HAL
open science

Résolution du problème de la stéréophotométrie non calibrée par estimation de l'intensité des éclairages

Yvain Quéau, Jean-Denis Durou

To cite this version:

Yvain Quéau, Jean-Denis Durou. Résolution du problème de la stéréophotométrie non calibrée par estimation de l'intensité des éclairages. Congrès des jeunes chercheurs en vision par ordinateur (ORASIS 2013), Jun 2013, Cluny, France. hal-00829380

HAL Id: hal-00829380

https://hal.science/hal-00829380v1

Submitted on 5 Jun 2013

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.

Résolution du problème de la stéréophotométrie non calibrée par estimation de l'intensité des éclairages

Yvain Quéau Jean-Denis Durou

IRIT, UMR CNRS 5505, Toulouse

yvain.queau@enseeiht.fr durou@irit.fr

Résumé

Nous présentons une nouvelle méthode de résolution du problème de la stéréophotométrie non calibrée dans le cadre lambertien, fondée sur l'hypothèse que tous les éclairages ont la même intensité, couplée à l'estimation de cette intensité. Nous montrons comment cette hypothèse permet de lever les ambiguïtés inhérentes à la linéarité du modèle lambertien, et notamment de résoudre l'ambiguïté de bas-relief. Le problème devenant alors bien posé, nous proposons une méthode complète pour résoudre le problème de la stéréophotométrie non calibrée et ainsi estimer conjointement les conditions d'éclairage et le champ de normales. Nous validons notre méthode par la reconstruction 3D de visages à partir d'images réelles, et nous comparons son efficacité et sa précision aux techniques les plus récentes de la littérature.

Mots Clef

Reconstruction 3D, *shape-from-X*, stéréophotométrie non calibrée, estimation de l'éclairage, ambiguïté de bas-relief.

Abstract

In this paper we present a new method for solving the uncalibrated photometric stereo problem in the Lambertian framework, based on the hypothesis that all lightings share the same intensity, which we can estimate. We show how this estimation allows one to solve the ambiguities of the linear Lambertian model, especially the so-called generalized bas-relief ambiguity. As the problem becomes well-posed, we propose a new full method for solving the uncalibrated photometric stereo problem, and thus estimate both the lighting conditions and the normal field. Our method is validated by reconstructing shapes of faces from real-world images, and we compare its efficiency and accuracy to the most state-of-the-art existing techniques.

Keywords

3D-reconstruction, shape-from-X, uncalibrated photometric stereo, lighting estimation, bas-relief ambiguity.

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 place font que cette technique de reconstruction 3D, qui n'est pas nouvelle puisqu'elle a été introduite dès le début des années 1980 [23], connaît depuis plusieurs années un regain d'intérêt. Elle présente l'avantage, contrairement à la plupart des autres techniques de reconstruction 3D, d'estimer également 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 le champ de normales 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 [11, 18, 9].

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, le réalisme des hypothèses conditionne la qualité des reconstructions 3D effectuées à partir d'images réelles.

Après avoir montré dans la partie 3 pourquoi l'hypothèse d'intensité des éclairages constante n'est généralement pas correctement formulée, nous montrons dans la partie 4 qu'en supposant les sources lumineuses de même intensité, on peut lever les ambiguïtés et résoudre le problème de la stéréophotométrie non calibrée, tout en estimant l'intensité commune aux différents éclairages. Nous détaillons la méthode proposée dans la partie 5 et comparons, sur images réelles, son efficacité aux techniques actuellement les plus performantes.

2 Travaux antérieurs

2.1 Stéréophotométrie

La stéréophotométrie, introduite par Woodham dans [23], permet de retrouver la normale et l'albédo en chaque pixel visible d'un objet à partir de $m \geqslant 3$ 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 [6, 15] 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* [17], 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 $m \geqslant 3$ 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, $N_p = [N_x, N_y, N_z](p)$ le vecteur normé normal à la surface en p, et en représentant l'éclairage dans la $i^{\text{ème}}$ image, en norme et direction, par le vecteur $S^i = [S_x^i, S_y^i, S_z^i]^\top$ orienté dans le sens opposé à la propagation de la lumière, le modèle lambertien s'écrit :

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

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 carrées 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 I_p correspondant à tous les pixels p situés à l'intérieur du masque Ω de l'objet ($|\Omega|$ 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 peut alors réécrire $\widehat{M}_p = I_p S^+$ pour obtenir directement :

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

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 $\widehat{N}$.

