



HAL
open science

Reconstruction 3D d'un insecte piégé dans l'ambre

Steve Bellart, Sarah-Myriam Ben Seddik, Matthew Cassidy, Jérôme Dolbecq, Noah Guitton, Jean Mélou, Yvain Quéau, François Lauze, Jean-Denis Durou

► **To cite this version:**

Steve Bellart, Sarah-Myriam Ben Seddik, Matthew Cassidy, Jérôme Dolbecq, Noah Guitton, et al.. Reconstruction 3D d'un insecte piégé dans l'ambre. Congrès Reconnaissance des Formes, Image, Apprentissage et Perception - RFIAP 2020, Sep 2020, Vannes, France. hal-02568259

HAL Id: hal-02568259

<https://normandie-univ.hal.science/hal-02568259v1>

Submitted on 8 May 2020

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.

Reconstruction 3D d'un insecte piégé dans l'ambre

Steve BELLART¹ Sarah-Myriam BEN SEDDIK¹ Matthew CASSIDY¹
Jérôme DOLBECQ¹ Noah GUITTON¹ Jean MÉLOU²
Yvain QUÉAU³ François LAUZE⁴ Jean-Denis DUROU²

¹ ENSEEIHT, Toulouse

² IRIT, UMR CNRS 5505, Toulouse

³ GREYC-CNRS, UNICAEN, ENSICAEN, Caen

⁴ DIKU, Copenhague, Danemark

Résumé

Dans cet article, nous nous attachons à reconstruire en 3D un objet opaque plongé dans un milieu réfringent. Plus précisément, le cas qui nous intéresse est celui d'un insecte piégé dans l'ambre. Cette matière, certes transparente, est colorée et réfringente, ce qui provoque des distorsions dans les images. Dans un tel cas de figure, l'estimation de la pose par structure-from-motion a été étudiée en détail, mais il n'en va pas de même de la reconstruction 3D par stéréoscopie multi-vues. Cette étude de faisabilité est validée par des tests sur images de synthèse, puis sur les images réelles d'un insecte coulé dans la résine. La qualité de ces premiers résultats est très encourageante.

Mots-clés

Reconstruction 3D, stéréoscopie multi-vues, réfraction.

Abstract

In this article, we focus on the 3D-reconstruction of an opaque object immersed in a refractive medium. More precisely, the case we are interested in is that of an insect trapped in amber. This material, although transparent, is coloured and refractive, which causes distortions in the images. The camera pose estimation by structure-from-motion has been studied in detail in this case, but the same cannot be said for 3D-reconstruction by multi-view stereo. This feasibility study is validated by tests on synthetic images, then on real images of an insect cast in the resin. The quality of these first results is very encouraging.

Keywords

3D-reconstruction, multi-view stereo, refraction.

1 Introduction

Ce travail vise à aider les entomologistes à obtenir une description 3D précise d'insectes préhistoriques dont il ne subsiste que des spécimens conservés dans l'ambre, une matière dans laquelle ils ont été piégés dans des ères géologiques anciennes. Les spécimens conservés dans l'*ambre baltique* (cf. figure 1) ont entre 40 et 50 millions d'années

et sont proches d'espèces actuelles. Des spécimens plus anciens piégés dans l'*ambre birmane*, qui ont plus de 90 millions d'années et représentent des espèces totalement disparues, constituent une source d'information inestimable aux yeux des spécialistes de l'évolution. S'il est impossible d'extraire l'insecte de sa gangue d'ambre, on peut néanmoins retailler le bloc d'ambre afin de lui donner une forme particulière, ce qui pourra faciliter la reconstruction 3D.



FIGURE 1 – Insecte préhistorique piégé dans l'ambre (\approx 40 millions d'années) provenant du Muséum d'Histoire Naturelle de Copenhague, photographié au microscope.

Cette situation est similaire à celle, très courante en médecine, d'un objet *opaque* situé à l'intérieur d'un autre objet *transparent*, les notions d'opacité et de transparence étant dépendantes du type d'ondes utilisées et de leur fréquence. L'exemple le plus connu est celui des rayons X, qui sont plus fortement atténués par les os que par les muscles ou les tissus.

Si le mot « scanner » désigne le plus souvent un instrument permettant d'imager l'intérieur du corps humain, il s'agit plus généralement de tout procédé permettant de modéliser un objet en 3D. La très grande variété d'objets à scanner explique pourquoi il existe autant de techniques de scannage 3D, sans que l'une d'elles surpasse les autres dans toutes les situations possibles.

Dans le cas qui nous intéresse, il serait envisageable d'utiliser un scanner médical pour obtenir un modèle 3D de l'insecte, mais cela serait long et coûteux. Pour donner un ordre d'idée, le Muséum d'Histoire Naturelle de Copenhague, pour qui cette étude de faisabilité a été menée, possède une collection d'environ 50 000 insectes préhistoriques conservés dans l'ambre. Il semble donc plus raisonnable d'opter pour le scannage par photogrammétrie, dont la mise en œuvre ne requiert qu'un appareil photographique et des algorithmes de reconstruction 3D.

Le recours à la photogrammétrie nécessite néanmoins de surmonter plusieurs difficultés. Tout d'abord, si le milieu réfringent que constitue l'ambre dévie la trajectoire des rayons lumineux qui le traversent, la forme du *dioptré*, c'est-à-dire de la surface de séparation entre air et ambre, influe fortement sur l'apparence de l'objet à reconstruire. En l'absence de connaissance a priori, il ne s'agit donc pas de résoudre un problème de reconstruction 3D, mais bien deux. Fort heureusement, il n'est pas interdit de tailler le bloc d'ambre pour lui donner une forme particulière, tant que cela ne nuit pas à l'intégrité de l'insecte. Nous supposons donc, pour simplifier cette étude, que le matériau réfringent a une forme parallélépipédique, ce qui facilitera l'estimation de la forme du dioptré.

Une autre difficulté à prévoir provient de la *dispersion de la lumière* à la traversée d'un milieu réfringent. Cet effet, également connu sous le nom d'*effet de prisme*, cause de sérieux problèmes aux fabricants d'objectifs photographiques, car il survient à la traversée de chaque dioptré air/verre ou verre/air, et provoque des artéfacts colorés (*aberration chromatique*) qu'il est possible de corriger grâce à des assemblages de lentilles convergentes et divergentes, ou par le biais de post-traitements numériques. Dans cette étude, nous simplifions le problème en supposant que l'*indice de réfraction* du milieu réfringent est indépendant de la fréquence de l'onde lumineuse.

