Comment transformer un appareil photographique en scanner à plat

Transform a digital camera into a flatbed scanner

Frédéric Courteille Ivan Pétillot Jean-Denis Durou Pierre Gurdjos
IRIT, Toulouse, France
Mél : {courteille,petillot,durou,gurdjos}@irit.fr

Résumé

Du fait que les appareils photographiques numériques deviennent de plus en plus performants et de moins en moins coûteux, nous proposons de les doter d'une nouvelle fonctionnalité, qui consiste à pouvoir les transformer en scanner à plat. Dans cet article, nous montrons que deux techniques de reconstruction tridimensionnelle (shape from shading and shape from contour) peuvent permettre de corriger tous les défauts, géométriques et photométriques, qui apparaissent dans les photographies de documents gauches.

Mots Clef

shape from shading, *shape from contour*, reconstruction 3D, documents, numérisation.

Abstract

As digital cameras are more affordable and powerful, we propose to extend their use to a new application, which consists in transforming a digital camera into a flatbed scanner. In this paper, we show that two different techniques of shape from X (shape from shading and shape from contour) can correct all the geometric and photometric distortions that appear in the digital images of warped documents.

Keywords

shape from shading, shape from contour, 3D reconstruction, documents, digitization.

1 Introduction

La numérisation de documents connaît actuellement une vogue considérable, probablement due aux possibilités offertes par Internet en termes de consultation à distance. Le procédé classique de numérisation requiert l'utilisation d'un scanner à plat, mais présente deux défauts lorsque le document à numériser est un livre. Primo, la manipulation répétitive consistant à retourner le livre, puis à exercer une pression afin de plaquer les pages contre la vitre du scanner, peut s'avérer relativement fastidieuse. Secundo, un certain nombre de défauts apparaissent sur l'image numérique, en particulier près de la reliure : caractères flous ou déformés, effets de parallaxe, niveau de gris non uniforme pour les zones non encrées. Pour éviter ces deux défauts, un grand nombre de systèmes spécifiques ont été développés, comme les scanners à « livre ouvert », mais il ne s'agit pas de produits « grand public ».

Une alternative consiste à simuler la mise à plat des documents gauches à l'aide de techniques de vision par ordinateur, en remplaçant le scanner, dont le système optique est difficile à modéliser, par un appareil photographique numérique. C'est dans ce cadre que se place notre travail. Afin de simuler une image corrigée qui soit conforme aux modèles, géométrique et photométrique, de l'imagerie photographique, il est nécessaire de calculer le relief du document au moment de la prise de vue à partir de la photographie initiale (*cf.* figure 1), puis d'utiliser ce relief pour simuler l'image du document mis à plat.

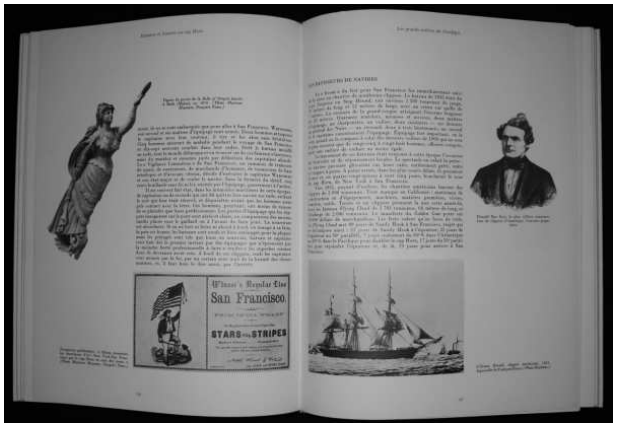


FIG. 1 – *Photographie d'un livre ouvert utilisée comme donnée de notre chaîne de traitement.*

L'article est organisé comme suit : dans les paragraphes 2 et 3, nous détaillons les corrections (géométrique et photométrique) souhaitées ; dans le paragraphe 4, deux approches différentes permettant de calculer le relief sont comparées ; nous tirons un certain nombre de conclusions sur ce travail dans le paragraphe 5.

2 Correction géométrique

2.1 État de l'art

Méthodes 2D-2D Une correction géométrique très simple des images de documents gauches a été proposée dans [15] : l'orientation des caractères situés près de la reliure est évaluée, ce qui permet ensuite de les redresser. Malheureusement, les résultats sont plutôt décevants. Dans [12], une déformation 2D-2D judicieuse est proposée, qui utilise le fait que la page est rectangulaire : les résultats sont spectaculaires, mais une mire (*paper checkerboard pattern*) doit être glissée sous la page afin de calculer son relief. Toutes les autres méthodes requièrent la connaissance explicite du relief du document, comme cela est montré sur la figure 2. Pour ce faire, plusieurs techniques de reconstruction tridimensionnelle (R3D) ont été utilisées. Nous allons détailler ces méthodes dans le prochain paragraphe.

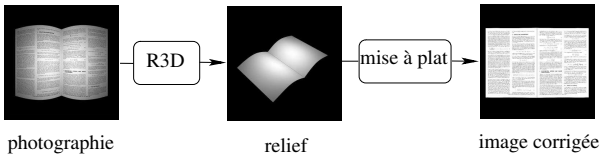


FIG. 2 – *Correction géométrique : principe des méthodes 2D-3D-2D.*

Méthodes 2D-3D-2D Afin de calculer le relief d'un document, certains auteurs ont besoin de plusieurs images. Dans [14], deux photographies sont prises sous deux angles de vue différents, tandis que dans [3], les deux photographies sont prises avec deux éclairages différents. Dans les deux cas, un calibrage précis (géométrique pour [14], photométrique pour [3]) est nécessaire. Dans [1], la scène est éclairée avec une lumière structurée. Plusieurs photographies sont alors requises pour obtenir le relief. Cette dernière technique, originellement proposée dans [5], semble plutôt réservée à la numérisation du patrimoine culturel, vu le coût relativement élevé du matériel mis en jeu.
D'autres auteurs se contentent d'une seule image. Dans [2], l'analyse de la déformation des lignes de texte permet d'obtenir des résultats assez satisfaisants. Une autre technique, proposée dans [8], analyse les contours de la page. Ces deux techniques sont analogues, dans la mesure où elles supposent toutes deux que les lignes détectées (les lignes de texte dans [2], les bords supérieur et inférieur de la page dans [8]) doivent être orthogonales à la reliure, pour un document mis à plat. Par ailleurs, le *shape from shading* a été utilisé, d'abord sur des images de scanner à plat [13], puis plus récemment sur des photographies [3, 17, 4]. Signalons enfin qu'à notre connaissance, personne n'a utilisé la relation entre orientation des caractères et relief, alors que l'orientation des caractères est évaluée à d'autres fins dans [15], comme nous l'avons déjà dit.
Si un assez grand nombre de techniques de reconstruction tridimensionnelle ont été utilisées, la deuxième étape des méthodes 2D-3D-2D, c'est-à-dire la simulation de la mise à plat, est plus ou moins toujours la même. En fait, si le document peut être considéré comme une surface « réglée développable », il est aisé de le « dégauchir » [17, 14, 7]. Signalons néanmoins une toute autre technique, proposée dans [1], qui utilise une analogie avec la physique, et qui semble particulièrement judicieuse pour dégauchir des documents non réglés développables, comme par exemple des manuscrits anciens.
Enfin, la simulation de la mise à plat déplace les pixels de telle sorte qu'ils ne forment plus une grille régulière. Par conséquent, une interpolation est nécessaire pour fournir l'image corrigée. Quelques auteurs [3] ont développé des méthodes d'interpolation spécifiques, tandis que les autres utilisent généralement l'algorithme de Smythe [10].

