



Reconstruction tridimensionnelle par stéréophotométrie

Yvain Quéau

► To cite this version:

Yvain Quéau. Reconstruction tridimensionnelle par stéréophotométrie. Traitement des images [eess.IV]. Institut National Polytechnique de Toulouse - INPT, 2015. Français. <NNT : 2015INPT0123>. <tel-01261526>

HAL Id: tel-01261526

<https://theses.hal.science/tel-01261526v1>

Submitted on 25 Jan 2016

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

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



HAL Authorization



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Image, Information et Hypermédia

Présentée et soutenue par :

M. YVAIN QUEAU

le jeudi 26 novembre 2015

Titre :

RECONSTRUCTION TRIDIMENSIONNELLE PAR
STEREOPHOTOMETRIE

Ecole doctorale :

Mathématiques, Informatique, Télécommunications de Toulouse (MITT)

Unité de recherche :

Institut de Recherche en Informatique de Toulouse (I.R.I.T.)

Directeur(s) de Thèse :

M. JEAN DENIS DUROU

Rapporteurs :

M. ADRIEN BARTOLI, UNIVERSITE CLERMONT-FERRAND 1

M. DANIEL CREMERS, TECHNISCHE UNIVERSITAT MUNICH

M. MAURIZIO FALCONE, UNIV. DEGLI STUDI DE ROME SAPIENZA

Membre(s) du jury :

M. VINCENT CHARVILLAT, INP TOULOUSE, Président

M. JEAN DENIS DUROU, UNIVERSITE TOULOUSE 3, Membre

M. JEAN-FRANCOIS AUJOL, UNIVERSITE DE BORDEAUX, Membre

M. OLIVIER AUBRETON, UNIVERSITE DE BOURGOGNE, Membre

M. PIERRE GURDJOS, INP TOULOUSE, Membre

À ma famille.

Kentoc'h mervel eget evañ dour.
[Dur Dabla]

Remerciements

Je remercie tout d’abord le jury d’évaluation, et plus particulièrement Adrien Bartoli, Daniel Cremers et Maurizio Falcone, pour avoir accepté de rédiger un rapport sur le manuscrit. Leurs critiques pertinentes et constructives ont été précieuses. Je remercie également Olivier Aubreton et Jean-François Aujol pour leur présence au sein du jury, et pour avoir apporté leurs points de vue sur les différents aspects de la thèse, de l’inspection visuelle de surface à l’utilisation des méthodes variationnelles. Enfin, merci à Pierre Gurdjos pour avoir participé au jury, et surtout pour les bons moments passés à discuter au cours de la thèse, que ce soit de géométrie projective ou bien de football.

Je souhaite ensuite remercier chaleureusement tous les membres de l’équipe pédagogique de l’ENSEEIH, qui m’ont permis d’acquérir les connaissances préalables à la réalisation de la thèse. C’est la qualité de la double formation que j’y ai reçue, en mathématiques appliquées et en informatique, qui m’a donné envie de poursuivre mes études jusqu’au doctorat. J’adresse un remerciement particulier à Vincent Charvillat, à la fois pour son implication dans la formation des ingénieurs de l’ENSEEIH, pour avoir accepté de présider le jury de thèse, pour sa gestion dynamique de notre équipe de recherche, et surtout pour l’ensemble des conseils qu’il m’a fournis au cours de la thèse. Ces conseils m’ont permis de cerner très rapidement ce qui était attendu d’un doctorant, et de faire les choix en conséquence.

Je remercie également tous les membres de l’IRIT, permanents, thésards, ou stagiaires, avec qui j’ai passé du temps au cours de ces trois années. La bonne ambiance qui règne dans l’équipe VORTEX a contribué à créer un environnement propice au bon déroulement de la thèse. Nous avons également la chance d’être accompagnés, pour les missions et formalités diverses, par les Sylvies, Audrey et Annabelle, dont la bonne humeur n’a d’égal que leur efficacité. Je les en remercie sincèrement.

Je remercie aussi l’ensemble des personnes avec qui j’ai eu la chance de collaborer scientifiquement, et en particulier Michael Breuss et son équipe à Cottbus, Roberto Mecca à Bologne, et Xavier Descombes à INRIA Sophia-Antipolis. Je les remercie pour la qualité de leur accueil, et pour avoir, grâce à un regard différent, débloqué certaines situations et levé de nouveaux problèmes. Je remercie également les gens qui rendent « tangibles » nos recherches, notamment les sociétés Fitting Box et Pixience : la collaboration avec ces sociétés a permis non seulement de donner un cadre applicatif aux travaux présentés dans ce mémoire, mais également d’identifier de nouvelles problématiques scientifiques. Je remercie tout particulièrement la SATT de Midi-Pyrénées, TTT, pour avoir mené les projets de transfert technologique vers ces sociétés, et surtout favorisé le dialogue entre les mondes académique et industriel, qui ne sont pas aussi hermétiques que l’on voudrait le laisser croire. Je remercie également Gilles Azzaro pour nous permettre de construire une nouvelle passerelle, vers le monde artistique.

Je tiens à exprimer ma gratitude envers mes proches qui m’ont soutenu pendant ces années, et en particulier ma compagne Cathie, qui a su faire preuve de patience et m’a soutenu dans les moments les plus difficiles. Merci également à tous mes amis

pour m'avoir permis de décompresser lorsque cela était nécessaire. Merci, surtout, à ma famille, pour m'avoir permis de réaliser des longues études, et pour m'avoir transmis des valeurs comme le sérieux dans le travail et le respect de la notion de service public.

Je remercie enfin les personnes qui ont encadré ce travail. François Lauze, qui a eu la gentillesse de m'accueillir à Copenhague lors de mon stage de fin d'études à l'ENSEEIH, est devenu au fur et à mesure bien plus qu'un collaborateur. Il a, dans les faits, co-encadré cette thèse et a toujours été un interlocuteur disponible et de qualité pour discuter de sciences, mais aussi d'autres choses. Je le remercie sincèrement pour avoir partagé avec moi une partie de ses connaissances en mathématiques, sans lesquelles cette thèse aurait probablement été très différente. Enfin, je ne saurais assez remercier Jean-Denis Durou pour la manière dont il a encadré la thèse. Jamais avare de discussion autour d'une tasse de café, ou de temps à consacrer à la relecture, il a réussi à faire preuve d'exigence sur la clarté du raisonnement scientifique, tout en me faisant prendre conscience de l'importance de cette exigence. Il a été, d'une part, le principal artisan de la construction d'un manuscrit cohérent, mais il m'a surtout transmis le goût d'une certaine approche de la recherche, fondée sur l'analyse critique et la retransmission claire du savoir. Je crois avoir, grâce à lui, beaucoup gagné en maturité scientifique, et lui en suis extrêmement reconnaissant.

Résumé

Cette thèse traite de la reconstruction 3D par stéréophotométrie, qui consiste à utiliser plusieurs photographies d'une scène prises sous le même angle, mais sous différents éclairages. Nous nous intéressons dans un premier temps à des techniques robustes pour l'estimation des normales à la surface, et pour leur intégration en une carte de profondeur. Nous étudions ensuite deux situations où le problème est mal posé : lorsque les éclairages sont inconnus, ou lorsque seuls deux éclairages sont utilisés. La troisième partie est consacrée à l'étude de modèles plus réalistes, à la fois en ce qui concerne les éclairages et la réflectance de la surface. Ces trois premières parties nous amènent aux limites de la formulation classique de la stéréophotométrie : nous introduisons finalement, dans la partie 4, une reformulation variationnelle et différentielle du problème qui permet de dépasser ces limites.

Abstract

This thesis tackles the photometric stereo problem, a 3D-reconstruction technique consisting in taking several pictures of a scene under different lightings. We first focus on robust techniques for estimating the normals to the surface, and for integrating these normals into a depth map. Then, we study two situations where the problem is ill-posed: when lightings are unknown and when only two images are used. Part 3 is devoted to more realistic models, in terms of lightings and of surface reflectance. These first three parts bring us to the limits of the usual formulation of photometric stereo: we eventually introduce in Part 4 a variational and differential reformulation of this problem which allows us to overcome these limits.

Table des matières

Introduction	1
I Résolution robuste de la stéréophotométrie	9
1 Estimation robuste des normales	13
1.1 Limites de la stéréophotométrie classique	13
1.1.1 Modèle de formation des images	13
1.1.2 Approximation linéaire du modèle lambertien	18
1.1.3 Considérations statistiques	19
1.1.4 Sources d'écarts au modèle lambertien linéaire	21
1.1.5 Biais induit par les écarts au modèle lambertien linéaire	28
1.2 Prétraitement des données	32
1.2.1 Changement d'espace colorimétrique	32
1.2.2 Projection sur l'espace des matrices de rang 3	34
1.2.3 Minimisation parcimonieuse du rang	35
1.2.4 Éclairages quasi-parallèles, non uniformes	37
1.3 Stéréophotométrie calibrée robuste	40
1.3.1 Prise en compte de l'éclairage « ambiant »	40
1.3.2 Sélection des niveaux de gris	43
1.3.3 Moindres carrés pondérés	43
1.3.4 M-estimateurs	45
1.3.5 Approximation différentiable de l'opérateur max	49
1.4 Algorithme pour la stéréophotométrie calibrée robuste	51
1.5 Cas des images colorées	51
1.6 Conclusion du chapitre	54
2 Intégration robuste des normales	55
2.1 Normale, gradient et intégrabilité	55
2.1.1 Équations de l'intégration des normales	55
2.1.2 Intégrabilité d'un champ de normales	58
2.2 Difficultés de l'intégration des normales	59
2.2.1 Robustesse au bruit et aux données aberrantes	59
2.2.2 Topologie du domaine de reconstruction	61
2.2.3 Condition au bord	61
2.2.4 Robustesse aux discontinuités de profondeur	62
2.2.5 Complexité algorithmique	64
2.2.6 Prise en compte d'un a priori	66
2.3 Intégration en moindres carrés	67

2.3.1	Discrétisation de la condition d'optimalité du problème (2.21) . . .	67
2.3.2	Résolution du problème discret associé à (2.21)	71
2.3.3	Extensions de l'intégration en moindres carrés	73
2.3.4	Évaluation quantitative	75
2.4	Nouvelles méthodes d'intégration robustes	77
2.4.1	Moindres carrés pondérés	77
2.4.2	Modèle L^2 -TV	81
2.4.3	M-estimateurs	85
2.4.4	Une méthode d'intégration optimale ?	89
2.5	Conclusion du chapitre	91
 II Deux problèmes de stéréophotométrie mal posés		93
3	Stéréophotométrie non calibrée	97
3.1	Un problème mal posé	98
3.1.1	Ambiguïté linéaire	98
3.1.2	Contrainte d'intégrabilité	99
3.1.3	Résolution du problème (3.8)	100
3.1.4	Ambiguïté GBR	101
3.1.5	Régularisation par variation totale	104
3.2	Désambiguïsation par régularisation	106
3.2.1	Motivations	106
3.2.2	Influence d'une transformation GBR sur la régularité de ρ , z et $\mathbf{m}$.	107
3.2.3	Modèle proposé	111
3.2.4	Estimation de μ et ν	112
3.2.5	Estimation de λ	112
3.2.6	Régularisation TV de la profondeur	115
3.3	Validation expérimentale	117
3.3.1	Résultats sur données synthétiques	117
3.3.2	Résultats sur données réelles	119
3.4	Conclusion du chapitre	124
4	Stéréophotométrie à deux images	127
4.1	Mise en évidence des ambiguïtés du problème PS2	128
4.1.1	Formulation non différentielle du problème PS2	128
4.1.2	Deux formulations différentielles du problème PS2	129
4.1.3	Exemple de problème PS2	131
4.1.4	Problème PS2 avec albédo inconnu	132
4.1.5	Exemple de problème PS2 ρ	133
4.2	Prédiction du nombre de solutions du problème PS2	134
4.2.1	Points singuliers	134
4.2.2	Utilisation des points singuliers pour la construction d'une condi- tion au bord	136

4.2.3	Une ambiguïté résiduelle	137
4.2.4	Contrainte d'intégrabilité	139
4.3	Résolution numérique du problème PS2	140
4.3.1	Reformulation du problème (4.1)	140
4.3.2	Solutions explicites du problème (4.34)	142
4.3.3	Désambiguïsation du problème par coupure de graphe	143
4.3.4	Régularisation « minimale »	145
4.3.5	Une méthode efficace de résolution du problème PS2	146
4.4	Applications	147
4.4.1	Séréophotométrie robuste avec seulement $m = 3$ images	149
4.4.2	Séréophotométrie couleur	150
4.4.3	Éclairages coplanaires	154
4.5	Conclusion du chapitre	158
 III Utilisation de modèles réalistes pour la stéréophotométrie		161
 5 Modèles d'éclairage non directionnels		165
5.1	Quelques modèles d'éclairage	165
5.1.1	De la luminance au vecteur d'éclairage	166
5.1.2	Source primaire lambertienne	167
5.1.3	Source sphérique	168
5.1.4	Source plane élémentaire	169
5.1.5	Source plane étendue	170
5.1.6	Champ d'éclairage non paramétré	178
5.2	Étalonnage photométrique	181
5.2.1	Estimation des paramètres d'une source lumineuse	181
5.2.2	Étalonnage du champ d'éclairage	186
5.3	Méthodes numériques de reconstruction 3D par stéréophotométrie sous des éclairages non directionnels	186
5.4	Conclusion du chapitre	191
 6 Modèles de réflectance non lambertiens		193
6.1	Séréophotométrie par apprentissage	194
6.1.1	Principe de l'approche par apprentissage	194
6.1.2	Extensions de l'approche par apprentissage	198
6.1.3	Conclusions sur l'approche par apprentissage	207
6.2	Réflectance diffuse non lambertienne	208
6.2.1	Modèle d'Oren-Nayar	208
6.2.2	Utilisation du modèle d'Oren-Nayar en stéréophotométrie	210
6.2.3	Estimation conjointe du relief, de l'albédo et de la rugosité	211
6.3	Réflectance spéculaire	212
6.3.1	Modèle de Ward	213
6.3.2	Inversion du modèle de Ward par une approche proximale	216

6.3.3	Résultats	219
6.4	Conclusion du chapitre	221
IV	Approche variationnelle de la stéréophotométrie	223
7	Résolution variationnelle de la stéréophotométrie différentielle	227
7.1	Résolution variationnelle en moindres carrés	228
7.1.1	Linéarisation de la stéréophotométrie : un obstacle à l'optimalité de la reconstruction 3D	228
7.1.2	Projection orthographique	230
7.1.3	Projection perspective	231
7.1.4	Modèle générique	232
7.1.5	Première variation	233
7.1.6	Schéma de type point fixe	238
7.1.7	Cas où l'albédo est inconnu	240
7.2	Estimation robuste du relief et de l'albédo	242
7.2.1	Prise en compte explicite des ombres propres	242
7.2.2	Modèle L^1 -TV	244
7.2.3	Itérations ADMM pour le problème L^1 -TV (7.62)	247
7.2.4	Itérations ADMM pour le problème L^1 - L^2 (7.82)	249
7.2.5	Optimisation alternée pour le problème L^2 - L^2 (7.87)	250
7.2.6	Itérations de point fixe pour le problème L^2 - L^2 non linéaire (7.90)	251
7.2.7	Analyse des paramètres	254
7.3	Conclusion du chapitre	257
8	Extensions de l'approche variationnelle	259
8.1	Stéréophotométrie colorée	260
8.1.1	Surface colorée éclairée par des sources blanches	261
8.1.2	Surface colorée éclairée par des sources colorées	262
8.2	Stéréophotométrie non calibrée, robuste et doublement colorée	265
8.2.1	La stéréophotométrie non calibrée : un problème bien posé	265
8.2.2	Une solution au problème de la stéréophotométrie non calibrée robuste	268
8.2.3	Stéréophotométrie non calibrée doublement colorée	269
8.3	Retour sur les modèles d'éclairage non directionnels	271
8.3.1	Modèle d'éclairage primaire explicite	273
8.3.2	Modèle d'éclairage sans paramètre	274
8.3.3	Stéréophotométrie semi-calibrée	276
8.4	Conclusion du chapitre	281
	Conclusion et perspectives	283
	Bibliographie	287

Introduction

La mesure tridimensionnelle consiste à obtenir une représentation en trois dimensions d'un objet ou d'une scène. Les situations où cette mesure est nécessaire sont variées. Un bon indicateur de cette diversité est le nombre élevé de solutions technologiques qui ont été mises en place, depuis la chaîne d'arpenteur jusqu'aux IRM de diffusion, en passant par les sonars, les télémètres laser ou encore les palpeurs mécaniques. Grâce à l'évolution constante de la qualité, de la résolution et du coût des appareils photographiques numériques, la mesure tridimensionnelle à partir de photographies constitue aujourd'hui une alternative, crédible et accessible, à ces outils. Elle est un problème inverse classique de vision par ordinateur, dont l'objectif est de retrouver la dimension qui a été perdue lors de l'acquisition de la photographie.

Lorsqu'une scène est photographiée par un appareil numérique, une partie des rayons lumineux qui l'atteignent sont renvoyés vers le capteur photosensible de l'appareil. Ce capteur mesure le nombre de photons reçus par chacun de ses photosites (*pixels*) pour former une *image* 2D de la scène 3D photographiée. Au cours de ce processus, une des trois dimensions est donc effectivement perdue : la *reconstruction 3D* consiste à retrouver cette dimension. Cette thèse décrit une technique particulière de reconstruction 3D photographique : la *stéréophotométrie* (« *photometric stereo* »), qui a été introduite à la fin des années 70 par Woodham [268], signifie littéralement l'analyse des volumes (*stereo*) par mesure (*metron*) de la lumière (*photos*).

Reconstruction 3D et scannage 3D

La finalité de la reconstruction 3D est la création de *modèles 3D*. Disposer d'un modèle 3D est utile, voire nécessaire, dans de nombreuses applications [182], comme par exemple l'inspection visuelle de surfaces [284], l'aide au diagnostic médical [243] ou encore la réalité augmentée [66]. Un modèle 3D est constitué d'un ensemble d'informations *géométriques* (position, orientation, etc.) et *photométriques* (couleur, texture, etc.). La connaissance de ces deux informations permet notamment de créer des images de synthèse, en simulant à la fois l'éclairage et l'appareil photographique. La trajectoire des rayons lumineux peut être calculée par ordinateur depuis la ou les sources lumineuses jusqu'à l'appareil photographique, après réflexions (éventuellement multiples) sur la surface des objets constituant la scène. Le *scannage 3D* est le problème dual de la synthèse d'images : à partir de photographies d'une scène, on cherche à caractériser géométriquement et photométriquement sa surface, en parcourant la trajectoire des rayons lumineux en sens *inverse*. La notion de scannage 3D englobe donc la reconstruction 3D de la scène (géométrie) et la mesure de sa *réflectance* (photométrie).

Les techniques de reconstruction 3D à partir de photographies sont regroupées sous le terme générique de *shape-from-X*, X signifiant que différents indices peuvent être utilisés (ombres, contours, etc.). Les principales techniques de *shape-from-X* sont regroupées

	Techniques géométriques	Techniques photométriques
Techniques mono-oculaires ($m = 1$ image)	Projection de lumière structurée [97] <i>Shape-from-shadows</i> [234] <i>Shape-from-contours</i> [36] <i>Shape-from-texture</i> [263] <i>Shape-from-template</i> [22]	<i>Shape-from-shading</i> [125]
Techniques multi-oculaires ($m \geq 2$ images)	<i>Structure-from-motion</i> [180] Stéréoscopie [113] <i>Shape-from-silhouettes</i> [116] <i>Shape-from-focus</i> [188]	Stéréophotométrie

TABLE 1 – Principales techniques de *shape-from-X*.

dans la table 1. Pour les classer, on doit se demander quel type d’indices (géométriques ou photométriques) et combien de photographies elles utilisent. Les techniques de *shape-from-X* actuellement les plus en vogue sont probablement la projection de lumière structurée et le *structure-from-motion*, pour lesquelles des solutions commerciales « clé en main » sont disponibles : à titre d’exemples, la Kinect V1 de Microsoft et le logiciel Photoscan utilisent, respectivement, la projection de lumière structurée et la technique du *structure-from-motion*.

Comme ne l’indique pas la table 1, les frontières entre les différentes techniques de *shape-from-X* sont floues : elles sont souvent combinées, afin de tirer parti des avantages de chacune. Par exemple, la quasi-totalité de ces techniques sont des techniques de *reconstruction* 3D. La seule qui permette également de retrouver les caractéristiques photométriques de la scène est la stéréophotométrie, qui est donc une technique de *scannage* 3D. Celle-ci est souvent couplée à d’autres techniques de reconstruction 3D [10, 118, 258]. En effet, comme nous le verrons au cours des huit chapitres de ce mémoire, une utilisation naïve de la stéréophotométrie induit un certain nombre de biais, qui ont conduit une partie de la communauté de vision par ordinateur à sous-estimer les possibilités de cette technique, en arguant qu’elle n’était pas capable de reconstruire avec précision les « basses fréquences ». Un des objectifs de cette thèse est de réhabiliter la stéréophotométrie en tant que technique de reconstruction 3D de précision. Pour cela, nous nous proposons d’identifier et de corriger les différentes sources de biais, à la fois dans la modélisation du problème et dans sa résolution.

Reconstruction 3D à partir de l’ombrage

Les techniques géométriques de *shape-from-X* reposent sur l’identification et l’analyse de points ou de régions d’intérêt dans l’image. En revanche, les techniques photométriques reposent sur l’analyse de la quantité de lumière reçue par chaque photosite du capteur de l’appareil photographique. Si ce dernier est étalonné photométriquement, alors le niveau de gris (ou le niveau de couleur) est proportionnel à la luminance émise

par la scène, qui est caractéristique de son relief. La reconstruction 3D photométrique est fondée sur l’analyse de l’*ombrage* (*shading*), un concept qui regroupe « l’ensemble des phénomènes à cause desquels un même élément de surface matérielle apparaît plus ou moins clair sur une photographie » [79]. Le terme de *shape-from-shading* (SFS) désigne la reconstruction 3D à partir de l’ombrage. Cette technique a été développée dans les années 70 au MIT, sous l’impulsion de Horn [125].

La luminance émise par une surface peut être modélisée par une fonction faisant intervenir l’éclairage, la réflectance de la surface et sa géométrie [78]. Le SFS, qui vise à inverser cette fonction pour reconstruire le relief de la scène, est un problème fondamentalement mal posé, selon le principe du trompe-l’œil qui est bien illustré par la « métaphore de l’atelier » d’Adelson et Pentland [4] (cf. figure 1). Si l’on demande à un peintre, à un sculpteur et à un « régisseur lumière » de réaliser une œuvre expliquant une image, ils apporteront chacun une réponse très différente, mais plausible. Le peintre supposera la surface plane et l’éclairage uniforme, les changements de clarté étant expliqués par des changements de couleur. Le sculpteur expliquera les changements de clarté uniquement par l’ombrage, qui résulte des variations locales d’orientation de la surface, l’éclairage et la réflectance étant supposés uniformes. Enfin, le régisseur lumière proposera une configuration d’éclairages sophistiquée placée face à une surface plane de réflectance uniforme.

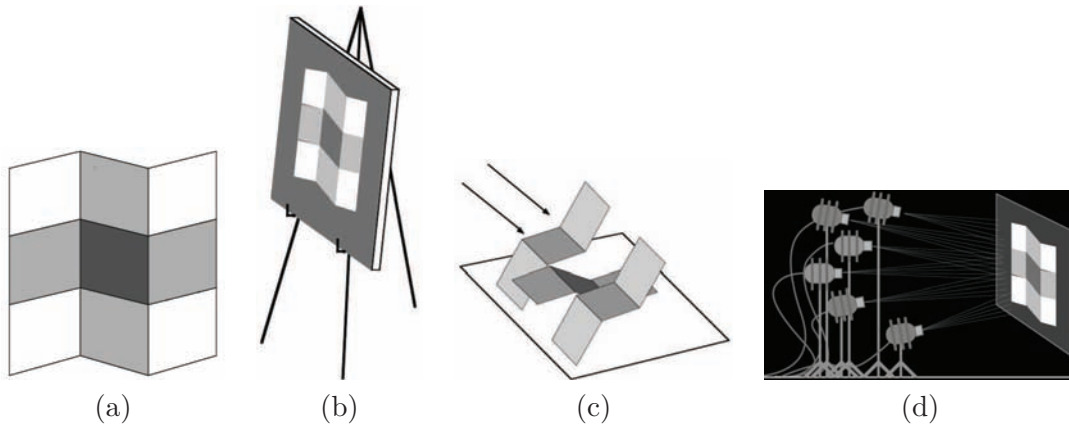


FIGURE 1 – La « métaphore de l’atelier » [4]. Pour expliquer l’image (a) en termes de réflectance, de relief et d’éclairage, (b) un peintre, (c) un sculpteur ou (d) un régisseur lumière apporteront trois réponses différentes, mais toutes trois plausibles. L’estimation de la contribution du sculpteur à partir d’une seule photographie d’une œuvre d’art combinant ces trois solutions constitue le problème du SFS. Elle n’est possible qu’en connaissant les contributions des deux autres artistes. Ces connaissances a priori ne sont pas nécessaires avec la stéréophotométrie : cette technique permet de déterminer sans ambiguïté les contributions du sculpteur et celles du peintre, pourvu que le régisseur lumière accepte de faire varier les éclairages.

Les prix de revient de la peinture, de la matière et des sources d'éclairage étant différents, la création d'une œuvre d'art combinant les compétences des trois artistes peut être formulée comme un problème de minimisation d'une fonction de coût combinant l'écart entre l'image à expliquer et la « reprojection » de l'œuvre, et le prix de revient total de l'œuvre. Les contraintes propres à chaque artiste constituent des a priori (au sens bayésien) sur chacun des éléments de l'œuvre (vus comme des variables aléatoires) : il est plus facile pour le peintre de réaliser des variations très fines de clarté, par des traits de crayon, que pour le régisseur lumière.

Cette métaphore est représentative du SFS, car estimer a posteriori la contribution du sculpteur à l'œuvre n'est possible qu'en disposant de connaissances a priori sur les contributions des deux autres artistes. Pour résoudre le SFS, on suppose ainsi généralement que la réflectance et l'éclairage sont connus, voire uniformes. On peut se passer de ces connaissances a priori avec la stéréophotométrie : en contrôlant l'éclairage, on peut déterminer sans ambiguïté les contributions du sculpteur et du peintre, i.e. le relief et la réflectance de la surface photographiée.

La fiabilité des connaissances a priori sur la réflectance et l'éclairage est déterminante pour reconstruire de façon précise le relief par SFS. Comme l'illustre la figure 2, tirée d'une contribution récente sur l'utilisation des imprimantes 3D en traitement mathématique des images [69], une hypothèse erronée sur la réflectance et l'éclairage se traduit par une interprétation erronée du relief.

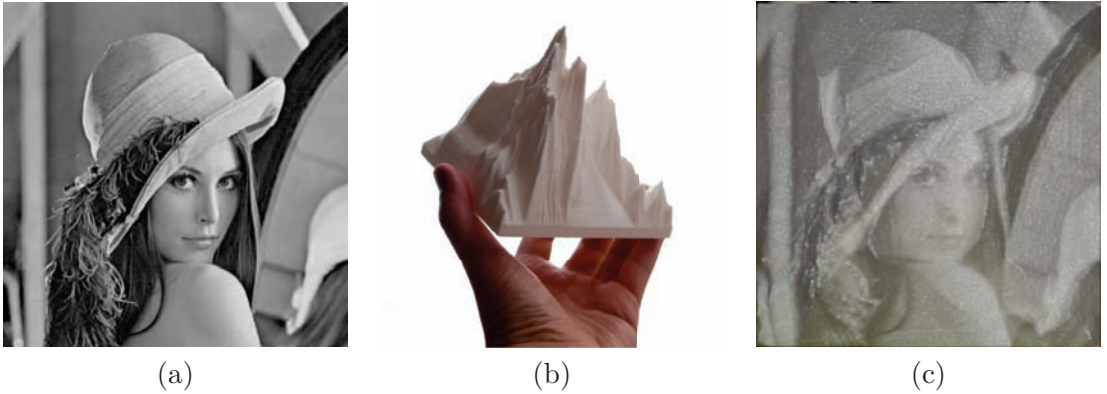


FIGURE 2 – Principe du trompe-l'œil appliqué au SFS [69]. (a) Exemple bien connu de photographie en niveaux de gris. (b) Relief estimé en appliquant la technique du SFS à l'image (a), en supposant (à tort) la réflectance uniforme et l'éclairage frontal. La reconstruction semble totalement décorrélée du relief réel. Pourtant, en prenant une photographie du relief (b) de dessus, et en utilisant le flash comme unique source de lumière, on obtient l'image (c), qui est très similaire à (a). Le relief (b) constitue donc effectivement une explication plausible de l'image (a) : ce sont les connaissances a priori sur la réflectance et l'éclairage qui sont erronées.

Afin de se passer de la connaissance a priori de la réflectance et de l'éclairage, il a été suggéré dans [19] de formuler le problème de l'estimation simultanée du relief, de l'albédo et de l'éclairage comme un problème d'inférence bayésienne. Le problème à résoudre comportant plus d'inconnues que de données, la pondération effective des connaissances a priori est cependant extrêmement délicate. De plus, même en connaissant parfaitement l'éclairage et la réflectance, le SFS reste un problème mal posé. En dehors de certains points singuliers, il existe en effet une infinité de normales qui expliquent un niveau de gris. Si la surface est diffusante (*lambertienne*), ces normales forment un cône de révolution autour de la direction d'éclairage. Il est donc très difficile de résoudre *localement* le SFS. Les premières approches du SFS, qui étaient locales, ont toutefois permis de se convaincre de l'intérêt de l'analyse de l'ombrage, comme en témoigne l'exemple de la figure 3, tiré de [201]. En observant que l'œil humain, tout comme un algorithme de SFS, était parfaitement capable d'estimer un relief tel que celui de la figure 3 uniquement à partir de l'ombrage, Pentland prédisait dès 1984 l'intérêt des techniques photométriques pour la compréhension de la vision humaine. Ce n'est toutefois que très récemment que des applications ont été proposées.

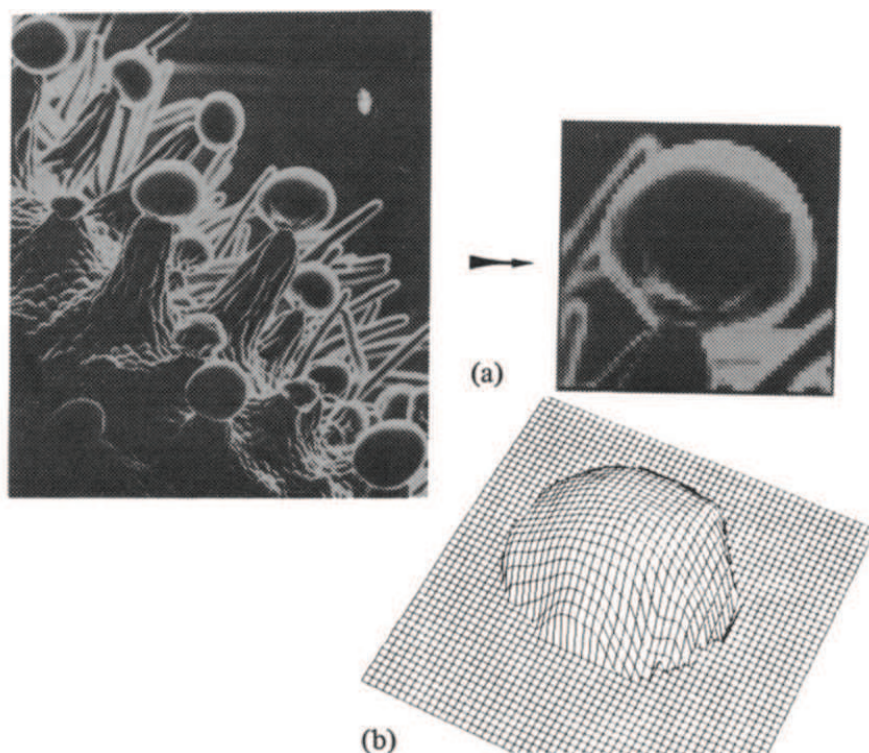


FIGURE 3 – (a) Image, obtenue par un microscope électronique, d'un trichrome d'une plante de *cannabis sativa*. (b) Reconstruction 3D à partir de l'ombrage. Source des images originales : [232]. Si l'intérêt applicatif des techniques photométriques de reconstruction 3D a été identifié très tôt [201], leur mise en œuvre effective est bien plus récente.

Un pas important vers la résolution du SFS a été réalisé lorsque les problèmes inverses de la vision par ordinateur ont attiré l'attention des mathématiciens [165]. Des approches différentielles efficaces ont alors été suggérées pour résoudre l'équation *eikonale*, qui est l'équation fondamentale du SFS. Cette équation lie la pente de la surface au niveau de gris de l'image. La résolution de cette équation aux dérivées partielles étant par nature *globale*, les ambiguïtés du SFS sont réduites par rapport à l'approche locale. Elles ne sont toutefois pas éliminées, car l'ambiguïté concave/convexe ne peut être levée sans connaissance a priori sur la surface. Le SFS reste donc, malgré de nombreux efforts, un problème mal posé. On peut par conséquent se demander s'il ne s'agit pas d'un simple « exercice de style ».

En fait, les applications du SFS sont récentes, et concernent des cas très particuliers où la géométrie de la surface à reconstruire est contrainte. Par exemple, si l'on sait par avance que la surface photographiée est une surface *développable* (surface de courbure gaussienne nulle en tout point, que l'on peut « dérouler » comme une feuille de papier), le SFS devient bien posé. Une application naturelle concerne la simulation de la mise à plat de documents [68, 203]. Mais à part ces quelques cas particuliers, la seule vraie solution pour lever les ambiguïtés du SFS consiste à utiliser plusieurs photographies, prises sous différents éclairages (cf. figure 4) : c'est le principe de la stéréophotométrie.

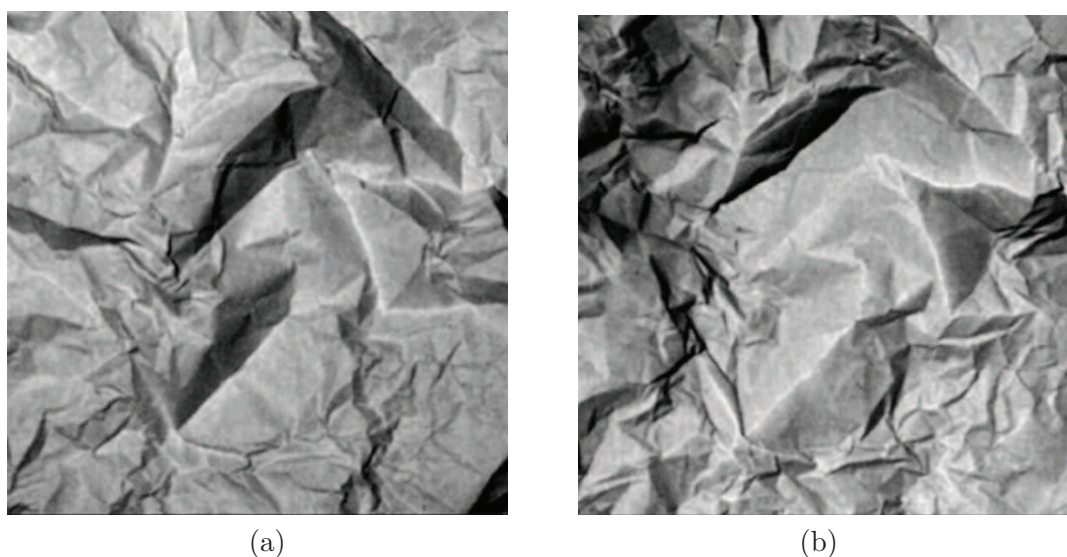


FIGURE 4 – Il est extrêmement difficile d'interpréter l'image (a) en termes de relief, même en sachant que la source lumineuse est placée « en haut à gauche ». L'interprétation du relief est facilitée si l'on dispose d'une image telle que (b), prise sous un autre éclairage (source lumineuse placée « en bas à droite »). Source : [148].

La stéréophotométrie

Même en admettant qu'il existe une formulation bien posée du SFS, la robustesse de méthodes numériques permettant sa résolution est impossible à garantir. Si la résolution *exacte* d'une équation liant le niveau de gris à la pente est déjà difficile, tout écart au modèle dans les données dégradera l'estimation du relief. La parade naturelle consiste à prendre davantage de photographies : on peut raisonnablement espérer que l'estimation du relief sera à la fois mieux posée et plus robuste que si l'on utilise une seule image (cf. partie I). Nous verrons dans le chapitre 1 que l'on peut effectivement déterminer de façon exacte le relief à partir de $m = 3$ photographies prises sous trois éclairages directionnels non coplanaires. L'estimation conjointe, en chaque point, de la normale et de l'albédo (qui est caractéristique de la réflectance lambertienne) peut en effet être formulée comme la résolution d'un système de $m = 3$ équations linéaires à trois inconnues. De plus, l'estimation robuste est possible dès lors que $m > 3$. Il est donc logique de choisir l'approche locale pour résoudre la stéréophotométrie, au contraire du SFS où l'approche globale permet de limiter les ambiguïtés. Cependant, la connaissance des normales ne suffit pas à caractériser géométriquement la scène photographiée. Pour obtenir une représentation 3D complète, il reste à *intégrer* le champ de normales estimé, pour en déduire la profondeur : c'est uniquement lors de cette seconde étape qu'une formulation différentielle, donc globale, est utilisée. La résolution de ce problème n'est pas aussi simple que pourrait le laisser penser la simplicité de sa formulation. Son étude fait l'objet du chapitre 2, où nous proposons notamment une méthode d'intégration qui est à la fois robuste et rapide.

La stéréophotométrie peut cependant être mal posée. Comme nous le montrons dans la partie II, il existe deux situations où elle ne peut être résolue localement. Nous verrons dans le chapitre 3 que, lorsque les éclairages sont inconnus (stéréophotométrie *non calibrée*), l'estimation locale des normales constitue un problème mal posé. Ce cas d'usage est donc plus proche du SFS que de la stéréophotométrie « classique ». Comme pour le SFS [126], il faut formuler le problème d'estimation des normales de façon globale et imposer la contrainte d'intégrabilité [93]. Mais même ainsi, le problème reste mal posé et requiert une connaissance supplémentaire. Nous verrons que l'introduction d'un critère de régularité de type *variation totale* sur la profondeur réduit l'ambiguïté à une ambiguïté dite « de bas-relief », qui peut être levée par régularisation. La stéréophotométrie est également mal posée dans la situation intermédiaire entre $m = 1$ (SFS) et $m \geq 3$. Il existe généralement deux normales qui expliquent une paire de niveaux de gris, lorsque les éclairages et l'albédo sont connus. Le problème de la stéréophotométrie devient alors combinatoire : nous proposons dans le chapitre 4 une approche markovienne de la stéréophotométrie à $m = 2$ images, reformulée comme un problème d'étiquetage. Le critère permettant de lever l'ambiguïté dans l'estimation est une nouvelle fois l'intégrabilité du champ de normales.

L'étude de modèles non linéaires d'éclairage et de réflectance a permis de lever certaines ambiguïtés du SFS. De tels modèles peuvent bien entendu être également utilisés dans le contexte de la stéréophotométrie, comme cela est illustré dans la partie III.

Contrairement au SFS [208], l'introduction d'un modèle d'éclairage ne sert pas en stéréophotométrie à limiter les ambiguïtés, mais uniquement à atteindre une certaine précision dans les applications. L'étude de modèles d'éclairage non directionnels et de leur utilisation en stéréophotométrie fait l'objet du chapitre 5. L'introduction de modèles de réflectance plus réalistes que le modèle lambertien a été suggérée récemment pour améliorer les résultats du SFS [253]. De tels modèles peuvent aussi être utilisés en stéréophotométrie, comme nous le montrons dans le chapitre 6, au travers d'une approche où le modèle est appris et d'une étude de deux modèles de réflectance explicites plus réalistes que le modèle lambertien.

L'utilisation de modèles réalistes fait toutefois perdre à la formulation locale de la stéréophotométrie son caractère linéaire, sur lequel reposent en grande partie les méthodes d'estimation robuste existantes. Pour cette raison, et pour rendre mieux posés les problèmes étudiés dans la partie II, nous proposons dans la partie IV une nouvelle approche de la stéréophotométrie. Comme cela a été fait pour le SFS, nous suggérons de préférer la formulation différentielle (globale) à la formulation non différentielle (locale). Après avoir introduit dans le chapitre 7 un ensemble de méthodes numériques, fondées sur le calcul des variations, pour résoudre la formulation différentielle de la stéréophotométrie, nous proposons un certain nombre d'extensions de cette approche dans le chapitre 8.

Les contributions compilées dans ce mémoire portent donc sur les aspects théoriques, numériques et applicatifs de la stéréophotométrie. Au travers de l'analyse de méthodes d'estimation robuste, de problèmes mal posés et de modèles réalistes, les trois premières parties visent à identifier et à limiter les sources de biais, afin de faire de la stéréophotométrie une technique de reconstruction 3D fiable. Elles nous permettront également de mettre en évidence certaines limites de cette technique, auxquelles la dernière partie apporte quelques éléments de réponse.

Première partie

Résolution robuste de la stéréophotométrie

Avant-propos

Dans cette partie, nous proposons une analyse du problème de la stéréophotométrie dans les conditions les plus favorables, i.e. en supposant que la surface à reconstruire est lambertienne et qu'elle est vue sous $m \geq 3$ éclairages directionnels, non coplanaires, connus. Ces hypothèses permettent de formuler le problème de façon relativement élémentaire, et garantissent son caractère bien posé. En effet, contrairement au problème du *shape-from-shading* (SFS), l'estimation conjointe de la normale et de l'albédo est un problème linéaire très facile à résoudre. La notion de voisinage, qui est centrale en SFS pour lever les ambiguïtés, n'intervient qu'une fois les normales estimées, lorsque l'on cherche à en déduire le relief.

Les applications de la stéréophotométrie semblent toutefois limitées, car la résolution s'avère peu robuste à des phénomènes difficiles à éviter, comme les ombres ou les reflets brillants. Nous montrons, dans le chapitre 1, que l'utilisation de méthodes d'estimation robuste et/ou de prétraitements appropriés permet de pallier aux limites que présente l'estimation des normales et de l'albédo par une approche plus élémentaire. Une fois les normales estimées, il reste à les intégrer pour en déduire la profondeur. Cette seconde étape fait l'objet d'une étude approfondie dans le chapitre 2, où nous montrons notamment comment tenir compte des discontinuités de profondeur.

Les méthodes présentées dans cette première partie visent donc à éliminer les sources de biais dans la formulation classique de la stéréophotométrie, afin de pouvoir l'utiliser en conditions réalistes. Il nous semble en effet important de savoir jusqu'à quel point la stéréophotométrie, en nous permettant de mesurer le relief et la couleur d'une scène à partir d'un seul point de vue, peut prétendre au statut de technique de métrologie 3D, si bien sûr elle est utilisée en conditions très contrôlées, e.g. en laboratoire.

Estimation robuste des normales

Dans ce chapitre, nous étudions différentes méthodes d'estimation des normales par stéréophotométrie. Le modèle utilisé par l'approche classique est rappelé dans le paragraphe 1.1 : si les éclairages sont supposés directionnels et connus, et que la surface est supposée lambertienne, alors l'estimation des normales, conjointement à celle de l'albédo, constitue un problème linéaire facile à résoudre. Par un raisonnement statistique, nous montrons cependant les limites de cette approche, et notamment le biais induit par la présence d'écarts au modèle. Nous étudions ensuite, dans le paragraphe 1.2, divers prétraitements permettant d'éviter que la présence de tels écarts ne dégrade les résultats. Ces prétraitements compensent effectivement certains des écarts au modèle, mais on doit leur préférer les méthodes d'estimation robuste, que nous présentons dans le paragraphe 1.3.

1.1 Limites de la stéréophotométrie classique

1.1.1 Modèle de formation des images

Dans cette partie, nous nous intéressons à la modélisation la plus simple de la stéréophotométrie, qui correspond au cas d'une surface régulière, lambertienne, éclairée successivement par $m \geq 3$ faisceaux lumineux parallèles (mêmes directions d'éclairage en tous points de la surface) et uniformes (mêmes intensités).

Projection orthographique et projection perspective

Soit $\mathcal{S}$ la surface observée, qui consiste en un ensemble de points $\mathbf{x} = [x, y, z]^\top \in \mathbb{R}^3$ dont les coordonnées sont définies relativement à un repère tridimensionnel direct $\mathbf{O}xyz$, ayant pour origine $\mathbf{O}$ le centre optique de l'appareil photographique, dont les axes $\mathbf{O}x$ et $\mathbf{O}y$ sont parallèles au plan image, et dont l'axe $\mathbf{O}z$ est parallèle à l'axe optique (par convention, $z < 0$ pour un point placé à l'avant de l'appareil, cf. figure 1.1). Par projection (orthographique ou perspective), chaque point visible de la surface est en bijection avec un point image $(u, v) \in \mathbb{R}^2$. Notons que les coordonnées (u, v) d'un point image sont définies dans un repère bidimensionnel dont les axes sont parallèles aux axes $\mathbf{O}x$ et $\mathbf{O}y$, et dont l'origine sera précisée plus loin. Nous assimilons les pixels à des carrés, et choisissons comme unité de référence la distance euclidienne entre les centres de deux pixels voisins. On peut alors paramétrer la surface $\mathcal{S}$ sur un domaine $\Omega \subset \mathbb{R}^2$, appelé *domaine de reconstruction*, défini comme la projection de la surface sur le plan image :

$$\mathcal{S} = \left\{ \mathbf{x}(u, v) = [x(u, v), y(u, v), z(u, v)]^\top, (u, v) \in \Omega \right\} \quad (1.1)$$

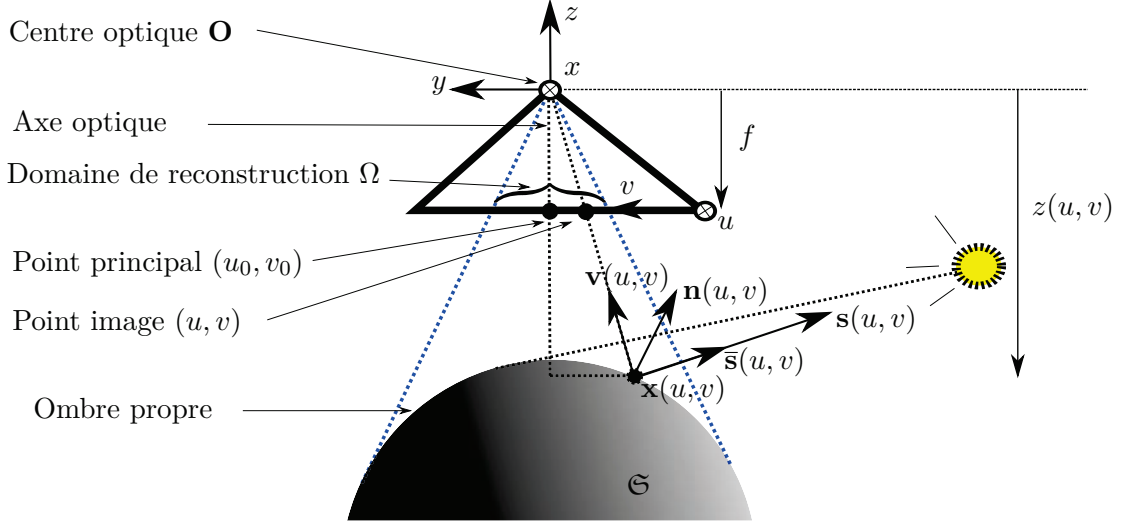


FIGURE 1.1 – Sous l’hypothèse lambertienne, le niveau de gris $I(u, v)$ en un point image (u, v) est proportionnel au produit scalaire de la normale $\mathbf{n}(u, v)$ au point $\mathbf{x}(u, v) \in \mathfrak{S}$ avec le vecteur $\mathbf{s}(u, v) = \phi(\mathbf{x}(u, v))\bar{\mathbf{s}}(\mathbf{x}(u, v))$, qui est colinéaire au vecteur unitaire $\bar{\mathbf{s}}(\mathbf{x}(u, v))$ orienté vers la source lumineuse, et dont la norme $\phi(\mathbf{x}(u, v))$ représente l’intensité lumineuse. Si $\mathbf{n}(u, v) \cdot \mathbf{s}(u, v) < 0$, alors le point $\mathbf{x}(u, v)$ est dans l’ombre (propre), et le niveau de gris $I(u, v)$ est nul. L’approche classique de la stéréophotométrie consiste à d’abord estimer $\mathbf{n}(u, v)$ en tout point image $(u, v) \in \Omega$ (domaine de reconstruction), puis à intégrer les normales estimées afin de retrouver la carte de profondeur z , qui caractérise le relief de la surface $\mathfrak{S}$.

Rappelons qu’avec le modèle de projection orthographique¹ :

$$\mathbf{x}(u, v) = [u - u_0, v - v_0, z(u, v)]^\top \quad (1.2)$$

où (u_0, v_0) sont les coordonnées du *point principal* de l’appareil photographique, qui se trouve à l’intersection de l’axe optique et du plan image (cf. figure 1.1). On peut connaître u_0 et v_0 par *étalonnage* de l’appareil photographique, mais cela n’est pas indispensable sous l’hypothèse orthographique. En effet, en translatant de $[u_0, v_0, 0]^\top$ l’ensemble des points $\mathbf{x}(u, v)$ définis par (1.2), on ne fait que translater la surface $\mathfrak{S}$ « en bloc », ce qui ne lui cause aucune déformation. On peut donc simplifier l’écriture (1.2) sous la forme suivante :

$$\mathbf{x}(u, v) = [u, v, z(u, v)]^\top \quad (1.3)$$

sans qu’il soit nécessaire de positionner l’origine du repère 2D en un point particulier².

1. Cette écriture est en fait celle de la projection orthogonale, qui est une version simplifiée de la projection orthographique lorsque le grandissement est égal à 1.

2. Par défaut, cette origine est généralement placée dans le coin supérieur gauche de l’image.

Cette simplification est impossible avec le modèle de projection perspective, pour laquelle (1.2) doit être remplacé par :

$$\mathbf{x}(u, v) = z(u, v) \left[\frac{u - u_0}{f}, \frac{v - v_0}{f}, 1 \right]^\top \quad (1.4)$$

car les deux premières coordonnées de $\mathbf{x}(u, v)$ dépendent de la profondeur $z(u, v)$. Dans l'équation (1.4), f désigne la *distance focale* (image) de l'appareil photographique. Avec l'orientation de l'axe $\mathbf{O}z$ indiquée sur la figure 1.1, $f < 0$.

Pour ces deux modèles de projection, la connaissance de $z(u, v)$ permet donc de déterminer sans ambiguïté le point objet $\mathbf{x}(u, v)$ (sous réserve que les *paramètres internes* (u_0, v_0, f) de l'appareil photographique soient connus, dans le cas de la projection perspective), de telle sorte que la partie visible de la surface est entièrement décrite par le *carreau de Monge* :

$$\mathcal{M} = \left\{ [u, v, z(u, v)]^\top, (u, v) \in \Omega \right\} \quad (1.5)$$

où z constitue la *carte de profondeur*³, qui associe à chaque point image (u, v) la troisième coordonnée cartésienne du point objet qui lui est conjugué. Le problème de la reconstruction 3D consiste à reconstituer ce carreau de Monge, i.e. à associer à chaque point image (u, v) une profondeur $z(u, v)$.

Normale à la surface

Contrairement au *shape-from-shading*, il est plus facile de résoudre la stéréophotométrie en commençant par estimer les normales, puis en intégrant les normales estimées afin d'obtenir une carte de profondeur z , plutôt qu'en estimant z directement (bien que nous remettons cette affirmation en question dans la partie IV). Nous supposons donc, dans un premier temps, que z est une fonction différentiable sur Ω , afin de pouvoir définir, en tout point $\mathbf{x} \in \mathfrak{S}$, le vecteur normal à la surface, unitaire, sortant. En utilisant le paramétrage de la surface par les coordonnées image (u, v) , le *champ de normales* est défini de la façon suivante sur le domaine de reconstruction Ω :

$$\begin{aligned} \mathbf{n}: \Omega \subset \mathbb{R}^2 &\longrightarrow \mathbb{S}^2 \subset \mathbb{R}^3 \\ (u, v) &\longmapsto \mathbf{n}(u, v) = [n_1(u, v), n_2(u, v), n_3(u, v)]^\top \end{aligned} \quad (1.6)$$

où $\mathbb{S}^2$ désigne la sphère unité de $\mathbb{R}^3$ ($\|\mathbf{n}(u, v)\| = 1$, par définition). Signalons au passage que, pour une surface *régulière*, la normale $\mathbf{n}(u, v)$ au point $\mathbf{x}(u, v)$ s'écrit :

$$\mathbf{n}(u, v) = \pm \frac{\partial_u \mathbf{x}(u, v) \times \partial_v \mathbf{x}(u, v)}{\|\partial_u \mathbf{x}(u, v) \times \partial_v \mathbf{x}(u, v)\|} \quad (1.7)$$

de telle sorte que, d'après (1.3) et (1.4), ce vecteur est une fonction du gradient $\nabla z = [\partial_u z, \partial_v z]^\top$ de la profondeur z . Cela explique pourquoi l'estimation de z , connaissant $\mathbf{n}$, est un problème d'*intégration* (cf. chapitre 2).

3. Le terme de « carte » est une traduction abusive, fréquemment utilisée en vision par ordinateur, de l'anglais *map*. On devrait lui préférer celui de « fonction ».

Vecteur d'éclairage

Nous décrivons l'éclairage incident au point $\mathbf{x} \in \mathfrak{S}$ par un vecteur $\mathbf{s}(\mathbf{x})$, de direction $\bar{\mathbf{s}}(\mathbf{x}) = \frac{\mathbf{s}(\mathbf{x})}{\|\mathbf{s}(\mathbf{x})\|}$, et dont la norme $\phi(\mathbf{x}) = \|\mathbf{s}(\mathbf{x})\|$ représente l'intensité lumineuse⁴. Par convention, ce vecteur est orienté vers la source. Nous noterons abusivement $\mathbf{s}(u, v) = \mathbf{s}(\mathbf{x}(u, v))$:

$$\begin{aligned} \mathbf{s}: \Omega \subset \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (u, v) &\longmapsto \mathbf{s}(u, v) = \phi(\mathbf{x}(u, v))\bar{\mathbf{s}}(\mathbf{x}(u, v)) \end{aligned} \quad (1.8)$$

En supposant la source lumineuse infiniment distante de la scène, l'éclairage est *directionnel*, i.e. son intensité et sa direction sont les mêmes en tout point de la surface :

$$\mathbf{s}(u, v) = \mathbf{s} \quad (1.9)$$

Notons que $\mathbf{s}(u, v)$ dépend a priori de la fréquence lumineuse. Nous supposons dans un premier temps que la source est « blanche », i.e. nous négligeons cette dépendance spectrale (nous reviendrons sur ce point dans le paragraphe 1.5).

Surface lambertienne

Comme cela est montré dans [78], un appareil photographique photométriquement étalonné peut servir d'instrument de mesure de la luminance. En effet, si l'on a corrigé les effets de *vigneting* et d'assombrissement de l'image « en $\cos^4 \alpha$ » de l'appareil photographique [5, 78, 127], l'éclairement $E_i(u, v)$ au point image (u, v) est proportionnel à la *luminance* émise par le point $\mathbf{x}(u, v)$ dans la direction d'observation $\mathbf{v}(u, v)$ ⁵ (cf. figure 1.1) :

$$E_i(u, v) \propto L(\mathbf{x}(u, v), \mathbf{v}(u, v)) \quad (1.10)$$

De plus, après correction des non linéarités de la courbe de réponse du capteur photosensible, le niveau de gris est proportionnel à l'éclairement :

$$I(u, v) \propto E_i(u, v) \quad (1.11)$$

Une surface *lambertienne* est une surface pour laquelle la luminance émise ne dépend pas de la direction d'observation $\mathbf{v}(u, v)$. On peut montrer dans ce cas que :

$$L(\mathbf{x}(u, v)) = \frac{\rho(\mathbf{x}(u, v))}{\pi} E(\mathbf{x}(u, v)) \quad (1.12)$$

Dans cette écriture, l'*albédo* $\rho(\mathbf{x}(u, v))$ mesure la proportion d'énergie lumineuse réémise par la surface au point $\mathbf{x}(u, v)$: il s'agit donc d'une grandeur sans unité comprise entre 0 et 1. D'autre part, $E(\mathbf{x}(u, v))$ désigne l'éclairement reçu par la surface au point $\mathbf{x}(u, v)$, qui ne doit pas être confondu avec l'éclairement $E_i(u, v)$ reçu par l'image au point (u, v) .

4. En réalité, le terme exact pour désigner $\phi(\mathbf{x})$ est la *densité de flux lumineux*.

5. Du moins, tant que le photosite n'est pas saturé. Dans la suite de ce mémoire, les phénomènes de saturation seront traités comme des données aberrantes.

Si l'éclairage, supposé directionnel, est caractérisé par le vecteur $\mathbf{s}$, alors cet éclairage s'écrit :

$$E(\mathbf{x}(u, v)) = \mathbf{s} \cdot \mathbf{n}(u, v) \quad (1.13)$$

On déduit de (1.10), (1.11), (1.12) et (1.13) :

$$I(u, v) \propto \rho(u, v) \mathbf{s} \cdot \mathbf{n}(u, v) \quad (1.14)$$

en utilisant la notation abusive $\rho(u, v) \equiv \rho(\mathbf{x}(u, v))$.

Comme le niveau de gris fourni par un appareil photographique est positif, la relation de proportionnalité (1.14) n'est en fait valide que si $\mathbf{n}(u, v) \cdot \mathbf{s} \geq 0$, i.e. si le point $\mathbf{x}(u, v)$ ne se situe pas dans une *ombre propre*. Nous devons donc remplacer (1.14) par :

$$I(u, v) \propto \rho(u, v) \max\{0, \mathbf{s} \cdot \mathbf{n}(u, v)\} \quad (1.15)$$

Dorénavant, nous écrirons la relation (1.15) sous forme d'égalité :

$$I(u, v) = \rho(u, v) \max\{0, \mathbf{s} \cdot \mathbf{n}(u, v)\} \quad (1.16)$$

tout en sachant qu'il s'agit en fait d'une relation de proportionnalité. Cela a comme conséquence que $\rho(u, v)$ représente dorénavant le produit de l'albédo par le coefficient de la relation de proportionnalité (1.15) : dans la suite de ce mémoire, $\rho(u, v)$ ne sera donc plus tenu à être compris entre 0 et 1, mais nous continuerons à l'appeler albédo.

De même que pour l'éclairage, l'albédo d'une surface est généralement fonction de la longueur d'onde. L'albédo devrait donc être défini par une courbe de réponse, ou au moins par un ensemble de valeurs échantillonnées sur le spectre visible. Nous supposons d'abord, dans un souci de simplicité, que l'appareil photographique mesure des niveaux de gris et que l'albédo est un scalaire (positif ou nul).

À partir de m images en niveaux de gris $I^1, \dots, I^m$ prises sous m éclairages directionnels $\mathbf{s}^1, \dots, \mathbf{s}^m$, la première étape de l'approche classique de la stéréophotométrie consiste à résoudre un système de m équations correspondant au modèle (1.16) :

$$\begin{cases} I^1(u, v) = \rho(u, v) \max\{0, \mathbf{s}^1 \cdot \mathbf{n}(u, v)\} \\ \vdots \\ I^m(u, v) = \rho(u, v) \max\{0, \mathbf{s}^m \cdot \mathbf{n}(u, v)\} \end{cases} \quad (1.17)$$

Dans ce système, les niveaux de gris $I^i(u, v)$, $i \in [1, m]$, constituent les données, tandis que les inconnues $\mathbf{n}(u, v)$ sont caractéristiques du relief de la surface observée. L'albédo $\rho(u, v)$, qui caractérise la *réflectance* d'un matériau lambertien, est en général également inconnu : contrairement au SFS, la stéréophotométrie permet effectivement d'estimer l'albédo de la surface observée. Enfin, les vecteurs $\mathbf{s}^i$ sont souvent supposés connus (*stéréophotométrie calibrée*), mais nous verrons dans le chapitre 3 comment traiter le cas où les éclairages sont inconnus (*stéréophotométrie non calibrée*).

1.1.2 Approximation linéaire du modèle lambertien

Dans le système (1.17), l'opérateur « max » et la contrainte $\|\mathbf{n}(u, v)\| = 1$ représentent deux difficultés du point de vue de la résolution, à cause de leur non linéarité. Cependant, en négligeant les ombres propres, et en introduisant l'inconnue $\mathbf{m}(u, v) = \rho(u, v)\mathbf{n}(u, v)$ qui, au contraire de $\mathbf{n}(u, v)$, est un vecteur de $\mathbb{R}^3$ non contraint, on peut transformer le système (1.17) en un système linéaire, qui s'écrit sous forme matricielle :

$$\underbrace{\begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix}}_{\mathbf{i}(u, v) \in \mathbb{R}^m} = \underbrace{\begin{bmatrix} \mathbf{s}^{1\top} \\ \vdots \\ \mathbf{s}^{m\top} \end{bmatrix}}_{\mathbf{S} \in \mathbb{R}^{m \times 3}} \underbrace{\begin{bmatrix} \rho(u, v)\mathbf{n}(u, v) \end{bmatrix}}_{\mathbf{m}(u, v) \in \mathbb{R}^3} \quad (1.18)$$

L'estimation de l'albédo ne complique donc pas la résolution de la stéréophotométrie. En réalité, l'introduction de cette « inconnue supplémentaire » n'a pas d'autre objectif que de rendre le problème linéaire⁶.

L'existence et l'unicité de la solution du système linéaire (1.18) dépend du nombre m d'éclairages et du rang de la matrice $\mathbf{S}$:

- Si $m = 3$ et $\text{rang}(\mathbf{S}) = 3$, alors le système admet une unique solution exacte :

$$\mathbf{m}(u, v) = \mathbf{S}^{-1}\mathbf{i}(u, v) \quad (1.19)$$

- Si $m > 3$ et $\text{rang}(\mathbf{S}) = 3$, alors le système est surcontraint et admet une unique solution approchée en moindres carrés, sur laquelle nous reviendrons.
- Si $\text{rang}(\mathbf{S}) < 3$, alors le système admet une infinité de solutions exactes. Cette configuration *mal posée* sera étudiée dans le chapitre 4.

Ces trois cas correspondent à des configurations très différentes. On se retrouve dans le premier cas lorsqu'on prend une seule photographie en couleur et qu'on utilise $m = 3$ sources lumineuses colorées en rouge, vert et bleu [117]. Le deuxième cas correspond à la situation la plus courante. Quant au dernier cas, il concerne la stéréophotométrie à $m = 2$ images ou à $m \geq 3$ images prises sous des éclairages coplanaires, e.g. sous le soleil à différentes heures de la journée [231].

Nous nous intéressons dans ce chapitre aux cas où $m \geq 3$ et $\text{rang}(\mathbf{S}) = 3$, i.e. au premier et au deuxième des trois cas énumérés ci-dessus : le système linéaire (1.18) peut alors être résolu de façon exacte ou approchée, ce qui fournit une estimation du vecteur $\mathbf{m}(u, v)$, de laquelle on déduit ensuite très simplement la normale et l'albédo :

$$\mathbf{n}(u, v) = \frac{\mathbf{m}(u, v)}{\|\mathbf{m}(u, v)\|} \quad (1.20)$$

$$\rho(u, v) = \|\mathbf{m}(u, v)\| \quad (1.21)$$

6. Dans la partie IV, nous reviendrons sur ce choix qui, d'une certaine façon, limite l'intérêt de la stéréophotométrie.

Aussi naïve que puisse paraître cette approche, elle permet de retrouver certains des détails les plus fins d'une surface, comme le montre l'exemple de la figure 1.2, qui est obtenu à partir de $m = 3$ images réelles⁷.

La linéarisation du système (1.17) a cependant un prix : tout écart au modèle lambertien linéaire (1.18) induit un biais. Sur l'exemple de la figure 1.2, l'objet photographié est une statue en plâtre, dont l'albédo *réel* est uniforme. Ce biais se traduit par le fait que l'albédo *estimé* n'est pas uniforme (cf. figure 1.2-e). Pour éviter ces problèmes, dus à la présence de données statistiquement aberrantes (*outliers*), on peut :

- soit limiter les écarts au modèle en l'affinant ;
- soit modifier les données afin de les faire « coller » au modèle ;
- soit enfin changer d'estimateur, afin de garantir une certaine *robustesse* aux données aberrantes.

Nous étudierons la première approche dans la partie III. Plus précisément, des modèles d'éclairage non directionnels seront étudiés dans le chapitre 5, et des modèles de surface non lambertiens dans le chapitre 6. Dans le présent chapitre, nous nous intéressons à la deuxième et à la troisième des approches énumérées ci-dessus.

1.1.3 Considérations statistiques

En l'absence de donnée aberrante, le choix d'un estimateur particulier a peu d'importance : il est alors raisonnable de choisir l'estimateur le plus simple, qui est généralement l'estimateur en moindres carrés ordinaires. En présence de données aberrantes, ce choix est déterminant car il influe directement sur la solution [256]⁸.

Supposons par exemple que les données diffèrent du modèle lambertien linéaire (1.18) par un bruit additif :

$$\mathbf{i}(u, v) = \mathbf{S}\mathbf{m}(u, v) + \boldsymbol{\epsilon}(u, v) \quad (1.22)$$

où $\boldsymbol{\epsilon}(u, v) = [\epsilon^1(u, v), \dots, \epsilon^m(u, v)]^\top$ est un vecteur aléatoire de dimension m . On suppose donc que le bruit ne porte que sur les données, et pas sur l'étalonnage des éclairages. Si :

- les erreurs sont *non corrélées*, i.e. $\text{cov}(\epsilon^i(u, v), \epsilon^j(u, v)) = 0, \forall (i, j) \in [1, m]^2, i \neq j$,
- le bruit est *exogène*, i.e. $E[\epsilon^i(u, v)] = 0, \forall i \in [1, m]$,
- les composantes du bruit sont *homoscédastiques*, i.e. $\text{var}(\epsilon^i(u, v)) = \sigma(u, v)^2, \forall i \in [1, m]$ (les m variances sont toutes égales),

alors l'estimateur en moindres carrés ordinaires⁹ :

$$\mathbf{m}(u, v) = \underset{\mathbf{m} \in \mathbb{R}^3}{\text{argmin}} \|\mathbf{i}(u, v) - \mathbf{S}\mathbf{m}\|^2 = \underset{\mathbf{m} \in \mathbb{R}^3}{\text{argmin}} \sum_{i=1}^m [I^i(u, v) - \mathbf{s}^i \cdot \mathbf{m}]^2 \quad (1.23)$$

7. <http://www.ece.ncsu.edu/imaging/Archives/ImageDataBase/Industrial/>

8. Tukey est probablement le premier à avoir mis en évidence l'intérêt des statistiques robustes [255], avant la formalisation mathématique de Huber [130].

9. Si l'estimation des éclairages était, elle aussi, bruitée, il faudrait remplacer les moindres carrés ordinaires par les moindres carrés totaux.

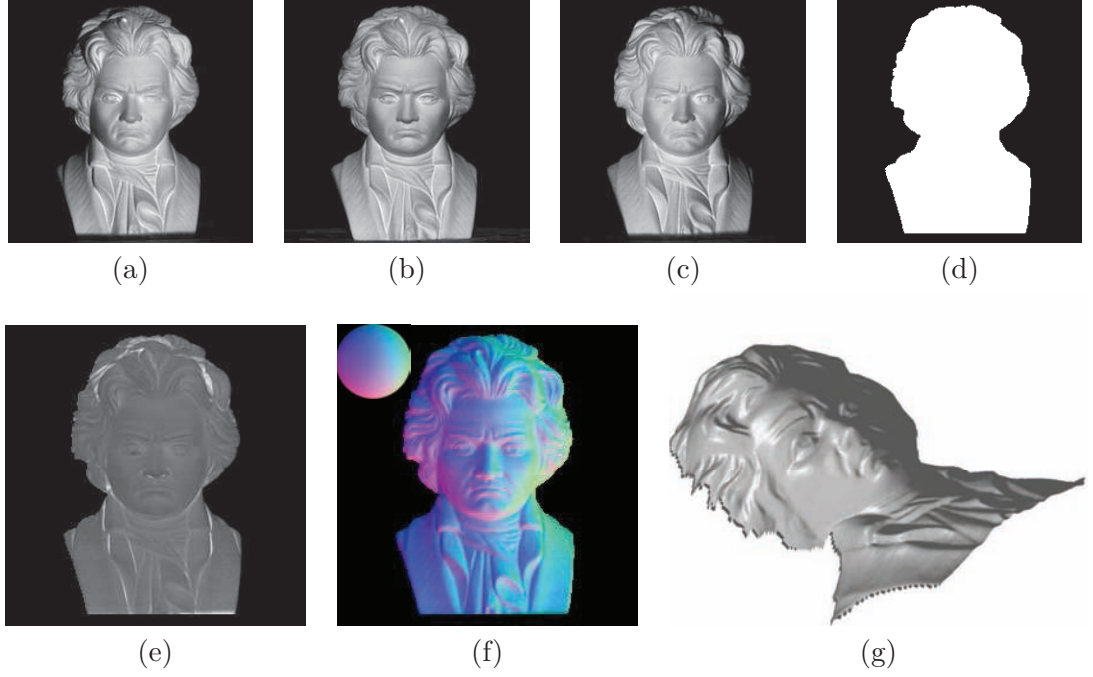


FIGURE 1.2 – Résolution exacte d'un problème de stéréophotométrie à $m = 3$ images. (a)-(b)-(c) Trois photographies d'un buste en plâtre de Beethoven (les vecteurs d'éclairage respectifs valent $\mathbf{s}^1 = [0, 08; 0, 26; 1]^\top$, $\mathbf{s}^2 = [-0, 20; -0, 02; 0, 97]^\top$ et $\mathbf{s}^3 = [0, 11; -0, 28; 1, 04]^\top$). (d) Domaine de reconstruction Ω . (e) Albédo estimé. (f) Champ de normales estimé (en fausses couleurs). (g) Relief obtenu par intégration des normales (cf. chapitre 2). Les écarts au modèle lambertien linéaire (1.18) ayant été négligés (en particulier, les ombres), l'estimation est biaisée, ce qui se traduit notamment par le fait que l'albédo estimé n'est pas uniforme.

est l'estimateur optimal (théorème de Gauss-Markov) : parmi tous les estimateurs non biaisés de $\mathbf{m}(u, v)$, il s'agit de celui dont la variance est minimale (estimateur *efficace*)¹⁰. En plus d'être optimal sous les hypothèses ci-dessus, cet estimateur présente l'intérêt de fournir une solution explicite :

$$\mathbf{m}(u, v) = \underbrace{(\mathbf{S}^\top \mathbf{S})^{-1} \mathbf{S}^\top}_{\mathbf{S}^+} \mathbf{i}(u, v) \quad (1.24)$$

où $\mathbf{S}^+$ est la pseudo-inverse de Moore-Penrose de $\mathbf{S}$.

Lorsque l'homoscédasticité n'est pas vérifiée, i.e. $\text{var}(\epsilon^i(u, v)) = \sigma^i(u, v)^2$ (composantes du bruit *hétéroscédastiques*, i.e. dont les variances ne sont pas toutes égales),

10. Rappelons que cette propriété ne vient pas du caractère i.i.d. gaussien des variables aléatoires $\epsilon^i(u, v)$, mais de leur homoscédasticité. Toutefois, si ces variables sont i.i.d. gaussiennes, alors l'homoscédasticité est garantie : on parle de *bruit blanc gaussien*.

l'estimateur en moindres carrés ordinaires n'est plus optimal, bien qu'il soit encore sans biais. L'estimateur efficace est alors l'estimateur en moindres carrés pondérés :

$$\mathbf{m}(u, v) = \operatorname{argmin}_{\mathbf{m} \in \mathbb{R}^3} \sum_{i=1}^m w^i(u, v) \left[I^i(u, v) - \mathbf{s}^i \cdot \mathbf{m} \right]^2 \quad (1.25)$$

où $w^i(u, v) = \frac{1}{\sigma^i(u, v)^2}$. Les variances $\sigma^i(u, v)^2$ ne sont généralement pas connues a priori, mais elles peuvent être estimées empiriquement si l'on dispose de suffisamment de données. En pratique, cela nécessiterait d'utiliser des centaines d'éclairages différents, ce qui représente une difficulté de mise en œuvre non négligeable¹¹.

Par ailleurs, si un modèle de bruit additif aléatoire exogène semble adapté à la modélisation de certains phénomènes comme la quantification ou le bruit d'acquisition, ces phénomènes ne sont pas les seules sources d'écarts au modèle lambertien linéaire. Les ombres, notamment, qui sont des phénomènes déterministes, provoquent des écarts bien plus importants. En présence de données aberrantes, l'estimateur en moindres carrés ordinaires est biaisé. L'estimateur en moindres carrés pondérés est généralement biaisé lui aussi, à moins de faire en sorte que $w^i(u, v) = 0$ si $I^i(u, v)$ est une donnée aberrante, ce qui suppose que l'on soit capable de détecter ces données (cf. paragraphe 1.2.2).

1.1.4 Sources d'écarts au modèle lambertien linéaire

Dans ce paragraphe, nous nous interrogeons sur les sources d'écarts au modèle lambertien linéaire (1.18), i.e. nous énumérons les différents types de données aberrantes.

Ombres propres et ombres portées

Comme nous l'avons déjà dit, le modèle lambertien linéaire (1.18) n'est valide que si le produit scalaire $\mathbf{s}^i \cdot \mathbf{n}(u, v)$ est positif, car un niveau de gris est forcément positif. On définit l'*ombre propre* de l'éclairage $\mathbf{s}^i$ comme l'ensemble des points $\mathbf{x}(u, v) \in \mathfrak{S}$ tels que $\mathbf{s}^i \cdot \mathbf{n}(u, v) < 0$ (cf. figure 1.3-a). Si la surface $\mathfrak{S}$ est régulière, toute limite entre une partie éclairée et une ombre propre, qui est constituée de points $\mathbf{x}(u, v)$ tels que $\mathbf{n}(u, v) \cdot \mathbf{s}^i = 0$, s'appelle un *terminateur*. La prise en compte des ombres propres dans (1.17) fait perdre au modèle lambertien son caractère linéaire, sur lequel repose l'estimation en moindres carrés qui a été décrite dans le paragraphe 1.1.3.

Il n'est pas nécessaire que $\mathbf{s}^i \cdot \mathbf{n}(u, v) < 0$ pour qu'un point $\mathbf{x}(u, v) \in \mathfrak{S}$ soit dans l'ombre. Effectivement, il existe de nombreux points tels que $\mathbf{s}^i \cdot \mathbf{n}(u, v) > 0$, mais qui ne sont pas éclairés par $\mathbf{s}^i$ car la lumière qui devrait les atteindre est arrêtée par un obstacle (cf. figure 1.3-b). Ces points constituent l'*ombre portée* de l'éclairage $\mathbf{s}^i$. Contrairement aux ombres propres, qui peuvent être caractérisées *localement*, les ombres portées ne peuvent être caractérisées que *globalement* (l'inégalité $\mathbf{s}^i \cdot \mathbf{n}(u, v) > 0$ ne suffit pas à les caractériser). En synthèse d'images, les ombres portées sont plus difficiles à calculer que les ombres propres, car elles dépendent de la géométrie globale de la scène.

11. Certains systèmes de reconstruction 3D par stéréophotométrie sont néanmoins équipés de plusieurs milliers de sources lumineuses [258].

Mais la reconstruction 3D d'une scène, qui constitue le problème inverse de la synthèse d'images, accentue considérablement cette différence : comme le montre le modèle lambertien (1.16), il est possible de tenir compte explicitement des ombres propres en stéréophotométrie, alors que cela constitue encore un problème ouvert pour les ombres portées. Dans ce mémoire, nous nous contenterons donc de tenir les ombres portées pour des données aberrantes.

Alors que les ombres propres et les ombres portées sont d'origines très différentes, elles se traduisent de la même façon dans une image¹², par un niveau de gris égal à 0 (en théorie du moins). Qui plus est, il arrive souvent qu'une zone d'ombre, c'est-à-dire une partie connexe de Ω de niveau de gris égal à 0, contienne à la fois les deux types d'ombre : sur l'exemple de la figure 1.3-b, l'attache de l'anse de la théière constitue effectivement une ligne de séparation entre ombre propre et ombre portée. Il semble donc difficile de distinguer ces deux types d'ombre. Or, comme les ombres propres s'expliquent par les caractéristiques locales du relief, contrairement aux ombres portées, le traitement réservé à chaque type d'ombre est différent : les ombres propres pourront être prises en compte explicitement (cf. paragraphe 1.3.5), tandis que les ombres portées restent des données aberrantes.

Reflets brillants

Le modèle lambertien est un modèle de *réflectance* qui caractérise bien les matériaux mats (« diffusants »). Bien entendu, aucun matériau réel n'est parfaitement lambertien, mais la plupart des matériaux mats sont plutôt bien modélisés par la loi de Lambert non linéaire (1.16). De nombreux autres modèles de réflectance ont néanmoins été proposés¹³, notamment pour modéliser l'aspect *brillant* de certains matériaux. Le modèle le plus connu de réflectance partiellement *spéculaire* est celui de Phong [206] (cf. figure 1.3-b). Ce modèle superpose à la *composante diffuse*, caractéristique du modèle lambertien (1.16), une *composante spéculaire* censée modéliser les reflets brillants :

$$I(u, v) = \underbrace{\rho(u, v) \max\{0, \mathbf{s} \cdot \mathbf{n}(u, v)\}}_{\text{Composante diffuse}} + \underbrace{\rho_s(u, v) (\max\{0, \mathbf{r}(u, v) \cdot \mathbf{v}(u, v)\})^\alpha}_{\text{Composante spéculaire}} \quad (1.26)$$

où $\mathbf{v}(u, v)$ est un vecteur unitaire pointant vers l'appareil photographique, $\mathbf{r}(u, v)$ indique la direction et le sens du faisceau lumineux après réflexion sur la surface selon la première loi de Descartes (cf. figure 1.4), et $\rho_s(u, v)$ constitue une nouvelle inconnue. Enfin, le paramètre $\alpha \geq 0$ caractérise la taille des lobes spéculaires : lorsque $\rho \equiv 0$ et $\alpha \rightarrow +\infty$, le modèle (1.26) décrit un miroir idéal, pour lequel la réflectance est une fonction de Dirac, tandis que la taille des lobes spéculaires croît lorsque α décroît.

Le modèle de Phong est certes empirique, mais il illustre bien le fait que le niveau de gris enregistré par l'appareil photographique peut être très supérieur à celui que prédit

12. On désigne indifféremment par « ombre » la zone non éclairée de l'espace 3D ou sa projection dans l'image.

13. Voir [179] pour une comparaison des modèles de réflectance les plus significatifs, et le chapitre 6 pour un aperçu de l'utilisation de différents modèles de réflectance en stéréophotométrie.



FIGURE 1.3 – (a) Valeur absolue de l'image d'une théière sous un éclairage directionnel, simulée en utilisant le modèle lambertien linéaire (1.18). Les niveaux de gris à l'intérieur de l'ombre propre sont négatifs. (b) Image de la même scène, simulée en utilisant le modèle de Phong (1.26) et en tenant compte à la fois des ombres propres et des ombres portées. Les trois premiers types d'écarts au modèle lambertien linéaire (ombres propres, ombres portées et reflets brillants) sont donc illustrés sur cette image.

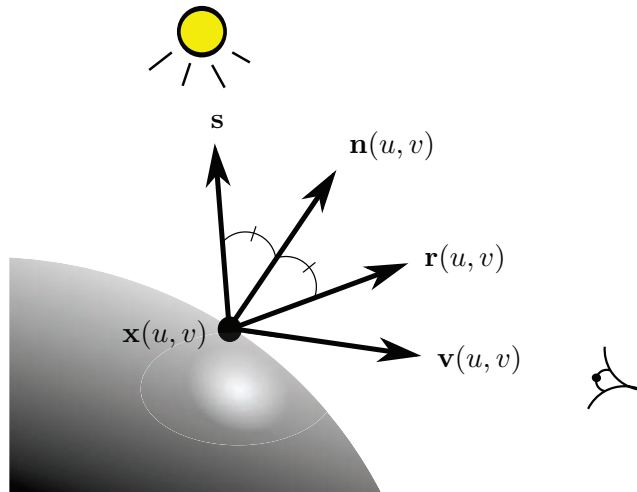


FIGURE 1.4 – Signification des vecteurs intervenant dans le modèle de Phong (1.26).

le modèle lambertien. La figure 1.3-b illustre simultanément les trois types d'écarts au modèle lambertien linéaire déjà cités (ombres propres, ombres portées et reflets brillants) sur l'image de synthèse d'une théière. Chaque type d'écarts affecte une ou plusieurs zones très localisées de la théière. Par conséquent, une grande partie des niveaux de gris de la figure 1.3-b sont conformes au modèle lambertien linéaire. Les données aberrantes sont donc minoritaires, ce qui justifie l'utilisation de méthodes d'estimation robuste.

Arêtes et discontinuités de profondeur

Nous avons pour l'instant supposé que la surface observée était régulière, afin que la normale puisse être définie en tout point. Cependant, les scènes réelles comportent souvent des arêtes, voire des discontinuités de profondeur. En de tels points, la normale n'est pas définie puisque la profondeur est non différentiable, voire non continue. Néanmoins, la question se pose autrement pour les images numériques, où la notion de point est remplacée par celle de *pixel*. En particulier, un appareil photographique ne mesure pas la luminance émise par chaque point objet, mais la somme des luminances émises par l'ensemble des points objets conjugués d'un même pixel.

Or, comme un pixel est conjugué d'une région entière de la surface, et non d'un point objet unique, il ne peut être entièrement situé sur une partie non différentiable : tout au plus peut-il chevaucher une ligne $\mathcal{C}$ de non différentiabilité, qui coupe le pixel en deux parties d'aires $\mathcal{A}_1$ et $\mathcal{A}_2$, caractérisées par des normales $\mathbf{n}_1$ et $\mathbf{n}_2$ potentiellement très différentes (cf. figure 1.5). En un tel pixel (u, v) , le niveau de gris $I(u, v)$ sous l'éclairage $\mathbf{s}$ est égal à la moyenne pondérée des niveaux de gris correspondant à chacune des deux parties. En appliquant le modèle lambertien linéaire (1.18) et en supposant, pour simplifier le raisonnement, que l'albédo ρ_0 est le même sur l'ensemble du pixel, le niveau de gris $I(u, v)$ s'écrit donc :

$$I(u, v) = \rho_0 \mathbf{s} \cdot (\pi_1 \mathbf{n}_1 + \pi_2 \mathbf{n}_2) \quad (1.27)$$

où $\pi_1 = \mathcal{A}_1 / (\mathcal{A}_1 + \mathcal{A}_2)$ et $\pi_2 = 1 - \pi_1$.

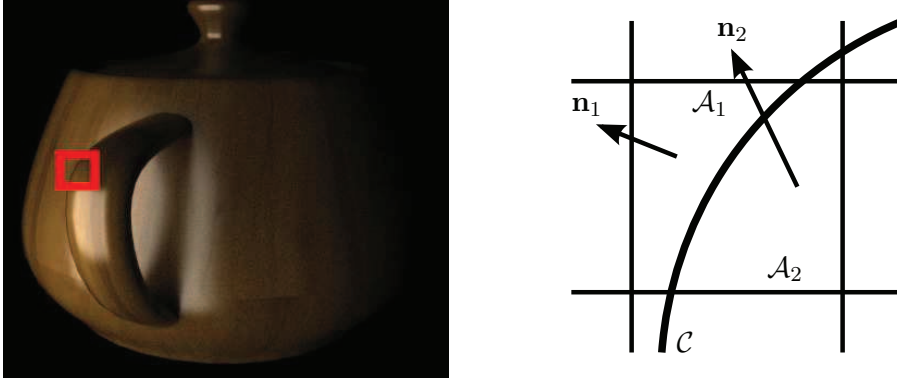


FIGURE 1.5 – Exemple d'un pixel composé de deux parties d'aires $\mathcal{A}_1$ et $\mathcal{A}_2$, caractérisées par des normales $\mathbf{n}_1$ et $\mathbf{n}_2$ potentiellement très différentes. La ligne de séparation $\mathcal{C}$ peut correspondre à une arête ou, comme ici, à une discontinuité de profondeur.

En quoi le niveau de gris $I(u, v)$ constitue-t-il un écart au modèle lambertien linéaire (1.18), puisque son écriture en découle directement ? Pour comprendre d'où provient cet écart, il nous faut préciser la notion de « normale » du pixel (u, v) . La moyenne pondérée $\bar{\mathbf{n}} = \pi_1 \mathbf{n}_1 + \pi_2 \mathbf{n}_2$ des normales $\mathbf{n}_1$ et $\mathbf{n}_2$ semble en donner une définition raisonnable, si ce n'est que $\bar{\mathbf{n}}$ n'est pas unitaire. En effet, il est facile de vérifier, grâce aux

égalités $\|\mathbf{n}_1\| = 1$, $\|\mathbf{n}_2\| = 1$ et $\pi_1 + \pi_2 = 1$, que $\|\bar{\mathbf{n}}\| < 1$ dès lors que π_1 et π_2 sont non nuls et que $\mathbf{n}_1$ et $\mathbf{n}_2$ ne sont pas colinéaires. Par conséquent, le vecteur $\bar{\mathbf{n}}$ doit être normalisé, ce qui donne comme définition de la « normale » en (u, v) :

$$\mathbf{n}(u, v) = \frac{\pi_1 \mathbf{n}_1 + \pi_2 \mathbf{n}_2}{\|\pi_1 \mathbf{n}_1 + \pi_2 \mathbf{n}_2\|} \quad (1.28)$$

Le niveau de gris prédit par le modèle lambertien linéaire (1.18) s'écrit donc :

$$I_{\text{prédit}}(u, v) = \rho_0 \mathbf{s} \cdot \frac{\pi_1 \mathbf{n}_1 + \pi_2 \mathbf{n}_2}{\|\pi_1 \mathbf{n}_1 + \pi_2 \mathbf{n}_2\|} \quad (1.29)$$

qui diffère effectivement de (1.27) à cause du facteur de normalisation $\|\bar{\mathbf{n}}\|$. En d'autres termes, le niveau de gris mesuré $I(u, v)$ est strictement inférieur au niveau de gris $I_{\text{prédit}}(u, v)$ prédit par le modèle lambertien linéaire. Par exemple, pour un pixel divisé en deux parties égales par une arête à angle droit, les vecteurs $\mathbf{n}_1$ et $\mathbf{n}_2$ sont orthogonaux, donc le rapport $I_{\text{prédit}}(u, v)/I(u, v)$ est égal à $\sqrt{2}$. Le niveau de gris d'un tel pixel constitue donc bien une donnée aberrante, et ce quel que soit l'éclairage $\mathbf{s}$.

Réflexions mutuelles

Les techniques photométriques sont bien adaptées à la reconstruction 3D de surfaces convexes. Les concavités sont plus difficiles à reconstruire, car chaque point de la surface se comporte lui-même comme une source lumineuse *secondaire* (par opposition à la source *primaire*, qui émet le flux lumineux principal). Dans le cas où la surface présente des concavités, ce flux secondaire éclaire d'autres points de la surface. Pour tenir compte de ces flux secondaires, le modèle (1.16) devrait donc être modifié de la façon suivante¹⁴ :

$$I(u, v) = \underbrace{\rho(u, v) \max\{0, \mathbf{s} \cdot \mathbf{n}(u, v)\}}_{\text{Contribution du flux primaire}} + \underbrace{\sum_{\mathbf{x}'} \rho(u, v) \max\{0, \mathbf{s}_{\mathbf{x}'}(\mathbf{x}) \cdot \mathbf{n}(u, v)\}}_{\text{Contributions des flux secondaires}} \quad (1.30)$$

Le terme modélisant les flux secondaires devrait en fait prendre la forme d'une intégrale, mais la somme discrète est plus facile à interpréter. Elle sous-entend que : la surface est constituée d'un ensemble de *facettes* ; la somme porte sur les centres $\mathbf{x}'$ des facettes visibles depuis $\mathbf{x}$; le vecteur $\mathbf{s}_{\mathbf{x}'}(\mathbf{x})$ caractérise l'éclairage au point $\mathbf{x}$ en provenance de la facette de centre $\mathbf{x}'$. Le modèle (1.30) semble malheureusement difficile à utiliser en stéréophotométrie, car les points $\mathbf{x}'$ font partie des inconnues (une difficulté similaire a déjà été évoquée à propos des ombres portées).

En réalité, la facette de centre $\mathbf{x}$ éclaire à son tour les facettes de centres $\mathbf{x}'$, ce qui explique pourquoi on parle de réflexions *mutuelles*. Le calcul exact de la luminance émise par chaque point de la surface $\mathfrak{S}$, après un nombre quelconque de « rebonds » successifs, est un problème délicat qui consiste à résoudre une *équation de Fredholm* [78], connue

14. Il faut prendre garde à ne pas inverser l'ordre de la somme et de l'opérateur max dans (1.30), car ce dernier est non linéaire.

en synthèse d'images sous le nom d'*équation du rendu* (*rendering equation*) ou *équation de radiosité*.

En tout cas, comme les réflexions mutuelles sont négligées dans le modèle lambertien linéaire (1.18), on peut s'attendre à ce que les niveaux de gris mesurés soient supérieurs aux valeurs prédites par ce modèle, en particulier au voisinage des arêtes concaves, comme cela est illustré sur l'exemple de la figure 1.6.

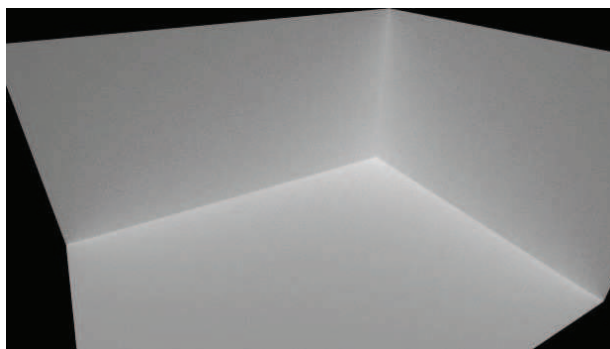


FIGURE 1.6 – Cube concave uniformément blanc : le niveau de gris observé au voisinage des arêtes concaves est supérieur au niveau de gris attendu, à cause des réflexions mutuelles.

Si certains points éclairés peuvent paraître plus clairs que ne le prédit le modèle lambertien linéaire (1.18), à cause des réflexions mutuelles, cela peut également concerner des points situés dans l'ombre. On parle alors de phénomène de *pénombre*. Les réflexions mutuelles se traduisent donc, globalement, par un éclaircissement des images. Un tel effet est souvent modélisé de façon empirique en synthèse d'images, de manière à rendre les images « plus réalistes ». Il est une des causes de l'*éclairage ambiant*, un concept sur lequel nous reviendrons dans le paragraphe 1.3.1.

Éclairages non directionnels

D'autres écarts au modèle lambertien linéaire (1.18) peuvent provenir de l'éclairage, qui est censé être *directionnel*. En pratique, il est difficile de garantir le caractère parallèle et uniforme d'un tel éclairage (à l'exception notable du soleil). Faute de mieux, on utilise soit une source ponctuelle, par exemple une *diode électroluminescente* (*light-emitting diode*, notée LED), soit une source *étendue*, par exemple un tube fluorescent (cf. figure 1.7). Ces sources doivent être éloignées de la scène, afin de garantir au mieux l'hypothèse directionnelle, mais comme leur intensité lumineuse est inversement proportionnelle au carré de la distance, il convient de trouver un compromis pour suffisamment éclairer la scène. Dans le cas d'une source étendue, cela rend flous les contours des ombres portées (cf. figure 1.7-b). Cet effet, par lequel on passe progressivement, et non pas brutalement, de la partie éclairée à la partie sombre, est également appelé *pénombre*.

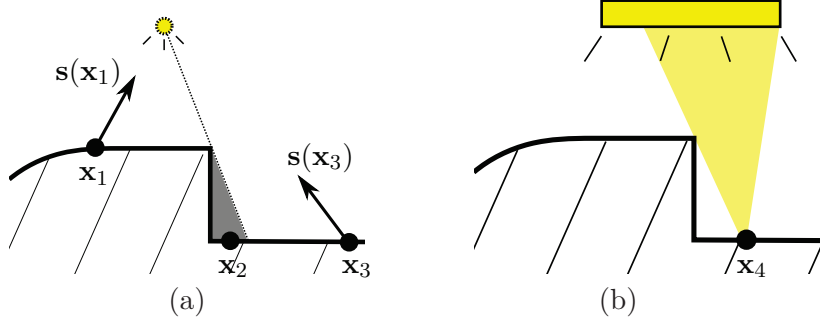


FIGURE 1.7 – Source ponctuelle et source étendue. (a) Lorsqu’une source ponctuelle est placée à proximité de la scène, l’éclairage varie d’un point à l’autre ($s(x_1)$ et $s(x_3)$ sont très différents), mais les ombres portées ont des contours nets (cf. figure 1.8-b) : la source ponctuelle est soit totalement visible (par exemple, depuis x_1 ou x_3), soit totalement occultée (par exemple, depuis x_2). (b) Avec une source étendue, la notion d’ombre doit être remplacée par celle de pénombre (cf. figure 1.8-c) : vue depuis x_4 , la source étendue est *partiellement* occultée.

La figure 1.8 montre que les images simulées avec une source ponctuelle ou avec une source étendue sont très différentes de l’image simulée avec un éclairage directionnel. Ces écarts au modèle lambertien linéaire (1.18) ne constituent pas réellement des données aberrantes, mais plutôt un biais systématique du modèle, qui ne prend pas en compte le fait que la source est proche. Comme nous le verrons dans le chapitre 5, il est possible de modéliser explicitement l’éclairage provenant d’une source proche, mais cela nécessite d’utiliser des méthodes numériques spécifiques. Pour appliquer l’approche classique de la stéréophotométrie, il faut prétraiter les données pour compenser le caractère non directionnel de l’éclairage, comme nous le verrons dans le paragraphe 1.2.

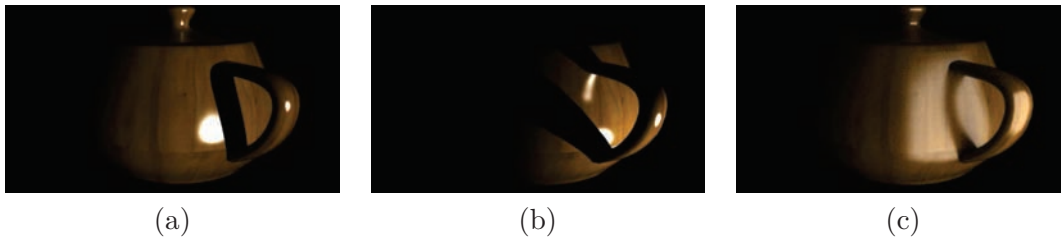


FIGURE 1.8 – Trois images d’une théière en bois verni, simulées sous différents éclairages : (a) éclairage directionnel, (b) source ponctuelle et (c) source étendue. Sur cette dernière image, on observe un phénomène de pénombre, qui apparaît aux points de la surface pour lesquels une partie de la source étendue est occultée (cf. figure 1.7-b).

1.1.5 Biais induit par les écarts au modèle lambertien linéaire

La stéréophotométrie « classique », c'est-à-dire telle que nous l'avons présentée dans le paragraphe 1.1.2, consiste à linéariser le modèle lambertien et à utiliser $m \geq 3$ éclairages directionnels, non coplanaires, connus. Le système linéaire (1.18) doit être résolu en chaque pixel (u, v) du domaine de reconstruction Ω . Sa matrice $\mathbf{S}$, qui est commune à l'ensemble des pixels, est de rang maximal (*full rank*). Sa solution approchée (1.24), au sens des moindres carrés, est donc unique. Si la présence d'écarts au modèle lambertien linéaire (1.18) dans les images semble inévitable, comment de tels écarts se traduisent-ils in fine dans l'estimation (1.20) de la normale et dans l'estimation (1.21) de l'albédo ? Nous illustrons maintenant les manifestations du biais de ces estimations sur trois exemples.

Biais induit par les ombres

La figure 1.2 montre les estimations de la normale et de l'albédo par stéréophotométrie classique, selon la procédure que nous venons de rappeler. Après intégration du champ des normales estimées, le relief obtenu (cf. figure 1.2-g) est *qualitativement* correct. Il est difficile de faire une évaluation *quantitative* de cette reconstruction 3D, car nous ne disposons pas de la vérité terrain. Néanmoins, un indice nous montre que les estimations comportent un biais : l'albédo estimé devrait être uniforme. Or, on observe très nettement sur la figure 1.2-e que l'albédo estimé est très élevé dans les zones où une image au moins (parmi $m = 3$) comporte une ombre. Peut-on en conclure que le biais induit par les ombres a systématiquement tendance à surestimer l'albédo ?

Pour simplifier le raisonnement, plaçons-nous dans le cas où $m = 3$ et supposons que les vecteurs d'éclairage $(\mathbf{s}^1, \mathbf{s}^2, \mathbf{s}^3)$ forment une base orthonormée $\mathcal{B}$. D'après le modèle lambertien linéaire (1.18), les niveaux de gris $(I^1(u, v), I^2(u, v), I^3(u, v))$ du pixel (u, v) constituent les coordonnées du vecteur $\mathbf{m}(u, v)$ dans la base $\mathcal{B}$. Il est donc facile de prédire l'effet d'un écart à ce modèle sur l'estimation de $\mathbf{m}(u, v)$. Prenons le cas d'un pixel se trouvant dans l'ombre (propre ou portée) de l'éclairage $\mathbf{s}^1$: l'estimation de $\mathbf{m}(u, v)$ est égale à sa projection sur le plan orthogonal à $\mathbf{s}^1$. Il s'ensuit que l'albédo est sous-estimé, ce qui semble en contradiction avec l'exemple de la figure 1.2-e. L'explication vient de ce que les vecteurs $(\mathbf{s}^1, \mathbf{s}^2, \mathbf{s}^3)$ de cet exemple ne constituent pas une base orthonormée (cela aurait accentué les ombres), mais forment deux à deux des angles de l'ordre de 30° . Pour de tels angles, la figure 1.9 montre pourquoi l'albédo est effectivement surestimé.

La figure 1.9 montre aussi que la normale estimée en (u, v) est biaisée. En effet, la seule façon d'expliquer l'égalité $I^1(u, v) = 0$ est d'orienter $\mathbf{n}(u, v)$ orthogonalement à l'éclairage $\mathbf{s}^1$. Nous pouvons néanmoins remarquer que, malgré ce biais, la solution approchée (1.24) est exacte lorsque $m = 3$. Par conséquent, quel que soit le nombre de données aberrantes, il existe forcément un *modèle 3D* (i.e., un relief et une carte d'albédo) qui permet d'expliquer exactement les images, dans le cas particulier où $m = 3$. Dans le cas général, ce modèle 3D peut être très éloigné de la réalité. Mais lorsque les écarts au modèle sont uniquement des ombres, il se trouve que le relief estimé est facile à interpréter. Comme cela est illustré sur la figure 1.10, il s'agit de la surface définie par l'union de la scène réelle et des ombres des différents éclairages¹⁵.

15. Falcone et al. ont montré un résultat équivalent pour le SFS [88].

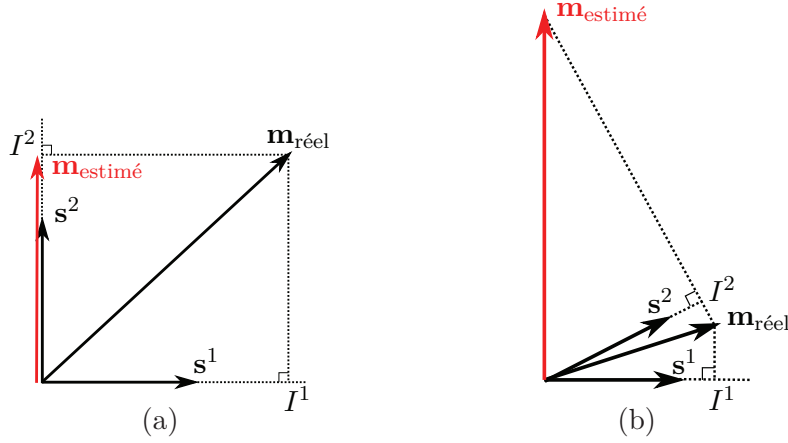


FIGURE 1.9 – Si un pixel se trouve dans l’ombre de l’éclairage $\mathbf{s}^1$, alors le vecteur $\mathbf{m}$ est projeté sur le plan $I^1 = 0$, orthogonalement au plan $\pi_{2,3}$ porté par les vecteurs $\mathbf{s}^2$ et $\mathbf{s}^3$. Cette projection cause un biais dans l’estimation de $\mathbf{m}$, mais deux cas se présentent : (a) si $\mathbf{s}^1$ est orthogonal au plan $\pi_{2,3}$, alors $\rho = \|\mathbf{m}_{\text{estimé}}\| < \|\mathbf{m}_{\text{réel}}\|$, donc l’albédo est sous-estimé ; (b) si l’angle entre $\mathbf{s}^1$ et le plan $\pi_{2,3}$ est inférieur à 45° (comme dans l’exemple de la figure 1.2), alors $\rho = \|\mathbf{m}_{\text{estimé}}\| > \|\mathbf{m}_{\text{réel}}\|$ est au contraire surestimé.

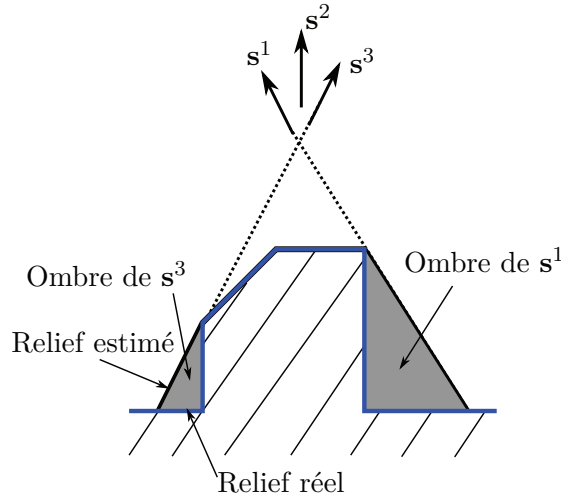


FIGURE 1.10 – Relief estimé par stéréophotométrie classique, dans le cas où $m = 3$, pour une scène provoquant des ombres sous les éclairages $\mathbf{s}^1$ et $\mathbf{s}^3$. Sans faire aucun traitement particulier pour les ombres, malgré le biais de l’estimation de la normale que cela implique, le relief estimé correspond *exactement* à la surface définie par l’union de la scène réelle et des ombres des différents éclairages.

Biais induit par les réflexions mutuelles

La figure 1.11 illustre une manifestation du biais induit par les réflexions mutuelles. L'image 1.11-a montre un moulage de sphère, i.e. une hémisphère concave. Visuellement, on aurait tendance à interpréter cette image comme celle d'une hémisphère convexe, à cause de l'ambiguïté concave/convexe bien connue dans le domaine du SFS [39]. La stéréophotométrie classique permet de lever cette ambiguïté (cf. figure 1.11-b), mais les réflexions mutuelles causent un biais qui se traduit par un aplatissement du relief estimé à proximité du bord de l'hémisphère. La figure 1.12 donne une explication graphique de ce phénomène, en tout point similaire à celle de la figure 1.9.



FIGURE 1.11 – (a) Une image d'un moulage de sphère éclairé de face. Avec cette seule image, il est impossible de lever l'ambiguïté concave/convexe. (b) Reconstruction 3D obtenue par stéréophotométrie, à partir de $m = 8$ images. L'ambiguïté a bien été levée, mais les pentes sont sous-estimées à proximité du bord de l'hémisphère.

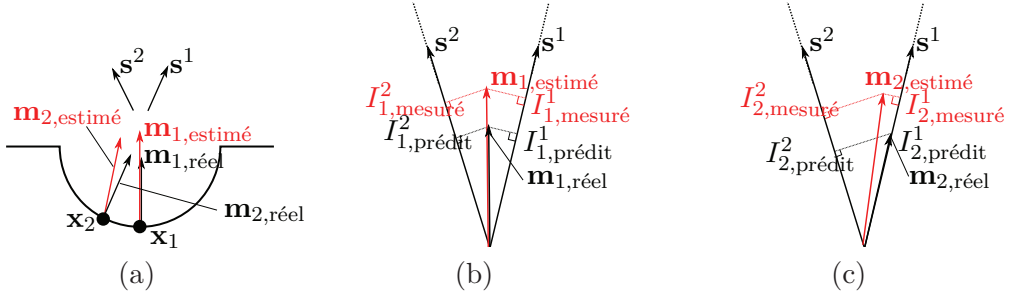


FIGURE 1.12 – (a) Hémisphère concave vue en coupe : les points $\mathbf{x}_1$ et $\mathbf{x}_2$ sont caractérisés par les vecteurs $\mathbf{m}_{1,\text{réel}}$ et $\mathbf{m}_{2,\text{réel}}$. À cause des réflexions mutuelles à l'intérieur de l'hémisphère, les niveaux de gris sont supérieurs aux valeurs prédites par le modèle lambertien linéaire. (b)-(c) En supposant que $E[I_{\text{mesuré}}^i - I_{\text{prédit}}^i]$ est indépendant de i (nous utiliserons cette hypothèse dans le paragraphe 1.3.1 pour modéliser l'éclairage « ambiant ») : $\mathbf{m}_{1,\text{estimé}}$ est colinéaire à $\mathbf{m}_{1,\text{réel}}$, donc l'estimation de la normale $\mathbf{n}_1$ n'est pas biaisée ; en revanche, $\mathbf{m}_{2,\text{estimé}}$ est plus proche de la verticale que $\mathbf{m}_{2,\text{réel}}$, donc l'estimation de la normale $\mathbf{n}_2$ est biaisée. Ce biais croît à mesure que l'on se rapproche du bord de l'hémisphère, ce qui explique l'aspect aplati du relief estimé près du bord.

Biais induit par l'utilisation d'éclairages non directionnels

La figure 1.13 montre une manifestation du biais induit par l'utilisation d'éclairages non directionnels. La société Pixience a mis au point un appareil photographique dédié à la dermatologie et à la cosmétique, appelé *dermoscope*. Afin de rendre les images les plus lisibles possible, les ombres sont éliminées en provoquant volontairement des réflexions mutuelles à l'intérieur du dermoscope (voir les figures 1.13-a et 1.13-b, ainsi que la figure 5.10 pour un schéma du dispositif). Par conséquent, l'éclairage incident est tout sauf directionnel ! Si l'on ne tient pas compte de cet écart au modèle lambertien linéaire, le relief reconstruit est très nettement gauchi (cf. figure 1.13-c). Grâce à l'estimation du flux lumineux incident, selon une procédure qui sera décrite dans le paragraphe 5.2, on parvient à « dégauchir » le relief (cf. figure 1.13-d).



(a)



(b)



(c)



(d)

FIGURE 1.13 – (a)-(b) Deux images (parmi $m = 15$) d'une pièce de monnaie, peinte en blanc mat pour limiter les reflets brillants. Les photographies sont acquises en utilisant un *dermoscope* conçu pour maximiser les réflexions mutuelles à l'intérieur de l'appareil (cf. figure 5.10), de façon à créer un éclairage diffus qui élimine les ombres. Reconstructions 3D par stéréophotométrie obtenues en utilisant : (c) un modèle d'éclairage directionnel ; (d) une estimation du flux lumineux incident (cf. paragraphe 5.2). Le relief (c) est très nettement gauchi.

1.2 Prétraitement des données

Une première façon de s'affranchir de l'influence des données aberrantes consiste à les *corriger*, c'est-à-dire à concevoir un procédé permettant, par exemple, de rendre négatif le niveau de gris d'un pixel situé dans une ombre propre, positif celui d'un pixel situé dans une ombre portée, et ainsi de suite, afin de forcer tous les niveaux de gris à vérifier le modèle lambertien linéaire (1.18). Ce procédé présente l'avantage de n'écarter aucun pixel, y compris un pixel dont tous les niveaux de gris sont manifestement erronés.

1.2.1 Changement d'espace colorimétrique

La sensibilité de l'œil humain à la lumière est fonction de la longueur d'onde. Elle est caractérisée par une « courbe de réponse », ou plutôt par trois courbes de réponse, une pour chaque type de cônes. Chacune de ces courbes présente un maximum qui se situe dans le rouge, le vert ou le bleu. De la même manière, les récepteurs photosensibles des appareils photographiques mesurent des quantités de lumière dans les longueurs d'onde situées autour du rouge, du vert et du bleu, de façon à produire des images à trois canaux, appelées images RGB (*red, green, blue*). Ces images RGB peuvent bien sûr être converties en niveaux de gris pour les besoins d'une application, e.g. la reconstruction 3D par stéréophotométrie, mais une partie de l'information n'est alors pas exploitée. Nous verrons dans le paragraphe 1.5 comment utiliser les images RGB en stéréophotométrie sans les convertir en niveaux de gris.

Il a été montré dans [185] que l'utilisation de capteurs multi- ou hyper-spectraux permettait de corriger les effets des réflexions mutuelles, qui dépendent de la couleur de la surface [163]. Cela semble indiquer que les réflexions spéculaires peuvent être supprimées en combinant judicieusement les informations photométriques mesurées dans les différents canaux. Cette idée a été suivie dans [167, 289], où l'espace RGB est remplacé par un nouvel espace colorimétrique appelé SUV. En désignant par $\mathbf{i}_{RGB} = [I_R, I_G, I_B]^\top$ le vecteur formé par les trois *niveaux de couleur*, ce nouvel espace colorimétrique est défini par une rotation :

$$\mathbf{i}_{SUV} = \mathbf{R} \mathbf{i}_{RGB} \quad (1.31)$$

La matrice de rotation $\mathbf{R}$ doit vérifier :

$$\mathbf{R} \mathbf{c} = [1, 0, 0]^\top \quad (1.32)$$

où le vecteur $\mathbf{c}$, qui doit bien sûr être unitaire, encode la couleur du flux lumineux incident dans l'espace RGB (on suppose que $\mathbf{c}$ est soit déterminé par une procédure d'étalonnage, soit estimé empiriquement).

Le canal S du vecteur $\mathbf{i}_{SUV}$ défini par (1.31) est censé encoder l'ensemble de la composante spéculaire, tandis que les canaux U et V sont supposés être exempts de reflets brillants. Une image en niveaux de gris censée être « purement diffuse », i.e. conforme au modèle lambertien (1.16), est finalement obtenue par :

$$I_D = \sqrt{I_U^2 + I_V^2} \quad (1.33)$$

La figure 1.14 montre un exemple de résultat obtenu par cette approche : on constate effectivement que les reflets brillants sont nettement atténués.

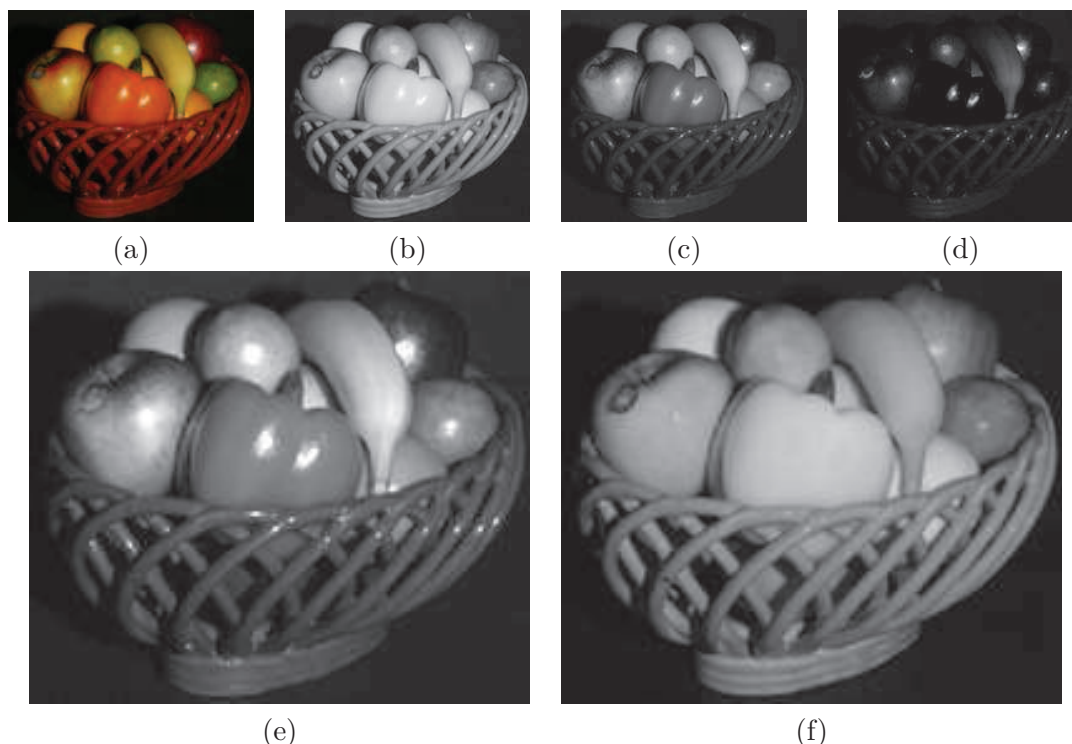


FIGURE 1.14 – Suppression des reflets brillants par changement d’espace colorimétrique. (a) Exemple d’image couleur. (b)-(c)-(d) Composantes de cette image dans les canaux R, G, B. (e) Composante spéculaire et (f) composante diffuse, obtenues par rotation de l’espace colorimétrique RGB. La couleur $\mathbf{c}$ de la source a été déterminée visuellement, en cherchant à minimiser les reflets brillants sur l’image (f).

Cette approche pose cependant un certain nombre de questions. Tout d’abord, il est implicitement supposé que l’éclairage n’est pas monochromatique de couleur rouge, verte ou bleue, car dans ce cas, la totalité de l’information serait contenue dans le canal S après transformation, et l’image diffuse de niveau de gris I_D serait totalement noire. Il semble donc que la méthode ne puisse pas s’appliquer à la stéréophotométrie avec éclairages colorés, qui présente pourtant un certain nombre d’avantages [117]. D’autre part, la couleur de la source doit être connue. Or, il est difficile de déterminer de façon précise cette couleur avec comme seule donnée I_{RGB} , sans connaître la couleur de la scène : nous nous trouvons alors confrontés à un problème de causalité difficile à résoudre. Ces différents arguments nous amènent à penser que, aussi séduisante soit-elle, l’idée de corriger des phénomènes complexes par une simple transformation linéaire est bien trop naïve, et que des méthodes plus spécifiques doivent être recherchées.

Les auteurs de [167, 289] justifient cette approche en expliquant, à juste titre, qu'elle ne repose sur aucun modèle photométrique explicite. Ceci peut être intéressant pour certaines applications de vision par ordinateur fondées sur la détection d'indices géométriques, comme la détermination du flot optique ou la reconstruction 3D par stéréoscopie¹⁶. Cependant, dans le cadre de la reconstruction 3D photométrique, un modèle photométrique est généralement postulé a priori¹⁷. Il semble donc raisonnable de nous fonder sur ce modèle pour identifier les données aberrantes. En effet, nous allons voir que la plupart des méthodes de prétraitement adaptées à la stéréophotométrie corrigent les données aberrantes à partir de l'ensemble des images, et non pas image par image.

1.2.2 Projection sur l'espace des matrices de rang 3

L'approximation linéaire (1.18) du modèle lambertien permet de reformuler, sous forme matricielle, l'estimation du champ $\mathbf{m}$ pour l'ensemble des pixels de Ω :

$$\underbrace{\begin{bmatrix} \mathbf{i}(u_1, v_1) & \dots & \mathbf{i}(u_{|\Omega|}, v_{|\Omega|}) \end{bmatrix}}_{\mathbf{I} \in \mathbb{R}^{m \times |\Omega|}} = \mathbf{S} \underbrace{\begin{bmatrix} \mathbf{m}(u_1, v_1) & \dots & \mathbf{m}(u_{|\Omega|}, v_{|\Omega|}) \end{bmatrix}}_{\mathbf{M} \in \mathbb{R}^{3 \times |\Omega|}} \quad (1.34)$$

où $\mathbf{S} \in \mathbb{R}^{m \times 3}$ et $|\Omega|$ désigne le nombre de pixels du domaine de reconstruction Ω .

D'après les propriétés sur le rang du produit de deux matrices, il s'ensuit que le rang de $\mathbf{I}$ doit être au plus égal à $\min\{\text{rang}(\mathbf{S}), \text{rang}(\mathbf{M})\} \leq 3$. En pratique, à cause des écarts au modèle, le rang de la matrice $\mathbf{I}$ est maximal, i.e. égal à sa plus petite dimension, qui est en général le nombre m d'éclairages. On peut modifier cette matrice¹⁸, de façon à forcer son rang à être égal à 3. Ceci peut être réalisé en projetant $\mathbf{I}$ sur l'espace des matrices de rang 3, par exemple par résolution du problème en moindres carrés suivant :

$$(\bar{\mathbf{U}}, \bar{\mathbf{W}}, \bar{\mathbf{V}}) = \underset{\substack{\mathbf{U} \in \mathbb{R}^{m \times 3} \\ \mathbf{W} \in \mathbb{R}^{3 \times 3} \\ \mathbf{V} \in \mathbb{R}^{|\Omega| \times 3}}}{\text{argmin}} \|\mathbf{I} - \mathbf{U}\mathbf{W}\mathbf{V}^\top\|_F^2 \quad (1.35)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius. Le problème (1.35) admet une solution explicite, qui peut être calculée par *décomposition en valeurs singulières* (*singular value decomposition*, notée SVD) tronquée à l'ordre 3. La matrice $\bar{\mathbf{I}}$ de rang 3 la plus proche de $\mathbf{I}$ (au sens de la norme de Frobenius) s'écrit alors :

$$\bar{\mathbf{I}} = \bar{\mathbf{U}} \bar{\mathbf{W}} \bar{\mathbf{V}}^\top \quad (1.36)$$

où $\bar{\mathbf{W}} \in \mathbb{R}^{3 \times 3}$ est une matrice diagonale contenant les trois plus grandes valeurs singulières de $\mathbf{I}$. Il est notable que, contrairement à $\mathbf{I}$, la matrice corrigée $\bar{\mathbf{I}}$ peut contenir des éléments négatifs. Cela est particulièrement intéressant pour la prise en compte des ombres propres (cf. figure 1.15).

16. En réalité, ces deux applications reposent *implicitement* sur l'hypothèse lambertienne.

17. Modérons toutefois cette affirmation en remarquant que l'analyse du gradient des images peut permettre d'identifier les courbes de niveau de la surface, indépendamment de tout modèle [52].

18. Une telle modification ne peut être effective que si $m > 3$, puisque sinon $\text{rang}(\mathbf{I}) = 3$.

L'approche décrite dans ce paragraphe, qui est fondée sur une SVD tronquée, a été étendue par Miyazaki et al. dans [178] au moyen de techniques de SVD plus robustes. Nous avons également montré, dans [77, 214], que la factorisation par SVD tronquée constituait un excellent *détecteur* de données aberrantes : les points où l'écart entre $\mathbf{I}$ et $\bar{\mathbf{I}}$ est élevé correspondent en général aux données aberrantes.

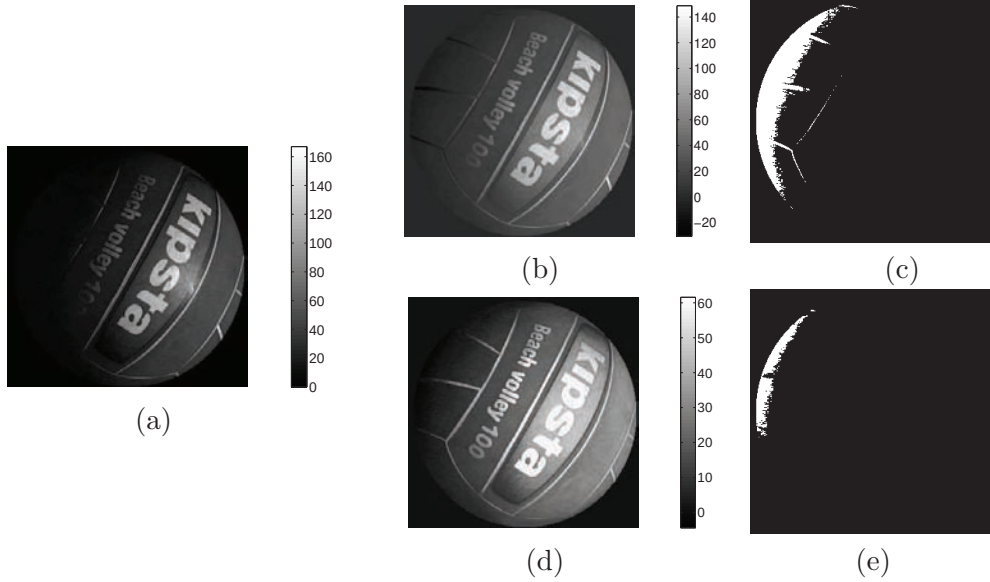


FIGURE 1.15 – (a) Une image (parmi $m = 8$) d'un ballon, comportant des ombres propres. (b) Correction des ombres par projection sur l'espace des matrices de rang 3. (c) Points où (b) est négatif. (d) Correction des ombres par la méthode de Wu et al. [270]. (e) Points où (d) est négatif. Par rapport à la projection sur l'espace des matrices de rang 3, la minimisation parcimonieuse du rang [270] permet une meilleure restauration des niveaux de gris situés dans les ombres propres.

1.2.3 Minimisation parcimonieuse du rang

Plutôt que de forcer explicitement le rang de la matrice $\mathbf{I}$ à être égal à 3, on peut chercher à *minimiser* son rang, tout en s'assurant que sa correction soit *parcimonieuse* (*sparse*). Wu et al., qui s'inspirent de travaux sur les méthodes « de rang faible » (« *low-rank* ») présentés dans [164], recommandent de décomposer la matrice $\mathbf{I}$ sous la forme de la somme d'une matrice $\mathbf{I}'$ de rang minimal et d'une matrice d'erreur $\mathbf{E} = \mathbf{I} - \mathbf{I}'$ la plus *creuse* (terme qui se traduit lui aussi par *sparse*) possible [270], ce qui revient à résoudre le problème d'optimisation suivant :

$$\begin{cases} \min_{\substack{\mathbf{I}' \in \mathbb{R}^{m \times |\Omega|} \\ \mathbf{E} \in \mathbb{R}^{m \times |\Omega|}}} \text{rang}(\mathbf{I}') + \lambda \|\mathbf{E}\|_0 \\ \text{s.c. } \mathbf{E} = \mathbf{I} - \mathbf{I}' \end{cases} \quad (1.37)$$

où $\|\mathbf{E}\|_0$ est égal au nombre d'éléments non nuls de la matrice $\mathbf{E}$ (nombre qu'on cherche à minimiser pour rendre $\mathbf{E}$ creuse), et $\lambda > 0$ est un hyper-paramètre. La semi-norme 0 étant non convexe, le problème (1.37) est difficile à résoudre. Wu et al. proposent d'optimiser la relaxation convexe suivante du problème (1.37) :

$$\begin{cases} \min_{(\mathbf{I}', \mathbf{E})} \|\mathbf{I}'\|_* + \lambda \|\mathbf{E}\|_* \\ \text{s.c. } \mathbf{E} = \mathbf{I} - \mathbf{I}' \end{cases} \quad (1.38)$$

où la norme $\|\cdot\|_*$ est égale à la somme des valeurs singulières. D'autres techniques de relaxation convexe ont été récemment proposées pour résoudre le problème de l'approximation d'une matrice par une matrice de rang faible : dans [242], la semi-norme 0 est remplacée par la norme 1, tandis qu'une approche plus récente consiste à résoudre le problème (1.37) par des itérations proximales (convexes) [259].

Au regard des résultats présentés sur les figures 1.16 et 1.17, les approches consistant à minimiser de façon parcimonieuse le rang de $\mathbf{I}$ surpassent la projection sur l'espace des matrices de rang 3¹⁹ même si, comme le montre le résultat de la figure 1.17, la correction des ombres portées est parfois décevante. Ceci semble montrer que la correction « à l'aveugle » des images de stéréophotométrie, même si elle constitue une piste intéressante, n'est vraiment satisfaisante que pour une faible proportion de données aberrantes.

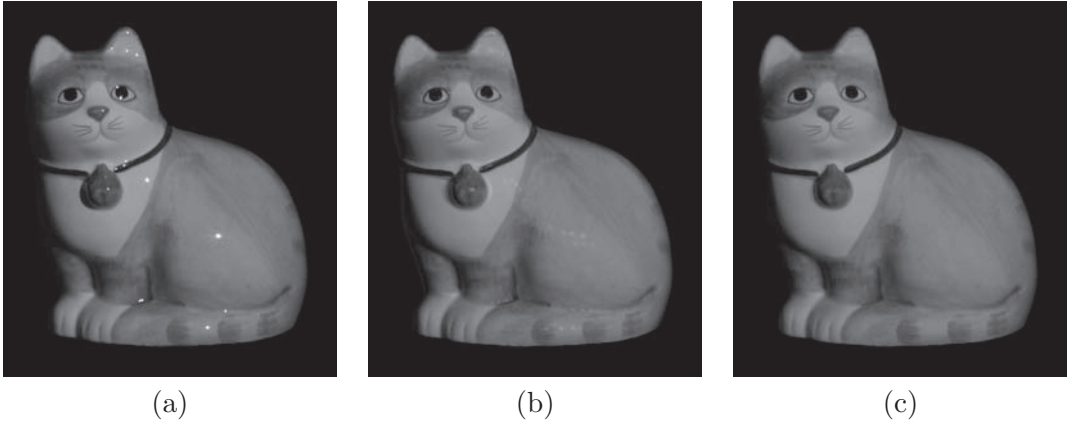


FIGURE 1.16 – Correction des reflets brillants. (a) Une image (parmi $m = 22$) d'un chat en porcelaine [122]. (b) Correction des niveaux de gris de (a) par projection sur l'espace des matrices de rang 3. (c) Correction par minimisation parcimonieuse du rang [270]. Comme les reflets brillants n'affectent qu'une faible proportion des niveaux de gris, ils sont assez faciles à corriger.

19. Cela nécessite quand même que λ soit convenablement choisi. En ce sens, λ constitue donc un paramètre « critique ».



FIGURE 1.17 – Correction des ombres portées. (a) Une image (parmi $m = 22$) d'un visage [84]. (b) Correction des niveaux de gris de (a) par projection sur l'espace des matrices de rang 3. (c) Correction par minimisation parcimonieuse du rang [270]. (d)-(e) Idem avec seulement $m = 5$ images. Même si l'approche parcimonieuse est plus performante que la projection sur l'espace des matrices de rang 3, elle ne fonctionne bien qu'en présence d'une faible proportion de données aberrantes.