Nous supposons enfin que le milieu réfringent est *homogène*, c'est-à-dire que son indice de réfraction est le même en tout point. Sans cette hypothèse, il nous faudrait tenir compte de l'*effet de mirage*, qui survient lorsque les trajectoires des rayons lumineux ne sont pas rectilignes. Or, si la prise en compte de cet effet ne présente pas de difficulté particulière du point de vue numérique, l'estimation d'un indice de réfraction variable d'un point à l'autre du bloc d'ambre constitue, en revanche, un problème difficile.

Dans le paragraphe 2, nous passons en revue les principaux travaux de vision 3D tenant compte de la réfraction. Nous rappelons le principe de la stéréoscopie multi-vues, en l'absence de dioptré (cf. paragraphe 3) puis en présence d'un dioptré (cf. paragraphe 4). Nous montrons dans le paragraphe 5 que le découpage de l'ambre en voxels permet de généraliser la résolution. Les paragraphes 6 et 7 nous permettent de valider cette étude, sur données de synthèse puis sur données réelles. Nous concluons en énumérant les difficultés qu'il nous reste à surmonter pour répondre aux besoins des entomologistes du Muséum de Copenhague.

2 État de l'art

Les rayons lumineux issus d'une source ponctuelle ont naturellement tendance à diverger. Par réfraction, une lentille permet de dévier leurs trajectoires de manière à les faire converger en un *point image*. La réfraction constitue donc la clé des instruments d'optique, à ceci près que la fabrication d'un objectif photographique nécessite d'aligner très précisément les lentilles qui le constituent, afin de limiter les effets non désirés, appelés *aberrations*.

Or, si la scène elle-même comporte des objets transparents, c'est-à-dire réfringents, cela modifie l'apparence des objets opaques, qui semblent distordus. Un certain nombre d'articles ont traité le problème de la correction de ces distorsions lorsque l'objet transparent, par exemple une vitre, est fixé à l'avant de la caméra, auquel cas les distorsions peuvent être calibrées.

À titre d'exemple, s'il est pertinent d'utiliser la photogrammétrie pour étudier la convection dans une cuve remplie d'eau, Maas montre dans [21] comment tenir compte de la réfraction à travers les deux faces du verre pour améliorer la précision des mesures. Dans [20], Luczynski et al. montrent comment corriger les images acquises par une paire de caméras sous-marines calibrées, protégées par une plaque de verre, afin de restaurer la géométrie épipolaire. Un autre exemple où la prise en compte de la réfraction est requise est celui de la reconstruction 3D des fonds marins à partir d'images aériennes [23].

A contrario, d'autres travaux, qui peuvent être regroupés sous le terme de *réfraction active*, tirent parti de la réfraction pour permettre la reconstruction 3D à partir d'une seule prise de vue, grâce au dédoublement de l'image, en utilisant soit un bi-prisme [18, 25], soit une plaque de verre rotative autour de l'axe optique [11].

Un problème faisant intervenir la réfraction, mais dont le but diffère du nôtre, est celui de l'estimation de la surface d'un objet transparent. La solution proposée par Morris et Kutulakos consiste à mettre en correspondance les points d'intérêt vus par transparence [22]. Une autre approche est celle de Ben-Ezra et Nayar [2], qui jouent sur les paramètres d'un modèle de surface de manière à reproduire l'image distordue d'un objet de géométrie connue.

Plus récemment, Chari et Sturm ont montré dans [6] comment étendre la géométrie épipolaire au cas où la caméra est séparée de la scène 3D par un dioptré plan, à l'aide d'une matrice fondamentale de taille 12×12 .

La prise en compte explicite de la réfraction dans l'estimation de la pose par *structure-from-motion* (SfM) a également été étudiée par Jordt et al. [16, 15], puis par Chadebecq et al. [4]. Dans ces différents articles, la validation est effectuée sur des images sous-marines.

Enfin, des extensions de l'algorithme PMVS (*patch-based multi-view stereo*) de Furukawa et Ponce [10] ont été proposées par Kang et al. [17] et par Chang et Chen [5], mais le procédé d'extension n'est pas détaillé et les résultats ne sont guère convaincants. Un des buts de notre travail est de combler ces lacunes.

3 Stéréoscopie multi-vues (MVS)

La stéréoscopie multi-vues (MVS, pour *multi-view stereo*) fournit une reconstruction 3D dense par pose, sous la forme d'une fonction de profondeur. Une vue d'ensemble des pratiques liées au MVS est fournie dans [9]. La figure 2 montre comment un point 3D $\mathbf{P} = [x, y, z]^T$, situé sur une surface opaque, se projette en un point \mathbf{p} de l'image. La projection centrale, de centre \mathbf{C} , sur le plan image, définie par $\mathbf{p} = \pi(\mathbf{P})$, est inversible si l'on connaît la fonction de profondeur z . Dans ce cas, et parce que la surface est opaque, il existe une bijection entre les points visibles de la scène 3D et leurs images, qui nous permet d'écrire $\mathbf{P} = \pi_z^{-1}(\mathbf{p})$.

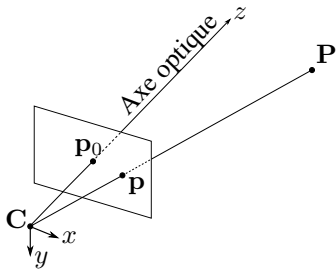


FIGURE 2 – Modèle sténopé : le point 3D $\mathbf{P} = [x, y, z]^T$ se projette en $\mathbf{p} = \pi(\mathbf{P})$, par projection centrale de centre \mathbf{C} sur le plan image.

