELLAPPA conservent des niveaux de SNR largement supérieurs aux autres. Par conséquent, seules les analyses de robustesse face au bruit et du temps de calcul peuvent les départager.

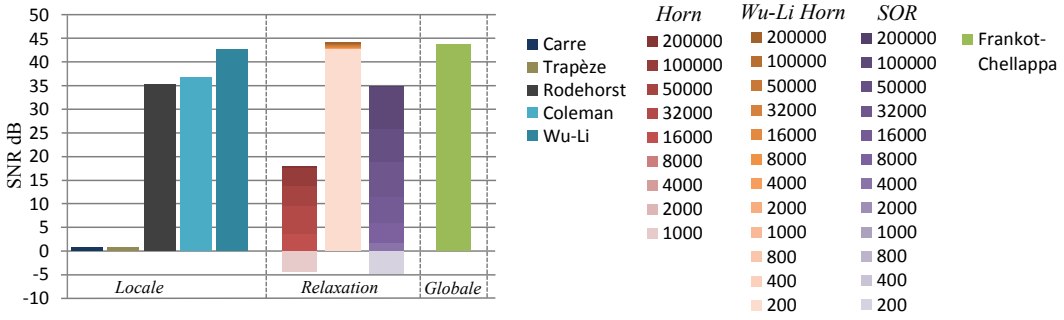


FIGURE A.8 – *Comparaison des reconstructions 3d (SNR) en fonction des méthodes d'intégration pour la demi-sphère micro-rugueuse.*

A.3.2 Comparaison 2 : robustesse face au bruit

Avec l'ajout du bruit les chutes des SNR sont plus importantes (figure A.9 et A.10).

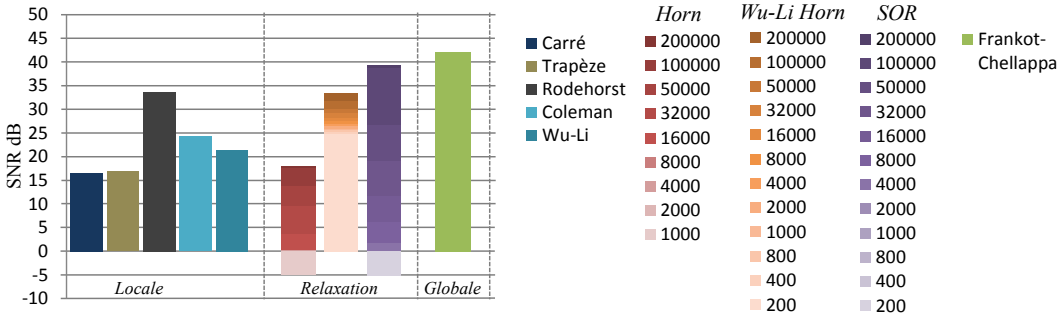


FIGURE A.9 – *Comparaison des reconstructions 3d (SNR) en fonction des méthodes d'intégration pour la demi-sphère lisse et la perturbation des champs de gradients avec un bruit blanc gaussien de variance* 12%.

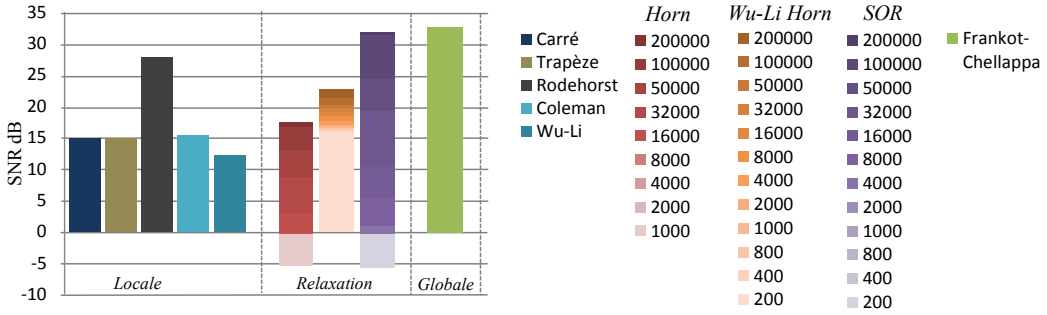


FIGURE A.10 – *Comparaison des reconstructions 3d (SNR) en fonction des méthodes d'intégration pour la demi-sphère lisse et la perturbation des champs de gradients avec un bruit blanc gaussien de variance* 25%.

Parmi les méthodes locales, seule celle de RODEHORST maintient un SNR acceptable, $33dB$ pour une variance de 12% et $27,8dB$ pour une variance de 25%. La chute de la méthode WU-LI entraîne aussi celle de la méthode de relaxation WU-LI HORN, le SNR est de $22,8dB$ malgré les 200000 itérations. La méthode SOR suit les mêmes tendances et maintient des SNR élevés après un grand nombre d'itérations. La méthode globale enregistre une légère chute pour un niveau de bruit important mais elle est la plus robuste comparée aux autres (SNR de $42,3dB$ pour un bruit de variance 25%).

A.3.3 Comparaison 3 : temps de calcul

La plupart des méthodes locales nécessitent un balayage de l'ensemble $U \times V$ *pixels* de la carte des hauteurs estimé ce qui les rend extrêmement rapides (tableau A.4). La méthode globale nécessite plusieurs transformées de Fourrier mais grâce aux algorithmes de FFT (Fast Fourrier Transform) le temps d'intégration équivaut à un balayage complet de la carte des hauteurs. Par conséquent elle est aussi rapide que les méthodes locales.

Méthodes	Temps (s)	Itérations	SNR (dB)
Carré	0.421	unique	24.64
Trapèze	0.421	unique	27.02
Rodehorst	0.421	unique	35.39
Coleman	0.421	unique	35.57
Wu-Li	0.421	unique	39.68
Horn	15.834	2000	-4.29
	124.535	16000	3.46
	1557.520	200000	17.84
Wu-Li + Horn	2.059	200	43.79
	18.689	2000	43.86
	147.654	16000	44.33
	1787.630	200000	45.75
SOR	1.638	200	-5.27
	12.980	2000	-1.71
	105.066	16000	11.36
	1333.840	200000	40.72
Frankot-Chellappa	0.421	unique	46.54

TABLE A.4 – *Comparaison des temps d'intégration entre les différentes méthodes pour la demi-sphère de synthèse de résolution* 800×800 *pixels.*

Il en est tout autre pour les méthodes basées sur la relaxation, chaque itération équivaut à un balayage de la carte. Ainsi, le nombre d'itérations définit le temps d'intégration. Comme nous l'avons constaté lors des comparaisons 1 et 2, le nombre d'itérations doit être très élevé pour que ces méthodes atteignent des qualités de reconstruction rivalisant avec les autres méthodes. Ainsi les temps d'intégration sont très élevés, de l'ordre de plusieurs minutes (figure A.11). Par exemple, la méthode SOR pour 200000 itérations nécessite $21,42$ *minutes* pour une carte de résolution 800×800 *pixels* et un SNR qui dépasse tout juste la meilleure méthode locale sans atteindre celui de la méthode globale.

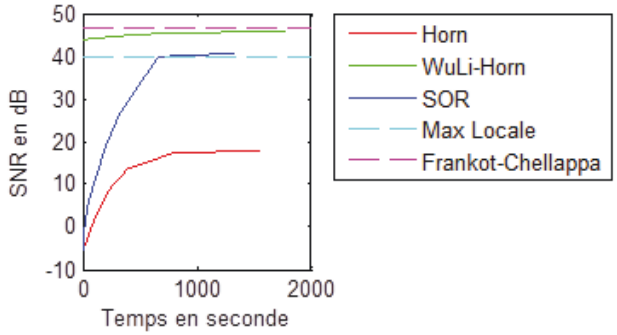


FIGURE A.11 – *Comparaison des temps d'intégration en fonction des SNR atteint pour les méthodes basées sur la relaxation, les valeurs des SNR Max Locale et* FRANKOT *et* CHELLAPPA *sont prolongés à titre de comparaison car le temps d'intégration est trop faible pour être visible.*

De plus, la résolution influe sur le temps, plus elle est grande plus le temps nécessaire est important. Nous ne présentons pas d'analyse car l'écart est déjà significatif entre les méthodes de relaxation et les autres.

A.4 Étude comparative des méthodes de calibration des positions des sources

Au cours de cette étude, nous évaluons les performances des méthodes de calibration en fonction de leur principe de base. Chaque méthode est employée sur des images d'une ou plusieurs demi-sphères de synthèse dont les géométries sont connues. Ces différentes images sont obtenues par un logiciel de lancer de rayons (POV-Ray). La direction des sources lumineuses est paramétrable suivant les deux angles zénith et azimut. Les sources peuvent être suffisamment éloignées de la surface pour être considérées comme directionnelles ou dans le cas contraire ponctuelles.

A.4.1 Étude des méthodes basées sur l'analyse statistique des intensités

Afin d'évaluer les estimateurs nous effectuons différentes séries de test :

- la première série (*azimut1*, *azimut2*) permet d'évaluer les performances d'estimation de l'angle azimut car seul l'angle azimut varie (de 0^o à 360^o par pas de 1^o) pour un angle zénith de 15^o ou 45^o.
- la seconde série (*zenith1*, *zenith2*, *zenith3*) permet d'évaluer les performances d'estimation de l'angle zénith car il varie (de 4^o à 90^o par pas de 1^o) pour un angle azimut de 45^o, 266^o et 0^o.
- la troisième série (*schema1*, *schema2*) permet une évaluation générale car les deux angles varient selon un schéma préalablement défini (figure A.12).

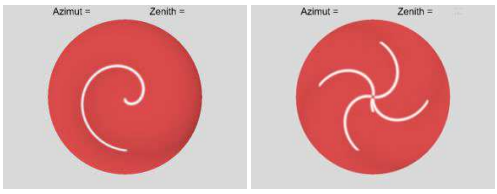
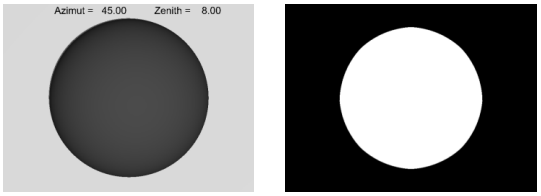


FIGURE A.12 – *Schémas d'éclairages (à gauche) schema1, (à droite) schema2.*

Comme les méthodes basées sur l'analyse statistique des intensités peuvent uniquement déterminer les angles azimuts et zéniths, les sources lumineuses utilisées sont directionnelles.