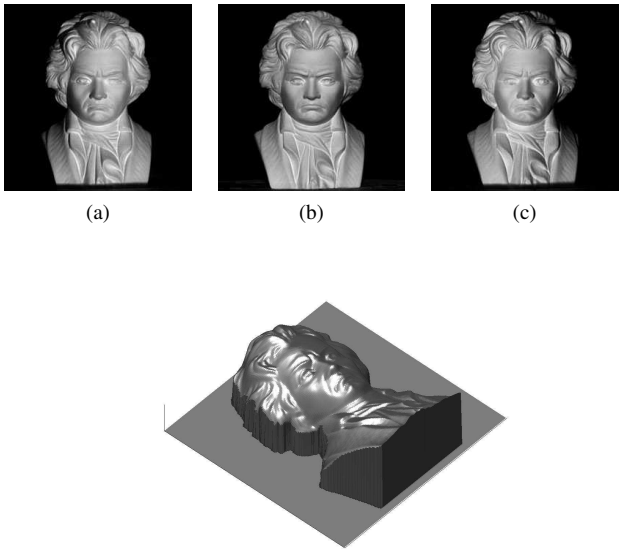

FIGURE 1 – Exemple de reconstruction par stéréophotométrie calibrée à partir de 3 images d'un buste de Beethoven (a-b-c) disponibles sur le web. On voit sur (d) que les détails les plus fins sont bien retrouvés.

Les directions d'éclairage sont généralement supposées non coplanaires pour garantir l'hypothèse rang$(S) = 3$, même si quelques travaux récents [1, 2] se sont intéressés au cas contraire, qui se produit en particulier pour des photographies prises en extérieur, vu que le Soleil se déplace dans un plan, relativement au référentiel terrestre. Signalons enfin que la stéréophotométrie à $m = 2$ images est plus délicate [19], et que le cas où $m = 1$, qui correspond au *shape-from-shading*, l'est encore bien plus.

2.2 Stéréophotométrie non calibrée

Lorsque les directions et les normes des vecteurs d'éclairage sont inconnues, le problème devient mal posé. En factorisant la matrice des niveaux de gris I, Hayakawa décrit dans [14] une méthode qui permet d'obtenir, à une transformation linéaire près, le champ de normales et les vecteurs d'éclairage. On cherche maintenant à estimer à la fois $M(p) = \rho(p)N(p)$ en chaque pixel p et une matrice d'éclairage S de taille $3 \times m$.

Par SVD, I se factorise en $I = UWV^T$, avec $U \in \mathbb{R}^{|\Omega|\times|\Omega|}$, $W \in \mathbb{R}^{|\Omega|\times m}$ et $V \in \mathbb{R}^{m\times m}$. Les éclairages et les normales étant des vecteurs de $\mathbb{R}^3$, le rang de I devrait être égal à 3 et seules les trois plus grandes valeurs singulières devraient être non nulles (en pratique, les autres ne sont pas exactement nulles, à cause de divers écarts au modèle). On peut donc restreindre W à sa sous-matrice principale de dimension 3, et U et V à leurs 3 premières colonnes, c'est-à-dire que $U \in \mathbb{R}^{|\Omega|\times 3}$, $W \in \mathbb{R}^{3\times 3}$ et $V \in \mathbb{R}^{m\times 3}$, de telle sorte que $I \approx UWV^T$.

On obtient alors la solution sous la forme suivante :

$$\begin{cases} \widehat{M} = UP^T \\ \widehat{S} = QV^T \end{cases} \tag{5}$$

où P et Q sont deux matrices 3×3 a priori quelconques telles que :

$$P^T Q = W \tag{6}$$

Néanmoins, le problème est mal posé puisqu'il existe une infinité de tels couples (P, Q). Yuille et Snow montrent dans [26] qu'en imposant la *contrainte d'intégrabilité* sur le champ de normales à reconstruire, l'ambiguïté est réduite à *l'ambiguïté de bas-relief*, qui a été décrite pour la première fois dans [4]. Cette contrainte s'écrit [11] :

$$\frac{\partial}{\partial y}\frac{N_x}{N_z} - \frac{\partial}{\partial x}\frac{N_y}{N_z} = 0 \tag{7}$$

et s'écrit de la même manière en remplaçant N par M. Yuille et Snow montrent que l'on peut estimer six des neuf coefficients de P^{-1} en développant l'équation (7) :

$$P^{-1} = \begin{pmatrix} a_1 & a_2 & b_1 \\ a_3 & a_4 & b_2 \\ a_5 & a_6 & b_3 \end{pmatrix} \tag{8}$$

en l'occurrence les coefficients $a_i,\ i = 1 \ldots 6$. 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 3×3 à trois paramètres (μ, ν, λ), $\lambda \neq 0$, qui s'écrivent :

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\mu}{\lambda} & -\frac{\nu}{\lambda} & \frac{1}{\lambda} \end{pmatrix} \tag{9}$$

Un exemple de transformation de bas-relief sur un relief hémisphérique est montré sur la figure 2.

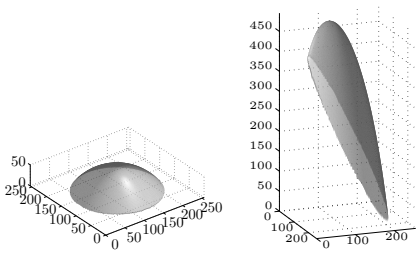

FIGURE 2 – À gauche : exemple de relief hémisphérique. À droite : effet de la transformation de bas-relief correspondant à $\mu = -2$, $\nu = 1$ et $\lambda = 5$.

