



**HAL**  
open science

## Stéréophotométrie non calibrée en présence d'écarts au modèle lambertien

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

► **To cite this version:**

Yvain Quéau, Jean-Denis Durou, Bastien Durix, Vincent Charvillat. Stéréophotométrie non calibrée en présence d'écarts au modèle lambertien. *Traitement du Signal*, 2012, 22 (1), pp.1-33. 10.3166/TS.22.1-33 . hal-02118500

**HAL Id: hal-02118500**

**<https://hal.science/hal-02118500>**

Submitted on 15 May 2019

**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.

---

# Stéréophotométrie non calibrée en présence d'écarts au modèle lambertien

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

*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 reflets spéculaires. Nous montrons d'abord comment détecter de tels écarts afin d'estimer de façon plus précise les paramètres caractéristiques des sources lumineuses, que l'on suppose de même intensité afin de lever l'ambiguïté de bas-relief inhérente au problème. Nous montrons ensuite comment estimer le champ de normales 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 par la reconstruction 3D d'objets 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 the normal field and the albedo in every pixel in presence of outliers, which is made possible thanks to the robust estimation of the light sources. Our method is validated on 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, ombres, reflets spéculaires.*

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

---

DOI:10.3166/TS.22.1-33 © 2012 Lavoisier

## 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 n'est pas nouvelle puisqu'elle a été introduite en 1980 par Woodham (Woodham, 1980), connaît depuis quelques années un regain d'intérêt. Contrairement à la plupart des autres techniques de reconstruction 3D, elle présente l'avantage de fournir à la fois une estimation du relief et une estimation de la couleur (ou *albédo*) des objets à reconstruire. En ce sens, plus qu'une simple technique de reconstruction 3D, elle constitue une technique de *scannage 3D*.

Le principe de la stéréophotométrie est d'utiliser plusieurs photographies d'un objet prises sous le même point de vue, mais sous différents éclairages. Lorsque les conditions d'éclairage sont connues, on parle de *stéréophotométrie calibrée*. Sous l'hypothèse lambertienne, on peut alors estimer très simplement la normale et l'albédo en chaque pixel. Pour obtenir le relief, il reste à *intégrer* le champ de normales obtenu. Cette étape, loin d'être triviale, ne sera pas traitée ici. Nous renvoyons le lecteur à la littérature existante (Frankot, Chellappa, 1988 ; Horn, 1990 ; Durou *et al.*, 2009).

Lorsque les conditions d'éclairage sont inconnues, il s'agit de *stéréophotométrie non calibrée*. L'estimation des normales et de l'albédo est alors beaucoup plus délicate, car le problème devient mal posé. Il est nécessaire de rajouter des hypothèses à l'hypothèse lambertienne. Différentes stratégies ont été suggérées pour rendre le problème bien posé (cf. partie 2). Bien entendu, la validité des hypothèses conditionne la qualité des reconstructions 3D obtenues.

Enfin, l'hypothèse lambertienne correspond à un modèle idéal de réflectance qui décrit bien les objets parfaitement diffusants, mais exclut tout écart à ce modèle tel que les ombres (propres ou portées) ou les reflets spéculaires.

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 d'une 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 permettent d'élaborer une méthode complète de stéréophotométrie non calibrée en présence d'écarts au modèle lambertien, que nous testons dans la partie 5.

## 2. Principe de la stéréophotométrie

### 2.1. Stéréophotométrie calibrée

La stéréophotométrie, introduite par Woodham dans (Woodham, 1980), permet de retrouver la normale et l'albédo en chaque point visible d'un objet à partir de  $m \geq 3$  photographies. Au contraire de la stéréoscopie, les différentes images sont acquises sous le même point de vue, ce qui ne permet de reconstruire que la partie visible d'un objet, même s'il a été montré dans (Chang *et al.*, 2007 ; Hernández *et al.*, 2008) que la stéréophotométrie pouvait être couplée à des techniques multi-vues pour obtenir le relief complet d'un objet.

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 images obtenues sous des conditions d'éclairage (connues) différentes, et en supposant la surface lambertienne et la lumière directionnelle, Woodham montre que ce caractère mal posé disparaît si l'on dispose de  $m \geq 3$  images.

En effet, en notant  $I_p^i$  le niveau de gris du pixel  $p$  dans la  $i^{\text{ème}}$  image,  $i \in [1, m]$ ,  $\rho_p$  l'albédo au pixel  $p$ ,  $N_p = [N_x, N_y, N_z](p)$  le vecteur ligne normé normal à la surface en  $p$ , et en représentant l'éclairage dans la  $i^{\text{ème}}$  image, en norme et en direction, par un vecteur colonne  $S^i = [S_x^i, S_y^i, S_z^i]^T$  appelé vecteur d'éclairage, orienté dans le sens opposé à la propagation de la lumière, le modèle lambertien s'écrit :

$$I_p^i = \rho_p N_p S^i \quad (1)$$

En notant  $I_p = [I_p^1, \dots, I_p^m]$ ,  $S = [S^1, \dots, S^m]$  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 :

$$\widehat{N}_p = \frac{\widehat{M}_p}{\|\widehat{M}_p\|} \quad \widehat{\rho}_p = \|\widehat{M}_p\| \quad (2)$$

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  $I = [I_1^T, \dots, I_{|\Omega|}^T]^T$  par « empilement » des vecteurs lignes  $I_p$  correspondant à tous les pixels  $p$  situés à l'intérieur du domaine de reconstruction  $\Omega$  correspondant à l'objet (on désigne par  $|\Omega|$  le cardinal de  $\Omega$ ). De la même façon, on construit  $M = [M_1^T, \dots, M_{|\Omega|}^T]^T$ , de sorte que la loi de Lambert s'écrit :

$$I = MS \quad (3)$$

On obtient alors directement :

$$\widehat{M} = IS^+ \quad (4)$$

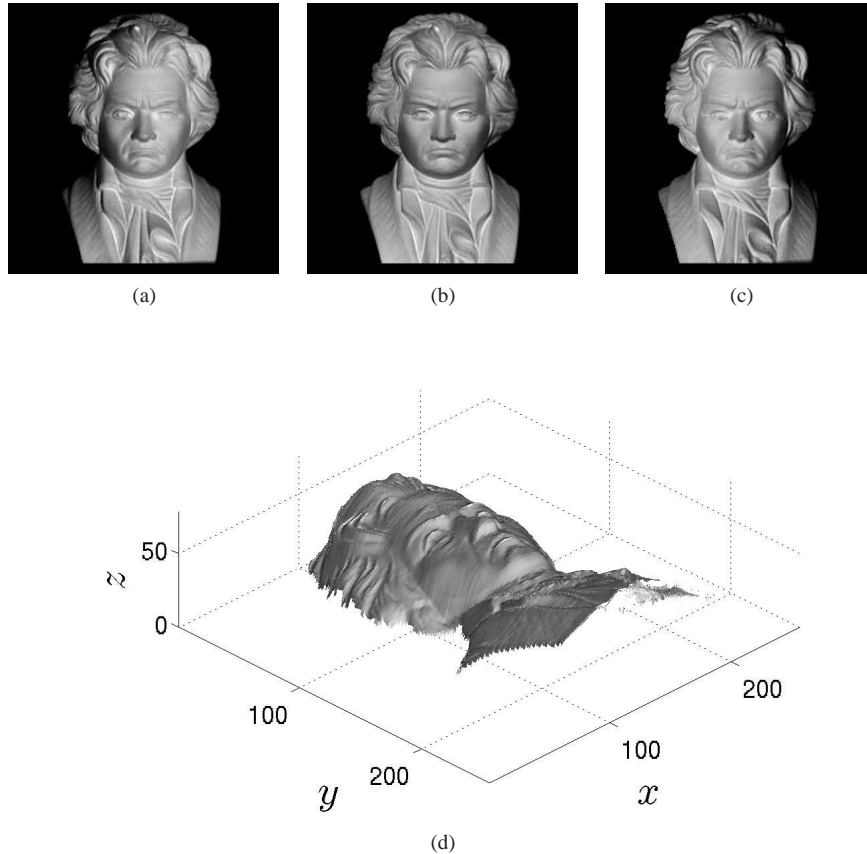


Figure 1. Exemple de reconstruction 3D par stéréophotométrie calibrée à partir de  $m = 3$  images d'un buste de Beethoven (a-b-c) disponibles sur le web. On observe sur le relief (d) que les détails les plus fins sont bien retrouvés.

L'albédo et le champ de normales peuvent ensuite être calculés de la même manière que dans l'équation (2). La figure 1 montre un exemple de résultat obtenu par cette méthode, après intégration du champ de normales  $\hat{N}$  sur  $\Omega$ .

Les directions d'éclairage sont généralement supposées 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, relativement au référentiel terrestre. La stéréophotométrie à  $m = 2$  images, qui est plus délicate, ne peut être résolue que si l'albédo est connu (Onn, Bruckstein, 1990; Kozera, 1991). Il en va de même dans le cas  $m = 1$ , qui correspond au *shape-from-shading*, mais le problème devient alors encore plus compliqué.

## 2.2. Stéréophotométrie non calibrée

La matrice  $I \in \mathbb{R}^{|\Omega| \times m}$  des niveaux de gris est telle que chaque colonne  $I^j$  de  $I$ ,  $j = 1 \dots m$ , représente l'image numéro  $j$  vectorisée, restreinte à  $\Omega$ . Le problème est d'estimer  $M \in \mathbb{R}^{|\Omega| \times 3}$  et  $S \in \mathbb{R}^{3 \times m}$  tels que  $I = MS$ . Cette décomposition n'est pas unique puisque  $\forall A \in \text{GL}(3)$ ,  $I = MS = MAA^{-1}S$ . De telles décompositions sont largement étudiées en statistiques dans le domaine de l'analyse factorielle (Mulaik, 1972).

En supposant que chaque image est idéalement 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 :

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

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

Les éclairages et les normales étant des vecteurs de  $\mathbb{R}^3$ , le rang de  $I$  devrait être au plus égal à 3 et au plus les trois plus grandes valeurs singulières devraient être non nulles (en pratique, les autres ne sont pas exactement nulles, à cause des divers écarts au modèle et du bruit, qui ne suivent pas réellement une loi gaussienne). Hayakawa propose dans (Hayakawa, 1994) de restreindre  $\widehat{W}$  à sa sous-matrice principale  $\bar{W}$  de dimension 3, et  $\widehat{U}$  et  $\widehat{V}$  à leurs 3 premières colonnes, c'est-à-dire que  $\bar{U} \in \mathbb{R}^{|\Omega| \times 3}$ ,  $\bar{W} \in \mathbb{R}^{3 \times 3}$  et  $\bar{V} \in \mathbb{R}^{m \times 3}$ , de telle sorte que  $I \approx \bar{U} \bar{W} \bar{V}^T$ .

