Stéréophotométrie non calibrée en présence d'écarts au modèle lambertien

Yvain Quéau, Jean-Denis Durou, Bastien Durix, Vincent Charvillat

Université de Toulouse,
IRIT, UMR CNRS 5505
2 rue Camichel
Toulouse, France
yvain.queau@enseeiht.fr,durou@irit.fr,durix@enseeiht.fr,charvi@enseeiht.fr

RÉSUMÉ. *Nous présentons une méthode robuste de résolution du problème de la stéréophotométrie non calibrée dans le cadre lambertien, en présence d'écarts à ce modèle tels que les ombres ou les taches spéculaires. Nous montrons d'abord comment détecter de tels écarts afin d'estimer de façon plus précise les paramètres caractéristiques des sources lumineuses, qui sont supposées de même intensité afin de lever l'ambiguïté de bas-relief généralisée inhérente au problème. Nous montrons ensuite comment estimer la normale et l'albédo en chaque point en prenant en compte les écarts au modèle, ce qui est facilité par l'estimation précise des éclairages. Nous validons notre méthode qualitativement et quantativement par la reconstruction 3D d'objets synthétiques et réels, et nous comparons son efficacité et sa précision aux meilleures méthodes actuelles.*

ABSTRACT. *In this paper we present a robust method for solving the uncalibrated photometric stereo problem in the Lambertian framework, in presence of outliers such as shadows or specular highlights. We show how to detect such outliers so as to accurately estimate the light sources, which are assumed to have uniform magnitude in order to solve for the generalized bas-relief ambiguity. We then propose a way to recover both the normal and the albedo in every pixel in presence of outliers, which is made possible thanks to the accurate estimation of the light sources. Our method is validated both qualitatively and quantitatively on synthetic and real-world datasets, and we compare its efficiency and accuracy to the most state-of-the-art existing techniques.*

MOTS-CLÉS : *Reconstruction 3D, shape-from-X, stéréophotométrie non calibrée, estimation de l'éclairage, ambiguïté de bas-relief généralisée, ombres, taches spéculaires.*

KEYWORDS: *3D-reconstruction, shape-from-X, uncalibrated photometric stereo, lighting estimation, generalized bas-relief ambiguity, shadows, specular highlights.*

DOI:10.3166/TS.22.1-**??** © 2012 Lavoisier

Résumé

Cet article traite du problème de la reconstruction 3D à partir de $m \geqslant 3$ images, prises depuis le même point de vue mais sous m éclairages différents. Ce problème, appelé *stéréophotométrie*, est un problème inverse classique de vision par ordinateur (Woodham, 1980), qui peut être résolu facilement lorsque :

– les éclairages sont supposés parallèles et uniformes ;

– la surface de l'objet à reconstruire est lambertienne (pas de taches spéculaires) ;

– les vecteurs modélisant les éclairages sont parfaitement connus, en norme et en direction ;

– les phénomènes d'ombres (propres ou portées) sont négligés.

Dans cet article, nous remettons en question les trois dernières hypothèses. Pour cela, nous combinons les méthodes décrites dans (Quéau, Durou, 2013) et dans (Durix *et al.*, 2013) afin de proposer une méthode de reconstruction 3D robuste, dans le cas où les éclairages sont inconnus.

Lorsqu'on ne connaît pas les éclairages, le problème est celui de la *stéréophotométrie non calibrée*, pour lequel les normales, l'albédo et les éclairages doivent être estimés simultanément. Malheureusement, ce problème est mal posé : on ne peut estimer ces caractéristiques qu'à une ambiguïté près. Sous l'hypothèse que le champ de normales soit suffisamment régulier (hypothèse d'intégrabilité), cette ambiguïté se réduit à *l'ambiguïté de bas-relief généralisée*, qui ne dépend que de trois paramètres. Estimer ces trois paramètres nécessite l'introduction d'une contrainte supplémentaire : nous complétons ici l'étude menée dans (Quéau, Durou, 2013) et montrons comment lever l'ambiguïté de bas-relief généralisée si l'on suppose que les éclairages ont tous la même intensité. Cette méthode est évaluée par comparaison avec la stéréophotométrie calibrée.

La qualité de la reconstruction 3D obtenue par stéréophotométrie (calibrée ou non) dépend cependant très fortement de l'absence d'ombres et de taches spéculaires. Poursuivant les travaux décrits dans (Durix *et al.*, 2013), nous nous attachons donc aussi à proposer une méthode de résolution robuste de la stéréophotométrie non calibrée. Pour cela, nous procédons en trois temps, ce qui nous permet de nous ramener au problème de l'estimation robuste des normales et de l'albédo dans le cadre de la stéréophotométrie calibrée. Les divers écarts au modèle lambertien sont d'abord identifiés et exclus du processus d'estimation des éclairages (à l'ambiguïté déjà mentionnée près), puis nous proposons de réestimer les normales et l'albédo de façon robuste à partir des éclairages estimés (toujours à cette ambiguïté près), avant de finalement lever l'ambiguïté en supposant que les sources lumineuses ont toutes la même intensité. Nous évaluons la pertinence de cette approche en montrant des exemples de reconstructions 3D d'objets soumis à des phénomènes d'ombres et de spécularités, et la comparons aux approches existantes les plus sophistiquées.

1. Introduction

La stéréophotométrie permet de reconstruire les reliefs jusqu'aux détails les plus fins. Le niveau de détail obtenu et la relative simplicité de sa mise en œuvre font que cette technique de reconstruction 3D, qui a été introduite en 1980 par Woodham (Woodham, 1980), connaît depuis quelques années un regain d'intérêt. Au contraire de la stéréoscopie où plusieurs images sont acquises sous différents points de vue, mais sous le même éclairage, la stéréophotométrie consiste à utiliser plusieurs photographies d'un objet prises sous le même point de vue, mais sous différents éclairages. Lorsque les éclairages sont connus, on parle de *stéréophotométrie calibrée*. Sous l'hypothèse d'une réflectance lambertienne, on peut alors estimer très simplement la normale et l'albédo en chaque pixel. Il reste ensuite à intégrer le champ de normales pour calculer une carte de profondeur.

La stéréophotométrie (notée SP) présente plusieurs avantages significatifs sur la stéréoscopie. En particulier, elle permet d'estimer simultanément le relief et l'albédo de l'objet photographié. En ce sens, plus qu'une simple technique de reconstruction 3D, elle constitue une technique de *scannage 3D*. Un autre avantage vient de ce que l'appareil photographique prend toutes les vues sous le même angle. Il n'est donc pas nécessaire d'apparier des indices visuels homologues entre images, car un point physique correspond au même point dans les différentes images. En outre, comme nous le verrons par la suite, le calcul de la solution est élémentaire sous l'hypothèse d'une réflectance lambertienne et peut être mené de front pour tous les pixels. Il s'ensuit que la résolution de la SP est bien plus rapide que celle de la stéréoscopie. Un problème fréquent en stéréoscopie est que certains indices visuels ne peuvent pas être appariés, à cause soit d'une occultation, soit d'un manque de texture sur certaines parties de la surface. Il en résulte que les cartes de profondeur obtenues en stéréoscopie peuvent comporter des trous, ce qui n'est pas le cas en SP. Enfin, on peut mettre en œuvre la SP avec des images provenant d'un appareil photographique non calibré, ce qui est plus délicat en stéréoscopie. D'une part, aucun calibrage externe n'est requis, puisque l'appareil photographique est fixe. D'autre part, nous n'avons à aucun moment précisé si le modèle de projection était perspectif ou non, ce qui semble indiquer que le calibrage interne non plus n'est pas requis. Effectivement, le choix d'un modèle de projection n'a pas d'influence sur l'estimation de l'albédo et des normales. En revanche, la connaissance de la matrice de projection est quand même nécessaire pour calculer la carte de profondeur à partir des normales estimées (Durou *et al.*, 2009).

Si la SP présente autant d'avantages vis-à-vis de la stéréoscopie, comment se fait-il que cette dernière reste encore aujourd'hui la technique de référence en matière de reconstruction 3D photographique ? Primo, la SP ne peut être mise en œuvre qu'en conditions très contrôlées : la scène doit être rigide et immobile, et l'appareil photographique doit être posé sur un pied. Il a bien été montré dans (Hernández *et al.*, 2011) que l'utilisation de $m \geqslant 3$ sources de lumière colorées permettait d'appliquer la SP à des scènes déformables. Néanmoins, la SP reste surtout une technique de laboratoire. Secundo, l'intégration du champ de normales estimé n'est pas triviale (Frankot, Chellappa, 1988 ; Horn, 1990 ; Durou *et al.*, 2009). D'une part, elle repose sur l'inté-

grabilité de la surface, qui n'est pas garantie. D'autre part, un biais dans l'estimation du champ de normales ou des éclairages peut provoquer une déformation importante du relief (Horovitz, Kiryati, 2004). C'est pourquoi il a été suggéré d'utiliser la SP en complément d'autres techniques de reconstruction 3D qui estiment directement le relief, comme la stéréoscopie (Nehab *et al.*, 2005). Il a déjà été dit que la SP procurait des cartes de profondeur sans trou. Cependant, puisque l'objet à reconstruire est photographié sous un seul angle, si l'on souhaite en effectuer une reconstruction 3D complète, il faut soit fusionner plusieurs cartes de profondeur acquises sous différents angles (Chang *et al.*, 2007), soit utiliser la SP en complément d'une autre technique de reconstruction 3D, comme par exemple le *shape-from-silhouettes* (Hernández *et al.*, 2008). Enfin, outre ces difficultés chroniques de la SP, la mise en oeuvre de cette technique repose sur la connaissance précise des éclairages, et sur un certain nombre d'hypothèses plus ou moins réalistes qui sont rappelées ci-après :

– Hypothèses sur l'appareil photographique : la réponse du récepteur photosensible est linéaire ; l'aberration de vignettage de l'objectif est négligée ; le facteur $\cos^4 \alpha$ de l'équation de luminance (*image irradiance equation*, cf. (Horn, 1989)) est assimilé à 1 ; les images ne comportent pas de flou de mise au point, même s'il a été montré dans (Joshi, Chaudhuri, 2004) que cette hypothèse n'était pas fondamentale.

– Hypothèses sur les éclairages : chaque éclairage consiste en un faisceau parallèle et uniforme ; le nombre m de sources lumineuses est supérieur ou égal à 3 ; au moins 3 directions d'éclairage sont non coplanaires ; les réflexions secondaires de la lumière sur l'objet lui-même, ou sur d'autres objets voisins, sont négligées.

– Hypothèses sur la scène : l'objet à reconstruire ne présente ni arête ni occultation, si bien que le vecteur normal est défini en tout point ; la surface de l'objet à reconstruire est opaque ; sa réflectance est lambertienne.

Dans le cas où les éclairages sont inconnus, on parle de *stéréophotométrie non calibrée*. L'estimation des normales et de l'albédo est beaucoup plus délicate, car le problème devient mal posé. Il est alors nécessaire de mieux contraindre le problème. De la validité des contraintes choisies dépend bien sûr la qualité des résultats.

Dans cet article, après avoir rappelé dans la partie 2 les équations de base de la stéréophotométrie, nous proposons dans la partie 3 une méthode pour résoudre le problème de la stéréophotométrie non calibrée dans le cas de la réflexion lambertienne idéale, étude initialement menée dans (Quéau, Durou, 2013). Reprenant les idées proposées dans (Durix *et al.*, 2013), nous proposons dans la partie 4 une méthode de résolution du problème qui tient compte de la présence d'écarts au modèle lambertien. Ces différentes contributions nous permettent d'élaborer une méthode complète de stéréophotométrie non calibrée en présence d'écarts au modèle lambertien, que nous évaluons dans la partie 5.

2. Équations de la stéréophotométrie

2.1. Stéréophotométrie calibrée

La stéréophotométrie a été introduite par Woodham dans (Woodham, 1980). Cette technique permet de retrouver la normale et l'albédo en chaque point visible d'un objet à partir de $m \geqslant 3$ photographies prises sous le même point de vue. Il s'agit d'une extension du problème du *shape-from-shading* (Horn, 1989), qui est connu pour son caractère mal posé. En utilisant plusieurs photographies prises sous différents éclairages (connus), et en supposant la surface lambertienne et la lumière directionnelle, Woodham montre que ce caractère mal posé disparaît si l'on dispose de $m \geqslant 3$ images. En effet, en notant I_p^i le niveau de gris du pixel p dans la $i^{\text{ème}}$ image, $i \in [1, m]$, ρ_p l'albédo au pixel p, $N_p = [N_x(p), N_y(p), N_z(p)]$ le vecteur ligne normé normal à la surface en p, et en caractérisant l'éclairage dans la $i^{\text{ème}}$ image, en norme et en direction, par un vecteur colonne $S^i = [S_x^i, S_y^i, S_z^i]^\top$ appelé *vecteur d'éclairage*, orienté dans le sens opposé à la propagation de la lumière, le modèle lambertien s'écrit :

$$I_p^i = \rho_p N_p S^i \tag{1}$$

Dorénavant, lorsqu'il sera fait référence au « modèle lambertien », il sera entendu qu'il s'agit de l'équation (1).