1.2.4 Éclairages quasi-parallèles, non uniformes

Il existe une situation où les données aberrantes sont relativement faciles à corriger par prétraitement, si l'on accepte une intervention de l'utilisateur (approche « supervisée ») : il s'agit des écarts au modèle lambertien linéaire provoqués par un éclairage quasi-directionnel, non uniforme. Supposons que l'éclairage ne soit pas parfaitement directionnel, i.e. que $\mathbf{s}$ varie spatialement. Supposons que sa direction $\bar{\mathbf{s}}$ soit uniforme et connue, mais que son intensité varie, i.e. $\mathbf{s}(u, v) = \phi(u, v)\bar{\mathbf{s}}$. Supposons de plus que la scène à reconstruire soit plane, de normale $\mathbf{n}(u, v) \equiv \mathbf{n}_{\text{réf}}$ connue, de couleur uniforme $\rho(u, v) \equiv \rho_{\text{réf}}$. Le i -ème niveau de gris, $i \in [1, m]$, en un pixel (u, v) s'écrit alors :

$$I_{\text{réf}}^i(u, v) = \rho_{\text{réf}} \phi^i(u, v) \bar{\mathbf{s}}^i \cdot \mathbf{n}_{\text{réf}} \quad (1.39)$$

Dans l'égalité (1.39), tous les termes sont connus, à l'exception de l'albédo $\rho_{\text{réf}}$ et de

l'intensité $\phi^i(u, v)$. Cette dernière peut être estimée, connaissant l'albédo, par :

$$\phi^i(u, v) = \frac{I_{\text{réf}}^i(u, v)}{\rho_{\text{réf}} \mathbf{s}^i \cdot \mathbf{n}_{\text{réf}}} \quad (1.40)$$

En pratique, on peut utiliser, pour effectuer cette opération, une feuille de papier blanc orientée orthogonalement à l'axe optique, i.e. de telle sorte que $\mathbf{n}_{\text{réf}} = [0, 0, 1]^\top$. On peut également fixer arbitrairement $\rho_{\text{réf}} = 1$.

En répétant l'estimation (1.40) pour chacun des m éclairages, les m intensités $\phi^i(u, v)$, $i \in [1, m]$ sont estimées. Pour estimer le champ $\mathbf{m}$ d'une nouvelle scène, il suffit de remplacer le système linéaire (1.18) par :

$$\underbrace{\begin{bmatrix} I^1(u, v)/\phi^1(u, v) \\ \vdots \\ I^m(u, v)/\phi^m(u, v) \end{bmatrix}}_{\bar{\mathbf{i}}(u, v) \in \mathbb{R}^m} = \underbrace{\begin{bmatrix} \bar{\mathbf{s}}^1^\top \\ \vdots \\ \bar{\mathbf{s}}^m^\top \end{bmatrix}}_{\bar{\mathbf{S}} \in \mathbb{R}^{m \times 3}} \underbrace{\begin{bmatrix} \rho(u, v) \mathbf{n}(u, v) \end{bmatrix}}_{\mathbf{m}(u, v) \in \mathbb{R}^3} \quad (1.41)$$

On se ramène ainsi au problème initial, tout en ayant *compensé* la non uniformité des éclairages. Cette idée très simple d'uniformisation du champ d'éclairage (appelée *flat-fielding* dans [12]) a été initialement proposée pour segmenter les caractères manuscrits sur une feuille de papier, e.g. pour lire automatiquement des formulaires : les caractères imprimés au laser ne présentent pas de relief, mais un albédo faible, tandis que l'écriture manuscrite provoque un léger relief [170]. Elle a été reprise dans [12, 244], puis plus récemment dans [129] pour la numérisation du patrimoine, où il est montré qu'elle permet d'éviter l'aspect « aplati » des reconstructions 3D obtenues sous l'hypothèse directionnelle (cf. figure 1.18-d).

En réalité, cette idée n'est valide que si l'intensité lumineuse du flux lumineux incident, en un point $\mathbf{x}(u, v)$ de la surface à reconstruire, est proche de celle au point $\mathbf{x}_{\text{réf}}(u, v)$ de la feuille de papier correspondant au même pixel, i.e. si la scène à reconstruire est proche du plan de la feuille. Pour pouvoir être plus précis, il faudrait étalonner l'intensité en chaque point de l'espace 3D. Ceci constitue une perspective intéressante, qui semble néanmoins difficile à mettre en œuvre. D'autre part, l'hypothèse d'une direction uniforme, i.e. $\bar{\mathbf{s}}(u, v) = \bar{\mathbf{s}}$, n'a aucune raison d'être vérifiée (cf. figure 1.7-a). Nous verrons dans le chapitre 5 comment prendre en compte des champs d'éclairage très généraux.

Avant de conclure ce paragraphe sur les prétraitements, rappelons les résultats que nous avons établis :

- Les effets d'un éclairage non uniforme peuvent être compensés en mesurant l'intensité de chaque éclairage en chaque pixel, au moyen d'une scène de relief et d'albédo connus.
- Les ombres portées, les reflets brillants et les réflexions mutuelles constituent des écarts au modèle lambertien linéaire (1.18) difficiles à corriger, pour lesquelles le prétraitement par changement d'espace colorimétrique [167, 289], par projection de la matrice des niveaux de gris sur les matrices de rang 3 [114, 282, 283], ou par son approximation par une matrice de rang faible [242, 259, 270] s'avèrent décevantes.

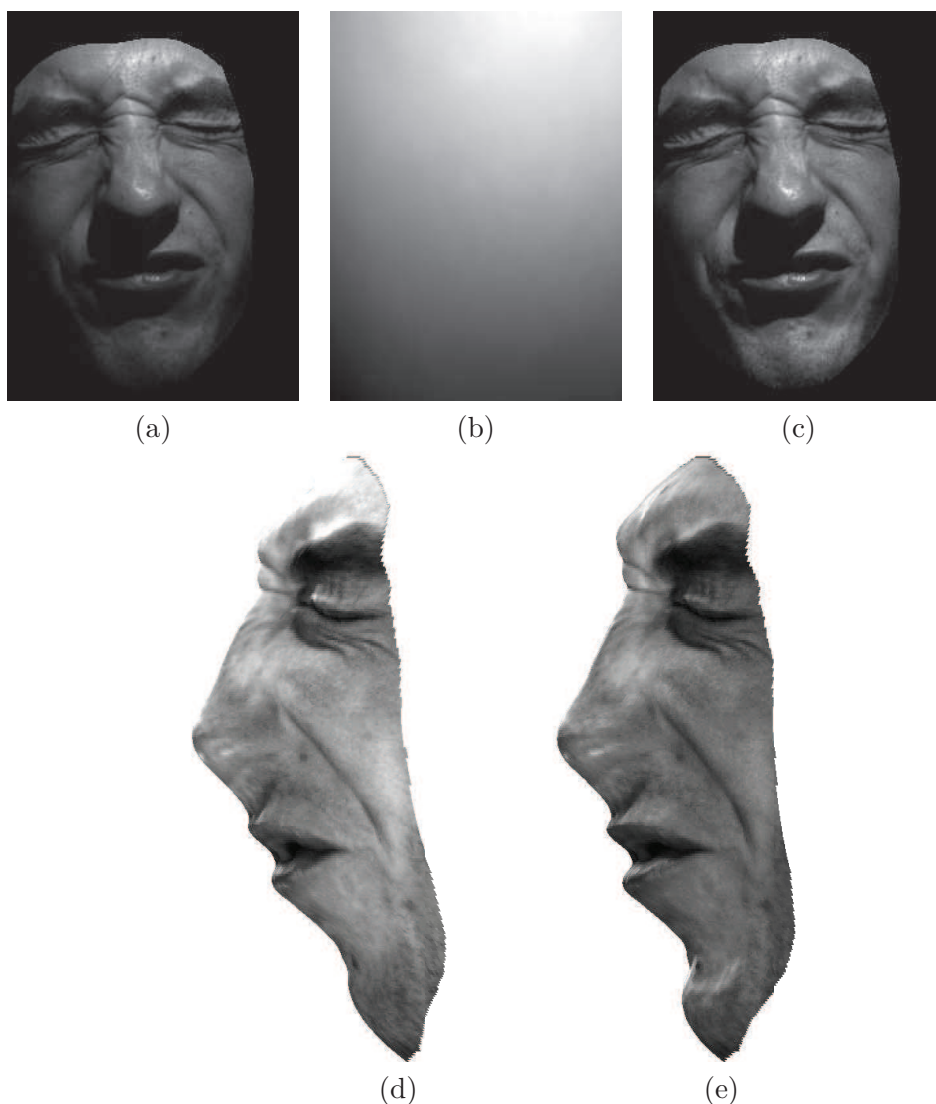


FIGURE 1.18 – Correction de la non uniformité des éclairages. (a) Une image réelle (parmi $m = 10$) d'un visage éclairé par une source ponctuelle proche. Le front, le bout du nez, les joues et le menton ont à peu près le même albédo et la même normale $\mathbf{n} \approx [0, 0, 1]^T$. Le niveau de gris serait donc sensiblement identique en tous ces points, si l'éclairage était parallèle et uniforme, ce qui n'est pas le cas. (b) Photographie d'une feuille de papier blanc orientée frontalement, prise sous le même éclairage. (c) Division pixel à pixel de l'image (a) par l'image (b). Reconstructions 3D par stéréophotométrie à partir : (d) des images d'origine telles que (a) ; (e) des images corrigées telles que (c). La correction préalable des images permet d'obtenir un relief moins fuyant et un albédo plus uniforme.

Dans la suite de ce chapitre, nous supposons que les effets d'éclairages non uniformes ont déjà été compensés. Nous nous intéressons plus particulièrement à des méthodes d'estimation robuste de la normale et de l'albédo, en présence d'ombres, de reflets brillants ou de réflexions mutuelles.

1.3 Stéréophotométrie calibrée robuste

Les prétraitements proposés dans le paragraphe 1.2 sont tous *non supervisés*, dans le sens où les éclairages $\mathbf{s}^i$ n'ont pas besoin d'être connus a priori, à l'exception du dernier prétraitement cité²⁰ (cf. paragraphe 1.2.4). Si de tels prétraitements semblent donc adaptés au problème de la stéréophotométrie *non calibrée* (cf. chapitre 3), il est dommage de ne pas profiter de la connaissance des éclairages dans le cas de la stéréophotométrie *calibrée*. Ne pourrait-on pas résoudre le système linéaire (1.18) en tenant compte *implicitement* de la présence de données aberrantes ? Ceci semble effectivement possible, en remplaçant l'estimation en moindres carrés $\mathbf{m} = \mathbf{S}^+ \mathbf{i}$ par une estimation robuste aux données aberrantes.

1.3.1 Prise en compte de l'éclairage « ambient »

Certains écarts au modèle lambertien linéaire (1.18), provoqués par des éclairages parasites ou par les réflexions mutuelles, sont souvent regroupés sous le concept d'éclairage « ambient ». Il convient de prendre en compte explicitement ces phénomènes, qui induisent un biais systématique de l'estimation que l'on peut modéliser ainsi :

$$I^i(u, v) = \rho(u, v) \mathbf{s}^i \cdot \mathbf{n}(u, v) + \phi_a^i(u, v) \quad (1.42)$$

Il est quasiment impossible d'exprimer le terme $\phi_a^i(u, v)$ d'éclairage ambient en fonction des caractéristiques de la scène, car il provient entre autres des réflexions de la lumière dans la pièce où sont prises les photographies. En outre, l'estimer automatiquement n'aurait pas de sens, car ce terme assimilerait l'ensemble des écarts au modèle lambertien linéaire.

Une façon relativement simple de tenir compte de ce biais consiste à utiliser un modèle d'harmoniques sphériques pour l'éclairage [23]. L'utilisation de trois harmoniques sphériques est équivalente au modèle d'éclairage directionnel que nous avons utilisé jusqu'ici, mais l'utilisation d'harmoniques sphériques supplémentaires permet de modéliser des éclairages plus généraux. En particulier, Or-El et al. ont montré récemment dans [193] qu'un modèle à quatre harmoniques sphériques permettait d'appliquer la technique du *shape-from-shading* sous un éclairage « naturel ».

20. Dans [170], ce prétraitement est effectué sans connaître les éclairages, grâce à l'approximation suivante : chaque image originale I^i , $i \in [1, m]$, est divisée pixel à pixel par l'image de référence $I_{\text{réf}}^i$, sans tenir compte du dénominateur $\rho_{\text{réf}} \mathbf{s}^i \cdot \mathbf{n}_{\text{réf}}$ de l'estimation (1.40) de $\phi^i(u, v)$, même si ce dénominateur dépend en réalité de i .

Une autre possibilité consiste à supposer que le terme d'éclairage ambiant est un champ scalaire non uniforme, mais qui ne dépend pas de i :

$$I^i(u, v) = \rho(u, v) \mathbf{s}^i \cdot \mathbf{n}(u, v) + \phi_a(u, v) \quad (1.43)$$

Bien sûr, il est difficile de supposer que les réflexions mutuelles soient indépendantes de l'éclairage. Cette approximation est donc assez grossière, et ne modélise en fait que l'éclairage ambiant « moyen ». Elle reste toutefois intéressante du point de vue des statistiques : estimer ϕ_a revient à estimer, en chaque pixel (u, v) , l'espérance de l'écart au modèle. Or, nous avons déjà mentionné qu'à cause des ombres portées, des reflets brillants ou encore des réflexions mutuelles, le bruit était endogène (biais systématique). En estimant ϕ_a , nous tenons compte de cet aspect, ce qui limite fortement le biais de l'estimation.

Le nouveau problème, qui doit être résolu en chaque pixel, est un problème linéaire :

$$\underbrace{\begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix}}_{\mathbf{i}(u, v) \in \mathbb{R}^m} = \underbrace{\begin{bmatrix} \mathbf{s}^{1\top} & 1 \\ \vdots & \\ \mathbf{s}^{m\top} & 1 \end{bmatrix}}_{\mathbf{S}' \in \mathbb{R}^{m \times 4}} \underbrace{\begin{bmatrix} \rho(u, v) \mathbf{n}(u, v) \\ \phi_a(u, v) \end{bmatrix}}_{\mathbf{m}'(u, v) \in \mathbb{R}^4} \quad (1.44)$$

dont la solution est unique dès lors que $m \geq 4$ et $\text{rang}(\mathbf{S}') = 4$.

Au vu du résultat de la figure 1.19, il semble que la stéréophotométrie soit une technique prometteuse pour séparer la texture de la géométrie et de l'éclairage, ce qui est un problème classique de traitement d'images, en particulier pour la décomposition *cartoon-texture* [16].

Remarquons que l'hypothèse $\text{rang}(\mathbf{S}') = 4$ peut ne pas être vérifiée en pratique. Un exemple de configuration dégénérée, i.e. pour laquelle $\text{rang}(\mathbf{S}') < 4$, est celui des éclairages « en anneaux » (*ring-light*) [237, 287], i.e. dont les directions se situent sur un cône de révolution ayant pour axe la direction d'observation. Pour de telles configurations dégénérées, il existe une infinité de solutions $\mathbf{m}'(u, v)$ au problème (1.44). Parmi toutes ces solutions, la SVD permet de déterminer explicitement celle dont la norme est minimale, i.e. de résoudre le problème :

$$\begin{cases} \min_{\mathbf{m}'} \|\mathbf{i}(u, v) - \mathbf{S}'\mathbf{m}'\| \\ \text{s.c. } \|\mathbf{m}'\| = \min \end{cases} \quad (1.45)$$

Puisque $\|\mathbf{m}'(u, v)\| = \sqrt{\rho(u, v)^2 + \phi_a(u, v)^2}$, de par la définition (1.44) de $\mathbf{m}'(u, v)$, imposer à cette norme d'être minimale a une conséquence directe sur l'estimation de ρ et de ϕ_a : en général, cette solution « mélange » ρ et ϕ_a et induit un biais important dans l'estimation des normales (cf. figure 1.19). Ce mélange vient directement du fait que $\text{rang}(\mathbf{S}') < 4$ (caractère mal posé).

Quoi qu'il en soit, il est très facile d'étendre l'approche classique de la stéréophotométrie à la prise en compte du terme additif $\phi_a(u, v)$ dans (1.43). Dorénavant, nous supposerons toutefois, afin d'alléger les notations, que $\phi_a \equiv 0$.

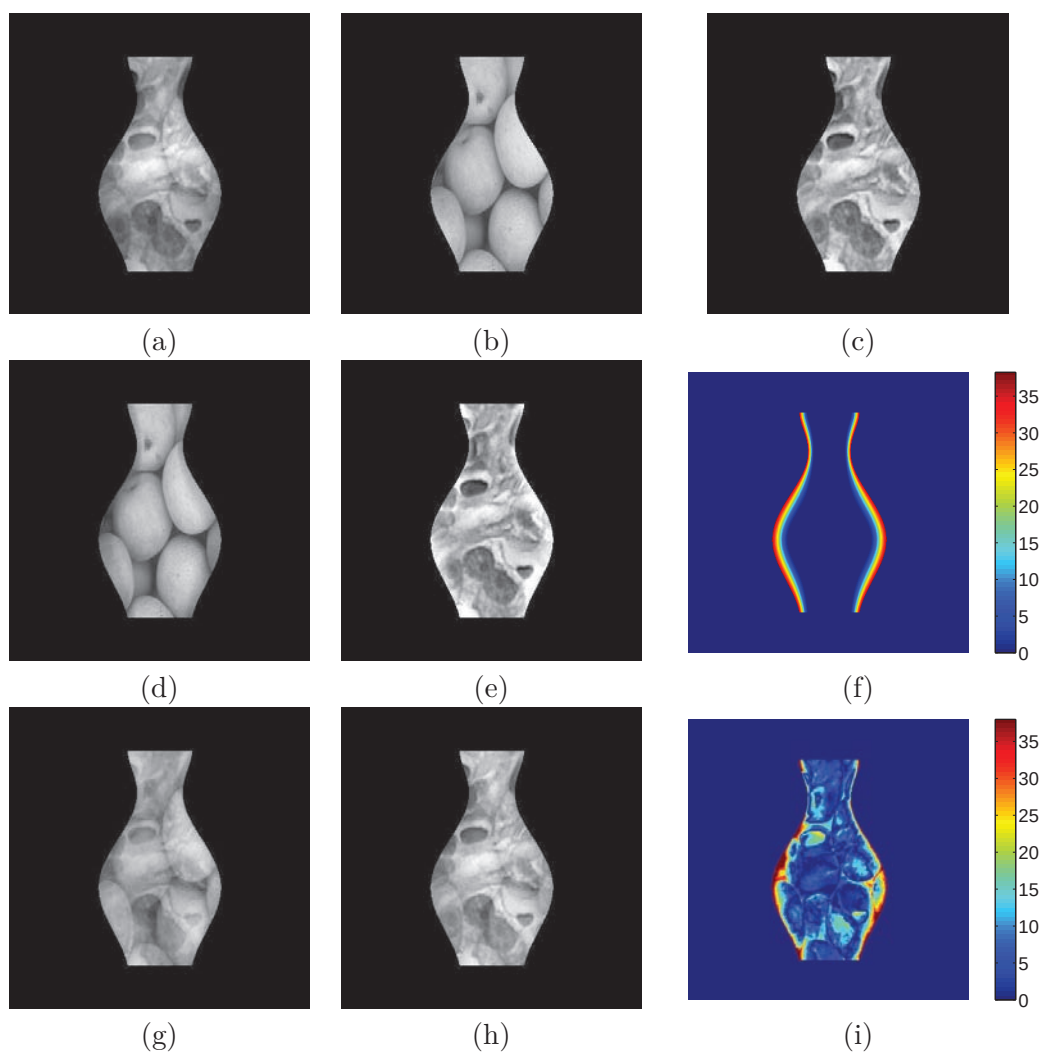


FIGURE 1.19 – Estimation de l'éclairage « ambiant ». (a) Une image de synthèse d'un vase (parmi $m = 50$), calculée avec le modèle (1.43), en utilisant l'image (b) comme albédo ρ et l'image (c) comme terme additif ϕ_a . (d)-(e) Estimations de ρ et ϕ_a en moindres carrés linéaires, lorsque les $m = 50$ éclairages sont répartis sur deux anneaux. (f) Écart angulaire entre la normale estimée et la normale réelle (l'écart angulaire moyen vaut $\text{EAM} = 5,48^\circ$). L'estimation est parfaite en dehors des bords du vase (ombres propres). (g)-(h)-(i) Mêmes résultats, lorsque les $m = 50$ éclairages sont répartis sur un seul anneau : l'estimation en moindres carrés linéaires mélange ρ et ϕ_a , et biaise l'estimation de la normale ($\text{EAM} = 8,76^\circ$). Si l'on veut estimer automatiquement l'éclairage ambiant, certaines configurations d'éclairage, comme celles étudiées dans [237, 287], doivent être proscrites.

1.3.2 Sélection des niveaux de gris

L'approche la plus naturelle pour résoudre le système linéaire (1.18) de façon robuste consiste à éliminer les niveaux de gris non conformes au modèle lambertien linéaire : nous avons déjà signalé dans le paragraphe 1.1.3 que l'estimateur en moindres carrés pondérés, construit en attribuant un poids nul aux niveaux de gris non conformes au modèle, était l'estimateur optimal (estimateur efficace).

Cette approche est celle que suivent Coleman et Jain dans [63] : si l'on dispose de $m = 4$ images, un niveau de gris corrompu par une ombre ou par un reflet brillant peut être éliminé en chaque pixel, et la solution peut être calculée à partir des trois niveaux de gris restants. Elle a été étendue aux images RGB par Barsky et Petrou dans [20, 204]. En utilisant $m > 3$ images, Chandraker et al. montrent que les ombres peuvent être automatiquement détectées par la technique de la coupure de graphe [51]. Enfin, mentionnons les heuristiques proposées par Sun et al. dans [245], et par Bringier et al. dans [40], qui permettent de déterminer, en chaque pixel, les niveaux de gris les plus conformes au modèle.

Au vu de la discussion menée dans le paragraphe 1.2.3, il semble que les données aberrantes puissent également être localisées par simple comparaison de la matrice $\mathbf{I}$ des niveaux de gris à son approximation par une matrice $\bar{\mathbf{I}}$ de rang 3, construite selon les équations (1.35) et (1.36). Nous avons effectivement proposé dans [77, 214] d'identifier les données aberrantes par simple seuillage de la différence $\mathbf{I} - \bar{\mathbf{I}}$. La figure 1.20 montre un exemple de résultat obtenu en utilisant cette approche.

1.3.3 Moindres carrés pondérés

L'approche décrite dans le paragraphe 1.3.2 repose sur une classification binaire. Cependant, il est parfois difficile de déterminer de façon binaire si une donnée est valide ou aberrante, par exemple à la périphérie des reflets brillants. Dans de tels cas, il est plus intéressant de définir, pour chaque pixel de chaque image, la probabilité que le niveau de gris soit valide. On peut alors estimer le champ $\mathbf{m}$ au sens des moindres carrés pondérés (cf. problème (1.25)), en choisissant cette probabilité comme poids.

Cette approche est proposée par Clark dans [60], où les poids sont déterminés par une estimation empirique de la variance du bruit. D'autres mécanismes de détermination de cette probabilité, fondés sur l'algorithme d'*espérance-maximisation* (*expectation-maximisation*, noté EM) [70] ont été proposés plus récemment dans [257, 272].

Dans [214], nous utilisons une pondération fondée sur le fait que les ombres (propres ou portées) correspondent à des niveaux de gris très faibles, et les reflets brillants à des niveaux de gris très élevés. Dans les deux cas, les écarts au modèle peuvent donc être déterminés par simple comparaison à la moyenne des niveaux de gris. Nous proposons une pondération gaussienne autour de cette valeur moyenne :

$$w^i(u, v) = \begin{cases} \exp \left\{ -\lambda_w \frac{(I^i(u, v) - E[I^i])^2}{2(E[I^i])^2} \right\} & \text{si } I^i(u, v) \in [0, E[I^i]] \\ \exp \left\{ -\lambda_w \frac{(I^i(u, v) - E[I^i])^2}{2(I_{\max} - E[I^i])^2} \right\} & \text{si } I^i(u, v) \in [E[I^i], I_{\max}] \end{cases} \quad (1.46)$$

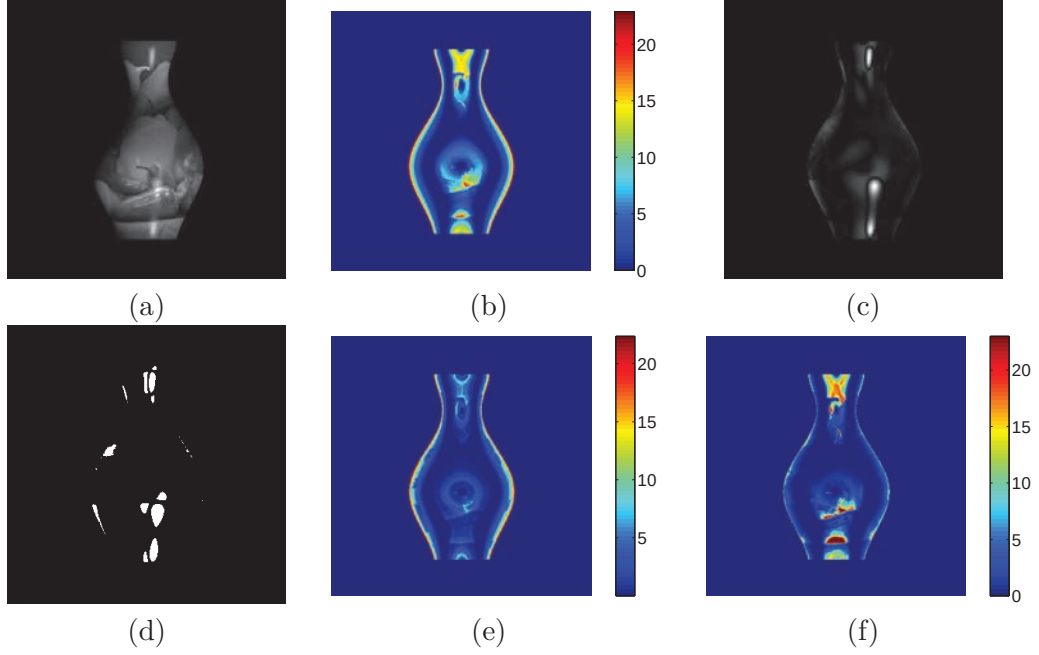


FIGURE 1.20 – (a) Une image de synthèse (parmi $m = 100$) d'un vase, calculée avec le modèle de réflectance spéculaire de Ward [260]. (b) Écart angulaire entre la normale estimée en moindres carrés et la normale réelle ($\text{EAM} = 4,05^\circ$). L'estimation de la normale est fortement biaisée aux endroits où un reflet brillant affecte une ou plusieurs images. (c) Valeur absolue de la différence entre (a) et l'image corrigée par projection de la matrice $\mathbf{I}$ sur l'espace des matrices de rang 3. (d) Détection des données aberrantes de l'image (a) par seuillage de (c) (le seuil est fixé à 10% du niveau de gris maximal de (a)). (e) Écart angulaire entre la normale réelle et son estimation à partir des données non aberrantes ($\text{EAM} = 2,48^\circ$). (f) Idem pour l'estimation en moindres carrés pondérés (cf. paragraphe 1.3.3), avec les poids (1.46) ($\text{EAM} = 2,35^\circ$).

où $E[I^i]$ est le niveau de gris moyen de la i -ème image, et λ_w un paramètre de la pondération permettant de contrôler l'étalement de la gaussienne (cf. figure 1.21). Dans les tests, nous choisissons empiriquement $\lambda_w = 9$.

La figure 1.20-f montre un exemple de résultat obtenu avec l'approche en moindres carrés pondérés. Le gain par rapport à l'approche en moindres carrés ordinaires (cf. figure 1.20-b) est important. Le résultat est qualitativement moins bon qu'en utilisant des poids binaires, bien que l'écart angulaire moyen (EAM) soit légèrement inférieur (mais la valeur relativement élevée de l'EAM, pour le résultat de la figure 1.20-e, est quasi-exclusivement due aux pixels du bord du vase). En effet, nous avons mentionné dans le paragraphe 1.1.3 que l'estimateur en moindres carrés pondérés était biaisé en présence de données aberrantes, si les poids n'étaient pas binaires.

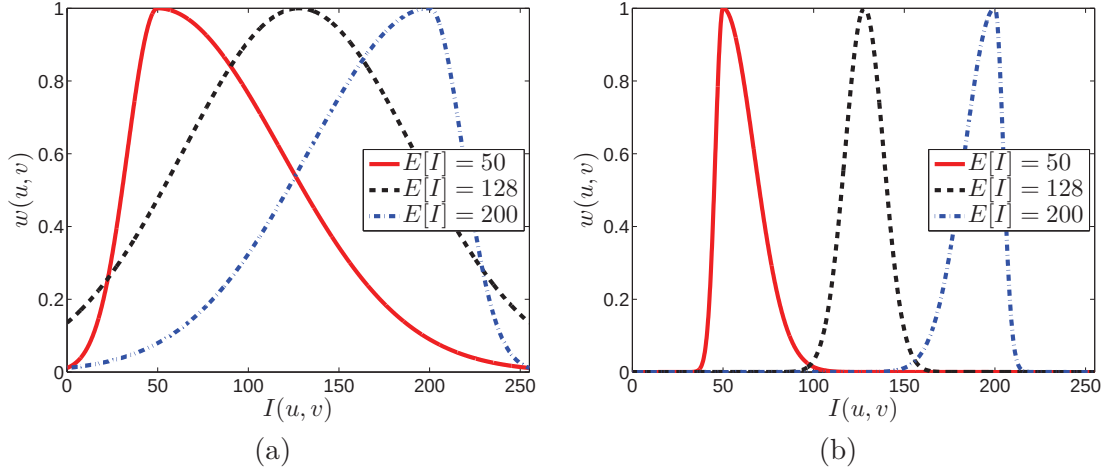


FIGURE 1.21 – Pondération (1.46) du niveau de gris, avec : (a) $\lambda_w = 9$; (b) $\lambda_w = 144$. Nous supposons ici que $I_{\max} = 255$.

Les approches fondées sur la sélection des niveaux de gris, tout comme celles fondées sur leur pondération, restent compliquées à utiliser car la modélisation d'une donnée aberrante est difficile. Plutôt que de chercher à modéliser les données aberrantes, il est plus simple de supposer qu'elles sont *minoritaires*, comme nous allons le faire dans le prochain paragraphe.

1.3.4 M-estimateurs

Une dernière classe d'estimateurs statistiques robustes est constituée par les estimateurs dits *parcimonieux* : l'objectif est alors de définir un estimateur maximisant le nombre de données conformes au modèle, plutôt que de minimiser l'écart quadratique moyen. De tels estimateurs sont adaptés aux données aberrantes, comme les ombres ou les reflets brillants, qui affectent en général un faible nombre de niveaux de gris par pixel.

L'estimateur parcimonieux est celui de la semi-norme 0 :

$$\min_{\mathbf{m}} \|\mathbf{i} - \mathbf{S}\mathbf{m}\|_0 \quad (1.47)$$

mais cette minimisation est difficile à réaliser numériquement, à cause de la non convexité de la semi-norme 0 (le problème est en fait NP-complet [42, 186]). Dans la plupart des applications, une relaxation convexe du problème (1.47) est utilisée, comme par exemple le problème de la régression en norme 1²¹ :

$$\min_{\mathbf{m}} \|\mathbf{i} - \mathbf{S}\mathbf{m}\|_1 \quad (1.48)$$

21. En statistiques, la régression en norme 1 correspond à un modèle de distribution laplacienne et à une minimisation de l'écart absolu médian, tandis que les approches en moindres carrés correspondent à une distribution gaussienne et à une minimisation de l'écart quadratique moyen.

Cette minimisation peut être réalisée efficacement par la méthode du *lagrangien augmenté* [164], qui converge plus rapidement que les moindres carrés repondérés [43]. Par un raisonnement bayésien, Ikehata et al. ont également présenté dans [131] d'autres estimateurs parcimonieux censés surpasser la régression en norme 1. Les résultats présentés dans [131] montrent qu'avec un faible nombre d'images ($m < 10$), la régression en norme 1 fournit des résultats supérieurs à leurs extensions, tandis qu'à partir de $m = 10$, ces extensions n'améliorent que très légèrement les résultats. L'intérêt de ces estimateurs est donc discutable, d'autant qu'ils nécessitent de choisir avec précaution plusieurs paramètres.

Il est également possible de proposer une interprétation fonctionnelle de l'estimateur en norme 1. Intéressons-nous pour cela au problème suivant, qui généralise (1.47) :

$$\min_{\mathbf{m}} \sum_{i=1}^m \Phi(|I^i - \mathbf{s}^i \cdot \mathbf{m}|) \quad (1.49)$$

L'estimateur parcimonieux (1.47) correspond à $\Phi_{L0}(s) = 1 - \delta(s)$, où δ est la fonction de Dirac. Choisir $\Phi_{LS}(s) = s^2$ revient à effectuer l'estimation en moindres carrés, $\Phi_{WLS}(s) = ws^2$ correspond aux moindres carrés pondérés, et $\Phi_{L1}(s) = s$ à l'estimation en norme 1. Pour obtenir une estimation robuste, on doit choisir une fonction Φ présentant une croissance au moins linéaire en 0, de façon à amplifier l'influence des résidus faibles correspondant aux « bonnes » données, ce qui est le cas de Φ_{LS} , Φ_{WLS} et Φ_{L1} , mais pas de Φ_{L0} , et au plus linéaire en $+\infty$, de façon à limiter l'influence des données aberrantes, ce qui est le cas de Φ_{L0} et Φ_{L1} , mais pas des fonctions quadratiques Φ_{LS} et Φ_{WLS} . Enfin, la non différentiabilité en 0 ne constitue pas un obstacle majeur, car elle peut être contournée en utilisant des méthodes proximales [199]. Les propriétés respectives des principaux estimateurs robustes sont synthétisées dans la table 1.1.

	$\Phi(s)$	Sur-linéarité en 0	Sous-linéarité en $+\infty$	Convexité	Différentiabilité
Φ_{L0}	$1 - \delta(s)$	Non	Oui	Non	Non
Φ_{L1}	$ s $	Oui	Oui	Oui	Non
Φ_{LS}	s^2	Oui	Non	Oui	Oui
Φ_{M1}	$2\sqrt{1+s^2} - 2$	Oui	Non	Oui	Oui
Φ_{M2}	$\ln(1+s^2)$	Oui	Oui	Non	Oui
Φ_{M3}	$\frac{s^2}{1+s^2}$	Oui	Oui	Non	Oui

TABLE 1.1 – Propriétés de plusieurs M-estimateurs. L'estimateur en norme 1 constitue un compromis très acceptable, car sa non différentiabilité n'est pas rédhibitoire. Toutefois, des résultats supérieurs sont souvent obtenus avec les estimateurs non convexes, même si leur non convexité représente une difficulté non négligeable.

Dans le vocabulaire des statisticiens, ce type d'estimation robuste est appelé M-estimation²² [130]. Mis à part les estimateurs en normes 0 et 1 discutés ci-dessus, les trois M-estimateurs les plus souvent rencontrés sont les suivants :

$$\Phi_{M1}(s) = 2\sqrt{1+s^2} - 2 \quad (1.50)$$

$$\Phi_{M2}(s) = \ln(1+s^2) \quad (1.51)$$

$$\Phi_{M3}(s) = \frac{s^2}{1+s^2} \quad (1.52)$$

présentés, respectivement, dans [55], [115] et [96]. Le lecteur intéressé trouvera une analyse détaillée de ces estimateurs dans [56].

Les différents estimateurs de la table 1.1 sont illustrés sur la figure 1.22 : l'estimateur Φ_{L1} est, parmi les estimateurs convexes, le seul qui admette une croissance sur-linéaire en 0 et sous-linéaire en $+\infty$, ce qui explique qu'il soit souvent utilisé. Notons toutefois qu'en dépit des difficultés numériques que pose la non convexité, il a été démontré récemment que les estimateurs non convexes étaient particulièrement adaptés à certains problèmes comme le débruitage d'images [158] ou l'intégration des normales [80].

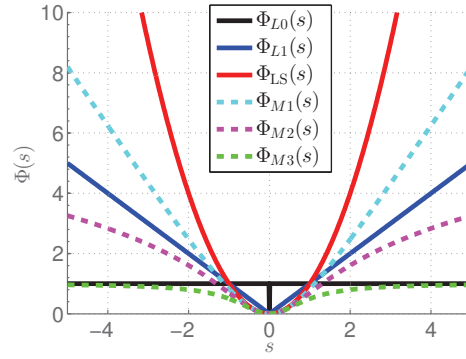


FIGURE 1.22 – Graphes des différents M-estimateurs de la table 1.1. La sur-linéarité en 0 est essentielle pour garantir la robustesse de l'estimation à un bruit modéré : à part l'estimateur Φ_{L0} , tous les estimateurs étudiés vérifient cette propriété. La sous-linéarité en $+\infty$ est, elle, nécessaire pour garantir la robustesse aux données aberrantes : Φ_{LS} et Φ_{M1} ne vérifient pas cette propriété. Enfin, la convexité est une propriété essentielle du point de vue numérique : Φ_{L0} , Φ_{M2} et Φ_{M3} ne sont pas des estimateurs convexes. Pour résumer, l'estimateur Φ_{L1} semble constituer le meilleur compromis.

22. Le M est l'abréviation de *Maximum de vraisemblance*. Chacun de ces estimateurs est en effet l'estimateur du maximum de vraisemblance pour une fonction de vraisemblance bien choisie. Par exemple, l'estimateur Φ_{L1} correspond à une distribution laplacienne, l'estimateur Φ_{LS} à une distribution gaussienne, et l'estimateur Φ_{M2} à une distribution de Cauchy.

La figure 1.23 montre les résultats obtenus par les estimateurs Φ_{L1} , Φ_{M1} , Φ_{M2} et Φ_{M3} sur le même jeu de données que celui de la figure 1.20. Dans ces tests, la minimisation est réalisée par itérations de type quasi-Newton (BFGS), en choisissant comme solution initiale la solution en moindres carrés.

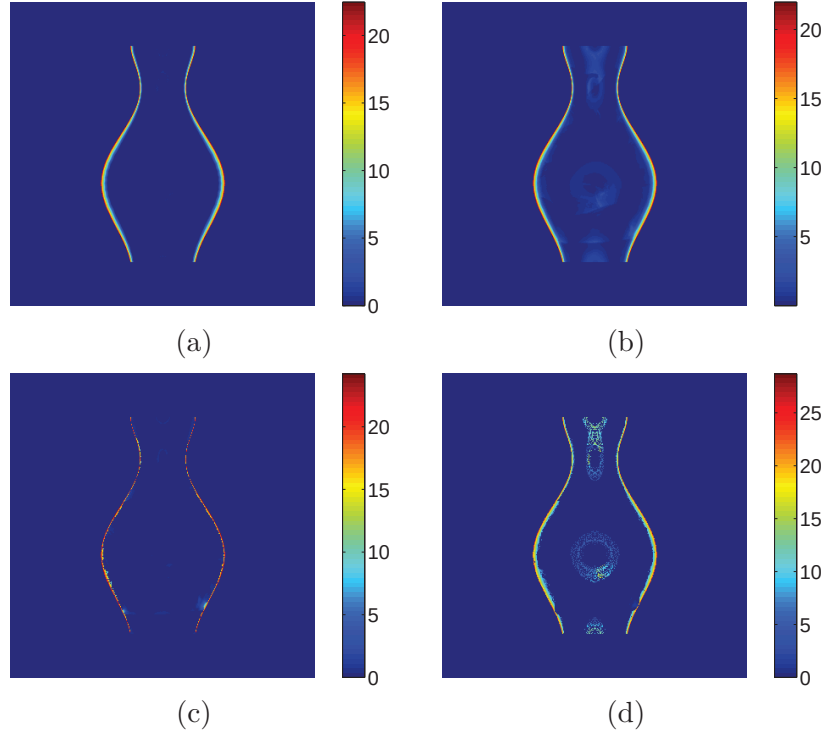


FIGURE 1.23 – Résultats obtenus sur le même jeu de données que celui de la figure 1.20 : (a) Φ_{L1} (EAM = 1,09°); (b) Φ_{M1} (EAM = 1,39°); (c) Φ_{M2} (EAM = 0,58°); (d) Φ_{M3} (EAM = 1,58°).

Ces résultats confirment l'intuition selon laquelle l'estimateur Φ_{L1} constitue un excellent compromis entre sur-linéarité en 0, sous-linéarité en $+\infty$ et convexité. L'estimateur non convexe Φ_{M2} offre cependant un niveau de précision encore supérieur, ce qui confirme les résultats expérimentaux présentés dans [158] dans le cadre du débruitage d'images. Il est surprenant que l'estimateur Φ_{M3} soit moins performant : nous verrons dans le chapitre 2 que cet estimateur robuste est plus adapté à l'intégration des normales. Remarquons enfin que le bord du vase, où les ombres propres sont les plus présentes, pose toujours problème. Nous présentons dans le paragraphe suivant une méthode permettant de limiter ce biais.

1.3.5 Approximation différentiable de l'opérateur max

Rappelons que la plupart du temps, une donnée n'est aberrante que vis-à-vis d'un certain modèle. En particulier, il peut sembler discutable de traiter les ombres propres comme des données aberrantes, alors que l'on dispose du modèle (1.16) pour les expliquer. Comme nous l'avons déjà mentionné, la principale difficulté de ce modèle réside dans sa *non linéarité*. La *non différentiabilité* en 0 de l'opérateur max peut représenter une difficulté supplémentaire. Face à ces deux difficultés, la parade la plus simple (et la plus utilisée) consiste à « oublier » l'opérateur max, i.e. à faire l'approximation :

$$\max\{0, x\} \approx x \quad (1.53)$$

qui bien entendu est erronée si x est négatif, i.e. dans le cas qui nous intéresse (celui des ombres propres).

Nous proposons d'utiliser une autre approximation de l'opérateur max, permettant de conserver le caractère différentiable du modèle (1.18), tout en modélisant explicitement les ombres propres (cf. figure 1.24) :

$$\max\{0, x\} \approx \frac{\ln(1 + \exp(\delta x))}{\delta} \quad (1.54)$$

où $\delta > 0$ est fixé a priori. Remarquons que l'approximation (1.54) est d'autant plus précise que δ est grand : dans les tests, nous utilisons $\delta = 5$.

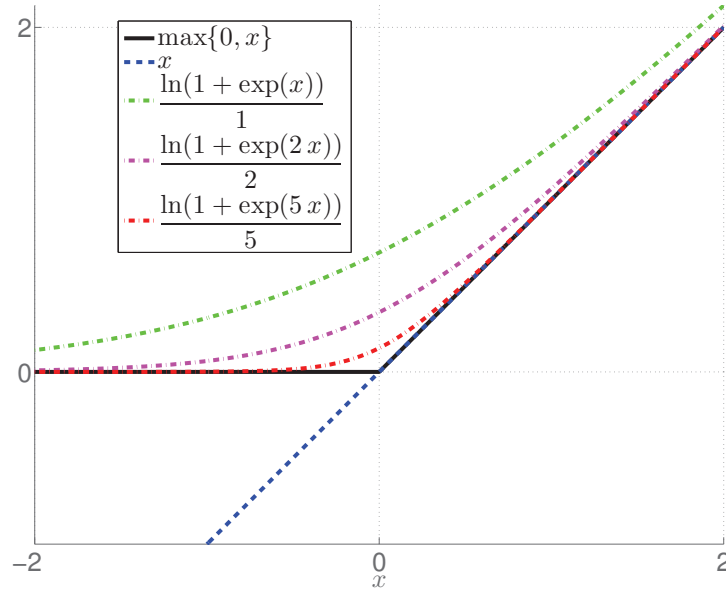


FIGURE 1.24 – Approximation différentiable (1.54) de l'opérateur max lorsque $\delta = 1$, $\delta = 2$ ou $\delta = 5$.

En utilisant l'approximation (1.54) de l'opérateur max, l'équation (1.16) devient :

$$I(u, v) = \frac{\ln(1 + \exp\{\delta \mathbf{s} \cdot \mathbf{m}(u, v)\})}{\delta} \quad (1.55)$$

La résolution d'un système formé par des équations du type (1.55) est plus aisée, car la non différentiabilité de l'opérateur max en 0 a disparu. Un algorithme de type descente de gradient peut donc être utilisé pour la résolution. On peut par exemple résoudre ce système d'équations au sens des moindres carrés :

$$\min_{\mathbf{m} \in \mathbb{R}^3} \sum_{i=1}^m \left[I^i(u, v) - \frac{\ln(1 + \exp\{\delta \mathbf{s}^i \cdot \mathbf{m}\})}{\delta} \right]^2 \quad (1.56)$$

en utilisant la méthode de Levenberg-Marcquardt. Le problème en moindres carrés non linéaires (1.56) est un problème convexe et différentiable. La descente de gradient converge donc très rapidement vers le minimum global. En outre, le gradient de la fonction objectif (1.56) est facile à calculer analytiquement, ce qui garantit des performances optimales.

Comme cela est illustré sur la figure 1.25, la prise en compte explicite des ombres propres permet d'annuler le biais dans l'estimation de la normale sur les fortes pentes. Par comparaison avec la figure 1.23, on constate qu'il est plus efficace de modéliser les ombres propres, plutôt que de les considérer comme des données aberrantes dans un processus d'estimation robuste. Rappelons toutefois que les ombres *portées*, au contraire des ombres propres, ne peuvent pas être modélisées localement. En présence d'ombres portées, de reflets brillants ou d'autres types d'écarts au modèle, il convient de remplacer les moindres carrés, dans le problème (1.56), par un estimateur robuste tel que ceux qui ont été présentés dans le paragraphe 1.3.4.

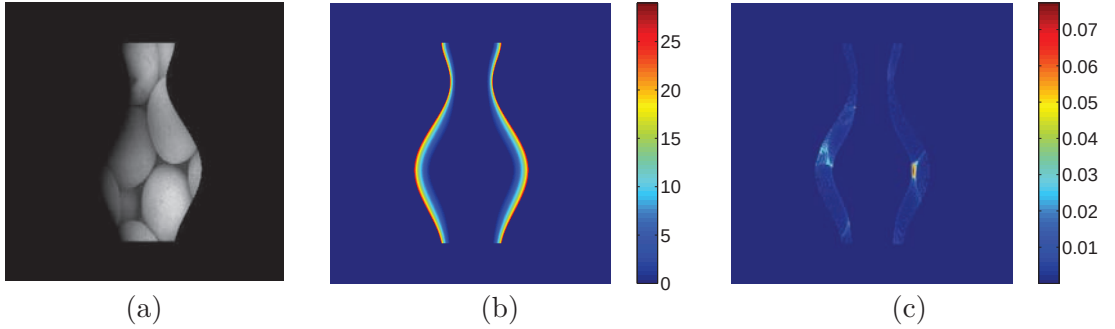


FIGURE 1.25 – (a) Une image de synthèse d'un vase (parmi $m = 100$). (b) Écart angulaire (en degrés) sur les normales estimées par résolution en moindres carrés, en utilisant le modèle lambertien linéaire ($EAM = 3, 11^\circ$). (c) Écart angulaire en utilisant l'approximation (1.54) de l'opérateur max ($EAM = 0, 001^\circ$).

1.4 Algorithme pour la stéréophotométrie calbrée robuste

Au vu des résultats des paragraphes précédents, nous proposons finalement l'algorithme 1 pour estimer la normale et l'albédo par stéréophotométrie calbrée robuste. La non uniformité des éclairages est compensée (cf. paragraphe 1.2.4), et les ombres propres sont modélisées explicitement (cf. paragraphe 1.3.5), tout comme l'éclairage ambiant (cf. paragraphe 1.3.1). Enfin, l'estimation est robuste aux données aberrantes grâce au choix d'un M-estimateur (cf. paragraphe 1.3.4).

Algorithme 1: Estimation robuste de la normale et de l'albédo.

Entrée :

- Images en niveaux de gris I^i , $i \in [1, m]$.
- Domaine de reconstruction Ω .
- Directions unitaires $\bar{\mathbf{s}}^i \in \mathbb{R}^3$ des éclairages, $i \in [1, m]$.
- Intensités $\phi^i : \Omega \rightarrow \mathbb{R}$ des éclairages, $i \in [1, m]$.
- Paramètre δ ($\delta = 5$ dans nos tests).
- Estimateur robuste $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ ($\Phi = \Phi_{M2}$ dans nos tests).

Sortie :

- Champ de normales $\mathbf{n} : \Omega \rightarrow \mathbb{R}^3$.
- Albédo $\rho : \Omega \rightarrow \mathbb{R}$.

Pour tout $(u, v) \in \Omega$, **faire** :

$$(\mathbf{m}(u, v), \phi_a(u, v)) = \underset{\mathbf{m} \in \mathbb{R}^3, \phi_a \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^m \Phi \left(\left| \frac{I^i(u, v)}{\phi^i(u, v)} - \frac{\ln \left(1 + \exp \left\{ \delta \bar{\mathbf{s}}^i \cdot \mathbf{m} \right\} \right)}{\delta} - \phi_a \right| \right) \quad (1.57)$$

$$\mathbf{n}(u, v) = \frac{\mathbf{m}(u, v)}{\|\mathbf{m}(u, v)\|} \quad (1.58)$$

$$\rho(u, v) = \|\mathbf{m}(u, v)\| \quad (1.59)$$

Fin Pour

1.5 Cas des images colorées

Avant de terminer ce chapitre sur l'estimation robuste des normales et de l'albédo, interrogeons-nous sur la façon d'étendre les méthodes présentées au cas des images colorées. Dans le modèle lambertien linéaire, deux grandeurs ont une dépendance spectrale : l'éclairage $\mathbf{s}$ et l'albédo ρ . Nous noterons $\mathbf{s}_\star^i = \phi_\star^i \bar{\mathbf{s}}_\star^i$ l'éclairage correspondant à l'image $I_\star^i$, $(i, \star) \in [1, m] \times \{R, G, B\}$, et $\rho_\star$ l'albédo relativement au canal $\star \in \{R, G, B\}$. En

supposant les éclairages directionnels, le modèle lambertien linéaire devient²³ :

$$I_\star^i(u, v) = \phi_\star \rho_\star(u, v) \overline{\mathbf{s}_\star^i} \cdot \mathbf{n}(u, v) \quad (1.60)$$

Malheureusement, il est impossible d'écrire le problème de l'estimation des inconnues $(\mathbf{n}, \rho_R, \rho_G, \rho_B)$ comme un problème linéaire, sous une forme similaire à (1.18). Nous verrons dans la partie IV comment découpler l'estimation de l'albédo de celle de la normale, ce qui nous permettra de traiter le cas général où la surface et les éclairages sont simultanément colorés. En suivant l'approche classique, il nous faut supposer que la surface est blanche et les éclairages colorés [117] ou bien, inversement, que la surface est colorée et les éclairages blancs [20].

Une première utilisation intéressante de la couleur en stéréophotométrie consiste à supposer les éclairages colorés, mais l'albédo identique dans les trois canaux, i.e. $\rho_\star = \rho$. Ce cas correspond à des éclairages colorés placés face à une surface blanche, ou peinte en blanc [150], qui permet par exemple de réaliser la reconstruction 3D en temps réel à partir d'une seule image RGB [117]. Nous verrons dans le paragraphe 4.4.1 que la résolution de ce problème se ramène exactement au cas des images en niveaux de gris.

Étudions maintenant la configuration inverse, où l'albédo est coloré, mais l'éclairage est indépendant de toute notion de couleur. En négligeant les dépendances en (u, v) , le système d'équations linéaires à résoudre s'écrit :

$$\underbrace{\begin{bmatrix} I_R^1 & I_G^1 & I_B^1 \\ \vdots & \vdots & \vdots \\ I_R^m & I_G^m & I_B^m \end{bmatrix}}_{\mathbf{I} \in \mathbb{R}^{m \times 3}} = \underbrace{\begin{bmatrix} \mathbf{s}^1{}^\top \\ \vdots \\ \mathbf{s}^m{}^\top \end{bmatrix}}_{\mathbf{S} \in \mathbb{R}^{m \times 3}} \underbrace{\mathbf{n} [\rho_R, \rho_G, \rho_B]}_{\mathbf{M} \in \mathbb{R}^{3 \times 3}} \quad (1.61)$$

Sachant que, de par sa définition, la matrice $\mathbf{M}$ est de rang 1, nous devons donc résoudre le problème en moindres carrés linéaires sous contrainte suivant :

$$\begin{cases} \min_{\mathbf{M} \in \mathbb{R}^{3 \times 3}} \|\mathbf{I} - \mathbf{S}\mathbf{M}\|_F^2 \\ \text{s.c. rang}(\mathbf{M}) = 1 \end{cases} \quad (1.62)$$

En notant :

$$\mathbf{M}_0 = \mathbf{S}^\top \mathbf{I} = (\mathbf{S}^\top \mathbf{S})^{-1} \mathbf{S}^\top \mathbf{I} \quad (1.63)$$

le problème (1.62) peut être réécrit :

$$\begin{cases} \min_{\mathbf{M} \in \mathbb{R}^{3 \times 3}} \|\mathbf{M} - \mathbf{M}_0\|_F^2 \\ \text{s.c. rang}(\mathbf{M}) = 1 \end{cases} \quad (1.64)$$

de telle sorte qu'il s'agit d'un problème d'approximation de rang 1 d'une matrice. Par SVD tronquée à l'ordre 1, on peut approcher $\mathbf{M}_0$ par :

$$\mathbf{M}_0 \approx \mathbf{u}_0 \mathbf{w} \mathbf{v}_0^\top \quad (1.65)$$

23. Cette écriture n'est théoriquement justifiée qu'en assimilant la réponse du récepteur photosensible à la superposition de trois fonctions δ de Dirac.

où $w > 0$, et $(\mathbf{u}_0, \mathbf{v}_0)$ sont deux vecteurs unitaires de $\mathbf{R}^3$. La solution en $\mathbf{n}$ et $\boldsymbol{\rho} = [\rho_R, \rho_G, \rho_B]^\top$ s'écrit alors :

$$(\mathbf{n}, \boldsymbol{\rho}) = \pm(\mathbf{u}_0, w\mathbf{v}_0) \quad (1.66)$$

où l'ambiguïté sur le signe est facilement levée, puisque la troisième coordonnée de $\mathbf{n}$ doit être positive. Ces deux étapes (estimation (1.63) de $\mathbf{M}_0$, puis « séparation » (1.66) de $\mathbf{M}_0$ entre normale et albédo) doivent être réalisées pixel par pixel. Comme cela est illustré sur la figure 1.26, cette procédure permet effectivement d'estimer l'albédo d'une surface colorée.

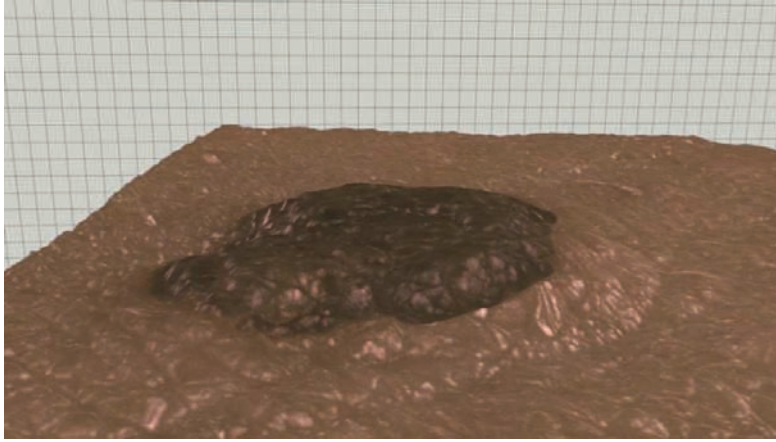


FIGURE 1.26 – Reconstruction 3D d'un « naevus », à partir de $m = 15$ images RGB acquises par le dermoscope de la société Pixience. Après avoir estimé la matrice $\mathbf{M}_0$ en chaque point par (1.63), l'albédo coloré et les normales sont obtenus en utilisant (1.66).

L'expérience de la figure 1.26 montre que, si les variations de couleur sont souvent corrélées aux variations de relief, cela n'est pas toujours vrai. L'utilisation de l'albédo comme variable auxiliaire est donc surtout motivée par la volonté de linéariser le problème de la stéréophotométrie. Nous verrons dans les prochaines parties que cette linéarisation limite en réalité les applications de la stéréophotométrie, plus qu'elle n'en facilite la résolution.

Quoi qu'il en soit, il est relativement facile d'étendre les méthodes adaptées aux images en niveaux de gris à des images RGB, pourvu que soit les éclairages, soit les albédos, soient indépendants de la couleur. Le modèle lambertien linéaire ne permet cependant pas de modéliser les dépendances spectrales de l'albédo et des éclairages sous la forme d'un problème linéaire : il s'agit d'une limite de l'approche classique de la stéréophotométrie. Nous verrons dans la partie IV qu'en découplant les estimations des caractéristiques géométriques et photométriques de la surface, grâce à une approche variationnelle, ce problème peut cependant être résolu.

1.6 Conclusion du chapitre

Dans ce chapitre, nous avons proposé une analyse statistique de la méthode d'estimation classique de la normale et de l'albédo par stéréophotométrie, qui est fondée sur le modèle lambertien linéaire (1.18). Nous avons mis en évidence la non robustesse de cette approche à la présence d'écarts importants au modèle, et recensé un certain nombre de phénomènes induisant ces écarts. Les effets d'éclairages non directionnels peuvent être fortement atténués au moyen de prétraitements relativement simples. D'autres prétraitements permettent de limiter les effets des ombres et des reflets brillants, mais ces méthodes non supervisées sont généralement moins performantes que des méthodes d'estimation robuste supervisées, telles que la régression parcimonieuse ou certains M-estimateurs non convexes. Enfin, l'éclairage « ambiant » est très facile à estimer automatiquement.

Ces prétraitements et ces estimateurs permettent d'estimer le champ de normales de manière robuste, i.e. d'appliquer la formulation classique de la stéréophotométrie « en conditions réalistes ». Si la surface à reconstruire est régulière, le champ de normales estimé de façon robuste a toutes les raisons d'être lui-même lisse : remarquons que nous n'avons pas eu besoin, pour cela, d'introduire de régularisation *spatiale*, la normale étant estimée pixel par pixel. Jusqu'à présent, nous ne nous sommes intéressés qu'à l'estimation de caractéristiques locales de la surface. Cependant, notre objectif final est de retrouver une surface expliquant les images, et non pas seulement les normales à cette surface. La prochaine étape consistera donc à *intégrer* le champ de normales estimé, afin d'obtenir une carte de profondeur. Nous allons voir que cette étape d'intégration requiert également l'utilisation d'estimateurs robustes, afin d'éviter l'apparition d'un biais à proximité des discontinuités de profondeur.

Intégration robuste des normales

La seconde étape de la résolution classique de la stéréophotométrie est l'*intégration* des normales estimées, qui permet de retrouver la carte de profondeur z caractéristique du relief de la surface $\mathfrak{S}$. Dans ce chapitre, nous menons une étude approfondie de ce problème. Après avoir rappelé dans le paragraphe 2.1 les équations qui le caractérisent, nous présentons dans le paragraphe 2.2 les principales propriétés que l'on est en droit d'attendre d'une « bonne » méthode d'intégration. Nous menons ensuite, dans le paragraphe 2.3, une analyse approfondie des méthodes d'intégration en moindres carrés. Sur la base de cette analyse, de nouvelles méthodes d'intégration adaptées aux discontinuités de profondeur sont introduites dans le paragraphe 2.4. Plusieurs d'entre elles ont déjà été décrites dans [210, 211].

2.1 Normale, gradient et intégrabilité

Le problème de l'intégration d'un champ de normales consiste à estimer, en chaque pixel $(u, v) \in \Omega$, le point $\mathbf{x}(u, v) = [x(u, v), y(u, v), z(u, v)]^\top$ de la surface $\mathfrak{S}$, connaissant sa normale $\mathbf{n}(u, v)$. Nous avons vu, dans le chapitre 1, comment estimer le champ de normales de la surface $\mathfrak{S}$. Pour en déduire une estimation du relief de la surface, i.e. de la profondeur $z(u, v)$, il nous faut d'abord expliciter le lien entre ces différentes variables.

2.1.1 Équations de l'intégration des normales

Commençons par établir l'équation qui lie la normale $\mathbf{n}(u, v)$ au gradient de profondeur $\nabla z(u, v)$. Nous allons voir que cette équation est quasiment la même pour les deux types de projection : orthographique ou perspectif.

Projection orthographique

Sous l'hypothèse d'une projection orthographique (cf. (1.3)) :

$$\mathbf{x}(u, v) = \begin{bmatrix} u \\ v \\ z(u, v) \end{bmatrix} \quad (2.1)$$

Le plan tangent à la surface est porté par les deux vecteurs $\partial_u \mathbf{x}(u, v)$ et $\partial_v \mathbf{x}(u, v)$,

dont le produit vectoriel s'écrit :

$$\partial_u \mathbf{x}(u, v) \times \partial_v \mathbf{x}(u, v) = \begin{bmatrix} -\partial_u z(u, v) \\ -\partial_v z(u, v) \\ 1 \end{bmatrix} \quad (2.2)$$

Sachant que ce vecteur est colinéaire à la normale à la surface, son produit vectoriel avec $\mathbf{n}(u, v) = [n_1(u, v), n_2(u, v), n_3(u, v)]^\top$ est nul, soit :

$$\begin{cases} n_3(u, v) \partial_u z(u, v) = -n_1(u, v) \\ n_3(u, v) \partial_v z(u, v) = -n_2(u, v) \\ n_2(u, v) \partial_u z(u, v) - n_1(u, v) \partial_v z(u, v) = 0 \end{cases} \quad (2.3)$$

ce qui constitue un système linéaire de trois équations à deux inconnues, qui sont les dérivées partielles $\partial_u z(u, v)$ et $\partial_v z(u, v)$ de la profondeur. Or, la troisième équation du système (2.3) se déduit des deux premières par combinaison linéaire, pour peu que $n_3(u, v) \neq 0$. Le cas où $n_3(u, v) = 0$ survient lorsque la normale $\mathbf{n}(u, v)$ est parallèle au plan image, i.e. lorsque la « pente » $\|\nabla z(u, v)\| = \sqrt{\partial_u z(u, v)^2 + \partial_v z(u, v)^2}$ est infinie. On ne peut donc pas calculer par intégration la profondeur en de tels points, qui forment dans l'image des *contours occultants*.

Si le point (u, v) ne se trouve pas sur un contour occultant, i.e. si $n_3(u, v) \neq 0$, alors la troisième équation du système (2.3) peut être ignorée, puisqu'elle n'est pas informative. Le déterminant du système linéaire formé par les deux premières équations de (2.3) est égal à $n_3(u, v)^2$. Comme cette valeur est non nulle par hypothèse, la détermination de $\nabla z(u, v) = [\partial_u z(u, v), \partial_v z(u, v)]^\top$ à partir de $\mathbf{n}(u, v)$ constitue un problème bien posé. L'équation liant $z(u, v)$ à $\mathbf{n}(u, v)$ est l'équation aux dérivées partielles (EDP) linéaire suivante :

$$\nabla z(u, v) = \begin{bmatrix} -\frac{n_1(u, v)}{n_3(u, v)} \\ -\frac{n_2(u, v)}{n_3(u, v)} \end{bmatrix} \quad (2.4)$$

En introduisant les notations usuelles :

$$\underbrace{\begin{bmatrix} p(u, v) \\ q(u, v) \end{bmatrix}}_{\mathbf{g}_o(u, v)} = \begin{bmatrix} -\frac{n_1(u, v)}{n_3(u, v)} \\ -\frac{n_2(u, v)}{n_3(u, v)} \end{bmatrix} \quad (2.5)$$

l'EDP (2.4) se réécrit sous la forme suivante :

$$\nabla z(u, v) = \mathbf{g}_o(u, v) \quad (2.6)$$

où $\mathbf{g}_o(u, v)$ est une estimation du gradient de z au point (u, v) , obtenue par exemple par stéréophotométrie. Comme, d'après (2.1), $\mathbf{x}(u, v)$ est entièrement déterminé par $z(u, v)$, la résolution de l'équation (2.6) permet bien de reconstruire le relief de la surface observée. Notons que cette résolution n'est possible qu'à une constante additive près, qui peut être fixée empiriquement ou grâce à une connaissance a priori sur la surface.

Projection perspective

Sous l'hypothèse d'une projection perspective (cf. (1.4)) :

$$\mathbf{x}(u, v) = z(u, v) \begin{bmatrix} \frac{u-u_0}{f} \\ \frac{v-v_0}{f} \\ 1 \end{bmatrix} \quad (2.7)$$

où (u_0, v_0) désignent les coordonnées du *point principal* de l'appareil photographique.

Le produit vectoriel de $\partial_u \mathbf{x}(u, v)$ et $\partial_v \mathbf{x}(u, v)$ s'écrit maintenant :

$$\partial_u \mathbf{x}(u, v) \times \partial_v \mathbf{x}(u, v) = \frac{z(u, v)}{f^2} \begin{bmatrix} -f \partial_u z(u, v) \\ -f \partial_v z(u, v) \\ z(u, v) + (u - u_0) \partial_u z(u, v) + (v - v_0) \partial_v z(u, v) \end{bmatrix} \quad (2.8)$$

Comme ce vecteur est colinéaire à la normale, son produit vectoriel avec $\mathbf{n}(u, v)$ est nul, soit :

$$\begin{cases} f n_3(u, v) \partial_u z(u, v) + n_1(u, v) [z(u, v) + (u - u_0) \partial_u z(u, v) + (v - v_0) \partial_v z(u, v)] = 0 \\ f n_3(u, v) \partial_v z(u, v) + n_2(u, v) [z(u, v) + (u - u_0) \partial_u z(u, v) + (v - v_0) \partial_v z(u, v)] = 0 \\ n_2(u, v) \partial_u z(u, v) - n_1(u, v) \partial_v z(u, v) = 0 \end{cases} \quad (2.9)$$

En introduisant le changement de variable suivant (on rappelle que $z < 0$, par convention) :

$$\tilde{z} = \ln|z| \quad (2.10)$$

le système (2.9) devient linéaire en $\partial_u \tilde{z}(u, v)$ et $\partial_v \tilde{z}(u, v)$:

$$\begin{cases} [f n_3(u, v) + (u - u_0) n_1(u, v)] \partial_u \tilde{z}(u, v) + (v - v_0) n_1(u, v) \partial_v \tilde{z}(u, v) = -n_1(u, v) \\ (u - u_0) n_2(u, v) \partial_u \tilde{z}(u, v) + [f n_3(u, v) + (v - v_0) n_2(u, v)] \partial_v \tilde{z}(u, v) = -n_2(u, v) \\ n_2(u, v) \partial_u \tilde{z}(u, v) - n_1(u, v) \partial_v \tilde{z}(u, v) = 0 \end{cases} \quad (2.11)$$

La matrice du système linéaire (2.11) est de rang 2, sauf si ses trois déterminants sont simultanément nuls :

$$\begin{cases} f n_3(u, v) [u - u_0, v - v_0, f]^\top \cdot \mathbf{n}(u, v) = 0 \\ -n_1(u, v) [u - u_0, v - v_0, f]^\top \cdot \mathbf{n}(u, v) = 0 \\ -n_2(u, v) [u - u_0, v - v_0, f]^\top \cdot \mathbf{n}(u, v) = 0 \end{cases} \quad (2.12)$$

Sachant que $\mathbf{n}(u, v) \neq \mathbf{0}$, cela ne peut se produire que si :

$$[u - u_0, v - v_0, f]^\top \cdot \mathbf{n}(u, v) = 0 \quad (2.13)$$

i.e., si (u, v) se situe sur un contour occultant. Dans les autres cas, on obtient finalement l'EDP linéaire suivante :

$$\nabla \tilde{z}(u, v) = \underbrace{\begin{bmatrix} -\frac{n_1(u, v)}{[u - u_0, v - v_0, f]^\top \cdot \mathbf{n}(u, v)} \\ -\frac{n_2(u, v)}{[u - u_0, v - v_0, f]^\top \cdot \mathbf{n}(u, v)} \end{bmatrix}}_{\mathbf{g}_p(u, v)} \quad (2.14)$$

et, une fois cette équation résolue, la profondeur $z(u, v)$ s'obtient par la transformation inverse (le signe $-$ vient de ce que $z < 0$) :

$$z(u, v) = -\exp(\tilde{z}(u, v)) \quad (2.15)$$

La stricte similarité entre les équations (2.6) et (2.14) montre qu'une méthode d'intégration conçue pour la projection orthographique peut être facilement étendue à la projection perspective, pourvu que les paramètres internes (u_0, v_0, f) soient connus, en effectuant le changement de variable (2.10)¹. Nous pouvons donc nous contenter de résoudre l'équation générique suivante :

$$\nabla z(u, v) = \mathbf{g}(u, v) \quad (2.16)$$

où il est implicitement entendu que $(z, \mathbf{g}) \equiv (z, \mathbf{g}_o)$ pour la projection orthographique, et $(z, \mathbf{g}) \equiv (\tilde{z}, \mathbf{g}_p)$ pour la projection perspective. Dans les deux cas, nous noterons pour simplifier :

$$\mathbf{g}(u, v) = \begin{bmatrix} p(u, v) \\ q(u, v) \end{bmatrix} \quad (2.17)$$

2.1.2 Intégrabilité d'un champ de normales

Soit $(\bar{u}, \bar{v})$ un pixel où la profondeur est connue. La solution de l'EDP linéaire (2.16) peut être obtenue en *intégrant* la donnée $\mathbf{g} = [p, q]^\top$ le long d'un chemin reliant $(\bar{u}, \bar{v})$ à (u, v) :

$$z(u, v) = z(\bar{u}, \bar{v}) + \int_{(\bar{u}, \bar{v})}^{(u, v)} \mathbf{g}(s, t) \cdot [ds, dt]^\top \quad (2.18)$$

S'il n'existe aucun point $(\bar{u}, \bar{v})$ de profondeur connue, on peut toujours choisir arbitrairement $(\bar{u}, \bar{v})$ et $z(\bar{u}, \bar{v})$: la profondeur z sera alors déterminée à une constante additive près en tout point $(u, v) \in \Omega$ (cette ambiguïté devient une ambiguïté multiplicative dans le cas de la projection perspective, en vertu de la transformation (2.15)).

En général, le chemin choisi pour calculer l'intégrale (2.18) influe sur la valeur de $z(u, v)$, ce qui est bien entendu préjudiciable. Pour garantir l'indépendance de la valeur de $z(u, v)$ vis-à-vis du chemin d'intégration, il faut s'assurer que le champ $\mathbf{g} = [p, q]^\top$ soit *intégrable*, ce qui se traduit par l'égalité :

$$\partial_v p = \partial_u q \quad (2.19)$$

Cette contrainte est appelée *contrainte d'intégrabilité* [93, 126]. Elle correspond au théorème de Schwartz : si une fonction z est de classe $\mathcal{C}^2$, alors ses dérivées secondes croisées sont égales :

$$\frac{\partial^2 z}{\partial_u \partial_v} = \frac{\partial^2 z}{\partial_v \partial_u} \quad (2.20)$$

Imposer *explicitement* la contrainte d'intégrabilité (2.19) peut aider à rendre mieux posés certains problèmes, comme la stéréophotométrie non calibrée [197, 282, 283] ou la

1. Cela a également été remarqué par Tankus et Kiryati dans [249].

stéréophotométrie à deux images [192] (cf. partie II). Pour le problème de l'intégration des normales, nous préconisons de ne pas imposer explicitement l'intégrabilité, i.e. de ne pas modifier le champ $\mathbf{g}$, contrairement à ce qui est suggéré dans [224] : si ce champ a été bien estimé (cf. chapitre 1), les points où la contrainte d'intégrabilité n'est pas vérifiée correspondent à des points particuliers de la surface, qu'il convient de traiter de façon appropriée.

Il nous semble préférable d'imposer cette contrainte *implicitement*, en cherchant la solution z dans un espace de fonctions bien choisi pour lequel l'égalité (2.20) est vérifiée. Si l'on cherche une fonction z deux fois différentiable sur Ω , le choix de l'espace $\mathcal{C}^2(\Omega)$ semble justifié. Mais en présence d'arêtes ou de discontinuités de profondeur, la contrainte d'intégrabilité (2.19) ne peut être imposée que *presque partout*. Par analogie avec ce qui a été proposé dans le domaine du traitement d'images, nous préférons pour finir l'espace $BV(\Omega)$ des fonctions à variation bornée².

2.2 Difficultés de l'intégration des normales

Nous détaillons maintenant les principales difficultés à surmonter pour résoudre l'équation (2.16) et effectuons un bref état de l'art.

2.2.1 Robustesse au bruit et aux données aberrantes

La première propriété que l'on est en droit d'attendre d'une méthode d'intégration des normales est sa robustesse aux erreurs numériques survenues lors de l'estimation du champ de normales $\mathbf{n}$.

Même lorsque $\mathbf{n}$ est estimé de façon robuste, il est très probable que le champ $\mathbf{g}$ qui en découle diffère de la valeur réelle de ∇z et qu'il ne soit pas intégrable. Les méthodes utilisant une approximation numérique de l'intégrale dans (2.18), comme celle de Wu et Li [274], ou les méthodes par propagation de front [95, 123], sont donc mal adaptées à ce défaut, car elles tendent à accumuler les erreurs le long des chemins d'intégration (cf. figure 2.1)³. Ces méthodes présentent cependant l'avantage d'être très rapides, et peuvent donc servir d'initialisation à des algorithmes plus robustes.

Le moyen le plus simple pour tenir compte d'un éventuel biais dans l'estimation des normales consiste à résoudre l'équation (2.16) de manière approchée, sur l'ensemble du domaine Ω ⁴, en cherchant à résoudre le problème d'optimisation en moindres carrés suivant :

$$\min_z \iint_{(u,v) \in \Omega} \|\nabla z(u,v) - \mathbf{g}(u,v)\|^2 du dv \quad (2.21)$$

2. Nous renvoyons le lecteur à [14] pour une analyse détaillée de ces espaces fonctionnels, et à [15] pour des illustrations dans le domaine du traitement d'images.

3. Remarquons cependant que des efforts ont été récemment consacrés à l'amélioration de la robustesse de l'approche par *fast marching* [18].

4. Nous utiliserons abusivement la même notation Ω pour décrire le domaine de reconstruction discret (bord compris) et un ouvert de $\mathbb{R}^2$ dans les problèmes continus.

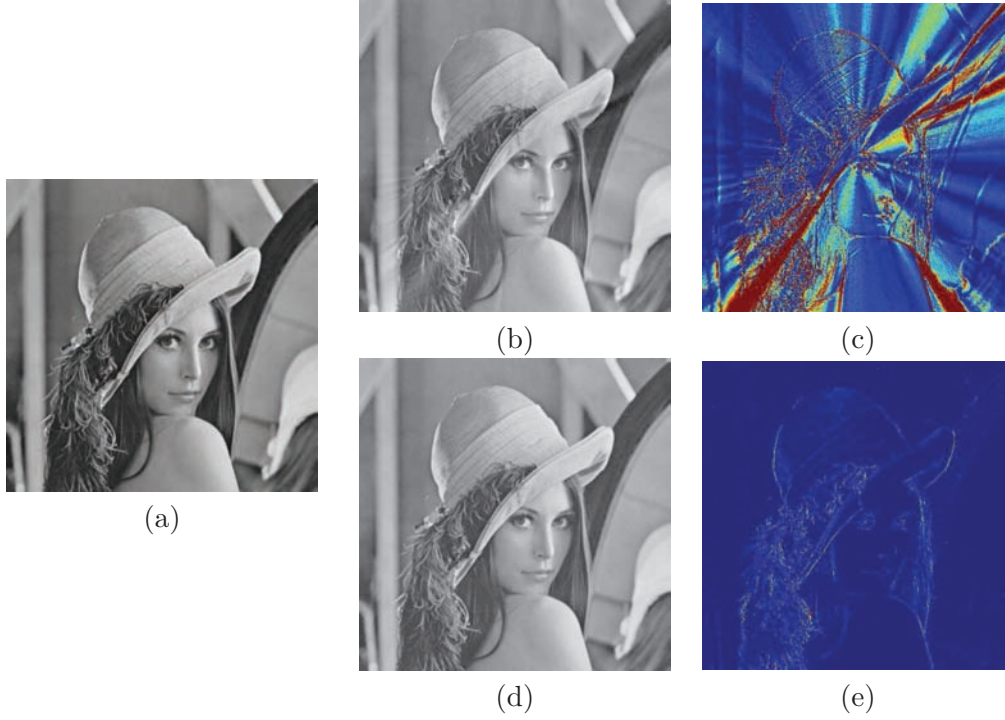


FIGURE 2.1 – (a) Exemple d'image en niveaux de gris. Après avoir calculé par différences finies le gradient de cette image, on cherche à la reconstruire par intégration. (b) Résultat de la reconstruction par propagation, en utilisant la méthode du *fast marching* [95], qui propage un front depuis un point initial, choisi ici au centre de l'image. (c) Écart quadratique, en fausses couleurs, entre (a) et (b). L'accumulation des erreurs lors de la propagation du front est clairement visible. (d) Résultat de l'approche en moindres carrés [239]. (e) Écart quadratique entre (a) et (d), en utilisant la même échelle de couleurs que pour (c). Le biais est concentré au niveau des discontinuités.

Cette approche, proposée initialement par Horn et Brooks [126] et récemment revisitée par Harker et O'Leary dans [109], est optimale si l'erreur d'estimation de $\mathbf{g}$ est gaussienne. Malheureusement, si l'on peut raisonnablement supposer que l'erreur sur l'estimation de la normale est gaussienne, la définition (2.5) de $\mathbf{g}$ indique que l'erreur sur $\mathbf{g}$ est le *quotient* de deux distributions gaussiennes : elle est donc distribuée selon une loi de Cauchy, ce qui nécessite l'utilisation de l'estimateur non convexe Φ_{M2} (cf. table 1.1). Nous reviendrons, dans le paragraphe 2.4.3, sur cette estimation, qui est un cas particulier du problème général suivant :

$$\min_z \iint_{(u,v) \in \Omega} \Phi(\|\nabla z(u,v) - \mathbf{g}(u,v)\|) du dv \quad (2.22)$$

pour lequel certains choix de l'estimateur Φ ont été discutés dans [6, 17, 76, 80, 211].

2.2.2 Topologie du domaine de reconstruction

Une « bonne » méthode d'intégration doit pouvoir s'adapter à un domaine de reconstruction Ω de forme quelconque⁵. En effet, dans le contexte de la stéréophotométrie, la surface à reconstruire occupe rarement tout le champ de l'image. Le domaine Ω est généralement déterminé par une segmentation préalable (cf. figure 1.2-d). Certes, le champ $\mathbf{g}$ peut toujours être complété par des valeurs nulles en dehors de Ω , de façon à ce qu'il soit défini sur l'image entière, c'est-à-dire sur une grille rectangulaire, mais cette « astuce » doit précisément être évitée si l'on prétend à une certaine précision car, à de rares exceptions près, cela rend le champ $\mathbf{g}$ fortement non intégrable sur la frontière $\partial\Omega$ du domaine Ω et induit des *phénomènes de Gibbs* (cf. figure 2.6). Ce risque est également souligné par Harker et O'Leary dans [111], qui conseillent de plutôt extrapoler les valeurs de $\mathbf{g}$ en dehors de Ω , pour les méthodes d'intégration qui imposent au domaine Ω d'être rectangulaire.

Malheureusement, en dehors des approches par propagation dont nous avons déjà discuté les limites, la plupart des méthodes d'intégration rapides existantes imposent au domaine Ω d'être rectangulaire [93, 109, 110, 111, 112, 153, 239, 261]. Nous verrons que lorsque Ω est rectangulaire, le système linéaire qui découle du problème (2.21) a effectivement une structure bien particulière qui permet une résolution très rapide. Lorsque Ω n'est pas rectangulaire, cette structure est malheureusement perdue.

2.2.3 Condition au bord

Une approche classique pour résoudre le problème (2.21) consiste à utiliser le calcul des variations. La condition d'optimalité associée au problème variationnel (2.21) est l'équation d'Euler-Lagrange :

$$\underbrace{\partial_z \mathcal{L}(u, v, z, \nabla z)}_{= 0} - \nabla \cdot \underbrace{\partial_{\nabla z} \mathcal{L}(u, v, z, \nabla z)}_{= 2[\nabla z(u, v) - \mathbf{g}(u, v)]} = 0 \quad (2.23)$$

où $\mathcal{L}$ désigne le *lagrangien* du problème (2.21) :

$$\mathcal{L}(u, v, z, \nabla z) = \|\nabla z(u, v) - \mathbf{g}(u, v)\|^2 \quad (2.24)$$

et où $\nabla \cdot$ désigne l'opérateur divergence, qui est l'opposé de l'adjoint de l'opérateur gradient. Puisque $\nabla \cdot \nabla = \Delta$, où Δ est l'opérateur laplacien, la condition d'optimalité associée au problème (2.21) est une *équation de Poisson* :

$$\Delta z(u, v) = \nabla \cdot \mathbf{g}(u, v) \quad (2.25)$$

qui doit être résolue en tout point (u, v) de l'ouvert Ω .

L'équation (2.25) n'admet pas une solution unique : une fonction *harmonique*, i.e. une solution de l'équation $\Delta z = 0$, peut être ajoutée à une solution pour former une

5. Nous supposons toutefois, par souci de simplicité, que Ω est connexe, puisque toute méthode d'intégration permet de traiter chaque composante connexe de Ω de façon indépendante.

nouvelle solution. Pour lever cette ambiguïté, il faut imposer une contrainte sur le bord $\partial\Omega$. Or, la solution de l'équation de Poisson (2.25) n'est la solution du problème (2.21) que si elle est complétée par la *condition naturelle* au bord $\partial\Omega$. On peut établir, par le calcul des variations, que cette condition au bord s'écrit [126] :

$$[\nabla z(u, v) - \mathbf{g}(u, v)] \cdot \boldsymbol{\mu}(u, v) = 0 \quad (2.26)$$

où $\boldsymbol{\mu}(u, v)$ est un vecteur normal au bord $\partial\Omega$ au point (u, v) . Cette condition au bord est donc de type Neumann.

Remplacer (2.26) par une autre condition au bord, par exemple une condition de type Dirichlet ($z = z_0$ sur $\partial\Omega$) ou Neumann homogène ($\nabla z \cdot \boldsymbol{\mu} = 0$ sur $\partial\Omega$ [6]), induit généralement un biais dans la reconstruction [109] : on préférera les méthodes d'intégration utilisant, explicitement ou implicitement, la condition naturelle au bord.

Il est remarquable que la méthode d'intégration la plus souvent citée, due à Frankot et Chellappa [93], utilise justement une condition au bord différente de la condition naturelle, à savoir une condition *périodique* (cf. figure 2.2). Pour un domaine rectangulaire $\Omega = [1, n_1] \times [1, n_c]$, cela signifie que $z(1, \cdot) = z(n_1, \cdot)$ et $z(\cdot, 1) = z(\cdot, n_c)$. Simchony et al. ont montré dans [239] comment étendre cette méthode à une condition au bord de type Dirichlet ou à la condition naturelle au bord. Comme l'ont fait remarquer Harker et O'Leary dans [112], ces extensions de la méthode de Frankot et Chellappa sont malheureusement biaisées. En effet, les différences finies utilisées pour discrétiser l'opérateur laplacien dans (2.25) ont une précision d'ordre 2, alors que la précision des différences finies qui permettent de discrétiser l'opérateur gradient dans (2.26) est seulement d'ordre 1. Ces observations ont conduit Harker et O'Leary, dans [109], à rejeter l'approche variationnelle (2.25) : « Toute approche fondée sur l'équation d'Euler-Lagrange n'est valide que pour quelques cas particuliers »⁶. Cette affirmation est néanmoins exagérée : en imposant la condition naturelle au bord, et en faisant un choix approprié de différences finies, l'approche variationnelle peut être non biaisée.

En réalité, le moyen le plus simple pour ne pas induire de biais au bord du domaine Ω consiste justement à *ne pas* imposer de condition au bord, ce qui est possible si l'on minimise directement la fonctionnelle discrète associée à (2.21), comme cela a été montré par Durou et Courteille dans [81] et par Harker et O'Leary dans [109].

2.2.4 Robustesse aux discontinuités de profondeur

Nous avons déjà mentionné que la robustesse d'une méthode d'intégration à une estimation erronée des normales était une propriété importante, quoique théoriquement superflue si cette estimation a été menée de façon robuste. Cependant, comme cela a été discuté dans le chapitre 1, il existe un cas où, quels que soient les efforts déployés lors de l'étape d'estimation des normales, le vecteur $\mathbf{g}(u, v)$ constitue une donnée aberrante, pour la bonne et simple raison que la normale $\mathbf{n}(u, v)$ n'est pas définie : il s'agit des pixels (u, v) pour lesquels $\mathbf{x}(u, v)$ est situé sur une arête ou une discontinuité de profondeur. Rappelons (cf. paragraphe 1.1.4) qu'un pixel est associé à une région de la surface et non

6. « Any approach based on the Euler-Lagrange equation is only valid for a few special cases ».

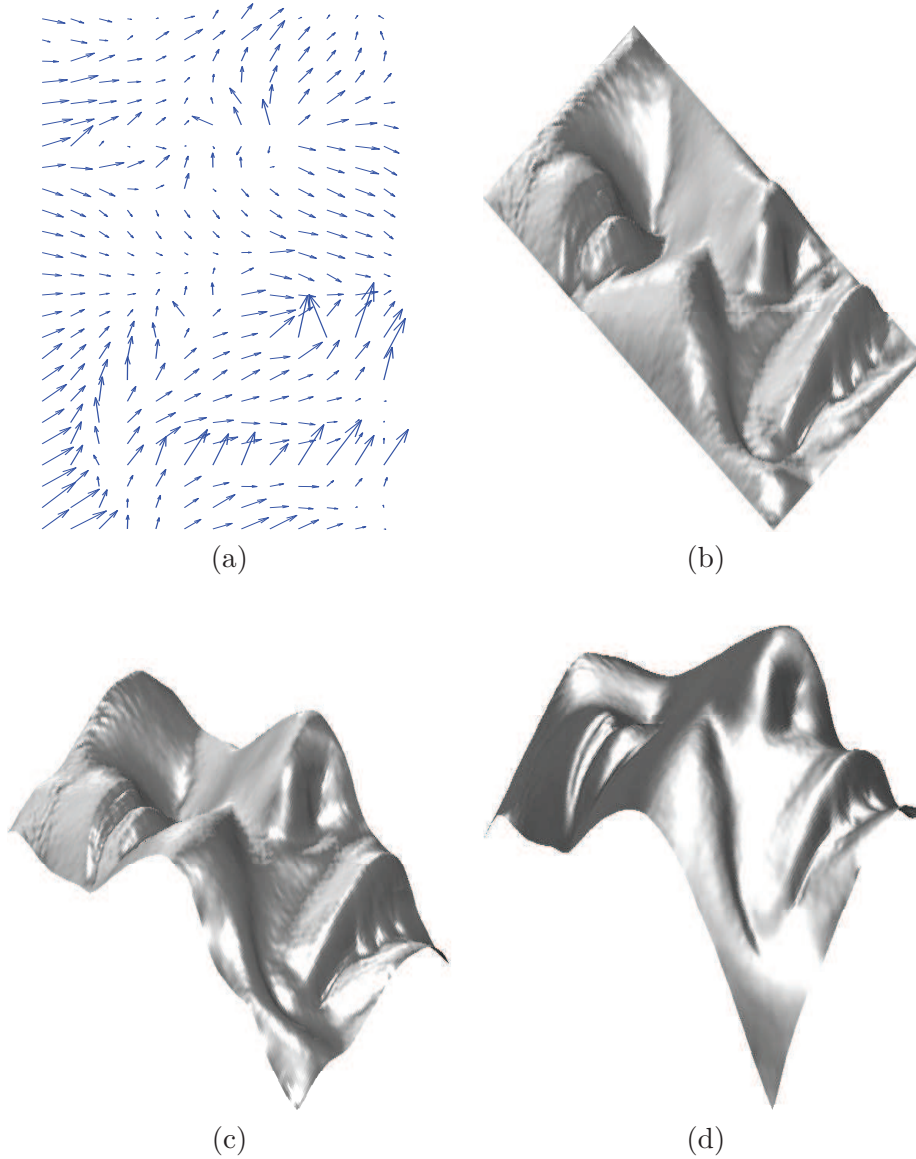


FIGURE 2.2 – Intégration d'un même champ de normales en moindres carrés en utilisant différentes conditions au bord. (a) Champ de gradient estimé par stéréophotométrie. (b) Intégration de ce champ de gradient avec une condition de type Dirichlet [239]. Comme nous ne disposons d'aucun a priori sur la surface, nous avons utilisé ici une condition homogène. (c) Intégration avec une condition périodique [93]. La profondeur est exactement la même sur les bords gauche et droit, c'est-à-dire au niveau de la joue et au niveau du nez, ce qui constitue un biais très important. (d) Intégration avec la condition naturelle au bord [239].

pas à un point : il est par conséquent impossible d'éviter la présence de discontinuités de profondeur, même en augmentant la résolution de l'appareil photographique. Le traitement de ces pixels requiert donc une attention particulière [143] : on observe en effet (cf. figure 2.1-e) que l'erreur est très souvent concentrée au niveau des discontinuités de profondeur.

En l'absence de biais sur l'estimation de $\mathbf{g}$, de tels points se situent généralement là où le terme d'intégrabilité :

$$\mathcal{I}(u, v) = |\partial_v p(u, v) - \partial_u q(u, v)| \quad (2.27)$$

est strictement positif. Ces points peuvent donc être déterminés automatiquement par seuillage de $\mathcal{I}$, et retirés du domaine Ω avant de procéder à l'intégration, pourvu que l'on soit capable de traiter des domaines Ω de forme quelconque.

Pour éviter une telle détection préliminaire, on peut utiliser un estimateur robuste tel que (2.22) : cette approche sera détaillée dans le paragraphe 2.4. En revanche, l'estimateur en moindres carrés (2.21) doit impérativement être exclu, sous peine d'induire un biais important, comme cela est illustré sur la figure 2.3.

2.2.5 Complexité algorithmique

Une propriété importante de toute méthode numérique est sa vitesse d'exécution. Les méthodes de résolution par propagation [18, 95, 123, 274], quoique généralement peu robustes, sont sans conteste les plus rapides, car elles permettent d'estimer le relief en ne « visitant » chaque pixel de Ω qu'une fois. La complexité algorithmique est alors en $O(|\Omega|)$ (complexité linéaire) et ce, quelle que soit la forme du domaine.

Lorsque le domaine de reconstruction Ω est rectangulaire, la discrétisation de l'équation de Poisson (2.25) fournit un système linéaire dont la matrice est de type Toeplitz, ce qui signifie que les éléments d'une même diagonale sont tous égaux. Ce système peut être résolu avec une complexité en $O(|\Omega| \ln |\Omega|)$ au moyen de la transformation de Fourier rapide (*fast Fourier transform*, notée FFT), comme l'ont fait Frankot et Chellappa dans [93] en complément d'une condition au bord périodique. S'inspirant des travaux de Buzbee et al. [41], Simchony et al. ont proposé des extensions de la méthode de Frankot et Chellappa à des conditions au bord de type Dirichlet ou Neumann qui utilisent, respectivement, la *transformation en sinus discrète* (*discrete sine transform*, notée DST) ou la *transformation en cosinus discrète* (*discrete cosine transform*, notée DCT) [239]. Le lecteur intéressé trouvera une présentation très claire de ces différentes techniques, accompagnée de codes Matlab, dans le chapitre 4 de [105].

Pour un domaine Ω de forme rectangulaire, Harker et O'Leary ont proposé dans [109] une résolution matricielle du problème discret associé à (2.21). En notant $\mathbf{Z}$ la matrice formée par les valeurs inconnues de z sur l'ensemble de la grille, et $(\mathbf{P}, \mathbf{Q})$ les matrices contenant les deux coordonnées du champ $\mathbf{g}$, ce problème peut être écrit sous forme matricielle :

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_c}} \|\mathbf{D}_u \mathbf{Z} - \mathbf{P}\|_F^2 + \|\mathbf{Z} \mathbf{D}_v - \mathbf{Q}\|_F^2 \quad (2.28)$$

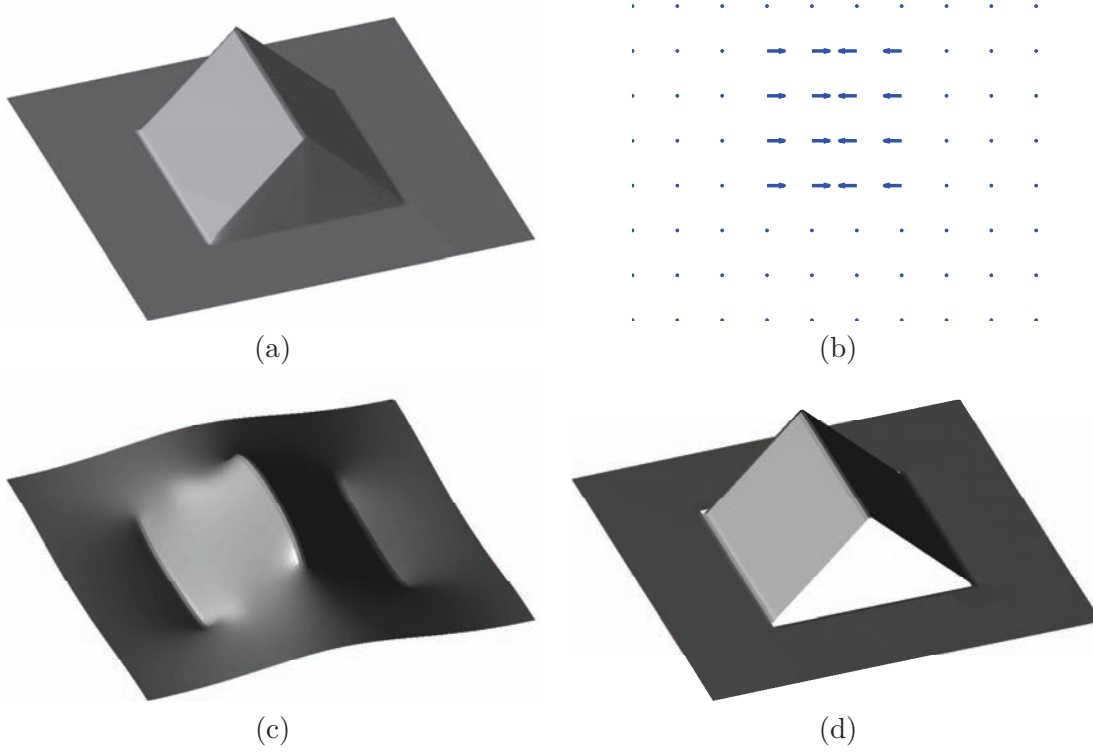


FIGURE 2.3 – Robustesse aux discontinuités de profondeur. (a) Relief en forme de « tente canadienne sur sol plat », pour lequel $\mathbf{g}$ est parfaitement connu, mais qui n'est pas intégrable sur le bord de la tente. (b) Champ de gradient associé au relief (a). (c) La reconstruction en moindres carrés par la méthode de Durou et Courteille [81] est très biaisée, à cause des discontinuités de profondeur. (d) Le biais est corrigé en retirant du domaine de reconstruction les points (u, v) tels que $\mathcal{I}(u, v) > \epsilon$ (le seuil ϵ est fixé arbitrairement à 0,1).

où $\|\cdot\|_F$ désigne la norme de Frobenius, et $\mathbf{D}_u$ et $\mathbf{D}_v$ sont deux matrices permettant de calculer les différences finies, donc telles que :

$$\partial_u z(u, v) \approx (\mathbf{D}_u \mathbf{Z})_{u,v} \quad (2.29)$$

$$\partial_v z(u, v) \approx (\mathbf{Z} \mathbf{D}_v)_{u,v} \quad (2.30)$$

L'intérêt de la formulation matricielle (2.28) est qu'elle permet de définir très facilement des différences finies *cohérentes*, i.e. précises au même ordre sur l'ensemble du domaine de reconstruction, y compris au bord. Ceci en fait une méthode non biaisée, au contraire par exemple de la méthode fondée sur la DCT [239].

L'équation d'optimalité (matricielle) associée au problème (2.28) est une *équation de Sylvester*, qui peut être résolue par l'algorithme de Bartels-Stewart [21] avec une complexité algorithmique en $O(|\Omega|^{3/2})$.

Lorsque Ω n'est pas rectangulaire, la matrice du système linéaire découlant du problème discret associé à (2.21), tout comme celle qui résulte de la discrétisation de l'équation de Poisson (2.25), ne sont pas de type Toeplitz. Ces systèmes linéaires ne peuvent alors être résolus que par des méthodes génériques. La résolution directe, par exemple par factorisation LU, est de complexité en $O(|\Omega|^3)$, ce qui peut être problématique pour des domaines de taille importante, ainsi d'ailleurs que les besoins en mémoire de la factorisation. On peut alors leur préférer les méthodes de résolution itératives, comme les méthodes de Jacobi [81, 126], de Gauss-Seidel ou de *sur-relaxation successive* (*successive over-relaxation*, notée SOR) [128], qui sont beaucoup moins gourmandes en mémoire. Les méthodes de type Jacobi peuvent de plus être implémentées selon une approche *multi-grilles* (*multigrid*) et « portées sur GPU » [146]. Toutefois, les performances obtenues dépendent beaucoup du programmeur, qui doit maîtriser des outils liés à l'architecture matérielle⁷. Nous proposons dans le paragraphe 2.3 de nouvelles méthodes de résolution présentant une complexité en $O(|\Omega| \ln |\Omega|)$, i.e. aussi rapides que les méthodes fondées sur la transformation de Fourier, mais valides pour un domaine de reconstruction Ω de forme quelconque.

Nous n'avons évoqué dans ce paragraphe que les méthodes d'intégration par propagation et celles qui sont fondées sur une minimisation en moindres carrés. Si l'on remplace l'approche en moindres carrés par une approche plus robuste telle que (2.22), cela induit bien sûr une complexité accrue. La condition d'optimalité qui en découle n'est généralement plus linéaire, et doit être résolue par une méthode itérative. Il est alors difficile de prédire la complexité, qui dépend de la précision recherchée, mais ces approches sont nécessairement plus lentes que celles des deux types précédemment décrits.

Notons enfin que, pour les méthodes fondées sur la résolution d'un système linéaire, les complexités que nous avons citées ne concernent que la résolution de ce système. Il faut également tenir compte de sa construction même, ce qui nécessite de parcourir au préalable l'ensemble des pixels de Ω .

2.2.6 Prise en compte d'un a priori

Une propriété très appréciable des méthodes d'intégration est leur capacité à pouvoir prendre en compte une connaissance a priori sur la surface à reconstruire. Il a été remarqué dans [128, 146] que l'utilisation de *points de contrôle* permettait de corriger le biais induit par un mauvais étalonnage des éclairages. Il est également noté dans [110] que l'ajout d'un terme de régularisation dans (2.21) permet d'améliorer la robustesse de la reconstruction. Si l'estimation de $\mathbf{g}$ est légèrement biaisée, il peut être effectivement intéressant de *régulariser* la surface pour limiter les variations de profondeur, ou pour forcer la solution à « ressembler » à un a priori obtenu par exemple par un capteur externe (e.g., une caméra de profondeur de type *Kinect*). Ceci constitue donc un moyen de fusionner les données photométriques avec des données géométriques obtenues par d'autres techniques de reconstruction 3D [10, 107, 193, 251, 281, 290].

7. Au contraire, la méthode de Frankot et Chellappa peut être implémentée en cinq lignes de Matlab !

Il est difficile d'étendre les méthodes d'intégration par propagation, qui reconstruisent la surface de proche en proche, à la prise en compte d'un tel a priori. En revanche, les approches fondées sur l'optimisation sont particulièrement adaptées à de telles extensions, qui peuvent être effectuées en remplaçant le problème (2.22) par le problème régularisé :

$$\min_z \iint_{(u,v) \in \Omega} \Phi(\|\nabla z(u,v) - \mathbf{g}(u,v)\|) + \Phi_r(\|\nabla^k(z - z_0)(u,v)\|) \, du \, dv \quad (2.31)$$

où z_0 représente l'a priori, ∇^k l'opérateur de différenciation à l'ordre k (en général, $k \in \{0, 1, 2\}$) et Φ_r est choisie en fonction de la fiabilité de l'a priori. Dans le cas où la connaissance a priori concerne seulement une partie des pixels de Ω (par exemple, s'il s'agit de points de contrôle), le terme de régularisation ne doit bien sûr être calculé que pour les pixels concernés. Si Φ et Φ_r sont quadratiques, alors l'équation d'Euler-Lagrange associée au problème (2.31) est une EDP linéaire constituant une variante de l'équation de Poisson (2.25). Mais dans le cas général, la condition d'optimalité peut être non linéaire. Il faut alors utiliser une méthode de résolution itérative de type descente de gradient. Ces deux cas seront explorés plus en détail dans le paragraphe 2.4.

En faisant le bilan de ce paragraphe, nous constatons qu'à l'exception de la robustesse aux discontinuités de profondeur, l'intégration en moindres carrés satisfait à l'ensemble des propriétés requises. C'est pourquoi nous consacrons le prochain paragraphe à l'étude de cette approche.

2.3 Intégration en moindres carrés

Dans ce paragraphe, nous détaillons la résolution du problème en moindres carrés (2.21). Nous avons déjà dit qu'il existait principalement deux approches : l'approche variationnelle et l'approche discrète.

Nous avons mentionné dans le paragraphe 2.2.3 que le problème en moindres carrés sur un domaine Ω rectangulaire pouvait être résolu par DCT, avec une complexité en $O(|\Omega| \ln|\Omega|)$ [239], au prix d'un biais sur le bord du domaine Ω . D'autre part, une solution non biaisée peut être obtenue par résolution d'une équation de Sylvester, avec une complexité en $O(|\Omega|^{3/2})$ [109]. Si le domaine n'est pas rectangulaire, la résolution directe par factorisation matricielle est caractérisée par une complexité en $O(|\Omega|^3)$, et des méthodes itératives (Jacobi, Gauss-Seidel ou SOR) peuvent lui être préférées [81, 128, 126]. Il semble donc qu'il n'existe aucune méthode qui soit à la fois rapide, sans biais, et adaptée à un domaine de reconstruction de forme quelconque.

Dans ce paragraphe, nous montrons comment résoudre le problème (2.21) sur un domaine Ω de forme quelconque, avec une complexité en $O(|\Omega| \ln|\Omega|)$.

2.3.1 Discrétisation de la condition d'optimalité du problème (2.21)

La façon la plus classique de résoudre le problème (2.21) consiste à utiliser le calcul des variations. Dans cette approche, la condition d'optimalité est alors l'équation d'Euler-Lagrange, qui est en l'occurrence l'équation de Poisson (2.25), accompagnée de

la condition naturelle au bord (2.26). La résolution numérique peut être effectuée par discrétisation, en utilisant soit des différences finies, soit des éléments finis. Comme Ω est supporté par une grille régulière, nous choisissons les différences finies. L'approche fondée sur les éléments finis sera abordée, de façon détournée, dans le paragraphe 2.3.2.

Discrétisation sur l'intérieur du domaine

La notion de continuité est remplacée, dans le cas discret, par celle de voisinage. Nous étudions ici le cas de la 4-connexité. L'intérieur $\overset{\circ}{\Omega}$ du domaine discret Ω ⁸ est alors défini comme l'ensemble des pixels de Ω dont les quatre plus proches voisins sont dans Ω :

$$\overset{\circ}{\Omega} = \left\{ (u, v) \in \Omega / ((u, v-1), (u-1, v), (u+1, v), (u, v+1)) \in \Omega^4 \right\} \quad (2.32)$$

Il nous faut discrétiser l'équation de Poisson (2.25) sur $\overset{\circ}{\Omega}$. Afin d'éviter tout biais, il convient de discrétiser les opérateurs laplacien et divergence avec le même ordre de précision. On choisit généralement des différences finies centrées d'ordre 2. Le laplacien de la profondeur z et la divergence du champ $\mathbf{g} = [p, q]^\top$ s'écrivent alors (en choisissant comme unité de longueur la distance entre pixels voisins) :

$$\Delta z(u, v) = z(u, v-1) + z(u-1, v) + z(u+1, v) + z(u, v+1) - 4z(u, v) \quad (2.33)$$

$$\nabla \cdot \mathbf{g}(u, v) = \frac{p(u+1, v) - p(u-1, v)}{2} + \frac{q(u, v+1) - q(u, v-1)}{2} \quad (2.34)$$

L'équation de Poisson (2.25) se réécrit :

$$4z(u, v) - z(u, v-1) - z(u-1, v) - z(u+1, v) - z(u, v+1) = \frac{p(u-1, v) - p(u+1, v)}{2} + \frac{q(u, v-1) - q(u, v+1)}{2} \quad (2.35)$$

Condition au bord de type Dirichlet

Si une condition de type Dirichlet est imposée sur le bord du domaine, alors les équations (2.35) obtenues pour $(u, v) \in \overset{\circ}{\Omega}$ forment un système linéaire de $|\overset{\circ}{\Omega}|$ équations à $|\overset{\circ}{\Omega}|$ inconnues $\{z(u, v), (u, v) \in \overset{\circ}{\Omega}\}$, que nous écrivons :

$$\mathbf{L}_D \mathbf{z} = \mathbf{d}_D \quad (2.36)$$

La matrice $\mathbf{L}_D$ est symétrique, à diagonale dominante, mais pas à diagonale *strictement* dominante : chaque terme diagonal est égal à la somme des termes non diagonaux situés sur la même ligne, si les quatre pixels voisins du pixel correspondant à cette ligne sont dans $\overset{\circ}{\Omega}$. En revanche, pour un pixel dont un des quatre voisins au moins appartient à

8. Nous rappelons que dans les problèmes continus traités jusqu'ici, Ω décrivait temporairement un ouvert. Nous revenons dorénavant à la définition discrète de l'ensemble de reconstruction.

$\partial\Omega$, le terme diagonal est strictement dominant. Cela suffit à rendre la matrice définie positive, donc inversible. Pour un domaine rectangulaire, la forme particulière (de type Toeplitz) de la matrice permet une résolution en $O(|\Omega| \ln|\Omega|)$ grâce à la DST [239]. Pour un domaine de forme quelconque, des travaux récents [152] ont montré que l'on pouvait trouver une solution $\mathbf{z}$ « ϵ -proche » de la solution réelle $\hat{\mathbf{z}}$, i.e.

$$\|\mathbf{z} - \hat{\mathbf{z}}\| < \epsilon \|\hat{\mathbf{z}}\| \quad (2.37)$$

avec une complexité en :

$$O(\text{nnz}(\mathbf{L}_D) \ln|\overset{\circ}{\Omega}| \ln(1/\epsilon)) = O(|\Omega| \ln|\Omega|) \quad (2.38)$$

où $\text{nnz}(\mathbf{L}_D)$ désigne le nombre d'éléments non nuls de $\mathbf{L}_D$, qui est en $O(5|\overset{\circ}{\Omega}|)$. La complexité est donc du même ordre que celle de la solution utilisant la DST [239], qui n'est valide que pour un domaine rectangulaire.

Condition naturelle au bord

Le cas de la condition naturelle au bord (2.26) est plus délicat, car sa discrétisation dépend du type de pixel de bord. Il existe a priori $2^4 - 1 = 15$ configurations différentes pour le voisinage d'un pixel de bord (nombre de parties d'un ensemble de cardinal 4, moins le cas où les quatre voisins sont dans Ω). Parmi ces pixels, nous retirons de Ω les pixels pour lesquels seuls deux voisins opposés sont dans Ω (soit les voisins de gauche et de droite, soit les voisins de dessus et de dessous). En effet, pour de tels pixels, il serait impossible de calculer, par différence finie, la composante du gradient de profondeur dans la direction orthogonale au « bord » de Ω . Nous éliminons également les pixels pour lesquels aucun voisin ne se trouve dans Ω . Il reste donc « seulement » douze types de pixels de bord, qui sont illustrés sur la figure 2.4 :

$$\Omega_1 = \{(u, v) \in \Omega / (u, v-1) \in \Omega, (u-1, v) \notin \Omega, (u+1, v) \notin \Omega, (u, v+1) \notin \Omega\} \quad (2.39)$$

$$\Omega_2 = \{(u, v) \in \Omega / (u, v-1) \in \Omega, (u-1, v) \in \Omega, (u+1, v) \in \Omega, (u, v+1) \notin \Omega\} \quad (2.40)$$

$$\Omega_3 = \{(u, v) \in \Omega / (u, v-1) \notin \Omega, (u-1, v) \notin \Omega, (u+1, v) \notin \Omega, (u, v+1) \in \Omega\} \quad (2.41)$$

$$\Omega_4 = \{(u, v) \in \Omega / (u, v-1) \notin \Omega, (u-1, v) \in \Omega, (u+1, v) \in \Omega, (u, v+1) \in \Omega\} \quad (2.42)$$

$$\Omega_5 = \{(u, v) \in \Omega / (u, v-1) \notin \Omega, (u-1, v) \in \Omega, (u+1, v) \notin \Omega, (u, v+1) \notin \Omega\} \quad (2.43)$$

$$\Omega_6 = \{(u, v) \in \Omega / (u, v-1) \in \Omega, (u-1, v) \in \Omega, (u+1, v) \notin \Omega, (u, v+1) \in \Omega\} \quad (2.44)$$

$$\Omega_7 = \{(u, v) \in \Omega / (u, v-1) \notin \Omega, (u-1, v) \notin \Omega, (u+1, v) \in \Omega, (u, v+1) \notin \Omega\} \quad (2.45)$$

$$\Omega_8 = \{(u, v) \in \Omega / (u, v-1) \in \Omega, (u-1, v) \notin \Omega, (u+1, v) \in \Omega, (u, v+1) \in \Omega\} \quad (2.46)$$

$$\Omega_9 = \{(u, v) \in \Omega / (u, v-1) \in \Omega, (u-1, v) \in \Omega, (u+1, v) \notin \Omega, (u, v+1) \notin \Omega\} \quad (2.47)$$

$$\Omega_{10} = \{(u, v) \in \Omega / (u, v-1) \notin \Omega, (u-1, v) \in \Omega, (u+1, v) \notin \Omega, (u, v+1) \in \Omega\} \quad (2.48)$$

$$\Omega_{11} = \{(u, v) \in \Omega / (u, v-1) \in \Omega, (u-1, v) \notin \Omega, (u+1, v) \in \Omega, (u, v+1) \notin \Omega\} \quad (2.49)$$

$$\Omega_{12} = \{(u, v) \in \Omega / (u, v-1) \notin \Omega, (u-1, v) \notin \Omega, (u+1, v) \in \Omega, (u, v+1) \in \Omega\} \quad (2.50)$$

Le bord du domaine peut être défini comme l'union de ces douze ensembles (qui sont disjoints) : $\partial\Omega = \bigcup_{i=1}^{12} \Omega_i$.

En discrétisant l'opérateur gradient par des différences finies d'ordre 1, avant ou arrière suivant le type de bord⁹, les discrétisations de la condition naturelle au bord (2.26) correspondant aux douze types de pixels de bord s'écrivent :

$$z(u, v) - z(u, v - 1) = q(u, v), \quad (u, v) \in \Omega_1 \cup \Omega_2 \quad (2.51)$$

$$z(u, v) - z(u, v + 1) = -q(u, v), \quad (u, v) \in \Omega_3 \cup \Omega_4 \quad (2.52)$$

$$z(u, v) - z(u - 1, v) = p(u, v), \quad (u, v) \in \Omega_5 \cup \Omega_6 \quad (2.53)$$

$$z(u, v) - z(u + 1, v) = -p(u, v), \quad (u, v) \in \Omega_7 \cup \Omega_8 \quad (2.54)$$

$$2z(u, v) - z(u, v - 1) - z(u - 1, v) = p(u, v) + q(u, v), \quad (u, v) \in \Omega_9 \quad (2.55)$$

$$2z(u, v) - z(u - 1, v) - z(u, v + 1) = p(u, v) - q(u, v), \quad (u, v) \in \Omega_{10} \quad (2.56)$$

$$2z(u, v) - z(u, v - 1) - z(u + 1, v) = -p(u, v) + q(u, v), \quad (u, v) \in \Omega_{11} \quad (2.57)$$

$$2z(u, v) - z(u, v + 1) - z(u + 1, v) = -p(u, v) - q(u, v), \quad (u, v) \in \Omega_{12} \quad (2.58)$$

On constate que l'équation est la même pour Ω_1 et Ω_2 , car la normale unitaire sortante est la même, et de même pour Ω_3 et Ω_4 , pour Ω_5 et Ω_6 et pour Ω_7 et Ω_8 (cf. figure 2.4).

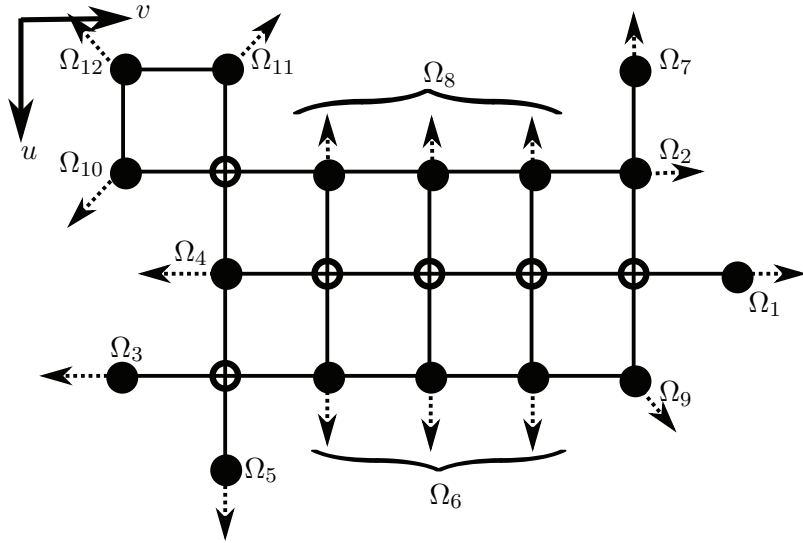


FIGURE 2.4 – Exemple de domaine Ω illustrant les douze types de pixels de bord. Les cercles blancs désignent les pixels de l'intérieur $\overset{\circ}{\Omega}$, les cercles noirs ceux du bord $\partial\Omega$. Chaque ensemble Ω_i , $i = 1 \dots 12$, contient un seul pixel, à l'exception de Ω_6 et Ω_8 qui en contiennent trois chacun. Les flèches représentent les normales sortantes.

9. Pour éviter tout biais, il serait préférable d'utiliser des différences finies d'ordre deux, mais cela serait nettement plus difficile à implémenter, car il existerait encore beaucoup plus de types de pixels de bord à distinguer.

Les équations (2.35) correspondant aux pixels de $\overset{\circ}{\Omega}$, complétées par les équations (2.51) à (2.58) correspondant aux pixels de $\partial\Omega$, forment un système linéaire :

$$\mathbf{L}_N \mathbf{z} = \mathbf{d}_N \quad (2.59)$$

de $|\Omega|$ équations à $|\Omega|$ inconnues (ces inconnues sont les valeurs $\{z(u, v), (u, v) \in \Omega\}$). La matrice $\mathbf{L}_N$ est symétrique, semi-définie positive, de rang $|\Omega| - 1$. Son noyau, qui est de dimension 1 d'après le théorème du rang, est engendré par le vecteur de $\mathbb{R}^{|\Omega|}$ dont toutes les coordonnées sont égales à 1. Cela est lié au fait que toute solution de l'équation initiale (2.16) ne peut être définie qu'à une constante additive près.

Lorsque Ω est rectangulaire, la DCT permet de calculer une des solutions du système (2.59) avec une complexité en $O(|\Omega| \ln |\Omega|)$. Pour un domaine Ω de forme quelconque, comme la matrice $\mathbf{L}_N$ n'est pas inversible, il est impossible d'utiliser les techniques de résolution directes comme par exemple la factorisation LU. Ce type d'approche n'est de toute façon pas très pertinent, car sa complexité est en $O(|\Omega|^3)$. Les méthodes itératives de type Jacobi, Gauss-Seidel ou SOR peuvent être employées, mais comme $\mathbf{L}_N$ n'est pas de rang maximal, leur convergence est extrêmement lente [112].

La matrice $\mathbf{L}_N$ est symétrique à diagonale dominante (pas strictement dominante), donc semi-définie positive. Il s'agit d'une matrice *laplacienne*, i.e. dont la somme des lignes et des colonnes est nulle. Elle contient une description du domaine Ω , vu comme un graphe, sous une forme très facile à interpréter : les éléments diagonaux indiquent le degré de chaque nœud du graphe (4 pour un pixel de $\overset{\circ}{\Omega}$, 1, 2 ou 3 pour un pixel de $\partial\Omega$), tandis que les éléments non diagonaux décrivent la connexité du graphe (0 pour deux nœuds non adjacents, -1 pour deux nœuds adjacents). Il a été montré récemment que de nombreuses EDP pouvaient être résolues très efficacement en suivant une approche par graphe [247, 250]. En particulier, malgré le caractère non inversible d'une matrice laplacienne, un système linéaire dont la matrice est de ce type peut être préconditionné par la méthode décrite dans [152] (préconditionnement *combinatorial multigrid*, noté CMG), avant d'être résolu par la méthode du gradient conjugué ou toute autre méthode de Krylov [227]. Le nombre d'éléments non nuls de $\mathbf{L}_N$ étant égal à :

$$\begin{aligned} \text{nnz}(\mathbf{L}_N) &= |\Omega| + 4 \left| \overset{\circ}{\Omega} \right| + 3 \sum_{i \in \{2,4,6,8\}} |\Omega_i| + 2 \sum_{i \in \{9,10,11,12\}} |\Omega_i| + \sum_{i \in \{1,3,5,7\}} |\Omega_i| \\ &= 5|\Omega| - \sum_{i \in \{2,4,6,8\}} |\Omega_i| - 2 \sum_{i \in \{9,10,11,12\}} |\Omega_i| - 3 \sum_{i \in \{1,3,5,7\}} |\Omega_i| \end{aligned} \quad (2.60)$$

qui est inférieur à $5|\Omega| = O(|\Omega|)$, la résolution du système linéaire (2.59) peut être réalisée en $O(|\Omega| \ln |\Omega|)$ [152]. On peut ainsi atteindre des performances comparables à l'approche utilisant la DCT [239], mais sans imposer à Ω d'être rectangulaire.

2.3.2 Résolution du problème discret associé à (2.21)

Nous avons vu, dans le paragraphe 2.3.1, que la condition naturelle au bord (2.26) était difficile à manipuler. Harker et O'Leary ont également souligné dans [112] que

l'utilisation de différences finies centrées d'ordre 2 à l'intérieur du domaine, et de différences finies avant ou arrière d'ordre 1 sur le bord, induisait un biais. Harker et O'Leary contournent cette difficulté en introduisant des opérateurs de différenciation *matriciels*. Cette approche est cependant limitée à des domaines rectangulaires. Pour un domaine de forme quelconque, on peut écrire la condition d'optimalité associée à chaque inconnue $z(u, v)$ du problème d'optimisation discret associé à (2.21). La discrétisation de ce problème, telle qu'elle est recommandée dans [81, 126], s'écrit :

$$\min_{\{z(u,v), (u,v) \in \Omega\}} \sum_{(u,v) \in \Omega^{u+}} \left[\partial_u^+ z(u, v) - \bar{p}(u, v) \right]^2 + \sum_{(u,v) \in \Omega^{v+}} \left[\partial_v^+ z(u, v) - \bar{q}(u, v) \right]^2 \quad (2.61)$$

où :

$$\bar{p}(u, v) = \frac{p(u+1, v) + p(u, v)}{2} \quad (2.62)$$

$$\bar{q}(u, v) = \frac{q(u, v+1) + q(u, v)}{2} \quad (2.63)$$

$$\partial_u^+ z(u, v) = z(u+1, v) - z(u, v) \quad (2.64)$$

$$\partial_v^+ z(u, v) = z(u, v+1) - z(u, v) \quad (2.65)$$

$$\Omega^{u+} = \{(u, v) \in \Omega / (u+1, v) \in \Omega\} \quad (2.66)$$

$$\Omega^{v+} = \{(u, v) \in \Omega / (u, v+1) \in \Omega\} \quad (2.67)$$

En procédant de cette façon, toutes les différences finies ont une précision d'ordre 1. Le moyennage des valeurs de p et q permet en outre de compenser le choix de différences finies avant, afin d'atteindre une précision d'ordre 2 sur l'ensemble du domaine.

La condition d'optimalité associée à chaque inconnue $z(u, v)$ s'écrit :

$$\begin{aligned} & [\chi(u, v-1) + \chi(u-1, v) + \chi(u+1, v) + \chi(u, v+1)] z(u, v) \\ & - \chi(u, v-1) z(u, v-1) - \chi(u-1, v) z(u-1, v) \\ & - \chi(u+1, v) z(u+1, v) - \chi(u, v+1) z(u, v+1) \\ & = \chi(u-1, v) \bar{p}(u-1, v) - \chi(u+1, v) \bar{p}(u, v) \\ & + \chi(u, v-1) \bar{q}(u, v-1) - \chi(u, v+1) \bar{q}(u, v) \end{aligned} \quad (2.68)$$

où χ est la fonction caractéristique de Ω . Sans surprise, le système :

$$\mathbf{Lz} = \mathbf{d} \quad (2.69)$$

formé par les $|\Omega|$ équations linéaires (2.68) est très similaire au système (2.59) obtenu par discrétisation de l'équation de Poisson (2.25) et de la condition naturelle au bord¹⁰. Comme cela est discuté dans le paragraphe 2.3.1, nous préconisons le préconditionnement CMG [152], puis l'algorithme du gradient conjugué.

10. En réalité, l'approche discrète proposée peut s'interpréter comme la résolution numérique du problème continu par éléments finis : il est donc logique de retrouver une condition d'optimalité similaire à la discrétisation par différences finies.

En revanche, la discrétisation diffère légèrement pour les pixels du bord, selon que l'on choisisse de discrétiser l'objectif ou sa condition d'optimalité continue. En effet, les discrétisations (2.51) à (2.58) de la condition naturelle au bord impliquent un nombre de points inférieur à celui qu'implique l'équation (2.68) pour un pixel du bord. En impliquant davantage de points dans le calcul de ces termes de « laplacien » et de « divergence », le biais au bord est limité. Par exemple, si $(u, v) \in \Omega_4$, la condition au bord (2.52) (issue de la discrétisation de la condition d'optimalité continue) s'écrit :

$$z(u, v) - z(u, v + 1) = -q(u, v) \quad (2.70)$$

Un seul des trois voisins de (u, v) situés dans Ω est donc impliqué, tandis que les trois le sont avec l'écriture (2.68), qui devient :

$$3z(u, v) - z(u - 1, v) - z(u + 1, v) - z(u, v + 1) = \frac{p(u - 1, v) - p(u + 1, v)}{-\frac{q(u, v + 1) + q(u, v)}{2}} \quad (2.71)$$

L'approche continue et l'approche discrète présentent donc exactement la même complexité, mais il est plus facile de limiter le biais avec l'approche discrète. C'est donc elle que nous préférons utiliser dans la suite de ce chapitre. Si l'on se remémore les propriétés que l'on attend d'une « bonne » méthode d'intégration (cf. paragraphe 2.2), celle que nous préconisons (résolution du système linéaire (2.69) associé au problème (2.61), par gradient conjugué préconditionné par CMG [152]) est robuste au bruit (minimisation d'un critère en moindres carrés), indépendante de la topologie du domaine (qui peut être de forme quelconque) et de faible complexité algorithmique (en $O(|\Omega| \ln |\Omega|)$). Ceci en fait la « meilleure » méthode d'intégration en l'absence de discontinuité de profondeur. Intéressons-nous maintenant aux extensions possibles de cette méthode permettant la prise en compte d'une connaissance a priori.

2.3.3 Extensions de l'intégration en moindres carrés

Conformément à la discussion menée dans le paragraphe 2.2.6, nous montrons comment prendre en compte une éventuelle connaissance a priori sur le relief à reconstruire.

Points de contrôle

Si l'on souhaite imposer la valeur de la profondeur en un certain nombre de *points de contrôle*, le problème à résoudre reste inchangé, mais cette connaissance a priori doit être considérée comme une contrainte « dure » [128] :

$$\begin{cases} \min_{\{z(u,v), (u,v) \in \Omega\}} \sum_{(u,v) \in \Omega^{u+}} \sum_{(u,v) \in \Omega^{v+}} \left[\partial_u^+ z(u, v) - \bar{p}(u, v) \right]^2 + \sum_{(u,v) \in \Omega^{u+}} \sum_{(u,v) \in \Omega^{v+}} \left[\partial_v^+ z(u, v) - \bar{q}(u, v) \right]^2 \\ \text{s.c. } z = z_c \text{ sur } \Omega_c \end{cases} \quad (2.72)$$

où l'ensemble Ω_c des points de contrôle est inclus dans Ω .

Soit $\mathbf{z}_c \in \mathbb{R}^{|\Omega_c|}$ le vecteur ayant pour composantes les profondeurs des points de contrôle, $\mathbf{z}' \in \mathbb{R}^{|\Omega| - |\Omega_c|}$ le vecteur construit en retirant de $\mathbf{z}$ ces valeurs, $\mathbf{L}_c \in \mathbb{R}^{|\Omega| \times |\Omega_c|}$ la sous-matrice formée par les colonnes de $\mathbf{L}$ correspondant aux points de contrôle, et $\mathbf{L}' \in \mathbb{R}^{|\Omega| \times (|\Omega| - |\Omega_c|)}$ la sous-matrice de $\mathbf{L}$ complémentaire de $\mathbf{L}_c$. La condition d'optimalité associée au problème (2.72) est le système linéaire suivant :

$$\mathbf{L}'\mathbf{z}' = \mathbf{d} - \mathbf{L}_c\mathbf{z}_c \quad (2.73)$$

qui comporte $|\Omega|$ équations à $|\Omega| - |\Omega_c|$ inconnues. Ce système est de rang maximal, donc surcontraint, dès que $|\Omega_c| > 0$. Sa solution en moindres carrés est la solution (exacte) de l'équation normale :

$$\mathbf{L}'^\top \mathbf{L}'\mathbf{z}' = \mathbf{L}'^\top (\mathbf{d} - \mathbf{L}_c\mathbf{z}_c) \quad (2.74)$$

qui est un système de Cramer de matrice $\mathbf{L}'^\top \mathbf{L}'$. La largeur de bande de $\mathbf{L}$ étant de l'ordre de $\sqrt{|\Omega|}$, le calcul de $\mathbf{L}'^\top \mathbf{L}'$ est quasi-instantané. La matrice $\mathbf{L}'^\top \mathbf{L}'$ est elle-même creuse et de faible largeur de bande. Cette matrice est symétrique, *définie positive*, alors que la matrice $\mathbf{L}$ est symétrique, *semi-définie positive*. Cette différence provient de ce que le problème initial (2.16) ne peut être résolu qu'à une constante additive près, et que l'existence d'au moins un point de contrôle (i.e., $|\Omega_c| > 0$) suffit à fixer la valeur de cette constante, donc à rendre le problème bien posé.