2.2 Notre contribution

Dans cet article, nous nous intéressons au *shape from contour* (SFC) et au *shape from shading* (SFS), qui sont deux techniques de reconstruction tridimensionnelle utilisant une seule image.

Montage optique utilisé Nous utilisons un appareil photographique numérique à visée réflex (Canon EOS 300D) fixé sous un pied (*cf.* figure 3) et commandé par un ordinateur. La scène est éclairée par un flash annulaire (Sigma EM-140 DG). Afin de limiter le flou de mise au point, le diaphragme de l'objectif est ouvert le moins possible.

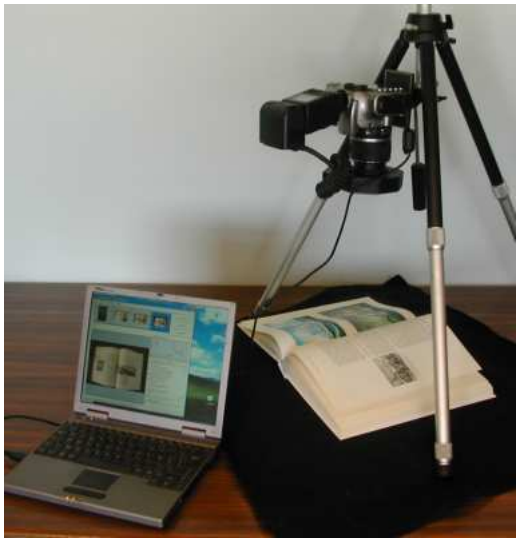


FIG. 3 – *Montage optique utilisé.*

Modélisations géométrique et photométrique Comme nous pouvons le voir sur la figure 4, le modèle de projection est le modèle sténopé, de centre de projection le centre optique C. Un repère orthonormé tridimensionnel $(C\mathbf{xyz})$ est attaché à la scène, coïncidant avec celui de l'appareil photographique. Nous notons $(A\mathbf{ij})$ le repère bidimensionnel utilisé dans l'image et introduisons la « matrice de ca-

librage » (*cf.* [6], page 156) :

$$\mathbf{K} = \begin{pmatrix} f & s & i_o \\ 0 & \tau f & j_o \\ 0 & 0 & 1 \end{pmatrix}, \qquad (1)$$

où f représente la distance focale (exprimée dans l'unité de l'axe des abscisses) et (i_o,j_o) les coordonnées du point principal ; τ et s représentent, respectivement, le « rapport d'échelle » et le « facteur d'obliquité » des pixels. Pour un pixel Q de coordonnées (i,j) dans le repère $(A\mathbf{ij})$, le point objet conjugué est un point P de coordonnées (x,y,z) dans le repère $(C\mathbf{xyz})$. L'inconnue du problème de reconstruction tridimensionnelle est la fonction u telle que la surface du document soit représentée par l'équation :

$$z = u(i,j). \qquad (2)$$

Sur la figure 4, le plan Π désigne le plan fronto-parallèle sur lequel la mise à plat sera simulée. Ce plan peut être

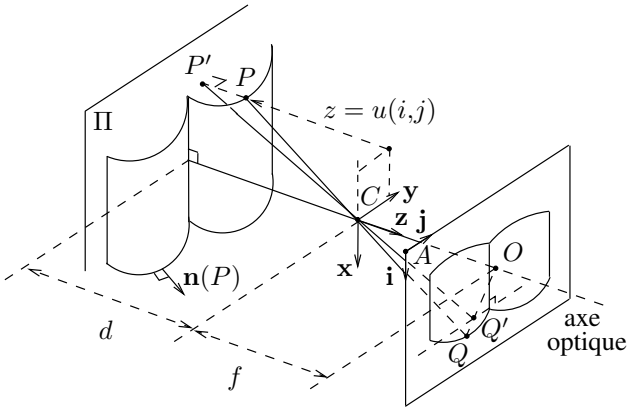


FIG. 4 – *Projection perspective et projection perspective faible.*

choisi à une distance d quelconque de C mais, comme nous supposons, dans le cadre de cet article, que la reliure du livre est orthogonale à l'axe optique (prise de vue « fronto-parallèle », hypothèse qui sera notée H1 par la suite), il est commode de choisir pour Π le plan fronto-parallèle contenant la reliure.

Shape from contour D'une part, l'équation de la projection perspective, qui lie les coordonnées $\mathbf{P} = (x\,y\,z)^T$ d'un point objet P aux coordonnées $\mathbf{Q} = (i\,j)^T$ de son point image conjugué Q, s'écrit (*cf.* [6], page 155) :

$$z \begin{pmatrix} \mathbf{Q} \\ 1 \end{pmatrix} = \mathbf{K} \left(\mathbf{I_3} \mid \mathbf{O_3} \right) \begin{pmatrix} \mathbf{P} \\ 1 \end{pmatrix} = \mathbf{KP}, \qquad (3)$$

où $\mathbf{I_3}$ représente la matrice unité d'ordre 3 et $\mathbf{O_3}$ le vecteur nul de dimension 3. D'autre part, l'équation de la projection perspective faible, qui lie la projection orthogonale P' de P sur Π à son point conjugué Q' dans l'image, de coordonnées $\mathbf{Q'} = (i'\,j')^T$, s'écrit (*cf.* [6], page 171) :

$$d \begin{pmatrix} \mathbf{Q'} \\ 1 \end{pmatrix} = \mathbf{K} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & d \end{pmatrix} \begin{pmatrix} \mathbf{P} \\ 1 \end{pmatrix} = \mathbf{K} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{P} + d\,\mathbf{K} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \qquad (4)$$

