

Numérisation 3D de documents par photogrammétrie

Pierre Cubaud – Jean-François Haas – Alexandre Topol

Centre d'études et de recherche en Informatique (CEDRIC)
Conservatoire national des arts & métiers (CNAM)
192 rue St-Martin F-75003 Paris

{cubaud, topol}@cnam.fr, jfhaas@jfhaas.com

Résumé : *Bien que d'application courante dans l'industrie et en muséologie, la numérisation 3D ne fait encore l'objet que de quelques applications dans le domaine des documents patrimoniaux. Après une revue des différentes techniques de numérisation 3D actuellement disponibles, nous décrivons la chaîne photogrammétrique développée et les résultats obtenus sur quelques ouvrages tests. Nous indiquons ensuite les applications possibles de cette technique : visualisation 3D, mise à plat et éclairage synthétique.*

Mots-clés : numérisation, 3D, photogrammétrie, document, livre relié, visualisation, mise à plat.

1 Introduction

La numérisation de fonds de bibliothèques et d'archives est aujourd'hui une pratique systématique, destinée à permettre leur accès à distance par un large public, tout en préservant l'intégrité des documents, qui risque d'être mise à mal par des manipulations répétées. Pour les ouvrages contemporains dont on peut sacrifier un exemplaire, la numérisation est pratiquée à une échelle quasi industrielle avec des machines de grande capacité, après massicotage et placement à plat dans des chargeurs automatiques. Sans massicotage apparaît un problème de déformation due à la perspective. En effet, la surface à numériser est rarement plane, ne serait-ce qu'à cause de la présence de la reliure. C'est pourquoi les ouvrages sont en général placés sous une vitre qui force la planéité. On conçoit cependant que cette technique ne soit pas appropriée pour des ouvrages précieux qu'il est hors de question de manipuler de la sorte. Par ailleurs, la nécessité de plaquer le livre contre la vitre à chaque page ralentit considérablement le processus et le rend difficilement automatisable. On utilise alors d'autres machines, qui numérisent les ouvrages tels quels, sans vitre. Certaines sont depuis peu équipées d'un tourne-page automatique (4DigitalBooks). Les déformations des images sont alors traitées par logiciel, selon diverses heuristiques : rectification des bords [BRO 03], des lignes de texte (logiciel I2S Book Restorer), inclinomètrie [WAD 95], etc.

Dès lors que l'objet livre devient lui-même centre d'intérêt, cette approche devient insuffisante. On peut citer ainsi les livres d'artistes, avec souvent de nombreux collages superposés, les livres « à systèmes » avec

parties mobiles, voire les herbiers. Il semble dès lors naturel de chercher à en capturer la géométrie par une numérisation 3D. Celle-ci pourrait également trouver des débouchés dans le commerce électronique de livres rares mais plus « conventionnels », actuellement en plein essor. Malgré le recours de plus en plus systématique à la numérisation 3D dans le domaine culturel [DEQ 00] et le développement d'un marché spécifique, il n'existe pas à l'heure actuelle de matériel dédié à la numérisation 3D de documents.

Après une revue des techniques de numérisation alternatives, nous décrivons une chaîne photogrammétrique développée à cette fin, puis nous discutons les résultats expérimentaux obtenus. Enfin, nous présentons les applications possibles de la numérisation 3D de documents. A partir de la connaissance de la surface 3D du document numérisé, on peut tenter de le déplier ou de le mettre à plat. C'est utile pour les manuscrits reliés (registres paroissiaux, ou manuscrits d'auteurs importants), pour lesquels les heuristiques de rectification de lignes de textes sont inopérantes. C'est historiquement la première application envisagée pour la numérisation 3D [BRO 00]. Les progrès effectués en matière de visualisation 3D temps-réel permettent toutefois de s'affranchir de ce procédé et d'offrir à l'utilisateur une interface de consultation purement 3D.

