

# Mise à plat de documents par *shape from shading* perspectif

## Setting flat of documents by perspective shape from shading

Frédéric Courteille

Alain Crouzil

Jean-Denis Durou

Pierre Gurdjos

Institut de Recherche en Informatique de Toulouse

Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4

### Résumé

*Cet article décrit une nouvelle modélisation du shape from shading, prenant en compte la perspective, et propose une nouvelle méthode de résolution pour l'équation qui en découle. Une application est présentée, qui consiste à corriger les défauts de photographies de documents gauches.*

### Mots Clef

*shape from shading, documents, projection perspective.*

### Abstract

*This paper describes a new modeling of the shape from shading problem taking perspective into account, and proposes a method of resolution for the new equation. An application is proposed, which consists in correcting the defects of photographs of skew i.e., non flat, documents.*

### Keywords

*shape from shading, documents, perspective projection.*

## 1 Introduction

Le *shape from shading* (SFS) consiste à reconstruire le relief d'une scène à partir d'une seule image en niveaux de gris, grâce à l'analyse des ombres propres. L'engouement qu'a connu le SFS par le passé semble être un peu retombé, probablement à cause de résultats relativement décevants sur images réelles [10, 3]. Néanmoins, plusieurs travaux récents [8, 9, 2] ont tenté de modéliser le SFS de manière plus réaliste, ce qui est indispensable pour envisager des applications. Dans cet article, nous souhaitons corriger les défauts de photographies de documents gauches. Le SFS doit nous permettre de calculer le relief des documents au moment de la prise de vue. Dans le paragraphe 2, nous rappelons les hypothèses classiques du SFS et la modélisation qui s'ensuit. Dans le paragraphe 3, l'approximation de la projection perspective par une projection orthogonale,

qui semble irréaliste vis-à-vis de l'application envisagée, est remise en question, ce qui nous amène à une modélisation un peu plus réaliste du SFS. Une méthode de résolution de l'équation obtenue est présentée dans le paragraphe 4. Enfin, l'application visée est détaillée, résultats à l'appui, dans le paragraphe 5.

## 2 Équation de luminance

Le problème du SFS étant relativement complexe, il est d'usage de le modéliser en faisant des hypothèses simplificatrices : [H1] la surface ne comporte ni arête ni partie cachée, donc la normale unitaire sortante  $\vec{n}(P)$  au point  $P$  est définie sans ambiguïté ; [H2] les réflexions secondaires sont négligeables ; [H3] l'image est nette et les différentes aberrations de l'objectif sont négligeables ; [H4] l'angle de champ  $\alpha$  (cf. figure 1) est négligeable devant 1 ; [H5] le récepteur photosensible est linéaire ; [H6] le matériau est lambertien ; [H7] la source lumineuse émet un faisceau parallèle et uniforme qui peut être décrit, en direction et en densité, par un vecteur  $\vec{S}$ . En faisant les hypothèses H1 à H5, le SFS est modélisé par « l'équation de luminance » :

$$I(P') = \frac{r t \pi D^2}{4 d'^2} L(P), \quad (1)$$

où  $L(P)$  est la luminance du point objet  $P$  dans la direction du centre optique  $C$ ,  $I(P')$  le niveau de gris au point image  $P'$  conjugué de  $P$ ,  $r$  le facteur de réponse du récepteur,  $t$  le facteur de transmission des lentilles,  $D$  le diamètre de la pupille d'entrée et  $d'$  la distance du centre optique au plan image  $\Pi'$ . Les hypothèses H6 et H7 permettent d'exprimer la luminance sous la forme  $L(P) = -[\rho/\pi] \vec{S} \cdot \vec{n}(P)$ , où  $\rho$  désigne l'albédo de la surface. En remplaçant dans (1), il vient :

$$I(P') = -\frac{r t \rho D^2}{4 d'^2} \vec{S} \cdot \vec{n}(P), \quad (2)$$

équation dont la résolution n'est possible que si l'on connaît la relation entre  $P$  et  $P'$ , c'est-à-dire si l'on précise le modèle de projection. Nous allons envisager