En notant $I_p = [I_p^1, \ldots, I_p^m]$, $S = [S^1, \ldots, S^m]$ et $M_p = \rho_p N_p$, on obtient alors le système d'équations linéaires $I_p = M_p S$. Si la matrice S est connue et de rang 3 (ce qui nécessite que 3 vecteurs d'éclairage au moins soient non coplanaires), une solution aux moindres carrés est fournie par la pseudo-inverse de Moore-Penrose $\widehat{M}_p = I_p S^+$, et l'on obtient la normale et l'albédo en p par :

$$\widehat{N}_p = \frac{\widehat{M}_p}{\|\widehat{M}_p\|} \qquad \widehat{\rho}_p = \|\widehat{M}_p\| \tag{2}$$

Ce résultat caractérise la normale et l'albédo en un pixel p quelconque. Il est également possible d'obtenir directement l'ensemble des normales et des albédos. Pour ce faire, on construit une matrice $I = [I_1^\top, \ldots, I_{|\Omega|}^\top]^\top$ par « empilement » des vecteurs lignes I_p correspondant à tous les pixels p situés à l'intérieur du domaine de reconstruction Ω ($|\Omega|$ désigne le cardinal de Ω). De la même façon, on construit $M = [M_1^\top, \ldots, M_{|\Omega|}^\top]^\top$, de sorte que la loi de Lambert s'écrit :

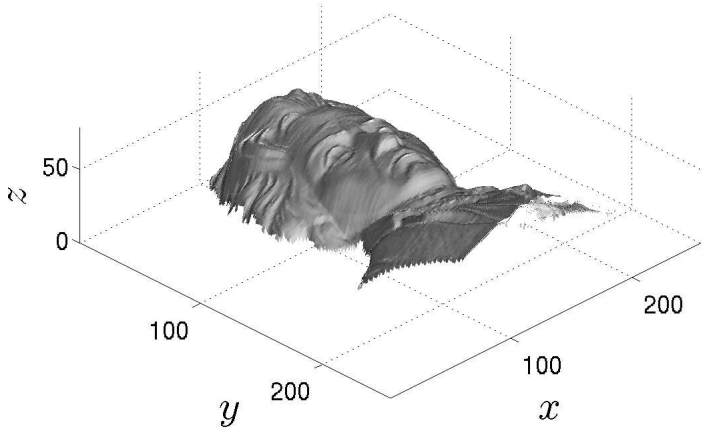
$$I = MS \tag{3}$$

On obtient alors directement :

$$\widehat{M} = IS^+ \tag{4}$$

Le champ de normales et l'albédo peuvent ensuite être calculés de la même manière qu'en (2). La figure 1 montre un exemple de résultat obtenu par ce procédé, après intégration du champ de normales $\widehat{N}$ sur Ω.

(a) (b) (c)



(d)

Figure 1. (d) Exemple de reconstruction 3D par stéréophotométrie calibrée à partir de $m = 3$ images d'un buste de Beethoven (a-b-c) disponibles sur le web. On observe sur ce relief que les détails les plus fins sont bien retrouvés.

Les vecteurs d'éclairage sont généralement supposés non coplanaires pour garantir l'hypothèse $\text{rang}(S) = 3$, même si quelques travaux récents (Abrams *et al.*, 2012 ; Ackermann *et al.*, 2012) se sont intéressés au cas $\text{rang}(S) = 2$, qui caractérise en particulier les photographies prises en extérieur, vu que le Soleil se déplace dans un plan au cours d'une même journée. La résolution de la stéréophotométrie à $m = 2$ images, qui est plus délicate, nécessite en général que l'albédo soit connu (Onn, Bruckstein, 1990 ; Kozera, 1991). Il en va de même pour le cas $m = 1$, qui correspond au *shape-from-shading*. Non seulement la reconstruction 3D photométrique devient moins fiable lorsque m décroît, mais au-dessous de $m = 3$, le problème devient de moins en moins bien posé, donc de plus en plus difficile à résoudre.

Par ailleurs, la résolution (4) du problème de la stéréophotométrie calibrée n'a de sens que si le modèle lambertien (1) est valide. Or, ce modèle suppose notamment

l'absence de phénomènes tels que les ombres (propres ou portées) ou les taches spéculaires. Il est possible de considérer les points où de tels phénomènes se produisent comme des données aberrantes, dans un processus d'estimation robuste tel que la méthode *Espérance-Maximisation* (Verbiest, Van Gool, 2008) ou la régression creuse (Ikehata *et al.*, 2012). Les taches spéculaires peuvent être prises en compte en considérant des modèles de réflectance plus sophistiqués (Georghiades, 2003) ou un objet de référence (Hertzmann, Seitz, 2005). Quant aux ombres, elles peuvent être localisées et rejetées de l'estimation (Barsky, Petrou, 2003), ce que réalise très efficacement la *coupure de graphe* (Chandraker *et al.*, 2005).

2.2. Stéréophotométrie non calibrée

La matrice $I \in \mathbb{R}^{|\Omega| \times m}$ des niveaux de gris est telle que chaque colonne I^i de I, $i = 1 \dots m$, représente l'image numéro i vectorisée, restreinte à Ω. Le problème de la stéréophotométrie non calibrée consiste à estimer deux matrices $M \in \mathbb{R}^{|\Omega| \times 3}$ et $S \in \mathbb{R}^{3 \times m}$ telles que $I = MS$. Cette décomposition n'est pas unique puisque, pour toute matrice $A \in \mathrm{GL}(3),\ MS = MAA^{-1}S$.

En supposant que chaque image soit obtenue par le modèle lambertien et l'addition d'un bruit blanc gaussien, il est raisonnable d'effectuer cette factorisation au sens des moindres carrés. La décomposition en valeurs singulières (SVD) permet d'estimer $\widehat{U} \in \mathbb{R}^{|\Omega| \times |\Omega|}$, $\widehat{W} \in \mathbb{R}^{|\Omega| \times m}$ et $\widehat{V} \in \mathbb{R}^{m \times m}$ telles que :

$$(\widehat{U}, \widehat{W}, \widehat{V}) = \underset{U,W,V}{\operatorname{argmin}} \|I - UWV^\top\|_F^2 \tag{5}$$

où $\|.\|_F$ désigne la norme de Frobenius.

Les vecteurs d'éclairage et les normales étant des vecteurs de $\mathbb{R}^3$, le rang de I devrait être au plus égal à 3, et seules les 3 plus grandes valeurs singulières devraient être non nulles (en pratique, les autres ne sont pas exactement nulles, à cause justement des écarts au modèle). Hayakawa propose dans (Hayakawa, 1994) de restreindre $\widehat{W}$ à sa sous-matrice principale $\bar{W} \in \mathbb{R}^{3 \times 3}$, et $\widehat{U}$ et $\widehat{V}$ à leurs 3 premières colonnes, c'est-à-dire à $\bar{U} \in \mathbb{R}^{|\Omega| \times 3}$ et $\bar{V} \in \mathbb{R}^{m \times 3}$, de telle sorte que $I \approx \bar{U}\bar{W}\bar{V}^\top$. On obtient alors la solution sous la forme suivante :

$$\begin{cases} \widehat{M} = \bar{U}P^\top \\ \widehat{S} = Q\bar{V}^\top \end{cases} \tag{6}$$

où P et Q sont deux matrices 3×3, a priori quelconques, telles que $P^\top Q = \bar{W}$. Néanmoins, le problème reste mal posé puisqu'il existe une infinité de tels couples (P, Q) : $P^\top Q = P^\top AA^{-1}Q$ pour toute matrice $A \in \mathrm{GL}(3)$. Résoudre le problème de la stéréophotométrie non calibrée nécessite de lever cette ambiguïté, ce qui n'est possible qu'en ajoutant des contraintes. Or, on sait par avance que le champ de normales $(x, y) \mapsto [N_x(x, y), N_y(x, y), N_z(x, y)]$ estimé par stéréophotométrie sera ensuite *in-*

tégré de façon à obtenir une carte de profondeur. Cela revient à résoudre l'équation suivante au sens des moindres carrés :

$$\nabla u = -\begin{bmatrix} N_x/N_z \\ N_y/N_z \end{bmatrix} \tag{7}$$

où ∇u est le gradient de la carte de profondeur $u : \Omega \to \mathbb{R}$. Toutefois, une telle formulation suppose implicitement que la fonction u soit de classe $\mathcal{C}^2$, donc que l'égalité de Schwarz, ou *contrainte d'intégrabilité* :

$$\frac{\partial}{\partial y}\frac{N_x}{N_z} - \frac{\partial}{\partial x}\frac{N_y}{N_z} = 0 \tag{8}$$

soit respectée (Frankot, Chellappa, 1988). Il est donc raisonnable d'imposer cette condition dès l'estimation des normales par stéréophotométrie. Yuille et Snow montrent dans (Yuille, Snow, 1997) qu'en imposant la contrainte d'intégrabilité au champ de normales inconnu, l'ambiguïté est réduite à l'ambiguïté *de bas-relief généralisée* (BRG), qui a été décrite pour la première fois dans (Belhumeur *et al.*, 1999). Ils montrent également qu'en projetant orthogonalement le champ de normales sur l'espace des fonctions intégrables, on peut estimer six des neuf coefficients de la matrice P^{-1} en développant l'équation (8) :

$$P^{-1} = \begin{bmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ a_5 & a_6 & b_3 \end{bmatrix} \tag{9}$$

Ces six coefficients correspondent aux deux premières colonnes de P^{-1}, c'est-à-dire aux coefficients $a_i,\ i = 1 \ldots 6$. Les trois autres coefficients b_1, b_2 et b_3, à ce stade inconnus, correspondent à l'ambiguïté BRG. Yuille et Snow proposent d'attribuer à ces trois coefficients des valeurs aléatoires, puis d'en déduire $\widehat{M}$ et $\widehat{S}$. La seule transformation résiduelle préservant à la fois l'hypothèse lambertienne et la contrainte d'intégrabilité est $\widehat{M}' = \widehat{M}G$ et $\widehat{S}' = G^{-1}\widehat{S}$, où G et G^{-1} sont des matrices 3×3 dépendant de trois paramètres réels (μ, ν, λ), $\lambda > 0$, qui s'écrivent :

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix} \quad G^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\mu}{\lambda} & -\frac{\nu}{\lambda} & \frac{1}{\lambda} \end{bmatrix} \tag{10}$$

L'effet d'une telle transformation BRG est illustré sur la figure 2. Notons que cette transformation affecte toutes les estimations : normales (et donc relief), albédos et éclairages.

Pour lever l'ambiguïté BRG (i.e., pour estimer μ, ν et λ), il est nécessaire de mieux contraindre le problème. Hayakawa propose dans (Hayakawa, 1994) soit d'identifier six pixels possédant le même albédo, soit de supposer l'intensité de l'éclairage identique dans six images au moins. Néanmoins, Yuille et Snow montrent dans (Yuille, Snow, 1997) que, sans la contrainte d'intégrabilité, ces contraintes ne suffisent pas

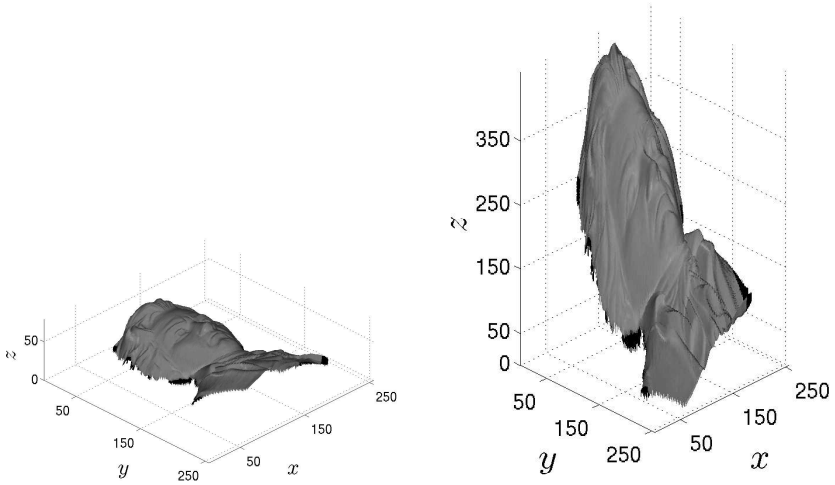


Figure 2. Effet d'une transformation de bas-relief généralisée (BRG) sur la reconstruction 3D du buste de Beethoven. À gauche : reconstruction par stéréophotométrie calibrée (cf. figure 1-d). À droite : effet de la transformation BRG correspondant aux valeurs $(\mu, \nu, \lambda) = (-2, 1, 5)$.