2.3 Résolution de l'ambiguïté de bas-relief

Pour lever l'ambiguïté de bas-relief (i.e., pour estimer μ, ν et λ), il faut rajouter au moins une hypothèse. Sans utiliser la condition d'intégrabilité, Hayakawa a proposé initialement dans [14] d'identifier six pixels possédant le même albédo ou de supposer l'intensité de l'éclairage identique dans six images au moins. Néanmoins, Yuille et Snow montrent dans [25] que, sans l'hypothèse d'intégrabilité, la première hypothèse ne suffit pas à lever toute l'ambiguïté, tandis que la seconde revient à supposer que l'objet est symétrique, ce qui est fort peu réaliste.

Yuille et Snow proposent dans [26] d'adapter cette seconde hypothèse en prenant en compte l'intégrabilité, ce qui leur permet d'identifier la transformation G. Néanmoins, la méthode qu'ils proposent pour lever l'ambiguïté de bas-relief suppose que l'intensité lumineuse est égale à une constante fixée arbitrairement à 1. Nous montrerons dans la partie 3 que ce choix arbitraire influe sur la qualité de la reconstruction 3D, voire sur sa faisabilité.

D'autres hypothèses pour résoudre l'ambiguïté de bas-relief ont été proposées : Alldrin *et al.* suggèrent dans [3] de minimiser l'entropie de la distribution d'albédo de façon à limiter le nombre de valeurs de l'albédo. Shi *et al.* montrent que l'ambiguïté est levée en utilisant trois canaux de couleur au lieu du seul niveau de gris [20]. En supposant que le matériau présente une composante additive non lambertienne, Tan et Mallick proposent dans [22] 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 [24], on peut également tirer parti de la présence d'ombres [21], de reflets spéculaires [7] ou de réflexions mutuelles [5]. Récemment, il a été proposé dans [10] 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, par exemple en remplaçant le modèle lambertien par celui de Torrance et Sparrow [12]. 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 [16].

Ainsi, il apparaît que l'estimation des trois paramètres μ, ν et λ, dans le cadre de la stéréophotométrie lambertienne non calibrée, reste un problème ouvert, particulièrement difficile car les images produites par les différents jeux d'éclairage et les champs de normales associés sont exactement les mêmes, et que les champs de normales sont tous également intégrables. Nous proposons une nouvelle méthode permettant de résoudre le problème de la stéréophotométrie non calibrée, qui ne nécessite pas d'estimer cette transformation de bas-relief, 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.

3 Pourquoi supposer les éclairages de même intensité est insuffisant

3.1 Méthode de Yuille et Snow

Yuille et Snow proposent d'estimer les six coefficients $a_i,\ i = 1 \dots 6$, puis de fixer b_1, b_2 et b_3 à des valeurs aléatoires. Ils obtiennent ainsi P^{-1} par l'équation (8). Reste à estimer la transformation de bas-relief G définie en (9).
Reprenant une idée de Hayakawa [14], ils supposent l'intensité des éclairages constante, *i.e.* $\forall i \in [1, m],\ {S^i}^T S^i = {S_0}^2$, avec $S_0 > 0$. Ils proposent en outre de **fixer arbitrairement** S_0 **à** 1, ce qui permet d'identifier $Q^T Q$ si $m \geqslant 6$. En effet, en notant V_i chaque ligne de V, et en utilisant (5), on obtient le système suivant :

$$V_i Q^T Q V_i^T = 1, \quad i = 1 \dots m \tag{10}$$

qui est linéaire vis-à-vis des coefficients inconnus de la matrice symétrique $Q^T Q$, qui sont au nombre de 6 puisque $Q^T Q$ est une matrice symétrique 3×3.
On obtient ensuite, par SVD de $Q^T Q$, une matrice Q qui convient, mais cette matrice n'est pas unique car pour toute matrice orthogonale R, $(RQ)^T(RQ) = Q^T Q$, donc RQ convient aussi.
À ce stade, U, W, V, Q et P sont estimés. Il reste alors à estimer G et R, *i.e.* à lever les deux ambiguïtés :

$$P \mapsto G^T P \quad \text{et} \quad Q \mapsto RQ \tag{11}$$

où R est orthogonale et G a la forme (9).
Si la matrice G ne peut être directement estimée, GG^T peut l'être car elle ne dépend que de matrices déjà estimées. En effet, la matrice $J \equiv GR$ peut être évaluée car on tire de (11) et de $P^T Q = W$ l'équation $P^T J Q = W$, dont la solution est $J = P^{-T} W Q^{-1}$. On a donc $JJ^T = GRR^T G^T = GG^T = P^{-T} W Q^{-1} Q^{-T} W P^{-1}$, que l'on peut écrire $GG^T = B$, où B est connue.
Or, d'après (9) :

$$GG^T = \begin{pmatrix} 1 & 0 & \mu \\ 0 & 1 & \nu \\ \mu & \nu & \mu^2 + \nu^2 + \lambda^2 \end{pmatrix} \tag{12}$$

ce qui donne par identification :

$$\begin{cases} \mu = B_{1,3} \\ \nu = B_{2,3} \\ \lambda^2 = B_{3,3} - B_{1,3}^2 - B_{2,3}^2 \end{cases} \tag{13}$$