On obtient alors la solution sous la forme suivante :

$$\begin{cases} \widehat{M} = \bar{U} P^T \\ \widehat{S} = Q \bar{V}^T \end{cases} \quad (6)$$

où  $P$  et  $Q$  sont deux matrices  $3 \times 3$ , a priori quelconques, telles que :

$$P^T Q = \bar{W} \quad (7)$$

Néanmoins, le problème reste mal posé puisqu'il existe une infinité de tels couples  $(P, Q) : P^T Q = P^T A A^{-1} Q$  pour toute matrice  $A \in \text{GL}(3)$ .

Yuille et Snow montrent dans (Yuille, Snow, 1997) qu'en imposant la *contrainte d'intégrabilité* sur le champ de normales inconnu, l'ambiguïté est réduite à l'*ambiguïté de bas-relief*, qui a été décrite pour la première fois dans (Belhumeur *et al.*, 1999). Cette contrainte s'écrit (Frankot, Chellappa, 1988):

$$\frac{\partial}{\partial y} \frac{N_x}{N_z} - \frac{\partial}{\partial x} \frac{N_y}{N_z} = 0 \quad (8)$$

Yuille et Snow montrent qu'en projetant orthogonalement le champ de normales sur l'espace des fonctions intégrables, on peut estimer six des neuf coefficients de  $P^{-1}$

en développant l'équation (8). Ces six coefficients correspondent aux deux premières colonnes de  $P^{-1}$  :

$$P^{-1} = \begin{pmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ a_5 & a_6 & b_3 \end{pmatrix} \quad (9)$$

en l'occurrence les coefficients  $a_i$ ,  $i = 1 \dots 6$ . Les trois autres coefficients ( $b_1, b_2, b_3$ ), à ce stade inconnus, correspondent à l'ambiguïté de bas-relief.

Yuille et Snow proposent d'attribuer aux trois coefficients  $b_1, b_2$  et  $b_3$  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  $3 \times 3$  à trois paramètres réels ( $\mu, \nu, \lambda$ ),  $\lambda > 0$ , qui s'écrivent :

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\mu}{\lambda} & -\frac{\nu}{\lambda} & \frac{1}{\lambda} \end{pmatrix} \quad (10)$$

L'effet d'une telle transformation de bas-relief est montré sur la figure 2. Notons que cette transformation affecte toutes les estimations : normales (et donc relief), albedo et sources lumineuses.

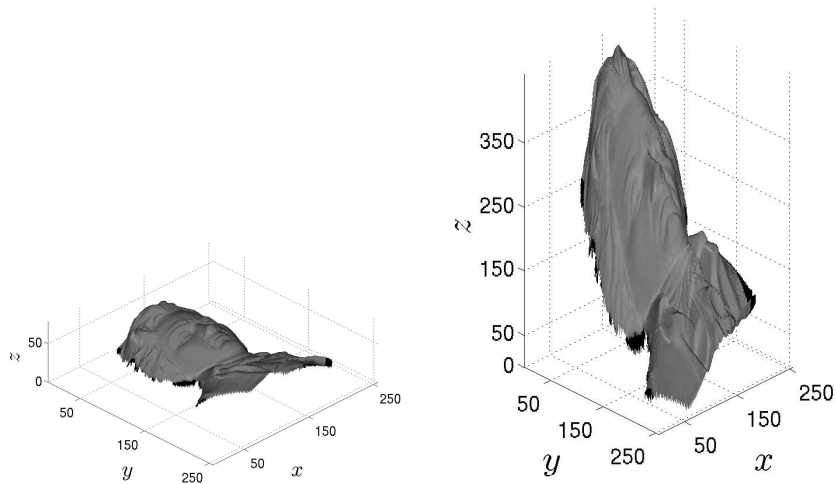


Figure 2. Effet d'une transformation de bas-relief sur la reconstruction 3D du buste de Beethoven. À gauche : reconstruction par stéréophotométrie calibrée. À droite : effet de la transformation de bas-relief correspondant aux valeurs  $(\mu, \nu, \lambda) = (-2, 1, 5)$ .

Pour lever l’ambiguïté de bas-relief (i.e., pour estimer  $\mu$ ,  $\nu$  et  $\lambda$ ), il faut rajouter au moins une hypothèse. Sans utiliser la contrainte d’intégrabilité, 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, la première méthode ne suffit pas à lever entièrement l’ambiguïté, tandis que la seconde revient à supposer que l’objet est symétrique, ce qui est fort peu réaliste. Ils conseillent donc dans (Yuille, Snow, 1997) d’adapter cette seconde hypothèse en prenant en compte l’intégrabilité, ce qui leur permet d’identifier la transformation  $G$ .

D’autres hypothèses pour résoudre l’ambiguïté de bas-relief ont été proposées : Alldrin *et al.* suggèrent dans (Alldrin *et al.*, 2007) de minimiser l’entropie de la distribution d’albédo, de façon à limiter le nombre de valeurs de l’albédo. Cette idée est reprise dans (Quéau *et al.*, 2013), où la variation totale remplace l’entropie. Shi *et al.* montrent que l’ambiguïté est levée en utilisant trois canaux de couleur au lieu du seul niveau de gris (Shi *et al.*, 2010). En supposant que le matériau présente une composante additive non lambertienne, Tan et Mallick proposent dans (Tan *et al.*, 2007) d’identifier les pixels pour lesquels les contraintes d’isotropie et de réciprocité sont respectées, ce qui leur permet de lever l’ambiguïté. 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 reflets spéculaires (Drbohlav, Chantler, 2005) ou de réflexions mutuelles (Chandraker *et al.*, 2005). Récemment, il a été proposé dans (Favaro, Papadimitri, 2012) d’utiliser les informations fournies par les pixels de niveau de gris maximal.

Pour obtenir une reconstruction 3D satisfaisante tout en s’affranchissant des hypothèses précédentes, qui peuvent être contraignantes, il est également possible de considérer des modèles d’éclairage 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 d’éclairage 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).

Ainsi, il apparaît que l’estimation des trois paramètres  $\mu$ ,  $\nu$  et  $\lambda$ , dans le cadre de la stéréophotométrie lambertienne non calibrée, reste un problème ouvert, car les critères traditionnels tels que l’erreur de reprojection ou l’intégrabilité du champ de normales ne peuvent être utilisés : il existe une infinité de couples  $(M, S)$  produisant les mêmes images et correspondant à un champ de normales intégrable. Nous proposerons dans la partie 3 une nouvelle façon de résoudre le problème de la stéréophotométrie non calibrée, qui ne nécessite pas d’estimer  $\mu$ ,  $\nu$  et  $\lambda$ , 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.



### 2.3. Avantages et inconvénients de la stéréophotométrie

Au vu des paragraphes précédents, la stéréophotométrie (PS pour *Photometric Stereo*) semble très simple à mettre en œuvre, que ce soit dans le cas calibré (CPS pour *Calibrated PS*) ou dans le cas non calibré (UPS pour *Uncalibrated PS*). Elle présente effectivement plusieurs avantages significatifs sur la stéréoscopie. En particulier, elle permet d'estimer simultanément le relief et l'albédo de l'objet photographié.

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 les points, car un point physique correspond au même pixel dans les différentes images. Il s'ensuit que la résolution de la PS est bien plus rapide que celle de la stéréoscopie. En outre, comme le montrent les équations (4) et (6), le calcul de la solution est élémentaire et peut être mené de front pour tous les pixels.

Un problème fréquent en stéréoscopie est que certains points 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 PS.

Enfin, on peut mettre en œuvre la PS avec des images provenant d'un appareil photographique non calibré, chose inconcevable 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 le relief à partir des normales estimées (Durou *et al.*, 2009).

Si la PS 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 PS 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  $m = 3$  sources de lumière colorées (en rouge, vert et bleu) permettait de prendre  $m = 3$  photographies simultanées d'une scène déformable. Néanmoins, cette technique reste surtout une technique de laboratoire.

Secundo, l'intégration du champ de normales estimé n'est pas triviale. D'une part, elle repose sur l'intégrabilité de la surface, qui n'est pas garantie. D'autre part, un biais dans l'estimation du champ de normales ou des sources peut provoquer une déformation importante du relief (Horovitz, Kiryati, 2004). C'est pourquoi il a été suggéré d'utiliser la PS 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 PS procurait des cartes de profondeur sans trou. Cependant, puisque l'objet à reconstruire est photographié sous un seul angle, il est impossible

d'en effectuer une reconstruction 3D complète. Pour cela, il faudrait soit fusionner plusieurs cartes de profondeur acquises sous différents angles (Chang *et al.*, 2007), soit là encore utiliser la PS 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 PS, la mise en œuvre de cette technique repose sur un certain nombre d'hypothèses plus ou moins réalistes, que nous allons maintenant détailler.

#### 2.4. Remise en question des hypothèses

Bien que le *shape-from-shading* soit mal posé et que l'utilisation de  $m \geq 3$  éclairages non coplanaires rende la PS bien posée, ces deux techniques de reconstruction 3D reposent fondamentalement sur les mêmes hypothèses (Daniel, 2000).

Les trois hypothèses suivantes servent à ne pas avoir un trop grand nombre d'étaonnages à effectuer :

- La réponse du récepteur photosensible est linéaire.
- Les aberrations de l'objectif sont négligées (distorsion radiale, vignetage, etc.).
- Le facteur  $\cos^4 \alpha$  de l'équation de luminance (*image irradiance equation*, cf. (Horn, 1989)) est assimilé à 1.

Quatre autres hypothèses semblent faciles à garantir en conditions de laboratoire :

- Les images sont nettes (pas de flou de mise au point). Il a d'ailleurs été montré que cette hypothèse n'était pas fondamentale (Joshi, Chaudhuri, 2004).
- La surface de l'objet à reconstruire est opaque.
- Le nombre  $m$  de sources lumineuses est tel que  $m \geq 3$ . Le cas où  $m = 2$  est mal posé, même si certains travaux en proposent une solution (Onn, Bruckstein, 1990 ; Kozera, 1991 ; Mecca, Durou, 2011).
- Au moins 3 vecteurs d'éclairage sont non coplanaires, de telle sorte que la matrice  $S$  soit de rang 3. Cette hypothèse semble impossible à garantir en extérieur puisque le Soleil se déplace dans un plan, relativement au référentiel terrestre. Néanmoins, des solutions au problème de stéréophotométrie en extérieur ont été proposées récemment (Abrams *et al.*, 2012 ; Ackermann *et al.*, 2012).