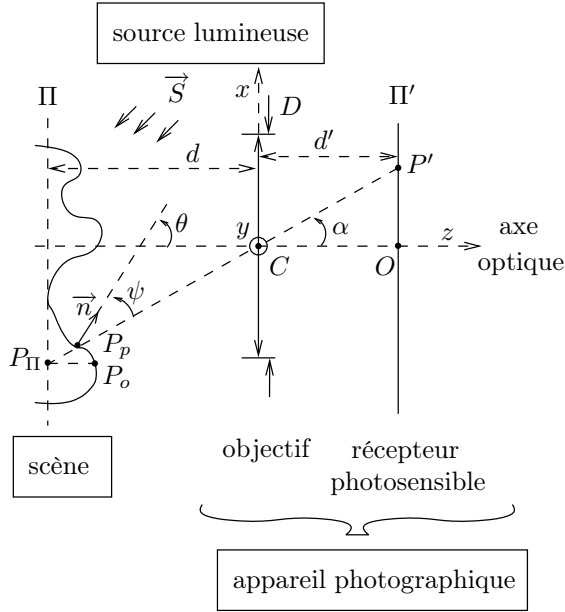


FIG. 1 – Conjugaison entre objet et image.

la projection orthogonale, puis la projection perspective. Dorénavant, nous supposons (hypothèse H8) que  $\vec{S} = (0,0, -S)$ . La généralisation à un  $\vec{S}$  quelconque est aisée.

### 3 Équation de l'eikonale

La projection orthogonale consiste à associer à  $P'$  le point  $P_o$  dont le projeté orthogonal  $P_{\Pi}$  sur le « plan de mise au point » de la scène  $\Pi$  a comme image  $P'$  par projection centrale, de centre  $C$ , sur le « plan image »  $\Pi'$ . Munissons l'espace tridimensionnel d'un repère orthonormé  $(Cxyz)$ , tel que l'axe  $Cz$  coïncide avec l'axe optique et le plan  $Cxy$  soit parallèle à  $\Pi'$ . La donnée du problème est la fonction  $i$  telle que  $i(x,y) = I(P')$ ,  $P'$  étant le point image de coordonnées  $(x,y,d')$ . L'inconnue du problème est la fonction  $u_o$  telle que  $P_o$  ait comme cote  $u_o(x,y)$ . Comme les coordonnées de  $P_o$  sont  $(x/g, y/g, u_o(x,y))$ , où  $g = -d'/d$  est le grandissement transversal,  $\vec{n}(P_o)$  peut s'exprimer en fonction des coordonnées de  $\vec{\nabla}u_o(x,y)$ :

$$\vec{n}(P_o) = \frac{(-g \partial_x u_o(x,y), -g \partial_y u_o(x,y), 1)}{\sqrt{g^2 \|\vec{\nabla}u_o(x,y)\|^2 + 1}}. \quad (3)$$

En posant  $i_{\max} = r t \rho D^2 S / (4 d'^2)$ , (2) se réécrit alors :

$$g^2 \|\vec{\nabla}u_o(x,y)\|^2 = \frac{i_{\max}^2}{i(x,y)^2} - 1. \quad (4)$$

Cette équation aux dérivées partielles, non linéaire du premier ordre, est appelée « équation de l'eikonale ». C'est l'équation la plus souvent rencontrée en SFS. Cependant, l'hypothèse de la projection orthogonale

n'est pas réaliste vis-à-vis de l'application envisagée, à savoir la correction de photographies de documents gauches, puisque les lignes de texte ne sont plus parallèles entre elles (cf. figure 2(b)).