(a) (b)

FIGURE A.13 – *Élimination des dérivées nulles, (a) exemple image de synthèse, (b) masque.*

Dans [Pen82], PENTLAND compare les performances de son estimateur par rapport aux estimations de l'homme, mais ne fournit pas une analyse des performances dans des conditions réelles. Comme des travaux postérieurs le montrent, le calcul des angles zéniths proposé ne donne pas d'estimations précises dans des cas réels. Si toutes les hypothèses de l'algorithme sont présentes dans un cas de test idéal (comme l'image d'une sphère), l'angle zénith doit

être estimé avec une grande précision. Lee et Rosenfeld [LR85] ainsi que Chojnacki et Al [CBG94] soulignent que l'estimation de l'angle zénith de Pentland ne permet pas d'obtenir un résultat correct. Lors de notre implémentation, nous prenons soin d'extraire par un masque (figure A.13b) la surface dont la somme des dérivées en hauteur est nulle afin de satisfaire les contraintes imposées par la méthode.

D'après les résultats (tableau A.5 et figure A.14) obtenus l'estimateur fournit des résultats correctes pour l'angle azimut l'erreur moyenne est d'environ 1^o et elle ne dépasse généralement pas 2^o. Par contre l'estimation de l'angle zénith ne fournit pas de résultats acceptables. Malgré des bonnes estimations dans de rare cas, l'erreur d'estimation moyenne est élevée $> 20^o$.

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2
ε azimut °	-0.05	-0.86	-0.90	-0.01	-0.003	-0.048	-0.13
etype	0.15	0.95	0.99	2.45	0.68	1.72	2.29
ε zenith °	-5.09	-9.93	-10.67	-2.63	-18.61	-17.46	-15.60
etype	29.91	19.33	26.82	1.39	3.45	16.21	14.00

Table A.5 – Pentland, *erreur moyenne d'estimation et l'écart type pour chaque série d'images tests.*

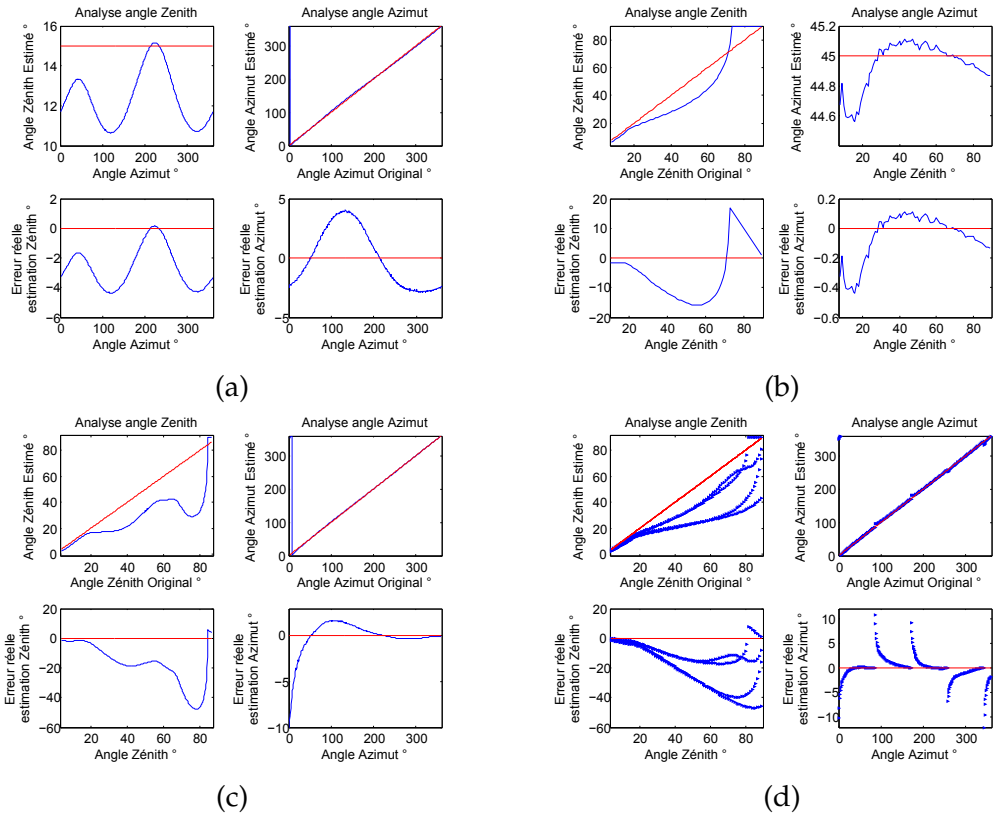


Figure A.14 – Pentland *(bleu), idéal (rouge), graphiques résultats (a) pour azimut* $(0 - 360^o)$ *et zénith* 15^o*, (b) azimut* 45^o *et zénith* $(4 - 89^o)$*, (c) pour éclairage schéma 1, (d) pour éclairage schéma 2.*

De plus nous pouvons remarquer que l'erreur est fortement corrélée avec l'angle azimut et aucun résultat ne fournit une bonne estimation lorsque l'angle zénith dépasse 40^o. Plusieurs études de l'estimateur sont arrivées aux mêmes conclusions : Funk dans [Fun04], Gibbins et Al dans [GBC91] ou Zheng et Chellappa dans [ZC91]. Les analyses menées montrent que l'estimateur ne doit être utilisé que sur des surfaces sphériques ou ellipsoïdes parfaitement

Lambertiennes, de plus il existe une corrélation entre la résolution image et l'erreur d'estimation. Plus le rayon de la sphère (ou demi-sphère, ou ellipsoïde) est important, plus l'erreur d'estimation de l'angle zénith augmente (figure A.15).

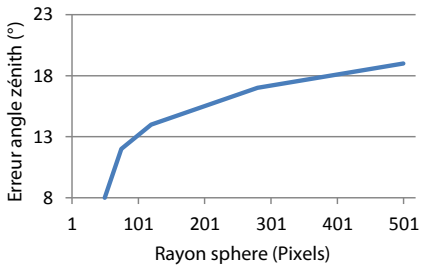
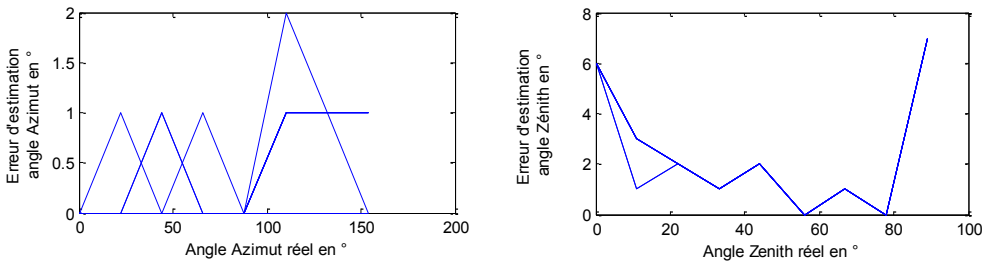


FIGURE A.15 – *Résultats de* FUNK *[Funo4] illustrant la corrélation entre la résolution des formes utilisées (ordonnés Pixel) et l'erreur d'estimation de l'angle zénith (abscisse degré) par l'algorithme de* PENTLAND.

Les résultats fournis (figure A.16) LEE et ROSENFELD dans [LR85], sont réalisés pour une sphère Lambertienne éclairée suivant 8 angles azimuts et 9 angles zéniths différents.



(a) (b)

FIGURE A.16 – *Résultats de* LEE *et* ROSENFELD *dans [LR85], (a) erreur angle azimut, (b) erreur angle zénith.*

Ces test restreints montrent une bonne estimation des angles azimuts ($\leq 2^o$ maximum) et une erreur limitée à 7^o pour l'estimation des angles zéniths. D'autres résultats présentent une erreur d'estimation maximum de 22^o pour l'angle zénith mais la surface utilisée est une ellipsoïde. D'après les auteurs, l'estimateur est plus performant que celui de PENTLAND. L'erreur d'estimation de l'angle zénith est fortement réduite (maximum 7^o alors que PENTLAND maximum 34^o). Les résultats obtenus lors de notre implémentation (tableau A.6 et figure A.17) démontrent que l'estimation de l'angle azimut peut être faite directement à partir des dérivées premières des intensités images suivant les directions x et y. L'erreur d'estimation azimut est identique à celle obtenue avec PENTLAND mais la complexité est réduite. La méthode proposée pour l'estimation de l'angle zénith apporte une forte amélioration, l'erreur d'estimation est réduite pour les angles zéniths supérieurs à 40^o et la corrélation entre l'erreur et l'angle azimut est limitée.

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2
ε azimut °	-0.05	-0.86	-0.89	-0.007	-0.001	-0.09	-0.14
etype	0.15	0.95	1.00	2.54	0.55	1.67	2.26
ε zenith °	-2.21	-2.25	-2.17	3.98	-1.05	-1.14	-1.30
etype	5.24	5.26	5.23	0.03	0.00	4.94	5.81

TABLE A.6 – LEE *et* ROSENFELD, *erreur moyenne d'estimation et l'écart type pour chaque série d'images tests.*

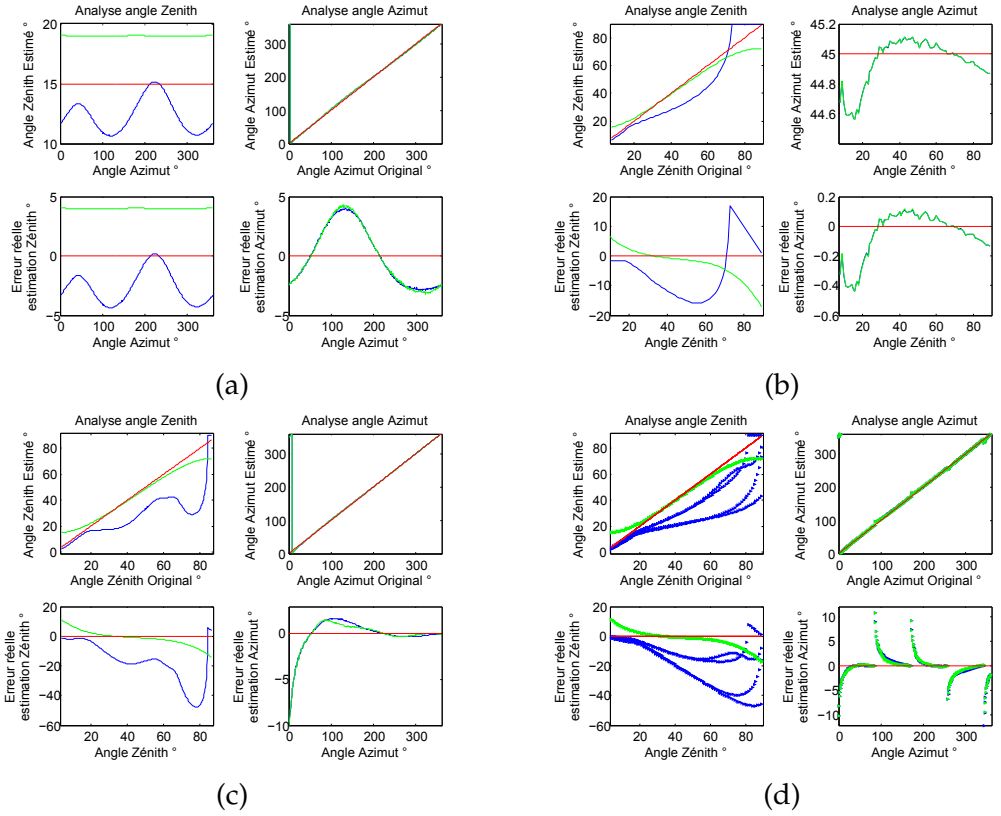


FIGURE A.17 – LEE *et* ROSENFELD *(vert),* PENTLAND *(bleu), idéal (rouge), graphiques résultats (a) pour azimut* $(0 - 360^o)$ *et zénith* 15^o*, (b) azimut* 45^o *et zénith* $(4 - 89^o)$*, (c) pour éclairage schéma 1, (d) pour éclairage schéma 2.*