En revanche, l'hypothèse suivante, qui est difficile à justifier, vise surtout à simplifier le problème :

- Les réflexions secondaires de la lumière sur l'objet lui-même, ou sur d'autres objets voisins, sont ignorées.

Aucune de ces hypothèses ne sera remise en question ici.

Par ailleurs, la modélisation de la PS sous la forme  $I = M S$  n'est possible qu'en faisant trois hypothèses supplémentaires :

- Chaque éclairage consiste en un faisceau parallèle et uniforme.
- L’objet à reconstruire ne présente ni arête ni occultation, si bien que la normale  $N$  est définie partout.
- La surface est lambertienne.

Se passer de la première de ces hypothèses reviendrait à une remise en question profonde du modèle, comme cela a d’ailleurs été fait pour le *shape-from-shading* lorsque la source lumineuse est ponctuelle (Prados, Faugeras, 2005 ; Breuss *et al.*, 2012).

La deuxième hypothèse ci-dessus est liée à la géométrie de l’objet photographié. On ne peut définir de normale qu’en un point où la *profondeur*, qui constitue une fonction des coordonnées de l’image, est de classe  $C^1$ , ce qui n’est pas le cas le long d’une arête ou d’une occultation. Par ailleurs, la contrainte d’intégrabilité (8) n’est valide que si la profondeur est une fonction de classe  $C^2$ . Par conséquent, un objet comportant des arêtes ou des occultations ne peut vérifier cette contrainte.

Dans cet article, nous concentrons notre attention sur la dernière hypothèse ci-dessus, qui suppose notamment l’absence de phénomènes tels que les ombres (propres ou portées) ou les reflets spéculaires. Dans le cas calibré, 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 *Expectation-Maximization* (Verbiest, Van Gool, 2008). Les reflets spéculaires peuvent aussi être pris en compte en considérant des modèles de réflectance plus sophistiqués (Georghiades, 2003), ou un objet de référence (Hertzmann, Seitz, 2005). Les ombres peuvent également être localisées et rejetées de l’estimation (Barsky, Petrou, 2003), ce qui peut être réalisé efficacement au moyen de la technique du *Graph Cuts* (Chandraker *et al.*, 2005). Le cas non calibré a été beaucoup moins étudié, une possibilité étant toutefois de réaliser un prétraitement des données pour tenter de corriger ces écarts au modèle lambertien (Wu *et al.*, 2010 ; Lin *et al.*, 2009).

La suite de cet article est organisée de la façon suivante :

- Dans la partie 3, nous présentons une méthode, initialement proposée dans (Quéau, Durou, 2013), pour résoudre le problème de la stéréophotométrie non calibrée dans le cadre lambertien, en supposant un cadre idéal dans lequel les hypothèses ci-dessus sont toutes vérifiées.
- Nous remettons en question la dernière hypothèse ci-dessus dans la partie 4, et montrons comment l’estimation des éclairages, du champ de normales et de l’albédo peut alors être réalisée, étendant ainsi l’étude menée dans (Durix *et al.*, 2013).
- Enfin, nous récapitulons la méthode ainsi construite dans la partie 5, et nous évaluons sa robustesse sur des jeux de données réelles.

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

En l'absence d'écart au modèle lambertien (1), nous avons rappelé dans la partie 2 comment  $S$  et  $M$  pouvaient être estimées par SVD de la matrice  $I$ , à une ambiguïté près. Pour lever l'ambiguïté, il « suffit » d'estimer les 9 coefficients de la matrice  $P^{-1}$  (équation (9)). Yuille et Snow proposent dans (Yuille, Snow, 1997) d'estimer les six coefficients  $a_i$ ,  $i = 1 \dots 6$ , de l'équation (9) en imposant l'intégrabilité du champ de normales, puis de fixer  $b_1$ ,  $b_2$  et  $b_3$  à des valeurs aléatoires. Reste à estimer la transformation de bas-relief  $G$  définie en (10).

Reprenant une idée de Hayakawa (Hayakawa, 1994), Yuille et Snow supposent l'intensité des éclairages constante et égale à  $S_0 = 1$ , *i.e.* :

$$\forall i \in [1, m], S^i \top S^i = S_0^2 = 1 \quad (11)$$

Dans (Quéau, Durou, 2013), il a été montré que le choix arbitraire  $S_0 = 1$  n'était pas sans conséquence sur la qualité de la reconstruction 3D, voire sur la faisabilité des calculs. L'expérience de la figure 3 montre que la qualité de la reconstruction 3D dépend du choix du paramètre  $S_0$ . Si ce paramètre est trop petit, la reconstruction 3D est impossible car les valeurs  $b_1$ ,  $b_2$  et  $b_3$  obtenues sont 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.

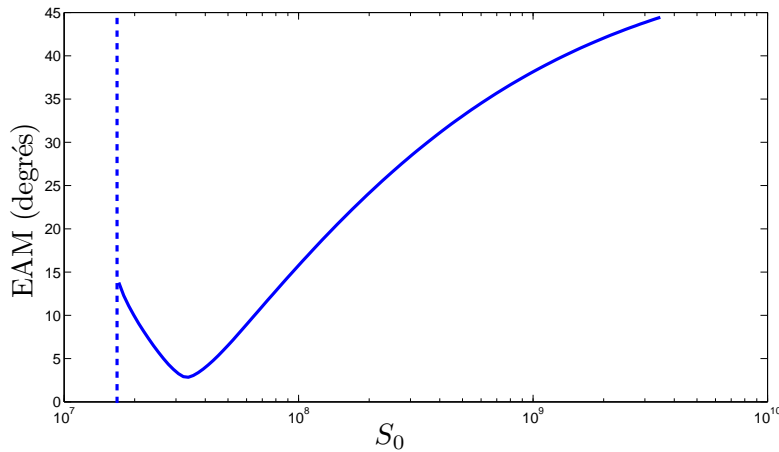


Figure 3. Influence du choix du paramètre  $S_0$  sur l'erreur angulaire moyenne (EAM) entre le champ de normales estimé par stéréophotométrie non calibrée et le champ estimé par stéréophotométrie calibrée : données de test YaleB07\_P00 (Georghiades et al., 2001). Lorsque  $S_0$  est inférieur à un seuil, 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 pourtant possible de procéder beaucoup plus simplement. D'après (6),  $S = Q\bar{V}^\top = P^{-T}\bar{W}\bar{V}^\top$ , donc  $S = P^{-T}A^\top$ , avec  $A = \bar{V}\bar{W}$ , puisque  $\bar{W}$  est diagonale donc  $\bar{W}^\top = \bar{W}$ , ce qui donne en utilisant (9) :

$$S = \begin{pmatrix} S_x \\ S_y \\ b^\top A^\top \end{pmatrix} \quad (12)$$

où  $S_x = [a_1 \ a_3 \ a_5]A^\top$ ,  $S_y = [a_2 \ a_4 \ a_6]A^\top$  et  $b = [b_1 \ b_2 \ b_3]^\top$ .

Rappelons que  $S = [S^1, \dots, S^m]$  est une matrice de taille  $3 \times m$ . Grâce à (12) et aux contraintes  $\|S^i\| = S_0$ , il vient :

$$Ab = \begin{pmatrix} \pm\sqrt{S_0^2 - S_x^2 - S_y^2} \\ \vdots \\ \pm\sqrt{S_0^2 - S_x^2 - S_y^2} \end{pmatrix} \quad (13)$$

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 termes de la troisième ligne de  $S$  sont positifs, donc fixer les signes dans (13) :

$$Ab = \begin{pmatrix} \sqrt{S_0^2 - S_x^2 - S_y^2} \\ \vdots \\ \sqrt{S_0^2 - S_x^2 - S_y^2} \end{pmatrix} = K(S_0) \quad (14)$$

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

$$\hat{b} = A^+ K(S_0) \quad (15)$$

Tous les coefficients de  $\hat{P}^{-1}$  étant maintenant identifiés (équation (9)), on obtient  $\hat{S}$  et  $\hat{M}$  par :

$$\begin{cases} \hat{S} = \hat{P}^{-T}\bar{W}\bar{V}^\top \\ \hat{M} = \bar{U}\hat{P}^\top \end{cases} \quad (16)$$

Cette approche n'utilise qu'une seule approximation aux moindres carrés, sans contrainte difficile à respecter, et 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'a de sens que pour une certaine plage de valeurs de  $S_0$ . Plutôt que de choisir une valeur de  $S_0$  suffisamment grande pour éviter les problèmes, 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 la norme du résidu de (14).

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

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

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

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

L'optimisation est efficace et permet d'obtenir  $\hat{b}$  et  $\hat{S}_0$  en quelques itérations. On obtient ensuite  $\hat{P}^{-1}$  par (9), puis  $\hat{M}$  et  $\hat{S}$  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) découlant d'un système linéaire de  $m$  équations à 4 inconnues, il a une solution unique dès lors que  $m \geq 4$ .

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 est alors mal posé. En effet, à chaque valeur de  $S_0$  correspond une solution  $\hat{b}$  exacte du système (13), telle que  $\hat{b} = A^{-1}K(S_0)$ . Il existe alors une infinité de solutions au problème (18). Pour lever l'ambiguïté, il faut donc rajouter une contrainte. Ceci est illustré sur la figure 4-a, 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 entre les vecteurs d'éclairage estimés et les valeurs réelles fournies (EAM-L, en degrés), et par l'écart angulaire moyen entre les normales estimées et les normales estimées par CPS (EAM-N, en degrés), qui sont tracés sur la figure 5 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 4-b, on voit que l'écart-type de l'albédo est minimal pour une certaine valeur de  $S_0$  qui correspond bien à la valeur qui minimise EAM-N.