à rendre le problème bien posé. Cependant, lorsque l'intégrabilité du champ de normales est imposée, ils montrent que l'égalité des intensités des éclairages permet de lever l'ambiguïté BRG, tandis que Shi *et al.* font de même en détectant des pixels de même albédo (Shi *et al.*, 2010). Alldrin *et al.* suggèrent dans (Alldrin *et al.*, 2007) de minimiser l'entropie de la distribution d'albédo. Cette idée est reprise dans (Quéau *et al.*, 2013), où la variation totale du champ M remplace l'entropie de la distribution d'albédo. Tan *et al.* proposent dans (Tan *et al.*, 2007) d'identifier les pixels pour lesquels les contraintes d'isotropie et de réciprocité sont respectées. Si l'on dispose d'un détecteur d'écarts au modèle lambertien (Wu *et al.*, 2010), on peut également tirer parti de la présence d'ombres (Sunkavalli *et al.*, 2010), de taches spéculaires (Drbohlav, Chantler, 2005) ou de réflexions mutuelles (Chandraker *et al.*, 2005). Récemment, il a été proposé dans (Favaro, Papadhimitri, 2012) d'utiliser les informations fournies par les pixels de niveau de gris maximal. Il est également possible de considérer des modèles de réflectance plus sophistiqués, en remplaçant par exemple le modèle lambertien par celui de Torrance et Sparrow (Georghiades, 2003). Une autre approche ne nécessitant aucune hypothèse sur le modèle de réflectance consiste à « apprendre » les normales à reconstruire sur un objet de référence constitué du même matériau que l'objet de relief inconnu (Hertzmann, Seitz, 2005).

3. Résolution de l'ambiguïté de bas-relief généralisée par estimation de l'intensité des éclairages

D'après ce qui précède, il semble que l'estimation des trois paramètres μ, ν et λ, dans le cadre de la stéréophotométrie lambertienne non calibrée, constitue encore un problème ouvert, car ni l'erreur de reprojection ni l'intégrabilité du champ de normales ne peuvent être utilisées : il existe une infinité de couples (M, S) correspondant à un champ de normales intégrable et produisant les mêmes images. Nous proposons maintenant une nouvelle façon de résoudre le problème de la stéréophotométrie non calibrée, qui ne nécessite pas d'estimer μ, ν et λ, mais qui estime directement les trois paramètres (b_1, b_2, b_3) de P^{-1} que la contrainte d'intégrabilité ne permet pas d'estimer.

Pour lever l'ambiguïté BRG, Yuille et Snow reprennent une idée de Hayakawa (Hayakawa, 1994), qui consiste à supposer que les intensités des éclairages sont toutes égales à une valeur connue S_0, *i.e.* :

$$\forall\, i \in [1, m],\ \|S^i\| = S_0 \tag{11}$$

Dans (Quéau, Durou, 2013), il a été montré que le choix arbitraire de S_0 n'était pas sans conséquence sur la qualité des résultats, voire sur la faisabilité des calculs. L'expérience de la figure 3 montre que la qualité de la reconstruction 3D dépend du choix de S_0. Si ce paramètre est trop petit, la reconstruction 3D est impossible car les valeurs b_1, b_2 et b_3 obtenues peuvent être complexes. S'il est trop élevé, la reconstruction 3D est possible mais de piètre qualité. Entre les deux, une valeur optimale semble exister, que nous nous proposons d'estimer.

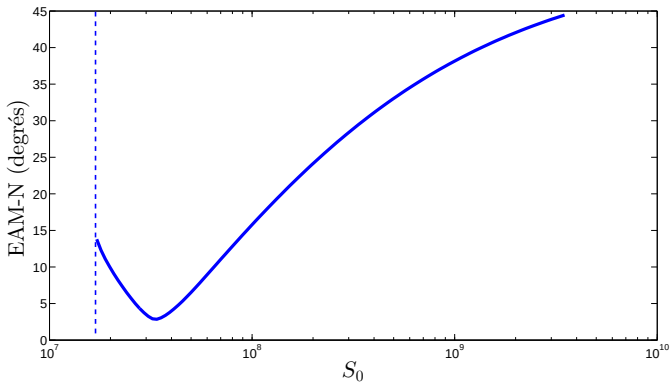


Figure 3. Influence du choix du paramètre S_0 sur l'erreur angulaire moyenne EAM-N entre les champs de normales estimés par stéréophotométrie non calibrée et par stéréophotométrie calibrée : données de test YaleB07_P00 (Georghiades et al.*, 2001). Lorsque S_0 est inférieur à un seuil, la reconstruction 3D est impossible.*

3.1. Contrainte sur les éclairages

Pour garantir la condition $||S^i|| = S_0$, $i = 1 \dots m$, il est possible de procéder beaucoup plus simplement. D'après (6), $S = Q\bar{V}^\top = P^{-T}\bar{W}\bar{V}^\top$, donc $S = P^{-T}A^\top$, en posant $A = \bar{V}\bar{W}^\top = \bar{V}\bar{W}$, puisque $\bar{W}$ est diagonale. Rappelons que $S = [S^1, \dots, S^m]$ est une matrice de taille $3 \times m$. Nous pouvons donc introduire trois vecteurs S_x, S_y et S_z de $\mathbb{R}^{1\times m}$ tels que $S = \left[{S_x}^\top, {S_y}^\top, {S_z}^\top\right]^\top$. En identifiant ces deux expressions de S et en utilisant l'expression (9) de P^{-1}, il vient :

$$\begin{cases} S_x = [a_1\ a_3\ a_5]A^\top \\ S_y = [a_2\ a_4\ a_6]A^\top \\ S_z = [b_1\ b_2\ b_3]A^\top \end{cases} \tag{12}$$

Grâce à (12) et aux contraintes $||S^i|| = S_0$, il vient, en posant $b = [b_1\ b_2\ b_3]^\top$:

$$Ab = {S_z}^\top = \begin{bmatrix} \pm\sqrt{{S_0}^2 - S_x^{1\,2} - S_y^{1\,2}} \\ \vdots \\ \pm\sqrt{{S_0}^2 - S_x^{m\,2} - S_y^{m\,2}} \end{bmatrix} \tag{13}$$

Si on suppose en outre que tous les éclairages forment avec la direction d'observation un angle inférieur à $\frac{\pi}{2}$, on peut conclure que tous les éléments de S_z sont positifs, donc déterminer les signes dans (13) :

$$Ab = \begin{bmatrix} \sqrt{{S_0}^2 - S_x^{1\,2} - S_y^{1\,2}} \\ \vdots \\ \sqrt{{S_0}^2 - S_x^{m\,2} - S_y^{m\,2}} \end{bmatrix} = K(S_0) \tag{14}$$

L'équation (14) constitue un système linéaire de m équations à 3 inconnues (b_1, b_2, b_3), qui est sur-contraint si $m > 3$. Ce système peut être résolu aux moindres carrés pour obtenir une solution unique rendant égales les normes des m vecteurs d'éclairage S^i :

$$\widehat{b} = A^+ K(S_0) \tag{15}$$

Tous les coefficients de P^{-1} étant identifiés, on obtient $\widehat{S}$ et $\widehat{M}$ par (6) :

$$\begin{cases} \widehat{S} = P^{-T}\bar{W}\bar{V}^\top \\ \widehat{M} = \bar{U}P^\top \end{cases} \tag{16}$$

Cette approche utilise une seule approximation aux moindres carrés, sans contrainte difficile à respecter, dont la solution est directement fournie par la pseudo-inverse de

$A = \bar{V}\bar{W}$. Néanmoins, la valeur de S_0 doit être fixée : nous allons maintenant proposer une méthode pour en réaliser l'estimation.

La remarque faite plus haut sur la validité de la méthode de Yuille et Snow est bien entendu toujours valable, puisque l'équation (14) n'admet de solution que pour une certaine plage de valeurs de S_0. Plutôt que de choisir arbitrairement une valeur de S_0 suffisamment grande, et puisqu'une valeur optimale de S_0 semble exister (cf. figure 3), nous nous proposons d'estimer S_0 en même temps que les 3 paramètres b_1, b_2 et b_3, de façon à minimiser le carré de la norme du résidu de l'équation (14).

Notons d'abord que, pour éviter d'obtenir des valeurs complexes, il suffit que ${S_0}^2 \geqslant \max_{i=1\ldots m} \{{S_x^i}^2 + {S_y^i}^2\}$. On peut donc se donner comme valeurs initiales :

$$\begin{cases} S_0^0 = \max_{i=1\ldots m} \{\sqrt{{S_x^i}^2 + {S_y^i}^2}\} \\ b^0 = A^+ K(S_0^0) \end{cases} \tag{17}$$

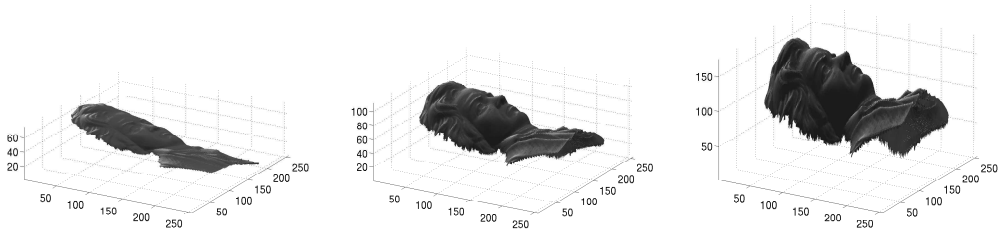
Ensuite, en partant de cette configuration initiale, on résout le problème suivant, qui est non linéaire, par la méthode de Levenberg-Marquardt :

$$\min_{(S_0,b)\in R^+\times\mathbb{R}^3} ||Ab - K(S_0)||^2 \tag{18}$$

L'optimisation est efficace et permet d'obtenir $\widehat{S}_0$ et $\widehat{b}$ en quelques itérations. On obtient ensuite P^{-1} par (9), puis $\widehat{M}$ et $\widehat{S}$ par (16).

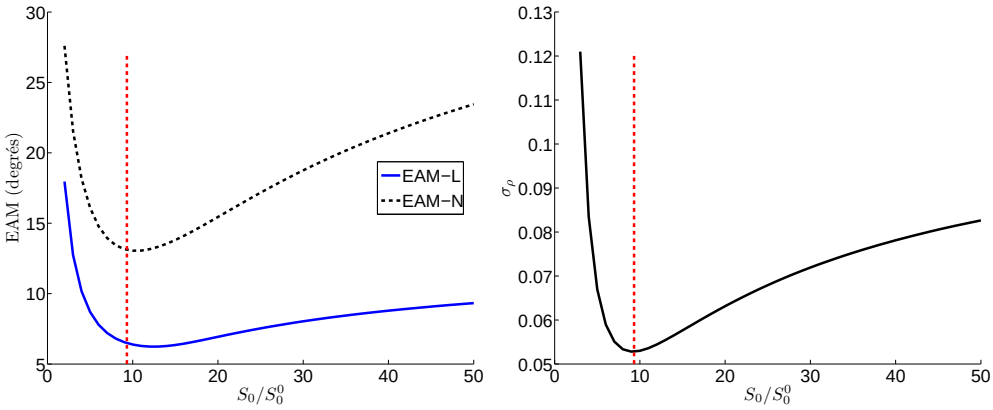
3.2. Influence du nombre d'images

Les méthodes de Hayakawa (Hayakawa, 1994) et de Yuille et Snow (Yuille, Snow, 1997) supposent qu'on dispose d'au moins $m = 6$ images. Dans notre cas, le problème (18) découle d'un système linéaire de m équations à 4 inconnues, donc admet une solution unique dès lors que $m \geqslant 4$. Les cas $m = 1$ (*shape-from-shading*) et $m = 2$ sortent du cadre de notre étude, mais nous fournissons également une solution dans le cas où $m = 3$, même si cette solution est non unique car le problème devient mal posé. En effet, à chaque valeur de S_0 correspond une solution $\widehat{b}$ *exacte* du système (14), telle que $\widehat{b} = A^{-1}K(S_0)$. Il existe alors une infinité de solutions au problème (18). Pour lever l'ambiguïté, il faut donc ajouter une contrainte. Ceci est illustré sur la figure 4, où nous montrons différentes reconstructions 3D du buste de Beethoven à partir des $m = 3$ images de la figure 1, en fonction du choix de S_0. La précision de la reconstruction 3D est évaluée par l'écart angulaire moyen EAM-L des vecteurs d'éclairage estimés avec les valeurs réelles fournies, et par l'écart angulaire moyen EAM-N des normales estimées avec les normales estimées par stéréophotométrie calibrée. Ces écarts sont tracés sur la figure 5-a en fonction de S_0. Le buste de Beethoven étant en plâtre, on peut toutefois faire l'hypothèse d'un albédo uniforme : sur la figure 5-b, on voit que l'écart-type de l'albédo est minimal pour une certaine valeur de S_0 qui minimise EAM-N. Dans la suite de cet article, nous supposerons toujours $m > 3$.