ZHENG et CHELLAPPA font une étude détaillée des performances de leurs méthodes, ils se comparent à la littérature ([Pen82] et [LR85]) pour des images avec ou sans bruits d'une sphère, ainsi que des images réelle (lenna, pepper, etc). D'après leurs résultats :

- Les méthodes d'estimation de l'angle azimut de PENTLAND, LEE et ROSENFELD ainsi que *"votting"* sont équivalentes pour une sphère parfaite, la méthode *"contour"* est trop dépendante de la bonne détection des contours et souffre d'imperfection selon la valeur de l'angle zénith. Les méthodes *"contour"* et *"votting"* semblent plus robustes au bruit et plus adaptées à des objets ellipsoïdaux.
- La nouvelle formulation de l'estimateur d'angle zénith semble apporter des améliorations par rapport aux autres méthodes surtout lorsque la valeur d'angle est importante (proche 90^o). La méthode de PENTLAND est très sensible au bruit ou à l'orientation d'une forme ellipsoïde alors que les méthodes LEE et ROSENFELD ainsi que celles proposées sont plus robustes à ce type de perturbations.
- L'estimation des positions des sources lumineuses sur images réelles semble fournir de meilleurs résultats avec les méthodes *"contour"* et *"votting"*, mais l'interprétation des résultats fournis est difficile car d'une part les méthodes proposées ne sont pas conçues pour être utilisées sur n'importe quel type de surface et d'autre part les positions réelles des sources lumineuses ne sont pas données.

Nous avons choisi d'implémenter uniquement la méthode *"voting"* pour estimer l'angle azimut. Ce choix permet une comparaison directe avec les méthodes PENTLAND ainsi que LEE

et Rosenfeld. Pour obtenir le système récursif, nous choisissons aléatoirement une série de point dont les directions par rapport au centre de la sphère sont équitablement distribuées (figure A.18).

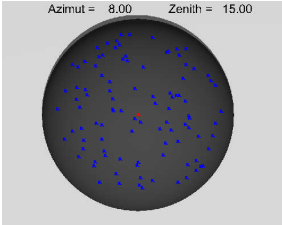
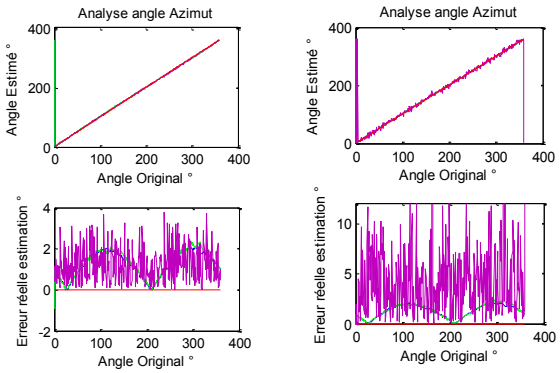


Figure A.18 – Zheng *et* Chellappa, *100 Points choisis aléatoirement dans toutes les directions pour construire le système récursif de la méthode 'voting'.*

Le nombre de points choisis influence les résultats d'estimation. Théoriquement 2 points suffisent à résoudre le système. Mais plus le nombre de point augmente, moins l'erreur d'estimation est importante. L'estimation de l'angle azimut est obtenue au facteur π près, ce facteur dépend des signe de $C(1)$ et $C(2)$. L'implémentation de l'estimateur d'angle zénith est identique à Lee et Rosenfeld car seule la fonction monolithique est modifiée.

Dans notre implémentation les mêmes séries d'images sont utilisées pour obtenir les résultats. L'estimation de l'angle azimut est fortement dépendante du nombre de points choisis (figure A.19), avec notre configuration l'erreur moyenne est proche de 5^o pour 1000 points et elle avoisine 2^o pour 10000 points. Avec un nombre de points suffisant l'erreur d'estimation de l'angle azimut est équivalente à la méthode de Pentland ou de Lee et Rosenfeld.



(a) (b)

Figure A.19 – Zheng *et* Chellappa, *comparaison de l'influence du nombre de point choisis pour l'estimation de l'angle azimut, (a)* 10000 *points, (b)* 1000 *points.*

Malgré les performances annoncées, l'estimation de l'angle zénith présente un biais très important. Après plusieurs remises en question de notre implémentation et du modèle utilisé pour produire les images de synthèse, nous ne sommes pas parvenus à un résultat satisfaisant. Le calcul de $\mathbb{E}(M)/\sqrt{\mathbb{E}(M^2)}$ ne permet pas d'obtenir la dynamique de la fonction d'estimation. Les résultats (figure A.20) présentés dans [GBC91], sont en concordance avec ceux obtenus lors de notre implémentations (tableau A.7 et figure A.21). Par conséquent soit la méthode présentée par Zheng et Chellappa est erronée, soit une astuce, un paramètre ou une condition n'est pas présentée dans leur document.

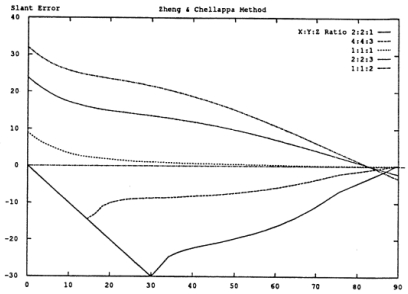


FIGURE A.20 – *Erreur d'estimation de l'angle zénith par la méthode de* ZHENG *et* CHELLAPPA. *Ces résultats sont obtenus dans [GBC91] pour des surfaces sphériques et ellipsoïdes. Attention l'angle zénith est inversé par rapport à notre implémentation (0^o <=> 90^o).*

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2
ε azimut °	0.02	-0.08	-0.26	-0.00	-0.01	0.17	-0.06
etype	0.77	0.57	0.80	1.34	0.42	1.12	1.25
ε zenith °	-16.27	-16.29	-16.25	0.13	-12.60	-14.26	-14.46
etype	13.23	13.23	13.23	0.03	0.00	12.68	13.81

TABLE A.7 – ZHENG *et* CHELLAPPA, *erreur moyenne d'estimation et l'écart type pour chaque série d'images tests.*

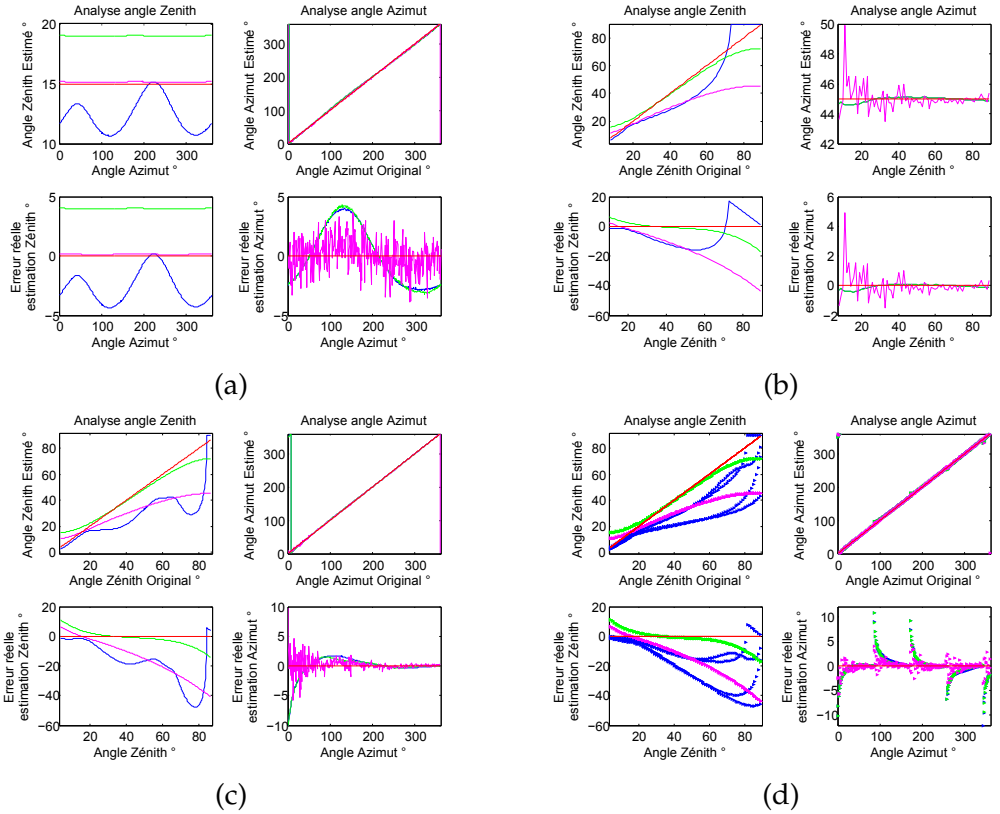
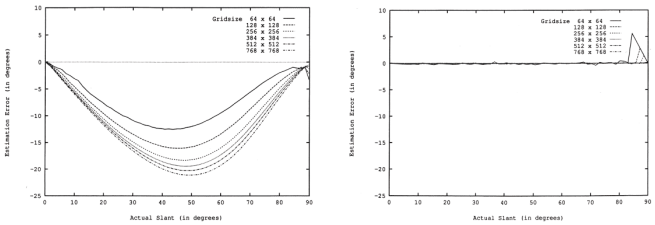


FIGURE A.21 – ZHENG *et* CHELLAPPA *(magenta),* LEE *et* ROSENFELD *(vert),* PENTLAND *(bleu), idéal (rouge), graphiques résultats (a) pour azimut (0 – 360^o) et zénith 15^o, (b) azimut 45^o et zénith (4 – 89^o), (c) pour éclairage schéma 1, (d) pour éclairage schéma 2.*

Dans leur version revisitée de l'estimateur de PENTLAND,CHOJNACKI et BROOK annoncent que l'erreur d'estimation de l'angle azimut reste identique, par contre les performances d'esti-

mation de l'angle zénith sont considérablement accrues. Les résultats présentés (figure A.22) pour une sphère Lambertienne illustre l'amélioration apportée.