2 Les techniques de numérisation 3D d'objets rapprochés

Lumière structurée

Un motif lumineux connu, par exemple une grille, est projeté sur la surface, et une prise de vue effectuée sous un angle différent. La déformation du motif sur l'image permet de déduire la surface (fig. 1a). La précision obtenue est liée à la finesse du motif. Les différentes lignes de la grille doivent pouvoir être identifiées sans ambiguïté. On peut par exemple leur donner des couleurs différentes. C'est le principe de l'appareil 3D1500 de Minolta, qui a fait l'objet d'un test au CEDRIC (CNAM) en juin 2000. Un résultat de ce test appliqué à une reliure in-12 assez typique du XVIIIème français (Nollet, Leçons de physique) est montré fig. 1b. Les ondulations du papier (de chiffon à l'époque, donc épais), ainsi que la

gorge de reliure, sont bien rendues. On notera cependant l'arrondi des angles et les ondulation au niveau des tranches dues à l'interférence entre les bandes et la liasse des feuilles. Cet appareil a également été utilisé dans le projet *Digital Atheneum* [BS 00] pour des manuscrits du moyen âge très gondolés. Le 3D1500 ne semble plus commercialisé.

La triangulation laser

Il s'agit d'appareils « tout en un », appelés « scanner 3D », constitués d'un laser à balayage associé à un appareil photographique numérique. Il s'agit d'un cas particulier de lumière structurée, où la ligne projetée est générée par le balayage d'un rayon laser, elle-même balayant l'objet. Nous avons pu tester le scanner Minolta VIVID 700, qui fonctionne selon ce principe, sur un exemple de livre ancien du même type que le précédent (Lalande, Almanach de physique, in-12 broché). La texture des caractères imprimés est interprétée comme un léger relief (fig. 2) : l'appareil ne fonctionne en effet qu'avec une radiométrie de l'objet uniforme. Cependant, la précision de restitution est très bonne dans les autres zones : inférieure au 1/10 mm.



FIG. 1 - Test du 3D1500 de Minolta. (a) projection de bandes de couleurs (b) surface obtenue

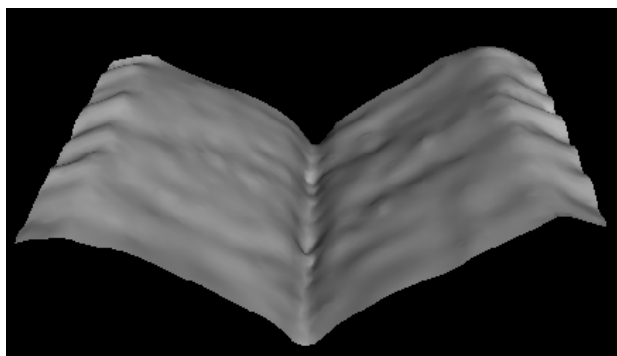


FIG. 2 - Surface obtenue avec le VIVID 700 (20 000 triangles)

La photogrammétrie

C'est la reconstruction de la surface à partir d'un couple de photographies stéréoscopiques [EGE 01]. Chaque couple de points identifié par corrélation entre les deux images donne les coordonnées du point objet (3D) correspondant. Par rapport aux techniques précédentes, la photogrammétrie offre les avantages suivants :

- Le couplage entre acquisition de la texture et acquisition de la surface : il n'y a pas à proprement parler d'acquisition spécifique de la surface, mais seulement un doublement de l'acquisition de la texture.

- La séparation totale entre la phase d'acquisition et la phase de calcul : le traitement informatique peut se faire ailleurs et plus tard, la seule information à transférer étant les clichés eux-mêmes.

- La compatibilité avec la prise de vues classique ou avec les systèmes d'acquisition à balayage.

- La rusticité de l'installation : un matériel photographique standard, orienté approximativement et non étalonné donne déjà des résultats.

- Le caractère instantané de l'acquisition, puisqu'il s'agit de deux clichés photographiques. Ceux-ci sont simultanés ou successifs. Dans le premier cas, qui nécessite deux appareils, l'immobilisation parfaite du livre n'est même pas nécessaire

En revanche, la photogrammétrie présente deux inconvénients :

- La nécessité de la présence d'une texture, qui exclut son utilisation pour des zones uniformes (marges, interlignes, etc.).

- Le risque de fausse corrélation, qui reste le problème majeur de la photogrammétrie si on l'envisage sans aucun contrôle humain.

L'acquisition par photogrammétrie est la solution retenue ici pour ses qualités, en particulier de rusticité mentionnées ci-dessus ; mais aussi parce qu'à notre connaissance, elle n'a pas fait à ce jour l'objet d'un travail spécifique à la numérisation de document.

3 La chaîne photogrammétrique

Malheureusement, il n'existe pas à notre connaissance de logiciel de photogrammétrie numérique de type « open source » réutilisable et nous avons dû en réaliser un, décrit plus complètement dans [HAA 03].