(a) (b) (c)

Figure 4. Reconstructions 3D du buste de Beethoven obtenues avec la méthode proposée de stéréophotométrie non calibrée, pour différentes valeurs de S_0 : (a) $S_0 = 2\ S_0^0$; (b) $S_0 = 10\ S_0^0$; (c) $S_0 = 30\ S_0^0$. Dans cet exemple, il reste encore une indétermination car $m = 3$. Cette indétermination peut être levée, par exemple en faisant une hypothèse sur l'albédo (cf. figure 5).



(a) (b)

Figure 5. (a) Écarts angulaires moyens entre les vecteurs d'éclairage estimés et les valeurs réelles (EAM-L), et entre les normales estimées par stéréophotométrie calibrée et non calibrée (EAM-N), à partir des $m = 3$ images du buste de Beethoven de la figure 1, en fonction de S_0. La méthode proposée ne permet pas de lever l'indétermination sur S_0 sur cet exemple car $m = 3$. Cependant, moyennant l'hypothèse d'un albédo uniforme, cette indétermination peut être levée. Nous montrons en (b) l'écart-type σ_ρ de l'albédo : le minimum de σ_ρ, matérialisé par les pointillés verticaux, correspond à la valeur qui minimise EAM-N, ce qui permet donc bien de lever l'indétermination.

3.3. Évaluation de la méthode proposée

La méthode que nous proposons, notée SPNC dans la suite (pour « stéréophotométrie non calibrée ») consiste en la séquence suivante :

1. Factoriser I par SVD tronquée à l'ordre 3, afin d'obtenir $\bar{U}$, $\bar{W}$ et $\bar{V}$ telles que $I \approx \bar{U}\bar{W}\bar{V}^\top$ (Hayakawa, 1994). Il reste à trouver P et Q telles que $\bar{W} = P^\top Q$.

2. Grâce à l'hypothèse d'intégrabilité du champ de normales, estimer les coefficients a_i, $i = 1 \ldots 6$, de P^{-1} par la méthode de Yuille et Snow (Yuille, Snow, 1997).

3. Faire une première estimation de b et S_0 comme indiqué en (17).

4. Résoudre le problème (18) par la méthode de Levenberg-Marquardt pour obtenir $\widehat{b}$ (et $\widehat{S_0}$, bien que l'on n'utilise pas cette estimation par la suite).

5. Calculer P^{-1} par (9), puis en déduire P et $Q = P^{-T}\bar{W}$. L'ambiguïté BRG est ainsi levée.

6. Enfin, calculer $\widehat{M} = \bar{U}P^\top$ et $\widehat{S} = Q\bar{V}^\top$.

Nous utilisons comme données de test les 10 premiers jeux d'images de visages de la base de données Yale Dataface B (Georghiades *et al.*, 2001). Une partie d'un de ces jeux d'images est montrée sur la figure 6. Nous obtenons une carte d'albédo et un champ de normales pour chaque visage (cf. figure 7). Après intégration des normales, nous obtenons un relief.

Pour chacun de ces 10 jeux de 22 images, nous comparons également notre méthode à deux méthodes proposées récemment, dont les codes sources en Matlab sont disponibles sur le web. La méthode d'Alldrin *et al.* (Alldrin *et al.*, 2007) vise à minimiser l'entropie de la distribution d'albédo. Celle de Favaro et Papadhimitri (Favaro, Papadhimitri, 2012) utilise les maxima locaux du niveau de gris pour identifier les directions des éclairages. Notre comparaison porte sur les temps de calcul (qui sont comparables, puisque toutes les implémentations sont réalisées en Matlab) et sur l'erreur angulaire moyenne EAM-N des normales estimées avec les normales estimées par stéréophotométrie calibrée (les vecteurs d'éclairage étant fournis).

L'ensemble des résultats est consigné dans le tableau 1, et quelques reconstructions 3D sont présentées sur la figure 8. Notre méthode s'avère incontestablement plus efficace en ce qui concerne le temps de calcul, tout en étant au moins aussi précise. L'entropie n'étant pas convexe, la méthode du minimum d'entropie (ME) utilise une méthode d'optimisation globale, ce qui explique sa lenteur. Le temps de calcul de la méthode des maxima diffus (MD) dépend du nombre de maxima d'intensité présents dans l'image, qui est imprévisible a priori. Au contraire, grâce à la minimisation d'un critère convexe, la méthode que nous proposons est très rapide, tout en étant au moins aussi précise que les deux méthodes concurrentes.

De plus, ces deux méthodes font des hypothèses fortes sur l'albédo de l'objet à reconstruire, alors que les distributions d'albédo peuvent varier énormément d'une scène à l'autre. Au contraire, la méthode que nous proposons fait une hypothèse sur les conditions d'acquisition, qui changent peu entre les différentes scènes reconstruites.

Elle présente donc une fiabilité accrue. Dans le cas du test Yale_B05_P00 cependant, les résultats obtenus par notre méthode sont en-deça des autres : la présence de larges zones noires dues au fond de la scène et aux cheveux de la personne photographiée biaise l'estimation des éclairages et des normales, alors que les méthodes ME et MD interprètent ces effets par une variation d'albédo.

Figure 6. Cinq images d'un visage provenant de la base Yale DataFaceB. Chaque jeu d'images utilisé contient 22 images d'un même visage.

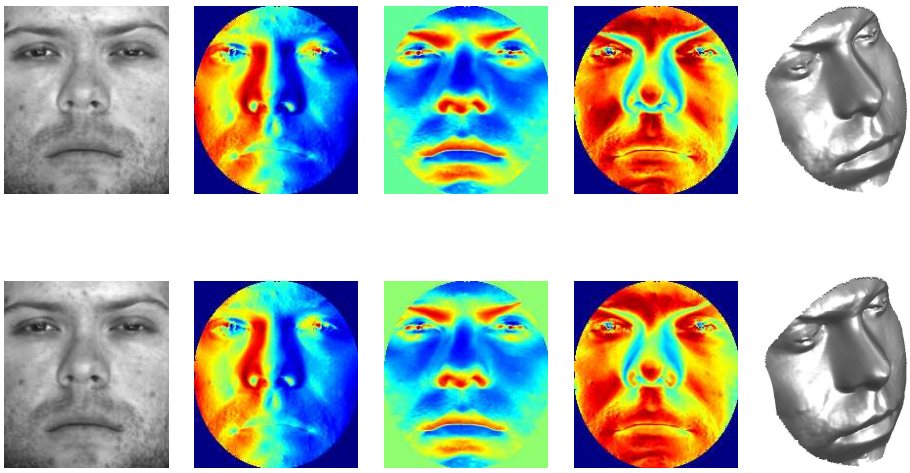


Figure 7. De gauche à droite : albédo estimé ; première, deuxième et troisième composantes du champ de normales estimé ; reconstruction 3D. Première ligne : stéréophotométrie non calibrée (notre méthode). Deuxième ligne : stéréophotométrie calibrée (résultat considéré comme la vérité terrain).

Toutefois, l'ensemble de la résolution proposée ici suppose l'absence d'écart au modèle lambertien. Il semble nécessaire de pouvoir reconstruire par stéréophotométrie non calibrée des objets constitués de matériaux plus complexes, et dès lors il faudra tenir compte d'écarts au modèle lambertien tels que les ombres et les taches spéculaires. Nous consacrons la prochaine partie à la prise en compte de tels écarts.

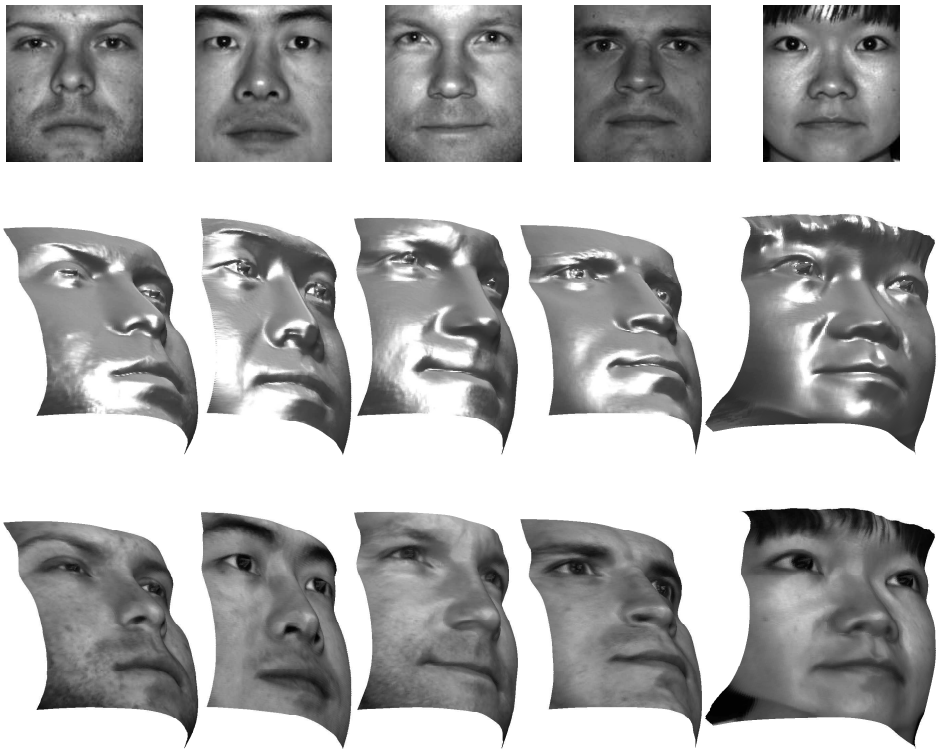


Figure 8. Les 5 premiers visages étudiés (parmi 10), et leurs reconstructions 3D, avec ou sans plaquage de l'albédo.

Données	YaleB01_P00	YaleB02_P00	YaleB03_P00	YaleB04_P00	YaleB05_P00
ME	13,59 (14,59 s)	18,99 (15,13 s)	12,58 (14,69 s)	10,88 (14,38 s)	**8,14** (14,04 s)
MD	9,83 (2,15 s)	17,11 (1,83 s)	8,33 (2,65 s)	11,44 (2,11 s)	8,30 (2,76 s)
SPNC	**6,95** (0,40 s)	**8,75** (0,44 s)	**6,51** (0,40 s)	**5,08** (0,42 s)	12,85 (0,44 s)

Données	YaleB06_P00	YaleB07_P00	YaleB08_P00	YaleB09_P00	YaleB10_P00
ME	12,22 (14,22 s)	10,33 (14,59 s)	20,92 (14,99 s)	16,91 (14,81 s)	8,24 (13,60 s)
MD	**6,92** (1,81 s)	6,27 (1,74 s)	10,25 (1,15 s)	9,06 (0,94 s)	8,45 (3,22 s)
SPNC	7,26 (0,40 s)	**5,88** (0,40 s)	**7,31** (0,40 s)	**6,17** (0,42 s)	**5,25** (0,40 s)

TABLE *1. Comparaison des performances entre la méthode du minimum d'entropie (ME), la méthode des maxima diffus (MD) et notre méthode (SPNC), qui suppose les éclairages de même intensité : erreur angulaire moyenne sur les normales (en degrés) et temps CPU.*

4. Résolution robuste de la stéréophotométrie non calibrée

4.1. Effets des écarts à la loi de Lambert

La loi de Lambert permet d'obtenir une modélisation linéaire de la SP, ce qui n'est pas le cas du *shape-from-shading* à cause de la contrainte non linéaire $\|N\| = 1$. En effet, lorsqu'on dispose d'une seule image ($m = 1$), il est nécessaire que l'albédo ρ soit connu, sans quoi le problème serait trop mal posé. Pour $m = 2$, si l'on suppose l'albédo connu, il reste à résoudre un système de $m = 2$ équations *non linéaires* à 2 inconnues. C'est donc seulement lorsque $m \geqslant 3$ que le modèle de la SP devient linéaire, sous l'hypothèse lambertienne : il n'est plus nécessaire de supposer l'albédo connu et la nouvelle inconnue du problème est $M = \rho N$, qui n'est pas un vecteur de norme 1.

En réalité, la loi (1) n'est valide que si le produit scalaire $N_p\, S^i$ est positif, car un niveau de gris est proportionnel à une énergie lumineuse, qui est forcément positive. Il faut donc remplacer la loi (1) par :

$$I_p^i = \rho_p \max\{N_p\, S^i, 0\} \tag{19}$$

Les points tels que $N_p\, S^i < 0$ constituent les *ombres propres* (*self-shadows*). En l'absence d'arêtes et d'occultations, la limite dans l'image entre une partie éclairée et une ombre propre est constituée de points où $N_p\, S^i = 0$ et s'appelle un *terminateur*.