La matrice $\mathbf{L}'^\top \mathbf{L}'$ étant à diagonale strictement dominante, l'utilisation de méthodes directes devient possible, à la place des méthodes itératives comme le gradient conjugué. En particulier, la factorisation de Cholesky de $\mathbf{L}'^\top \mathbf{L}'$ peut être calculée avec une complexité en $O((|\Omega| - |\Omega_c|)^2)$. Bien que ce coût soit supérieur à celui de l'approche itérative par gradient conjugué préconditionné, l'approche directe est aussi rapide en pratique, et plus facile à implémenter. En effet, des algorithmes de factorisation de Cholesky optimisés comme CHOLMOD¹¹ peuvent être employés. De tels algorithmes optimisent notamment les produits matriciels de type « matrice-matrice » (BLAS de niveau 3).

Régularisation à l'ordre 0

On dispose parfois d'un autre type d'a priori sur la surface à reconstruire. Par exemple, si l'on couple la stéréophotométrie à un capteur de profondeur, la profondeur est connue de façon approchée en chaque point de Ω ¹² : il a d'ailleurs été montré récemment qu'un tel a priori rendait possible la reconstruction 3D par SFS sous un éclairage « naturel » [193].

Pour prendre en compte cette connaissance a priori, il suffit d'introduire une réguli-

11. Il s'agit de l'algorithme invoqué par Matlab à l'exécution de la commande `A \ b` [58], pour une matrice $\mathbf{A}$ symétrique, définie positive.

12. En réalité, la carte de profondeur issue de ce type de capteurs comporte souvent des trous, mais la profondeur peut bien sûr être interpolée dans ces régions.

sation quadratique à l'ordre 0 de la profondeur, comme nous l'avons proposé dans [211] :

$$\begin{aligned} \min_{\{z(u,v), (u,v) \in \Omega\}} & \sum_{(u,v) \in \Omega^{u+}} \sum_{(u,v) \in \Omega^{v+}} \left[\partial_u^+ z(u,v) - \bar{p}(u,v) \right]^2 + \sum_{(u,v) \in \Omega^{v+}} \left[\partial_v^+ z(u,v) - \bar{q}(u,v) \right]^2 \\ & + \lambda \sum_{(u,v) \in \Omega} [z(u,v) - z_0(u,v)]^2 \end{aligned} \quad (2.75)$$

où le choix de $\lambda > 0$ dépend de la confiance que l'on porte à l'a priori z_0 . Si aucun a priori n'est disponible, on peut aussi choisir $z_0 \equiv 0$ et $\lambda \rightarrow 0^+$, de façon à limiter l'influence de cet « a priori ». En fait, faire tendre λ vers 0 revient, comme précédemment, à rendre le problème bien posé en fixant la constante additive de la solution, de telle sorte que $E[z] = E[z_0]$.

L'équation d'optimalité associée au problème (2.75) s'écrit :

$$(\mathbf{L} + \lambda \mathbf{I}) \mathbf{z} = \mathbf{d} + \lambda \mathbf{z}_0 \quad (2.76)$$

où la matrice $\mathbf{L}$ et le vecteur $\mathbf{d}$ sont définis en (2.69), $\mathbf{I}$ désigne la matrice identité d'ordre $|\Omega|$ et $\mathbf{z}_0$ le vecteur construit par concaténation des valeurs de z_0 .

Le système (2.76) est à diagonale *strictement dominante* dès lors que $\lambda > 0$. Comme pour l'intégration en présence de points de contrôle, cette propriété permet d'utiliser une méthode de résolution directe, e.g. la factorisation de Cholesky CHOLMOD [58], pour une complexité théorique en $O(|\Omega|^2)$, mais dont le coût est très faible en pratique grâce aux accélérations BLAS de niveau 3. La méthode du gradient conjugué préconditionné [152], de complexité en $O(|\Omega| \ln |\Omega|)$, reste bien entendu utilisable. L'approche directe par factorisation est légèrement plus rapide, mais nécessite davantage de mémoire. Le choix de l'une ou l'autre de ces méthodes dépend des ressources en mémoire dont on dispose.

Il est donc très facile d'adapter l'approche discrète décrite dans le paragraphe 2.3.2 pour prendre en compte d'éventuels points de contrôle ou un a priori sur la profondeur, sans que cela augmente le coût algorithmique. Si l'on excepte la robustesse aux discontinuités de profondeur, toutes les « bonnes propriétés » décrites dans le paragraphe 2.2 sont donc maintenant garanties.

2.3.4 Évaluation quantitative

La table 2.1 compile les résultats obtenus par différentes méthodes d'intégration en moindres carrés, ainsi que par la méthode du *fast marching* [95] (notée FM), dont nous avons déjà dit qu'elle était biaisée, mais qui peut servir à initialiser efficacement les méthodes itératives. Nous ne présentons dans cette table que les méthodes d'intégration en moindres carrés jugées les plus représentatives, à savoir :

- l'approche directe CHOLMOD [58] appliquée au problème régularisé à l'ordre 0 (cf. paragraphe 2.3.3), avec $z_0 \equiv 0$ et $\lambda = 10^{-9}$;
- le gradient conjugué préconditionné par CMG [152] (notée PCG-CMG) ;
- l'approche SOR, qui est représentative des schémas introduits dans [81, 128].

Il ressort de cette étude comparative que la méthode directe est la plus rapide, quoique la méthode itérative préconditionnée offre des performances quasi-équivalentes.

Les méthodes itératives présentent cependant l'avantage d'utiliser beaucoup moins d'espace mémoire que les méthodes directes. De plus, le préconditionnement ne concerne que la matrice du système à résoudre : si l'intégration doit être réalisée un grand nombre de fois sur le même domaine Ω , mais pour différentes estimations du gradient, il est préférable d'utiliser une méthode itérative. Comme nous le verrons, cela est important dans le cadre de l'intégration robuste aux discontinuités de profondeur. Nous recommandons finalement de plutôt utiliser une méthode itérative, à savoir le gradient conjugué, associée au préconditionnement CMG [152].

	$ \Omega \sim 10^4$		$ \Omega \sim 10^6$	
	RMSE	CPU	RMSE	CPU
Plan $z = 0$	91,05	-	365,07	-
DCT [239]	0,756	0,01	2,27	0,12
FM [95]	4,935	< 0,01	5,00	0,15
SOR [128]	1,78 *	155,2 *	171,0 *	2542,1 *
FM+SOR	0,345 *	141,8 *	2,60 *	2559,2 *
PCG-CMG	0,370	0,32	0,360	3,82
FM+PCG-CMG	0,370	0,18	0,360	3,04
Cholesky	0,370	0,09	0,371	2,16

TABLE 2.1 – Comparaison de la précision et de la durée d'exécution de plusieurs méthodes d'intégration en moindres carrés sur un domaine Ω de forme arbitraire (les mesures d'erreur RMSE sont exprimées en pixels, les temps CPU en secondes). Ces résultats sont obtenus pour la fonction **peaks** de Matlab, dont le gradient est connu sous forme analytique. Nous utilisons comme domaine de reconstruction le cercle inscrit dans un domaine carré de taille 256×256 ($|\Omega| \sim 10^4$) ou 1024×1024 ($|\Omega| \sim 10^6$). La solution par DCT est très rapide, mais biaisée si le domaine n'est pas rectangulaire (le champ $\mathbf{g}$ a été complété par des 0 en dehors de Ω , ce qui le rend non intégrable). La solution par *fast marching* (FM) est également très rapide, mais ses résultats sont peu fiables à cause de la propagation des erreurs inhérente à la méthode. Cependant, elle peut être utilisée comme initialisation (tout comme la solution par DCT) pour les méthodes en moindres carrés itératives (SOR, ou gradient conjugué préconditionné par CMG [152]), plutôt que l'initialisation arbitraire $z \equiv 0$. L'approche directe, par régularisation à l'ordre 0 et factorisation de Cholesky, est plus rapide que les méthodes itératives (mais plus gourmande en mémoire). Les méthodes itératives de type SOR (et a fortiori, de type Gauss-Seidel ou Jacobi [81, 128]) convergent très lentement, même avec une initialisation adéquate (l'astérisque * indique que le critère d'arrêt, en l'occurrence un écart relatif en norme 2 inférieur à 10^{-6} entre deux itérations, n'a pas été atteint en 100000 itérations).

2.4 Nouvelles méthodes d'intégration robustes

L'intégration en moindres carrés présente l'avantage de se traduire par des méthodes numériques très efficaces, grâce à la linéarité des systèmes d'équations qui en découlent. Toutefois, cette approche n'est pas robuste aux données aberrantes, qui ne peuvent malheureusement pas être évitées dès lors que la surface à reconstruire n'est régulière que *par morceaux*. Dans ce paragraphe, nous proposons plusieurs techniques d'intégration robustes spécialement conçues pour tenir compte de telles données.

2.4.1 Moindres carrés pondérés

Une première approche consiste à prendre en compte l'hétéroscédasticité de l'erreur d'estimation du champ de gradient $\mathbf{g}$, au moyen d'une pondération appropriée. L'approche en moindres carrés (2.21) peut être remplacée par une minimisation en moindres carrés pondérés :

$$\min_z \iint_{(u,v) \in \Omega} w(u,v) \|\nabla z(u,v) - \mathbf{g}(u,v)\|^2 du dv \quad (2.77)$$

où le poids $w(u,v)$, déterminé a priori, est censé refléter la confiance que l'on porte à l'estimation de $\mathbf{g}(u,v)$.

Cette approche a été proposée par Horovitz dans le cadre de la prise en compte de points de contrôle : les poids sont alors calculés en fonction de la distance euclidienne au point de contrôle le plus proche. Saracchini et al. [229, 228], ainsi que Kimmel et al. [146], utilisent également cette approche dans le cadre plus général de l'estimation robuste, mais ils supposent que la carte de poids est fournie, c'est-à-dire qu'elle constitue une donnée supplémentaire. Nous avons proposé dans [210, 211] deux façons de définir ces poids.

Pondération par l'intégrabilité

Une première possibilité pour résoudre le problème en moindres carrés pondérés (2.77), lorsqu'on ne dispose que de la donnée du champ estimé $\mathbf{g}$, est de choisir une pondération qui soit une fonction décroissante de l'intégrabilité $\mathcal{I}$, telle qu'elle a été définie en (2.27). Pour justifier le choix de cette pondération, écrivons la condition d'optimalité associée à (2.77).

En supposant $w > 0$, et en remarquant que $\frac{\nabla w}{w} = \nabla(\ln w)$, la condition d'optimalité associée au problème variationnel (2.77) est une équation de Poisson modifiée :

$$\Delta z(u,v) + \nabla(\ln w(u,v)) \cdot (\nabla z(u,v) - \mathbf{g}(u,v)) = \nabla \cdot \mathbf{g}(u,v) \quad (2.78)$$

qui peut être interprétée comme une *équation de diffusion anisotrope* [202].

Nous proposons d'utiliser la pondération suivante :

$$w(u,v) = \exp\left(-\gamma \mathcal{I}(u,v)^2\right) \quad (2.79)$$

où γ est un paramètre permettant de contrôler la pondération. Si $\gamma = 0$, $w(u, v) \equiv 1$ et on retrouve le problème en moindres carrés (2.21) (pondération uniforme). Ceci est également le cas si la surface est parfaitement régulière, i.e. si $\mathcal{I} \equiv 0$. En revanche, pour les points (u, v) situés sur une discontinuité de profondeur, la valeur $\mathcal{I}(u, v)$ de l'intégrabilité peut devenir très élevée, et donc $w(u, v) < 1$ pour peu que $\gamma > 0$. Une importance moindre est donc accordée à ces points dans le problème (2.77).

Pondération par le niveau de gris

Dans le contexte de la stéréophotométrie, on dispose de données bien plus riches que la seule estimation de $\mathbf{g}$. En particulier, les données aberrantes peuvent être détectées directement sur les images. Nous proposons une autre pondération, que nous avons déjà introduite dans [77], qui compare le vecteur $\mathbf{i}(u, v) = [I^1(u, v), \dots, I^m(u, v)]^\top$ des niveaux de gris du pixel (u, v) aux vecteurs $\mathbf{i}(\mathcal{V}_i(u, v)) = [I^1(\mathcal{V}_i(u, v)), \dots, I^m(\mathcal{V}_i(u, v))]^\top$, $i \in [1, 8]$, des niveaux de gris de ses huit plus proches voisins. D'après le modèle lambertien linéaire (1.18), $I^i(u, v) = \rho(u, v) \mathbf{s}^i \cdot \mathbf{n}(u, v)$. Il s'en suit que le vecteur $\mathbf{i}(u, v) / \|\mathbf{i}(u, v)\|$ est indépendant de l'albédo : si la surface est lisse autour du pixel (u, v) , chaque composante de ce vecteur normalisé est sensiblement égale à la composante correspondante dans le vecteur normalisé $\mathbf{i}(\mathcal{V}_i(u, v)) / \|\mathbf{i}(\mathcal{V}_i(u, v))\|$. Un changement brutal dans l'une des m composantes de ce vecteur ne peut donc s'expliquer que par une variation brutale de la normale entre (u, v) et $\mathcal{V}_i(u, v)$. Évaluer l'écart maximal entre les composantes de ces deux vecteurs, pour chacun des huit voisins, permet donc de détecter les potentielles discontinuités. Nous utilisons donc finalement comme pondération :

$$w(u, v) = \exp \left\{ -\gamma \mathcal{D}(u, v)^2 \right\} \quad (2.80)$$

où :

$$\mathcal{D}(u, v) = \max \left\{ \left\| \frac{\mathbf{i}(u, v)}{\|\mathbf{i}(u, v)\|} - \frac{\mathbf{i}(\mathcal{V}_i(u, v))}{\|\mathbf{i}(\mathcal{V}_i(u, v))\|} \right\|_\infty, i \in [1, 8] \right\} \quad (2.81)$$

La figure 2.5 illustre les pondérations (2.79) et (2.80) sur un exemple. Ces deux types de pondération semblent effectivement affecter un poids faible aux pixels situés près des discontinuités de profondeur. On note également que la pondération (2.79), qui est sensible aux erreurs d'estimation de $\mathbf{g}$, est plus « bruitée » que la pondération (2.80), qui est directement calculée à partir des images.

Résolution numérique

Comme pour l'intégration en moindres carrés, on peut utiliser deux approches pour résoudre le problème d'optimisation (2.77). La première consiste à discrétiser l'équation d'Euler-Lagrange (2.78). Rappelons que cette équation doit être accompagnée d'une condition au bord bien choisie, et qu'il existe douze types de pixels de bord à traiter avec minutie. Plutôt que d'utiliser cette approche, nous préférons résoudre le problème

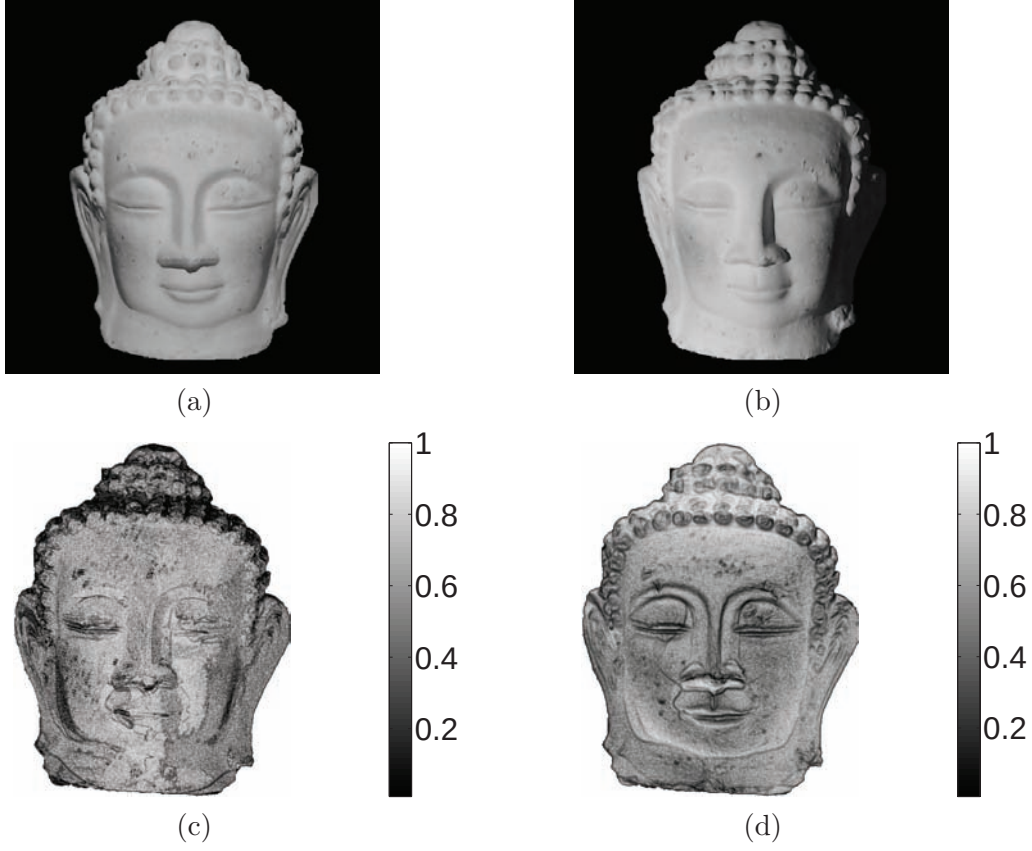


FIGURE 2.5 – Comparaison de deux pondérations permettant d'intégrer (en moindres carrés pondérés) des normales estimées par stéréophotométrie. (a)-(b) Deux photographies (parmi dix) de taille 1062×836 d'une tête de Buddha en plâtre. (c) Pondération (2.79). (d) Pondération (2.80). La valeur du paramètre γ a été choisie, dans chaque cas, de manière à obtenir le résultat visuellement jugé le meilleur. Si ces deux pondérations semblent effectivement limiter le poids des pixels situés près des discontinuités, la pondération (d) semble moins « bruitée » que la pondération (c), ce qui peut s'expliquer par le fait qu'elle est calculée à partir des images mêmes, et non à partir de l'estimation de $\mathbf{g}$ par stéréophotométrie, qui est potentiellement biaisée.

discret suivant :

$$\min_{\{z(u,v), (u,v) \in \Omega\}} \sum_{(u,v) \in \Omega^{u+}} w(u,v) \left[\partial_u^+ z(u,v) - \bar{p}(u,v) \right]^2 + \sum_{(u,v) \in \Omega^{v+}} w(u,v) \left[\partial_v^+ z(u,v) - \bar{q}(u,v) \right]^2 \quad (2.82)$$

La condition d'optimalité en un pixel (u, v) quelconque de Ω s'écrit :

$$\begin{aligned}
 & [\chi(u, v-1)w(u, v-1) + \chi(u-1, v)w(u-1, v) + \chi(u+1, v)w(u, v) + \chi(u, v+1)w(u, v)]z(u, v) \\
 & - \chi(u, v-1)w(u, v-1)z(u, v-1) - \chi(u-1, v)w(u-1, v)z(u-1, v) \\
 & - \chi(u+1, v)w(u, v)z(u+1, v) - \chi(u, v+1)w(u, v)z(u, v+1) \\
 & = \chi(u-1, v)w(u-1, v)\bar{p}(u-1, v) - \chi(u+1, v)w(u, v)\bar{p}(u, v) \\
 & + \chi(u, v-1)w(u, v-1)\bar{q}(u, v-1) - \chi(u, v+1)w(u, v)\bar{q}(u, v)
 \end{aligned} \tag{2.83}$$

Le système linéaire formé par les $|\Omega|$ équations (2.83) comporte $|\Omega|$ inconnues. Sa matrice est une matrice symétrique, semi-définie positive, mais dont les éléments non nuls ne présentent plus de régularité particulière dès lors que les poids ne sont pas uniformes (si $w \equiv 1$, on retrouve exactement les équations (2.68)). Les techniques discutées dans le paragraphe 2.3 peuvent cependant s'appliquer : nous préconisons soit la résolution directe par factorisation de Cholesky du problème régularisé à l'ordre 0 (cf. paragraphe 2.3.3), soit la résolution itérative par gradient conjugué, préconditionné par la méthode CMG [152].

Évaluation qualitative

Les figures 2.6, 2.7 et 2.8 illustrent les deux pondérations proposées (par l'intégrabilité et par le niveau de gris). Les reconstructions 3D sont obtenues par intégration du champ de gradient estimé sur un domaine non rectangulaire, complété par des valeurs nulles. En procédant de la sorte, le champ est fortement non intégrable sur le bord du domaine, ce qui provoque habituellement des phénomènes de Gibbs. En pondérant les moindres carrés, les phénomènes de Gibbs sont nettement atténués.

L'approche en moindres carrés pondérés repose sur une détection *a priori* des discontinuités. Cependant, il n'est pas toujours facile d'effectuer une telle détection. Il semble donc plus raisonnable de déterminer les poids automatiquement, en fonction de la norme du résidu $\|\nabla z - \mathbf{g}\|$ [6, 80], sachant que la condition d'optimalité ne sera plus linéaire et qu'une résolution itérative sera requise. Nous verrons, dans le prochain paragraphe, comment utiliser cette approche pour l'intégration robuste, au moyen d'une fonctionnelle inspirée du modèle L^2 -TV bien connu dans le domaine du traitement d'images [226]. Plus généralement, tout intégrateur robuste de la forme (2.22) semble pertinent. Afin de pouvoir utiliser indifféremment l'approche directe par factorisation de Cholesky ou la méthode du gradient conjugué préconditionné, nous nous attachons maintenant à résoudre le problème très général suivant :

$$\min_z \iint_{(u,v) \in \Omega} \Phi(\|\nabla z(u, v) - \mathbf{g}(u, v)\|) + \lambda [z(u, v) - z_0(u, v)]^2 \, du \, dv \tag{2.84}$$

pour différents choix de la fonction Φ .

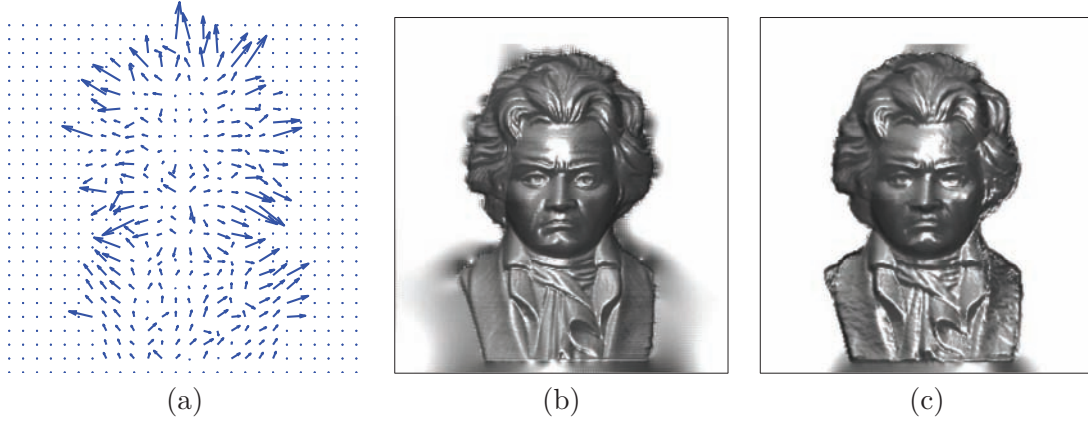


FIGURE 2.6 – Effet des discontinuités de profondeur sur la reconstruction 3D obtenue par intégration du champ de gradient estimé. (a) Champ de gradient $\mathbf{g}$ estimé à partir des $m = 3$ images du buste de Beethoven (cf. figure 1.2), complété par des valeurs nulles en dehors du buste. (b) Rééclairage frontal de la surface estimée en intégrant le champ $\mathbf{g}$ en moindres carrés. Des phénomènes de Gibbs sont visibles sur le bord du buste. (c) Rééclairage frontal de la surface estimée en moindres carrés pondérés, en utilisant la pondération (2.79), qui dépend de l'intégrabilité : ces phénomènes sont atténués.

2.4.2 Modèle L^2 -TV

Au vu de la discussion menée dans le paragraphe 1.3.4, nous pouvons prédire que le choix $\Phi(s) = |s|$ dans (2.84) constitue un compromis très acceptable entre convexité du problème et robustesse. Il a de plus été établi dans [76] et dans [80] (au moyen d'une approximation différentiable de la norme 1) que l'intégration en norme 1 était nettement plus robuste aux discontinuités de profondeur que celle en moindres carrés. Nous pouvons également justifier cette observation par la remarque suivante : il est bien connu en traitement d'images que la mesure de la *variation totale isotrope* :

$$\text{TV}(z) = \iint_{\Omega} \|\nabla z(u, v)\| \, du \, dv \quad (2.85)$$

possède des propriétés intéressantes relatives à la préservation des arêtes, et qu'elle tend à favoriser les solutions lisses par morceaux. En considérant les discontinuités de profondeur comme le pendant en 3D des contours d'une image en 2D, il semble raisonnable de chercher à obtenir un résidu $\nabla z - \bar{\mathbf{g}}$ qui soit lui-même lisse par morceaux. Cette remarque nous invite à adapter le modèle de Rudin, Osher et Fatemi [226] au problème de l'intégration. En choisissant $\Phi(s) = |s|$, le problème (2.84) devient :

$$\min_z \iint_{(u,v) \in \Omega} \|\nabla z(u, v) - \bar{\mathbf{g}}(u, v)\| + \lambda [z(u, v) - z_0(u, v)]^2 \, du \, dv \quad (2.86)$$

Nous présentons dans le prochain paragraphe un schéma numérique efficace pour résoudre (2.86).

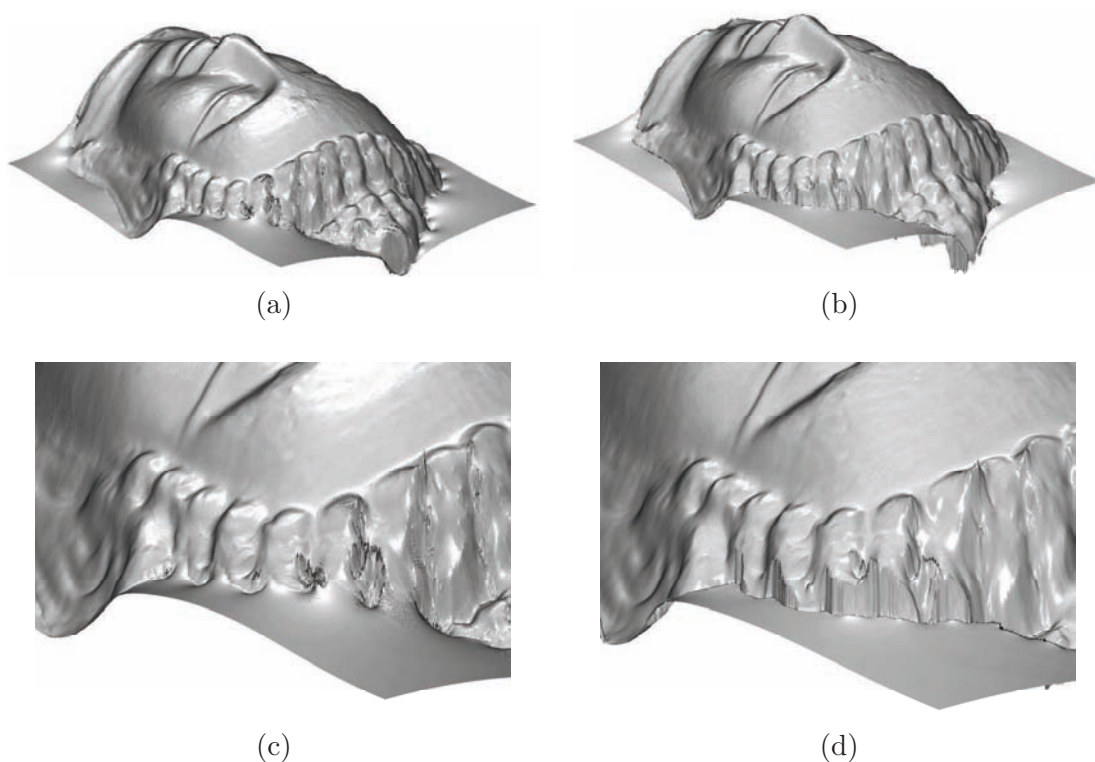


FIGURE 2.7 – Intégration en moindres carrés pondérés à partir des données de la figure 2.5, sur un domaine rectangulaire (le champ $\mathbf{g}$ est complété par des valeurs nulles en dehors du domaine de reconstruction). Reconstructions 3D (a) en moindres carrés [109] et (b) en moindres carrés pondérés, en utilisant la pondération (2.80), qui dépend des niveaux de gris. Les deux reliefs sont similaires, sauf à proximité des discontinuités, ce qui est illustré par les gros plans (c) et (d). Voir également la figure 2.8.

Résolution par ADMM

Afin de résoudre numériquement le problème (2.86), nous choisissons une méthode de type « lagrangien augmenté ». Nous discrétisons pour cela (2.86) par le problème sous contrainte suivant :

$$\begin{cases} \min_{z, \mathbf{r}} \sum_{(u,v) \in \Omega} \|\mathbf{r}(u,v)\| + \lambda [z(u,v) - z_0(u,v)]^2 \\ \text{s.c.} \quad \mathbf{r} = \nabla z - \bar{\mathbf{g}} \end{cases} \quad (2.87)$$

où désormais, nous supposons pour simplifier que $\Omega = \Omega_u^+ \cap \Omega_v^+$, nous remplaçons le champ $\mathbf{g}$ par le moyennage suivant :

$$\bar{\mathbf{g}}(u,v) = \frac{1}{2} \begin{bmatrix} \bar{p}(u,v) \\ \bar{q}(u,v) \end{bmatrix} \quad (2.88)$$

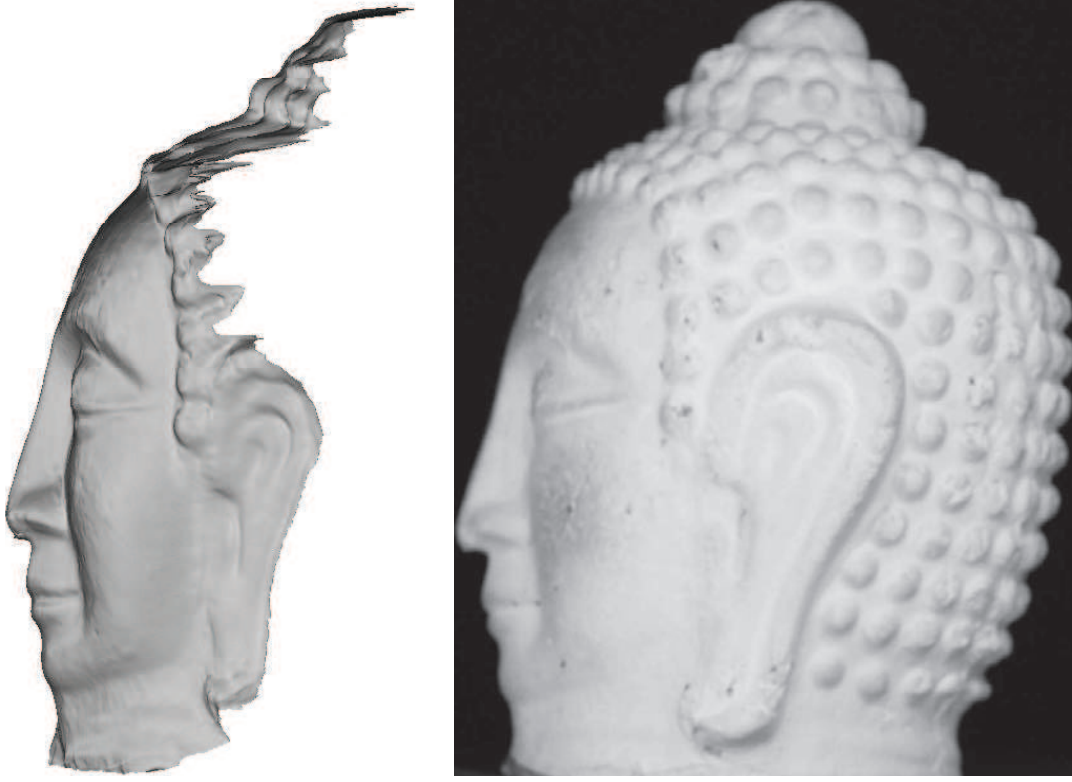


FIGURE 2.8 – Vue latérale de la reconstruction 3D en moindres carrés pondérés (pondération (2.80)), calculée sur un domaine Ω non rectangulaire (i.e., sans complétion de $\mathbf{g}$ par des valeurs nulles), et relief réel vu sous un angle similaire.

où $\bar{p}$ et $\bar{q}$ sont définis, respectivement, en (2.62) et (2.63), et nous notons abusivement :

$$\nabla z(u, v) = \begin{bmatrix} \partial_u^+ z(u, v) \\ \partial_v^+ z(u, v) \end{bmatrix} \quad (2.89)$$

En introduisant une variable auxiliaire $\mathbf{b}$ ¹³, le problème (2.87) peut être résolu itérativement par un schéma à *direction de descente alternée* (*alternating direction method*

13. Cette variable auxiliaire est fortement liée à la *distance de Bregman*. Les approches fondées sur le lagrangien augmenté, qui datent des années 1970 [225], sont en effet équivalentes aux méthodes plus récentes de type itérations de Bregman [280]. Ces deux méthodes sont, en outre, des cas particuliers d'algorithmes proximaux. De la même façon, les schémas ADMM [34] correspondent aux itérations de type split-Bregman [104]. La formulation en termes d'itérations de Bregman semble cependant offrir certains avantages du point de vue numérique [279]. Les liens existant entre les approches fondées sur le lagrangien augmenté et celles fondées sur les itérations de Bregman sont discutés dans [86].

of multipliers, noté ADMM) [34] :

$$\mathbf{r}^{k+1}(u, v) = \operatorname{argmin}_{\mathbf{r} \in \mathbb{R}^2} \left\| \mathbf{r} - \left(\nabla z^k(u, v) - \bar{\mathbf{g}}(u, v) - \mathbf{b}^k(u, v) \right) \right\|^2 \quad (2.90)$$

$$z^{k+1} = \operatorname{argmin}_z \lambda \|z - z_0\|_{L^2(\Omega)}^2 + \alpha \left\| \nabla z - \left(\mathbf{r}^{k+1} + \bar{\mathbf{g}} + \mathbf{b}^k \right) \right\|_{L^2(\Omega)}^2 \quad (2.91)$$

$$\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{r}^{k+1} - \nabla z^{k+1} + \bar{\mathbf{g}} \quad (2.92)$$

à partir d'une solution initiale $(\mathbf{r}^0, z^0, \mathbf{b}^0) = (\nabla z_0 - \bar{\mathbf{g}}, z_0, \mathbf{0})$. Ce schéma comporte deux paramètres : $\lambda \geq 0$ est un paramètre du modèle (2.86), fixé a priori ; en revanche, α est un paramètre du schéma, assimilable à une vitesse de descente.

La convergence des schémas de type ADMM peut être établie pour toute valeur de α [235] : le choix d'une valeur particulière pour α affecte uniquement la vitesse de convergence. On peut en outre montrer que le taux de convergence de ces schémas est ergodique, i.e. égal à $O(1/k)$, et la méthode peut être accélérée pour atteindre un taux de convergence de $O(1/k^2)$ [103].

Le problème (2.90) est un problème de « poursuite de base » (*base pursuit*) [57], dont la solution est donnée, en chaque pixel, par l'opérateur de « seuillage doux » (les dépendances en (u, v) ont été omises pour des raisons de lisibilité, et pour montrer que ce seuillage est facile à paralléliser) :

$$\mathbf{r}^{k+1} = \frac{\nabla z^k - \bar{\mathbf{g}} - \mathbf{b}^k}{\|\nabla z^k - \bar{\mathbf{g}} - \mathbf{b}^k\|} \max \left\{ \left\| \nabla z^k - \bar{\mathbf{g}} - \mathbf{b}^k \right\| - \frac{1}{\alpha}, 0 \right\} \quad (2.93)$$

tandis que le problème (2.91) est un problème d'intégration en moindres carrés, dont nous avons déjà discuté abondamment la résolution numérique.

Remarquons que l'étape d'intégration en moindres carrés (2.91) doit être répétée un certain nombre de fois sur le même domaine. C'est précisément dans un tel cas qu'il est intéressant de recourir à l'approche itérative par gradient conjugué préconditionné. En effet, seul le second membre du système à résoudre change au cours des itérations. Le préconditionnement peut donc être effectué une fois pour toutes, et réutilisé sans nouveau calcul à chaque itération. Cela garantit des performances optimales, à la fois en termes de complexité algorithmique et d'occupation de la mémoire [152].

Mentionnons enfin deux autres méthodes que l'algorithme ADMM¹⁴ pour résoudre le problème (2.87) : l'algorithme FISTA proposé par Beck et Teboulle [27], dont le taux de convergence est théoriquement équivalent, mais qui est généralement plus rapide et ne dépend pas du choix d'un paramètre de descente tel que α , et les approches *primales-duales* [47, 288].

14. Pour une étude plus générale de ces algorithmes, nous renvoyons le lecteur aux deux monographies proposées par Boyd et al. [34, 199] qui traitent, respectivement, des méthodes ADMM et des opérateurs proximaux, ainsi qu'à [235] pour les preuves de convergence de plusieurs schémas numériques.

Résultats expérimentaux

Nous avons évalué l'algorithme ADMM dans [211] sur des champs de gradient obtenus par stéréophotométrie à partir des jeux de données réelles *Scholar*¹⁵ (cf. figure 2.9) et *Beethoven*¹⁶ (cf. figure 2.10). Pour mettre en évidence la robustesse aux discontinuités de profondeur, nous avons appliqué la méthode sur le domaine rectangulaire entier, en complétant $\mathbf{g}$ par des valeurs nulles.

Sur ces deux exemples, un artéfact en forme d'escalier apparaît en arrière-plan. Cet effet *staircase* est bien connu dans le contexte du débruitage : en adaptant les schémas fondés sur la variation totale généralisée [37], nous pourrions certainement nous en débarrasser. Nous proposons dans le prochain paragraphe une autre piste, fondée sur la régularisation non convexe. Cependant, cet effet n'est visible que sur l'arrière-plan, et ne doit donc pas être considéré comme réellement dommageable, puisque l'intégration peut être réalisée sur des domaines non rectangulaires : pour l'éviter, il suffirait de segmenter (par exemple, manuellement) l'objet dont on cherche à estimer le relief.

2.4.3 M-estimateurs

L'approche ADMM permet d'étendre très facilement l'intégration robuste aux estimateurs Φ_{M1} , Φ_{M2} et Φ_{M3} dont nous avons donné les définitions en (1.50), (1.51) et (1.52). Ce type d'estimateurs a déjà été adapté par Durou et al. [80] à l'intégration des normales, en modifiant la méthode itérative de Jacobi conçue par Horn et Brooks pour l'intégration en moindres carrés [126]. Plus généralement, le problème de l'intégration robuste peut s'écrire sous la forme suivante, qui généralise (2.87) :

$$\begin{cases} \min_{z, \mathbf{r}} \sum_{(u,v) \in \Omega} \Phi(\|\mathbf{r}(u,v)\|) + \lambda [z(u,v) - z_0(u,v)]^2 \\ \text{s.c.} \quad \mathbf{r} = \nabla z - \bar{\mathbf{g}} \end{cases} \quad (2.94)$$

et être résolu par le schéma numérique suivant :

$$\mathbf{r}^{k+1}(u,v) = \operatorname{argmin}_{\mathbf{r} \in \mathbb{R}^2} \mathcal{E}_\Phi(\mathbf{r}, u, v) \quad (2.95)$$

$$z^{k+1} = \operatorname{argmin}_z \lambda \|z - z_0\|_{L^2(\Omega)}^2 + \alpha \left\| \nabla z - (\mathbf{r}^{k+1} + \bar{\mathbf{g}} + \mathbf{b}^k) \right\|_{L^2(\Omega)}^2 \quad (2.96)$$

$$\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{r}^{k+1} - \nabla z^{k+1} + \bar{\mathbf{g}} \quad (2.97)$$

où :

$$\mathcal{E}_\Phi(\mathbf{r}, u, v) = \Phi(\|\mathbf{r}\|) + \alpha \left\| \mathbf{r} - (\nabla z^k(u,v) - \bar{\mathbf{g}}(u,v) - \mathbf{b}^k(u,v)) \right\|^2 \quad (2.98)$$

La convergence (ergodique) de ce schéma est garantie dès lors que Φ est convexe, ce qui n'est le cas ni de Φ_{M2} , ni de Φ_{M3} . Comme cela a été montré expérimentalement dans [80], des résultats satisfaisants peuvent toutefois être obtenus dans le cas non

15. <http://vision.seas.harvard.edu/qsfs/Data.html>

16. <http://www.ece.ncsu.edu/imaging/Archives/ImageDataBase/Industrial/>

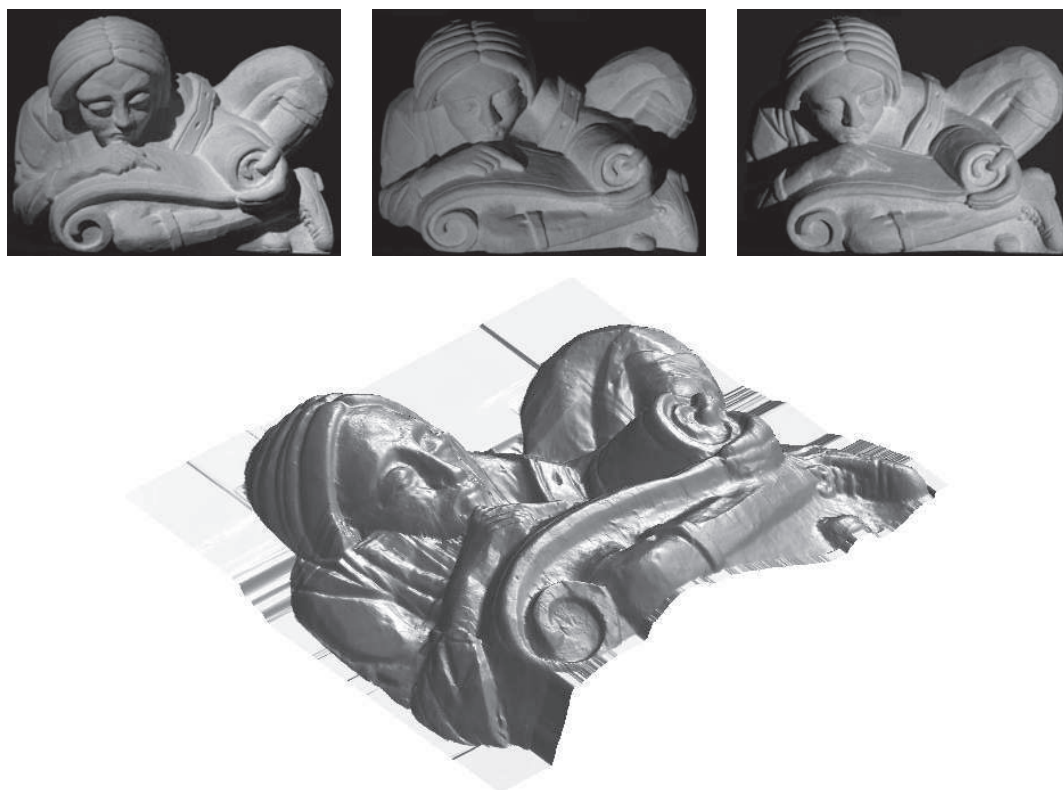


FIGURE 2.9 – Résultat de l’approche L^2 -TV (2.86) sur le jeu de données *Scholar*. Première ligne : trois des $m = 20$ photographies prises sous différents éclairages. Deuxième ligne : reconstruction 3D obtenue en intégrant les normales à l’aide du schéma ADMM. Un escalier est clairement visible sur cet exemple. Comme il affecte uniquement l’arrière-plan, il aurait suffi de segmenter le premier-plan pour l’éviter.

convexe, pourvu que la solution soit initialisée de façon appropriée, par exemple par la solution du problème en moindres carrés (2.61). Lorsque la fonction Φ n’est pas convexe, la convergence de l’algorithme ADMM vers le minimum global dépend du choix de α , qui doit être choisi suffisamment grand pour éviter les minima locaux [124].

Il est malheureusement difficile de *prédire* la convergence. Cela constitue à notre connaissance un problème ouvert, même si des efforts ont été déployés récemment dans ce sens [166]. Par ailleurs, le cas de fonctions Φ non convexes a été étudié très récemment par Lanza et al. dans [158] dans le cadre du débruitage d’images, qui est un problème très similaire à celui de l’intégration. Cette étude tend à confirmer l’intérêt qu’il y a à utiliser des fonctions Φ non convexes, sachant que le caractère « simple » des sous-problèmes (2.95) et (2.96) est la propriété essentielle qui permet au schéma ADMM d’être convergent (sans que cette propriété soit pour autant garantie). Le sous-problème (2.96) ayant déjà été étudié dans le cadre de l’intégration en moindres carrés,

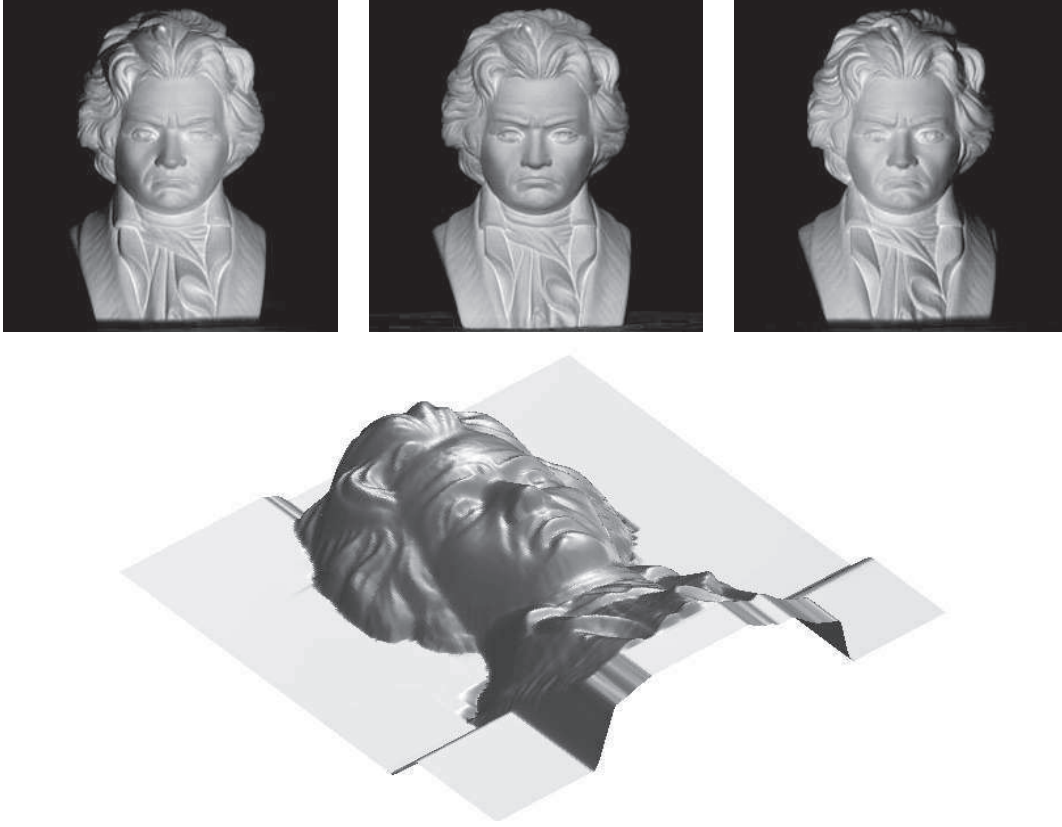


FIGURE 2.10 – Résultat de l’approche L^2 -TV (2.86) sur le jeu de données *Beethoven*. Première ligne : les $m = 3$ photographies acquises sous différents éclairages. Deuxième ligne : reconstruction 3D obtenue en intégrant les normales à l’aide du schéma ADMM. Sur cet exemple également, un escalier apparaît sur l’arrière-plan.

focalisons-nous maintenant sur la recherche d’une méthode de résolution efficace du sous-problème (2.95).

Lorsque la fonction Φ du problème (2.94) est telle que $\Phi(s) = |s|$ ou $\Phi(s) = s^2$, le problème (2.95) est non seulement convexe, mais admet même une solution analytique. Le cas où $\sum \sum_{(u,v) \in \Omega} \Phi(\|\mathbf{r}(u,v)\|)$ constitue une norme L^p , $0 < p < 1$, qui est un cas non convexe, a également été traité par Badri et al. [17]. Dans le cas général, il est toujours possible d’approcher la solution de (2.95) par des itérations de type Gauss-Newton, en remarquant que le gradient en $\mathbf{r}$ de la fonction $\mathcal{E}_\Phi(\mathbf{r}, u, v)$ définie en (2.98) s’écrit :

$$\nabla_{\mathbf{r}} \mathcal{E}_\Phi(\mathbf{r}, u, v) = \frac{\Phi'(\|\mathbf{r}\|)}{\|\mathbf{r}\|} \mathbf{r} + 2\alpha \left[\mathbf{r} - \left(\nabla z^k(u, v) - \bar{\mathbf{g}}(u, v) - \mathbf{b}^k(u, v) \right) \right] \quad (2.99)$$

mais il serait bien sûr préférable de disposer d’une solution explicite du problème (2.95), ce que nous allons faire pour les trois M-estimateurs suivants, dont les définitions géné-

ralisent celles du paragraphe 1.3.4 :

$$\Phi_1(s) = \sqrt{\beta^2 + s^2} - 1 \quad (2.100)$$

$$\Phi_2(s) = \frac{1}{2} \ln(\beta^2 + s^2) \quad (2.101)$$

$$\Phi_3(s) = \frac{1}{2\beta^2} \frac{s^2}{\beta^2 + s^2} \quad (2.102)$$

Leurs dérivées s'écrivent :

$$\Phi_1'(s) = \frac{s}{\sqrt{\beta^2 + s^2}} \quad (2.103)$$

$$\Phi_2'(s) = \frac{s}{\beta^2 + s^2} \quad (2.104)$$

$$\Phi_3'(s) = \frac{s}{(\beta^2 + s^2)^2} \quad (2.105)$$

La forte similarité entre ces trois expressions suggère d'utiliser la même stratégie pour les trois estimateurs. En effet, le gradient (2.99) se simplifie pour ces trois fonctions :

$$\nabla_{\mathbf{r}} \mathcal{E}_{\Phi}(\mathbf{r}, u, v) = \frac{1}{(\beta^2 + \|\mathbf{r}\|^2)^p} \mathbf{r} + 2\alpha \left[\mathbf{r} - \left(\nabla z^k(u, v) - \bar{\mathbf{g}}(u, v) - \mathbf{b}^k(u, v) \right) \right] \quad (2.106)$$

avec $p = 1/2$ pour $\Phi = \Phi_1$, $p = 1$ pour $\Phi = \Phi_2$ et $p = 2$ pour $\Phi = \Phi_3$.

Dans [80], le problème (2.94) est résolu par un schéma semi-implicite, qui consiste à rendre la partie linéaire implicite mais à garder la partie non linéaire explicite, i.e. à faire l'approximation suivante de (2.106) :

$$\nabla_{\mathbf{r}} \mathcal{E}_{\Phi}(\mathbf{r}, u, v) \approx \underbrace{\frac{1}{(\beta^2 + \|\mathbf{r}^k(u, v)\|^2)^p}}_{\text{Facteur constant}} \mathbf{r} + 2\alpha \left[\mathbf{r} - \left(\nabla z^k(u, v) - \bar{\mathbf{g}}(u, v) - \mathbf{b}^k(u, v) \right) \right] \quad (2.107)$$

de sorte que la mise à jour (2.95) s'écrit explicitement :

$$\mathbf{r}^{k+1}(u, v) = \frac{2\alpha \left(\beta^2 + \|\mathbf{r}^k(u, v)\|^2 \right)^p}{1 + 2\alpha \left(\beta^2 + \|\mathbf{r}^k(u, v)\|^2 \right)^p} \left(\nabla z^k(u, v) - \bar{\mathbf{g}}(u, v) - \mathbf{b}^k(u, v) \right) \quad (2.108)$$

Nous avons observé expérimentalement que l'algorithme ADMM semblait converger pour les fonctions non convexes Φ_2 et Φ_3 , dès lors que α est choisi suffisamment grand (il est fixé à 0,5 dans nos tests). Contrairement aux observations de Durou et al. dans [80], nous avons également remarqué que l'initialisation n'avait pas d'influence sur le résultat, mais uniquement sur la vitesse de convergence : outre de meilleures performances en termes de temps de calcul, l'algorithme ADMM que nous proposons semble donc éviter les minima locaux. Quant au paramètre β , sa valeur a une influence directe sur le résultat. Il semble que le choix d'une valeur faible pour β permette d'obtenir des arêtes plus saillantes, tout en accélérant la convergence (cf. figure 2.11).

β	L^2 -TV	Φ_1	Φ_2	Φ_3
0,1	 50 it. (2,76 s)	 13 it. (0,79 s)	 102 it. (4,71 s)	 63 it. (3,34 s)
0,5		 13 it. (0,78 s)	 175 it. (7,32 s)	 81 it. (3,94 s)
1		 14 it. (0,85 s)	 23 it. (1,34 s)	 188 it. (7,86 s)

FIGURE 2.11 – Reconstruction du buste de Beethoven par l'algorithme ADMM, à partir de la solution initiale $z^0 = z_0 \equiv 0$, avec $\alpha = 0,5$. La solution L^2 -TV, qui fait apparaître un escalier, est montrée en guise de comparaison. La fonction convexe Φ_1 ne permet pas de retrouver les discontinuités, contrairement aux fonctions non convexes Φ_2 et Φ_3 . La fonction Φ_3 offre une meilleure robustesse que Φ_2 au choix du paramètre β , mais le choix de Φ_2 peut être justifié par un raisonnement statistique (cf. paragraphe 2.2.1).

2.4.4 Une méthode d'intégration optimale ?

Les résultats des paragraphes précédents nous permettent d'établir que :

- La méthode optimale d'intégration en moindres carrés consiste à utiliser l'algorithme du gradient conjugué, accompagné d'un préconditionnement CMG [152].
- L'intégration par adaptation du modèle L^2 -TV requiert une minimisation itérative, qui peut être réalisée efficacement par l'algorithme ADMM, pourvu qu'un a priori quadratique sur la solution soit introduit : les discontinuités de profondeur sont bien retrouvées, mais des artéfacts en forme d'escalier apparaissent.
- Les estimateurs non convexes permettent eux aussi de bien retrouver les discontinuités de profondeur, sans artéfact en forme d'escalier, mais nécessitent d'effectuer un plus grand nombre d'itérations.

Au vu de ces remarques, nous proposons donc une nouvelle méthode d'intégration en trois étapes, qui semble être optimale à la fois en termes de temps de calcul, d'occupation de la mémoire et de précision :

1. Effectuer l'intégration en moindres carrés en utilisant un préconditionnement CMG du système linéaire (2.69), qui peut être résolu par gradient conjugué. La solution de cette première étape est notée z_{LS} .

2. En choisissant $z_0 = z^0 = z_{LS}$ et $\lambda = 10^{-9}$, résoudre le problème L^2 -TV (2.86) par l'algorithme ADMM, avec $\alpha = 0,5$. Ceci implique de résoudre plusieurs fois le problème en moindres carrés (2.91) : le préconditionnement effectué à l'étape 1 peut être réutilisé, ce qui permet une résolution très efficace. La solution de cette deuxième étape est notée z_{TV} .
3. En choisissant enfin $z_0 = z^0 = z_{TV}$, résoudre le problème non convexe (2.94), en choisissant $\Phi = \Phi_2$ et $\beta = 1$, par l'algorithme ADMM, avec $\alpha = 0,5$.

Pour les algorithmes itératifs (gradient conjugué et ADMM), nous utilisons comme critère d'arrêt un écart relatif moyen inférieur à 10^{-3} . La convergence du schéma est garantie à l'étape 2, mais pas à l'étape 3. Cependant, puisque l'étape 3 utilise une initialisation suffisamment proche de la solution, la convergence vers le minimum global n'est généralement pas mise en défaut.

Cet algorithme est illustré sur la figure 2.12, qui montre le relief estimé à l'issue de chaque étape, accompagné du temps de calcul cumulé.

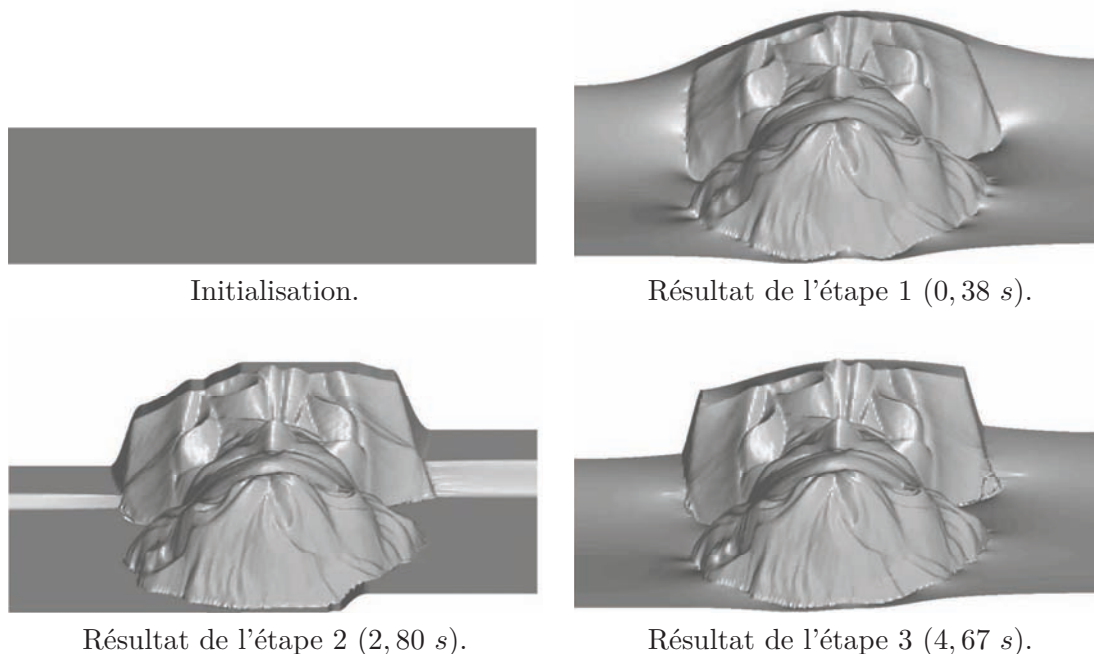


FIGURE 2.12 – Illustration de la méthode d'intégration optimale que nous préconisons : relief estimé à l'issue de chaque étape, accompagné du temps de calcul cumulé.

2.5 Conclusion du chapitre

Nous avons proposé dans ce chapitre une analyse, aussi exhaustive que possible, du problème de l'intégration d'un champ de normales sur un domaine Ω de forme quelconque (à ceci près que les points contenus dans Ω sont répartis sur une grille régulière). Après avoir énoncé les principales propriétés que doit vérifier un intégrateur « optimal », nous avons analysé différentes méthodes de résolution du problème en moindres carrés. Ces méthodes sont très rapides, mais ne sont pas robustes aux discontinuités de profondeur. Nous avons donc ensuite étudié différentes approches robustes aux discontinuités de profondeur, et proposé une méthode de résolution efficace par l'algorithme ADMM.

En couplant cette méthode d'intégration « optimale » avec les méthodes d'estimation robuste de la normale qui ont été discutées dans le chapitre 1, nous disposons dorénavant d'une chaîne de traitement efficace qui nous permet d'appliquer la stéréophotométrie à des scènes réelles comportant de nombreux écarts au modèle lambertien linéaire (éclairages non directionnels, ombres propres et portées, discontinuités de profondeur, etc.). La figure 2.13 montre un résultat obtenu en appliquant cette chaîne de traitement à un jeu de $m = 2812$ images d'une pelote de laine, provenant de [271]. Le relief estimé par stéréophotométrie « classique » (estimation de la normale et de l'albédo en moindres carrés par (1.24), suivie d'une intégration par DCT [239], après complétion par des 0 du gradient estimé) est très aplati, à cause des écarts au modèle (ombres propres, ombres portées et éclairage « ambiant »). Cela peut suffire pour des applications de rééclairage (e.g., dans le cadre de la *réalité augmentée*), mais certainement pas à des fins de métrologie. En revanche, le résultat est bien meilleur si l'on utilise la chaîne de traitement préconisée : estimation robuste de la normale et de l'albédo, puis intégration robuste du champ de normales.

Nous avons donc montré, dans cette première partie, qu'en ayant recours à des méthodes d'estimation robuste, la résolution classique (en deux étapes) de la stéréophotométrie permettait d'obtenir des reconstructions 3D très réalistes. Rappelons que le choix de résoudre la stéréophotométrie en deux étapes provient de la linéarité et du caractère bien posé de chacun des deux sous-problèmes. Cela explique pourquoi cette approche est suivie dans la plupart des articles. Nous en montrerons cependant les limites dans les deux prochaines parties. Dans la partie II, nous allons d'abord nous intéresser à deux situations particulières pour lesquelles l'estimation de la normale constitue un problème mal posé : la stéréophotométrie non calibrée et la stéréophotométrie à $m = 2$ images.

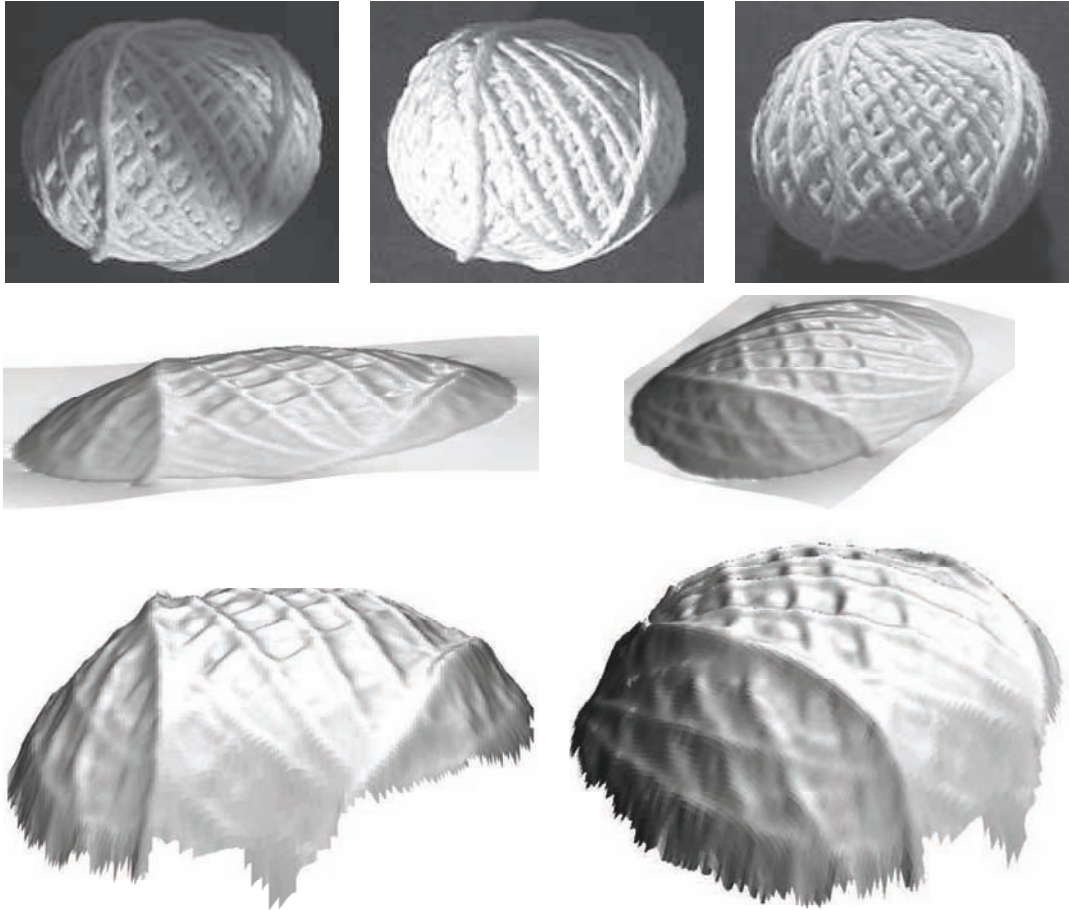


FIGURE 2.13 – Reconstruction 3D par stéréophotométrie robuste. Première ligne : trois images d’une pelote de laine (parmi $m = 2812$), de taille 144×171 . Deuxième ligne : reconstruction 3D par stéréophotométrie « classique ». Troisième ligne : reconstruction 3D par stéréophotométrie robuste. Même avec un nombre m d’images aussi élevé, l’approche classique est décevante : le relief est aplati, et les discontinuités de profondeur ne sont pas retrouvées. En utilisant une estimation robuste des normales, puis une intégration robuste aux discontinuités de profondeur, le relief est bien mieux restitué.

Deuxième partie

Deux problèmes de stéréophotométrie mal posés

Avant-propos

Nous avons montré dans la partie I que le modèle lambertien linéaire se prêtait particulièrement bien à la reconstruction 3D par stéréophotométrie. Ce modèle permet notamment de garantir l’unicité de la solution dès que $m \geq 3$ éclairages, non coplanaires et étalonnés, sont utilisés. Sa « naïveté » ne constitue pas un facteur limitant, si l’on a recours à des méthodes d’estimation robuste. On peut cependant se demander si l’unicité de la solution est toujours garantie lorsque les éclairages sont inconnus, c’est-à-dire en l’absence d’étalonnage, ou lorsqu’on ne dispose que de $m = 2$ éclairages. La linéarité de la formulation, qui constitue un avantage décisif pour la résolution de la stéréophotométrie, devient alors source d’ambiguïté. Comme nous allons le voir, l’unicité de la solution est en fait perdue, i.e. ces deux problèmes sont *mal posés*.

Deux approches peuvent être suivies pour surmonter cette difficulté. La première consiste à choisir un modèle plus complexe, comme cela a été proposé pour réduire les ambiguïtés du SFS [208]. Nous verrons dans la partie III que cette approche pose de sérieux problèmes numériques, et mène à des algorithmes très complexes, en comparaison des méthodes élémentaires que l’on peut déployer dans le cas lambertien linéaire. Nous suivons ici une autre approche, qui consiste à conserver le modèle lambertien linéaire, mais à introduire des contraintes supplémentaires sur la cohérence spatiale du champ de normales, de l’albédo et/ou de la profondeur, afin de garantir l’unicité de la solution. Qu’il s’agisse du problème de la stéréophotométrie non calibrée (*uncalibrated photometric stereo*, notée UPS), qui traite du cas où les éclairages sont inconnus, ou du problème de la stéréophotométrie à $m = 2$ images (*photometric stereo with two images*, notée PS2), la notion d’*intégrabilité*, qui a été introduite dans le chapitre 2 pour l’intégration des normales, permet de réduire les ambiguïtés.

Le problème UPS est étudié dans le chapitre 3. Reprenant les travaux que nous avons déjà présentés dans [77, 209, 214, 215, 217], nous montrons comment réduire l’ambiguïté inhérente à la formulation lambertienne linéaire, à la seule ambiguïté de *bas-relief généralisée* (*generalized bas-relief*, notée GBR), grâce à la contrainte d’intégrabilité. Nous montrons également qu’en ajoutant des contraintes de variation totale, le problème devient bien posé. Quant au problème PS2, son étude fait l’objet du chapitre 4. Nous montrons notamment que, même si l’albédo est connu, le niveau de gris peut être expliqué par deux normales en chaque point. Néanmoins, le champ de normales « le plus intégrable » peut être déterminé par *coupure de graphe*, comme nous l’avons déjà montré dans [218]. En outre, comme ce problème est similaire au cas de $m \geq 3$ éclairages coplanaires, nous montrons comment adapter la stéréophotométrie aux photographies prises en extérieur, grâce à une méthode originale qui a déjà été décrite dans [213].

Stéréophotométrie non calibrée

Nous avons montré dans le chapitre 1 comment estimer la normale et l'albédo par stéréophotométrie, en inversant le modèle lambertien linéaire (1.18) de façon robuste, lorsque les vecteurs d'éclairage $\mathbf{s}^i$ sont connus. Cependant, l'étalonnage des éclairages n'est pas une tâche facile (cf. chapitre 5), et il a été montré dans [128] qu'une erreur d'étalonnage, même infime, pouvait se traduire par un biais important dans les reconstructions 3D.

Une parade consiste à utiliser la stéréophotométrie *non calibrée* (UPS), i.e. à déterminer automatiquement les vecteurs d'éclairage, considérés comme des variables cachées du modèle. Cependant, il s'agit d'un problème mal posé, qui ne permet d'estimer le champ $\mathbf{m}$ qu'à une transformation linéaire près [114]. Si l'intégrabilité du champ de normales est imposée, et que l'appareil photographique est orthographique, l'ambiguïté linéaire se réduit à trois paramètres, et est connue sous le nom d'ambiguïté de *bas-relief généralisée* (GBR) [28].

Nous nous interrogeons, dans ce chapitre, sur la possibilité de lever l'ambiguïté du problème UPS au moyen d'une contrainte de régularité, par exemple la *variation totale*. Cela ouvrira la voie à la résolution variationnelle du problème UPS, sur laquelle nous reviendrons en détail dans la partie IV. Dans le présent chapitre, nous étudions le caractère bien posé du problème UPS régularisé. Tout ou partie de ce travail a déjà été présenté dans plusieurs articles [77, 209, 214, 215, 217].

Dans le paragraphe 3.1, nous rappelons pourquoi le problème UPS est un problème mal posé, puis nous montrons dans le paragraphe 3.2 comment lever l'ambiguïté par régularisation de la surface. En particulier, la variation totale du champ $\mathbf{m} = \rho \mathbf{n}$, ou bien de la profondeur z , permet de ramener l'ambiguïté GBR à une ambiguïté de type *bas-relief* (BR), au prix de l'introduction d'un a priori selon lequel la « meilleure » surface est celle orientée, en moyenne, vers l'appareil photographique. L'ambiguïté BR est plus facile à résoudre car elle dépend d'un seul paramètre, au lieu de trois paramètres pour l'ambiguïté GBR. Nous montrons de plus que les différentes régularisations par variation totale proposées reviennent à la minimisation de fonctions convexes très simples, grâce à quoi les coûts algorithmiques sont nettement inférieurs à ceux des méthodes de l'état de l'art. Mais surtout, la méthode de désambiguïsation du problème UPS par régularisation que nous proposons ouvrira la voie à une nouvelle manière de résoudre la stéréophotométrie au moyen de méthodes variationnelles (cf. partie IV) : l'intégrabilité sera garantie de façon naturelle en estimant directement la profondeur z .

3.1 Un problème mal posé

Dans ce chapitre, nous supposons que nous disposons de $m \geq 3$ photographies prises sous des éclairages directionnels non coplanaires, mais que les vecteurs d'éclairage $\mathbf{s}^i$, $i \in [1, m]$, sont inconnus. Ce problème est la version non calbrée de la stéréophotométrie, que nous désignerons par UPS (*uncalibrated photometric stereo*). Nous supposons également que le modèle lambertien linéaire (1.18) est vérifié, i.e. nous négligeons les phénomènes tels que les ombres ou les reflets brillants. Dans les tests sur données réelles, il nous faudra donc prétraiter les données de façon à les rendre conformes au modèle (1.18). Au vu de la discussion menée dans le paragraphe 1.2.3, nous utiliserons de préférence les prétraitements de type *low-rank* [270]. Enfin, nous supposons que la surface à reconstruire est régulière, au moins par morceaux, afin que l'intégrabilité du champ de normales soit garantie.

3.1.1 Ambiguïté linéaire

Dans l'équation (1.18), le vecteur $\mathbf{m}(u, v)$ constitue une inconnue vectorielle *en chaque pixel*, tandis que la matrice $\mathbf{S}$ constitue une inconnue matricielle *globale*. Plutôt que d'écrire une équation (1.18) par pixel et par éclairage pour expliquer les niveaux de gris, il est commode d'utiliser une seule équation matricielle pour décrire de façon compacte l'ensemble du processus de formation des images.

Pour cela, nous concaténons les vecteurs $\mathbf{i}(u, v) \in \mathbb{R}^m$ en une matrice $\mathbf{I} \in \mathbb{R}^{m \times |\Omega|}$, et les vecteurs $\mathbf{m}(u, v)$ en une matrice $\mathbf{M} \in \mathbb{R}^{3 \times |\Omega|}$, de telle sorte que le modèle devient :

$$\underbrace{\begin{bmatrix} I^1(u_1, v_1) & \dots & I^1(u_{|\Omega|}, v_{|\Omega|}) \\ \vdots \\ I^m(u_1, v_1) & \dots & I^m(u_{|\Omega|}, v_{|\Omega|}) \end{bmatrix}}_{\mathbf{I} \in \mathbb{R}^{m \times |\Omega|}} = \underbrace{\begin{bmatrix} \mathbf{s}^{1\top} \\ \vdots \\ \mathbf{s}^{m\top} \end{bmatrix}}_{\mathbf{S} \in \mathbb{R}^{m \times 3}} \underbrace{\begin{bmatrix} \mathbf{m}(u_1, v_1) & \dots & \mathbf{m}(u_{|\Omega|}, v_{|\Omega|}) \end{bmatrix}}_{\mathbf{M} \in \mathbb{R}^{3 \times |\Omega|}} \quad (3.1)$$

Dans cette équation, la matrice $\mathbf{I}$ constitue la donnée, et les matrices $\mathbf{S}$ et $\mathbf{M}$ les inconnues du problème UPS, qui est donc un problème de factorisation matricielle, connu en statistiques sous le nom de *séparation aveugle de source* (*blind source separation*). Bien que les variables liées aux éclairages puissent être considérées comme les variables cachées (ou latentes) d'un algorithme d'*espérance-maximisation* (EM), nous utilisons dans ce chapitre des outils d'algèbre linéaire élémentaire plutôt que des raisonnements statistiques. Le lecteur intéressé trouvera un aperçu des méthodes statistiques appropriées dans le chapitre 12 de [32], et nous reviendrons dans la partie IV sur la résolution du problème par une approche variationnelle. Quoi qu'il en soit, l'estimation de $\mathbf{S}$ et de $\mathbf{M}$ à partir de l'équation (3.1) est un problème mal posé [84, 114]. En effet, sa résolution n'est possible qu'à une ambiguïté linéaire près, puisque pour toute matrice $\mathbf{A} \in GL(3)$:

$$\mathbf{SM} = \mathbf{SAA}^{-1}\mathbf{M} \quad (3.2)$$

Nous allons voir qu'il est possible de limiter cette ambiguïté en imposant l'intégrabilité des normales (cf. paragraphe 2.1.2).

3.1.2 Contrainte d'intégrabilité

Dans le reste de cette partie, nous supposons l'appareil photographique orthographique. On rappelle (cf. chapitre 2) que la normale $\mathbf{n}$ est définie par :

$$\mathbf{n}(u, v) = \frac{1}{\sqrt{\|\nabla z(u, v)\|^2 + 1}} \begin{bmatrix} -p(u, v) \\ -q(u, v) \\ 1 \end{bmatrix} \quad (3.3)$$

où $\nabla z(u, v) = [p(u, v), q(u, v)]^\top$ est le gradient de la profondeur z . Si la profondeur z , qui caractérise la géométrie de la surface, est supposée deux fois différentiable, alors d'après le théorème de Schwarz, la contrainte d'intégrabilité :

$$\frac{\partial^2 z}{\partial_u \partial_v} = \frac{\partial^2 z}{\partial_v \partial_u} \quad (3.4)$$

doit être satisfaite. Or, nous savons d'après (2.5) que :

$$\begin{cases} \frac{\partial z}{\partial u} = p = -\frac{n_1}{n_3} \\ \frac{\partial z}{\partial v} = q = -\frac{n_2}{n_3} \end{cases} \quad (3.5)$$

Comme $\mathbf{m} = \rho \mathbf{n}$, nous déduisons de (3.4) et (3.5) que la contrainte d'intégrabilité revient à annuler la troisième composante du rotationnel $\text{rot}(\mathbf{n}/n_3) = \text{rot}(\mathbf{m}/m_3)$, notée $\overline{\text{rot}} \mathbf{m}$:

$$\overline{\text{rot}} \mathbf{m} = \frac{\partial}{\partial v} \left(\frac{m_1}{m_3} \right) - \frac{\partial}{\partial u} \left(\frac{m_2}{m_3} \right) = 0 \quad (3.6)$$

Yuille et Snow proposent dans [282] de minimiser l'écart au modèle lambertien linéaire, tout en imposant la contrainte (3.6). Le problème UPS s'écrit alors comme l'estimation conjointe de la matrice $\mathbf{S}$ et du champ $\mathbf{m}$:

$$\begin{cases} \min_{\substack{\mathbf{S} \in \mathbb{R}^{m \times 3} \\ \mathbf{m}: \Omega \rightarrow \mathbb{R}^3}} \|\mathbf{S}\mathbf{m} - \mathbf{i}\|_{L^2(\Omega)}^2 \\ \text{s.c. } \overline{\text{rot}} \mathbf{m} = 0 \end{cases} \quad (3.7)$$

D'un point de vue pratique, plutôt que de traiter le problème continu (3.7), il est plus facile de résoudre le problème discret associé. Nous verrons en effet que, dans le cas discret, l'ensemble des solutions admissibles peut être déterminé par des techniques de factorisation matricielle. En utilisant les définitions (3.1) des matrices $\mathbf{M}$ et $\mathbf{I}$, nous nous référerons donc parfois au problème discret suivant :

$$\begin{cases} \min_{\substack{\mathbf{S} \in \mathbb{R}^{m \times 3} \\ \mathbf{M} \in \mathbb{R}^{3 \times |\Omega|}}} \|\mathbf{S}\mathbf{M} - \mathbf{I}\|_F^2 \\ \text{s.c. } \overline{\text{rot}} \mathbf{M} = 0 \end{cases} \quad (3.8)$$

Dans cette écriture, la notation abusive $\overline{\text{rot}} \mathbf{M}$ est entendue comme une approximation discrète de l'opération définie en (3.6), calculée par différences finies sur le champ discret $\mathbf{m}$ formé par les colonnes de la matrice $\mathbf{M}$:

$$\mathbf{m}(u, v) = \mathbf{M}(:, \#(u, v)) \quad (3.9)$$

où $\#(u, v)$ désigne le numéro de colonne du pixel (u, v) dans la matrice $\mathbf{M}$.

3.1.3 Résolution du problème (3.8)

La SVD de $\mathbf{I}$ nous permet de trouver trois matrices $\hat{\mathbf{U}} \in \mathbb{R}^{m \times m}$, $\hat{\mathbf{W}} \in \mathbb{R}^{m \times |\Omega|}$ et $\hat{\mathbf{V}} \in \mathbb{R}^{|\Omega| \times |\Omega|}$ telles que :

$$(\hat{\mathbf{U}}, \hat{\mathbf{W}}, \hat{\mathbf{V}}) = \underset{\mathbf{U}, \mathbf{W}, \mathbf{V}}{\operatorname{argmin}} \|\mathbf{U}\mathbf{W}\mathbf{V}^\top - \mathbf{I}\|_F^2 \quad (3.10)$$

Comme nous l'avons déjà expliqué dans le paragraphe 1.2.2, le rang de la matrice $\mathbf{I}$ devrait être au plus égal à 3, donc seules ses trois plus grandes valeurs singulières devraient être non nulles¹. Hayakawa propose dans [114] d'utiliser une SVD tronquée à l'ordre 3, i.e. de restreindre $\hat{\mathbf{W}}$ à sa sous-matrice principale $\overline{\mathbf{W}} \in \mathbb{R}^{3 \times 3}$, et $\hat{\mathbf{U}}$ et $\hat{\mathbf{V}}$ à leurs trois premières colonnes, c'est-à-dire à $\overline{\mathbf{U}} \in \mathbb{R}^{m \times 3}$ et $\overline{\mathbf{V}} \in \mathbb{R}^{|\Omega| \times 3}$, de telle sorte que $\mathbf{I} \approx \overline{\mathbf{U}} \overline{\mathbf{W}} \overline{\mathbf{V}}^\top$. Les solutions du problème :

$$\min_{\substack{\mathbf{S} \in \mathbb{R}^{m \times 3} \\ \mathbf{M} \in \mathbb{R}^{3 \times |\Omega|}}} \|\mathbf{S}\mathbf{M} - \mathbf{I}\|_F^2 \quad (3.11)$$

s'écrivent donc :

$$\begin{cases} \mathbf{S}^0 = \overline{\mathbf{U}} \mathbf{P}^\top \\ \mathbf{M}^0 = \mathbf{Q} \overline{\mathbf{V}}^\top \end{cases} \quad (3.12)$$

où $\mathbf{P}$ et $\mathbf{Q}$ sont deux matrices 3×3 telles que :

$$\mathbf{P}^\top \mathbf{Q} = \overline{\mathbf{W}} \quad (3.13)$$

Néanmoins, le problème reste mal posé puisqu'il existe une infinité de tels couples $(\mathbf{P}, \mathbf{Q})$: pour toute matrice $\mathbf{A} \in \text{GL}(3)$, $\mathbf{P}^\top \mathbf{Q} = \mathbf{P}^\top \mathbf{A} \mathbf{A}^{-1} \mathbf{Q}$. Résoudre le problème de la stéréophotométrie non calbrée nécessite de lever cette ambiguïté, ce qui n'est possible qu'en ajoutant des contraintes.

Yuille et Snow montrent dans [282, 283] que l'intégrabilité du champ de normales peut être utilisée pour identifier les matrices $\mathbf{P}$ et $\mathbf{Q}$ dans (3.12). En effet, la contrainte d'intégrabilité (3.6) se traduit, après quelques manipulations algébriques, par un système

1. En pratique, les autres valeurs singulières ne sont pas exactement nulles, à cause des écarts au modèle : Miyazaki et Ikeuchi suggèrent dans [178] d'effectuer une SVD robuste ; il est également possible de prétraiter les données par une approche *low-rank* [270].

linéaire sur-contraint ayant pour inconnues six des neuf cofacteurs de la matrice $\mathbf{Q}$, en l'occurrence :

$$\begin{cases} \delta_{1,1} = \mathbf{Q}_{2,2}\mathbf{Q}_{3,3} - \mathbf{Q}_{3,2}\mathbf{Q}_{2,3} \\ \delta_{1,2} = -(\mathbf{Q}_{2,1}\mathbf{Q}_{3,3} - \mathbf{Q}_{3,1}\mathbf{Q}_{2,3}) \\ \delta_{1,3} = \mathbf{Q}_{2,1}\mathbf{Q}_{3,2} - \mathbf{Q}_{3,1}\mathbf{Q}_{2,2} \\ \delta_{2,1} = -(\mathbf{Q}_{1,2}\mathbf{Q}_{3,3} - \mathbf{Q}_{3,2}\mathbf{Q}_{1,3}) \\ \delta_{2,2} = \mathbf{Q}_{1,1}\mathbf{Q}_{3,3} - \mathbf{Q}_{3,1}\mathbf{Q}_{1,3} \\ \delta_{2,3} = -(\mathbf{Q}_{1,1}\mathbf{Q}_{3,2} - \mathbf{Q}_{3,1}\mathbf{Q}_{1,2}) \end{cases} \quad (3.14)$$

qui peuvent donc être estimés au sens des moindres carrés. Les trois autres cofacteurs $(\delta_{3,1}, \delta_{3,2}, \delta_{3,3})$ n'ont aucune influence ni sur la valeur de $\|\mathbf{S}\mathbf{M} - \mathbf{I}\|_F^2$, ni sur l'intégrabilité du champ discret $\mathbf{m}$, qui est défini à partir de $\mathbf{M}$ comme indiqué en (3.9). On peut donc fixer ces trois autres cofacteurs, par exemple, à $(0, 0, 1)$. Si la matrice $\mathbf{Q}$ est inversible, on sait que $\mathbf{Q}^{-1}$ est proportionnelle à la transposée de sa comatrice :

$$\mathbf{Q} \propto \begin{bmatrix} \delta_{1,1} & \delta_{2,1} & \delta_{3,1} \\ \delta_{1,2} & \delta_{2,2} & \delta_{3,2} \\ \delta_{1,3} & \delta_{2,3} & \delta_{3,3} \end{bmatrix}^{-1} \quad (3.15)$$

La matrice $\mathbf{Q}$ est donc entièrement déterminée en fixant à 1 le coefficient de proportionnalité dans (3.15), et $\mathbf{P}$ se déduit de $\mathbf{Q}$ en utilisant (3.13). Dans ce qui suit, nous noterons $(\mathbf{S}^0, \mathbf{M}^0)$ la solution du problème (3.8) obtenue en calculant (3.12) avec les valeurs de $\mathbf{P}$ et $\mathbf{Q}$ ainsi déterminées. Nous allons voir que ce problème admet une infinité d'autres solutions.

3.1.4 Ambiguïté GBR

Il est prouvé dans [28, 282, 283] que les solutions du problème (3.8) s'écrivent :

$$\begin{cases} \mathbf{S} = \mathbf{S}^0 \mathbf{G}(\mu, \nu, \lambda) \\ \mathbf{M} = \mathbf{G}(\mu, \nu, \lambda)^{-1} \mathbf{M}^0 \end{cases} \quad (3.16)$$

où $(\mathbf{S}^0, \mathbf{M}^0)$ désigne la solution du problème (3.8) déterminée dans le paragraphe 3.1.3, et où la matrice $\mathbf{G}(\mu, \nu, \lambda)$, définie comme suit :

$$\mathbf{G}(\mu, \nu, \lambda) = \begin{pmatrix} 1 & 0 & -\mu/\lambda \\ 0 & 1 & -\nu/\lambda \\ 0 & 0 & 1/\lambda \end{pmatrix} \quad (3.17)$$

dépend de trois paramètres $(\mu, \nu, \lambda) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ a priori inconnus. En effet, il a été montré par une recherche exhaustive [283] qu'il s'agissait des seules transformations laissant inchangés les six cofacteurs (3.14) de la matrice $\mathbf{Q}$. Il est facile de vérifier que l'inverse de $\mathbf{G}(\mu, \nu, \lambda)$ s'écrit :

$$\mathbf{G}(\mu, \nu, \lambda)^{-1} = \begin{pmatrix} 1 & 0 & \mu \\ 0 & 1 & \nu \\ 0 & 0 & \lambda \end{pmatrix} \quad (3.18)$$

L'indétermination des paramètres (μ, ν, λ) dans l'expression (3.16) de la solution correspond à l'ambiguïté de *bas-relief généralisée* (GBR) [28]. Seul le signe de λ est prévisible, car la troisième composante du champ $\mathbf{m}$ est nécessairement positive. Nous supposons dorénavant, sans perte de généralité, que $\lambda > 0$.

Le problème (3.8) reste donc mal posé, puisque l'on peut appliquer n'importe quelle transformation GBR (3.16) sans changer ni la valeur de l'attache aux données $\|\mathbf{SM} - \mathbf{I}\|_F^2$, ce qui est une conséquence directe de (3.2), ni la contrainte d'intégrabilité puisque, comme cela est montré dans [28, 282, 283] :

$$\overline{\text{rot}} \mathbf{m} = 0 \iff \overline{\text{rot}}(\mathbf{G}(\mu, \nu, \lambda)^{-1} \mathbf{m}) = 0 \quad (3.19)$$

Il est montré dans [28] qu'une transformation GBR préserve également les ombres (propres et portées). On pourrait donc se demander s'il était vraiment nécessaire de supposer les images exemptes d'ombres, comme nous l'avons fait au début de ce chapitre. Or, bien que l'estimation des paramètres (μ, ν, λ) ne soit pas affectée par la présence d'ombres dans les images, l'estimation de la solution initiale $(\mathbf{S}^0, \mathbf{M}^0)$ effectuée selon la procédure décrite dans le paragraphe 3.1.3 serait biaisée. Cela explique donc nos hypothèses de travail.

En termes de profondeur, il est montré dans [28] qu'une transformation GBR modifie $z(u, v)$ de la façon suivante :

$$\bar{z}(u, v) = \frac{z(u, v) - \mu u - \nu v}{\lambda} \quad (3.20)$$

La figure 3.1 illustre les effets d'une telle transformation sur l'albédo, sur les normales et sur le relief. L'expression (3.20) montre d'autre part que le paramètre λ contrôle l'échelle du relief dans la direction de l'axe z : cette dernière opération correspond en fait à une transformation dite « de bas-relief » (BR).

La résolution de l'ambiguïté GBR a fait l'objet d'un nombre très élevé de publications, ce qui peut surprendre vu son apparente simplicité. Hayakawa propose dans [114] d'identifier six pixels possédant le même albédo, ou d'utiliser au moins six éclairages dont les intensités sont supposées égales. Néanmoins, Yuille et Snow montrent dans [282] que, sans la contrainte d'intégrabilité, ceci ne contraint pas suffisamment le problème. Lorsque l'intégrabilité du champ de normales est imposée, ils montrent aussi que l'égalité des intensités des éclairages permet de lever l'ambiguïté GBR. Les images réelles s'écartant généralement du modèle lambertien linéaire (3.1) (cf. chapitre 1), la détection d'écarts au modèle tels que les reflets brillants [73, 74] ou les réflexions mutuelles [54] permet également de lever l'ambiguïté GBR, tout comme la détection des points singuliers dans les images², i.e. des points où le niveau de gris est localement maximal [89, 196].

Toutes ces tentatives pour lever l'ambiguïté GBR reposent explicitement sur une étape de détection. Non seulement cette étape peut être délicate, mais l'existence même de points remarquables n'est pas garantie : il paraît donc plus raisonnable d'introduire un a priori global sur le relief [99] ou sur sa réflectance. Il a notamment été montré

2. Ces points sont également très appréciés dans le contexte du SFS [132], puisqu'il s'agit des rares points où le problème n'est pas ambigu. Nous y reviendrons dans le chapitre 4.

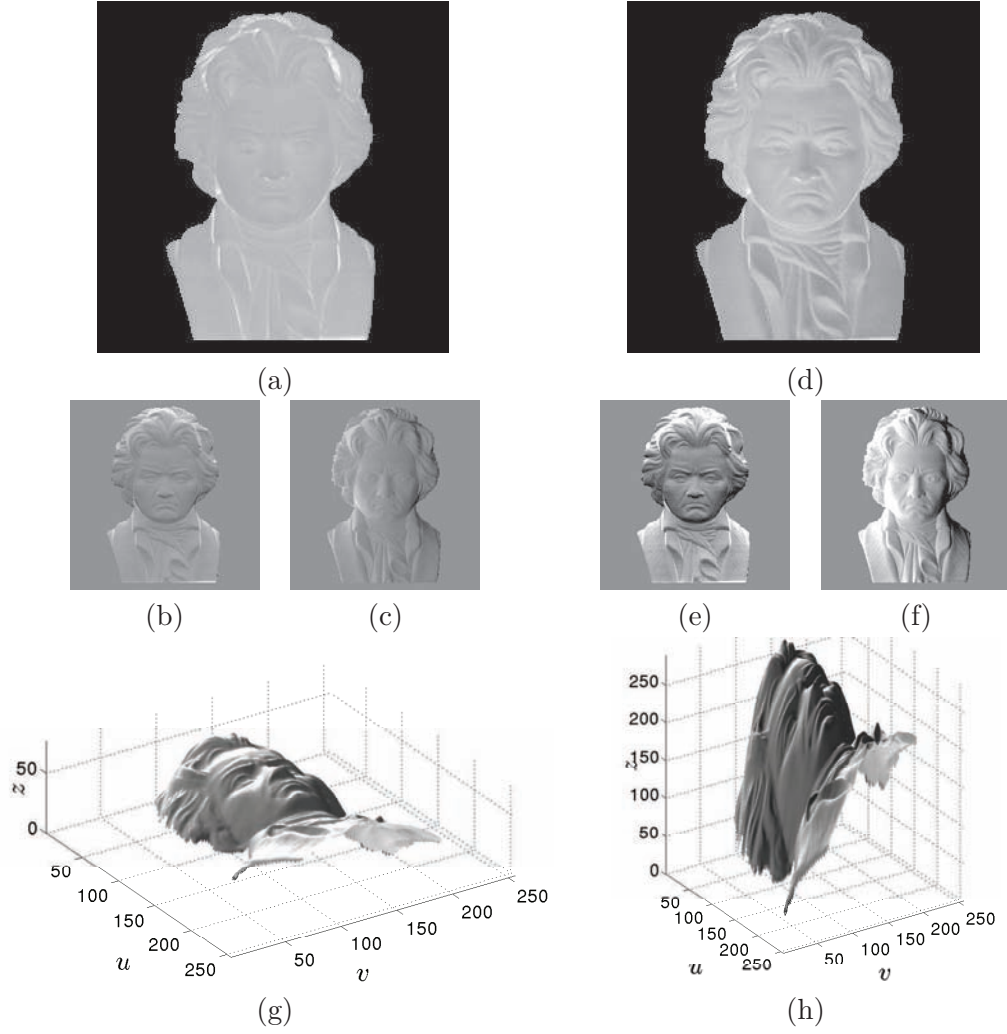


FIGURE 3.1 – Effet d’une transformation GBR. (a)-(b)-(c) Estimation de ρ , p et q par stéréophotométrie calibrée, à partir des $m = 3$ images de la figure 1.2. (d)-(e)-(f) Résultat d’une transformation GBR, avec $\mu = 1/5$, $\nu = 1/2$ et $\lambda = 1/3$. (g)-(h) Reliefs correspondants, obtenus par intégration [239]. La solution du problème non calibré (UPS) située sur la colonne de droite est a priori tout aussi plausible que celle de la colonne de gauche. Néanmoins, l’albédo de la solution de droite est beaucoup moins uniforme que celui de la solution de gauche. Nous observons également que les variations de la profondeur et de son gradient sont plus élevées après transformation GBR. Il semble donc que la minimisation de ces variations puisse nous aider à lever l’indétermination.

récemment que, pour certains modèles de réflectance non lambertiens, le problème UPS pouvait être mieux posé [98]. Un moyen de se passer à la fois de la connaissance des éclairages et d’un modèle explicite de réflectance consiste à « apprendre » les éclairages

et la réflectance sur les images d'un objet de référence [121, 122], comme cela sera détaillé dans le paragraphe 6.1. Dans le cas lambertien, Alldrin et al. proposent de minimiser l'entropie de la distribution d'albédo [9], puisqu'une transformation GBR a tendance à « étaler » l'histogramme de l'albédo (cf. figures 3.1-a et 3.1-d). Ils motivent leur travail par la volonté de favoriser des matériaux présentant une certaine « homogénéité », i.e. constitués d'un faible nombre de composantes homogènes. Cette méthode de résolution du problème UPS revient en fait à résoudre séquentiellement le couple de problèmes suivant :

$$\begin{cases} (\mathbf{S}^0, \mathbf{m}^0) = \underset{\mathbf{S}, \mathbf{m}}{\operatorname{argmin}} \|\mathbf{S}\mathbf{m} - \mathbf{i}\|_{L^2(\Omega)}^2 \\ \text{s.c. } \overline{\operatorname{rot}} \mathbf{m} = 0 \\ (\hat{\mu}, \hat{\nu}, \hat{\lambda}) = \underset{\mu, \nu, \lambda}{\operatorname{argmin}} \mathcal{E}(\|\mathbf{G}(\mu, \nu, \lambda)^{-1} \mathbf{m}^0\|) \end{cases} \quad (3.21)$$

où $\mathcal{E}(\|\mathbf{G}(\mu, \nu, \lambda)^{-1} \mathbf{m}^0\|)$ désigne l'entropie de la distribution de l'albédo affecté par une transformation GBR (on rappelle que $\rho = \|\mathbf{m}\|$). Le premier sous-problème de (3.21) n'est autre que (3.7), le deuxième vise à lever l'ambiguïté GBR. Finalement, une fois le problème (3.21) résolu, la solution du problème UPS s'écrit, en utilisant (3.16) :

$$\begin{cases} \hat{\mathbf{S}} = \mathbf{S}^0 \mathbf{G}(\hat{\mu}, \hat{\nu}, \hat{\lambda}) \\ \hat{\mathbf{m}} = \mathbf{G}(\hat{\mu}, \hat{\nu}, \hat{\lambda})^{-1} \mathbf{m}^0 \end{cases} \quad (3.22)$$

3.1.5 Régularisation par variation totale

La *variation totale* (*total variation*, notée TV) d'une fonction est une mesure très utilisée dans le contexte de la régularisation. Pour une fonction $f : \Omega \subset \mathbb{R}^p \rightarrow \mathbb{R}$ différentiable presque partout, elle s'écrit :

$$\operatorname{TV}(f) = \int_{\Omega} \|\nabla f(\mathbf{x})\| \, d\mathbf{x} \quad (3.23)$$

et s'étend plus généralement aux fonctions à *variation bornée* [15].

Nous fixons dorénavant $p = 2$, puisque nous traitons des images 2D. La variation totale a été introduite dans la communauté du traitement d'images par Rudin, Osher et Fatemi [226] : étant donnée une image en niveaux de gris bruitée $I_0 = I + \eta$, où η est la réalisation d'un processus gaussien de moyenne nulle et de variance σ^2 connue, le modèle ROF (Rudin, Osher, Fatemi) assimile l'image débruitée à la solution du problème suivant :

$$\begin{cases} \min_I \operatorname{TV}(I) \\ \text{s.c. } \|I - I_0\|_{L^2(\Omega)}^2 \leq \sigma^2 \end{cases}$$

Chambolle et Lions ont montré dans [46] que ce problème était équivalent à la minimisation de la fonctionnelle :

$$\mathcal{F}(I) = \|I - I_0\|_{L^2(\Omega)}^2 + \gamma_{\sigma} \operatorname{TV}(I) \quad (3.24)$$

pour une certaine valeur du paramètre de régularisation γ_σ . Des algorithmes efficaces de résolution du problème (3.24) ont été proposés ces dernières années, parmi lesquels les plus fréquemment rencontrés en traitement d'images sont les algorithmes de type *primal-dual* [45], la descente de gradient améliorée [27] ou les approches par lagrangien augmenté [34, 104]. Ces dernières seront utilisées pour la résolution numérique du problème UPS dans la partie IV. Dans le cadre d'une étude plus théorique sur le caractère bien posé du problème UPS, nous allons voir ci-après que la résolution peut être effectuée à l'aide de méthodes élémentaires.

Nous nous intéressons dans ce chapitre à la régularisation de la profondeur z et de l'albédo ρ , mais également à celle du champ vectoriel $\mathbf{m} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$. Il nous faut donc définir la *variation totale vectorielle*. Lorsque $\mathbf{f}$ est une fonction à valeurs dans $\mathbb{R}^{n_d}$, avec $n_d > 1$, plusieurs définitions peuvent être utilisées³, mais la définition la plus répandue est probablement celle de la variation totale *isotrope* :

$$\text{TV}(\mathbf{f}) = \iint_{\Omega} \|\mathbf{J}_{\mathbf{f}}(u, v)\|_F du dv \quad (3.25)$$

où $\|\mathbf{J}_{\mathbf{f}}(u, v)\|_F$ est la norme de Frobenius de la matrice jacobienne $\mathbf{J}_{\mathbf{f}}$ de $\mathbf{f} = [f_1 \dots f_{n_d}]^\top$ au point (u, v) . Cette définition, qui est invariante par rotation des composantes, couple leurs dérivées partielles, ce qui est intéressant par exemple pour le débruitage d'images. Une autre définition courante de la variation totale vectorielle est la suivante :

$$\text{TV}(\mathbf{f}) = \sum_{i=1}^{n_d} \iint_{\Omega} \|\nabla f_i(u, v)\| du dv = \sum_{i=1}^{n_d} \text{TV}(f_i) \quad (3.26)$$

qui est *anisotrope*, c'est-à-dire non invariante par rotation, puisque les régularisations des n_d composantes sont découplées.

Contrairement au problème de l'intégration, où nous avons choisi une régularisation similaire à la définition isotrope (3.25) (cf. paragraphe 2.4.2), l'invariance par rotation n'est pas une propriété intéressante pour le problème UPS. En effet, l'ordre des trois composantes m_1 , m_2 et m_3 est important vis-à-vis de la contrainte d'intégrabilité (3.6), et chaque paramètre d'une transformation GBR affecte une seule composante de $\mathbf{m}$. Nous préférons utiliser la définition anisotrope (3.26), qui est plus facile à manipuler. Dans la suite de ce chapitre, la variation totale du champ $\mathbf{m} = [m_1, m_2, m_3]^\top$ sera donc définie par :

$$\text{TV}(\mathbf{m}) = \text{TV}(m_1) + \text{TV}(m_2) + \text{TV}(m_3) \quad (3.27)$$

Nous verrons que la minimisation de la variation totale de la profondeur z ou du champ $\mathbf{m}$ ramène l'ambiguïté GBR à une ambiguïté BR, et que la minimisation de la variation totale de l'albédo ρ permet de lever cette ambiguïté résiduelle. Nous proposerons dans la partie IV une résolution variationnelle « classique » du problème UPS, permettant de déterminer la profondeur tout en garantissant sa régularité. Dans le présent chapitre, nous nous intéressons surtout à prouver le caractère bien posé du problème

3. Nous renvoyons le lecteur à [38] et [100] pour une discussion plus approfondie.

UPS régularisé : nous n'utilisons donc pas la variation totale comme un outil de « lissage » des éventuels biais de l'estimation, mais comme un moyen de rendre le problème bien posé.

Il nous reste à déterminer, parmi les solutions (3.16) du problème (3.8), qui dépendent des trois paramètres (μ, ν, λ) de l'ambiguïté GBR, celle qui minimise le critère de régularité choisi. On peut donc s'attendre à ce que ce problème d'optimisation à trois variables réelles soit relativement facile à analyser et à résoudre numériquement.

3.2 Désambiguïstation par régularisation

3.2.1 Motivations

Au vu des résultats de la figure 3.1, une transformation GBR semble augmenter les variations d'un certain nombre de champs caractéristiques de la surface à reconstruire, ce qui suggère d'introduire un critère de *régularisation* pour résoudre le problème UPS. Dans ce paragraphe, nous illustrons plus clairement l'influence d'une transformation GBR sur les variations de l'albédo ρ , de la profondeur z et de son gradient $\nabla z = [p, q]^\top = -[n_1/n_3, n_2/n_3]^\top$.

Ce travail s'inspire de la méthode de Alldrin et al., qui montrent dans [9] que les paramètres GBR peuvent être estimés en minimisant l'entropie de la distribution d'albédo. Or, l'entropie de l'albédo n'étant pas convexe vis-à-vis des paramètres (μ, ν, λ) , sa minimisation nécessite de prendre garde à la possible existence de minima locaux : Alldrin et al. utilisent une approche de type « glouton » (*brute force*), en échantillonnant l'ensemble des valeurs possibles de μ , ν et λ , ce qui se traduit par un algorithme extrêmement lent. Même si des méthodes numériques assez efficaces ont été récemment introduites [241], la minimisation de l'entropie reste un problème difficile. De plus, l'entropie est indépendante de toute notion de cohérence spatiale, comme cela est illustré sur la figure 3.2.

Lorsqu'on cherche à favoriser des zones « homogènes », on peut postuler que des points spatialement proches ont des albédos similaires, ce que ne favorise nullement l'entropie. En revanche, la variation totale permet de minimiser les variations locales de l'albédo, qui doivent être faibles dans les zones homogènes. De plus, la minimisation de la variation totale étant nettement plus facile que celle de l'entropie, on n'a pas à utiliser un algorithme glouton et on peut donc s'attendre à de meilleures performances CPU.

Une transformation GBR modifiant non seulement l'albédo de la surface, mais également son relief, les variations de ρ ne sont pas les seules affectées : celles de la profondeur z et de son gradient $\nabla z = [p, q]^\top$ le sont également. Nous montrons sur la figure 3.3 l'influence d'une transformation GBR sur ces variations.

Il semble donc que, pour lever l'ambiguïté GBR, on puisse minimiser non seulement les variations de ρ , mais également celles de z et de ∇z . Or, puisque :

$$\mathbf{m} = \frac{\rho}{\sqrt{\|\nabla z\|^2 + 1}} \begin{bmatrix} -\nabla z \\ 1 \end{bmatrix} \quad (3.28)$$

pourquoi ne pas directement minimiser les variations du champ $\mathbf{m}$?

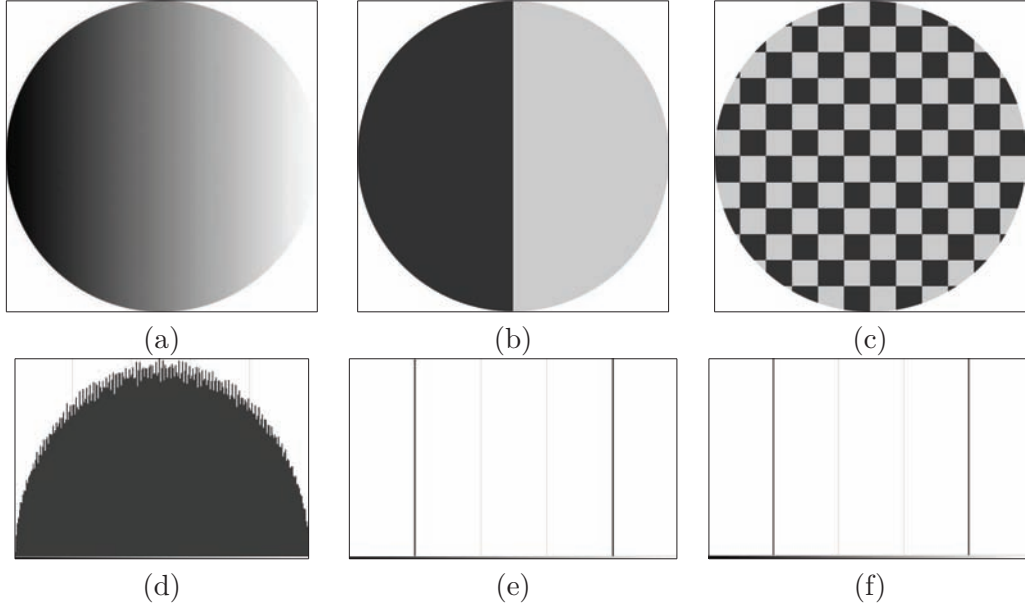


FIGURE 3.2 – Trois configurations d'albédo et leurs histogrammes respectifs : l'entropie de (a) est élevée, cf. l'histogramme (d) ; (b) et (c) ont la même entropie (faible), cf. les histogrammes (e) et (f), mais les distributions spatiales de l'albédo sont très différentes dans les deux cas. La variation totale tend à favoriser des situations telles que (b).

3.2.2 Influence d'une transformation GBR sur la régularité de ρ , z et $\mathbf{m}$

Sur les figures 3.4 et 3.5, nous illustrons l'influence des paramètres μ et ν sur les variations totales de l'albédo ρ , de la profondeur z et du champ $\mathbf{m}$ estimés à partir des images du buste de Beethoven. Pour effectuer ces mesures, nous avons choisi comme vérité terrain la solution du problème calibré, à laquelle nous avons appliqué différentes transformations GBR, en fixant λ à 1 : les valeurs théoriques des paramètres μ et ν sont donc $(\mu, \nu) = (0, 0)$. Ces courbes montrent que la minimisation de la variation totale de ρ , de z ou de $\mathbf{m}$ permet effectivement de lever l'ambiguïté en μ et ν .

Le cas de λ est très différent. L'effet de ce paramètre sur les variations totales, lorsque μ et ν sont fixés à leurs valeurs théoriques, est illustré sur la figure 3.6 : $\text{TV}(z)$ est minimale pour $\lambda = +\infty$ (cf. paragraphe 3.2.6 pour l'explication de ce phénomène), et $\text{TV}(\mathbf{m})$ est minimale pour $\lambda = 0$ (cf. paragraphe 3.2.5), deux cas qui correspondent à un relief dégénéré (cf. équation (3.20)). Seule la minimisation de la variation totale de l'albédo $\text{TV}(\rho)$ semble capable de lever l'indétermination sur la valeur de λ (ambiguïté BR). En pratique, les résultats peuvent néanmoins être décevants : l'expérience reportée sur la figure 3.6-a est biaisée par le caractère uniforme de l'albédo (le buste est en plâtre). En utilisant le jeu de données *Doll* de [9], pour lequel l'albédo n'est pas rigoureusement uniforme, on obtient les résultats de la figure 3.7 : l'indétermination est bien levée, mais la valeur de λ qui minimise $\text{TV}(\rho)$ est très éloignée de sa valeur réelle $\lambda = 1$.

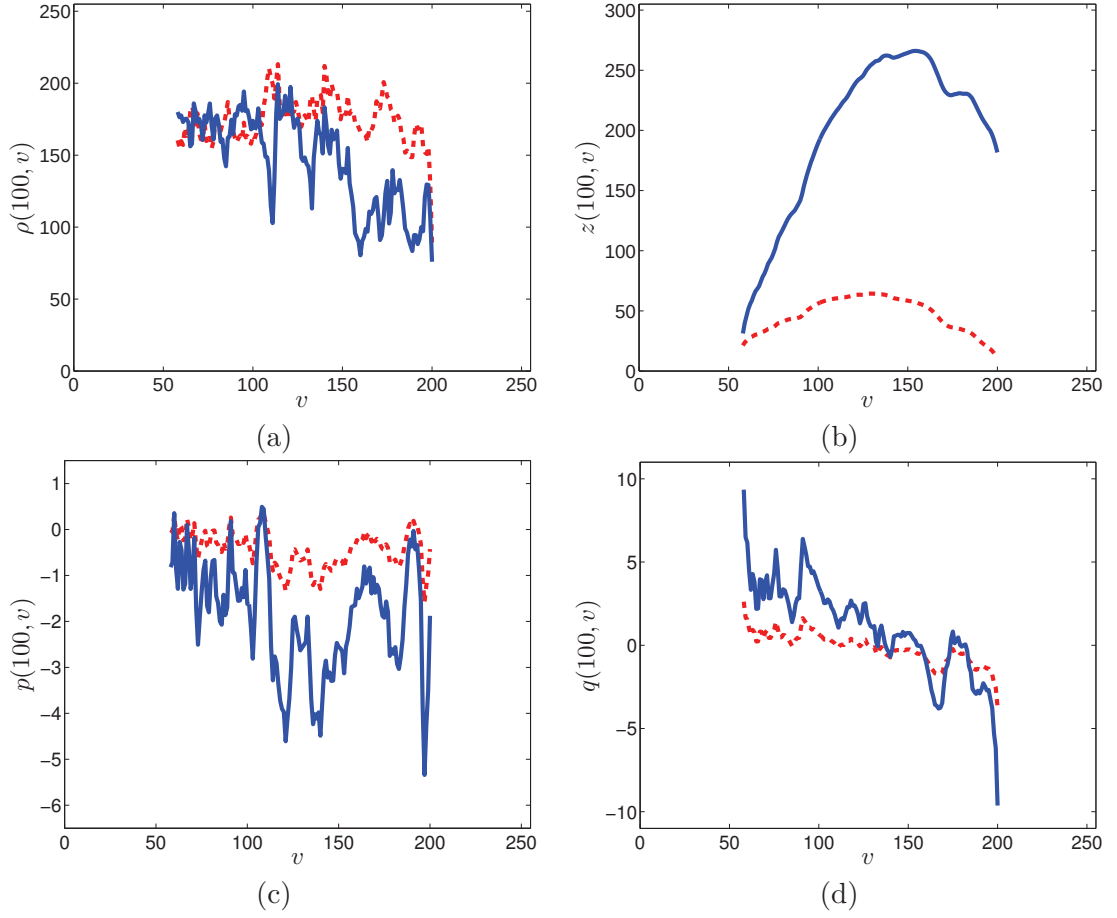


FIGURE 3.3 – Influence d’une transformation GBR sur les variations de l’albédo ρ , de la profondeur z et de son gradient ∇z . Nous représentons ici les valeurs de (a) ρ , (b) z , (c) p et (d) q , le long de la ligne $u = 100$, pour les valeurs de v telles que $(100, v) \in \Omega$, issues des estimations de la figure 3.1. Il apparaît clairement que les courbes en trait plein bleu (après transformation GBR) présentent des variations bien plus marquées que les courbes en pointillés rouges (avant transformation GBR).

Il semble donc que la variation totale de l’albédo ne permette pas de lever l’ambiguïté BR de manière fiable, même si elle garantit l’unicité de la solution. Le moyen le plus simple pour lever l’ambiguïté autrement consiste à adapter une autre méthode de résolution de l’ambiguïté GBR en la simplifiant : au lieu d’estimer les trois paramètres (μ, ν, λ) de l’ambiguïté GBR, nous n’utilisons cette autre méthode que pour estimer λ . En pratique, nous utilisons soit la méthode de Alldrin et al., qui repose sur la minimisation de l’entropie [9], soit une autre méthode que nous décrivons dans le paragraphe 3.2.5, qui suppose que les intensités des éclairages sont égales [114, 214, 282].

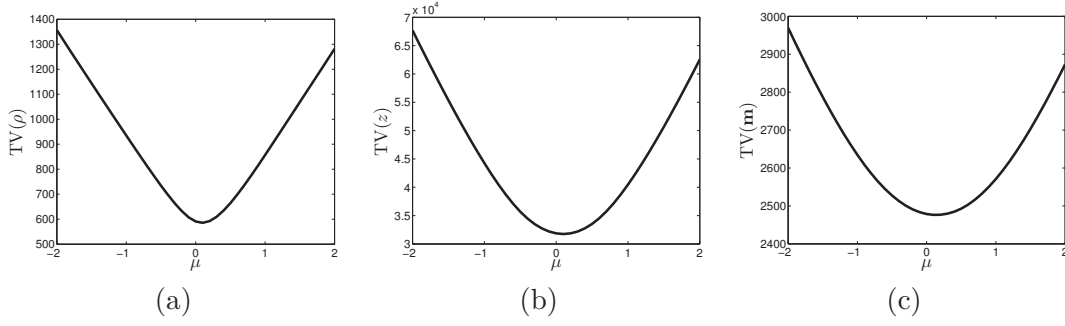


FIGURE 3.4 – Influence du paramètre μ sur les variations totales (a) de l'albédo ρ , (b) de la profondeur z et (c) du champ $\mathbf{m}$ calculés en appliquant une transformation GBR à la reconstruction du buste de Beethoven par stéréophotométrie calibrée (figure 3.1). Les trois variations totales sont minimales pour des valeurs très proches de la valeur théorique $\mu = 0$.

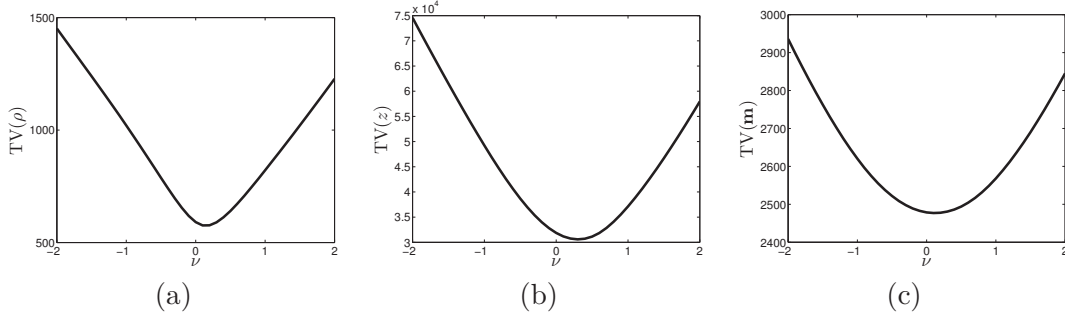


FIGURE 3.5 – Influence du paramètre ν sur les variations totales (a) de l'albédo ρ , (b) de la profondeur z et (c) du champ $\mathbf{m}$, sur le même jeu de données que la figure 3.4. Comme pour μ , les trois variations totales sont minimales pour des valeurs très proches de la valeur théorique $\nu = 0$.

L'étude de la variation totale $TV(\nabla z)$ du gradient de profondeur est laissée en perspective. Pourtant, la minimisation de $TV(\nabla z)$, couplée à celle de $TV(\rho)$, permettrait de contrôler implicitement les variations du champ $\mathbf{m}$. En effet, comme nous l'avons déjà montré dans [217], les variations totales de ρ , de ∇z et de $\mathbf{m}$ vérifient l'inégalité suivante :

$$TV(\mathbf{m}) \leqslant \underbrace{3 \iint_{\Omega} \|\nabla \rho(u, v)\| du dv}_{TV(\rho)} + \underbrace{4 \int_{\Omega} \rho(u, v) (\|\nabla p(u, v)\| + \|\nabla q(u, v)\|) du dv}_{\leqslant \rho_{\max} TV(\nabla z)} \quad (3.29)$$

en notant $\rho_{\max}$ la valeur maximale de l'albédo. Remarquons au passage que, dans l'in-

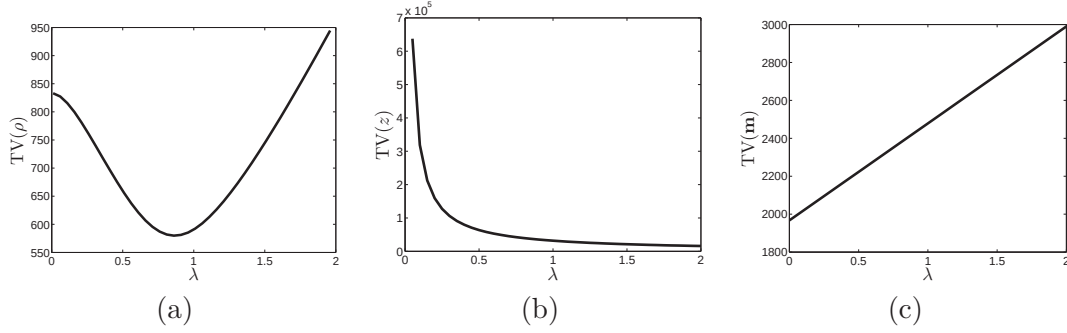


FIGURE 3.6 – Influence du paramètre λ sur la variation totale (a) de l'albédo ρ , (b) de la profondeur z et (c) du champ $\mathbf{m}$, sur le même jeu de données que la figure 3.4. Seule la variation totale de l'albédo admet un minimum dans $]0; +\infty[$, alors que $TV(z)$ est inversement proportionnel à λ et que $TV(\mathbf{m})$ est proportionnel à λ .

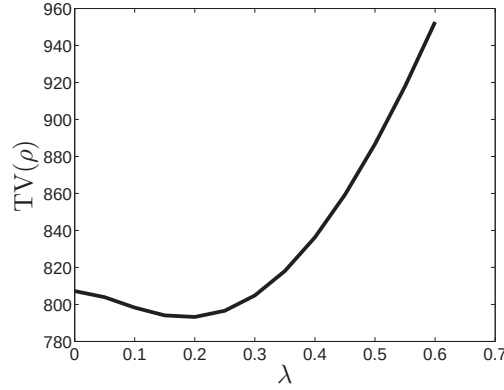


FIGURE 3.7 – Variation totale de l'albédo en fonction de λ , sur le jeu de données *Doll*. La valeur de λ qui minimise $TV(\rho)$ est très éloignée de sa valeur réelle $\lambda = 1$.

égalité (3.29), $\iint_{\Omega} \rho (\|\nabla p\| + \|\nabla q\|)$ constitue une pondération de la variation totale de ∇z par l'albédo. Cette pondération permet de « relâcher » la contrainte sur la variation totale lorsque l'albédo de la surface est faible, i.e. là où le matériau est sombre. Sachant que si $\rho = 0$, le problème UPS est mal posé, cette pondération semble intéressante pour limiter le biais de la reconstruction 3D.

Mais si l'utilisation de $TV(\nabla z)$ constitue une piste intéressante, elle comporte toutefois une difficulté numérique qui lui est spécifique. En effet, la minimisation de $TV(\nabla z)$ revient à une hypothèse de régularité des dérivées d'ordre 2 de la profondeur z . La condition d'optimalité correspondante serait donc d'ordre 4.

Nous allons maintenant montrer comment résoudre le problème UPS au moyen de la régularisation TV. Nous cherchons pour cela, parmi toutes les solutions en $(\mathbf{S}, \mathbf{M})$ du problème (3.8), celle qui minimise un critère de type TV.

3.2.3 Modèle proposé

Intéressons-nous d'abord à la régularisation $\text{TV}(\mathbf{m})$. Une première approche, inspirée par les modèles de débruitage d'images, consisterait à modifier le problème (3.7) de la façon suivante :

$$\begin{cases} \min_{\mathbf{S}, \mathbf{m}} \|\mathbf{S}\mathbf{m} - \mathbf{i}\|_{L^2(\Omega)}^2 + \gamma \text{TV}(\mathbf{m}) \\ \text{s.c. } \overline{\text{rot}} \mathbf{m} = 0 \end{cases} \quad (3.30)$$

où $\gamma > 0$. Un tel problème serait par exemple utile pour lisser les artéfacts dus aux écarts au modèle lambertien linéaire, mais il ne permettrait pas de lever l'ambiguïté BR. Nous préférons pour l'instant coller au plus près du modèle lambertien linéaire (puisque les écarts au modèle sont négligés) et donc continuer à minimiser le seul terme d'attache aux données. De plus, l'opérateur $\overline{\text{rot}} \mathbf{m}$ étant non linéaire, la contrainte $\overline{\text{rot}} \mathbf{m} = 0$ est difficile à imposer. Nous reviendrons dans la partie IV sur cette approche, et montrerons qu'en choisissant comme inconnues le triplet $(\mathbf{S}, \rho, z)$ plutôt que le couple $(\mathbf{S}, \mathbf{m})$, une solution variationnelle du problème UPS peut être déterminée, la contrainte d'intégrabilité étant alors *implicitement* imposée.

Pour cette étude plus « théorique » sur le caractère bien posé du problème UPS, nous supposons que les données sont conformes au modèle, à un bruit blanc gaussien près. Nous pouvons donc faire tendre γ vers 0 dans (3.30), ce qui revient à chercher, parmi toutes les solutions qui minimisent l'attache aux données $\|\mathbf{S}\mathbf{m} - \mathbf{i}\|_{L^2(\Omega)}^2$, celle dont la variation totale est minimale. Cela se réécrit de la façon suivante :

$$\begin{cases} \min_{\mathbf{m}} \text{TV}(\mathbf{m}) \\ \text{s.c. } \begin{cases} \|\mathbf{S}\mathbf{m} - \mathbf{i}\|_{L^2(\Omega)}^2 = \min \\ \overline{\text{rot}} \mathbf{m} = 0 \end{cases} \end{cases} \quad (3.31)$$

où l'ensemble admissible des couples $(\mathbf{S}, \mathbf{m})$ satisfaisant la contrainte est donné, dans le cas discret, par (3.16).

Pour résoudre un problème de la forme (3.31), nous exploitons la connaissance de toutes les solutions de l'équation (3.8), dont nous savons qu'elles sont liées par la transformation GBR. Nous proposons une solution très simple, fondée sur la solution explicite de Yuille et Snow [282, 283] et l'estimation des paramètres GBR par minimisation de la variation totale. En effet, nous avons montré dans [217] que le problème (3.31) était strictement équivalent à :

$$(\hat{\mathbf{S}}, \hat{\mathbf{m}}) = (\mathbf{S}^0 \mathbf{G}(\hat{\mu}, \hat{\nu}, \hat{\lambda}), \mathbf{G}(\hat{\mu}, \hat{\nu}, \hat{\lambda})^{-1} \mathbf{m}^0) \quad (3.32)$$

où $(\mathbf{S}^0, \mathbf{m}^0)$ est la solution explicite fournie par la méthode de Yuille et Snow (cf. paragraphe 3.1.3) et :

$$(\hat{\mu}, \hat{\nu}, \hat{\lambda}) = \underset{\substack{(\mu, \nu) \in \mathbb{R}^2 \\ \lambda \in \mathbb{R}^+ \setminus \{0\}}}{\text{argmin}} \text{TV}(\mathbf{G}(\mu, \nu, \lambda)^{-1} \mathbf{m}^0) \quad (3.33)$$

qui est une variante du deuxième sous-problème de (3.21).

Détaillons maintenant la résolution du problème (3.33). D'après la définition (3.26) de la variation totale et l'expression (3.18) d'une transformation GBR, sa fonction objectif s'écrit :

$$\begin{aligned}
 \mathcal{E}(\mu, \nu, \lambda) &= \text{TV}(\mathbf{G}(\mu, \nu, \lambda)^{-1} \mathbf{m}^0) \\
 &= \text{TV}([1, 0, \mu] \mathbf{m}^0) + \text{TV}([0, 1, \nu] \mathbf{m}^0) + \text{TV}([0, 0, \lambda] \mathbf{m}^0) \\
 &= \text{TV}(m_1^0 + \mu m_3^0) + \text{TV}(m_2^0 + \nu m_3^0) + \text{TV}(\lambda m_3^0)
 \end{aligned} \tag{3.34}$$

Les estimations des trois paramètres (μ, ν, λ) peuvent donc être découplées. Ce découplage, qui est une conséquence du choix de la définition anisotrope (3.26) de la TV vectorielle, ne se serait pas produit avec la définition isotrope (3.25).

3.2.4 Estimation de μ et ν

On montre facilement, en différenciant la fonction $\mathcal{E}(\mu, \nu, \cdot)$, sa convexité vis-à-vis de μ et ν , et donc l'unicité de la solution en (μ, ν) . D'un point de vue pratique, vu cette convexité, nous utilisons une méthode de descente de type « quasi-Newton », initialisée par la solution du problème quadratique suivant :

$$\begin{cases} \hat{\mu}_0 = \underset{\mu}{\operatorname{argmin}} \iint_{\Omega} \|\nabla(m_1^0(u, v) + \mu m_3^0(u, v))\|^2 du dv \\ \hat{\nu}_0 = \underset{\nu}{\operatorname{argmin}} \iint_{\Omega} \|\nabla(m_2^0(u, v) + \nu m_3^0(u, v))\|^2 du dv \end{cases} \tag{3.35}$$

qu'il est facile de calculer explicitement :

$$\begin{cases} \hat{\mu}_0 = -\frac{\iint_{\Omega} \nabla m_1^0(u, v)^{\top} \nabla m_3^0(u, v) du dv}{\iint_{\Omega} \|\nabla m_3^0(u, v)\|^2 du dv} \\ \hat{\nu}_0 = -\frac{\iint_{\Omega} \nabla m_2^0(u, v)^{\top} \nabla m_3^0(u, v) du dv}{\iint_{\Omega} \|\nabla m_3^0(u, v)\|^2 du dv} \end{cases} \tag{3.36}$$

Intéressons-nous maintenant à l'estimation de λ , qui influe sur l'échelle en z de la reconstruction 3D (ambiguïté BR).