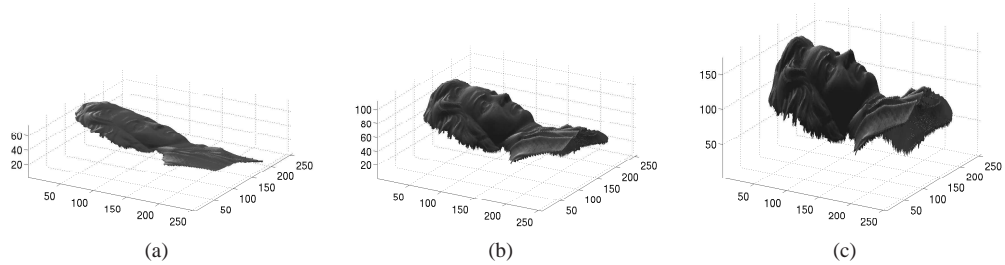


Figure 4. Reconstructions 3D du buste de Beethoven obtenues avec la méthode proposée, pour différentes valeurs de  $S_0$  : (a)  $S_0 = 2 S_0^0$  ; (b)  $S_0 = 10 S_0^0$  ; (c)  $S_0 = 30 S_0^0$ . Dans cet exemple, il reste encore une indétermination car  $m = 3$ .

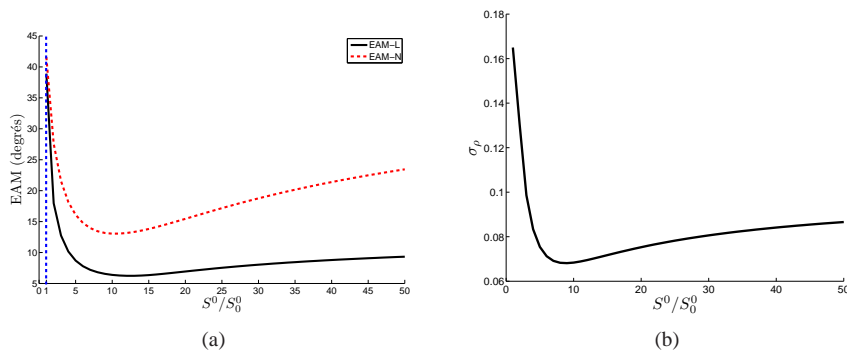


Figure 5. (a) Écart angulaire moyen sur l'estimation des sources lumineuses (EAM-L) par rapport aux valeurs réelles, et sur les normales (EAM-N) entre la stéréophotométrie calibrée (CPS) et la stéréophotométrie non calibrée (UPS), à partir des  $m = 3$  images du buste de Beethoven de la figure 1, en fonction de  $S^0$  (rappelons que si  $S^0 < S_0^0$ , au moins un coefficient de la matrice  $S$  est complexe). La méthode proposée ne permet pas de lever l'indétermination sur  $S_0$  dans ce cas car  $m = 3$ . Cependant, moyennant la connaissance a priori que l'albédo est uniforme, cette indétermination peut être levée: nous montrons en (b) l'écart-type  $\sigma_\rho$  de l'albédo. Le minimum de  $\sigma_\rho$  correspond à la valeur qui minimise EAM-N, ce qui permet donc bien de lever l'indétermination.

Notons enfin que nous chercherons dans la partie 4 à identifier les points où la loi de Lambert n'est pas vérifiée, par comparaison de la matrice  $I$  à sa reprojection. Or si  $m = 3$ , le rang de  $I$  est égal à 3 et donc  $I$  est parfaitement égale à sa reprojection, ce qui empêchera une telle comparaison. Nous supposerons donc dorénavant  $m > 3$ .

### 3.3. Évaluation de la méthode proposée

Rappelons d'abord les hypothèses sur lesquelles se fonde notre méthode :

- Le modèle lambertien est supposé parfaitement valide, au bruit et à la quantification près. Ceci permet d'assurer que la factorisation par SVD n'induit pas de biais.
- Le champ de normales estimé doit être intégrable. Cette hypothèse permet de réduire l'ambiguïté  $I = MS = MAA^{-1}S$  à l'ambiguïté de bas-relief  $A \equiv G$ , où  $G$  est définie par (10).
- Les sources lumineuses ont toutes la même intensité  $S_0$ . Cette hypothèse permet de résoudre l'ambiguïté de bas-relief, *i.e.* de déterminer la matrice  $G$ .

La méthode que nous proposons consiste en la séquence suivante :

1. Factoriser  $I$  par SVD, afin d'obtenir  $\bar{U}$ ,  $\bar{W}$  et  $\bar{V}$  telles que  $I = \bar{U}\bar{W}\bar{V}^T$  (Hayakawa, 1994). Il reste à trouver  $P$  et  $Q$  tels que  $\bar{W} = P^T Q$ .
2. Grâce à l'hypothèse d'intégrabilité du champ de normales, estimer les 6 coefficients  $a_i$ ,  $i = 1 \dots 6$ , de  $P^{-1}$  par la méthode de Yuille et Snow (Yuille, Snow, 1997).
3. Faire une première estimation de  $b$  et  $S_0$  par (17).
4. Résoudre le problème (18) par la méthode de Levenberg-Marquardt pour obtenir  $\hat{b}$ .
5. Calculer  $\hat{P}^{-1}$  par (9), puis en déduire  $\hat{P}$  et  $\hat{Q} = \hat{P}^{-T} \bar{W}$ . L'ambiguïté de bas-relief est ainsi levée.
6. Enfin, calculer  $\hat{M} = \bar{U} \hat{P}^T$  et  $\hat{S} = \hat{Q} \bar{V}^T$ .

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

Pour chacun de ces 10 jeux de 22 images, nous comparons également notre méthode à deux travaux récents dont les codes sources en Matlab sont disponibles sur le web. La méthode de Alldrin *et al.* (Alldrin *et al.*, 2007) vise à minimiser l'entropie de la distribution d'albédo. Celle de Favaro et Papadimitri (Favaro, Papadimitri, 2012) utilise les maxima locaux du niveau de gris pour identifier les directions des sources lumineuses. 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) entre le champ de normales estimé et le champ de normales obtenu par stéréophotométrie calibrée (les conditions d'éclairage étant fournies).

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 temps de calcul, tout en étant au moins aussi précise.





Figure 6. 5 images d'un visage provenant de la base Yale DataFaceB. Chaque jeu d'images utilisé contient 22 images d'un même visage.

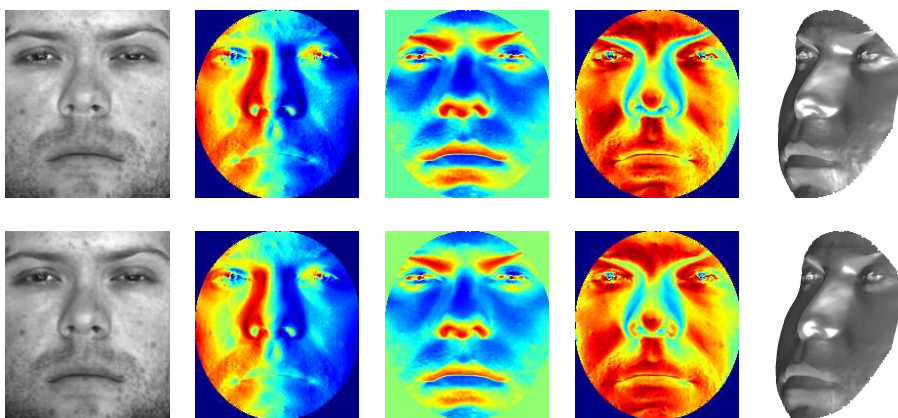


Figure 7. De gauche à droite : albédo estimé ; première, deuxième et troisième composantes du champ de normales estimé ; reconstruction 3D. Première ligne : stéréophotométrie non calibrée (notre méthode). Deuxième ligne : stéréophotométrie calibrée (résultat considéré comme la vérité terrain).

Il est cependant clair que l'hypothèse sur laquelle repose notre méthode n'est pas valide pour toutes les données de stéréophotométrie. Si lors de l'acquisition des images, l'intensité de l'éclairage variait fortement d'une image à l'autre, alors il faudrait changer d'hypothèse. Mais la vraie limite de notre méthode vient de l'hypothèse d'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 de défauts tels que les ombres et les taches spéculaires. Nous proposons pour cela de procéder en deux temps, en estimant d'abord les éclairages à partir de pixels où les niveaux de gris sont conformes au modèle lambertien, puis en estimant la normale en chaque pixel. Nous consacrons la prochaine partie à la prise en compte de tels écarts au modèle lambertien.



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)	<b>8,14</b> (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)
EI	<b>6,95</b> (0,40 s)	<b>8,75</b> (0,44 s)	<b>6,51</b> (0,40 s)	<b>5,08</b> (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	<b>6,92</b> (1,81 s)	6,27 (1,74 s)	10,25 (1,15 s)	9,06 (0,94 s)	8,45 (3,22 s)
EI	7,26 (0,40 s)	<b>5,88</b> (0,40 s)	<b>7,31</b> (0,40 s)	<b>6,17</b> (0,42 s)	<b>5,25</b> (0,40 s)

TABLE 1. Comparaison des performances entre notre méthode, qui suppose les éclairages de même intensité (EI), la méthode du minimum d'entropie (ME) et la méthode des maxima diffus (MD). Nous donnons l'erreur angulaire moyenne sur les normales (en degrés) et le 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 PS, ce qui n'est pas le cas en *shape-from-shading* : la normale  $N$  est normée, elle doit vérifier la contrainte non linéaire  $\|N\| = 1$ . Or, lorsqu'on dispose d'une seule image ( $m = 1$ ), il n'est pas raisonnable de supposer l'albédo  $\rho$  inconnu, car cela rendrait le problème encore plus mal posé. Pour  $m = 2$ , il est encore nécessaire de supposer l'albédo connu et de résoudre un système de  $m = 2$  équations non linéaires à 2 inconnues. C'est donc seulement lorsque  $m \geq 3$  que le modèle de la PS 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 normé.

Cependant, 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 :

$$I_p^i = \max\{\rho_p N_p S^i, 0\} \quad (19)$$

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 pixels 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 laquelle repose en grande partie la résolution « classique » des problèmes CPS et UPS (cf. partie 2). Pour restaurer la linéarité, on peut soit modifier les termes concernés de la matrice  $I$ , en acceptant qu'ils deviennent négatifs, par exemple par factorisation SVD, soit les retirer purement et simplement de la matrice  $I$ , ce que nous proposerons par la suite. Notons que Hertzmann et al. 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 al. 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 ont été proposés, parmi lesquels le modèle de Phong est un des plus souvent utilisés (Phong, 1975). Ce modèle rajoute à la loi de Lambert une composante spéculaire censée tenir compte des reflets brillants :

$$I_p^i = \max\{\rho_p N_p S^i + \rho'_p (V R^i)^\alpha, 0\} \quad (20)$$

où  $V$  désigne la direction d'observation,  $R^i$  caractérise le faisceau lumineux  $S^i$  après réflexion sur la surface (selon la première loi de Descartes),  $\alpha$  est un paramètre du modèle et  $\rho'_p$  constitue une nouvelle inconnue en  $p$ .