La prise en compte des ombres propres dans (19) fait perdre à la loi de Lambert son caractère linéaire, sur lequel repose en grande partie la résolution « classique » des problèmes de la stéréophotométrie calibrée (SPC) et non calibrée (SPNC) (cf. partie 2). Pour retrouver la linéarité, on peut soit corriger les termes concernés dans la matrice I, en acceptant qu'ils deviennent négatifs, par exemple par factorisation SVD, soit retirer ces termes de la matrice I, ce que nous proposerons par la suite. Notons que Hertzmann et Seitz ont montré dans (Hertzmann, Seitz, 2005) comment s'accommoder de la non linéarité de la loi (19), en intégrant dans la scène une sphère recouverte du même matériau que l'objet à reconstruire, ce qui permet d'apprendre la relation entre normales et niveaux de gris, y compris d'ailleurs pour un matériau non lambertien.

Un deuxième type d'écarts à la loi de Lambert linéaire (1) provient des *ombres portées* (*cast-shadows*). Contrairement aux ombres propres, les ombres portées ne peuvent pas être caractérisées localement, donc l'approche de Hertzmann et Seitz ne peut y remédier.

Enfin, on est en droit de se demander si une loi aussi simple que la loi de Lambert peut modéliser la réémission de la lumière par des objets réels. Aussi surprenant que cela puisse paraître, les matériaux mats sont « plutôt bien modélisés » par la loi de Lambert non linéaire (19). Néanmoins, de nombreux autres modèles de réflectance ont été proposés, parmi lesquels le modèle de Phong est un des plus souvent utilisés

(Phong, 1975). Ce modèle de réflectance ajoute à la loi de Lambert une composante spéculaire censée tenir compte des taches spéculaires :

$$I_p^i = \rho_p \max\left\{N_p S^i, 0\right\} + \rho'_p \left(\max\left\{(V R_p^i, 0)\right\}\right)^\alpha \tag{20}$$

où V désigne la direction d'observation, R_p^i la direction du faisceau lumineux après réflexion de S^i sur la surface (selon la première loi de Descartes), α est un paramètre du modèle et ρ'_p constitue une nouvelle inconnue en chaque pixel p.

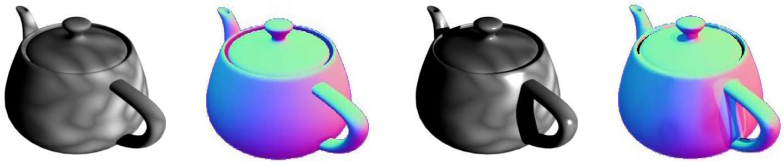
La figure 9 met en évidence les écarts entre le modèle lambertien idéal (1) sans ombre portée (cf. image 9-a) et le modèle de Phong (20) avec ombres portées (cf. image 9-b). Il est notable qu'une très grande partie des pixels héritent du même niveau de gris avec les deux modèles, ce qui permet de considérer les écarts au modèle lambertien comme des *données aberrantes* (*outliers*).

(a) (b) (c)

Figure 9. Écarts entre le modèle lambertien et le modèle de Phong : (a) champ scalaire $|M\ S|$ calculé pour un modèle 3D de théière, afin de mettre en évidence les ombres propres ; (b) image de la même théière simulée sous le même éclairage, en utilisant le modèle de Phong (20) et en prenant en compte les ombres portées (les trois types d'écarts au modèle lambertien (1) *sont donc présents sur cette image) ; (c) valeur absolue des écarts entre (a) et (b).*

L'estimation de M et S au sens des moindres carrés, par exemple en utilisant la décomposition en valeurs singulières (cf. partie 3), n'est robuste qu'au bruit blanc additif gaussien. Yuille *et al.* montrent dans (Yuille *et al.*, 1999) que pour un grand nombre d'images m, il est effectivement possible d'assimiler l'effet des ombres et des taches spéculaires à un bruit blanc additif gaussien. Lorsqu'on ne dispose pas d'une quantité suffisante d'images, s'il paraît toujours raisonnable de modéliser de la sorte le bruit induit par le capteur et par la quantification du niveau de gris, ce choix paraît en revanche beaucoup plus discutable pour la modélisation des ombres et des taches spéculaires, qui ne sont pas par nature des phénomènes stochastiques, et dont la distribution dépend de l'objet à reconstruire.

Dès lors, tout écart à la loi de Lambert (1) dégrade les résultats obtenus par SPC et a fortiori par SPNC. Cela est illustré sur l'exemple de la figure 10, où l'on constate effectivement que les normales sont très mal estimées là où la loi de Lambert est mise en défaut, notamment sous l'anse de la théière.



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

Figure 10. Résolution de la stéréophotométrie non calibrée en présence d'écarts au modèle lambertien : (a) une image de synthèse (parmi 8) d'une théière lambertienne sans ombre portée, mais avec des ombres propres ; (b) champ de normales estimé, codé dans les trois canaux RVB ; (c-d) idem lorsque la théière comporte des taches spéculaires et des ombres portées. Le résultat (d) est très dégradé par rapport à (b), qui sert de référence, même si dans les deux cas (b) et (d), le résultat est légèrement biaisé au niveau des ombres propres.

4.2. Présentation de la méthode proposée

Nous avons cité dans la partie 2 plusieurs articles qui proposent de rendre robuste la résolution de la stéréophotométrie calibrée. Le cas de la stéréophotométrie non calibrée robuste a été nettement moins étudié. Une stratégie pour s'affranchir de l'influence des niveaux de gris ne suivant pas le modèle lambertien (1), qui constituent des données aberrantes, consiste à corriger les images *a priori* comme cela a été proposé dans (Wu *et al.*, 2010) : la méthode SPNC de la partie 3 peut ensuite être appliquée aux images corrigées.

Toutefois, cette correction est fondée sur des critères purement algébriques, et ne tient pas compte du modèle de formation des images. Nous proposons au contraire d'appliquer la méthode SPNC à une matrice $\tilde{I}$ correspondant aux images corrigées en tenant compte du modèle lambertien. Nous calculons $\tilde{I}$ par le produit $\tilde{I} = MS$, où M et S sont des estimations du champ ρN et de la matrice des éclairages, calculées de la façon suivante :

1. S est d'abord estimée, à une ambiguïté près, à partir d'un sous-ensemble Ω_2 de points conformes au modèle lambertien dans chacune des images ;

2. M est ensuite calculée par stéréophotométrie calibrée robuste, à partir de la matrice S estimée précédemment.

Nous obtenons ainsi des estimations de M et S à une ambiguïté près. Puisque $MS = (MA)(A^{-1}S)$, le produit $\tilde{I} = MS$ est indépendant de cette ambiguïté. Nous pouvons donc ensuite appliquer la méthode décrite dans la partie 3 à cette matrice $\tilde{I}$.

4.3. Estimation robuste de la matrice des éclairages S

Si chaque pixel du domaine de reconstruction Ω est conforme au modèle lambertien (1) dans toutes les images, alors la matrice I des niveaux de gris est nécessairement de rang 3. Soit $I = UWV^\top$ sa décomposition en valeurs singulières, et soit $\bar{U}$, $\bar{W}$ et $\bar{V}$ les sous-matrices de U, W et V correspondant aux trois plus grandes valeurs singulières : $\bar{U}$ et $\bar{V}$ ont pour tailles respectives $|\Omega| \times 3$ et $m \times 3$, et W est une matrice diagonale de taille 3×3. Soit $\bar{I} = \bar{U}\bar{W}\bar{V}^\top$ cette SVD tronquée à l'ordre 3. Dans le cas où I est de rang 3, il s'ensuit que $I = \bar{I}$. Par ailleurs, $I = MS$, donc la matrice S des éclairages est égale à $\bar{V}^\top$, à une multiplication près par une matrice 3×3.

Supposons maintenant qu'il existe seulement un sous-ensemble $\Omega_2 \subset \Omega$ de points conformes au modèle lambertien dans toutes les images. Soit I_{Ω_2} la sous-matrice de I obtenue en ne conservant que les lignes de I correspondant aux pixels de Ω_2, et soit $\bar{I}_{\Omega_2} = \bar{U}_{\Omega_2}\bar{W}_{\Omega_2}\bar{V}_{\Omega_2}^\top$ sa SVD tronquée à l'ordre 3. D'après le raisonnement précédent, il s'ensuit que $I_{\Omega_2} = \bar{I}_{\Omega_2}$ et donc S est égale à $\bar{V}_{\Omega_2}^\top$, à une multiplication près par une matrice 3×3. La détermination d'un tel ensemble Ω_2 permet donc de résoudre, à cette ambiguïté près, le problème de l'estimation de S. De plus, il est alors facile de déterminer si le modèle lambertien est vérifié dans toutes les images pour un point p de $\Omega \backslash \Omega_2$, en évaluant la racine carrée de l'écart quadratique moyen entre les niveaux de gris en p dans les m images d'origine et les niveaux de gris des images reprojetées selon le modèle estimé. Puisque la matrice des éclairages est assimilée à $\bar{V}_{\Omega_2}^\top$, il vient immédiatement, d'après (1), que M_p peut être estimé par $I_p \left(\bar{V}_{\Omega_2}^\top\right)^+$, et donc la reprojection est obtenue en calculant $I_p \left(\bar{V}_{\Omega_2}^\top\right)^+ \left(\bar{V}_{\Omega_2}^\top\right)$. On définit ainsi l'erreur d'ajustement d'un point p à l'ensemble Ω_2 :

$$E_{\mathrm{a}}(p, \Omega_2) = \sqrt{\frac{1}{m} \left\| I_p - I_p \left(\bar{V}_{\Omega_2}^\top\right)^+ \left(\bar{V}_{\Omega_2}^\top\right) \right\|^2} \tag{21}$$

Enfin, supposons que Ω_2 contienne à la fois des points conformes au modèle lambertien et des « données aberrantes ». Dans ce cas, I_{Ω_2} et son approximation $\bar{I}_{\Omega_2}$ de rang 3 seront différentes. Ceci peut être vérifié en évaluant la racine carrée de l'erreur quadratique moyenne entre I_{Ω_2} et $\bar{I}_{\Omega_2}$, définie par :

$$\mathrm{REQM}(I_{\Omega_2}, \bar{I}_{\Omega_2}) = \sqrt{\frac{1}{m\,|\Omega|} \left\| I_{\Omega_2} - \bar{I}_{\Omega_2} \right\|_F^2} \tag{22}$$

De ces trois cas de figure, nous déduisons que :

– La conformité des données au modèle lambertien (1) sur un sous-ensemble Ω_2 de Ω peut être évaluée par $\mathrm{REQM}(I_{\Omega_2}, \bar{I}_{\Omega_2})$.

– La matrice S est égale à la matrice $\bar{V}_{\Omega_2}^\top$ de la SVD tronquée de I_{Ω_2}, à une multiplication près par une matrice 3×3, pourvu que l'ensemble Ω_2 soit conforme au modèle lambertien.

– La qualité de l'ajustement au modèle estimé d'un point p de $\Omega \backslash \Omega_2$ peut être évaluée par $E_{\mathrm{a}}(p, \Omega_2)$.

Afin de garantir une certaine stabilité numérique, il nous faut donc déterminer le plus grand sous-ensemble Ω_2 de Ω sur lequel le modèle lambertien (1) est valide.

4.3.1. Tirages aléatoires (algorithme RANSAC)

La façon la plus naturelle pour déterminer Ω_2 est d'utiliser l'algorithme RANSAC (Fischer, Bolles, 1981) qui, étant donné un modèle connu (ici la loi de Lambert) estime les paramètres du modèle (la matrice S) et, en même temps, rejette les points non conformes au modèle.

Il s'agit d'un algorithme itératif dont l'adaptation à notre problème consiste en la boucle suivante :

1. Tirer ω points aléatoirement dans Ω, qui forment l'ensemble courant Ω_2. Former la sous-matrice I_{Ω_2} de I en ne conservant que les lignes correspondant aux points de Ω_2.

2. Calculer $\bar{I}_{\Omega_2} = \bar{U}_{\Omega_2}\bar{W}_{\Omega_2}\bar{V}_{\Omega_2}^\top$ la SVD tronquée à l'ordre 3 de I_{Ω_2}, puis en déduire $S = \bar{V}_{\Omega_2}^\top$.

3. Pour chaque point p de $\Omega\backslash\Omega_2$, ajouter p à Ω_2 si $E_{\mathrm{a}}(p, \Omega_2) < \tau$, où τ est un seuil fixé par l'utilisateur.

4. Si le nombre de points dans Ω_2 est supérieur à un seuil d fixé par l'utilisateur, redéfinir I_{Ω_2}, recalculer $\bar{I}_{\Omega_2}$ par SVD tronquée à l'ordre 3 et en déduire S.

