

Reconstruction 3D passive par fusion de données polarimétriques et radiométriques

E. JOUBERT, P. MICHÉ, R. DEBRIE

La3i / LCIA, INSA de ROUEN, F-76131 Mont Saint Aignan Cedex, FRANCE

Tél. : (33) 35 52 84 07 - Fax : (33) 35 52 84 83 - Email : joubert@aigle.insa-rouen.fr

Résumé : Cette communication présente une nouvelle approche de la reconstruction de surface en trois dimensions, selon une configuration passive. Notre méthode propose l'utilisation de l'analyse de l'état de polarisation des faisceaux lumineux renvoyés par la scène pour obtenir une information d'orientation. Ce type d'information n'étant pas autonome, les données recueillies sont fusionnées avec des données issues d'un système radiométrique pour achever la reconstruction. Des résultats sont présentés pour quelques scènes réelles.

Abstract : This paper presents a new approach for passive 3D surface reconstruction. In fact, our method proposes to use the analysis of polarization state of light beams issued from the scene, so as to provide an orientation information. Since this type of information is not autonomous, the obtained data are fused with radiometric data provided by a radiometric system, to achieve the reconstruction. Results for some real scenes are presented.

1) Introduction

Différentes techniques envisagent le problème de la reconstruction de surfaces en trois dimensions. Parmi les plus répandues, les méthodes dites passives sont relativement intéressantes pour un certain nombre d'applications car elles n'exigent pas de conditions de prises de vue spécifiques, à l'inverse des méthodes actives qui sont très précises mais qui sont destinées à des applications bien délimitées [1][2]. Les différentes méthodes passives se distinguent surtout par le nombre de points de vue mis en oeuvre.

Dans cette classification, les méthodes fondées sur la stéréovision sont les seules qui peuvent proposer une information de distance absolue [3][4] ; les méthodes qui n'utilisent qu'un seul point de vue ne procurent, par propriété géométrique, qu'une information de forme [5].

Néanmoins, les méthodes fondées sur la stéréovision passive n'apportent en général qu'une information tridimensionnelle limitée aux contours des objets observés[3]. Différentes solutions mathématiques permettent de retrouver les points intermédiaires et ainsi de proposer des cartes de profondeur denses, moyennant un certain nombre d'hypothèses concernant la nature des surfaces observées ou leur texture [4].

La méthode que nous proposons représente une alternative intéressante par le fait que la reconstruction mise en oeuvre dépend d'une mesure physique concrète, en l'occurrence l'état de polarisation des faisceaux lumineux réfléchis par les objets de la scène. En effet, ce paramètre de la lumière contient une information riche [6][7][8] en terme de mesure d'orientation de surface, qui nécessite cependant un traitement global particulier. Des résultats de ce traitement sont présentés pour quelques scènes réelles et montrent l'intérêt de ce type d'analyse pour la reconstruction 3D.

2) Nature du phénomène physique et information géométrique

Le principe physique qui est exploité dans notre méthode est la polarisation d'un front d'onde non polarisé par réflexion

sur une surface diélectrique. En effet, un front de ce type est réfléchi sous forme d'un front partiellement polarisé, pour peu que la surface rencontrée présente une composante spéculaire non négligeable [7].

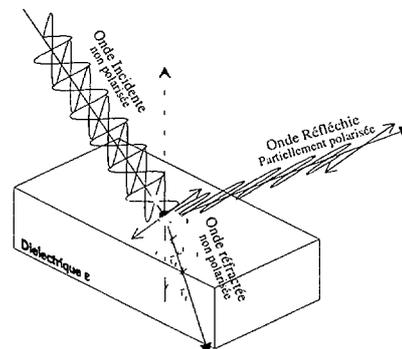


Figure 1 : Polarisation par réflexion

Les équations de Fresnel montrent que le front réfléchi possède deux composantes de polarisation ; une de ces composantes est parallèle au plan d'incidence, l'autre est perpendiculaire à ce même plan. Les importances relatives de ces composantes en fonction de l'angle d'incidence sont traduites, pour l'exemple du verre, par les courbes suivantes:

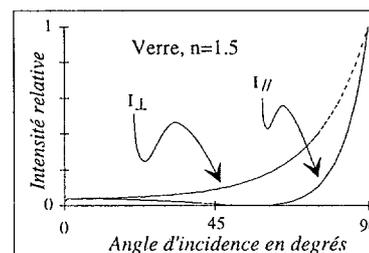


Figure 2 : Prépondérance du champ perpendiculaire, pour le verre

Ces courbes mettent notamment en évidence la prépondérance du champ perpendiculaire au plan d'incidence,



en dehors des cas limites $i=0^\circ$ et $i=90^\circ$. Cette propriété reste vraie pour un large éventail d'indices de réfraction du matériau rencontré, même si cet indice possède une partie imaginaire non-nulle, comme c'est le cas pour les métaux.