Connaissant $t + 1$ images de la même scène 3D et les $t + 1$ poses correspondantes de la caméra, nous choisissons la première pose comme référence. Pour un point 3D \mathbf{P} visible dans toutes les images, nous notons $\mathbf{p} = \pi(\mathbf{P})$ la projection de \mathbf{P} dans l'image de référence et $\mathbf{p}_j = \pi_j(\mathbf{P})$, $j \in \{1, \dots, t\}$, ses projections dans les t autres images, dites *images témoins*. L'hypothèse lambertienne s'écrit :

$$I_j(\mathbf{p}_j) = I(\mathbf{p}), \quad j \in \{1, \dots, t\} \quad (1)$$

où I_j et I désignent les fonctions de niveau de gris de la $j^{\text{ème}}$ image témoin et de l'image de référence. Soit Ω l'ensemble des points de l'image de référence visibles dans toutes les images témoins. En un point $\mathbf{p} \in \Omega$:

$$I_j \circ \pi_j \circ \pi_z^{-1}(\mathbf{p}) = I(\mathbf{p}), \quad j \in \{1, \dots, t\} \quad (2)$$

La méthode MVS consiste à chercher la fonction de profondeur z de la pose de référence qui maximise la *cohérence photométrique* avec les t images témoins. Plutôt que de chercher une solution exacte des équations (2), il est préférable de résoudre ce système en moindres carrés :

$$\hat{z}(\mathbf{p}) = \underset{z(\mathbf{p}) \in \mathbb{R}}{\operatorname{argmin}} \sum_{j=1}^t [I_j \circ \pi_j \circ \pi_z^{-1}(\mathbf{p}) - I(\mathbf{p})]^2 \quad (3)$$

Dans (3), la cohérence photométrique se réduit à la comparaison en moindres carrés de deux niveaux de gris. En pratique, cette comparaison se fait entre les voisinages de deux pixels, à l'aide d'un estimateur robuste (cf. le panorama très complet présenté dans [9, chapitre 2]).

La fonction à minimiser peut se révéler non linéaire, non dérivable et/ou non convexe, ce qui rend l'optimisation potentiellement difficile. Pour cette raison, la minimisation est généralement effectuée par une recherche exhaustive (*brute-force*) dans une liste de valeurs prédéfinies de la profondeur. Cette stratégie, apparemment simpliste, s'avère très efficace pour la reconstruction 3D de scènes dont la surface est suffisamment texturée [12].

La figure 3 montre un exemple de reconstruction 3D dense obtenue avec la suite logicielle libre AliceVision [1]. Une fois qu'une fonction de profondeur a été estimée pour chaque pose, un modèle 3D complet est obtenu par fusion des fonctions de profondeur [3, 14], sous la forme d'un maillage 3D. Une texture peut ensuite être plaquée sur ce maillage [19]. La qualité d'un maillage 3D tel que celui de la figure 3-b montre que les pipelines de photogrammétrie fondés sur le MVS sont devenus très performants.

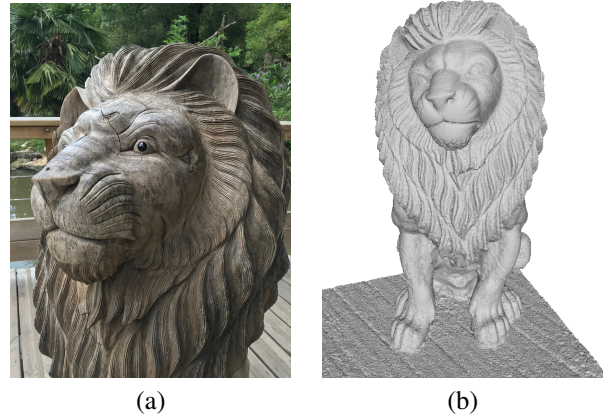


FIGURE 3 – Exemple de reconstruction 3D dense obtenue par MVS (78 images) : (a) image de référence ; (b) maillage 3D obtenu avec la suite logicielle AliceVision [1].

4 MVS en présence d'un dioptre

La présence d'un dioptre entre l'objet à reconstruire et la caméra a pour effet de déformer les images. Entre autres conséquences, la géométrie épipolaire d'une paire d'images est perdue. Au lieu de correspondre à une *droite épipolaire* dans la deuxième image, un point de la première image correspond maintenant à une courbe, dont la forme dépend évidemment de la position du dioptre, de sa géométrie, et des indices de réfraction mis en jeu.

Chari et Sturm ont montré dans [6] que l'on pouvait généraliser le formalisme matriciel de la géométrie épipolaire avec une matrice fondamentale de taille 12×12 , mais la géométrie épipolaire sert avant tout à l'estimation de la pose par *structure-from-motion* (SfM). Or, dans cette étude de faisabilité, nous ne nous préoccupons pas d'adapter le SfM à la présence d'un dioptre, vu que ce problème a déjà été étudié dans plusieurs articles [15, 4]. Nous nous attachons en revanche à la généralisation du MVS.

Cette généralisation revient à résoudre un problème similaire à (3) en chaque point \mathbf{p} de l'image de référence, en procédant à quelques adaptations :

- Pour calculer $\mathbf{P} = \pi_z^{-1}(\mathbf{p})$ en présence d'un dioptre, il faut tracer le rayon issu de \mathbf{C} et passant par \mathbf{p} . Alors que ce rayon est rectiligne en l'absence de dioptre, il forme maintenant une ligne brisée.
- A contrario, le calcul de $\mathbf{p}_j = \pi_j(\mathbf{P})$ est nettement plus difficile s'il y a un dioptre, car la fonction π_j n'est plus une projection centrale sur l'image témoin I_j . Nous verrons un peu plus loin que ce calcul nécessite de résoudre un problème de *plus court chemin optique*.
- Comme le montre l'exemple de la figure 11-b, un même point 3D \mathbf{P} peut se projeter en plusieurs points de l'image témoin I_j (ici, deux), qui correspondent à autant de minima locaux du chemin optique entre \mathbf{P} et \mathbf{C}_j , selon le *principe de Fermat*.

La principale difficulté de la réfraction consiste donc à calculer $\mathbf{p}_j = \pi_j(\mathbf{P})$. Si l'on note i l'angle que le rayon lumineux incident fait avec la normale au dioptre, dans un milieu d'indice de réfraction n_1 (l'ambre), la première loi de Descartes de la réfraction nous dit que le rayon réfracté se trouve dans le *plan d'incidence*, défini par le rayon incident et la normale au point de réfraction $\bar{\mathbf{P}}$ (cf. figure 4). Pour cette raison, il est possible de représenter le phénomène de la réfraction dans un plan.

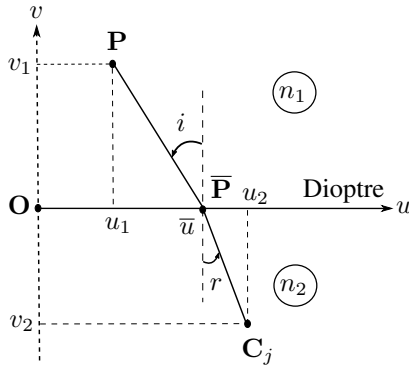


FIGURE 4 – Deuxième loi de Descartes de la réfraction.