La prise en compte de la projection perspective en SFS n'avait fait l'objet, jusqu'à très récemment, que de très peu de travaux [7, 6, 4], et aucun de ces travaux n'avait proposé de nouvelle modélisation du SFS. Récemment, trois groupes d'auteurs ont établi simultanément cette nouvelle modélisation [8, 9, 2]. Le conjugué de  $P'$  est maintenant le point  $P_p$  ayant  $P'$  comme image par projection centrale, de centre  $C$ , sur  $\Pi'$ . Le point  $P_p$  a comme coordonnées  $(x u_p(x,y)/d', y u_p(x,y)/d', u_p(x,y))$ , la fonction  $u_p$  étant la nouvelle inconnue du problème. Après quelques calculs, on peut exprimer le vecteur  $\vec{n}(P_p)$  en fonction de  $u_p(x,y)$  et des coordonnées de  $\vec{\nabla}u_p(x,y)$ :

$$\vec{n}(P_p) = \frac{(-\hat{g}(x,y) \partial_x u_p(x,y), -\hat{g}(x,y) \partial_y u_p(x,y), 1)}{\sqrt{\hat{g}(x,y)^2 \|\vec{\nabla}u_p(x,y)\|^2 + 1}}, \quad (5)$$

où  $\hat{g}(x,y) = d'/\hat{u}_p(x,y)$ , avec  $\hat{u}_p(x,y) = u_p(x,y) + x \partial_x u_p(x,y) + y \partial_y u_p(x,y)$ . L'équation (2) se réécrit alors :

$$\frac{d'^2}{\hat{u}_p(x,y)^2} \|\vec{\nabla}u_p(x,y)\|^2 = \frac{i_{\max}^2}{i(x,y)^2} - 1. \quad (6)$$

Cette équation, que nous appelons « équation de l'eikonale perspective », est très similaire à (4), mais elle fait intervenir l'inconnue  $u_p$  elle-même, et non plus seulement son gradient. On peut donc craindre que sa résolution ne soit plus compliquée que celle de (4). Dans [8], (6) est résolue par le biais de la recherche de ses solutions de viscosité. Nous abordons la résolution de (6) sous un angle nettement différent.

### 4 Pseudo-équation de l'eikonale

Comme de nombreux auteurs [5], nous rajoutons comme inconnue la normale  $\vec{n}(P)$ . Ce vecteur unitaire ayant deux degrés de liberté, cela est équivalent à introduire deux fonctions inconnues, par exemple  $p$  et  $q$  telles que :

$$\vec{n}(P) = \frac{(-p(x,y), -q(x,y), 1)}{\sqrt{p(x,y)^2 + q(x,y)^2 + 1}}. \quad (7)$$

L'équation (2) se réécrit alors :

$$p(x,y)^2 + q(x,y)^2 = \frac{i_{\max}^2}{i(x,y)^2} - 1. \quad (8)$$

Cette équation, qui ressemble à (4) et à (6), n'est pas une équation aux dérivées partielles. Pour cette raison, nous proposons de l'appeler « pseudo-équation de l'eikonale ». Sa résolution est évidemment un problème mal posé, puisqu'en chaque  $(x,y)$  il y a une équation pour deux inconnues, mais différentes stratégies

peuvent être envisagées pour le rendre bien posé [6]. Dans le cadre de notre application, des connaissances *a priori* sur la surface de la scène contraignent suffisamment le problème pour nous permettre de calculer  $p$  et  $q$  sans ambiguïté. La résolution du problème consiste ensuite à calculer le relief à partir de  $p$  et  $q$ , par intégration. Sous l'hypothèse de la projection orthogonale, on déduit de (3) et (7) que  $\overrightarrow{\nabla u_o}(x,y) = (p(x,y)/g, q(x,y)/g)$ . Sous l'hypothèse de la projection perspective, on déduit de (5) et (7) que  $\overrightarrow{\nabla u_p}(x,y) = (p(x,y)/\hat{g}(x,y), q(x,y)/\hat{g}(x,y))$ , mais cette relation ne peut pas être utilisée telle quelle, puisqu'elle équivaut au système linéaire suivant pour les composantes de  $\overrightarrow{\nabla u_p}(x,y)$  (les dépendances en  $(x,y)$  ont été omises) :

$$\begin{cases} [xp - d'] \partial_x u_p + yp \partial_y u_p = -p u_p, \\ xq \partial_x u_p + [yq - d'] \partial_y u_p = -q u_p, \end{cases} \quad (9)$$