3.2.5 Estimation de λ

Commençons par calculer le troisième terme de l'expression (3.34) de la fonction objectif $\mathcal{E}(\mu, \nu, \lambda)$:

$$\text{TV}(\lambda m_3^0) = \iint_{\Omega} \|\nabla(\lambda m_3^0(u, v))\| du dv \tag{3.37}$$

$$= \lambda \text{TV}(m_3^0) \tag{3.38}$$

La dépendance linéaire de $\mathcal{E}(\mu, \nu, \lambda)$ en λ explique a posteriori l'allure de la courbe de la figure 3.6-c. Si l'on cherche à minimiser $\mathcal{E}(\mu, \nu, \lambda)$ et que l'on veut éviter la solution $\lambda = 0^+$, qui correspond à un relief dégénéré (cf. (3.20)), alors il est nécessaire de remplacer

la contrainte $\lambda > 0$ par $\lambda \geq \epsilon$, avec $\epsilon > 0$. Sous cette contrainte, la valeur de λ qui minimise $\mathcal{E}(\mu, \nu, \lambda)$ est bien sûr :

$$\hat{\lambda} = \epsilon \quad (3.39)$$

Le choix de la valeur de ϵ est donc important en pratique, puisque ce choix détermine l'échelle, i.e. lève l'ambiguïté BR. Un tel choix peut paraître très arbitraire. Néanmoins, comme la minimisation en (μ, ν) est indépendante du choix de ϵ , l'ambiguïté BR peut être levée a posteriori, quitte à fixer arbitrairement ϵ dans un premier temps. À la manière d'un hyper-paramètre, ϵ peut donc être choisi manuellement pour obtenir une surface « acceptable ».

Nous proposons donc de résoudre les deux problèmes suivants en séquence :

$$\begin{cases} (\hat{\mu}, \hat{\nu}) = \underset{\mu, \nu}{\operatorname{argmin}} \operatorname{TV}(\mathbf{G}(\mu, \nu, 1)^{-1} \mathbf{m}^0) \\ \hat{\lambda} = \underset{\lambda \geq \epsilon}{\operatorname{argmin}} \operatorname{TV}(\mathbf{G}(\hat{\mu}, \hat{\nu}, \lambda)^{-1} \mathbf{m}^0) \end{cases} \quad (3.40)$$

en fixant empiriquement la valeur de ϵ . Cette formulation du problème permet de comprendre pourquoi l'estimation de λ correspond à la résolution de l'ambiguïté BR puisque, de la définition (3.18) de $\mathbf{G}(\mu, \nu, \lambda)^{-1}$, on tire facilement :

$$\mathbf{G}(\hat{\mu}, \hat{\nu}, \lambda)^{-1} \mathbf{m}^0 = \mathbf{G}(0, 0, \lambda)^{-1} (\mathbf{G}(\hat{\mu}, \hat{\nu}, 1)^{-1} \mathbf{m}^0) \quad (3.41)$$

et que $\mathbf{G}(0, 0, \lambda)$ est une matrice de transformation BR [28]. L'expression (3.41) correspond donc à une transformation BR du vecteur :

$$\mathbf{m}^1 = \mathbf{G}(\hat{\mu}, \hat{\nu}, 1)^{-1} \mathbf{m}^0 \quad (3.42)$$

Il s'ensuit que l'influence de λ est très facile à interpréter : $\lambda \rightarrow +\infty$ tend à « aplatiser » le relief tout en accentuant les variations d'albédo (principe du « trompe-l'œil ») ; inversement, $\lambda \rightarrow 0^+$ accentue les pentes tout en réduisant les variations d'albédo. Plus formellement, λ influe sur la profondeur z de la façon suivante, d'après (3.20) :

$$\bar{z}(u, v) = \frac{z(u, v)}{\lambda} \quad (3.43)$$

tandis que, comme cela est montré dans [28], ce paramètre influe sur l'albédo ρ comme suit :

$$\bar{\rho}(u, v) = \rho(u, v) \left(\frac{\lambda^2 + \|\nabla z(u, v)\|^2}{1 + \|\nabla z(u, v)\|^2} \right)^{1/2} \quad (3.44)$$

La figure 3.8 montre l'effet d'une transformation BR sur le relief et l'albédo du buste de Beethoven.⁴

Pour résoudre *effectivement* l'ambiguïté BR, nous avons utilisé deux méthodes de l'état de l'art : la méthode de minimisation de l'entropie de l'albédo [9], qui a déjà

4. D'autres liens entre la géométrie d'une surface et son albédo, sous l'effet d'une transformation BR, sont détaillés dans [147].

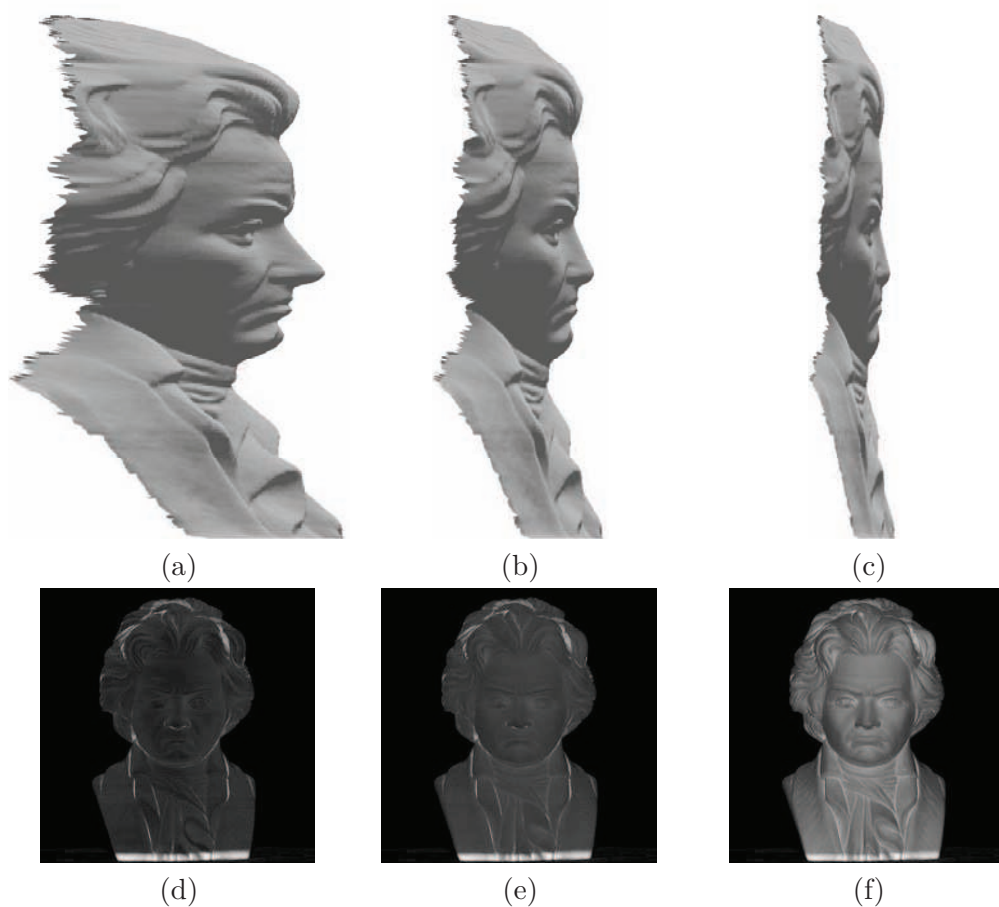


FIGURE 3.8 – Effet d’une transformation de *bas-relief* : reliefs estimés avec (a) $\lambda = 1/2$, (b) $\lambda = 1$ (relief original) et (c) $\lambda = 2$; (d)-(e)-(f) albédos correspondant aux mêmes valeurs de λ .

été mentionnée à plusieurs reprises, et une méthode fondée sur l’hypothèse d’égalité des intensités des éclairages, introduite dans [114, 282], que nous avons étudiée dans [209, 214] et que nous allons détailler maintenant. Étant donné que ces méthodes utilisent des a priori de natures très différentes, on ne peut pas dire que l’une soit supérieure à l’autre : cela dépend des données.

Imposer l’égalité des intensités des éclairages comme a priori pour lever l’ambiguïté GBR est relativement difficile [282]. En revanche, cette contrainte se traduit très simplement, dans le cas de l’ambiguïté BR, par un système linéaire de m équations à deux inconnues. Le problème consiste à trouver le paramètre λ tel que :

$$\|\mathbf{s}^{0,i\top} \mathbf{G}(\hat{\mu}, \hat{\nu}, \lambda)\|^2 = \phi_0^2, \quad \forall i \in [1, m] \quad (3.45)$$

où l'intensité ϕ_0 est inconnue et $\mathbf{s}^{0,i\top}$ constitue la i -ème ligne de $\mathbf{S}^0$ (cf. (3.1)). Or :

$$\mathbf{G}(\hat{\mu}, \hat{\nu}, \lambda) = \mathbf{G}(\hat{\mu}, \hat{\nu}, 1) \mathbf{G}(0, 0, \lambda) \quad (3.46)$$

de telle sorte que, en notant $\mathbf{s}^{1,i\top} = [s_1^{1,i}, s_2^{1,i}, s_3^{1,i}]$ la i -ème ligne de $\mathbf{S}^1 = \mathbf{S}^0 \mathbf{G}(\hat{\mu}, \hat{\nu}, 1)$, la contrainte (3.45) se réécrit :

$$\|\mathbf{s}^{1,i\top} \mathbf{G}(0, 0, \lambda)\|^2 = \phi_0^2, \quad \forall i \in [1, m] \quad (3.47)$$

soit, d'après la forme (3.17) d'une transformation GBR :

$$\left(s_1^{1,i}\right)^2 + \left(s_2^{1,i}\right)^2 + \frac{1}{\lambda^2} \left(s_3^{1,i}\right)^2 = \phi_0^2 \quad (3.48)$$

qui est une équation linéaire à deux inconnues scalaires $1/\lambda^2$ et ϕ_0^2 . Ces inconnues peuvent être estimées, au sens des moindres carrés, par résolution du système des m équations de la forme (3.48). De là, on déduit λ^2 , puis λ sachant que $\lambda > 0$ par hypothèse.

3.2.6 Régularisation TV de la profondeur

Nous venons de voir comment ramener l'ambiguïté GBR à une ambiguïté BR, en minimisant la variation totale du champ $\mathbf{m}$. Cette méthode nécessite toutefois de calculer la jacobienne de $\mathbf{m}$, qui peut être approchée par différences finies. Afin d'éviter tout problème numérique dû au bruit ou à la présence de discontinuités, il est nécessaire d'appliquer un filtrage gaussien au champ $\mathbf{m}^0$ avant de calculer les différences finies⁵. On peut éviter le recours à de telles approximations numériques en préférant régulariser la profondeur z . Comme nous allons le voir, cela permet effectivement d'éviter toute approximation numérique.

Rappelons que, si la profondeur z est supposée deux fois différentiable (au moins par morceaux), alors la condition d'intégrabilité est implicitement satisfaite. Le problème (3.31) peut donc être reformulé comme la recherche d'une fonction z régulière solution de :

$$\begin{cases} \min_z \text{TV}(z) \\ \text{s.c.} \quad \left\| \rho \mathbf{S} \frac{[-\partial_u z, -\partial_v z, 1]^\top}{\sqrt{\|\nabla z\|^2 + 1}} - \mathbf{i} \right\|_{L^2(\Omega)}^2 = \min \end{cases} \quad (3.49)$$

Ce type de problème variationnel sera étudié dans la partie IV. Pour l'instant, nous exploitons à nouveau la forme explicite de l'ambiguïté GBR.

Soit z^0 la profondeur obtenue par intégration du champ de normales $\mathbf{n}^0 = \mathbf{m}^0 / \|\mathbf{m}^0\|$ estimé par la méthode de Yuille et Snow (il n'est pas nécessaire de calculer explicitement cette profondeur). Cette solution appartient à l'ensemble admissible du problème (3.49), les autres solutions admissibles étant reliées à z^0 par l'équation (3.20). En notant $\mathbf{u}^0 = [u, v, z^0(u, v)]^\top$ le carreau de Monge correspondant à z^0 , l'ensemble des fonctions z admissibles s'écrit donc :

$$\left\{ \mathbf{g}(\mu, \nu, \lambda) \mathbf{u}^0, (\mu, \nu, \lambda) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ \setminus \{0\} \right\} \quad (3.50)$$

5. La même « astuce » est utilisée dans [9] pour imposer l'intégrabilité.

où le vecteur $\mathbf{g}(\mu, \nu, \lambda) = \frac{1}{\lambda}[-\mu, -\nu, 1]^\top$, qui dépend des paramètres GBR, correspond à la troisième colonne de $\mathbf{G}(\mu, \nu, \lambda)$.

La solution de (3.49) s'écrit donc :

$$\hat{z} = \mathbf{g}(\hat{\mu}, \hat{\nu}, \hat{\lambda}) \mathbf{u}^0 \quad (3.51)$$

où :

$$(\hat{\mu}, \hat{\nu}, \hat{\lambda}) = \underset{\substack{(\mu, \nu) \in \mathbb{R}^2 \\ \lambda \in \mathbb{R}^+ \setminus \{0\}}}{\operatorname{argmin}} \operatorname{TV}(\mathbf{g}(\mu, \nu, \lambda) \mathbf{u}^0) \quad (3.52)$$

Or, le gradient de la fonction de profondeur transformée $\bar{z} = \mathbf{g}(\mu, \nu, \lambda) \mathbf{u}^0$ s'écrit :

$$\nabla \bar{z} = -\frac{1}{\lambda} \begin{bmatrix} \frac{m_1^0}{m_3^0} + \mu \\ \frac{m_2^0}{m_3^0} + \nu \end{bmatrix} \quad (3.53)$$

L'estimation de λ est donc à nouveau problématique, puisqu'elle aboutit à la solution $\lambda = +\infty$, ce qui explique a posteriori l'allure de la courbe de la figure 3.6-b. Comme nous l'avons déjà fait dans le paragraphe 3.2.5, nous imposons une contrainte, en l'occurrence $\lambda \leq \epsilon'$, avec $\epsilon' > 0$, pour éviter cette solution dégénérée. On montre aisément [217] que l'estimation de μ et ν est indépendante de celle de λ , i.e. que l'on peut fixer dans un premier temps $\lambda = 1$ pour estimer μ et ν , avant de lever l'ambiguïté BR résiduelle en minimisant soit la variation totale de l'albédo, soit l'entropie de l'albédo, ou encore en imposant l'égalité des intensités des éclairages.

L'estimation de μ et ν constitue un problème convexe que nous résolvons par une méthode de type quasi-Newton, à partir de la solution $(\hat{\mu}_0, \hat{\nu}_0)$ du problème quadratique :

$$(\hat{\mu}_0, \hat{\nu}_0) = \underset{\mu, \nu}{\operatorname{argmin}} \iint_{\Omega} \left\| \nabla(\mathbf{g}(\mu, \nu, 1) \mathbf{u}^0(u, v)) \right\|^2 du dv \quad (3.54)$$

qu'il est facile de calculer explicitement :

$$\begin{cases} \hat{\mu}_0 = \bar{p}^0 \\ \hat{\nu}_0 = \bar{q}^0 \end{cases} \quad (3.55)$$

où $\bar{p}^0$ et $\bar{q}^0$ désignent les valeurs moyennes de $p^0 = -m_1^0/m_3^0$ et de $q^0 = -m_2^0/m_3^0$. En réalité, résoudre (3.54) revient à forcer le champ de normales $\mathbf{n}$ à être égal, en moyenne, à $[0, 0, 1]^\top$, i.e. orienté vers la caméra, ce qui paraît raisonnable : cette contrainte a également été utilisée comme a priori dans un algorithme EM de résolution robuste de la stéréophotométrie calbrée [257].

En résumé, minimiser la variation totale de la profondeur z ou du champ $\mathbf{m}$ permet de réduire l'ambiguïté GBR à une ambiguïté BR, qui peut être levée a posteriori, par exemple en minimisant la variation totale de l'albédo ou l'entropie de l'albédo, ou encore en imposant l'égalité des intensités des éclairages. Nous allons voir qu'en plus de l'intérêt théorique de cette étude, qui suggère en particulier une résolution variationnelle du problème UPS (cf. partie IV), l'approche proposée fournit la plupart du temps de meilleurs résultats, qualitativement et quantitativement, que les méthodes de l'état de l'art.

3.3 Validation expérimentale

Dans ce paragraphe, nous évaluons la précision des reconstructions 3D obtenues en levant l'ambiguïté du problème UPS par variation totale. Certains auteurs [89] évaluent cette précision en calculant l'*écart angulaire moyen* (EAM) entre le champ de normales estimé et celui obtenu par stéréophotométrie calibrée. Alldrin note cependant dans [9] que les résultats obtenus dans le contexte non calibré pourraient être meilleurs que ceux de la stéréophotométrie calibrée, étant donné que le recours à un a priori peut compenser un étalonnage approximatif des éclairages. Nous menons d'abord l'évaluation sur des jeux de données synthétiques, pour lesquels les normales réelles sont connues.

3.3.1 Résultats sur données synthétiques

Pour évaluer la qualité de l'estimation de μ et ν par régularisation TV, nous utilisons les images de synthèse d'un vase, simulées de la façon suivante :

- La profondeur $z(u, v)$ étant une fonction polynomiale de u et v (cf., entre autres, [82]), l'expression analytique des normales est facile à établir. La profondeur z est représentée sur la figure 3.9-a, et ses dérivées p et q sur les figures 3.9-b et 3.9-c.
- La matrice $\mathbf{S}$, de taille 22×3 ($m = 22$), est constituée de vecteurs dont les directions sont uniformément réparties sur l'hémisphère nord de la sphère unité, et dont les intensités sont toutes égales.
- La carte d'albédo est choisie arbitrairement (voir plus bas).
- Enfin, les $m = 22$ images, de taille 256×256 , sont générées à partir du modèle lambertien, avec prise en compte des ombres propres : $\mathbf{i} = \max\{\mathbf{S}\mathbf{m}, 0\}$.

Puisque les intensités des éclairages sont parfaitement égales pour ces données synthétiques, c'est bien sûr cette hypothèse qui est utilisée pour lever l'ambiguïté BR résiduelle, i.e. pour estimer λ . Les méthodes d'estimation des paramètres GBR seront donc notées TV- $\mathbf{m}$ -MI et TV- z -MI ($\mathbf{m}$ ou z indique quelle variation totale est utilisée pour estimer μ et ν , et MI, pour *même intensité*, indique la méthode utilisée pour estimer λ).

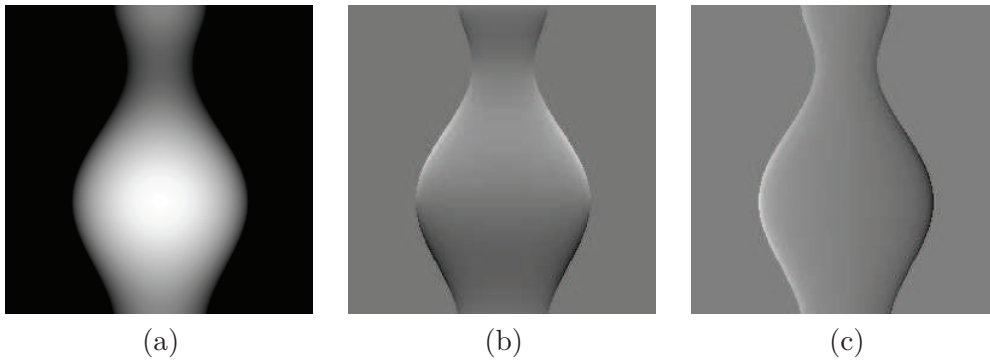


FIGURE 3.9 – Données synthétiques (a) z , (b) p et (c) q utilisées dans les tests.

Nous comparons nos reconstructions 3D à celles obtenues en utilisant les méthodes [9] et [89], dont les codes Matlab sont disponibles en ligne. Nos codes ont également été développés en Matlab, et les temps CPU sont mesurés sur un processeur I7 à 2,9 GHz. Rappelons que :

- La méthode du *minimum de l'entropie* (ME) [9] estime les paramètres de la transformation GBR en minimisant l'entropie de l'albédo.
- La méthode des *maxima diffus* (MD) [89, 196] suppose que les points singuliers de l'image, i.e. ceux où le niveau de gris est maximal, sont localement orientés vers la source lumineuse. Il est alors possible d'en déduire analytiquement les paramètres de la transformation GBR.

Les caractéristiques de ces méthodes étant très différentes, nous les avons comparées sur plusieurs exemples :

- Cas 1 : lorsque l'albédo est uniforme (cf. figure 3.10-a), l'entropie est très faible, et l'orientation des maxima diffus correspond effectivement à la direction d'éclairage. Ce cas devrait donc être favorable aux méthodes ME [9] et MD [89].
- Cas 2 : lorsque l'albédo est uniforme par morceaux (cf. figure 3.10-b), l'entropie reste faible, mais certains maxima diffus sont détectés, à tort, sur des zones d'albédo élevé. Ce cas de figure devrait donc favoriser la méthode ME [9].
- Cas 3 : lorsque l'albédo suit une distribution de type gradient radial (cf. figure 3.10-c), l'entropie est élevée, et les maxima d'intensité peuvent à nouveau correspondre à des valeurs élevées de l'albédo.

La figure 3.10 montre une image de chacun de ces jeux de données.

Pour évaluer la précision des différentes méthodes, nous calculons l'erreur relative entre les paramètres GBR estimés et leurs valeurs réelles, ainsi que l'erreur angulaire entre le champ de normales estimé et le champ de normales réel. Tous ces résultats sont consignés dans la table 3.1.

Comme on pouvait s'y attendre, les méthodes ME et MD procurent de très bons résultats dans le cas 1, mais sont toutes deux mises en échec lorsque l'albédo varie continûment (cas 3). Au contraire, la régularisation TV se comporte aussi bien dans les trois cas. Comme cette régularisation ne repose sur l'identification d'aucun point particulier, cela la rend plus générale (au moins pour l'estimation de μ et ν).

La méthode ME utilise une optimisation globale pour minimiser l'entropie de la distribution de l'albédo, qui est non convexe. Ceci se traduit par des temps CPU très élevés. Pour la méthode MD, le temps de calcul est très difficile à prévoir, puisqu'il dépend du nombre de maxima diffus détectés dans les images. Quant aux méthodes TV proposées, elles procèdent à une optimisation convexe à trois paramètres et sont donc très rapides.

Commentons enfin les légères différences entre les résultats obtenus avec les méthodes TV-**m**-MI et TV-**z**-MI : TV-**m**-MI nécessite de calculer une jacobienne par approximations numériques, ce qui est complètement biaisé par la non intégrabilité des normales sur le bord du vase. C'est pour cette raison que la méthode TV-**z**-MI fournit de meilleurs résultats. Si l'on applique un filtrage gaussien avant les approximations par différences

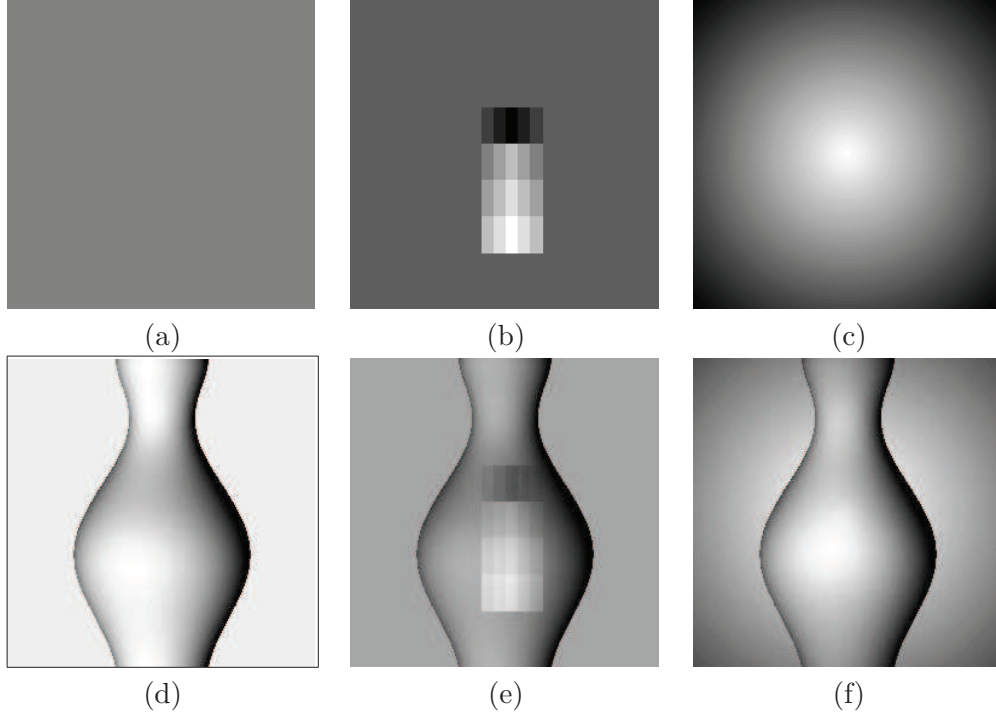


FIGURE 3.10 – Différents types d'albédos : (a) uniforme, (b) uniforme par morceaux, (c) de type gradient radial. (d)-(e)-(f) Images correspondant à ces différents albédos.

finies, ce biais est toutefois réduit. Cette version modifiée de TV- $\mathbf{m}$ -MI est notée TV- $\mathbf{m}$ -MI-Gauss dans la table 3.1. Étant donné que ce filtrage améliore sensiblement les résultats, il sera systématiquement appliqué, et TV- $\mathbf{m}$ -MI fera dorénavant référence à cette version de la méthode avec filtrage.

3.3.2 Résultats sur données réelles

Pour des données réelles, il est difficile d'évaluer la précision de la reconstruction 3D, en l'absence de vérité terrain. Nous comparons donc nos résultats à ceux obtenus par stéréophotométrie calibrée. Rappelons que cette comparaison est fournie à titre purement indicatif, puisque la reconstruction 3D calibrée est potentiellement biaisée si l'étalonnage des éclairages n'a pas été correctement réalisé.

De plus, rappelons que les méthodes proposées, qui visent à résoudre l'ambiguïté GBR, ne servent qu'à illustrer la pertinence de l'approche par régularisation : en réalité, le résultat majeur de cette étude est de montrer que le problème UPS devient bien posé si l'on a recours à la régularisation TV de la profondeur z (ou du champ $\mathbf{m}$) et à celle de l'albédo ρ . Nous présenterons dans la partie IV des algorithmes variationnels bien plus efficaces pour résoudre le problème UPS.

		$\frac{ \nu - \hat{\nu} }{ \nu }$	$\frac{ \mu - \hat{\mu} }{ \mu }$	$\frac{ \lambda - \hat{\lambda} }{ \lambda }$	EAM	Max-EA	Min-EA	σ_{EA}	CPU
Cas 1	ME [9]	1,54	0,00	0,05	0,63	16,41	0,03	1,10	26,41
	MD [89]	5,36	0,28	0,01	0,98	16,04	0,01	0,95	2,79
	TV- m -MI	93,72	0,27	0,57	10,09	24,77	0,03	3,14	0,13
	TV- m -MI-Gauss	0,61	0,02	0,07	0,57	16,75	0,00	1,28	0,26
	TV- <i>z</i> -MI	0,61	0,02	0,07	0,57	16,75	0,00	1,28	0,57
Cas 2	ME [9]	1,19	0,02	0,04	0,59	16,35	0,00	1,09	28,54
	MD [89]	24,11	1,59	0,11	4,82	14,97	0,06	1,25	3,34
	TV- m -MI	53,38	0,11	0,36	7,42	21,53	0,11	2,11	0,07
	TV- m -MI-Gauss	0,50	0,02	0,07	0,56	16,77	0,00	1,27	0,26
	TV- <i>z</i> -MI	0,50	0,02	0,07	0,56	16,77	0,00	1,27	0,50
Cas 3	ME [9]	8,01	18,22	0,21	15,59	18,89	2,03	3,64	30,82
	MD [89]	8,77	20,74	0,10	15,53	18,40	3,03	3,45	0,56
	TV- m -MI	102,61	0,89	0,20	4,43	18,76	0,07	1,35	0,06
	TV- m -MI-Gauss	2,70	0,03	0,07	0,75	16,53	0,01	1,26	0,15
	TV- <i>z</i> -MI	2,41	0,02	0,07	0,75	16,53	0,01	1,26	0,23

TABLE 3.1 – Comparaison de la méthode du minimum d’entropie (ME), de la méthode des maxima diffus (MD) et des méthodes que nous proposons : TV-**m**-MI, TV-**m**-MI-Gauss (application d’un filtrage gaussien avant l’approximation numérique de la jacobienne de $\mathbf{m}^0$) et TV-*z*-MI. Nous donnons les valeurs de l’erreur relative de l’estimation des paramètres μ , ν et λ , l’écart angulaire moyen sur les normales (EAM), l’écart angulaire maximal (Max-EA), l’écart angulaire minimal (Min-EA), l’écart-type de l’écart angulaire (σ_{EA}) ainsi que le temps CPU (les valeurs angulaires sont exprimées en degrés, les temps CPU en secondes). Les méthodes proposées surpassent celles de l’état de l’art, sont plus rapides, et permettent de prendre en compte des situations où les méthodes de l’état de l’art tendent à échouer, comme dans le cas où l’albédo varie continûment.

Jeux de données

Nous avons utilisé pour nos tests :

- Dix jeux de données d’objets « presque lambertiens ». Les jeux de données *Buddha*, *Cat*, *Horse*, *Owl* et *Rock* fournis par Dan Goldman et Steven Seitz⁶ comportent chacun douze images de taille 320×300 (sauf *Horse*, qui est composé d’images de taille 270×350). Les jeux de données *Octopus*, *Redfish* et *Korean Doll* sont disponibles sur la page personnelle de Neil Alldrin⁷ : ils sont constitués de cinq images, de tailles respectives 321×281 , 351×301 et 1321×521 . Le jeu de données *Doll*, qui a été proposé dans [89], est composé de quinze images de taille 405×250 . Enfin, le jeu de données *Beethoven* est constitué de trois images de taille 256×256 .

6. <http://courses.cs.washington.edu/courses/cse455/10wi/projects/>

7. <http://vision.ucsd.edu/~nalldrin/research/>

- Les vingt premiers jeux de la base de données *Extended Yale DataFace B* [99, 159]. Nous avons sélectionné, parmi les 64 images disponibles pour chaque jeu, les 27 images correspondant aux vecteurs d'éclairage présentant les angles d'élévation les plus faibles, de façon à limiter les ombres. Chaque image a pour taille 192×168 .

Les images réelles ne sont bien entendu pas exemptes d'écarts au modèle, tels que les ombres ou les reflets brillants. Comme cela a été discuté dans le paragraphe 1.2, les approches de type « prétraitements » sont particulièrement adaptées au problème UPS, car leur mise en œuvre est la même si l'on ne connaît pas les éclairages. Nous avons donc appliqué le prétraitement proposé dans [270], impliquant un hyper-paramètre que nous avons fixé aux valeurs suggérées dans [89], i.e. à 1,7 si $m \geq 12$ et à 3 si $m < 12$. La figure 3.11 montre une image prétraitée des quinze premiers jeux de données testés. Bien entendu, le prétraitement n'est pas un outil « magique » : il est impossible de retirer certains effets très marqués, comme par exemple les reflets brillants sur les yeux de la chouette dans le jeu de données *Owl*.

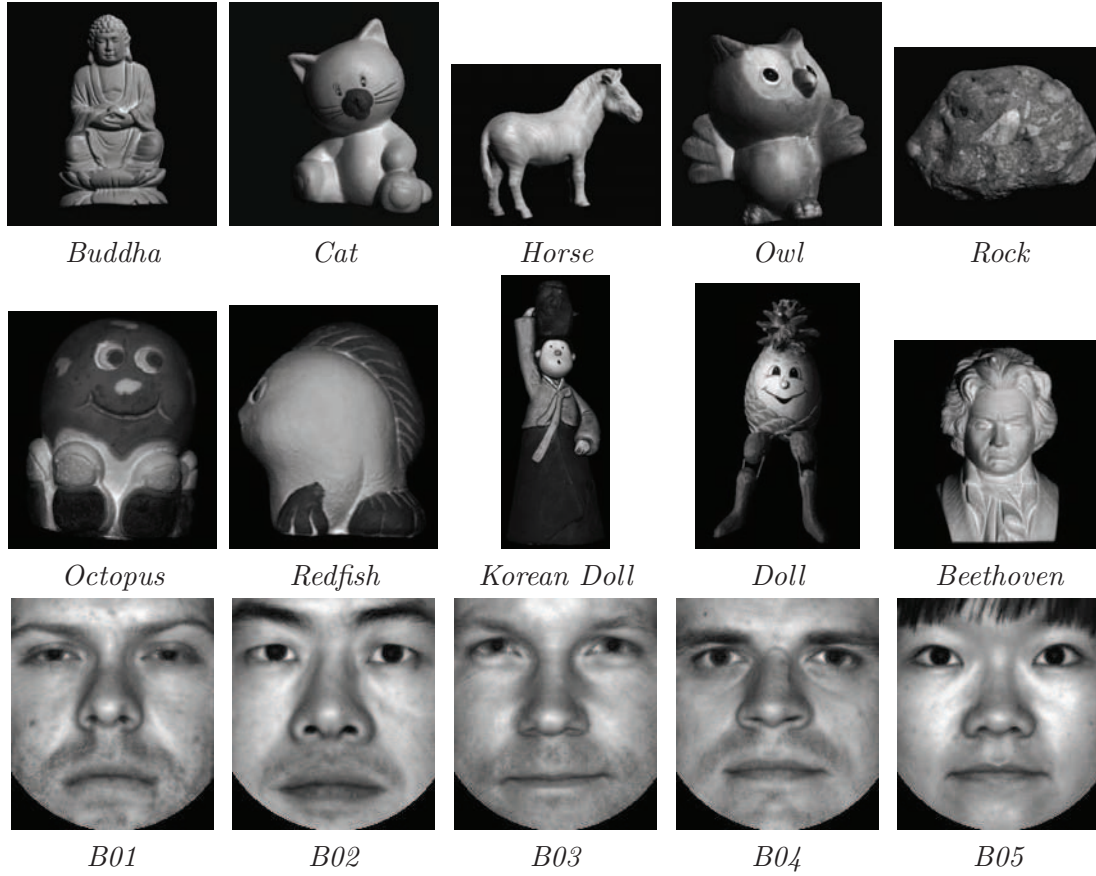


FIGURE 3.11 – Une image (prétraitée) de chacun des quinze premiers jeux de données réelles testés.

Résultats

L'erreur relative sur les paramètres GBR estimés, l'EAM entre les champs de normales estimées par stéréophotométrie non calbrée et par stéréophotométrie calbrée, ainsi que les temps de calcul, sont consignés dans les tables 3.2 et 3.3. Comme on peut le constater, les méthodes proposées atteignent la plupart du temps des performances comparables à l'état de l'art en termes de reconstruction 3D, mais sont bien plus rapides grâce à la convexité des problèmes à résoudre.

	$\frac{ \nu - \hat{\nu} }{ \nu }$	$\frac{ \mu - \hat{\mu} }{ \mu }$	$\frac{ \lambda - \hat{\lambda} }{ \lambda }$	EAM	CPU
<i>Buddha</i>					
ME	0,81	0,22	0,06	6,38	13,87
MD	0,07	0,06	0,12	4,99	1,91
TV-m-MI	0,91	0,35	0,21	10,92	0,06
TV-z-MI	1,03	0,51	1,34	18,85	0,03
TV-z-ME	1,03	0,51	0,00	14,79	0,35
<i>Cat</i>					
ME	0,48	0,49	0,17	14,37	17,82
MD	0,95	0,05	0,08	5,37	0,57
TV-m-MI	1,00	0,10	0,27	6,16	0,06
TV-z-MI	0,13	0,54	5,94	35,57	0,04
TV-z-ME	0,13	0,54	0,18	16,04	0,45
<i>Horse</i>					
ME	11,33	0,28	0,09	10,88	14,21
MD	14,25	0,20	0,03	4,80	1,25
TV-m-MI	11,64	0,11	0,21	6,68	0,07
TV-z-MI	34,83	0,34	0,64	18,25	0,03
TV-z-ME	34,83	0,34	0,13	15,59	0,34
<i>Owl</i>					
ME	0,41	0,63	0,25	15,67	22,71
MD	0,35	0,00	0,07	6,63	0,43
TV-m-MI	0,50	0,23	0,13	8,42	0,08
TV-z-MI	0,15	0,13	0,08	5,54	0,05
TV-z-ME	0,15	0,13	0,15	7,54	0,54
<i>Rock</i>					
ME	14,77	0,75	0,23	18,37	32,30
MD	5,35	0,39	0,05	11,61	0,61
TV-m-MI	3,13	0,88	-	-	0,08
TV-z-MI	9,43	1,19	-	-	0,07
TV-z-ME	9,43	1,19	0,19	27,56	0,76
<i>Octopus</i>					
ME	0,06	0,07	0,07	3,06	22,49
MD	0,10	0,16	0,19	6,64	0,26
TV-m-MI	0,03	0,14	-	-	0,07
TV-z-MI	0,11	0,29	-	-	0,07
TV-z-ME	0,11	0,29	0,28	9,41	0,67
<i>Redfish</i>					
ME	7,06	0,06	0,03	4,38	21,52
MD	8,85	0,18	0,05	5,60	0,79
TV-m-MI	20,83	0,06	3,48	25,92	0,10
TV-z-MI	3,28	3,75	2,93	26,07	0,05
TV-z-ME	3,28	3,75	0,17	14,17	0,61
<i>Korean Doll</i>					
ME	0,04	0,19	0,08	3,39	151,99
MD	0,20	0,80	0,27	17,74	2,28
TV-m-MI	0,11	0,46	-	-	0,49
TV-z-MI	0,09	0,50	-	-	0,34
TV-z-ME	0,09	0,50	0,15	10,36	3,87
<i>Doll</i>					
ME	0,05	2,64	0,47	26,24	17,13
MD	0,03	0,81	0,34	12,15	0,59
TV-m-MI	0,05	1,81	-	-	0,06
TV-z-MI	0,04	1,82	-	-	0,03
TV-z-ME	0,04	1,82	0,44	19,84	0,35
<i>Beethoven</i>					
ME	10,50	6,69	0,57	63,83	16,39
MD	0,02	0,25	0,41	10,01	1,31
TV-m-MI	0,15	0,34	0,65	16,44	0,05
TV-z-MI	0,06	0,31	0,47	13,70	0,03
TV-z-ME	0,06	0,31	0,07	13,30	0,32

TABLE 3.2 – Comparaison des résultats obtenus par les méthodes ME [9], MD [89] et les méthodes utilisant la régularisation TV, sur dix objets « presque » lambertiens. Les valeurs angulaires sont exprimées en degrés, les temps CPU en secondes.

Il est notable que, pour certains jeux de données, aucune erreur n'est indiquée dans la table 3.2 : il s'agit de cas où la résolution des équations (3.48) fournit une solution négative pour λ^2 , ce qui indique que la contrainte d'égalité des intensités lumineuses n'est pas appropriée, contrairement aux images de synthèse, où cette contrainte a été imposée. Cependant, puisque l'estimation de μ et ν est indépendante de celle de λ , nous pouvons

3.3. VALIDATION EXPÉRIMENTALE

	$\frac{ \nu - \hat{\nu} }{ \nu }$	$\frac{ \mu - \hat{\mu} }{ \mu }$	$\frac{ \lambda - \hat{\lambda} }{ \lambda }$	EAM	CPU		$\frac{ \nu - \hat{\nu} }{ \nu }$	$\frac{ \mu - \hat{\mu} }{ \mu }$	$\frac{ \lambda - \hat{\lambda} }{ \lambda }$	EAM	CPU
<i>B01</i>						<i>B11</i>					
ME	18,83	5,26	0,26	14,78	17,03	ME	2,29	4,62	0,13	19,10	17,99
MD	33,92	5,17	0,06	14,78	9,05	MD	2,17	2,32	0,16	12,06	1,67
TV-m-MI	35,45	3,27	0,75	16,23	0,03	TV-m-MI	0,14	1,96	0,68	16,49	0,03
TV-z-MI	24,45	4,23	0,75	16,22	0,03	TV-z-MI	2,02	1,52	0,51	14,36	0,03
TV-z-ME	24,45	4,23	0,22	12,69	0,26	TV-z-ME	2,02	1,52	0,33	12,79	0,43
<i>B02</i>						<i>B12</i>					
ME	2,27	13,24	0,30	17,08	16,92	ME	4,33	22,20	0,46	25,51	21,27
MD	3,80	15,51	0,16	18,56	4,34	MD	0,83	10,92	0,36	14,92	1,06
TV-m-MI	3,95	3,45	1,02	15,41	0,14	TV-m-MI	5,52	5,30	0,04	7,92	0,04
TV-z-MI	1,12	8,86	1,15	16,66	0,12	TV-z-MI	6,70	1,34	0,03	6,63	0,03
TV-z-ME	1,12	8,86	0,11	9,37	0,26	TV-z-ME	6,70	1,34	0,36	11,19	0,38
<i>B03</i>						<i>B13</i>					
ME	7,70	3,25	0,23	11,56	23,57	ME	0,84	1,03	0,08	5,21	22,26
MD	10,79	2,86	0,10	9,23	4,67	MD	2,21	1,89	0,10	8,36	11,63
TV-m-MI	22,32	0,64	0,51	12,95	0,04	TV-m-MI	0,39	0,43	0,41	7,13	0,03
TV-z-MI	10,18	0,61	0,33	9,01	0,03	TV-z-MI	2,99	3,29	3,53	26,68	0,03
TV-z-ME	10,18	0,61	0,16	5,41	0,28	TV-z-ME	2,99	3,29	0,09	14,82	0,52
<i>B04</i>						<i>B15</i>					
ME	77,65	2,90	0,18	10,41	20,85	ME	22,59	3,06	0,16	11,47	19,11
MD	73,36	4,69	0,01	14,23	4,44	MD	38,75	4,45	0,09	16,38	2,62
TV-m-MI	121,13	2,11	0,63	15,69	0,03	TV-m-MI	35,21	2,06	0,63	16,25	0,03
TV-z-MI	49,40	4,22	0,94	19,15	0,03	TV-z-MI	30,25	4,07	1,17	22,55	0,04
TV-z-ME	49,40	4,22	0,14	12,84	0,39	TV-z-ME	30,25	4,07	0,15	14,80	0,35
<i>B05</i>						<i>B16</i>					
ME	0,40	0,12	0,26	7,50	18,54	ME	4,89	2,16	0,26	8,55	21,61
MD	0,04	0,39	0,13	3,38	5,65	MD	8,42	4,12	0,07	9,96	6,34
TV-m-MI	2,38	2,91	0,59	13,93	0,03	TV-m-MI	7,53	0,68	0,34	8,71	0,04
TV-z-MI	0,86	2,51	0,58	13,08	0,03	TV-z-MI	7,42	1,57	0,38	9,65	0,03
TV-z-ME	0,86	2,51	0,02	8,87	0,28	TV-z-ME	7,42	1,57	0,21	7,25	0,33
<i>B06</i>						<i>B17</i>					
ME	3,32	3,73	0,19	12,89	19,23	ME	3,35	2,72	0,19	10,15	21,45
MD	4,53	2,93	0,19	11,35	3,11	MD	4,86	3,25	0,02	10,84	3,68
TV-m-MI	1,28	1,58	0,30	9,05	0,03	TV-m-MI	5,06	2,00	0,53	12,99	0,03
TV-z-MI	4,96	2,02	0,55	13,43	0,03	TV-z-MI	2,25	2,59	0,64	14,61	0,03
TV-z-ME	4,96	2,02	0,21	9,85	0,32	TV-z-ME	2,25	2,59	0,18	9,56	0,32
<i>B07</i>						<i>B18</i>					
ME	1,94	3,54	0,24	11,21	18,53	ME	2,01	1,98	0,18	10,12	20,21
MD	2,29	2,62	0,15	7,98	11,48	MD	5,14	1,89	0,12	11,19	1,39
TV-m-MI	2,28	2,00	0,53	11,97	0,03	TV-m-MI	4,53	0,18	0,27	9,38	0,03
TV-z-MI	3,63	3,19	0,83	16,41	0,03	TV-z-MI	4,16	1,52	0,61	14,64	0,03
TV-z-ME	3,63	3,19	0,21	10,73	0,35	TV-z-ME	4,16	1,52	0,13	9,33	0,43
<i>B08</i>						<i>B19</i>					
ME	157,37	4,66	0,34	19,70	18,31	ME	24,39	5,47	0,18	16,24	21,73
MD	127,10	3,69	0,25	14,75	2,90	MD	32,28	5,71	0,24	17,80	3,18
TV-m-MI	136,65	0,87	0,33	10,01	0,03	TV-m-MI	23,29	2,45	0,45	13,26	0,10
TV-z-MI	154,10	1,13	0,38	10,96	0,03	TV-z-MI	37,93	3,61	0,73	17,61	0,08
TV-z-ME	154,10	1,13	0,22	7,90	0,40	TV-z-ME	37,93	3,61	0,18	12,50	0,82
<i>B09</i>						<i>B20</i>					
ME	5,50	6,99	0,30	17,98	23,36	ME	1,83	6,90	0,33	17,20	18,05
MD	7,24	7,22	0,14	15,98	1,10	MD	3,89	5,17	0,21	11,86	5,69
TV-m-MI	7,76	2,11	0,30	10,95	0,03	TV-m-MI	2,76	1,90	0,30	8,38	0,03
TV-z-MI	8,24	0,88	0,21	9,36	0,03	TV-z-MI	5,91	1,40	0,31	8,71	0,03
TV-z-ME	8,24	0,88	0,30	9,38	0,32	TV-z-ME	5,91	1,40	0,22	7,10	0,32
<i>B10</i>						<i>B21</i>					
ME	12,34	217,43	0,15	6,92	19,22	ME	43,56	1,80	0,27	9,98	16,76
MD	16,48	459,18	0,14	11,34	5,32	MD	90,43	2,64	0,18	11,93	2,35
TV-m-MI	6,22	340,09	0,22	9,10	0,03	TV-m-MI	112,68	1,05	0,55	14,06	0,03
TV-z-MI	25,54	87,17	0,17	7,83	0,03	TV-z-MI	103,96	1,82	0,67	15,67	0,03
TV-z-ME	25,54	87,17	0,13	6,26	0,33	TV-z-ME	103,96	1,82	0,16	10,38	0,48

TABLE 3.3 – Comparaison des résultats obtenus avec les méthodes ME [9], MD [89] et les méthodes utilisant la régularisation TV, sur les vingt premiers jeux de la base de données Yale DataFace B.

remplacer cette dernière étape par une autre méthode de résolution de l’ambiguïté BR. Nous fournissons donc également les résultats obtenus en estimant λ (et uniquement ce paramètre) par ME. Cette méthode, notée TV- z -ME, s’écrit formellement :

$$\begin{cases} (\hat{\mu}, \hat{\nu}) = \underset{\mu, \nu}{\operatorname{argmin}} \operatorname{TV}(\mathbf{g}(\mu, \nu, 1)\mathbf{u}^0) \\ \hat{\lambda} = \underset{\lambda}{\operatorname{argmin}} \mathcal{E}(\|\mathbf{G}(\hat{\mu}, \hat{\nu}, \lambda)^{-1}\mathbf{m}^0\|) \end{cases} \quad (3.56)$$

où la seconde étape est effectuée par recherche exhaustive sur un échantillonnage des valeurs possibles pour λ . Remarquons que la taille de l’espace à échantillonner est considérablement plus petite que pour la méthode ME proposée par Alldrin et al. [9], puisque μ et ν sont déjà estimés. L’exemple *Korean Doll* (cf. table 3.3) est en ce sens très informatif : le temps de calcul passe de plus de deux minutes pour la méthode ME à quelques secondes pour la méthode TV- z -ME proposée. Ceci confirme que la méthode d’estimation de μ et ν peut être utilisée pour accélérer toute méthode de résolution de l’ambiguïté GBR. Mais surtout, ces tests valident le choix d’une régularisation de type TV, qui semble adaptée à une grande variété d’objets.

Enfin, nous présentons sur les figures 3.12 et 3.13 les modèles 3D reconstruits, obtenus en intégrant [239] le champ de normales estimé par la méthode TV qui fournit la valeur la plus faible de l’EAM, puis en plaquant l’albédo estimé sur le relief, pour les quinze jeux de données de la figure 3.11. Le rééclairage de ces modèles 3D est généré à partir du modèle de Phong [206] et de trois sources lumineuses.

En certains points où les images diffèrent fortement du modèle lambertien linéaire (e.g. l’œil en céramique de la chouette), même après prétraitement, le champ de normales estimé est fortement non intégrable. La reconstruction 3D est donc biaisée à proximité de ces points, ce qui se traduit par des artéfacts relativement importants. Ces artéfacts pourraient facilement être « lissés » par une régularisation de la reconstruction 3D. Ceci suggère l’utilisation d’une méthode variationnelle car, en plus de garantir le caractère bien posé du problème UPS, la régularisation servirait à éliminer les artéfacts dus aux écarts au modèle. Cette idée sera reprise et approfondie dans la partie IV.

3.4 Conclusion du chapitre

Nous avons montré dans ce chapitre qu’une transformation GBR affectait les variations de la profondeur, de son gradient, de l’albédo et des normales. En particulier, régulariser par variation totale la profondeur z ou le champ $\mathbf{m}$ permet de ramener l’ambiguïté GBR à une ambiguïté BR, laquelle est bien plus facile à résoudre. Régulariser également l’albédo ρ par variation totale rend le problème UPS bien posé, même si nous avons vu que pour certaines surfaces dont l’albédo n’est pas suffisamment régulier, on devait préférer d’autres méthodes de résolution de l’ambiguïté BR.

Le choix d’une régularisation de type TV reste cependant empirique. Bien qu’il ait été validé a posteriori sur un grand nombre de données, il resterait à mener une étude statistique de la distribution des fonctions ρ , z et ∇z , qui pourrait permettre de comprendre pourquoi ces critères TV fournissent d’aussi bons résultats.

Rappelons également que nous nous sommes restreints au cas où la réflectance est lambertienne, la caméra est orthographique et les éclairages sont directionnels. Avec des modèles aussi simples, voire simplistes, le problème UPS est un problème mal posé. Avec un modèle de projection perspective, il a été montré dans [197] que l'intégrabilité des normales suffisait à rendre le problème bien posé. Des modèles de réflectance non lambertiens peuvent également être pris en compte pour le problème UPS [98], ainsi que des modèles d'éclairage non directionnels [198]. En contrepartie, il faut s'attendre à des difficultés numériques résultant de la perte de la linéarité. Nous proposerons dans la partie IV une nouvelle méthode permettant de parer à ces difficultés numériques.

Enfin, l'hypothèse faite jusqu'ici sur le nombre d'images est que $m \geq 3$. Dans le prochain chapitre, nous allons revenir au problème calibré, pour étudier le cas $m = 2$. Nous allons voir qu'il s'agit également d'un problème mal posé, qui peut lui aussi tirer bénéfice de la contrainte d'intégrabilité du champ de normales.



FIGURE 3.12 – Rééclairage des modèles 3D estimés par la méthode TV qui procure la valeur la plus faible de l'EAM, pour les six derniers jeux de données de la figure 3.11.

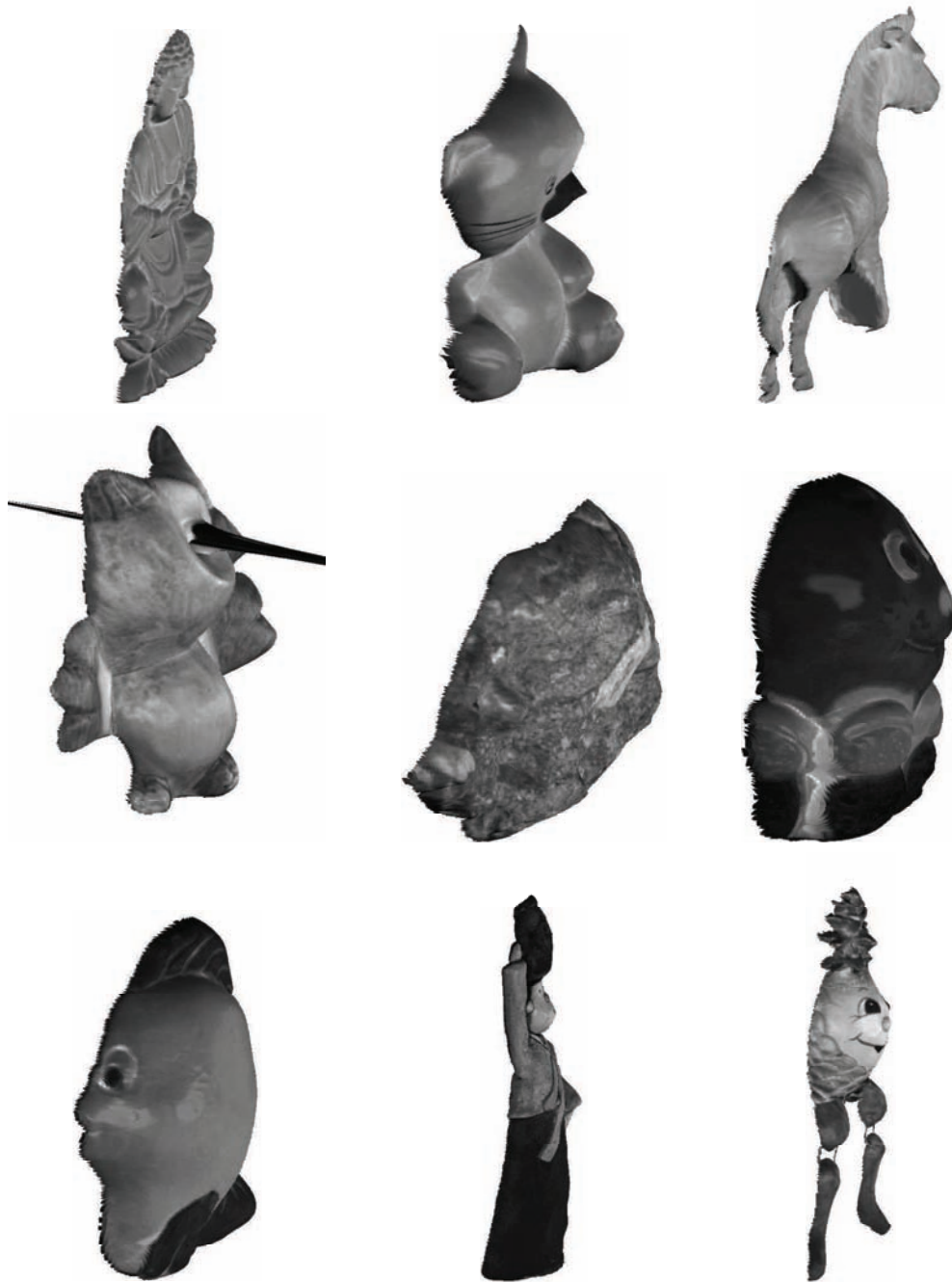


FIGURE 3.13 – Rééclairage des modèles 3D estimés par la méthode TV qui procure la valeur la plus faible de l'EAM, pour les neufs premiers jeux de données de la figure 3.11.

Stéréophotométrie à deux images

Dans le paragraphe 1.1.2, nous avons montré que le problème de la stéréophotométrie calibrée sous des éclairages directionnels admettait une solution unique si les $m \geq 3$ éclairages n'étaient pas tous coplanaires : la normale et l'albédo peuvent être déterminés sans ambiguïté, sans qu'il soit nécessaire de recourir à une hypothèse supplémentaire. Dans le cas du SFS, qui correspond à $m = 1$, le problème est intrinsèquement mal posé car il est impossible de séparer l'albédo de l'ombrage. Il est donc nécessaire de supposer l'albédo connu, mais cette hypothèse n'est pas suffisante : il faut généralement une hypothèse supplémentaire, par exemple sur la régularité de la surface à reconstruire.

Dans ce chapitre, nous décrivons des travaux déjà présentés dans [218] qui concernent le cas intermédiaire $m = 2$ entre le *shape-from-shading* (SFS) et la stéréophotométrie à $m \geq 3$ éclairages non coplanaires (*photometric stereo with three images*, notée PS3). Le cas $m = 2$ se généralise facilement au cas de $m \geq 3$ éclairages coplanaires, que nous avons déjà traité dans [213] et qui permet de décrire, par exemple, l'éclairage provenant du soleil au cours d'une même journée : dans les deux cas, la matrice $\mathbf{S}$ du système linéaire (1.18) n'est plus de rang maximal (*full rank*). Nous verrons dans le paragraphe 4.1 que le cas $m = 2$ est bien plus proche du SFS que de la PS3 : l'albédo doit être connu, ce qui rend le problème non linéaire ; même sous cette hypothèse très forte, le problème reste mal posé sans hypothèse supplémentaire. Nous désignerons dorénavant par PS2 la stéréophotométrie à $m = 2$ images, dans le cas où l'albédo ρ est connu en tout point.

Le problème PS2 a été étudié dans relativement peu d'articles. Onn et Bruckstein prouvent dans [192] que la détermination de la normale est un problème a priori ambigu, et montrent comment éliminer cette ambiguïté en utilisant la contrainte d'intégrabilité. Yang et al. étudient dans [277] le cas particulier des objets convexes. Dans [231], Sato et Ikeuchi utilisent la méthode de résolution proposée par Onn et Bruckstein pour résoudre le problème de la stéréophotométrie à partir de $m \geq 3$ images prises en extérieur, i.e. d'une scène éclairée par le soleil [267]. Dans [155], Kozera parvient aux mêmes conclusions que Onn et Bruckstein à partir d'une formulation différentielle du problème PS2. Au cours de la décennie suivante, seul Ikeda s'est intéressé au problème PS2, mais en considérant la seconde image comme un moyen de résoudre plus précisément le problème SFS, tout comme Mecca et al. dans [173]. Plus récemment, le problème de la stéréophotométrie en extérieur a été revisité dans [1, 3, 142]. Enfin, une version préliminaire de notre travail a déjà été publiée dans [172].

Nous montrons dans le paragraphe 4.2 que l'approche classique, qui consiste à estimer la normale, reste utilisable pourvu que, comme pour le problème UPS (cf. chapitre 3), on trouve un moyen de forcer l'intégrabilité du champ de normales. On peut bien sûr se demander si le problème PS2 n'est pas un problème purement formel, dans la mesure où

la tendance actuelle est d'utiliser bien plus que deux images, comme dans [151] où une séquence vidéo entière est utilisée, dans [258] où les images sont acquises dans une salle équipée de plus de 1200 sources lumineuses, ou bien dans [1, 3] où des séquences d'images acquises sur une période de plusieurs années sont traitées. Outre l'intérêt théorique d'étudier le nombre de solutions de la stéréophotométrie, nous avons mentionné dans le chapitre 1 le problème des ombres (propres et portées) et celui des reflets brillants, qui ont été beaucoup étudiés, notamment dans plusieurs articles récents [20, 51, 119, 131, 245, 257, 270, 272]. Lorsque $m > 3$, les niveaux de gris situés à l'intérieur d'une ombre ou d'une spécularité peuvent être considérés comme des données aberrantes et ignorés, explicitement [20, 51, 245] ou implicitement [131, 257, 270]. Lorsque $m = 3$, Hernández et al. proposent dans [119] un traitement spécial pour les points qui ne sont éclairés que par deux sources lumineuses, i.e. là où le problème PS2 doit être résolu, afin d'améliorer la précision de la reconstruction 3D : nous étudions cette possibilité dans le paragraphe 4.4.1. Enfin, nous étudions le cas de photographies prises en extérieur dans le paragraphe 4.4.3.

4.1 Mise en évidence des ambiguïtés du problème PS2

Nous commençons par montrer, et par illustrer sur des exemples simples, pourquoi le problème de la stéréophotométrie à $m = 2$ images est mal posé. Bien que l'albédo soit supposé connu, nous montrons notamment que cela ne suffit pas à garantir le caractère bien posé du problème PS2.

4.1.1 Formulation non différentielle du problème PS2

Afin d'alléger les notations, nous supposons temporairement que les vecteurs d'éclairage $\mathbf{s}^i$ sont unitaires, i.e. $\phi^i = 1$. En choisissant $m = 2$ dans (1.18), et en notant les vecteurs d'éclairage $\mathbf{s}^i = [s_1^i, s_2^i, s_3^i]^\top$ (supposés connus), et la normale $\mathbf{n}(u, v) = [n_1(u, v), n_2(u, v), n_3(u, v)]^\top$, la formulation non différentielle du problème PS2 s'écrit :

$$\begin{cases} \rho(u, v) [s_1^1 n_1(u, v) + s_2^1 n_2(u, v) + s_3^1 n_3(u, v)] = I^1(u, v) \\ \rho(u, v) [s_1^2 n_1(u, v) + s_2^2 n_2(u, v) + s_3^2 n_3(u, v)] = I^2(u, v) \\ n_1(u, v)^2 + n_2(u, v)^2 + n_3(u, v)^2 = 1 \end{cases} \quad (4.1)$$

Si $\rho(u, v)$ est supposé connu, le problème (4.1) consiste en chaque point (u, v) en un système de trois équations à trois inconnues $(n_1(u, v), n_2(u, v), n_3(u, v))$. À cause de la non linéarité de la troisième équation, il se pourrait que la solution ne soit pas unique, comme cela a été remarqué par Ikeuchi et Horn dans [132] : « Bien sûr, les équations linéaires ci-dessus peuvent admettre plus d'une solution, auquel cas une information supplémentaire (telle qu'une troisième image) peut être requise pour garantir l'unicité de la solution »¹. En effet, nous allons voir que la plupart du temps, l'équation (4.1) n'admet pas une solution unique en $\mathbf{n}$, mais deux.

1. « *Naturally, the above nonlinear equations [...] may have more than one solution, in which case additional information (such as a third image) may be needed to find a unique answer.* »

Pour résoudre le système (4.1), nous définissons l'ensemble $\mathcal{S}$ des normales unitaires $\mathbf{n}$ deux fois éclairées, i.e. celles pour lesquelles $\mathbf{s}^1 \cdot \mathbf{n} > 0$ et $\mathbf{s}^2 \cdot \mathbf{n} > 0$. Cet ensemble est une partie ouverte de la sphère de Gauss $\mathcal{G}$, limitée par deux plans passant par l'origine (cf. figure 4.1-a) :

$$\begin{cases} \pi^1 : & s_1^1 n_1 + s_2^1 n_2 + s_3^1 n_3 = 0 \\ \pi^2 : & s_1^2 n_1 + s_2^2 n_2 + s_3^2 n_3 = 0 \end{cases} \quad (4.2)$$

Les plans π^1 et π^2 sont orthogonaux, respectivement, à $\mathbf{s}^1$ et à $\mathbf{s}^2$. En chaque point (u, v) deux fois éclairé, caractérisé par les niveaux de gris $I^1(u, v)$ et $I^2(u, v)$ et par l'albédo $\rho(u, v) > 0$, supposé connu, le système (4.1) admet deux solutions, notées $\mathbf{n}^+(u, v)$ et $\mathbf{n}^-(u, v)$, qui sont les intersections de $\mathcal{S}$ et de deux plans obtenus par translations de π^1 et π^2 de $I^1(u, v)/\rho(u, v)$ et $I^2(u, v)/\rho(u, v)$ dans les directions de $\mathbf{s}^1$ et $\mathbf{s}^2$, respectivement (cf. figure 4.1-b) ². Notons π le plan supporté par $\mathbf{s}^1$ et $\mathbf{s}^2$ et passant par l'origine de $\mathcal{G}$. On voit sur la figure 4.1-b que $\mathbf{n}^+(u, v)$ et $\mathbf{n}^-(u, v)$ sont symétriques par rapport à π .

Si $I^1(u, v)$ et $I^2(u, v)$ sont parfaitement conformes au modèle lambertien linéaire, alors le système (4.1) admet soit deux solutions, soit une seule solution dans le cas limite où $\mathbf{n}^+(u, v) = \mathbf{n}^-(u, v)$. Mais si tel n'est pas le cas, par exemple à cause du bruit impossible à éviter dès lors que l'on utilise des images réelles, il existera probablement des points (u, v) où le problème n'admet aucune solution exacte réelle. Une façon de prendre en compte ces différents cas sera proposée dans le paragraphe 4.3.

4.1.2 Deux formulations différentielles du problème PS2

En utilisant le modèle de projection orthographique, le lien entre la normale $\mathbf{n}$ et le gradient de profondeur ∇z s'écrit :

$$\mathbf{n}(u, v) = \frac{1}{\sqrt{\|\nabla z(u, v)\|^2 + 1}} \left[-\nabla z(u, v)^\top, 1 \right]^\top \quad (4.3)$$

En remplaçant la normale par cette expression dans (4.1), et en ajoutant une condition au bord de type Dirichlet, une première formulation différentielle du problème PS2 comporte deux équations aux dérivées partielles (EDP) non linéaires en z :

$$\begin{cases} \rho(u, v) \frac{-\tilde{\mathbf{s}}^1 \cdot \nabla z(u, v) + s_3^1}{\sqrt{1 + \|\nabla z(u, v)\|^2}} = I^1(u, v) & \forall (u, v) \in \Omega \\ \rho(u, v) \frac{-\tilde{\mathbf{s}}^2 \cdot \nabla z(u, v) + s_3^2}{\sqrt{1 + \|\nabla z(u, v)\|^2}} = I^2(u, v) & \forall (u, v) \in \Omega \\ z(u, v) = g(u, v) & \forall (u, v) \in \partial\Omega \end{cases} \quad (4.4)$$

où $\mathbf{s}^i = [\tilde{\mathbf{s}}^i^\top, s_3^i]^\top$, $i = 1, 2$. Le problème (4.4) sera probablement plus difficile à résoudre que le système (4.1) à cause de sa non linéarité, mais son principal inconvénient concerne le recours à une condition au bord : outre les images (I^1, I^2) et les vecteurs d'éclairage ($\mathbf{s}^1, \mathbf{s}^2$), la fonction g constitue une information supplémentaire, qui est rarement disponible en pratique.

2. Une ambiguïté similaire apparaît dans d'autres problèmes de vision par ordinateur [22, 67].

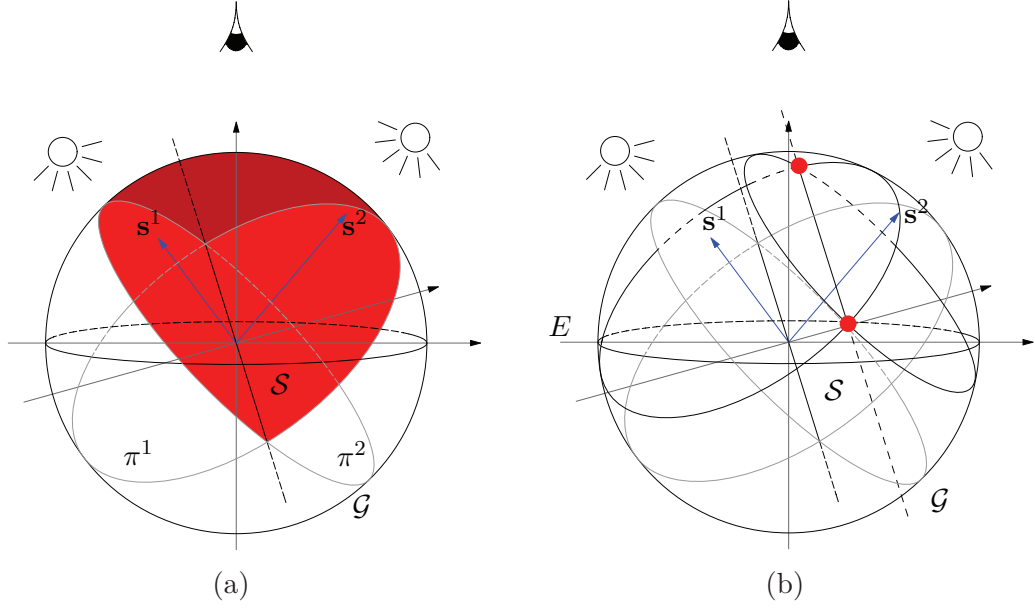


FIGURE 4.1 – (a) L'ensemble $\mathcal{S}$ des normales deux fois éclairées, mis en évidence en rouge, est la partie de la sphère de Gauss $\mathcal{G}$ limitée par les deux plans π^1 et π^2 , qui sont orthogonaux aux vecteurs d'éclairage $\mathbf{s}^1$ et $\mathbf{s}^2$. (b) Un exemple où le système (4.1) admet deux solutions en $\mathbf{n}$ (mises en évidence par des points rouges).

Néanmoins, comme un même facteur $\rho(u, v) / \sqrt{1 + \|\nabla z(u, v)\|^2}$ apparaît dans les deux EDP de (4.4), on peut combiner ces EDP de façon à éliminer simultanément l'albédo $\rho(u, v)$ et la non linéarité (il suffit de supposer que $\rho(u, v) \neq 0$) :

$$\left[I^2(u, v) \tilde{\mathbf{s}}^1 - I^1(u, v) \tilde{\mathbf{s}}^2 \right] \cdot \nabla z(u, v) = I^2(u, v) s_3^1 - I^1(u, v) s_3^2 \quad (4.5)$$

En utilisant la même condition au bord que dans (4.4), nous obtenons alors le problème linéaire suivant :

$$\begin{cases} \mathbf{b}(u, v) \cdot \nabla z(u, v) = f(u, v) & \forall (u, v) \in \Omega \\ z(u, v) = g(u, v) & \forall (u, v) \in \partial\Omega \end{cases} \quad (4.6)$$

où :

$$\begin{cases} \mathbf{b}(u, v) = I^2(u, v) \tilde{\mathbf{s}}^1 - I^1(u, v) \tilde{\mathbf{s}}^2 \\ f(u, v) = I^2(u, v) s_3^1 - I^1(u, v) s_3^2 \end{cases} \quad (4.7)$$

Cette seconde formulation différentielle du problème PS2 permet de propager l'information $g(u, v)$ depuis le bord $\partial\Omega$ vers le reste du domaine Ω , le long des lignes du champ $\mathbf{b}(u, v)$. Il est prouvé dans [173] que le problème (4.6) admet une unique solution $z(u, v)$ lipschitzienne continue, pour peu que $\mathbf{b}$, f et g soient « suffisamment régulières », dès lors que $\partial\Omega$ ne constitue pas une caractéristique. Dans ce même article, un schéma numérique est proposé pour approcher la solution, dont la convergence est prouvée.

4.1.3 Exemple de problème PS2

Afin de mettre en évidence la cohérence entre les différentes formulations du problème PS2, et les ambiguïtés associées, prenons l'exemple d'une surface plane d'équation $z(u, v) = u$, éclairée par les vecteurs :

$$\begin{cases} \mathbf{s}^1 &= [0, 0, 1]^\top \\ \mathbf{s}^2 &= \frac{1}{2}[1, 1, \sqrt{2}]^\top \end{cases} \quad (4.8)$$

Si $\rho \equiv 1$, les images de cette surface sont uniformes, de niveaux de gris :

$$\begin{cases} I^1(u, v) &= \frac{1}{\sqrt{2}} \\ I^2(u, v) &= \frac{\sqrt{2}-1}{2\sqrt{2}} \end{cases} \quad (4.9)$$

Commençons par résoudre ce problème PS2 au moyen de la formulation non différentielle (4.1), qui s'écrit ici :

$$\begin{cases} n_3(u, v) = \frac{1}{\sqrt{2}} \\ \frac{n_1(u, v) + n_2(u, v) + \sqrt{2}n_3(u, v)}{2} = \frac{\sqrt{2}-1}{2\sqrt{2}} \\ n_1(u, v)^2 + n_2(u, v)^2 + n_3(u, v)^2 = 1 \end{cases} \quad (4.10)$$

Il est facile de vérifier que, pour tout (u, v) , le système (4.10) admet exactement deux solutions :

$$\begin{cases} \mathbf{n}^+ = \frac{1}{\sqrt{2}} [-1, 0, 1]^\top \\ \mathbf{n}^- = \frac{1}{\sqrt{2}} [0, -1, 1]^\top \end{cases} \quad (4.11)$$

Si la surface à reconstruire est supposée différentiable partout, il existe donc deux champs de normales possibles $\mathbf{n}(u, v) = \mathbf{n}^+$ ou $\mathbf{n}(u, v) = \mathbf{n}^-$. Il s'ensuit, d'après (4.3), qu'il existe deux valeurs possibles pour $\nabla z(u, v)$:

$$\begin{cases} \nabla z^+(u, v) = \frac{-1}{n_3^+(u, v)} \begin{bmatrix} n_1^+(u, v) \\ n_2^+(u, v) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \nabla z^-(u, v) = \frac{-1}{n_3^-(u, v)} \begin{bmatrix} n_1^-(u, v) \\ n_2^-(u, v) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases} \quad (4.12)$$

En imposant arbitrairement $z(0, 0) = 0$, nous obtenons finalement les deux solutions $z^+(u, v) = u$ et $z^-(u, v) = v$. Nous constatons sans surprise que l'une de ces solutions est la surface d'origine.

Écrivons maintenant la première formulation différentielle (4.4), en supposant de nouveau que $\rho \equiv 1$, sachant qu'aucune condition au bord n'est disponible :

$$\begin{cases} \frac{1}{\sqrt{1 + \|\nabla z(u, v)\|^2}} = \frac{1}{\sqrt{2}} \\ \frac{-\partial_u z(u, v) - \partial_v z(u, v) + \sqrt{2}}{2\sqrt{1 + \|\nabla z(u, v)\|^2}} = \frac{\sqrt{2}-1}{2\sqrt{2}} \end{cases} \quad (4.13)$$

Le système (4.13) est équivalent à :

$$\begin{cases} \partial_u z(u, v) + \partial_v z(u, v) = 1 \\ \partial_u z(u, v)^2 + \partial_v z(u, v)^2 = 1 \end{cases} \quad (4.14)$$

On vérifie facilement que le système (4.14) admet exactement les mêmes solutions (4.12) en $\nabla z(u, v)$ que (4.10). Les deux formulations sont donc équivalentes.

Enfin, la deuxième formulation différentielle (4.6) du problème se caractérise ici par les expressions suivantes de $\mathbf{b}(u, v)$ et $f(u, v)$:

$$\begin{cases} \mathbf{b}(u, v) = - \left[\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4} \right]^\top \\ f(u, v) = - \frac{\sqrt{2}}{4}. \end{cases} \quad (4.15)$$

et s'écrit donc, en l'absence de condition au bord, sous la forme d'une EDP linéaire :

$$\partial_u z(u, v) + \partial_v z(u, v) = 1 \quad (4.16)$$

Cette EDP est exactement la même que la première équation du système (4.14). Or, elle admet une infinité de solutions. Par exemple, toutes les fonctions de la forme $z(u, v) = au + (1 - a)v + w(u - v)$, pour tout $a \in \mathbb{R}$ et pour toute fonction différentiable $w : \mathbb{R} \rightarrow \mathbb{R}$, sont des solutions de (4.16). Comment expliquer que la deuxième formulation différentielle de ce problème PS2 ne nous amène pas aux mêmes conclusions que (4.10) ou (4.14) ? Il suffit de se rappeler qu'en l'absence de condition au bord, la formulation différentielle (4.6) du problème PS2 est une condition *nécessaire, mais non suffisante*.

4.1.4 Problème PS2 avec albédo inconnu

Il nous semble intéressant de mentionner ici une variante du problème PS2, notée PS2 ρ , où l'albédo de la surface est également inconnu. La présence d'une inconnue supplémentaire $\rho(u, v)$ en chaque point $(u, v) \in \Omega$ accroît forcément la complexité du problème. Une des propriétés remarquables de la stéréophotométrie, dans sa version classique, est de pouvoir estimer l'albédo [269], ce qui n'est pas le cas du SFS puisque, même lorsque l'albédo est connu, le problème reste généralement mal posé [191]. Qu'en est-il du problème PS2 ρ ?

Commençons par proposer une interprétation géométrique de l'équation (4.5). Puisque $\mathbf{s}^i = [\tilde{\mathbf{s}}^i, s_3^i]^\top$, $i = 1, 2$, et $\mathbf{n}(u, v)$ est parallèle à $[-\nabla z(u, v)^\top, 1]^\top$, cette équation se réécrit :

$$[I^2(u, v) \mathbf{s}^1 - I^1(u, v) \mathbf{s}^2] \cdot \mathbf{n}(u, v) = 0 \quad (4.17)$$

Comme cela a été remarqué dans [119], l'équation (4.17) peut être directement obtenue à partir des deux premières équations du système (4.1). Elle signifie que $\mathbf{n}(u, v)$ est contenu dans un plan $\chi(u, v)$ orthogonal au vecteur $I^2(u, v) \mathbf{s}^1 - I^1(u, v) \mathbf{s}^2$, qui est lui-même contenu dans le plan π , puisque π est supporté par les vecteurs $\mathbf{s}^1$ et $\mathbf{s}^2$. On peut remarquer que l'égalité (4.17) est vérifiée pour toute valeur $\rho(u, v)$ de l'albédo, car l'équation (4.5) a été obtenue, à partir du problème (4.4), par élimination de $\rho(u, v)$.

Nous avons établi dans le paragraphe 4.1.1 que, pour toute valeur connue $\rho(u, v)$, le problème (4.1) admettait deux solutions $\mathbf{n}^+(u, v)$ et $\mathbf{n}^-(u, v)$ symétriques par rapport au plan π . Soit θ l'angle entre π et les normales. L'équation (4.17) nous dit que, pour chaque paire $(I^1(u, v), I^2(u, v))$ de niveaux de gris, et pour chaque valeur de l'albédo $\rho(u, v)$, les deux vecteurs $\mathbf{n}^+(u, v)$ et $\mathbf{n}^-(u, v)$ sont contenus dans le plan $\chi(u, v)$. En outre, ces vecteurs dépendent de la valeur de $\rho(u, v)$: il est facile de vérifier que l'angle θ entre le plan π et les normales $\mathbf{n}^+(u, v)$ et $\mathbf{n}^-(u, v)$ diminue progressivement jusqu'à s'annuler, lorsque $\rho(u, v)$ décroît de 1 vers une valeur minimale $\rho_{\inf}(u, v)$ qui correspond au cas limite $\mathbf{n}^+(u, v) = \mathbf{n}^-(u, v)$. Il existe donc, pour tout $(u, v) \in \Omega$, un intervalle $[\rho_{\inf}(u, v), 1]$ de valeurs possibles pour $\rho(u, v)$. Si $\mathbf{s}^1 \neq \mathbf{s}^2$, on montre facilement que la borne inférieure de cet intervalle s'écrit :

$$\rho_{\inf}(u, v) = \sqrt{\frac{I^1(u, v)^2 + I^2(u, v)^2 - 2 I^1(u, v) I^2(u, v) (\mathbf{s}^1 \cdot \mathbf{s}^2)}{1 - (\mathbf{s}^1 \cdot \mathbf{s}^2)^2}} \quad (4.18)$$

Par conséquent, le problème PS2 ρ est mal posé. Toutefois, lorsqu'une condition au bord est disponible, le problème (4.6) peut être résolu sans connaissance de l'albédo, et l'albédo $\rho(u, v)$ peut être calculé a posteriori, en utilisant une des deux EDP du système (4.4).

4.1.5 Exemple de problème PS2 ρ

Reprenons l'exemple du paragraphe 4.1.3, sous l'hypothèse d'un albédo $\rho(u, v)$ inconnu. Le problème non différentiel (4.10) s'écrit :

$$\begin{cases} \rho(u, v) n_3(u, v) = \frac{1}{\sqrt{2}} \\ \rho(u, v) \frac{n_1(u, v) + n_2(u, v) + \sqrt{2} n_3(u, v)}{2} = \frac{\sqrt{2} - 1}{2\sqrt{2}} \\ n_1(u, v)^2 + n_2(u, v)^2 + n_3(u, v)^2 = 1 \end{cases} \quad (4.19)$$

Pour chaque valeur de $\rho(u, v)$, les deux premières équations de (4.19) admettent une infinité de solutions dépendant d'un paramètre $t \in \mathbb{R}$:

$$\mathbf{n}(u, v) = \frac{1}{\rho(u, v)} \left(\frac{1}{\sqrt{2}} [-1, 0, 1]^\top + t [-1, 1, 0]^\top \right) \quad (4.20)$$

En remplaçant $\mathbf{n}(u, v)$ par (4.20) dans la troisième équation de (4.19), on trouve en quelques lignes de calcul les deux solutions suivantes :

$$\mathbf{n}(u, v) = \frac{1}{2\sqrt{2}\rho(u, v)} \left[-1 - \epsilon \sqrt{4\rho(u, v)^2 - 3}, -1 + \epsilon \sqrt{4\rho(u, v)^2 - 3}, 2 \right]^\top \quad (4.21)$$

où $\epsilon = \pm 1$. Dans (4.21), la valeur $\rho(u, v)$ de l'albédo peut être fixée arbitrairement, pourvu que $4\rho(u, v)^2 - 3 \geq 0$. Ceci signifie que $\rho(u, v) \geq \sqrt{3}/2$, valeur qui correspond bien à l'expression (4.18) de la valeur limite $\rho_{\inf}(u, v)$. Cet exemple montre le caractère

mal posé du problème PS2 ρ , puisqu'en chaque point $(u, v) \in \Omega$, pour chaque valeur $\rho(u, v) \in]\sqrt{3}/2, 1]$, il existe deux normales possibles, qui correspondent aux deux valeurs possibles de ϵ dans l'expression (4.21).

Quant au problème différentiel (4.13), il s'écrit maintenant :

$$\begin{cases} \rho(u, v) \frac{1}{\sqrt{1 + \|\nabla z(u, v)\|^2}} = \frac{1}{\sqrt{2}} \\ \rho(u, v) \frac{-\partial_u z(u, v) - \partial_v z(u, v) + \sqrt{2}}{2\sqrt{1 + \|\nabla z(u, v)\|^2}} = \frac{\sqrt{2} - 1}{2\sqrt{2}} \end{cases} \quad (4.22)$$

ce qui est équivalent au système suivant :

$$\begin{cases} \partial_u z(u, v) + \partial_v z(u, v) = 1 \\ \partial_u z(u, v)^2 + \partial_v z(u, v)^2 = 2\rho(u, v)^2 - 1 \end{cases} \quad (4.23)$$

Il s'agit à nouveau d'un problème mal posé. La première équation de (4.23), qui n'est autre que l'EDP linéaire (4.16), ne dépend pas de $\rho(u, v)$. Comme cela a déjà été mentionné, le champ vectoriel $\mathbf{b}(u, v)$ indique dans quelle direction l'information de profondeur est propagée depuis le bord, mais cette propagation a lieu de façon indépendante le long de chaque caractéristique, ce qui peut créer un effet de « lignage ». Ce phénomène a été signalé dans [119] par Hernández et al., qui utilisent un terme de régularisation pour le gommer. Cependant, aucune régularisation n'est nécessaire en théorie, si la profondeur est connue sur la frontière $\partial\Omega$: résoudre la première équation de (4.23) par propagation depuis le bord fournit non seulement la profondeur en chaque point, mais également l'albédo $\rho(u, v)$, qui peut être estimé a posteriori en utilisant la seconde équation de (4.23).

Pour conclure, il n'y a aucune incohérence entre la formulation non différentielle et la première formulation différentielle du problème PS2, et ceci est également vrai pour sa variante PS2 ρ . Cela devrait nous permettre de prédire le même nombre de solutions globales, en utilisant indifféremment l'une ou l'autre de ces formulations. Cette prédiction fait l'objet du prochain paragraphe.

4.2 Prédiction du nombre de solutions du problème PS2

En chaque point $(u, v) \in \Omega$ deux fois éclairé, nous avons montré dans le paragraphe 4.1.1 que le problème PS2 admettait une ou deux solutions en $\mathbf{n}(u, v)$. Bien sûr, pour prédire le nombre de champs de normales possibles, les points où la normale peut être déterminée sans ambiguïté sont d'une importance primordiale. De tels *points singuliers* ont été étudiés en détail dans le contexte du SFS [132].

4.2.1 Points singuliers

Comme cela a déjà été mentionné dans le paragraphe 4.1.1, la première situation où l'unicité de la normale peut être prouvée est celle où les deux solutions $\mathbf{n}^+(u, v)$ et $\mathbf{n}^-(u, v)$ coïncident. Dans ce cas, la solution unique est contenue dans le plan π .

L'ensemble $\mathcal{S}_R$ de ces normales constitue une géodésique de $\mathcal{G}$, obtenue par intersection de $\mathcal{S}$ et du plan π (cf. figure 4.2-a). Un second type de points singuliers est celui où l'une des deux solutions de l'équation (4.1) n'est pas visible, car elle pointe vers l'hémisphère sud de la sphère de Gauss $\mathcal{G}$. Notons $\mathcal{S}_Y$ cet ensemble de normales. Il apparaît alors clairement, sur la figure 4.2-b, que la seule solution visible pointe vers l'ensemble $\mathcal{S}_G$ en bijection avec $\mathcal{S}_Y$ (cf. figure 4.2-b). Notons Ω_R et Ω_G les ensembles de points singuliers pour lesquels la normale appartient à $\mathcal{S}_R$ ou à $\mathcal{S}_G$, respectivement.

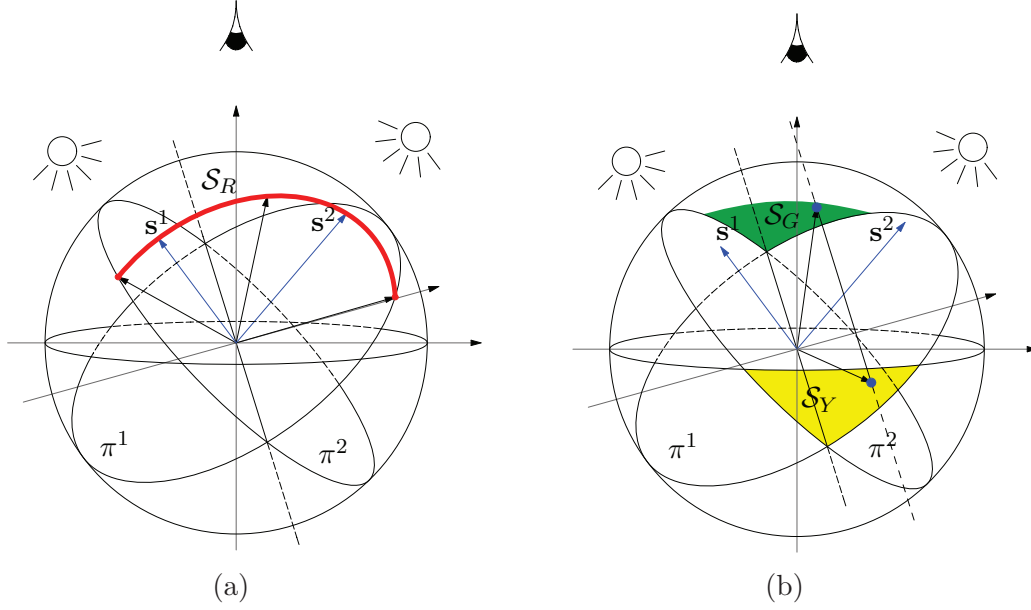


FIGURE 4.2 – Points singuliers du problème PS2. (a) La géodésique rouge $\mathcal{S}_R$ est l'intersection de $\mathcal{S}$ (ensemble des normales deux fois éclairées) et de π (plan porté par $\mathbf{s}^1$ et $\mathbf{s}^2$ passant par l'origine). Chaque normale estimée qui pointe vers $\mathcal{S}_R$ est connue sans ambiguïté. (b) Ceci est également vrai pour toute normale estimée pointant vers $\mathcal{S}_G$, puisque l'autre normale possible pointe vers la partie non visible $\mathcal{S}_Y$ de $\mathcal{S}$.

Sachant qu'un problème PS2 est constitué d'une paire de problèmes de type SFS (pourvu que la pose de la caméra soit la même), pourquoi ne pas utiliser les points singuliers de chaque problème SFS? Ces points sont ceux pour lesquels $\mathbf{n} = \mathbf{s}^1$ dans la première image, et $\mathbf{n} = \mathbf{s}^2$ dans la seconde. Cependant, ces deux valeurs de la normale pointent vers la géodésique $\mathcal{S}_R$, puisque $\mathbf{s}^1$ et $\mathbf{s}^2$ supportent le plan π . Ceci nous montre que les points singuliers du problème PS2 englobent ceux des deux problèmes SFS sous-jacents.

Le champ de normales est continu dès lors que la surface est supposée de classe $\mathcal{C}^1$. Sous cette hypothèse, est-il possible de propager la connaissance de la normale depuis un point singulier vers ses voisins non singuliers? La réponse à cette question dépend du type de point singulier considéré. Soit $P_R \in \Omega_R$ un point singulier du premier type, i.e.

pour lequel la normale pointe vers $\mathcal{S}_R$. La normale en un point non singulier $\bar{P}_R$, voisin de P_R , peut se situer des deux côtés de $\mathcal{S}_R$, i.e. dans chacun des deux sous-ensembles $\mathcal{S}_U$ et $\mathcal{S}_B$ de $\mathcal{S}$ situés, respectivement, au-dessus ou en dessous de la géodésique $\mathcal{S}_R$ (cf. figure 4.3). En d'autres termes, il y a une ambiguïté résiduelle sur la normale au point $\bar{P}_R$. En revanche, soit $P_G \in \Omega_G$ un point singulier du second type, dont la normale pointe vers $\mathcal{S}_G$. En tout point non singulier $\bar{P}_G$, voisin de P_G , on peut déduire, grâce à la continuité du champ de normales, que la normale pointe vers l'ensemble $\mathcal{S}_U$. Pour conclure, la normale peut être déterminée sans ambiguïté pour tous les points de $\Omega \setminus \Omega_R$ connexes à Ω_G .

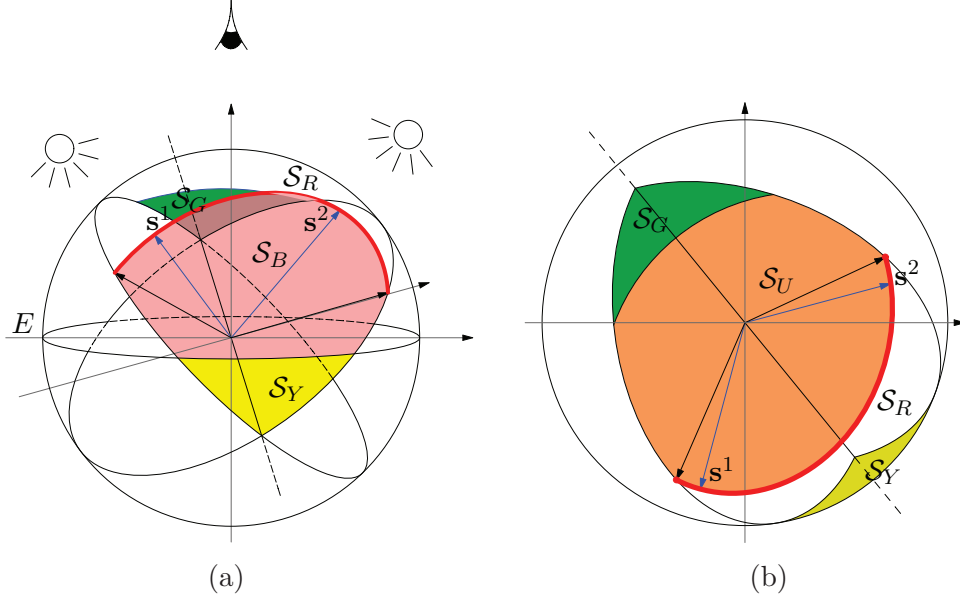


FIGURE 4.3 – (a) L'ensemble rose $\mathcal{S}_B$ est le sous-ensemble de $\mathcal{S}$ borné par l'équateur E et la géodésique $\mathcal{S}_R$. (b) Vu sous un autre angle, l'ensemble orange $\mathcal{S}_U$ est le sous-ensemble de $\mathcal{S}$ situé entre $\mathcal{S}_G$ et $\mathcal{S}_R$.

4.2.2 Utilisation des points singuliers pour la construction d'une condition au bord

En guise d'exemple, calculons deux images de la surface lisse représentée sur la figure 4.4-a, en supposant $\rho \equiv 1$ et en choisissant les vecteurs d'éclairage de telle sorte que tous les points soient éclairés deux fois (cf. figures 4.4-b et 4.4-c). Les ensembles de points singuliers Ω_R et Ω_G , qui sont estimés numériquement selon la procédure décrite dans [172], sont superposés à ces images. Nous faisons une prédiction similaire à celle de Onn et Bruckstein dans [192] pour compter les champs de normales possibles. En chacune des trois parties connexes de $\Omega \setminus \Omega_R$, il y a deux solutions : l'une pointe vers $\mathcal{S}_U$, l'autre vers $\mathcal{S}_B$. Ceci nous fournit donc $2^3 = 8$ champs de normales continus. Une diffé-

rence entre notre comptage et celui proposé par Onn et Bruckstein provient cependant de la détection de Ω_G : puisque l'une des parties connexes de $\Omega \setminus \Omega_R$ contient Ω_G , on peut affirmer que le champ de normales dans cette partie est contenu dans $\mathcal{S}_U$. Pour conclure, le nombre de champs de normales possibles, pour cet exemple de problème PS2, est égal à $2^2 = 4$.

Sur chaque ensemble connexe de points singuliers, la solution z peut être calculée par intégration, à une constante près. Sachant que la formulation différentielle (4.6) nécessite une condition au bord pour être bien posée, voilà un moyen très simple d'en construire une. Puisque le bord de l'image est connexe à Ω_G à l'intérieur de $\Omega \setminus \Omega_R$ (cf. figures 4.4-b et 4.4-c), la profondeur peut y être calculée (à une constante près). D'après [173], ceci nous permet de prédire une solution unique, qui est bien entendu la surface d'origine de la figure 4.4-a. En nous fondant sur cet exemple, nous pourrions conclure que les formulations non différentielle et différentielle du problème PS2 sont complémentaires pour la prédiction du nombre de solutions, dans la mesure où la condition au bord requise par la seconde approche peut apparemment être fournie par la première.

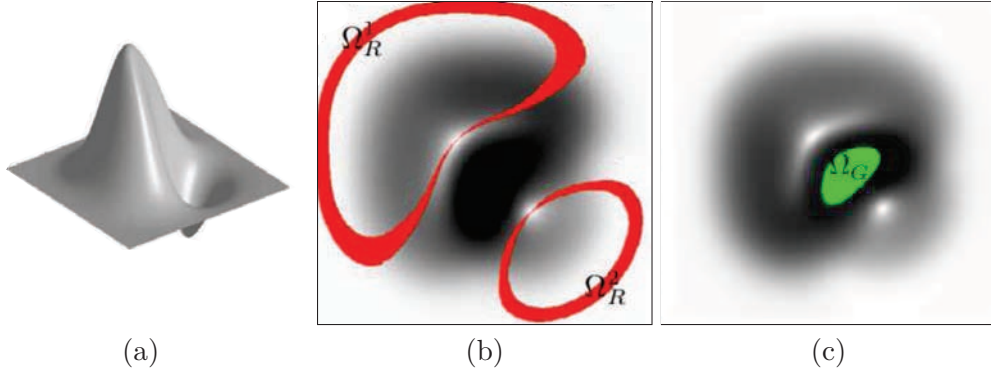


FIGURE 4.4 – Points singuliers d'une surface synthétique. (b)-(c) Deux images de synthèse de la surface lisse (a) telles que tous les points soient éclairés deux fois, sur lesquelles sont superposés les ensembles $\Omega_R = \Omega_R^1 \cup \Omega_R^2$ (en rouge) et Ω_G (en vert).

4.2.3 Une ambiguïté résiduelle

La conclusion du paragraphe précédent n'est pas toujours vraie. Par exemple, pour un problème PS2 tel que celui du paragraphe 4.1.3, où tous les points ont la même normale, il n'existe aucun point singulier. Un autre exemple de problème PS2 pour lequel une ambiguïté subsiste a été suggéré par Kozera [154]. Il consiste à éclairer la surface d'équation $z = uv$ et d'albédo uniforme $\rho \equiv 1$ par les vecteurs $\mathbf{s}^1 = [s, s, c]^\top$ et $\mathbf{s}^2 = [-s, -s, c]^\top$, où $s = \sqrt{2}/2 \sin \alpha$ et $c = \cos \alpha$, pour $\alpha \in]0, \pi/2[$. Les deux images

simulées ainsi ont pour niveaux de gris :

$$\begin{cases} I^1(u, v) = \frac{s(-u - v) + c}{\sqrt{1 + u^2 + v^2}} \\ I^2(u, v) = \frac{-s(-u - v) + c}{\sqrt{1 + u^2 + v^2}} \end{cases} \quad (4.24)$$

Les deux premières équations du système (4.1) s'écrivent alors :

$$\begin{cases} s[n_1(u, v) + n_2(u, v)] + cn_3(u, v) = \frac{s(-u - v) + c}{\sqrt{1 + u^2 + v^2}} \\ -s[n_1(u, v) + n_2(u, v)] + cn_3(u, v) = \frac{-s(-u - v) + c}{\sqrt{1 + u^2 + v^2}} \end{cases} \quad (4.25)$$

Puisque, par définition, s et c sont non nuls, le système (4.25) est équivalent à :

$$\begin{cases} n_1(u, v) + n_2(u, v) = \frac{-u - v}{\sqrt{1 + u^2 + v^2}} \\ n_3(u, v) = \frac{1}{\sqrt{1 + u^2 + v^2}} \end{cases} \quad (4.26)$$

En utilisant (4.26), la troisième équation de (4.1) se réécrit :

$$n_1(u, v)^2 + \frac{u + v}{\sqrt{1 + u^2 + v^2}} n_1(u, v) + \frac{uv}{1 + u^2 + v^2} = 0 \quad (4.27)$$

On montre facilement que cette équation de degré 2 admet toujours deux solutions réelles en $n_1(u, v)$, qui sont égales lorsque $v = u$ et uniquement dans ce cas. Ces deux solutions fournissent deux normales possibles en chaque point $(u, v) \in \Omega$:

$$\begin{cases} \mathbf{n}^+(u, v) = \frac{1}{\sqrt{1 + u^2 + v^2}} \begin{bmatrix} -v \\ -u \\ 1 \end{bmatrix} \\ \mathbf{n}^-(u, v) = \frac{1}{\sqrt{1 + u^2 + v^2}} \begin{bmatrix} -u \\ -v \\ 1 \end{bmatrix} \end{cases} \quad (4.28)$$

On déduit de (4.28) et (4.3) deux valeurs possibles pour $\nabla z(u, v)$:

$$\begin{cases} \nabla z^+(u, v) = \frac{-1}{n_3^+(u, v)} \begin{bmatrix} n_1^+(u, v) \\ n_2^+(u, v) \end{bmatrix} = \begin{bmatrix} v \\ u \end{bmatrix} \\ \nabla z^-(u, v) = \frac{-1}{n_3^-(u, v)} \begin{bmatrix} n_1^-(u, v) \\ n_2^-(u, v) \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} \end{cases} \quad (4.29)$$

En intégrant ces deux gradients de profondeur, on obtient finalement les deux solutions suivantes :

$$\begin{cases} z^+(u, v) = u v \\ z^-(u, v) = \frac{u^2 + v^2}{2} \end{cases} \quad (4.30)$$

Il reste donc une ambiguïté.

D'autre part, Ω_G est vide, mais il est facile de déduire de (4.28) que l'ensemble Ω_R est la droite d'équation $v = u$. La profondeur peut donc être calculée sans ambiguïté le long de cette droite, à une constante additive près. En utilisant un raisonnement similaire à celui du paragraphe 4.2.2, nous pourrions conclure que la solution est unique, ce qui contredirait le résultat (4.30). Il est facile d'expliquer cette apparente contradiction : la connaissance de la normale peut être propagée, depuis Ω_R , le long des caractéristiques. Or, dans cet exemple, l'ensemble Ω_R constitue précisément une caractéristique. La propagation de l'information depuis Ω_R est donc impossible.

4.2.4 Contrainte d'intégrabilité

Le champ vectoriel $[p, q]^\top = [-n_1/n_3, -n_2/n_3]^\top$ se déduit facilement du champ de normales estimé, mais il n'est pas garanti que la contrainte d'intégrabilité [126] :

$$\frac{\partial p}{\partial v} = \frac{\partial q}{\partial u} \quad (4.31)$$

soit vérifiée, alors que cette contrainte est requise pour pouvoir intégrer le champ $[p, q]^\top$ en une fonction z de classe $\mathcal{C}^2$, par résolution de l'équation $\nabla z = [p, q]^\top$ (cf. chapitre 2). Onn et Bruckstein font remarquer dans [192] que « la plupart du temps », dans chaque partie connexe $\mathcal{P}$ de $\Omega \setminus \Omega_R$, l'un des deux champs de normales possibles peut être éliminé car il n'est pas intégrable. Pour procéder à cette élimination, ils utilisent un critère déduit de (4.31) :

$$\iint_{(u,v) \in \mathcal{P}} \left[\frac{\partial p}{\partial v}(u, v) - \frac{\partial q}{\partial u}(u, v) \right]^2 du dv = \min \quad (4.32)$$

Onn et Bruckstein caractérisent également les « cas rares » où (4.32) est satisfaite par plusieurs champs de normales, c'est-à-dire où le problème PS2 admet plusieurs solutions (il se trouve que les exemples des paragraphes 4.1.3 et 4.2.3 sont de tels « cas rares »).

Revenons maintenant à l'exemple de la figure 4.4. En utilisant l'approche non différentielle, nous avons trouvé $2^2 = 4$ champs de normales possibles, tandis qu'un raisonnement plus complexe fondé sur la topologie des ensembles Ω_R et Ω_G nous a permis de prédire l'unicité de la solution. Il semble donc que, parmi les quatre champs de normales possibles, un seul soit intégrable.

Cependant, il est difficile d'étendre la prédiction fondée sur la topologie de Ω_R et Ω_G au cas discret, puisque la notion de continuité est perdue. De plus, calculons les ensembles Ω_R et Ω_G lorsqu'un bruit gaussien de variance égale à 15% du niveau de gris maximal est ajouté aux images des figures 4.4-b et 4.4-c. La figure 4.5 montre que, même si la topologie de Ω_G est robuste, celle de Ω_R est très sensible au bruit. Il est donc impossible

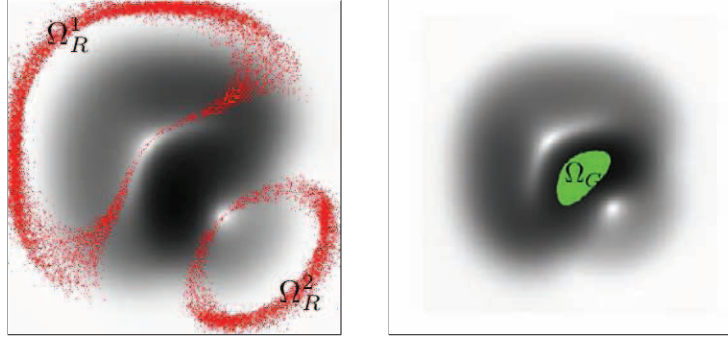


FIGURE 4.5 – Ajout d’un bruit gaussien de variance 0,15 aux deux images de la figure 4.4. L’ensemble Ω_G reste connexe, mais cela n’est pas le cas de l’ensemble $\Omega_R = \Omega_R^1 \cup \Omega_R^2$.

de prédire le nombre de solutions en appliquant le raisonnement du paragraphe 4.2.2 à un ensemble Ω_R aussi fragmenté.

Afin de proposer une nouvelle prédiction du nombre de solutions qui puisse s’étendre au cas discret, il nous faut revenir à une analyse locale du problème PS2, i.e. à sa formulation non différentielle. Nous savons (cf. paragraphe 4.1.1) que le problème PS2 admet deux normales possibles en chaque point non singulier $(u, v) \in \Omega$. Nous pouvons donc construire $2^{|\Omega|}$ champs de normales discrets différents. Comme nous allons le voir, la notion d’intégrabilité, qui s’étend aisément au cas discret, va nous permettre de sélectionner en chaque point la « meilleure » normale.

4.3 Résolution numérique du problème PS2

4.3.1 Reformulation du problème (4.1)

Le problème (4.1) admet deux solutions en $\mathbf{n}$ en chaque point (cf. figure 4.6). Dans ce paragraphe, nous nous proposons de déterminer explicitement ces deux normales par une méthode algébrique. En chaque point $(u, v) \in \Omega$, les deux premières équations du problème (4.1) s’écrivent (nous omettons les dépendances en (u, v) pour alléger les notations, et les vecteurs d’éclairage ne sont plus supposés unitaires) :

$$\underbrace{\begin{bmatrix} \mathbf{s}^{1\top} \\ \mathbf{s}^{2\top} \end{bmatrix}}_{\mathbf{S} \in \mathbb{R}^{2 \times 3}} \mathbf{n} = \underbrace{\begin{bmatrix} \frac{I^1}{I^2} \\ \frac{\rho}{\rho} \end{bmatrix}}_{\mathbf{i} \in \mathbb{R}^2} \quad (4.33)$$

La résolution de l’équation (4.33) au sens des moindres carrés consiste à reformuler le problème (4.1) de la façon suivante :

$$\begin{cases} \min_{\mathbf{n}} \|\mathbf{S}\mathbf{n} - \mathbf{i}\|^2 \\ \text{s.c. } \|\mathbf{n}\| = 1 \end{cases} \quad (4.34)$$

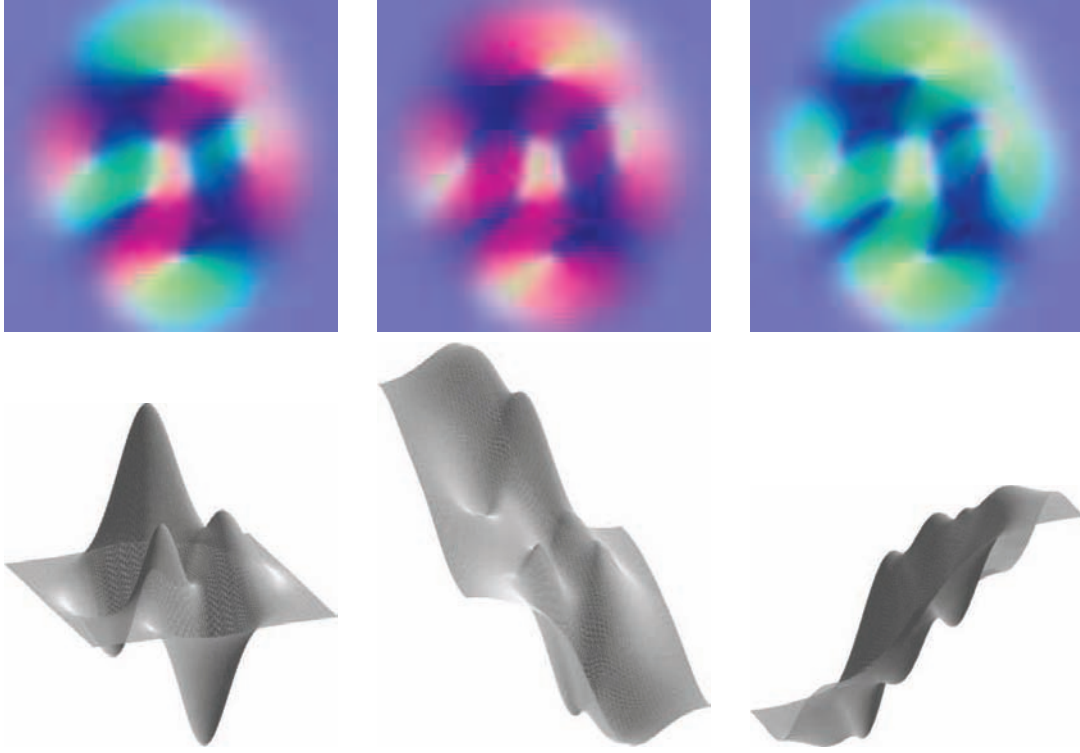


FIGURE 4.6 – Illustration des ambiguïtés inhérentes au problème PS2. Première ligne : normales en fausses couleurs. Deuxième ligne : reliefs 3D obtenus en intégrant ces normales par la méthode [239]. De gauche à droite : vérité terrain, solution $\mathbf{n}^+$ et solution $\mathbf{n}^-$. Les reliefs 3D correspondant aux champs de normales $\mathbf{n}^+$ et $\mathbf{n}^-$ ne sont que deux solutions parmi $2^{|\Omega|}$, puisque toute combinaison de ces champs est également plausible.

La décomposition en valeurs singulières de $\mathbf{S}$ s'écrit :

$$\mathbf{S} = \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}}_{\mathbf{U} \in \mathbb{R}^{2 \times 2}} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}}_{\mathbf{\Sigma} \in \mathbb{R}^{2 \times 3}} \underbrace{\begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \mathbf{v}_3^\top \end{bmatrix}}_{\mathbf{V}^\top \in \mathbb{R}^{3 \times 3}} \quad (4.35)$$

où $\sigma_1 \geq \sigma_2 > 0$ sont les deux valeurs singulières de $\mathbf{S}$ ($\mathbf{S}$ est de rang 2 si $\mathbf{s}^1$ et $\mathbf{s}^2$ sont non colinéaires). Par ailleurs, $\mathbf{u}_1$ et $\mathbf{u}_2$ sont des vecteurs de $\mathbb{R}^2$, et $\mathbf{v}_1$, $\mathbf{v}_2$ et $\mathbf{v}_3$ des vecteurs de $\mathbb{R}^3$, tels que les matrices $\mathbf{U}$ et $\mathbf{V}$ soient orthogonales. La matrice suivante :

$$\mathbf{\Sigma}^+ = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \\ 0 & 0 \end{bmatrix} \quad (4.36)$$

permet de définir la pseudo-inverse de Moore-Penrose de $\mathbf{S}$:

$$\mathbf{S}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^\top \quad (4.37)$$

On montre [200] que le vecteur :

$$\mathbf{n}_0 = \mathbf{S}^+\mathbf{i} = \frac{\mathbf{u}_1^\top \mathbf{i}}{\sigma_1} \mathbf{v}_1 + \frac{\mathbf{u}_2^\top \mathbf{i}}{\sigma_2} \mathbf{v}_2 \quad (4.38)$$

est l'unique solution du problème en moindres carrés suivant :

$$\begin{cases} \min_{\mathbf{n}} \|\mathbf{S}\mathbf{n} - \mathbf{i}\|^2 \\ \text{s.c. } \|\mathbf{n}\| = \min \end{cases} \quad (4.39)$$

Les problèmes (4.34) et (4.39) sont différents, quoique similaires, mais nous allons voir que l'on peut déduire de la solution (4.38) du problème (4.39) les solutions du problème (4.34).

4.3.2 Solutions explicites du problème (4.34)

Intéressons-nous maintenant à la résolution du problème (4.34). Trois cas se présentent, selon la valeur de $\|\mathbf{n}_0\|$:

1. Si $\|\mathbf{n}_0\| = 1$, alors $\mathbf{n}^+ = \mathbf{n}^- = \mathbf{n}_0$ est la solution unique et exacte de (4.34).
2. Si $\|\mathbf{n}_0\| < 1$, alors l'ensemble des solutions du problème (4.39) est donné par :

$$\mathbf{n}_0 + \mathbb{R}\mathbf{v}_3 \quad (4.40)$$

où $\mathbf{v}_3$ est un vecteur unitaire du noyau de $\mathbf{S}$, dont on sait d'après le théorème du rang qu'il est de dimension 1. Les deux solutions unitaires suivantes :

$$\begin{cases} \mathbf{n}^+ = \mathbf{n}_0 + \sqrt{1 - \|\mathbf{n}_0\|^2} \mathbf{v}_3 \\ \mathbf{n}^- = \mathbf{n}_0 - \sqrt{1 - \|\mathbf{n}_0\|^2} \mathbf{v}_3 \end{cases} \quad (4.41)$$

sont les deux seules solutions du problème (4.34). Notons au passage que les expressions (4.41) donnent, a posteriori, la signification des exposants + et -.

3. Enfin, si $\|\mathbf{n}_0\| > 1$, le problème n'admet aucune solution exacte. Il est tentant de choisir la solution approchée $\mathbf{n}^+ = \mathbf{n}^- = \frac{\mathbf{n}_0}{\|\mathbf{n}_0\|}$ [213], qui n'est pourtant pas la solution de (4.34). Le calcul de la solution exacte, inspiré du chapitre 6 de [105], est détaillé ci-après.

Nous supposons dorénavant que $\|\mathbf{n}_0\| > 1$. En introduisant un multiplicateur de Lagrange λ , nous transformons le problème (4.34) en un problème non contraint :

$$\min_{\mathbf{n}, \lambda} \|\mathbf{S}\mathbf{n} - \mathbf{i}\|^2 + \lambda (\|\mathbf{n}\|^2 - 1) \quad (4.42)$$

La condition d'optimalité associée au problème (4.42) s'écrit :

$$\begin{cases} (\mathbf{S}^\top \mathbf{S} + \lambda \mathbf{I}_3) \mathbf{n} = \mathbf{S}^\top \mathbf{i} \\ \|\mathbf{n}\|^2 = 1 \end{cases} \quad (4.43)$$

En remplaçant $\mathbf{S}$ dans (4.43) par sa décomposition en valeurs singulières (4.35), on montre [92] que la résolution du système (4.43) revient à la recherche des zéros de la fonction suivante :

$$h(\lambda) = \left(\frac{\sigma_1 \mathbf{u}_1^\top \mathbf{i}}{\sigma_1^2 + \lambda} \right)^2 + \left(\frac{\sigma_2 \mathbf{u}_2^\top \mathbf{i}}{\sigma_2^2 + \lambda} \right)^2 - 1 \quad (4.44)$$

Grâce aux observations suivantes :

1. la fonction h est de classe C^1 sur $\mathbb{R}^+$,
2. $h(0) = \|\mathbf{n}_0\| - 1 > 0$,
3. $h' < 0$ sur $\mathbb{R}^+$,
4. $h(\lambda)$ tend vers -1 quand λ tend vers $+\infty$,

on peut affirmer que h admet une seule racine positive λ_0 . Quand $\|\mathbf{n}_0\| > 1$, une solution exacte de (4.34) s'obtient en reportant cette valeur de λ dans la première équation du système (4.43). On trouve finalement :

$$\mathbf{n} = \frac{\sigma_1 \mathbf{u}_1^\top \mathbf{i}}{\sigma_1^2 + \lambda_0} \mathbf{v}_1 + \frac{\sigma_2 \mathbf{u}_2^\top \mathbf{i}}{\sigma_2^2 + \lambda_0} \mathbf{v}_2 \quad (4.45)$$

Il est également prouvé dans [105] que les racines négatives de h conduisent à des solutions de (4.34) « moins optimales ». Pour calculer λ_0 , il nous faut déterminer la racine positive (unique) de l'équation :

$$\left(\frac{\sigma_1 \mathbf{u}_1^\top \mathbf{i}}{\sigma_1^2 + \lambda} \right)^2 + \left(\frac{\sigma_2 \mathbf{u}_2^\top \mathbf{i}}{\sigma_2^2 + \lambda} \right)^2 - 1 = 0 \quad (4.46)$$

Cette *équation séculaire* se réécrit aisément sous la forme d'une équation algébrique de degré 4 en λ , dont les solutions peuvent être obtenues par les formules de Ferrari-Cardan, mais il est bien plus rapide de résoudre l'équation (4.46) par la méthode de Newton. Il faut toutefois garder à l'esprit qu'une équation de ce type doit être résolue en chaque point (u, v) où $\|\mathbf{n}_0\| > 1$.

4.3.3 Désambiguïsation du problème par coupure de graphe

Le cas où $\|\mathbf{n}_0\| < 1$ étant de loin le plus fréquent, on peut construire « presque » $2^{|\Omega|}$ champs de normales, qui tous sont solutions du problème (4.1). Dans [266], il est proposé de contraindre davantage le problème en supposant que les normales sont distribuées selon une loi de Laplace, mais cette hypothèse n'est réaliste que pour des surfaces présentant une certaine symétrie. Comme cela a été mentionné dans le paragraphe 4.2.4, nous conseillons plutôt d'utiliser la contrainte d'intégrabilité du champ de normales pour

la désambiguïsation, ce qui est beaucoup moins restrictif : la surface est simplement supposée de classe $\mathcal{C}^2$, au moins par morceaux. Nous décrivons maintenant un algorithme efficace pour lever l'ambiguïté du problème (4.1), qui utilise cette contrainte.

Pour déterminer le champ de normales « le plus intégrable », une étiquette $l \in \{+, -\}$ est attribuée à chaque pixel, de façon à indiquer si la normale est égale à $\mathbf{n}^+$ ou à $\mathbf{n}^-$. Il nous faut donc résoudre le problème suivant :

$$\min_l \sum_{(u,v) \in \Omega} \left[\frac{\partial p^l}{\partial v}(u, v) - \frac{\partial q^l}{\partial u}(u, v) \right]^2 \quad (4.47)$$

où $[p^l, q^l]^\top = [-n_1^l/n_3^l, -n_2^l/n_3^l]^\top$. Remarquons que le critère retenu pour le problème (4.47) n'est rien d'autre qu'une version discrète du critère (4.32).

Pour calculer les dérivées spatiales du critère (4.47) à l'aide de différences finies d'ordre 1, il nous faut distinguer les quatre familles de *cliques* d'ordre 3 suivantes :

$$\begin{cases} \mathcal{C}_1^3 = \{ \{(u, v), (u-1, v), (u, v-1)\} \in \Omega^3 \} \\ \mathcal{C}_2^3 = \{ \{(u, v), (u+1, v), (u, v-1)\} \in \Omega^3 \} \\ \mathcal{C}_3^3 = \{ \{(u, v), (u-1, v), (u, v+1)\} \in \Omega^3 \} \\ \mathcal{C}_4^3 = \{ \{(u, v), (u+1, v), (u, v+1)\} \in \Omega^3 \} \end{cases} \quad (4.48)$$

Si nous notons (l_1, l_2, l_3) les étiquettes de chacun des trois pixels de chaque clique, dans l'ordre où ils sont définis dans (4.48), le problème (4.47) se réécrit de la façon suivante :

$$\min_l \sum_{c^3 \in \mathcal{C}^3} V_{c^3}^{\text{int}}(l_1, l_2, l_3) \quad (4.49)$$

où $\mathcal{C}^3 = \cup_{i=1}^4 \mathcal{C}_i^3$, et où le *potentiel* $V_{c^3}^{\text{int}}(l_1, l_2, l_3)$ mesure l'intégrabilité locale de la clique courante c^3 pour l'étiquetage courant l . Par exemple, si $c_1^3 \in \mathcal{C}_1^3$:

$$V_{c_1^3}^{\text{int}}(l_1, l_2, l_3) = \left[(p^{l_1}(u, v) - p^{l_3}(u, v-1)) - (q^{l_1}(u, v) - q^{l_2}(u-1, v)) \right]^2 \quad (4.50)$$

Le problème (4.49) est un problème d'étiquetage binaire, où le potentiel local dépend du pixel courant et de deux de ses voisins. De tels problèmes d'optimisation combinatoire ont été étudiés par Kolmogorov dans [149], où il est montré que l'algorithme de la « coupure de graphe » (*graph cut*) [35] peut être utilisé pour la minimisation de l'énergie $\sum_{c^3 \in \mathcal{C}^3} V_{c^3}^{\text{int}}(l_1, l_2, l_3)$, pourvu que sa *régularité* soit garantie, ce qui s'écrit ici :

$$\begin{cases} V_{c^3}^{\text{int}}(+, +, l_3) + V_{c^3}^{\text{int}}(-, -, l_3) \leq V_{c^3}^{\text{int}}(+, -, l_3) + V_{c^3}^{\text{int}}(-, +, l_3) \\ V_{c^3}^{\text{int}}(+, l_2, +) + V_{c^3}^{\text{int}}(-, l_2, -) \leq V_{c^3}^{\text{int}}(+, l_2, -) + V_{c^3}^{\text{int}}(-, l_2, +) \\ V_{c^3}^{\text{int}}(l_1, +, +) + V_{c^3}^{\text{int}}(l_1, -, -) \leq V_{c^3}^{\text{int}}(l_1, +, -) + V_{c^3}^{\text{int}}(l_1, -, +) \end{cases} \quad (4.51)$$

pour tout $c^3 \in \mathcal{C}^3$ et tout $(l_1, l_2, l_3) \in \{+, -\}^3$. Bien entendu, ces inégalités n'ont aucune raison d'être vérifiées.

4.3.4 Régularisation « minimale »

Pour garantir la condition de régularité, nous transformons le problème (4.49) en :

$$\min_l \sum_{c^3 \in \mathcal{C}^3} V_{c^3}^{\text{int}}(l_1, l_2, l_3) + \sum_{c^2 \in \mathcal{C}^2} V_{c^2}^{\text{Ising}}(l_1, l_2) \quad (4.52)$$

où $\mathcal{C}^2 = \cup_{i=1}^4 \mathcal{C}_i^2$ regroupe les quatre familles de cliques d'ordre 2 suivantes :

$$\begin{cases} \mathcal{C}_1^2 = \{ \{(u, v), (u-1, v)\} \in \Omega^2 \} \\ \mathcal{C}_2^2 = \{ \{(u, v), (u, v-1)\} \in \Omega^2 \} \\ \mathcal{C}_3^2 = \{ \{(u, v), (u-1, v-1)\} \in \Omega^2 \} \\ \mathcal{C}_4^2 = \{ \{(u, v), (u+1, v-1)\} \in \Omega^2 \} \end{cases} \quad (4.53)$$

et où $V_{c^2}^{\text{Ising}}(l_1, l_2)$ est défini ci-après, en notant (l_1, l_2) les étiquettes des deux pixels de chaque clique $c^2 \in \mathcal{C}^2$, dans l'ordre où ils apparaissent dans (4.53) :

$$V_{c^2}^{\text{Ising}}(l_1, l_2) = \beta_{c^2} \delta(l_1 \neq l_2) \quad (4.54)$$

Dans (4.54), β_{c^2} est un coefficient positif ou nul défini *localement*, i.e. dont la valeur doit être définie pour chaque clique $c^2 \in \mathcal{C}^2$. En imposant la condition de régularité à l'énergie du problème (4.52), on déduit une borne inférieure pour chaque β_{c^2} . Par exemple, si $c_3^2 = \{(u, v), (u-1, v-1)\} \in \mathcal{C}_3^2$:

$$\beta_{c_3^2} \geq \frac{1}{2} \max \left\{ 0, \max_{(l_3, l_4) \in \{+, -\}^2} \left\{ \Delta V_{c_3^2}^{\text{int}}(l_3), \Delta V_{c_2^2}^{\text{int}}(l_4) \right\} \right\} \quad (4.55)$$

où $c_3^3 = c_3^2 \cup (u, v-1)$, $c_2^3 = c_3^2 \cup (u-1, v)$, et (l_3, l_4) désignent les étiquettes de $(u, v-1)$ et de $(u-1, v)$, respectivement. Enfin, pour $j \in \{2, 3\}$ et $k \in \{+, -\}$:

$$\Delta V_{c_j^3}^{\text{int}}(k) = V_{c_j^3}^{\text{int}}(k, +, +) + V_{c_j^3}^{\text{int}}(k, -, -) - V_{c_j^3}^{\text{int}}(k, +, -) - V_{c_j^3}^{\text{int}}(k, -, +) \quad (4.56)$$

Du point de vue markovien, notre approche consiste à utiliser un *modèle d'Ising non stationnaire* [71], de même par exemple que pour les modèles de type « verre de spins » (*spin glasses*) [133]. Grâce aux inégalités telles que (4.55), il est possible de prédire, pour chaque clique $c^2 \in \mathcal{C}^2$, la plus petite valeur du coefficient β_{c^2} permettant de garantir la condition de régularité. Ces coefficients ne doivent donc pas être considérés comme des paramètres, ce qui constitue un avantage supplémentaire. Ceci nous permet de limiter la régularisation de façon à simplement garantir la régularité au sens de Kolmogorov [149], en évitant tout sur-lissage. En effet, si les coefficients β_{c^2} ne sont pas fixés à leurs valeurs minimales, l'étiquetage est sur-lissé, ce qui biaise la reconstruction 3D. Ce phénomène est illustré sur la figure 4.7.

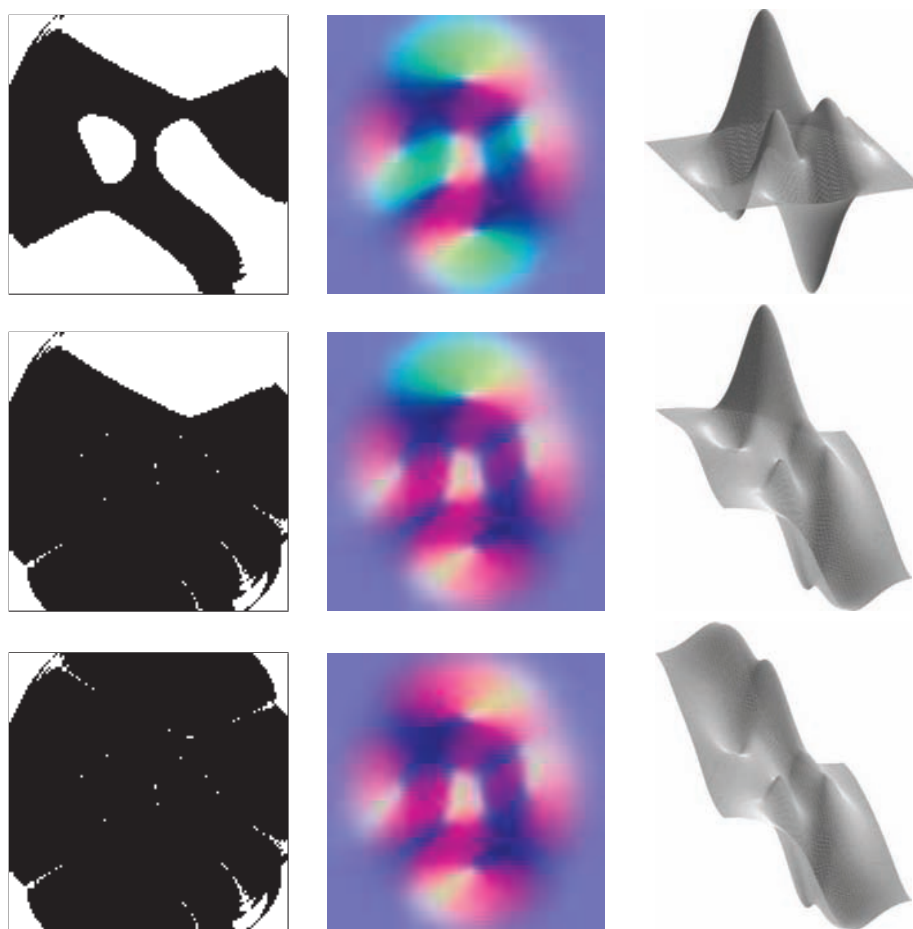


FIGURE 4.7 – Influence des paramètres de régularisation : un sur-lissage apparaît lorsque des valeurs élevées des coefficients β_{c2} sont utilisées, sur l'exemple de la figure 4.6. De gauche à droite : étiquetages obtenus (noir pour $\mathbf{n}^+$, blanc pour $\mathbf{n}^-$), champs de normales estimés et reliefs 3D, en fixant les coefficients β_{c2} , de haut en bas, à 5 fois, 10 fois ou 100 fois les valeurs minimales permettant de garantir la condition de régularité.