Quant à la deuxième loi de Descartes de la réfraction, elle nous dit que l'angle r entre le rayon réfracté et la normale, dans un milieu d'indice n_2 (l'air), vérifie l'égalité :

$$n_1 \sin i = n_2 \sin r \quad (4)$$

Intéressons-nous au cas d'un dioptre plan, qui fait partie de nos hypothèses, et élevons au carré les deux membres de l'égalité (4). Avec les notations de la figure 4, il vient :

$$n_1^2 \frac{(u_1 - \bar{u})^2}{(u_1 - \bar{u})^2 + v_1^2} = n_2^2 \frac{(u_2 - \bar{u})^2}{(u_2 - \bar{u})^2 + v_2^2} \quad (5)$$

La détermination du point de réfraction $\bar{\mathbf{P}}$ nécessite donc

de résoudre une équation polynomiale de degré 4, d'inconnue l'abscisse \bar{u} de $\bar{\mathbf{P}}$, qui se déduit facilement de (5) :

$$a_4 \bar{u}^4 + a_3 \bar{u}^3 + a_2 \bar{u}^2 + a_1 \bar{u} + a_0 = 0 \quad (6)$$

et dont les coefficients s'écrivent, en notant $\alpha = n_2/n_1$:

$$\begin{cases} a_4 = \alpha^2 - 1 \\ a_3 = -2 a_4 (u_1 + u_2) \\ a_2 = a_4 (u_1^2 + u_2^2 + 4 u_1 u_2) + \alpha^2 v_1^2 - v_2^2 \\ a_1 = a_3 u_1 u_2 - 2 \alpha^2 v_1^2 u_2 + 2 v_2^2 u_1 \\ a_0 = a_4 u_1^2 u_2^2 + \alpha^2 v_1^2 u_2^2 - v_2^2 u_1^2 \end{cases} \quad (7)$$

Par conséquent, le calcul de $\mathbf{p}_j = \pi_j(\mathbf{P})$ doit être effectué en deux temps. Il faut d'abord résoudre l'équation (6) pour déterminer le point de réfraction $\bar{\mathbf{P}}$, ce qui revient à chercher le plus court chemin optique entre \mathbf{P} et \mathbf{C}_j , puis effectuer ensuite la projection centrale, de centre \mathbf{C}_j , de $\bar{\mathbf{P}}$ sur l'image témoin I_j .

Il est entendu que la présence d'un dioptre entre la caméra et un objet 3D déforme l'image de ce dernier, mais nous sommes en droit de nous demander si la transformation π_j conserve l'alignement des points, c'est-à-dire si l'image d'une droite par π_j est une droite, comme cela est le cas en l'absence de dioptre. Si tel est le cas, alors, pour résoudre le MVS en présence d'un dioptre, il suffira de chercher le point $\hat{\mathbf{p}}_j$ correspondant à un point \mathbf{p} de l'image de référence le long d'une droite de l'image témoin I_j , cette droite pouvant être déterminée en calculant les images \mathbf{p}_j^1 et \mathbf{p}_j^2 de n'importe quelle paire de points \mathbf{P}^1 et \mathbf{P}^2 situés sur le rayon lumineux (\mathbf{Cp}) après réfraction.

Lorsqu'une règle est plongée dans l'eau, une observation rapide pourrait laisser penser que l'image de la partie immergée, qui semble « cassée », est néanmoins rectiligne. Or, non seulement la géométrie épipolaire n'est pas préservée en présence d'un dioptre, mais il a été montré par Chari et Sturm dans [6] que, dans le cas d'un dioptre plan, à un point \mathbf{p} de l'image de référence correspond, dans une image témoin, non pas une droite, mais une courbe dont l'équation se déduit d'une matrice de taille 12×12 .

Cette équation n'est malheureusement guère utile pour résoudre le problème (3) en présence d'un dioptre. En effet, nous avons déjà dit, dans le paragraphe 3, que ce problème était potentiellement non linéaire, non dérivable et/ou non convexe, et que, pour ces différentes raisons, il était recommandé de mener la résolution par une recherche exhaustive (*brute-force*) dans une liste de valeurs prédéfinies de la profondeur z du point 3D \mathbf{P} conjugué de \mathbf{p} (cf. figure 5).

Une première stratégie de résolution du MVS en présence d'un dioptre consiste donc à procéder exactement comme en l'absence de dioptre, si ce n'est que la complexité de la résolution est nettement accrue. En effet, pour chaque pixel \mathbf{p} de l'image de référence, et pour chaque valeur prédéfinie de la profondeur z de \mathbf{P} , le calcul de l'argument du problème (3) nécessite de résoudre t équations polynomiales de degré 4 telles que (6), alors qu'en l'absence de dioptre, il suffisait de projeter \mathbf{P} sur les t images témoins.

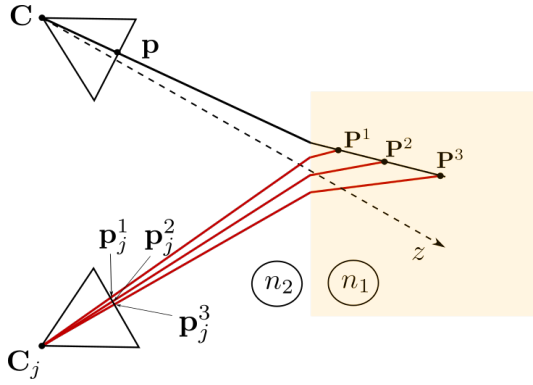


FIGURE 5 – Résolution du MVS en présence d'un dioptre. Pour chacun des points P^k , $k \in \{1, 2, 3, \dots\}$, situés sur le rayon réfracté issu de C et passant par un point p de l'image de référence, il est nécessaire de trouver son image p_j^k dans l'image témoin I_j . Le point P^k retenu est celui dont l'image p_j^k est la plus similaire à p .

Pour donner un ordre de grandeur de la complexité de cette stratégie de résolution, le test de la figure 12-a est effectué avec $t = 4$ images témoins. L'image de référence comporte environ $1,5 \cdot 10^5$ pixels. Sachant que la profondeur z à l'intérieur de l'ambre a été discrétisée en 200 valeurs, la résolution complète du problème (3) nécessite de résoudre de l'ordre de 10^8 équations de degré 4, ce qui prend environ une heure de temps de calcul sur un PC standard, en utilisant soit une méthode de résolution analytique, soit une méthode de résolution par dichotomie.

Outre qu'elle est très gourmande en ressources de calcul, cette stratégie de résolution présente un autre défaut bien plus gênant : elle est difficilement généralisable à un dioptre de forme quelconque. Or, ces deux défauts peuvent être surmontés en optant pour une stratégie de résolution discrète, qui consiste à découper l'ambre en voxels.