5. Calculer REQM$(I_{\Omega_2}, \bar{I}_{\Omega_2})$: si cette erreur est inférieure à l'erreur courante, accepter l'estimation de S, sinon garder l'estimation précédente. Revenir ensuite à l'étape 1.

Cet algorithme garantit une très bonne solution en un nombre *infini* d'itérations. En pratique, il faut évidemment fixer le nombre d'itérations ou choisir un autre critère d'arrêt.

Les paramètres à fournir à l'algorithme sont :

– le nombre N_{iter} d'itérations,

– le nombre ω de points tirés aléatoirement à chaque itération,

– le seuil τ pour accepter un point dans l'ensemble courant,

– le nombre minimal d de points que doit contenir l'ensemble courant.

Moyennant des connaissances a priori sur la distribution des données aberrantes, une valeur théorique du nombre d'itérations peut être calculée. Dans les tests, nous avons choisi arbitrairement d'itérer $N_{\text{iter}} = 5000$ fois.

Pour choisir la valeur de ω, remarquons que nous disposons à chaque étape 2 de ωm données (I_{Ω_2} est de taille $\omega \times m$) et que nous cherchons à estimer les $3m$ éléments de S ainsi que les 3ω éléments de M. Il faut donc garantir la condition $\omega m \geqslant 3m + 3\omega$, soit $\omega \geqslant 3m/(m-3)$ (nous avons supposé $m > 3$, ce qui est de toute façon nécessaire pour pouvoir comparer I à son approximation de rang 3). En pratique, nous choisissons pour ω la plus petite valeur vérifiant cette condition, notée $\lceil 3m/(m-3) \rceil$, et nous choisissons arbitrairement $d = 5\omega$.

Il reste à fixer la valeur de τ, qui représente l'écart moyen maximal entre le vecteur des niveaux de gris d'un pixel p et le vecteur de ses niveaux de gris recalculés d'après le modèle estimé. Nous choisissons arbitrairement $\tau = 5$ (les niveaux de gris étant, dans nos expériences, compris entre 0 et 255).

Cet algorithme fournit donc un sous-ensemble Ω_2 de Ω censé contenir les données non aberrantes, ainsi qu'une estimation de S (à l'ambiguïté $S \mapsto A^{-1}S$ près) qui n'est pas influencée par les données aberrantes. Un exemple de résultat obtenu sur images réelles est montré sur la figure 11. Les ombres portées et les taches spéculaires semblent correctement détectées.

(a) (b) (c)

Figure 11. Détection des données aberrantes par l'algorithme RANSAC. (a-b) Deux images d'un visage (parmi 22). (c) Points correspondant au modèle (en blanc) lorsque $\tau = 5$, $N = 5000$, $\omega = 4$ et $d = 5\omega$. Les points situés près du nez (ombres portées), des yeux et du front (taches spéculaires) ont bien été éliminés.

4.3.2. Détection directe par comparaison de I à son approximation de rang 3

L'approche précédente offre d'excellents résultats, mais au prix du réglage empirique de nombreux paramètres, et surtout d'un temps de calcul très élevé : en théorie, un nombre infini d'itérations est nécessaire, et chaque itération implique le calcul d'au moins une décomposition en valeurs singulières, qui peut être coûteuse pour des images de grande taille.

Nous proposons donc d'évaluer directement l'écart entre I et son approximation $\bar{I}$ de rang 3, calculée par SVD tronquée. Pour chaque pixel p, nous pouvons ensuite évaluer l'adéquation de p au modèle en calculant $E_{\mathrm{a}}(p, \Omega)$. Nous observons expérimentalement que les valeurs élevées de $E_a(p, \Omega)$ correspondent aux ombres et aux taches spéculaires, comme le montre l'exemple de la figure 12-b.

Nous pouvons alors définir l'ensemble Ω_2 des points conformes au modèle par un simple seuillage :

$$\Omega_2 = \{p \in \Omega,\ E_{\mathrm{a}}(p, \Omega) < \tau\} \tag{23}$$

(a) (b) (c)

Figure 12. Détection des données aberrantes par comparaison de I à son approximation de rang 3. (a) Une image d'un visage (parmi 22). (b) Valeurs de $E_a(p, \Omega)$. (c) Ensemble Ω_2 déterminé par le critère (23)*, lorsque $\tau = 5$. Les écarts au modèle* (1) *les plus importants (ombres portées autour du nez, taches spéculaires sur le front) sont bien éliminés. Ce résultat est proche de celui de la figure 11-c.*

Un exemple de résultat obtenu avec $\tau = 5$ est présenté sur la figure 12-c. L'ensemble estimé est comparable à celui obtenu par l'algorithme RANSAC pour la même valeur du paramètre τ (cf. figure 11-c). Toutefois, il conserve davantage de points : sur la joue gauche notamment, certains points situés dans l'ombre n'ont pas été détectés. La méthode directe est donc un peu moins précise que l'approche itérative, mais elle ne nécessite le calcul que d'une seule décomposition en valeurs singulières, et dépend d'un seul paramètre τ. Enfin, nous montrons sur la figure 13 l'influence du choix de τ sur l'ensemble estimé Ω_2 : bien sûr, plus la valeur de τ est élevée, plus l'ensemble Ω_2 est étendu. Le choix optimal de la valeur de τ correspond donc à un compromis entre la taille de Ω_2 et la fiabilité des points que cet ensemble contient.

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

Figure 13. Influence de la valeur de τ sur la détection des données aberrantes : ensemble Ω_2 estimé avec (a) $\tau = 3$, (b) $\tau = 5$, (c) $\tau = 10$ et (d) $\tau = 20$.

4.4. Estimation robuste du champ M

Nous supposons maintenant que S est connue à l'ambiguïté $S \mapsto A^{-1}S$ près, et cherchons à estimer le champ M en tenant compte des ombres et des spécularités. Nous décrivons ici une méthode très simple et très rapide, fondée sur une estimation au sens des moindres carrés pondérés. Notons que n'importe quel algorithme d'estimation robuste du champ M par stéréophotométrie calibrée aurait pu être utilisé, comme l'algorithme Espérance-Maximisation (Verbiest, Van Gool, 2008), RANSAC (Mukaigawa *et al.*, 2007) ou la régression creuse (Ikehata *et al.*, 2012), qui sont des méthodes plus robustes que la méthode proposée, mais beaucoup plus lentes.

Pour estimer le champ M à l'ambiguïté $M \mapsto MA$ près, nous nous retrouvons maintenant dans le cas calibré. En effet, nous cherchons à résoudre le système suivant, d'inconnue M :

$$I = MS \tag{24}$$

où S est connue à une ambiguïté près. Évidemment, la solution $M = IS^+$ n'est pas acceptable, car elle induirait un biais important en tout point contaminé par une ombre ou une spécularité. Afin de prendre en compte les écarts au modèle lambertien, il faut revenir à la formulation locale :

$$I_p = M_p S \tag{25}$$

Rappelons la formulation locale du problème de l'estimation de M_p au sens des moindres carrés :

$$\widehat{M_p} = \underset{M_p \in \mathbb{R}^{1\times 3}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{m} \left(I_p^i - M_p S^i \right)^2 \right\} \tag{26}$$

Cette solution doit être rejetée, puisqu'elle ne tient pas compte des écarts au modèle lambertien. Une solution plus robuste consiste à donner moins de poids aux données que l'on juge corrompues. Pour ce faire, on peut modifier l'argument de (26) en pondérant ses différents termes. On doit alors résoudre le problème aux *moindres carrés pondérés* suivant :

$$\widehat{M_p} = \underset{M_p \in \mathbb{R}^{1\times 3}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{m} {w_p^i}^2 \left(I_p^i - M_p S^i \right)^2 \right\} \tag{27}$$

où les poids w_p^j sont choisis de façon à limiter l'influence des données corrompues.

En désignant par W_p la matrice diagonale de taille $m \times m$, dont la diagonale contient les poids $w_p^i,\ i = 1 \ldots m$, on peut réécrire le problème aux moindres carrés pondérés (27) sous la forme suivante :

$$\widehat{M_p} = \underset{M_p \in \mathbb{R}^{1\times 3}}{\operatorname{argmin}} \left\{ \left\| I_p W_p - M_p S W_p \right\|^2 \right\} \tag{28}$$

dont la solution peut être obtenue en calculant la pseudo-inverse $(SW_p)^+$ de (SW_p) :

$$\widehat{M_p} = (I_p W_p)\,(SW_p)^+ \tag{29}$$

-

Intéressons nous maintenant au choix des poids $w_p^i,\ i = 1 \dots m$. Pour définir ces poids, remarquons que les écarts au modèle considérés affectent les images de deux façons :

– Les ombres sont caractérisées par des niveaux de gris très faibles.

– Les taches spéculaires sont caractérisées par des niveaux de gris élevés.

On souhaiterait donc idéalement que w_p^i soit faible à la fois pour les valeurs très faibles et pour les valeurs très élevées de I_p^i, et élevé pour les valeurs intermédiaires. Toute pondération satisfaisant ces propriétés pourrait être choisie. Les niveaux de gris très faibles ou très élevés étant caractérisés par des écarts importants à la valeur moyenne des niveaux de gris, nous proposons par exemple le choix suivant :

$$w_p^i = \max\left\{\epsilon, 1 - \frac{I_p^i}{\langle I^i \rangle}\right\} \tag{30}$$

où $\langle I^i \rangle$ est le niveau de gris moyen dans la i-ème image, et $\epsilon > 0$ un paramètre fixé par l'utilisateur, qui permet d'assurer que W_p soit inversible (dans les tests, nous utilisons $\epsilon = 10^{-9}$). Ce paramètre n'étant pas critique, la méthode proposée d'estimation robuste de M peut être considérée comme non paramétrique, au contraire des approches existantes (Verbiest, Van Gool, 2008 ; Ikehata *et al.*, 2012).

Il est notable que le produit matriciel $\widehat{M_p}S$ « corrige » les niveaux de gris. En effet, aux points où l'angle entre la normale et la direction d'éclairage est supérieur à $\frac{\pi}{2}$, le produit scalaire $\widehat{M_p}S$ est strictement négatif. On obtient donc des niveaux de gris corrigés négatifs à l'intérieur des ombres propres, alors que le niveau de gris original est nul, au bruit près (cf. équation (19)). Les spécularités sont également corrigées : on obtient aux pixels concernés un niveau de gris inférieur au niveau de gris d'origine. Enfin, les ombres portées sont supprimées. Puisque nous utilisons un modèle a priori (la loi de Lambert), la méthode de correction des images que nous proposons est plus réaliste que celle de (Wu *et al.*, 2010), qui est effectuée uniquement grâce à des considérations sur le rang de I, sans prendre en compte ni la géométrie de l'objet ni les éclairages, c'est-à-dire ni M ni S.

Soit M la matrice obtenue en « empilant » les vecteurs lignes $\widehat{M_p}$. Comme S est connue à l'ambiguïté $S \mapsto A^{-1}S$ près, le champ M est lui aussi connu à l'ambiguïté $M \mapsto MA$ près, puisque $MS = MAA^{-1}S$. En revanche, le produit $\tilde{I} = MS$ ne dépend pas de cette ambiguïté. Nous pouvons donc appliquer la méthode SPNC proposée dans la partie 3 à la matrice des niveaux de gris corrigés $\tilde{I}$. Rappelons que la méthode SPNC réduit cette ambiguïté à une ambiguïté BRG grâce à la contrainte d'intégrabilité du champ de normales. L'ambiguïté BRG est ensuite levée en ajoutant une contrainte (en l'occurrence, nous faisons l'hypothèse que les éclairages ont tous la même intensité).

5. Résultats

Les différentes contributions de cet article nous permettent de proposer une nouvelle méthode de résolution de la stéréophotométrie non calibrée en présence d'écarts au modèle lambertien, qui consiste en la séquence suivante :

1. Détection des points conformes au modèle lambertien, et estimation des éclairages S à une ambiguïté près (cf. partie 4.3.2).

2. Estimation du champ M à une ambiguïté près, grâce à l'estimation des éclairages effectuée à l'étape précédente (cf. partie 4.4).

3. Utilisation de l'intégrabilité du champ de normales et de l'égalité des intensités des éclairages pour lever l'ambiguïté sur S et sur M, en appliquant la méthode de la partie 3 à la matrice $\tilde{I}$ des niveaux de gris corrigés.

5.1. Images de synthèse

Pour valider quantitativement notre méthode, nous utilisons un jeu de $m = 22$ images d'un vase de synthèse, pour lequel les normales sont calculées analytiquement. La simulation des images est effectuée comme suit :

– L'albédo est choisi uniforme, et les intensités lumineuses égales.

– Les m directions des éclairages sont les mêmes que dans (Georghiades *et al.*, 2001).

– Le niveau de gris est calculé par le modèle de Phong (20), de façon à faire apparaître des ombres propres et des taches spéculaires.