(a) (b)

FIGURE A.22 – *Comparaison de l'erreur d'estimation de l'angle zénith pour la méthode de* PENTLAND *(a) et la méthode "disk method" (b) de* CHOJNACKI *et* BROOK. *Ces résultats sont issus de [CBG94].*

Notre implémentation change peu car pour PENTLAND nous utilisons déjà un masque circulaire pour ne conserver que les points de la demi-sphère. La taille du disque semble être un paramètre important mais malheureusement les auteurs ne font aucune précision. Deux implémentations peuvent être réalisées :

- le rayon du disque reste fixe quel que soit l'ombre sur la demi-sphère.
- le rayon du disque varie en fonction de l'ombre présente sur la demi-sphère.

Nous avons donc procédé aux deux implémentations, la première implémentation est la plus simple nous avons fixé arbitrairement le rayon du disque, plusieurs valeurs $a = [0.25; 0.5; 0.75; \ldots]$ ont été testées. Pour la seconde implémentation le rayon du disque est estimé en fonction de la détection de l'ombre propre sur la demi-sphère. Pour détecter la limite de l'ombre propre nous utilisons un filtre détecteur de contour de type *"Canny"* ou *"Passe bas"*. L'estimation de l'angle zénith est par conséquent dépendante de la détection de l'ombre propre. Comme le démontrent les résultats (figure A.23), le choix du rayon du disque influx fortement sur les résultats d'estimation d'angle zénith. Une valeur a constante n'améliore pas les résultats car les ombres propres de la demi-sphère ne sont pas prises en compte. Lorsque α évolue en fonction de la limite de l'ombre propre, l'estimation dépend de la performance de détection. La méthode *"canny"* est la moins performante car à partir de 40^o d'angle zénith la limite de l'ombre propre devient difficile à détecter. La détection par filtre passe bas est plus performante, l'estimation est seulement affectée pour les fortes valeurs d'angle zénith car le rayon du disque devient trop faible. L'erreur d'estimation reste importante ($> 10^o$ à partir d'un angle zénith $> 40^o$).

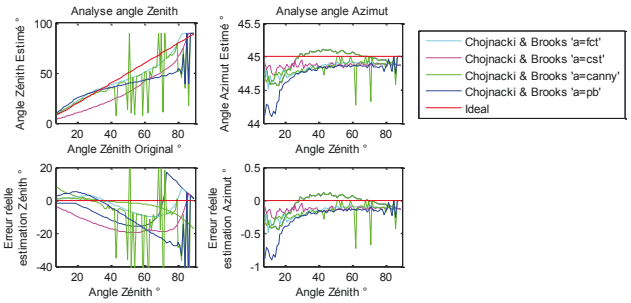


FIGURE A.23 – CHOJNACKI *et* BROOK, *influence de a sur l'estimation de l'angle zénith.*

Afin d'obtenir de meilleurs résultats d'estimation, nous avons cherché à déterminer l'évolution optimale pour a. La fonction d'évolution présentée figure A.24 (fonction polynomial) est déterminée à partir d'un polynôme d'ordre 3 estimé par régression linéaire. Cette évolution de a permet de réduire le biais d'estimation mais la méthode proposée n'apporte pas une réelle amélioration des performances comparé à celle de Pentland.

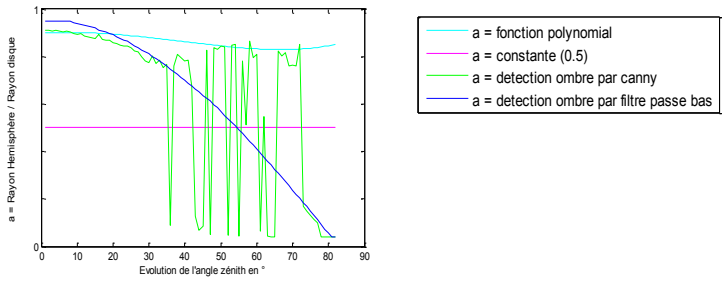


Figure A.24 – Chojnacki *et* Brook, *évolution de a en fonction de l'angle zénith réel.*

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2
ε azimut °	-0.25	-0.04	0.00	0.04	0.00	-0.08	-0.001
etype	0.21	0.18	0.00	0.61	0.11	0.25	0.19
ε zenith °	-9.28	-21.68	-13.46	2.03	-14.22	-22.15	-24.46
etype	12.72	24.16	13.36	1.46	5.38	25.82	28.27

Table A.8 – Chojnacki *et* Brook, *erreur moyenne d'estimation et l'écart type pour les séries d'images tests.*

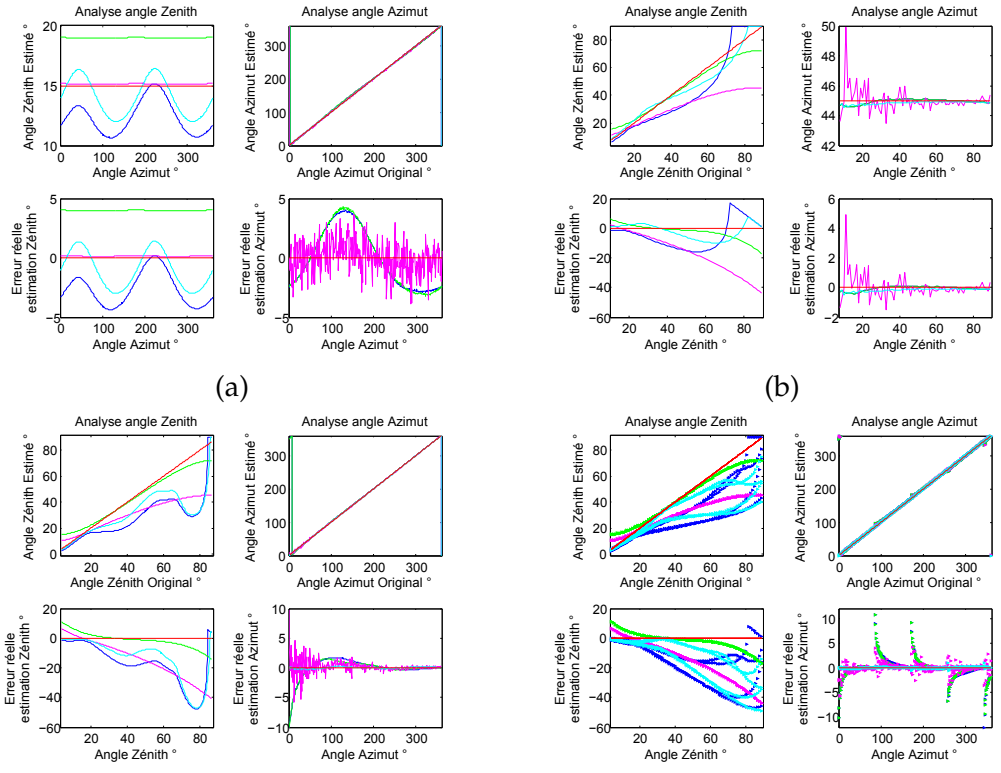


Figure A.25 – Chojnacki *et* Brook *(cyan),* Zheng *et* Chellappa *(magenta),* Lee *et* Rosenfeld *(vert),* Pentland *(bleu), idéal (rouge), graphiques résultats (a) pour azimut* $(0 - 360^o)$ *et zénith* 15^o*, (b) azimut* 45^o *et zénith* $(4 - 89^o)$*, (c) pour éclairage schéma 1, (d) pour éclairage schéma 2.*

D'après les résultats obtenus tableau A.8 et figure A.25, les performances d'estimation de l'angle azimut sont identiques à la méthode de Lee et Rosenfeld et très proche de Pentland. Par contre, l'erreur d'estimation de l'angle zénith est très différente de celle présentée par les auteurs. On note une légère amélioration par rapport à l'estimateur de Pentland et on retrouve une forte corrélation entre l'angle azimut et l'erreur d'estimation de l'angle zénith.

A.4.2 Etude des méthodes basées sur l'analyse des ombres

D'après les résultats présentés par Zhang et Yang, la méthode [ZY01] est capable d'estimer correctement 8 sources lumineuses distinctes avec un biais moyen de 0.1^o pour l'angle zénith, de 0.5^o pour l'angle azimut et 1.1% pour l'intensité. Les images utilisées sont sur 8 Bits pour une résolution de 180×180 *pixels*. Elles sont obtenues de façon synthétique par la modélisation d'une sphère et le modèle Lambertien. L'étude précise que la résolution des images n'influe pas sur les performances d'estimation, de plus les résultats sont donnés pour plusieurs configurations où le nombre et l'intensité des sources varie. Par contre, aucune information n'est donnée sur le nombre de direction ou le schéma d'éclairement testé. La méthode est tout de même très sensible au bruit, ce dernier point est très largement remis en question par Wei dans [Wei03] où il propose une méthode de recherche des points limites et un regroupement beaucoup plus robuste au bruit.

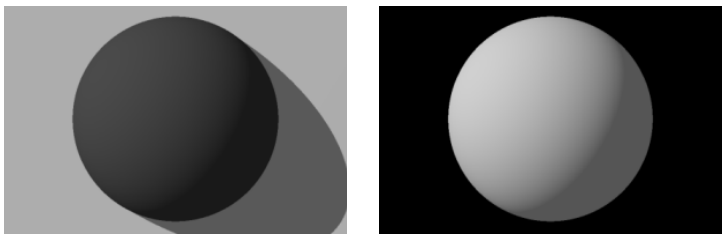
Notre implémentation repose sur le principe de Zhang et Yang mais elle est adaptée à la détection d'une source lumineuse unique afin de pouvoir comparer le principe général de l'estimation par l'analyse des ombres aux méthodes d'analyse statistique des intensités. L'adaptation à une seule source lumineuse permet de simplifier significativement les étapes de détection et de regroupement des points limites. L'étape de création et de résolution du système basé sur l'équation d'irradiance est elle aussi simplifiée car seule l'estimation de la direction de la source est prise en compte. L'algorithme implémenté comporte 4 étapes :

1. estimation de la position et du rayon de la demi-sphère par détection des contours et transformée de Hough [Hou62].
2. recherche des points limites par filtrage passe bas et morphologie mathématique. L'étape de regroupement des points en arc n'est pas nécessaire. Mais nous conservons un affinage par comparaison des intensités afin d'éliminer les points n'appartenant pas réellement à la limite.
3. calcul des normales à la surface en chaque point et élaboration du système linéaire.
4. résolution du système par la méthode des moindres carrés et estimation de l'angle zénith et azimut à partir du vecteur $\vec{l}$.