En déduisant $\mathbf{P}$ de (3), et en reportant dans (4), nous obtenons :

$$d \begin{pmatrix} \mathbf{Q'} \\ 1 \end{pmatrix} = z\,\mathbf{K} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{K}^{-1} \begin{pmatrix} \mathbf{Q} \\ 1 \end{pmatrix} + d \begin{pmatrix} \mathbf{Q}_o \\ 1 \end{pmatrix}, \qquad (5)$$

où $\mathbf{Q}_o = (i_o\,j_o)^T$ est le vecteur des coordonnées du point principal. Pour toute matrice $\mathbf{K}$ de la forme (1), il est facile de montrer que :

$$\mathbf{K} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{K}^{-1} = \begin{pmatrix} 1 & 0 & -i_o \\ 0 & 1 & -j_o \\ 0 & 0 & 0 \end{pmatrix}. \qquad (6)$$

Par conséquent, la relation entre $\mathbf{Q}$ et $\mathbf{Q'}$, c'est-à-dire entre projections perspective et perspective faible, s'écrit :

$$\frac{d}{z} \begin{pmatrix} \mathbf{Q'} \\ 1 \end{pmatrix} - \begin{pmatrix} \mathbf{Q} \\ 1 \end{pmatrix} = (\frac{d}{z} - 1) \begin{pmatrix} \mathbf{Q}_o \\ 1 \end{pmatrix}. \qquad (7)$$

Nous en déduisons les deux égalités :

$$z = d\,\frac{i' - i_o}{i - i_o} \quad ; \quad z = d\,\frac{j' - j_o}{j - j_o}. \qquad (8)$$

La double égalité (8) correspond à une expression de la « cote » z plus générale que celle donnée dans [7], où le point principal est supposé coïncider avec l'origine A du repère de l'image. Remarquons d'abord que notre approche analytique, qui s'oppose à l'approche purement géométrique de [7], met en évidence le fait que les seuls paramètres intrinsèques qui soient impliqués dans (8) sont les coordonnées (i_o,j_o) du point principal. Rappelons que d peut être choisi de manière arbitraire, ce qui signifie que u ne peut être calculée qu'à un facteur près (c'est-à-dire que le relief ne peut être reconstruit qu'à un facteur d'échelle près).
Une deuxième remarque est la suivante : on peut facilement déduire, de l'équation (7), que les paires de points (Q,Q') définissent des droites formant un faisceau dans l'image, dont le centre est le point principal O (*cf.* figure 5). En effet, nous savons que toute combinaison linéaire des vecteurs $(\mathbf{Q}^T 1)^T$ et $(\mathbf{Q'}^T 1)^T$ représente un point de l'image situé sur la droite passant par Q et Q'. Puisque Q et Q' sont les points conjugués dans l'image d'un point quelconque de la scène, toutes les droites considérées contiennent nécessairement O.
Remarquons enfin que, pour un pixel donné Q, les deux équations (8) comportent trois inconnues (i',j',z) si les coordonnées (i_o,j_o) du point principal sont connues, et cinq

inconnues sinon. Dans les deux cas de figure, (8) est donc un système sous-contraint. De manière générale, il n'existe pas d'autre équation liant ces inconnues : le problème est mal posé, c'est-à-dire que nous ne savons pas exprimer la valeur de la cote z associée à un pixel.

Pour rendre le problème bien posé, nous devons rajouter des hypothèses. Une première hypothèse consiste à supposer que la surface du document n'est pas une surface gauche quelconque, mais une surface cylindrique, de génératrices parallèles à la reliure (hypothèse H2). Pour un pixel Q, appelons Q_b et Q_h les pixels situés aux intersections de la droite passant par Q, parallèle à la reliure R, et des bords inférieur B_b et supérieur B_h de la page (*cf.* figure 5). Grâce aux hypothèses H1-H2, les trois pixels Q, Q_b et Q_h sont associés à trois points de la scène ayant la même cote z. En retranchant membre à membre les deux équations de la forme (7) obtenues pour les pixels Q_b et Q_h, nous obtenons l'égalité vectorielle :

$$\frac{d}{z}\,(\mathbf{Q}'_b - \mathbf{Q}'_h) = \mathbf{Q}_b - \mathbf{Q}_h, \tag{9}$$

qui est équivalente à un système de deux équations à cinq inconnues (les coordonnées des pixels Q'_b et Q'_h, plus la cote z). Ce dernier système devient bien contraint grâce à une nouvelle hypothèse disant que le contour de la page mise à plat est rectangulaire (hypothèse H3). Sachant que l'image de la reliure R n'est pas modifiée par la mise à plat, il découle des hypothèses H1-H2-H3 que les pixels de B_b sont déplacés, par mise à plat, sur la droite D_b orthogonale à R et passant par le pixel R_b le plus bas de R, de même que les pixels de B_h sont déplacés sur la droite D_h orthogonale à R et passant par le pixel R_h le plus haut de R (*cf.* figure 5). Si les coordonnées de O sont connues, alors

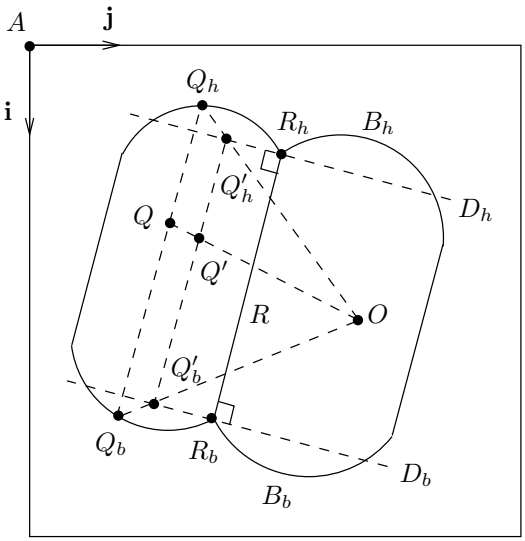


FIG. 5 – *Principe de la méthode SFC.*

il suffit d'observer la figure 5 pour constater que les pixels Q'_b et Q'_h sont définis sans ambiguïté. Par conséquent, le système (9) devient sur-contraint et n'a de solution exacte que si la droite (Q'_b,Q'_h) est parallèle à R. Mais il est encore plus important de remarquer que, si les coordonnées de O sont inconnues, le système est quand même bien contraint. En effet, en projetant l'égalité vectorielle (9) sur un axe parallèle à la reliure, il vient facilement :

$$z = d\,\frac{\|\mathbf{R}_b - \mathbf{R}_h\|}{\|\mathbf{Q}_b - \mathbf{Q}_h\|}. \tag{10}$$

Sous les hypothèses H1-H2-H3, nous venons donc de proposer une expression de la cote z qui est indépendante du point principal, contrairement à [7].