La figure 9 illustre ces trois types d'écarts au modèle lambertien linéaire (ombres propres, ombres portées et taches spéculaires) sur l'image de synthèse d'une théière. Chacun des écarts affecte une zone très localisée de la théière. Par conséquent, une grande partie des niveaux de gris des deux premières images de la figure 9 sont identiques.

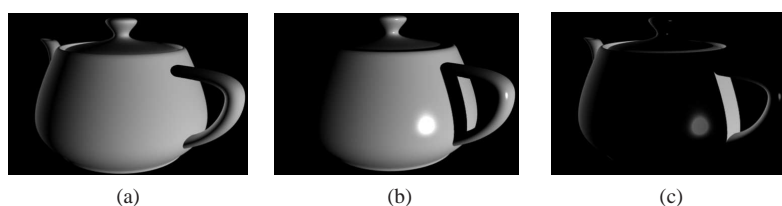


Figure 9. Différence entre le modèle lambertien et le modèle de Phong : (a) champ scalaire  $|MS|$  calculé sur un modèle 3D de théière, afin de mettre en évidence les ombres propres ; (b) image de la même théière simulée sous le même éclairage, en utilisant le modèle de Phong (20) et en prenant en compte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

On doit bien sûr s'attendre à ce que tout écart à la loi de Lambert linéaire dégrade les résultats obtenus par PS et a fortiori par UPS, où ces écarts se traduiront à la fois par une mauvaise estimation des éclairages et par une mauvaise estimation des normales. 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.

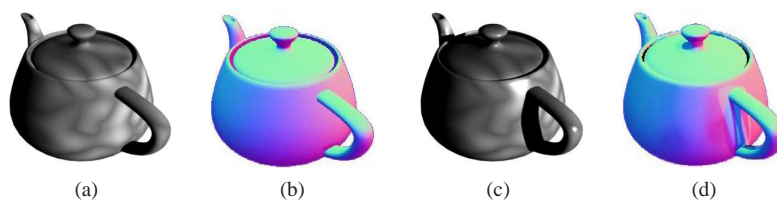


Figure 10. Résolution de la stéréophotométrie non calibrée en présence d'écarts au modèle lambertien : (a) une image de synthèse (parmi 8) d'une théière lambertienne sans ombre portée ; (b) champ de normales estimé, codé dans les trois canaux (R,V,B) ; (c-d) idem lorsque la théière comporte des taches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

En effet, la parade classique à la présence de telles *données aberrantes (outliers)* consiste à estimer  $M$  et  $S$  au sens des moindres carrés, en modélisant les écarts au modèle par un bruit blanc additif :

$$\forall p \in [1, |\Omega|], I_p = M_p S + \eta_p \quad (21)$$

qui suit une distribution gaussienne centrée :

$$\forall p \in [1, |\Omega|], \eta_p \sim \mathcal{N}(0_{\mathbb{R}^{m \times 1}}, \Sigma) \quad (22)$$

où  $\Sigma$  est une matrice  $m \times m$  diagonale (les coefficients non diagonaux sont nuls car le bruit dans chaque image est supposé indépendant des autres images). Si les éclairages ont tous la même intensité, on peut même écrire  $\Sigma = \sigma^2 I_m$ , où  $I_m$  désigne la matrice identité de taille  $m \times m$ .

Sous l'hypothèse lambertienne, l'estimateur aux moindres carrés obtenu par SVD constitue une bonne estimation (il est d'ailleurs égal à l'estimateur du maximum de vraisemblance). S'il paraît raisonnable de modéliser de la sorte le bruit induit par le capteur et l'approximation liée à la quantification du niveau de gris, ce choix paraît en revanche beaucoup plus discutable pour la modélisation des ombres et des reflets brillants, qui ne sont pas par nature des phénomènes stochastiques, et dont la distribution dépend de l'objet à reconstruire. Yuille *et al.* ont largement étudié dans (Yuille *et al.*, 1999) l'effet de ces phénomènes sur la factorisation par SVD. Il ressort de leur étude que cette factorisation ne peut corriger que de faibles écarts au modèle, si le nombre d'images est très élevé, la loi des grands nombres permettant effectivement sous certaines hypothèses d'identifier la distribution des écarts à une loi gaussienne.

Une stratégie pour s'affranchir de l'influence des niveaux de gris ne suivant pas le modèle lambertien linéaire (1), qui constituent des données aberrantes, consiste à corriger les images *a priori* comme cela a été proposé dans (Wu *et al.*, 2010). Nous proposons une approche différente, qui consiste simplement à les retirer de la matrice  $I$  pour procéder à l'estimation des éclairages, puis à estimer *a posteriori* le champ  $M$ .

En effet, remarquons que pour résoudre le problème UPS, on cherche à estimer des paramètres globaux (la matrice  $S$ ) et des paramètres locaux ( $M_p$  en chaque pixel  $p$ ). La valeur de  $S$  étant la même pour des points conformes au modèle ou présentant des écarts au modèle, il semble raisonnable de commencer par estimer  $S$ , en considérant que certaines des lignes de  $I$  constituent des données aberrantes : on estime  $S$  à partir d'une sous-matrice de  $I$  obtenue en excluant ces lignes (qui correspondent aux points jugés suspects). Nous proposons donc de résoudre le problème UPS de façon robuste ainsi :

1. Détection des points de  $\Omega$  conformes au modèle, et factorisation des données correspondantes par SVD, pour obtenir  $S$  à l'ambiguïté  $S \mapsto A^{-1}S$  près,  $A \in \text{GL}(3)$ . Ce point fait l'objet de la prochaine partie.
2. Estimation de  $M$  en tout point de  $\Omega$ , à l'ambiguïté  $M \mapsto MA$  près (cf. partie 4.3).
3. Résolution de l'ambiguïté  $M \mapsto MA$ , *i.e.* détermination de  $A$  (ce point a été traité dans la partie 3).

#### 4.2. Détection des points conformes au modèle

Soit  $I \in \mathbb{R}^{|\Omega| \times m}$  la matrice des niveaux de gris, où  $\Omega$  est le domaine de reconstruction et  $m$  le nombre d'images. Soit  $\bar{\Omega} \subset \Omega$  un sous-ensemble de points de  $\Omega$ , et soit  $\bar{I} \in \mathbb{R}^{|\bar{\Omega}| \times m}$  la sous-matrice de  $I$  obtenue en éliminant les lignes correspondant aux points de  $\Omega \setminus \bar{\Omega}$ . Soient alors  $\bar{M}$  et  $S$  le champ et la matrice d'éclairage obtenus par SVD de  $\bar{I}$  puis projection sur l'espace des matrices de rang 3.

On cherche le plus grand sous-ensemble  $\bar{\Omega}$  sur lequel le modèle lambertien (1) est valide, *i.e.* tel que  $\bar{I} = \bar{M}S$ . Pour mesurer l'écart entre  $\bar{I}$  et  $\bar{M}S$ , nous utilisons la racine carrée de l'erreur quadratique moyenne (REQM) entre  $\bar{I}$  et son approximation  $\bar{M}S$  par une matrice de rang 3, que nous définissons localement en chaque pixel  $p$  par :

$$\text{REQM}_{\text{loc}}(\bar{I}_p, \bar{M}_p S) = \sqrt{\frac{1}{m} \sum_{j=1 \dots m} (\bar{I}_p^j - \bar{M}_p S^j)^2} \quad (23)$$

et de façon globale par :

$$\text{REQM}(\bar{I}, \bar{M}S) = \sqrt{\frac{1}{m|\bar{\Omega}|} \sum_{p \in \bar{\Omega}} \sum_{j=1 \dots m} (\bar{I}_p^j - \bar{M}_p S^j)^2} \propto \|\bar{I} - \bar{M}S\|_F \quad (24)$$

Le problème consiste donc à trouver un ensemble  $\bar{\Omega}$  contenant suffisamment de points pour permettre l'estimation des paramètres (ce nombre est calculé dans le paragraphe suivant) et minimisant cette erreur globale, en rejetant les données aberrantes (les niveaux de gris corrompus par des ombres ou des reflets spéculaires). Ce faisant, on dispose d'une estimation robuste de  $S$ , à l'ambiguïté  $S \mapsto A^{-1}S$  près.

##### 4.2.1. Tirages aléatoires (approche RANSAC)

La façon la plus naturelle de réaliser cette estimation est d'utiliser l'algorithme RANSAC (Fischer, Bolles, 1981), qui étant donné un modèle connu (ici la loi de Lambert) estime dans le même temps les paramètres du modèle (la matrice  $S$ ) et 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  $w$  points aléatoirement dans  $\Omega$ , qui constituent l'ensemble courant.
2. Estimer  $S$  et  $\bar{M}$  par SVD de la sous-matrice  $\bar{I}$  de  $I$  correspondante.
3. Pour tous les autres points  $p$  de  $\Omega$ , estimer  $M_p = I_p S^+$ , et ajouter  $p$  à l'ensemble courant si  $\text{REQM}_{\text{loc}}(I_p, M_p S) < \tau$ , où  $\tau$  est un seuil qui, idéalement, est lié à la variance du bruit dans les images (malheureusement inconnue).
4. Si le nombre d'éléments de l'ensemble courant est suffisant, réestimer  $S$  et  $\bar{M}$  sur cet ensemble, et calculer  $\text{REQM}(\bar{I}, \bar{M}S)$  par (24) : si cette erreur est inférieure à l'erreur courante, accepter l'estimation de  $S$ . En pratique, nous réestimons  $S$  dès que le cardinal de l'ensemble courant est supérieur à un seuil  $d$  arbitraire.

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 donc :

- le nombre d'itérations,
- le nombre  $\omega$  de points tirés aléatoirement à chaque itération,
- le seuil  $\tau$  pour accepter un point dans l'ensemble courant,
- le nombre minimal  $d$  de points que doit contenir l'ensemble courant pour être accepté.

Moyennant des connaissances a priori sur la distribution des données aberrantes, une valeur théorique du nombre d'itérations à effectuer peut être calculée. Dans les tests, nous avons choisi arbitrairement d'itérer 5000 fois.