Etant donnée cette prépondérance, il apparaît donc qu'une grandeur optique dépendant géométriquement du plan d'incidence local émerge souvent d'une réflexion ; le fait d'analyser l'orientation de cette grandeur optique particulière permet d'obtenir une information intéressante concernant la localisation du plan d'incidence, et par la même, l'orientation de l'élément de surface observé.

Dans le cas d'une configuration faisant intervenir une source d'éclairage étendue et homogène de couleur blanche (donc non-polarisée ; ce type d'éclairage est celui qui est le plus fréquemment rencontré), un élément de surface répondant aux critères prédéfinis, et un système d'observation, il est donc possible de caractériser un plan d'incidence vectoriel local qui contient par propriété un vecteur normal à l'élément de surface considéré, pour un grand nombre des points d'une scène. En effet, le type d'éclairage retenu permet par propriété de définir une réflexion spéculaire si les orientations de surface répondent à quelques conditions restrictives, en l'occurrence que les incidences ne soient ni rasantes ni normales, c'est-à-dire que les surfaces ne soient ni parallèles ni perpendiculaires à l'axe optique du système d'observation.

Partant de ce principe, un plan d'incidence vectoriel peut-être défini par le vecteur porté depuis le centre optique des lentilles de l'objectif jusqu'au point image, et par le vecteur caractérisé par l'angle formé par le plan de polarisation prépondérant par rapport à une référence fixe. Ce plan correspond concrètement à la figure suivante :

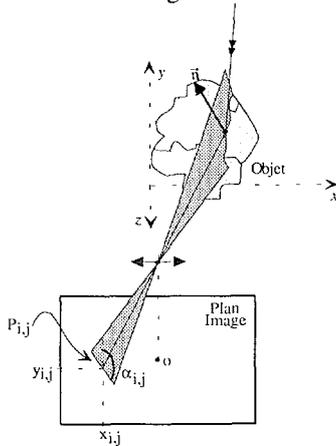


Figure 3 : Information géométrique recueillie

Dans le repère (x,y,z) défini sur la figure (3), cette particularité permet de définir un plan vectoriel π_{ij} pour tout point p_{ij} de l'image :

$$\pi_{i,j} = \langle \mathbf{u}_{i,j}, \mathbf{v}_{i,j} \rangle \text{ avec } \mathbf{u}_{i,j} = \begin{pmatrix} x_{i,j} \\ y_{i,j} \\ f \end{pmatrix} \text{ et } \mathbf{v}_{i,j} = \begin{pmatrix} \sin(\alpha_{i,j}) \\ \cos(\alpha_{i,j}) \\ 0 \end{pmatrix} \quad (1)$$

Dans cette définition, f désigne la distance entre le centre optique des lentilles et la plan image, et α_{ij} l'angle de polarisation obtenu au point considéré. Cette relation apporte donc une description vectorielle des orientations dans la scène, qui servira d'outil de base dans la suite des traitements. Ce principe sous-entend cependant que l'angle de polarisation α_{ij} soit connu pour un maximum de points de l'image. La partie qui suit envisage ce problème dans le cadre d'un système de traitement d'images.

3) Estimation des paramètres de polarisation à partir d'un système de traitement d'image

Les caméras électroniques et dans une moindre mesure les systèmes biologiques ne sont pas sensibles aux paramètres de polarisation de la lumière. Cette particularité implique la nécessité d'adjoindre un processeur optique pour obtenir une sensibilité à ce paramètre.