Shape from shading La méthode SFS fait l'hypothèse que le relief de la scène est suffisamment régulier pour que la normale sortante $\mathbf{n}(P)$ soit définie sans ambiguïté en chaque point P (*cf.* figure 4). Pour simplifier les notations, nous supposons dans ce paragraphe que le repère $(A\mathbf{ij})$ est orthonormé et que son origine A coïncide avec le point principal O. Il a été établi récemment, par trois groupes d'auteurs simultanément [9, 11, 4], l'expression des coordonnées de $\mathbf{n}(P)$ sous l'hypothèse de la projection perspective :

$$\mathbf{n}(P) = \frac{\big(-h(i,j)\,\partial_i u(i,j) \quad -h(i,j)\,\partial_j u(i,j) \quad 1\big)^T}{\sqrt{h(i,j)^2\,\|\nabla u(i,j)\|^2 + 1}}, \tag{11}$$

où $h(i,j) = f/v(i,j)$ et $v(i,j) = u(i,j) + i\,\partial_i u(i,j) + j\,\partial_j u(i,j)$. L'équation de base du SFS lie le niveau de gris d'un pixel Q à la normale $\mathbf{n}(P)$ en son point conjugué P, mais cela n'est possible que si l'éclairage incident et la réflectance sont connus en chaque point P de la scène, sans quoi le problème du SFS serait mal posé. Les hypothèses les plus souvent rencontrées sont les suivantes : la scène est éclairée par un faisceau lumineux parallèle, uniforme et frontal (hypothèse H4) ; la surface est lambertienne (hypothèse H5). Afin de satisfaire au mieux l'hypothèse H4, nous éclairons la scène à l'aide du flash annulaire dont il a déjà été question. Quant à l'hypothèse H5, elle est plus difficile à contrôler. En particulier, elle est assez mal vérifiée s'il s'agit de papier glacé. En outre, la réflectance n'est pas totalement connue, même si l'hypothèse H5 est valide, puisqu'une autre caractéristique photométrique de la surface, à savoir son albédo ρ, doit être connue : ρ est égal à $(1,1,1)$ pour du papier blanc non encré, à $(0,0,0)$ pour du papier encré à l'encre noire, et peut prendre toutes les valeurs du domaine $[0,1]^3$, correspondant à toutes les couleurs possibles dans la représentation RVB. Par conséquent, nous ne pouvons calculer $\mathbf{n}(P)$ que pour les pixels Q non encrés. Sous les hypothèses H4 et H5, il a été montré dans [9, 11, 4] que le problème du SFS peut être modélisé par une équation aux dérivées partielles du premier ordre, appelée « équation de l'eikonale perspective ». Dans [9, 11], la résolution de cette équation est menée avec u comme inconnue. Dans [4], deux inconnues intermédiaires

p et q sont introduites, définies par :

$$\mathbf{n} = \frac{\begin{pmatrix} -p(i,j) & -q(i,j) & 1 \end{pmatrix}^T}{\sqrt{p(i,j)^2 + q(i,j)^2 + 1}}, \tag{12}$$

car un vecteur normé a bien deux degrés de liberté. Il est facile de montrer que l'introduction de ces deux inconnues intermédiaires permet de réécrire l'équation de l'eikonale perspective sous la forme suivante :

$$p(i,j)^2 + q(i,j)^2 = \frac{{g_{\max}}^2}{g(i,j)^2} - 1, \tag{13}$$

où $g(i,j)$ désigne le niveau de gris au pixel (i,j), et $g_{\max}$ la valeur maximale de g. L'équation (13) comportant deux inconnues $(p(i,j),q(i,j))$ par pixel, cela exprime le fait que le SFS est, de manière générale, un problème mal posé. Néanmoins, sous l'hypothèse H2 que la surface est cylindrique, et après l'éventuel changement de repère dans l'image dont il a déjà été question un peu plus haut, il est facile d'établir que $p(i,j)$ est uniformément nul, ce qui rend le problème bien posé. En utilisant en outre la connaissance *a priori* relative à la convexité de chaque page, $q(i,j)$ peut être calculé sans ambiguïté sur le signe, pour chaque pixel, grâce à l'équation (13).
Il reste ensuite à calculer $u(i,j)$ à partir de $q(i,j)$. En utilisant (11) et (12), il est montré dans [4] que :

$$\nabla u(i,j) = \frac{u(i,j) \quad \begin{pmatrix} p(i,j) & q(i,j) \end{pmatrix}^T}{f - i\,p(i,j) - j\,q(i,j)}, \tag{14}$$

qui est une nouvelle équation aux dérivées partielles. Cette équation est néanmoins nettement plus simple à résoudre que l'équation de l'eikonale perspective, dans la mesure où elle linéaire.

Différence fondamentale entre SFC et SFS Il existe une différence fondamentale entre les deux méthodes de reconstruction tridimensionnelle que nous venons de décrire : alors que, pour la méthode SFC, la connaissance d'aucun paramètre intrinsèque n'est requise, il n'en va pas de même pour la méthode SFS. D'une part, la distance focale f apparaît explicitement dans l'équation (14). D'autre part, nous avons supposé le repère $(A\mathbf{ij})$ orthonormé et choisi son origine A au point principal O. Dans la mesure où les coordonnées (i,j) du pixel apparaissent explicitement dans (14), il faut que la matrice de calibrage $\mathbf{K}$ de l'appareil photographique soit connue.

Simulation de la mise à plat Comme nous l'avons déjà dit, la simulation de la mise à plat comporte deux étapes : une étape de dégauchissage suivie d'une étape d'interpolation. Nous utilisons la même méthode de dégauchissage que dans la plupart des articles où l'hypothèse H2 de surfaces gauches cylindriques est faite [17, 14, 7].
Pour la deuxième étape, nous avons comparé trois méthodes d'interpolation : l'algorithme de Smythe [10], par l'intermédiaire du logiciel libre `xmorph` ; la triangulation de Delaunay, par l'intermédiaire de la fonction `griddata` de MATLAB ; enfin, une méthode disponible dans la bibliothèque `OpenGL`, optimisée pour l'affichage graphique. Dans la mesure où cette dernière méthode s'est avérée la plus rapide des trois et la plus simple à mettre en œuvre, c'est elle que nous avons retenue.