Dans un logiciel de photogrammétrie, la partie délicate est l'identification des couples de points homologues sur les deux vues. Pour chaque point d'intérêt de la vue gauche par exemple, on cherche à identifier son homologue sur la vue droite par corrélation automatique. La description d'une surface nécessitant des dizaines de milliers de points, le temps de calcul peut devenir rédhibitoire, et ceci d'autant plus qu'on cherche une fiabilité (absence de fausses corrélations) élevée.

Nous utilisons dans ce logiciel une approche itérative destinée à minimiser ce temps de calcul. En effet, la surface d'une page de livre a des propriétés de continuité

qu'il convient d'exploiter : la hauteur d'un point n'est jamais très éloignée de celles de ses voisins. Or la connaissance à priori de la plage Z1Z2 des hauteurs possibles du point objet M correspondant à un point de l'image de gauche mG donné permet de limiter considérablement le domaine de recherche m1m2 de son homologue sur l'image de droite (fig. 3).

Nous procédons donc ainsi :

Un premier semis de points est sélectionné sur l'image de gauche. Le calcul de leurs homologues à droite n'est pas très long puisqu'ils sont peu nombreux (environ 60 points). Une approximation de la surface est ensuite obtenue par triangulation des points objets alors calculés.

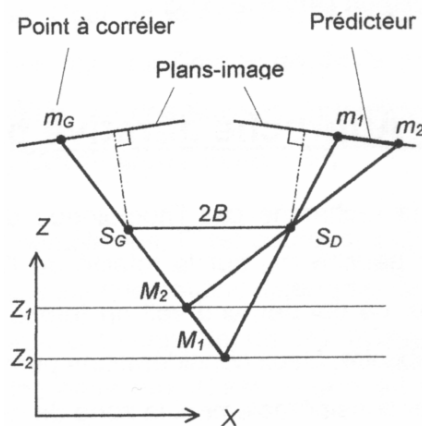


FIG. 3 - Recherche de points homologues

Un second semis de points, quatre fois plus dense, est sélectionné à gauche. Pour chacun d'eux, la surface approximative obtenue précédemment permet de prédire la plage de hauteurs possibles du point objet correspondant. La recherche de l'homologue à droite se fait alors sur un segment ("prédicteur"), comme expliqué précédemment, ce qui réduit le temps de calcul de corrélation pour chacun des points. Une surface plus proche de la réalité est obtenue par triangulation entre ces nouveaux points.

On recommence ainsi le processus : A chaque boucle, le nombre de points à corrélérer est plus grand, mais la surface triangulée étant plus précise, l'approximation de la hauteur des points-objet est meilleure, et les segments de recherche sont plus courts.

L'algorithme général du logiciel est le suivant :

```
Chargement image et paramètres d'orientation
approximatifs
répéter
  Production des points d'intérêt à gauche
  (coins de Harris)
```

```
Sélection d'un jeu de points d'intérêt
Recherche des points homologues à droite
Affinage de l'orientation et calcul des
points objets
Ajout de points à la triangulation
jusqu'à nombre de points suffisants
Production du fichier VRML, de la texture et de
l'ortho-image
```

Sa mise en œuvre représente 5600 lignes de code Java, sous Windows. Il est complété par une feuille de calcul Excel où sont entrés les paramètres géométriques du banc de prise de vue et les paramètres extrinsèques des caméras, dont on déduit la résolution à attendre du modèle 3D.

4 Résultats expérimentaux

Le banc expérimental a été monté sur un bâti de prises de vues professionnel. Sur le socle, deux appareils photographiques (Konika Digital Revio KD500Z et Canon PowerShot S45, 4Mpix) ont été fixés à l'aide de deux « bras magiques » et visent le document à 30° par rapport à la verticale. Le banc est équipé de quatre projecteurs tungstène, deux de chaque côté, orientés à 45°. La résolution attendue est de l'ordre de 0,1 mm. Nous avons fait l'impasse sur l'étalonnage des appareils. Celui-ci aurait nécessité une petite étude à part entière, en particulier la réalisation d'un objet étalon aux dimensions parfaitement connues. Cet aspect n'a pas été jugé prioritaire, dans la mesure où l'on peut faire de la restitution 3D non métrique (c'est-à-dire à un facteur d'échelle près) en se passant des paramètres d'orientation interne (focale et coordonnées du point principal d'autocollimation). Dans les faits, une échelle approximative est déterminée par l'entrée manuelle dans le logiciel de la longueur de la base stéréoscopique, mesurée au double-décimètre. En ce qui concerne la distorsion, sa non-prise en compte (même si celle-ci est prévue dans le logiciel) conduit à rechercher chaque point homologue non plus strictement sur un segment, mais sur une plage autour de ce segment, ainsi qu'à une légère erreur dans le positionnement des points-objet.