4.3.5 Une méthode efficace de résolution du problème PS2

Pour résumer, la méthode de résolution du problème PS2 que nous préconisons comporte deux étapes :

1. Calcul des champs de normales $\mathbf{n}^+$ et $\mathbf{n}^-$ selon la procédure indiquée dans le paragraphe 4.3.2.
2. Désambiguïsation du problème en utilisant le critère d'intégrabilité, un modèle d'Ising, les valeurs minimales des coefficients β_{c2} permettant de garantir la condition de régularité, et l'algorithme de la coupure de graphe.

En pratique, ces deux étapes sont très rapides : les solutions initiales $\mathbf{n}^+$ et $\mathbf{n}^-$ sont explicites, et la complexité de l'optimisation par coupure de graphe est polynomiale.

Le résultat de la figure 4.8 a été obtenu en appliquant cette méthode à l'exemple de la figure 4.6. Le champ de normales estimé, tout comme le relief correspondant, sont très similaires à ceux de la vérité terrain (cf. figure 4.6). Le temps CPU pour cet exemple (de taille 128×128) est inférieur à 1 *ms* (les tests sont réalisés en C++ sur un processeur I7 cadencé à 2,9 *GHz*), ce qui confirme l'efficacité de la méthode proposée.

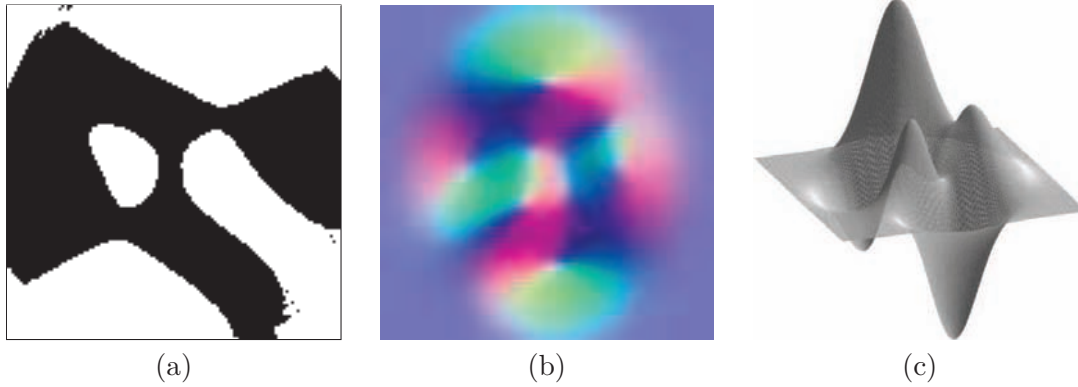


FIGURE 4.8 – En minimisant l'intégrabilité à partir des champs $\mathbf{n}^+$ et $\mathbf{n}^-$ de la figure 4.6, on obtient l'étiquetage (a). Le champ de normales (b) correspondant à cet étiquetage donne le relief (c) après intégration.

L'introduction de termes de régularisation correspond cependant à un a priori sur la régularité de l'étiquetage l . Si l'étiquetage est effectué comme cela est préconisé dans le paragraphe 4.3.2, il sera régulier dès lors que la surface est régulière, puisqu'il s'agit d'un processus déterministe (cf. figure 4.8-a). Le terme d'a priori du modèle d'Ising est donc justifié. Cependant, si nous simulons une hypothétique distribution moins lisse des étiquetages entre les deux solutions $\mathbf{n}^+$ et $\mathbf{n}^-$, en échangeant aléatoirement une proportion r d'étiquettes, le résultat se dégrade très nettement (cf. figures 4.9 et 4.10).

4.4 Applications

Nous montrons maintenant pourquoi la résolution numérique efficace du problème PS2 n'est pas un simple exercice de style, en exhibant deux situations où la méthode proposée peut être utile : la stéréophotométrie à partir d'une seule image RGB sous $m = 3$ éclairages colorés, et le cas où les images sont prises en extérieur lors d'une journée ensoleillée.

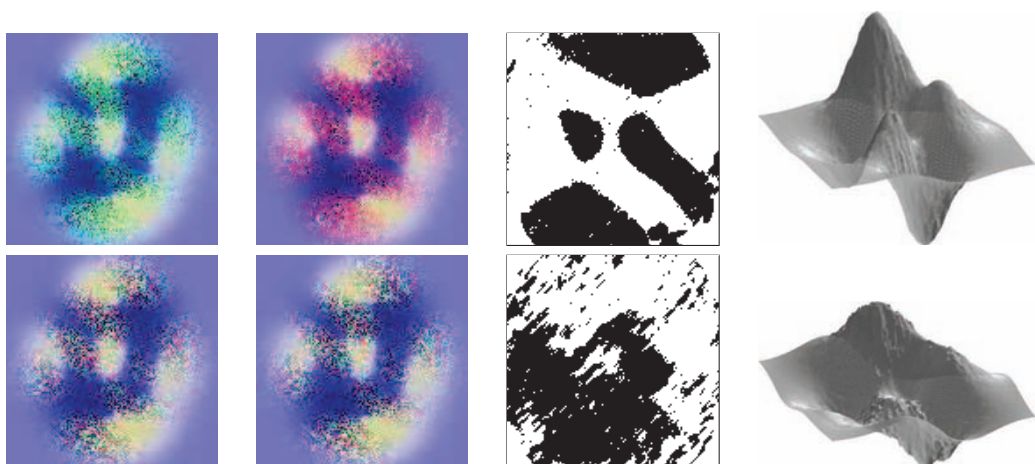


FIGURE 4.9 – En échangeant une proportion $r = 0,2$ (première ligne) ou $r = 0,5$ (deuxième ligne) de valeurs entre $\mathbf{n}^+$ (première colonne) et $\mathbf{n}^-$ (deuxième colonne), les résultats se dégradent très nettement (voir aussi la figure 4.10). En effet, même si l'influence des termes de régularisation est limitée en choisissant les valeurs minimales des paramètres β_{c2} , les potentiels d'Ising s'opposent à un étiquetage qui ne soit pas uniforme par morceaux, et donc lissent excessivement l'étiquetage (troisième colonne), ce qui induit un biais important dans la reconstruction 3D (quatrième colonne). Il est donc important de favoriser un étiquetage lisse, ce qui est le cas de la méthode algébrique proposée dans le paragraphe 4.3.2.

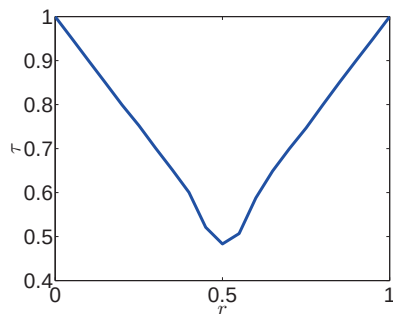


FIGURE 4.10 – Proportion τ d'étiquetages corrects en fonction de la proportion r d'échanges entre $\mathbf{n}^+$ et $\mathbf{n}^-$ (cf. figure 4.9). Les résultats sont excellents lorsque les deux champs de normales $\mathbf{n}^+$ et $\mathbf{n}^-$ sont lisses : dans les cas $r = 0$, correspondant à la donnée illustrée sur la figure 4.6, et $r = 1$, correspondant à un échange intégral des champs $\mathbf{n}^+$ et $\mathbf{n}^-$, on trouve $\tau = 1$. Lorsque l'on échange aléatoirement $\mathbf{n}^+$ et $\mathbf{n}^-$, la proportion d'étiquetages corrects décroît linéairement, jusqu'à atteindre la valeur $\tau = 0,5$ lorsque $r = 0,5$ (ce cas correspond à la deuxième ligne de la figure 4.9).

4.4.1 Stéréophotométrie robuste avec seulement $m = 3$ images

Reprenons l'exemple de la reconstruction 3D du buste de Beethoven de la figure 1.2, que nous présentons à nouveau sur la figure 4.11. La reconstruction 3D exacte obtenue en inversant en chaque point le problème (1.18) semble satisfaisante (figure 4.11-d), mais l'estimation de l'albédo, qui devrait être uniforme puisque le matériau est homogène, est décevante (cf. figure 4.11-e).

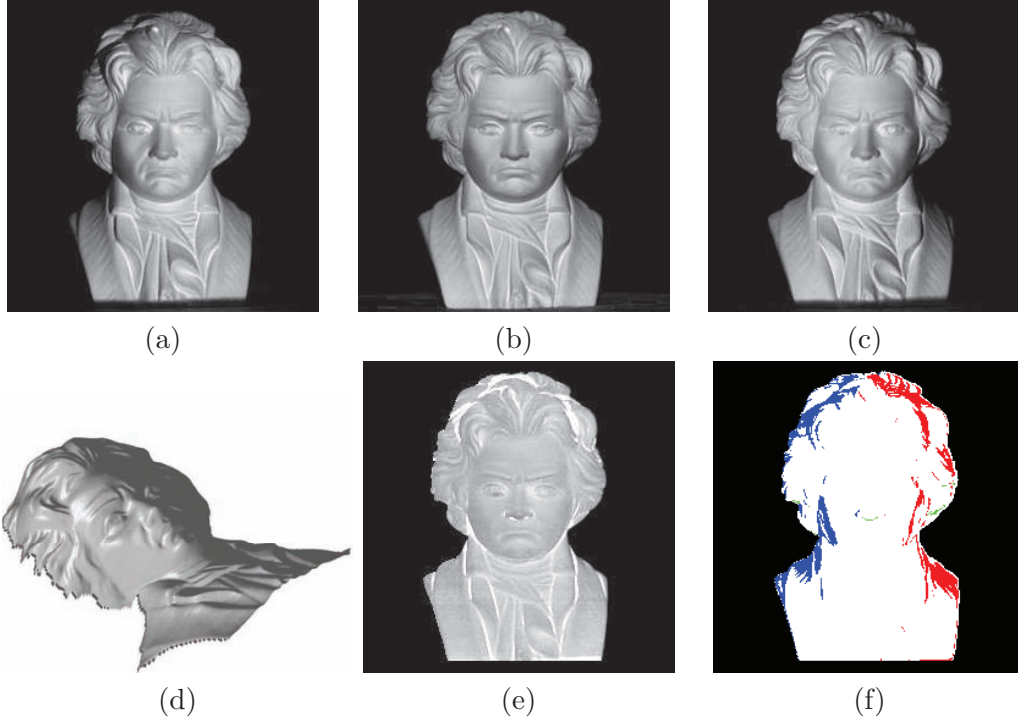


FIGURE 4.11 – (a)-(b)-(c) Trois photographies d'un buste de Beethoven en plâtre. (d)-(e) Relief et albédo estimés par la méthode PS3. (f) Partition du domaine Ω en quatre sous-domaines, suivant le problème à résoudre : $\Omega_3^{1,2,3}$ (blanc), $\Omega_2^{1,2}$ (rouge), $\Omega_2^{1,3}$ (vert), $\Omega_2^{2,3}$ (bleu), en choisissant un seuil égal à 20% du niveau de gris maximal.

Les points où l'estimation est biaisée correspondent aux zones d'ombre. Le problème des ombres en stéréophotométrie est bien connu. Nous avons présenté dans le chapitre 1 plusieurs méthodes d'estimation robuste permettant de le contourner lorsque $m > 3$ [20, 51, 131, 245, 257, 270, 272], mais ces méthodes ne sont pas adaptées au cas $m = 3$. Le seul travail proposant une estimation robuste pour ce cas est [119], à l'aide d'une régularisation qui permet de lisser les phénomènes de lignage dus à la propagation de l'information de profondeur le long des caractéristiques (cf. figure 4.15).

Lorsque l'albédo est uniforme, comme cela est le cas pour l'exemple de la figure 4.11, l'histogramme de son estimation par PS3 présente un pic important, situé ici à $\rho = 0,74$. Il semble raisonnable de considérer cette valeur comme la valeur « réelle » de l'albédo

sur l'ensemble de la surface. On peut alors appliquer la méthode PS2 dans les zones d'ombre : à partir d'un simple seuillage des niveaux de gris, la figure 4.11-f montre comment partitionner le domaine Ω en quatre sous-domaines : $\Omega_3^{1,2,3}$ correspond aux points éclairés dans les trois images, $\Omega_2^{1,2}$ à ceux éclairés dans la première et la deuxième image, et idem pour $\Omega_2^{1,3}$ et $\Omega_2^{2,3}$. Une telle combinaison des approches PS3 et PS2 améliore indiscutablement la finesse de la reconstruction 3D (cf. figure 4.12). Remarquons cependant que, pour imposer l'intégrabilité du champ de normales sur Ω , il est préférable d'appliquer l'algorithme de la coupure de graphe sur le domaine Ω tout entier, plutôt que de façon indépendante sur chaque sous-domaine. Pour procéder ainsi, il suffit de dupliquer la normale estimée en chaque point de $\Omega_3^{1,2,3}$, c'est-à-dire de traiter ces points comme s'il s'agissait de points singuliers.

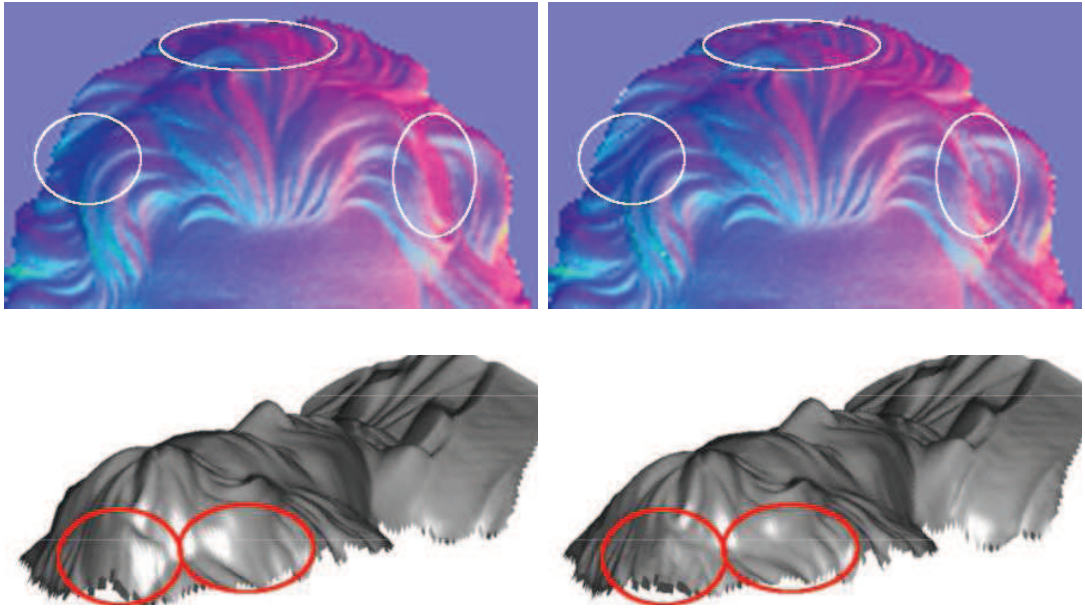


FIGURE 4.12 – Normales estimées près des cheveux (première ligne) et relief 3D complet (seconde ligne), obtenus à partir des trois photographies de la première ligne de la figure 4.11, en utilisant soit l'approche PS3 (à gauche), soit une combinaison de PS3 et PS2 (à droite), qui permet de mieux retrouver les détails du relief.

4.4.2 Stéréophotométrie couleur

Hernández et al. font l'observation suivante dans [119] : « L'utilisation de la stéréophotométrie avec seulement trois images peut apparaître comme une restriction déraisonnable. Il existe toutefois une situation particulière où l'on ne dispose que de trois images. Cette technique est connue sous le nom de stéréophotométrie couleur »³. En effet, la

3. « *Using photometric stereo on just three images may seem like an unreasonably hard restriction. There is, however, a particular situation when only three images are available. This technique is known as color photometric stereo.* »

stéréophotométrie sous des éclairages colorés constitue l'application la plus naturelle du problème PS2, puisque l'on est nécessairement confronté au cas $m = 3$.

Kontsevich et al. montrent dans [150] comment reconstruire une surface blanche éclairée par $m = 3$ sources lumineuses colorées. En effet, la plupart des images en couleur sont constituées de trois canaux. En assimilant chaque canal à une image en niveaux de gris, il est possible d'appliquer la technique PS3 en utilisant une seule image RGB. Une scène déformable telle qu'un visage peut ainsi être reconstruite en temps réel, même s'il n'est pas immobile. Remarquons cependant que l'albédo doit être le même dans chaque canal, ce qui nécessite de maquiller préalablement la scène.

Hernández et al. montrent dans [117] « comment un éclairage à spectres séparés permet surtout d'acquérir trois images (avec des directions d'éclairage différentes) en une seule prise de vue, rendant ainsi possible la reconstruction 3D par stéréophotométrie image par image »⁴. Un exemple d'une telle image RGB, extraite d'une séquence vidéo⁵ est montré sur la figure 4.13, ainsi que les trois canaux de couleur associés (rouge, vert et bleu).

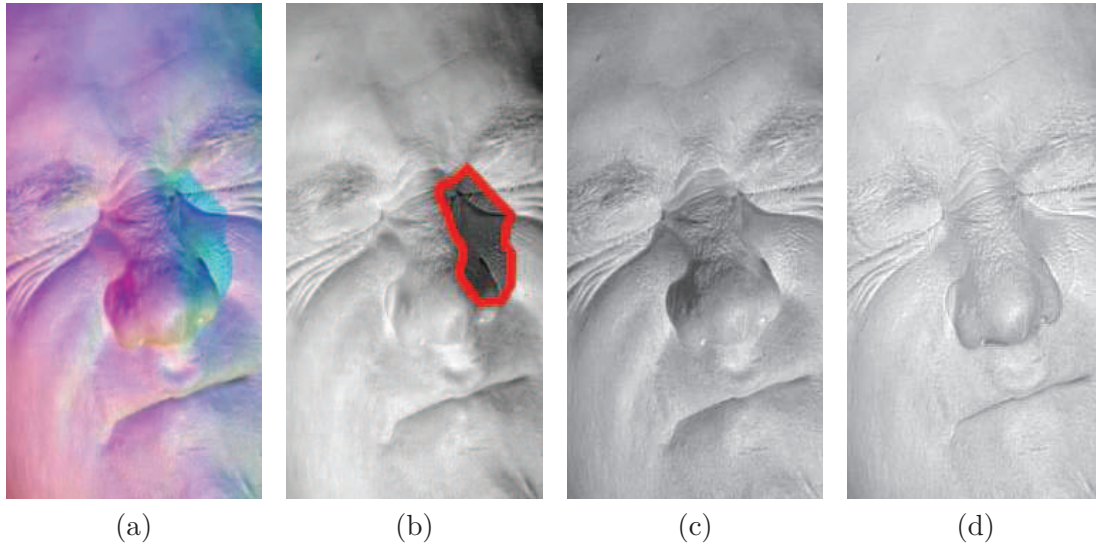


FIGURE 4.13 – (a) Image RGB d'un visage éclairé par trois éclairages directionnels, non coplanaires et colorés, de directions et d'intensités connues. (b)-(c)-(d) Décomposition de cette image en trois canaux (rouge, vert et bleu). La forme de l'ombre portée, indiquée sur la figure (b), peut être comparée à celle de la reconstruction 3D du nez sur la figure 4.14-a.

4. « how multispectral lighting allows one to essentially capture three images (each with a different light direction) in a single snapshot, thus making per-frame photometric reconstruction possible. »

5. <http://mi.eng.cam.ac.uk/research/projects/VideoNormals/>

On peut appliquer la technique PS3 à l'image de la figure 4.13-a mais, sans un traitement spécifique, le résultat de la reconstruction 3D est biaisé autour du nez (cf. figure 4.14-a), où une ombre rend le canal rouge inutilisable (cf. figure 4.13-b). On peut toutefois appliquer la technique PS2 dans cette zone, pourvu que l'albédo soit connu, puisque les canaux vert et bleu sont utilisables. Comme nous l'avons déjà suggéré, l'albédo « réel » (supposé uniforme) peut être estimé à partir de l'histogramme des estimations de ρ par PS3. Nous pouvons donc appliquer la technique PS2 à cette zone⁶ : le résultat est bien meilleur (cf. figure 4.14-b).

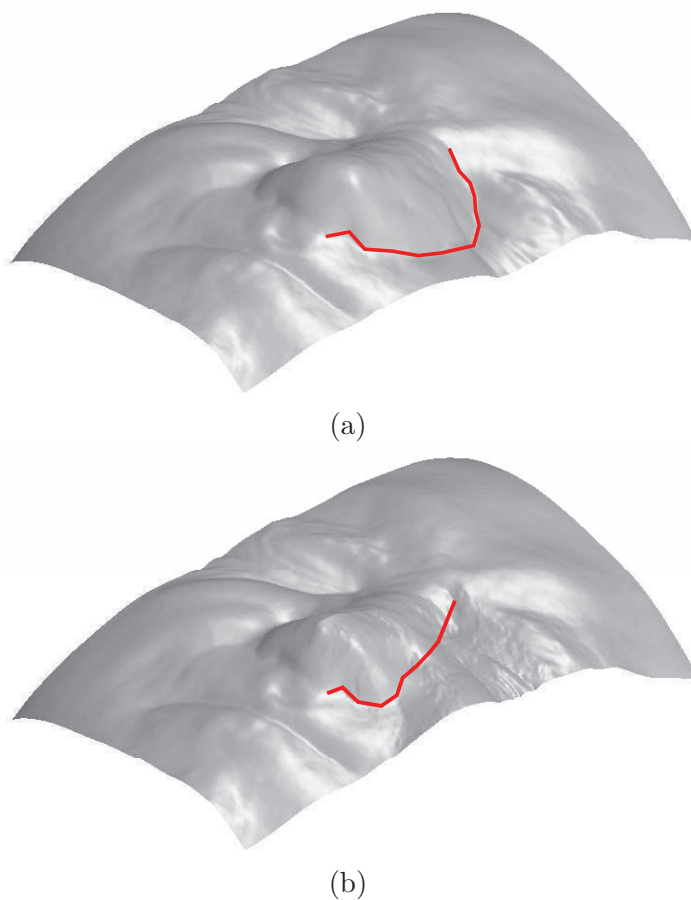


FIGURE 4.14 – Reconstructions 3D obtenues à partir de l'image RGB de la figure 4.13-a, en utilisant : (a) PS3 seul ; (b) une combinaison de PS3 et PS2. La reconstruction 3D autour du nez est de bien meilleure qualité en utilisant cette dernière approche.

6. Pour des raisons de robustesse, nous appliquons PS3 aux points pour lesquels deux des trois niveaux de couleur sont inférieurs au seuil de détection des ombres, qui a été fixé pour cet exemple à 47% du niveau de couleur maximal.

Le problème que constituent les ombres en stéréophotométrie couleur a également été étudié dans [119]. La parade proposée consiste à éliminer le canal incriminé, i.e. à résoudre un problème PS2. Les auteurs résolvent ce problème via la formulation différentielle, en recourant à une régularisation afin d'éviter les phénomènes de lignage (cf. figure 4.15) : nous venons toutefois de prouver qu'en théorie, il n'est pas nécessaire d'introduire d'autre lissage que la contrainte d'intégrabilité.

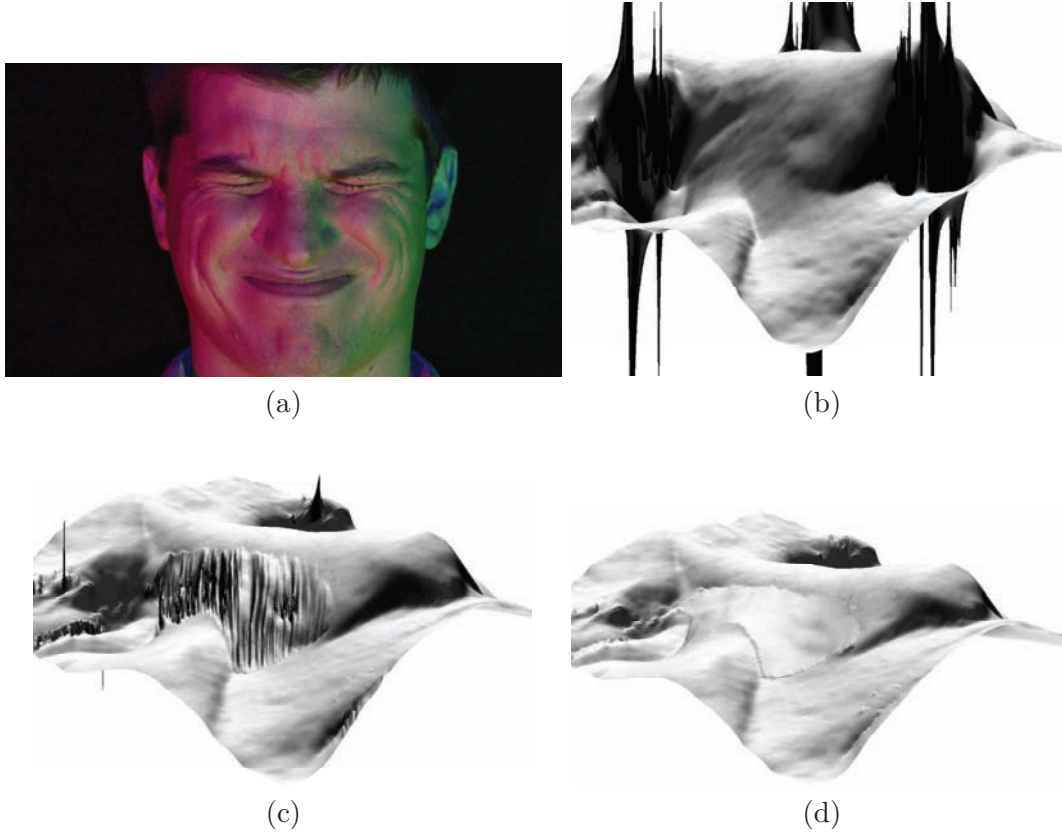


FIGURE 4.15 – (a) L'image RGB d'une surface éclairée par trois sources colorées permet d'effectuer (b) la reconstruction 3D par stéréophotométrie [150]. L'estimation des normales est cependant biaisée aux points où l'une des sources colorées est occultée, ce qui se traduit après intégration par des pics. (c) L'utilisation de la technique PS2 en ces points permet de limiter ce biais, mais des phénomènes de lignage apparaissent, car l'information de profondeur est propagée le long des caractéristiques (qui sont, ici, des droites parallèles). (d) Pour gommer ces phénomènes indésirables, on peut régulariser la profondeur dans la direction orthogonale aux caractéristiques, ce qui nécessite d'introduire un hyper-paramètre permettant de contrôler le poids de la régularisation [119].

4.4.3 Éclairages coplanaires

Photographies prises en extérieur

Les données de nombreuses webcams sont aujourd’hui disponibles à la consultation sur Internet, quasiment en temps réel. Le projet AMOS (*the Archive of Many Outdoor Scenes*) [135] totalisait, au moment où l’article [213] a été écrit, plus de 650 millions d’images acquises par près de 25000 caméras réparties en divers endroits du globe, par exemple au centre Pompidou à Paris, sous la Statue de la Liberté à New York (cf. figure 4.16) ou sur les étendues gelées de l’Antarctique. Chaque image est accompagnée des coordonnées GPS de la webcam, ainsi que de la date et de l’heure de prise de vue.

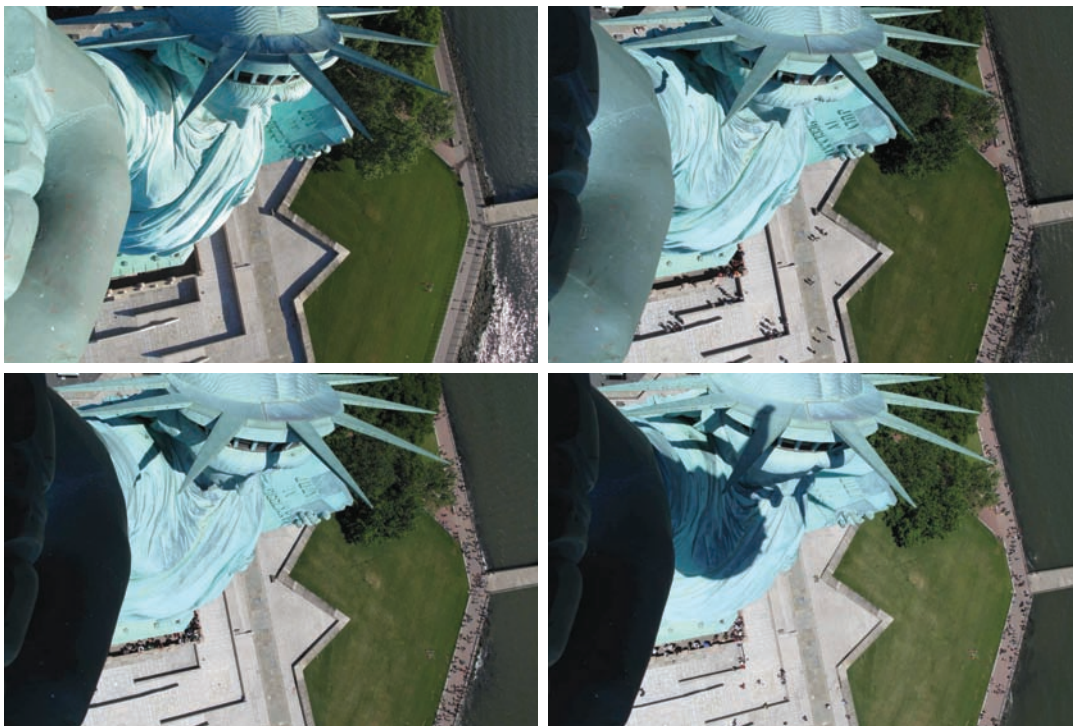


FIGURE 4.16 – Images de la Statue de la Liberté prises le 5 juillet 2014 entre 11h30 et 17h30 (source : projet AMOS). La course du soleil permet de fournir des images de la scène prises sous différents éclairages coplanaires, qui constituent les données d’un problème PS2.

Parmi les applications envisageables, il a été proposé dans [1, 3] d’utiliser ces immenses bases de données pour calculer le relief et la réflectance des scènes observées. Dans ces articles, il est montré comment mettre en œuvre la stéréophotométrie, en utilisant les variations d’éclairage sur de très longues périodes (plusieurs saisons, voire plusieurs années). En effet, à partir des coordonnées GPS de la webcam, de la date et de l’heure de prise de vue, il est possible d’estimer précisément la position du soleil re-

lativement au référentiel terrestre [223]. La pose de la webcam peut être estimée, par exemple, en ajustant aux images les paramètres d'un modèle statistique décrivant le ciel [157]. Une fois ces étapes accomplies, les m directions d'éclairage $\mathbf{s}^i$, $i \in [1, m]$, sont parfaitement connues dans un repère lié à la caméra, mais cela n'est pas le cas des m intensités $\phi^i \in \mathbb{R}$, qui dépendent de phénomènes météorologiques difficiles à prédire.

L'estimation des inconnues $\mathbf{n}$, ρ et ϕ^i , à partir d'images en niveaux de gris, est reformulée dans [1, 3] comme un problème d'optimisation en moindres carrés pondérés :

$$\min_{\substack{\mathbf{n}: \Omega \rightarrow \mathbb{S}^2 \\ \rho: \Omega \rightarrow \mathbb{R} \\ \phi^i \in \mathbb{R}^+, i \in [1, m]}} \sum_{(u,v) \in \Omega} \sum_{i=1}^m w^i(u,v) \left[I^i(u,v) - \rho(u,v) \phi^i \overline{\mathbf{s}^i} \cdot \mathbf{n}(u,v) \right]^2 \quad (4.57)$$

où les w^i sont des poids permettant de filtrer les images « corrompues » par des conditions météorologiques non favorables, qui sont déterminés de façon heuristique en analysant par exemple la couleur du ciel. La minimisation est réalisée au moyen d'algorithmes génériques d'optimisation en moindres carrés non linéaires.

L'idée sous-jacente à ces méthodes est que les directions d'éclairage $\overline{\mathbf{s}^i}$ présentent une variabilité suffisante pour garantir le caractère bien posé du problème. En pratique, un grand nombre d'images sont utilisées (plusieurs milliers), et la qualité des résultats obtenus force à reconnaître que cette approche de type *big data* est pertinente.

Nous pensons cependant qu'elle pourrait tirer bénéfice d'une étude plus approfondie du cas où les photographies ont toutes été prises le même jour, i.e. où toutes les directions d'éclairage sont coplanaires. La technique PS2 doit alors être appliquée puisque ce cas de figure est dégénéré. En effet, quel que soit $m \geq 2$, la matrice d'éclairage $\mathbf{S}$ est de rang 2, ce qui rend le problème (1.18) mal posé, même si l'albédo est connu. En théorie, comme le plan contenant la direction des rayons du soleil varie d'un jour à l'autre, le rang de $\mathbf{S}$ est égal à 3 dès que ces images sont prises sur une période de plusieurs jours, mais la matrice $\mathbf{S}$ correspondante est très mal conditionnée s'il s'agit d'une période trop courte. On comprend alors mieux pourquoi seuls des jeux d'images prises en différentes saisons sont utilisés dans [1, 3] : cela permet de rendre le problème (1.18) à la fois bien posé et bien conditionné, sans introduire d'hypothèse supplémentaire sur la scène observée.

Cependant, il est parfois impossible d'utiliser des images prises en différentes saisons, par exemple si le relief de la scène est fixe à l'échelle d'une journée, mais susceptible d'évoluer notablement à l'échelle de plusieurs mois. Cela est le cas, par exemple, d'une webcam observant un glacier. Réaliser la reconstruction 3D à partir de séquences prises au cours d'une seule journée permettrait de relever quotidiennement le relief des glaciers, afin d'affiner les relevés stéréoscopiques [205]. L'approche proposée constitue un premier pas vers l'utilisation de la stéréophotométrie pour résoudre de tels problèmes. L'albédo ρ doit être connu, ou du moins homogène, une hypothèse qui, justement, semble réaliste dans le cas des glaciers.

Résolution numérique de la stéréophotométrie sous des éclairages coplanaires

A priori, il suffit, pour résoudre le problème de la stéréophotométrie à $m \geq 2$ éclairages coplanaires, d'adapter la méthode PS2 décrite dans le paragraphe 4.3.5. Cependant, pour des photographies réelles prises en extérieur, les intensités ϕ^i sont a priori inconnues. Nous pouvons donc modéliser ce nouveau problème sous la forme suivante (on rappelle que l'albédo est supposé connu, et nous supposons pour l'instant que les images sont en niveaux de gris) :

$$\begin{cases} \min_{\mathbf{n}, \{\phi^i\}} \sum_{(u,v) \in \Omega} \sum_{i=1}^m \left[I^i(u,v) - \rho(u,v) \phi^i \bar{\mathbf{s}}^i \cdot \mathbf{n}(u,v) \right]^2 \\ \text{s.c.} \quad \begin{cases} \|\mathbf{n}(u,v)\| = 1, \forall (u,v) \in \Omega \\ \frac{\partial}{\partial v} \left(-\frac{n_1(u,v)}{n_3(u,v)} \right) - \frac{\partial}{\partial u} \left(-\frac{n_2(u,v)}{n_3(u,v)} \right) = 0, \forall (u,v) \in \Omega \end{cases} \end{cases} \quad (4.58)$$

Il faut donc déterminer la normale $\mathbf{n}$ en chaque point (u,v) , et l'intensité lumineuse ϕ^i pour chaque image, de façon à minimiser l'écart quadratique entre les données et les images « reprojetées ». Parmi l'infinité de champs de normales permettant de minimiser ce critère, nous retenons le plus intégrable.

L'optimisation peut être effectuée itérativement, en alternant les estimations en $\mathbf{n}$ et en ϕ^i , à partir d'une solution initiale $\mathbf{n}^0 = [0, 0, 1]^\top$ (plan horizontal) et $\phi^{i,0} = 1$, $i \in [1, m]$ (éclairages de même intensité). La mise à jour des normales s'effectue au moyen de l'algorithme de résolution du problème PS2 décrit dans le paragraphe 4.3.5. La condition d'optimalité du problème (4.58) vis-à-vis des variables ϕ^i s'écrit, pour $i \in [1, m]$:

$$\phi^{i,k+1} = \frac{\sum_{(u,v) \in \Omega} I^i(u,v) \rho(u,v) \bar{\mathbf{s}}^i \cdot \mathbf{n}^k(u,v)}{\sum_{(u,v) \in \Omega} \left[\rho(u,v) \bar{\mathbf{s}}^i \cdot \mathbf{n}^k(u,v) \right]^2} \quad (4.59)$$

Afin de vérifier expérimentalement si ce schéma converge, et si oui, à quelle vitesse, nous simulons les images du relief synthétique de la figure 4.6 sous $m = 100$ éclairages dont les directions sont contenues dans un plan vertical. L'angle de la direction d'éclairage avec la verticale varie uniformément entre $-\pi/6$ et $\pi/6$, c'est-à-dire que le soleil est au zénith à midi, afin de limiter les ombres, qui sont toutefois simulées en seillant à zéro les valeurs négatives du niveau de gris. Les intensités lumineuses ϕ^i sont tirées aléatoirement selon une loi uniforme sur $[0, 1]$.

Nous évaluons l'évolution de l'erreur de reprojection :

$$E_r = \sqrt{\frac{1}{m|\Omega|} \sum_{(u,v) \in \Omega} \sum_{i=1}^m \left[I^i(u,v) - \rho(u,v) \phi^i \bar{\mathbf{s}}^i \cdot \mathbf{n}(u,v) \right]^2} \quad (4.60)$$

et celle de l'erreur d'intégrabilité :

$$E_i = \sqrt{\frac{1}{|\Omega|} \sum_{(u,v) \in \Omega} \left[\frac{\partial}{\partial v} \left(-\frac{n_1(u,v)}{n_3(u,v)} \right) - \frac{\partial}{\partial u} \left(-\frac{n_2(u,v)}{n_3(u,v)} \right) \right]^2} \quad (4.61)$$

en fonction du nombre d'itérations k . Les résultats de la figure 4.17 montrent une convergence très rapide de ces deux critères.

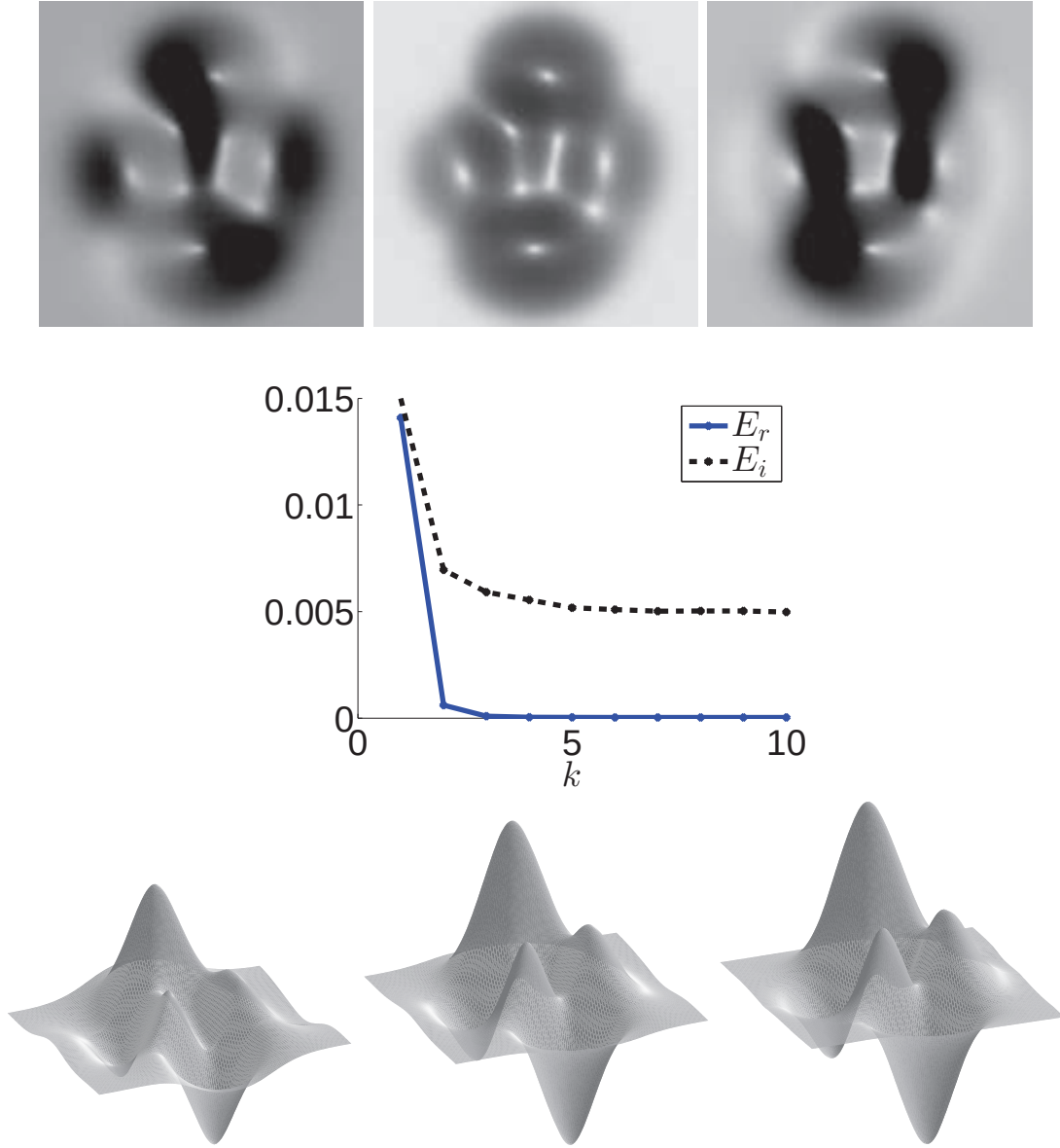


FIGURE 4.17 – Première ligne : trois des $m = 100$ images simulées. Deuxième ligne : évolution de E_r et E_i en fonction du nombre d'itérations k . Troisième ligne : relief estimé pour $k = 1$, $k = 2$ et $k = 10$.

Éclairages colorés

Nous avons, pour simplifier, supposé dans le paragraphe 4.4.3 que les images étaient en niveaux de gris. Il est également possible d'utiliser des images RGB, en supposant que la couleur observée dépend uniquement du spectre incident, et non de la texture de l'objet. Toute notion de couleur est alors entendue relativement à la scène, qui est supposée uniformément blanche et définit le blanc de référence. Ceci restreint toutefois notre étude à des objets non texturés, le cas d'objets texturés étant laissé en perspective. Sous cette hypothèse, le nouveau modèle lambertien s'écrit :

$$\left[I_R^i(u, v), I_G^i(u, v), I_B^i(u, v) \right] = \rho(u, v) \bar{\mathbf{s}}^i \cdot \mathbf{n}(u, v) \left[\phi_R^i, \phi_G^i, \phi_B^i \right] \quad (4.62)$$

où ϕ_R^i , ϕ_G^i et ϕ_B^i désignent les intensités lumineuses des canaux R, G et B.

On peut ainsi retrouver l'évolution du spectre au cours de la journée, en même temps que la géométrie de la scène, car l'algorithme précédent s'adapte sans difficulté : l'estimation des intensités est effectuée canal par canal, et le système linéaire (1.18) devient un système de $3m$ équations correspondant aux trois canaux (l'estimation de $\mathbf{n}$ n'en est que plus robuste). Ceci est particulièrement utile dans le cas des webcams car, en général, la balance des blancs s'adapte automatiquement à la couleur de l'éclairage, i.e. les rapports $\frac{\phi_R^i}{\phi_G^i}$, $\frac{\phi_R^i}{\phi_B^i}$ et $\frac{\phi_G^i}{\phi_B^i}$ sont susceptibles de varier avec i .

Un exemple de reconstruction 3D d'une partie de la Statue de la Liberté est montré sur la figure 4.18. Les intensités colorées estimées, relativement à la couleur de la statue, correspondent bien, du moins qualitativement, aux variations de couleur perceptibles sur les images (cf. la ligne du haut de la figure 4.18). Par exemple, l'image I^{12} comporte une ombre qui provoque une chute brutale des intensités. Sur cette image, la couleur dominante passe également du vert au bleu, ce qui a bien été retrouvé par notre méthode de résolution (cf. la courbe du bas de la figure 4.18).

4.5 Conclusion du chapitre

Dans ce chapitre, nous avons présenté une étude théorique du problème de la reconstruction 3D par stéréophotométrie dans le cas dégénéré où $m = 2$ images sont utilisées et où l'albédo est connu. Cette étude s'étend au cas de $m \geq 3$ éclairages coplanaires. Nous avons choisi la formulation non différentielle du problème PS2 pour effectuer sa résolution numérique : en chaque pixel, deux normales expliquent les niveaux de gris. Nous avons proposé une façon efficace de déterminer le champ de normales le plus intégrable. Notre principale contribution consiste à avoir adapté l'algorithme de la coupure de graphe pour résoudre un problème combinatoire où le critère à optimiser est l'intégrabilité, et à ajouter un terme d'Ising non stationnaire pour satisfaire la condition de régularité de l'énergie. Nous avons montré que cette approche permettait d'estimer la normale à partir d'une seule photographie RGB, et qu'elle permettait aussi d'estimer les intensités des éclairages colorés et la géométrie d'une scène à partir d'images en couleur prises par une webcam au cours d'une journée ensoleillée.

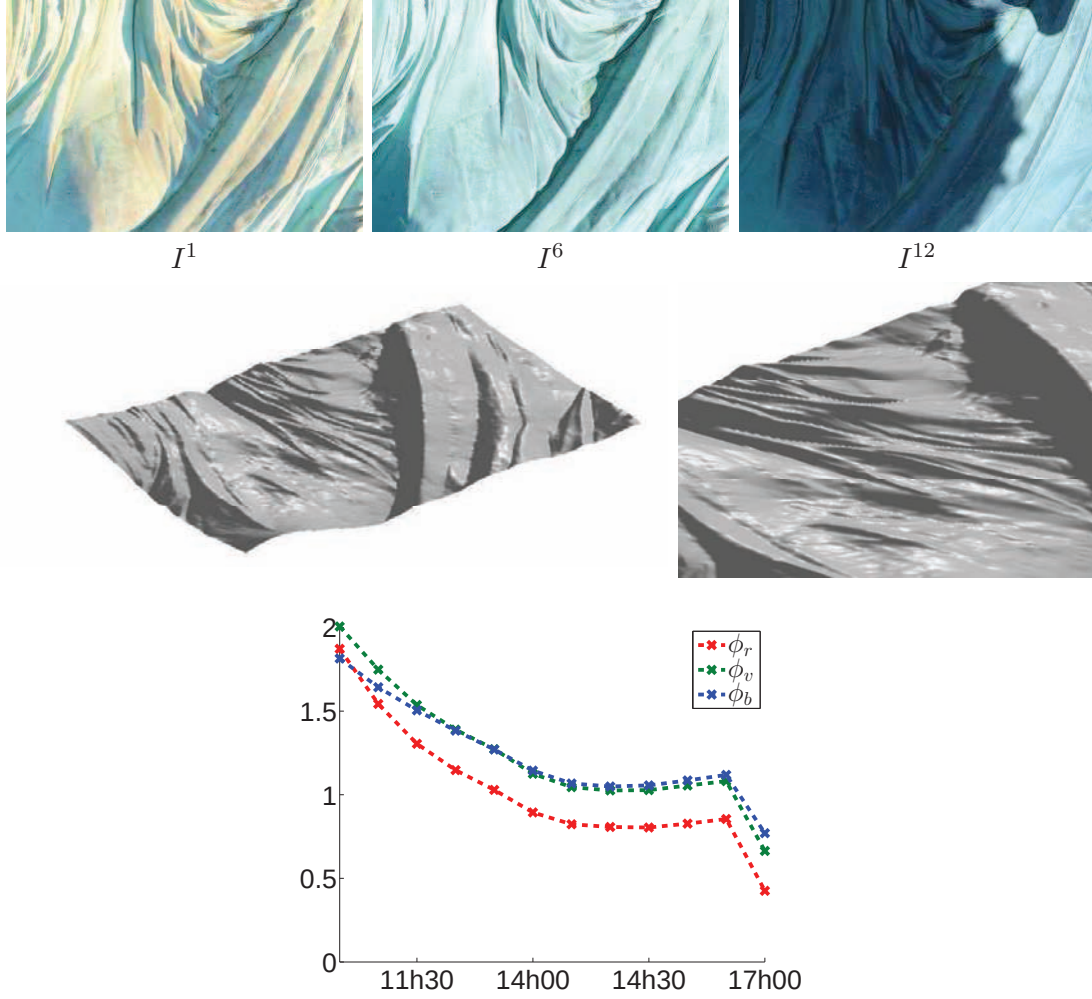


FIGURE 4.18 – Reconstruction 3D d’une partie de la Statue de la Liberté, de couleur uniforme, utilisant $m = 12$ images de taille 231×356 , prises entre 10h30 et 17h00, le 5 juillet 2014. Première ligne : I^1 , I^6 et I^{12} . La couleur dominante varie beaucoup d’une image à l’autre, à cause de l’effet combiné du changement de spectre lumineux au cours de la journée, de la balance automatique des blancs de la webcam, et de la présence de nuages. Ces effets peuvent être pris en compte en utilisant les trois canaux RGB. Deuxième ligne : reconstruction 3D obtenue, et gros plan sur les « plis » de la Statue de la Liberté (temps de calcul pour 5 itérations : moins d’une seconde sur un processeur I7 récent). Une telle reconstruction 3D pourrait servir à enrichir des modèles 3D grossiers. Par exemple, le modèle 3D de la Statue de la Liberté disponible dans Google Earth ne comporte aucun des détails très fins retrouvés ici. Troisième ligne : estimation des intensités lumineuses. Les variations d’intensité correspondent à des changements effectivement perceptibles sur les images.

Toutefois, nous avons supposé l'albédo connu, ou au moins uniforme, ce qui limite en pratique notre méthode à des données synthétiques, à des objets réels non texturés (exemple : les glaciers) ou à des scènes peintes (exemple : un visage maquillé). Une extension naturelle de ce travail consisterait à traiter le cas d'un albédo inconnu et non uniforme. Il faudra alors probablement utiliser un a priori sur la distribution de l'albédo, à la manière de ce qui est fait dans [9] pour résoudre les ambiguïtés de la stéréophotométrie non calibrée dans le cas d'éclairages non coplanaires (cf. chapitre 3). Nous avons également fait l'hypothèse d'un modèle de caméra orthographique. Une autre piste consiste à étendre ce travail à une caméra perspective, ce qui implique une relation légèrement différente entre normale et profondeur : il faudra alors adapter la définition du critère d'intégrabilité, comme cela a déjà été fait dans le cas de la stéréophotométrie non calibrée [197]. Une autre piste intéressante consiste à renforcer la robustesse au bruit grâce à une régularisation du champ de normales. Ceci est toutefois plus difficile qu'il n'y paraît, à cause de la contrainte d'unitarité des normales : il faut dans ce cas résoudre un problème variationnel sur la sphère unité, qui est une variété riemannienne. Lellmann et al. ont montré récemment [161] comment résoudre ce problème difficile par des techniques de relaxation convexe. Les développements récents des algorithmes proximaux [262] et semi-quadratiques [29] constituent également une perspective intéressante pour la résolution de tels problèmes. L'autre piste, qui a été suivie dans [119], consiste à revenir à la formulation différentielle du problème PS2 : l'intégrabilité est implicitement garantie, et il est relativement aisé de régulariser la profondeur, qui est non contrainte.

Nous rejoignons donc les conclusions du chapitre 3 : à cause de la linéarité du modèle, certains problèmes de stéréophotométrie, comme les problèmes UPS et PS2, sont mal posés. Pour réduire les ambiguïtés du problème du *shape-from-shading*, une solution consiste à utiliser une approche différentielle [87, 82], ce qui permet d'imposer implicitement l'intégrabilité : nous étudierons dans la partie IV les avantages de cette approche dans le contexte de la stéréophotométrie.

Les deux premières parties de ce mémoire sont fondées sur les hypothèses suivantes : la surface est lambertienne et les éclairages sont directionnels. Ces hypothèses peuvent sembler excessivement réductrices, mais nous avons montré dans la partie I que la simplicité du modèle pouvait être compensée par la robustesse des méthodes numériques employées. Il semble toutefois nécessaire, à ce stade, de nous interroger sur l'intérêt d'utiliser des modèles plus réalistes. Dans la prochaine partie, nous nous intéressons donc à des modèles plus généraux de réflectance et d'éclairage. Nous verrons que les modèles d'éclairage plus généraux sont assez faciles à prendre en compte, mais font explicitement intervenir l'inconnue z . Ceci constitue un argument supplémentaire en faveur de l'approche différentielle, sur laquelle nous reviendrons dans la partie IV. Au contraire, les méthodes numériques à employer pour les modèles de réflectance non lambertiens sont compliqués, ce qui tend à discréditer les approches fondées sur de tels modèles, en comparaison de l'efficacité des méthodes robustes appliquées au modèle lambertien linéaire.

Troisième partie

Utilisation de modèles réalistes pour la stéréophotométrie

Avant-propos

Les deux premières parties de ce mémoire reposent explicitement sur le modèle de réflectance lambertienne et sur des éclairages directionnels. Nous avons proposé, dans la partie I, plusieurs façons d'étendre la méthode de résolution classique de la stéréophotométrie à certains cas où ces hypothèses sont partiellement mises en défaut. Nous avons ensuite mis en évidence, dans la partie II, les limites du modèle lambertien linéaire (1.18). Nous devons maintenant nous demander s'il est pertinent d'utiliser des modèles plus réalistes, que ce soit en termes de réflectance ou d'éclairages.

Dans le chapitre 5, nous analysons en détail un modèle de source lumineuse lambertienne que nous avons présenté initialement dans [212]. Nous dérivons plusieurs versions approchées de ce modèle, qui nous permettent de retrouver le modèle d'éclairage directionnel utilisé jusqu'ici, le modèle de source ponctuelle et celui de source plane étendue [220]. Nous montrons également que la prise en compte des non linéarités dans le modèle d'éclairage est assez aisée du point de vue numérique, et que les méthodes robustes étudiées dans la partie I sont faciles à adapter, ce qui nous permet d'appliquer la stéréophotométrie dans des conditions opératoires très variées.

Nous nous intéressons ensuite, dans le chapitre 6, aux modèles de réflectance. Nous commençons par proposer une analyse d'une méthode de résolution de la stéréophotométrie due à Hertzmann et Seitz [122, 121], qui ne repose sur aucun modèle explicite de réflectance : ceci est possible en « apprenant » la réflectance et les directions d'éclairage sur un objet de référence de relief connu. Nous avons montré dans [83] comment accélérer cette méthode et comment améliorer sa robustesse aux variations de réflectance et aux ombres portées. Aussi séduisante que paraisse cette idée, elle est cependant limitée par la nécessité de disposer d'un objet de référence, et ne peut pas s'étendre à des éclairages non directionnels. Dans le cas général, il est nécessaire d'utiliser un modèle explicite de réflectance. Nous nous intéressons particulièrement à deux modèles de réflectance non linéaires : le modèle d'Oren-Nayar et celui de Ward. Il apparaît que les méthodes simples déployées pour traiter le modèle lambertien deviennent nettement plus compliquées.

Modèles d'éclairage non directionnels

Nous avons vu que l'extension de la formulation classique de la stéréophotométrie calibrée au cas d'une caméra perspective ne posait pas de difficulté particulière. Nous avons également montré comment appliquer cette technique à des surfaces régulières seulement par morceaux (cf. chapitre 2). Nous verrons dans le chapitre 6 que son extension à des modèles de réflectance non lambertienne est plus compliquée. Dans le présent chapitre, nous remettons en question l'hypothèse d'éclairages directionnels, pour nous intéresser à des modèles d'éclairage qui, tout en étant plus réalistes, se prêtent bien aux méthodes d'estimation robuste présentées dans la partie I.

En revenant à la définition d'une surface lambertienne, vue comme une source lumineuse primaire, nous montrons dans le paragraphe 5.1 comment établir les modèles d'éclairage directionnel, ponctuel ou étendu. Cette étude a été partiellement présentée dans [212], et le cas particulier des sources étendues a été traité dans [219, 220]. Avec ces modèles d'éclairage explicites, la dépendance de s vis-à-vis de la profondeur z peut devenir compliquée : pour parer à cette difficulté, nous introduisons un modèle plus général, dans lequel nous considérons s comme un champ vectoriel quelconque [216], ce qui nous permet de modéliser les trois cas particuliers précédents, ainsi que les réflexions mutuelles. Nous proposons ensuite, dans le paragraphe 5.2, des protocoles expérimentaux permettant d'étalonner les éclairages. Enfin, nous montrons dans le paragraphe 5.3 comment étendre les méthodes numériques déjà présentées dans la partie I à des éclairages non directionnels.

5.1 Quelques modèles d'éclairage

La prise en compte d'éclairages non directionnels constitue un axe de recherche important pour la reconstruction 3D photométrique. Le modèle de source ponctuelle proche (de la scène) a d'abord été étudié avec un certain succès dans les années 1990 [60, 134]. Il a même été prouvé que ce modèle permettait de rendre le problème du *shape-from-shading* mieux posé [208]. Le cas d'une source étendue a également été étudié, principalement par Clark [61, 62]. Ces deux modèles étant en apparence très différents, nous détaillons ci-après le raisonnement permettant d'obtenir la formulation explicite de l'éclairage en fonction des paramètres caractéristiques de la source. Cela nécessite de revenir à la définition d'une source lambertienne, et de manipuler les concepts de luminance et de réflectance [78, 189].

5.1.1 De la luminance au vecteur d'éclairage

Rappelons que le niveau de gris d'un appareil photographique peut, après divers étalonnages, être supposé proportionnel à la luminance de la scène (cf. paragraphe 1.1.1). Les techniques photométriques assimilent la scène photographiée à une source lumineuse émettant dans la direction d'observation $\mathbf{v}$. La luminance émise par un point objet $\mathbf{x}$, de normale $\mathbf{n}(\mathbf{x})$, dans la direction d'émission $\mathbf{v}$ caractérisée par les angles (θ_e, ϕ_e) , relativement à la normale $\mathbf{n}(\mathbf{x})$, s'écrit alors [78] :

$$L(\mathbf{x}, \theta_e, \phi_e) = \int_{\theta_i=0}^{\pi/2} \int_{\phi_i=0}^{2\pi} \kappa(\mathbf{x}, \theta_i, \phi_i, \theta_e, \phi_e) \bar{L}(\theta_i, \phi_i) \sin \theta_i \cos \theta_i d\theta_i d\phi_i \quad (5.1)$$

où κ désigne la *fonction de distribution de la réflectance bidirectionnelle* (*bidirectional reflectance distribution function*, ou BRDF) du matériau, que l'on appelle également sa *réflectance*, et $\bar{L}(\theta_i, \phi_i)$ désigne la *luminance incidente* dans la direction caractérisée par les angles (θ_i, ϕ_i) , relativement à la normale $\mathbf{n}(\mathbf{x})$.

Dans le cas général, l'expression (5.1) est difficile à manipuler. Si la surface est supposée lambertienne, alors la réflectance ne dépend ni de la direction incidente, ni de la direction d'émission. On montre qu'elle s'écrit alors très simplement en fonction de l'albédo ρ :

$$\kappa(\mathbf{x}, \theta_i, \phi_i, \theta_e, \phi_e) = \frac{\rho(\mathbf{x})}{\pi} \quad (5.2)$$

De (5.1) et (5.2), on tire :

$$L(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\pi} \int_{\theta_i=0}^{\pi/2} \int_{\phi_i=0}^{2\pi} \bar{L}(\theta_i, \phi_i) \sin \theta_i \cos \theta_i d\theta_i d\phi_i \quad (5.3)$$

En désignant par $\mathbf{u}_i$ le vecteur unitaire indiquant le sens et la direction de la lumière incidente, caractérisée par les angles (θ_i, ϕ_i) , relativement à la normale $\mathbf{n}(\mathbf{x})$, et en observant que $\cos \theta_i = \mathbf{u}_i \cdot \mathbf{n}(\mathbf{x})$, on peut réécrire (5.3) :

$$L(\mathbf{x}) = \frac{\rho}{\pi} \mathbf{n}(\mathbf{x}) \cdot \underbrace{\int_{\theta_i=0}^{\pi/2} \int_{\phi_i=0}^{2\pi} \bar{L}(\theta_i, \phi_i) \sin \theta_i \mathbf{u}_i d\theta_i d\phi_i}_{\mathbf{s}(\mathbf{x})} \quad (5.4)$$

Le vecteur $\mathbf{s}(\mathbf{x})$ caractérise le flux lumineux total incident en $\mathbf{x}$, qui inclut aussi bien les flux en provenance directe des différentes sources primaires visibles depuis $\mathbf{x}$, que les flux éclairant $\mathbf{x}$ après un nombre quelconque de réflexions sur la surface elle-même (ou sur d'autres surfaces alentour). Dans le cas d'une source unique, ponctuelle et infiniment distante, et si l'on néglige les phénomènes de réflexions mutuelles, $\mathbf{s}(\mathbf{x})$ est un champ parallèle et uniforme, i.e. $\mathbf{s}(\mathbf{x}) = \mathbf{s}$, ce qui correspond au modèle directionnel utilisé jusqu'à présent. Nous nous consacrons, dans ce chapitre, à l'étude de situations plus générales, et montrons notamment comment obtenir une expression analytique du flux primaire dans différentes configurations. Pour cela, nous revenons à la définition d'une source primaire lambertienne.

5.1.2 Source primaire lambertienne

Soit une source lumineuse élémentaire située en $\mathbf{x}_s \in \mathbb{R}^3$, constituée d'un élément de surface $d\Sigma_s$ de normale unitaire sortante $\mathbf{n}(\mathbf{x}_s)$. Notons $d^2\Phi$ le flux lumineux élémentaire émis par cette source à l'intérieur d'un cône de sommet $\mathbf{x}_s$, d'angle solide $d\omega$ et de direction $\mathbf{u}_e$. La luminance de la source $\mathbf{x}_s$ est définie comme le flux lumineux par unité d'angle solide et par unité de *surface apparente* dans la direction $\mathbf{u}_e$. Elle s'écrit donc :

$$\bar{L}_{\mathbf{x}_s}(\mathbf{u}_e) = \frac{d^2\Phi}{d\omega d\Sigma'_s} \quad (5.5)$$

où :

$$d\Sigma'_s = d\Sigma_s (\mathbf{n}(\mathbf{x}_s) \cdot \mathbf{u}_e) \quad (5.6)$$

est la surface apparente de $d\Sigma_s$ dans la direction $\mathbf{u}_e$.

Nous nous intéressons dans ce chapitre aux sources lumineuses primaires *lambertiennes*, i.e. pour lesquelles $\bar{L}_{\mathbf{x}_s}(\mathbf{u}_e)$ est indépendant de $\mathbf{u}_e$:

$$\bar{L}_{\mathbf{x}_s}(\mathbf{u}_e) = \bar{L}_{\mathbf{x}_s} \quad (5.7)$$

Autant l'hypothèse lambertienne est discutable pour modéliser la réflectance des surfaces à reconstruire (cf. chapitre 6), autant il s'agit d'un modèle de source lumineuse très réaliste, puisque les fabricants essaient généralement de garantir l'isotropie de l'éclairage, pour le confort de l'utilisateur.

Supposons maintenant que la source élémentaire située en $\mathbf{x}_s$ éclaire un élément de surface $d\Sigma$ situé autour d'un point $\mathbf{x}$ de la surface à reconstruire, de normale unitaire sortante $\mathbf{n}(\mathbf{x})$. Comme cela est illustré sur la figure 5.1, l'angle solide du cône $\mathcal{C}$ de sommet $\mathbf{x}_s$, s'appuyant sur $d\Sigma$, s'écrit :

$$d\omega = \frac{d\Sigma (-\mathbf{u}_e \cdot \mathbf{n}(\mathbf{x}))}{\|\mathbf{x} - \mathbf{x}_s\|^2} \quad (5.8)$$

L'éclairement (*irradiance*) élémentaire en $\mathbf{x}$ dû à cette source élémentaire est défini comme le flux lumineux incident reçu par unité de surface :

$$dE(\mathbf{x}) = \frac{d^2\Phi}{d\Sigma} \quad (5.9)$$

La direction d'éclairage $\mathbf{u}_e$ s'écrit, en outre :

$$\mathbf{u}_e = \frac{\mathbf{x} - \mathbf{x}_s}{\|\mathbf{x} - \mathbf{x}_s\|} \quad (5.10)$$

Des égalités (5.5) à (5.10), nous tirons finalement :

$$dE(\mathbf{x}) = d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \quad (5.11)$$

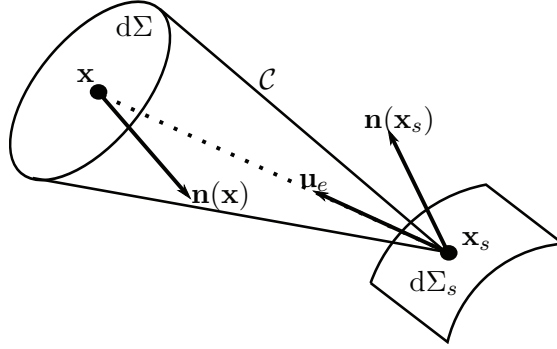


FIGURE 5.1 – La source lambertienne élémentaire située en $\mathbf{x}_s$, de surface $d\Sigma_s$ et de normale sortante $\mathbf{n}(\mathbf{x}_s)$, éclaire un élément de surface situé en $\mathbf{x}$, de surface $d\Sigma$ et de normale sortante $\mathbf{n}(\mathbf{x})$.

où le vecteur d'éclairage élémentaire $d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x})$ est défini comme suit :

$$d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) = \underbrace{\frac{\bar{L}_{\mathbf{x}_s} d\Sigma_s}{\|\mathbf{x}_s - \mathbf{x}\|^2}}_{\text{Atténuation en } 1/\text{distance}^2} \underbrace{\frac{\mathbf{n}(\mathbf{x}_s) \cdot (\mathbf{x} - \mathbf{x}_s)}{\|\mathbf{x}_s - \mathbf{x}\|}}_{\text{Anisotropie de la source}} \underbrace{\frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|}}_{\text{Direction d'éclairage}} \quad (5.12)$$

En intégrant (5.11) sur la partie $\Omega_s(\mathbf{x})$ de la source lumineuse visible depuis $\mathbf{x}$, et en utilisant l'expression (5.12) de $d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x})$, l'éclairement total au point objet $\mathbf{x}$ s'écrit :

$$E(\mathbf{x}) = \underbrace{\iint_{\mathbf{x}_s \in \Omega_s(\mathbf{x})} \frac{\bar{L}_{\mathbf{x}_s}}{\|\mathbf{x}_s - \mathbf{x}\|^2} \frac{\mathbf{n}(\mathbf{x}_s) \cdot (\mathbf{x} - \mathbf{x}_s)}{\|\mathbf{x}_s - \mathbf{x}\|} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|} d\Sigma_s \cdot \mathbf{n}(\mathbf{x})}_{\mathbf{s}(\mathbf{x})} \quad (5.13)$$

qui est une généralisation de (1.13).

Reprenant les travaux présentés dans [212], nous présentons dans ce qui suit les différentes approximations de $\mathbf{s}(\mathbf{x})$ conduisant aux modèles d'éclairage les plus fréquemment rencontrés en stéréophotométrie.

5.1.3 Source sphérique

En dehors du cas directionnel, le modèle d'éclairage le plus fréquemment rencontré est celui de *source sphérique*. Ce modèle a été introduit dans [134] pour le *shape-from-shading* (il a d'ailleurs servi à prouver le caractère mieux posé du problème dans [208]), et dans [60] pour la stéréophotométrie. Plus récemment, il a également été utilisé dans [177, 198, 276].

L'éclairage total reçu au point $\mathbf{x}$, et provenant d'une source sphérique de rayon R , s'obtient en sommant les vecteurs d'éclairage élémentaires $d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x})$ correspondant à la partie visible de la sphère depuis $\mathbf{x}$, qui est une hémisphère $\mathcal{H}$:

$$\mathbf{s}(\mathbf{x}) = \iint_{\mathbf{x}_s \in \mathcal{H}} d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) \quad (5.14)$$

Supposons que la luminance de la source soit uniforme, i.e. $\bar{L}_{\mathbf{x}_s} \equiv \bar{L}$. Supposons également que le rayon R de la sphère soit négligeable devant la distance source-surface, i.e. $R \ll \|\mathbf{x}_s - \mathbf{x}\|$, de telle sorte que le vecteur $\mathbf{x}_s - \mathbf{x}$ soit le même pour tous les points $\mathbf{x}_s$ de l'hémisphère $\mathcal{H}$. Sous cette dernière hypothèse, nous pouvons assimiler $\mathbf{x}_s$ au centre de la sphère. En utilisant l'expression (5.12) de $\mathrm{d}\mathbf{s}_{\mathbf{x}_s}(\mathbf{x})$ dans (5.14), nous obtenons :

$$\mathbf{s}(\mathbf{x}) = \frac{\bar{L}}{\|\mathbf{x}_s - \mathbf{x}\|^2} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|} \iint_{\mathbf{x}_s \in \mathcal{H}} \frac{\mathbf{n}(\mathbf{x}_s) \cdot (\mathbf{x} - \mathbf{x}_s)}{\|\mathbf{x}_s - \mathbf{x}\|} \mathrm{d}\Sigma_s \quad (5.15)$$

Or, l'intégrale du membre droit n'est autre que la surface apparente $\sigma_{\mathcal{H}}$ de $\mathcal{H}$, qui en l'occurrence est égale à πR^2 . Nous pouvons donc écrire le vecteur d'éclairage d'une source sphérique, en un point $\mathbf{x}$, sous la forme suivante :

$$\mathbf{s}(\mathbf{x}) = \frac{\bar{L} \sigma_{\mathcal{H}}}{\|\mathbf{x}_s - \mathbf{x}\|^2} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|} \quad (5.16)$$

Dans cette expression, on reconnaît le facteur d'atténuation en $1/\text{distance}^2$ caractéristique d'une source « ponctuelle ».

Le modèle d'éclairage le plus simple est obtenu en supposant que, dans le modèle de source sphérique, la source est située en un point $\mathbf{x}_s$ infiniment éloigné. Dans ce cas, on peut considérer que le vecteur $\mathbf{x}_s - \mathbf{x}$ est indépendant de $\mathbf{x}$. Le modèle de source sphérique (5.16) se simplifie alors :

$$\mathbf{s}(\mathbf{x}) = \phi \bar{\mathbf{s}} \quad (5.17)$$

et correspond bien au modèle utilisé jusqu'au chapitre 4, et à l'immense majorité des modèles d'éclairage utilisés pour les techniques de reconstruction 3D photométrique (SFS et stéréophotométrie).

5.1.4 Source plane élémentaire

Lorsque la source est un rectangle élémentaire d'aire $\mathrm{d}\Sigma_s$, situé en $\mathbf{x}_s$, de normale unitaire sortante $\mathbf{n}_0$ et de luminance $\bar{L}_{\mathbf{x}_s}$, le vecteur d'éclairage élémentaire $\mathrm{d}\mathbf{s}_{\mathbf{x}_s}(\mathbf{x})$ en un point $\mathbf{x}$ est défini par (5.12), que nous réécrivons sous une forme très similaire à (5.16) :

$$\mathrm{d}\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) = \frac{\bar{L} \mathrm{d}\Sigma_s a_{\text{Lambert}}(\mathbf{x}, \mathbf{x}_s)}{\|\mathbf{x}_s - \mathbf{x}\|^2} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|} \quad (5.18)$$

où :

$$a_{\text{Lambert}}(\mathbf{x}, \mathbf{x}_s) = \frac{\mathbf{n}_0 \cdot (\mathbf{x} - \mathbf{x}_s)}{\|\mathbf{x}_s - \mathbf{x}\|} \quad (5.19)$$

est un facteur d'anisotropie caractéristique d'une source lambertienne. Si l'on remarque que le produit $\mathrm{d}\Sigma_s a_{\text{Lambert}}(\mathbf{x}, \mathbf{x}_s)$ est égal à la surface apparente $\sigma_{\mathrm{d}\Sigma_s}$ du rectangle dans la direction du point $\mathbf{x}$, on obtient une expression parfaitement similaire à (5.16) :

$$\mathrm{d}\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) = \frac{\bar{L} \sigma_{\mathrm{d}\Sigma_s}}{\|\mathbf{x}_s - \mathbf{x}\|^2} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|} \quad (5.20)$$

Le facteur d'anisotropie a_{Lambert} de la source lumineuse a été récemment remis en question par Mecca et al., qui ont utilisé dans [175] des modèles d'anisotropie de la forme :

$$a(\mathbf{x}, \mathbf{x}_s) = \left(\frac{\mathbf{n}_0 \cdot (\mathbf{x} - \mathbf{x}_s)}{\|\mathbf{x}_s - \mathbf{x}\|} \right)^\mu \quad (5.21)$$

où $\mu \geq 0$. Ce modèle inclut le cas des sources isotropes ($\mu = 0$), des sources lambertiennes ($\mu = 1$) et des sources « lambertiennes imparfaites » ($\mu > 1$). Moreno et al. montrent dans [181] que le modèle de source lambertienne imparfaite peut être utilisé pour décrire l'émission des LED, en choisissant :

$$\mu = \frac{-\ln(2)}{\ln(\cos \theta_{1/2})} \quad (5.22)$$

où $\theta_{1/2}$ désigne l'angle pour lequel la luminance est égale à la moitié de sa valeur maximale (cet angle est généralement fourni par les fabricants).

5.1.5 Source plane étendue

Une autre configuration d'éclairage ayant connu un certain succès est celle de source *étendue*, constituée d'une infinité de sources élémentaires formant une surface éclairante [62]. Nous nous intéressons dans ce paragraphe au cas d'une surface plane. En l'occurrence, nous utilisons l'écran d'un dispositif multimédia comme source lumineuse, dans le but reconstruire en 3D une scène placée face à l'écran et filmée par une caméra de type *webcam*¹. Nous avons déjà présenté ces travaux dans [219, 220].

Chaque source élémentaire (i.e., chaque pixel de l'écran) est assimilée à un rectangle d'aire $d\Sigma_s = dx_s dy_s$, de même normale $\mathbf{n}_0$. Le passage du repère lié à la source à un repère lié à la webcam peut se faire au moyen d'une procédure d'étalonnage spécifique que nous décrivons dans [220], mais qui sort du cadre de cette thèse consacrée à la photométrie. Notons simplement qu'en positionnant un miroir plan face à l'écran, de façon à ce que la webcam voie l'écran, on peut mettre en correspondance des points particuliers de l'écran avec leurs images par réflexion sur le miroir. En utilisant au moins deux poses du miroir et trois correspondances, la détermination de la pose de l'écran, relativement au repère de la webcam, est un problème bien posé. Nous pouvons donc supposer, sans perte de généralité, que le repère de référence est celui de l'écran. Dans ce repère, $\mathbf{n}_0 = [0, 0, 1]^\top$ et $\mathbf{x}_s = [x_s, y_s, 0]^\top$.

Nous désignons par $\mathcal{E}$ la surface éclairante, qui est constituée d'un ensemble de sources élémentaires caractérisées par (5.12). L'éclairage total en un point $\mathbf{x}$ est donc obtenu par sommation :

$$\mathbf{s}(\mathbf{x}) = \iint_{\mathbf{x}_s \in \mathcal{T}(\mathbf{x})} d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) \quad (5.23)$$

Dans cette expression, $\mathcal{T}(\mathbf{x}) \subset \mathcal{E}$ désigne l'ensemble des points $\mathbf{x}_s \in \mathcal{E}$ qui sont *visibles* depuis $\mathbf{x}$. Le fait que le domaine d'intégration $\mathcal{T}(\mathbf{x})$ dépende de $\mathbf{x}$ est une cause de *pénombre*. Ce phénomène est illustré sur la figure 5.2.

1. Ce travail a été initié par Richard Modrzejewski au cours de son stage de première année de master à l'IRIT.

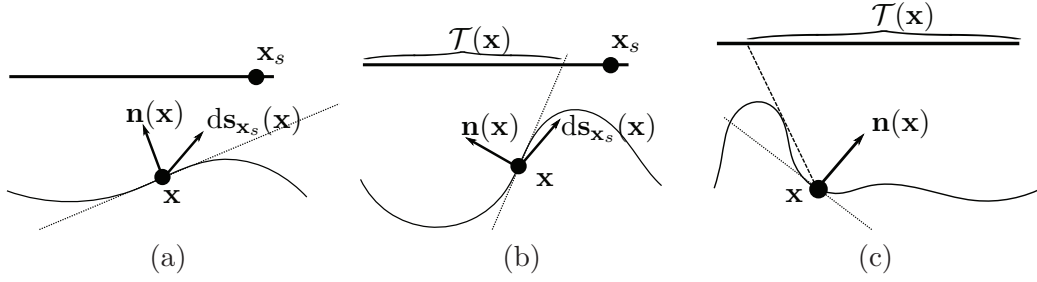


FIGURE 5.2 – Occultation partielle d’une source étendue. (a) Lorsque toute la source étendue $\mathcal{E}$ est visible depuis $\mathbf{x}$, toutes les sources élémentaires $\mathbf{x}_s \in \mathcal{E}$ contribuent à l’éclairage total en $\mathbf{x}$. Cela n’est pas le cas en (b), où seuls les points situés à l’intérieur du domaine $\mathcal{T}(\mathbf{x})$ sont visibles depuis $\mathbf{x}$, ce qui est cause de pénombre. Il est cependant possible de définir explicitement $\mathcal{T}(\mathbf{x})$, par intersection du plan tangent à la surface au point $\mathbf{x}$ et de la surface éclairante. En revanche, dans le cas (c), l’occultation partielle de la surface éclairante est due à un objet extérieur (ici, la surface elle-même). Ce cas est particulièrement difficile à modéliser, car il nécessite de connaître la géométrie *globale* de la surface.

En utilisant comme repère de référence le repère de la source plane, on a $z_s = 0$ pour tout $\mathbf{x}_s = [x_s, y_s, z_s]^\top \in \mathcal{E}$, de telle sorte que $\mathcal{T}(\mathbf{x})$ est caractérisé par un domaine 2D, noté $\Omega_s(\mathbf{x})$. En remplaçant $ds_{\mathbf{x}_s}(\mathbf{x})$ par son expression (5.12) dans (5.23), on obtient l’expression suivante de l’éclairage total au point $\mathbf{x} = [x, y, z]^\top$:

$$\mathbf{s}(\mathbf{x}) = z \iint_{(x_s, y_s) \in \Omega_s(\mathbf{x})} \bar{L}_{\mathbf{x}_s} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|^4} dx_s dy_s \quad (5.24)$$

Remarques sur la visibilité partielle

L’expression (5.24) peut être utilisée pour la modélisation de n’importe quelle source plane lambertienne, pourvu que la luminance $L_{\mathbf{x}_s}$ soit connue, ce qui est précisément le cas des écrans puisque cette quantité est alors proportionnelle (en théorie du moins) au niveau de gris ou au niveau de couleur affiché en chaque pixel. Les coordonnées $\mathbf{x}_s$ étant également connues, une approximation discrète de l’intégrale dans l’expression (5.24) peut être calculée numériquement. Ceci est suffisant pour le rendu en synthèse d’images : la géométrie $\mathbf{x} = [x, y, z]^\top$ de la scène étant parfaitement connue, le domaine $\mathcal{T}(\mathbf{x})$, et donc $\Omega_s(\mathbf{x})$, peuvent être déterminés par lancer de rayons.

En revanche, dans les applications dédiées à la reconstruction 3D, la géométrie de la scène est justement l’inconnue. La visibilité pourrait donc être estimée automatiquement au cours d’un processus itératif, en utilisant à chaque itération l’estimation précédente de $\mathbf{x}$ et de $\mathbf{n}(\mathbf{x})$ pour estimer la visibilité courante. Cependant, une telle démarche n’est pas réaliste car ce processus devrait être répété pour chaque point $\mathbf{x}$ de la scène, ce qui conduirait à un algorithme extrêmement lent, même avec les ordinateurs récents. Il peut en effet y avoir autant de points $\mathbf{x}$ que de pixels sur le récepteur photosensible de

l'appareil photographique, et la surface éclairante visible $\Omega_s(\mathbf{x})$ peut contenir un nombre très élevé de points (égal à la résolution de l'écran dans cette application).

Aussi, par souci de simplicité, nous cherchons dorénavant une expression analytique de l'intégrale dans (5.24), de façon à obtenir une estimation de l'éclairage en un temps raisonnable. Pour parvenir à cet objectif, nous négligeons dorénavant les effets de pénombre, et laissons leur étude en perspective. Par conséquent, notre modélisation ne sera valide que pour des surfaces présentant de faibles pentes, en l'absence d'occultation.

Domaines plans de forme arbitraire et de luminance uniforme

Pour simplifier davantage encore l'intégrale dans l'expression (5.24), il nous faut expliciter la dépendance de la luminance $L_{\mathbf{x}_s}$ de la surface éclairante. Commençons par le cas le plus simple d'une luminance uniforme $L_{\mathbf{x}_s} = L_0$. L'expression (5.24) se réécrit :

$$\mathbf{s}(\mathbf{x}) = z L_0 \iint_{(x_s, y_s) \in \Omega_s} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|^4} dx_s dy_s \quad (5.25)$$

ou encore :

$$\mathbf{s}(\mathbf{x}) = -\frac{L_0}{2} [F_1(\mathbf{x}), F_2(\mathbf{x}), F_3(\mathbf{x})]^\top \quad (5.26)$$

avec :

$$\begin{cases} F_1(\mathbf{x}) = -2z \iint_{(x_s, y_s) \in \Omega_s} \frac{x_s - x}{[(x_s - x)^2 + (y_s - y)^2 + z^2]^2} dx_s dy_s \\ F_2(\mathbf{x}) = -2z \iint_{(x_s, y_s) \in \Omega_s} \frac{y_s - y}{[(x_s - x)^2 + (y_s - y)^2 + z^2]^2} dx_s dy_s \\ F_3(\mathbf{x}) = 2z^2 \iint_{(x_s, y_s) \in \Omega_s} \frac{1}{[(x_s - x)^2 + (y_s - y)^2 + z^2]^2} dx_s dy_s \end{cases} \quad (5.27)$$

En notant $r = x_s - x$ et $s = y_s - y$, ces fonctions se réécrivent :

$$\begin{cases} F_1(\mathbf{x}) = -2z \iint_{(r, s) \in \Omega_s - (x, y)} \frac{r}{(r^2 + s^2 + z^2)^2} dr ds \\ F_2(\mathbf{x}) = -2z \iint_{(r, s) \in \Omega_s - (x, y)} \frac{s}{(r^2 + s^2 + z^2)^2} dr ds \\ F_3(\mathbf{x}) = 2z^2 \iint_{(r, s) \in \Omega_s - (x, y)} \frac{1}{(r^2 + s^2 + z^2)^2} dr ds \end{cases} \quad (5.28)$$

Soit $\mathcal{C} \subset \mathbb{R}^2$ un domaine plan, de contour $\partial\mathcal{C}$ orienté dans le sens direct, et C^1 par morceaux. Pour toute paire (P, Q) de fonctions continues $\mathcal{C} \rightarrow \mathbb{R}$, la formule de Green-Riemann s'écrit :

$$\iint_{(r, s) \in \mathcal{C}} \left(\frac{\partial Q}{\partial r} - \frac{\partial P}{\partial s} \right) dr ds = \oint_{(r, s) \in \partial\mathcal{C}} (P dr + Q ds) \quad (5.29)$$

Il est facile de vérifier les identités suivantes :

$$\begin{cases} \frac{\partial}{\partial r} \left(\frac{1}{r^2 + s^2 + z^2} \right) = -\frac{2r}{(r^2 + s^2 + z^2)^2} \\ -\frac{\partial}{\partial s} \left(\frac{1}{r^2 + s^2 + z^2} \right) = \frac{2s}{(r^2 + s^2 + z^2)^2} \\ \frac{\partial}{\partial r} \left(\frac{r}{r^2 + s^2 + z^2} \right) - \frac{\partial}{\partial s} \left(\frac{-s}{r^2 + s^2 + z^2} \right) = \frac{2z^2}{(r^2 + s^2 + z^2)^2} \end{cases} \quad (5.30)$$

Des équations (5.28), (5.29) et (5.30), nous déduisons aisément de nouvelles expressions pour les intégrales doubles de (5.28), qui sont transformées en intégrales curvilignes :

$$\begin{cases} F_1(\mathbf{x}) = z \oint_{(r,s) \in \partial\Omega_s - (x,y)} \frac{ds}{r^2 + s^2 + z^2} \\ F_2(\mathbf{x}) = -z \oint_{(r,s) \in \partial\Omega_s - (x,y)} \frac{dr}{r^2 + s^2 + z^2} \\ F_3(\mathbf{x}) = \oint_{(r,s) \in \partial\Omega_s - (x,y)} \frac{r ds - s dr}{r^2 + s^2 + z^2} \end{cases} \quad (5.31)$$

Pour peu que le contour $\partial\Omega_s$ ait une forme « simple », il est effectivement envisageable que ces trois intégrales curvilignes s'expriment analytiquement. Comme contours « simples », nous pouvons citer les motifs rectangulaires étudiés dans [31, 94, 190, 233, 265], les motifs circulaires [31], et même les sources partiellement occultées dans un cas comme celui de la figure 5.2-b, puisque l'ensemble Ω_s peut alors être déterminé explicitement (rappelons que ce cas est laissé en perspective).

L'originalité des expressions (5.31) réside dans leur généralité, puisqu'elles fournissent un cadre pour établir explicitement l'éclairage induit par un domaine de forme quelconque. En guise d'exemple, étudions le cas où Ω_s est rectangulaire. En effectuant une partition rectangulaire non uniforme d'une image, cet exemple simple nous permettra de modéliser de façon approchée, mais très réaliste, l'éclairage émis par une image affichée sur un écran : nous avons montré dans [220] que cela permettrait d'appliquer la stéréophotométrie aux images d'un utilisateur placé face à un diaporama, prises par la webcam interne d'un système multimédia.

Expression analytique pour un domaine rectangulaire

Si Ω_s est un rectangle $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$ (on rappelle que la luminance de la surface éclairante est supposée uniforme dans ce rectangle), alors les intégrales curvilignes de (5.31) se simplifient et le champ d'éclairage total au point $\mathbf{x}$ est donné par (5.26),

avec :

$$\left\{ \begin{array}{l} F_1(\mathbf{x}) = \left[\left[\frac{z}{\sqrt{r^2 + z^2}} \tan^{-1} \left(\frac{s}{\sqrt{r^2 + z^2}} \right) \right]_{s=y_{\min}-y}^{y_{\max}-y} \right]_{r=x_{\min}-x}^{x_{\max}-x} \\ F_2(\mathbf{x}) = \left[\left[\frac{z}{\sqrt{s^2 + z^2}} \tan^{-1} \left(\frac{r}{\sqrt{s^2 + z^2}} \right) \right]_{s=y_{\min}-y}^{y_{\max}-y} \right]_{r=x_{\min}-x}^{x_{\max}-x} \\ F_3(\mathbf{x}) = \left[\left[\frac{r \tan^{-1} \left(\frac{s}{\sqrt{r^2 + z^2}} \right)}{\sqrt{r^2 + z^2}} + \frac{s \tan^{-1} \left(\frac{r}{\sqrt{s^2 + z^2}} \right)}{\sqrt{s^2 + z^2}} \right]_{s=y_{\min}-y}^{y_{\max}-y} \right]_{r=x_{\min}-x}^{x_{\max}-x} \end{array} \right. \quad (5.32)$$

Ces expressions analytiques pourraient notamment servir à améliorer la qualité de la reconstruction 3D par stéréophotométrie en utilisant des rectangles uniformes affichés sur un écran [31, 94, 190, 233, 265]. En comparaison du modèle de Clark [61], notre modèle prend en compte l'anisotropie de chaque pixel, alors que Clark considère les pixels de l'écran comme des sources isotropes, ce qui limite l'utilisation de sa technique à des petits objets : « Nous supposons que les pixels d'un écran LCD constituent des sources isotropes, ce qui n'est pas le cas. L'hypothèse d'isotropie est plus acceptable dans notre montage expérimental, où l'objet est petit »². Nous verrons que la prise en compte de l'anisotropie améliore nettement la précision du modèle de source étendue.

Partition rectangulaire

Nous étudions un deuxième cas simple : celui d'une luminance uniforme par morceaux (figure 5.3), qui correspond par exemple à l'approximation d'une image par une partition rectangulaire non uniforme, comme cela a été suggéré par Clark dans [61]. Nous supposons que la source est constituée d'une partition $\cup_{i=1}^n \Omega_s^i$, où $\Omega_s^i = [x_{\min}^i, x_{\max}^i] \times [y_{\min}^i, y_{\max}^i]$, et que la luminance est uniforme à l'intérieur de chaque rectangle Ω_s^i , de valeur L_0^i .

Le champ d'éclairage total reçu en $\mathbf{x}$ s'écrit alors, d'après (5.26) :

$$\mathbf{s}(\mathbf{x}) = - \sum_{i=1}^n \frac{L_0^i}{2} \left[F_1^i(\mathbf{x}), F_2^i(\mathbf{x}), F_3^i(\mathbf{x}) \right]^\top \quad (5.33)$$

en utilisant les notations de (5.32). L'expression (5.33) fournit une approximation réaliste de l'éclairage provenant d'une image affichée sur un écran, en effectuant par exemple la décomposition « *quadtree* » de l'image [90].

2. « We assume that the LCD pixels are isotropic illuminants, which is not the case [...]. The assumption of isotropy is made more palatable [...] in our experimental setup, where the object is small [...]. »

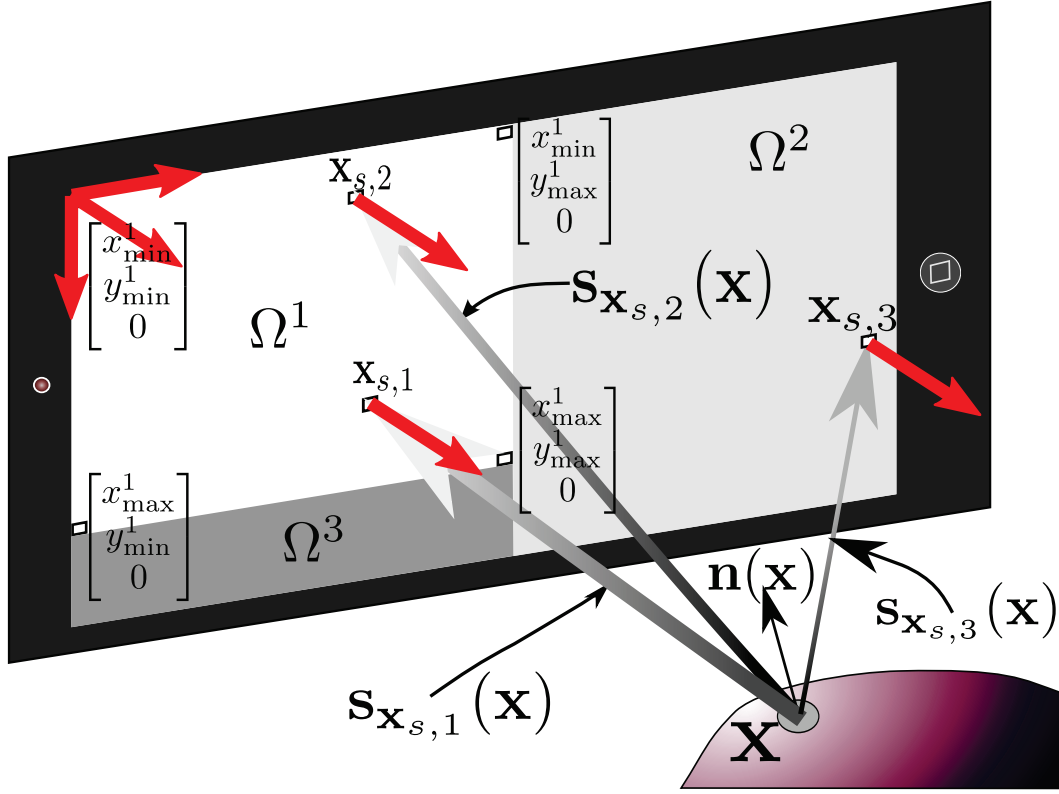


FIGURE 5.3 – Flux lumineux émis par une partition rectangulaire non uniforme, comportant trois motifs rectangulaires Ω^1 , Ω^2 et Ω^3 . L'éclairage total est la somme des contributions de toutes les sources infinitésimales $\mathbf{x}_s$. La largeur des flèches grises représente « l'intensité » de la source lumineuse infinitésimale associée, qui est une fonction décroissante de la distance source-surface, de l'angle entre la direction d'éclairage et la normale $[0, 0, 1]^\top$ (anisotropie), et de la luminance (représentée par le niveau de gris affiché à l'écran).

Validation expérimentale

Nous évaluons expérimentalement la validité du modèle de source rectangulaire uniforme défini par (5.26) et (5.32), pour des motifs rectangulaires Ω_s de surface variable (luminance uniforme), et celle du modèle de source rectangulaire non uniforme (5.33) pour plusieurs images naturelles (luminance non uniforme). Nous utilisons pour cela une feuille de papier blanc collée sur un support plan disposé face à l'écran (figure 5.4).

Méthodologie Nous supposons que la feuille de papier constitue une surface lambertienne, et que sa pose est connue (en pratique, elle est collée sur un damier) : nous notons $\mathbf{n}$ sa normale. Nous avons déjà signalé que, pour une surface lambertienne, la luminance est proportionnelle à son éclairage, le facteur de proportionnalité étant égal à ρ/π . La



FIGURE 5.4 – Dispositif expérimental pour la validation du modèle d'éclairage par un écran. En haut : écran (ordinateur portable HP EliteBook8570w) affichant un motif rectangulaire (à gauche) ou une image en niveaux de gris (à droite), placée en face d'une feuille de papier blanc. En bas : images réelles utilisées dans les expériences, de taille 1600×900 (résolution de l'écran).

luminance émise par la feuille de papier s'écrit donc, en un point $\mathbf{x}$ de sa surface :

$$l(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\pi} \mathbf{s}(\mathbf{x}) \cdot \mathbf{n} \quad (5.34)$$

où $\rho(\mathbf{x})$ désigne l'albédo, qui est un scalaire si nous supposons pour simplifier que la webcam mesure des niveaux de gris, et que l'écran affiche également une image en niveaux de gris³. L'albédo du papier étant uniforme ($\rho(\mathbf{x}) = \rho$), il s'ensuit que :

$$\sqrt{\sum_{\mathbf{x}} l(\mathbf{x})^2} = \frac{\rho}{\pi} \sqrt{\sum_{\mathbf{x}} (\mathbf{s}(\mathbf{x}) \cdot \mathbf{n})^2} \quad (5.35)$$

où les deux sommes sont calculées pour tous les points $\mathbf{x}$ de la feuille de papier. En divisant membre à membre (5.34) et (5.35), il vient :

$$\frac{l(\mathbf{x})}{\sqrt{\sum_{\mathbf{x}} l(\mathbf{x})^2}} = \frac{\mathbf{s}(\mathbf{x}) \cdot \mathbf{n}}{\sqrt{\sum_{\mathbf{x}} (\mathbf{s}(\mathbf{x}) \cdot \mathbf{n})^2}} \quad (5.36)$$

3. En utilisant les méthodes présentées dans le paragraphe 1.5, on peut utiliser la stéréophotométrie avec une source étendue colorée éclairant une surface d'albédo monochromatique [233, 190], ou une source étendue monochromatique éclairant une surface colorée [220].

Cette normalisation permet d'éliminer l'albédo inconnu ρ . La luminance $l(\mathbf{x})$ est proportionnelle au niveau de gris mesuré par l'appareil photographique. Le membre de gauche de l'équation (5.36) peut donc être directement mesuré et constitue la donnée. La précision d'un modèle d'éclairage peut alors être mesurée en comparant cette valeur à la valeur théorique du membre de droite.

Motifs rectangulaires uniformes Nous utilisons d'abord des motifs rectangulaires de taille variable (cf. figure 5.5). Comme nous pouvions nous y attendre, le modèle de source ponctuelle anisotrope [94] est adapté à des petits motifs, tandis que le modèle de source étendue isotrope [61] est plus adapté à des motifs de grande taille. Quant au modèle que nous proposons, il est aussi bien adapté aux deux types de configurations.

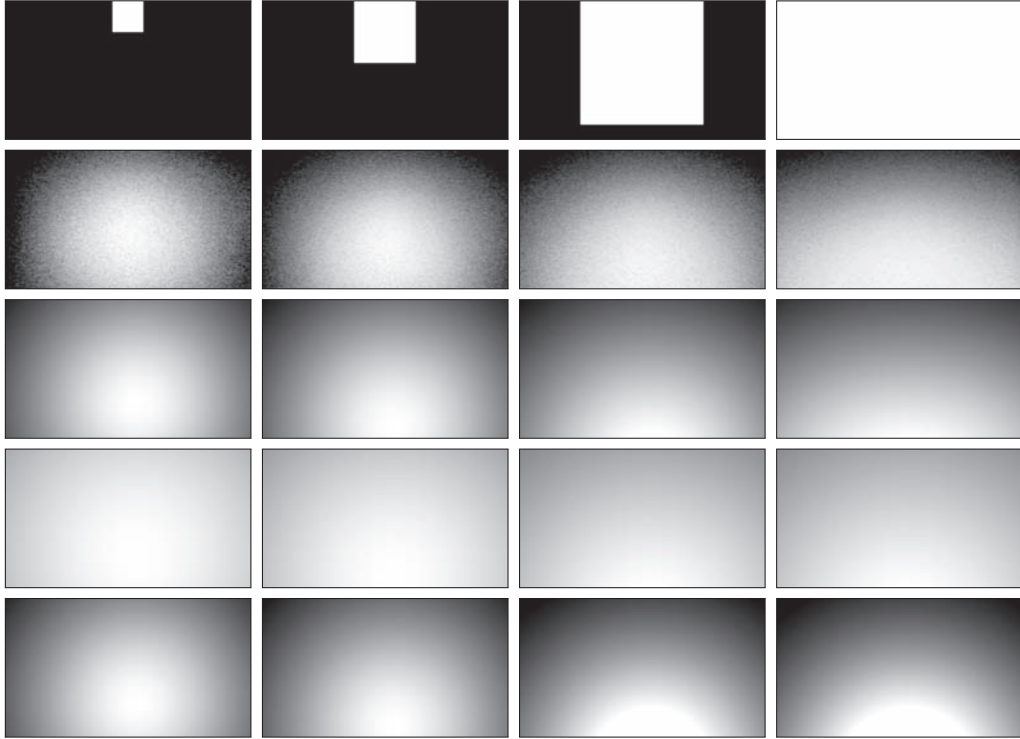


FIGURE 5.5 – Évaluation qualitative du modèle de source rectangulaire uniforme, pour des motifs Ω_s de taille variable. De haut en bas : motifs affichés à l'écran, de tailles respectives 200×200 , 400×400 , 800×800 et 1600×900 (pleine résolution) ; donnée mesurée sur la feuille de papier (membre de gauche de l'équation (5.36)) ; modèle (membre de droite de l'équation (5.36)) en utilisant, de haut en bas, notre modèle (équations (5.26) et (5.32)), un modèle de source étendue isotrope [61] et un modèle de source ponctuelle anisotrope [94].

Images réelles Nous modélisons maintenant la lumière émise par une image en niveaux de gris affichée sur un écran, ou plutôt par son approximation « *quadtree* » [90]. Cela nous fournit une partition rectangulaire non uniforme de l'image en n rectangles (cf. figure 5.6). L'expression de l'éclairage est alors (5.33), où la valeur choisie pour L_0^i est le niveau de gris moyen du rectangle i . Ce modèle est expérimentalement comparé, qualitativement et quantitativement, à celui de Clark [61] sur la figure 5.6 et dans la table 5.1.

Image	$n = 4$		$n = 64$		$n = 256$	
	[61]	Proposé	[61]	Proposé	[61]	Proposé
<i>Chien</i>	1,14	1,09	1,12	0,98	1,10	0,98
<i>Montagne</i>	2,20	0,78	2,19	0,67	2,19	0,66
<i>Soleil</i>	1,46	1,17	1,55	1,14	1,56	1,15
<i>Lac</i>	1,76	1,11	1,62	0,97	1,62	0,95
<i>Mur</i>	1,34	0,94	1,53	0,94	1,55	0,98
<i>Cimetière</i>	1,05	1,04	1,15	0,82	1,17	0,79
<i>Bâtiment</i>	1,36	0,96	1,42	0,88	1,43	0,88
<i>Ruelle</i>	1,43	1,13	1,66	0,57	1,70	0,55
<i>Fleur</i>	1,90	1,03	1,95	0,97	1,96	0,98
<i>Grotte</i>	3,05	1,26	2,76	1,02	2,76	1,01

TABLE 5.1 – RMSE (multipliée par 10^4) entre la luminance normalisée mesurée sur la feuille de papier et celles simulées avec le modèle de Clark [61] ou avec notre modèle, pour les dix images de la figure 5.4. Notre modèle surpasse systématiquement celui de Clark, ce qui confirme l'importance de la prise en compte de l'anisotropie dans le modèle.

Ces expériences confirment que le modèle proposé permet de simuler de façon réaliste divers éclairages non directionnels. Nous avons constaté que $n = 64$ représentait un bon compromis entre la précision du modèle ($n = 1$ correspond à l'approximation grossière de la luminance par le niveau de gris moyen de l'image) et la rapidité d'évaluation (si n est égal au nombre de pixels de l'écran, nous obtenons l'approximation discrète de l'intégrale (5.24), qui est inutilisable pour de grandes images).

5.1.6 Champ d'éclairage non paramétré

L'ensemble des modèles d'éclairage que nous avons dérivés du modèle de source primaire lambertienne (5.12) sont *explicites*, mais ce caractère explicite pose plusieurs questions.

Tout d'abord, les modèles explicites dépendent des paramètres de la source : par exemple, la position du centre d'une source sphérique, tout comme la forme d'une source étendue, influent directement sur l'expression du champ vectoriel $\mathbf{s}(\mathbf{x})$. Ce champ dépend donc d'un ensemble de paramètres qui doivent être soit étalonnés (ce qui est le principe de la stéréophotométrie *calibrée*), soit estimés automatiquement au cours de la reconstruction 3D (stéréophotométrie *non calibrée* : ce problème a été étudié pour des éclairages directionnels, cf. chapitre 3).

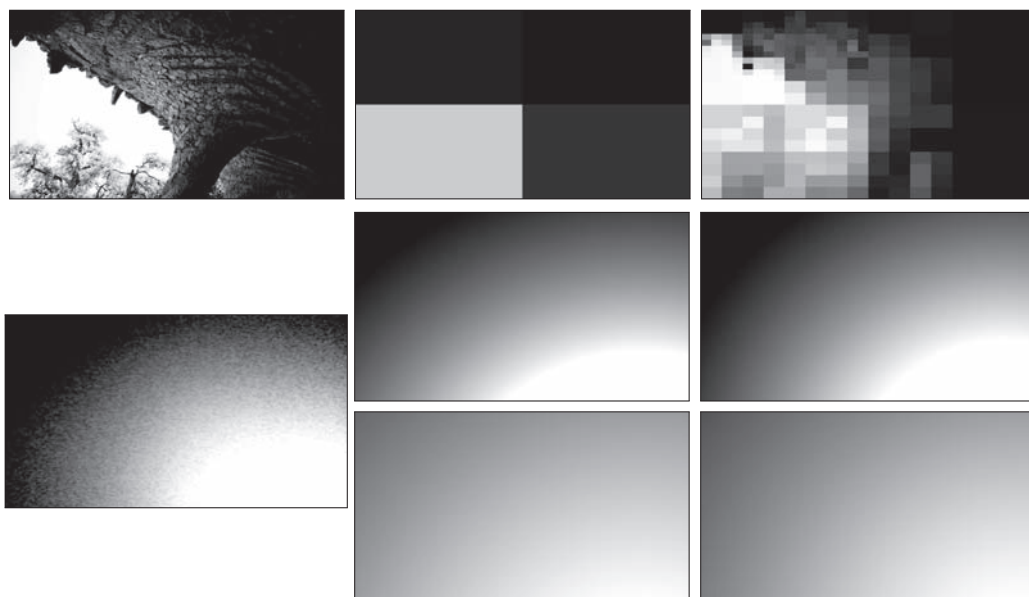


FIGURE 5.6 – Évaluation qualitative du modèle de luminance pour une source étendue non uniforme. À gauche : image affichée à l'écran (en haut) et luminance normalisée mesurée sur la feuille de papier (en bas). Comme l'écran et la webcam sont tous deux orientés vers la feuille de papier, une zone claire sur la partie gauche de l'image affichée se traduit par une zone claire sur la partie droite de la feuille. Au milieu : approximation de l'image par $n = 4$ rectangles (en haut) ; luminance simulée en utilisant soit notre modèle (au milieu), soit le modèle de Clark [61] (en bas). À droite : idem pour $n = 256$.

D'autre part, un modèle explicite dépend des coordonnées du point $\mathbf{x}$, qui constituent justement l'inconnue dans tout problème de reconstruction 3D. L'utilisation de tels modèles en stéréophotométrie augure donc de difficultés, tant du point de vue numérique que de celui du caractère bien posé du problème. Nous reviendrons sur ces questions dans le paragraphe 5.3.

Nous avons également mis en évidence le problème des ombres portées, ainsi que celui de la pénombre dans le cas de sources étendues : ces phénomènes sont difficiles à expliquer car ils dépendent de la géométrie globale de la scène. Aussi précis soit-il, un modèle explicite peut donc être partiellement ou totalement mis en défaut par la complexité de la géométrie de la scène observée.

Enfin, s'il semble raisonnable de chercher à modéliser le flux primaire, le flux lumineux réel reçu par la surface est égal à la somme du flux primaire et de flux *secondaires* dus aux réflexions mutuelles (cf. figure 1.6). Nous avons déjà signalé que ces phénomènes étaient relativement faciles à simuler par lancer de rayons, mais extrêmement complexes à prendre en compte en reconstruction 3D car, comme les ombres et la pénombre, ils dépendent de la géométrie globale de la scène.

Coupler la stéréophotométrie avec les techniques modernes de rendu, afin d'intégrer le calcul de la visibilité au processus de reconstruction 3D, constitue une piste de recherche prometteuse. Nous avons choisi une autre approche, qui consiste à *abstraire* le champ d'éclairage $\mathbf{s}(\mathbf{x})$: cette approche présente l'avantage de bien mieux se prêter à l'optimisation numérique. On peut toujours écrire que le vecteur d'éclairage total est obtenu en sommant les contributions de toutes les sources élémentaires :

$$\mathbf{s}(\mathbf{x}) = \iint_{\mathbf{x}_s \in \text{sources}} d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) \quad (5.37)$$

quitte à supposer que, si une source élémentaire est masquée, alors $d\mathbf{s}_{\mathbf{x}_s}(\mathbf{x}) = \mathbf{0}$, de façon à ce que cette source élémentaire ait une contribution nulle dans l'intégrale. Puisque nous ne cherchons pas à expliciter cette intégrale, rien ne nous empêche non plus de supposer que les sources élémentaires *secondaires* sont également prises en compte dans le calcul. Par conséquent, que le flux lumineux incident soit purement primaire ou non, et en présence ou non d'ombres (ou de pénombre), on pourra toujours écrire que l'éclairage est un champ quelconque de $\mathbb{R}^3$ dans $\mathbb{R}^3$:

$$\mathbf{s}(\mathbf{x}) = \phi(\mathbf{x})\overline{\mathbf{s}(\mathbf{x})} \quad (5.38)$$

et considérer que $\mathbf{s}(\mathbf{x})$ représente le flux lumineux incident *effectif* au point $\mathbf{x}$ de la surface. Cette approche nécessite toutefois de connaître parfaitement le champ $\mathbf{s}$, qui est un champ vectoriel dense de l'espace à trois dimensions. En pratique, il faut connaître un vecteur de $\mathbb{R}^3$ en chaque point de l'espace 3D et pour chaque configuration d'éclairage, ce qui constitue un problème. Comme nous ne souhaitons pas expliciter sa dépendance en $\mathbf{x}$, nous préférons paramétrer le champ $\mathbf{s}$ par les coordonnées (u, v) :

$$\begin{aligned} \mathbf{s} : \Omega \subset \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (u, v) &\longmapsto \mathbf{s}(u, v) = \mathbf{s}(\mathbf{x}(u, v)) \end{aligned} \quad (5.39)$$

Pour représenter de façon plus compacte les champs $\mathbf{s}^i$, on peut les supposer suffisamment réguliers et les décomposer en harmoniques sphériques [23, 25]. Il a été montré dans [24, 26] que dans le contexte de la stéréophotométrie, la reconstruction 3D sous des éclairages très généraux devenait possible à partir de neuf harmoniques sphériques : 9m scalaires suffisent donc pour représenter de façon assez précise chaque éclairage.

Cependant, et comme cela est illustré sur la figure 5.7, un champ d'éclairage n'est décomposable en harmoniques sphériques que s'il est régulier. En effet, les harmoniques sphériques constituent une base orthonormée de l'espace de Hilbert $L^2(\mathbb{S}^2)$ contenant les fonctions sphériques de carré intégrable, définies sur la sphère unité $\mathbb{S}^2 \subset \mathbb{R}^3$. Les travaux de Basri et al. [26, 23, 24, 25] montrent qu'en général, le flux lumineux peut être effectivement assimilé à une fonction de ce type, mais ceci n'est plus vrai en présence d'occultations. En effet, les phénomènes d'ombres et de réflexions mutuelles violent la contrainte de régularité du champ d'éclairage : le champ $\mathbf{s}$, s'il n'est pas continu, n'est plus décomposable en harmoniques sphériques. Nous avons montré dans [216] qu'il pouvait cependant être supposé régulier *par morceaux* : on peut par exemple supposer seulement qu'il s'agit d'un champ à variations bornées, pouvant comporter des « sauts » importants près des discontinuités de la surface, mais lisse ailleurs.

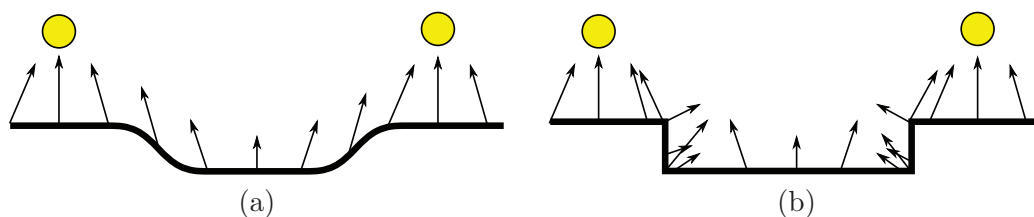


FIGURE 5.7 – (a) Si deux sources ponctuelles sont positionnées à proximité d’une surface régulière, de telle sorte qu’aucune ne soit occultée, le champ d’éclairage est régulier, donc décomposable en harmoniques sphériques. (b) En revanche, les discontinuités de profondeur d’une surface créent des phénomènes d’ombre ou de pénombre, ainsi que des réflexions mutuelles : l’apparition de ces phénomènes brise la régularité du champ d’éclairage, qui n’est plus lisse que par morceaux. Un tel champ n’est pas décomposable en harmoniques sphériques.

5.2 Étalonnage photométrique

Pour pouvoir utiliser les modèles d’éclairage explicites dans le contexte de la stéréophotométrie calibrée, il faut étalonner leurs paramètres au moyen d’une procédure spécifique. Les paramètres du modèle d’éclairage directionnel sont très faciles à estimer en utilisant une sphère spéculaire (la direction indiquée par le centre du reflet brillant est directement liée à la direction d’éclairage) ou mate (la normale à la sphère étant facile à déterminer, on peut inverser le modèle lambertien linéaire pour déterminer l’intensité et la direction moyenne de l’éclairage). Le problème de l’étalonnage dans un cadre plus général représente cependant un problème plus difficile.

5.2.1 Estimation des paramètres d’une source lumineuse

Nous avons montré dans [220] que, pour une source étendue, plane, de luminance non uniforme, comme par exemple un écran, il suffit d’étalonner géométriquement la caméra par rapport à l’écran pour pouvoir étalonner photométriquement la source, de façon à expliciter l’orientation commune des sources élémentaires (les pixels) et leurs coordonnées 3D, ce qui simplifie de beaucoup l’étalonnage, en comparaison de méthodes dédiées telles que celle qui est décrite dans [285].

L’étalonnage d’une source ponctuelle est un problème bien connu, mais la plupart des méthodes existantes [2, 13, 108, 207, 236, 248] ne prennent en compte ni l’atténuation en $1/\text{distance}^2$, ni l’anisotropie de la source, ce qui peut conduire à des résultats très imprécis. À notre connaissance, la seule méthode tenant compte de façon explicite de ces phénomènes dans la procédure d’étalonnage est celle proposée par Xie et al. dans [275], qui repose cependant sur plusieurs étapes difficiles, notamment la détection des spécularités et l’estimation de la distance à la mire. Mentionnons également le travail de Zhou et al., qui permet d’estimer les paramètres d’un modèle d’éclairage très général utilisable avec des techniques de lancer de rayons [286].

Nous présentons dans ce qui suit une nouvelle méthode d'étalonnage d'un éclairage par inversion du modèle photométrique⁴, qui présente l'avantage de fournir en même temps les paramètres internes de l'appareil photographique. Pour cela, nous utilisons un motif de damier, imprimé en noir et blanc sur une feuille de papier collée sur un support plan. L'albédo des cases blanches est supposé égal à 1 (valeur arbitraire qui permet de simplifier les notations). L'albédo des cases noires n'est probablement pas nul, mais sa valeur est sans importance, car nous n'utiliserons pas les niveaux de gris correspondants. D'après la loi de Lambert, le niveau de gris de l'image prise sous un éclairage $\mathbf{s}$ s'écrit, pour un pixel d'une case blanche $(u, v) \in \Omega_b$, où Ω_b désigne l'ensemble des cases blanches :

$$I(u, v) = \mathbf{s}_\theta(\mathbf{x}(u, v)) \cdot \mathbf{n}(u, v) \quad (5.40)$$

où $\theta \in \mathbb{R}^{m_s}$ est le vecteur des paramètres de la source, et $\mathbf{x}(u, v)$ le point 3D du damier conjugué de $(u, v) \in \Omega_b$. Rappelons que $m_s = 3$ pour une source directionnelle (deux degrés de liberté pour la direction, un pour l'intensité), $m_s = 4$ pour une source ponctuelle lambertienne isotrope (trois degrés de liberté pour la position 3D, un pour l'intensité), et $m_s = 7$ pour une source ponctuelle lambertienne imparfaite (trois degrés de liberté pour la position 3D, un pour l'intensité, deux pour l'orientation de la source, un pour le coefficient d'anisotropie μ défini en (5.21)).

Puisque tous les points de la surface sont situés sur un même plan (celui du damier), ils partagent la même normale, i.e. $\mathbf{n}(u, v) = \mathbf{n}$, de sorte que l'équation (5.40) ne contraint pas suffisamment l'estimation de θ , même si $\mathbf{n}$ est connu. Pour parer à ce problème, nous utilisons m_{poses} poses du damier, et notons :

$$I^j(u, v) = \mathbf{s}_\theta(\mathbf{x}^j(u, v)) \cdot \mathbf{n}^j \quad (5.41)$$

le niveau de gris du pixel $(u, v) \in \Omega_b^j$, où Ω_b^j désigne l'ensemble des cases blanches pour la j -ème pose du damier. Dans (5.41), $\mathbf{x}^j(u, v)$ désigne le point 3D conjugué de (u, v) pour la j -ème pose du damier.

Les algorithmes de détection de damiers étant très éprouvés, il est facile de déterminer, pour chaque pose, l'ensemble Ω_b^j des cases blanches. De plus, à partir de $m_{\text{poses}} \geq 3$ poses, de nombreux algorithmes permettent d'estimer sans ambiguïté les coordonnées 3D des points $\mathbf{x}^j(u, v)$ du damier. Ce type d'algorithmes permet également d'estimer automatiquement les normales $\mathbf{n}^j$, ainsi que les paramètres internes de la caméra. Pour réaliser ces opérations, nous utilisons la boîte à outils *Computer Vision* de Matlab, sachant que les mêmes fonctions sont librement accessibles dans la bibliothèque *OpenCV*.

Dans (5.41), il ne reste donc plus que les paramètres θ comme inconnues. Ces paramètres étant indépendants de la pose du damier, la solution du problème de minimisation :

$$\min_{\theta \in \mathbb{R}^{m_s}} \sum_{j=1}^{m_{\text{poses}}} \sum_{(u,v) \in \Omega_b^j} \left[I^j(u, v) - \mathbf{s}_\theta(\mathbf{x}^j(u, v)) \cdot \mathbf{n}^j \right]^2 \quad (5.42)$$

4. Cette méthode a été développée par Jade Boumaza et Tom Lucas au cours de leur stage de première année de master à l'IRIT.

constitue une excellente estimation du vecteur de paramètres $\boldsymbol{\theta}$. Il s'agit d'un problème d'estimation en moindres carrés de $m_s < 10$ paramètres à partir de $\sum_j |\Omega_b^j| \gg 10$ équations : le théorème central limite nous garantit qu'il s'agit de l'estimateur efficace. Dans le cas du modèle directionnel, il s'agit d'un problème en moindres carrés linéaires qui peut être résolu analytiquement. Dans le cas général, le problème est non linéaire, mais tellement surcontraint qu'une minimisation de type Levenberg-Marquardt, à partir d'une initialisation grossière des paramètres, offre une estimation très robuste. Cette estimation est d'autant plus robuste que le nombre de poses m_{poses} est plus grand.

Par exemple, le problème de l'étalonnage d'une source ponctuelle de type LED, modélisée comme une source lambertienne imparfaite ($m_s = 7$) s'écrit :

$$\min_{\substack{\mathbf{x}_s, \phi_0 \\ \mathbf{n}_0, \mu}} \sum_{j=1}^{m_{\text{poses}}} \sum_{(u,v) \in \Omega_b^j} \left\{ I^j(u,v) - \frac{((\phi_0)^{1/\mu} \mathbf{n}_0 \cdot [\mathbf{x}^j(u,v) - \mathbf{x}_s])^\mu}{\|\mathbf{x}^j(u,v) - \mathbf{x}_s\|^{3+\mu}} \left[(\mathbf{x}_s - \mathbf{x}^j(u,v)) \cdot \mathbf{n}^j \right] \right\}^2 \quad (5.43)$$

où $\mathbf{x}_s$ désigne la position de la source, ϕ_0 son intensité, $\mathbf{n}_0$ son orientation, et μ son coefficient d'anisotropie. En imposant $\mu = 0$, le modèle est indépendant de $\mathbf{n}_0$ (source isotrope), tandis que $\mu \geq 1$ correspond au modèle de source ponctuelle anisotrope.

La résolution du problème (5.43) est difficile si μ est inconnu, mais facile si μ est connu. Plutôt que d'estimer ϕ_0 et $\mathbf{n}_0$ séparément, ce qui nécessiterait de respecter la contrainte $\|\mathbf{n}_0\| = 1$, il est plus raisonnable d'estimer le vecteur $\mathbf{m}_0 = (\phi_0)^{1/\mu} \mathbf{n}_0$, duquel on peut extraire $\phi_0 = \|\mathbf{m}_0\|^\mu$ et $\mathbf{n}_0 = \mathbf{m}_0 / \|\mathbf{m}_0\|$. La méthode de Levenberg-Marquardt⁵ converge alors très rapidement (une dizaine d'itérations dans nos tests).

On peut soit faire confiance à la valeur de μ fournie par le fabricant, soit l'estimer en appliquant la procédure qui vient d'être décrite pour plusieurs valeurs de μ . En effet, comme la valeur minimale du critère (5.43) dépend fortement de μ , la valeur qui minimise ce minimum peut être choisie comme estimation (cf. figure 5.9). Pour l'exemple des trois LED de la figure 5.8⁶, le fabricant⁷ indique que $\theta_{1/2} \approx 60^\circ$, soit $\mu \approx 1$ d'après (5.22), pour les trois LED. En réalité, la LED rouge semble légèrement plus anisotrope ($\mu \approx 4$). Notons au passage que ce montage nous a permis de tester la stéréophotométrie colorée (cf. paragraphe 4.4.1), et de coupler cette technique avec un capteur de profondeur. Ces travaux seront présentés dans un travail à venir.

5. Notons quand même que les coordonnées 3D de la source et son intensité constituent des grandeurs physiques totalement différentes, et donc d'ordres de grandeur potentiellement très différents : il est nécessaire d'utiliser un préconditionnement approprié pour garantir la convergence.

6. Contrairement à ce que montre la photographie du montage, et conformément à la procédure proposée, il nous faut allumer les LED une par une pour pouvoir les étalonner, les LED n'étant pas parfaitement monochromatiques.

7. <http://www.luxeonstar.com/lxml-pb01-0040-blue-luxeon-rebel-led-470nm>

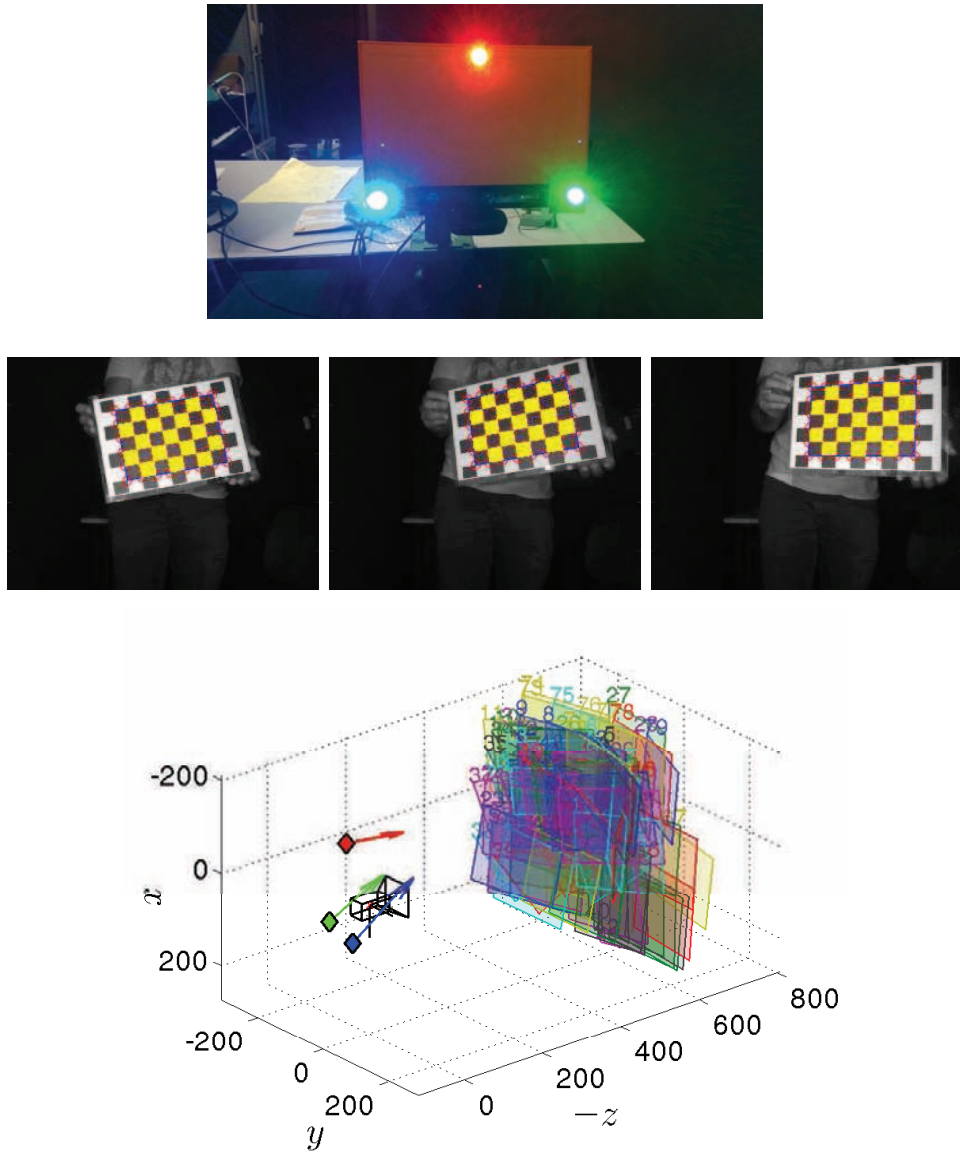


FIGURE 5.8 – Étalonnage des paramètres d'un modèle de source ponctuelle lambertienne imparfaite. Première ligne : trois LED colorées fixées sur une caméra RGBD. Deuxième ligne : trois poses du damier, avec détection automatique des cases blanches. En bas : reconstruction 3D de l'environnement complet de l'étalonnage (poses du damier et caméra étalonnée), accompagnée des paramètres du modèle d'éclairage, ici les positions et orientations 3D de trois LED colorées situées autour de la caméra RGBD. L'unité est le millimètre, et la position de chaque LED est désignée par un losange, son orientation par une flèche dont la longueur est proportionnelle à l'intensité.