2.3 Résultats

Afin de valider nos deux méthodes de correction géométrique, nous avons choisi les deux photographies représentées sur la colonne de gauche de la figure 6. La première (figure 6a) est la photographie d'une grille régulière, tandis que la deuxième est la partie gauche de la photographie déjà montrée sur la figure 1. Les résultats obtenus avec la première image sont présentés sur les figures 6b (correction par SFC) et 6c (correction par SFS). Ces résultats sont particulièrement intéressants, car ils montrent que les deux méthodes permettent de retrouver la structure euclidienne de la page mise à plat, dans la mesure où les lignes de la grille deviennent quasi-rectilignes et régulièrement espacées, sans que cela ait été utilisé comme connaissance *a priori*. Les résultats obtenus avec la deuxième image, qui sont présentés sur les figures 6b' (correction par SFC) et 6c' (correction par SFS), sont eux aussi plutôt satisfaisants. Une comparaison rapide entre ces deux derniers résultats montre que le SFC effectue une correction légèrement meilleure que le SFS. Cela est dû au fait que, pour ces deux images, la seule hypothèse spécifique au SFC (H3 : page mise à plat supposée rectangulaire) est plus facile à garantir, en pratique, que les deux hypothèses spécifiques au SFS (H4 : éclairage parallèle, uniforme et frontal ; H5 : surface lambertienne). En particulier, l'hypothèse H4 est rendue partiellement caduque au voisinage de la reliure, à cause des réflexions secondaires d'une page sur l'autre (dans [13], l'effet des réflexions secondaires est partiellement pris en compte).

3 Correction photométrique

3.1 État de l'art

Une fois qu'elles sont corrigées géométriquement, les images comportent encore un défaut photométrique, dans la mesure où le papier non encré n'a pas un niveau de gris uniforme (*cf.* figures 6b-c-b'-c'). Pour corriger ce défaut, différentes techniques ont été proposées, mais la plupart d'entre elles sont relativement naïves : elles consistent soit à seuiller l'image [15, 2], soit à interpoler les niveaux de gris des bords de la page [12], soit même à n'effectuer aucune correction [5, 14, 7, 1].
Quelques auteurs [16, 13] simulent une « image de luminance » (*shading image*) en utilisant l'équation (13) et la connaissance du relief obtenue lors de l'étape de correction géométrique (cette simulation, qui revient à effectuer du « *shading from shape* », sera notée SFS^{-1} par la suite). Puisque nous connaissons le relief du document, ainsi que la réflectance de la surface et l'éclairage de la scène (hypothèses H4-H5), il semble effectivement intéressant de calculer l'image de luminance d'un document non encré pos-

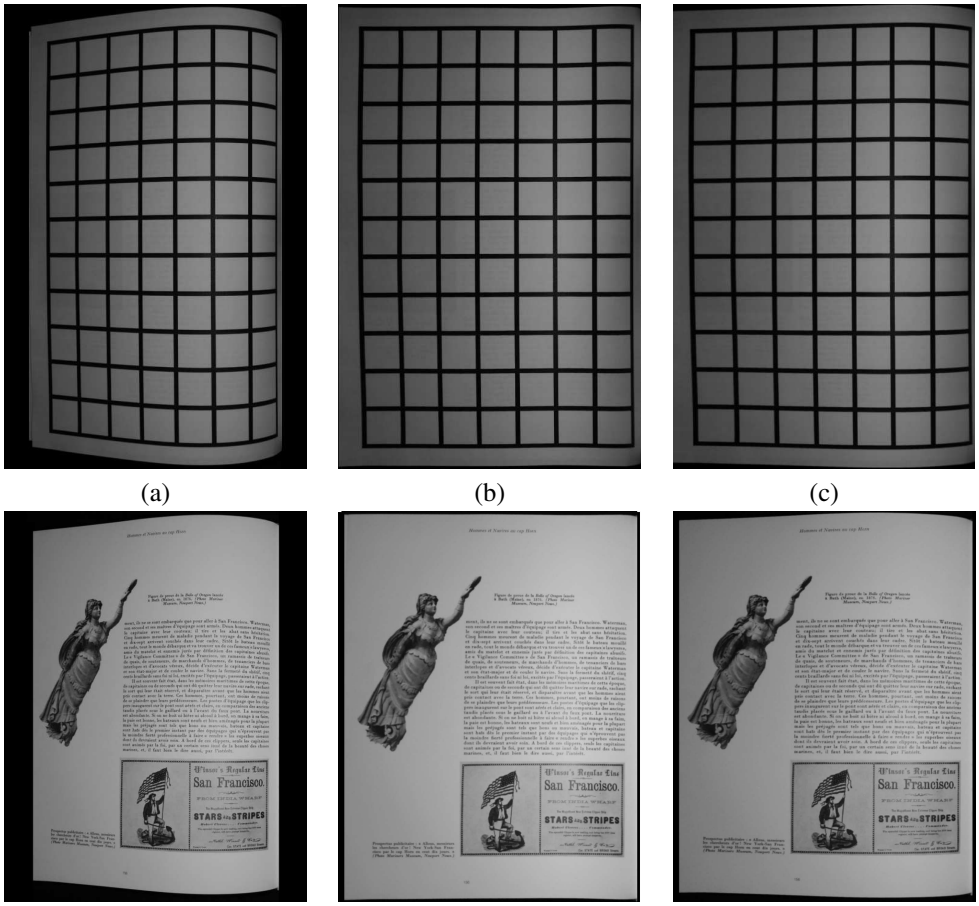


FIG. 6 – *(a-a') Photographies originales et corrigées géométriquement, en utilisant (b-b') le SFC ou (c-c') le SFS.*

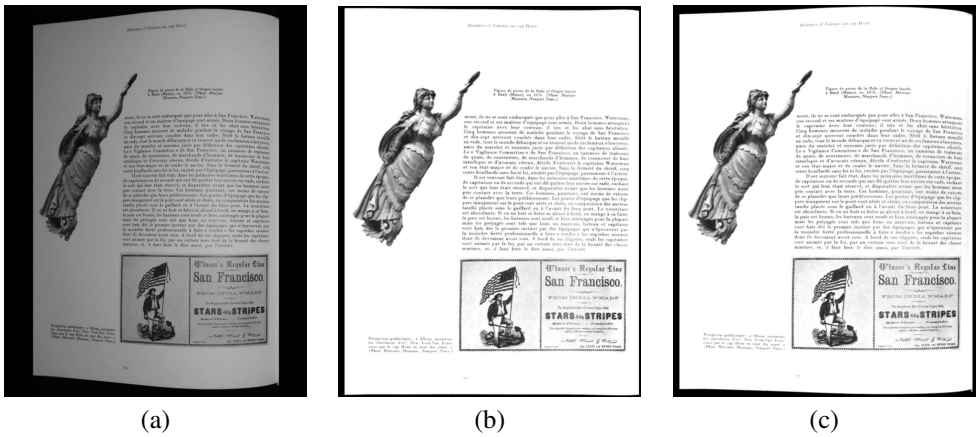


FIG. 7 – *(a) Photographie originale et images corrigées géométriquement et photométriquement, en utilisant (b) le SFC ou (c) le SFS.*

sédant le même relief. La figure 8 montre comment cette image de luminance est utilisée dans la chaîne complète de traitement : il faut noter que la photographie est corrigée photométriquement avant d'être mise à plat.