– Des carrés noirs sont superposés aux images pour simuler des ombres portées. Une telle simulation peut paraître peu réaliste, mais puisque la détection des ombres n'utilise pas de critère de cohérence spatiale, n'importe quelle région noire peut être choisie pour simuler des ombres portées.

– Enfin, un bruit gaussien d'écart-type égal à 1% du niveau de gris maximal est ajouté à chaque image.

Le relief original, ainsi que trois des 22 images simulées par ce procédé, sont montrés sur la figure 14.

Notre méthode est censée retrouver des éclairages et des normales bien orientés, ainsi que des éclairages de même intensité et des albédos aussi uniformes que possible. Nous utilisons plusieurs critères pour évaluer les résultats :

– l'écart angulaire moyen sur les normales EAM-N ;

– l'écart angulaire moyen sur les éclairages EAM-L ;

– l'écart-type $\sigma_{||S||}$ des intensités lumineuses (mises à l'échelle entre 0 et 1) ;

– l'écart-type σ_ρ des albédos (mis à l'échelle entre 0 et 1) ;

– le temps de calcul t.

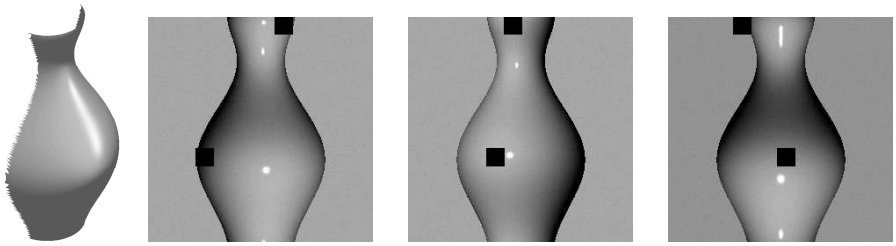


Figure 14. Relief d'un vase, et trois images de ce vase (parmi 22*) simulées sous différents éclairages, en utilisant le modèle de Phong* (20)*. Des carrés noirs ont été superposés aux images pour simuler des ombres portées.*

Les résultats de notre méthode, notée R-SPNC, sont comparés à trois autres méthodes :

– La méthode SPC correspond à la stéréophotométrie calibrée non robuste (cf. équation (4)). Cette méthode est souvent utilisée comme référence pour évaluer les résultats des méthodes de stéréophotométrie non calibrée. Mais, puisque nous tenons compte des écarts au modèle lambertien, nous espérons obtenir de meilleurs résultats avec la méthode proposée.

– La méthode SPNC correspond à la stéréophotométrie non calibrée non robuste : factorisation par (Yuille, Snow, 1997), puis résolution de l'ambiguïté BRG comme présenté dans la partie 3.

– La méthode LRA (*Low-Rank Approximation*) correspond à la méthode de (Wu *et al.*, 2010). Cette méthode est, dans son principe, très similaire à la méthode que nous proposons. Toutefois, dans la méthode LRA, la matrice des niveaux de gris est corrigée sur des considérations algébriques et non géométriques.

Afin de pouvoir comparer les temps de calcul, chaque méthode est implémentée dans le même langage (Matlab) et testée sur la même machine. De plus, les trois méthodes non calibrées utilisent la même méthode pour lever l'ambiguïté BRG (méthode proposée dans la partie 3, qui est la plus adaptée dans nos tests puisque nous avons pris soin, à la création des images, de choisir des éclairages de même intensité). De même, les reliefs sont tous obtenus en utilisant la même méthode d'intégration (Harker, O'Leary, 2008), et l'intégrabilité du champ de normales est imposée de la même manière, comme indiqué dans (Yuille, Snow, 1997).

L'ensemble des résultats est regroupé dans le tableau 2. Les reliefs reconstruits par ces quatre méthodes sont montrés sur la figure 15, et les différentes estimations de l'albédo sur la figure 16. Il est notable que le relief estimé par SPNC est extrêmement déformé : cela provient de la mauvaise factorisation de I par SVD, qui fournit un champ M tellement dégradé qu'il est impossible d'imposer la contrainte d'intégrabilité. Ces résultats prouvent clairement l'intérêt de notre approche, puisque les différentes inconnues du modèle sont mieux estimées : les éclairages estimés sont plus

proches des éclairages originaux, et l'estimation de l'albédo et des normales n'est pas biaisée par la présence d'écarts au modèle lambertien.

	EAM-N (degrés)	EAM-L (degrés)	$\sigma_{\|\|S\|\|}$	σ_ρ	$t\ (s)$
SPC	6,11	0	0	0,06	0,24
SPNC	55,95	45,30	0,05	0,25	**0,90**
LRA	11,57	12,09	**0,003**	0,06	4,47
R-SPNC	**1,54**	**1,55**	0,004	**0,01**	1,91

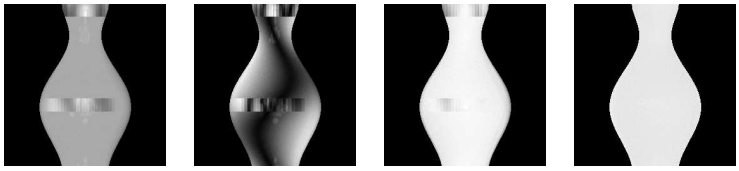
TABLE 2. *Comparaison des performances de quatre méthodes de stéréophotométrie sur les images de synthèse du vase. La méthode SPNC est totalement mise en défaut : les normales, l'albédo et les éclairages sont mal estimés à cause des écarts au modèle. La méthode SPC, malgré la connaissance des éclairages, induit un biais sur l'estimation des normales. La méthode LRA dépasse de beaucoup la méthode SPNC, mais est elle-même dépassée par la méthode R-SPNC, qui trouve un albédo quasiment uniforme (cf. figure 16-d), ainsi que des vecteurs d'éclairage et des normales très proches des valeurs exactes.*

(a) SPC (b) SPNC (c) LRA (d) R-SPNC

Figure 15. Reconstructions 3D du vase. Les reconstructions (a) et (b) sont manifestement très influencées par la présence d'écarts au modèle. Le résultat (c) est bien meilleur que (a), mais un léger artéfact est toujours visible. La méthode proposée permet de retrouver un relief (d) très proche du relief original. Les différences entre les deux premiers et les deux derniers résultats valident la démarche qui consiste à estimer les éclairages de façon robuste, puis à estimer les normales et l'albédo en tenant compte des écarts au modèle lambertien.

5.2. Images réelles

L'évaluation numérique de la méthode sur images réelles est plus délicate, puisque les valeurs réelles des normales et de l'albédo ne sont généralement pas connues. En revanche, les vecteurs d'éclairage étant disponibles dans le cas des images de (Georghiades *et al.*, 2001), nous donnons dans le tableau 3 les valeurs obtenues pour le critère EAM-L. Dans la plupart des cas, l'estimation robuste que nous proposons permet de gagner en précision. Ce résultat est intéressant car il montre que le fait de ne pas



(a) SPC (b) SPNC (c) LRA (d) R-SPNC

Figure 16. Albédos estimés à partir des images du vase. Seule la méthode proposée permet de retrouver un albédo quasiment uniforme.

connaître les éclairages n'est pas nécessairement un obstacle insurmontable : l'estimation des éclairages à partir des images permet de retrouver, à très peu de chose près, les données de la stéréophotométrie calibrée, technique pour laquelle un étalonnage préalable des éclairages est requis (une sphère brillante est généralement introduite dans la scène, ce qui permet d'estimer la direction de chaque éclairage).

Données	YaleB01_P00	YaleB02_P00	YaleB03_P00	YaleB04_P00	YaleB05_P00
SPNC	7,43	6,09	6,46	5,06	11,89
LRA	4,12	6,85	6,38	4,32	12,64
R-SPNC	**3,57**	**5,99**	**5,08**	**3,34**	**8,86**
Données	YaleB06_P00	YaleB07_P00	YaleB08_P00	YaleB09_P00	YaleB10_P00
SPNC	7,20	**4,74**	8,14	4,98	6,04
LRA	5,93	5,38	6,81	5,65	**5,81**
R-SPNC	**5,37**	6,33	**5,05**	**4,25**	8,88

TABLE *3. Valeurs du critère EAM-L (en degrés) pour les méthodes SPNC, LRA et R-SPNC appliquées à des données réelles de visages. Les temps de calcul sont du même ordre de grandeur que ceux du tableau 2. La méthode LRA étant purement algébrique, elle s'avère souvent moins précise que la méthode proposée, qui exploite explicitement le modèle lambertien.*

Nous montrons également deux résultats obtenus à partir d'images prises dans notre laboratoire. Le dispositif expérimental consiste en une salle obscure dans laquelle l'objet à reconstruire est placé face à un appareil photographique, et où huit sources lumineuses sont allumées successivement. Trois images de deux jeux de huit images ainsi acquises sont montrées sur les deux premières lignes de la figure 17. Les reconstructions 3D obtenues sont présentées sur les figures 18 et 19. Les ombres portées présentes sur les images n'influent pas sur le relief reconstruit, et les détails les plus fins, comme les granularités du bas-relief, sont bien retrouvés. Enfin, nous avons appliqué notre méthode aux jeux d'images « Frog » et « Korean Doll » (deux dernières lignes de la figure 17) utilisés dans (Chandraker *et al.*, 2007), où une méthode est proposée qui vise à éliminer les ombres par la technique du *Graph Cut* : les résultats obtenus par notre méthode (figures 20 et 21) sont qualitativement équivalents à ceux présentés dans (Chandraker *et al.*, 2007), où des données calibrées sont utilisées.

6. Conclusion et perspectives

Dans cet article, après avoir proposé dans la partie 3 une solution simple pour résoudre la stéréophotométrie non calibrée, nous avons montré dans la partie 4 comment l'adapter à la prise en compte de différents écarts au modèle lambertien tels que les ombres ou les taches spéculaires.

Plus précisément, ces écarts affectant à la fois l'estimation des éclairages et celle du champ de normales, nous conseillons de détecter les pixels où la loi de Lambert n'est pas vérifiée, par simple comparaison de la matrice des niveaux de gris à son approximation de rang 3. Les éclairages peuvent être estimés à partir d'un sous-ensemble de pixels conformes au modèle lambertien. Dans un deuxième temps, le champ de normales et l'albédo sont estimés grâce à une formulation aux moindres carrés pondérés. La méthode proposée est évaluée à la fois quantitativement et qualitativement, et se révèle au moins aussi efficace que les méthodes actuellement les plus performantes.

La méthode proposée ne peut toutefois être appliquée que si l'objet à reconstruire ne présente pas « trop » d'écarts au modèle lambertien : par exemple, les objets très brillants comme certains métaux sont exclus. Dans un tel cas, au lieu de chercher à éliminer les points suspects, il faudrait peut-être chercher à adapter la méthode SPNC à des modèles de réflectance plus sophistiqués, ce qui nécessiterait la mise en place de méthodes d'optimisation spécifiques.

Tout au long de l'exposé, nous avons supposé que les éclairages consistaient en des faisceaux parallèles et uniformes, ce qui n'est pas facile à garantir en pratique. Une extension naturelle de ce travail consistera à remettre en question cette hypothèse, en considérant des sources ponctuelles plutôt que des éclairages directionnels. Il faudra alors également prendre en compte l'atténuation de l'intensité lumineuse, ce qui rendra le problème non linéaire.

Enfin, il est notable que les différentes étapes de l'estimation robuste sont indépendantes. Plus qu'une nouvelle méthode, l'étude menée ici consiste en un cadre de travail adapté à la stéréophotométrie non calibrée. Pour améliorer les résultats, il est tout à fait envisageable de proposer des solutions plus efficaces pour la détection des données aberrantes, la factorisation des données, l'estimation du champ M ou la résolution de l'ambiguïté de bas-relief généralisée. Chacun de ces sous-problèmes étant indépendant des autres, l'assemblage de solutions efficaces pour les différents sous-problèmes doit permettre, à terme, d'obtenir une méthode très performante de résolution robuste de la stéréophotométrie non calibrée.

Bibliographie

Abrams A., Hawley C., Pless R. (2012). Heliometric Stereo: Shape from Sun Position. In *ECCV 2012, LNCS 7573*, p. 357–370.

Ackermann J., Langguth F., Fuhrmann S., Goesele M. (2012). Photometric Stereo for Outdoor Webcams. In *CVPR 2012*, p. 262–269.

Alldrin N. G., Mallick S. P., Kriegman D. J. (2007). Resolving the Generalized Bas-relief Ambiguity by Entropy Minimization. In *CVPR 2007*.

Barsky S., Petrou M. (2003). The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *PAMI*, vol. 25, n° 10, p. 1239–1252.

Belhumeur P. N., Kriegman D. J., Yuille A. L. (1999). The Bas-Relief Ambiguity. *IJCV*, vol. 35, n° 1, p. 33–44.