Trois ouvrages anciens ont été utilisés pour l'expérimentation, dans chacun desquels nous avons numérisé une ou quelques doubles pages. Les clichés ont été utilisés en niveaux de gris, tel quels dans un premier temps. Il s'est cependant révélé nécessaire d'effacer le fond (le plan de travail), pour éviter les recherches de corrélation inutiles. Les résultats de mesure qui suivent sont donnés pour les premières pages du volume représenté fig. 4 :

Phase	Grandeur	Début	Boucles						Fin
			1	2	3	4	5	6	
Orient. modèle	itérations/angle	4 / 26,3°							
Points d'intérêt	nombre	30 936							
Points sélect.	distance minimale	(pixels)	100	50	25	12	6	3	
	nombre		85	209	695	1900	4548	8748	
Corrélation	abandons		52	76	185	360	872	1039	
	taux		61%	36%	27%	19%	19%	12%	
	Incohérences		0	1	4	7	14	36	
	faux		0,0%	0,5%	0,6%	0,4%	0,3%	0,4%	
	fausses corrélét.		0	0	2	7	20	50	
	faux		0,0%	0,0%	0,0%	0,3%	0,4%	0,6%	
Orient. objet	itérations/angle	3 / 0,19°							
Triangulation	sommets ajoutés		33	132	506	1533	3662	7671	
	chemin moyen	(triangles)	3	6	12	19	27	32	
Points objet	incertitude	(mm)	0,05	0,12	0,07	0,08	0,07	0,17	
Script VRML	nombre triangles								27 075
	nombre sommets								13 537

À chaque boucle une partie croissante des 30 936 points d'intérêt détectés sont sélectionnés pour être corrélés sur l'image de droite. Leur nombre est en progression géométrique de facteur environ 2,5. Le taux de non-corrélations (abandons) est élevé au départ, et ne cesse de décroître. Ceci confirme que le prédicteur fait de mieux en mieux son travail, et que le segment de recherche étant de plus en plus petit (fig. 5), il y a de moins en moins de points candidats et donc d'ambiguïté.

Dans le cas d'une page de texte, le détecteur de coins de Harris fournit une grande quantité de points d'intérêt assez bien répartis. Dans le cas d'une gravure, de grandes zones sont dépourvues de points d'intérêt. Il faudrait ajouter des points par simple détection de gradient horizontal. D'autres méthodes sont encore à évaluer [SCH 00]. Le nombre de fausses corrélations indiqué dans le tableau ci-dessus a été relevé visuellement. Il est donc très approximatif. On compte quelques dizaines de fausses corrélations parmi des milliers de points. Celles-ci sont cependant très gênantes, car elles provoquent des aberrations de surface inacceptables. À l'exception de rares cas, les fausses corrélations se concentrent sur les zones contenant des lignes, contours et motifs répétitifs (fig. 6) parallèles à la base stéréoscopique. Dans ce cas, le segment de recherche leur est parallèle, et l'ambiguïté dans l'identification de l'homologue est grande. Ce problème peut être contourné par une rotation du document. Les bords de page posent également problème du fait de l'absence de texture. La triangulation produite (fig. 7 et 8) est calculée en moins de 5 min. sur un ordinateur de type Pentium IV cadencé à 2 GHz. Ce temps devrait pouvoir être considérablement diminué par optimisation de l'algorithme, ne serait-ce que sa parallélisation. Nous ne nous sommes pas penchés sur cet aspect.