5 Découpage de l'ambre en voxels

Le principe de Huygens-Fresnel permet de prédire, de proche en proche, les surfaces d'onde d'un flux lumineux, sachant que les normales à ces surfaces sont tangentes aux rayons lumineux. Une version discrète du calcul des surfaces d'onde est fournie par l'algorithme de Dijkstra [7], qui permet de trouver le plus court chemin entre deux nœuds dans un graphe. Si r est un nœud d'un graphe \mathcal{G} connexe, pondéré et non orienté, cet algorithme construit un sous-graphe de \mathcal{G} sous la forme d'un arbre \mathcal{A} de racine r , tel que toute branche menant de r à un nœud n définit le chemin de poids cumulé minimal entre r et n .

Le principe de Fermat nous dit que la lumière se propage d'un point à un autre le long d'une trajectoire dont la durée est localement minimale. Pour calculer le trajet des rayons lumineux, nous découpons donc l'ambre en c voxels cubiques. Les centres v_l des voxels, $l \in \{1, \dots, c\}$, constituent les nœuds d'un graphe \mathcal{G} dont les arêtes découlent du système de voisinage « des 26 plus proches voisins ». Le

poids p_{v_1, v_2} de l'arête entre nœuds voisins v_1 et v_2 est égal au temps de parcours de la lumière le long de cette arête :

$$p_{v_1, v_2} = n_1 \|v_1 - v_2\|/c \quad (8)$$

où n_1 désigne l'indice de réfraction de l'ambre et c la vitesse de la lumière dans le vide. Pour connaître le temps de parcours de la lumière entre un voxel v_l et le centre de projection C_j , $j \in \{1, \dots, t\}$, nous construisons un nouveau graphe \mathcal{G}_j obtenu en ajoutant à \mathcal{G} le nœud $r_j \equiv C_j$. Seuls les nœuds v_l de \mathcal{G} situés sur la surface de l'ambre et faisant face à la $j^{\text{ème}}$ caméra témoin sont reliés au nœud r_j par une arête dont le poids p_{v_l, r_j} est égal au temps de parcours de la lumière entre v_l et r_j :

$$p_{v_l, r_j} = n_2 \|v_l - r_j\|/c \quad (9)$$

où l'indice de réfraction de l'air n_2 est assimilable à 1. La figure 6 montre l'exemple d'un tel graphe $\mathcal{G}_j = \mathcal{G} \cup \{r_j\}$.

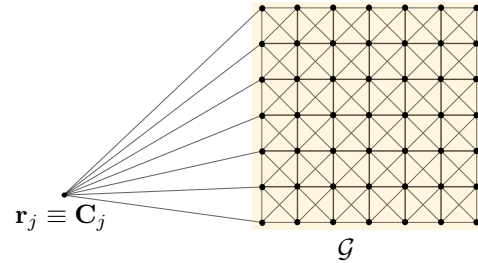


FIGURE 6 – Exemple de graphe $\mathcal{G}_j = \mathcal{G} \cup \{r_j\}$ construit par découpage de l'ambre en voxels, en utilisant le système de voisinage « des 26 plus proches voisins » (sur ce schéma en 2D, seuls les 8 plus proches voisins ont été représentés), complété par le nœud $r_j \equiv C_j$.

En lançant l'algorithme de Dijkstra sur le graphe \mathcal{G}_j de la figure 6, depuis le nœud r_j , nous obtenons un arbre \mathcal{A}_j de racine r_j , représenté en rouge sur la figure 7, dont les branches constituent une approximation discrète du trajet des rayons lumineux entre tous les centres de voxels v_l , $l \in \{1, \dots, c\}$, et le centre de projection C_j .

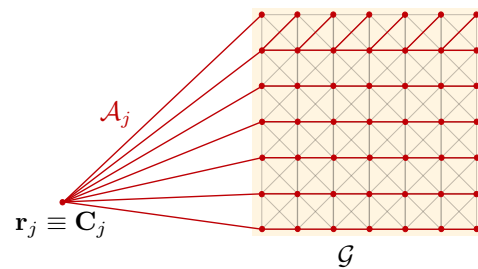


FIGURE 7 – En partant du nœud r_j du graphe \mathcal{G}_j de la figure 6, l'algorithme de Dijkstra construit un arbre \mathcal{A}_j , représenté en rouge, dont les branches indiquent le trajet des rayons lumineux entre les centres de voxels v_l et le centre de projection C_j .

Pour chaque point \mathbf{P}^k correspondant à une valeur prédéfinie de la profondeur z (cf. figure 5), il suffit maintenant de suivre le trajet du rayon lumineux depuis le voxel contenant \mathbf{P}^k , jusqu'à « couper » l'image témoin I_j . Rappelons néanmoins que l'algorithme de Dijkstra ne fournit qu'une approximation discrète du trajet réel de la lumière, qui est d'autant plus précise que les voxels sont plus petits.

En réalité, il est inutile de tester toutes les valeurs prédéfinies de la profondeur z . En effet, certains des points \mathbf{P}^k se projettent dans l'image témoin I_j en dehors de la silhouette de l'objet piégé dans l'ambre, voire même en dehors de l'image. Une fois l'arbre \mathcal{A}_j construit, on peut donc se contenter de conserver un sous-arbre $\bar{\mathcal{A}}_j$ obtenu en supprimant les branches coupant l'image I_j en dehors de cette silhouette (cf. figure 8), et attribuer à chaque voxel $\mathbf{v}_l \in \mathcal{G}$ un booléen b_l indiquant si ce voxel se projette ou non à l'intérieur de la silhouette de l'objet à reconstruire en 3D.

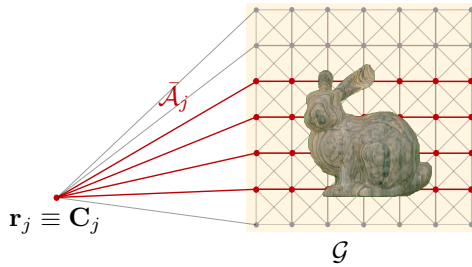


FIGURE 8 – Parmi les branches de l'arbre \mathcal{A}_j (cf. figure 7), celles qui coupent l'image témoin I_j en dehors de la silhouette de l'objet piégé dans l'ambre sont supprimées. Les branches résiduelles constituent un sous-arbre $\bar{\mathcal{A}}_j$ de \mathcal{A}_j .