Notre choix s'est porté vers un filtre polarisant rectiligne. Les lois de Malus définissent, pour la situation polarimétrique décrite précédemment, la fonction d'éclairement suivante pour un point p_{ij} :

$$\tilde{I}_{i,j}(\theta) = A_{i,j} \cdot \cos^2(\phi_{i,j} - \theta) + B_{i,j} \quad (2)$$

où A_{ij} et B_{ij} intègrent des termes liés à l'amplitude du phénomène de polarisation et θ représente l'angle que forme l'axe privilégié du filtre polarisant par rapport à une référence verticale. L'intensité perçue est donc modulée par l'angle de polarisation recherché ϕ_{ij} qui correspond à l'angle α_{ij} augmenté de $\pi/2$ (le plan d'incidence est perpendiculaire au plan de polarisation prépondérant).

Partant de ce modèle, une méthode de recherche de l'angle de polarisation ϕ_{ij} fondée sur un critère à moindres carrés permet d'obtenir un compromis intéressant entre la précision souhaitée et les temps de calcul. Ce critère J_{ij} s'écrit donc :

$$J_{i,j} = \sum_{k=0}^n (I_{k,i,j} - \tilde{I}_{i,j}(\theta_k))^2 \quad (3)$$

où $I_{k,i,j}$ désigne l'intensité lumineuse perçue pour un angle θ_k . Pour simplifier cette expression, une loi de progression a été définie pour le filtre polarisant. Ce filtre possédant une périodicité de π , la progression suivante procure distribution suffisamment étendue :

$$\theta_k = \frac{k \cdot \pi}{n+1} \quad (4)$$

Dès lors, la minimisation de J_{ij} apporte la meilleure solution pour ϕ_{ij} :

$$\frac{dJ_{i,j}}{d\phi_{i,j}} = 0 \Rightarrow \phi_{i,j} = \frac{1}{2} \text{Arctg} \left(\frac{\sum_{k=0}^n I_{k,i,j} \cdot \sin\left(\frac{2 \cdot k \cdot \pi}{n+1}\right)}{\sum_{m=0}^n I_{m,i,j} \cdot \cos\left(\frac{2 \cdot m \cdot \pi}{n+1}\right)} \right) \quad (5)$$

Ce critère propose donc de manière efficace une valeur d'angle de polarisation (angle existant entre le plan de polarisation prépondérant et une référence verticale) pour chacun des points d'une scène. Dans certains cas limites, cette valeur n'a cependant pas de sens concret. C'est notamment le cas des surfaces lambertiennes qui renvoient un front d'onde non-polarisé. La valeur ϕ_{ij} est donc validée par un critère d'exploitabilité. Ce critère a été choisi comme étant l'amplitude du phénomène de polarisation. Si l'écart entre les $n+1$ valeurs I_{ijk} est faible, le résultat est médiocre, compte tenu de la quantification et du bruit lié à la chaîne de traitement. Le paramètre A_{ij} est fonction de cette valeur. L'annulation de la dérivée de J_{ij} par A_{ij} donne la meilleure valeur pour A_{ij} . D'où l'expression suivante :

$$|A_{i,j}| = \frac{4 \cdot \sqrt{\left(\sum_{k=0}^n I_{k,i,j} \cdot \cos\left(\frac{2 \cdot k \cdot \pi}{n+1}\right) \right)^2 + \left(\sum_{m=0}^n I_{m,i,j} \cdot \sin\left(\frac{2 \cdot m \cdot \pi}{n+1}\right) \right)^2}}{n+1} \quad (6)$$

Par la suite, cette valeur est comparée à un seuil de crédibilité commun à toute l'image, et qui intègre les paramètres réducteurs du système.

Dans le cas de notre méthode, la valeur de n choisie est 3, ce qui correspond à l'acquisition de quatre images. Cette valeur procure, une précision acceptable et des temps de calcul raisonnables.

En effet, ce traitement appliqué à une scène de test réelle constituée elle-même d'un filtre polarisant rétroéclairé par une source diffuse, montre une erreur maximale commise de l'ordre de 2° sur une échelle allant de 0° à 180° , avec un temps de calcul de 1,8s sur un système MATRA Péricolor 3100, pour des images 512×512 . Les quatre images acquises par les caméras sont filtrées par un masque moyennneur avant d'être introduites dans le processus de calcul, de manière à réduire le bruit dû à la caméra. La précision globale est donc relativement correcte et adaptée aux traitements ultérieurs. La carte résultant de ce traitement est appelée *Carte des angles de polarisation*, et sert de carte de base pour la fusion avec les données radiométriques.