La solution est unique à la transformation $\lambda \mapsto -\lambda$ près, qui correspond à l'ambiguïté concave/convexe et ne peut être levée sans connaissance supplémentaire sur la forme de l'objet.
Il reste enfin à obtenir R, qui est la solution du problème aux moindres carrés sous contrainte suivant :

$$\begin{cases} \min\limits_{R \in \mathbb{R}^{3\times 3}} ||J - GR||^2 \\ \text{s.c.} \quad R^T R = I \end{cases} \tag{14}$$

Finalement, on obtient $M = UP^T G$ et $S = RQV^T$.

3.2 Validité de la méthode de Yuille et Snow

Dans l'approche précédente, on a supposé que l'intensité de chaque éclairage $||S^i||$ était égale à 1. Ce choix, proposé par Hayakawa dans [14] et par Yuille et Snow dans [26], est loin d'être un choix anodin. Il pourrait même s'avérer désastreux : si l'on décidait de fixer ces normes à une valeur quelconque $||S^i|| = S_0$, $i = 1 \dots m$, l'équation (10) devient :

$$V_i Q^T Q V_i^T = {S_0}^2, \quad i = 1 \dots m \tag{15}$$

Le reste du raisonnement est ensuite identique en remplaçant $Q^T Q$ par ${S_0}^2 Q^T Q$, ce qui aboutit finalement aux identifications suivantes :

$$\begin{cases} \mu = \frac{B_{1,3}}{{S_0}^2} \\ \nu = \frac{B_{2,3}}{{S_0}^2} \\ \lambda^2 = \frac{B_{3,3}}{{S_0}^2} - (\frac{B_{1,3}}{{S_0}^2})^2 - (\frac{B_{2,3}}{{S_0}^2})^2 \end{cases} \tag{16}$$

où les coefficients $B_{i,j}$ sont exactement les mêmes que ceux calculés plus haut.
Il est alors immédiat de constater que si S_0 est suffisamment petit, l'équation $\lambda^2 = \frac{B_{3,3}}{{S_0}^2} - (\frac{B_{1,3}}{{S_0}^2})^2 - (\frac{B_{2,3}}{{S_0}^2})^2$ peut n'admettre aucune solution dans $\mathbb{R}$. Des coefficients complexes dans G n'ayant aucun sens, le choix arbitraire $||S^i|| = 1$ doit donc être remis en question.
L'expérience de la figure 3 montre que la qualité de la reconstruction 3D dépend du choix du paramètre S_0. S'il est trop faible, la reconstruction 3D est impossible car les coefficients obtenus 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 dans la partie 4.

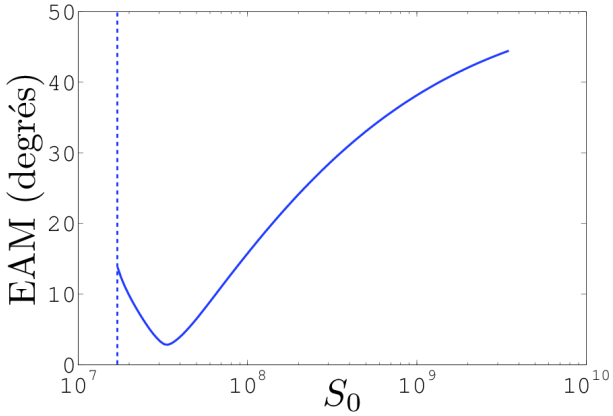

FIGURE 3 – Influence du choix de la constante S_0 sur l'erreur angulaire moyenne (EAM) entre le champ de normales estimé par stéréophotométrie non calibrée et le champ estimé par stéréophotométrie calibrée (données de test : YaleB07_P00). Lorsque S_0 est faible, la reconstruction 3D est impossible car la valeur obtenue pour λ est complexe.