Pour choisir la valeur de  $\omega$ , remarquons que l'on dispose à chaque étape 2 de  $\omega m$  données ( $\bar{I}$  est de taille  $\omega \times m$ ) et que l'on cherche à estimer les  $3m$  éléments de  $S$  ainsi que les  $3\omega$  éléments de  $\bar{M}$ . Il faut donc garantir la condition  $\omega m \geq 3m + 3\omega$  soit  $\omega \geq 3m/(m - 3)$  (on suppose  $m > 3$ , ce qui est de toute façon nécessaire pour pouvoir comparer  $I$  à sa projection sur l'espace des matrices de rang 3, obtenue par SVD). En pratique, nous choisissons pour  $\omega$  la plus petite valeur vérifiant cette condition, et nous choisissons arbitrairement  $d = 5\omega$ .

Reste à fixer la valeur de  $\tau$ . Nous utilisons ici arbitrairement  $\tau = 20$ .

Cet algorithme fournit donc un sous-ensemble  $\bar{\Omega}$  de  $\Omega$  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.

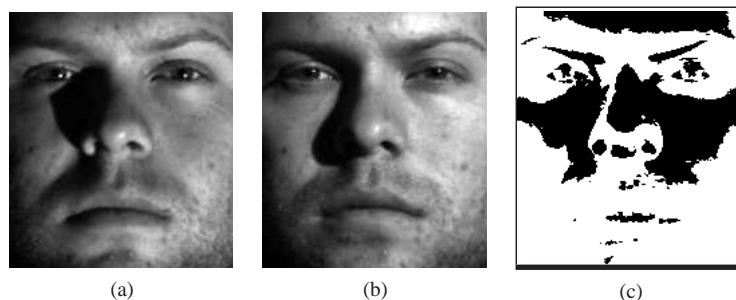


Figure 11. Détection des données aberrantes par l'algorithme RANSAC. (a-b) Deux images d'un visage (parmi 22). (c) Points correspondant au modèle (en blanc) avec les valeurs de paramètres indiquées ci-dessus. Les points situés près du nez (ombres portées), des yeux et du front (reflets spéculaires) ont bien été éliminés.

#### 4.2.2. Détection directe

L'approche précédente offre d'excellents résultats, mais au prix du réglage empirique de nombreux paramètres, et surtout avec un temps de calcul très élevé : en théorie, un nombre infini d'itérations est nécessaire. Nous proposons donc une approche plus efficace en terme de temps de calcul. Plutôt que de procéder à une infinité de tirages aléatoires pour minimiser l'erreur (24), nous faisons l'approximation :

$$\text{REQM}(\bar{I}, \bar{M}S) \approx \sqrt{\frac{1}{m|\bar{\Omega}|} \sum_{p \in \bar{\Omega}} \sum_{j=1 \dots m} (I_p^j - M_p S^j)^2} \propto \|\bar{I} - \bar{M}S\|_F \quad (25)$$

où  $M$  et  $S$  sont obtenues par SVD de  $I$  : on calcule une seule SVD au lieu d'une par itération dans l'approche précédente. Cette approximation est d'autant plus justifiée que la proportion de données aberrantes est plus faible : on perd en robustesse ce qu'on gagne en efficacité algorithmique. Toutefois, puisque la stéréophotométrie ne peut se pratiquer qu'en conditions contrôlées, il est raisonnable de supposer que les images ont été prises de façon à limiter l'apparition de phénomènes comme les ombres. On peut ainsi, pour chaque pixel  $p$ , évaluer l'écart entre  $I_p$  et sa projection sur l'espace des matrices de rang 3 :

$$\text{REQM}_{\text{loc}}(I_p, M_p S) = \sqrt{\frac{1}{m} \sum_{j=1 \dots m} (I_p^j - M_p S^j)^2} \quad (26)$$

On choisit alors de rejeter les pixels  $p$  de  $\Omega$  tels que  $\text{REQM}_{\text{loc}}(I_p, M_p S) > \tau'$  (ici,  $\tau' = 5$ ). On estime ensuite  $S$  par SVD de la matrice  $\bar{I}$  ainsi obtenue. En procédant de la sorte, il n'y a qu'un seul paramètre  $\tau'$  à régler, et surtout l'algorithme n'est pas itératif : on ne calcule qu'une seule SVD et un seuillage. Un exemple de résultat est montré sur la figure 12.

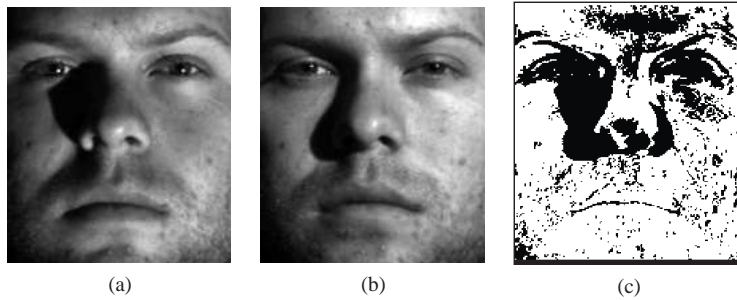


Figure 12. Détection des données aberrantes par SVD. (a-b) Deux images d'un visage (parmi 22). (c) Points correspondant au modèle pour  $\tau' = 5$ . Ce résultat est assez proche de celui de la figure 11, mais il est obtenu beaucoup plus rapidement !



### 4.3. Estimation robuste du champ $M$

Nous supposons maintenant que  $S$  est connue à l'ambiguïté  $S \mapsto A^{-1}S$  près, et nous décrivons une méthode pour estimer le champ  $M$  en tenant compte des ombres et des spécularités. À ce stade, il serait possible d'utiliser un algorithme robuste comme *Expectation-Maximization* (Verbiest, Van Gool, 2008) ou RANSAC (Mukaigawa *et al.*, 2007), mais ces méthodes étant très gourmandes en ressources, nous proposons ici une méthode plus simple, fondée sur une pondération découlant directement des images.

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

$$I = MS \quad (27)$$

où  $S$  est connue à une ambiguïté près.

Évidemment, la solution  $M = IS^+$  ne peut être acceptée, car elle induirait un biais important aux points contaminés par une ombre ou une spécularité. De plus, tous les points n'étant pas affectés de la même manière par ces défauts, il n'est pas raisonnable d'envisager un traitement global. Il faut donc revenir à la formulation locale :

$$I_p = M_p S, \quad p \in [1, |\Omega|] \quad (28)$$

Si l'on modélise les écarts au modèle par un bruit blanc gaussien, l'équation (28) doit être remplacée par :

$$I_p = M_p S + \eta_p, \quad p \in [1, |\Omega|] \quad (29)$$

où  $\eta_p \sim \mathcal{N}(0_{\mathbb{R}^m \times 1}, \Sigma_p)$  et  $\Sigma_p$  est une matrice  $m \times m$  diagonale, proportionnelle à la matrice identité puisqu'on suppose tous les éclairages de même intensité.

Sous cette hypothèse, une solution robuste peut être obtenue aux moindres carrés :

$$\widehat{M}_p = \operatorname{argmin}_{M_p \in \mathbb{R}^{3 \times 1}} \left\{ \sum_{j=1 \dots m} (I_p^j - M_p S^j)^2 \right\}, \quad p \in [1, |\Omega|] \quad (30)$$

ce qui fournit la solution suivante :

$$\widehat{M}_p = I_p S^+, \quad p \in [1, |\Omega|] \quad (31)$$

La résolution est donc locale, ce qui permet de ne pas faire l'hypothèse que le bruit suit la même distribution en tout point. En contrepartie, (31) nécessite de calculer  $|\Omega|$  produits matriciels, donc on perd en efficacité, à moins de paralléliser le calcul, puisque cette formulation est utilisable en chaque pixel, indépendamment des autres pixels.

Toutefois, comme nous l'avons signalé dans la partie 4.1, il n'est pas très réaliste de modéliser les ombres et les reflets spéculaires par un bruit blanc gaussien. Une solution

plus robuste consiste à donner moins de poids aux données que l'on juge corrompues. On peut donc modifier l'équation (30) pour obtenir le problème aux *moindres carrés pondérés* suivant :

$$\widehat{M}_p = \operatorname{argmin}_{M_p \in \mathbb{R}^{3 \times 1}} \left\{ \sum_{j=1 \dots m} w_p^{j^2} (I_p^j - M_p S^j)^2 \right\}, p \in [1, |\Omega|] \quad (32)$$

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

La solution de ce problème aux moindres carrés pondérés est donnée par :

$$\widehat{M}_p = (I_p W_p) (S W_p)^+, p \in [1, |\Omega|] \quad (33)$$

où  $W_p$  est la matrice diagonale formée par les poids  $w_p^j$ ,  $j = 1 \dots m$ .

Reste maintenant à définir ces poids. Pour cela, remarquons que les écarts au modèle considérés affectent les images de deux façons :

- Les ombres sont caractérisées par des niveaux de gris faibles.
- Les reflets spéculaires sont caractérisés des niveaux de gris élevés.

Pour pénaliser simultanément ces deux types d'écarts, on peut donc choisir un poids  $w_p^j$  dépendant de  $|I_p^j - 128|$ , si les niveaux de gris varient entre 0 et 255. Nous proposons le choix suivant :

$$w_p^j = 128 - |I_p^j - 128| + \epsilon, p \in [1, |\Omega|], j \in [1, m] \quad (34)$$

où  $\epsilon$  est choisi très petit de façon à garantir  $w_p^j > 0$ ,  $j \in [1, m]$ ,  $p \in [1, |\Omega|]$ .

Notons également que le calcul du produit matriciel  $M_p S$  fournit des niveaux de gris « corrigés ». En effet, aux points où l'angle entre la normale et la direction d'éclairage est supérieur à  $\frac{\pi}{2}$ , on a  $M_p S < 0$  : on obtient donc des niveaux de gris corrigés négatifs à l'intérieur des ombres propres (où le niveau de gris original est nul, au bruit près, cf. équation (19)). En ne prenant pas en compte de terme spéculaire (le terme additif de l'équation (20), par exemple), on obtient aux pixels concernés un niveau de gris inférieur au niveau de gris initial. Enfin, les ombres portées sont supprimées. Puisque nous corrigeons les images grâce à un modèle a priori (la loi de Lambert), cette correction d'images est plus réaliste que celle proposée dans (Wu *et al.*, 2010), qui est effectuée uniquement grâce à des considérations sur le rang de  $I$ , sans prendre en compte la géométrie de l'objet ni les conditions d'éclairage, c'est-à-dire ni  $M$  ni  $S$ .