4) Fusion des données polarimétriques et radiométriques

L'état de polarisation procure la caractérisation d'un plan d'incidence vectoriel (Eq. 1). Cette donnée n'est pas suffisante pour caractériser entièrement l'orientation locale. Aussi est-il nécessaire d'au moins la doubler pour obtenir une information d'orientation complète. Le fait de doubler cette information revient à prendre un autre point de vue et donc à utiliser une conformation stéréovision, ce qui nécessite un certain nombre de choix concernant la configuration physique du système, et plus particulièrement la position des caméras. Dans notre cas, ce choix s'est porté vers des caméras séparées d'une distance c sur l'axe des x , et dont les plans images sont coplanaires ; ainsi, les domaines épipolaires se trouvent être des lignes droites parallèles qui, par une calibration adaptée, portent le même numéro dans les deux images. Dès lors, en supposant tout problème de mise en correspondance résolu (ce problème est abordé dans la suite), deux plans d'incidence vectoriels sont définis pour un même point de la scène. Leur intersection porte un vecteur normal local de la surface observée.

$$\langle \mathbf{u}_{g,i,j}, \mathbf{v}_{g,i,j} \rangle \cap \langle \mathbf{u}_{d,i,j}, \mathbf{v}_{d,i,j} \rangle = \langle \mathbf{n}_{i,j} \rangle \quad (7)$$

où $\langle \mathbf{n}_{i,j} \rangle$ représente la droite vectorielle normale à la surface, au point en position (i,j) , et $\mathbf{u}_g, \mathbf{v}_g, \mathbf{u}_d, \mathbf{v}_d$ les vecteurs définis dans l'équation (1) ramenés aux points de vue gauche et droit. Cette égalité engendre une formulation spécifique possible pour \mathbf{n}_{ij} , dans un repère $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ conforme à la figure 3:

$$\mathbf{n}_{i,j} = K_{i,j} \cdot \begin{pmatrix} -p_{i,j} \\ -q_{i,j} \\ 1 \end{pmatrix} \quad (8)$$

Les expressions de p_{ij} et q_{ij} sont alors les suivantes :

$$\left\{ \begin{array}{l} p_{i,j} = \frac{-\sin(\alpha_{g,i,j}) \cdot x_{d,i,j} \cdot \cos(\alpha_{d,i,j}) + \sin(\alpha_{d,i,j}) \cdot x_{g,i,j} \cdot \cos(\alpha_{g,i,j})}{f \cdot \sin(\alpha_{d,i,j} - \alpha_{g,i,j})} \\ q_{i,j} = \frac{\sin(\alpha_{d,i,j} - \alpha_{g,i,j}) \cdot y_{i,j} + \cos(\alpha_{d,i,j}) \cdot \cos(\alpha_{g,i,j}) \cdot (x_{g,i,j} - x_{d,i,j})}{f \cdot \sin(\alpha_{d,i,j} - \alpha_{g,i,j})} \end{array} \right. \quad (9)$$

Par propriété, p_{ij} et q_{ij} représentent les pentes locales de la profondeur z selon x et y . Ces pentes caractérisent entièrement

l'orientation en P_{ij} , et permettent d'obtenir l'ensemble des profondeurs notées z_{ij} le long d'un chemin, selon l'équation :

$$z_{i,j} = z_{i_0,j_0} + \sum_{(k,l)=(i_0,j_0)}^{(i,j)} p_{k,l} \cdot \Delta x + q_{k,l} \cdot \Delta y \quad (10)$$

Cette équation soulève cependant deux problèmes particuliers : d'une part elle sous-entend la connaissance d'un point initiateur P_{i_0,j_0} , et d'autre part elle suppose que les pentes p et q sont toujours définies, alors que leur définition présuppose que la mise en correspondance est effectuée (et donc que la profondeur est connue). Cette dualité est résolue dans notre cas par la fusion des données d'orientation avec des grandeurs initiatrices provenant d'un système radiométrique associé, et par le recours à un processus récursif qui permet la détermination de tous les paramètres de façon progressive.