Chandraker M. K., Agarwal S., Kriegman D. (2007). Shadowcuts: Photometric stereo with shadows. In *CVPR 2007*.

Chandraker M. K., Kahl F., Kriegman D. J. (2005). Reflections on the Generalized Bas-Relief Ambiguity. In *CVPR 2005*, vol. 1, p. 788–795.

Chang J., Lee K., Lee S. (2007). Multiview normal field integration using level set methods. In *CVPR 2007*.

Drbohlav O., Chantler M. (2005). Can Two Specular Pixels Calibrate Photometric Stereo? In *ICCV 2005*, vol. 2, p. 1850–1857.

Durix B., Quéau Y., Durou J.-D., Charvillat V. (2013). Quels prétraitements pour la stéréophotométrie non calibrée ? In *ORASIS 2013*.

Durou J.-D., Aujol J.-F., Courteille F. (2009). Integration of a Normal Field in the Presence of Discontinuities. In *EMMCVPR 2009, LNCS 5681*, p. 261–273.

Favaro P., Papadhimitri T. (2012). A Closed-Form Solution to Uncalibrated Photometric Stereo via Diffuse Maxima. In *CVPR 2012*, p. 821–828.

Fischer M., Bolles R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, vol. 24, n° 6, p. 381–395.

Frankot R. T., Chellappa R. (1988). A Method for Enforcing Integrability in Shape from Shading Algorithms. *PAMI*, vol. 10, n° 4, p. 439–451.

Georghiades A. S. (2003). Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *ICCV 2003*, vol. 2, p. 816–823.

Georghiades A. S., Kriegman D. J., Belhumeur P. N. (2001). From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *PAMI*, vol. 23, n° 6, p. 643–660.

Harker M., O'Leary P. (2008). Least squares surface reconstruction from measured gradient fields. In *CVPR 2008*.

Hayakawa H. (1994). Photometric stereo under a light-source with arbitrary motion. *JOSA A*, vol. 11, n° 11, p. 3079–3089.

Hernández C., Vogiatzis G., Brostow G. J., Stenger B., Cipolla R. (2008). Multiview Photometric Stereo. *PAMI*, vol. 30, n° 3, p. 548–554.

Hernández C., Vogiatzis G., Stenger B., Cipolla R. (2011). Video normals from colored lights. *PAMI*, vol. 33, n° 10, p. 2104–2114.

Hertzmann A., Seitz S. M. (2005). Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. *PAMI*, vol. 27, n° 8, p. 1254–1264.

Horn B. K. P. (1989). Obtaining Shape from Shading Information. In *Shape from shading*, p. 123–171. MIT Press.

Horn B. K. P. (1990). Height and Gradient from Shading. *IJCV*, vol. 5, n° 1, p. 37–75.

Horovitz I., Kiryati N. (2004). Depth from gradient fields and control points: Bias correction in photometric stereo. *Image and Vision Computing*, vol. 22, n° 9, p. 681–694.

Ikehata S., Wipf D., Matsushita Y., Aizawa K. (2012). Robust photometric stereo using sparse regression. In *CVPR 2012*, p. 318–325.

Joshi M. V., Chaudhuri S. (2004). Photometric stereo under blurred observations. *Pattern Recognition*, vol. 3, n° 23-26, p. 169–172.

Kozera R. (1991). Existence and uniqueness in photometric stereo. *Applied Mathematics and Computation*, vol. 44, n° 1, p. 1–103.

Mukaigawa Y., Ishii Y., Shakunaga T. (2007). Analysis of photometric factors based on photometric linearization. *JOSA A*, vol. 24, n° 10, p. 3326–3334.

Nehab D., Rusinkiewicz S., Davis J., Ramamoorthi R. (2005). Efficiently combining positions and normals for precise 3D geometry. In *SIGGRAPH 2005*, p. 536–543.

Onn R., Bruckstein A. M. (1990). Integrability Disambiguates Surface Recovery in Two-Image Photometric Stereo. *IJCV*, vol. 5, n° 1, p. 105–113.

Phong B. T. (1975). Illumination for computer generated pictures. In *Communications of ACM 1975*, p. 311–317.

Quéau Y., Durou J.-D. (2013). Résolution du problème de la stéréophotométrie non calibrée par estimation de l'intensité des éclairages. In *ORASIS 2013*.

Quéau Y., Lauze F., Durou J.-D. (2013). Solving the uncalibrated photometric stereo problem using total variation. In *SSVM 2013*, p. 270–281.

Shi B., Matsushita Y., Wei Y., Xu C., Tan P. (2010). Self-calibrating Photometric Stereo. In *CVPR 2010*, p. 1118–1125.

Sunkavalli K., Zickler T., Pfister H. (2010). Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. In *ECCV 2010, LNCS 6312*, p. 251–264.

Tan P., Mallick S. P., Quan L., Kriegman D. J., Zickler T. (2007). Isotropy, reciprocity and the generalized bas-relief ambiguity. In *CVPR 2007*.

Verbiest F., Van Gool L. (2008). Photometric stereo with coherent outlier handling and confidence estimation. In *CVPR 2008*.

Woodham R. J. (1980). Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering*, vol. 19, n° 1, p. 139–144.

Wu L., Ganesh A., Shi B., Matsushita Y., Wang Y., Ma Y. (2010). Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery. In *ACCV 2010, LNCS 6494*, p. 703–717.

Yuille A. L., Snow D. (1997). Shape and Albedo from Multiple Images using Integrability. In *CVPR 1997*, p. 158–164.

Yuille A. L., Snow D., Epstein R., Belhumeur P. N. (1999). Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability. *IJCV*, vol. 35, n° 3, p. 203–222.

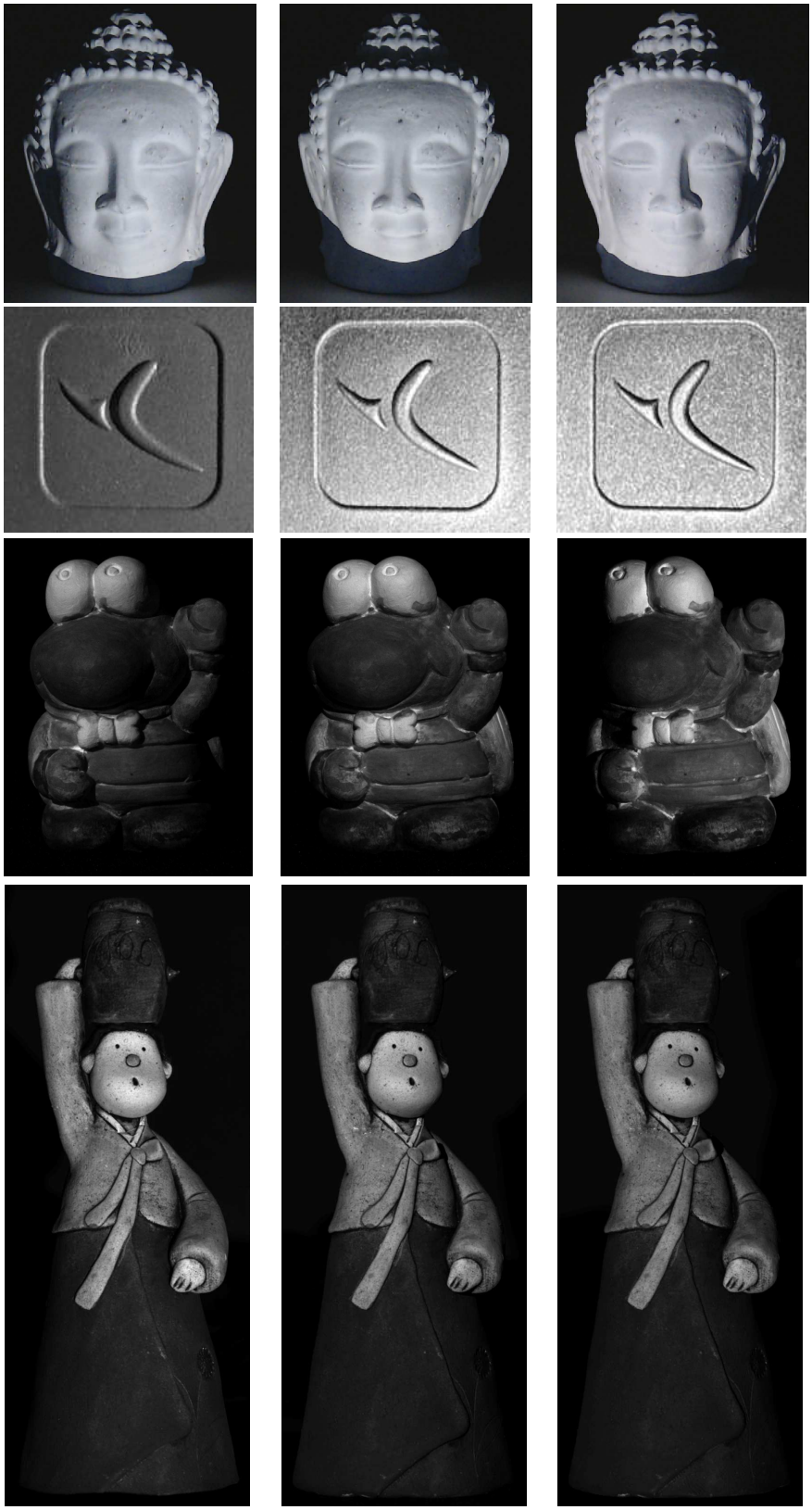


Figure 17. Images réelles. Première ligne : buste de Buddha, sur lesquelles de nombreuses ombres sont visibles. Deuxième ligne : bas-relief présentant des ombres portées et des taches spéculaires. Troisième et quatrième lignes : deux figurines présentant des ombres.

Figure 18. Reconstruction 3D du buste de Buddha, sous différents angles. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D (avec plaquage de l'albédo).

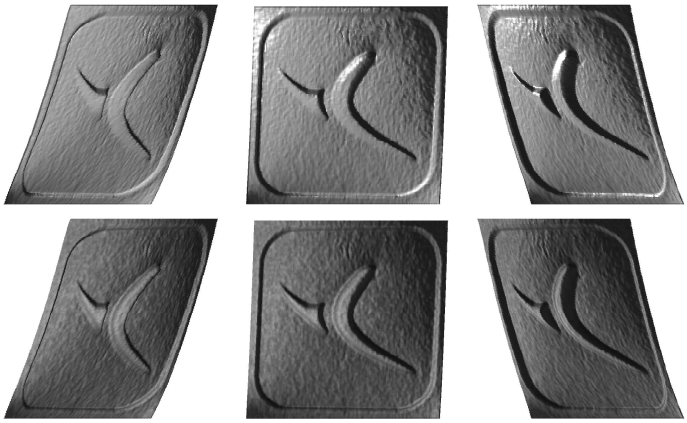


Figure 19. Reconstruction 3D du bas-relief. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D.

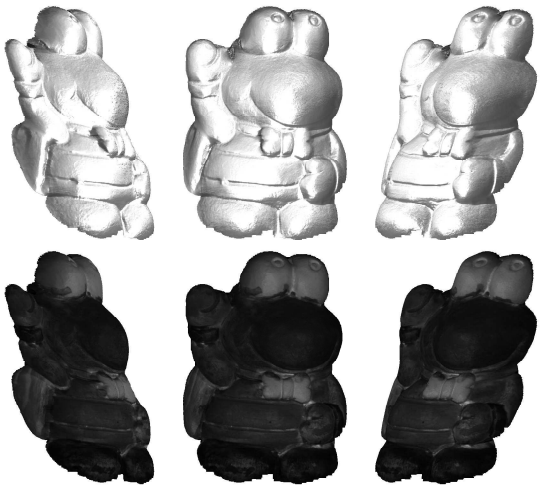


Figure 20. Reconstruction 3D de la première figurine de la figure 17. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D. Les nombreuses ombres n'influent pas sur la reconstruction 3D.

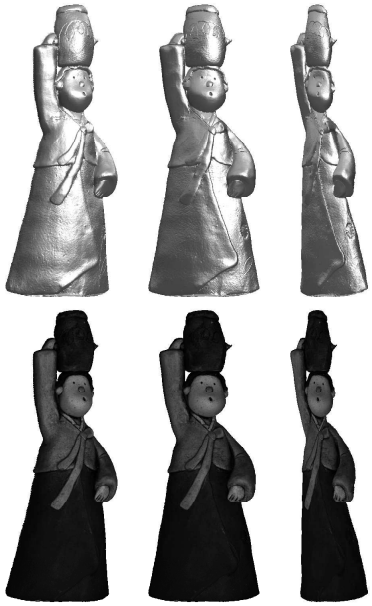


Figure 21. Reconstruction 3D de la seconde figurine de la figure 17. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D. Sur la première ligne, on observe que le motif très fin gravé sur le vase est correctement reconstruit.