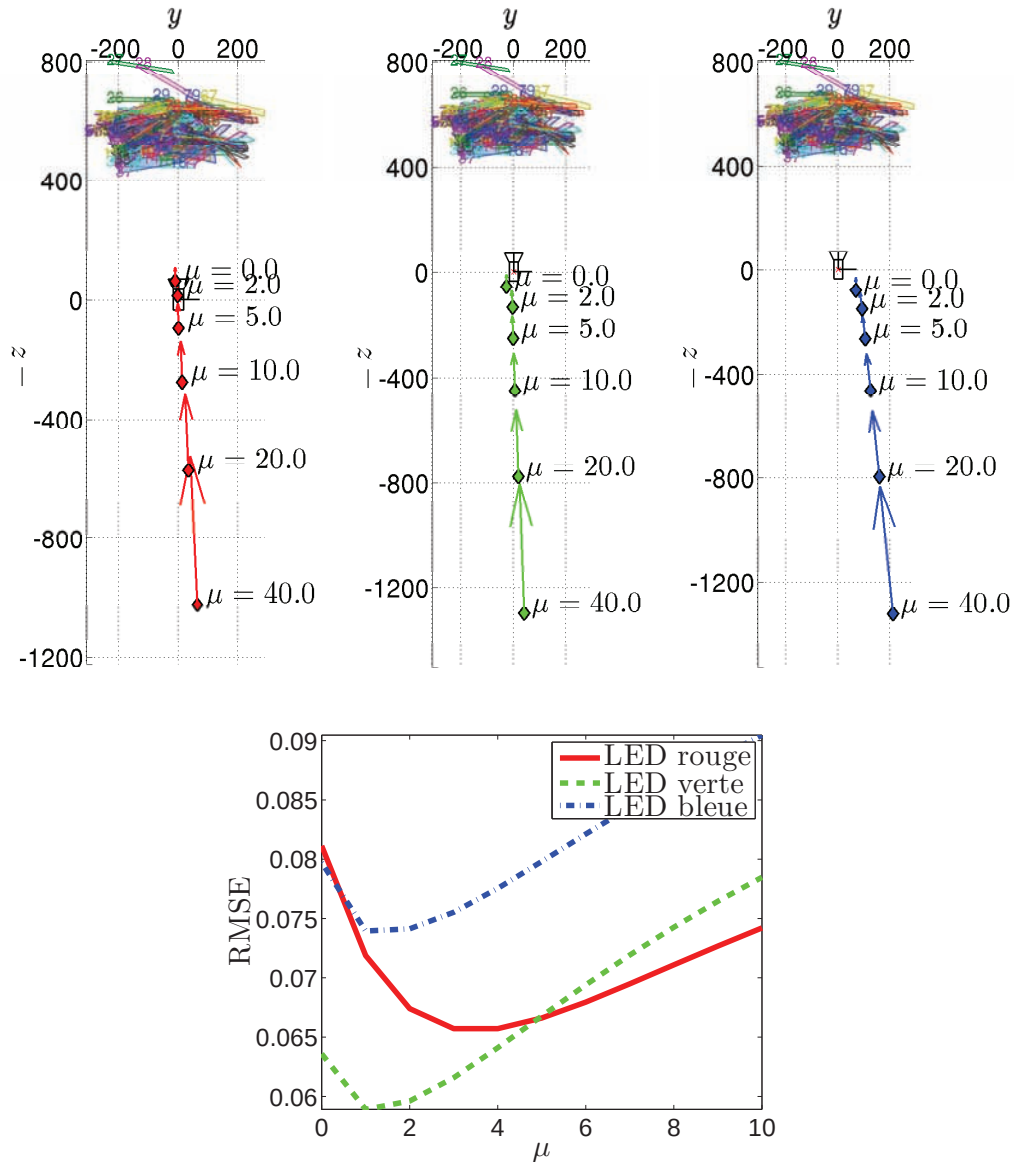


FIGURE 5.9 – Influence du paramètre d'anisotropie μ sur l'étalonnage photométrique. En haut : positions et orientations du damier et positions, orientations et intensités des sources qui sont visibles sur la première ligne de la figure 5.8, pour différentes valeurs de μ . Si ce paramètre d'anisotropie est trop élevé, cela se traduit par un éloignement des sources. En bas : racine carrée de la valeur moyenne du critère défini en (5.43) (RMSE). Les valeurs fournies par le fabricant sont $\mu = 1$ pour les trois LED : la LED rouge semble en réalité plus anisotrope que ce qui est annoncé.

5.2.2 Étalonnage du champ d'éclairage

Il est également possible de directement étalonner le flux incident $\mathbf{s} : \Omega \rightarrow \mathbb{R}^3$, en inversant le modèle lambertien linéaire en chaque point $\mathbf{x}$. Ceci suppose cependant que l'on dispose d'un grand nombre de poses de la mire en chaque point de l'espace 3D, ce qui semble très difficile à mettre en œuvre.

Si l'on ne tient compte que du flux primaire, et que l'on suppose que l'intensité de l'éclairage varie spatialement, mais que sa direction est à peu près uniforme, l'étalonnage du champ $\mathbf{s}$ se ramène à la technique de « compensation » qui a été décrite dans le paragraphe 1.2.

Afin de garantir un étalonnage des éclairages plus fiable que par cette procédure somme toute simpliste, et pour éviter un étalonnage exhaustif du champ $\mathbf{s}$ sur tout l'espace 3D, nous avons mis en place une procédure d'étalonnage intermédiaire, dans le cadre d'un projet de transfert technologique en collaboration avec la société Pixience, piloté par la SATT de Toulouse Midi-Pyrénées (TTT, pour *Toulouse Tech Transfer*). Cette procédure vise à estimer le vecteur d'éclairage en un nombre restreint $m_{\text{échantillons}}$ d'échantillons $\mathbf{x}(u_0^e, v_0^e)$ de l'espace 3D, $e \in [1, m_{\text{échantillons}}]$, répartis sur le plan moyen de la surface à reconstruire (cf. figure 5.10). Ceci est effectué par résolution, au voisinage de chaque point d'échantillonnage (u_0^e, v_0^e) , d'un problème de minimisation en moindres carrés linéaires de la forme (5.42), en choisissant (localement) le modèle d'éclairage directionnel, i.e. $\boldsymbol{\theta}^i \in \mathbb{R}^3$. En utilisant une mire 3D (usinée à partir de son modèle), on peut disposer de suffisamment d'orientations différentes de la surface au voisinage de chaque point (u_0^e, v_0^e) ⁸. Remarquons qu'une idée similaire a été proposée par Johnson et al. dans [138], pour étalonner les harmoniques sphériques du flux lumineux incident. Cette procédure permet d'étalonner le flux incident sur une surface de référence. Bien sûr, le flux incident effectif sur une nouvelle surface sera légèrement différent, notamment à cause des réflexions mutuelles. Nous verrons dans la partie IV comment affiner l'estimation de ce flux au cours de la reconstruction 3D.

5.3 Méthodes numériques de reconstruction 3D par stéréophotométrie sous des éclairages non directionnels

Les méthodes numériques d'estimation de la normale et de l'albédo doivent être adaptées, afin de prendre en compte le caractère non directionnel des éclairages. Avec un éclairage non directionnel décrit par le champ vectoriel $\mathbf{s} : (u, v) \rightarrow \mathbf{s}(\mathbf{x}(u, v))$, le système d'équations induit par le modèle lambertien linéaire s'écrit, en chaque pixel $(u, v) \in \Omega$:

$$\underbrace{\begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix}}_{\mathbf{i}(u, v) \in \mathbb{R}^m} = \underbrace{\begin{bmatrix} \mathbf{s}^1(\mathbf{x}(u, v))^\top \\ \vdots \\ \mathbf{s}^m(\mathbf{x}(u, v))^\top \end{bmatrix}}_{\mathbf{S}(\mathbf{x}(u, v)) \in \mathbb{R}^{m \times 3}} \underbrace{\rho(u, v) \mathbf{n}(u, v)}_{\mathbf{m}(u, v) \in \mathbb{R}^3} \quad (5.44)$$

8. Ce protocole d'étalonnage a été mis en place par Mathieu Pizenberg, dans le cadre de son stage de première année de master à l'IRIT.

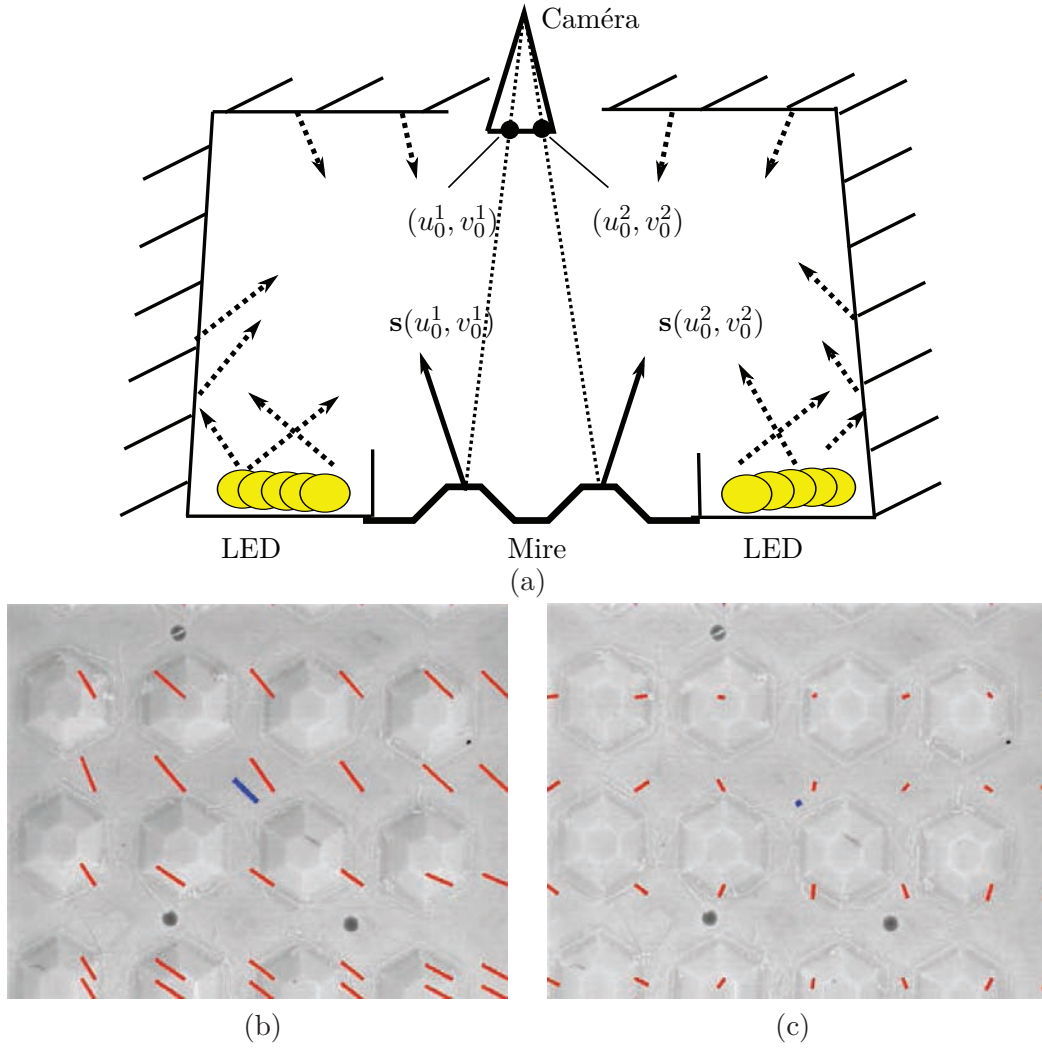


FIGURE 5.10 – Étalonnage du flux lumineux incident sur un plan moyen. (a) Le *dermoscope numérique* (appareil permettant d’observer la peau à très haute résolution) de la société Pixience est équipé d’un ensemble de LED qui éclairent la surface indirectement, de façon à créer un éclairage diffus sur la surface observée, et ce afin d’éviter les ombres. Le champ lumineux atteignant la surface est donc difficile à modéliser explicitement. En revanche, il est facile d’étalonner le flux incident en un nombre fini de points de la surface d’un objet de relief connu, comme cela est illustré en (b) et (c), où l’éclairage est désigné, en direction et en intensité, par des flèches rouges (la flèche bleue illustre l’estimation obtenue avec un modèle d’éclairage directionnel). On peut ensuite interpoler ces échantillons pour densifier l’estimation du flux incident.

Dans (5.44), si $\mathbf{x}(u, v)$ est connu, alors la solution en moindres carrés en $\mathbf{m}(u, v)$ est obtenue en calculant, en chaque pixel $(u, v) \in \Omega$, la pseudo-inverse de la matrice $\mathbf{S}(u, v)$. Une estimation robuste peut être obtenue en utilisant un M-estimateur, comme cela est décrit dans le chapitre 1. Ceci nous a permis, par exemple, d'obtenir la reconstruction 3D de la figure 5.11, après étalonnage du flux incident selon la procédure décrite dans le paragraphe 5.2.2.

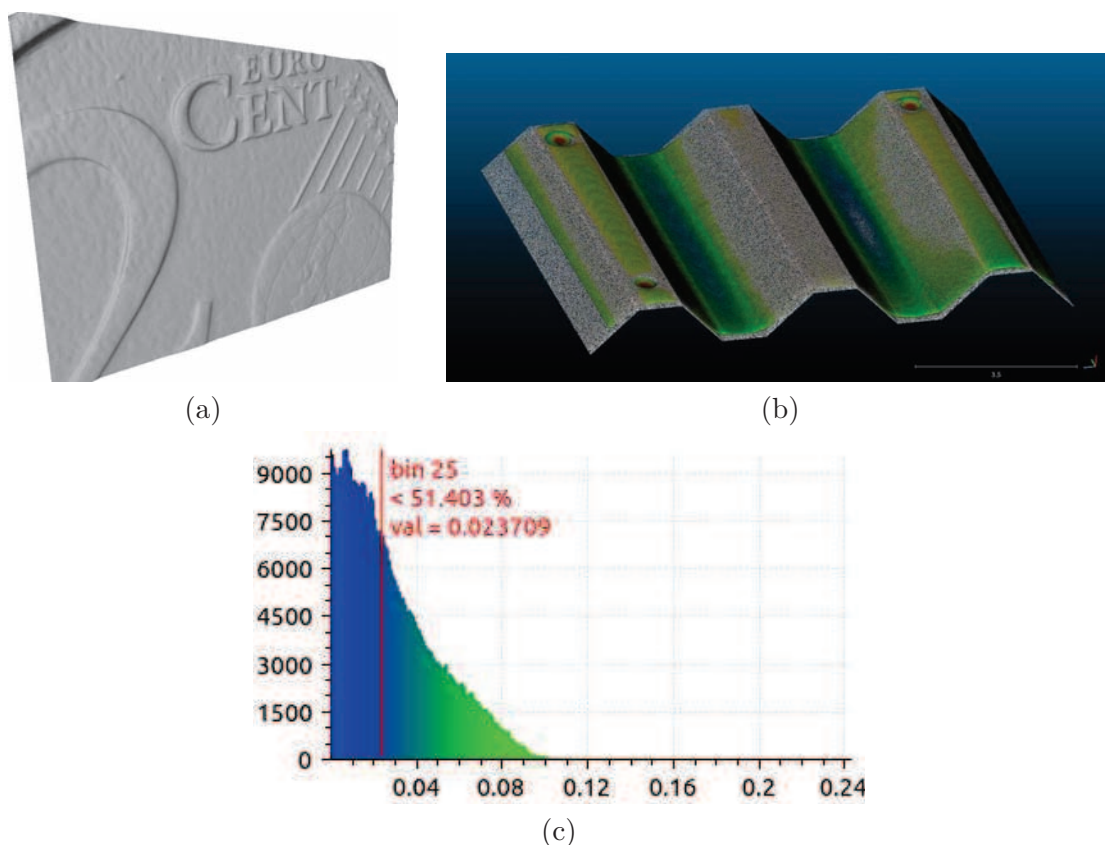


FIGURE 5.11 – Reconstruction 3D à partir du flux incident étalonné selon la procédure schématisée sur la figure 5.10. Nous avons transformé le dermoscope numérique de la société Pixience en un *scanner 3D*, de façon à estimer le relief de surfaces localement planes, comme la pièce de monnaie (a), que nous avons peinte avec une peinture mate afin de garantir une réflectance diffuse fidèle au modèle lambertien. (b) Nous avons également pu comparer le résultat de la reconstruction 3D d'une mire en plastique mat avec le modèle 3D ayant servi à son usinage : le modèle 3D de référence est affiché en gris, le modèle estimé est affiché en couleur. (c) L'écart moyen de la reconstruction 3D avec le modèle 3D est de l'ordre de 25 microns, l'écart absolu étant inférieur à 25 microns pour plus de 50% des points, et à 100 microns pour plus de 99% des points.

5.3. MÉTHODES NUMÉRIQUES DE RECONSTRUCTION 3D PAR STÉRÉOPHOTOMÉTRIE SOUS DES ÉCLAIRAGES NON DIRECTIONNELS

Lorsque l'on considère un modèle paramétrique explicite, e.g. un modèle de source ponctuelle lambertienne imparfaite, les coordonnées $\mathbf{x}$ intervenant dans la valeur de l'éclairage $\mathbf{s}$ sont les inconnues du problème de la reconstruction 3D. Des méthodes heuristiques permettent de déterminer de façon approchée les coordonnées des points $\mathbf{x}$, et d'en déduire $\mathbf{S}(\mathbf{x})$: Ahmad et al. ont montré dans [7] que ceci pouvait être réalisé en identifiant les maxima locaux diffus dans les images. L'utilisation d'un capteur de profondeur peut également permettre de résoudre ce problème. Dans un deuxième temps, l'estimation de la normale et de l'albédo peut être effectuée pixel par pixel, en utilisant les méthodes décrites dans le chapitre 1.

Une autre possibilité pour prendre en compte la dépendance explicite de l'éclairage vis-à-vis des points de la surface consiste à se rappeler que, si la normale à la surface est connue, alors les coordonnées 3D peuvent être obtenues par *intégration* des normales, comme cela est décrit dans le chapitre 2. Rappelons que, outre l'équation du niveau de gris (5.44), nous disposons également des deux équations :

$$\nabla z(u, v) = \mathbf{g}(\mathbf{n}(u, v)) \quad (5.45)$$

$$\mathbf{x}(u, v) = \mathbf{h}(z(u, v)) \quad (5.46)$$

et que pour une caméra orthographique :

$$\mathbf{g}_o(\mathbf{n}(u, v)) = \begin{bmatrix} -\frac{n_1(u, v)}{n_3(u, v)} \\ -\frac{n_2(u, v)}{n_3(u, v)} \end{bmatrix} \quad (5.47)$$

$$\mathbf{h}_o(z(u, v)) = \begin{bmatrix} u \\ v \\ z(u, v) \end{bmatrix} \quad (5.48)$$

alors que pour une caméra perspective :

$$\mathbf{g}_p(\mathbf{n}(u, v)) = \begin{bmatrix} -\frac{n_1(u, v)}{[u-u_0, v-v_0, f]^\top \cdot \mathbf{n}(u, v)} \\ -\frac{n_2(u, v)}{[u-u_0, v-v_0, f]^\top \cdot \mathbf{n}(u, v)} \end{bmatrix} \quad (5.49)$$

$$\mathbf{h}_p(z(u, v)) = z(u, v) \begin{bmatrix} \frac{u-u_0}{f} \\ \frac{v-v_0}{f} \\ 1 \end{bmatrix} \quad (5.50)$$

On peut introduire ces contraintes dans le problème, en cherchant à minimiser un critère de la forme :

$$\begin{cases} \min_{\rho, \mathbf{n}, z, \mathbf{x}} \sum_{(u, v) \in \Omega} \Phi_n (\|\mathbf{i}(u, v) - \rho(u, v) \mathbf{S}(\mathbf{x}(u, v)) \mathbf{n}(u, v)\|) \\ \quad + \sum_{(u, v) \in \Omega} \Phi_i (\|\nabla z(u, v) - \mathbf{g}(\mathbf{n}(u, v))\|) \\ \text{s.c. } \mathbf{x}(u, v) = \mathbf{h}(z(u, v)) \end{cases} \quad (5.51)$$

où Φ_n et Φ_i sont deux M-estimateurs choisis de façon à garantir la robustesse à la fois de l'estimation de la normale (cf. chapitre 1) et de l'intégration des normales (cf. chapitre 2), et où nous notons abusivement ∇ l'approximation de l'opérateur gradient par différences finies. Toutefois, rappelons-nous que le problème de l'intégration n'admet généralement de solution qu'à une constante près (additive dans le cas orthographique, multiplicative dans le cas perspectif) : on peut par exemple fixer cette constante arbitrairement, en fonction d'une estimation empirique de la distance moyenne à la surface [198], ou la traiter comme une inconnue supplémentaire du problème⁹. On peut également utiliser une régularisation à l'ordre 0 de la profondeur, comme cela est expliqué dans le chapitre 2, en ajoutant à la fonction objectif du problème (5.51) un terme de la forme $\lambda(z - z_0)^2$. Le problème est bien posé dès lors que $\lambda > 0$.

Le problème (5.51) pourrait probablement être résolu de façon très efficace par un algorithme de type ADMM (cf. chapitre 7). Nous proposons ici une solution plus simple, en explicitant les parties non linéaires et en rendant implicites les parties linéaires, de sorte que l'albédo, les normales et la profondeur sont estimés par les itérations suivantes :

$$(\rho^{k+1}, \mathbf{n}^{k+1}) = \underset{\rho, \mathbf{n}}{\operatorname{argmin}} \sum_{(u,v) \in \Omega} \Phi_n \left(\left\| \mathbf{i}(u, v) - \rho(u, v) \mathbf{S}(\mathbf{x}^k(u, v)) \mathbf{n}(u, v) \right\| \right) \quad (5.52)$$

$$z^{k+1} = \underset{z}{\operatorname{argmin}} \sum_{(u,v) \in \Omega} \Phi_i \left(\left\| \nabla z(u, v) - \mathbf{g}(\mathbf{n}^{k+1}(u, v)) \right\| \right) + \lambda [z(u, v) - z_0(u, v)]^2 \quad (5.53)$$

$$\mathbf{x}^{k+1}(u, v) = \mathbf{h}(z^{k+1}(u, v)) \quad (5.54)$$

qui consistent à alterner une estimation locale de la surface (orientation et réflectance) et une estimation globale par intégration (on rappelle que si $\lambda = 0$, il faut ajouter une étape intermédiaire de détermination de la constante d'intégration). On se ramène ainsi aux deux problèmes étudiés dans la partie I, et un modèle non linéaire de réflectance est facile à introduire : il suffit de remplacer l'étape (5.52) par une des méthodes qui seront proposées dans le chapitre 6. La figure 5.12 montre un exemple de reconstruction 3D d'un visage obtenue avec cet algorithme et un modèle de source ponctuelle isotrope¹⁰.

Rappelons que ce schéma itératif est, à quelques détails près, le même que celui qui est décrit, entre autres, dans [7, 198, 212]. Enfin, Papadimitri et Favaro [198], ainsi que Huang et al. [129], ont montré que les positions 3D de sources ponctuelles (leurs travaux peuvent probablement être très facilement étendus à d'autres modèles d'éclairage) pouvaient également être affinées au cours de l'estimation, en ajoutant une étape d'étalonnage au schéma explicité ci-dessus.

9. C'est cette dernière procédure qui a été implémentée sur une borne de démonstration de l'IRIT conçue et réalisée par Bastien Durix et Benjamin Rollet, dans le cadre d'un projet de transfert technologique en partenariat avec la société Fitting Box, piloté par Toulouse Tech Transfer.

10. Ce résultat a été obtenu par Bastien Durix dans le cadre du projet déjà mentionné.

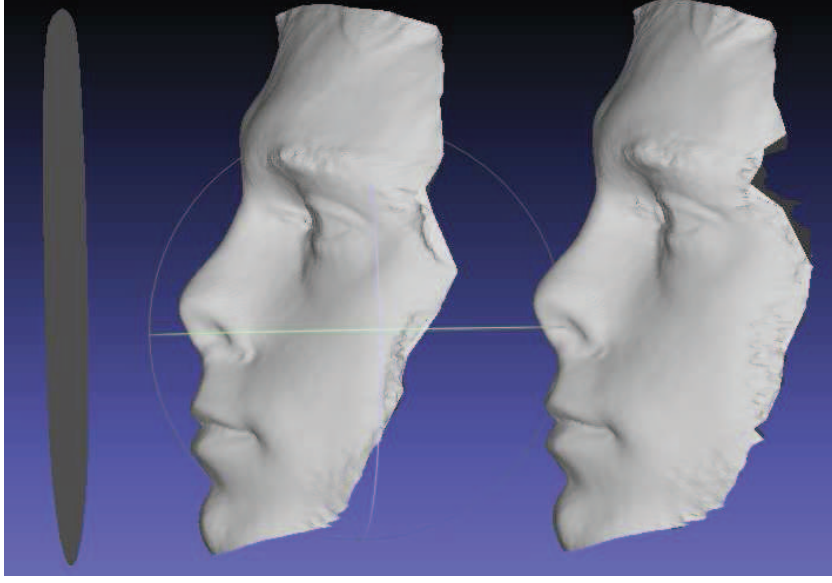


FIGURE 5.12 – Reconstruction 3D itérative par stéréophotométrie, en présence d’éclairages ponctuels isotropes. En partant d’une solution initiale plane (à gauche), et en alternant les phases d’estimation robuste de la normale et d’intégration robuste, l’algorithme converge très rapidement (ici, en trois itérations) vers une solution qualitativement satisfaisante.

5.4 Conclusion du chapitre

Dans ce chapitre, nous avons montré comment modéliser de façon très précise le flux lumineux émis par plusieurs types de sources d’éclairage, et comment s’abstraire des paramètres du modèle d’éclairage. Nous avons ensuite proposé de nouvelles méthodes permettant d’étalonner les paramètres du modèle, ou de mesurer directement le flux incident. Enfin, nous avons montré comment étendre les méthodes numériques de reconstruction 3D présentées dans la partie I au cas d’éclairages non directionnels.

Cette modélisation de l’éclairage permet de formuler le problème de la reconstruction 3D de façon suffisamment réaliste, pour des applications où le caractère directionnel de l’éclairage est difficile à garantir : nous pensons notamment à l’éclairage provenant d’une source ponctuelle à bas coût de type LED, ou au champ d’éclairage produit par un dispositif complexe tel qu’un dermoscope.

La modélisation précise de l’éclairage est d’autant plus nécessaire que l’éclairage primaire peut être totalement contrôlé dans les applications, au contraire de la surface à reconstruire. En particulier, tout comme nous avons remis en question le modèle d’un éclairage directionnel, il nous faut maintenant nous interroger sur la pertinence du modèle lambertien pour décrire la réflectance de la surface. Cela fait précisément l’objet du prochain chapitre.

Modèles de réflectance non lambertiens

Dans ce chapitre, nous commençons par revisiter la résolution de la stéréophotométrie par apprentissage, telle que l’ont proposée Hertzmann et Seitz dans [122, 121]. En utilisant un objet de géométrie connue, il est possible de reconstruire le relief de n’importe quel objet de même réflectance, pourvu que les deux objets soient photographiés dans les mêmes conditions, i.e. du même point de vue et sous les mêmes éclairages. Cette approche permet d’appliquer la stéréophotométrie à des matériaux de réflectance quelconque, et non plus seulement aux matériaux lambertiens. Elle revient à utiliser un modèle de réflectance appris sur les images de l’objet de référence.

Pour éviter les contraintes de l’approche par apprentissage, nous nous intéressons ensuite à des modèles de réflectance *explicites* autres que le modèle lambertien. De tels modèles, qui sont surtout utilisés dans le contexte de la synthèse d’images, peuvent être classés en trois catégories¹ : les modèles « empiriques » (e.g. le modèle de Phong, que nous avons déjà décrit dans le chapitre 1), les modèles « physiques » et les modèles « expérimentaux ». Nous nous attachons plus particulièrement à décrire deux modèles de réflectance non empiriques, à savoir le modèle d’Oren-Nayar (« physique ») et le modèle de Ward (« expérimental »). Nous avons choisi d’étudier deux modèles explicites, mais des modèles *implicites* peuvent également être utilisés [8, 52]. Rappelons enfin qu’il est possible d’appliquer la stéréophotométrie à des matériaux partiellement diffusants, en séparant les composantes diffuse et spéculaire de la réflectance (cf. paragraphe 1.2.1), puis en estimant le relief à partir de la seule composante diffuse [278].

Dans le paragraphe 6.2, nous montrons comment appliquer la stéréophotométrie au modèle d’Oren-Nayar, qui est censé tenir compte de la « rugosité » de certaines surfaces diffuses. Dans le paragraphe 6.3, nous faisons de même pour le modèle de Ward, qui considère la réflectance comme la somme d’une composante diffuse et d’une composante spéculaire. Nous supposons dans les deux cas que la scène est constituée d’un seul matériau, sachant qu’il est généralement possible de segmenter les différents matériaux par des approches markoviennes [102, 101], puis d’appliquer la stéréophotométrie à chaque partie segmentée.

1. Nous renvoyons le lecteur à [179] pour une catégorisation récente des modèles de réflectance.

6.1 Stéréophotométrie par apprentissage

Nous revisitons la résolution de la stéréophotométrie par apprentissage et proposons notamment, dans le paragraphe 6.1.1, une illustration graphique de cette approche, qui permet de clarifier ses limites et les conditions dans lesquelles elle doit être utilisée. Nous étudions ensuite, dans le paragraphe 6.1.2, diverses extensions de la méthode de Hertzmann et Seitz [122, 121] permettant de réduire sa complexité algorithmique et d'améliorer sa robustesse, notamment en présence d'ombres portées et de texture. Ces contributions ont déjà été présentées dans [83].

6.1.1 Principe de l'approche par apprentissage

L'idée d'introduire un objet de référence dans la scène pour faciliter la reconstruction 3D par stéréophotométrie est due à Woodham [268]. Elle a été exploitée pour la première fois par Silver, qui suggère dans [238] de construire une table de correspondance entre les triplets d'intensité mesurés sur une sphère mate blanche, sous trois éclairages directionnels, et le gradient de profondeur de la sphère. En cherchant dans la table le triplet le plus proche de celui mesuré sur un objet mat blanc à reconstruire, on peut déterminer en chaque point le gradient de profondeur. Probablement à cause de la simplicité de résolution de la stéréophotométrie dans sa version « classique » (cf. partie I), cette approche « par apprentissage » n'a d'abord pas suscité beaucoup d'intérêt. Il a fallu attendre les travaux de Hertzmann et Seitz pour qu'elle connaisse un certain succès [122, 121]. En particulier, elle a été généralisée récemment à des séquences d'images issues de vidéos [151], et il a été montré, grâce à l'utilisation conjointe de techniques multi-vues, qu'il était possible de se passer de l'objet de référence [3].

Hertzmann et Seitz font dans [122, 121] le constat suivant, qu'ils appellent la « consistance de l'orientation » : deux points d'une surface ayant la même orientation, i.e. partageant la même normale, réfléchissent la lumière vers l'observateur exactement de la même façon, pourvu qu'ils soient photographiés sous les mêmes éclairages directionnels, que la réflectance soit la même en ces deux points, et qu'aucun des deux points ne se trouve dans une ombre portée. Comme l'illustre l'exemple de la figure 6.1, ce constat est valide pour de nombreuses réflectances, et pas seulement pour le modèle lambertien, ce qui justifie le choix de placer cette approche « par apprentissage » dans un chapitre consacré aux modèles de réflectance.

Illustration graphique par les isophotes

Nous proposons d'illustrer la résolution de la stéréophotométrie par apprentissage en utilisant la notion d'*isophotes*. Un isophote est formé par les points d'une image qui ont un même niveau de gris. Un isophote est généralement une courbe, sauf dans le cas d'une ombre (isophote 0), qui constitue une partie de l'image d'aire non nulle.

Supposons par exemple que nous souhaitions estimer la normale au point (u_1, v_1) du vase repéré par une croix noire sur la figure 6.3-a. Chaque niveau de gris de (u_1, v_1) , lu sur une des images du bas de la figure 6.2, est associé à un isophote dans l'image de



FIGURE 6.1 – Principe de la stéréophotométrie par apprentissage. La sphère constitue l’objet de référence. La bouteille est un objet de même réflectance, de relief inconnu. Les vecteurs de niveaux de gris $\mathbf{i} = [I^1, \dots, I^m]^\top$ sont les mêmes en deux points de même orientation (croix jaunes), si les éclairages sont directionnels et ne provoquent pas d’ombre portée. L’orientation de la surface en un point (u', v') de la surface à reconstruire peut donc être estimée en cherchant le point (u, v) de la sphère pour lequel le vecteur $\mathbf{i}(u, v)$ est le plus proche de $\mathbf{i}(u', v')$.

la sphère correspondant au même éclairage. Chaque éclairage contraint donc le point (u_1, v_1) à avoir la même normale qu’un des points de la sphère situé sur l’isophote correspondant. Si l’intersection de tous les isophotes ainsi définis est unique, alors on a bien trouvé un moyen de lever l’ambiguïté sur la normale. La superposition des cinq isophotes ainsi recueillis montre qu’il en va bien ainsi (cf. figure 6.3-b, où les couleurs utilisées sont les mêmes que celles des vecteurs d’éclairage indiqués sur la figure 6.2). La normale en ce point de la sphère nous donne donc une estimation de la normale au point (u_1, v_1) du vase.

L’estimation de la normale par intersection d’isophotes n’est pas biaisée par la présence d’ombres propres. Sur la figure 6.3-c, nous avons reporté les $m = 5$ isophotes correspondant au point (u_2, v_2) du vase repéré par un cercle jaune sur la figure 6.3-a. Comme ce point se situe dans une ombre propre sous l’éclairage $\mathbf{s}^2$, l’isophote correspondant à $\mathbf{s}^2$ (de couleur verte) a une aire non nulle, mais les isophotes ont encore une intersection unique.

L’intersection d’isophotes permet enfin d’illustrer différentes situations connues pour être ambiguës. Le cas $m = 1$ (cf. figure 6.4-a) illustre l’ambiguïté fondamentale du *shape-from-shading*. Le cas $m = 2$ (cf. figure 6.4-b) montre qu’on passe d’une infinité de solutions à deux solutions (cf. chapitre 4), c’est-à-dire que l’ajout d’un deuxième éclairage suffit à rendre le problème « presque bien posé » [154, 172, 192, 218]. Enfin, le cas où l’on dispose de $m \geq 3$ éclairages coplanaires (cf. figure 6.4-c) reste mal posé (cf. paragraphe 4.4.3), ce qui s’illustre par le fait que les $m \geq 3$ isophotes n’ont pas une intersection unique.

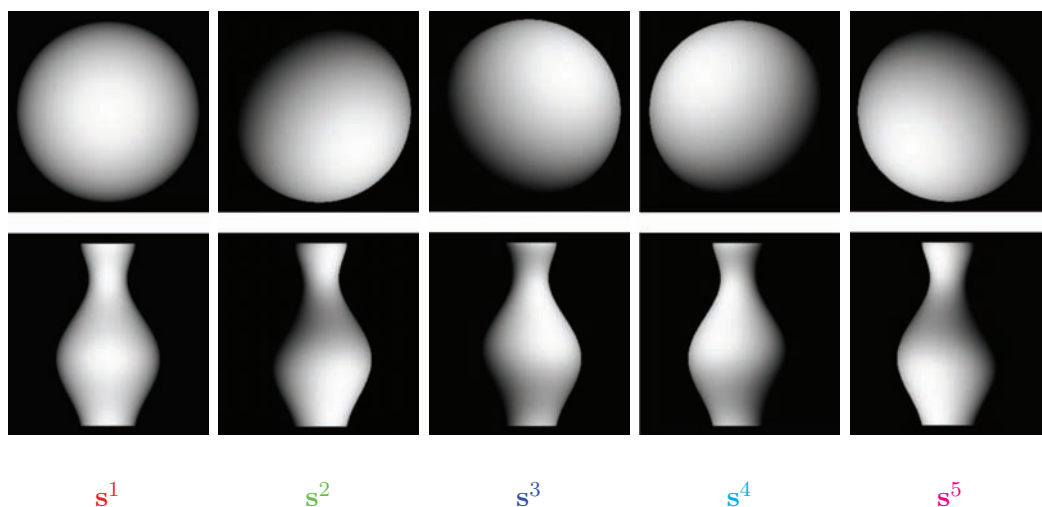


FIGURE 6.2 – Images de synthèse (de taille 256×256) d'une sphère et d'un vase, vus sous les mêmes $m = 5$ éclairages. Les deux objets ayant la même réflectance, en l'occurrence lambertienne, le vase peut être reconstruit par « apprentissage » sur la sphère, qui constitue l'objet de référence.

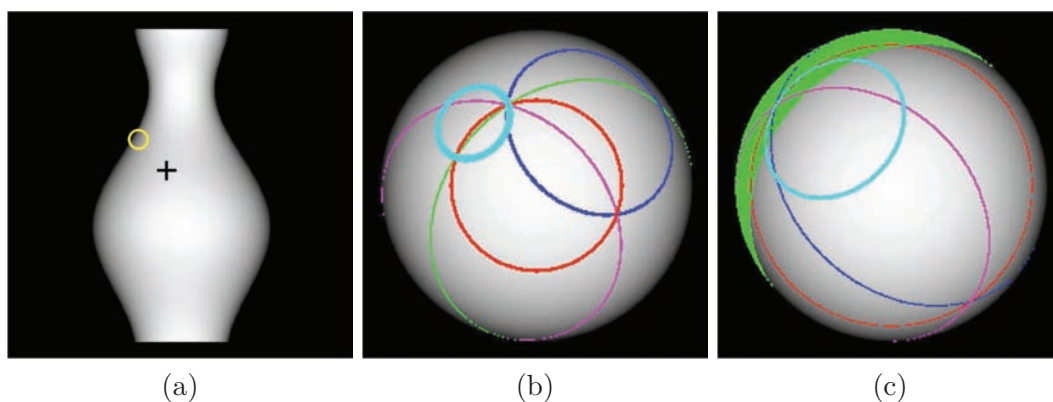


FIGURE 6.3 – (a) Le point (u_1, v_1) du vase repéré par une croix noire est éclairé dans les cinq images, alors que le point (u_2, v_2) repéré par un cercle jaune est dans l'ombre sous l'éclairage s^2 , c'est-à-dire dans la deuxième image du bas de la figure 6.2. (b) Les cinq isophotes des images de la sphère correspondant aux cinq niveaux de gris du point (u_1, v_1) sont concourants en un point unique, ce qui nous donne un moyen d'estimer la normale en (u_1, v_1) . (c) Ce résultat est encore vrai pour le point (u_2, v_2) , bien que le deuxième niveau de gris de (u_2, v_2) soit nul, et que l'isophote correspondant (de couleur verte) soit une ombre propre.

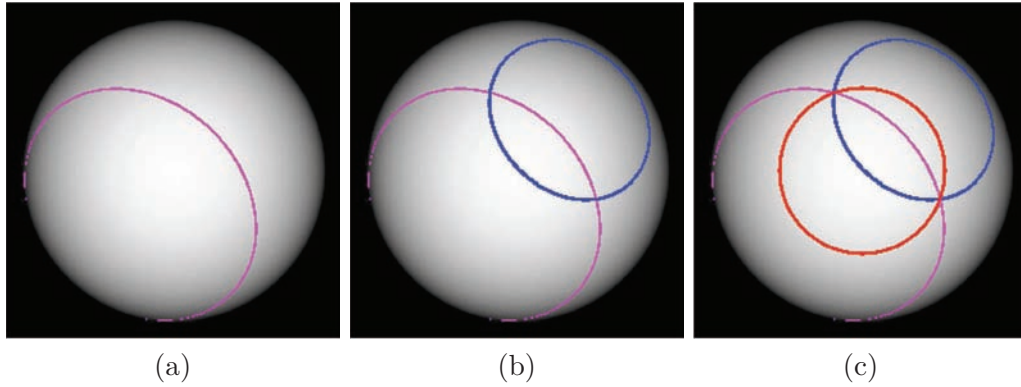


FIGURE 6.4 – (a) Le cas $m = 1$ illustre l’ambiguïté du *shape-from-shading*. (b) Si $m = 2$, il y a deux intersections entre les isophotes, donc deux normales possibles. (c) Avec $m = 3$ éclairages coplanaires, cette ambiguïté persiste.

Méthode de Hertzmann et Seitz

L’idée de résoudre la stéréophotométrie par apprentissage est due à Woodham, qui la suggère dès ses premiers travaux [268]. Néanmoins, sa mise en œuvre est récente [3, 121, 151]. En fait, l’estimation de la normale par intersection d’isophotes, telle que nous venons de la présenter, est avant tout pédagogique. Comment faire, en pratique, pour trouver de telles intersections ? La solution très intuitive proposée par Hertzmann et Seitz dans [122, 121] consiste à rechercher, pour chaque point (u, v) de la *surface de test* (ici, le vase) caractérisé par m niveaux de gris formant un *vecteur de test* $\mathbf{i}_t \in \mathbb{R}^m$, le point de la *surface d’apprentissage* (ici, la sphère) dont le *vecteur d’apprentissage* $\mathbf{i}_a \in \mathbb{R}^m$ est le plus proche (au sens de la distance euclidienne dans $\mathbb{R}^m$), et à attribuer la normale correspondant à $\mathbf{i}_a$ au point (u, v) . Comme le nombre de points situés sur la surface d’apprentissage est fini, puisque ces points sont en fait des pixels, il est envisageable d’effectuer une recherche exhaustive du point le plus proche dans $\mathbb{R}^m$.

Cette approche par apprentissage se passe donc de modèle de réflectance *explicite*. En l’occurrence, c’est l’objet de référence qui sert de « modèle », au sens premier du terme : « ce qui est donné pour être reproduit ». Cela signifie que la façon dont le matériau réémet la lumière est lue sur les images de l’objet de référence. Nous pouvons donc parler de modèle de réflectance *appris*. L’idée d’apprendre la réflectance est la même que celle qui consiste à mesurer la réflectance à l’aide d’un instrument appelé *goniomètre*. En mesurant la luminance émise par un matériau donné, et en échantillonnant la direction d’éclairage incident et la direction d’observation, on peut effectivement « apprendre » sa réflectance. Dans l’approche qui nous intéresse, la procédure est considérablement simplifiée : on n’utilise que m éclairages, et sous chacun de ces éclairages, le nombre d’échantillons est égal au nombre de pixels situés sur l’objet de référence. Nous ne pouvons donc acquérir qu’une connaissance très partielle de la réflectance. Néanmoins, cette connaissance est suffisante pour la reconstruction 3D par stéréophotométrie.

Chaque m -uplet de niveaux de gris de l'objet de test doit être comparé à chacun des m -uplets de l'objet de référence, ce qui est en pratique extrêmement coûteux. Hertzmann et Seitz utilisent comme accélération l'algorithme rapide ANN (*approximate nearest neighbour*). Dans le cas de $m = 8$ images d'apprentissage de taille 332×328 et de $m = 8$ images de test de taille 1176×398 (cf. figure 6.5), le temps de calcul sur un processeur 2 GHz Intel Xeon passe ainsi d'un jour, pour une recherche exhaustive, à environ 5 minutes. Dans [230], l'utilisation de la distance euclidienne dans un espace de dimension m potentiellement élevée est remise en question. Une distance géodésique sur la nappe passant par les vecteurs $\mathbf{i}_a$ lui est préférée. Si cette nuance est probablement valide lorsque le nombre d'images est très élevé, comme dans [230] et [151] où toutes les images d'une séquence vidéo sont utilisées, la méthode de Hertzmann et Seitz semble néanmoins valide, au vu des tests, pour un nombre d'images relativement faible.



FIGURE 6.5 – Images réelles d'une sphère (de taille 332×328) et d'une bouteille (de taille 1176×398) vues sous les mêmes $m = 8$ éclairages, provenant de <http://grail.cs.washington.edu/projects/sam/>.

6.1.2 Extensions de l'approche par apprentissage

Accélération de la mise en correspondance par ACP

Notre première contribution consiste à réduire la dimension m pour accélérer la recherche du plus proche voisin, qui constitue l'étape cruciale de la méthode de Hertzmann et Seitz. Nous effectuons donc l'analyse en composantes principales (ACP) des vecteurs d'apprentissage $\mathbf{i}_a$ et des vecteurs de test $\mathbf{i}_t$, afin de ne garder que les $p < m$ premières composantes principales (on suppose dorénavant que $m \geq 3$), qui forment des « vecteurs réduits » $\bar{\mathbf{i}}_a$ et $\bar{\mathbf{i}}_t$. Il est difficile de prédire si cette réduction de dimension est appropriée :

on sait que les données d'apprentissage doivent se répartir dans $\mathbb{R}^m$ sur une « nappe », puisque la normale (normée) possède deux degrés de liberté, mais la forme de cette nappe dépend beaucoup du matériau. La validation par les tests nous dira si la démarche était justifiée ou non. Un premier test a été mené sur les images de synthèse de la figure 6.2. La table 6.1 donne le temps CPU (en secondes) et la RMSE (en pixels) des reconstructions 3D obtenues par la méthode de Hertzmann et Seitz sans ACP, ou après réduction de la dimension aux $p = 4, 3, 2$ ou 1 premières composantes principales. Les temps CPU indiqués ne prennent pas en compte l'apprentissage sur les $m = 5$ images de la sphère.

Méthode	CPU	RMSE
Hertzmann et Seitz	0,66	1,58
ACP ($p = 4$)	0,57	1,58
ACP ($p = 3$)	0,47	1,59
ACP ($p = 2$)	0,37	4,98
ACP ($p = 1$)	0,28	6,92

TABLE 6.1 – Reconstructions 3D du vase en appliquant l'approche par apprentissage de la stéréophotométrie aux images de synthèse de la figure 6.2 : temps CPU (en secondes) et RMSE (en pixels) sans ACP (méthode de Hertzmann et Seitz) ou avec ACP, en utilisant les $p = 4, p = 3, p = 2$ ou $p = 1$ premières composantes principales.

Ces résultats montrent que la réduction de dimension de l'espace des configurations ne dégrade pas la précision du résultat tant que $p \geq 3$. Pour $p < 3$, la situation est similaire à l'une de celles de la figure 6.4 : l'estimation de la normale devient mal posée, ce qui se traduit par des valeurs très élevées de la RMSE. En revanche, la réduction de dimension permet bien de réduire le temps CPU de façon significative.

Un deuxième test a été mené sur les images réelles de la figure 6.5. Une sphère et une bouteille recouvertes de la même peinture ont été photographiées sous les mêmes $m = 8$ éclairages. La peinture utilisée possède une composante spéculaire non négligeable. Les images comportent des ombres propres, mais quasiment aucune ombre portée. La table 6.2 confirme les enseignements tirés du premier test. Les valeurs de la RMSE affichées dans cette table ne reflètent pas l'écart au relief réel de la bouteille (que nous ne connaissons pas), mais l'écart au résultat obtenu par la méthode de Hertzmann et Seitz, censée être la plus précise. Les surfaces reconstruites en intégrant les normales estimées sont montrées sur la figure 6.6. Ces résultats confirment l'intérêt de la stéréophotométrie par apprentissage, pour une surface qui est loin d'être lambertienne. Les résultats de la figure 6.6 sont qualitativement satisfaisants². En particulier, certains détails très fins de la surface semblent correctement retrouvés.

Nous disposons donc d'une méthode de résolution de la stéréophotométrie quasiment aussi précise que la méthode décrite dans [122, 121], mais plus rapide : d'après la table 6.2, la méthode de Hertzmann et Seitz requiert 608 secondes (cf. figure 6.6-a), contre seule-

2. À l'exception toutefois du cas $p = 2$ (cf. figure 6.6-d), dont nous savions par avance qu'il ne pourrait pas procurer de bons résultats.

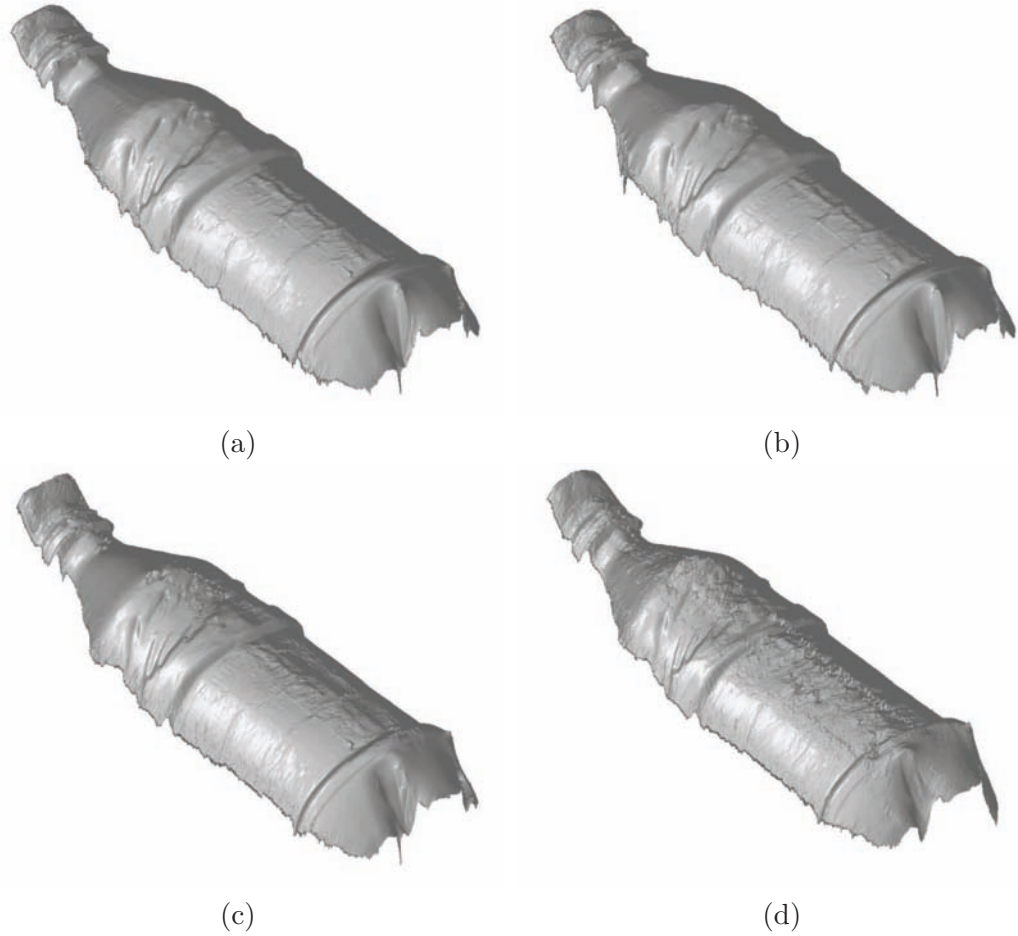


FIGURE 6.6 – Reconstructions 3D de la bouteille en appliquant l’approche par apprentissage de la stéréophotométrie aux images de la figure 6.5 : (a) sans ACP (méthode de Hertzmann et Seitz) ou avec ACP, en utilisant les (b) $p = 7$, (c) $p = 3$ ou (d) $p = 2$ premières composantes principales.

ment 238 secondes en utilisant les $p = 3$ premières composantes principales de l’ACP (cf. figure 6.6-c). Deux obstacles limitent néanmoins le potentiel de cette méthode. Primo, les ombres portées ne sont pas prises en compte. Plus généralement, l’estimation de la normale se dégrade dès lors qu’un niveau de gris est erroné. Secundo, et cet obstacle est le plus limitant, on ne peut pas envisager de systématiquement peindre les objets à reconstruire !

Méthode	CPU	RMSE
Hertzmann et Seitz	608	0
ACP ($p = 7$)	518	0,116
ACP ($p = 6$)	470	0,150
ACP ($p = 5$)	392	0,164
ACP ($p = 4$)	312	0,214
ACP ($p = 3$)	238	0,265
ACP ($p = 2$)	191	0,400
ACP ($p = 1$)	170	0,911

TABLE 6.2 – Reconstructions 3D de la bouteille en appliquant l’approche par apprentissage de la stéréophotométrie aux images de la figure 6.5 : temps CPU (en secondes) et RMSE (en pixels) sans ACP (méthode de Hertzmann et Seitz) ou avec ACP, en utilisant les $p = 7, \dots, p = 1$ premières composantes principales. Le résultat de la méthode de Hertzmann et Seitz constitue la surface de référence.

Prise en compte des ombres portées et de la texture

Reprenons l’exemple de la figure 6.2 pour illustrer l’effet des ombres portées. Supposons que le point (u_1, v_1) du vase, repéré par une croix noire sur la figure 6.3-a, se trouve dans une ombre portée sous l’éclairage $\mathbf{s}^4$ (cela n’est pas le cas sur la figure 6.2). L’isophote (de couleur cyan) correspondant à ce niveau de gris devient erroné, et l’intersection des isophotes devient vide (cf. figure 6.7-a).

D’autres difficultés peuvent survenir si l’on ne peut pas peindre la surface de l’objet à reconstruire. Nous n’aborderons pas ce problème difficile dans le cas général, et renvoyons le lecteur intéressé à [122], où la méthode proposée utilise une base de données d’objets d’apprentissage, plutôt qu’un seul objet. Nous nous restreignons au cas où les deux objets ont la même réflectance, mais où l’objet de test est « texturé », contrairement à l’objet d’apprentissage. Nous entendons ici par « texture » un coefficient multiplicatif pouvant varier d’un point à l’autre de la surface. Par analogie avec le cas lambertien, où ce coefficient est l’albédo, nous l’appelons abusivement « albédo » et le notons ρ (nous reviendrons sur ce point dans le paragraphe 6.2.1). Par exemple, si le point (u_1, v_1) du vase a un « albédo » égal à 0,8, alors ses $m = 5$ niveaux de gris sont assombris d’un même facteur 0,8, et les cinq isophotes sont simultanément erronés (cf. figure 6.7-b), puisque la surface d’apprentissage est supposée sans texture, i.e. $\rho \equiv 1$.

Ces deux situations (ombre portée et texture) présentent plusieurs points communs. D’une part, un certain nombre de niveaux de gris ont des valeurs différentes de leurs valeurs théoriques. D’autre part, ces niveaux de gris erronés rendent erronés les isophotes correspondants, en conséquence de quoi les isophotes ne sont généralement plus concourants. On peut quand même distinguer ces deux cas de figure, et apporter à chacun un remède approprié. En présence d’ombres portées, on sait que les niveaux de gris erronés ont des valeurs faibles (mais pas forcément nulles, à cause de l’effet de pénombre). Il suffit

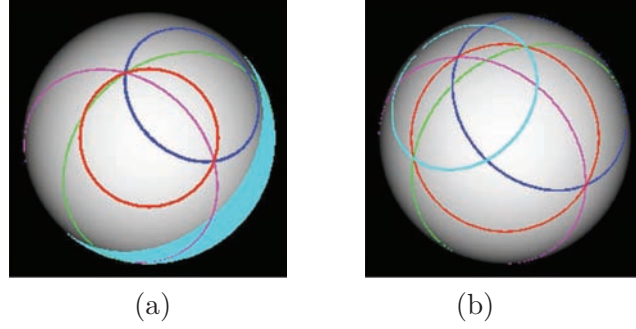


FIGURE 6.7 – (a) Si le point (u_1, v_1) de la figure 6.3-a se trouve dans une ombre portée sous l'éclairage $\mathbf{s}^4$, l'isophote correspondant (en cyan) est erroné, et l'intersection des isophotes est vide. (b) Si l'albédo de (u_1, v_1) est égal à 0,8, tous les isophotes sont erronés, et seuls les isophotes des éclairages coplanaires se croisent encore (cette situation correspond au problème PS2 ρ , cf. paragraphe 4.1.4).

donc de retirer les isophotes incriminés pour que les isophotes restants soient à nouveau concourants. En présence d'une texture, tous les niveaux de gris sont simultanément erronés, donc tous les isophotes aussi, mais en considérant ρ comme une nouvelle inconnue du problème, on peut y remédier : il suffit en effet de chercher la valeur de ρ telle que les isophotes correspondant aux m niveaux de gris divisés par ρ soient concourants. Il est bien sûr impossible de procéder ainsi en pratique.

Nous proposons de tirer à nouveau parti de l'ACP, en ne nous contentant plus de rechercher, dans l'espace de dimension réduite $\mathbb{R}^p$, le vecteur d'apprentissage réduit $\bar{\mathbf{i}}_a$ le plus proche d'un vecteur de test réduit $\bar{\mathbf{i}}_t$. La distance minimale entre ces deux vecteurs réduits étant caractéristique de la fiabilité des niveaux de gris, nous l'utilisons pour détecter les ombres portées (ou tout autre écart au modèle appris). Nous appliquons donc un traitement spécifique aux pixels de l'image de test pour lesquels cette distance minimale est supérieure à un seuil.

En guise de premier test, nous utilisons les images de synthèse d'une sphère et d'une théière de même réflectance, vues sous les mêmes $m = 8$ éclairages (cf. figure 6.8)³.

En utilisant les $p = 5$ premières composantes principales de l'ACP, on peut effectivement calculer, pour chaque vecteur de test réduit $\bar{\mathbf{i}}_t$ de la théière, la distance minimale à un vecteur d'apprentissage réduit $\bar{\mathbf{i}}_a$ (cf. figure 6.9-a). Cette distance constitue une mesure de la fréquence d'apparition d'une ombre portée. Si l'on applique la méthode d'estimation des normales par recherche du plus proche voisin dans $\mathbb{R}^5$, sans tenir compte de cette mesure, on obtient un champ de normales erroné dans les zones sujettes aux ombres portées, comme par exemple à proximité de l'anse de la théière (cf. figure 6.10-a). En utilisant cette distance pour détecter et éliminer les ombres portées, avant d'appliquer la reconstruction par apprentissage aux niveaux de gris conservés, nous parvenons à améliorer la cohérence du champ de normales près de la discontinuité (cf. figure 6.10-b).

3. Ces images nous ont été gracieusement fournies par la société FittingBox.

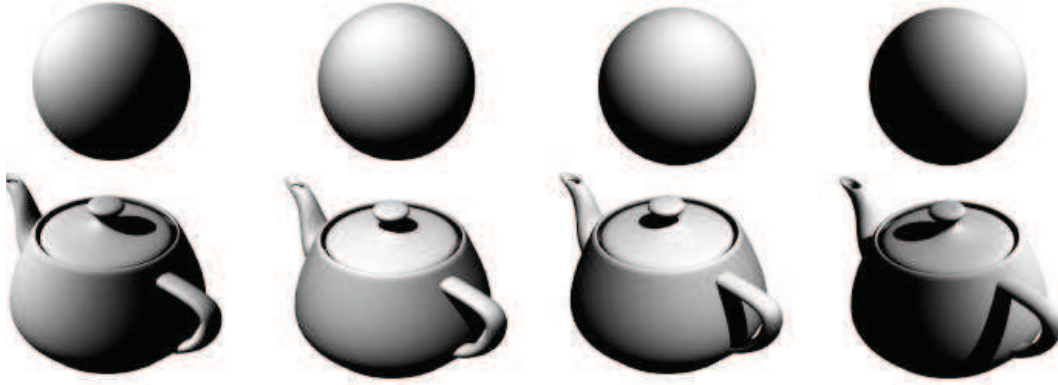


FIGURE 6.8 – Quatre images de synthèse (parmi $m = 8$) d'une sphère et d'une théière vues sous les mêmes éclairages. Les images de la théière comportent des ombres propres et des ombres portées.

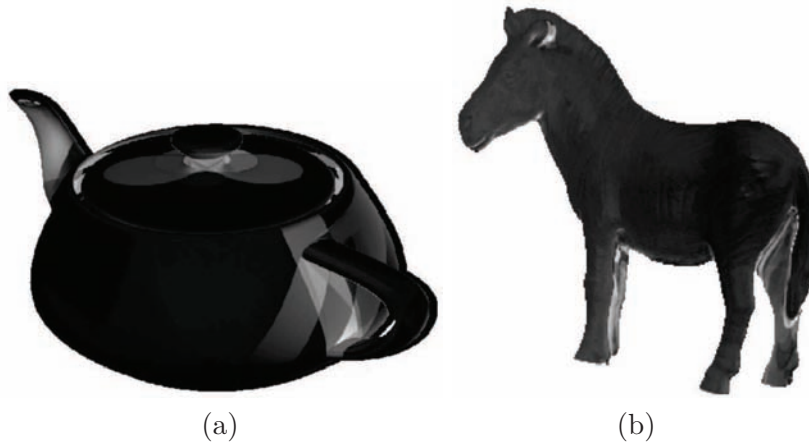
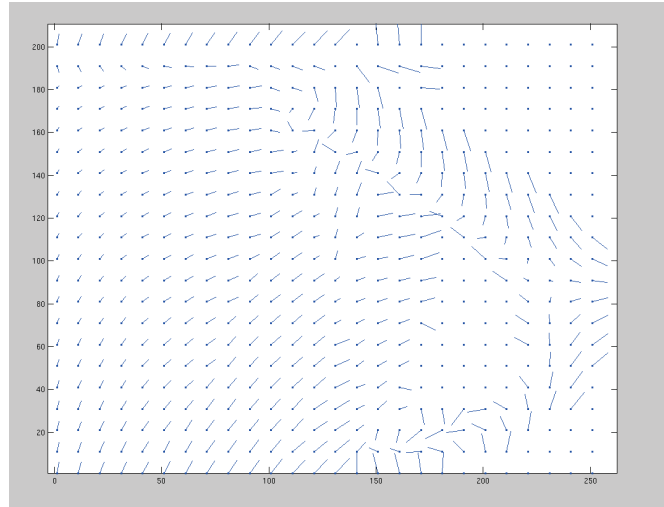
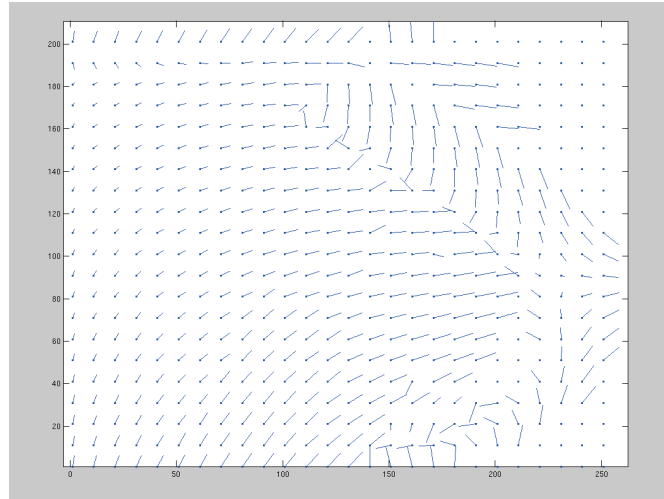


FIGURE 6.9 – Fréquence d'apparition d'une ombre portée (affichée en niveaux de gris) calculée en utilisant les $p = 5$ premières composantes principales de l'ACP, pour les images : (a) de la figure 6.8 ; (b) de la figure 6.11.

Effectuons un deuxième test sur les images réelles de la figure 6.11. Une sphère et un cheval miniature, recouverts de la même peinture, sont vus sous les mêmes $m = 12$ éclairages. En utilisant les $p = 5$ premières composantes principales de l'ACP, nous mesurons la fréquence d'apparition d'une ombre portée (cf. figure 6.9-b). La figure 6.12 montre les reconstructions 3D obtenues en utilisant ou non cette fréquence d'apparition. L'amélioration du résultat dans une zone sujette aux ombres portées se voit très nettement, par exemple au niveau des pattes avant du cheval.



(a)



(b)

FIGURE 6.10 – Estimations du champ de normales au voisinage de l'anse de la théière, en appliquant l'approche par apprentissage de la stéréophotométrie aux images de la figure 6.8, en utilisant les $p = 5$ premières composantes principales de l'ACP : (a) approche par apprentissage classique ; (b) approche par apprentissage avec détection des ombres portées (cf. figure 6.9-a). L'estimation (a) du champ de normales est biaisée à proximité de l'anse, ce qui se traduit par une discontinuité du champ estimé dans une zone où la surface est pourtant régulière. En revanche, le champ de normales estimé (b) est continu, sauf bien sûr au niveau de la discontinuité de profondeur.

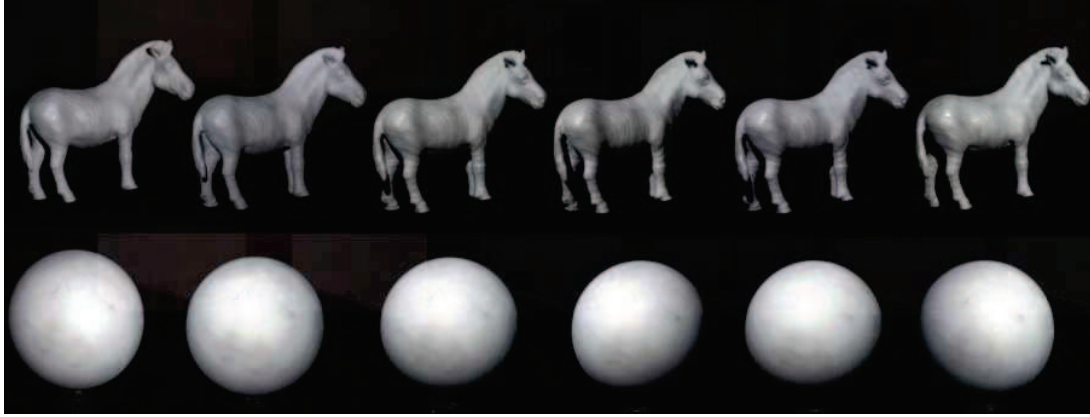


FIGURE 6.11 – Six images réelles (parmi $m = 12$) d'une sphère et d'un cheval miniature présentant une réflectance diffuse non lambertienne, provenant de <http://www.cs.washington.edu/education/courses/csep576/05wi/projects/project3/project3.htm>. Les images du cheval comportent des ombres propres et des ombres portées (près des pattes).

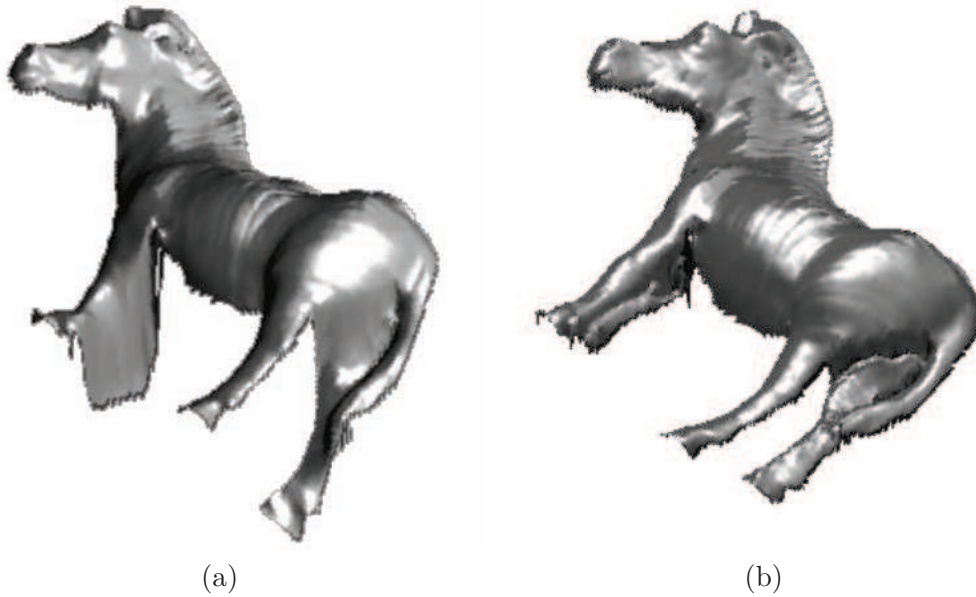


FIGURE 6.12 – Reconstructions 3D du cheval en appliquant l'approche par apprentissage de la stéréophotométrie aux images de la figure 6.11, en utilisant les $p = 5$ premières composantes principales de l'ACP : (a) approche par apprentissage classique ; (b) approche par apprentissage avec détection des ombres portées (cf. figure 6.9-b).

Effectuons un dernier test sur les images de synthèse d'une sphère et d'une théière vues sous les mêmes $m = 5$ éclairages (cf. figure 6.13). La théière est revêtue d'une texture de type « mur de briques ». Elle n'a donc pas un « albédo » uniforme. Bien que les théières soient rarement en briques, cet exemple donne une idée d'une version « non intrusive » de notre méthode, dans laquelle la surface à reconstruire est photographiée telle quelle. Pour chaque point de la théière de vecteur de test réduit $\bar{\mathbf{i}}_t$, on cherche la valeur de ρ telle que $\bar{\mathbf{i}}_t/\rho$ se trouve « au plus près » d'un vecteur d'apprentissage réduit $\bar{\mathbf{i}}_a$. La figure 6.14 montre la carte des « albédos » et le relief estimés ainsi.

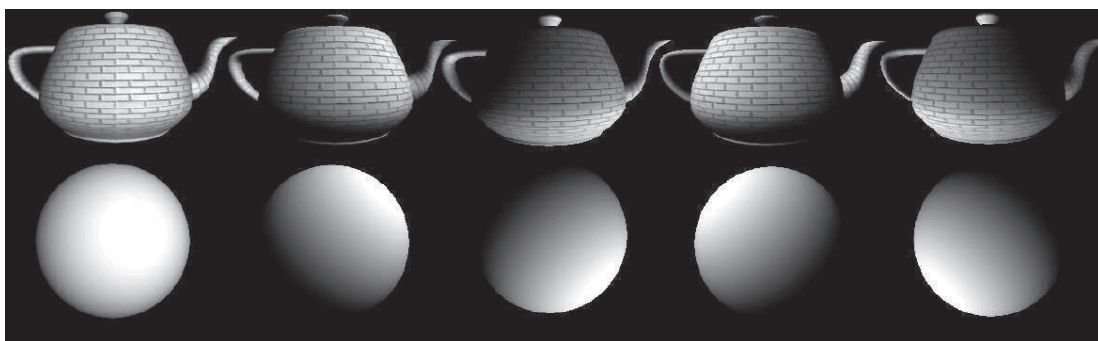


FIGURE 6.13 – Cinq images de synthèse d'une sphère et d'une théière texturée, vues sous les mêmes éclairages.

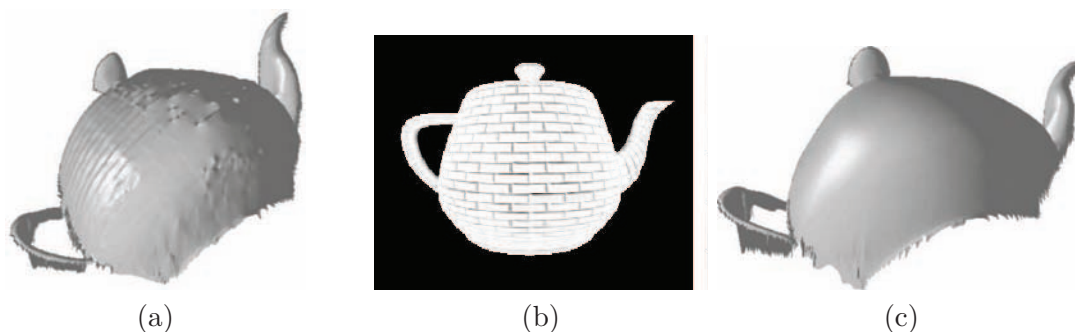


FIGURE 6.14 – (a) Relief obtenu en appliquant la méthode de Hertzmann et Seitz aux images de la figure 6.13. Comme cette méthode ne tient pas compte d'une éventuelle texture, les lignes sombres à la surface de la théière sont interprétées, non pas comme une texture, mais comme un réseau de fines stries (ce phénomène est en fait lié à l'ambiguïté de trompe-l'œil, que nous avons déjà évoquée). (b) Albédo et (c) relief obtenus avec l'extension de la méthode de Hertzmann et Seitz que nous proposons, en utilisant les $p = 3$ premières composantes principales de l'ACP.

6.1.3 Conclusions sur l'approche par apprentissage

Dans ce paragraphe, nous avons repris une idée de Woodham, mise en œuvre par Hertzmann et Seitz [122, 121], qui consiste à résoudre la stéréophotométrie par apprentissage. Cela permet de se libérer d'un certain nombre de contraintes imposées par la résolution classique du problème. En particulier, il n'est plus nécessaire de connaître un modèle de réflectance *explicite* de la surface à reconstruire, ni même d'estimer les paramètres d'un modèle, puisque le modèle est appris sur les images d'un objet de relief connu. Un autre avantage est que les vecteurs d'éclairage ne doivent pas être estimés. Enfin, les ombres propres ne constituent pas des données aberrantes. De même que leur prise en compte améliore sensiblement les résultats de l'approche classique (cf. paragraphe 1.3.5), les ombres propres fournissent à l'approche par apprentissage une information pertinente sur l'orientation locale de la surface.

Nous avons montré que l'analyse en composantes principales des vecteurs d'apprentissage et des vecteurs de test permettait d'accélérer la méthode de Hertzmann et Seitz sans perdre en précision, et avons proposé deux généralisations de cette méthode. D'une part, l'utilisation de la distance au vecteur d'apprentissage réduit le plus proche permet de détecter les ombres portées et d'exclure les niveaux de gris concernés. D'autre part, si l'objet de test est « texturé », la minimisation de cette distance minimale permet d'estimer son « albédo », si celui-ci est considéré comme une nouvelle inconnue.

Ce travail ouvre autant de perspectives qu'il ne résout de problèmes. Primo, nous n'avons pas traité le cas d'une surface texturée avec ombres portées. Comme ces deux types d'écarts au modèle appris sont détectés dans le même espace de dimension réduite $\mathbb{R}^p$, il ne sera pas forcément facile de les distinguer. Secundo, il arrive souvent que les ombres portées se concentrent près des discontinuités. Dans l'exemple de la figure 6.8, certains points situés sous l'anse de la théière ne sont éclairés que dans deux images, voire dans une seule (cf. figure 6.9-a). Il nous faudrait proposer une estimation de la normale dans de tels cas de figure, en utilisant par exemple l'approche décrite dans le chapitre 4. Enfin, nous nous sommes restreints dans ce travail à des images en niveaux de gris, contrairement au travail de Hertzmann et Seitz [122, 121] qui tire parti de la couleur : la prise en compte de la couleur ne représente pas une difficulté insurmontable, car il suffit d'augmenter la dimension de l'espace de recherche. Pour des images RGB, nous pourrions ainsi utiliser les $3m$ niveaux de couleur de chaque pixel de l'objet de test. À ce propos, nous avons vu dans le paragraphe 1.5 qu'en utilisant un modèle de réflectance explicite, en l'occurrence le modèle lambertien, nous ne savions faire profiter de l'information de couleur à la stéréophotométrie que si les éclairages et la surface n'étaient pas simultanément colorés. Or, cela ne pose pas de problème particulier pour l'approche par apprentissage : il s'agit là d'un point fort supplémentaire de cette approche.

Toutefois, la principale limitation de l'approche par apprentissage est de devoir disposer d'un objet de référence. Pour éviter cette contrainte, la seule possibilité est d'utiliser un modèle de réflectance explicite, comme nous l'avons déjà fait avec le modèle lambertien. Nous allons donc montrer, dans le prochain paragraphe, comment étendre la stéréophotométrie à des modèles de réflectance plus réalistes que le modèle lambertien.

6.2 Réflectance diffuse non lambertienne

Outre bien sûr le modèle lambertien, le modèle explicite de réflectance diffuse le plus utilisé est sans doute celui d'Oren-Nayar [194]. Nous avons vu dans le paragraphe 1.1.1 que, sous l'hypothèse lambertienne, la luminance émise par la surface pouvait être modélisée sous la forme suivante (cf. équation (1.12)) :

$$L(\mathbf{x}(u, v)) = \frac{\rho(\mathbf{x}(u, v))}{\pi} E(\mathbf{x}(u, v)) \quad (6.1)$$

où $E(\mathbf{x}(u, v))$ désigne l'éclairement reçu par la surface au point $\mathbf{x}(u, v)$. Intéressons-nous au modèle d'Oren-Nayar, qui constitue une généralisation du modèle lambertien.

6.2.1 Modèle d'Oren-Nayar

Le modèle de réflectance diffuse d'Oren-Nayar permet de prendre en compte les « rétro-réflexions » [194, 195]. La « rugosité » est une nouvelle caractéristique de la surface, qui découle de sa modélisation par un ensemble de microfacettes supposées lambertiennes. Oren et Nayar montrent dans [194] que, sous cette hypothèse, la luminance émise par la surface au point $\mathbf{x}(u, v)$, dans la direction $\mathbf{v}(u, v)$, s'écrit :

$$L(\mathbf{x}(u, v), \mathbf{v}(u, v)) = \underbrace{\frac{\rho(\mathbf{x}(u, v))}{\pi} \chi_{\text{ON}}(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))}_{\kappa_{\text{ON}}(\mathbf{x}(u, v), \bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))} E(\mathbf{x}(u, v)) \quad (6.2)$$

Dans ce modèle, qui généralise le modèle lambertien (6.1), κ_{ON} constitue la réflectance. Le facteur correctif χ_{ON} s'écrit :

$$\chi_{\text{ON}}(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) = 1 - \frac{\sigma^2}{2(\sigma^2 + 0,57)} + \frac{0,45 \sigma^2 \max \{0, \gamma(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))\}}{\sigma^2 + 0,09} \times C(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) \quad (6.3)$$

où :

$$\gamma(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) = [\bar{\mathbf{s}} - (\bar{\mathbf{s}} \cdot \mathbf{n}(u, v)) \mathbf{n}(u, v)] \cdot [\mathbf{v}(u, v) - (\mathbf{v}(u, v) \cdot \mathbf{n}(u, v)) \mathbf{n}(u, v)] \quad (6.4)$$

$$C(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) = \sin \left(\max \left\{ \cos^{-1}(\bar{\mathbf{s}} \cdot \mathbf{n}(u, v)), \cos^{-1}(\mathbf{v}(u, v) \cdot \mathbf{n}(u, v)) \right\} \right) \times \tan \left(\min \left\{ \cos^{-1}(\bar{\mathbf{s}} \cdot \mathbf{n}(u, v)), \cos^{-1}(\mathbf{v}(u, v) \cdot \mathbf{n}(u, v)) \right\} \right) \quad (6.5)$$

et où σ représente l'écart-type de l'orientation des facettes, i.e. la *rugosité* de la surface. Lorsque $\sigma = 0$, l'expression (6.2) de la luminance se ramène effectivement à (6.1), puisque $\chi_{\text{ON}} \equiv 1$.

Nous avons vu dans le paragraphe 1.1.1 que le niveau de gris au point (u, v) était proportionnel à la luminance émise par le point $\mathbf{x}(u, v)$. Puisque, par définition, l'éclairement reçu s'écrit $E(\mathbf{x}(u, v)) = \mathbf{s} \cdot \mathbf{n}(u, v)$ (cf. (1.13)), on obtient finalement, par analogie

avec l'équation (1.16), l'équation :

$$I(u, v) = \rho(u, v) \max\{0, \chi_{\text{ON}}(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) \mathbf{s} \cdot \mathbf{n}(u, v)\} \quad (6.6)$$

La figure 6.15 montre plusieurs exemples d'images de synthèse calculées en utilisant (6.6) : le niveau de gris à proximité des contours d'occultation, i.e. lorsque le produit scalaire $\mathbf{n}(u, v) \cdot \mathbf{v}(u, v) \rightarrow 0$, est supérieur à celui que prédit le modèle lambertien. Ce phénomène, qui provient des *rétro-réflexions* sur les microfacettes, compense l'effet d'ombrage à proximité des terminateurs. Vu sous un éclairage frontal, le relief semble donc plus aplati que pour une surface lambertienne.

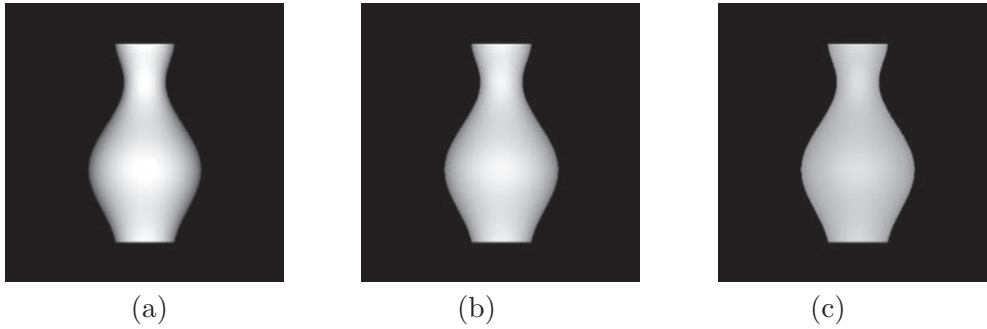


FIGURE 6.15 – Images de synthèse d'un vase texturé, calculées avec le modèle d'Oren-Nayar (6.6) : (a) $\sigma = 0$ (modèle lambertien) ; (b) $\sigma = 0,25$; (c) $\sigma = 0,5$. Plus la valeur de σ est élevée, plus le relief semble aplati.

La présence des fonctions trigonométriques $\sin$, $\tan$ et $\cos^{-1}$ dans le modèle (6.6), ainsi que celle des opérateurs $\min$ et $\max$, rend la résolution numérique de la stéréophotométrie particulièrement ardue. Remarquons toutefois que ce modèle a connu un certain succès dans le cadre du *shape-from-shading* [253]. Il a notamment été montré comment résoudre l'équation (6.6), en supposant l'albédo connu, au moyen de méthodes de type *fast-marching* [140, 141] ou semi-lagrangiennes [254].

Revenons momentanément sur une des extensions de la stéréophotométrie par apprentissage que nous avons proposée dans le paragraphe 6.1.2, lorsque l'objet de test, qui par hypothèse est de même réflectance que l'objet d'apprentissage, possède une « texture ». Nous avons défini cette texture comme un *coefficient multiplicatif* pouvant varier d'un point à l'autre de la surface. L'expression (6.6) du modèle d'Oren-Nayar justifie une telle définition, si l'on assimile $\rho(u, v)$ à ce coefficient. Remarquons toutefois que, si deux points $\mathbf{x}(u_1, v_1)$ et $\mathbf{x}(u_2, v_2)$ sont éclairés par le même éclairage directionnel $\mathbf{s}$ et possèdent la même normale $\mathbf{n}(u_1, v_1) = \mathbf{n}(u_2, v_2)$, le rapport $I(u_2, v_2)/I(u_1, v_1)$ des niveaux de gris prédits par le modèle (6.6) n'est pas strictement égal au rapport $\rho(u_2, v_2)/\rho(u_1, v_1)$ de leurs albédos, car les directions d'observation $\mathbf{v}(u_1, v_1)$ et $\mathbf{v}(u_2, v_2)$ ne sont pas nécessairement égales. Remarquons enfin qu'il est abusif d'appeler « albédo » le coefficient $\rho(u, v)$ de l'expression (6.6), car l'albédo est défini comme la proportion de lumière réémise par la surface, ce qui n'est pas le cas ici, dès lors que $\sigma > 0$.

6.2.2 Utilisation du modèle d'Oren-Nayar en stéréophotométrie

Résoudre la stéréophotométrie avec le modèle d'Oren-Nayar nécessite de résoudre, en tout point $(u, v) \in \Omega$, un système de m équations non linéaires :

$$I^i(u, v) = \max\{0, \chi_{\text{ON}}(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) \mathbf{s}^i \cdot \mathbf{m}(u, v)\}, \quad \forall i \in [1, m] \quad (6.7)$$

où nous utilisons à nouveau le vecteur $\mathbf{m}(u, v) = \rho(u, v)\mathbf{n}(u, v)$. Le modèle d'Oren-Nayar a été très peu étudié dans le contexte de la stéréophotométrie, mis à part pour des configurations d'éclairage particulières [171]. Pourtant, il n'est pas beaucoup plus difficile d'utiliser ce modèle que le modèle lambertien. En effet, si χ_{ON} est connu, alors l'estimation du vecteur $\mathbf{m}$ au sens des moindres carrés est une simple généralisation de l'estimation (1.24) du cas lambertien :

$$\mathbf{m}(u, v) = \mathbf{S}^+ \begin{bmatrix} \frac{I^1(u, v)}{\chi_{\text{ON}}(\bar{\mathbf{s}}^1, \mathbf{n}(u, v), \mathbf{v}(u, v))} \\ \vdots \\ \frac{I^m(u, v)}{\chi_{\text{ON}}(\bar{\mathbf{s}}^m, \mathbf{n}(u, v), \mathbf{v}(u, v))} \end{bmatrix} \quad (6.8)$$

Dans (6.8), le coefficient multiplicatif $\chi_{\text{ON}}(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))$ dépend de la normale, qui est inconnue. Pour faire face à cette dépendance, nous suggérons d'utiliser un schéma semi-implicite pour résoudre le système (6.7), i.e. de procéder itérativement en rendant explicite le coefficient multiplicatif $\chi_{\text{ON}}(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))$, qui comporte des non linéarités. À chaque itération, il nous faut donc résoudre un problème strictement similaire à celui de la stéréophotométrie lambertienne, en « compensant » les images par les rétro-réflexions estimées à l'itération précédente. Nous avons observé expérimentalement que ce schéma numérique très simple était convergent. Il présente de plus l'avantage de se ramener au cas lambertien, pour lequel nous disposons de méthodes de résolution robuste. En présence d'ombres portées ou d'autres types de données aberrantes, il suffit de remplacer l'estimation en moindres carrés par une méthode de régression robuste, en utilisant par exemple un estimateur parcimonieux (cf. paragraphe 1.3).

Si l'on s'en tient aux moindres carrés et que l'on linéarise le système (6.7), une itération de cet algorithme s'écrit de la façon suivante :

$$\mathbf{m}^{k+1}(u, v) = \mathbf{S}^+ \begin{bmatrix} \frac{I^1(u, v)}{\chi_{\text{ON}}(\bar{\mathbf{s}}^1, \mathbf{n}^k(u, v), \mathbf{v}(u, v))} \\ \vdots \\ \frac{I^m(u, v)}{\chi_{\text{ON}}(\bar{\mathbf{s}}^m, \mathbf{n}^k(u, v), \mathbf{v}(u, v))} \end{bmatrix} \quad (6.9)$$

en partant d'une solution initiale $\mathbf{m}^0(u, v)$ qui peut être, par exemple, la solution du problème lambertien, i.e. du système (6.7) lorsque $\chi_{\text{ON}} \equiv 1$. La figure 6.16 montre un exemple de résultat obtenu à partir de $m = 10$ images de synthèse, calculées avec le modèle d'Oren-Nayar lorsque $\sigma = 0, 5$.

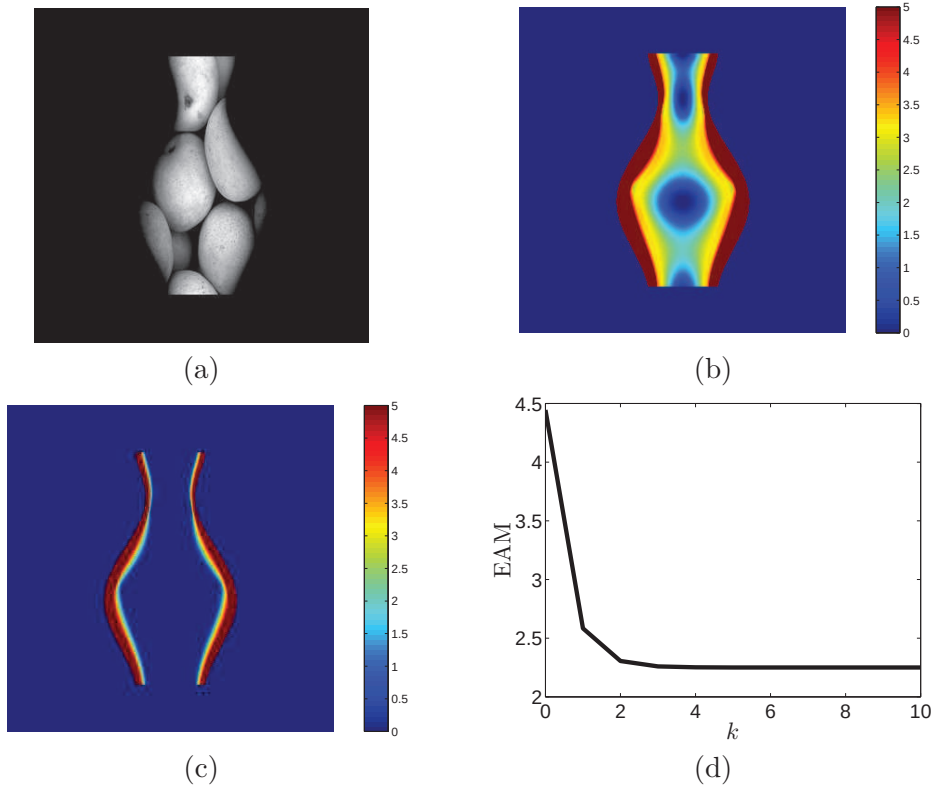


FIGURE 6.16 – Évaluation du schéma numérique visant à estimer simultanément les normales et l'albédo d'une surface, sur un jeu de $m = 10$ images de synthèse d'un vase texturé, calculées avec le modèle d'Oren-Nayar lorsque $\sigma = 0,5$. (a) Une des $m = 10$ images. (b) Écart angulaire (en degrés) entre le champ de normales réel et celui estimé en supposant le modèle lambertien. L'écart angulaire moyen vaut $\text{EAM} = 4,44^\circ$. (c) Idem, après $k = 10$ itérations du schéma numérique proposé. L'écart angulaire moyen vaut $\text{EAM} = 2,25^\circ$. (d) L'évolution de l'écart angulaire moyen en fonction du nombre d'itérations montre que ce schéma converge très rapidement.

6.2.3 Estimation conjointe du relief, de l'albédo et de la rugosité

La méthode de résolution proposée dans le paragraphe précédent comporte un point faible : le paramètre de rugosité σ doit être connu. On peut toutefois échantillonner les valeurs de σ , puis estimer les normales et l'albédo par le schéma (6.9) pour chaque échantillon, et enfin choisir l'échantillon qui minimise la RMSE entre les images reprojetées (avec le modèle (6.6)) et les images d'origine. Comme cela est montré sur la figure 6.17, l'échantillon qui minimise la RMSE est effectivement très proche de la valeur réelle de σ . On voit aussi, sur cette figure, que la minimisation de l'EAM permettrait d'estimer encore plus précisément la valeur de la rugosité, mais l'EAM ne peut être calculée que si l'on dispose de la vérité terrain, contrairement à la RMSE.

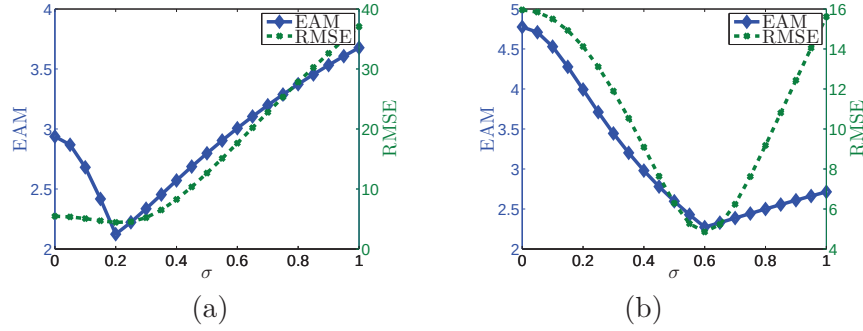


FIGURE 6.17 – Pour estimer la rugosité, il suffit d'utiliser le schéma itératif (6.9) pour différentes valeurs de σ , et de sélectionner l'estimation qui minimise la RMSE entre les images reprojetées et les images d'origine. On constate sur ces deux graphes que la RMSE est minimale pour une valeur : (a) très proche de la valeur réelle $\sigma = 0,2$; (b) égale à la valeur réelle $\sigma = 0,6$. Nous avons également porté sur ces graphes l'évolution de l'EAM entre les champs de normales estimé et réel, en fonction de σ , dont le minimum coïncide parfaitement avec la valeur réelle de σ . L'EAM peut donc sembler plus pertinente que la RMSE pour estimer la rugosité, mais en général, l'EAM n'est pas calculable car on ne dispose pas de la vérité terrain.

Pour conclure, nous avons montré qu'il était relativement aisé de prendre en compte les phénomènes de rétro-réflexions, qui constituent le fondement du modèle d'Oren-Nayar, en appliquant itérativement les méthodes de résolution décrites dans le chapitre 1.3. Il suffit, à chaque itération, de « compenser » les images par les rétro-réflexions estimées à l'itération précédente, puis d'estimer les normales et l'albédo en utilisant le modèle lambertien. Le modèle d'Oren-Nayar permet effectivement d'appliquer la stéréophotométrie à un grand nombre de surfaces mates. En revanche, ce modèle de réflectance diffuse ne permet pas de modéliser les reflets brillants.

6.3 Réflectance spéculaire

Les modèles de réflectance spéculaire ont connu un intérêt croissant au cours des dernières années : plutôt que de considérer les reflets brillants comme des données aberrantes, il peut sembler plus pertinent de chercher à les expliquer, à la manière de ce que nous avons proposé dans le paragraphe 1.3.5 pour les ombres propres. De tels phénomènes sont en effet très informatifs sur la géométrie de la scène. Dans le cas d'une surface purement spéculaire (e.g., un miroir), la lumière est réémise dans une seule direction selon la loi de Descartes pour la réflexion. En un point brillant, la normale est donc donnée par la bissectrice de l'angle formé par le vecteur d'éclairage $\mathbf{s}$ et la direction d'observation $\mathbf{v}$. Si la reconstruction 3D de surfaces spéculaires transparentes est un problème particulièrement ardu (voir notamment [85, 75, 176]), nous allons voir que la stéréophotométrie permet la reconstruction 3D de surfaces spéculaires opaques.

Le premier modèle de réflectance spéculaire utilisé dans le contexte de la reconstruction 3D photométrique est probablement celui proposé par Nayar et al. dans [187]. Ce modèle a été adapté à des objets présentant des reflets brillants très localisés (la fonction modélisant le terme spéculaire est alors une fonction de Dirac). Des reflets brillants de plus grande taille peuvent être décrits par le modèle de Torrance-Sparrow [252], comme cela est fait dans [98], ou par le modèle empirique de Blinn-Phong [33], comme cela a été récemment proposé dans [144, 174]. Pour faciliter l'utilisation du modèle de Torrance-Sparrow en stéréophotométrie, Georghiades a proposé dans [98] un certain nombre d'hypothèses simplificatrices plus ou moins justifiées. Le modèle de Blinn-Phong est plus facile à manipuler en stéréophotométrie que celui de Torrance-Sparrow, mais il s'agit d'un modèle empirique potentiellement moins réaliste, qui ne repose pas sur une modélisation microscopique de la surface par des microfacettes.

6.3.1 Modèle de Ward

Le modèle de Ward [260] constitue un bon compromis entre réalisme et simplicité. Ce modèle a été utilisé dans [59] pour reconstruire par stéréophotométrie le relief de surfaces présentant des reflets brillants de grande taille. Il ne s'agit ni d'un modèle physique, ni d'un modèle empirique, mais d'un modèle « expérimental », i.e. déterminé par des mesures réalisées sur des surfaces. Comme le modèle d'Oren-Nayar, le modèle de Ward modélise les surfaces par un ensemble de microfacettes dont les orientations suivent une distribution gaussienne d'écart-type σ , mais les microfacettes sont supposées réfléchissantes, et non pas lambertiennes. La réflectance de ce modèle s'écrit :

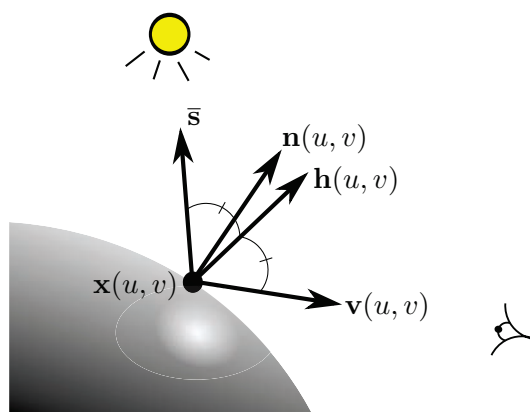
$$\kappa_W(\mathbf{x}(u, v), \bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) = \frac{\rho(\mathbf{x}(u, v))}{\pi} + \frac{\rho_s(\mathbf{x}(u, v))}{4\pi\sigma^2} \frac{\exp\left\{-\frac{\tan^2 \zeta}{\sigma^2}\right\}}{\sqrt{\bar{\mathbf{s}} \cdot \mathbf{n}(u, v)} \sqrt{\mathbf{v}(u, v) \cdot \mathbf{n}(u, v)}} \quad (6.10)$$

où ζ désigne l'angle entre la normale $\mathbf{n}(u, v)$ et le vecteur unitaire $\mathbf{h}(u, v)$ (cf. figure 6.18), qui désigne la bissectrice de l'angle entre $\bar{\mathbf{s}}$ et $\mathbf{v}(u, v)$ ($\mathbf{h}$ joue un rôle similaire au vecteur $\mathbf{r}$ du modèle de Phong, cf. figure 1.4) :

$$\mathbf{h}(u, v) = \frac{1}{\|\bar{\mathbf{s}} + \mathbf{v}(u, v)\|} [\bar{\mathbf{s}} + \mathbf{v}(u, v)] \quad (6.11)$$

Dans (6.10), les paramètres ρ_s et σ désignent, respectivement, l'*albédo spéculaire* et la *rugosité* du matériau (ce dernier paramètre contrôle la taille des reflets brillants). Lorsque l'éclairage est directionnel, la luminance reçue $\bar{L}$ est une fonction de Dirac, donc l'expression générale (5.1) de la luminance émise est égale au produit de la réflectance κ par l'éclairement $\mathbf{s} \cdot \mathbf{n}$. Cela permet à Chung et al., dans [59], d'écrire le niveau de gris sous la forme suivante, pour une surface dont la réflectance suit le modèle de Ward :

$$I(u, v) = \rho(u, v) \max\{0, \mathbf{s} \cdot \mathbf{n}(u, v)\} + \rho_s(u, v) \max\{0, \mathbf{s} \cdot \mathbf{n}(u, v)\} \underbrace{\frac{\exp\left\{-\frac{\tan^2(\cos^{-1}(\mathbf{h}(u, v) \cdot \mathbf{n}(u, v)))}{\sigma^2}\right\}}{\sigma^2 \sqrt{(\bar{\mathbf{s}} \cdot \mathbf{n}(u, v)) (\mathbf{v}(u, v) \cdot \mathbf{n}(u, v))}}}_{\chi_W(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))} \quad (6.12)$$


 FIGURE 6.18 – Signification du vecteur $\mathbf{h}$ du modèle de Ward.

Dans cette expression, les constantes de normalisation π et 4π ont été « absorbées », respectivement, par l'albédo ρ et l'albédo spéculaire ρ_s (cf. paragraphe 1.1.1).

Le niveau de gris (6.12) prédit par le modèle de Ward sous un éclairage directionnel est une généralisation du cas lambertien. Cela est également le cas du niveau de gris (6.6) prédit par le modèle d'Oren-Nayar, mais au lieu d'être une *multiplication* par un « facteur correctif », la modification consiste ici en l'*addition* d'un terme spéculaire. La figure 6.19 montre que les images simulées avec le modèle (6.12) sont plutôt réalistes.

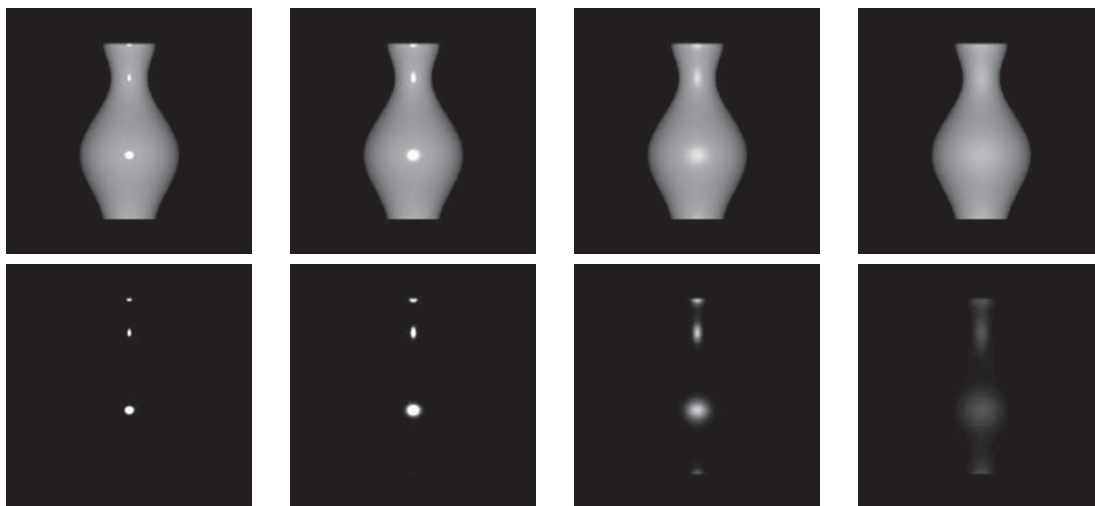


FIGURE 6.19 – Première ligne : images de synthèse calculées avec le modèle de Ward (6.12), pour différentes valeurs de σ (de gauche à droite) : 0,05 ; 0,1 ; 0,2 ; 0,4. Deuxième ligne : composantes spéculaires correspondantes (deuxième ligne de l'équation (6.12)).

Comme cela est illustré sur la figure 6.20, lorsque $\sigma \in [0; 0, 1]$, les reflets brillants sont très localisés, i.e. ils n'affectent qu'une partie infime de l'image : le terme dominant dans (6.12) est le terme lambertien. Il convient dans ce cas de traiter les reflets brillants comme des écarts *parcimonieux* au modèle lambertien, en utilisant une des méthodes robustes présentées dans le chapitre 1, plutôt que de chercher à inverser le modèle spéculaire. D'autre part, lorsque la valeur de σ est élevée ($\sigma \geq 0, 4$), le matériau est quasi-lambertien, auquel cas il ne semble pas raisonnable d'utiliser un modèle aussi complexe. Le seul cas où l'utilisation d'un modèle spéculaire pourrait être intéressante est donc le cas intermédiaire⁴ ($\sigma \approx 0, 2$). L'intérêt d'utiliser un modèle de réflectance spéculaire explicite paraît donc, de manière générale, plutôt limité.

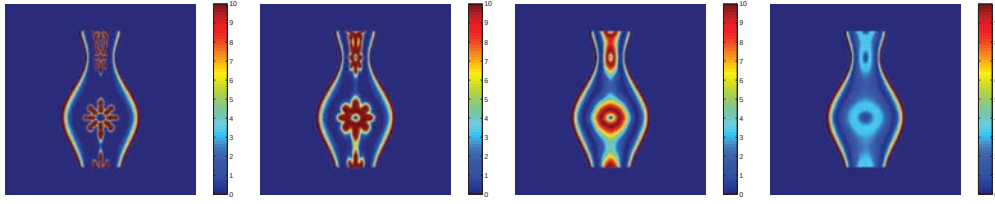


FIGURE 6.20 – Écart angulaire (en degrés) entre la normale réelle et celle estimée en utilisant le modèle lambertien, à partir de $m = 24$ images calculées avec le modèle de Ward (6.12), pour différentes valeurs de σ (de gauche à droite) : 0, 05 ; 0, 1 ; 0, 2 ; 0, 4. Les erreurs sont localisées aux emplacements des reflets brillants lorsque la valeur de σ est faible, et deviennent négligeables lorsque cette valeur est élevée.

Contrairement au modèle d'Oren-Nayar qui est le produit de deux facteurs, le modèle de Ward (6.12) est une somme pondérée de deux termes, l'un diffus et l'autre spéculaire. Chung et al. ont proposé dans [59] une heuristique pour estimer l'ensemble des paramètres du modèle : en utilisant des techniques de *shape-from-shadow*, la normale est identifiée sans ambiguïté en certains points, et les paramètres de réflectance σ et ρ_s (supposés uniformes) peuvent être estimés à partir de ces points. Une fois ces paramètres estimés, il est proposé d'alterner des phases d'estimation de l'albédo (diffus) ρ , connaissant la normale $\mathbf{n}$ (moindres carrés linéaires), et d'estimation de $\mathbf{n}$ connaissant ρ . Cette dernière étape, qui n'est pas triviale puisqu'il faut tenir compte de la contrainte $\|\mathbf{n}\| = 1$, n'est malheureusement pas détaillée. Au vu des non linéarités du modèle de Ward et de la contrainte $\|\mathbf{n}\| = 1$, il nous semble utile de proposer un algorithme permettant d'estimer la normale. Nous décrivons donc, dans le prochain paragraphe, une méthode moins empirique que celle de Chung et al., ne nécessitant pas d'hypothèse sur ρ_s , permettant de se passer du recours à une technique de reconstruction 3D annexe, et fondée uniquement sur des estimations en moindres carrés linéaires, grâce à l'introduction d'un algorithme proximal. De plus, l'approche présentée est généralisable à tout modèle de réflectance défini comme la somme pondérée de deux termes (modèles *dichromatiques*).

4. C'est justement ce cas intermédiaire qui fait l'objet de l'étude présentée dans [59].

6.3.2 Inversion du modèle de Ward par une approche proximale

En introduisant les variables auxiliaires $\mathbf{m}(u, v) = \rho(u, v)\mathbf{n}(u, v)$ et $\mathbf{m}_s(u, v) = \rho_s(u, v)\mathbf{n}(u, v)$, l'estimation de $\mathbf{n}$, ρ et ρ_s peut être reformulée comme l'estimation de $\mathbf{m}$ et $\mathbf{m}_s$ à partir du système d'équations quasi-linéaires suivant :

$$\mathbf{s} \cdot \mathbf{m}(u, v) = I(u, v) - \chi_W(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v)) \max \{0, \mathbf{s} \cdot \mathbf{m}_s(u, v)\} \quad (6.13)$$

$$\mathbf{s} \cdot \mathbf{m}_s(u, v) = \frac{I(u, v) - \max \{0, \mathbf{s} \cdot \mathbf{m}(u, v)\}}{\chi_W(\bar{\mathbf{s}}, \mathbf{n}(u, v), \mathbf{v}(u, v))} \quad (6.14)$$

$$\mathbf{m}_s(u, v) = \frac{\rho_s(u, v)}{\rho(u, v)} \mathbf{m}(u, v) \quad (6.15)$$

En rendant explicites les parties non linéaires et implicites les parties linéaires, la solution de ce problème est obtenue en itérant la résolution du problème sous contrainte suivant :

$$\begin{cases} (\mathbf{m}^{k+1}(u, v), \mathbf{m}_s^{k+1}(u, v)) = \underset{\mathbf{m}, \mathbf{m}_s}{\operatorname{argmin}} \left\| \mathbf{S}\mathbf{m} - \mathbf{b}^k(u, v) \right\|^2 + \left\| \mathbf{S}\mathbf{m}_s - \mathbf{b}_s^k(u, v) \right\|^2 \\ \text{s.c.} \quad \mathbf{m}_s = \frac{\rho_s^k(u, v)}{\rho^k(u, v)} \mathbf{m} \end{cases} \quad (6.16)$$

où :

$$\mathbf{b}^k(u, v) = \begin{bmatrix} I^1(u, v) - \chi_W(\bar{\mathbf{s}}^1, \mathbf{n}^k(u, v), \mathbf{v}(u, v)) \max \{0, \mathbf{s}^1 \cdot \mathbf{m}_s^k(u, v)\} \\ \vdots \\ I^m(u, v) - \chi_W(\bar{\mathbf{s}}^m, \mathbf{n}^k(u, v), \mathbf{v}(u, v)) \max \{0, \mathbf{s}^m \cdot \mathbf{m}_s^k(u, v)\} \end{bmatrix} \quad (6.17)$$

$$\mathbf{b}_s^k(u, v) = \begin{bmatrix} \frac{I^1(u, v) - \max \{0, \mathbf{s}^1 \cdot \mathbf{m}^k(u, v)\}}{\chi_W(\bar{\mathbf{s}}^1, \mathbf{n}^k(u, v), \mathbf{v}(u, v))} \\ \vdots \\ \frac{I^m(u, v) - \max \{0, \mathbf{s}^m \cdot \mathbf{m}^k(u, v)\}}{\chi_W(\bar{\mathbf{s}}^m, \mathbf{n}^k(u, v), \mathbf{v}(u, v))} \end{bmatrix} \quad (6.18)$$

On peut par exemple initialiser l'itération (6.16) de la manière suivante :

- $\mathbf{m}^0(u, v)$ est la solution du problème lambertien ;
- $\rho_s^0(u, v)$ est la solution (en moindres carrés linéaires) du système formé par les m équations (6.12) correspondant aux m éclairages $\mathbf{s}^i$, $i \in [1, m]$, dans lesquelles on fixe $\rho(u, v)$ à $\|\mathbf{m}^0(u, v)\|$ et $\mathbf{n}(u, v)$ à $\mathbf{m}^0(u, v)/\|\mathbf{m}^0(u, v)\|$;
- $\mathbf{m}_s^0(u, v) = \rho_s^0(u, v) \mathbf{m}^0(u, v)/\|\mathbf{m}^0(u, v)\|$.

La solution du problème ADMM (6.16) s'obtient par les limites :

$$\mathbf{m}^{k+1} = \mathbf{m}^{k, \infty} \quad (6.19)$$

$$\mathbf{m}_s^{k+1} = \mathbf{m}_s^{k, \infty} \quad (6.20)$$

des itérations suivantes :

$$\mathbf{m}^{k,j+1}(u, v) = \underset{\mathbf{m}}{\operatorname{argmin}} \left\| \mathbf{S}\mathbf{m} - \mathbf{b}^k(u, v) \right\|^2 + \theta \left\| \frac{\rho_s^k(u, v)}{\rho^k(u, v)} \mathbf{m} - \mathbf{m}_s^{k,j}(u, v) + \mathbf{m}_{\text{aux}}^{k,j}(u, v) \right\|^2 \quad (6.21)$$

$$\mathbf{m}_s^{k,j+1}(u, v) = \underset{\mathbf{m}_s}{\operatorname{argmin}} \left\| \mathbf{S}\mathbf{m}_s - \mathbf{b}_s^k(u, v) \right\|^2 + \theta \left\| \frac{\rho_s^k(u, v)}{\rho^k(u, v)} \mathbf{m}^{k,j+1}(u, v) - \mathbf{m}_s + \mathbf{m}_{\text{aux}}^{k,j}(u, v) \right\|^2 \quad (6.22)$$

$$\mathbf{m}_{\text{aux}}^{k,j+1}(u, v) = \mathbf{m}_{\text{aux}}^{k,j}(u, v) + \frac{\rho_s^k(u, v)}{\rho^k(u, v)} \mathbf{m}^{k,j+1}(u, v) - \mathbf{m}_s^{k,j+1}(u, v) \quad (6.23)$$

où la valeur de $\theta > 0$ influe principalement sur le taux de convergence [235]. L'initialisation des trois itérations (6.21), (6.22) et (6.23) s'effectue de la manière suivante :

- $\mathbf{m}^{k,0} = \mathbf{m}^k$;
- $\mathbf{m}_s^{k,0} = \mathbf{m}_s^k$;
- $\mathbf{m}_{\text{aux}}^{k,0} = \frac{\rho_s^k}{\rho^k} \mathbf{m}^k - \mathbf{m}_s^k$.

Les sous-problèmes (6.21) et (6.22) admettent des solutions explicites :

$$\mathbf{m}^{k,j+1}(u, v) = \left(\mathbf{S}^\top \mathbf{S} + \theta \frac{\rho_s^k(u, v)}{\rho^k(u, v)} \mathbf{I}_3 \right)^{-1} \left[\mathbf{S}^\top \mathbf{b}^k(u, v) + \theta \left(\mathbf{m}_s^{k,j}(u, v) - \mathbf{m}_{\text{aux}}^{k,j}(u, v) \right) \right] \quad (6.24)$$

$$\mathbf{m}_s^{k,j+1}(u, v) = \left(\mathbf{S}^\top \mathbf{S} + \theta \mathbf{I}_3 \right)^{-1} \left[\mathbf{S}^\top \mathbf{b}_s^k(u, v) + \theta \left(\frac{\rho_s^k(u, v)}{\rho^k(u, v)} \mathbf{m}^{k,j+1}(u, v) + \mathbf{m}_{\text{aux}}^{k,j}(u, v) \right) \right] \quad (6.25)$$

On remarque que la matrice à inverser dans le problème (6.25) est la même en tout point (u, v) , ce qui permet une résolution extrêmement rapide. Cela n'est pas vrai, en revanche, pour le problème (6.24). Plutôt que d'inverser une à une $|\Omega|$ matrices 3×3 (qui sont toutes symétriques, définies positives), on peut inverser la matrice suivante :

$$\begin{bmatrix} \mathbf{S}^\top \mathbf{S} + \theta \frac{\rho_s^k(u_1, v_1)}{\rho^k(u_1, v_1)} \mathbf{I}_3 & \mathbf{0}_{3,3} & \dots & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{S}^\top \mathbf{S} + \theta \frac{\rho_s^k(u_2, v_2)}{\rho^k(u_2, v_2)} \mathbf{I}_3 & \dots & \mathbf{0}_{3,3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{3,3} & \dots & \mathbf{0}_{3,3} & \mathbf{S}^\top \mathbf{S} + \theta \frac{\rho_s^k(u_{|\Omega|}, v_{|\Omega|})}{\rho^k(u_{|\Omega|}, v_{|\Omega|})} \mathbf{I}_3 \end{bmatrix} \quad (6.26)$$

dont les blocs diagonaux correspondent aux différentes matrices des problèmes (6.24) associés aux pixels $(u, v) \in \Omega$. La matrice (6.26) étant toujours symétrique, définie positive, le système linéaire formé par l'union des $|\Omega|$ systèmes (6.24) est résolu très rapidement par factorisation de Cholesky de la matrice (6.26).

Enfin, comme pour le modèle d'Oren-Nayar, on peut également estimer le paramètre de réflectance σ . Plutôt que d'échantillonner les valeurs possibles de σ et de comparer les RMSE, nous avons observé qu'il était préférable d'estimer σ au cours des itérations, par une optimisation en moindres carrés non linéaires. La méthode de résolution complète est décrite par l'algorithme 2, dans lequel les deux problèmes non triviaux (6.31) et (6.27) sont résolus, respectivement, par l'algorithme ADMM décrit ci-dessus et par une optimisation en moindres carrés non linéaires de type Levenberg-Marquardt.

Algorithme 2: Une itération de l'algorithme permettant d'appliquer la stéréophotométrie à des surfaces dont la réflectance suit le modèle de Ward. La dépendance explicite en σ de la fonction χ_W signifie que la rugosité est considérée comme une nouvelle inconnue du problème.

$$\sigma^k = \underset{\sigma}{\operatorname{argmin}} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[I^i(u,v) - \rho^k(u,v) \max \left\{ 0, \mathbf{s}^i \cdot \mathbf{n}^k(u,v) \right\} - \rho_s^k(u,v) \max \left\{ 0, \mathbf{s}^i \cdot \mathbf{n}^k(u,v) \right\} \chi_W(\bar{\mathbf{s}}^i, \mathbf{n}^k(u,v), \mathbf{v}(u,v), \sigma) \right]^2 \quad (6.27)$$

$$\chi_W^{i,k}(u,v) = \frac{\exp \left\{ -\frac{\tan^2(\cos^{-1}(\mathbf{h}^i(u,v) \cdot \mathbf{n}^k(u,v)))}{(\sigma^k)^2} \right\}}{(\sigma^k)^2 \sqrt{(\bar{\mathbf{s}}^i \cdot \mathbf{n}^k(u,v)) (\mathbf{v}(u,v) \cdot \mathbf{n}^k(u,v))}}, \quad \forall i \in [1, m] \quad (6.28)$$

$$\mathbf{b}^k(u,v) = \begin{bmatrix} I^1(u,v) - \chi_W^{1,k}(u,v) \max \left\{ 0, \mathbf{s}^1 \cdot \mathbf{m}_s^k(u,v) \right\} \\ \vdots \\ I^m(u,v) - \chi_W^{m,k}(u,v) \max \left\{ 0, \mathbf{s}^m \cdot \mathbf{m}_s^k(u,v) \right\} \end{bmatrix} \quad (6.29)$$

$$\mathbf{b}_s^k(u,v) = \begin{bmatrix} \frac{I^1(u,v) - \max \left\{ 0, \mathbf{s}^1 \cdot \mathbf{m}^k(u,v) \right\}}{\chi_W^{1,k}(u,v)} \\ \vdots \\ \frac{I^m(u,v) - \max \left\{ 0, \mathbf{s}^m \cdot \mathbf{m}^k(u,v) \right\}}{\chi_W^{m,k}(u,v)} \end{bmatrix} \quad (6.30)$$

$$\begin{aligned} \left(\mathbf{m}^{k+1}(u,v), \mathbf{m}_s^{k+1}(u,v) \right) &= \underset{\mathbf{m}, \mathbf{m}_s}{\operatorname{argmin}} \left\| \mathbf{S}\mathbf{m} - \mathbf{b}^k(u,v) \right\|^2 + \left\| \mathbf{S}\mathbf{m}_s - \mathbf{b}_s^k(u,v) \right\|^2 \\ \text{s.c.} \quad \mathbf{m}_s &= \frac{\rho_s^k(u,v)}{\rho^k(u,v)} \mathbf{m} \end{aligned} \quad (6.31)$$

$$\rho^{k+1}(u,v) = \left\| \mathbf{m}^{k+1}(u,v) \right\| \quad (6.32)$$

$$\rho_s^{k+1}(u,v) = \left\| \mathbf{m}_s^{k+1}(u,v) \right\| \quad (6.33)$$

$$\mathbf{n}^{k+1}(u,v) = \frac{\mathbf{m}^{k+1}(u,v)}{\left\| \mathbf{m}^{k+1}(u,v) \right\|} \quad (6.34)$$

6.3.3 Résultats

La figure 6.21 montre l'évolution de l'écart angulaire moyen (EAM) sur la normale estimée au cours des premières itérations de l'algorithme 2, lorsque σ est soit connu, soit estimé. Nous testons trois valeurs du nombre m d'images (9, 12 et 96) et deux valeurs de σ (0,1 et 0,2). Rappelons que, pour σ faible, il serait également possible de coupler une méthode d'estimation parcimonieuse au modèle lambertien, tandis que pour σ élevé, l'erreur induite par le modèle lambertien deviendrait négligeable. Ces résultats montrent que l'approche est prometteuse, mais qu'elle doit être appliquée, en pratique, à un grand nombre d'images.

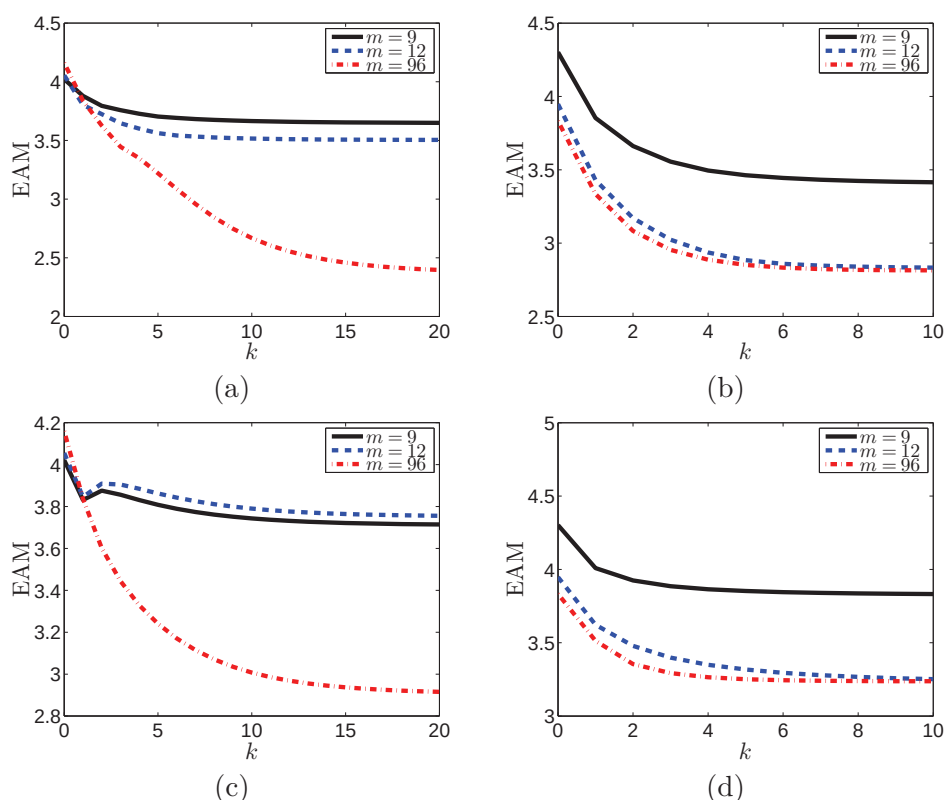


FIGURE 6.21 – Convergence de l'algorithme 2, qui permet d'appliquer la stéréophotométrie aux modèles de réflectance spéculaire. Première ligne : σ est fixé à sa valeur réelle ($\sigma = 0,1$ à gauche, $\sigma = 0,2$ à droite). Deuxième ligne : σ est estimé. Ces résultats montrent que l'approche est prometteuse, car elle permet d'améliorer nettement l'estimation des normales par rapport au modèle lambertien (qui correspond à l'itération 0). Cependant, la qualité de l'estimation croît lorsque le nombre m d'images croît, ou lorsque les reflets brillants sont localisés, i.e. lorsque σ décroît (auquel cas il est préférable d'utiliser le modèle lambertien couplé à une méthode d'estimation robuste).

La figure 6.22 montre par ailleurs que l'initialisation joue un rôle très important : si, au lieu d'estimer ρ_s^0 au sens des moindres carrés linéaires comme dans le paragraphe 6.3.2, on utilise la valeur réelle $\rho_s^0 = \rho_s$, la reconstruction 3D est de bien meilleure qualité.

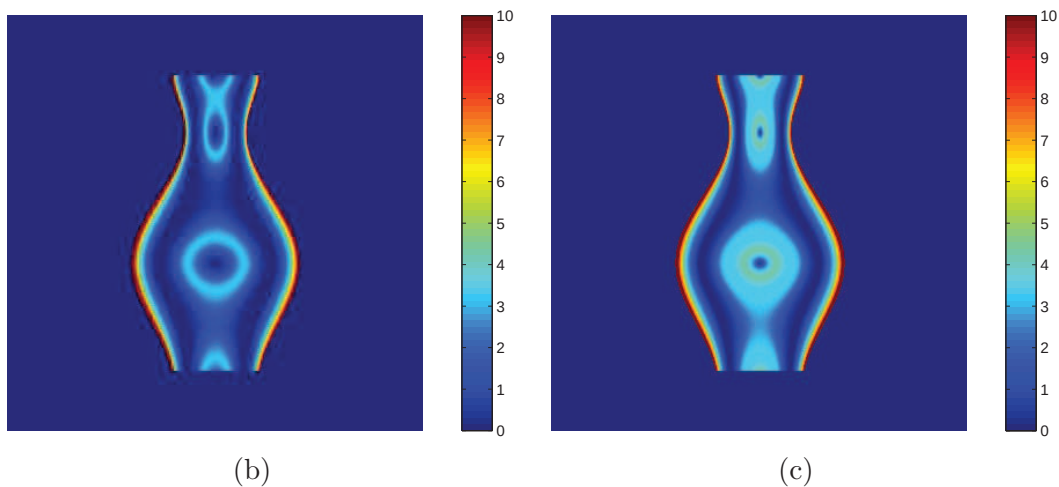
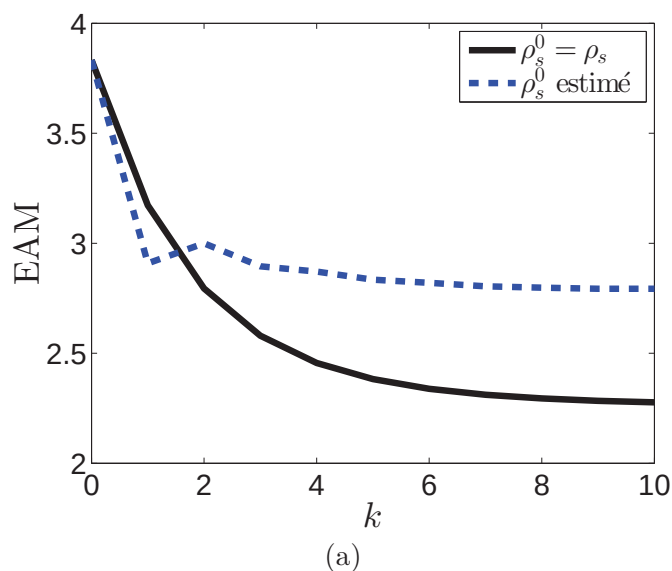


FIGURE 6.22 – (a) Évolution de l'écart angulaire moyen en fonction des itérations, à partir de $m = 24$ images, en fixant σ à sa valeur réelle 0,2 : soit avec l'initialisation de ρ_s par estimation en moindres carrés linéaires, soit en utilisant la valeur réelle $\rho_s^0 = \rho_s$. (b) Écart angulaire à $k = 10$ lorsque $\rho_s^0 = \rho_s$. (c) Écart angulaire à $k = 10$ lorsque ρ_s^0 est estimé. L'estimation de ρ_s semble fortement dégrader la solution.

6.4 Conclusion du chapitre

Après avoir présenté une méthode de résolution de la stéréophotométrie permettant de se passer d'un modèle de réflectance explicite, nous avons montré comment étendre les méthodes de résolution conçues pour les surfaces lambertiennes à des modèles de réflectance plus réalistes, comme le modèle d'Oren-Nayar. Nous avons ensuite proposé un algorithme ADMM permettant d'estimer la normale et l'albédo de façon couplée, lorsque la luminance est modélisée par la somme de deux termes, ce qui permet de traiter des modèles *dichromatiques* tels que celui de Ward. Nous pourrions étendre sans difficulté l'algorithme ADMM proposé à tout modèle de réflectance dichromatique, e.g. au modèle de Torrance-Sparrow.

Cette étude des modèles de réflectance non lambertiens a mis en évidence une faiblesse de l'approche classique de la stéréophotométrie, qui consiste à estimer *simultanément* les normales et l'albédo. Nous avons vu que la qualité de la reconstruction 3D dépendait fortement de l'estimation initiale de l'albédo. Or, c'est précisément pour faciliter la résolution du système linéaire (1.17) que l'albédo est généralement considéré comme une inconnue supplémentaire : grâce à cette astuce, la contrainte $\|\mathbf{n}\| = 1$ n'a plus cours. Pour des modèles plus complexes, la linéarité étant de toute façon perdue, il serait peut-être plus judicieux de découpler l'estimation de la normale de celle de l'albédo.

L'effort à fournir pour modéliser les reflets brillants, puis pour résoudre la stéréophotométrie en utilisant des modèles nécessairement plus complexes que le modèle lambertien, n'est-il pas disproportionné avec le bénéfice que l'on peut espérer en tirer ? Nous avons vu que, sous l'hypothèse lambertienne, il était aisé de concevoir des méthodes d'estimation robuste permettant de prendre en compte les reflets brillants, en les considérant comme des écarts au modèle. Nous avons également vu qu'il était nécessaire de disposer d'un grand nombre d'images pour pouvoir utiliser des modèles de réflectance réalistes comme les modèles dichromatiques. Ceci semble plaider en faveur d'un « retour » au modèle lambertien. Rappelons en outre que nous avons discrédité l'approche par apprentissage à cause de la nécessité d'utiliser un objet de référence. Un reproche moindre, mais du même type, peut être fait à l'encontre de l'utilisation de modèles de réflectance sophistiqués. En effet, sans connaissance a priori sur la scène, comment faire pour choisir un modèle de réflectance plutôt qu'un autre ?

Forts de ces conclusions, nous présentons dans la prochaine partie une nouvelle approche de la stéréophotométrie, fondée sur l'hypothèse lambertienne, mais qui permet de dépasser les limites que nous avons mises en évidence dans les trois premières parties de ce mémoire, à savoir :

- L'estimation robuste de la normale (cf. chapitre 1) est *locale*, et ne garantit donc pas la régularité du champ estimé.
- La surface est reconstruite en deux étapes (cf. chapitre 2), avec le risque de propagation de biais que cela implique.
- Dans certaines situations comme les problèmes non calibré (cf. chapitre 3) ou à deux images (cf. chapitre 4), l'estimation locale de la normale est ambiguë.

- Une modélisation fine de l'éclairage (cf. chapitre 5) fait intervenir les coordonnées des points de la surface, qui sont inconnues, ce qui rompt la linéarité du modèle.

Nous allons montrer que la formulation variationnelle de la stéréophotométrie différentielle permet d'apporter une solution à chacune de ces limitations apparentes.