La majeure partie des problèmes rencontrés lors de l'implémentation concerne la détection des points limites car les performances d'estimation sont corrélées aux performances de détection. Afin de faciliter la localisation des points limites, nous avons modifié l'albédo de la demi-sphère et celui de la surface plane (figure A.26). Par contre les schémas d'éclairage restent identiques à ceux présentés lors des tests précèdent.

Les résultats (figure A.27) démontrent les bonnes performances d'estimation de l'angle zénith, le biais est considérablement réduit comparé aux méthodes d'analyse statistique des intensités. Le tableau A.9, montre que l'erreur zénith moyenne ne dépasse pas 2^o pour un écart type maximum de 1.7^o. Seuls les angles zéniths proches de 0^o génèrent des erreurs (figure A.27a), ce point est abordé par les auteurs et une formule permet de calculer l'angle

minimum zénith détectable. Dans notre cas l'angle zénith minimum est d'environ 10^o car un angle plus faible ne génère pas assez d'ombres portées sur la sphère du fait de la projection perspective.



(a) (b)

FIGURE A.26 – *Images de synthèses (a) albédo(s) tests précèdent, (b) nouveaux albédos pour la méthode* ZHANG *et* YANG.

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2
ε azimut °	1.37	10.56	-4.03	0.003	0.007	-1.46	1.51
etype	34.65	34.08	33.66	6.37	2.38	14.64	40.46
ε zenith °	-0.06	-1.13	-0.14	-0.74	-0.91	-1.08	-1.25
etype	0.50	1.70	0.57	0.55	0.22	0.93	1.49

TABLE A.9 – ZHANG *et* YANG, *erreur moyenne d'estimation et l'écart type pour chaque série d'images tests.*

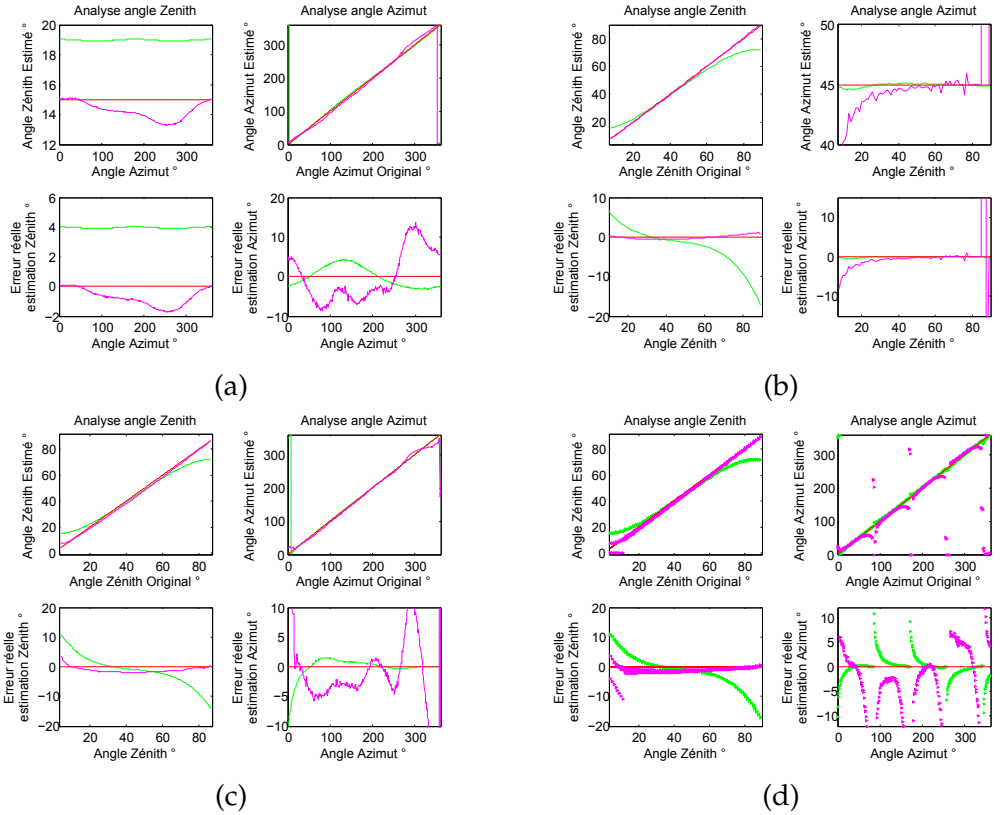
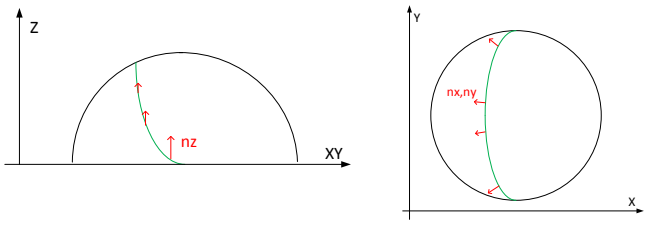


FIGURE A.27 – ZHANG *et* YANG *(magenta),* LEE *et* ROSENFELD *(vert), idéal (rouge), graphiques résultats (a) pour azimut* $(0 - 360^o)$ *et zénith* 15^o*, (b) azimut* 45^o *et zénith* $(4 - 89^o)$*, (c) pour éclairage schéma 1, (d) pour éclairage schéma 2.*

Malheureusement le biais d'estimation de l'angle azimut est globalement plus important, en fonction du schéma d'éclairage (tableau A.9) l'erreur présente un écart type maximum de 40^o. Cet écart de performance est dû à la méthode d'estimation du vecteur $\vec{l}$, pour comprendre il est nécessaire de se placer dans un contexte géométrique. L'ensemble des points choisis présentent obligatoirement une élévation suivant l'axe z ($n_z > 0$), par conséquent aucun point n'est opposé lors de l'estimation de l_z. Il n'en est pas de même pour l'estimation de l_x et l_y car les points choisis peuvent présenter des normales n_x et n_y diamétralement opposées. Pour résumer simplement, chaque point limite travaille ensemble pour l'estimation de l'angle zénith alors qu'il s'oppose pour l'estimation de l'angle azimut (figure A.28). Par conséquent, les performances d'estimation de l'angle azimut sont étroitement liées à la répartition des points limites sur l'ensemble de la sphère.



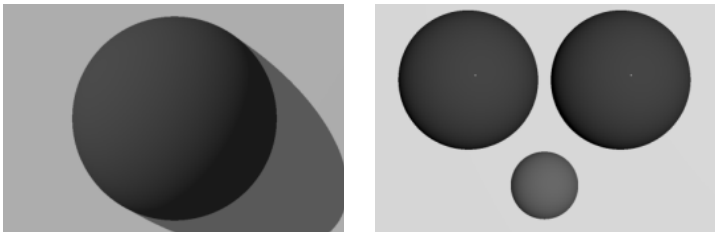
(a) (b)

FIGURE A.28 – *Démonstration, (a) les normales nz aux points critiques présentent toutes une élévation, (b) les normales nx ny aux points critiques sont diamétralement opposées en cas d'oubli ou d'erreur sur l'une d'entre elle l'équilibre est rompu et les erreurs d'estimation apparaissent.*

A.4.3 Etude des méthodes basées sur l'analyse des points spéculaires

POWELL fournit dans [PSG01] des résultats sur cas réels pour 51 positions de source dont les angles zéniths varient entre 10 et 50^o et les distances varient entre 0.5 et 2 m, l'angle azimut reste fixe. Les résultats obtenus montrent une erreur angulaire moyenne de 2.7^o avec un écart type de 0.4^o et une erreur de distance moyenne de 13 cm avec un écart type de 2 cm.

Pour vérifier les performances de la méthode, nous générons comme précédemment des images de synthèse à l'aide du logiciel de lancer de rayon (POV-Ray). La scène simulée (figure A.29b) est similaire à l'expérimentation de POWELL. Deux demi-sphères spéculaires d'un rayon de 55 mm sont placées à égale distance de la caméra et de part et d'autre de l'axe optique. Une troisième sphère Lambertienne plus petite est placée à une distance plus faible de la caméra. Les sources lumineuses suivent les mêmes schémas d'éclairage que les tests précèdent mais elles sont maintenant ponctuelles. La distance de la source est fixée à 1 m.



(a) (b)

FIGURE A.29 – *Images de synthèses (a) tests précèdent, (b) tests* POWELL.

Nous ajoutons un schéma d'éclairage (*schema3*) où les angles sont fixes et seule la distance de

la source varie entre 0.5 et 2 *m* afin d'évaluer les performances d'estimation de la distance. Le centre des sphères, le rayon et les points spéculaires sont obtenus par extraction des contours et la transformée de Hough [Hou62]. Pour faciliter l'extraction, les ombres portées sont éliminée et le coefficient de réflexion des sphères spéculaires est adapté pour que la tâche produite soit la petite possible.

Les résultats obtenus tableau A.10 et figure A.30 démontrent les performances de l'estimation angulaire, comparés aux méthodes d'estimation par analyse des intensités ou des contours, l'erreur moyenne et l'écart type pour les schémas d'éclairage complet sont inférieurs à 1^o contre $1-2^o$ avec des écart type $> 15^o$. Les erreurs d'estimation les plus importantes proviennent des configurations spécifiques telles que des angles azimuts à 0^o,180^o (schéma d'éclairage *Zenith2*). Cette méthode est dépendante de l'extraction des sphères et de la détection des points spéculaires, contrairement à [Pen82], plus la résolution des images augmente plus la méthode est précise. Nous avons testé de multiplier par 2 la résolution, l'erreur d'estimation angulaire est alors divisée par 2.