4 Estimer la norme des éclairages lève l'ambiguïté

4.1 Amélioration algorithmique

La méthode précédente consiste à attribuer à b_1, b_2 et b_3 des valeurs aléatoires, puis à corriger ce choix en estimant la « meilleure » transformation de bas-relief. Pour ce faire, on réalise une décomposition en valeurs singulières, une identification assez grossière des coefficients μ, ν et λ dans (12), les valeurs censées être nulles ne l'étant pas vraiment en pratique, et une optimisation (14) non triviale où le compromis entre la satisfaction de la contrainte et la minimisation de l'attache aux données est très délicat à trouver. Pour assurer $||S^i|| = S_0$, $i = 1 \dots m$, il est pourtant possible de procéder beaucoup plus simplement. Introduisons les notations $S_x = [a_1, a_3, a_5]WV^T = [S_x^1 \dots S_x^m]$, $S_y = [a_2, a_4, a_6]WV^T = [S_y^1 \dots S_y^m]$, $A = VW$ et $b = [b_1, b_2, b_3]^T$. D'après (5), $S = QV^T = P^{-T}WV^T = P^{-T}A^T$, ce qui donne en utilisant (8) :

$$S = \begin{pmatrix} S_x \\ S_y \\ b^T A^T \end{pmatrix} \tag{17}$$

Grâce à (17) et aux contraintes $||S^i|| = S_0$, il vient :

$$Ab = \begin{pmatrix} \pm\sqrt{{S_0}^2 - {S_x^1}^2 - {S_y^1}^2} \\ \vdots \\ \pm\sqrt{{S_0}^2 - {S_x^m}^2 - {S_y^m}^2} \end{pmatrix} \tag{18}$$

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 (18) :

$$Ab = \begin{pmatrix} \sqrt{{S_0}^2 - {S_x^1}^2 - {S_y^1}^2} \\ \vdots \\ \sqrt{{S_0}^2 - {S_x^m}^2 - {S_y^m}^2} \end{pmatrix} = K(S_0) \tag{19}$$

Il s'agit d'un système linéaire de m équations à 3 inconnues, 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 :

$$\widehat{b} = A^+ K(S_0) \tag{20}$$

On identifie enfin les coefficients manquants de $\widehat{P}^{-1}$ dans (8), et on obtient $\widehat{S}$ et $\widehat{M}$ par :

$$\begin{cases} \widehat{S} = \widehat{P}^{-T} W V^T \\ \widehat{M} = U \widehat{P}^T \end{cases} \tag{21}$$

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 $A = VW$, tout en assurant les mêmes propriétés que la méthode précédente.

4.2 Estimation de l'intensité lumineuse

La remarque formulée dans la partie 3.2 sur la validité de la méthode de Yuille et Snow est bien entendu toujours valable, puisque l'équation (18) n'a de sens que pour une certaine plage de valeurs de S_0. Plutôt que de choisir une valeur suffisamment grande pour éviter les problèmes, et puisqu'une valeur optimale de ce coefficient S_0 semble exister (cf. figure 3), nous nous proposons d'estimer ce coefficient 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 (19).

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

$$\begin{cases} S_0^0 = \max\{\sqrt{{S_x^i}^2 + {S_y^i}^2}, i = 1 \dots m\} \\ b^0 = A^+ K(S_0^0) \end{cases} \tag{22}$$

$K(S_0^0)$ étant défini par (19).

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

$$\min_{(b,S_0)\in\mathbb{R}^3\times\mathbb{R}^+} ||Ab - K(S_0)||^2 \tag{23}$$

L'optimisation est efficace et permet d'obtenir $\widehat{S_0}$ et $\widehat{b}$ en quelques itérations. On obtient ensuite $\widehat{P}^{-1}$ par (8), puis $\widehat{M}$ et $\widehat{S}$ par (21).

4.3 Influence du nombre d'images

Rappelons que les méthodes d'Hayakawa [14] et de Yuille et Snow [26] supposent qu'on dispose d'au moins $m = 6$ images, afin de pouvoir identifier la matrice Q^TQ en résolvant le système (10). Dans notre approche, comme nous nous passons de l'estimation de Q^TQ, nous n'avons pas besoin de tant d'images. Le problème (23) étant un système linéaire de m équations à 4 inconnues, il a 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 est alors mal posé. En effet, à chaque valeur de S_0 correspond une solution $\widehat{b}$ *exacte* du système (18), telle que $\widehat{b} = A^{-1}K(S_0)$. Il existe alors une infinité de solutions au problème (23).

Néanmoins, considérant qu'il n'est plus problématique aujourd'hui d'acquérir et de traiter de grandes quantités d'images (pour nous, typiquement, une douzaine d'images est largement suffisante), nous n'essaierons pas ici de mieux estimer S_0 dans le cas $m = 3$. Nous nous contenterons dans ce cas de l'estimation (22), qui est une solution exacte de (23) pour $m = 3$, le résidu étant alors nul.

5 Une méthode complète de résolution de la stéréophotométrie non calibrée

5.1 Exposé de la méthode

Supposons que nous disposions comme données de $m \geqslant 3$ images et d'un masque Ω. Nous utilisons les deux hypothèses suivantes :

- Le champ de normales estimé doit être intégrable. Cette hypothèse permet de réduire l'ambiguïté linéaire $I = MS = MAA^{-1}S$ à l'ambiguïté de bas-relief $A \equiv G$, où G est définie par (9).
- Les sources lumineuses ont toutes la même intensité S_0. Cette hypothèse permet de résoudre l'ambiguïté de bas-relief, *i.e.* de fixer G.