Quatrième partie

Approche variationnelle de la stéréophotométrie

Avant-propos

La résolution « classique » de la stéréophotométrie opère en deux étapes : estimation des normales, puis intégration des normales. Cela permet de décomposer le problème en deux sous-problèmes linéaires. Néanmoins, nous avons mis en évidence, dans le chapitre 5, les limites de cette approche lorsque la profondeur z intervient explicitement dans le modèle de formation des images. Il devient alors nécessaire de répéter l'estimation des normales et l'intégration des normales en boucle. Si l'initialisation de z est suffisamment proche de la solution, cette méthode de résolution converge généralement en quelques itérations (cf. figure 5.12), mais dans le cas contraire, elle peut également diverger. Une façon d'éviter de genre de difficultés numériques consiste à *estimer directement la profondeur*, i.e. à cesser de manipuler $\mathbf{n}$, comme s'il s'agissait d'une variable indépendante de z . Puisque $\mathbf{n}$ dépend de ∇z , il nous faut maintenant résoudre la formulation *différentielle* de la stéréophotométrie.

À vrai dire, nous nous sommes déjà intéressés à la formulation différentielle du problème PS2 (cf. chapitre 4), mais uniquement dans un cadre théorique. Contrairement au *shape-from-shading*, où la formulation différentielle a fait l'objet d'un nombre relativement important de travaux (voir, e.g., [82] pour un état de l'art récent), en particulier sur la question du nombre de solutions, la formulation différentielle de la stéréophotométrie a été très peu étudiée. Citons toutefois le travail de Chabrowski et al. [44], dans lequel le caractère bien posé du problème PS2 est prouvé par un raisonnement variationnel, les articles de Chandraker et al. [52, 53], qui utilisent cette formulation pour mettre en évidence un certain nombre d'invariants, et ceux de Mecca et al., qui ont montré récemment [174, 175] qu'elle se prêtait particulièrement bien à l'étude du caractère bien posé de la stéréophotométrie en présence d'éclairages non directionnels (cf. chapitre 5) ou de surfaces spéculaires (cf. chapitre 6).

Cependant, l'ensemble de ces travaux sont plutôt d'ordre théorique, les seules méthodes de résolution numérique proposées étant fondées sur des techniques de propagation : de tels schémas numériques sont appropriés pour calculer une solution exacte, mais ils propagent également les erreurs d'estimation, inévitables avec des données réelles. Pour dépasser le cap de la simple étude théorique, et proposer des méthodes de résolution numérique adaptées à des situations complexes, il nous faut recourir à une approche *variationnelle* de la stéréophotométrie différentielle, afin d'estimer une solution approchée du problème, optimale au sens d'un certain critère qui, s'il est judicieusement choisi, garantira la robustesse de l'estimation. À notre connaissance, la formulation de la stéréophotométrie comme un problème variationnel unique n'a pas encore été étudiée. Ce type de formulation a toutefois été utilisé par Yin et al. pour résoudre le problème de l'estimation conjointe du relief, de l'éclairage et de l'albédo, supposé uniforme [137] ou uniforme par morceaux [136], à partir de plusieurs photographies d'un objet prises sous différents angles. Yin et al. montrent que, dans ce cas, une modélisation variationnelle naïve conduit à des schémas numériques instables. Nous pouvons donc également nous attendre à des difficultés numériques, dans le cas de la stéréophotométrie.

Dans cette partie, nous montrons comment résoudre la stéréophotométrie différentielle à l'aide d'une approche variationnelle. Dans le chapitre 7, nous revenons sur le « cas d'école » du chapitre 1 ($m \geq 3$ sources directionnelles, non coplanaires, connues). Nous proposons une nouvelle résolution de ce problème, où la profondeur est directement estimée par minimisation d'un critère L^1 -TV inspiré de méthodes dont l'efficacité pour le traitement d'images est bien établie [49]. Cette nouvelle méthode de résolution repousse certaines limites de la résolution classique en deux étapes : il n'est plus nécessaire d'imposer explicitement l'intégrabilité ; la robustesse de l'estimation est relativement facile à garantir ; les discontinuités de profondeur peuvent être retrouvées. Nous montrons ensuite, dans le chapitre 8, que l'approche variationnelle de la stéréophotométrie différentielle ouvre de nouvelles perspectives pour la résolution de problèmes difficiles, comme le cas où les éclairages sont directionnels mais inconnus (problème UPS, cf. chapitre 3), celui où les éclairages sont non directionnels (cf. chapitre 5), ou encore la stéréophotométrie « doublement colorée » (surface et éclairages simultanément colorés).

Résolution variationnelle de la stéréophotométrie différentielle

Dans ce chapitre, nous proposons une reformulation de la stéréophotométrie visant à estimer directement la profondeur, sans passer par l'étape intermédiaire d'estimation des normales. Nous allons voir que les difficultés rencontrées en utilisant cette approche sont très différentes de celles qui ont pu être observées jusqu'ici. Avec l'approche classique de la stéréophotométrie, la résolution du « cas d'école » ($m \geq 3$ éclairages directionnels, non coplanaires, connus) peut être effectuée en utilisant des méthodes relativement simples (cf. partie I). Lorsque l'une de ces hypothèses est mise en défaut (cf. parties II et III), la résolution devient nettement plus ardue. Avec l'approche différentielle, la difficulté est déplacée : il est relativement difficile de résoudre le problème sous les hypothèses les plus favorables, mais l'extension de cette approche à des configurations a priori moins favorables ne présente que peu de difficultés.

Les estimateurs robustes proposés dans le chapitre 1 ont été utilisés pixel par pixel, sans aucune notion de cohérence spatiale. Il semble assez évident que l'introduction d'un critère de cohérence spatiale ne peut qu'améliorer la robustesse de la stéréophotométrie : les méthodes variationnelles offrent un cadre naturel à l'introduction d'un a priori sur la régularité des fonctions estimées. Le niveau de robustesse ainsi atteint pourra notamment permettre d'utiliser la stéréophotométrie dans des environnements peu contrôlés, comme par exemple celui de l'imagerie endoscopique [64]. Un autre intérêt des méthodes variationnelles consiste en leur flexibilité. Il est en effet très facile d'intégrer des opérateurs modélisant certains effets, comme les parties manquantes [30, 169] ou le flou de mise au point [50]. Le modèle ROF [226], initialement dédié au débruitage, s'étend par exemple très naturellement aux problèmes combinés de débruitage, de défloutage, et de zoom (voir, e.g., [48]). Certains auteurs ont tenté récemment de coupler la stéréophotométrie avec des techniques de défloutage [72, 139] : voilà un exemple de problème où l'approche variationnelle serait particulièrement adaptée.

Dans le paragraphe 7.1, nous reformulons le problème en moindres carrés du « cas d'école » de la stéréophotométrie, en cherchant à estimer directement la profondeur grâce à une formulation variationnelle. Nous introduisons un schéma numérique efficace pour résoudre ce problème, fondé sur la combinaison d'un algorithme de point fixe et d'un algorithme ADMM. Nous adaptons ensuite, dans le paragraphe 7.2, certains des estimateurs robustes déjà présentés dans le chapitre 1, et les couplons avec des termes de régularisation par variation totale, qui augmentent encore la robustesse de l'estimation. Ceci nous permet finalement de proposer un cadre théorique et numérique original pour la résolution robuste de la stéréophotométrie.

7.1 Résolution variationnelle en moindres carrés

Nous revenons d'abord sur le problème de stéréophotométrie le plus simple i.e., l'estimation en moindres carrés du relief et de l'albédo d'une surface lambertienne, à partir de $m \geq 3$ photographies en niveaux de gris, prises sous m éclairages directionnels, non coplanaires, ayant été préalablement étalonnés. Avec l'approche classique de la stéréophotométrie, deux problèmes linéaires doivent être résolus séquentiellement : l'estimation des normales et de l'albédo en chaque pixel, puis l'intégration des normales. Avec l'approche différentielle, au contraire, un seul problème est résolu, ce qui permet d'éviter la propagation de biais inhérente à la résolution en deux étapes successives.

7.1.1 Linéarisation de la stéréophotométrie : un obstacle à l'optimalité de la reconstruction 3D

Afin de motiver la résolution variationnelle de la stéréophotométrie différentielle, attardons-nous un instant sur l'optimalité de la reconstruction 3D (au sens du maximum de vraisemblance). Pour simplifier, nous négligeons dans ce paragraphe les ombres propres, et nous supposons que l'appareil photographique est orthographique. Nous allons montrer qu'il est impossible d'obtenir une estimation du relief qui soit à la fois optimale et linéaire : afin de garantir l'optimalité de la reconstruction 3D, il faut soit utiliser des estimateurs non linéaires (ce que nous avons proposé dans la partie I), soit accepter la non linéarité du modèle.

Approche classique. Le problème que nous avons étudié dans le chapitre 1 est celui de l'estimation d'une valeur réelle $\rho \in \mathbb{R}$ et d'un vecteur $\mathbf{n} \in \mathbb{R}^3$, *unitaire*, tel que le couple $(\rho(u, v), \mathbf{n}(u, v))$ constitue une solution approchée du système d'équations $I^i(u, v) = \rho(u, v) \mathbf{s}^i \cdot \mathbf{n}(u, v)$, $i \in [1, m]$. La contrainte d'unitarité peut être éliminée en introduisant la variable $\mathbf{m}(u, v) = \rho(u, v) \mathbf{n}(u, v)$: l'estimateur en moindres carrés $\min_{\mathbf{m}} \sum_{i=1}^m \|I^i(u, v) - \mathbf{s}^i \cdot \mathbf{m}\|^2$ admet une solution analytique en chaque pixel (u, v) , de laquelle on déduit facilement la normale $\mathbf{n}(u, v) = \mathbf{m}(u, v) / \|\mathbf{m}(u, v)\|$. Supposons que l'écart entre la normale estimée et la normale réelle soit gaussien, de moyenne nulle (ceci n'est pas le cas en pratique, comme cela sera discuté dans le prochain paragraphe). Il reste à *intégrer* le champ de normales estimé $\mathbf{n} = [n_1, n_2, n_3]^\top$. Une estimation du gradient de la carte de profondeur est donnée par $\mathbf{g} = [-n_1/n_3, -n_2/n_3]^\top$. L'erreur sur l'estimation du gradient est donc le quotient de deux gaussiennes de moyenne nulle : il s'agit d'une distribution de Cauchy. La surface obtenue par résolution du problème en moindres carrés linéaires $\min_z \|\nabla z - \mathbf{g}\|_{L^2(\Omega)}^2$ n'est donc pas l'estimation « optimale », au sens du maximum de vraisemblance. Pour obtenir l'estimation optimale, il faut remplacer l'estimateur en moindres carrés $\Phi_{LS}(s) = s^2$ par l'estimateur non convexe $\Phi_2(s) = \frac{1}{2} \ln(\beta^2 + s^2)$, ce qui complique fortement la résolution (cf. chapitre 2). La linéarité de l'approche classique est donc illusoire, car cette approche ne permet d'obtenir une estimation optimale qu'en introduisant des estimateurs non linéaires.

Ratios non-différentiels. La première étape de l'approche classique permet d'estimer la variable $\mathbf{m} = \rho \mathbf{n}$ au sens du maximum de vraisemblance. Elle ne garantit toutefois pas l'optimalité de l'estimation de la normale. Afin de garantir une estimation optimale de la normale, l'albédo ρ peut être éliminé des inconnues au moyen des ratios d'images (cf. paragraphe 4.1.2). En effet, pour une surface lambertienne d'albédo non nul, nous pouvons écrire, en un point (u, v) quelconque de Ω :

$$\frac{I^i(u, v)}{I^j(u, v)} = \frac{\rho(u, v) \mathbf{s}^i \cdot \mathbf{n}(u, v)}{\rho(u, v) \mathbf{s}^j \cdot \mathbf{n}(u, v)}, \forall (i, j) \in [1, m]^2, i \neq j \quad (7.1)$$

en supposant que les dénominateurs ne s'annulent pas, et donc :

$$\left[I^j(u, v) \mathbf{s}^i - I^i(u, v) \mathbf{s}^j \right] \cdot \mathbf{n}(u, v) = 0, \forall (i, j) \in [1, m]^2, i \neq j \quad (7.2)$$

ce qui généralise l'équation (4.17) au cas de m éclairages. Le système formé par les C_m^2 équations (7.2) est un système linéaire homogène, dans lequel l'albédo a été éliminé et le vecteur inconnu $\mathbf{n}(u, v)$ est unitaire. Ce système peut donc être résolu analytiquement au sens des moindres carrés, par décomposition en valeurs singulières [271]. Cela permet d'estimer très facilement la normale indépendamment de l'albédo. Cependant, si les images diffèrent du modèle théorique par un bruit additif gaussien, d'espérance nulle (exogène) et de variance identique pour toutes les images (homoscédastique), alors les ratios réels diffèrent du modèle (7.1) par un bruit distribué selon une loi de Cauchy : la résolution au sens des moindres carrés du système (7.2) ne fournit donc pas l'estimation optimale de la normale. Estimer la normale à l'aide des ratios d'images ne permet donc pas non plus d'obtenir une estimation à la fois linéaire et optimale.

Ratios différentiels. Afin d'éviter la propagation de biais entre l'étape d'estimation de la normale et l'étape d'intégration, il a été proposé très récemment dans [240] de combiner l'utilisation des ratios d'images avec une approche différentielle de la stéréophotométrie. Pour cela, la normale $\mathbf{n}$ dans (7.2) est remplacée par sa définition différentielle (cf. paragraphe 2.1), ce qui permet d'obtenir un système d'équations linéaires en ∇z de la forme : $\mathbf{A}(u, v) \nabla z(u, v) = \mathbf{b}(u, v)$, où $\mathbf{A} : \Omega \rightarrow \mathbb{R}^{C_m^2 \times 2}$ est un champ matriciel, et $\mathbf{b} : \Omega \rightarrow \mathbb{R}^m$ est un champ vectoriel. Dans [240], ce système d'EDP est résolu de manière *globale*, en cherchant une carte de profondeur z solution du problème variationnel en moindres carrés associé. Malheureusement, comme cela a été discuté dans le paragraphe précédent, les ratios réels diffèrent du modèle théorique par une erreur distribuée selon une loi de Cauchy : la résolution du problème variationnel en moindres carrés $\min_z \|\mathbf{A} \nabla z - \mathbf{b}\|_{L^2(\Omega)}^2$ ne fournit donc pas la surface optimale¹. Si l'utilisation des ratios d'images est intéressante d'un point de vue théorique, par exemple pour l'étude d'invariants du problème sous l'hypothèse d'une réflectance isotrope [52, 53], pour limiter le nombre de variables à estimer dans les problèmes mal posés [119, 160, 264], ou pour la recherche d'une preuve d'existence et d'unicité d'une solution en z , indépendamment de l'albédo [173, 175], il est impossible de garantir son optimalité sans recourir à l'utilisation de l'estimateur non convexe Φ_2 .

1. Mais c'est la meilleure estimation linéaire non biaisée (*BLUE*, pour *best linear unbiased estimator*).

On constate donc, quelle que soit l'approche suivie, que l'estimation optimale est non linéaire. On peut donc s'interroger sur la pertinence du choix de linéariser à tout prix le modèle, en introduisant une variable auxiliaire ou en utilisant les ratios d'images. Au contraire, nous préconisons d'estimer l'albédo et le relief comme deux entités indépendantes, en acceptant dès le départ la non linéarité du modèle. En choisissant le couple (ρ, z) comme inconnue, plutôt que le couple $(\rho, \mathbf{n})$, l'étape d'intégration des normales est évitée : le problème est formulé directement comme un système d'EDP non linéaires, que nous suggérons de résoudre par une approche variationnelle. Avec cette méthodologie, on peut notamment garantir que l'estimation de (ρ, z) en moindres carrés est optimale, au sens du maximum de vraisemblance, en présence d'un bruit additif gaussien exogène sur les images. Un schéma d'optimisation alternée permettant d'approcher itérativement cette estimation sera proposé dans le paragraphe 7.1.7. Il s'agit, à notre connaissance, de la première formulation à la fois non biaisée et optimale de la stéréophotométrie. Nous proposerons de plus, dans le paragraphe 7.2, un modèle L^1 -TV permettant d'approcher l'estimateur du maximum a posteriori, lorsque les images diffèrent du modèle par un bruit endogène (prise en compte des ombres propres et de l'éclairage ambiant), distribué selon une loi de Laplace (prise en compte des ombres portées, des reflets brillants et des discontinuités de profondeur), et que la surface est supposée régulière par morceaux.

Dans un premier temps, nous supposons que l'albédo est connu, de telle sorte que la seule inconnue du problème est la profondeur z . Rappelons que l'estimation de la normale (cf. chapitre 1) est indépendante du modèle de projection (orthographique ou perspective), et que le choix d'un modèle de projection n'a d'importance, dans l'approche classique, que pour l'étape d'intégration (cf. chapitre 2). Au contraire, la modélisation différentielle dépend explicitement du modèle de projection. Nous étudions donc séparément la projection orthographique et la projection perspective.

7.1.2 Projection orthographique

Commençons par le cas de la projection orthographique. Rappelons que, dans ce cas, la normale $\mathbf{n}$ est liée à la profondeur z par :

$$\mathbf{n}(u, v) = \frac{1}{\sqrt{\|\nabla z(u, v)\|^2 + 1}} \begin{bmatrix} -\nabla z(u, v) \\ 1 \end{bmatrix} \quad (7.3)$$

Sous l'hypothèse d'un éclairage directionnel, caractérisé par le vecteur d'éclairage $\mathbf{s} = [s_1, s_2, s_3]^\top \in \mathbb{R}^3$, le modèle lambertien linéaire $I(u, v) = \rho(u, v) \mathbf{s} \cdot \mathbf{n}(u, v)$ s'écrit donc sous la forme d'une équation aux dérivées partielles (EDP) en z :

$$I(u, v) = \rho(u, v) \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}}_{\mathbf{s}} \cdot \underbrace{\frac{1}{\sqrt{\|\nabla z(u, v)\|^2 + 1}} \begin{bmatrix} -\nabla z(u, v) \\ 1 \end{bmatrix}}_{\mathbf{n}(u, v)} \quad (7.4)$$

qui se ramène à l'équation *eikonale* lorsque $\mathbf{s} = [0, 0, 1]^\top$ et $\rho(u, v)$ est connu.

Pour estimer la profondeur z par stéréophotométrie, il nous faut résoudre le système d'EDP suivant :

$$\begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix} = \frac{\rho(u, v)}{\sqrt{\|\nabla z(u, v)\|^2 + 1}} \left(\begin{bmatrix} -s_1^1 & -s_2^1 \\ \vdots & \vdots \\ -s_1^m & -s_2^m \end{bmatrix} \nabla z(u, v) + \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \right) \quad (7.5)$$

On remarque que toute solution z de (7.5) n'est définie qu'à une constante additive près. Comme dans le chapitre 2, nous contournons cette difficulté en introduisant une connaissance a priori z_0 sur la profondeur, de telle sorte que la solution en moindres carrés du système (7.5) sur Ω^2 , avec l'a priori de profondeur, est la solution du problème variationnel :

$$\min_z \left\| \begin{bmatrix} I^1 \\ \vdots \\ I^m \end{bmatrix} - \frac{\rho}{\sqrt{\|\nabla z\|^2 + 1}} \left(\begin{bmatrix} -s_1^1 & -s_2^1 \\ \vdots & \vdots \\ -s_1^m & -s_2^m \end{bmatrix} \nabla z + \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \right) \right\|_{L^2(\Omega)}^2 + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \quad (7.6)$$

où $\|\cdot\|_{L^2(\Omega)}$ désigne la norme de l'espace des fonctions de carré Lebesgue-intégrable : plutôt que de résoudre un problème d'optimisation en chaque point, comme cela est le cas avec l'approche classique, le problème de reconstruction 3D est formulé comme un unique problème d'optimisation, qui est global. Ceci ne nécessite qu'une hypothèse très faible sur z : il suffit, pour pouvoir écrire (7.6), que l'espace de recherche de z soit l'espace de Sobolev $\mathcal{W}^{1,2}$ des fonctions z telles que $\nabla z \in L^2(\Omega)$.

7.1.3 Projection perspective

Si la projection est perspective, la normale s'écrit :

$$\mathbf{n}(u, v) = \frac{1}{d_z(u, v)} \left[\begin{array}{c} -\nabla z(u, v) \\ \frac{z(u, v)}{f} + \left[\frac{u-u_0}{f}, \frac{v-v_0}{f} \right]^\top \cdot \nabla z(u, v) \end{array} \right] \quad (7.7)$$

où :

$$d_z(u, v) = \sqrt{\|\nabla z(u, v)\|^2 + \left(\frac{z(u, v)}{f} + \left[\frac{u-u_0}{f}, \frac{v-v_0}{f} \right]^\top \cdot \nabla z(u, v) \right)^2} \quad (7.8)$$

Le système d'EDP à résoudre est donc le suivant :

$$\begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix} = \frac{\rho(u, v)}{d_z(u, v)} \left(\begin{bmatrix} -s_1^1 + \frac{s_3^1(u-u_0)}{f} & -s_2^1 + \frac{s_3^1(v-v_0)}{f} \\ \vdots & \vdots \\ -s_1^m + \frac{s_3^m(u-u_0)}{f} & -s_2^m + \frac{s_3^m(v-v_0)}{f} \end{bmatrix} \nabla z(u, v) + \frac{z(u, v)}{f} \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \right) \quad (7.9)$$

2. Le domaine de reconstruction Ω est temporairement considéré comme un ouvert de $\mathbb{R}^2$.

CHAPITRE 7. RÉSOLUTION VARIATIONNELLE DE LA STÉRÉOPHOTOMÉTRIE DIFFÉRENTIELLE

Toute solution de (7.9) n'est définie qu'à une constante multiplicative près. Nous contournons à nouveau cette difficulté en introduisant une connaissance a priori z_0 sur la profondeur, et en estimant z comme la solution de :

$$\min_z \left\| \begin{bmatrix} I^1 \\ \vdots \\ I^m \end{bmatrix} - \frac{\rho}{d_z} \left(\begin{bmatrix} -s_1^1 + \frac{s_3^1(u-u_0)}{f} & -s_2^1 + \frac{s_3^1(v-v_0)}{f} \\ \vdots & \vdots \\ -s_1^m + \frac{s_3^m(u-u_0)}{f} & -s_2^m + \frac{s_3^m(v-v_0)}{f} \end{bmatrix} \nabla z + \frac{z}{f} \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \right) \right\|_{L^2(\Omega)}^2 + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \quad (7.10)$$

7.1.4 Modèle générique

Que la projection soit orthographique (équation (7.5)) ou perspective (équation (7.9)), le système d'équations à résoudre est un système d'EDP quasi-linéaires de la forme :

$$\mathbf{A}_z \nabla z = \mathbf{b}_z \quad \text{sur } \Omega \quad (7.11)$$

où :

$$\begin{cases} \mathbf{A}_z(u, v) = \frac{\rho(u, v)}{d_z(u, v)} \begin{bmatrix} -s_1^1 & -s_2^1 \\ \vdots & \vdots \\ -s_1^m & -s_2^m \end{bmatrix} \\ \mathbf{b}_z(u, v) = \begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix} - \frac{\rho(u, v)}{d_z(u, v)} \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \\ d_z(u, v) = \sqrt{\|\nabla z(u, v)\|^2 + 1} \end{cases} \quad (7.12)$$

si la projection est orthographique, tandis que si elle est perspective :

$$\begin{cases} \mathbf{A}_z(u, v) = \frac{\rho(u, v)}{d_z(u, v)} \begin{bmatrix} -s_1^1 + \frac{s_3^1(u-u_0)}{f} & -s_2^1 + \frac{s_3^1(v-v_0)}{f} \\ \vdots & \vdots \\ -s_1^m + \frac{s_3^m(u-u_0)}{f} & -s_2^m + \frac{s_3^m(v-v_0)}{f} \end{bmatrix} \\ \mathbf{b}_z(u, v) = \begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix} - \frac{\rho(u, v)z(u, v)}{f d_z(u, v)} \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \\ d_z(u, v) = \sqrt{\|\nabla z(u, v)\|^2 + \left(\frac{z(u, v)}{f} + \left[\frac{u-u_0}{f}, \frac{v-v_0}{f} \right]^\top \cdot \nabla z(u, v) \right)^2} \end{cases} \quad (7.13)$$

Les problèmes (7.6) et (7.10) peuvent donc tous deux être réécrits sous la forme compacte suivante :

$$\min_z \|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^2(\Omega)}^2 + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \quad (7.14)$$

La résolution du problème en moindres carrés (7.14) n'est pas triviale, à cause de la dépendance de $\mathbf{A}_z$ et $\mathbf{b}_z$ en z qui le rend non linéaire. Avant de proposer un algorithme de type point fixe permettant de résoudre ce problème variationnel, nous analysons d'abord sa condition d'optimalité au premier ordre.

7.1.5 Première variation

Nous étudions uniquement le problème en perspective i.e., nous choisissons les définitions (7.13) pour le problème (7.14), sachant que le problème orthographique est légèrement plus simple.

Nous réécrivons pour cela le problème (7.14) sous forme intégrale :

$$\min_z \overbrace{\iint_{(u,v) \in \Omega} \underbrace{\|\mathbf{A}_z(u,v) \nabla z(u,v) - \mathbf{b}_z(u,v)\|^2 + \lambda [z(u,v) - z_0(u,v)]^2}_{\mathcal{L}(u,v,z,\nabla z)} du dv}^{\mathcal{E}(z)} \quad (7.15)$$

où $\mathcal{L}$ est le lagrangien du problème en moindres carrés (7.15), défini par :

$$\mathcal{L}(a,b,g,[p,q]^\top) = \left\| \mathbf{A}_g(a,b) [p(a,b), q(a,b)]^\top - \mathbf{b}_g(a,b) \right\|^2 + \lambda [g(a,b) - z_0(a,b)]^2 \quad (7.16)$$

avec les notations standard $\nabla g = [p, q]^\top$. Le gradient de $\mathcal{E}$ en z , dans la direction w , est défini par :

$$\langle \nabla \mathcal{E}(z), w \rangle = \left. \frac{d}{dt} \underbrace{\mathcal{E}(z + tw)}_{h(t)} \right|_{t=0} \quad (7.17)$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire canonique de $L^2(\Omega)$. Si z est un point critique de $\mathcal{E}$, alors h est minimale pour $t = 0$, et donc $h'(0) = 0$. On peut donc déterminer la condition d'optimalité au premier ordre $\nabla \mathcal{E}(z) = 0$ en écrivant $h'(0)$ comme un produit scalaire.

Nous notons :

$$\mathcal{E}(z) = \mathcal{D}(z) + \mathcal{R}(z) \quad (7.18)$$

où :

$$\mathcal{D}(z) = \iint_{(u,v) \in \Omega} \|\mathbf{A}_z(u,v) \nabla z(u,v) - \mathbf{b}_z(u,v)\|^2 du dv \quad (7.19)$$

et :

$$\mathcal{R}(z) = \lambda \iint_{(u,v) \in \Omega} [z(u,v) - z_0(u,v)]^2 du dv \quad (7.20)$$

de telle sorte que :

$$\left. \frac{d}{dt} \mathcal{E}(z + tw) \right|_{t=0} = \left. \frac{d}{dt} \mathcal{D}(z + tw) \right|_{t=0} + \left. \frac{d}{dt} \mathcal{R}(z + tw) \right|_{t=0} \quad (7.21)$$

et donc, d'après (7.17) :

$$\nabla \mathcal{E}(z) = \nabla \mathcal{D}(z) + \nabla \mathcal{R}(z) \quad (7.22)$$

CHAPITRE 7. RÉSOLUTION VARIATIONNELLE DE LA STÉRÉOPHOTOMÉTRIE DIFFÉRENTIELLE

Intéressons-nous dans un premier temps au terme d'attache aux données, que nous réécrivons de la façon suivante :

$$\mathcal{D}(z) = \iint_{\mathbf{p} \in \Omega} \sum_{i=1}^m \underbrace{\left[I^i(\mathbf{p}) - \rho(\mathbf{p}) \mathbf{s}^i \cdot \mathbf{n}_z(\mathbf{p}) \right]^2}_{L^i(\mathbf{n}_z)} d\mathbf{p} \quad (7.23)$$

où la normale $\mathbf{n}_z$ est indiquée par z pour rappeler la dépendance :

$$\mathbf{n}_z(\mathbf{p}) = \frac{1}{\underbrace{\sqrt{\|\nabla z(\mathbf{p})\|^2 + T_z(\mathbf{p})^2}}_{d_z(\mathbf{p})}} \begin{bmatrix} -\nabla z(\mathbf{p}) \\ T_z(\mathbf{p}) \end{bmatrix} \quad (7.24)$$

avec :

$$\mathbf{p} = \begin{bmatrix} u \\ v \end{bmatrix} \quad (7.25)$$

et :

$$T_z(\mathbf{p}) = \frac{z(\mathbf{p}) + \left(\mathbf{p} - [u_0, v_0]^\top \right) \cdot \nabla z(\mathbf{p})}{f} \quad (7.26)$$

Pour alléger les notations, nous oublierons dorénavant les dépendances en $\mathbf{p}$ lorsqu'elles sont évidentes.

Nous cherchons à calculer la première variation de $\mathcal{D}$ qui s'écrit, d'après la règle de dérivation des fonctions composées :

$$\left. \frac{d}{dt} \mathcal{D}(z + tw) \right|_{t=0} = \iint_{\Omega} \sum_{i=1}^m d_{\mathbf{n}_z} L^i(\mathbf{n}_z) \cdot \left. \frac{d}{dt} \mathbf{n}_{z+tw} d\mathbf{p} \right|_{t=0} \quad (7.27)$$

On obtient finalement :

$$\left. \frac{d}{dt} \mathbf{n}_{z+tw} \right|_{t=0} = -\frac{\nabla z \cdot \nabla w + T_z T_w}{d_z^2} \mathbf{n}_z + \frac{1}{d_z} \begin{bmatrix} -\nabla w \\ T_w \end{bmatrix} \quad (7.28)$$

où T_w est défini par analogie à T_z (cf. équation (7.26)).

La différentielle de $L^i(\mathbf{n}_z)$ s'écrit, quant à elle :

$$d_{\mathbf{n}_z} L^i(\mathbf{n}_z) = -2\rho \left(I^i - \tilde{I}^i \right) \mathbf{s}^i \quad (7.29)$$

où :

$$\tilde{I}^i = \rho \mathbf{s}^i \cdot \mathbf{n}_z \quad (7.30)$$

est l'image « reprojetée ».

En combinant les expressions (7.27), (7.28) et (7.29), et en développant l'expression de T_w , nous obtenons finalement :

$$\begin{aligned}
 \frac{d}{dt} \mathcal{D}(z + tw) \Big|_{t=0} &= 2 \iint_{\Omega} \underbrace{\left[\sum_{i=1}^m \frac{\tilde{I}^i (I^i - \tilde{I}^i)}{d_z^2} \right]}_{f_z} \nabla z \cdot \nabla w \, d\mathbf{p} \\
 &+ 2 \iint_{\Omega} \underbrace{\left[\sum_{i=1}^m \frac{(I^i - \tilde{I}^i)}{d_z} \left(\left(\frac{\tilde{I}^i T_z}{f d_z} - \frac{\rho s_3^i}{f} \right) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \rho \begin{bmatrix} s_1^i \\ s_2^i \end{bmatrix} \right) \right]}_{\mathbf{g}_z} \cdot \nabla w \, d\mathbf{p} \\
 &+ 2 \iint_{\Omega} \underbrace{\left[\sum_{i=1}^m \frac{(I^i - \tilde{I}^i)}{f d_z} \left(\frac{\tilde{I}^i T_z}{d_z} - \rho s_3^i \right) \right]}_{h_z} w \, d\mathbf{p} \tag{7.31}
 \end{aligned}$$

D'après la formule de Green, (7.31) se réécrit³ :

$$\frac{d}{dt} \mathcal{D}(z + tw) \Big|_{t=0} = 2 \iint_{\Omega} w [h_z - \nabla \cdot (f_z \nabla z + \mathbf{g}_z)] \, d\mathbf{p} + 2 \oint_{\partial\Omega} w (f_z \nabla z + \mathbf{g}_z) \cdot \boldsymbol{\mu} \, d\mathbf{p} \tag{7.32}$$

où $\partial\Omega$ est la frontière de l'ouvert Ω , et $\boldsymbol{\mu}$ est la normale à $\partial\Omega$, unitaire et sortante. En utilisant la définition (7.17), on peut identifier $\nabla \mathcal{D}$ dans (7.32) :

$$\nabla \mathcal{D}(z) = \begin{cases} 2 [h_z - \nabla \cdot (f_z \nabla z + \mathbf{g}_z)] & \text{sur } \Omega \\ 2 (f_z \nabla z + \mathbf{g}_z) \cdot \boldsymbol{\mu} & \text{sur } \partial\Omega \end{cases} \tag{7.33}$$

Le calcul du gradient du terme de régularisation (7.20) est classique, et donne :

$$\nabla \mathcal{R}(z) = 2\lambda (z - z_0) \quad \text{sur } \Omega \tag{7.34}$$

D'après (7.22), (7.33) et (7.34) :

$$\nabla \mathcal{E}(z) = \begin{cases} 2 [h_z - \nabla \cdot (f_z \nabla z + \mathbf{g}_z)] + 2\lambda (z - z_0) & \text{sur } \Omega \\ 2 (f_z \nabla z + \mathbf{g}_z) \cdot \boldsymbol{\mu} & \text{sur } \partial\Omega \end{cases} \tag{7.35}$$

Pour éliminer l'intégrale de bord dans (7.32), il faut soit faire en sorte que la fonction $(f_z \nabla z + \mathbf{g}_z) \cdot \boldsymbol{\mu}$ soit orthogonale à toute fonction w (forme faible du principe variationnel), soit faire en sorte que cette fonction s'annule sur $\partial\Omega$ (forme forte), ce qui est possible soit en imposant $z \equiv 0$ (condition de type Dirichlet), soit en imposant imposer la *condition naturelle* au bord oblique :

$$(f_z \nabla z + \mathbf{g}_z) \cdot \boldsymbol{\mu} = 0 \quad \text{sur } \partial\Omega \tag{7.36}$$

3. Pour pouvoir écrire (7.32), il faut une hypothèse plus forte sur la régularité de z que la simple appartenance à l'espace de Sobolev $\mathcal{W}^{1,2}(\Omega)$, par exemple $z \in \mathcal{C}^2(\Omega)$. Nous remplacerons dans le paragraphe 7.2.2 cette hypothèse trop forte par une contrainte sur le caractère borné des variations de z .

CHAPITRE 7. RÉSOLUTION VARIATIONNELLE DE LA STÉRÉOPHOTOMÉTRIE DIFFÉRENTIELLE

Cette condition au bord, de type Neumann, est difficile à manipuler car elle dépend de z de façon non linéaire. Dans les schémas numériques présentés plus loin, traitant du cas discret, la condition au bord naturelle est cependant beaucoup plus simple.

La condition d'optimalité au premier ordre s'écrit donc finalement :

$$\begin{cases} \nabla \cdot (f_z \nabla z) - \lambda z = h_z - \nabla \cdot \mathbf{g}_z - \lambda z_0 & \text{sur } \Omega \\ (f_z \nabla z + \mathbf{g}_z) \cdot \boldsymbol{\mu} = 0 & \text{sur } \partial\Omega \end{cases} \quad (7.37)$$

Lorsque f_z est de signe constant, l'équation (7.37) est une EDP non linéaire décrivant l'état stationnaire d'une équation de réaction-diffusion. Malheureusement, cette propriété n'est pas garantie, ce qui suggère l'utilisation d'une méthode fondée sur une descente de gradient. Dans [216], nous avons utilisé une descente de type quasi-Newton (BFGS) pour résoudre l'équation (7.37). Il est difficile, et même impossible, de montrer la convexité du problème en moindres carrés (7.15), donc la convergence de la descente n'est pas garantie [168]. Cependant, en déterminant le pas de descente par une recherche linéaire avec retour arrière (*backtracking line search*), la convergence n'est généralement pas mise en défaut [162], comme cela est illustré sur la figure 7.2, où nous avons appliqué cette descente aux images de la figure 7.1, en choisissant une condition au bord de type Dirichlet. Pour comprendre pourquoi la descente se comporte bien dans le cas de notre problème, il faudrait étudier la seconde variation de $\mathcal{E}$, par exemple en calculant la condition faible de Legendre. Cette étude est laissée en perspective. Nous choisissons plutôt de nous concentrer, dans le prochain paragraphe, sur l'utilisation de schémas numériques efficaces pour résoudre le problème (7.14). Nous allons voir qu'une approche fondée sur un algorithme de point fixe permet d'éliminer les non linéarités.

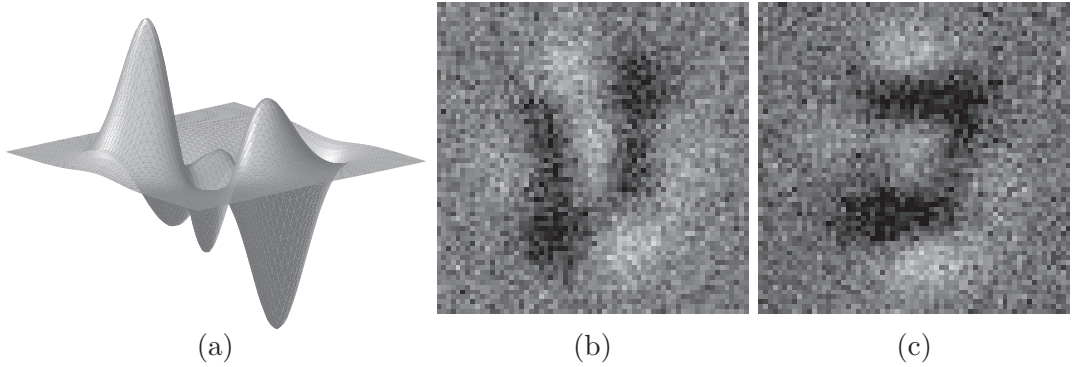


FIGURE 7.1 – (a) Surface *peaks* de Matlab, de fonction de profondeur z_p . (b)-(c) Deux images (parmi $m = 50$), de taille 64×64 , de cette surface vue en perspective ($f = -1000$). Pour simuler les images, nous avons fixé $\rho \equiv 255$ et choisi aléatoirement $m = 50$ directions d'éclairage uniformément réparties sur l'hémisphère nord de la sphère unité. Nous avons ensuite ajouté aux images un bruit gaussien, de moyenne nulle et d'écart-type σ égal à 10% de la valeur maximale $\|\mathbf{i}\|_\infty = 255$ des niveaux de gris.

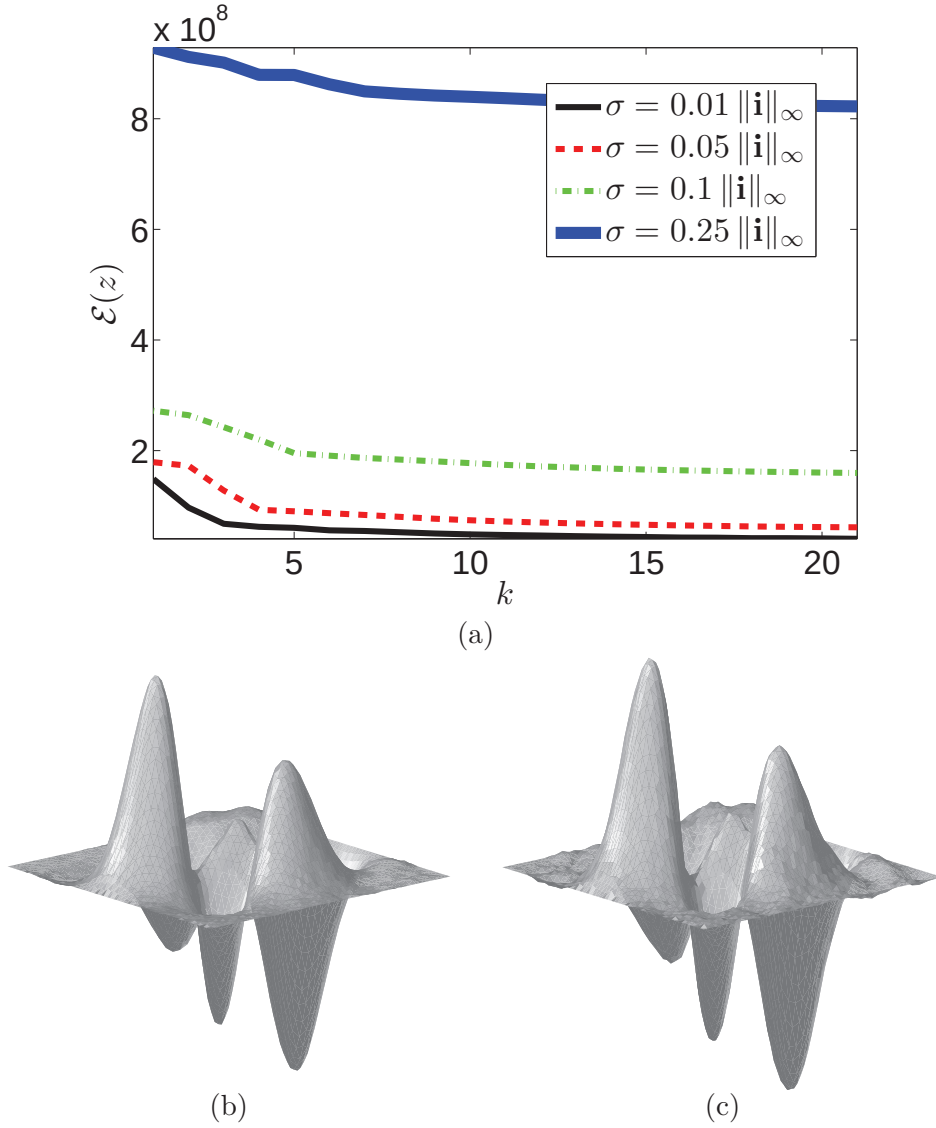


FIGURE 7.2 – (a) Évolution de l'énergie $\mathcal{E}(z)$ (cf. (7.15)) au cours des itérations de la descente (BFGS), pour plusieurs valeurs de l'écart-type σ du bruit gaussien sur les images, en appliquant la stéréophotométrie en moindres carrés aux images de la figure 7.1, avec une condition au bord de type Dirichlet homogène. Dans chaque test, la descente converge très rapidement. (b)-(c) Reliefs estimés à la convergence de la descente, pour $\sigma = 0,1\|\mathbf{i}\|_\infty$ et $\sigma = 0,25\|\mathbf{i}\|_\infty$. La reconstruction variationnelle en moindres carrés s'avère très robuste au bruit : elle est optimale au sens du maximum de vraisemblance, contrairement aux approches de la littérature.

7.1.6 Schéma de type point fixe

Les difficultés du problème (7.14) étant principalement liées à la division par d_z dans $\mathbf{A}_z$ et $\mathbf{b}_z$, nous proposons une résolution itérative où, à l'itération $k + 1$, d_z est fixé à la valeur estimée à l'itération k i.e., à d_{z^k} . Ceci se traduit par le schéma⁴ semi-implicite (point fixe) décrit dans l'algorithme 3 où, pourvu que la solution initiale se trouve dans un « bon » bassin de convergence, la suite $\{z^k\}_k$ converge vers la solution du problème en moindres carrés (7.14).

Algorithme 3: Itérations de point fixe pour la résolution du problème (7.14). La mise à jour (7.41) est effectuée par l'algorithme 4. La suite $\{z^k\}_k$ converge vers la solution du problème en moindres carrés (7.14).

$$d^k(u, v) = \sqrt{\|\nabla z^k(u, v)\|^2 + \left(\frac{z^k(u, v)}{f} + \left[\frac{u - u_0}{f}, \frac{v - v_0}{f} \right]^\top \cdot \nabla z^k(u, v) \right)^2} \quad (7.38)$$

$$\mathbf{A}^k(u, v) = \frac{\rho(u, v)}{d^k(u, v)} \begin{bmatrix} -s_1^1 + s_3^1 \frac{u - u_0}{f} & -s_2^1 + s_3^1 \frac{v - v_0}{f} \\ \vdots & \vdots \\ -s_1^m + s_3^m \frac{u - u_0}{f} & -s_2^m + s_3^m \frac{v - v_0}{f} \end{bmatrix} \quad (7.39)$$

$$\mathbf{b}^k(u, v) = \begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix} - \frac{\rho(u, v) z^k(u, v)}{f d^k(u, v)} \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \quad (7.40)$$

$$z^{k+1} = \underset{z}{\operatorname{argmin}} \sum_{(u, v) \in \Omega} \|\mathbf{A}^k(u, v) \nabla z(u, v) - \mathbf{b}^k(u, v)\|^2 + \lambda [z(u, v) - z_0(u, v)]^2 \quad (7.41)$$

L'intérêt de ce schéma de type point fixe est qu'il remplace la résolution d'un problème non linéaire par celle d'une suite de problèmes linéaires. En effet, puisque $\mathbf{A}^k$ et $\mathbf{b}^k$ sont indépendants de z dans (7.41), la condition d'optimalité associée à ce problème est linéaire. De plus, la prise en compte de la condition au bord non linéaire (7.36) peut être évitée en découpant à nouveau le problème (7.41) en deux problèmes plus simples : une estimation du gradient point par point, et l'intégration de ce gradient. Pour cela, nous introduisons la variable auxiliaire $\mathbf{g}$, et écrivons le problème (7.41) comme un problème

4. Dans l'ensemble de cette partie, nous notons, abusivement, de la même façon l'opérateur gradient ∇ dans les problèmes variationnels, et son approximation par différences finies avant à l'ordre 1 dans les schémas numériques. Le même symbole Ω est également utilisé pour décrire un ensemble ouvert dans les problèmes variationnels et le domaine de reconstruction discret dans les schémas numériques. Enfin, la notation z décrit tantôt une fonction réelle sur le domaine continu Ω , tantôt l'ensemble des valeurs prises par cette fonction sur le domaine de reconstruction discret Ω .

d'optimisation sous contrainte :

$$\begin{cases} (z^{k+1}, \mathbf{g}^{k+1}) = \underset{z, \mathbf{g}}{\operatorname{argmin}} \sum_{(u,v) \in \Omega} \|\mathbf{A}^k(u, v) \mathbf{g}(u, v) - \mathbf{b}^k(u, v)\|^2 + \lambda [z(u, v) - z_0(u, v)]^2 \\ \text{s.c. } \mathbf{g} = \nabla z \end{cases} \quad (7.42)$$

La solution de (7.42) peut alors être calculée par un schéma ADMM⁵, comme cela est indiqué dans l'algorithme 4.

Algorithme 4: Itérations de type ADMM pour résoudre le problème (7.42). La mise à jour (7.43) est donnée explicitement par (7.46), tandis que (7.44) est un problème d'intégration (cf. chapitre 2). La suite $\{z^{k,l}\}_l$ converge vers la solution z^{k+1} du problème en moindres carrés (7.41).

$$\mathbf{g}^{k,l+1}(u, v) = \underset{\mathbf{g}}{\operatorname{argmin}} \|\mathbf{A}^k(u, v) \mathbf{g} - \mathbf{b}^k(u, v)\|^2 + \alpha \|\mathbf{g} - (\nabla z^{k,l}(u, v) - \mathbf{e}^{k,l}(u, v))\|^2 \quad (7.43)$$

$$\begin{aligned} z^{k,l+1} = \underset{z}{\operatorname{argmin}} \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 \\ + \alpha \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - (\mathbf{g}^{k,l+1}(u, v) + \mathbf{e}^{k,l}(u, v)) \right\|^2 \end{aligned} \quad (7.44)$$

$$\mathbf{e}^{k,l+1}(u, v) = \mathbf{e}^{k,l}(u, v) + \mathbf{g}^{k,l+1}(u, v) - \nabla z^{k,l+1}(u, v) \quad (7.45)$$

La mise à jour (7.43) s'écrit sous forme explicite :

$$\begin{aligned} \mathbf{g}^{k,l+1}(u, v) = \underbrace{(\mathbf{A}^k(u, v)^\top \mathbf{A}^k(u, v) + \alpha \mathbf{I}_2)^{-1}}_{\mathbf{B}^k(u, v)} \left[\mathbf{A}^k(u, v)^\top \mathbf{b}^k(u, v) \right. \\ \left. + \alpha (\nabla z^{k,l}(u, v) - \mathbf{e}^{k,l}(u, v)) \right] \end{aligned} \quad (7.46)$$

où la matrice $\mathbf{B}^k(u, v)$ peut être déterminée analytiquement, puisqu'elle est l'inverse d'une matrice 2×2 .

Le problème (7.44) est un problème d'intégration, dont nous avons discuté la résolution dans le chapitre 2. La matrice du système linéaire correspondant à ce problème d'intégration étant la même à chaque itération l , on a intérêt à choisir la résolution par gradient conjugué préconditionné.

Contrairement à l'approche présentée dans le paragraphe 7.1.5, le seul moment où une condition au bord doit être prise en compte est celui de la résolution du problème

5. On rappelle que ce schéma converge avec un taux ergodique pour toute valeur du paramètre α [235].

d'intégration. En réalité, comme cela a été discuté dans le chapitre 2, on peut ne pas utiliser explicitement de condition au bord pour résoudre le sous-problème de l'intégration, en écrivant la condition d'optimalité associée au problème *discret* (7.44)⁶, et en appliquant la méthode présentée dans le paragraphe 2.3.3.

7.1.7 Cas où l'albédo est inconnu

Si l'albédo est connu, l'algorithme 3 converge vers une solution z du problème en moindres carrés (7.14). Pour estimer simultanément l'albédo et la normale, il faut remplacer (7.14) par :

$$\min_{z, \rho} \|\mathbf{A}_{z, \rho} \nabla z - \mathbf{b}_{z, \rho}\|_{L^2(\Omega)}^2 + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \quad (7.47)$$

où $\mathbf{A}_{z, \rho}$ et $\mathbf{b}_{z, \rho}$ sont deux champs, respectivement matriciel et vectoriel, qui dépendent de l'albédo inconnu :

$$\begin{cases} \mathbf{A}_{z, \rho}(u, v) = \frac{\rho(u, v)}{d_z(u, v)} \begin{bmatrix} -s_1^1 + \frac{s_3^1(u-u_0)}{f} & -s_2^1 + \frac{s_3^1(v-v_0)}{f} \\ \vdots & \vdots \\ -s_1^m + \frac{s_3^m(u-u_0)}{f} & -s_2^m + \frac{s_3^m(v-v_0)}{f} \end{bmatrix} \\ \mathbf{b}_{z, \rho}(u, v) = \begin{bmatrix} I^1(u, v) \\ \vdots \\ I^m(u, v) \end{bmatrix} - \frac{\rho(u, v)z(u, v)}{f d_z(u, v)} \begin{bmatrix} s_3^1 \\ \vdots \\ s_3^m \end{bmatrix} \end{cases} \quad (7.48)$$

Le problème (7.47) est convexe vis-à-vis de ρ : pour résoudre (7.47), nous pouvons donc utiliser un schéma d'optimisation à direction de descente bi-alternée (voir e.g., le paragraphe 9.2 dans [34]). Ceci revient à alterner des phases d'estimation de ρ à z fixé, et de z à ρ fixé. Lorsque z est fixé, ρ est aisément déterminé à partir de la condition nécessaire⁷ d'optimalité en ρ associée à (7.47). Il suffit ainsi d'ajouter, à la fin de l'algorithme 3, l'étape suivante :

$$\rho^{k+1}(u, v) = d^k(u, v) \frac{\sum_{i=1}^m I^i(u, v) \mathbf{s}^i \cdot \left[\frac{z^{k+1}(u, v)}{f} + \left[\frac{u-u_0}{f}, \frac{v-v_0}{f} \right]^\top \cdot \nabla z^{k+1}(u, v) \right]}{\sum_{i=1}^m \left(\mathbf{s}^i \cdot \left[\frac{z^{k+1}(u, v)}{f} + \left[\frac{u-u_0}{f}, \frac{v-v_0}{f} \right]^\top \cdot \nabla z^{k+1}(u, v) \right] \right)^2} \quad (7.49)$$

Pour valider expérimentalement le schéma numérique proposé, nous l'appliquons au jeu de données *Beethoven* (cf. figure 7.3). Les paramètres internes de l'appareil photographique n'étant pas connus, nous avons arbitrairement fixé le point principal au centre de l'image, et la focale à $f = -2000$ (l'unité étant le pixel). Nous avons également fixé

6. Ceci peut s'interpréter en termes de discrétisation par *éléments* finis, plutôt que par *différences* finies.

7. Cette condition est ici suffisante, le problème (7.47) étant convexe vis-à-vis de ρ .

$z_0 = 2f$, $\lambda = 10^{-9}$ (faible importance donnée à l'a priori sur la profondeur), $\alpha = 0,1$ (paramètre ADMM), et avons choisi comme critère d'arrêt pour les algorithmes itératifs un écart relatif moyen sur z , entre deux itérations successives, inférieur à 10^{-3} . Dans cette expérience, les sous-itérations ADMM convergent au plus en $l = 3$ itérations, et la convergence globale est atteinte en $k = 4$ itérations de point fixe (moins d'une seconde sur un processeur I7 récent).

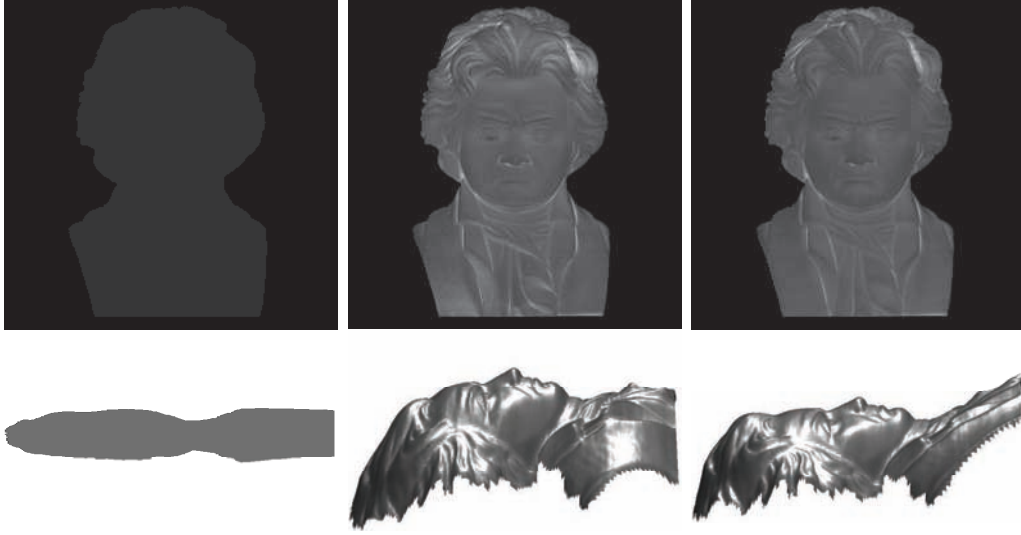


FIGURE 7.3 – Estimations de l'albédo (première ligne) et du relief (deuxième ligne), aux itérations $k = 0$, $k = 2$ et $k = 4$ (convergence) de la variante de l'algorithme 3 appliquée aux trois images du buste de Beethoven de la figure 1.2. L'approche variationnelle en moindres carrés proposée fournit des résultats comparables à ceux obtenus avec l'approche classique (cf. figures 1.2-e et 1.2-g).

Le schéma numérique proposé est donc très rapide, et fournit des résultats qualitativement comparables à l'approche « classique », qui consiste à d'abord estimer les normales et l'albédo en moindres carrés, puis à intégrer (toujours en moindres carrés) le champ de normales estimé. Avec cette approche, la donnée du problème d'intégration est le résultat de l'étape d'estimation des normales, qui peut être biaisée. L'approche proposée, qui ne minimise que l'erreur de reprojection, permet d'éviter toute propagation de biais inhérente aux algorithmes en plusieurs étapes. Outre qu'elle garantit « l'optimalité » des estimations de la profondeur et de l'albédo, un avantage important de l'approche variationnelle est sa flexibilité. En effet, nous allons voir dans le prochain paragraphe qu'il est aisé d'étendre la formulation en moindres carrés pour améliorer la robustesse de l'estimation. De plus, on peut garantir la régularité « par morceaux » du relief et de l'albédo, en cherchant ces fonctions dans l'espace des fonctions à variations bornées $BV(\Omega)$, plutôt que dans l'espace de Sobolev $\mathcal{W}^{1,2}(\Omega)$.

7.2 Estimation robuste du relief et de l'albédo

Conformément au raisonnement que nous avons mené dans le chapitre 1, nous suggérons pour garantir une certaine robustesse :

- de modéliser explicitement les ombres propres au moyen d'une approximation différentiable de l'opérateur max (cf. paragraphe 1.3.5) ;
- d'estimer l'éclairage « ambient » (cf. paragraphe 1.3.1) ;
- de remplacer les moindres carrés par un estimateur robuste, e.g. la norme 1 : robustesse aux ombres portées et aux reflets brillants (cf. paragraphe 1.3.4), ainsi qu'aux discontinuités de profondeur (cf. paragraphe 2.4.2).

En procédant de la sorte, nous souhaitons parvenir à reformuler le problème d'estimation robuste du relief et de l'albédo présenté dans la partie I comme un seul et unique problème variationnel. Les outils introduits pour l'estimation robuste des normales et de l'albédo (prise en compte des ombres propres et de l'éclairage ambient, et estimation robuste par la norme 1), présentés dans le chapitre 1, doivent être associés aux méthodes d'intégration qui ont été présentées dans le chapitre 2.

Nous allons voir que l'approche variationnelle permet d'aller encore plus loin, puisqu'il est naturel d'y introduire des connaissances a priori sur la régularité spatiale des inconnues, en utilisant une régularisation appropriée.

7.2.1 Prise en compte explicite des ombres propres

Rappelons que le modèle lambertien linéaire $I = \rho \mathbf{s} \cdot \mathbf{n}$ n'est valide que si l'angle entre $\mathbf{s}$ et $\mathbf{n}$ est inférieur à $\pi/2$: lorsque cela n'est pas le cas, il est nécessaire d'utiliser la formulation non linéaire $I = \max\{0, \rho \mathbf{s} \cdot \mathbf{n}\}$ i.e., de modéliser les ombres propres. Dans la majorité des travaux traitant des estimateurs robustes en stéréophotométrie [273, 270, 131], les ombres propres sont traitées comme des données aberrantes. Or, nous avons vu dans le paragraphe 1.3.5 qu'il était possible de tenir explicitement compte des ombres propres, au moyen d'une approximation différentiable de l'opérateur max. Cette approximation peut être utilisée dans le modèle variationnel, en remplaçant le système d'EDP quasi-linéaires (7.9) par le système d'EDP non linéaires suivant :

$$I^i(u, v) = \rho(u, v) \underbrace{M\left(\tilde{\mathbf{c}}_z^i(u, v) \cdot \nabla z(u, v) + c_{z,3}^i(u, v)\right)}_{M_z^i(u, v)}, \forall i \in [1, m] \quad (7.50)$$

où M est l'approximation différentiable de l'opérateur max :

$$M(x) = \frac{\ln(1 + \exp(\delta x))}{\delta} \quad (7.51)$$

et :

$$\begin{cases} \tilde{\mathbf{c}}_z^i(u, v) = \frac{1}{d_z(u, v)} \begin{bmatrix} -s_1^i + \frac{s_3^1(u-u_0)}{f} \\ -s_2^i + \frac{s_3^i(v-v_0)}{f} \end{bmatrix} \\ c_{z,3}^i(u, v) = \frac{s_3^i z(u, v)}{f d_z(u, v)} \end{cases} \quad (7.52)$$

Le problème d'estimation en moindres carrés de z et ρ s'écrit donc :

$$\min_{z, \rho} \sum_{i=1}^m \left\| I^i - \rho M_z^i \right\|_{L^2(\Omega)}^2 + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \quad (7.53)$$

où M_z^i est défini en (7.50).

Le calcul de la première variation de l'énergie à minimiser dans le problème non linéaire (7.53) est sensiblement le même que celui mené dans le paragraphe 7.1.5. Il suffit en effet de remplacer la différentielle de $L^i(\mathbf{n}_z)$, définie en (7.29), par :

$$d_{\mathbf{n}_z} L^i(\mathbf{n}_z) = -2\rho \left(I^i - \rho M_z^i \right) \frac{\exp(\delta \mathbf{s}^i \cdot \mathbf{n}_z)}{1 + \exp(\delta \mathbf{s}^i \cdot \mathbf{n}_z)} \mathbf{s}^i \quad (7.54)$$

ce qui conduit finalement à la condition nécessaire d'optimalité suivante :

$$\begin{cases} \nabla \cdot (\bar{f}_z \nabla z) - \lambda z = \bar{h}_z - \nabla \cdot \bar{\mathbf{g}}_z - \lambda z_0 & \text{sur } \Omega \\ z = z_0 & \text{sur } \partial\Omega \end{cases} \quad (7.55)$$

où les notations suivantes sont utilisées :

$$\bar{f}_z = \sum_{i=1}^m \frac{\tilde{I}^i (I^i - \bar{I}^i) \text{Sigm}(\delta \mathbf{s}^i \cdot \mathbf{n}_z)}{d_z^2} \quad (7.56)$$

$$\bar{\mathbf{g}}_z = \sum_{i=1}^m \frac{(I^i - \bar{I}^i) \text{Sigm}(\delta \mathbf{s}^i \cdot \mathbf{n}_z)}{d_z} \left(\left(\frac{\tilde{I}^i T_z}{f d_z} - \frac{\rho s_3^i}{f} \right) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \rho \begin{bmatrix} s_1^i \\ s_2^i \end{bmatrix} \right) \quad (7.57)$$

$$\bar{h}_z = \sum_{i=1}^m \frac{(I^i - \bar{I}^i) \text{Sigm}(\delta \mathbf{s}^i \cdot \mathbf{n}_z)}{f d_z} \left(\frac{\tilde{I}^i T_z}{d_z} + \rho s_3^i \right) \quad (7.58)$$

où :

$$\bar{I}^i = \rho M_z^i \quad (7.59)$$

$$\tilde{I}^i = \rho \mathbf{s}^i \cdot \mathbf{n}_z \quad (7.60)$$

sont les images reprojetées, respectivement, avec ou sans prise en compte des ombres propres, et où :

$$\text{Sigm}(x) = \frac{\exp(x)}{1 + \exp(x)} \quad (7.61)$$

est une fonction sigmoïde qui pondère l'influence des données susceptibles de correspondre à des ombres propres.

Une première solution pour résoudre l'EDP non linéaire (7.55) consiste à utiliser une descente de type quasi-Newton (BFGS). Comme cela est illustré sur la figure 7.4, cette approche montre l'intérêt de la prise en compte des ombres propres : la valeur de l'énergie est considérablement diminuée par rapport à l'approche présentée dans le paragraphe 7.1.5, ce qui indique que les images sont « mieux expliquées ». La descente est cependant relativement lente, et repose de plus sur une condition au bord de type Dirichlet. Pour surmonter ces difficultés, nous allons de nouveau recourir à un algorithme de point fixe.

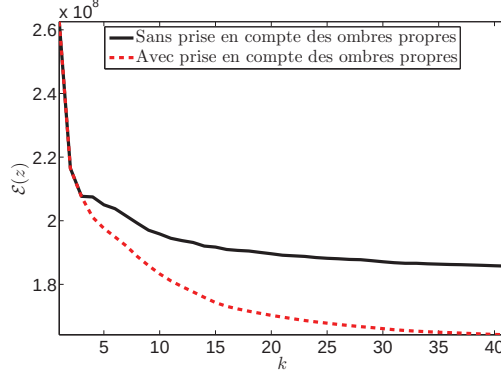


FIGURE 7.4 – Évolution de l’erreur de reprojection au cours des itérations de la descente (BFGS), sur le même exemple que celui de la figure 7.1 (avec $\sigma = 0, 1 \|\mathbf{i}\|_\infty$), avec ou sans prise en compte des ombres propres. Les images sont « mieux expliquées » en tenant compte des ombres propres.

7.2.2 Modèle L^1 -TV

Pour améliorer la robustesse, il nous faut tenir compte du caractère exogène du bruit (éclairage « ambient »), ainsi que des données aberrantes (ombres portées et reflets brillants). En présence d’un bruit très important sur les images, il peut être également intéressant d’introduire une régularisation spatiale des champs estimés. Nous allons pour cela étudier une variante régularisée du problème (7.53), fondée sur la résolution du problème variationnel suivant⁸ :

$$\min_{z, \rho, \phi_a} \underbrace{\frac{1}{m} \sum_{i=1}^m \left\| \underbrace{I^i - \underbrace{\phi_a}_{\text{Éclairage ambient}} - \underbrace{\rho M(\tilde{\mathbf{c}}_z^i \cdot \nabla z + c_{z,3}^i)}_{\text{Ombres propres}} \right\|_{L^1(\Omega)}}_{\text{Attache aux données}} + \underbrace{\lambda \|z - z_0\|_{L^2(\Omega)}^2}_{\text{Connaissance a priori}} + \underbrace{\gamma_z \text{TV}(z) + \gamma_\rho \text{TV}(\rho)}_{\text{Régularisations}} \quad (7.62)$$

où nous rappelons que la variation totale d’une fonction $g \in BV(\Omega)$ peut être définie par :

$$\text{TV}(g) = \iint_{(u,v) \in \Omega} \|\nabla g(u, v)\| du dv \quad (7.63)$$

Le problème (7.62) garantit la robustesse de l’estimation du relief et de l’albédo grâce à la prise en compte explicite de l’éclairage ambient et des ombres propres, couplée à l’introduction d’une connaissance a priori permettant d’éviter les dérives, à des termes

8. Nous moyennons dorénavant l’attache aux données, afin que le choix des paramètres λ , γ_z et γ_ρ soit indépendant du nombre m d’images.

de régularisation TV garantissant un aspect lisse par morceaux, et à un terme d'attache aux données L^1 offrant une certaine robustesse aux données aberrantes (ombres portées, reflets brillants, discontinuités de profondeur). Ceci n'est bien entendu possible qu'au prix de l'introduction de paramètres (δ , λ , γ_z , et γ_ρ), mais nous verrons dans le paragraphe 7.2.7 qu'aucun de ces paramètres n'est critique.

Nous pouvons donner une interprétation bayésienne au problème (7.62), en utilisant un raisonnement inspiré des travaux de Mumford [183]. Pour cela, nous proposons d'estimer (z, ρ, ϕ_a) comme le maximum a posteriori de la distribution :

$$\mathcal{P}(z, \rho, \phi_a | \mathbf{i}) = \frac{\mathcal{P}(\mathbf{i} | z, \rho, \phi_a) \mathcal{P}(z, \rho, \phi_a)}{\mathcal{P}(\mathbf{i})} \quad (7.64)$$

où $\mathcal{P}(\mathbf{i})$ est la probabilité marginale de $\mathbf{i}$, qui peut être négligée puisqu'elle est constante et ne joue donc aucun rôle dans l'estimation par maximum a posteriori (MAP). La probabilité a priori $\mathcal{P}(z, \rho, \phi_a)$ se factorise en $\mathcal{P}(z, \rho, \phi_a) = \mathcal{P}(\phi_a | z, \rho) \mathcal{P}(z, \rho)$. Le relief de la surface étant indépendant du matériau dont elle est constituée⁹, on peut écrire :

$$\mathcal{P}(z, \rho) = \mathcal{P}(z) \mathcal{P}(\rho) \quad (7.65)$$

Contrairement à z et à ρ , ϕ_a n'a pas de signification physique : cette variable représente seulement l'espérance du bruit de mesure en chaque pixel. On peut donc, pour simplifier, supposer qu'elle est indépendante de z et de ρ ¹⁰, de telle sorte que $\mathcal{P}(\phi_a | z, \rho) = \mathcal{P}(\phi_a)$, et que finalement :

$$\mathcal{P}(z, \rho, \phi_a) = \mathcal{P}(z) \mathcal{P}(\rho) \mathcal{P}(\phi_a) \quad (7.66)$$

En calculant l'opposé du logarithme de l'expression (7.64), la maximisation de la probabilité a posteriori est équivalente à la minimisation de :

$$\mathcal{E}(z, \rho, \phi_a) = -\ln \mathcal{P}(\mathbf{i} | z, \rho, \phi_a) - \ln \mathcal{P}(z) - \ln \mathcal{P}(\rho) - \ln \mathcal{P}(\phi_a) \quad (7.67)$$

Le modèle $I^i = \phi_a + \max\{0, \rho \mathbf{s}^i \cdot \mathbf{n}\}$, $i \in [1, m]$, n'est bien sûr pas exactement vérifié avec des images réelles : le bruit inhérent à l'acquisition des images empêche l'égalité. Nous supposons dans la suite que les résidus en chaque pixel sont indépendants et identiquement distribués selon une loi de Laplace centrée en ϕ_a . Ceci permet d'autoriser certaines données aberrantes telles que les ombres portées ou les reflets brillants. La vraisemblance est donc proportionnelle à¹¹ :

$$\mathcal{P}(\mathbf{i} | z, \rho, \phi_a) \propto \exp \left\{ -\frac{1}{m} \sum_{i=1}^m \left\| I^i - \phi_a - \rho M \left(\tilde{\mathbf{c}}_z^i \cdot \nabla z + c_{z,3}^i \right) \right\|_{L^1(\Omega)} \right\} \quad (7.68)$$

9. Ceci peut sembler évident, mais il s'agit bien d'une des clés permettant de dépasser un certain nombre de limites de l'approche classique de la stéréophotométrie.

10. Cela n'est pas le cas en réalité, car ϕ_a modélise, entre autres choses, les réflexions mutuelles provoquées par les concavités de la surface. Cependant, dans notre cas, cette variable ne joue qu'un rôle de variable « cachée » : il est donc assez raisonnable de la découpler des autres variables.

11. Nous négligeons les constantes de normalisation dans les probabilités, puisqu'elles ne jouent pas de rôle dans l'estimation MAP.

CHAPITRE 7. RÉSOLUTION VARIATIONNELLE DE LA STÉRÉOPHOTOMÉTRIE DIFFÉRENTIELLE

La surface observée peut présenter des arêtes, voire des discontinuités de profondeur. On connaît de plus une approximation z_0 de la profondeur. Il est donc naturel de choisir comme a priori sur la profondeur une régularisation de type variation totale (TV), couplée à un a priori gaussien z_0 :

$$\mathcal{P}(z) \propto \exp \left\{ - \left[\gamma_z \text{TV}(z) + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \right] \right\} \quad (7.69)$$

Nous supposons également que l'albédo est lisse par morceaux (texture de type « Arlequin »), ce qui est un a priori souvent rencontré sous la forme d'une contrainte sur l'entropie de sa distribution [9], voire sur le caractère fini de l'ensemble des valeurs qu'il peut prendre [11]. Le caractère lisse par morceaux correspond également à un a priori TV :

$$\mathcal{P}(\rho) \propto \exp \{ -\gamma_\rho \text{TV}(\rho) \} \quad (7.70)$$

Enfin, ϕ_a est une variable auxiliaire qui compense le caractère endogène du bruit. On peut difficilement faire une hypothèse a priori sur cette variable, donc nous choisissons un a priori impropre (*improper prior*) :

$$\mathcal{P}(\phi_a) = \text{constante} \quad (7.71)$$

D'après (7.67), (7.68), (7.69), (7.70) et (7.71), l'estimateur du maximum a posteriori de (z, ρ, ϕ_a) est effectivement la solution du problème L^1 -TV (7.62).

Bien qu'il semble offrir une solution optimale en termes de robustesse, le problème (7.62) peut paraître difficile à résoudre de prime abord. En réalité, la stratégie consistant à « diviser pour mieux régner », qui vise à s'approcher itérativement de la solution de (7.62) en séparant l'estimation en sous-problèmes plus faciles à résoudre, peut être mise en place assez aisément. Nous proposons pour cela une combinaison d'itérations ADMM et de type point fixe, schématisée par les quatre étapes suivantes :

1. Le problème L^1 -TV (7.62) est d'abord transformé en un problème L^1 - L^2 par ADMM (cf. paragraphe 7.2.3).
2. Ce problème L^1 - L^2 est ensuite lui-même transformé en un problème L^2 - L^2 , toujours par ADMM (cf. paragraphe 7.2.4).
3. Ce problème L^2 - L^2 est à nouveau séparé en trois sous-problèmes, où les estimations de z , ρ et ϕ_a sont découplées. Les mises à jour de ρ et de ϕ_a sont alors effectuées explicitement (cf. paragraphe 7.2.5).
4. Le problème L^2 - L^2 non linéaire d'estimation de z est finalement séparé, via un algorithme de point fixe couplé à des itérations ADMM, en un problème d'estimation non linéaire à deux paramètres, que l'on peut résoudre par la méthode de Levenberg-Marquardt, et un problème quadratique fournissant une mise à jour explicite de z (cf. paragraphe 7.2.6).

7.2.3 Itérations ADMM pour le problème L^1 -TV (7.62)

Nous commençons par introduire des variables auxiliaires pour transformer la régularisation TV, qui est une régularisation globale, en une régularisation locale. Nous remplaçons pour cela le problème (7.62) en un problème sous contraintes¹² :

$$\left\{ \begin{array}{l} \min_{z, \rho, \phi_a} \frac{1}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left| I^i(u, v) - \phi_a(u, v) - \rho(u, v) M_z^i(u, v) \right| \\ \quad + \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 + \gamma_z \sum_{(u,v) \in \Omega} \|\mathbf{g}_z(u, v)\| + \gamma_\rho \sum_{(u,v) \in \Omega} \|\mathbf{g}_\rho(u, v)\| \\ \text{s.c.} \quad \begin{cases} \mathbf{g}_z = \nabla z \\ \mathbf{g}_\rho = \nabla \rho \end{cases} \end{array} \right. \quad (7.72)$$

La solution du problème (7.72) peut être estimée au moyen d'un schéma ADMM, consistant en la suite d'itérations suivante :

$$\mathbf{g}_z^{k+1}(u, v) = \min_{\mathbf{g}_z} \gamma_z \|\mathbf{g}_z\| + \alpha_z \|\mathbf{g}_z - \nabla z^k(u, v) + \mathbf{e}_z^k(u, v)\|, \forall (u, v) \in \Omega \quad (7.73)$$

$$\mathbf{g}_\rho^{k+1}(u, v) = \min_{\mathbf{g}_\rho} \gamma_\rho \|\mathbf{g}_\rho\| + \alpha_\rho \|\mathbf{g}_\rho - \nabla \rho^k(u, v) + \mathbf{e}_\rho^k(u, v)\|, \forall (u, v) \in \Omega \quad (7.74)$$

$$\begin{aligned} (z^{k+1}, \rho^{k+1}, \phi_a^{k+1}) = \operatorname{argmin}_{z, \rho, \phi_a} & \frac{1}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left| I^i(u, v) - \phi_a(u, v) - \rho(u, v) M_z^i(u, v) \right| \\ & + \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}_z^{k+1}(u, v) - \mathbf{e}_z^k(u, v) \right\|^2 \\ & + \alpha_\rho \sum_{(u,v) \in \Omega} \left\| \nabla \rho(u, v) - \mathbf{g}_\rho^{k+1}(u, v) - \mathbf{e}_\rho^k(u, v) \right\|^2 \end{aligned} \quad (7.75)$$

$$\mathbf{e}_z^{k+1}(u, v) = \mathbf{e}_z^k(u, v) + \mathbf{g}_z^{k+1}(u, v) - \nabla z^{k+1}(u, v), \forall (u, v) \in \Omega \quad (7.76)$$

$$\mathbf{e}_\rho^{k+1}(u, v) = \mathbf{e}_\rho^k(u, v) + \mathbf{g}_\rho^{k+1}(u, v) - \nabla \rho^{k+1}(u, v), \forall (u, v) \in \Omega \quad (7.77)$$

pour toutes valeurs positives de α_z et α_ρ , qui influencent principalement la vitesse de convergence.

En remarquant que la solution d'un problème de la forme :

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\| + \frac{1}{2\eta} \|\mathbf{x} - \mathbf{a}\|^2 \quad (7.78)$$

est donnée explicitement par l'opérateur de « seuillage doux » (*shrinkage*) :

$$\operatorname{Shr}_\eta(\mathbf{a}) = \max \left\{ 0, 1 - \frac{\eta}{\|\mathbf{a}\|} \right\} \mathbf{a} \quad (7.79)$$

dont l'effet est illustré sur la figure 7.5, le schéma de résolution est finalement donné par l'algorithme 5, où la difficulté principale consiste en un problème L^1 - L^2 global, dont nous détaillons la résolution dans les prochains paragraphes.

12. Dans cette partie consacrée aux schémas numériques, nous utiliserons les formulations discrètes des problèmes variationnels.

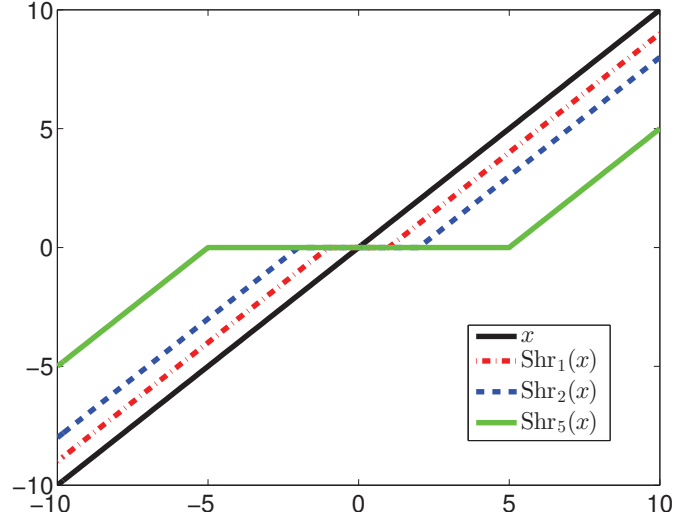


FIGURE 7.5 – Effet de l'opérateur de seuillage doux (7.79).

Algorithme 5: Schéma ADMM pour la résolution du problème L^1 -TV (7.72). Les mises à jour (7.80) et (7.81) sont réalisées par seuillage doux (cf. (7.79)). Le problème L^1 - L^2 (7.82) peut être résolu par l'algorithme 6. La suite $\{(z^k, \rho^k, \phi_a^k)\}_k$ converge vers la solution (z, ρ, ϕ_a) de (7.62) lorsque $k \rightarrow \infty$.

$$\mathbf{g}_z^{k+1}(u, v) = \text{Shr}_{\frac{\gamma_z}{2\alpha_z}} \left(\nabla z^k(u, v) - \mathbf{e}_z^k(u, v) \right) \quad (7.80)$$

$$\mathbf{g}_\rho^{k+1}(u, v) = \text{Shr}_{\frac{\gamma_\rho}{2\alpha_\rho}} \left(\nabla \rho^k(u, v) - \mathbf{e}_\rho^k(u, v) \right) \quad (7.81)$$

$$\begin{aligned} (z^{k+1}, \rho^{k+1}, \phi_a^{k+1}) = \operatorname{argmin}_{z, \rho, \phi_a} & \frac{1}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left| I^i(u, v) - \phi_a(u, v) - \rho(u, v) M_z^i(u, v) \right| \\ & + \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}_z^{k+1}(u, v) - \mathbf{e}_z^k(u, v) \right\|^2 \\ & + \alpha_\rho \sum_{(u,v) \in \Omega} \left\| \nabla \rho(u, v) - \mathbf{g}_\rho^{k+1}(u, v) - \mathbf{e}_\rho^k(u, v) \right\|^2 \end{aligned} \quad (7.82)$$

$$\mathbf{e}_z^{k+1}(u, v) = \mathbf{e}_z^k(u, v) + \mathbf{g}_z^{k+1}(u, v) - \nabla z^{k+1}(u, v) \quad (7.83)$$

$$\mathbf{e}_\rho^{k+1}(u, v) = \mathbf{e}_\rho^k(u, v) + \mathbf{g}_\rho^{k+1}(u, v) - \nabla \rho^{k+1}(u, v) \quad (7.84)$$

7.2.4 Itérations ADMM pour le problème L^1 - L^2 (7.82)

Le problème L^1 - L^2 (7.82) peut être simplifié en étant transformé en un problème d'optimisation sous contrainte :

$$\left\{ \begin{array}{l} (z^{k+1}, \rho^{k+1}, \phi_a^{k+1}, \mathbf{r}^{k+1}) = \underset{z, \rho, \phi_a}{\operatorname{argmin}} \frac{1}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} |r^i(u,v)| + \lambda \sum_{(u,v) \in \Omega} [z(u,v) - z_0(u,v)]^2 \\ \quad + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u,v) - \mathbf{g}_z^{k+1}(u,v) - \mathbf{e}_z^k(u,v) \right\|^2 \\ \quad + \alpha_\rho \sum_{(u,v) \in \Omega} \left\| \nabla \rho(u,v) - \mathbf{g}_\rho^{k+1}(u,v) - \mathbf{e}_\rho^k(u,v) \right\|^2 \\ \text{s.c. } r^i(u,v) = I^i(u,v) - \phi_a(u,v) - \rho(u,v)M_z^i(u,v) \end{array} \right. \quad (7.85)$$

Le problème (7.85) peut lui aussi être résolu par un schéma ADMM (cf. l'algorithme 6). Il suffit pour cela d'alterner des opérations de seuillage doux local (7.86) avec des estimations globales en moindres carrés (7.87), qui sont décrites dans le prochain paragraphe.

Algorithme 6: Schéma ADMM pour transformer le problème L^1 - L^2 (7.85) en un problème L^2 - L^2 (7.87). La suite $\{(z^{k,l}, \rho^{k,l}, \phi_a^{k,l})\}_l$ converge vers la solution $(z^{k+1}, \rho^{k+1}, \phi_a^{k+1})$ de (7.82) lorsque $l \rightarrow \infty$.

$$r^{i,l+1}(u,v) = \operatorname{Shr}_{\frac{1}{2m\alpha_{L^1}}} \left(I^i(u,v) - \phi_a^{k,l}(u,v) - \rho^{k,l}(u,v)M_{z^{k,l}}^i(u,v) \right) \quad (7.86)$$

$$\begin{aligned} (z^{k,l+1}, \rho^{k,l+1}, \phi_a^{k,l+1}) = \underset{z, \rho, \phi_a}{\operatorname{argmin}} & \lambda \sum_{(u,v) \in \Omega} [z(u,v) - z_0(u,v)]^2 \\ & + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u,v) - \mathbf{g}_z^{k+1}(u,v) - \mathbf{e}_z^k(u,v) \right\|^2 \\ & + \alpha_\rho \sum_{(u,v) \in \Omega} \left\| \nabla \rho(u,v) - \mathbf{g}_\rho^{k+1}(u,v) - \mathbf{e}_\rho^k(u,v) \right\|^2 \\ & + \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[r^{i,l+1}(u,v) - I^i(u,v) + \phi_a(u,v) + \rho(u,v)M_z^i(u,v) + s^{i,l}(u,v) \right]^2 \end{aligned} \quad (7.87)$$

$$s^{i,l+1}(u,v) = s^{i,l}(u,v) + r^{i,l+1}(u,v) - I^i(u,v) + \phi_a^{k,l+1}(u,v) + \rho^{k,l+1}(u,v)M_{z^{k,l+1}}^i(u,v) \quad (7.88)$$

7.2.5 Optimisation alternée pour le problème L^2 - L^2 (7.87)

Le problème (7.87) est convexe en ρ et ϕ_a . On peut donc utiliser un algorithme d'optimisation alternée, comme cela est décrit dans l'algorithme 7, où, pour alléger les notations, nous avons utilisé la variable auxiliaire :

$$t^{i,l}(u, v) = I^i(u, v) - r^{i,l+1}(u, v) - s^{i,l}(u, v) \quad (7.89)$$

Algorithme 7: Optimisation alternée pour la résolution du problème (7.87). Le problème quadratique (7.91) est résolu par gradient conjugué préconditionné [152], le problème (7.92) admet la solution analytique (7.93), et le problème (7.90) est résolu par l'algorithme 8. La suite $\{(z^{k,l,p+1}, \rho^{k,l,p+1}, \phi_a^{k,l,p+1})\}_p$ converge vers la solution $(z^{k,l+1}, \rho^{k,l+1}, \phi_a^{k,l+1})$ de (7.87) lorsque $p \rightarrow \infty$.

$$\begin{aligned} z^{k,l,p+1} = \operatorname{argmin}_z & \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 \\ & + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}_z^{k+1}(u, v) - \mathbf{e}_z^k(u, v) \right\|^2 \\ & + \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[\rho^{k,l,p}(u, v) M_z^i(u, v) - t^{i,l}(u, v) + \phi_a^{k,l,p}(u, v) \right]^2 \end{aligned} \quad (7.90)$$

$$\begin{aligned} \rho^{k,l,p+1} = \operatorname{argmin}_\rho & \alpha_\rho \sum_{(u,v) \in \Omega} \left\| \nabla \rho(u, v) - \mathbf{g}_\rho^{k+1}(u, v) - \mathbf{e}_\rho^k(u, v) \right\|^2 \\ & + \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[\rho M_{z^{k,l,p+1}}^i(u, v) - t^{i,l}(u, v) + \phi_a^{k,l,p}(u, v) \right]^2 \end{aligned} \quad (7.91)$$

$$\phi_a^{k,l,p+1} = \operatorname{argmin}_{\phi_a} \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[\phi_a - t^{i,l}(u, v) + \rho^{k,l,p+1}(u, v) M_{z^{k,l,p+1}}^i(u, v) \right]^2 \quad (7.92)$$

La condition d'optimalité associée au problème (7.92) d'estimation de l'éclairage ambiant admet la solution analytique suivante :

$$\phi_a^{k,l,p+1} = \frac{1}{m} \sum_{i=1}^m \left(t^{i,l} - \rho^{k,l,p+1} M_{z^{k,l,p+1}}^i \right) \quad (7.93)$$

Le problème (7.91) (estimation de l'albédo) est un problème quadratique similaire au problème de l'intégration en moindres carrés avec a priori à l'ordre 0 (cf. chapitre 2).

La condition d'optimalité associée s'écrit :

$$\Delta\rho - \left(\frac{\alpha_{L^1}}{m\alpha_\rho} \sum_{i=1}^m \left(M_{z^{k,l,p+1}}^i \right)^2 \right) \rho = \nabla \cdot \left(\mathbf{g}_\rho^{k+1} + \mathbf{e}_\rho^k \right) - \frac{\alpha_{L^1}}{m\alpha_\rho} \sum_{i=1}^m \left(t^{i,l} - \phi_a^{k,l,p} \right) M_{z^{k,l,p+1}}^i \quad (7.94)$$

avec une condition naturelle au bord, qui est ici une condition de type Neumann similaire à la condition naturelle de l'intégration.

La matrice du système linéaire résultant de la discrétisation de (7.94) est creuse, symétrique, définie positive. Le système peut donc être résolu efficacement par gradient conjugué préconditionné [152].

7.2.6 Itérations de point fixe pour le problème L^2 - L^2 non linéaire (7.90)

Afin de transformer le problème en moindres carrés non linéaires (7.90) en un problème en moindres carrés linéaires, nous utilisons un schéma de point fixe, qui est décrit dans l'algorithme 8.

Algorithme 8: Itérations de point fixe pour la résolution du problème (7.90). Les fonctions $\tilde{\mathbf{c}}_z^i$ et $c_{z,3}^i$ sont définies en (7.52), et la mise à jour (7.97) est détaillée dans l'algorithme 9.

$$\tilde{\mathbf{c}}^{i,q+1}(u, v) = \tilde{\mathbf{c}}_{z^{k,l,p,q}}^i(u, v) \quad (7.95)$$

$$c_3^{i,q+1}(u, v) = c_{z^{k,l,p,q},3}^i(u, v) \quad (7.96)$$

$$\begin{aligned} z^{k,l,p,q+1} = \operatorname{argmin}_z \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 \\ + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}_z^{k+1}(u, v) - \mathbf{e}_z^k(u, v) \right\|^2 \\ + \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[\rho^{k,l,p}(u, v) M \left(\tilde{\mathbf{c}}^{i,q+1}(u, v) \cdot \nabla z(u, v) + c_3^{i,q+1}(u, v) \right) \right. \\ \left. - t^{i,l}(u, v) + \phi_a^{k,l,p}(u, v) \right]^2 \end{aligned} \quad (7.97)$$

À chaque itération de point fixe, on peut enfin séparer une dernière fois le problème en deux sous-problèmes plus faciles à résoudre : un problème d'estimation de deux paramètres réels en moindres carrés non linéaires, et un problème d'intégration en moindres carrés d'un champ de gradient. Pour réaliser cette séparation, le problème (7.97) est transformé en un problème d'optimisation sous contrainte, qui est résolu au moyen d'un

nouveau (et dernier) schéma ADMM (cf. algorithme 9) :

$$\left\{ \begin{array}{l} \min_{z, \mathbf{g}} \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}_z^{k+1}(u, v) - \mathbf{e}_z^k(u, v) \right\|^2 \\ \quad + \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \sum_{(u,v) \in \Omega} \left[\rho^{k,l,p}(u, v) M \left(\tilde{\mathbf{c}}^{i,q+1}(u, v) \cdot \mathbf{g}(u, v) + c_3^{i,q+1}(u, v) \right) \right. \\ \quad \quad \left. - t^{i,l}(u, v) + \phi_a^{k,l,p}(u, v) \right]^2 \\ \text{s.c. } \mathbf{g} = \nabla z \end{array} \right. \quad (7.98)$$

Algorithme 9: Schéma ADMM pour la résolution du problème (7.98). La mise à jour (7.99) est réalisée par une descente de type Levenberg-Marquardt, et (7.100) par gradient conjugué préconditionné.

$$\begin{aligned} \mathbf{g}^{r+1}(u, v) = \operatorname{argmin}_{\mathbf{g}} \frac{\alpha_{L^1}}{m} \sum_{i=1}^m \left[\frac{\rho^{k,l,p}(u, v)}{\delta} \ln \left(1 + \exp \left\{ \delta \tilde{\mathbf{c}}^{i,q+1}(u, v) \cdot \mathbf{g} \right. \right. \right. \\ \left. \left. \left. + \delta c_3^{i,q+1}(u, v) \right\} \right) - t^{i,l}(u, v) + \phi_a^{k,l,p}(u, v) \right]^2 \\ + \alpha_{\mathbf{g}} \left\| \mathbf{g} - \nabla z(u, v) + \mathbf{s}^{r+1}(u, v) \right\|^2 \end{aligned} \quad (7.99)$$

$$\begin{aligned} z^{k,l,p,q,r+1} = \operatorname{argmin}_z \lambda \sum_{(u,v) \in \Omega} [z(u, v) - z_0(u, v)]^2 \\ + \alpha_z \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}_z^{k+1}(u, v) - \mathbf{e}_z^k(u, v) \right\|^2 \\ + \alpha_{\mathbf{g}} \sum_{(u,v) \in \Omega} \left\| \nabla z(u, v) - \mathbf{g}^{r+1}(u, v) - \mathbf{s}^{r+1}(u, v) \right\|^2 \end{aligned} \quad (7.100)$$

$$\mathbf{s}^{r+1}(u, v) = \mathbf{s}^r(u, v) + \mathbf{g}^{r+1}(u, v) - \nabla z^{k,l,p,q,r+1}(u, v) \quad (7.101)$$

Chaque sous-problème dans l'algorithme 9 peut enfin être résolu directement. Le problème (7.99) peut être résolu itérativement en chaque point par une descente de type Levenberg-Marquardt, tandis que le problème (7.100) est un problème quadratique dont la condition d'optimalité fournit un système linéaire de matrice symétrique, définie positive. Nous résolvons à nouveau ce système par la méthode du gradient conjugué préconditionné [152].

Remarquons qu'à chaque itération de l'algorithme 9, une estimation du gradient de profondeur est effectuée, puis ce gradient est intégré :

- L'estimation du gradient tient compte à la fois des données et du résultat des estimations précédentes : l'influence respective de ces deux termes est pondérée par les paramètres α_{L^1} et $\alpha_{\mathbf{g}}$.
- L'intégration est réalisée en combinant la connaissance a priori (pondérée par λ), l'estimation précédente (pondérée par $\alpha_{\mathbf{g}}$) et la régularisation TV (pondérée par α_z). Le terme $\mathbf{e}_z^k$ représente de plus la parcimonie de l'estimation, qui permet de retrouver les discontinuités de profondeur.

Remarquons que cette façon de procéder donne un cadre formel moins empirique à la méthode présentée dans [276], qui alterne une phase d'estimation locale du gradient (« *local shaping* ») et de propagation/intégration de cette estimation locale (« *global blending* »). Dans [276], le problème est formulé au moyen de concepts issus de la géométrie discrète (sommets, arêtes, etc.) : autant ce formalisme est intéressant pour le rendu, ou pour la reconstruction multi-vues, autant il est avantageusement remplacé par celui des EDP, dans le cas où le domaine de reconstruction est une grille discrète régulière.

La figure 7.6 confirme expérimentalement les bonnes propriétés de convergence du schéma numérique proposé, qui combine point fixe et itérations ADMM. De plus, comme cela est illustré sur la figure 7.7, la combinaison des différentes façons d'améliorer la robustesse fournit des résultats nettement supérieurs à l'estimation en moindres carrés.

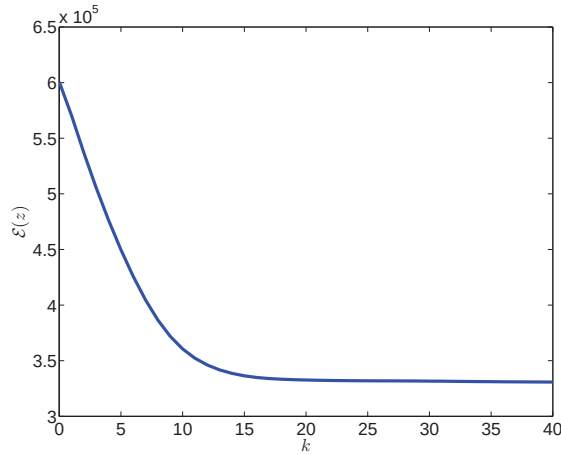


FIGURE 7.6 – Évolution de l'énergie du problème variationnel (7.62) au cours des itérations de point fixe, pour le test de la figure 7.7. Le schéma proposé semble converger à une vitesse raisonnable (les 40 itérations sont effectuées en deux minutes sur un processeur I7 récent).



FIGURE 7.7 – Vues de profil des reconstructions 3D variationnelles en moindres carrés (à gauche) et avec l’approche L^1 -TV proposée (à droite), à partir de $m = 141$ photographies d’une pelote de laine (cf. figure 2.13 : les $m = 141$ images sont choisies aléatoirement parmi les 2812 photographies du jeu de données utilisé dans [273]). Nous avons utilisé $\gamma_z = 10^{-7}$, $\gamma_\rho = 10^{-2}$, $f = -1000$, $z_0 \equiv 2f$ et $\lambda = 10^{-9}$. Ce résultat est comparable à celui de la figure 2.13, mais en utilisant vingt fois moins d’images ! La nouvelle approche, qui est itérative, est cependant beaucoup plus lente (deux minutes pour traiter ces 141 images, contre moins d’une minute pour traiter 2812 images avec l’approche classique).

7.2.7 Analyse des paramètres

L’algorithme 5, qui décrit la résolution du problème L^1 -TV (7.62), comporte un certain nombre de paramètres, qu’il peut sembler difficile d’ajuster. En réalité, aucun de ces paramètres n’est critique :

- γ_z, γ_ρ et sont des *hyper-paramètres* choisis par l’utilisateur, qui contrôlent les régularisations TV des champs scalaires z et ρ . Nous avons vu, dans la première partie de ce mémoire, que la résolution robuste de la stéréophotométrie était possible sans régularisation. La plupart du temps, ces paramètres peuvent donc être fixés à des valeurs très faibles, ce qui suffira à éliminer les artéfacts.
- λ est un autre hyper-paramètre, qui contrôle l’influence de la connaissance a priori z_0 . Si l’on ne dispose d’aucune connaissance a priori, on peut choisir un a priori arbitraire et fixer λ à une valeur très faible.
- $\alpha_z, \alpha_\rho, \alpha_{L^1}$ et α_g sont les paramètres des schémas ADMM. Ils n’ont d’influence que sur la vitesse de convergence. Dans nos tests, ils sont tous fixés à 0, 1. Le choix de paramètres permettant d’optimiser la vitesse de convergence est laissé en perspective, tout comme le remplacement des itérations ADMM par des algorithmes sans paramètre e.g., FISTA [27].
- δ est un paramètre qui contrôle la précision de l’approximation de l’opérateur max. Comme l’illustre la figure 1.25, la précision de l’approximation est d’autant meilleure que δ est plus grand. Cependant, une valeur trop élevée de δ induit des erreurs numériques (l’argument de l’exponentielle est proportionnel à δ). Nous avons expérimentalement constaté que $\delta = 5$ constituait un compromis « acceptable ».

Afin d'illustrer l'influence des hyper-paramètres de régularisation, ainsi que le gain de l'approche L^1 -TV par rapport aux moindres carrés, nous avons appliqué le schéma numérique proposé au jeu de données *Toys* (cf. figure 7.8) utilisé dans [273]¹³ :

- La scène à reconstruire est une boîte à jouet ouverte, qui génère beaucoup d'ombres portées.
- Le jouet est enveloppé dans un film plastique, ce qui provoque de nombreux reflets brillants.
- Les phénomènes de réflexions mutuelles sont très marqués.
- Le relief comporte de nombreuses discontinuités de profondeur.



FIGURE 7.8 – Première ligne : trois photographies (parmi $m = 27$), de taille 209×212 , d'une boîte à jouet [273]. Deuxième ligne : relief et albédo estimés par l'approche variationnelle en moindres carrés du paragraphe 7.1.7. Les pentes élevées sont aplaties, les discontinuités sont lissées et l'albédo contient beaucoup d'informations géométriques parasites.

13. Merci à Manohar Kuse et aux auteurs de [273] de nous avoir gracieusement fourni ces données.

Malgré ces difficultés, la reconstruction 3D obtenue avec notre algorithme L^1 -TV (cf. figure 7.9) est très satisfaisante. Elle est beaucoup plus réaliste que la reconstruction 3D en moindres carrés (cf. figure 7.8), et même que celle de [273]¹⁴, ce qui montre bien que l'approche variationnelle présente un réel intérêt par rapport à l'approche classique, ne serait-ce qu'en termes de robustesse. Comme le montre la figure 7.10, il est de plus possible de contrôler la régularité des estimations en choisissant convenablement les hyper-paramètres. Ces résultats nous montrent que les détails les plus fins sont perdus si γ_z est trop élevé. Une valeur faible de γ_z suffit à éliminer les éventuels artéfacts, tout en préservant les détails. En revanche, il faut choisir une valeur relativement élevée du paramètre γ_ρ pour garantir l'aspect uniforme par morceaux de l'albédo.



FIGURE 7.9 – Trois rééclairages du modèle 3D estimé par l'algorithme L^1 -TV, avec $(\gamma_z, \gamma_\rho) = (10^{-7}, 10^{-1})$. Même pour une scène complexe comme celle-ci, le résultat de l'approche préconisée peut être utilisé pour des applications de rééclairage.

14. Notons que nous n'avons utilisé que $m = 27$ images, choisis par échantillonnage uniforme des 3153 images fournies par les auteurs de [273]. Les résultats obtenus avec si peu d'images surpassent qualitativement ceux de [273], ce qui prouve la robustesse de l'approche L^1 -TV.

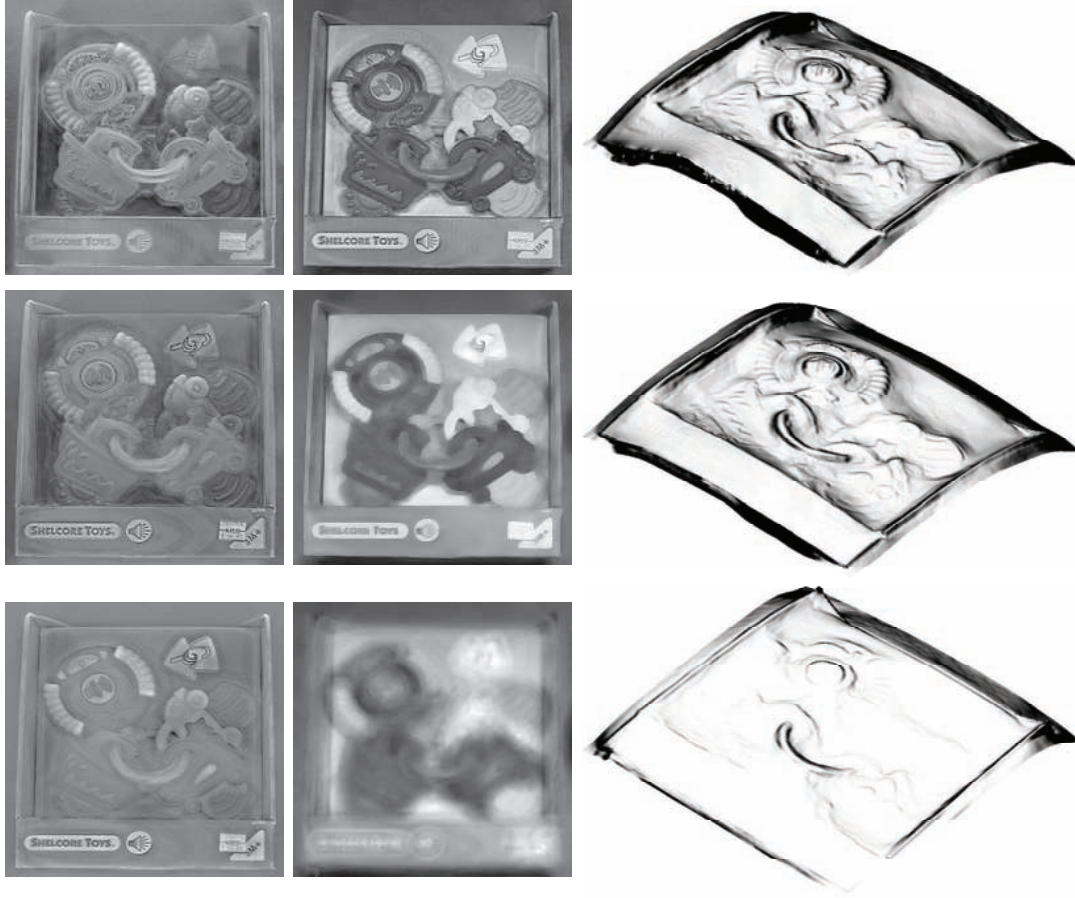


FIGURE 7.10 – De gauche à droite : estimations de l’éclairage ambiant, de l’albédo et du relief par l’approche L^1 -TV préconisée, en augmentant, de haut en bas, l’influence des régularisations TV de l’albédo et de la profondeur (respectivement, $(\gamma_z, \gamma_\rho) = (0, 0)$, $(\gamma_z, \gamma_\rho) = (10^{-4}, 10^{-1})$ et $(\gamma_z, \gamma_\rho) = (10^{-3}, 1)$, les autres paramètres étant constants). Les discontinuités sont bien retrouvées, et la reconstruction des plus fortes pentes est meilleure qu’avec l’approche en moindres carrés (figure 7.8), pourvu que l’hyperparamètre de régularisation de la profondeur γ_z ne soit pas trop élevé, ce qui tend à provoquer un sur-lissage. En augmentant la valeur de γ_ρ , on favorise de plus un albédo uniforme par morceaux, les variations étant absorbées par le relief et l’éclairage ambiant.

7.3 Conclusion du chapitre

Dans ce chapitre, nous avons proposé une méthode de résolution variationnelle du problème de la stéréophotométrie calibrée avec $m \geq 3$ éclairages directionnels, non coplanaires, connus. En particulier, nous avons introduit un cadre variationnel pour la résolution numérique, fondé sur la stratégie consistant à « diviser pour mieux régner »,

qui combine un algorithme de point fixe avec des itérations ADMM. Nous avons vu que cette stratégie pouvait être aisément étendue pour prendre en compte les diverses améliorations suggérées dans la partie I. Enfin, nous avons introduit une nouvelle méthode de résolution robuste du problème de la stéréophotométrie calibrée, fondée sur la minimisation du critère L^1 -TV (7.62).

Cette méthode répond à tous les critères de robustesse que nous avons évoqués dans la partie I :

- prise en compte de l'éclairage ambiant (estimation d'un champ additif $\phi_a : \Omega \rightarrow \mathbb{R}$ indépendant de i) ;
- prise en compte des ombres propres (approximation de l'opérateur max) ;
- prise en compte éventuelle d'une connaissance a priori sur la profondeur (régularisation à l'ordre 0) ;
- estimation locale robuste aux ombres portées et aux reflets brillants (estimation en norme 1) ;
- estimation globale robuste aux discontinuités de profondeur (estimation en norme 1) ;
- lissage des artefacts par régularisation (variation totale).

Grâce aux schémas ADMM et de type point fixe que nous avons introduits, la méthode présente de plus les avantages très appréciables suivants :

- Aucun des paramètres n'est critique.
- L'estimation est « assez rapide » : chaque sous-problème similaire au problème d'intégration présente une complexité en $O(|\Omega|\ln|\Omega|)$ (cf. paragraphe 2.3), et l'estimation locale en moindres carrés non linéaires ne porte que sur deux paramètres scalaires.

L'approche proposée fournit donc un cadre théorique et numérique original, qui permet de revisiter certains problèmes de stéréophotométrie connus pour être difficiles. Nous allons nous intéresser, dans le prochain chapitre, à plusieurs extensions de la formulation variationnelle. Nous verrons notamment que cela apporte des éléments de réponse aux problèmes évoqués dans les parties II et III.

Extensions de l'approche variationnelle

Dans ce dernier chapitre, nous illustrons un certain nombre d'extensions de l'approche variationnelle présentée dans le chapitre 7. Nous montrons notamment que plusieurs situations, identifiées comme « difficiles » dans les trois premières parties de ce mémoire, se prêtent particulièrement bien à la résolution variationnelle. Les difficultés de ces problèmes proviennent principalement de la formulation classique de la stéréophotométrie, qui prend la forme d'un problème linéaire : si cette approche se traduit effectivement par des méthodes numériques élémentaires, elle est en revanche très limitée. En effet, de nombreux phénomènes physiques intervenant dans la formation des images sont par nature non linéaires, et la formulation linéaire du problème ne permet de prendre en compte ces phénomènes qu'au prix d'approximations plus ou moins justifiées.

L'exemple des éclairages non directionnels est à ce titre très représentatif. Le flux primaire incident est en réalité non uniforme, dès lors que la surface est éclairée par une source qui n'est pas positionnée à l'infini. La parade la plus naturelle consiste à tout simplement négliger les dépendances spatiales de l'éclairage, en utilisant une approximation par un éclairage directionnel (cf. parties I et II). Nous avons toutefois montré dans le chapitre 5 qu'une modélisation plus fine de l'éclairage pouvait être utilisée, mais que cela faisait perdre au problème son caractère linéaire : une modélisation réaliste du flux incident fait intervenir explicitement les coordonnées 3D des points de la surface, qui sont inconnues, au même titre que les normales et l'albédo. Pour tenir compte de cette dépendance, nous avons proposé dans le paragraphe 5.3 une méthode visant à alterner l'estimation des normales et de l'albédo avec l'intégration des normales. Ce découplage des normales et de la profondeur est empirique, et n'a d'autre but que de se ramener à l'étude d'un problème de stéréophotométrie « classique » i.e., linéaire.

Avec l'approche variationnelle, chaque inconnue du problème peut être découplée des autres inconnues. En particulier, l'albédo et les normales, qui sont des grandeurs physiques indépendantes¹, ne sont plus mélangés au sein d'une variable auxiliaire introduite pour rendre le problème linéaire. Bien sûr, ceci induit un certain nombre de difficultés numériques, mais nous avons vu dans le chapitre 7 comment mettre au point des techniques de résolution efficaces. Nous montrons d'abord, dans le paragraphe 8.1, que si les éclairages et la surface sont simultanément colorés, le problème n'est pas plus difficile à

1. L'albédo dépend principalement du matériau, et les normales du relief macroscopique, mais l'indépendance entre albédo et normales n'est pas si nette, car la réflectance dépend aussi du relief à l'échelle microscopique (cf. chapitre 6).

résoudre que pour des images en niveaux de gris. Dans le paragraphe 8.2, nous revenons sur le problème de la stéréophotométrie non calibrée (UPS). La question de l'intégrabilité des normales ne se posant pas si l'on cherche à estimer directement la profondeur, nous verrons que le problème UPS est bien mieux posé avec l'approche différentielle. Nous montrons ensuite, dans le paragraphe 8.3, que la prise en compte d'éclairages non directionnels est très naturelle avec l'approche variationnelle.

La principale limite de l'approche suggérée, et plus généralement de la stéréophotométrie, est l'hypothèse selon laquelle le flux incident à la surface est le flux *primaire* i.e., que les réflexions mutuelles peuvent être négligées. Aussi précis soit-il, le modèle de flux primaire ne tiendra jamais compte de ces phénomènes. Nous terminons donc ce mémoire sur une perspective qui nous semble intéressante pour leur prise en compte.

8.1 Stéréophotométrie colorée

Comme nous l'avons déjà dit dans le paragraphe 1.5, deux grandeurs ont une dépendance spectrale : l'éclairage $\mathbf{s}$ et l'albédo ρ . Avec le modèle lambertien linéaire, le système d'équations à résoudre s'écrit, dans le cas le plus général :

$$I_{\star}^i(u, v) = \int_0^{+\infty} c_{\star}(\lambda) \rho(u, v, \lambda) \mathbf{s}^i(\lambda) \cdot \mathbf{n}(u, v) d\lambda \quad (8.1)$$

où λ désigne la longueur d'onde et $c_{\star}$ la réponse du récepteur dans le canal de couleur $\star$. Le seul cas où l'équation (8.1) peut s'écrire sous la forme du produit d'un facteur caractéristique du matériau par un facteur caractéristique de l'éclairage est celui où chaque réponse $c_{\star}(\lambda)$, $\star \in \{R, G, B\}$ est une fonction de Dirac. Cette hypothèse est valide en imagerie hyper-spectrale², mais elle est discutable pour les appareils photographiques usuels. Kim et al. sont toutefois parvenus dans [145] à des résultats plausibles sous cette hypothèse. Sans perdre de vue qu'il ne s'agit que d'une approximation, nous nous intéressons donc dorénavant au modèle suivant :

$$I_{\star}^i(u, v) = \rho_{\star}(u, v) \mathbf{s}_{\star}^i \cdot \mathbf{n}(u, v), \quad (i, \star) \in [1, m] \times \{R, G, B\} \quad (8.2)$$

sans faire l'hypothèse, comme cela est souvent le cas, que la surface est blanche ou que les sources lumineuses sont blanches.

Le premier cas de figure ($\rho_R = \rho_G = \rho_B$) se ramène à la stéréophotométrie classique (cf. paragraphe 4.4.1), tandis que le deuxième cas de figure ($\mathbf{s}_R^i = \mathbf{s}_G^i = \mathbf{s}_B^i$, $i \in [1, m]$) peut être résolu en deux temps (cf. paragraphe 1.5). Cependant, le cas où la surface et les éclairages sont simultanément colorés reste, à notre connaissance, un problème ouvert. Nous allons voir, dans le paragraphe 8.1.1, que l'approche variationnelle permet de simplifier la résolution dans le cas d'une surface colorée éclairée par une source blanche, et qu'il n'est pas beaucoup plus difficile d'étendre le modèle variationnel au cas où les sources sont elles aussi colorées (cf. paragraphe 8.1.2).

2. Les images hyper-spectrales présentent des propriétés intéressantes pour la stéréophotométrie, vis-à-vis de la suppression des réflexions mutuelles [185].

8.1.1 Surface colorée éclairée par des sources blanches

Nous supposons dans ce paragraphe que les sources lumineuses sont blanches i.e., que $\mathbf{s}_R^i = \mathbf{s}_G^i = \mathbf{s}_B^i$, $i \in [1, m]$. Le problème à résoudre est celui de l'estimation des albédos ρ_R , ρ_G et ρ_B , qui sont caractéristiques de la réflectance, et de la carte de profondeur z , qui est caractéristique du relief, à partir de m images RGB.

Comme dans le chapitre précédent, nous modélisons explicitement les ombres propres et l'éclairage ambiant, pour garantir une certaine robustesse. Nous supposons donc que les niveaux de couleur sont donnés par le modèle statistique :

$$I_\star^i(u, v) = \rho_\star(u, v) \max \left\{ 0, \mathbf{s}^i \cdot \mathbf{n}(u, v) \right\} + \mathfrak{L}_{\phi_{a,\star}}(u, v) \quad (8.3)$$

où $\mathfrak{L}_{\phi_{a,\star}}(u, v)$ est une variable aléatoire décrivant un bruit additif distribué selon une loi de Laplace centrée en $\phi_{a,\star}$. Notons que le bruit peut être distribué différemment entre les différents canaux.

Nous utilisons dorénavant le modèle de projection perspective pour l'appareil photographique, et remplaçons l'opérateur $\max$ par son approximation différentiable, de telle sorte que (8.3) se réécrit :

$$I_\star^i(u, v) = \rho_\star(u, v) \underbrace{M \left(\tilde{\mathbf{c}}_z^i(u, v) \cdot \nabla z(u, v) + c_{z,3}^i(u, v) \right)}_{M_z^i(u, v)} + \mathfrak{L}_{\phi_{a,\star}}(u, v) \quad (8.4)$$

où nous utilisons les définitions de M , $\tilde{\mathbf{c}}_z^i$ et $c_{z,3}^i$ données dans le paragraphe 7.2.1.

Les discontinuités de l'albédo sont généralement les mêmes dans les trois canaux, ce qui suggère l'utilisation d'un a priori couplé pour ρ_R , ρ_G et ρ_B [38, 100]. Ceci est toutefois laissé en perspective. Par souci de simplicité, nous négligeons la corrélation entre les albédos³. Nous supposons simplement que chaque canal d'albédo est à variations bornées, de sorte qu'en adaptant le raisonnement du paragraphe 7.2.2, on montre que l'estimateur MAP de $(z, \rho_R, \rho_G, \rho_B, \phi_{a,R}, \phi_{a,G}, \phi_{a,B})$ est la solution du problème L^1 -TV :

$$\begin{aligned} \min_{\substack{z \\ \rho_R, \rho_G, \rho_B \\ \phi_{a,R}, \phi_{a,G}, \phi_{a,B}}} & \frac{1}{3m} \sum_{i=1}^m \sum_{\star \in \{R, G, B\}} \left\| I_\star^i - \phi_{a,\star} - \rho_\star M_z^i \right\|_{L^1(\Omega)} + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \\ & + \gamma_z \text{TV}(z) + \gamma_\rho \sum_{\star \in \{R, G, B\}} \text{TV}(\rho_\star) \end{aligned} \quad (8.5)$$

On constate que le problème (8.5) est très similaire au problème (7.62), qui décrit la résolution robuste de la stéréophotométrie avec des images en niveaux de gris. Il suffit donc d'apporter quelques modifications mineures à l'algorithme 5 pour résoudre le problème (8.5).

Les modifications à apporter à l'algorithme de résolution du modèle L^1 -TV du chapitre 7 étant évidentes, nous ne les détaillons pas dans ce paragraphe. Remarquons simplement que la mise à jour des albédos $\rho_\star$ et des éclairages ambiants $\phi_{a,\star}$ n'utilise

3. Nous utilisons donc une définition *anisotrope* de la variation totale.

comme données que les m images $\{I_{\star}^i\}_i$ (tout se passe comme si on effectuait l'estimation canal par canal), alors que la mise à jour de la profondeur utilise la totalité des images, ce qui se traduit par une robustesse accrue pour l'estimation du relief (tout se passe comme si l'on disposait de $3m$ images en niveaux de gris).

Pour tester cette nouvelle approche, nous avons utilisé les mêmes $m = 27$ images que pour les expériences des figures 7.9 et 7.10, qui proviennent du jeu de données *Toys* utilisé dans [273], mais nous utilisons cette fois-ci les images en couleur. La couleur permet de mieux se rendre compte des phénomènes de réflexions mutuelles (cf. figure 8.1) : on distingue en effet le relief dans l'ombre portée de l'image du milieu (effet de pénombre). Malgré ces difficultés, le schéma L^1 -TV fournit des reconstructions 3D plutôt réalistes, en tous cas largement suffisantes pour des applications de rééclairage, comme cela est montré sur la figure 8.2.



FIGURE 8.1 – Trois photographies couleur (parmi $m = 27$) provenant du jeu de données *Toys* [273]. Sur les images en couleur, les reflets brillants sont plus marqués, et on se rend mieux compte des phénomènes de réflexions mutuelles que sur les images en niveaux de gris (cf. la première ligne de la figure 7.8).

8.1.2 Surface colorée éclairée par des sources colorées

Pour utiliser l'approche variationnelle, nous avons dû renoncer à la linéarité du modèle lambertien vis-à-vis de la variable « artificielle » $\mathbf{m} = \rho \mathbf{n}$. Même si cela induit un certain nombre de difficultés techniques, que nous sommes néanmoins parvenus à surmonter dans le chapitre 7, ce choix rend possible la résolution de certains problèmes qu'il est impossible de formuler sous forme linéaire. En particulier, nous avons vu dans le paragraphe 1.5 que la stéréophotométrie en présence d'une surface colorée éclairée par des sources blanches n'était pas beaucoup plus difficile à résoudre que la stéréophotométrie classique. Le cas où les éclairages et la surface sont simultanément colorés est cependant nettement plus difficile (cf. figure 8.3). En effet, il n'est plus possible d'écrire le problème d'estimation simultanée des albédos et des normales comme un problème linéaire. Pour cette raison, le problème de la stéréophotométrie « doublement colorée » a été très peu étudié dans la littérature : les méthodes existantes sont fondées sur des



FIGURE 8.2 – Deux rééclairages de la reconstruction 3D obtenue à partir des $m = 27$ photographies du jeu de données *Toys* [273] (cf. figure 8.1).

hypothèses très fortes sur le caractère fini de l'ensemble des valeurs d'albédos [11], sur la complémentarité des couleurs des sources [221], ou sur l'utilisation conjointe de la stéréophotométrie et d'un capteur de profondeur [10].

Grâce au simple découplage de l'albédo et des normales, il est possible de résoudre le problème (8.2) sans introduire aucun de ces a priori, et de façon robuste. Il suffit de modifier le modèle L^1 -TV (8.5), en remplaçant M_z^i par $M_{\star,z}^i$:

$$\begin{aligned} \min_{\substack{z \\ \rho_R, \rho_G, \rho_B \\ \phi_{a,R}, \phi_{a,G}, \phi_{a,B}}} & \frac{1}{3m} \sum_{i=1}^m \sum_{\star \in \{R,G,B\}} \left\| I_{\star}^i - \phi_{a,\star} - \rho_{\star} M_{\star,z}^i \right\|_{L^1(\Omega)} + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \\ & + \gamma_z \text{TV}(z) + \gamma_{\rho} \sum_{\star \in \{R,G,B\}} \text{TV}(\rho_{\star}) \end{aligned} \quad (8.6)$$

où l'approximation de l'opérateur max dépend du canal de couleur :

$$M_{\star,z}(u, v) = M \left(\tilde{\mathbf{c}}_{z,\star}^i(u, v) \cdot \nabla z(u, v) + c_{\star,z,3}^i(u, v) \right) \quad (8.7)$$

et où les définitions des champs $\tilde{\mathbf{c}}_{z,\star}^i$ et $c_{z,\star,3}^i$ se déduisent de (7.52) en tenant compte de la dépendance spectrale des éclairages. Cette dépendance n'est pas un problème, puisqu'en pratique (cf. algorithme 8), les champs $\tilde{\mathbf{c}}_{z,\star}^i$ et $c_{z,\star,3}^i$ sont « gelés » au sein d'un algorithme de point fixe. Les modifications à apporter à l'algorithme de résolution du modèle L^1 -TV du chapitre 7 étant évidentes, nous ne les détaillons pas dans ce paragraphe. Remarquons simplement que la mise à jour de $(\rho_{\star}, \phi_{\star})$ n'utilise comme données que les m images $\{I_{\star}^i\}_i$ et les m éclairages $\{\mathbf{s}_{\star}^i\}_i$ du même canal de couleur (tout se passe comme si on effectuait l'estimation canal par canal), alors que la mise à jour de la profondeur utilise la totalité des images et des éclairages, ce qui se traduit par une robustesse encore accrue pour l'estimation du relief (tout se passe comme si l'on disposait de $3m$ images en niveaux de gris).

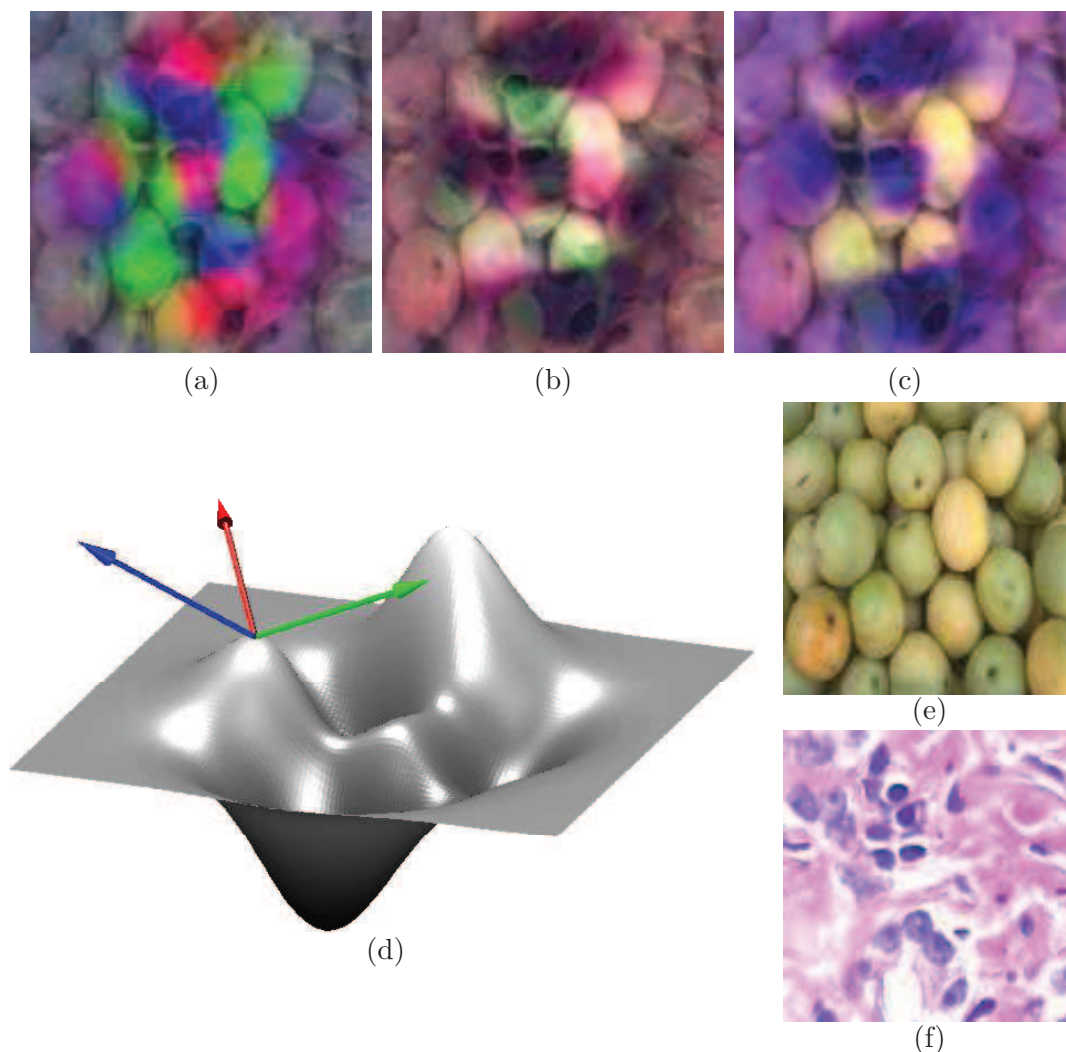


FIGURE 8.3 – (a)-(b)-(c) Trois images (parmi $m = 50$) de la surface (d), sur laquelle est plaqué l'albédo (e). Un éclairage « ambient » (f) est ajouté aux images synthétisées par le modèle (8.2). Trois sources directionnelles colorées en rouge, vert, et bleu, sont utilisées simultanément pour générer chaque image. Par exemple, les flèches colorées de (d) correspondent aux éclairages utilisés pour simuler l'image (a). L'interprétation des images (a)-(b)-(c) en termes de relief est extrêmement difficile pour l'œil humain. Est-elle possible pour la stéréophotométrie ? Si les éclairages colorés sont étalonnés (stéréophotométrie calibrée), la réponse est oui. Sinon (stéréophotométrie non calibrée), le relief peut être déterminé sans ambiguïté, alors que l'albédo ne peut être que partiellement déterminé.

La figure 8.4 montre que la conversion de toutes les données (images et éclairages) en niveaux de gris par simple moyennage des trois canaux, afin d'appliquer un algorithme « simple » de stéréophotométrie (e.g., la résolution en moindres carrés du paragraphe 7.1.7), n'est satisfaisante ni en termes de relief estimé, ni en termes d'albédo estimé. Au contraire, le résultat obtenu avec l'approche proposée confirme que le cas de figure où les éclairages et la surface sont simultanément colorés est moins difficile qu'il n'y paraît, du moins si l'on utilise le modèle approché (8.2). Il suffit d'étalonner les éclairages et de découpler l'estimation du relief de celle de l'albédo.

8.2 Stéréophotométrie non calibrée, robuste et doublement colorée

Rappelons que l'intégrabilité des normales joue un rôle crucial pour limiter les ambiguïtés du problème UPS [282] (cf. chapitre 3). Un avantage majeur de la formulation différentielle provient du fait que la question de l'intégrabilité ne se pose pas, puisque la profondeur est estimée directement. La stéréophotométrie avec éclairages inconnus est donc forcément mieux posée avec la formulation différentielle. Commençons par étudier le cas des images en niveaux de gris.

8.2.1 La stéréophotométrie non calibrée : un problème bien posé

Papadimitri et Favaro montrent dans [197] que dans le cas de la projection perspective, imposer l'intégrabilité des normales suffit à garantir l'unicité du triplet (éclairages, albédo, normales) expliquant les images, pourvu que les paramètres internes de la caméra soient connus (nous avons fait cette hypothèse dans tout le mémoire). Leur raisonnement présente toutefois un certain nombre de défauts, puisqu'il faut d'abord factoriser la matrice des niveaux de gris pour séparer le relief et l'albédo des éclairages (robustesse aux données aberrantes difficile à garantir), puis approcher par différences finies le gradient du champ $\mathbf{m}$ issu de la factorisation matricielle (robustesse aux discontinuités de profondeur non garantie), avant d'identifier par une optimisation en moindres carrés la transformation linéaire rendant ce champ « le plus intégrable ». Cette approche est donc difficile à appliquer en conditions réelles. Cependant, plus que par les outils numériques proposés, le travail de Papadimitri et Favaro est intéressant car le résultat important suivant y est prouvé : il existe un seul champ de normales intégrable qui explique un ensemble de $m \geq 3$ photographies d'une surface lambertienne, vue en perspective sous des éclairages directionnels inconnus. Cependant, il subsiste toujours une ambiguïté : si le couple $(\rho, \{\mathbf{s}^i\})$ est solution, alors $(\frac{1}{\tau}\rho, \{\tau\mathbf{s}^i\})$ l'est également, quel que soit $\tau > 0$.