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2	Schema3
ε azimut °	0.545	-0.882	-1.188	0.005	0.002	0.208	0.052	1.095
etype	0.479	2.767	3.800	1.144	0.540	0.852	0.841	0.159
ε zenith °	-0.024	-1.36	-2.654	0.129	-0.187	-0.199	-0.251	0.231
etype	0.158	3.682	7.121	0.150	0.262	0.398	0.489	0.069

Table A.10 – Powell, *erreur moyenne d'estimation angulaire et l'écart type pour chaque série d'images tests.*

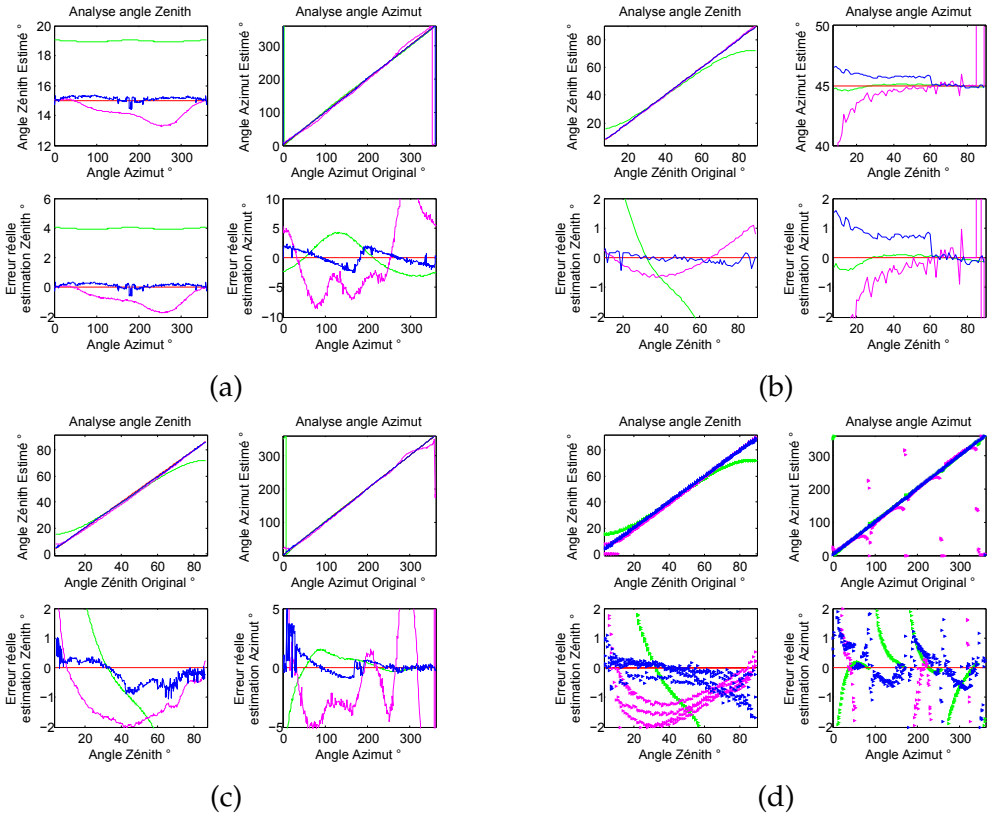


Figure A.30 – Powell *(bleu)*, Zhang *et* Yang *(magenta)*, Lee *et* Rosenfeld *(vert), idéal (rouge), graphiques résultats (a) pour azimut* $(0-360^o)$ *et zénith* 15^o*, (b) azimut* 45^o *et zénith* $(4-89^o)$*, (c) pour éclairage schéma 1, (d) pour éclairage schéma 2.*

Les performances d'estimation de la distance focale suivent celle de l'estimation angulaire, l'erreur moyenne est très faible car inférieure au millimètre avec un écart type proche de 0 *cm* (tableau A.11 et figure A.31). Par contre l'estimation de la distance des sources est loin des performances annoncées, d'après les résultats l'erreur d'estimation moyenne est proche de 20 *cm* avec un écart type important (10 *cm*). Ce biais est généré par l'erreur de détection de la position des sphères et des points spéculaires, une erreur de position de 1 *pixel* (selon la résolution) a peu d'influence sur l'estimation angulaire par contre l'influence sur la triangulation est considérable. De plus, la résolution fixe la capacité d'estimation de la distance ainsi que la position angulaire, avec la configuration utilisée la position du point spéculaire reste identique quelle que soit la distance si elle dépasse 250 *cm*.

Série :	Zenith1	Zenith2	Zenith3	Azimut1	Azimut2	Schema1	Schema2	Schema3
ε focal cm	-0.008	-0.009	-0.006	-0.001	-0.002	-0.003	-0.003	-0.013
etype	0.005	0.002	0.014	0.021	0.018	0.016	0.016	0.000
ε distance cm	73.54	-30.36	-20.28	8.56	-24.33	-19.27	-21.21	-47.788
etype	34.47	17.62	8.21	4.55	4.11	5.96	6.59	29.870

TABLE A.11 – POWELL, *erreur moyenne d'estimation des distances et l'écart type pour chaque série d'images tests.*

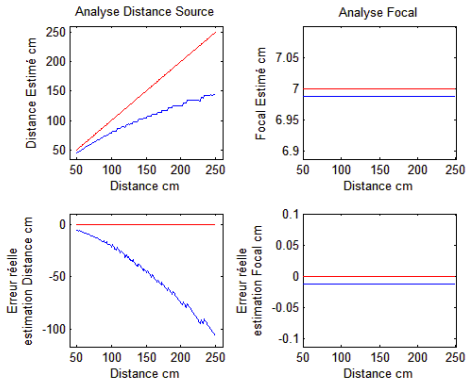


FIGURE A.31 – POWELL, *erreur d'estimation des distances pour une série d'images, la distance de la source varie et les angles restent fixes (azimut* 45^o*, zénith* 15^o*), erreur moyenne distance* -47.78 *cm - écart type* 29.87 *cm, erreur moyenne focal* -0.013 *cm - écart type* 0.00 *cm.*

Bibliographie

[Aa85] Haruhisa Ando and al. Design consideration and performance of a new mos imaging device. *Electron Devices, IEEE Transactions on*, 32(8) :1484–1489, 1985. (Cité page 53.)

[Allo4] David Alleysson. 30 ans de démosaïçage. *Traitement du signal*, 21(6) :561–581, 2004. (Cité page 62.)

[APB10] Vasileios Argyriou, Maria Petrou, and Svetlana Barsky. Photometric stereo with an arbitrary number of illuminants. *Computer Vision and Image Understanding*, 114(8) :887–900, 2010. (Cité pages 82, 84, 87, 88, 90 et 147.)

[AZK08] Neil Alldrin, Todd Zickler, and David Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. (Cité pages 132 et 151.)

[Bay76] Bryce E Bayer. Color imaging array. *US Patent, 3,971,065*, 1976. (Cité page 60.)

[BBK11] Alexandre Bony, Benjamin Bringier, and Majdi Khoudeir. Sur-quantification d'images par variation de lumière pour la reconstruction 3d de surfaces rugueuses. *GRETSI 2011*, 2011. (Cité page 141.)

[BBK12] Benjamin Bringier, Alexandre Bony, and Majdi Khoudeir. Specularity and shadow detection for the multisource photometric reconstruction of a textured surface. *JOSA A*, 29(1) :11–21, 2012. (Cité pages 85 et 147.)

[BBK13a] Alexandre Bony, Benjamin Bringier, and Majdi Khoudeir. Accurate image quantization adapted to multisource photometric reconstruction for rough textured surface analysis. *JOSA A*, 30(3) :518–526, 2013. (Cité page 141.)

[BBK13b] Alexandre Bony, Benjamin Bringier, and Majdi Khoudeir. Stéréo-photométrie itérative dans le cas de sources lumineuses ponctuelles non uniformes. *GRETSI 2013*, 2013. (Cité page 124.)

[BBK13c] Alexandre Bony, Benjamin Bringier, and Majdi Khoudeir. Tridimensional reconstruction by photometric stereo with near spot light sources. *EUSIPCO 2013*, 2013. (Cité page 124.)

[Bey92] Horst A Beyer. Accurate calibration of ccd-cameras. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on*, pages 96–101. IEEE, 1992. (Cité page 37.)

[BH85] Michael J. Brooks and Berthold K. P. Horn. Shape and source from shading. In *Proceedings of the 9th international joint conference on Artificial intelligence - Volume 2*, IJCAI'85, pages 932–936, San Francisco, CA, USA, 1985. Morgan Kaufmann Publishers Inc. (Cité page 106.)

[BHK08] Benjamin Bringier, David Helbert, and Majdi Khoudeir. Photometric reconstruction of a dynamic textured surface from just one color image acquisition. *JOSA A*, 25(3) :566–574, 2008. (Cité page 74.)

[Bli77] James F. Blinn. Models of light reflection for computer synthesized pictures. *SIGGRAPH Comput. Graph.*, 11(2) :192–198, July 1977. (Cité page 25.)

[BP03] Svetlana Barsky and Maria Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(10) :1239–1252, 2003. (Cité pages 82, 84, 85, 87 et 147.)

[Bro71] Duane C Brown. Close-range camera calibration. *Photogrammetric engineering*, 37(8) :855–866, 1971. (Cité page 37.)

[BS87] Petr Beckmann and Andre Spizzichino. The scattering of electromagnetic waves from rough surfaces. *Norwood, MA, Artech House, Inc., 1987, 511 p.*, 1, 1987. (Cité page 26.)

[CBG94] Wojciech Chojnacki, Michael J Brooks, and Danny Gibbins. Revisiting pentland's estimator of light source direction. *JOSA A*, 11(1) :118–124, 1994. (Cité pages 106, 110, 168 et 173.)

[CCDG07] Frédéric Courteille, Alain Crouzil, Jean-Denis Durou, and Pierre Gurdjos. Shape from shading for the digitization of curved documents. *Machine Vision and Applications*, 18(5) :301–316, 2007. (Cité page 69.)

[CJ82] E. North Coleman and Ramesh Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4) :309 – 328, 1982. (Cité pages 74, 76 et 82.)

[CKK+96] Peter Cheeseman, Bob Kanefsky, Richard Kraft, John Stutz, and Robin Hanson. Super-resolved surface reconstruction from multiple images. In *Maximum Entropy and Bayesian Methods*, pages 293–308. Springer, 1996. (Cité page 132.)

[Cla12] RN Clark. Pixel size, iso and noise in digital cameras. `http://www.clarkvision.com/articles/pixel.size.and.iso/`, 2012. (Cité page 135.)

[CM02] Maureen Clerc and Stephane Mallat. The texture gradient equation for recovering shape from texture. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4) :536–549, 2002. (Cité page 69.)

[CMB96] Joseph Calderer, Mauricio Moreno, and Marco Braam. Integration of phototransistors in cmos circuits. *Sensors and materials*, 8(4) :199–208, 1996. (Cité page 47.)