Le problème (3) devient alors très facile à résoudre. Pour chaque pixel \mathbf{p} de l'image de référence, et pour chaque point \mathbf{P}^k correspondant à une valeur prédéfinie de la profondeur z , il faut trouver le voxel \mathbf{v}_l contenant \mathbf{P}^k , et ne projeter \mathbf{P}^k dans l'image témoin I_j que si $b_l = 1$.

Rappelons enfin qu'un rayon lumineux emprunte un trajet de durée *localement* minimale. Pour traiter le cas de figure où l'image est dédoublée (cf. l'exemple de la figure 11-b), il faut donc lancer l'algorithme de Dijkstra deux fois, en reliant \mathbf{r}_j aux nœuds d'une seule face (cf. figure 9).

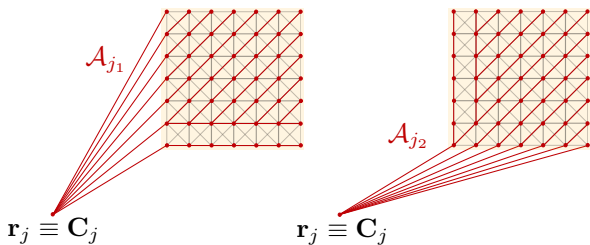


FIGURE 9 – Dans le cas où deux faces sont visibles depuis le point \mathbf{C}_j , il convient de lancer l'algorithme de Dijkstra sur deux graphes différents, afin de modéliser le dédoublement de l'image illustré sur l'exemple de la figure 11-b.

Effectivement, en reliant \mathbf{r}_j aux nœuds des deux faces visibles, l'algorithme de Dijkstra calcule un seul rayon par voxel (cf. figure 10). Pour un voxel donné \mathbf{v}_l , ce rayon correspond, entre les deux rayons de la figure 9, au *minimum absolu* du temps de parcours de la lumière entre \mathbf{v}_l et \mathbf{C}_j .

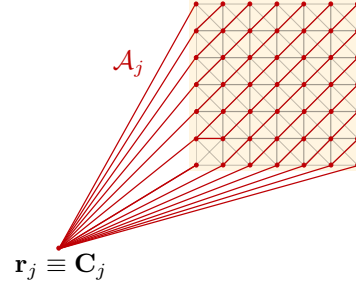


FIGURE 10 – Si \mathbf{r}_j est relié aux nœuds des deux faces visibles, l'algorithme de Dijkstra détermine le rayon de la figure 9 qui correspond au temps de parcours minimal.

6 Résultats sur données de synthèse

Nous simulons des images du modèle 3D *Stanford Bunny* coulé dans un bloc de verre incolore de forme parallélépipédique, grâce aux fonctionnalités de rendu par lancer de rayons du logiciel libre *Blender* : une texture de bois est plaquée sur le lapin ; la valeur 1,5 correspond à un indice de réfraction réaliste pour le verre ; enfin, le bloc de verre est placé devant un fond vert afin de faciliter la segmentation fond/forme. La figure 11 montre deux des vingt-neuf images du lapin de Stanford ainsi calculées. Nous avons déjà signalé que, sous certains angles, l'image du lapin pouvait être dédoublée, du fait de la réfraction.

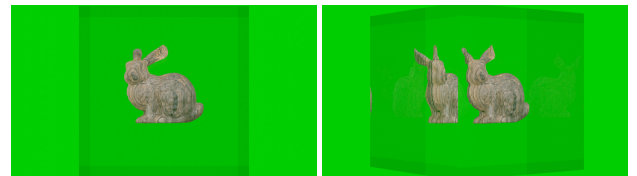


FIGURE 11 – Deux images (parmi vingt-neuf) du *lapin de Stanford* plongé dans un bloc de verre de forme parallélépipédique, simulées avec le logiciel libre *Blender*. Sous certains angles, l'image du lapin est dédoublée.

Avant d'expérimenter la méthode de reconstruction 3D par MVS décrite dans le paragraphe 5, nous procédons préalablement à deux tests. La figure 12 montre deux reconstructions 3D effectuées à partir de $t + 1 = 5$ images prises à travers la face avant du parallélépipède. Alors que le résultat de la figure 12-a tient compte de la réfraction, cela n'est pas le cas du résultat de la figure 12-b. Une comparaison « à l'œil nu » montre que ce dernier résultat est excessivement aplati. Dans les deux cas, l'allure « poussiéreuse » du nuage vient de ce que, sans image prise de côté, il est difficile de « contrôler » la profondeur des points 3D.



FIGURE 12 – Reconstruction 3D par MVS avec (à gauche) ou sans (à droite) prise en compte de la réfraction. Le deuxième nuage de points 3D est excessivement aplati.

Un deuxième test, illustré sur la figure 13, montre qu’il est effectivement possible d’utiliser le dédoublement de l’image (par réfraction à travers deux faces du bloc de verre) pour mettre en œuvre la *stéréoscopie mono-vue* [18, 25]. Bien qu’elle fournisse sur cet exemple un résultat très imprécis, cette technique diffère des autres techniques de reconstruction 3D mono-vue telles que le *shape-from-shading* [8], dans la mesure où elle utilise la *triangulation*.

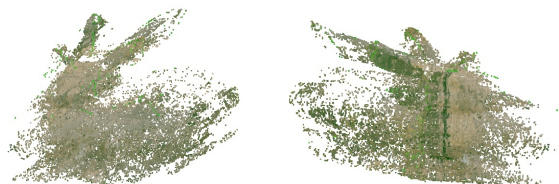


FIGURE 13 – Reconstruction 3D par stéréoscopie mono-vue : deux vues du nuage de points 3D obtenu à partir d’une seule image, en l’occurrence l’image de la figure 11-b.

La figure 14 montre le nuage de points 3D obtenu à partir de vingt-quatre images telles que celles de la figure 11, par fusion de quatre nuages correspondant aux quatre faces verticales du parallélépipède. Chacun de ces nuages est estimé à partir de six images : les cinq images prises à travers une même face sont complétées par une image prise de côté. En éliminant les points 3D qui se reprojettent en dehors de la silhouette du lapin dans cette image témoin supplémentaire, il est possible de gommer la « poussière » des nuages de la figure 12.

Les valeurs de la RMSE correspondant à ces quatre nuages de points 3D, dont la fusion constitue le nuage complet de la figure 14, sont compilées dans la table 1.

Face	Avant	Arrière	Gauche	Droite
RMSE (<i>mm</i>)	0,28	0,29	0,34	0,38

TABLE 1 – RMSE de chacun des nuages de points 3D obtenus à partir de six images : cinq images prises à travers une même face, complétées par une image prise de côté. L’unité est le millimètre, sachant que la profondeur du modèle 3D se situe entre 90 *mm* et 110 *mm*.