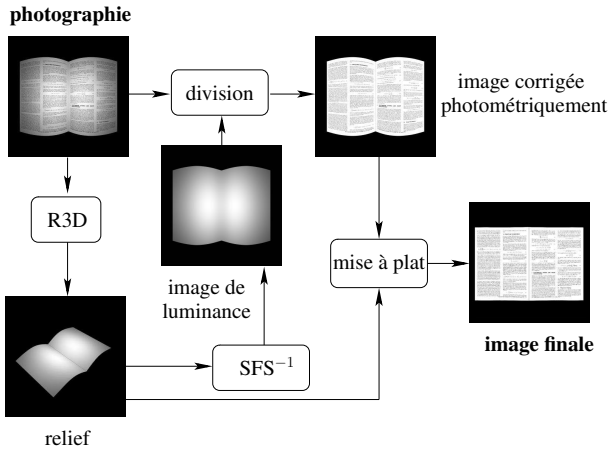


FIG. 8 – *Chaîne complète de traitement.*

3.2 Résultats

Les images corrigées géométriquement et photométriquement correspondant à la photographie représentée sur la figure 7a sont présentées sur les figures 7b-c. On peut constater que les parties non encrées sont devenues uniformément blanches, cependant que les zones de texte et les images ont bien retrouvé leur apparence d'origine. On peut également remarquer que ces deux images sont quasi-identiques.

4 Discussion

4.1 Lisibilité du texte

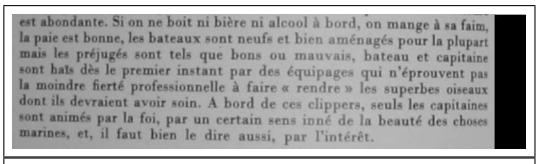
Sur la figure 9, on a reporté trois imagettes extraites des figures 7a-b-c, correspondant à une même zone de texte. Afin de valider les résultats de manière quantitative, nous avons testé le logiciel libre `gocr`[1] de reconnaissance de caractères (OCR) sur ces trois imagettes, ainsi que sur l'image scannée. Les résultats, qui ont été reportés dans le tableau 1, montrent que les deux corrections effectuées permettent à l'OCR de décider plus souvent, et de manière plus fiable, que sur la photographie originale.

	(a)	(b)	(c)	(d)
succès	92 %	2 %	61 %	55 %
échec	6 %	70 %	24 %	28 %
abstention	2 %	28 %	15 %	17 %

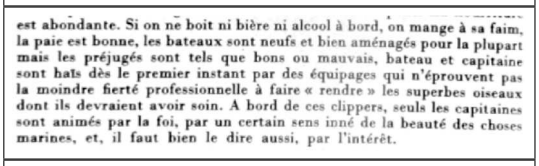
TAB. 1 – *Résultats de l'OCR sur une même zone de texte extraite des quatre images suivantes : (a) image scannée du document à plat ; (b) photographie du document gauche ; (c) image corrigée en utilisant le SFC ; (d) image corrigée en utilisant le SFS.*

1. http://jocr.sourceforge.net/

(a)



(b)



(c)

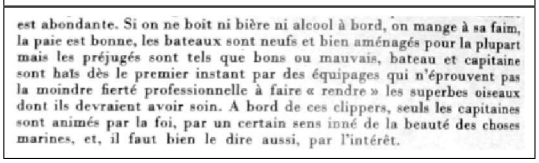


FIG. 9 – *(a-b-c) Zooms sur une même zone de texte des images (a-b-c) de la figure 7.*

4.2 Limites des deux méthodes

Comme nous l'avons déjà remarqué, les deux méthodes de correction proposées fournissent des résultats très comparables. Cependant, elles se distinguent par le fait qu'elles ne peuvent pas être appliquées aux mêmes images.
Par exemple, la photographie de la figure 10a ne se prête pas à la correction par SFC, puisque aucun bord de la page n'est visible, tandis que le SFS permet d'obtenir un bon résultat, présenté sur la figure 10c. En revanche, le SFS ne peut pas être utilisé pour les photographies des figures 10a'-a", puisque soit l'albédo ne vérifie pas l'hypothèse H4 (figure 10a'), soit l'éclairage ne vérifie pas l'hypothèse H5 (figure 10a"). Sur les figures 10b'-b", on peut observer les bons résultats obtenus sur ces deux mêmes photographies en utilisant le SFC, mais on remarque néanmoins que la correction photométrique de notre chaîne de traitement n'a pas pu être effectuée, puisque l'étape de SFS^{-1} n'est pas applicable.

4.3 Relâchement des hypothèses

Nous avons déjà signalé une disymétrie importante entre les méthodes SFC et SFS : alors que la méthode SFC peut fonctionner avec un appareil photographique non calibré, cela est impossible pour la méthode SFS. Afin de mettre en évidence cette disymétrie, nous avons pris deux photographies, représentées sur les figures 11a-a', à l'aide d'un appareil photographique numérique « grand public » (Nikon Coolpix 775) non calibré, sans utiliser de montage optique spécifique. Par conséquent, l'hypothèse H1 de fronto-parallélisme n'est plus parfaitement vérifiée, et les paramètres internes sont inconnus. Les résultats, présentés sur les figures 11b-b', sont néanmoins de très bonne qualité.