La méthode que nous proposons consiste à effectuer les étapes suivantes, dans l'ordre :

1. Factoriser I par SVD, afin d'obtenir U, W et V telles que $I = UWV^T$ [14], puis décomposer W en $W = P^TQ$.
2. En assurant l'intégrabilité du champ de normales, estimer les 6 coefficients a_i, $i = 1 \ldots 6$, de P^{-1} par la méthode de Yuille et Snow [26].
3. Faire une première estimation de b et S_0 par (22).
4. Si $m > 3$, résoudre le problème (23) par la méthode de Levenberg-Marquardt pour obtenir $\widehat{b}$.
5. Calculer $\widehat{P}^{-1}$ par (8), puis en déduire $\widehat{P}$ et $\widehat{Q} = \widehat{P}^{-T}W$. L'ambiguïté de bas-relief est ainsi levée.
6. Enfin, calculer $\widehat{M} = U\widehat{P}^T$ et $S = \widehat{Q}V^T$.

5.2 Quelques résultats

Nous utilisons comme données les 10 premières séries d'images de visages de la base de données Yale Dataface B [13]. Une partie de l'un de ces jeux d'images est montrée sur la figure 4. Nous obtenons ainsi un champ de normales et une carte d'albédo pour chaque visage (cf. figure 5). Après intégration, nous obtenons les reliefs correspondants (cf. figure 6).

Pour chacun des 10 visages de cette base de données, nous comparons également notre méthode à deux travaux récents dont les codes sources sont disponibles sur le web. La méthode de Alldrin *et al.* [3] vise à minimiser l'entropie de la distribution d'albédo, celle de Favaro et Papadhimitri [10] extrait une information des maxima locaux diffus. Notre comparaison porte sur le temps de calcul et sur l'erreur angulaire moyenne (EAM) entre le champ de normales reconstruit et le champ de normales obtenu par stéréophotométrie calibrée. L'ensemble des résultats est consigné dans le tableau 1, et quelques reconstructions 3D sont représentées sur la figure 7. Notre méthode s'avère clairement plus efficace en temps de calcul, tout en étant au moins aussi précise. Nous présentons également dans le tableau 2 les valeurs de S_0 estimées et les comparons aux valeurs optimales.

FIGURE 4 – 5 images d'un visage provenant de [13]. Chaque jeu de données utilisé contient 22 images.

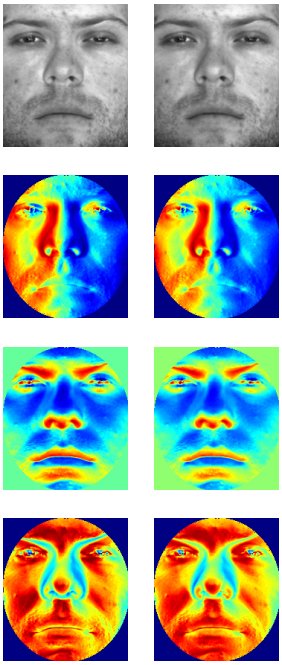

FIGURE 5 – Albédo, première, deuxième et troisième composantes du champ de normales estimé. Première colonne : notre méthode (stéréophotométrie non calibrée). Deuxième colonne : stéréophotométrie calibrée (résultat considéré comme la vérité terrain).

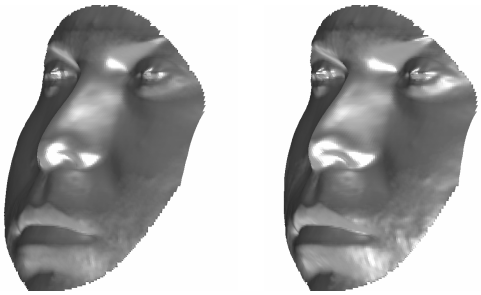

FIGURE 6 – Reconstructions 3D obtenues à partir de 22 images d'un visage provenant de [13]. Première colonne : notre méthode (stéréophotométrie non calibrée). Deuxième colonne : stéréophotométrie calibrée.