dont le déterminant vaut  $d'(d' - xp - yq)$ . Or, en utilisant les notations de la figure 1, le produit scalaire  $\overrightarrow{n}(P_p) \cdot \overrightarrow{CP}'$  vaut : d'une part,  $(d' - xp - yq) \cos \theta$  en utilisant (7) et en remarquant que la troisième coordonnée de  $\overrightarrow{n}(P_p)$  vaut  $\cos \theta$  ; d'autre part,  $d' \cos \psi / \cos \alpha$ . Donc  $d' - xp - yq = d' \cos \psi / (\cos \alpha \cos \theta)$ , qui ne s'annule que si  $\psi = \pi/2$ . Sinon, c'est-à-dire si  $P'$  n'est pas sur une « silhouette », alors la solution de (9) est :

$$\overrightarrow{\nabla u_p}(x,y) = \frac{u_p(x,y)}{d' - xp(x,y) - yq(x,y)} \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}. \quad (10)$$

Les deux types de projection sont donc relativement semblables vis-à-vis de l'intégration et toute méthode d'intégration utilisable pour la projection orthogonale est généralisable au cas perspectif. Néanmoins, à partir d'un couple  $(p,q)$ , les deux intégrations fourniront bien sûr deux reliefs différents. Par ailleurs, l'intégration orthogonale fournira le relief à une constante près, alors que l'intégration perspective fournira le relief à un facteur d'échelle près. Cette différence, qui transparaît dans les équations (4) et (6), a déjà été remarquée dans [8, 9].

## 5 Application à la correction de photographies de documents gauches

Le problème de la correction de photographies de livres n'est pas nouveau. Dans [1], un système de numérisation spécifique est élaboré, mais la modélisation du SFS et la méthode de résolution utilisées sont classiques. Il était nécessaire de prendre en compte les effets de la perspective, très visibles sur les photographies prises à faible distance.

Sur la figure 2, nous pouvons voir un exemple de résultat  $I_2^1$  fourni par notre algorithme à partir d'une photographie  $I_1^1$ . Nous rappelons au lecteur que seules les

zones non encrées de  $I_1^1$  sont analysées ; ainsi, le quasi-parallélisme des lignes de texte dans  $I_2^1$  n'est pas utilisé comme connaissance *a priori* sur le document analysé. De la même façon, le cadre au bas de l'image  $I_2^1$  a retrouvé sa forme rectangulaire originelle. Notre technique de mise à plat semble donc plutôt satisfaisante. Du point de vue de la correction photométrique, il apparaît que certains caractères de l'image corrigée, en particulier ceux qui sont situés au voisinage de la zone de pente nulle, sont moins « contrastés » que d'autres. Ceci provient probablement du fait que les hypothèses du SFS ne sont pas parfaitement vérifiées, et que la correction photométrique (qui est censée gommer les écarts aux hypothèses  $H_4, H_5, H_6$  et  $H_7$ ) est moins performante en ces points. Cependant, ce défaut est relativement léger et il existe de nombreux opérateurs de traitement d'images qui seraient capables d'uniformiser le contraste de l'image  $I_2^1$ .

La figure 3 montre trois imagerie extraites des images  $I_0^1, I_1^1$  et  $I_2^1$ , qui correspondent à une même zone de texte. À l'œil nu, il semble que la correction soit de très bonne qualité. Pour valider plus rigoureusement

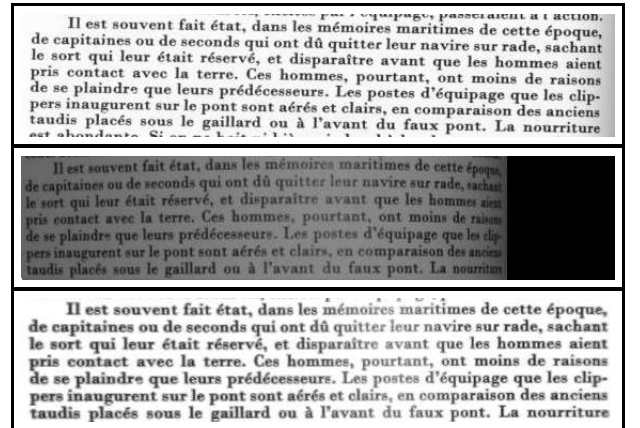


FIG. 3 – Extraits des images  $I_0^1, I_1^1$  et  $I_2^1$ .