Avec la formulation différentielle, imposer l'intégrabilité des normales se traduit par une contrainte sur la régularité de la solution z . Si l'on suppose que $z \in \mathcal{W}^{1,2}(\Omega)$ (on rappelle que $\mathcal{W}^{1,2}(\Omega)$ est un espace de Sobolev), alors le champ de normales correspondant

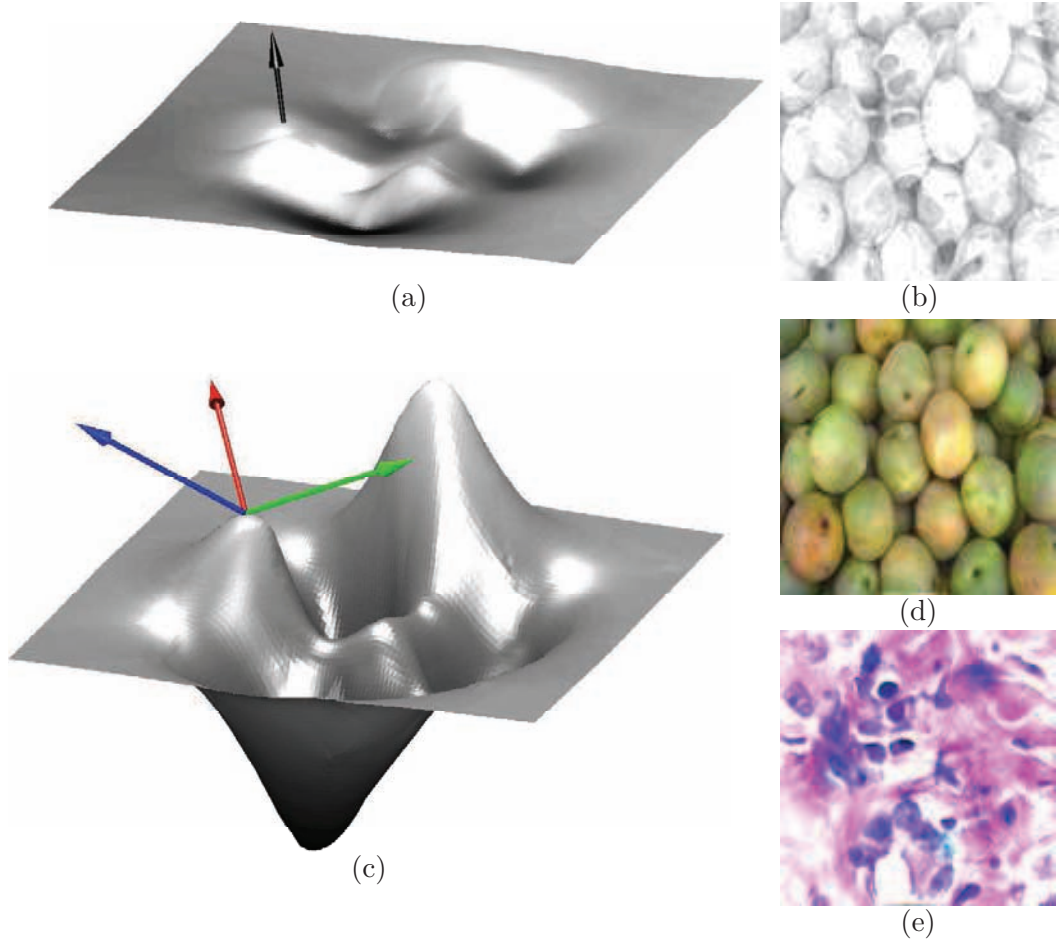


FIGURE 8.4 – Interprétation des images de la figure 8.3 en termes (a) de relief et (b) d'albédo (non coloré), par stéréophotométrie en moindres carrés. L'algorithme du paragraphe 7.1.7 a été appliqué sur des images en « niveaux de gris ». Chacune de ces images est obtenue en calculant la moyenne des trois canaux de couleur d'une image couleur, et en utilisant comme éclairage la moyenne des trois éclairages colorés (la flèche noire indique la direction moyenne d'éclairage pour l'image de la figure 8.3-a, i.e. la moyenne des trois vecteurs de la figure 8.3-d. Ni le relief estimé, ni l'albédo estimé ne sont satisfaisants. (c) Relief estimé par stéréophotométrie « doublement colorée », en estimant simultanément le relief ainsi que : (d) les trois canaux de l'albédo et (e) ceux de l'éclairage ambiant, connaissant les directions et intensités des éclairages colorés (modélisés par les trois flèches de couleur). La connaissance de la dépendance spectrale des éclairages permet de beaucoup mieux séparer le relief de l'albédo et de l'éclairage ambiant.

à la surface z solution du problème en moindres carrés :

$$\min_{\substack{z, \rho \\ \{s^i \in \mathbb{R}^3, i \in [1, m]\}}} \frac{1}{m} \sum_{i=1}^m \left\| I^i - \frac{\rho}{dz} \left(\begin{bmatrix} -s_1^i + s_3^i \frac{u-u_0}{f} \\ -s_2^i + s_3^i \frac{v-v_0}{f} \end{bmatrix} \cdot \nabla z + \frac{zs_3^i}{f} \right) \right\|_{L^2(\Omega)}^2 \quad (8.8)$$

est un champ intégrable. Le résultat de Papadimitri et Favaro prouve donc que le problème variationnel (8.8) admet une solution unique en z , à la constante d'intégration près. Bien sûr, en l'absence de bruit, d'écarts au modèle lambertien linéaire et de discontinuités de profondeur, cette solution unique est exactement la solution déterminée algébriquement dans [197]. Les étapes intermédiaires de l'approche classique (recherche d'un champ de normales non forcément intégrable, puis projection de cette solution sur l'espace des champs intégrables, avant intégration) ne sont pas nécessaires avec l'approche variationnelle, ce qui évite la propagation de biais.

Cependant, cette approche en trois étapes présente l'inconvénient suivant : même si la solution de (8.8) est unique, la convergence d'un schéma numérique approchant itérativement cette solution n'est pas garantie. En effet, si nous avons constaté expérimentalement la convergence du schéma numérique proposé dans le chapitre 7, à partir d'initialisations assez grossières, et en dépit de l'apparente non convexité du problème variationnel, cela n'est plus vrai dans le cas non calbré. Dans ce cas, les schémas numériques peuvent diverger si l'estimation initiale des éclairages est très éloignée des éclairages réels. En revanche, pour peu qu'une initialisation « raisonnable » des éclairages soit choisie, nous avons observé un comportement convergent : l'estimation initiale est importante pour déterminer un bassin de convergence approprié.

En réalité, cette apparente limite de l'approche variationnelle n'en est pas vraiment une. Même si nous avons évoqué son manque de robustesse, la solution proposée par Papadimitri et Favaro peut être utilisée pour obtenir une estimation initiale des vecteurs d'éclairage. Dans certains jeux de données, e.g. ceux fournis dans [273], un étalonnage partiel des éclairages est même fourni, sous la forme d'une approximation de la direction d'éclairage mais sans l'intensité associée. L'intensité peut être fixée arbitrairement à 1 dans un premier temps (c'est ce que nous avons fait pour les expériences des figures 7.7 et 7.8), pour initialiser les vecteurs d'éclairage, qui seront affinés lors de la résolution du problème UPS⁴.

Quoi qu'il en soit, cette discussion nous montre qu'il n'est pas difficile d'obtenir une approximation grossière des vecteurs d'éclairage, suffisante pour se placer dans un bassin de convergence où le critère à minimiser dans (8.8) soit localement convexe. Partant de cette estimation initiale des éclairages (et, e.g., d'une approximation empirique de z et de ρ par des fonctions constantes), la minimisation itérative du problème variationnel (8.8) converge dans tous les tests que nous avons effectués. À cause de l'absence de garantie sur la convexité de l'objectif à minimiser, on ne peut bien sûr rien dire sur le caractère global de ce minimum, mais les résultats expérimentaux semblent indiquer qu'il s'agit d'une solution « acceptable ».

4. À notre avis, ceci constitue l'application la plus intéressante du problème UPS. En effet, la stéréophotométrie étant une technique de laboratoire, il est déraisonnable de supposer les éclairages totalement inconnus. Cependant, même si l'étalonnage est très précis, l'éclairage incident « réel » sera toujours légèrement différent de l'éclairage théorique, à cause notamment des phénomènes de réflexions mutuelles. Affiner l'estimation initiale des éclairages par résolution du problème UPS permet de tenir partiellement compte de ces phénomènes.

8.2.2 Une solution au problème de la stéréophotométrie non calibrée robuste

L'intérêt de formuler le problème UPS sous la forme du problème variationnel (8.8), outre la garantie de l'intégrabilité et de l'unicité de la solution en z , est qu'il est facile d'étendre l'approche pour améliorer la robustesse aux ombres (propres et portées), aux reflets brillants et aux discontinuités de profondeur. Ceci permet d'éviter tout prétraitement [270] pour la résolution du problème UPS (cf. chapitre 3).

Plutôt que de chercher z dans l'espace de Sobolev $\mathcal{W}^{1,2}(\Omega)$, on peut par exemple le chercher dans $BV(\Omega)$, comme la solution du problème L^1 -TV :

$$\begin{aligned} \min_{\substack{z, \rho, \phi_a \\ \{\mathbf{s}^i \in \mathbb{R}^3, i \in [1, m]\}}} & \frac{1}{m} \sum_{i=1}^m \left\| I^i - \phi_a - \frac{\rho}{\delta} \ln \left(1 + \exp \left\{ \frac{\delta}{d_z} \begin{bmatrix} -s_1^i + s_3^i \frac{u-u_0}{f} \\ -s_2^i + s_3^i \frac{v-v_0}{f} \end{bmatrix} \cdot \nabla z + \frac{\delta z s_3^i}{f d_z} \right\} \right) \right\|_{L^1(\Omega)} \\ & + \lambda \|z - z_0\|_{L^2(\Omega)}^2 + \gamma_z \text{TV}(z) + \gamma_\rho \text{TV}(\rho) \end{aligned} \quad (8.9)$$

qui est une variante de (7.62) comportant l'estimation des vecteurs d'éclairage $\mathbf{s}^i$. La résolution numérique du problème non calibré (8.9) est donc sensiblement la même que celle du problème calibré (7.62). En effet, il suffit d'ajouter les vecteurs $\mathbf{s}^i$ aux inconnues du problème (7.87), ce qui revient à ajouter une étape d'estimation de chaque vecteur $\mathbf{s}^i$ à la fin de l'algorithme 7, qui est décrite par le problème d'optimisation suivant :

$$\begin{aligned} \mathbf{s}^{i,l,p+1} = \underset{[s_1^i, s_2^i, s_3^i]^\top}{\text{argmin}} & \sum_{(u,v) \in \Omega} \sum \left[\phi_a^{k,l,p+1}(u,v) - t^{i,l}(u,v) + \frac{\rho^{k,l,p+1}(u,v)}{\delta} \ln \left(1 \right. \right. \\ & \left. \left. + \exp \left\{ \frac{\delta}{d_{z^{k,l,p+1}}(u,v)} \begin{bmatrix} -s_1^i + s_3^i \frac{u-u_0}{f} \\ -s_2^i + s_3^i \frac{v-v_0}{f} \end{bmatrix} \cdot \nabla z^{k,l,p+1}(u,v) + \frac{\delta z^{k,l,p+1}(u,v) s_3^i}{f d_{z^{k,l,p+1}}(u,v)} \right\} \right) \right]^2 \end{aligned} \quad (8.10)$$

qui est un problème d'optimisation convexe, visant à estimer trois paramètres réels à partir de $|\Omega| \gg 3$ données. À chaque itération, la mise à jour des estimations des éclairages peut être réalisée en résolvant le problème (8.10) par une descente de type Levenberg-Marquardt qui converge très rapidement, étant donnée la convexité du problème, le faible nombre de variables à estimer, et le fait que le gradient du problème (8.10) peut être déterminé analytiquement.

Au vu du grand nombre de données, en regard du nombre d'inconnues, l'estimation des éclairages $\mathbf{s}^i$ en moindres carrés est naturellement robuste (théorème central limite). Ceci est d'autant plus vrai qu'elle tient compte des ombres propres (approximation de l'opérateur max), des données aberrantes et des discontinuités de profondeur (estimation en norme 1), et de l'éclairage ambiant.

L'approche variationnelle de la stéréophotométrie offre donc une solution très simple au problème de la stéréophotométrie non calibrée robuste. Par rapport aux approches existantes, elle présente l'avantage de n'être guidée que par les données et non par une identification préalable des écarts au modèle lambertien linéaire [246, 178]. En outre,

8.2. STÉRÉOPHOTOMÉTRIE NON CALBRÉE, ROBUSTE ET DOUBLEMENT COLORÉE

les régularisations TV permettent d'appliquer la méthode avec peu de données, ou avec des données très bruitées. Elles permettent d'une part de garantir le caractère (presque) bien posé, même dans le cas de la projection orthographique : nous avons vu dans le chapitre 3 que dans ce cas, le problème UPS était bien posé dès lors que l'on bornait les variations de la profondeur et de l'albédo par des régularisations TV. Pourvu que γ_z et γ_ρ soient suffisamment grands, et que la solution initiale soit « raisonnable », les régularisations empêchent d'autre part les dérives.

Cette approche s'étend naturellement aux images RGB, alors que la stéréophotométrie non calbrée colorée reste un problème ouvert en utilisant l'approche classique. Nous allons voir que, même lorsque les éclairages et la surface sont simultanément colorés, le relief peut être déterminé sans ambiguïté.

8.2.3 Stéréophotométrie non calbrée doublement colorée

Avec des images RGB, la résolution de la stéréophotométrie non calbrée est un problème encore ouvert. La première étape de l'approche classique du problème UPS à partir d'images en niveaux de gris [9, 89, 114, 197, 217, 283] consiste en effet à *factoriser* la matrice des niveaux de gris pour obtenir une première estimation des éclairages d'un côté, et des normales multipliées par l'albédo de l'autre. Avec des images RGB, cette factorisation semble impossible.

Au contraire, si l'on accepte d'emblée la non linéarité du modèle, la prise en compte de la couleur est très facile avec l'approche variationnelle, pour laquelle les éclairages peuvent même être colorés ! Il n'est en effet pas très difficile d'adapter la résolution du problème UPS en niveaux de gris (8.9) à celle du problème plus général où l'albédo et les éclairages sont colorés, en plus d'être inconnus :

$$\begin{aligned}
 \min_{\substack{z \\ \rho_R, \rho_G, \rho_B \\ \phi_{a,R}, \phi_{a,G}, \phi_{a,B} \\ \{s_\star^i, (i, \star) \in [1, m] \times \{R, G, B\}\}}} & \frac{1}{3m} \sum_{i=1}^m \sum_{\star \in \{R, G, B\}} \left\| I_\star^i - \phi_{a, \star} - \frac{\rho_\star}{\delta} \ln \left(1 + \right. \right. \\
 & \left. \left. \exp \left\{ \frac{\delta}{dz} \left[\begin{array}{c} -s_{\star,1}^i + s_{\star,3}^i \frac{u-u_0}{f} \\ -s_{\star,2}^i + s_{\star,3}^i \frac{v-v_0}{f} \end{array} \right] \cdot \nabla z + \frac{\delta z s_{\star,3}^i}{f dz} \right\} \right) \right\|_{L^1(\Omega)} \\
 & + \lambda \|z - z_0\|_{L^2(\Omega)}^2 + \gamma_z \text{TV}(z) + \gamma_\rho \sum_{\star \in \{R, G, B\}} \text{TV}(\rho_\star)
 \end{aligned} \tag{8.11}$$

Les modifications à apporter au schéma numérique du paragraphe précédent sont mineures : il suffit de répéter la procédure d'estimation des éclairages pour chaque canal de couleur.

Dans l'expérience de la figure 8.5, nous avons utilisé cette approche sur le jeu de données de la figure 8.3, en choisissant comme estimation initiale des éclairages le résultat de la première ligne de la figure 8.4 i.e, l'éclairage initial est la moyenne des trois éclairages colorés réels. L'estimation initiale est donc relativement éloignée de la solution, mais cela suffit pour se placer dans un « bon » bassin de convergence du schéma L^1 -TV. Sur l'exemple de la figure 8.5, le relief estimé est presque identique à celui obtenu par stéréophotométrie calbrée (cf. figure 8.4-c).

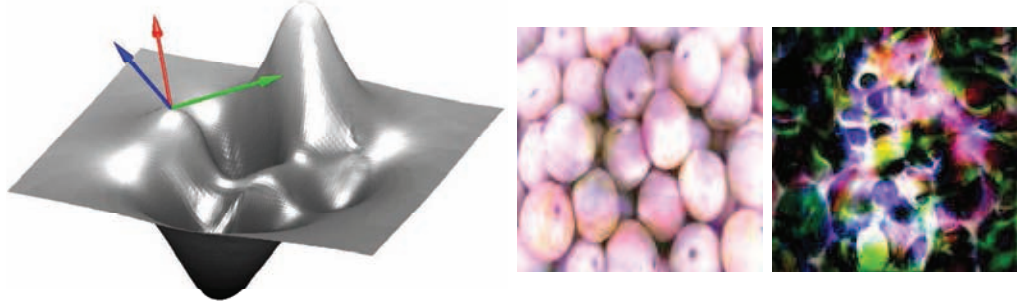


FIGURE 8.5 – Résultat de la stéréophotométrie non calibrée, robuste, colorée, sur les données de la figure 8.3. De gauche à droite : relief estimé, sur lequel les éclairages colorés estimés sont représentés par des flèches colorées, albédo estimé et éclairage ambiant estimé (multiplié par 5). L'estimation du relief est très satisfaisante, tout comme celle des directions d'éclairage et la séparation des caractéristiques géométriques entre la composante « ambiante » et l'albédo. Cependant, les couleurs ne sont pas satisfaisantes : si aucune information sur l'intensité des éclairages n'est disponible, il reste encore une ambiguïté.

On constate cependant sur cette figure que l'estimation de l'albédo est décevante. On peut facilement l'expliquer. Nous avons déjà mentionné que le problème non calibré avec des images en niveaux de gris était ambigu, car on peut multiplier l'albédo estimé par une constante, et diviser les éclairages par cette même constante, sans changer les images. Cette ambiguïté n'est pas vraiment problématique avec un seul canal, mais elle est plus gênante avec des images RGB, car une telle ambiguïté est présente dans chaque canal, ce qui explique l'albédo de la figure 8.5. Il est impossible de lever cette ambiguïté sans hypothèse supplémentaire. Dans le cas général, elle est déterminée implicitement par l'initialisation : si les sources sont « presque blanches », comme sur l'exemple de la figure 8.6, alors les intensités des éclairages sont plus faciles à ajuster, et la couleur de l'albédo estimé est plus réaliste (cf. figure 8.7).

On constate de plus, sur l'exemple de la figure 8.6, que les directions d'éclairage estimées par stéréophotométrie non calibrée sont légèrement différentes des directions étalonnées dans [273]. En effet, la direction fournie dans [273] est celle de l'éclairage incident *primaire* moyen, tandis que les trois flux (correspondant aux trois canaux) que nous estimons par stéréophotométrie non calibrée sont les flux *totaux* moyens, qui prennent en compte plus de phénomènes, notamment les réflexions mutuelles.

En outre, comme nous le montrons sur la figure 8.7, il n'est pas nécessaire de disposer d'une initialisation très précise pour résoudre ce problème non calibré. Une initialisation grossière du relief par un plan, de l'albédo par la moyenne des m images (canal par canal), et des éclairages par des vecteurs unitaires, suffit sur cet exemple à garantir la convergence vers une solution très réaliste.



FIGURE 8.6 – Trois images du jeu de données *Toys*, éclairages colorés estimés (représentés par des flèches colorées) et relief estimé par stéréophotométrie non calibrée, robuste et doublement colorée. Chaque flèche noire indique l'estimation initiale de l'éclairage, qui est la direction fournie dans [273]. Dans ce jeu de données, les intensités des éclairages ne sont pas fournies : la stéréophotométrie non calibrée permet de traiter ces données incomplètes, en estimant les intensités des éclairages, et en affinant l'estimation du flux lumineux total incident dans chaque canal de couleur.

8.3 Retour sur les modèles d'éclairage non directionnels

Nous venons de voir qu'avec très peu d'informations sur la scène, une reconstruction 3D « plutôt réaliste » est obtenue par stéréophotométrie non calibrée colorée, ce qui est suffisant pour des applications de rééclairage.

Si l'éclairage est bien contrôlé, la stéréophotométrie peut également être utilisée comme une technique de *métrologie* et d'*inspection de surface*. Nous revenons donc maintenant sur la prise en compte explicite d'éclairages non directionnels, qui sont généralement utilisés dans les dispositifs permettant d'envisager ce type d'applications. Nous allons voir que l'approche variationnelle offre une solution très naturelle à ce problème habituellement considéré comme difficile [175, 212, 276]. Comme nous l'avons montré dans le chapitre 5, pour une surface lambertienne et un éclairage non directionnel, la définition de $\mathbf{s}$ doit tenir compte du caractère non uniforme de l'éclairage :

$$I(u, v) = \rho(u, v) \max \{0, \mathbf{s}(\mathbf{x}(u, v)) \cdot \mathbf{n}(u, v)\} \quad (8.12)$$

La prise en compte de la dépendance de $\mathbf{s}$ en $\mathbf{x}$, qui est l'inconnue, est difficile avec l'approche classique de la stéréophotométrie. Il n'en est rien avec l'approche variationnelle.



FIGURE 8.7 – Relief, albédo et éclairage ambiant (multiplié par 5) estimés, de gauche à droite, après les itérations $k = 1$, $k = 5$ et $k = 11$ (convergence) du schéma de résolution de la stéréophotométrie non calibrée robuste et doublement colorée. L'initialisation de chaque éclairage consiste en un éclairage blanc, d'intensité identique pour les différentes images et les trois canaux, et de direction égale à la direction moyenne du flux primaire, qui a été étalonnée dans [273].

8.3.1 Modèle d'éclairage primaire explicite

Puisque la normale $\mathbf{n}$ dépend du gradient de profondeur ∇z , on peut réécrire le système de m équations issues du modèle lambertien (8.12) sous forme différentielle :

$$I^i(u, v) = \rho(u, v) \max \left\{ 0, \underbrace{\frac{1}{d_z(u, v)} \left[-s_{z,1}^i(u, v) + s_{z,3}^i(u, v) \frac{u-u_0}{f} \right]}_{\tilde{\mathbf{c}}_z^i(u, v)} \cdot \nabla z(u, v) + \underbrace{\frac{z(u, v) s_{z,3}^i(u, v)}{f d_z(u, v)}}_{c_{z,3}^i(u, v)} \right\} \quad (8.13)$$

où les éclairages $\mathbf{s}_z^i$ varient spatialement et dépendent de la profondeur z inconnue.

Si l'on néglige les réflexions mutuelles, l'éclairage incident se réduit à l'éclairage primaire, dont la dépendance en z peut être modélisée explicitement pour une grande variété de sources (voir le paragraphe 5.1, qui étend la discussion initiée dans [212]). Par exemple, l'éclairage primaire provenant d'une source ponctuelle anisotrope située en $\mathbf{x}_s^i$, de direction principale d'émission $\mathbf{n}_0^i$ et de coefficient d'anisotropie μ^i , s'écrit, d'après (5.18) et (5.21) :

$$\mathbf{s}_z^i(u, v) = \frac{\phi_0^i}{\|\mathbf{x}_s^i - \mathbf{x}_z(u, v)\|^2} \left(\frac{\mathbf{n}_0^i \cdot [\mathbf{x}_z(u, v) - \mathbf{x}_s^i]}{\|\mathbf{x}_s^i - \mathbf{x}_z(u, v)\|} \right)^{\mu^i} \frac{\mathbf{x}_s^i - \mathbf{x}_z(u, v)}{\|\mathbf{x}_s^i - \mathbf{x}_z(u, v)\|} \quad (8.14)$$

où $\mathbf{x}_z$ est défini par (cf. paragraphe 1.1.1) :

$$\mathbf{x}_z(u, v) = z(u, v) \left[\frac{u - u_0}{f}, \frac{v - v_0}{f}, 1 \right]^\top \quad (8.15)$$

La dépendance en z des éclairages peut donc être passablement compliquée et rendre la résolution numérique ardue.

Cependant, aussi compliquée soit-elle, il n'en reste pas moins que le système d'équations (8.13) est quasiment le même que le système d'équations (7.50), qui décrit l'approche différentielle de la stéréophotométrie. L'unique différence entre ces deux modèles est la dépendance des $s_{z,j}^i$ en (u, v) et en z dans (8.13), alors qu'il s'agissait de constantes dans le système (7.50). Or, nous avons montré dans le paragraphe 7.2 que les termes $\tilde{\mathbf{c}}_z^i$ et $c_{z,3}^i$, qui dépendent justement des s_j^i (cf. (7.52)), pouvaient être « gelés » au sein d'itérations de point fixe. Ceci reste vrai avec des éclairages non directionnels⁵. Il s'ensuit que l'algorithme proposé dans le paragraphe 7.2, fondé sur un modèle L^1 -TV associé à la prise en compte explicite des ombres propres et de l'éclairage ambiant, peut servir à résoudre la stéréophotométrie sous des éclairages non directionnels. Il suffit d'effectuer

5. La convergence est bien sûr « encore moins » garantie, mais nous avons expérimentalement observé qu'elle n'était pas mise en défaut.

une modification mineure de l'algorithme 8, en remplaçant les mises à jour (7.95) et (7.96) par :

$$\begin{cases} \tilde{c}^{i,q+1}(u, v) = \frac{1}{d_{z^k, l, p, q}(u, v)} \left[-s_{z^k, l, p, q, 1}^i(u, v) + s_{z^k, l, p, q, 3}^i(u, v) \frac{u-u_0}{f} \right] \\ c_3^{i,q+1}(u, v) = \frac{z^{k, l, p, q}(u, v) s_{z^k, l, p, q, 3}^i(u, v)}{f d_{z^k, l, p, q}(u, v)} \end{cases} \quad (8.16)$$

où les s_z^i sont maintenant entendus comme des *fonctions* de la profondeur et des paramètres d'éclairage (positions, intensités, etc.). Ainsi, grâce à la formulation variationnelle, la prise en compte des non linéarités de l'éclairage ne présente plus de réelle difficulté.

Rappelons de plus qu'avec cette approche, le relief est directement estimé, sans avoir à découpler les normales de la profondeur comme nous l'avons fait dans le paragraphe 5.3. L'approche variationnelle de la stéréophotométrie sous des éclairages non directionnels est à ce titre originale, dans le sens où un tel découplage est très souvent utilisé explicitement [198]. À notre connaissance, les deux seules approches de la stéréophotométrie sous des éclairages non directionnels visant à estimer directement la profondeur, sans passer par l'estimation des normales puis par leur intégration, sont celles proposées par Xie et al. dans [276], et par Mecca et al. dans [175]. Comme nous l'avons déjà mentionné, le formalisme de la géométrie discrète utilisé par Xie et al. peut être avantageusement remplacé par celui des EDP. Le formalisme adopté par Mecca et al. est également celui des EDP, mais les méthodes de résolution numérique proposées sont fondées sur des algorithmes de propagation. Bien qu'elles soient intrinsèquement liées, cette approche et l'approche variationnelle que nous proposons sont en réalité assez différentes. Les méthodes proposées par Mecca et al. sont des méthodes « exactes », qui permettent de prouver l'existence et l'unicité de la solution dans le cas continu. Or, nous pensons au contraire que l'approche variationnelle, dont on peut prouver l'optimalité par un raisonnement statistique, est plus adaptée aux images réelles car elle est naturellement robuste aux écarts au modèle.

8.3.2 Modèle d'éclairage sans paramètre

Lorsque les conditions d'éclairage sont trop complexes pour permettre une modélisation explicite du flux primaire, par exemple avec un éclairage « diffus » obtenu en orientant plusieurs sources vers un réflecteur (e.g., le dermoscope de la société Pixience schématisé sur la figure 5.10), nous avons vu dans le paragraphe 5.2.2 comment étalonner le flux incident sur le plan moyen de la scène. Contrairement aux modèles explicites du paragraphe précédent, nous disposons non pas d'un ensemble réduit de paramètres, mais d'un échantillonnage du flux incident sur Ω , i.e. de champs :

$$\mathbf{s}^i : \Omega \rightarrow \mathbb{R}^3 \quad (8.17)$$

L'éclairage incident est donc supposé indépendant de z ⁶, et les m niveaux de gris au point (u, v) peuvent être modélisés par :

$$I^i(u, v) = \rho(u, v) \max \left\{ 0, \frac{1}{d_z(u, v)} \left[\begin{array}{l} -s_1^i(u, v) + s_3^i(u, v) \frac{u-u_0}{f} \\ -s_2^i(u, v) + s_3^i(u, v) \frac{v-v_0}{f} \end{array} \right] \cdot \nabla z(u, v) + \frac{z(u, v) s_3^i(u, v)}{f d_z(u, v)} \right\} \quad (8.18)$$

La résolution de ce système d'EDP non linéaires peut être à nouveau effectuée par minimisation du problème L^1 -TV (7.62), en adaptant les définitions de $\tilde{\mathbf{c}}_z^i$ et de $\mathbf{c}_{z,3}^i$. De plus, comme nous l'avons montré dans le paragraphe 8.1.2, la prise en compte de surfaces et d'éclairages simultanément colorés⁷ est naturelle. Nous présentons sur la figure 8.8 un jeu de données de stéréophotométrie particulièrement difficile à traiter, qui consiste en $m = 4$ photographies couleur d'une pièce de 1 euro italienne (représentant l'homme de Vitruve), prises avec le dermoscope de la société Pixience. La pièce étant constituée d'un mélange de laiton et de nickel, elle est très brillante. Sa surface présente en outre des discontinuités de profondeur assez importantes, où se produit une forte concentration d'énergie lumineuse due aux réflexions mutuelles.

Malgré ces difficultés, la reconstruction 3D obtenue avec le modèle L^1 -TV (7.62) (cf. figure 8.9) est plutôt satisfaisante, en tous cas suffisante pour l'inspection des détails de la surface. Ceci confirme qu'il n'est pas forcément très utile d'introduire un modèle de réflectance sophistiqué pour décrire les surfaces spéculaires : un modèle aussi simple que le modèle lambertien, couplé à des éclairages très contrôlés, et à une estimation robuste, suffit à observer des détails très fins et, au besoin, à les mesurer.

Cependant, et malgré toutes les précautions que nous avons prises lors de l'étalonnage du flux incident, le relief présente un aspect globalement arrondi. Nous pourrions remettre en question, à tort, la procédure d'étalonnage du paragraphe 5.2.2, car en réalité ce problème est dû au matériau. En effet, la pièce étant métallique, elle réémet beaucoup plus de lumière vers les parois du dispositif d'acquisition que la surface mate utilisée pour l'étalonnage. Ces réflexions parasites sont à leur tour (partiellement) réémises. Le flux incident total est donc légèrement différent du flux incident étalonné, et ce phénomène est accentué sur les bords de la pièce, qui sont plus proches des parois du dispositif. Comme ces effets dépendent de l'éclairage, ils ne peuvent pas être pris en compte par le terme d'éclairage ambiant. Cependant, ils peuvent être atténués en utilisant les termes de régularisation. En augmentant le paramètre λ (connaissance a priori, en choisissant pour z_0 un plan) jusqu'à la valeur $\lambda = 10^{-5}$, on obtient la reconstruction 3D de la figure 8.10, qui est plus réaliste. L'utilisation d'une connaissance a priori sur la profondeur, qui est très facile à exprimer avec l'approche variationnelle, permet de corriger partiellement le biais visible sur le résultat de la figure 8.9.

6. L'étalonnage du flux dans l'espace est plus difficile [65], mais constitue une perspective intéressante.

7. Il faut étalonner au préalable le flux primaire dans chaque canal de couleur.

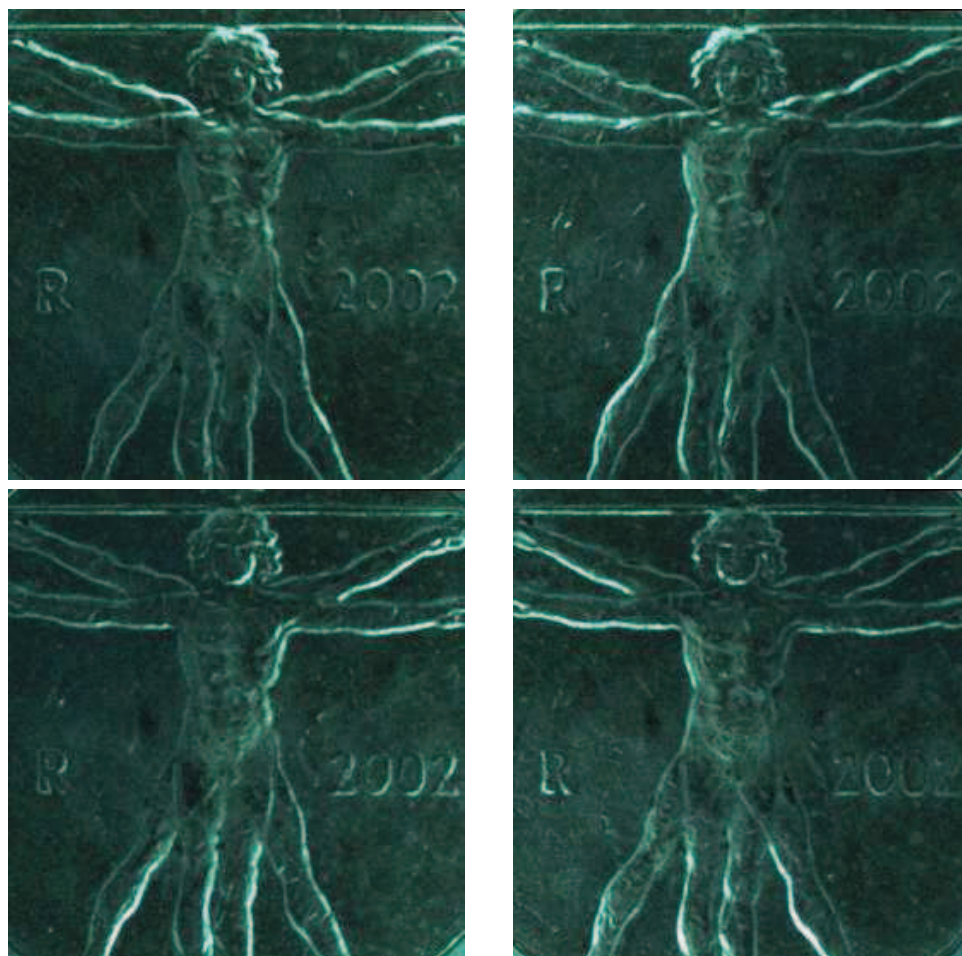


FIGURE 8.8 – Quatre photographies couleur des microgravures de l’avers d’une pièce de 1 euro italienne. Le flux lumineux incident est très complexe, mais peut être étalonné en suivant la procédure indiquée dans le paragraphe 5.2.2. La pièce étant en métal, elle a un aspect brillant, et les réflexions mutuelles au niveau des discontinuités de profondeur sont non négligeables.

8.3.3 Stéréophotométrie semi-calibrée

Dans le paragraphe précédent, nous avons montré les limites de la modélisation du flux incident *primaire*, et même de tout étalonnage de ce flux. Malheureusement, et comme nous l’avons mentionné dans le paragraphe 1.1.4, la modélisation explicite des phénomènes de réflexions mutuelles est difficile, car cela nécessiterait de résoudre une équation de Fredholm [78]. Nous présentons maintenant une piste, qui pourrait permettre, à terme, des reconstructions 3D par stéréophotométrie très précises, même en présence de phénomènes aussi complexes que les réflexions mutuelles.

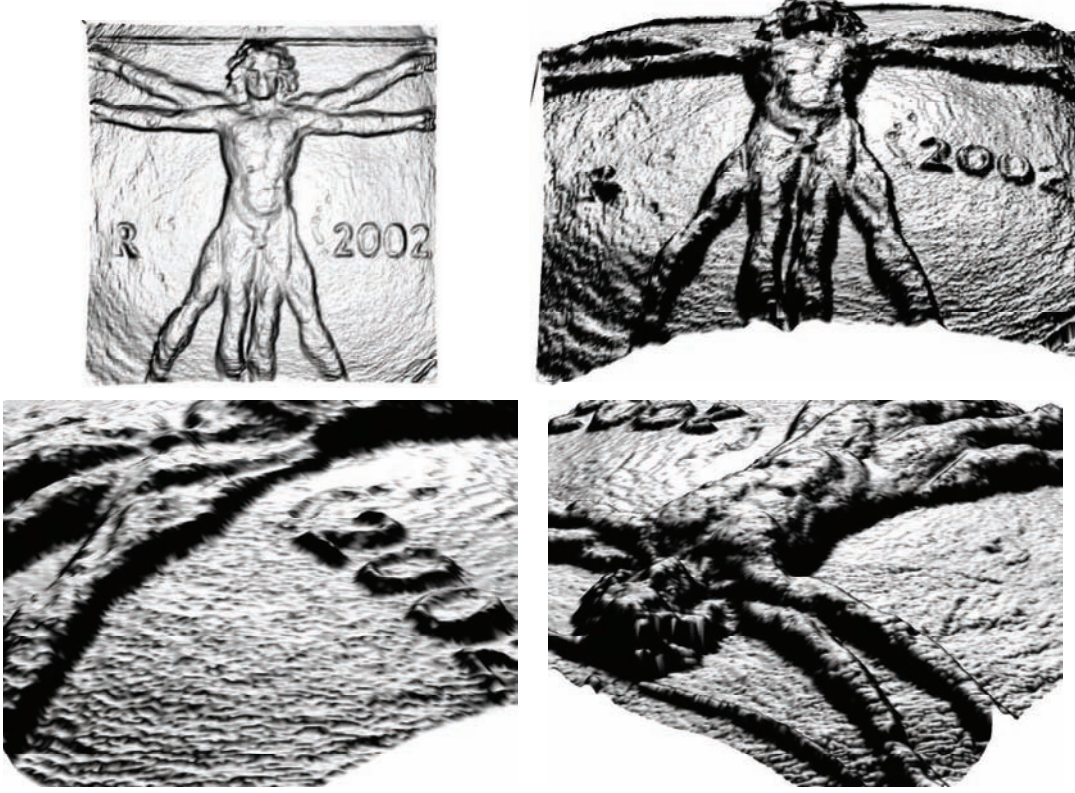


FIGURE 8.9 – Quatre vues de la reconstruction 3D des microgravures d’une pièce de 1 euro italienne, à partir des $m = 4$ images de la figure 8.8. L’approche proposée permet donc l’inspection de surfaces, y compris métalliques. Le relief présente toutefois un aspect globalement arrondi, similaire à celui de la figure 1.13-c (éclairages mal étalonnés) : la pièce étant très brillante, elle réfléchit beaucoup de lumière vers les parois du dispositif, qui est partiellement renvoyée vers la surface (réflexions mutuelles). Le flux incident total est donc sensiblement différent du flux incident étalonné, en particulier à proximité des parois (bords de l’image).

Remarquons que le modèle qui se prête le mieux au problème inverse est le modèle lambertien linéaire :

$$I^i(u, v) = \rho(u, v) \mathbf{s}^i(u, v) \cdot \mathbf{n}(u, v) \quad (8.19)$$

et qu’en réalité, ce modèle suffit à expliquer n’importe quel niveau de gris, si l’on ne contraint pas $\mathbf{s}^i$ à être égal au flux lumineux incident. Il suffit par exemple de choisir $\mathbf{s}^i(u, v)$ orthogonal à $\mathbf{n}(u, v)$ pour expliquer un niveau de gris nul à cause d’une ombre propre, ou de choisir $\|\mathbf{s}^i(u, v)\| = 0$ pour expliquer un niveau de gris nul à cause d’une ombre portée. Le cas des réflexions mutuelles n’est pas plus difficile : par exemple, si un flux primaire $\mathbf{s}^{i,1}$ et un flux secondaire $\mathbf{s}^{i,2}$ éclairent simultanément une surface, on peut

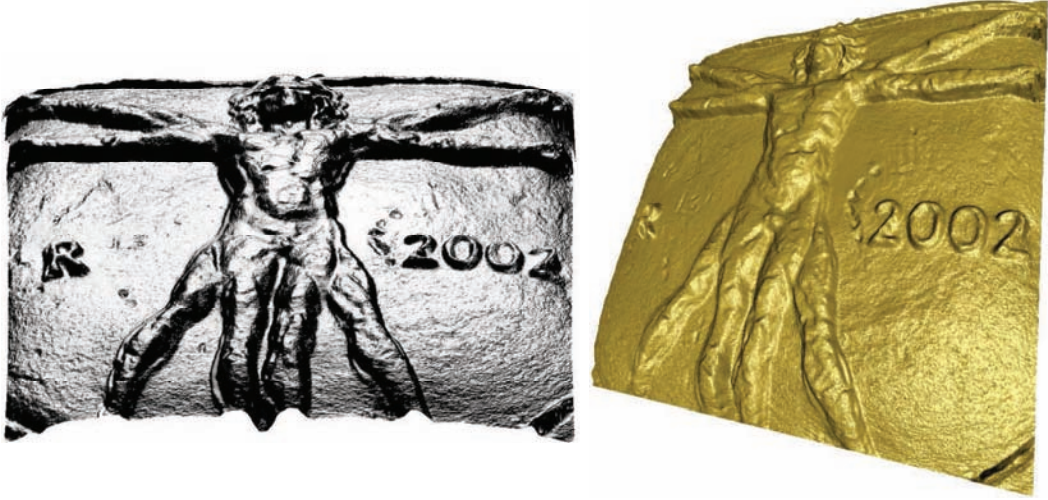


FIGURE 8.10 – Deux rééclairages de la reconstruction 3D de la pièce de 1 euro italienne, en utilisant le fait qu'elle est, en moyenne, assimilable à un plan, comme connaissance a priori sur la profondeur. La pièce est clairement moins « bombée » que sur la figure 8.9.

écrire :

$$I^i(u, v) = \rho(u, v) \underbrace{\left(\mathbf{s}^{i,1}(u, v) + \mathbf{s}^{i,2}(u, v) \right)}_{\mathbf{s}^i(u, v)} \cdot \mathbf{n}(u, v) \quad (8.20)$$

même s'il n'est pas possible d'expliciter $\mathbf{s}^{i,2}$, au contraire du flux primaire $\mathbf{s}^{i,1}$. Enfin, il a été montré très récemment dans [184] que le modèle (8.19) restait valide dans un milieu diffusant (« *scattering medium* ») tel que l'eau (la définition de $\mathbf{s}$ n'a dans ce cas plus rien à voir avec celle d'un flux incident primaire).

Ces différentes remarques plaident en faveur d'une « abstraction » de la notion d'éclairage. En effet, si $\mathbf{s}^i(u, v)$ n'est pas contraint, on pourra toujours le choisir de telle sorte que son produit scalaire avec $\rho \mathbf{n}(u, v)$ explique le niveau de gris $I^i(u, v)$, et ce quels que soient l'éclairage incident et la réflectance de la surface. Ceci nous invite à remplacer la notion d'éclairage par celle de *contrôle* : les « éclairages » deviennent des variables auxiliaires sans réelle signification physique. En utilisant la formulation différentielle, le problème le plus général peut s'écrire sous la forme suivante, dans le cas de la projection perspective :

$$I^i(u, v) = \rho(u, v) \mathbf{c}^i(u, v) \cdot \begin{bmatrix} \nabla z(u, v) \\ z(u, v) \end{bmatrix} \quad (8.21)$$

où $\mathbf{c}^i$ est un contrôle inconnu que l'on va chercher à optimiser. Le problème de la stéréophotométrie devient alors un problème d'estimation simultanée du relief, de l'albédo⁸ et

8. On pourrait tout aussi bien éliminer l'albédo des inconnues en posant $\mathbf{c}^i \leftarrow \rho \mathbf{c}^i$: le terme de contrôle modéliserait alors implicitement *n'importe* quel modèle de réflectance, ce qui généraliserait les approches présentées dans le chapitre 6.

du « meilleur » contrôle. Une idée similaire a été proposée dans le contexte du SFS [91]. L'estimation de ce terme de contrôle est toutefois extrêmement difficile dans le contexte du SFS, qui est un problème fondamentalement mal posé. Nous pensons que cette idée est plus adaptée à la stéréophotométrie, qui constitue un problème bien mieux posé.

Finalement, la formulation la plus générale du problème est étonnamment simple, puisque l'EDP (8.21) est linéaire. Estimer le relief et l'albédo à partir d'équations telles que (8.21) ne semble toutefois possible que si les champs $\mathbf{c}^i$ sont connus. Sinon, le problème est très mal posé : il faudrait estimer, en plus de la géométrie et de l'albédo, m vecteurs de $\mathbb{R}^3$ par point à partir de m niveaux de gris. En première approximation, nous pouvons cependant assimiler le contrôle au flux primaire, si l'on connaît sa forme explicite, voire au flux primaire étaloné :

$$\mathbf{c}^i(u, v) \approx \frac{1}{d_z(u, v)} \begin{bmatrix} -s_1^i(u, v) + s_3^1(u, v) \frac{(u-u_0)}{f} \\ -s_2^i(u, v) + s_3^i(u, v) \frac{(v-v_0)}{f} \\ \frac{s_3^i}{f} \end{bmatrix} \quad (8.22)$$

Pour initialiser la solution, nous pouvons résoudre le système d'équations (8.21) avec la forme explicite (8.22) des contrôles $\mathbf{c}^i$. Le système d'équations est alors le système d'équations quasi-linéaires auquel nous nous sommes déjà intéressés dans cette partie. Au vu des résultats que nous avons présentés, nous pouvons espérer que cette initialisation offre une solution relativement « proche » du relief réel : même si la reconstruction 3D de la figure 8.9 est indéniablement biaisée, elle constitue une approximation plus que raisonnable de la surface réelle.

Partant de cette initialisation « raisonnable », nous pouvons procéder, dans un deuxième temps, à l'estimation conjointe du contrôle et du relief. L'estimation du contrôle permettra de prendre en compte les phénomènes que nous n'avons pas réussi à modéliser avec le seul flux primaire. Nous espérons ainsi corriger le biais résiduel dans l'estimation du relief. Cette seconde étape, difficile, constitue une perspective intéressante, probablement la plus intéressante de toutes celles évoquées jusqu'à présent. Cette étude a été initiée dans [216], où nous avons proposé, pour parer au caractère mal posé du problème, d'introduire un a priori de régularité sur les contrôles $\mathbf{c}^i : \Omega \rightarrow \mathbb{R}^3$. Comme cela est illustré sur la figure 5.7, il est raisonnable de supposer que le flux incident total est un champ lisse, au moins par morceaux. Nous pouvons donc, par exemple, introduire un a priori TV sur ces champs, et résoudre une variante du problème variationnel L^1 -TV (7.62) (nous négligeons l'éclairage ambiant et les ombres propres) qui comporte une régularisation des contrôles :

$$\min_{\substack{z, \rho \\ \{\mathbf{c}^i, i \in [1, m]\}}} \frac{1}{m} \sum_{i=1}^m \left\| I^i - \rho \mathbf{c}^i \cdot \left[\nabla z^\top, z \right]^\top \right\|_{L^1(\Omega)} + \lambda \|z - z_0\|_{L^2(\Omega)}^2 + \gamma_z \text{TV}(z) + \gamma_\rho \text{TV}(\rho) + \gamma_c \sum_{i=1}^m \text{TV}(\mathbf{c}^i) \quad (8.23)$$

en utilisant une définition vectorielle de la variation totale pour la régularisation des contrôles [38, 100]. Ce type de problème peut à nouveau être résolu par une approche proximale [199].

Pour valider cette approche, nous nous sommes concentrés sur les microgravures de la pièce de monnaie peinte en blanc que dont nous avons déjà effectué la reconstruction 3D dans le chapitre 1 (cf. figure 1.13). Les bords des lettres 'C' et 'E' du mot 'Cent' constituent des discontinuités de profondeur autour desquelles le flux lumineux se concentre, ce qui se traduit par des niveaux de gris supérieurs à ceux attendus (cf. figure 8.11). D'autre part, le relief étant dans ce cas effectivement lisse par morceaux, on peut espérer que l'a priori TV sur la profondeur corrige le biais à proximité des discontinuités, et force les niveaux de gris trop élevés à être expliqués par une variation du contrôle, plutôt que par une variation de profondeur.

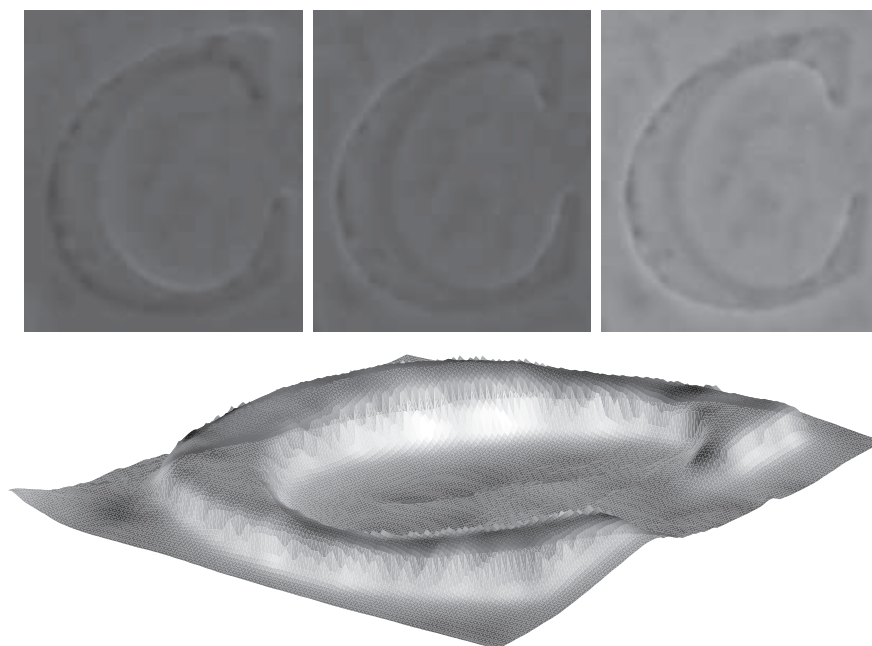


FIGURE 8.11 – Trois photographies (parmi $m = 15$) de la lettre C de la pièce de monnaie de la figure 1.13, et reconstruction 3D obtenue par la méthode décrite dans [216]. Malgré les phénomènes de réflexions mutuelles, l'a priori TV sur la profondeur permet de reconstruire des arêtes relativement nettes.

Dans les expériences des figures 8.11 et 8.12, nous choisissons comme initialisation du contrôle la moyenne du flux étalonné sur la zone observée, et comme initialisation de la surface un plan, avant de les affiner tous deux itérativement par une approche L^1 -TV où la régularisation porte à la fois sur la profondeur et sur les contrôles [216]. Les résultats sont encourageants (cf. figures 8.11 et 8.12), car ils montrent d'une part que l'a priori TV sur la profondeur permet de limiter le biais dû aux réflexions mutuelles, et d'autre part qu'en affinant l'estimation du contrôle au cours de la reconstruction, on prend effectivement en compte les phénomènes de réflexions mutuelles.

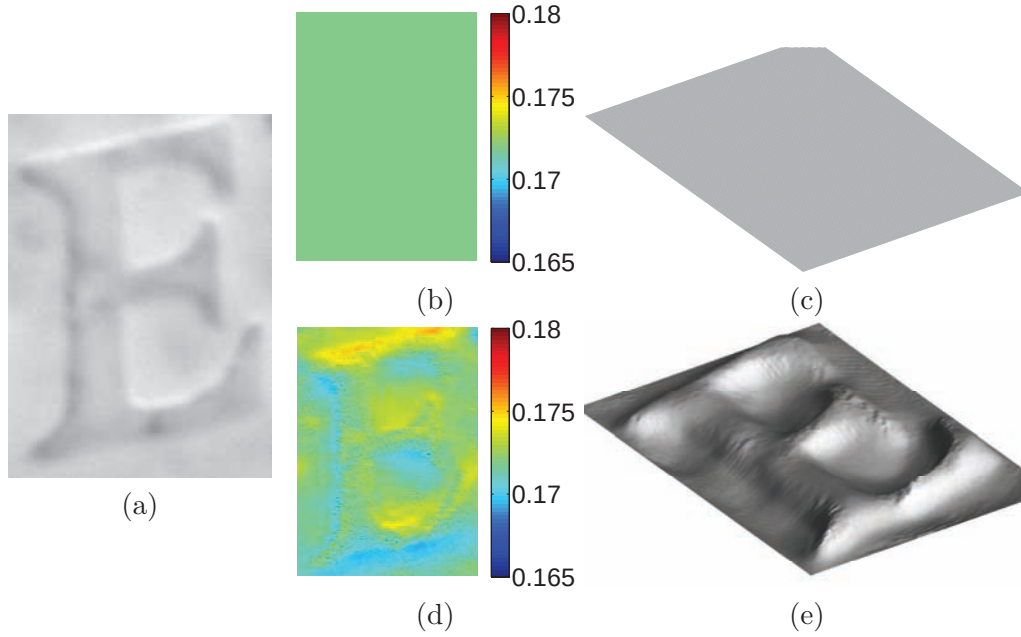


FIGURE 8.12 – Reconstruction 3D de la lettre E (cf. figure 1.13). (a) Une des $m = 15$ images, en l'occurrence I^1 . (b) Valeur initiale de $\|\mathbf{c}^1\|$. (c) Relief initial. (d) Valeur estimée de $\|\mathbf{c}^1\|$ à la convergence. (e) Relief estimé à la convergence. Le contrôle estimé prend bien en compte à la fois les réflexions mutuelles et les ombres.

8.4 Conclusion du chapitre

Dans ce chapitre, nous avons proposé plusieurs extensions de l'approche variationnelle de la stéréophotométrie. Nous avons d'abord montré qu'elle permettait de résoudre le problème si la surface et les éclairages étaient simultanément colorés. Nous avons ensuite reformulé le problème UPS sous forme variationnelle, ce qui permet à la fois de le rendre mieux posé et de le résoudre de manière robuste. Enfin, nous sommes revenus sur les modèles d'éclairage non directionnels, et avons montré que l'approche variationnelle était, dans ce cas également, particulièrement adaptée. Il est naturel, avec cette approche, de prendre en compte la dépendance spatiale de l'éclairage, qui peut même être abstrait sous la forme d'un terme de contrôle afin de modéliser implicitement les phénomènes de réflexions mutuelles.

Il est étonnant de constater que les méthodes variationnelles ont eu assez peu d'écho dans le contexte de la stéréophotométrie, à l'exception bien sûr des travaux liés au problème de l'intégration des normales. Pourtant, au vu des résultats présentés dans ce chapitre, il semble qu'elles soient particulièrement adaptées à la résolution de ce problème inverse, grâce notamment à leur flexibilité. En effet, il suffit généralement d'apporter quelques modifications mineures au modèle L^1 -TV pour prendre en compte des phénomènes complexes. L'approche variationnelle est donc très facile à « instancier »

en de multiples variantes, toutes similaires et pouvant être traitées avec les mêmes outils, qui permettent de résoudre des problèmes de difficulté croissante. Chaque extension de l'approche ne requiert que de changer une des « briques de base » du schéma numérique : on pourrait presque parler d'approche « orientée objet ».

Une autre extension possible de l'approche variationnelle consisterait à modéliser des phénomènes tels que la translucence, le flou de bougé ou le flou de mise au point [72, 139]. Il suffirait, du moins en théorie, d'inclure un opérateur de convolution K dans le terme d'attache aux données du problème (8.23) :

$$\min \frac{1}{m} \sum_{i=1}^m \left\| I^i - K * \left(\rho \mathbf{c}^i \cdot \left[\nabla z^\top, z \right]^\top \right) \right\|_{L^1(\Omega)} + \lambda \|z - z_0\|_{L^2(\Omega)}^2 \quad (8.24)$$

$$+ \gamma_z \text{TV}(z) + \gamma_\rho \text{TV}(\rho) + \gamma_c \sum_{i=1}^m \text{TV}(\mathbf{c}^i)$$

où K est un opérateur linéaire qui est soit connu (déconvolution *supervisée*), soit estimé automatiquement (déconvolution *aveugle*), cette dernière approche pouvant éventuellement être appliquée en ajoutant un a priori (e.g., de type TV) sur l'opérateur K [50].

Conclusion et perspectives

Conclusion

Nous avons présenté dans ce mémoire plusieurs contributions théoriques, numériques et applicatives au problème de la reconstruction 3D par stéréophotométrie.

La partie I nous a permis de revenir sur les aspects liés à la robustesse de l’approche « classique », qui opère en deux étapes : estimation robuste des normales (chapitre 1), puis intégration robuste des normales (chapitre 2). Dans ces deux chapitres, nous avons identifié un certain nombre de biais qu’il convient de traiter de manière appropriée, afin de pouvoir appliquer la stéréophotométrie en conditions réelles. Dans la partie II, nous avons étudié deux problèmes de stéréophotométrie mal posés : la stéréophotométrie non calibrée (chapitre 3) et la stéréophotométrie à $m = 2$ images (chapitre 4). L’intérêt de ces deux problèmes est qu’ils illustrent les limites de l’approche classique : l’estimation locale de la normale étant ambiguë, il est nécessaire de revenir à une formulation globale (cette même idée permet de limiter les ambiguïtés du *shape-from-shading*). Nous avons ensuite remis en question, dans la partie III, l’hypothèse d’une surface lambertienne éclairée par des sources directionnelles. Si une modélisation plus fine du flux incident a un réel intérêt pour limiter les sources de biais dans la reconstruction 3D (chapitre 5), il n’en va pas de même pour les modèles de réflectance (chapitre 6). À titre d’exemple, les reflets brillants constituent effectivement des écarts au modèle lambertien, mais l’effort à fournir pour utiliser un modèle de réflectance spéculaire est sans commune mesure avec une simple estimation des normales sous l’hypothèse lambertienne, pour peu que cette estimation soit robuste.

En présence de données aberrantes ou de modèles non linéaires, la résolution de la stéréophotométrie par l’approche classique requiert l’utilisation de méthodes qui, si elles ont le mérite de « bien fonctionner », risquent de s’apparenter à des « heuristiques ». Plutôt que de proposer des améliorations incrémentales de ces heuristiques, nous suggérons, dans la partie IV, une profonde remise en question de l’approche classique : d’après les conclusions des trois premières parties, il semble finalement qu’une « bonne » approche de la stéréophotométrie ne puisse être que *globale*. Cette observation, qui pourrait faire figure de conclusion en une seule phrase, semble aller à l’encontre de l’évidence, car la vogue actuelle de la stéréophotométrie s’explique en partie par la simplicité de sa formulation *locale*. Cependant, même si les méthodes de résolution proposées dans le chapitre 7 sont plus sophistiquées que celles de la partie I, le chapitre 8 nous a permis de montrer le gain effectif que l’on peut attendre d’une approche globale pour résoudre la stéréophotométrie : certains problèmes deviennent mieux posés, plus faciles à résoudre de façon robuste, voire tout simplement solubles.

Plus que l’identification et la correction des sources de biais, qui étaient annoncés dans l’introduction comme l’objectif principal de cette thèse, nous avons finalement mené une analyse critique de la stéréophotométrie, que nous espérons pertinente et constructive. Nous avons tenté d’identifier les problématiques les plus représentatives des travaux de

recherche actuels sur la stéréophotométrie, de suggérer de nouvelles pistes pour leur résolution, et de formuler un certain nombre de nouveaux problèmes. Alors que le *shape-from-shading* s'est révélé, au fil des années, être un problème difficile très prisé par les mathématiciens, mais aux applications extrêmement limitées, la stéréophotométrie est nettement plus simple à résoudre et peut rivaliser avec les techniques de reconstruction 3D les plus performantes. À ce titre, l'étude des problèmes mal posés et la recherche de schémas numériques efficaces occupent dans les publications une place somme toute mineure, en comparaison de travaux à visées plus applicatives. Dans cette thèse, qui se situe à la frontière entre théorie et applications, nous avons cherché, dans la mesure du possible, à éviter l'utilisation de critères ou de méthodes empiriques, mais sans prouver ni la convergence des schémas numériques, ni le caractère bien posé des problèmes abordés. Nous espérons avoir réussi à trouver le « juste milieu » entre mathématiques appliquées et applications des mathématiques.

Perspectives

Outre les perspectives mentionnées dans chaque conclusion de chapitre, plusieurs axes de recherche découlent assez naturellement des travaux présentés dans ce mémoire. Nous en avons sélectionné trois, qui correspondent aux trois aspects du problème : applicatif, numérique et théorique.

Perspective applicative : fusion de données RGBD

Nous avons vu, dans le paragraphe 2.3.3, que l'introduction d'un a priori sur le relief à reconstruire permettait de rendre le problème de l'intégration mieux posé, puisqu'il rend la fonctionnelle à minimiser coercive. L'approche variationnelle de la stéréophotométrie étant très liée au problème de l'intégration, nous avons réutilisé cet a priori dans le chapitre 7. Dans les deux cas, il s'agit d'un artifice pour rendre le problème bien posé, et le choix de l'a priori est empirique (en pratique, nous choisissons souvent comme a priori un plan).

Toutefois, un a priori n'a aucune raison d'être empirique. Par exemple, en utilisant une caméra RGBD (e.g., la Kinect de Microsoft), il serait possible d'associer aux images RGB une « image » de la profondeur en chaque point (canal D, pour *depth*). Une fois le système étalonné (ce qui est un problème non trivial, cf. [120, 222]), nous disposerions en chaque pixel (u, v) d'un quadruplet de valeurs, quitte à interpoler les données manquantes : les trois niveaux de couleur $I_R(u, v)$, $I_G(u, v)$ et $I_B(u, v)$, et une estimation de la profondeur $z_0(u, v)$. Cette estimation de la profondeur pourrait être choisie comme a priori, soit pour l'étape d'intégration des normales, soit directement dans l'approche variationnelle⁹.

9. Comme en témoigne l'image de la figure 5.8, nous avons déjà commencé à explorer cette piste, au cours des stages de première année de master de Jade Boumaza et de Tom Lucas.

Perspective numérique : utilisation des techniques de lancer de rayons au sein d'une boucle d'analyse par synthèse

Une seconde perspective, que nous avons évoquée à plusieurs reprises, consisterait à utiliser les techniques, bien connues en synthèse d'images, de lancer de rayons pour prendre en compte les ombres portées, les phénomènes de pénombre et les réflexions mutuelles. Ces trois effets sont en effet impossibles à modéliser localement : c'est la raison pour laquelle nous les avons traités comme des données aberrantes tout au long du mémoire.

Puisque la simulation de ces phénomènes est possible en synthèse d'images, il peut sembler paradoxal de ne pas se servir des mêmes outils pour résoudre le problème inverse, ce qui permettrait d'expliquer la totalité des données. Cependant, aussi séduisante qu'elle puisse paraître, cette idée doit être considérée avec précaution. Supposons que l'on dispose d'une estimation grossière du relief, de la réflectance et des éclairages. La génération de m images de synthèse par un moteur de rendu permettra facilement d'identifier les écarts entre les images d'origine et les images reprojetées, et donc les zones où l'estimation doit être remise en question. Cependant, comment modifier effectivement l'estimation courante pour la faire davantage « coller » aux données ? On pourrait certes générer aléatoirement des modèles 3D, selon une approche d'« analyse par synthèse » dont l'intérêt pour la reconstruction 3D de structures aussi complexes que des feuillages a été démontré dans [106]. Cependant, il faudrait dans notre scénario remettre en question non seulement les positions 3D, mais également la réflectance et les éclairages, ce qui semble irréaliste en termes de temps de calcul, même avec les cartes graphiques les plus récentes.

Cette piste, qui était déjà mentionnée par Frédéric Courteille en conclusion de sa thèse sur le *shape-from-shading* [68], constituait en réalité le sujet initial de ma thèse¹⁰. Force est de constater que j'ai été « détourné » de cet objectif initial. Il est probable qu'il faille encore attendre quelques années avant de pouvoir générer en un temps acceptable un nombre suffisamment élevé d'images de synthèse « réalistes », i.e., nécessitant le recours au lancer de rayons, et intégrer ces techniques au sein d'une boucle d'analyse par synthèse.

Perspective théorique : reformulation de la stéréophotométrie en termes de contrôle

La dernière perspective est liée à une piste que nous avons évoquée à la fin du chapitre 8. Sous l'hypothèse lambertienne, la luminance émise par la surface au point $\mathbf{x}(u, v)$ est proportionnelle au produit scalaire de la normale $\mathbf{n}(u, v)$ par le vecteur d'éclairage $\mathbf{s}(u, v)$. Ce vecteur d'éclairage, dans la plupart des travaux sur la reconstruction 3D photométrique, caractérise le flux primaire. Or, ce modèle *local* ne permet pas d'expliquer certaines données, à cause de phénomènes qui sont par nature *globaux*, comme les ombres

10. Si le « nous » a été utilisé jusqu'ici par « modestie », et surtout pour témoigner du caractère collectif des recherches présentées dans ce mémoire, il me faut néanmoins assumer le choix d'avoir fait « une autre » thèse que celle qui était initialement prévue.

portées, la pénombre ou les réflexions mutuelles. Si $\mathbf{s}(u, v)$ désigne le flux incident total au point $\mathbf{x}(u, v)$, alors toutes les données peuvent effectivement être expliquées par le modèle lambertien. Mais si nous cherchons une estimation précise de la normale, celle-ci ne doit pas dépendre de l'hypothèse lambertienne, qui n'est jamais parfaitement vérifiée. Nous proposons donc d'*abstraire* la notion d'éclairage en termes de contrôle, c'est-à-dire de remplacer le vecteur d'éclairage $\mathbf{s}(u, v)$ par un vecteur de contrôle $\mathbf{c}(u, v)$. Grâce à la redondance d'informations qui caractérise la stéréophotométrie, il doit être possible d'abstraire encore plus ces vecteurs en procédant à leur élimination. Nous espérons ainsi éviter tout biais dans l'estimation des normales.

Cette piste n'est pas totalement utopique. En effet, il a été montré récemment par Kunsberg et Zucker [156] que les outils de la géométrie différentielle permettaient, dans le contexte du *shape-from-shading*, de caractériser le relief sans recourir à aucun modèle d'éclairage. Puisque les niveaux de gris dépendent des deux dérivées partielles du premier ordre de la profondeur, le relief est décrit localement par une équation à deux inconnues par point. Les dérivées partielles de l'image fournissent des équations supplémentaires, mais également d'autres inconnues, qui sont les dérivées partielles de la profondeur d'ordres supérieurs. Kunsberg et Zucker ont montré comment reformuler le problème du « *shape-from-shading* non calibré » avec albédo uniforme, en utilisant de telles équations. Un ensemble d'équations indépendantes de l'éclairage est obtenu, au moyen d'outils de la géométrie riemannienne (dérivées covariantes, transport parallèle, etc.). Le problème reste mal posé, mais l'utilisation d'images supplémentaires, acquises sous différents éclairages, devrait suffire à lever l'ambiguïté. Ceci permettra, nous l'espérons, de prouver que la formulation *locale* de la stéréophotométrie non calibrée constitue un problème bien posé.

Bibliographie

- [1] A. ABRAMS, C. HAWLEY et R. PLESS : Heliometric Stereo: Shape from Sun Position. *In European Conference on Computer Vision (ECCV)*, volume 7573 de *Lecture Notes in Computer Science*, pages 357–370, Florence, Italie, 2012. [127](#), [128](#), [154](#), [155](#)
- [2] J. ACKERMANN, S. FUHRMANN et M. GOESELE : Geometric Point Light Source Calibration. *In Vision, Modeling & Visualization*, pages 161–168, Lugano, Suisse, 2013. [181](#)
- [3] J. ACKERMANN, M. RITZ, A. STORK et M. GOESELE : Removing the example from example-based photometric stereo. *In Trends and Topics in Computer Vision (ECCV Workshops)*, volume 6554 de *Lecture Notes in Computer Science*, pages 197–210. Florence, Italie, 2012. [127](#), [128](#), [154](#), [155](#), [194](#), [197](#)
- [4] E. H. ADELSON et A. P. PENTLAND : The perception of shading and reflectance. *In Perception as Bayesian inference*, pages 409–423. Cambridge University Press, 1996. [3](#)
- [5] M. AGGARWAL, H. HUA et N. AHUJA : On cosine-fourth and vignetting effects in real lenses. *In IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 472–479, Vancouver, Canada, 2001. [16](#)
- [6] A. AGRAWAL, R. RASKAR et R. CHELLAPPA : What is the range of surface reconstructions from a gradient field? *In European Conference on Computer Vision (ECCV)*, volume 3951 de *Lecture Notes in Computer Science*, pages 578–591. Graz, Autriche, 2006. [60](#), [62](#), [80](#)
- [7] J. AHMAD, J. SUN, L. SMITH et M. SMITH : An improved photometric stereo through distance estimation and light vector optimization from diffused maxima region. *Pattern Recognition Letters*, 50:15–22, 2014. [189](#), [190](#)
- [8] N. ALLDRIN, T. E. ZICKLER et D. J. KRIEGMAN : Photometric stereo with non-parametric and spatially-varying reflectance. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, États-Unis, 2008. [193](#)
- [9] N. G. ALLDRIN, S. P. MALLICK et D. J. KRIEGMAN : Resolving the Generalized Bas-Relief Ambiguity by Entropy Minimization. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, États-Unis, 2007. [104](#), [106](#), [107](#), [108](#), [113](#), [115](#), [117](#), [118](#), [120](#), [122](#), [123](#), [124](#), [160](#), [246](#), [269](#)
- [10] R. ANDERSON, B. STENGER et R. CIPOLLA : Augmenting depth camera output using photometric stereo. *In Machine Vision Applications*, pages 369–372, Nara, Japon, 2011. [2](#), [66](#), [263](#)
- [11] R. ANDERSON, B. STENGER et R. CIPOLLA : Color photometric stereo for multi-colored surfaces. *In IEEE International Conference on Computer Vision (ICCV)*, pages 2182–2189, Barcelone, Espagne, 2011. [246](#), [263](#)

- [12] M.-E. ANGELOPOULOU et M. PETROU : Uncalibrated flatfielding and illumination vector estimation for photometric stereo face reconstruction. *Machine Vision and Applications*, 25(5):1317–1332, 2014. [38](#)
- [13] T. AOTO, T. TAKETOMI, T. SATO, Y. MUKAIGAWA et N. YOKOYA : Position estimation of near point light sources using a clear hollow sphere. In *International Conference on Pattern Recognition (ICPR)*, pages 3721–3724, Tsukuba, Japon, 2012. [181](#)
- [14] H. ATTOUCH, G. BUTTAZZO et G. MICHAILLE : *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. SIAM, 2ème édition, 2014. [59](#)
- [15] G. AUBERT et P. KORNPORST : *Mathematical Problems in Image Processing*, volume 147 de *Applied Mathematical Sciences*. Springer New York, 2006. [59](#), [104](#)
- [16] J.-F. AUJOL, G. GILBOA, T. CHAN et S. OSHER : Structure-Texture Image Decomposition—Modeling, Algorithms, and Parameter Selection. *International Journal of Computer Vision*, 67(1):111–136, 2006. [41](#)
- [17] H. BADRI, H. YAHIA et D. ABOUTAJDINE : Robust Surface Reconstruction via Triple Sparsity. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2291–2298, Columbus, États-Unis, 2014. [60](#), [87](#)
- [18] M. BÄHR et M. BREUSS : An improved eikonal method for surface normal integration. In *German Conference on Pattern Recognition (GCPR)*, volume 9358 de *Lecture Notes in Computer Science*, pages 274–284. Aix-la-Chapelle, Allemagne, 2015. [59](#), [64](#)
- [19] J. T. BARRON et J. MALIK : Shape, albedo, and illumination from a single image of an unknown object. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, États-Unis, 2012. [5](#)
- [20] S. BARSKY et M. PETROU : The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003. [43](#), [52](#), [128](#), [149](#)
- [21] R. H. BARTELS et G. W. STEWART : Solution of the matrix equation $AX + XB = C$. *Communications of the ACM*, 15(9):820–826, 1972. [65](#)
- [22] A. BARTOLI, Y. GÉRARD, F. CHADEBECQ, T. COLLINS et D. PIZARRO : Shape-from-template. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(10):2099–2118, 2015. [2](#), [129](#)
- [23] R. BASRI et D. W. JACOBS : Lambertian reflectance and linear subspaces. In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 383–390, Vancouver, Canada, 2001. [40](#), [180](#)
- [24] R. BASRI et D. W. JACOBS : Photometric stereo with general, unknown lighting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 374–381, Kauai, États-Unis, 2001. [180](#)

-
- [25] R. BASRI et D. W. JACOBS : Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003. 180
 - [26] R. BASRI, D. W. JACOBS et I. KEMELMACHER : Photometric Stereo with General, Unknown Lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007. 180
 - [27] A. BECK et M. TEBoulLE : A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009. 84, 105, 254
 - [28] P. N. BELHUMEUR, D. J. KRIEGMAN et A. L. YUILLE : The Bas-Relief Ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999. 97, 101, 102, 113
 - [29] R. BERGMANN, R. H. CHAN, R. HIELSCHER, J. PERSCH et G. STEIDL : Restoration of Manifold-Valued Images by Half-Quadratic Minimization. *Preprint, ArXiv #1505.07029*, 2015. (soumis). 160
 - [30] M. BERTALMIO, G. SAPIRO, V. CASELLES et C. BALLESTER : Image inpainting. In *ACM SIGGRAPH*, pages 417–424, La Nouvelle-Orléans, États-Unis, 2000. 227
 - [31] L. BI, Z. SONG et L. XIE : A novel LCD based photometric stereo method. In *IEEE International Conference on Information Science and Technology (ICIST)*, pages 611–614, Kopaonik, Croatie, 2014. 173, 174
 - [32] C. M. BISHOP : *Pattern Recognition and Machine Learning*. Springer New York, 2006. 98
 - [33] J. F. BLINN : Models of light reflection for computer synthesized pictures. In *ACM SIGGRAPH*, volume 11, pages 192–198, San Jose, États-Unis, 1977. 213
 - [34] S. BOYD, N. PARIKH, E. CHU, B. PELEATO et J. ECKSTEIN : Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011. 83, 84, 105, 240
 - [35] Y. BOYKOV, O. VEKSLER et R. ZABIH : Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001. 144
 - [36] M. BRADY et A. L. YUILLE : An extremum principle for shape from contour. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(3):288–301, 1984. 2
 - [37] K. BREDIES, K. KUNISCH et T. POCK : Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010. 85
 - [38] X. BRESSON et T. CHAN : Fast Dual Minimization of the Vectorial Total Variation Norm and Applications to Color Image Processing. *Inverse Problems and Imaging*, 2(4):455–484, 2008. 105, 261, 279
 - [39] M. BREUSS, E. CRISTIANI, J.-D. DUROU, M. FALCONE et O. VOGEL : Perspective Shape from Shading: Ambiguity Analysis and Numerical Approximations. *SIAM Journal on Imaging Sciences*, 5(1):311–342, 2012. 30

- [40] B. BRINGIER, A. BONY et M. KHOUDEIR : Specularity and shadow detection for the multisource photometric reconstruction of a textured surface. *Journal of the Optical Society of America A*, 29(1):11–21, 2012. [43](#)
- [41] B. L. BUZBEE, G. H. GOLUB et C. W. NIELSON : On Direct Methods for Solving Poisson’s Equations. *SIAM Journal on Numerical Analysis*, 7(4):627–656, 1970. [64](#)
- [42] E. J. CANDÈS et T. TAO : Decoding by Linear Programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005. [45](#)
- [43] E. J. CANDÈS, M. B. WAKIN et S. P. BOYD : Enhancing Sparsity by Reweighted L1 Minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, 2008. [46](#)
- [44] J. CHABROWSKI et Z. KEWEI : On variational approach to photometric stereo. *Annales de l’Institut Henri Poincaré (C) Analyse non linéaire*, 10(4):363–375, 1993. [225](#)
- [45] A. CHAMBOLLE : An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision*, 20(1):89–97, 2004. [105](#)
- [46] A. CHAMBOLLE et P.-L. LIONS : Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997. [104](#)
- [47] A. CHAMBOLLE et T. POCK : A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2010. [84](#)
- [48] A. CHAMBOLLE, T. POCK, V. CASELLES, M. NOVAGA et D. CREMERS : An introduction to total variation for image analysis. *Theoretical Foundations and Numerical Methods for Sparse Recovery*, 9:263–340, 2010. [227](#)
- [49] T. F. CHAN et S. ESEDOGLU : Aspects of Total Variation Regularized L1 Function Approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817–1837, 2005. [226](#)
- [50] T. F. CHAN et C. K. WONG : Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3):370–375, 1998. [227](#), [282](#)
- [51] M. CHANDRAKER, S. AGARWAL et D. J. KRIEGMAN : ShadowCuts: Photometric Stereo with Shadows. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, États-Unis, 2007. [43](#), [128](#), [149](#)
- [52] M. CHANDRAKER, J. BAI et R. RAMAMOORTHY : A theory of differential photometric stereo for unknown isotropic BRDFs. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2505–2512, Colorado Springs, États-Unis, 2011. [34](#), [193](#), [225](#), [229](#)
- [53] M. CHANDRAKER, J. BAI et R. RAMAMOORTHY : On differential photometric reconstruction for unknown, isotropic BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2941–2955, 2013. [225](#), [229](#)

-
- [54] M. K. CHANDRAKER, F. KAHL et D. J. KRIEGMAN : Reflections on the Generalized Bas-Relief Ambiguity. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 788–795, Pékin, Chine, 2005. 102
 - [55] P. CHARBONNIER : *Reconstruction d'image : Régularisation avec prise en compte des discontinuités*. Thèse de doctorat, Sophia Antipolis, 1994. 47
 - [56] P. CHARBONNIER, L. BLANC-FERAUD, G. AUBERT et M. BARLAUD : Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2):298–311, 1997. 47
 - [57] S. S. CHEN, D. L. DONOHO et M. A. SAUNDERS : Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998. 84
 - [58] Y. CHEN, T. A. DAVIS, W. W. HAGER et S. RAJAMANICKAM : Algorithm 887: CHOLMOD, Supernodal Sparse Cholesky Factorization and Update/Downdate. *ACM Transactions on Mathematical Software*, 35(3):1–14, 2008. 74, 75
 - [59] H. J. CHUNG et J. JIA : Efficient photometric stereo on glossy surfaces with wide specular lobes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, États-Unis, 2008. 213, 215
 - [60] J. J. CLARK : Active photometric stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 29–34, Champaign, États-Unis, 1992. 43, 165, 168
 - [61] J. J. CLARK : Photometric stereo using LCD displays. *Image and Vision Computing*, 28(4):704–714, 2010. 165, 174, 177, 178, 179
 - [62] J. J. CLARK et H. PEKAU : An integral formulation for differential photometric stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 119–124, Fort Collins, États-Unis, 1999. 165, 170
 - [63] E. N. COLEMAN et R. JAIN : Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, 1982. 43
 - [64] T. COLLINS et A. BARTOLI : 3D Reconstruction in Laparoscopy with Close-Range Photometric Stereo. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 7511 de *Lecture Notes in Computer Science*, pages 634–642. Nice, France, 2012. 227
 - [65] T. COLLINS et A. BARTOLI : Towards live monocular 3d laparoscopy using shading and specular information. In *Information Processing in Computer-Assisted Interventions (IPCAI)*, pages 11–21, Pise, Italie, 2012. 275
 - [66] T. COLLINS et A. BARTOLI : Realtime shape-from-template: System and applications. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 116–119, 2015. 1
 - [67] COLLINS, T. AND DUROU, J.-D. AND GURDJOS, P. AND BARTOLI, A. : Single-view perspective shape-from-texture with focal length estimation: A piecewise affine approach. Marseille, France, 2010. 129

- [68] F. COURTEILLE : *Vision monoculaire: contributions théoriques et application à la numérisation des documents*. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, 2006. 6, 285
- [69] E. CRISTIANI : 3D Printers: A new challenge for mathematical modeling. 2014. arXiv preprint 1409.1714. 4
- [70] A. P. DEMPSTER, N. M. LAIRD et D. B. RUBIN : Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1):1–38, 1977. 43
- [71] X. DESCOMBES, R. D. MORRIS, J. ZERUBIA et M. BERTHOD : Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood. *IEEE Transactions on Image Processing*, 8(7):954–963, 1999. 145
- [72] B. DONG, K.-D. MOORE, W. ZHANG et P. PEERS : Scattering Parameters and Surface Normals from Homogeneous Translucent Materials Using Photometric Stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2299–2306, Columbus, États-Unis, 2014. 227, 282
- [73] O. DRBOHLAV et M. CHANTLER : Can Two Specular Pixels Calibrate Photometric Stereo? In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1850–1857, Pékin, Chine, 2005. 102
- [74] O. DRBOHLAV et R. SARA : Specularities reduce ambiguity of uncalibrated photometric stereo. In *European Conference on Computer Vision (ECCV)*, volume 2351 de *Lecture Notes in Computer Science*, pages 46–60. Copenhagen, Danemark, 2002. 102
- [75] F. DROUET, C. STOLZ, O. LALIGANT et O. AUBRETON : 3D reconstruction of external and internal surfaces of transparent objects from polarization state of highlights. *Optics Letters*, 39(10):2955–2958, 2014. 212
- [76] Z. DU, A. ROBLES-KELLY et F. LU : Robust Surface Reconstruction from Gradient Field Using the L1 Norm. In *9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications (DICTA)*, pages 203–209, Adélaïde, Australie, 2007. 60, 81
- [77] B. DURIX, Y. QUÉAU, V. CHARVILLAT et J.-D. DUROU : Quels prétraitements pour la stéréophotométrie non calibrée? In *ORASIS*, Cluny, France, 2013. 35, 43, 78, 95, 97
- [78] J.-D. DUROU : *Shape from shading - Éclairages, réflexions et perspectives*. Habilitation à diriger des recherches, Université Paul Sabatier, Toulouse, France, 2007. 3, 16, 25, 165, 166, 276
- [79] J.-D. DUROU : Reconstruction 3D à partir des ombrages. In *Problèmes inverses en imagerie et en vision*, volume 2 de *Traité Signal et Image, IC2*, chapitre 10, pages 353–390. Lavoisier, 2009. 3
- [80] J.-D. DUROU, J.-F. AUJOL et F. COURTEILLE : Integrating the normal field of a surface in the presence of discontinuities. In *Energy Minimization Methods in*

- Computer Vision and Pattern Recognition (EMMCVPR)*, volume 5681 de *Lecture Notes in Computer Science*, pages 261–273. Bonn, Allemagne, 2009. 47, 60, 80, 81, 85, 88
- [81] J.-D. DUROU et F. COURTEILLE : Integration of a Normal Field without Boundary Condition. In *Proceedings of the First International Workshop on Photometric Analysis For Computer Vision (ICCV Workshops)*, Rio de Janeiro, Brésil, 2007. 62, 65, 66, 67, 72, 75, 76
- [82] J.-D. DUROU, M. FALCONE et M. SAGONA : Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding*, 109(1):22–43, 2008. 117, 160, 225
- [83] J.-D. DUROU, Y. QUÉAU et V. CHARVILLAT : Résolution de la stéréophotométrie par apprentissage. In *Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, Lyon, France, 2012. 163, 194
- [84] R. EPSTEIN, A. L. YUILLE et P. N. BELHUMEUR : Learning and Recognizing Objects Using Illumination Subspaces. In *International Workshop on Object Representation for Computer Vision (ECCV Workshops)*, Cambridge, Royaume-Uni, 1996. 37, 98
- [85] G. EREN, O. AUBRETON, F. MERIAUDEAU, L. A. SANCHEZ SECADES, D. FOFI, A. TEOMAN NASKALI, F. TRUCHETET et A. ERCIL : Scanning from heating: 3D shape estimation of transparent objects from local surface heating. *Optics Express*, 17(14):11457–11468, 2009. 212
- [86] E. ESSER : Applications of Lagrangian-based alternating direction methods and connections to split Bregman. UCLA CAM Report 09-31, 2009. 83
- [87] M. FALCONE et M. SAGONA : An algorithm for the global solution of the shape-from-shading model. In *Image Analysis and Processing*, volume 1310 de *Lecture Notes in Computer Science*, pages 596–603. Florence, Italie, 1997. 160
- [88] M. FALCONE, M. SAGONA et A. SEGhini : A global algorithm for the Shape-from-Shading model with “black shadows”. In *Numerical Mathematics and Advanced Applications - ENUMATH 2001*, pages 503–512. Ischia, Italie, 2003. 28
- [89] P. FAVARO et T. PAPADHIMITRI : A Closed-Form Solution to Uncalibrated Photometric Stereo via Diffuse Maxima. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 821–828, Providence, États-Unis, 2012. 102, 117, 118, 120, 121, 122, 123, 269
- [90] R. A. FINKEL et J. L. BENTLEY : Quad trees: a data structure for retrieval on composite keys. *Acta informatica*, 4(1):1–9, 1974. 174, 178
- [91] D. A. FORSYTH : Variable-Source Shading Analysis. *International Journal of Computer Vision*, 91(3):280–302, 2010. 279
- [92] G. E. FORSYTHE et G. H. GOLUB : On the stationary values of a second-degree polynomial on the unit sphere. *Journal of the Society for Industrial & Applied Mathematics*, 13(4):1050–1068, 1965. 143

- [93] R. T. FRANKOT et R. CHELLAPPA : A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):439–451, 1988. 7, 58, 61, 62, 63, 64
- [94] N. FUNK et Y.-H. YANG : Using a raster display for photometric stereo. In *Canadian Conference on Computer and Robot Vision (CRV)*, pages 201–207, Montréal, Canada, 2007. 173, 174, 177
- [95] S. GALLIANI, M. BREUSS et Y.-C. JU : Fast and Robust Surface Normal Integration by a Discrete Eikonal Equation. In *British Machine Vision Conference (BMVC)*, Surrey, Royaume-Uni, 2012. 59, 60, 64, 75, 76
- [96] S. GANAN et D. MCCLURE : Bayesian image analysis : An application to single photon emission tomography. In *Statistical Computation Section*, pages 12–18, Washington DC, États-Unis, 1985. 47
- [97] J. GENG : Structured-light 3D surface imaging; a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011. 2
- [98] A. S. GEORGHIADES : Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 816–823, Nice, France, 2003. 103, 125, 213
- [99] A. S. GEORGHIADES, D. J. KRIEGMAN et P. N. BELHUMEUR : From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001. 102, 121
- [100] B. GOLDLUECKE, E. STREKALOVSKIY et D. CREMERS : The Natural Vectorial Total Variation Which Arises from Geometric Measure Theory. *SIAM Journal on Imaging Sciences*, 5(2):537–563, 2012. 105, 261, 279
- [101] D. B. GOLDMAN, B. CURLESS, A. HERTZMANN et S. M. SEITZ : Shape and spatially-varying BRDFs from photometric stereo. In *IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 341–348, Pékin, Chine, 2005. 193
- [102] D. B. GOLDMAN, B. CURLESS, A. HERTZMANN et S. M. SEITZ : Shape and spatially-varying BRDFs from photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(6):1060–1071, 2010. 193
- [103] T. GOLDSTEIN, B. O'DONOGHUE, S. SETZER et R. BARANIUK : Fast Alternating Direction Optimization Methods. *SIAM Journal on Imaging Sciences*, 7(3):1588–1623, 2014. 84
- [104] T. GOLDSTEIN et S. OSHER : The Split Bregman Method for L1-Regularized Problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009. 83, 105
- [105] G. H. GOLUB et C. F. VAN LOAN : *Matrix Computations*. JHU Press, 4ème édition, 2012. 64, 142, 143

-
- [106] J. GUÉNARD, G. MORIN, F. BOUDON et V. CHARVILLAT : Reconstructing Plants in 3D from a Single Image Using Analysis-by-Synthesis. *In International Symposium on Visual Computing*, volume 8033 de *Lecture Notes in Computer Science*, pages 322–332. Rethymnon, Grèce, 2013. 285
 - [107] S. M. HAQUE, A. CHATTERJEE et V. M. GOVINDU : High quality photometric reconstruction using a depth camera. *In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2283–2290, Columbus, États-Unis, 2014. 66
 - [108] K. HARA, K. NISHINO et K. IKEUCHI : Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):493–505, 2005. 181
 - [109] M. HARKER et P. O’LEARY : Least squares surface reconstruction from measured gradient fields. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, États-Unis, 2008. 60, 61, 62, 64, 67, 82
 - [110] M. HARKER et P. O’LEARY : Least squares surface reconstruction from gradients: Direct algebraic methods with spectral, Tikhonov, and constrained regularization. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2529–2536, Colorado Springs, États-Unis, 2011. 61, 66
 - [111] M. HARKER et P. O’LEARY : MATLAB toolbox for the regularized surface reconstruction from gradients. *In International Conference on Quality Control by Artificial Vision (QCAV)*, volume 9534 de *Proceedings of SPIE*, Le Creusot, France, 2015. 61
 - [112] M. HARKER et P. O’LEARY : Regularized Reconstruction of a Surface from its Measured Gradient Field. *Journal of Mathematical Imaging and Vision*, 51(1):46–70, 2015. 61, 62, 71
 - [113] R. I. HARTLEY et A. ZISSERMAN : *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2ème édition, 2004. 2
 - [114] H. HAYAKAWA : Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America A*, 11(11):3079–3089, 1994. 38, 97, 98, 100, 102, 108, 114, 269
 - [115] T. HEBERT et R. LEAHY : A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Transactions on Medical Imaging*, 8(2):194–202, 1989. 47
 - [116] C. HERNÁNDEZ : *Stereo and Silhouette Fusion for 3D Object Modeling from Uncalibrated Images Under Circular Motion*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, 2004. 2
 - [117] C. HERNÁNDEZ, G. VOGIATZIS, G. J. BROSTOW, B. STENGER et R. CIPOLLA : Non-rigid Photometric Stereo with Colored Lights. *In IEEE International Conference on Computer Vision (ICCV)*, Rio de Janeiro, Brésil, 2007. 18, 33, 52, 151

- [118] C. HERNÁNDEZ, G. VOGIATZIS et R. CIPOLLA : Multiview Photometric Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):548–554, 2008. 2
- [119] C. HERNÁNDEZ, G. VOGIATZIS et R. CIPOLLA : Overcoming shadows in 3-source photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(2):419–426, 2011. 128, 132, 134, 149, 150, 153, 160, 229
- [120] C. D. HERRERA, J. KANNALA et J. HEIKKILÄ : Joint depth and color camera calibration with distortion correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10):2058–2064, 2012. 284
- [121] A. HERTZMANN et S. M. SEITZ : Shape and materials by example: a photometric stereo approach. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 533–540, Madison, États-Unis, 2003. 104, 163, 193, 194, 197, 199, 207
- [122] A. HERTZMANN et S. M. SEITZ : Example-based photometric stereo: shape reconstruction with general, varying BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005. 36, 104, 163, 193, 194, 197, 199, 201, 207
- [123] J. HO, J. LIM, M.-H. YANG et D. J. KRIEGMAN : Integrating surface normal vectors using fast marching method. In *European Conference on Computer Vision (ECCV)*, volume 3953 de *Lecture Notes in Computer Science*, pages 239–250. Graz, Autriche, 2006. 59, 64
- [124] M. HONG, Z.-Q. LUO et M. RAZAVIYAYN : Convergence Analysis of Alternating Direction Method of Multipliers for a Family of Nonconvex Problems. *Arxiv pre-print 1410.1390*, 2014. 86
- [125] B. K. P. HORN : *Shape from Shading: a Method for Obtaining the Shape of a Smooth Opaque Object from One View*. PhD Thesis, MIT, Cambridge, États-Unis, 1970. 2, 3
- [126] B. K. P. HORN et M. J. BROOKS : The variational approach to shape from shading. *Computer Vision, Graphics, and Image Processing*, 33(2):174–208, 1986. 7, 58, 60, 62, 66, 67, 72, 85, 139
- [127] B. K. P. HORN et R. W. SJOBERG : Calculating the reflectance map. *Applied Optics*, 18(11):1770–1779, 1979. 16
- [128] I. HOROVITZ et N. KIRYATI : Depth from gradient fields and control points: Bias correction in photometric stereo. *Image and Vision Computing*, 22(9):681–694, 2004. 66, 67, 73, 75, 76, 97
- [129] X. HUANG, M. WALTON, G. BEARMAN et O. COSSAIRT : Near Light Correction for Image Relighting and 3D Shape Recovery. In *Digital Heritage*, Grenade, Espagne, 2015. 38, 190
- [130] J. H. HUBER : *Robust Statistics*. Wiley, 1981. 19, 47

-
- [131] S. IKEHATA, D. WIPF, Y. MATSUSHITA et K. AIZAWA : Robust photometric stereo using sparse regression. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 318–325, Providence, États-Unis, 2012. [46](#), [128](#), [149](#), [242](#)
 - [132] K. IKEUCHI et B. K. P. HORN : Numerical Shape from Shading and Occluding Boundaries. *Artificial Intelligence*, 17(1–3):141–184, 1981. [102](#), [128](#), [134](#)
 - [133] K. H. ISCHER et J. A. HERTZ : *Spin glasses*. Cambridge University Press, 1991. [145](#)
 - [134] Y. IWAHORI, H. SUGIE et N. ISHII : Reconstructing shape from shading images under point light source illumination. *In IEEE International Conference on Pattern Recognition (ICPR)*, volume 1, pages 83–87, Atlantic City, États-Unis, 1990. [165](#), [168](#)
 - [135] N. JACOBS, N. ROMAN et R. PLESS : Consistent temporal variations in many outdoor scenes. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, États-Unis, 2007. [154](#)
 - [136] H. JIN, D. CREMERS, D. WANG, E. PRADOS, A. J. YEZZI et S. SOATTO : 3-D Reconstruction of Shaded Objects from Multiple Images Under Unknown Illumination. *International Journal of Computer Vision*, 76(3):245–256, 2008. [225](#)
 - [137] H. JIN, D. CREMERS, A. J. YEZZI et S. SOATTO : Shedding light on stereoscopic segmentation. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 36–42, 2004. [225](#)
 - [138] M. K. JOHNSON, F. COLE, A. RAJ et E. H. ADELSON : Microgeometry capture using an elastomeric sensor. *ACM Transactions on Graphics*, 30(4):1, 2011. [186](#)
 - [139] M.V. JOSHI et S. CHAUDHURI : Photometric stereo under blurred observations. *In IEEE International Conference on Pattern Recognition (ICPR)*, volume 3, pages 169–172, Cambridge, Royaume-Uni, 2004. [227](#), [282](#)
 - [140] Y.-C. JU, M. BREUSS, A. BRUHN et S. GALLIANI : Shape from Shading for Rough Surfaces: Analysis of the Oren-Nayar Model. *In British Machine Vision Conference (BMVC)*, Surrey, Royaume-Uni, 2012. [209](#)
 - [141] Y.-C. JU, S. TOZZA, M. BREUSS, A. BRUHN et A. KLEEFELD : Generalised Perspective Shape from Shading with Oren-Nayar Reflectance. *In British Machine Vision Conference (BMVC)*, Bristol, Royaume-Uni, 2013. [209](#)
 - [142] J. JUNG, J.-Y. LEE et I. S. KWEON : One-day outdoor photometric stereo via skylight estimation. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, États-Unis, 2015. [127](#)
 - [143] B. KARAÇALI et W. SNYDER : Reconstructing discontinuous surfaces from a given gradient field using partial integrability. *Computer Vision and Image Understanding*, 92(1):78–111, 2003. [64](#)
 - [144] M. KHANIAN, A. BOROUJERDI et M. BREUSS : Perspective photometric stereo beyond Lambert. *In International Conference on Quality Control by Artificial*

- Vision (QCAV)*, volume 9534 de *Proceedings of SPIE*, Le Creusot, France, 2015. 213
- [145] H. KIM, B. WILBURN et M. BEN-EZRA : Photometric stereo for dynamic surface orientations. In *European Conference on Computer Vision (ECCV)*, pages 59–72. Heraklion, Grèce, 2010. 260
- [146] R. KIMMEL et I. YAVNEH : An Algebraic Multigrid Approach for Image Analysis. *SIAM Journal on Scientific Computing*, 24(4):1218–1231, 2003. 66, 77
- [147] J. J. KOENDERINK et A. J. VAN DOORN : The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*, 23(3):217–234, 1997. 113
- [148] J. J. KOENDERINK et A. J. VAN DOORN : Shape and shading. *The visual neurosciences*, pages 1090–1105, 2003. 6
- [149] V. KOLMOGOROV et R. ZABIH : What Energy Functions Can Be Minimized via Graph Cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004. 144, 145
- [150] L. L. KONTSEVICH, A. P. PETROV et I. S. VERGELSKAYA : Reconstruction of shape from shading in color images. *Journal of the Optical Society of America A*, 11(3):1047–1052, 1994. 52, 151, 153
- [151] S. J. KOPPAL et S. G. NARASIMHAN : Appearance derivatives for isonormal clustering of scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(8):1375–1385, 2009. 128, 194, 197, 198
- [152] I. KOUTIS, G. L. MILLER et R. PENG : A Nearly-m log n Time Solver for SDD Linear Systems. In *IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 590–598, Palm Springs, États-Unis, 2011. 69, 71, 72, 73, 75, 76, 80, 84, 89, 250, 251, 252
- [153] P. KOVESI : Shapelets correlated with surface normals produce surfaces. In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 994–1001, Pékin, Chine, 2005. 61
- [154] R. KOZERA : Existence and uniqueness in photometric stereo. *Applied Mathematics and Computation*, 44(1):1–103, 1991. 137, 195
- [155] R. KOZERA : On Shape Recovery from Two Shading Patterns. *International Journal of Pattern Recognition and Artificial Intelligence*, 6(4):673–698, 1993. 127
- [156] B. KUNSBURG et S. W. ZUCKER : How Shading Constrains Surface Patches without Knowledge of Light Sources. *SIAM Journal on Imaging Sciences*, 7(2):641–668, 2014. 286
- [157] J.-F. LALONDE, S. G. NARASIMHAN et A. A. EFROS : What do the sun and the sky tell us about the camera? *International Journal of Computer Vision*, 88:24–51, 2010. 155
- [158] A. LANZA, S. MORIGI et F. SGALLARI : Convex image denoising via non-convex regularization. In *Scale Space and Variational Methods in Computer Vision (SSVM)*,

- volume 9087 de *Lecture Notes in Computer Science*, pages 666–677. Lège Cap-Ferret, France, 2015. 47, 48, 86
- [159] K. C. LEE, J. HO et D. J. KRIEGMAN : Acquiring Linear Subspaces for Face Recognition under Variable Lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5):684–698, 2005. 121
- [160] S. LEE et M. BRADY : Integrating stereo and photometric stereo to monitor the development of glaucoma. *Image and Vision Computing*, 9(1):39–44, 1991. 229
- [161] J. LELLMANN, E. STREKALOVSKIY, S. KOETTER et D. CREMERS : Total Variation Regularization for Functions with Values in a Manifold. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2944–2951, Sidney, Australie, 2013. 160
- [162] D.-H. LI et M. FUKUSHIMA : A modified BFGS method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics*, 129(1-2):15–35, 2001. 236
- [163] M. LIAO, X. HUANG et R. YANG : Interreflection removal for photometric stereo by using spectrum-dependent albedo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 689–696, Colorado Springs, États-Unis, 2011. 32
- [164] Z. LIN, M. CHEN et Y. MA : The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. Technical Report UILU-ENG-09-2215, 2010. 35, 46
- [165] P.-L. LIONS, E. ROUY et A. TOURIN : Shape-from-shading, viscosity solutions and edges. *Numerische Mathematik*, 64(1):323–353, 1993. 6
- [166] S. MAGNÚSSON, P. CHATHURANGA WEERADDANA, M. G. RABBAT et C. FISCHIONE : On the Convergence of Alternating Direction Lagrangian Methods for Nonconvex Structured Optimization Problems. *Arxiv preprint 1409.8033*, 2014. 86
- [167] S. P. MALLICK, T. E. ZICKLER, D. J. KRIEGMAN et P. N. BELHUMEUR : Beyond Lambert: Reconstructing Specular Surfaces Using Color. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 619–626, San Diego, États-Unis, 2005. 32, 34, 38
- [168] W. F. MASCARENHAS : The BFGS method with exact line searches fails for non-convex objective functions. *Mathematical Programming*, 99(1):49–61, 2004. 236
- [169] S. MASNOU et J.-M. MOREL : Level lines based disocclusion. In *IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 259–263, Chicago, États-Unis, 1998. 227
- [170] G. MCGUNNIGLE et M. J. CHANTLER : Resolving handwriting from background printing using photometric stereo. *Pattern Recognition*, 36(8):1869–1879, 2003. 38, 40

- [171] G. MCGUNNIGLE, J. DONG et X. WANG : Photometric stereo applied to diffuse surfaces that violate Lambert’s law. *Journal of the Optical Society of America A*, 29(4):627–636, 2012. [210](#)
- [172] R. MECCA et J.-D. DUROU : Unambiguous photometric stereo using two images. *In International Conference on Image Analysis and Processing (ICIAP)*, volume 6978 de *Lecture Notes in Computer Science*, pages 286–295. Ravenna, Italie, 2011. [127](#), [136](#), [195](#)
- [173] R. MECCA et M. FALCONE : Uniqueness and Approximation of a Photometric Shape-from-Shading Model. *SIAM Journal on Imaging Sciences*, 6(1):616–659, 2013. [127](#), [130](#), [137](#), [229](#)
- [174] R. MECCA, E. RODOLÀ et D. CREMERS : Realistic photometric stereo using partial differential irradiance equation ratios. *Computers & Graphics*, 51:8–16, 2015. [213](#), [225](#)
- [175] R. MECCA, A. WETZLER, A. M. BRUCKSTEIN et R. KIMMEL : Near Field Photometric Stereo with Point Light Sources. *SIAM Journal on Imaging Sciences*, 7(4):2732–2770, 2014. [170](#), [225](#), [229](#), [271](#), [274](#)
- [176] F. MERIAUDEAU, L. A. SANCHEZ SECADES, G. EREN, A. ERCIL, F. TRUCHETET, O. AUBRETON et D. FOFI : 3-D Scanning of Nonopaque Objects by Means of Imaging Emitted Structured Infrared Patterns. *IEEE Transactions on Instrumentation and Measurement*, 59(11):2898–2906, 2010. [212](#)
- [177] T. MIGITA, S. OGINO et T. SHAKUNAGA : Direct bundle estimation for recovery of shape, reflectance property and light position. *In European Conference on Computer Vision (ECCV)*, volume 5304 de *Lecture Notes in Computer Science*, pages 412–425. Marseille, France, 2008. [168](#)
- [178] D. MIYAZAKI et K. IKEUCHI : Photometric stereo under unknown light sources using robust SVD with missing data. *In IEEE International Conference on Image Processing (ICIP)*, pages 4057–4060, Hong-Kong, Chine, 2010. [35](#), [100](#), [268](#)
- [179] R. MONTES SOLDADO et C. UREÑA ALMAGRO : An Overview of BRDF Models. Technical Report, LSI-2012-001, 2012. [22](#), [193](#)
- [180] T. MOONS, L. VAN GOOL et M. VERGAUWEN : 3D Reconstruction from Multiple Images. *Foundations and Trends in Computer Graphics and Vision*, 4(4):287–404, 2008. [2](#)
- [181] I. MORENO, M. AVENDAÑO-ALEJO et R. I. TZONCHEV : Designing light-emitting diode arrays for uniform near-field irradiance. *Applied Optics*, 45(10):2265–2272, 2006. [170](#)
- [182] G. MORIN : *3D models for...* Habilitation à diriger des recherches, Université de Toulouse, Toulouse, France, 2014. [1](#)
- [183] D. MUMFORD : Geometry-Driven Diffusion in Computer Vision. Computational Imaging and Vision, chapitre Bayesian Rationale for the Variational Formulation, pages 135–146. Springer Netherlands, 1994. [245](#)

-
- [184] Z. MUREZ, T. TREIBITZ, R. RAMAMOORTHY et D. J. KRIEGMAN : Photometric stereo in a scattering medium. *In IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chili, 2015. 278
 - [185] G. NAM et M. H. KIM : Multispectral photometric stereo for acquiring high-fidelity surface normals. *IEEE Computer Graphics and Applications*, 34(6):57–68, 2014. 32, 260
 - [186] B. K. NATARAJAN : Sparse Approximate Solutions to Linear Systems. *SIAM Journal on Computing*, 24(2):227–234, 1995. 45
 - [187] S. K. NAYAR, K. IKEUCHI et T. KANADE : Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, 1990. 213
 - [188] S. K. NAYAR et Y. NAKAGAWA : Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):824–831, 1994. 2
 - [189] F. E. NICODEMUS, J. C. RICHMOND, J. J. HSIA, I. W. GINSBERG et T. LIMPERIS : Geometrical considerations and nomenclature for reflectance. NBS Monograph 160, National Bureau of Standards, Washington D.C., États-Unis, 1977. 165
 - [190] V. NOZICK : Pyramidal normal map integration for real-time photometric stereo. *In EAM Mechatronics*, pages 128–132, Yokohama, Japon, 2010. 173, 174, 176
 - [191] J. OLIENSIS : Uniqueness in Shape from Shading. *International Journal of Computer Vision*, 6(2):75–104, 1991. 132
 - [192] R. ONN et A. M. BRUCKSTEIN : Integrability disambiguates surface recovery in two-image photometric stereo. *International Journal of Computer Vision*, 5(1): 105–113, 1990. 59, 127, 136, 139, 195
 - [193] R. OR-EL, G. ROSMAN, A. WETZLER, R. KIMMEL et A. M. BRUCKSTEIN : RGBD-Fusion: Real-Time High Precision Depth Recovery. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, États-Unis, 2015. 40, 66, 74
 - [194] M. OREN et S. K. NAYAR : Generalization of Lambert’s reflectance model. *In ACM SIGGRAPH*, pages 239–246, Orlando, États-Unis, 1994. 208
 - [195] M. OREN et S. K. NAYAR : Generalization of the Lambertian model and implications for machine vision. *International Journal of Computer Vision*, 14(3):227–251, 1995. 208
 - [196] T. PAPADHIMITRI et P. FAVARO : A Closed-Form, Consistent and Robust Solution to Uncalibrated Photometric Stereo Via Local Diffuse Reflectance Maxima. *International Journal of Computer Vision*, 107(2):139–154, 2013. 102, 118
 - [197] T. PAPADHIMITRI et P. FAVARO : A New Perspective on Uncalibrated Photometric Stereo. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1474–1481, Portland, États-Unis, 2013. 58, 125, 160, 265, 267, 269
 - [198] T. PAPADHIMITRI et P. FAVARO : Uncalibrated Near-Light Photometric Stereo. *In British Machine Vision Conference (BMVC)*, Nottingham, Royaume-Uni, 2014. 125, 168, 190, 274

- [199] N. PARIKH et S. BOYD : Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):127–239, 2014. [46](#), [84](#), [279](#)
- [200] R. PENROSE et J. A. TODD : On best approximate solutions of linear matrix equations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 52(01):17–19, 1956. [142](#)
- [201] A. P. PENTLAND : Local Shading Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):170–187, 1984. [5](#)
- [202] P. PERONA et J. MALIK : Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990. [77](#)
- [203] M. PERRIOLLAT et A. BARTOLI : A computational model of bounded developable surfaces with application to image-based three-dimensional reconstruction. *Computer Animation and Virtual Worlds*, 24(5):459–476, 2013. [6](#)
- [204] M. PETROU et S. BARSKY : Shadows and highlights detection in 4-source colour photometric stereo. In *IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 967–970, Thessalonique, Grèce, 2001. [43](#)
- [205] H.-T. PHAM, H. HE, F. VERNIER, E. TROUVÉ, L. BENOIT, L. MOREAU et B. GIRARD : Analyse de "Time-Lapse" stéréo pour la mesure de déformation 3D, application au suivi du glacier d'Argentière. In *Reconnaissance de Formes et Intelligence Artificielle (RFIA)*, Reims, France, 2014. [155](#)
- [206] B. T. PHONG : Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975. [22](#), [124](#)
- [207] M. W. POWELL, S. SARKAR et D. GOLDFOG : Calibration of light sources. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 263–269, Hilton Head, États-Unis, 2000. [181](#)
- [208] E. PRADOS et O. FAUGERAS : Shape from Shading: A Well-Posed Problem? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 870–877, San Diego, États-Unis, 2005. [8](#), [95](#), [165](#), [168](#)
- [209] Y. QUÉAU et J.-D. DUROU : Résolution du problème de la stéréophotométrie non calibrée par estimation de l'intensité des éclairages. In *ORASIS*, Cluny, France, 2013. [95](#), [97](#), [114](#)
- [210] Y. QUÉAU et J.-D. DUROU : Intégration d'un champ de gradient rapide et robuste aux discontinuités - Application à la stéréophotométrie. In *Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, Rouen, France, 2014. [55](#), [77](#)
- [211] Y. QUÉAU et J.-D. DUROU : Edge-Preserving Integration of a Normal Field: Weighted Least Squares, TV and L1 Approaches. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 9087 de *Lecture Notes in Computer Science*, pages 576–588, Lège Cap-Ferret, France, 2015. [55](#), [60](#), [75](#), [77](#), [85](#)

-
- [212] Y. QUÉAU et J.-D. DUROU : Some Illumination Models for Industrial Applications of Photometric Stereo. In *International Conference on Quality Control by Artificial Vision (QCAV)*, volume 9534 de *Proceedings of SPIE*, Le Creusot, France, 2015. 163, 165, 168, 190, 271, 273
 - [213] Y. QUÉAU, J.-D. DUROU et X. DESCOMBES : Que peut-on apprendre d'une scène vue par une webcam à partir d'images prises au cours d'une journée ensoleillée ? In *ORASIS*, Amiens, France, 2015. 95, 127, 142, 154
 - [214] Y. QUÉAU, J.-D. DUROU, B. DURIX et V. CHARVILLAT : Stéréophotométrie non calibrée en présence d'écarts au modèle lambertien. *Traitement du Signal*, 31(1-2):107–141, 2014. 35, 43, 95, 97, 108, 114
 - [215] Y. QUÉAU, F. LAUZE et J.-D. DUROU : Solving the Uncalibrated Photometric Stereo Problem using Total Variation. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, numéro 7893 de LNCS, pages 270–281, Schloss Seggau, Autriche, 2013. 95, 97
 - [216] Y. QUÉAU, F. LAUZE et J.-D. DUROU : A L1-TV Algorithm for Robust Perspective Photometric Stereo with Spatially-Varying Lightings. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 9087 de *Lecture Notes in Computer Science*, pages 498–510, Lège Cap-Ferret, France, 2015. 165, 180, 236, 279, 280
 - [217] Y. QUÉAU, F. LAUZE et J.-D. DUROU : Solving Uncalibrated Photometric Stereo using Total Variation. *Journal of Mathematical Imaging and Vision*, 52(1):87–107, 2015. 95, 97, 109, 111, 116, 269
 - [218] Y. QUÉAU, R. MECCA, J.-D. DUROU et X. DESCOMBES : Photometric Stereo with Only Two Images: A Theoretical Study and Numerical Resolution. (*soumis*), 2015. 95, 127, 195
 - [219] Y. QUÉAU, R. MODRZEJEWSKI, P. GURDJOS et J.-D. DUROU : Transformation d'un dispositif multimédia webcam-écran en un scanner 3D. In *COmpression et REprésentation des Signaux Audiovisuels (CORESA)*, Reims, France, 2014. 165, 170
 - [220] Y. QUÉAU, R. MODRZEJEWSKI, P. GURDJOS et J.-D. DUROU : A Full Photometric and Geometric Model for Webcam + Matte Screen Devices. *Signal Processing: Image Communications*, 40:65–81, 2016. 163, 165, 170, 173, 176, 181
 - [221] S. RAHMAN, A. LAM, I. SATO et A. ROBLES-KELLY : Color photometric stereo using a rainbow light for non-lambertian multicolored surfaces. In *Asian Conference on Computer Vision (ACCV)*, volume 9003 de *Lecture Notes in Computer Science*, pages 335–350. Singapour, Singapour, 2014. 263
 - [222] C. RAPOSO, J. P. BARRETO et U. NUNES : Fast and accurate calibration of a kinect sensor. In *IEEE International Conference on 3D Vision (3DV)*, pages 342–349, Seattle, États-Unis, 2013. 284
 - [223] I. REDA et A. ANDREAS : Solar position algorithm for solar radiation applications. *Solar Energy*, 76(5):577–589, 2004. 155

- [224] D. REDDY, A. AGRAWAL et R. CHELLAPPA : Enforcing integrability by error correction using l^1 -minimization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2350–2357, Miami, États-Unis, 2009. [59](#)
- [225] R. T. ROCKAFELLAR : Augmented Lagrangians and Applications of the Proximal Point Algorithm in Convex Programming. *Mathematics of Operations Research*, 1(2):97–116, 1976. [83](#)
- [226] L. I. RUDIN, S. OSHER et E. FATEMI : Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992. [80](#), [81](#), [104](#), [227](#)
- [227] Y. SAAD : *Iterative methods for sparse linear systems*. SIAM, 4ème édition, 2003. [71](#)
- [228] R. F. V. SARACCHINI, J. STOLFI, H. LEITÃO, G. ATKINSON et M. SMITH : Multi-Scale Depth from Slope with Weights. In *British Machine Vision Conference (BMVC)*, Aberystwyth, Royaume-Uni, 2010. [77](#)
- [229] R. F. V. SARACCHINI, J. STOLFI, H. LEITÃO, G. ATKINSON et M. SMITH : A robust multi-scale integration method to obtain the depth from gradient maps. *Computer Vision and Image Understanding*, 116(8):882–895, 2012. [77](#)
- [230] I. SATO, T. OKABE, Q. YU et Y. SATO : Shape Reconstruction Based on Similarity in Radiance Changes under Varying Illumination. In *IEEE International Conference on Computer Vision (ICCV)*, Rio de Janeiro, Brésil, 2007. [198](#)
- [231] Y. SATO et K. IKEUCHI : Reflectance analysis under solar illumination. In *Workshop on Physics-Based Modeling in Computer Vision (ICCV Workshops)*, pages 180–187, Cambridge, États-Unis, 1995. [18](#), [127](#)
- [232] D. SCHARF : *Magnifications: Photography with the Scanning Electron Microscope*. Schocken Books, 1977. [5](#)
- [233] G. SCHINDLER : Photometric stereo via computer screen lighting for real-time surface reconstruction. In *International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, Atlanta, États-Unis, 2008. [173](#), [174](#), [176](#)
- [234] S. A. SHAFER et T. KANADE : Using shadows in finding surface orientations. *Computer Vision, Graphics, and Image Processing*, 22(1):145–176, 1983. [2](#)
- [235] R. SHEFI et M. TEBoulLE : Rate of Convergence Analysis of Decomposition Methods Based on the Proximal Method of Multipliers for Convex Minimization. *SIAM Journal on Optimization*, 24(1):269–297, 2014. [84](#), [217](#), [239](#)
- [236] H. L. SHEN : Calibrating light sources by using a planar mirror. *Journal of Electronic Imaging*, 20(1), 2011. [181](#)
- [237] R. SHIRADKAR, P. TAN et S.-H. ONG : Auto-calibrating photometric stereo using ring light constraints. *Machine Vision and Applications*, 25(3):801–809, 2014. [41](#), [42](#)
- [238] M. W. SILVER : Determining shape and reflectance using multiple images. Master thesis, MIT, 1980. [194](#)

-
- [239] T. SIMCHONY, R. CHELLAPPA et M. SHAO : Direct analytical methods for solving Poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):435–446, 1990. [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [67](#), [69](#), [71](#), [76](#), [91](#), [103](#), [124](#), [141](#)
 - [240] W. SMITH et F. FANG : Height from Photometric Ratio with Model-based Light Source Selection. *Computer Vision and Image Understanding*, 2015. (to appear). [229](#)
 - [241] M. SOUIAI, M. R. OSWALD, Y. KEE, J. KIM, M. POLLEFEYS et D. CREMERS : Entropy minimization for convex relaxation approaches. In *IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chili, 2015. [106](#)
 - [242] S. SUGIMOTO et M. OKUTOMI : Practical low-rank matrix approximation under robust l_1 -norm. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1410–1417, Providence, États-Unis, 2012. [36](#), [38](#)
 - [243] J. SUN, M. SMITH, L. SMITH, L. COUTTS, R. DABIS, C. HARLAND et J. BAMBER : Reflectance of human skin using colour photometric stereo: with particular application to pigmented lesion analysis. *Skin Research and Technology*, 14(2):173–179, 2008. [1](#)
 - [244] J. SUN, M. SMITH, L. SMITH et A. FAROOQ : Sampling Light Field for Photometric Stereo. *International Journal of Computer Theory and Engineering*, 5(1):14–18, 2013. [38](#)
 - [245] J. SUN, M. SMITH, L. SMITH, S. MIDHA et J. 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. [43](#), [128](#), [149](#)
 - [246] K. SUNKAVALLI, T. E. ZICKLER et H. PFISTER : Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. In *European Conference on Computer Vision (ECCV)*, volume 6312 de *Lecture Notes on Computer Science*, pages 251–264. Héraklion, Grèce, 2010. [268](#)
 - [247] V.-T. TA, A. ELMOATAZ et O. LÉZORAY : Nonlocal PDEs-based morphology on weighted graphs for image and data processing. *IEEE Transactions on Image Processing*, 20(6):1504–1516, 2011. [71](#)
 - [248] T. TAKAI, A. MAKI, K. NIINUMA et T. MATSUYAMA : Difference sphere: An approach to near light source estimation. *Computer Vision and Image Understanding*, 113(9):966–978, 2009. [181](#)
 - [249] A. TANKUS et N. KIRYATI : Photometric stereo under perspective projection. In *IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 611–616, Pékin, Chine, 2005. [58](#)
 - [250] D. TENBRINCK, F. LOZES et A. ELMOATAZ : Solving minimal surface problems on surfaces and point clouds. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 9087 de *Lecture Notes in Computer Science*, pages 601–612. Lège Cap-Ferret, France, 2015. [71](#)

- [251] C. TI, R. YANG, K. DAVIS et Z. PAN : Simultaneous Time-of-Flight Sensing and Photometric Stereo With a Single ToF Sensor. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, États-Unis, 2015. [66](#)
- [252] K. E. TORRANCE et E. M. SPARROW : Theory for Off-Specular Reflection From Roughened Surfaces. *Journal of the Optical Society of America A*, 57(9):1105–1114, 1967. [213](#)
- [253] S. TOZZA : *Analysis and Approximation of Non-Lambertian Shape-from-Shading Models*. Tesi, Università La Sapienza, Rome, 2015. [8](#), [209](#)
- [254] S. TOZZA et M. FALCONE : A Semi-Lagrangian Approximation of the Oren–Nayar PDE for the Orthographic Shape–from–Shading Problem. *In International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 711–716, Lisbonne, Portugal, 2014. [209](#)
- [255] J. W. TUKEY : Sampling from contaminated distributions. *In Contributions to probability and statistics*, pages 448–485. Stanford University Press, 1960. [19](#)
- [256] J. W. TUKEY : Robust techniques for the user. *In Robustness in Statistics*. Academic Press, New York, 1979. [19](#)
- [257] F. VERBIEST et L. VAN GOOL : Photometric stereo with coherent outlier handling and confidence estimation. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, États-Unis, 2008. [43](#), [116](#), [128](#), [149](#)
- [258] D. VLASIC, P. PEERS, I. BARAN, P. DEBEVEC, J. POPOVIĆ, S. RUSINKIEWICZ et W. MATUSIK : Dynamic shape capture using multi-view photometric stereo. *ACM Transactions on Graphics*, 28(5):1, 2009. [2](#), [21](#), [128](#)
- [259] Y.-X. WANG, C. M. LEE, L.-F. CHEONG et K.-C. TOH : Practical Matrix Completion and Corruption Recovery Using Proximal Alternating Robust Subspace Minimization. *International Journal of Computer Vision*, 111(3):315–344, 2015. [36](#), [38](#)
- [260] G. J. WARD : Measuring and modeling anisotropic reflection. *In ACM SIGGRAPH*, pages 265–272, Chicago, États-Unis, 1992. [44](#), [213](#)
- [261] T. WEI et R. KLETTE : Depth recovery from noisy gradient vector fields using regularization. *In Computer Analysis of Images and Patterns*, volume 2756 de *Lecture Notes in Computer Science*, pages 116–123. Springer Berlin Heidelberg, Groningen, Pays-Bas, 2003. [61](#)
- [262] A. WEINMANN, L. DEMARET et M. STORATH : Total Variation Regularization for Manifold-Valued Data. *SIAM Journal on Imaging Sciences*, 7(4):2226–2257, 2014. [160](#)
- [263] A. P. WITKIN : Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17(1):17–45, 1981. [2](#)
- [264] L. B. WOLFF et E. ANGELOPOULOU : 3-d stereo using photometric ratios. *In European Conference on Computer Vision (ECCV)*, volume 801 de *Lecture Notes in Computer Science*, pages 247–258. Stockholm, Suède, 1994. [229](#)

-
- [265] J. H. WON, M. H. LEE et I. K. PARK : Active 3D shape acquisition using smart-phones. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 29–34, Providence, États-Unis, 2012. 173, 174
 - [266] W. L. WOO et S. S. DLAY : 3D shape restoration using sparse representation and separation of illumination effects. *Signal Processing*, 103:258–272, 2014. 143
 - [267] R. J. WOODHAM : Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from a Single View. In *Proceedings of the 22nd SPIE Annual Technical Symposium*, volume 155 de *Proceedings of the International Society for Optical Engineering*, pages 136–143, San Diego, États-Unis, 1978. 127
 - [268] R. J. WOODHAM : Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings. Technical Report MIT AITR-457, 1978. 1, 194, 197
 - [269] R. J. WOODHAM : Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):134–144, 1980. 132
 - [270] L. WU, A. GANESH, B. SHI, Y. MATSUSHITA, Y. WANG et Y. MA : Robust photometric stereo via low-rank matrix completion and recovery. In *Asian Conference on Computer Vision (ACCV)*, volume 6494 de *Lecture Notes in Computer Science*, pages 703–717. Queenstown, Nouvelle-Zélande, 2010. 35, 36, 37, 38, 98, 100, 121, 128, 149, 242, 268
 - [271] T.-P. WU et C.-K. TANG : Dense Photometric Stereo Using a Mirror Sphere and Graph Cut. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 140–147, San Diego, États-Unis, 2005. 91, 229
 - [272] T.-P. WU et C.-K. TANG : Photometric stereo via expectation maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):546–560, 2010. 43, 128, 149
 - [273] T.-P. WU, K. L. TANG, C.-K. TANG et T. T. WONG : Dense Photometric Stereo: A Markov Random Field Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1830–1846, 2006. 242, 254, 255, 256, 262, 263, 267, 270, 271, 272
 - [274] Z. WU et L. LI : A line-integration based method for depth recovery from surface normals. *Computer Vision, Graphics, and Image Processing*, 43(1):53–66, 1988. 59, 64
 - [275] L. XIE, Z. SONG et X. HUANG : A novel method for the calibration of an LED-based photometric stereo system. In *IEEE International Conference on Information and Automation (ICIA)*, pages 780–783, Lodz, Pologne, 2013. 181
 - [276] W. XIE, C. DAI et C. C. L. WANG : Photometric Stereo With Near Point Lighting: A Solution by Mesh Deformation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4585–4593, Boston, États-Unis, 2015. 168, 253, 271, 274
 - [277] J. YANG, N. OHNISHI et N. SUGIE : Two image photometric stereo method. In *Intelligent Robots and Computer Vision XI: Biological, Neural Net, and 3D Methods*,

- volume 1826 de *Proceedings of the International Society for Optical Engineering*, pages 452–463, Boston, États-Unis, 1992. [127](#)
- [278] Q. YANG et N. AHUJA : Surface reflectance and normal estimation from photometric stereo. *Computer Vision and Image Understanding*, 116(7):793–802, 2012. [193](#)
- [279] W. YIN et S. OSHER : Error Forgetting of Bregman Iteration. *Journal of Scientific Computing*, 54(2-3):684–695, 2013. [83](#)
- [280] W. YIN, S. OSHER, D. GOLDFARB et J. DARBON : Bregman Iterative Algorithms for ℓ_1 -Minimization with Applications to Compressed Sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, 2008. [83](#)
- [281] L. F. YU, S. K. YEUNG, Y. W. TAI et S. LIN : Shading-Based Shape Refinement of RGB-D Images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1415–1422, Portland, États-Unis, 2013. [66](#)
- [282] A. L. YUILLE et D. SNOW : Shape and albedo from multiple images using integrability. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 158–164, San Juan, États-Unis, 1997. [38](#), [58](#), [99](#), [100](#), [101](#), [102](#), [108](#), [111](#), [114](#), [265](#)
- [283] A. L. YUILLE, D. SNOW, R. EPSTEIN et P. N. BELHUMEUR : Determining Generative Models of Objects Under Varying Illumination: Shape and Albedo from Multiple Images Using SVD and Integrability. *International Journal of Computer Vision*, 35(3):203–222, 1999. [38](#), [58](#), [100](#), [101](#), [102](#), [111](#), [269](#)
- [284] S. ZAMBAL, W. PALFINGER, M. STÖGER et C. EITZINGER : Accurate fibre orientation measurement for carbon fibre surfaces. *Pattern Recognition*, 48(11):3324–3332, 2015. [1](#)
- [285] W. ZHOU et C. KAMBHAMETTU : Estimation of the size and location of multiple area light sources. In *IEEE International Conference on Pattern Recognition (ICPR)*, volume 3, pages 214–217, Cambridge, Royaume-Uni, 2004. [181](#)
- [286] W. ZHOU et C. KAMBHAMETTU : A unified framework for scene illuminant estimation. *Image and Vision Computing*, 26(3):415–429, 2008. [181](#)
- [287] Z. ZHOU et P. TAN : Ring-light photometric stereo. In *European Conference on Computer Vision (ECCV)*, volume 6312 de *Lecture Notes in Computer Science*, pages 265–279. Héraklion, Grèce, 2010. [41](#), [42](#)
- [288] M. ZHU et T. CHAN : An efficient primal-dual hybrid gradient algorithm for total variation image restoration. pages 08–34, 2008. [84](#)
- [289] T. E. ZICKLER, S.P. MALLICK, D. J. KRIEGMAN et P. N. BELHUMEUR : Color Subspaces as Photometric Invariants. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 2000–2010, New York, États-Unis, 2006. [32](#), [34](#), [38](#)
- [290] M. ZOLLHÖFER, A. DAI, M. INNEMANN, C. WU, M. STAMMINGER, C. THEOBALT et M. NIESSNER : Shading-based refinement on volumetric signed distance functions. *ACM Transactions on Graphics*, 34(4):1–14, 2015. [66](#)