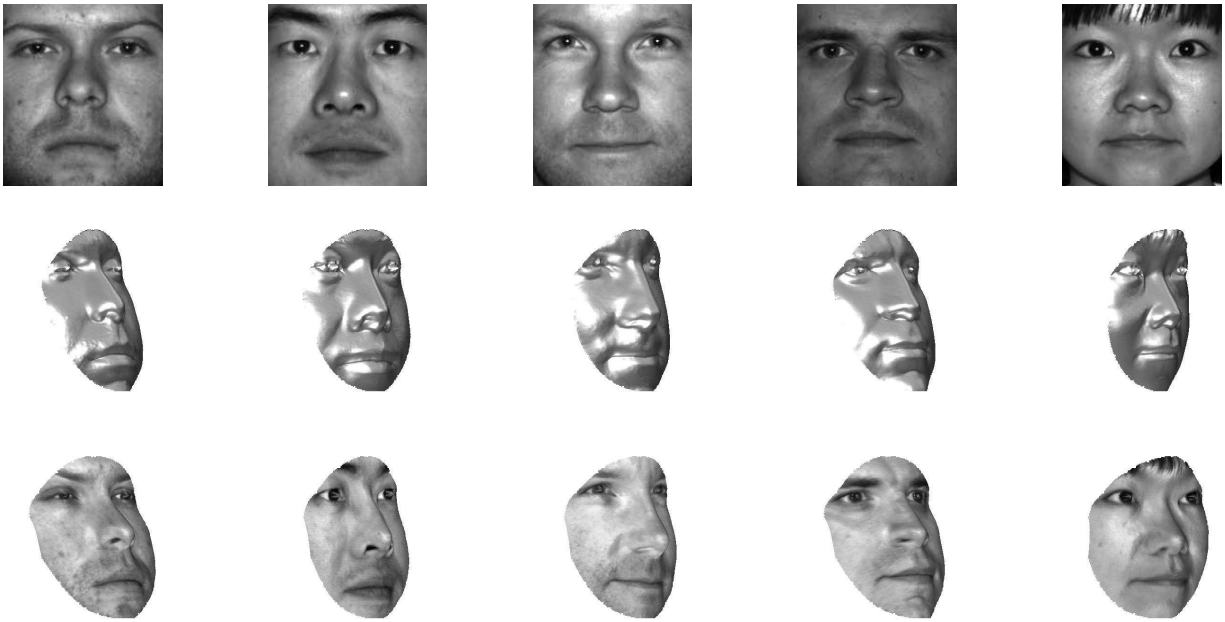

FIGURE 7 – Les 5 premiers visages étudiés (parmi 10), leurs reconstructions 3D vues de trois quart face, et leurs reconstructions 3D après plaquage de l'albédo.

Données	YaleB01_P00	YaleB02_P00	YaleB03_P00	YaleB04_P00	YaleB05_P00
ME	13.90 (14.59s)	23.64 (15.13 s)	13.26 (14.69 s)	12.89 (14.38s)	9.91 (14.04 s)
MD	11.00 (2.15 s)	20.17 (1.83 s)	9.76 (2.65 s)	13.59 (2.11 s)	**6.40** (2.76 s)
EI	**5.64** (0.40 s)	**10.17** (0.44 s)	**6.36** (0.40 s)	**4.45** (0.42 s)	9.03 (0.44 s)

Données	YaleB06_P00	YaleB07_P00	YaleB08_P00	YaleB09_P00	YaleB10_P00
ME	15.83 (14.22 s)	12.63 (14.59 s)	19.61 (14.99 s)	17.50 (14.81 s)	9.32 (13.60 s)
MD	11.46 (1.81 s)	8.60 (1.74 s)	11.49 (1.15 s)	13.10 (0.94 s)	10.21 (3.22 s)
EI	**5.50** (0.40 s)	**5.64** (0.40 s)	**6.96** (0.40 s)	**5.40** (0.42 s)	**5.34** (0.40 s)

TABLE 1 – Comparaison des performances entre notre méthode, en supposant les éclairages de même intensité (EI), la méthode du minimum d'entropie (ME) d'Alldrin *et al.* [3] et la méthode des maxima diffus (MD) de Favaro et Papadhimitri [10]. Nous donnons l'erreur angulaire moyenne sur les normales (en degrés) et le temps CPU.

Données	YaleB01_P00	YaleB02_P00	YaleB03_P00	YaleB04_P00	YaleB05_P00
S_0 optimale	4.52E7	3.59E7	5.76E7	2.38E7	4.98E7
S_0 estimée	4.40E7	3.39E7	5.79E7	2.36E7	4.16E7
Erreur relative	0.027	0.059	0.005	0.009	0.197

Données	YaleB01_P06	YaleB01_P07	YaleB01_P08	YaleB01_P09	YaleB10_P00
S_0 optimale	3.94E7	3.33E7	2.94E7	3.22E7	5.04E7
S_0 estimée	4.04E7	3.32E7	2.51E7	3.10E7	5.10E7
Erreur relative	0.025	0.003	0.171	0.039	0.012

TABLE 2 – Comparaison entre la valeur de S_0 estimée et celle qui minimise l'EAM (correspondant au minimum de la figure 3). Rappelons que la méthode de Yuille et Snow consiste à attribuer la valeur 1 à S_0, qui est pour ces exemples très éloignée de la valeur optimale.