	$I_0^1$	$I_1^1$	$I_2^1$
taux de succès	92 %	2 %	61 %
taux d'échec	6 %	70 %	24 %
taux d'abstention	2 %	28 %	15 %

TAB. 1 – Résultats de l'OCR.

notre résultat, nous avons effectué des tests de reconnaissance de caractères. Pour cela, nous avons utilisé l'OCR libre *gocr* sur les trois imagerie de la figure 3. Les résultats, reportés dans le tableau 1, montrent que le pourcentage de caractères reconnus avec succès a été amélioré d'un facteur 30 environ (61 % au lieu de 2 %), que le pourcentage d'échecs a été réduit d'un facteur 3, et que le taux d'abstention a été, quant à lui, divisé d'un facteur 2, ce qui nous per-

met de conclure que la rectification effectuée permet à l'OCR de décider plus souvent et de manière plus fiable.

Les autres exemples vont illustrer clairement le fait que notre algorithme n'utilise que les zones non encrées du document, indépendamment de sa structure géométrique, de sa mise en forme ou des éventuelles figures qu'il comporte. L'exemple de la figure 4 représente les images du calligramme « La colombe poignardée et le jet d'eau » de Guillaume Apollinaire : (a) scannée, (b) photographiée et (c) après correction. On peut remarquer que la photographie comporte des défauts géométrique et photométrique. Notamment, le mot « Jardins » (premier mot de la dernière ligne du texte) est complètement déformé et presque illisible, alors que le même mot, une fois mis à plat, est tout à fait lisible, tant au niveau de sa géométrie que de son contraste par rapport au fond blanc. Cependant, le résultat n'est pas parfait, puisque si l'on compare l'image du document scannée à plat et l'image corrigée obtenue par notre méthode, on peut s'apercevoir que, d'une part, le texte est légèrement déformé (différentes occurrences d'un même caractère n'ayant pas la même taille), et que d'autre part, le défaut photométrique déjà présent sur l'exemple 1 est encore visible.

La figure 5 est un autre exemple qui, cette fois-ci, est un document en couleurs ne comportant que des dessins, et dont le papier est légèrement glacé. La mise à plat d'un tel document serait difficilement réalisable par des méthodes n'utilisant pas le SFS. On peut remarquer que, même si les hypothèses du SFS ne sont pas complètement satisfaites (la scène n'étant pas lambertienne), le résultat obtenu est tout à fait satisfaisant. Si l'on observe de plus près les deux personnages situés sur les bords gauche et droit de l'image, la mise à plat est de très bonne qualité : les trois figurines sont très proches des originaux, tant au niveau de la forme qu'au niveau de la couleur. Cependant, le résultat n'est pas parfait puisque, comme on peut le voir sur la figure 5(c), la tête du personnage central a été dégradée. Le papier n'étant pas lambertien, le flash a manifestement produit un artéfact à cet endroit.

## 6 Conclusion

Nous avons proposé dans cet article une nouvelle modélisation du SFS qui prend en compte la perspective, ainsi qu'une nouvelle méthode de résolution adaptée à l'équation qui en découle. Nous avons ensuite utilisé cette méthode de résolution pour la correction de photographies de documents gauches. Les résultats obtenus sur images réelles montrent qu'il est tout à fait possible d'envisager des applications au SFS.

## Remerciements

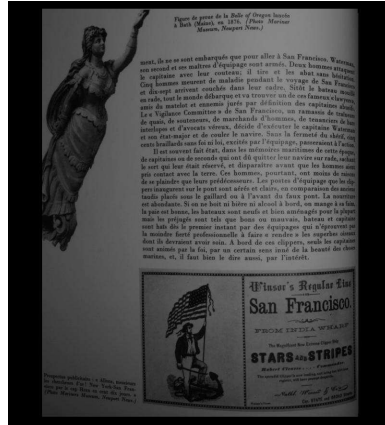
Ce travail est soutenu par EGIDE, dans le cadre du projet Galilée « PLATONOV » (Photocopies de Livres Anciens par Techniques d'Optimisation Numérique et Solutions de Viscosité).