[CT82] Robert L Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (TOG)*, 1(1) :7–24, 1982. (Cité page 26.)

[Dan00] Pascal Daniel. *Peut-on extraire le relief d'une seule image ?* PhD thesis, Université Paul Sabatier Toulouse, 2000. (Cité pages 68, 78 et 161.)

[DB24] Louis De Broglie. *Recherches sur la théorie des quanta*. PhD thesis, Migration-université en cours d'affectation, 1924. (Cité page 9.)

[DC07] Jean-Denis Durou and Frédéric Courteille. Integration of a normal field without boundary condition. In *Proceedings of the First International Workshop on Photometric Analysis For Computer Vision-PACV 2007*, 2007. (Cité page 78.)

[DCD+11] Jean-Denis Durou, Vincent Charvillat, Maxime Daramy, Pierre Gurdjos, et al. Résolution du shape-from-shading par apprentissage. In *ORASIS-Congrès des jeunes chercheurs en vision par ordinateur*, 2011. (Cité page 68.)

[Deb08] Paul Debevec. Rendering synthetic objects into real scenes : Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *ACM SIGGRAPH 2008 classes*, page 32. ACM, 2008. (Cité pages 106 et 113.)

[Des65] Rene Descartes. La dioptrique (1637). *Ders, : Oeuvres des Descartes*, 6 :79–228, 1965. (Cité page 6.)

[dl83] Commission Internationale de l'Eclairage. The basis of physical photometry. *CIE publication*, 18, 1983. (Cité page 17.)

[DLMF00] Yavuz Degerli, Francis Lavernhe, Pierre Magnan, and Jean A Farré. Analysis and reduction of signal readout circuitry temporal noise in cmos image sensors for low-light levels. *Electron Devices, IEEE Transactions on*, 47(5) :949–962, 2000. (Cité page 47.)

[DM97] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, pages 369–378, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co. (Cité pages 136, 138 et 140.)

[EGSSF11] Amin El Gendy, Ahmed Shalaby, Mohamed Saleh, and Gerardo W Flintsch. Stereo-vision applications to reconstruct the 3d texture of pavement surface. *International Journal of Pavement Engineering*, 12(03) :263–273, 2011. (Cité page 82.)

[Ein05] Albert Einstein. Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt. *Annalen der Physik*, 17 :132–148, 1905. (Cité page 9.)

[EM13] Moumen T El-Melegy. Radial lens distortion calibration from spheres : Theory and method. *Int. J. Com. Dig. Sys*, 2(1) :39–46, 2013. (Cité page 38.)

[FC88] Robert T. Frankot and Rama Chellappa. A method for enforcing integrability in shape from shading algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(4) :439–451, 1988. (Cité pages 74, 78, 79 et 91.)

[FS05] Paolo Favaro and Stefano Soatto. A geometric approach to shape from defocus. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(3) :406–417, 2005. (Cité page 70.)

[Fun04] Nathan Funk. A review of selected single light source detection methods. *University of Alberta, Edmonton, Alberta, Canada*, 2004. (Cité pages 168 et 169.)

[GAW+10] Miguel Granados, Boris Ajdin, Michael Wand, Christian Theobalt, H-P Seidel, and Hendrik PA Lensch. Optimal hdr reconstruction with linear digital cameras. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 215–222. IEEE, 2010. (Cité page 137.)

[GBC91] Danny Gibbins, Michael J. Brooks, and Wojciech Chojnacki. Light source direction from a single image : a performance analysis. *Australian Computer Journal*, 23(4) :165–174, 1991. (Cité pages 168, 171 et 172.)

[GCHS10] Dan B Goldman, Brian Curless, Aaron Hertzmann, and Steven M Seitz. Shape and spatially-varying brdfs from photometric stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(6) :1060–1071, 2010. (Cité pages 132 et 151.)

[Geo03] Athinodoros S Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 816–823. Ieee, 2003. (Cité page 82.)

[GN03] Michael D Grossberg and Shree K Nayar. High dynamic range from multiple images : Which exposures to combine ? In *Proc. ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV)*, 2003. (Cité pages 137 et 142.)

[HB86] Berthold KP Horn and Michael J Brooks. The variational approach to shape from shading. *Computer Vision, Graphics, and Image Processing*, 33(2) :174–208, 1986. (Cité page 77.)

[Hem03] Elsayed E Hemayed. A survey of camera self-calibration. In *Proceedings. IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.*, pages 351–357. IEEE, 2003. (Cité page 38.)

[HK04] Itsik Horovitz and Nahum Kiryati. Depth from gradient fields and control points : Bias correction in photometric stereo. *Image and Vision Computing*, 22(9) :681–694, 2004. (Cité page 78.)

[HMI10] Tomoaki Higo, Yasuyuki Matsushita, and Katsushi Ikeuchi. Consensus photometric stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1157–1164. IEEE, 2010. (Cité pages 27 et 82.)

[Hoo94] Friits N Hooge. 1/f noise sources. *Electron Devices, IEEE Transactions on*, 41(11) :1926–1935, 1994. (Cité page 47.)

[Hor75] Berthold KP Horn. Obtaining shape from shading information. *The psychology of computer vision*, pages 115–155, 1975. (Cité page 67.)

[Hor86] Berthold Klaus Paul Horn. *Robot vision*. the MIT Press, 1986. (Cité pages 41 et 134.)

[Hou62] Paul VC Hough. Method and means for recognizing complex patterns, 1962. (Cité pages 119, 175 et 178.)

[Huy67] Christian Huygens. *Traité de la lumière*. Culture et civilisation, 1967. (Cité page 7.)

[HVB+07] Carlos Hernández, George Vogiatzis, Gabriel J Brostow, Björn Stenger, and Roberto Cipolla. Non-rigid photometric stereo with colored lights. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007. (Cité page 74.)

[HVC08] Carlos Hernández, George Vogiatzis, and Roberto Cipolla. Shadows in three-source photometric stereo. In *Computer Vision–ECCV 2008*, pages 290–303. Springer, 2008. (Cité page 82.)

[INI83] Katsushi Ikeuchi, Memo No, and Katsushi Ikeuchi. Constructing a depth map from images. In *AI Memo AIM-744, Artificial Intelligence Laboratory, MIT*. Citeseer, 1983. (Cité pages 74 et 77.)

[IoNA87] American National Standards Institute and Illuminating Engineering Society of North America. *Nomenclature and Definitions for Illuminating Engineering*. American National Standard. Illuminating Engineering Society of North America, 1987. (Cité page 13.)

[ISO09] NF ISO. 25178. *Spécification géométrique des produits (GPS)-État de surface : surfacique*, 2009. (Cité page 148.)

[Ka80] Norio Koike and al. Mos area sensor : Part i - design consideration and performance of an n-p-n structure 484 x 384 element color mos imager. *Solid-State Circuits, IEEE Journal of*, 15(4) :741–746, 1980. (Cité page 53.)

[KA06] Kristian Kirk and Hans Jørgen Andersen. Noise characterization of weighting schemes for combination of multiple exposures. In *BMVC*, pages 1129–1138. Citeseer, 2006. (Cité page 137.)

[Lam60] Johann Heinrich Lambert. *Photometria*. sumptibus viduae E. Klett, 1760. (Cité page 22.)

[LLLS03] Yuanzhen Li, Stephen Lin, Hanqing Lu, and Heung-Yeung Shum. Multiple-cue illumination estimation in textured scenes. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1366–1373. IEEE, 2003. (Cité page 106.)

[LR85] Chia-Hoang Lee and Azriel Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *artificial Intelligence*, 26(2) :125–143, 1985. (Cité pages 106, 108, 111, 168, 169 et 170.)

[Max65] J Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155 :459–512, 1865. (Cité page 8.)

[ML99] Jean-Louis Meyzonnette and Thierry Lépine. *Bases de radiométrie optique*. Cépaduès, 1999. (Cité page 13.)

[MN99] Tomoo Mitsunaga and Shree K Nayar. Radiometric self calibration. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1. IEEE, 1999. (Cité page 137.)

[Mor05] Olivier Morel. *Environnement actif pour la reconstruction tridimensionnelle de surfaces métalliques spéculaires par imagerie polarimétrique*. PhD thesis, Université de Bourgogne, 2005. (Cité page 161.)

[MP95] S. Mann and R. Picard. Being 'undigital' with digital cameras : Extending dynamic range by combining differently exposed pictures. In *Proc. Annual Conf. IS&T*, pages 422–428, 1995. (Cité page 136.)

[N+22] Isaac Newton et al. *Traité d'optique sur les réflexions réfractions, inflexions et les couleurs de la lumière par Monsieur le Chevalier Newton, traduit [et préfacé] par M. Coste sur la seconde édition anglaise, augmentée par l'auteur*. chez Montalant, 1722. (Cité page 7.)

[Nay91] Shree K Nayar. Shape and reflectance from image intensities. In *Proceedings of the Third Annual Conference on Intelligent Robotic Systems for Space Exploration (IEEE, 1991)*, pages 81–98, 1991. (Cité page 82.)

[NIK90] Shree K Nayar, Katsushi Ikeuchi, and Takeo Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *Robotics and Automation, IEEE Transactions on*, 6(4) :418–431, 1990. (Cité page 26.)

[NN94] Shree K Nayar and Yasuo Nakagawa. Shape from focus. *Pattern analysis and machine intelligence, IEEE Transactions on*, 16(8) :824–831, 1994. (Cité page 70.)

[ON95] Michael Oren and Shree K Nayar. Generalization of the lambertian model and implications for machine vision. *International Journal of Computer Vision*, 14(3) :227–251, 1995. (Cité page 24.)

[Pa02] William H Press and al. Numerical recipes in c++. 2002. *Cambridge UP*, 2002. (Cité pages 74 et 78.)

[PCF05] James A Paterson, David Claus, and Andrew W Fitzgibbon. Brdf and geometry capture from extended inhomogeneous samples using flash photography. *Computer Graphics Forum (Special Eurographics Issue)*, 24(3) :383–391, 2005. (Cité pages 106 et 117.)

[Pen82] Alex P Pentland. Finding the illuminant direction. *JOSA*, 72(4) :448–455, 1982. (Cité pages 106, 107, 108, 110, 167, 170 et 178.)

[Pen84] Alex P Pentland. Local shading analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-6(2) :170–187, 1984. (Cité page 68.)

[Pen91] Michael A. Penna. Camera calibration : A quick and easy way to determine the scale factor. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(12) :1240–1245, 1991. (Cité page 38.)