6 Conclusion

Nous avons proposé une nouvelle méthode de résolution du problème de la stéréophotométrie non calibrée, à la fois précise et efficace, en faisant l'hypothèse que l'intensité de la source lumineuse est la même dans toutes les images et en l'estimant. Nous avons également montré expérimentalement, sur des images réelles de visages, la qualité des reconstructions 3D que notre méthode permet d'obtenir, et nous l'avons comparée à deux méthodes proposées récemment.

Cependant, il est 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 varie fortement d'une image à l'autre, il faudra changer d'hypothèse. Il serait d'ailleurs intéressant de pouvoir estimer *a priori* quelle hypothèse est la plus plausible, afin de choisir la « bonne » solution (éclairages uniformes, minimum d'entropie, maxima diffus...).

Par ailleurs, si les sources lumineuses ne sont pas suffisamment éloignées de la scène, il est probable que l'hypothèse d'un éclairage *directionnel* doive être reconsidérée. L'hypothèse d'un éclairage ponctuel permettrait peut-être alors de résoudre le problème de la stéréophotométrie non calibrée.

Mais la vraie limite de notre méthode vient de l'hypothèse d'un matériau lambertien. Il serait intéressant 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 taches spéculaires. Il faudrait soit rendre artificiellement lambertiennes les données pour appliquer la méthode proposée ici, soit proposer une méthode de stéréophotométrie non-calibrée adaptée à des modèles plus complexes, soit enfin 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 les normales en chaque pixel. La première étape pourrait être réalisée par exemple en utilisant la méthode décrite dans [8], et on se retrouverait pour la seconde dans le cadre de la stéréophotométrie calibrée avec un modèle non lambertien.

Références

[1] A. Abrams, C. Hawley, et R. Pless. Heliometric Stereo: Shape from Sun Position. In *ECCV 2012, LNCS 7573*, pages 357–370.

[2] J. Ackermann, F. Langguth, S. Fuhrmann, et M. Goesele. Photometric Stereo for Outdoor Webcams. In *CVPR 2012*, pages 262–269.

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

[4] P. N. Belhumeur, D. J. Kriegman, et A. L. Yuille. The Bas-Relief Ambiguity. *IJCV*, 35(1):33–44, 1999.

[5] M. K. Chandraker, F. Kahl, et D. J. Kriegman. Reflections on the Generalized Bas-Relief Ambiguity. In *CVPR 2005 (vol. I)*, pages 788–795.

[6] J. Chang, K. Lee, et S. Lee. Multiview normal field integration using level set methods. In *CVPR 2007*, pages 1–8.

[7] O. Drbohlav et M. Chantler. Can Two Specular Pixels Calibrate Photometric Stereo ? In *ICCV 2005 (vol. II)*, pages 1850–1857.

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

[9] J.-D. Durou, J.-F. Aujol, et F. Courteille. Integration of a Normal Field in the Presence of Discontinuities. In *EMMCVPR 2009, LNCS 5681*, pages 261–273.

[10] P. Favaro et T. Papadhimitri. A Closed-Form Solution to Uncalibrated Photometric Stereo via Diffuse Maxima. In *CVPR 2012*, pages 821–828.

[11] R. T. Frankot et R. Chellappa. A Method for Enforcing Integrability in Shape from Shading Algorithms. *PAMI*, 10(4):439–451, 1988.

[12] A. S. Georghiades. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *ICCV 2003 (vol. II)*, pages 816–823.

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

[14] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *JOSA A*, 11(11):3079–3089, 1994.

[15] C. Hernández, G. Vogiatzis, G. J. Brostow, B. Stenger, et R. Cipolla. Multiview Photometric Stereo. *PAMI*, 30(3):548–554, 2008.

[16] A. Hertzmann et S. M. Seitz. Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. *PAMI*, 27(8):1254–1264, 2005.

[17] B. K. P. Horn. Obtaining Shape from Shading Information. In *Shape from Shading*, pages 123–171. MIT Press, 1989.

[18] B. K. P. Horn. Height and Gradient from Shading. *IJCV*, 5(1):37–75, 1990.

[19] R. Onn et A. M. Bruckstein. Integrability Disambiguates Surface Recovery in Two-Image Photometric Stereo. *IJCV*, 5(1):105–113, 1990.

[20] B. Shi, Y. Matsushita, Y. Wei, C. Xu, et P. Tan. Self-calibrating Photometric Stereo. In *CVPR 2010*.

[21] K. Sunkavalli, T. Zickler, et H. Pfister. Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. In *ECCV 2010, LNCS 6312*, pages 251–264.

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

[23] R. J. Woodham. Photometric Method for Determining Surface Orientation from Multiple Images. *Opt. Engin.*, 19(1):139–144, 1980.

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

[25] A. L. Yuille. Mathematical Results on SVD for Lighting and Warps. In *Technical Report. Smith-Kettlewell Eye Research Institute, San Francisco, California*, 1996.

[26] A. L. Yuille et D. Snow. Shape and Albedo from Multiple Images using Integrability. In *CVPR 1997*, pages 158–164.