## Références

- [1] S. I. CHO, H. SAITO et S. OZAWA. « A Divide-and-conquer Strategy in Shape from Shading Problem ». Dans *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 413–419, San Juan, Porto Rico, juin 1997.
- [2] F. COURTEILLE, A. CROUZIL, J.-D. DUROU et P. GURDJOS. « Towards shape from shading under realistic photographic conditions ». Dans *Proceedings of the 17<sup>th</sup> International Conference on Pattern Recognition (volume II)*, pages 277–280, Cambridge, Royaume-Uni, août 2004.
- [3] J.-D. DUROU, M. FALCONE et M. SAGONA. « A Survey of Numerical Methods for Shape from Shading ». Rapport de Recherche 2004-2-R, Institut de Recherche en Informatique de Toulouse, Toulouse, France, janvier 2004.
- [4] J. K. HASEGAWA et C. L. TOZZI. « Shape from Shading with Perspective Projection and Camera Calibration ». *Computers and Graphics*, 20(3):351–364, mai-juin 1996.
- [5] B. K. P. HORN et M. J. BROOKS. « The Variational Approach to Shape From Shading ». *Computer Vision, Graphics, and Image Processing*, 33(2):174–208, février 1986.
- [6] K. M. LEE et C.-C. J. KUO. « Shape from Shading with Perspective Projection ». *Computer Vision, Graphics, and Image Processing: Image Understanding*, 59(2):202–212, mars 1994.
- [7] M. A. PENNA. « A Shape from Shading Analysis for a Single Perspective Image of a Polyhedron ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):545–554, juin 1989.
- [8] E. PRADOS et O. FAUGERAS. « “Perspective Shape from Shading” and Viscosity Solutions ». Dans *Proceedings of the 9<sup>th</sup> IEEE International Conference on Computer Vision (volume II)*, pages 826–831, Nice, France, octobre 2003.
- [9] A. TANKUS, N. SOCHEN et Y. YESHURUN. « A New Perspective [on] Shape-from-Shading ». Dans *Proceedings of the 9<sup>th</sup> IEEE International Conference on Computer Vision (volume II)*, pages 862–869, Nice, France, octobre 2003.
- [10] R. ZHANG, P.-S. TSAI, J. E. CRYER et M. SHAH. « Shape from Shading: A Survey ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):690–706, août 1999.



(a)

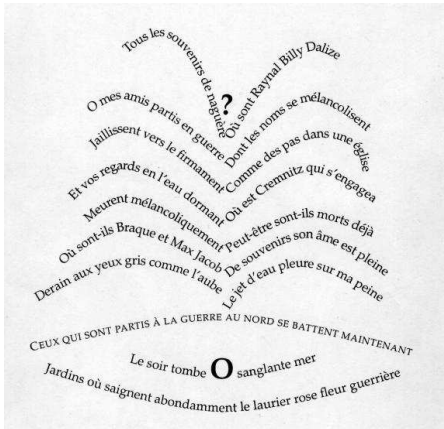


(b)

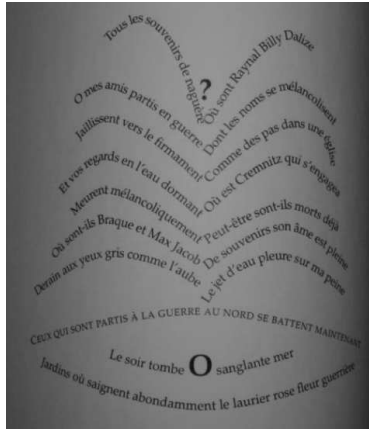


(c)

FIG. 2 – Exemple 1 : (a) image scannée  $I_0^1$ , (b) photographie  $I_1^1$  et (c) image corrigée  $I_2^1$ .



(a)



(b)



(c)

FIG. 4 – Exemple 2 : (a) image scannée  $I_0^2$ , (b) photographie  $I_1^2$  et (c) image corrigée  $I_2^2$ .



(a)



(b)



(c)

FIG. 5 – Exemple 3 : (a) image scannée  $I_0^3$ , (b) photographie  $I_1^3$  et (c) image corrigée  $I_2^3$ .