Le système de stéréovision radiométrique, dont les spécificités sont détaillées dans la partie *Résultats*, fournit des valeurs de disparités pour des points correspondant aux points de contour de la scène. Sa configuration physique est la même que celle définie précédemment, et permet de traiter le problème de mise en correspondance de façon monodimensionnelle. Cette particularité fixe un chemin particulier pour l'équation (10). Dans notre traitement, cette équation est donc limitée à l'équation suivante :

$$z_{i,j} = z_{i_0,j} + \sum_{k=i_0}^i p_{k,j} \cdot \Delta x \quad (11)$$

Dans la suite, la notion de disparité est préférée à la notion de profondeur, pour des raisons de compatibilité avec le système radiométrique. Entre ces deux grandeurs, une relation analytique est établie en fonction de la configuration retenue.

$$z_{i,j} = \frac{f \cdot c}{d_{i,j}} \quad (12)$$

Cette équation permet de définir l'équation (11) en fonction de la disparité et sous une forme récursive dont le terme générique est :

$$d_{i+1,j} = d_{i,j} + \left(\frac{1}{1 + \frac{p_{i,j} \cdot c}{f}} \right) \quad (13)$$

où e désigne la distance entre deux pixels au niveau plan image. Cette relation, qui n'a de sens que si une disparité a été fournie par le système radiométrique, réduit les conséquences de l'interdépendance des valeurs des pentes et de la disparité. En effet, l'utilisation successive des équations (13) et (9) au sein d'une ligne permet de déterminer de proche en proche chacun de ces paramètres pour chacun des points. Dès lors, cette méthode d'intégration s'identifie au développement d'une double suite. La convergence et la stabilité de cette suite ne peuvent être garanties a priori, étant donné le peu de corrélation des données mises en oeuvre. Cependant, des conditions de prise de vue adaptées (écartement suffisant des caméras) permettent de réduire efficacement les phénomènes d'instabilité.

En marge des termes génériques (9) et (13), un certain nombre de tests est effectué, de manière à prendre en compte d'éventuelles absences de données polarimétriques, et à proposer des données les plus vraisemblables en fonction du voisinage.

La partie qui suit présente des résultats de ce traitement particulier sur des scènes génériques réelles, qui montrent



l'intérêt de notre principe.

5) Résultats

Le processus que nous avons développé a été implémenté sur un système Matra Péricolor 3100. Le système de stéréovision radiométrique qui est utilisé pour la fusion possède des particularités remarquables comme l'utilisation des *déclivités* pour la phase de segmentation et la mise en oeuvre de la programmation dynamique pour l'étape de mise en correspondance[2] ; ces principes permettent d'obtenir des cartes d'une bonne précision de manière très rapide.

Les résultats proposés pour le processus complet concernent deux scènes génériques constituées d'une part d'une pyramide de 2 cm de hauteur, et d'autre part d'une sphère de bakélite de 2,5cm de diamètre. Les prises de vues sont effectuées en vue plongeante, les caméras sont espacées de 2,45cm et situées à 15 cm des objets. Les images résultats correspondent aux cartes de disparités obtenues. Ces images possèdent un codage particulier : un niveau de gris 0 correspond à une disparité de 150, un niveau de gris 255 correspond à une disparité de 210. La première image concerne la pyramide.

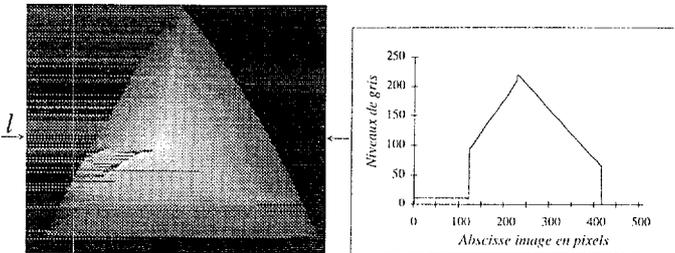


Figure 4 : Image résultat de la pyramide

Cette carte montre une bonne correspondance avec la scène originale. Cette correspondance est également vérifiée sur le profil de ligne effectué pour la ligne *l* : les portions correspondant aux faces de la pyramide sont bien droites. L'erreur maximale commise au sommet de la pyramide est de 1,5mm, ce qui est très acceptable au regard de la configuration matérielle mise en oeuvre.