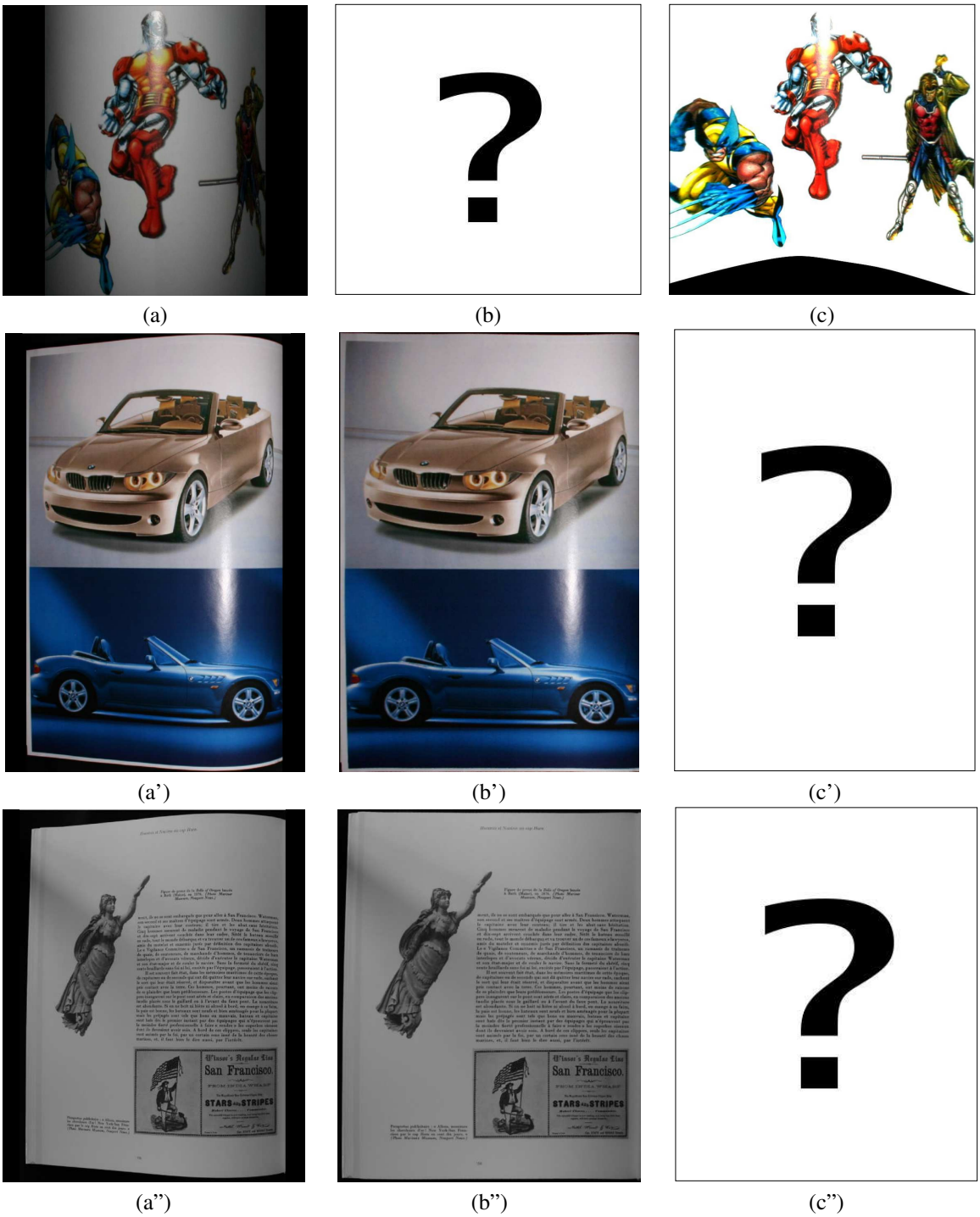
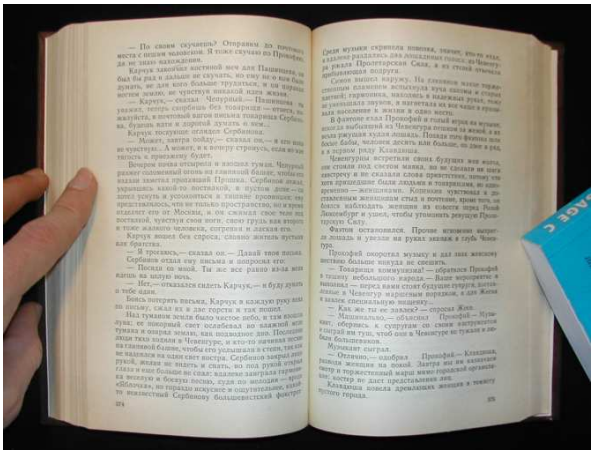
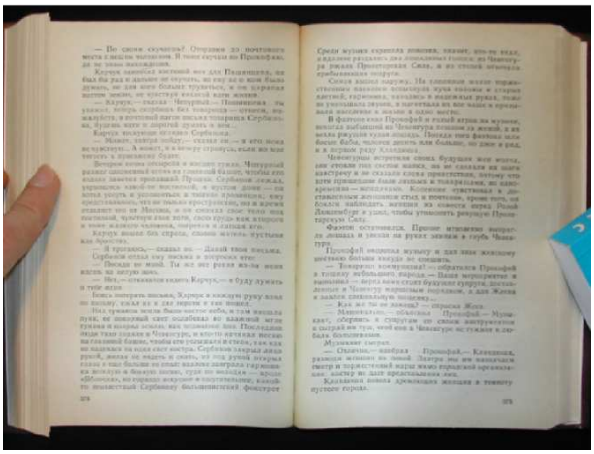


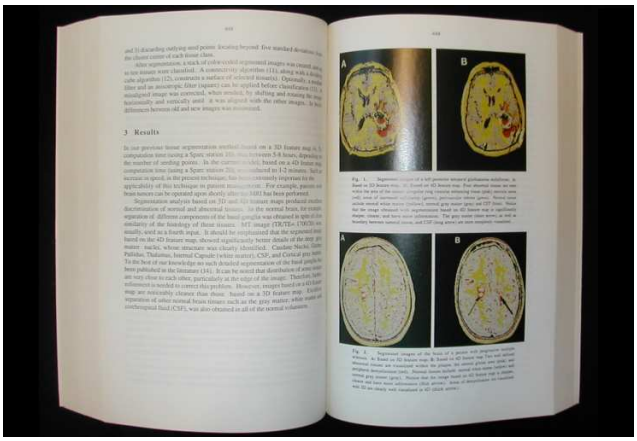
FIG. 10 – *(a) Photographie sans les bords de la page ; (b) le SFC n'est pas utilisable ; (c) image corrigée, géométriquement et photométriquement, en utilisant le SFS ; (a') photographie avec albédo inconnu ; (b') image corrigée géométriquement en utilisant le SFC ; (c') le SFS n'est pas utilisable ; (a") photographie prise en éclairage naturel ; (b") image corrigée géométriquement en utilisant le SFC ; (c") le SFS n'est pas utilisable.*