Puisque  $S$  est connue à l'ambiguïté  $S \mapsto A^{-1}S$  près, le champ  $M$  est de la même façon connu à l'ambiguïté  $M \mapsto MA$  près. En effet,  $MS = MAA^{-1}S$ . Il reste donc à lever cette ambiguïté par la méthode proposée dans la partie 3. Rappelons que l'ambiguïté est d'abord réduite à une ambiguïté de bas-relief en imposant la contrainte d'intégrabilité sur  $M$  (qui est estimée ligne par ligne, comme indiqué en (33)), puis l'ambiguïté de bas-relief est levée en rajoutant une contrainte (dans notre cas, nous faisons à nouveau l'hypothèse que les éclairages ont tous la même intensité).

## 5. Évaluation de la méthode de résolution proposée

Notre méthode de résolution de la stéréophotométrie non calibrée en présence d'écarts au modèle lambertien peut être résumée comme suit :

1. Détecter les points conformes au modèle lambertien. À partir de ces points, estimer les éclairages  $S$  à une ambiguïté près (cf. partie 4.2.2).
2. Estimer le champ  $M$  à une ambiguïté près, à partir des éclairages précédemment estimés (cf. partie 4.3).
3. Lever cette ambiguïté en imposant l'intégrabilité du champ de normales et l'uniformité des intensités lumineuses (cf. partie 3). Tout se passe comme si on appliquait cette méthode à la matrice corrigée  $\tilde{I} = MS$ .

À l'issue de ces 3 étapes, nous obtenons donc des estimations précises de  $S$ ,  $N$  et  $\rho$ , qui tiennent compte de la présence d'écarts au modèle lambertien. Évaluons maintenant la précision de ces estimations.

### 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 ainsi :

- L'albédo est choisi uniforme, de même que les intensités lumineuses.
- Le niveau de gris est calculé par le modèle de Phong (20), de façon à faire apparaître des ombres propres et des taches spéculaires.
- Pour simuler des ombres portées, des carrés noirs sont superposés aux images.
- Enfin, un faible bruit gaussien (d'écart-type égal à 1% du niveau de gris maximal) est ajouté à chaque image.

Le relief original, ainsi que 3 des 22 images simulées à partir de ce relief, sont montrés sur la figure 13.

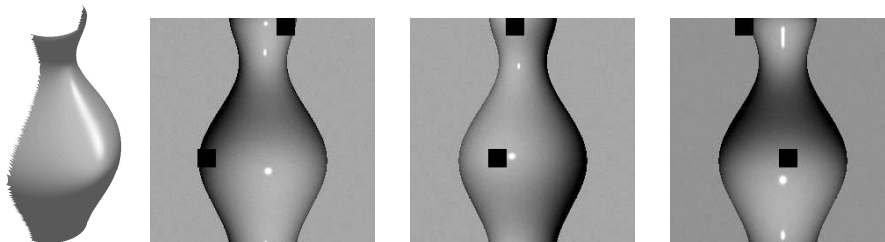


Figure 13. Relief représentant un vase, et trois images de ce vase simulées sous différents éclairages, en utilisant le modèle de Phong. Des carrés noirs ont été superposés aux images pour simuler des ombres portées.

Notre méthode est censée retrouver des éclairages et des normales bien orientés, ainsi que des albédos et des intensités lumineuses aussi uniformes que possible. Nous utilisons plusieurs critères pour évaluer chaque résultat :

- l'écart angulaire moyen sur les normales, noté EAM-N,
- l'écart angulaire moyen sur les éclairages, noté EAM-L,
- l'écart-type des intensités lumineuses (mises à l'échelle entre 0 et 1), noté  $\sigma_{||S||}$ ,
- l'écart-type de l'albédo (mis à l'échelle entre 0 et 1), noté  $\sigma_\rho$ ,
- le temps de calcul, noté  $t$ .

Les résultats obtenus par notre méthode, notée R-UPS, sont comparés à trois autres méthodes :

- La méthode CPS correspond à la stéréophotométrie calibrée non robuste : équation (4).
- La méthode UPS correspond à la stéréophotométrie non calibrée non robuste : factorisation par (Yuille, Snow, 1997), puis résolution de l'ambiguïté comme présenté dans la partie 3.
- La méthode LRA correspond à la méthode (Wu *et al.*, 2010) : correction a priori des images, puis méthode UPS.

La méthode CPS est souvent utilisée comme référence pour évaluer les résultats des méthodes UPS. Mais ici, puisqu'on tient compte des écarts au modèle lambertien, on espère obtenir de meilleurs résultats avec la méthode R-UPS. Notons aussi qu'afin de pouvoir comparer les temps de calcul, chaque méthode est implémentée dans le même langage (Matlab) et testée sur la même machine.

Afin de ne pas biaiser les comparaisons, les trois méthodes non calibrées utilisent bien sûr la même méthode pour lever l'ambiguïté de bas-relief (méthode proposée dans la partie 3, qui est la plus adaptée ici puisqu'on a pris soin, à la création des images, de choisir des éclairages de même intensité). De même, les différents 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 présenté dans le tableau 2. Les reliefs reconstruits par ces quatre méthodes sont montrés sur la figure 14, et les différentes estimations de l'albédo sur la figure 15. Il est notable que le relief estimé par UPS est extrêmement déformé : cela provient de la mauvaise factorisation de  $I$  par SVD, qui fournit un champ  $M$  tellement dégradé que la contrainte d'intégrabilité par la méthode (Yuille, Snow, 1997) n'est que très approximativement respectée.

Ces résultats prouvent clairement l'intérêt de notre approche, puisque les différents paramètres du modèle sont mieux estimés : 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)	$\sigma_{  S  }$	$\sigma_\rho$	$t$ (s)
CPS	6,11	0	0	0,06	0,24
UPS	55,95	45,30	0,05	0,25	<b>0,90</b>
LRA	11,57	12,09	<b>0,003</b>	0,06	4,47
R-UPS	<b>1,54</b>	<b>1,55</b>	0,004	<b>0,01</b>	1,91

TABLE 2. Comparaison des performances sur les images du vase. La méthode UPS est totalement mise en défaut. La méthode CPS, malgré la connaissance des sources, induit un biais sur l'estimation des normales. La méthode LRA améliore largement les résultats UPS, mais est elle-même dépassée par la méthode R-UPS, qui retrouve un albédo quasiment uniforme, ainsi que des sources lumineuses et des normales très proches des valeurs réelles.

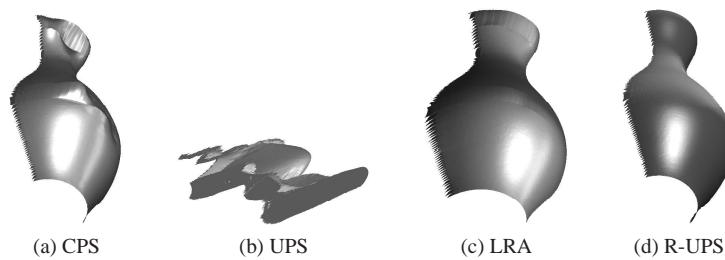


Figure 14. Reconstructions 3D du vase. Les reconstructions (a) et (b) sont manifestement très influencées par la présence d'écarts au modèle. La reconstruction (c) est bien meilleure que (b), mais un léger artefact est toujours visible. La méthode proposée permet de retrouver un relief (d) très proche 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.

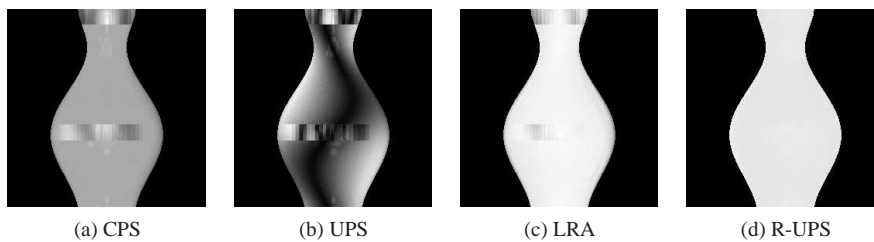


Figure 15. Albédos estimés pour le vase (mis à l'échelle entre 0 et 255 pour l'affichage). Parmi les quatre méthodes étudiées, seule la méthode proposée permet de retrouver un albédo quasiment uniforme.

## 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 disponibles. Les directions d'éclairage étant par contre disponibles dans le cas des images de (Georgiades *et al.*, 2001), nous donnons dans le tableau 3 les valeurs obtenues pour EAM-L. Dans la plupart des cas, l'estimation robuste que nous proposons permet de gagner en précision sur les éclairages. Ce résultat est intéressant car il montre que le fait de ne pas connaître les éclairages n'est plus nécessairement un obstacle insurmontable : l'estimation des sources lumineuses à partir des images permet de retrouver, à très peu de choses près, les données de la stéréophotométrie calibrée, pour laquelle l'estimation précise des sources lumineuses nécessite une intervention spécifique de l'utilisateur (en général, une sphère en chrome est introduite dans la scène, ce qui fournit la direction de l'éclairage).

Nous montrons ici quelques 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 une caméra. Huit sources lumineuses sont allumées successivement. Le domaine de reconstruction est déterminé à la main. Trois images de deux jeux d'images ainsi acquises sont montrées sur les deux premières lignes de la figure 16. Les reconstructions 3D obtenues sont présentées sur les figures 17 et 18. Les ombres portées présentes sur les images n'influent pas sur le relief reconstruit, et les détails les plus fins, comme les granularités sur le bas-relief, sont bien retrouvés.

Enfin, nous avons également appliqué notre méthode aux jeux d'images « Frog » et « Korean Doll » (deux dernières lignes de la figure 16) 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 19 et 20) sont qualitativement équivalents à ceux présentés dans (Chandraker *et al.*, 2007), où des données calibrées sont utilisées.

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

TABLE 3. Comparaison de EAM-L (en degrés) sur des données réelles de visage, entre les méthodes UPS, LRA et R-UPS. Les temps de calcul sont du même ordre de grandeur que ceux du tableau 2.