Ces très faibles valeurs de la RMSE, comparées à l’échelle du modèle 3D de référence (de l’ordre de 20 *mm*), suffisent à valider notre approche, qui n’a d’autre prétention que celle d’une étude de faisabilité.

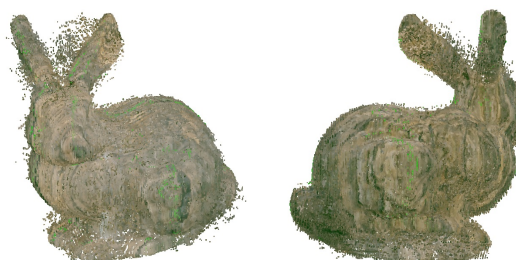


FIGURE 14 – Reconstruction 3D par MVS : deux vues du nuage de points 3D obtenu à partir de vingt-quatre images telles que celles de la figure 11. L’espace entre les deux oreilles du lapin reste « poussiéreux » à cause du fait que cet espace n’est visible que dans un petit nombre d’images.

7 Résultats sur données réelles

Étant donné que nous ne disposons pas encore d’un jeu d’images d’un même échantillon du Muséum d’Histoire Naturelle de Copenhague, nous nous sommes contentés d’un insecte contemporain provenant d’une collection privée. Il s’agit d’une sauterelle coulée dans un bloc de résine de forme parallélépipédique. Les photographies ont été prises avec un appareil de type *bridge*, dans des conditions opératoires peu contrôlées, en plaçant la scène devant un fond bleu pour faciliter la segmentation (cf. figure 15).



FIGURE 15 – Deux images réelles (parmi six) d’une sauterelle coulée dans un bloc de résine de forme parallélépipédique. La petite grille collée sur chaque face permet d’estimer la pose de la caméra et la position du dioptré.

La figure 16 montre la reconstruction 3D obtenue à partir de cinq images prises à travers la face avant, en utilisant la suite logicielle AliceVision [1]. Les poses sont estimées de manière cohérente entre elles, mais c’est le bloc de résine, et non l’insecte, qui est reconstruit en 3D, ce qui s’explique en partie par le fait qu’aucune segmentation n’a été effectuée au préalable. Néanmoins, même après segmentation, la réfraction n’est pas prise en compte par ce logiciel.

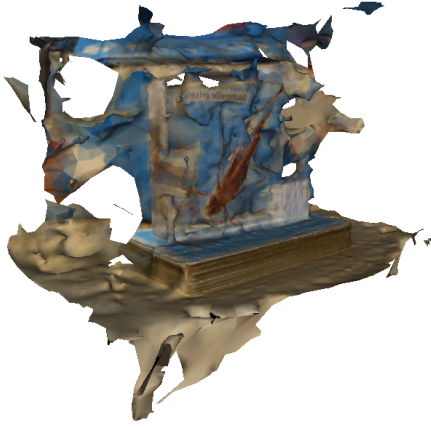


FIGURE 16 – Reconstruction 3D obtenue avec AliceVision [1], à partir de cinq images réelles prises à travers la face avant, l’image de droite de la figure 15 n’étant pas jugée par AliceVision suffisamment cohérente avec les cinq autres.

La principale difficulté de mise en œuvre de notre méthode de reconstruction 3D consiste à estimer les poses de la caméra. Nous aurions pu utiliser les poses estimées par AliceVision, ou recourir aux méthodes d’estimation de la pose par *refractive structure-from-motion* [4], mais il nous faut également connaître la position des faces du bloc de résine. Nous nous sommes donc contentés de coller une petite grille sur chaque face, ce qui nous a permis d’obtenir des estimations cohérentes entre elles.

La figure 17 montre deux vues de la reconstruction 3D de la sauterelle obtenue à partir des cinq images prises à travers la face avant, plus l’image prise de côté qui permet de « contrôler » la profondeur. Nous avons fixé l’indice de réfraction de la résine à la valeur plausible $n_1 = 1,6$, mais sans la valider avec un réfractomètre. Bien qu’imparfait, ce premier résultat nous paraît encourageant. En particulier, les pattes et les antennes reconstruites semblent réalistes, ce qui est sans doute un critère important pour déterminer si nous répondons ou non aux besoins des entomologistes.

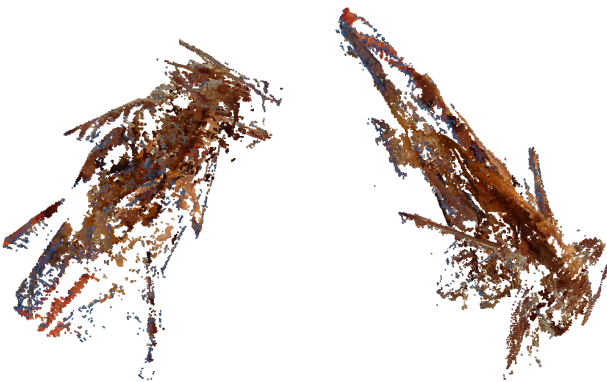


FIGURE 17 – Reconstruction 3D par MVS : deux vues du nuage de points 3D obtenu à partir de cinq images prises à travers la face avant, plus une prise de côté (cf. figure 15).

8 Conclusion et perspectives

Dans cet article, nous nous sommes attachés à reconstruire en 3D un insecte piégé dans l’ambre, en adaptant la stéréoscopie multi-vues au cas où l’objet d’intérêt, en l’occurrence l’insecte, est vu à travers un dioptre. Sachant que la réfraction déforme les images, il est nécessaire d’en tenir compte pour modéliser le chemin des rayons lumineux. Plus particulièrement, nous avons proposé une méthode de résolution discrète utilisant l’algorithme de Dijkstra. Nos premiers résultats sur données réelles sont encourageants, même s’il reste de nombreux obstacles à franchir avant que cette étude de faisabilité débouche sur une méthode opérationnelle utilisable par les entomologistes.

Parmi les tâches à mener en priorité, il est nécessaire d’améliorer l’estimation des poses de la caméra, afin de permettre la reconstruction 3D d’un grand nombre d’échantillons. Deux pistes peuvent être envisagées. La première consiste à estimer la pose « en ligne », pour chaque nouvelle acquisition de données. Nous devons alors nous tourner vers les méthodes d’estimation de la pose par *refractive structure-from-motion*, qui se fondent sur l’appariement entre points d’intérêt vus par transparence [6]. La deuxième possibilité, plus simple, consiste à standardiser le processus de prise de vues, de telle sorte que l’estimation préalable des poses puisse être menée « hors ligne », préalablement à la collecte des données.