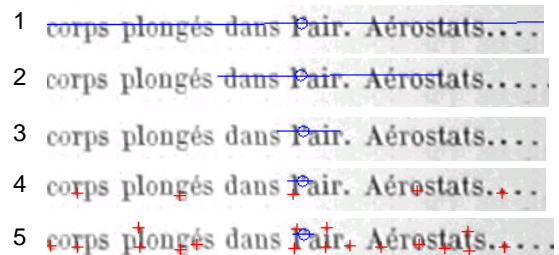
La robustesse, qu'il s'agisse d'éviter les fausses corrélations ou bien de les supprimer à posteriori, peut être considérablement améliorée par un certain nombre de techniques. Ceci constitue un sujet pour une étape ultérieure de ce travail. Parmi ces techniques, citons l'introduction :

- d'une troisième image,
- d'une corrélation croisée : l'homologue de l'homologue d'un point doit être ce point lui-même,

- d'une fenêtre de corrélation adaptative, c'est-à-dire tenant compte de la surface approximative locale déjà calculée,
- de contraintes sur les vecteurs normaux de la triangulation, afin de tenir compte des éventuelles propriétés de continuité et de dérivabilité de la surface.



FIG. 4 - Exemplaire très abîmé par l'humidité : Privat, Deschanel « Précis de physique », Paris, 1855, in-8°



A gauche, le numéro de la boucle de calcul.

FIG. 5 - Evolution du segment de recherche du point homologue

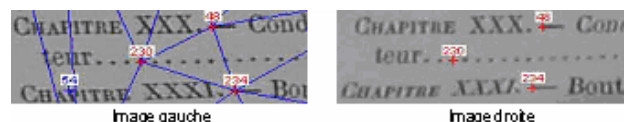


FIG. 6 - Exemple de fausse corrélation (point 230) sur un motif répétitif parallèle à la base



FIG. 7 - Vue gauche après triangulation (27075 triangles)

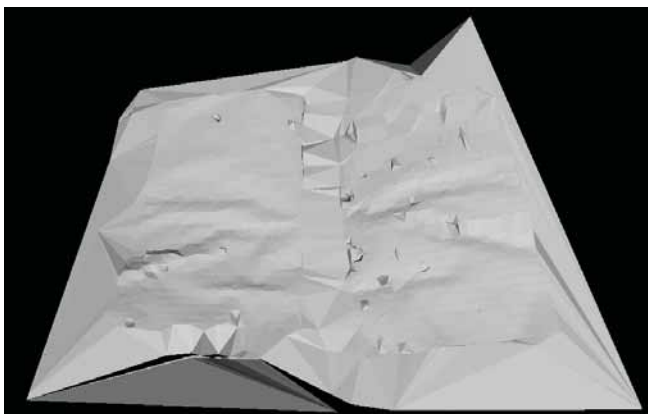


FIG. 8 - Surface 3D traduite en VRML, sans plaqué de texture pour faire ressortir les fausses corrélations

5 Applications de la numérisation 3D

Visualisation 3D

La visualisation du modèle 3D de 30 000 triangles avec les informations de textures peut se faire en temps réel avec une carte graphique standard. Cependant, l'utilisation d'un navigateur 3D généraliste (au format VRML, Metastream ou QTVR) s'avère vite fastidieux pour la manipulation du document. En effet, seules les actions de zooms et la rotation autour de l'axe défini par la gorge de reliure sont vraiment utiles. Pour la consultation de l'ensemble d'un même texte, on peut également vouloir offrir une possibilité de défilement des pages. Celle-ci peut être simulée de manière triviale par rotation d'un rectangle sur un bord (expérience *Turning the page* à la British Library [CAR 98], récemment reprise pour les manuscrits de Léonard de Vinci de l'Institut de France) ou prendre en compte de manière minimale le comportement physique d'une feuille de papier fléchie [WIT 03]. Enfin, l'objet 3D pourrait être plongé au sein d'une interface 3D de

bibliothèque numérique plus générale [CUB 01] et nous travaillons en ce sens.

Mise à plat du document

La surface d'un document relié peut être approchée par une surface cylindrique. Le dépliage de la surface s'apparente alors au redressement d'une courbe 2D. Cette méthode simple mais efficace donne de bons résultats lorsqu'on peut faire l'hypothèse d'une certaine régularité de la surface. Elle est exploitée commercialement (brevets Xerox US patent 5760925 et Minolta US patent 5084611, par ex.). Pour traiter les déformations de surface accidentelles, seules des méthodes numériques peuvent assurer un dépliage de la surface dans tous ses détails. La surface 3D est alors représentée par un maillage régulier, dont on déplace les nœuds de façon itérative en prenant soin de toujours garder une distance constante (ou modifiée le moins possible) entre eux. Une telle méthode, dite « simulation masses-ressorts » a été utilisée par [PIL 01] et [BRO 01]. Elle a l'avantage de s'appliquer aux surfaces qui peuvent être mise à plat sans déformation, telle une feuille de papier froissée, ou avec déformation, telle une feuille cloquée par l'humidité. Dans ce dernier cas, les déformations locales de la texture sont réparties « au mieux ».