## 6. Conclusion et perspectives

Dans cet article, après avoir proposé une solution simple pour résoudre l'ambiguïté de bas-relief (partie 3), nous avons montré dans la partie 4 comment adapter la méthode de résolution traditionnelle de la stéréophotométrie non calibrée pour prendre en compte différents écarts au modèle lambertien tels que les ombres ou les reflets spéculaires. Plus précisément, ces écarts affectant à la fois l'estimation des sources lumineuses et celle du champ de normales, nous conseillons de détecter les pixels où la loi de Lambert n'est pas vérifiée, par simple comparaison des données aux images obtenues par reprojection. L'estimation des éclairages peut alors se faire uniquement grâce à un sous-ensemble de pixels, avant d'estimer dans un second temps le champ de normales et l'albédo 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 les plus récentes.

La méthode proposée ne peut toutefois être appliquée que si l'objet à reconstruire ne présente pas « trop » d'écarts au modèle lambertien : les objets très brillants comme certains métaux sont par exemple exclus. Plutôt que de chercher à éliminer les points suspects, il faudrait dans ce cas plutôt chercher à adapter la méthode UPS à des modèles de réflectance plus appropriés. Néanmoins, la formulation du problème n'étant alors plus matricielle, cela nécessiterait la mise en place de méthodes d'optimisation spécifiques. Tout au long de l'exposé, nous avons également supposé que les éclairages consistaient en des faisceaux parfaitement parallèles, ce qui permet une formulation matricielle du problème, mais n'est pas facile à garantir en pratique. Une extension intéressante de ce travail consisterait à remettre en question cette hypothèse (en considérant des sources ponctuelles plutôt que directionnelles). Il faudrait néanmoins prendre en compte l'atténuation de l'intensité lumineuse, ce qui rendrait 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 méthodes 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. Chacun de ces sous-problèmes est indépendant des autres, et l'assemblage de solutions efficaces à chaque sous-problème dans le cadre proposé ici doit permettre, à terme, d'obtenir une méthode très performante de résolution 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.
- Breuss M., Cristiani E., Durou J.-D., Falcone M., Vogel O. (2012). Perspective Shape from Shading: Ambiguity Analysis and Numerical Approximations. *SIAM Journal on Imaging Sciences*, vol. 5, n° 1, p. 311–342.
- 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*.
- Daniel P. (2000). Peut-on extraire le relief d'une seule image ? *Thèse de l'Université Paul Sabatier, Toulouse*.
- 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., Papadimitri 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.
- Horowitz 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.
- 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.
- Lin Z., Chen M., Ma Y. (2009). *The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-rank Matrices*. Tech. Rep. n° UILU-ENG-09-2215, UIUC.
- Mecca R., Durou J.-D. (2011). Unambiguous photometric stereo using two images. In *ICIAP 2011*, p. 286–295.
- Mukaigawa Y., Ishii Y., Shakunaga T. (2007). Analysis of photometric factors based on photometric linearization. *JOSA A*, vol. 24, n° 10, p. 3326–3334.
- Mulaik S. A. (1972). *The foundations of factor analysis* (vol. 88). McGraw-Hill New York.
- 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.
- Prados E., Faugeras O. (2005). Shape from shading: a well-posed problem? In *CVPR 2005*, vol. 2, p. 870–877.
- 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*.
- 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.

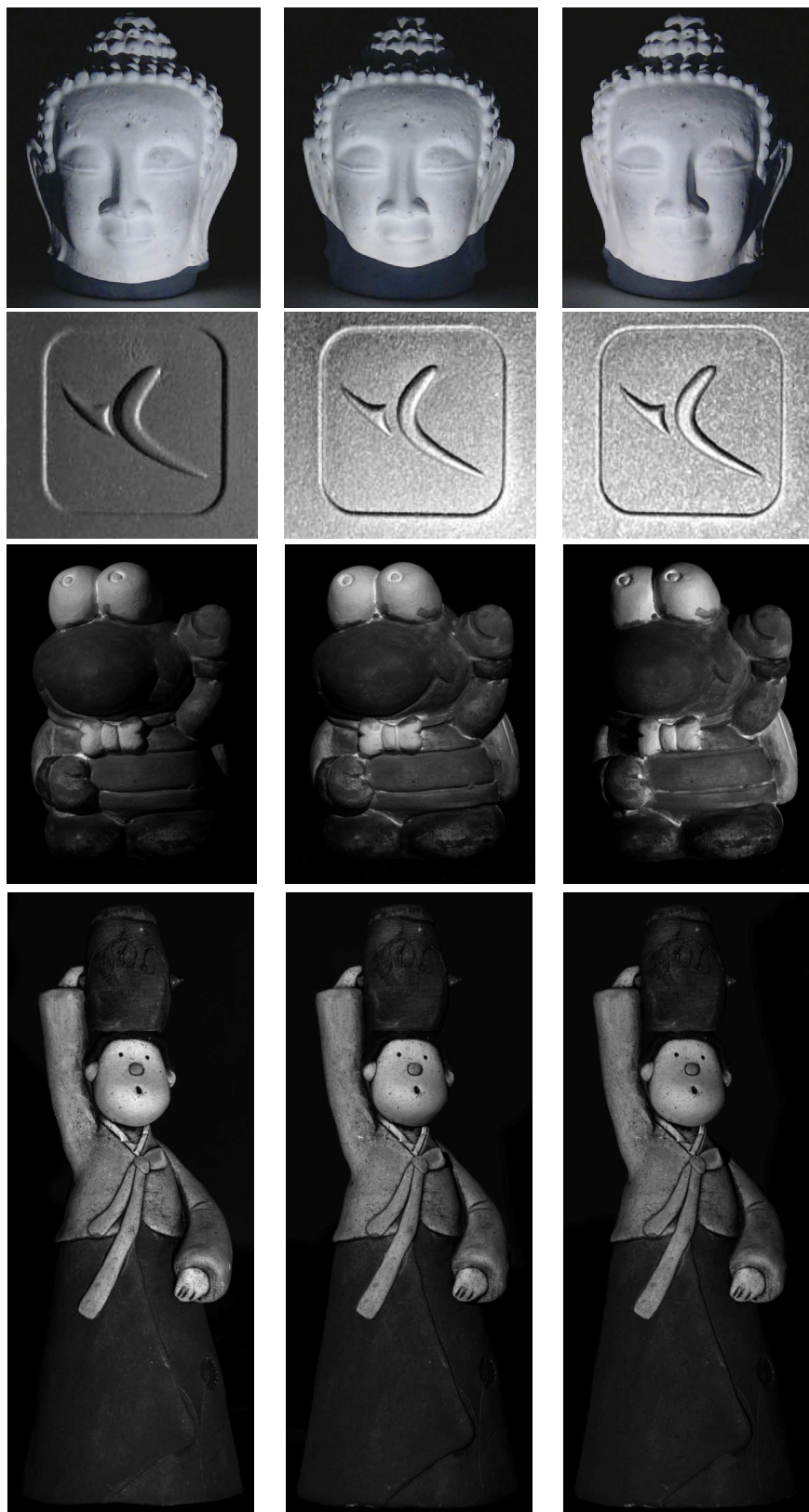


Figure 16. Images réelles. Première ligne : un buste de Buddha, sur lesquelles de nombreuses ombres sont visibles. Deuxième ligne : un bas-relief comportant des ombres portées et des reflets spéculaires. Troisième et quatrième lignes : deux figurines comportant des ombres.



Figure 17. Reconstruction 3D du buste de Buddha, sous différents angles. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D (avec plaquage de l'albedo).

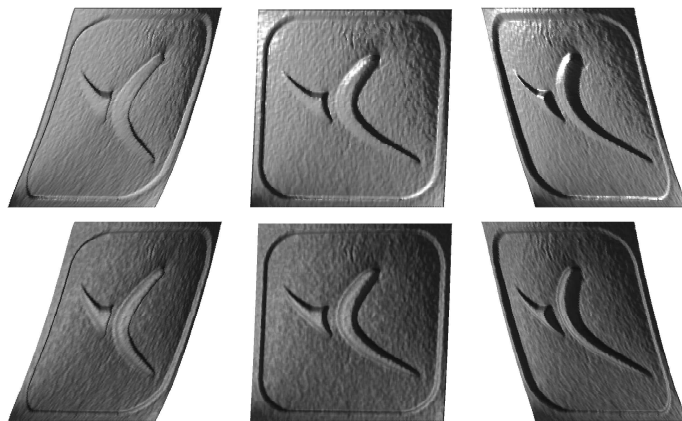


Figure 18. Reconstruction 3D du bas-relief. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D.

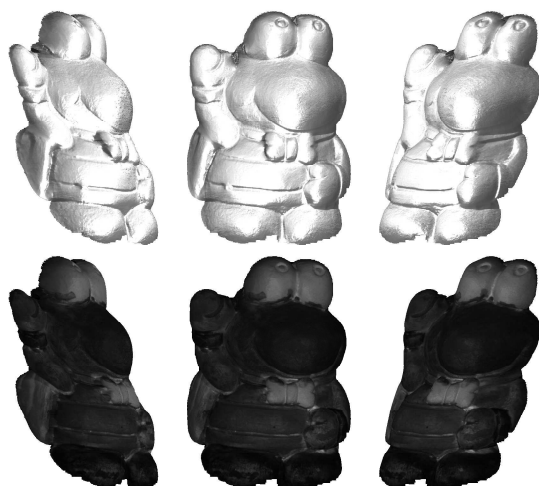


Figure 19. Reconstruction 3D de la première figurine de la figure 16. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D. Les nombreuses ombres n'influent pas sur la reconstruction 3D.

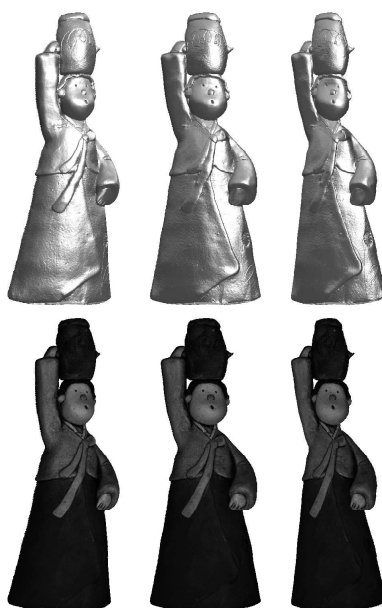


Figure 20. Reconstruction 3D de la seconde figurine de la figure 16. Première ligne : relief rééclairé. Deuxième ligne : modèle 3D. Sur la première ligne, on observe que le motif très fin gravé sur le vase est correctement reconstruit.