Par ailleurs, il est réducteur de supposer que le bloc d’ambre doit être forcément parallélépipédique, et donc à faces planes. Une autre amélioration de notre méthode consistera donc à estimer la forme du bloc d’ambre par *shape-from-silhouettes* [13]. En veillant à placer le bloc d’ambre devant un fond coloré, comme nous l’avons déjà fait (cf. figure 15), cette technique permet d’estimer le volume de n’importe quel objet convexe, pour peu qu’un nombre suffisant d’images soient utilisées.

Une autre caractéristique essentielle du milieu réfringent, dont l’estimation précise améliorerait les résultats, est son indice de réfraction, que nous avons fixé arbitrairement à la valeur 1,6 pour le test de la figure 17. Rappelons que le problème serait bien plus complexe si le matériau n’était pas homogène, du point de vue de l’indice de réfraction.

Enfin, à plus long terme, nous envisageons d’adapter à la réfraction les méthodes de reconstruction 3D photométriques que sont le *shape-from-shading* [8] et la stéréophotométrie [24]. À notre connaissance, ces problèmes ont été très peu abordés dans la littérature, et aucune solution opérationnelle n’a été proposée. Outre qu’il sera nécessaire de modéliser les phénomènes d’atténuation de la lumière à l’intérieur du milieu réfringent, selon la *loi de Beer-Lambert*, il faudra également tenir compte de la couleur de l’ambre, c’est-à-dire considérer que cette atténuation dépend de la longueur d’onde. Cela nous rappelle que, si nous avons délibérément ignoré la dispersion de la lumière, ce phénomène devra lui aussi être pris en compte dans le contexte de la reconstruction 3D photométrique.

Références

- [1] AliceVision. <https://alicevision.org/>.
- [2] M. BEN-EZRA et S. K. NAYAR : What does motion reveal about transparency? *In Proceedings of the IEEE International Conference on Computer Vision*, volume 2, pages 1025–1032, 2003.
- [3] Y. BOYKOV et V. KOLMOGOROV : An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9): 1124–1137, 2004.
- [4] F. CHADEBECQ, F. VASCONCELOS, R. LACHER, E. MANEAS, A. DESJARDINS, S. OURSELIN, T. VERCAUTEREN et D. STOYANOV : Refractive Two-View Reconstruction for Underwater 3D Vision. *International Journal of Computer Vision*, 2019.
- [5] Y. J. CHANG et T. CHEN : Multi-View 3D Reconstruction for Scenes under the Refractive Plane with Known Vertical Direction. *In Proceedings of the IEEE International Conference on Computer Vision*, pages 351–358, 2011.
- [6] V. CHARI et P. STURM : Multiple-View Geometry of the Refractive Plane. *In Proceedings of the British Machine Vision Conference*, pages 1–11, 2009.
- [7] E. W. DIJKSTRA : A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.
- [8] J.-D. DUROU, M. FALCONE et M. SAGONA : Numerical Methods for Shape-from-shading: A New Survey with Benchmarks. *Computer Vision and Image Understanding*, 109(1):22–43, 2008.
- [9] Y. FURUKAWA et C. HERNÁNDEZ : Multi-View Stereo: A Tutorial. *Foundations and Trends in Computer Graphics and Vision*, 9(1–2):1–148, 2015.
- [10] Y. FURUKAWA et J. PONCE : Accurate, Dense, and Robust Multiview Stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1362–1376, 2010.
- [11] C. GAO et N. AHUJA : A Refractive Camera for Acquiring Stereo and Super-resolution Images. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2316–2323, 2006.
- [12] M. GOESELE, B. CURLESS et S. M. SEITZ : Multi-View Stereo Revisited. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2402–2409, 2006.
- [13] C. HERNÁNDEZ et F. SCHMITT : Silhouette and stereo fusion for 3D object modeling. *Computer Vision and Image Understanding*, 96(3):367–392, 2004.
- [14] M. JANCOSEK et T. PAJDLA : Exploiting Visibility Information in Surface Reconstruction to Preserve Weakly Supported Surfaces. *International Scholarly Research Notices*, 2014:1–20, 2014.
- [15] A. JORDT, K. KÖSER et R. KOCH : Refractive 3D reconstruction on underwater images. *Methods in Oceanography*, 15–16:90–113, 2016.
- [16] A. JORDT-SEDLAZECK, D. JUNG et R. KOCH : Refractive Plane Sweep for Underwater Images. *In Proceedings of the German Conference on Pattern Recognition*, pages 333–342, 2013.
- [17] L. KANG, L. WU et Y. H. YANG : Two-View Underwater Structure and Motion for Cameras under Flat Refractive Interfaces. *In Proceedings of the European Conference on Computer Vision*, volume 7575 de *Lecture Notes in Computer Science*, pages 303–316, 2012.
- [18] D. H. LEE, I.-S. KWEON et R. CIPOLLA : A biprism-stereo camera system. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, page 87, 1999.
- [19] B. LEVY, S. PETITJEAN, N. RAY et J. MAILLOT : Least Squares Conformal Maps for Automatic Texture Atlas Generation. *ACM Transactions on Graphics*, 21(3):362–371, 2002.
- [20] T. LUCZYNSKI, M. PFINGSTHORN et A. BIRK : Image rectification with the pinax camera model in underwater stereo systems with verged cameras. *In OCEANS 2017*, pages 1–7, 2017.
- [21] H. G. MAAS : New developments in multimedia photogrammetry. *In Optical 3-D Measurement Techniques III*, 1995.
- [22] N. J. W. MORRIS et K. N. KUTULAKOS : Dynamic Refraction Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1518–1531, 2011.
- [23] T. MURASE, M. TANAKA, T. TANI, Y. MIYASHITA, N. OHKAWA, S. ISHIGURO, Y. SUZUKI, H. KAYANNE et H. YAMANO : A Photogrammetric Correction Procedure for Light Refraction Effects at a Two-Medium Boundary. *Photogrammetric Engineering and Remote Sensing*, 9(8):1129–1136, 2008.
- [24] R. J. WOODHAM : Photometric Method For Determining Surface Orientation From Multiple Images. *Optical Engineering*, 19(1):134–144, 1980.
- [25] A. YAMASHITA, Y. SHIRANE et T. KANEKO : Monocular Underwater Stereo – 3D Measurement Using Difference of Appearance Depending on Optical Paths. *In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3652–3657, 2010.