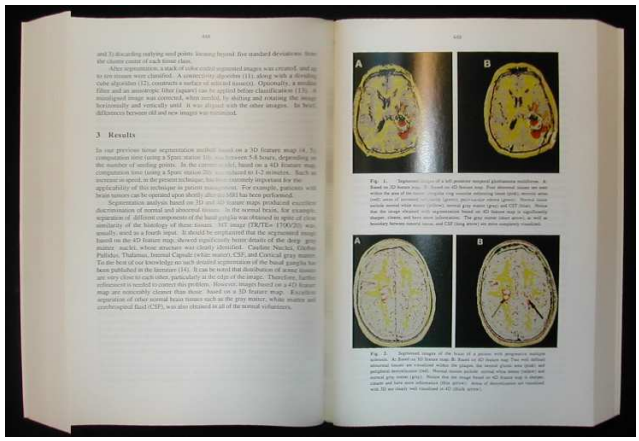
(a)



(b)



(a')



(b')

FIG. 11 – *(a-a') Photographies prises avec un appareil « grand public » non calibré et (b-b') images corrigées géométriquement en utilisant le SFC.*

5 Conclusion et perspectives

Dans cet article, nous avons proposé d'utiliser deux méthodes de reconstruction tridimensionnelle et de prendre en compte la perspective, afin de simuler la mise à plat d'un document gauche cylindrique, à partir d'une seule photographie de ce document. Les deux algorithmes qui en résultent ont été testés sur images réelles.
En guise de perspective, nous allons maintenant nous intéresser au cas de documents gauches non cylindriques. Cela nécessitera probablement de fusionner les deux techniques proposées, dans la mesure où, pour de tels reliefs, chacune des deux méthodes revient à résoudre un problème mal posé.

Remerciements
Ce travail a été partiellement financé par EGIDE, dans le cadre du projet Galilée « PLATONOV » (Photocopies de Livres Anciens par Techniques d'Optimisation Numérique et sOlutions de Viscosité).

Références

[1] M. S. BROWN et W. B. SEALES. « Image Restoration of Arbitrarily Warped Documents ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10):1295–1306, octobre 2004.

[2] H. CAO, X. DING et C. LIU. « A Cylindrical Surface Model to Rectify the Bound Document Image ». Dans *Proceedings of the 9th IEEE International Conference on Computer Vision (volume I)*, pages 228–233, Nice, France, octobre 2003.

[3] S. I. CHO, H. SAITO et S. OZAWA. « A Divide-and-conquer Strategy In Recovering Shape Of Book Surface From Shading ». Dans *Proceedings of the 9th International Conference on Image Analysis and Processing (volume II)*, volume 1311 de *Lecture Notes in Computer Science*, pages 262–269, Florence, Italie, septembre 1997.

[4] F. COURTEILLE, A. CROUZIL, J.-D. DUROU et P. GURDJOS. « Towards shape from shading under realistic photographic conditions ». Dans *Proceedings of the 17th International Conference on Pattern Recognition (volume II)*, pages 277–280, Cambridge, Royaume-Uni, août 2004.

[5] A. DONCESCU, A. BOUJU et V. QUILLET. « Former books digital processing: image warping ». Dans *Proceedings of the IEEE Workshop on Document Image Analysis*, pages 5–9, San Juan, Porto Rico, juin 1997.

[6] R. I. HARTLEY et A. ZISSERMAN. *Multiple View Geometry in Computer Vision*. Cambridge University Press, seconde édition, 2003.

[7] M. KASHIMURA, T. NAKAJIMA, N. ONDA, H. SAITO et S. OZAWA. « Practical Introduction of Image Processing Technology to Digital Archiving of Rare Books ». Dans *Proceedings of the International Conference on Signal Processing Application Technology*, Toronto, Canada, octobre 1999.

[8] S. OZAWA et M. KASHIMURA. « Application of Image Processing Technology to Digital Archiving and Study of Rare Books ». Dans *Proceedings of the Kyoto International Conference on Digital Libraries*, pages 49–56, Kyoto, Japon, novembre 2000.

[9] E. PRADOS et O. FAUGERAS. « "Perspective Shape from Shading" and Viscosity Solutions ». Dans *Proceedings of the 9th IEEE International Conference on Computer Vision (volume II)*, pages 826–831, Nice, France, octobre 2003.

[10] D. B. SMYTHE. « A Two-Pass Mesh Warping Algorithm for Object Transformation and Image Interpolation ». Technical Memo 1030, Industrial Light and Magic, Computer Graphics Department, Lucasfilm Ltd., 1990.

[11] A. TANKUS, N. SOCHEN et Y. YESHURUN. « A New Perspective [on] Shape-from-Shading ». Dans *Proceedings of the 9th IEEE International Conference on Computer Vision (volume II)*, pages 862–869, Nice, France, octobre 2003.

[12] Y.-C. TSOI et M. S. BROWN. « Geometric and Shading Correction for Images of Printed Materials: A Unified Approach Using Boundary ». Dans *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (volume I)*, pages 240–246, Washington, D.C., États-Unis, juin 2004.

[13] T. WADA, H. UKIDA et T. MATSUYAMA. « Shape from Shading with Interreflections under Proximal Light Source: 3D shape Reconstruction of Unfolded Book Surface from a Scanner Image ». Dans *Proceedings of the 5th IEEE International Conference on Computer Vision*, pages 66–71, Boston, Massachusetts, États-Unis, juin 1995.

[14] A. YAMASHITA, A. KAWARAGO, T. KANEKO et K. T. MIURA. « Shape Reconstruction and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System ». Dans *Proceedings of the 17th International Conference on Pattern Recognition (volume I)*, pages 482–485, Cambridge, Royaume-Uni, août 2004.

[15] Z. ZHANG et C. L. TAN. « Recovery of Distorted Document Images from Bound Volumes ». Dans *Proceedings of the 6th International Conference on Document Analysis and Recognition*, pages 429–433, Seattle, Washington, États-Unis, septembre 2001.

[16] Z. ZHANG, C. L. TAN et L. FAN. « Estimation of 3D Shape of Warped Document Surface for Image Restoration ». Dans *Proceedings of the 17th International Conference on Pattern Recognition (volume I)*, pages 486–489, Cambridge, Royaume-Uni, août 2004.

[17] Z. ZHANG, C. L. TAN et L. FAN. « Restoration of Curved Document Images through 3D Shape Modeling ». Dans *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (volume I)*, pages 10–15, Washington, D.C., États-Unis, juin 2004.