Eclairage synthétique

La prise en compte du relief du document permettrait de supprimer les traits qui résultent de l'ombre des bosses, ce que le dépliage ne saurait faire (fig. 9). L'ajout de lumières synthétiques permettrait également l'élimination d'ombres pour des prises de vues délicates, comme celles d'herbiers.

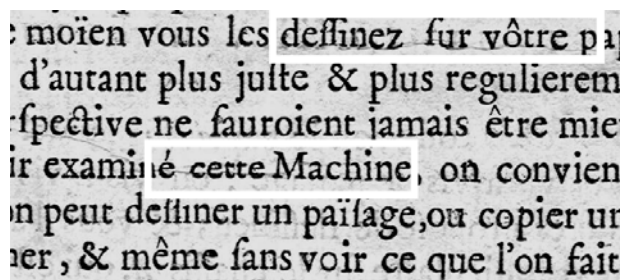


FIG. 9 : Pliures de papier et ombres résultantes. (extrait de Ramelli, Les diverses et artificioses machines, 1588
<http://cnum.cnam.fr>)

Les auteurs remercient A. Dequier (alors au service informatique du Louvre) pour les expérimentations avec le Minolta 3D1500, Minolta France pour le test du Minolta Vivid 700, S. Vilain du service Audiovisuel du CNAM pour son aide à la réalisation du banc expérimental, S. Barles (U. Paris XIII) pour le prêt d'appareil, G. Deblock (bib. centrale, CNAM) pour le prêt des ouvrages à numériser.

Références

- [BRO 00] BROWN, M.S., SEALES, W.B. « Beyond 2D images : Effective 3D imaging for library materials » Proc. of the 5th ACM conf. on Digital Libraries, san Antonio, June 2000, p. 27-36.
- [BRO 01] BROWN, M. S., SEALES, W.B. « Document Restauration Using 3D Shape : A General Deskewing Algorithm for Arbitrarily Warped Documents. » Proc. Int'l Conference On Computer Vision (ICCV'01), Vancouver, 7-14 juillet 2001, vol.2, pp.367-375.
- [BRO 03] BROWN, M. S., TSOI, D. « Correcting Common Distorsions in Camera-Imaged Library Materials » Proc. IEEE & ACM Joint Conference on Digital Libraries (JCDL'03), Houston, may 2003, p.367-368.
- [CAR 98] CARPENTER, L. SHAW, S., PRESCOTT, A. (eds) Towards the digital library : The british library initiatives for access programme. London, British lib., 1998 (p. 61 et suiv.)
- [CUB 01] CUBAUD, P., TOPOL, A., « A VRML-based user interface for an online digitalized antiquarian collection » Proc. ACM SIGGRAPH Web3D Symp, Padderborn, April 2001, pp. 51-59.
- [DEQ 00] DEQUIER A., La numérisation 3D dans le domaine culturel : Modélisation et diffusion d'œuvres d'art sur Internet. Rapport de stage DEA ESTC CNAM, Paris, sept 2000. 96 p.
- [EGE 01] EGELS, Yves, KASSER, Michel. Photogrammétrie numérique. Paris, Hermès Science Publications, 2001, 661 p.
- [HAA 03] HAAS, J.F., Numérisation tridimensionnelle de livres anciens par photogrammétrie. Rapport de stage Mastère de photogrammétrie numérique, ENSG, Noisy, sept. 2003. 81 p. En ligne sur <http://cedric.cnam.fr>.
- [PIL 01] PILU, Maurizio. « Undoing Paper Curl Distorsion Using Applicable Surfaces. » Computer Vision and Pattern Reconition (CVPR 01), Hawaii, dec 2001.
- [SCH 00] SCHMIDT, C., MOHR, R., BAUCKAGE Ch. « Evaluation of interest point detectors ». International Journal of Computer Vision 37(2):151-172, 2000.
- [WAD 95] WADA, T., UKIDA, H., MATSUYAMA, T. « Shape from shading with interreflections under proximal light source : 3D shape reconstruction of unfolded book surface from a scanner image » Proc. International Conference of Computer Vision (ICCV'95), 1995, pp 66-71.