La deuxième image résultat concerne la sphère de bakélite.

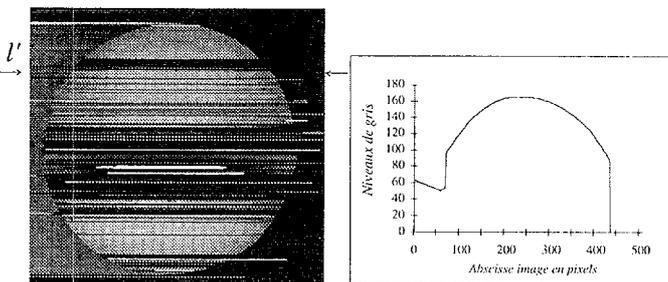


Figure 5 : Image résultat de la sphère de bakélite

Cette carte présente les mêmes particularités que la carte précédente. Dans ce cas, le profil de ligne correspond bien à la coupe d'une sphère. De même l'erreur commise au sommet de la sphère est de l'ordre de 2mm. Par contre, un certain nombre de défauts apparaissent de façon plus marquée que dans le cas de la pyramide : certaines lignes, voire certaines zones, ne sont pas reconstruites. Ces défauts proviennent principalement d'un manque de points de disparité initiale issus du système radiométrique (les contours d'une sphère correspondent aux limites de fonctionnement d'un tel système) et d'une trop forte ressemblance polarimétrique observée depuis les deux points

de vue. Des conditions de prise de vue mieux adaptées permettront dans un avenir proche de réduire sensiblement ce problème.

6) Conclusion

La méthode que nous avons développée et qui est fondée sur l'analyse de l'état de polarisation représente une alternative passive intéressante aux systèmes de reconstruction de surface purement radiométriques. Les résultats proposés montrent une réelle capacité à traiter à la fois des surfaces planes ou courbes. Cette particularité est relativement rare dans le cas d'un système purement passif.

Certes, un certain nombre de défauts subsistent dans la configuration actuelle. Une trop forte ressemblance polarimétrique engendre une absence de réponse en terme d'orientation, ce qui se traduit par des ruptures dans le processus final d'intégration.

Néanmoins, de prochains développements mettant en oeuvre notamment une augmentation du nombre de points de vue et l'utilisation d'un système radiométrique plus précis, doivent permettre à terme d'obtenir des résultats généraux de meilleure qualité, pour des scènes dont les orientations sont très variées.

7) Bibliographie :

- [1] E. Joubert, M. Ketata, P. Miché, R. Debric, *Utilisation de la polarisation de la lumière dans un système de vision*, CONGRES OPTO92, PARIS, pp345-349, 1992.
- [2] P. Miché et al., *Fast and self-adaptive stereovision algorithms for implementation in a 3-D vision sensor*, 3rd International Workshop on advanced motion control, Berkeley, USA, pp993-1001, 1994.
- [3] L. B. Wolff, *Surface orientation from two camera stereo with polarizers*, SPIE, Opt. Illumination and Image sensing for machine vision IV, Vol. 1194, pp287-297, 1989.
- [4] K. Koshikawa, Y. Shirai, *A model-based recognition of glossy objects using their polarimetric properties*, Advanced Robotics, Vol. 2, N° 2, pp137-147, 1987.
- [5] H. Sahli, J. Harthong, *Absolute measurement of 3-D object shapes and deformation by using Moiré topography*, CONFERENCE AUTOMATISATION INDUSTRIELLE, Montréal, 1-3 Juin 1992, Vol. 2, pp21.25-21.31, 1992.
- [6] W. Cheung, F.P. Ferric, G. Carayannis, J.B. Edwards, *Rockpile surface decomposition machine vision in mining*, CONFERENCE AUTOMATISATION INDUSTRIELLE, Montréal, 1-3 Juin 1992, Vol. 2, pp21.33-21.36, 1992.
- [7] B.K.P. Horn, *Shape from shading: a method for obtaining the shape of smooth opaque object from one view*, Ph.D. THESIS, Dep. of Elect. Eng. MIT, 1970.
- [8] F. Devernay, O. Faugeras, *Computing differential properties of 3-D shapes from stereoscopic images without 3-D models*, Rapport de recherche INRIA N° 2304, Juillet 94, 28 pages