[PFa04] Emmanuel Prados, Olivier Faugeras, and al. A rigorous and realistic shape from shading method and some of its applications. Technical Report RR-5133, INRIA, 2004. (Cité page 68.)

[PFa06] Emmanuel Prados, Olivier Faugeras, and al. Rôle clé de la modélisation en shape from shading. In *15ème Congrès de Reconnaissance des Formes et Intelligence Artificielle*, 2006. (Cité page 68.)

[Pho75] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6) :311–317, 1975. (Cité page 24.)

[Pla01] Max Planck. Über das gesetz der energieverteilung im normalspektrum. *Annalen der Physik*, 4 :698–706, 1901. (Cité page 9.)

[PPK03] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super-resolution image reconstruction : a technical overview. *Signal Processing Magazine, IEEE*, 20(3) :21–36, 2003. (Cité page 132.)

[PSG01] Mark W. Powell, Sudeep Sarkar, and Dmitry Goldgof. A simple strategy for calibrating the geometry of light sources. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(9) :1022–1027, 2001. (Cité pages 106, 114 et 177.)

[RBS03] Mark A Robertson, Sean Borman, and Robert L Stevenson. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging*, 12(2) :219–228, 2003. (Cité page 137.)

[RDLR91] Marc Richetin, Michel Dhome, Jean-Thierry Lapreste, and Gerard Rives. Inverse perspective transform using zero-curvature contour points : application to the localization of some generalized cylinders from a single view. *IEEE transactions on pattern analysis and machine intelligence*, 13(2) :185–192, 1991. (Cité page 69.)

[Rod93] Volker Rodehorst. Vertiefende analyse eines gestalts-constraints von aloimonos und shulman. *TechnischerBericht, CV-Bericht*, 8, 1993. (Cité pages 74 et 76.)

[RSSF02] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21(3) :267–276, July 2002. (Cité page 137.)

[RTG97] Holly Rushmeier, Gabriel Taubin, and André Guéziec. *Applying shape from lighting variation to bump map capture*. Springer, 1997. (Cité page 84.)

[SI96] Fredric Solomon and Katsushi Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(4) :449–454, 1996. (Cité page 82.)

[SN81] Simon M Sze and Kwok K Ng. *Physics of Semiconductor Devices*. Edition John Wiley and Sons, 1981. (Cité page 43.)

[SSS+07] Jiuai Sun, Melvyn Smith, Lyndon Smith, Sagar Midha, and Jeff Bamber. Object surface recovery using a multi-light photometric stereo technique for non-lambertian surfaces subject to shadows and specularities. *Image and Vision Computing*, 25(7) :1050–1057, 2007. (Cité pages 82, 84, 88 et 147.)

[SSSF10] Jiuai Sun, M Smith, L Smith, and A Farooq. Compensation of illumination radiance in photometric stereo. *ICMV 2011*, pages 40–44, 2010. (Cité pages 106 et 117.)

[Ter84] Demetri Terzopoulos. *Multilevel reconstruction of visual surfaces : Variational principles and finite-element representations*. Springer, 1984. (Cité page 78.)

[Ter88] Demetri Terzopoulos. The computation of visible-surface representations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(4) :417–438, 1988. (Cité page 78.)

[TFG01] Hui Tian, Boyd Fowler, and Abbas E Gamal. Analysis of temporal noise in cmos photodiode active pixel sensor. *Solid-State Circuits, IEEE Journal of*, 36(1) :92–101, 2001. (Cité page 47.)

[The95] Albert J Theuwissen. *Solid-state imaging with charge-coupled devices*, volume 1. Springer, 1995. (Cité page 54.)

[TN13] I. Takayanagi and J. Nakamura. High-resolution cmos video image sensors. *Proceedings of the IEEE*, 101(1) :61–73, 2013. (Cité page 132.)

[TS67] Kenneth E Torrance and Ephraim M Sparrow. Theory for off-specular reflection from roughened surfaces. *JOSA*, 57(9) :1105–1112, 1967. (Cité page 26.)

[VLR94] Lode KJ Vandamme, Xiaosong Li, and Dominique Rigaud. 1/f noise in mos devices, mobility or number fluctuations ? *Electron Devices, IEEE Transactions on*, 41(11) :1936–1945, 1994. (Cité page 47.)

[vO06] J.D. van Ouwerkerk. Image super-resolution survey. *Image and Vision Computing*, 24(10) :1039 – 1052, 2006. (Cité page 132.)

[WB92] Han Wang and Michael Brady. A structure-from-motion algorithm for robot vehicle guidance. In *Intelligent Vehicles' 92 Symposium., Proceedings of the*, pages 30–35. IEEE, 1992. (Cité page 71.)

[WCH92] Juyang Weng, Paul Cohen, and Marc Herniou. Camera calibration with distortion models and accuracy evaluation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(10) :965–980, 1992. (Cité page 38.)

[Wei90] D Weinshall. The shape and the direction of illumination from shading on occluding contours. *Artificial intelligence memo*, 1264, 1990. (Cité page 106.)

[Wei03] Jie Wei. Robust recovery of multiple light source based on local light source constant constraint. *Pattern recognition letters*, 24(1) :159–172, 2003. (Cité pages 113 et 175.)

[WK01] Tiangong Wei and Reinhard Klette. A new algorithm for gradient field integration. Technical report, CITR, The University of Auckland, New Zealand, 2001. (Cité page 79.)

[WK02] Tiangong Wei and Reinhard Klette. Height from gradient using surface curvature and area constraints. In *Third Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)(ICVGIP, 2002)*, pages 204–210. Citeseer, 2002. (Cité pages 74 et 79.)

[WL88] Zhongquan Wu and Lingxiao Li. A line-integration based method for depth recovery from surface normals. *Computer Vision, Graphics, and Image Processing*, 43(1) :53–66, 1988. (Cité pages 74 et 76.)

[Wol96] Lawrence B Wolff. Generalizing lambert's law for smooth surfaces. In *Computer Vision—ECCV'96*, pages 40–53. Springer, 1996. (Cité page 24.)

[Woo80] Robert J Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1) :191139–191139, 1980. (Cité page 69.)

[WS02] Yang Wang and Dimitris Samaras. Estimation of multiple illuminants from a single image of arbitrary known geometry. In *Computer Vision—ECCV 2002*, pages 272–288. Springer, 2002. (Cité page 112.)

[YY91] Yibing Yang and Alan Yuille. Sources from shading. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 534–539. IEEE, 1991. (Cité page 106.)

[ZC91] Qinfen Zheng and Rama Chellappa. Estimation of illuminant direction, albedo, and shape from shading. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 540–545. IEEE, 1991. (Cité pages 106, 109 et 168.)

[ZEAa93] Jeffrey J Zarnowski, Sayed I Eid, Frank S Arnold, and al. Performance of a large-format charge-injection device. In *IS&T/SPIE's Symposium on Electronic Imaging : Science and Technology*, pages 110–120. International Society for Optics and Photonics, 1993. (Cité page 54.)

[Zhaoo] Zhengyou Zhang. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11) :1330–1334, 2000. (Cité page 38.)

[ZK02] Wei Zhou and Chandra Kambhamettu. Estimation of illuminant direction and intensity of multiple light sources. *Lecture Notes in Computer Science*, pages 206–220, 2002. (Cité page 106.)

[ZWPa91] Jeffrey J Zarnowski, Bryn Williams, Matthew A Pace, and al. Selectable one-to-four-port, very high speed 512 x 512 charge-injection device. In *Medical Imaging'91, San Jose, CA*, pages 191–201. International Society for Optics and Photonics, 1991. (Cité page 53.)

[ZY01] Yufei Zhang and Y-H Yang. Multiple illuminant direction detection with application to image synthesis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(8) :915–920, 2001. (Cité pages 106, 112 et 175.)

Titre : Modélisation de l'interaction lumière/matière pour l'analyse de surfaces rugueuses texturées par stéréo photométrie

Résumé : Les techniques de reconstruction 3d sont devenues incontournables pour des applications telles que la caractérisation et l'analyse de surfaces. Les travaux réalisés au cours de cette thèse ont pour objectif d'améliorer la qualité des reconstructions 3d par stéréo-photométrie. Cette méthode repose sur deux principes, l'inversion d'un modèle d'interaction lumière/matière (BRDF) et la configuration d'un système d'éclairage et de prises de vues. Pour des surfaces diffuses, la stéréo-photométrie est réalisée à partir d'un minimum de trois images acquises d'un point de vue fixe pour des directions d'éclairages différentes. Son avantage est d'extraire simultanément les propriétés géométriques et colorimétriques des surfaces analysées même en cas de forte rugosité. Néanmoins, son application exige la formulation de plusieurs hypothèses qui sont difficilement respectables dans un contexte réel. Ceci génère des erreurs significatives dans les reconstructions. Pour les réduire, nous proposons différentes contributions qui s'articulent autour de la prise en compte globale de la chaine d'acquisition. Les apports de nos travaux se situent aux niveaux de la caractérisation et de la modélisation du système d'éclairage, du capteur d'acquisition et de l'amélioration de la qualité des images. Nous nous sommes aussi intéressés à l'optimisation des protocoles de prises de vues dans le cas de spécularité surfacique ou d'ombrage dus à la présence de rugosité. Les résultats obtenus montrent que la prise en compte de ces caractéristiques dans l'inversion d'un modèle de BRDF permet une nette amélioration des reconstructions et offre la possibilité de réduire la taille des systèmes d'acquisition.

Mots-clés : *reconstruction 3d, acquisition d'image, photométrie, stéréo-photométrie, brdf, imagerie.*

Title : Light-matter interaction modelling for analysing textured rough surfaces by photometric stereo

Abstract : Tridimensional reconstruction method has become essential for applications such as the characterization and analysis of surfaces. In this thesis, aims are to increase the quality of 3d reconstructions by photometric stereo. This method is based on two principles, reversing light-matter interaction model and configuration of a lighting system. With diffuse surfaces, the photometric stereo use three captured images from a fixed point of view for different illumination directions. Its main advantage is to extract the color and geometric properties for the textured rough surfaces. However, its application requires to make assumptions that are not credible in real cases. This problem generates significant errors in the reconstructions. To reduce them, we offer various solutions around the overall consideration of the acquisition chain. Our contribution focuses on the characterization and modeling of the lighting system, the acquisition sensor and improved image quality. We are also interested to optimize acquisition protocol in the case of specular surface or shading due to the surface geometry. Our results show that the inclusion of these features in the inversion of a BRDF model allows an improvement of 3d reconstructions as well as the possibility of reducing the size of the acquisition systems.

Keywords : *3d reconstruction, image acquisition, photometry, photometric stereo, BRDF, imaging.*